

Estadística para las ciencias del comportamiento

7^a EDICIÓN

Robert R. Pagano

University of Pittsburgh

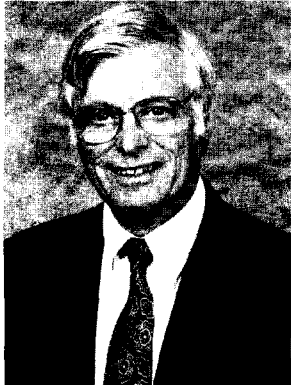
Revisión técnica:

Dr. Florente López Velázquez

Facultad de Psicología, UNAM

THOMSON
—★—™

Australia • Brasil • Canadá • España • Estados Unidos • México • Reino Unido • Singapur



SEMBLANZA DEL AUTOR

Robert R. Pagano estudió ingeniería eléctrica en el Instituto Politécnico Rensselaer en 1956 e hizo un doctorado en psicología biológica en la Universidad de Yale en 1965. Fue profesor asistente y profesor asociado del Departamento de Psicología en la Universidad de Washington en Seattle, Washington, de 1965 a 1989. Fue presidente asociado del Departamento de Neurociencia de la Universidad de Pittsburgh en Pittsburgh, Pennsylvania, de 1990 a junio de 2000. Además de sus restantes obligaciones como director de carrera, fue asesor departamental para especialidades profesionistas, impartió cursos de estadística a nivel de licenciatura y posgrado y se desempeñó como consultor de estadística para el profesorado del departamento. Asimismo, Bob fue director de los núcleos de estadística para dos subvenciones del centro NIH en esquizofrenia y enfermedad de Parkinson. Se retiró de la Universidad de Pittsburgh en junio de 2000. Los intereses de investigación de Bob se orientan a la psicobiología del aprendizaje y la memoria y a la psicología de la conciencia. Impartió cursos de la introducción a la estadística en la Universidad de Washington y en la Universidad de Pittsburgh por más de 30 años. Ha sido finalista entre los aspirantes al premio de la excelencia pedagógica en la Universidad de Washington, en reconocimiento de su labor docente en la asignatura de introducción a la estadística.

Bob está casado con Carol A. Eikleberry con quien tiene un hijo de 12 años, Robby. Bob tiene cinco hijas mayores, Renee, Laura, Maria, Elizabeth y Christina, y una nieta, Mikaela. La jubilación le abre nuevas oportunidades que complementan sus intereses por la enseñanza y la escritura. A Bob le encanta el tenis y actualmente se está entrenando para presentarse en Wimbledon (aunque la verdad es que, hasta ahora, le cuesta trabajo vencer a su hijo Robby). Le encantan también las actividades al aire libre, en particular el excursionismo y el café de la mañana.



CONTENIDO BREVE

parte uno **PANORAMA 1**

- 1 La estadística y el método científico 3

parte dos **ESTADÍSTICA DESCRIPTIVA 21**

- 2 Conceptos básicos de matemáticas y de medición 23
- 3 Distribuciones de frecuencias 38
- 4 Medidas de tendencia central y variabilidad 66
- 5 La curva normal y los puntajes estándar 86
- 6 Correlación 103
- 7 Regresión lineal 134

parte tres **ESTADÍSTICA INFERENCIAL 159**

- 8 Muestreo aleatorio y probabilidad 161
- 9 Distribución binomial 197
- 10 Introducción a la prueba de hipótesis mediante la prueba del signo 214
- 11 Potencia 243
- 12 Distribuciones muestrales, distribución muestral de la media, prueba de la desviación normalizada 263
- 13 Prueba t de Student para muestras simples 293
- 14 Prueba t de Student para grupos correlacionados e independientes 317
- 15 Introducción al análisis de varianza 354
- 16 Comparaciones múltiples 381
- 17 Introducción al análisis de varianza de dos factores o vías 395
- 18 Ji-cuadrada y otras pruebas no paramétricas 425
- 19 Repaso de estadística inferencial 465



CONTENIDO

parte uno PANORAMA 1

CAPÍTULO 1 La estadística y el método científico 3

- Introducción 3
- Métodos de adquisición del conocimiento 4
- Definiciones 6
 - Un ejemplo general: modo de presentación 7
- Investigación científica y estadística 9
- Muestreo aleatorio 9
- Estadística descriptiva e inferencial 10
- Uso de la computadora en estadística 10
- La estadística y el “mundo real” 11
- ¿CUÁL ES LA VERDAD? Datos, datos, ¿dónde están los datos? 12
- ¿CUÁL ES LA VERDAD? Las autoridades, están bien, pero . . . 13
- ¿CUÁL ES LA VERDAD? Datos, datos, ¿qué son los datos?—1 14
- ¿CUÁL ES LA VERDAD? Datos, datos, ¿qué son los datos?—2 15
- Resumen 17
- Términos importantes 17
- Preguntas y problemas 17

parte dos ESTADÍSTICA DESCRIPTIVA 21

CAPÍTULO 2 Conceptos básicos de matemáticas y de medición 23

- Sugerencias de estudio para el estudiante 23
- Notación matemática 24
- Sumatoria 25

Escalas de medición: Nominales, ordinales, de intervalos y de razón	27
Escalas nominales	27
Escalas ordinales	28
Escalas de intervalos	29
Escalas de razón	29
Escalas de medición en las ciencias del comportamiento	31
Variables continuas y discretas	31
Límites reales de una variable continua	32
Cifras significativas	33
Redondeo	34
Resumen	35
Términos importantes	35
Preguntas y problemas	35
Notas	36

CAPÍTULO 3 Distribuciones de frecuencias 38

Introducción: Distribuciones de frecuencias no agrupadas	38
Datos agrupados	40
Elaboración de una distribución de frecuencias de datos agrupados	41
Frecuencia relativa, frecuencia acumulada y distribuciones de porcentajes acumulados	45
Percentiles	46
Cálculo de los puntos percentiles	47
Rangos percentiles	50
Cálculo de rangos percentiles	50
Graficación de las distribuciones de frecuencias	53
La gráfica de barras	53
El histograma	55
El polígono de frecuencias	56
La curva del porcentaje acumulado	56
Forma de las curvas de frecuencia	57
Análisis exploratorio de los datos	58
Diagramas de tallo y hojas	58
¿CUÁL ES LA VERDAD? Al cambiar la escala, cambia el cuento	60
Resumen	62
Términos importantes	62
Preguntas y problemas	62

CAPÍTULO 4 Medidas de tendencia central y variabilidad 66

Introducción	66
Medidas de tendencia central: media aritmética, mediana y moda	67
La media aritmética	67
La media global	70

La mediana	72
La moda	74
Medidas de tendencia central y simetría	75
Medidas de la variabilidad: rango, desviación estándar y varianza	76
El rango	76
La desviación estándar	76
La varianza	82
Resumen	82
Términos importantes	82
Preguntas y problemas	82
Notas	85

CAPÍTULO 5 La curva normal y los puntajes estándar 86

Introducción	86
La curva normal	87
Área debajo de la curva normal	88
Puntajes estándar (puntajes z)	88
Características de los puntajes z	91
Determinación de las áreas correspondientes a cualquier dato en bruto	92
Determinación del dato en bruto correspondiente a un área dada	97
Resumen	100
Términos importantes	100
Preguntas y problemas	100

CAPÍTULO 6 Correlación 103

Introducción	103
Relaciones	104
Relaciones lineales	104
Relaciones positivas y negativas	107
Relaciones perfectas e imperfectas	108
Correlación	111
El coeficiente de correlación lineal r de Pearson	112
Otros coeficientes de correlación	120
Efecto del rango sobre la correlación	124
Efecto de los datos extremos	125
La correlación no implica causalidad	125
¿CUÁL ES LA VERDAD? "Un buen director = una buena escuela", ¿o no es así?	126
Resumen	127
Términos importantes	128
Preguntas y problemas	128

CAPÍTULO 7	Regresión lineal	134
	Introducción	134
	Predicción y relaciones imperfectas	135
	Construcción de la recta de regresión por mínimos cuadrados:	
	Regresión de Y sobre X	137
	Regresión de X sobre Y	143
	Medición de errores de predicción: el error estándar de la estimación	146
	Consideraciones al utilizar la regresión lineal para la predicción	149
	Relación entre las constantes de regresión y la r de Pearson	150
	Regresión múltiple y correlación múltiple	151
	Resumen	155
	Términos importantes	156
	Preguntas y problemas	156
parte tres	ESTADÍSTICA INFERENCIAL	159
CAPÍTULO 8	Muestreo aleatorio y probabilidad	161
	Introducción	161
	Muestreo aleatorio	162
	Técnicas para el muestreo aleatorio	164
	Muestreo con o sin reemplazo	165
	Probabilidad	166
	Algunos aspectos básicos relativos a los valores de probabilidad	167
	Cálculo de probabilidades	167
	La regla de la suma	168
	La regla del producto	173
	Reglas de la suma y el producto	183
	Probabilidad y variables continuas con distribución normal	186
	¿CUÁL ES LA VERDAD? "No soy culpable, soy una víctima de las coincidencias": ¿excusa descarada o verdad?	189
	¿CUÁL ES LA VERDAD? Declive del número de espermatozoides: ¿Falla del hombre o del muestreo?	190
	¿CUÁL ES LA VERDAD? La muestra de una muestra	191
	Resumen	193
	Términos importantes	193
	Preguntas y problemas	193
	Notas	196
CAPÍTULO 9	Distribución binomial	197
	Introducción	197
	Definición e ilustración de la distribución binomial	197

Generación de la distribución binomial a partir del desarrollo binomial	200
Uso de la tabla binomial	202
Resumen	211
Términos importantes	211
Preguntas y problemas	211
Notas	213

CAPÍTULO 10 **Introducción a la prueba de hipótesis mediante la prueba del signo 214**

Introducción	214
Un experimento: La marihuana y el tratamiento de pacientes con SIDA	215
Diseño de medidas repetidas	217
Hipótesis alternativa (H_1)	218
Hipótesis nula (H_0)	218
Regla de decisión (nivel α)	218
Evaluación del experimento de la marihuana mediante la distribución binomial	219
Errores de tipo I y de tipo II	220
El nivel alfa y el proceso de decisión	221
Evaluación de la cola de la distribución	223
Evaluaciones de probabilidad para una y dos colas	225
Magnitud del efecto: Significativo versus importante	232
¿CUÁL ES LA VERDAD? ¿Azar o efecto real?	232
¿CUÁL ES LA VERDAD? ¿Azar o efecto real?	234
¿CUÁL ES LA VERDAD? "Ningún producto es mejor que el nuestro"	235
¿CUÁL ES LA VERDAD? Informes anecdóticos contra investigación sistemática	236
Resumen	238
Términos importantes	239
Preguntas y problemas	239
Notas	242

CAPÍTULO 11 **Potencia 243**

Introducción	243
¿Qué es potencia?	244
P_{nula} y P_{real}	244
P_{real} : Una medida de la magnitud y la dirección del efecto real	245
Análisis de potencia del experimento sobre el SIDA	246
Efecto de N y magnitud del efecto real	247
Potencia y beta (β)	251
Potencia y alfa (α)	252
Alfa-beta y realidad	253
Interpretación de resultados no significativos	253

Cálculo de potencia	254
¿CUÁL ES LA VERDAD? Astrología y ciencia	259
Resumen	261
Términos importantes	261
Preguntas y problemas	261
Notas	262

CAPÍTULO 12 **Distribuciones muestrales, distribución muestral de la media, prueba de la desviación normalizada** 263

Introducción	263
Distribuciones muestrales	264
Generación de distribuciones muestrales	264
La prueba (z) de la desviación normalizada	268
<i>Un experimento: Evaluación del programa de lectura en las escuelas públicas locales</i>	268
Distribución muestral de la media	268
Revisión del experimento sobre la habilidad para la lectura	275
Solución alternativa utilizando z_{obt} y la región crítica para el rechazo de H_0	277
Condiciones en las cuales resulta adecuada la prueba z	282
La potencia y la prueba z	282
Resumen	290
Términos importantes	290
Preguntas y problemas	290

CAPÍTULO 13 **Prueba t de Student para muestras simples** 293

Introducción	293
Comparación entre las pruebas z y t	294
<i>Un experimento: Técnica para incrementar el habla temprana en los niños</i>	294
La distribución muestral de t	295
Grados de libertad	296
Comparación de las distribuciones t y z	297
Revisión del experimento de la emisión de las primeras palabras	298
Cálculo de t_{obt} a partir de los datos originales	299
Condiciones bajo las cuales la prueba t es adecuada	304
Distribución t : Intervalos de confianza para la media poblacional	304
Construcción de un intervalo de confianza al 95%	305
<i>Ejemplo: Estimación del CI promedio de los profesores de una universidad</i>	306
Ecuaciones generales para cualquier intervalo de confianza	307
Prueba de la significancia de la r de Pearson	309
Resumen	312

Términos importantes	313
Preguntas y problemas	313
Notas	315

CAPÍTULO 14 Prueba t de Student para grupos correlacionados e independientes 317

Introducción	318
Prueba t de Student para grupos correlacionados	319
<i>Un experimento: Estimulación cerebral y alimentación</i>	319
Comparación de la prueba t para grupos correlacionados y la prueba t para muestras únicas	320
Revisión y análisis del experimento de estimulación cerebral	321
Comparación de la prueba t para grupos correlacionados y la prueba de los signos	326
Supuestos subyacentes a la prueba t para grupos correlacionados	327
Pruebas z y t para grupos independientes	327
Diseño de grupos independientes	327
Prueba z para grupos independientes	328
<i>Un experimento: Hormona X y comportamiento sexual</i>	328
La distribución muestral de la diferencia entre las medias muestrales ($\bar{X}_1 - \bar{X}_2$)	329
<i>Un experimento: Revisión del experimento de la hormona X</i>	330
Prueba t de Student para grupos independientes	331
Comparación de las ecuaciones para z_{obt} y t_{obt}	331
Análisis del experimento de la hormona X con la prueba t para grupos independientes	333
Cálculo de t_{obt} cuando $n_1 = n_2$	334
Supuestos subyacentes a la prueba t para grupos independientes	338
Violación de los supuestos de la prueba t para grupos independientes	339
Magnitud del efecto	339
Potencia de la prueba t	341
Comparación de los diseños para grupos correlacionados y grupos independientes	342
Resumen	345
Términos importantes	346
Preguntas y problemas	347
Notas	352

CAPÍTULO 15 Introducción al análisis de varianza 354

Introducción: La distribución F	354
La prueba F y el análisis de varianza (ANOVA)	356
Panorama de la técnica del análisis de varianza de un factor o vía	357
Estimación de la varianza dentro de grupos, s_w^2	359

	Estimación de la varianza entre grupos, s_B^2	360
	La razón F	362
	Análisis de datos con la técnica del análisis de varianza	362
	<i>Un experimento: Situaciones diversas y tensión</i>	362
	Lógica subyacente al análisis de varianza de un factor	366
	La relación entre el análisis de varianza y la prueba t	370
	<i>Ejemplo: Citas. Las jóvenes de clubes femeninos</i>	
	estudiantiles versus las de dormitorios	370
	Supuestos subyacentes al análisis de varianza	372
	Magnitud del efecto	373
	Potencia del análisis de varianza	373
	Potencia y N	374
	Potencia y el efecto real de la variable independiente	374
	Potencia y la variabilidad de la muestra	374
	¿CUÁL ES LA VERDAD? Mucho ruido y pocas nueces	375
	Resumen	376
	Términos importantes	377
	Preguntas y problemas	377
CAPÍTULO 16	Multiple Comparisons	381
	Introducción	381
	Comparaciones <i>a priori</i> , o planeadas	382
	Comparaciones <i>a posteriori</i> o <i>post hoc</i>	384
	La prueba de diferencia honestamente	
	significativa (DHS) de Tukey	385
	La prueba de Newman-Keuls	386
	Pruebas DHS y Newman-Keuls con n distinta	391
	Contrastación entre las comparaciones planeadas, la prueba DHS	
	de Tukey y la prueba de Newman-Keuls	391
	Resumen	392
	Términos importantes	392
	Preguntas y problemas	393
	Notas	393
CAPÍTULO 17	Introducción al análisis de varianza	
	de dos factores o vías	395
	Introducción al análisis de varianza de dos	
	factores-Presentación cualitativa	395
	Presentación cuantitativa del análisis de varianza de dos factores	399
	Estimación de la varianza dentro de las celdas (s_W^2)	400
	Estimación de la varianza por renglón (s_R^2)	402
	Estimación de la varianza por columna (s_C^2)	404
	Estimación de la varianza por renglón \times columna (s_{RC}^2)	405
	Cálculo de las razones F	406

Análisis de un experimento con el ANOVA de dos factores	406
<i>Un experimento: Efecto del ejercicio sobre el sueño</i>	406
Interpretación de los resultados	410
Comparaciones múltiples	420
Supuestos subyacentes al ANOVA de dos factores	421
Resumen	421
Términos importantes	422
Preguntas y problemas	422

CAPÍTULO 18 Ji-cuadrada y otras pruebas no paramétricas 425

Introducción: Distinción entre las pruebas paramétricas y las no paramétricas	426
Ji-cuadrada (χ^2)	427
Experimentos con una variable	427
<i>Un experimento: Preferencia por las diversas marcas de cerveza ligera</i>	427
Prueba de independencia entre dos variables	431
<i>Un experimento: Afiliación política y actitud</i>	432
Supuestos subyacentes a χ^2	440
La prueba de rangos con signo y pares igualados de Wilcoxon	441
<i>Un experimento: Promoción de actitudes más favorables hacia la conservación de la vida silvestre</i>	441
Supuestos de la prueba de rangos con signo de Wilcoxon	444
La prueba <i>U</i> de Mann–Whitney	444
<i>Un experimento: Efecto de una dieta alta en proteínas sobre el desarrollo intelectual</i>	444
Rangos empatados	448
Supuestos subyacentes de la prueba <i>U</i> de Mann–Whitney	450
La Prueba de Kruskal–Wallis	450
<i>Un experimento: Evaluación de dos programas para adelgazar</i>	450
Supuestos subyacentes de la prueba de Kruskal–Wallis	454
¿CUÁL ES LA VERDAD? Estadística e investigación social aplicada: ¿Útiles o "abusivas"?	455
Resumen	457
Términos importantes	458
Preguntas y problemas	458
Notas	464

CAPÍTULO 19 Repaso de estadística inferencial 465

Introducción	465
Términos y conceptos	466
Proceso de la prueba de hipótesis	467
Diseños con una muestra	468
Prueba <i>z</i> para muestras simples	468
Prueba <i>t</i> para muestras simples	469

Prueba t para probar la significancia de la r de Pearson	469
Experimentos con dos condiciones:	
Diseño de grupos correlacionados	470
Prueba t para grupos correlacionados	470
Prueba de rangos con signo y pares acoplados de Wilcoxon	471
Prueba de los signos	471
Diseño de grupos independientes: Dos grupos	472
Prueba t para grupos independientes	472
Prueba U de Mann–Whitney	473
Experimentos con varios grupos	474
Análisis de varianza de un factor, prueba F	474
Análisis de varianza de un factor, prueba de Kruskal–Wallis	477
Análisis de varianza de dos factores, prueba F	477
Análisis de varianza de dos factores, prueba F	478
Análisis de datos nominales	479
Prueba χ^2 -cuadrada	479
Elección de la prueba adecuada	480
Preguntas y problemas	482

Apéndices 489

A. Repaso de los prerrequisitos de matemáticas	491
B. Ecuaciones	501
C. Respuestas a las preguntas y problemas de final de capítulo	509
D. Tablas	524

Glosario 549

Índice 552



MATERIAL DE APOYO EN LA RED*

- Introducción
- CAPÍTULO 1** La estadística y el método científico
Problemas de práctica adicionales y soluciones: 1.1
- CAPÍTULO 2** Conceptos básicos de matemáticas y de medición
Problemas de práctica adicionales y soluciones: 2.1, 2.2, 2.3
Problemas para computadora SPSS: 2.1, 2.2
- CAPÍTULO 3** Distribuciones de frecuencias
Problemas de práctica adicionales y soluciones: 3.1, 3.2
Problemas para computadora SPSS: 3.1, 3.2
- CAPÍTULO 4** Medidas de tendencia central y variabilidad
Problemas de práctica adicionales y soluciones: 4.1, 4.2, 4.3
Problemas para computadora SPSS: 4.1, 4.2, 4.3, 4.4
Experimentos y problemas del libro de texto resueltos por medio de ecuaciones conceptuales: 4.6, 4.7
- CAPÍTULO 5** La curva normal y los puntajes estándar: 5.1, 5.2, 5.3, 5.4, 5.5
Problemas de práctica adicionales y soluciones
Problemas para computadora SPSS: 5.1, 5.2, 5.3
- CAPÍTULO 6** Correlación
Problemas de práctica adicionales y soluciones: 6.1, 6.2
Problemas para computadora SPSS: 6.1, 6.2, 6.3, 6.4, 6.5, 6.6
Experimentos y problemas del libro de texto resueltos por medio de ecuaciones conceptuales: 6.1, 6.2 y
DATOS DE LA TABLA CONCEPTUAL 6.5

*Este material está disponible sólo en inglés en el sitio de Thomson, www.thomsonlearning.com.mx. Visite la página, escriba PAGANO en la casilla "Buscar" y haga clic en "Material de apoyo" o "Material complementario".

- CAPÍTULO 7** Regresión lineal
 Problemas de práctica adicionales y soluciones: 7.1, 7.2
 Problemas para computadora SPSS: 7.1, 7.2, 7.3, 7.4, 7.5, 7.6
- CAPÍTULO 8** Muestreo aleatorio y probabilidad
 Problemas de práctica adicionales y soluciones: 8.1, 8.2
- CAPÍTULO 9** Distribución binomial
 Problemas de práctica adicionales y soluciones: 9.1, 9.2
- CAPÍTULO 10** Introducción a la prueba de hipótesis mediante la prueba de los signos
 Problemas de práctica adicionales y soluciones: 10.1, 10.2
- CAPÍTULO 11** Potencia
 Problemas de práctica adicionales y soluciones: 11.1
- CAPÍTULO 12** Distribuciones muestrales, distribución muestral de la media, prueba de la desviación normalizada
 Problemas de práctica adicionales y soluciones: 12.1, 12.2
- CAPÍTULO 13** Prueba *t* de Student para muestras simples
 Problemas de práctica adicionales y soluciones: 13.1, 13.2
 Problemas para computadora SPSS: 13.1, 13.2, 13.3, 13.4
 Experimentos y problemas del libro de texto resueltos por medio de ecuaciones conceptuales
 EXPERIMENTO CONCEPTUAL DE LA EMISIÓN DE PALABRAS-DATOS DE PUNTAJES EN BRUTO 13.1, Parte C, 13.2, Parte C
- CAPÍTULO 14** Prueba *t* de Student para grupos correlacionados e independientes
 Problemas de práctica adicionales y soluciones: 14.1, 14.2, 14.3, 14.4
 Problemas para computadora SPSS: 14.1, 14.2, 14.3, 14.4, 14.5, 14.6
 Experimentos y problemas del libro de texto resueltos por medio de ecuaciones conceptuales
 EXPERIMENTO CONCEPTUAL DE ESTIMULACIÓN CEREBRAL, PARTE 3: 14.1 y 14.2 Parte C
 EXPERIMENTO CONCEPTUAL DE LA HORMONA X, PARTE 3: 14.3 y 14.4 Parte C
- CAPÍTULO 15** Introducción al análisis de varianza
 Problemas de práctica adicionales y soluciones: 15.1, 15.2
 Problemas para computadora SPSS: 15.1, 15.2, 15.3, 15.4
 Experimentos y problemas del libro de texto resueltos por medio de ecuaciones conceptuales
 EXPERIMENTO CONCEPTUAL DE TENSIÓN: 15.1
 EXPERIMENTO CONCEPTUAL DE CITAS
- CAPÍTULO 16** Comparaciones múltiples: 16.1
- CAPÍTULO 17** Introducción al análisis de varianza de dos factores o vías
 Problemas de práctica adicionales y soluciones: 17.1
 Problemas para computadora SPSS: 17.1, 17.2, 17.3, 17.4
- CAPÍTULO 18** Ji-cuadrada y otras pruebas no paramétricas
 Problemas de práctica adicionales y soluciones: 18.1, 18.2
 Para acceder a los problemas de cada capítulo, haga clic en el título del capítulo y los problemas aparecerán en la columna de la derecha.



PREFACIO

He impartido el curso de introducción a la estadística durante más de 30 años, primero en el Departamento de Psicología de la Universidad de Washington y después en el Departamento de Neurociencias de la Universidad de Pittsburgh. Este libro de texto es el principal apoyo de dicho curso. La mayoría de mis estudiantes se han graduado en psicología, pero muchos provienen de áreas como biología, comercio, educación, neurociencias, enfermería y otros campos de estudio. Debido a que, por lo general, estos estudiantes no sienten atracción por las matemáticas ni tienen mucha habilidad en la materia, he utilizado una aproximación informal en lugar de una estrictamente matemática. Esta aproximación sólo presupone el conocimiento de un curso de álgebra de bachillerato. Se basa en una presentación clara, una secuencia particularmente eficaz del material inferencial, descripciones verbales detalladas, ilustrativos e interesantes ejemplos, así como problemas prácticos resueltos en forma completa para que los estudiantes comprendan el material.

Mi curso de estadística ha tenido mucho éxito. Los estudiantes son capaces de asimilar el material, incluso los temas más complejos, como el de "potencia", y al mismo tiempo llegan a disfrutar del aprendizaje. Las calificaciones de los alumnos que toman el curso han sido bastante altas. Los alumnos califican al libro de texto aún mejor: destacan, entre otras cosas, la gran claridad de la obra y sus toques de humor. Esto les ayuda a comprender el material, que se presenta con gran detalle.

Al preparar esta séptima edición, me he guiado por los comentarios y sugerencias de los estudiantes y profesores. De nuevo, estoy satisfecho de que esta retroalimentación haya resultado provechosa, y que para la mayor parte del libro el consejo fue el de no modificar nada, debido a que el texto funciona muy bien así. Sin embargo, en la séptima edición hay un cambio importante y varias modificaciones menores que vale la pena mencionar. El primer cambio importante es la adición de una nueva sección, en el capítulo 15, titulada "Potencia del análisis de varianza". (Ésta es una sección que acompaña a la sección titulada "Potencia de la prueba t " del capítulo 14.) Allí se amplía la exposición de la potencia de la prueba de los signos, la prueba z y la prueba t para el análisis de varianza de un solo factor o vía. El segundo cambio consiste en que en los problemas del final de capítulo de la séptima edición se resaltan visualmente las áreas de la psico-

logía y otros campos afines a los que se aplica el problema. Esta identificación ayudará a los profesores a asignar los problemas por campos específicos y hará que también los estudiantes aprecien la amplitud de las aplicaciones de la estadística. El tercer cambio consiste en tres nuevas secciones de “¿Cuál es la verdad?”: dos al capítulo 1 y una al capítulo 10.

El último cambio se refiere al material de apoyo disponible sólo en inglés en el sitio de Thomson, www.thomsonlearning.com.mx. Visite la página, escriba PAGANO en la casilla “Buscar” y haga clic en “Material de apoyo” o “Material complementario” para descargarlo. Para consultar los problemas, haga clic en el capítulo correspondiente al tema de su interés. Tenga en cuenta que los complementos electrónicos son independientes del libro y están sujetos a cambios o supresión parcial o total sin previo aviso.

Intención del libro

Éste es un libro de texto introductorio que abarca la estadística descriptiva y la inferencial. Está dirigido a los estudiantes de la segunda mitad de las carreras de ciencias del comportamiento. Para muchos de los alumnos del área, la estadística es un tema que genera mucha angustia y tratan de evitarla en la medida de lo posible. Además, creo justo el decir que al terminar un curso normal de estadística, muchos alumnos no han comprendido gran parte del material de la estadística inferencial. Esto se debe, en parte, a la dificultad inherente de esos temas y a que los propios alumnos no son hábiles en matemáticas; pero también se debe, desde mi punto de vista, a que la mayoría de los libros de texto no explican en forma adecuada la estadística inferencial para este grupo de estudiantes. En esos textos se cometen, por lo general, las siguientes equivocaciones: 1) no están escritos de manera clara; 2) no contienen suficientes detalles; 3) presentan el material en una forma demasiado matemática; 4) ofrecen dicho material a un nivel muy bajo; 5) no proporcionan un número suficiente de problemas resueltos de principio a fin, los cuales puedan ser utilizados por el estudiante, y 6), en lo referente a la estadística inferencial, los temas se presentan en un orden inadecuado, pues se empieza con la distribución muestral de la media.

En ésta y en las seis ediciones anteriores he intentado corregir tales deficiencias mediante un estilo coloquial de redacción; una presentación escrita de manera clara, detallada y orientada a la teoría, que sólo requiere para su comprensión el álgebra de bachillerato; la inclusión de muchos problemas prácticos interesantes, totalmente resueltos, localizados justo después de la explicación de los temas correspondientes, y un ordenamiento del material de estadística inferencial más adecuado para los estudiantes a quienes va dirigido este libro.

Considero que un factor clave para comprender la estadística inferencial es el material presentado en los primeros capítulos relativos a la inferencia y su orden. Desde mi punto de vista, se logra un aprendizaje óptimo del material cuando se procede en el siguiente orden: muestreo aleatorio y probabilidad, distribución binomial, introducción a la prueba de hipótesis mediante la prueba de los signos, la potencia, distribuciones muestrales (incluida su generación empírica), distribución muestral de la media, prueba z para muestras simples, prueba t para muestras simples, intervalos de confianza, prueba t para grupos correlacionados e independientes, introducción al análisis de varianza, comparaciones múltiples, análisis de varianza, pruebas no paramétricas y finalmente un repaso de toda la estadística inferencial.

El concepto de distribución muestral está en el centro de la inferencia estadística. La primera distribución muestral que es analizada en la mayoría de los textos es la distribución muestral de la media. La dificultad en este caso es que ésta no se puede generar a partir de consideraciones probabilísticas simples, lo cual dificulta su comprensión. Este problema se complica por el hecho de que en la mayoría de los textos no se intenta ge-

nerar la distribución muestral de la media de una manera concreta, sino se la define de manera teórica, como la distribución de probabilidad que surgiría al tomar una infinidad de muestras aleatorias de tamaño N de la población y calcular la media de cada muestra. Esta definición es demasiado abstracta para los estudiantes, en particular cuando es su primer contacto con la idea de distribuciones muestrales. *Y cuando los alumnos no logran captar el concepto de distribución muestral, no pueden captar tampoco el resto de la estadística inferencial.* Lo que parece ocurrir en esos casos es que, como los alumnos no comprenden el material desde un punto de vista conceptual, no tienen más remedio que memorizar las ecuaciones y resolver los problemas de memoria. Así, con frecuencia, los estudiantes resuelven los problemas sin comprender realmente lo que están haciendo; y todo porque no entienden la esencia de las distribuciones muestrales.

Para propiciar una comprensión básica, creo que es mucho mejor comenzar con la prueba de los signos (capítulo 10), una prueba de inferencia simple cuya distribución muestral adecuada es la distribución binomial. Esta distribución es fácil de comprender y se puede deducir de las reglas básicas de la probabilidad desarrolladas en un capítulo anterior (el capítulo 8, Muestreo aleatorio y probabilidad). Depende por completo de consideraciones lógicas, por lo cual es sencillo entender cómo se genera. Además, también se puede originar con facilidad mediante el mismo proceso empírico que se utiliza más adelante, en el texto, para generar la distribución muestral de la media. Por lo tanto, sirve como un puente importante para comprender todas las distribuciones muestrales que serán analizadas posteriormente en el texto. La presentación de las pruebas de hipótesis junto con la prueba de los signos tiene otras ventajas: permite ilustrar todos los conceptos importantes relativos a las pruebas de hipótesis; por ejemplo, la hipótesis nula, la hipótesis alternativa, el nivel alfa, los errores tipo I y II, y la potencia. La prueba de los signos ilustra también el diseño antes-después (empleando medidas repetidas), lo cual es una forma excelente de comenzar, pues la mayoría de los estudiantes ya están familiarizados con este tipo de experimento, de modo que les resulta muy fácil seguir la lógica del diseño.

En el capítulo 11 se expone la potencia. Este tema no figura en muchos textos o, cuando mucho, se lo menciona hacia el final del libro. La potencia estadística es un tema complicado. La cuestión se simplifica cuando la prueba de los signos se utiliza como vehículo para el análisis de la potencia. Es necesario entender este concepto para captar la metodología misma de la investigación científica. Cuando los estudiantes empiezan a vislumbrar el concepto de potencia, pueden comprender por qué nos tomamos la molestia de examinar los errores tipo II. Además, logran ver por primera vez por qué nuestra conclusión fue “conservar H_0 ” como una explicación razonable de los datos, en lugar de “aceptar la validez de H_0 ” (esta distinción es muy importante). En la misma tónica, los estudiantes aprecian también el error en que se incurre al concluir que existe igualdad entre dos condiciones, a partir de datos que no son estadísticamente significativos. Así, la potencia es un tema que permite enfocar con nitidez toda la metodología para poner a prueba las hipótesis.

A estas alturas de la exposición, un estudiante diligente puede captar la idea de que el análisis de datos consiste básicamente en dos pasos: 1) calcular el estadístico apropiado y 2) evaluar el estadístico a partir de su distribución muestral. El momento es apropiado para presentar una exposición formal de las distribuciones muestrales y las maneras en que pueden ser generadas (capítulo 12). Después de esto se presenta la distribución muestral de la media. En lugar de depender de una definición teórica abstracta de la distribución muestral de la media, el texto explica la forma en que esa distribución muestral puede generarse empíricamente. Esto proporciona una comprensión mucho más concreta de la distribución muestral de la media.

Gracias a la experiencia anterior con una distribución muestral fácil de entender, la distribución binomial, y aplicando el enfoque empírico para la distribución muestral de

la media, la mayoría de los estudiantes dedicados adquieren una buena comprensión del significado de las distribuciones muestrales y la razón por la cual son esenciales para la estadística inferencial. Como las distribuciones muestrales subyacentes a la prueba t de Student y al análisis de varianza también se explican en términos de su generación empírica, los estudiantes pueden comprender conceptualmente el uso de estas pruebas, en vez de contentarse con resolver los problemas de memoria. Este enfoque les proporciona una idea de que todos los conceptos de las pruebas de hipótesis son iguales al pasar de un estadístico a otro; lo que varía en cada uno de los experimentos es el estadístico utilizado y su distribución muestral asociada. El escenario está listo para analizar las demás pruebas de inferencia.

Los capítulos 12, 13, 14 y 18 analizan, de una manera convencional, las pruebas z y t para muestras simples, la prueba t para grupos correlacionados e independientes, y la estadística no paramétrica. Sin embargo, estos capítulos difieren de los que suelen presentarse en otros textos por la claridad de su exposición, por sus muchos e interesantes problemas completamente resueltos, y por el uso de las distribuciones muestrales deducidas en forma empírica. Además, también existen diferencias específicas para cada prueba. Por ejemplo: 1) la prueba t para grupos correlacionados se presenta inmediatamente después de la prueba t para muestras simples y se desarrolla como un caso particular de la prueba t para muestras simples, utilizando datos de diferencia y no datos en bruto; 2) la prueba t para grupos correlacionados y la de los signos se comparan para ilustrar la diferencia en la potencia estadística que resulta del uso de cada una de estas pruebas; 3) se analizan los factores que influyen sobre la potencia de los experimentos en los que se usa la prueba t de Student, y 4) se contrasta la utilidad de los diseños con grupos correlacionados e independientes.

Los capítulos 15 y 17 tratan sobre el análisis de la varianza. En estos capítulos se utiliza de manera deliberada la notación con un solo subíndice y no con subíndices dobles, pues esta última confunde a los estudiantes. Desde mi punto de vista, la notación con un subíndice y las resultantes sumas sencillas funcionan mejor para los estudiantes de último año en psicología y los campos afines, porque son más sencillas y, en el caso de ese tipo de usuarios, contribuyen a la comprensión de este material de suyo complejo. En este texto he seguido parcialmente la notación utilizada por Edward Minium en *Statistical Reasoning in Psychology and Education*, 2ª ed., John Wiley and Sons, Nueva York, 1978. Estoy en deuda con el profesor Minium por esta contribución.

Vale la pena resaltar otras características de este texto. El capítulo 8, que trata sobre probabilidad, no profundiza en la teoría de probabilidades. Esto no es necesario, pues las bases matemáticas adecuadas para todas las pruebas de inferencia contenidas en este libro se pueden construir, como se aprecia en el capítulo 8, mediante las definiciones básicas de probabilidad, las reglas de la suma y del producto. El capítulo 16, que abarca las comparaciones planeadas y las *post hoc*, contiene dos pruebas de este último tipo: la prueba DHS de Tukey y la prueba de Newman-Keuls. El capítulo 17 es un capítulo independiente dedicado al análisis de varianza de dos factores y está dirigido a los profesores que deseen exponer este tema con mayor profundidad. Para los maestros que no dispongan del tiempo suficiente para revisar este material, al principio del capítulo 17 he descrito en forma cualitativa la técnica del análisis de varianza de dos factores, haciendo énfasis en los conceptos de efectos principales e interacciones. El capítulo 19 es un capítulo de repaso que reúne todas las pruebas de inferencia y que, a manera de práctica, permite determinar cuál es la prueba que se debe utilizar en el análisis de los datos de diseños experimentales diferentes y con diversas escalas. A los estudiantes les agrada en particular el diagrama de árbol de este capítulo como una ayuda para determinar cuál es la prueba adecuada. Por último, en diversos lugares del texto hay secciones llamadas “¿Cuál es la verdad?”; en esas secciones se presentan a los alumnos algunas aplicaciones prácticas de la estadística.

Es conveniente hacer algunos comentarios sobre la parte de este libro que corresponde a la estadística descriptiva. El material descriptivo ha sido escrito en un nivel que 1) sirve de base para los capítulos acerca de la inferencia y 2) permite a los estudiantes describir adecuadamente los datos para sus propios fines. En su mayoría, el material sobre estadística descriptiva se presenta en el formato tradicional porque así resulta bien. No obstante, el capítulo 1 es una excepción pues en él se analizan métodos para determinar la verdad y se establece el lugar de la estadística como parte del método científico, lo cual es bastante inusual en un libro de texto de estadística.

Cambios en la séptima edición

Material nuevo Como ya se mencionó, gracias a la retroalimentación positiva de los usuarios de la sexta edición, esta séptima edición no tiene grandes cambios. Sin embargo, hay algunas diferencias importantes y muchas otras de menor trascendencia. Los cambios son:

- ◆ **Problemas de final de capítulo con el área a la que pertenecen resaltada visualmente.** Los problemas de final identifican con otra tipografía ocho áreas de la psicología y campos afines (psicología cognitiva, psicología social, psicología del desarrollo, psicología biológica, psicología clínica, psicología industrial/organizacional, psicología de la salud, educación y otros). Esta identificación visual será útil para trabajar en problemas clasificados por áreas y ayudará a que los estudiantes aprecien la amplitud de las aplicaciones de la estadística.
- ◆ **Una nueva sección titulada “Potencia análisis de varianza” (en el capítulo 15).** Esta sección es paralela a la que se titula “Potencia de la prueba t ” (en el capítulo 14). En ella se amplía la cobertura de la potencia, pasando de la prueba de los signos, la prueba z y la prueba t , al análisis de varianza de grupos independientes de un factor o vía.
- ◆ **Preguntas y problemas nuevos o revisados de final de capítulo.** Más de 30 preguntas y problemas nuevos y revisados de final de capítulo han sido añadidos para ampliar y actualizar este material.
- ◆ **La presentación cualitativa del análisis de varianza de dos factores o vías fue transferida al capítulo 17.** Para unificar la presentación del análisis de varianza de dos factores, la presentación cualitativa de éste ha sido trasladada del final del capítulo dedicado al análisis de varianza de un factor al principio del capítulo 17. “Introducción al análisis de varianza de dos factores o vías”.
- ◆ **Tres nuevas secciones de “¿Cuál es la verdad?”.** Se han agregado dos nuevas secciones de “¿Cuál es la verdad?” al capítulo 1, y una tercera sección de “¿Cuál es la verdad?” ha sido añadida al capítulo 10. Se titulan respectivamente “Datos, datos, ¿qué son los datos? 1”, “Datos, datos, ¿qué son los datos? 2” y “¿Azar o efecto real? 2”. La primera de las nuevas secciones ilustra el argumento de que no basta que en un estudio se presenten datos, sino que la conclusión del mismo tiene que ser apropiada para dichos datos. La segunda de las nuevas secciones se refiere también a datos. Sin embargo, esta vez el enfoque no es si “¿existen datos?” o “¿son los datos adecuados para las afirmaciones de veracidad presentadas?”, sino más bien “¿qué tan precisos son los datos?”. La nueva sección final ilustra que el simple hecho de que un resultado sea estadísticamente significativo no implica de manera automática que el resultado corresponda a un efecto real. En este ejemplo, a pesar de que el resultado sea estadísticamente significativo, parece que puede explicarse mejor considerándolo como un efecto del azar.
- ◆ **Cambios de redacción y de formato.** En todo el libro se han introducido cambios menores de redacción y formato para mejorar la calidad y aumentar el interés.

- ◆ **Material de apoyo en la red.** Este material está disponible exclusivamente en inglés y contiene problemas de práctica complementarios totalmente resueltos (más o menos dos por capítulo) para los estudiantes que deseen más práctica guiada. También se han incluido ejemplos ilustrativos y problemas de práctica que en el libro de texto se resuelven por medio de ecuaciones computacionales. Sin embargo, estos problemas se resuelven en el material complementario utilizando ecuaciones conceptuales en lugar de las computacionales; esto será útil para los estudiantes cuando las clases son impartidas por profesores afectos a hacer que sus alumnos trabajen en muchos problemas empleando ecuaciones conceptuales en lugar, o además, de las de tipo computacional. El último conjunto de problemas contenidos en el sitio consiste en problemas específicos para los distintos capítulos, que deben resolverse en la computadora y resultan apropiados para utilizar el software SPSS de la edición para el estudiante. Esto será muy útil en las clases que incluyen un componente de computación SPSS. Por último, el material electrónico de apoyo incluye un capítulo completo sobre la prueba *U* de Mann-Whitney para utilizarse en las clases impartidas por profesores que prefieren un tratamiento profundo de esta prueba, en vez del estudio breve que se presenta en el capítulo 18 del libro de texto.
- ◆ El usuario debe tener en cuenta que, como todo el material electrónico de apoyo en la red, el de este libro está sujeto a cambios, adición o supresión parcial o total sin previo aviso.

Complementos

Este libro cuenta con complementos en inglés para el instructor y sólo se proporcionan a los profesores que adopten la presente obra como texto para sus cursos. Para mayor información, favor de comunicarse a las oficinas de nuestros representantes: clientes@thomsonlearning.com.mx.

Agradecimientos

Muchas personas ayudaron en el desarrollo y la producción de esta edición. En primer lugar, quiero reconocer el excelente trabajo del personal de Wadsworth. Aunque todos merecen mi gratitud, tres individuos destacan en forma especial. En primer lugar y ante todo está Vicki Knight, mi editora. Deseo agradecerle sus normas de trabajo tan elevadas, y el que haya sido tan accesible conmigo, a pesar de su carga de trabajo tan pesada; sus habilidades para el desarrollo y su buena disposición para hacer siempre un esfuerzo adicional a fin de elaborar un producto de alta calidad. A continuación, me gustaría agradecer a Brian Chaffee, gerente de proyecto de publicidad, por el maravilloso trabajo que llevó a cabo al escribir el texto publicitario para esta edición. La tercera persona que merece mi gratitud especial es Darin Derstine, gerente ejecutiva de proyectos de tecnología, quien tuvo a su cargo el aspecto técnico de la producción. Fue un placer trabajar con Darin, quien produjo el material electrónico de apoyo que, en mi opinión, resulta muy atractivo y funcional. Los demás miembros del personal de Wadsworth a quienes deseo expresar mi agradecimiento son: Jennifer Wilkinson, subeditora en jefe; Mónica Sarmiento, asistente editorial; Paul Wells, gerente ejecutivo de proyectos de producción; Vernon Boes, director creativo; Bill Stanton, diseñador de la portada; Beth Zuber, editora a cargo de la gestión de los permisos; Lori Grebe, gerente de mercadotecnia, y Laurel Anderson, asistente de mercadotecnia.

Estoy muy agradecido con Suzanne Kastner, editora administrativo de Graphic World Publishing Services, quien hizo posible la buena marcha de la producción del libro.

Deseo agradecer a los siguientes revisores por sus valiosos comentarios acerca de las ediciones anteriores:

Bryan Auday, Gordon College	James Ha, University of Washington
Michael Biderman, University of Tennessee at Chattanooga	George P. Knight, Arizona State University
Michael Brown, University of Washington	Deana Liddy, University of Nebraska
Dennis Doverspike, University of Akron	Jae Myung, Ohio State University
Thomas Eissenbert, Virginia Commonwealth University	Nona Phillips, University of Washington
Penny Fidler, California State University, Long Beach	Ladonna Rush, College of Wooster
Michael Gardner, California State University, Northridge	Anna Smith, Troy State University
William Gibson, Northern Arizona University	Paul Smith, Alverno College
Roderick Gillis, University of Miami	Eva Szeli, University of Miami
Tim Goldsmith, University of New Mexico	Mary Tallent-Runnels, Texas Tech University
	Evangeline Wheeler, Towson State University
	Todd Zakrajsek, Southern Oregon State College

Estoy especialmente agradecido con los revisores de la séptima edición:

David Bennett, Simmons College	Richard Hagen, Florida State University
James Diefendorff, Louisiana State University	Steve Hall, Embry-Riddle Aeronautical University
Virginia Diehl, Western Illinois University	Le Xuan Hy, Seattle University
Jean Doerr, SUNY Fredonia	Daniel Langmeyer, University of Cincinnati (special thanks)
David Falcone, LaSalle University	Susan Nolan, Seton Hall University
James Ha, University of Washington	Erwin Segal, SUNY Buffalo

Tengo una deuda de gratitud especial con las siguientes personas por la precisión con la que realizaron la revisión de los materiales:

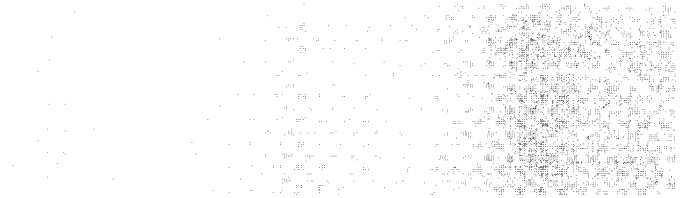
Eli White, estudiante de posgrado, University of Cincinnati, quien leyó el manuscrito completo del libro de texto y resolvió todos los problemas.
 Anthony Simmons, desarrollador del material de apoyo, quien revisó que todo el contenido estuviera correcto.
 Caroline Wood, University of Cincinnati, quien revisó material exclusivo de la edición en inglés.

Agradezco al ejecutor del testamento del difunto Sir Ronald A. Fisher, F.R.S.; y a Longman Group Ltd., Londres, por su autorización para reproducir las tablas III, IV y VII de su libro *Statistical Tables for Biological, Agricultural and Medical Research* (sexta edición, 1974).

El material expuesto en este libro de texto y en el sitio resulta adecuado para los estudiantes de nivel licenciatura en psicología o disciplinas afines de las ciencias del comportamiento. Creo que el enfoque que hemos seguido ayuda en gran medida a impartir esta materia y facilitar su comprensión. Agradeceré todos los comentarios que ayuden a mejorar la calidad de estos materiales.

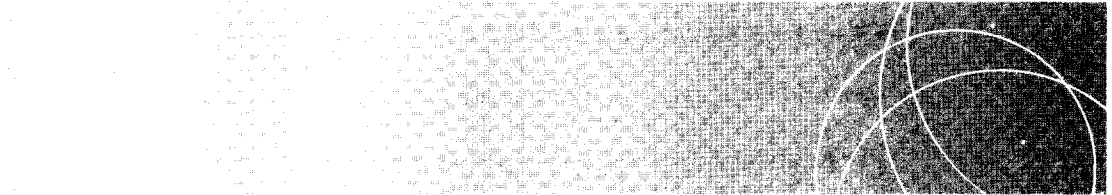
Robert R. Pagano

PANORAMA



1 La estadística y el método científico

La estadística y el método científico



Introducción

Métodos de adquisición del conocimiento

Definiciones

Un ejemplo general: modo de presentación y retención

Investigación científica y estadística

Muestreo aleatorio

La estadística descriptiva e inferencial

Uso de la computadora en estadística

Estadística y el "mundo real"

¿CUÁL ES LA VERDAD? Datos, datos, ¿dónde están los datos?

¿CUÁL ES LA VERDAD? Las autoridades están bien, pero . . .

¿CUÁL ES LA VERDAD? Datos, datos, ¿qué son los datos?-1

¿CUÁL ES LA VERDAD? Datos, datos, ¿qué son los datos?-2

Resumen

Términos importantes

Preguntas y problemas



INTRODUCCIÓN

Se ha preguntado usted alguna vez ¿en qué forma se llega a conocer la verdad? La mayoría de los estudiantes universitarios estaría de acuerdo en que el descubrimiento de lo verdadero, ya sea en el mundo, en nosotros mismos o en los demás, constituye una actividad muy importante. Un poco de reflexión nos revela que dedicamos gran parte del tiempo a esa tarea. Si estudiamos geografía, queremos saber lo que es *verdadero* acerca de las características geográficas de una región en particular. ¿Es la región montañosa o plana, agrícola o industrial? Si nos interesa el estudio de la naturaleza humana, queremos saber lo que es *verdadero* acerca de los humanos. ¿Realmente poseemos una naturaleza espiritual o en *verdad* somos reducibles a simples átomos y moléculas como afirmarían los reduccionistas? ¿Cómo piensan los humanos? ¿Qué ocurre en el cuerpo para

que se produzca una sensación o un movimiento? Cuando siento enojo, ¿es *cierto* que existe un patrón fisiológico subyacente único? ¿Cuál es ese patrón? ¿Ser profesor es mi *verdadero* propósito en la vida? ¿Es *cierto* que los animales piensan? Podríamos continuar citando de manera indefinida este tipo de ejemplos en virtud de que gran parte de la vida la transcurrimos en la búsqueda y el hallazgo de la verdad.



MÉTODOS DE ADQUISICIÓN DEL CONOCIMIENTO

Desde un punto de vista histórico, la humanidad ha empleado cuatro métodos para adquirir conocimiento: la autoridad, el racionalismo, la intuición y el método científico.

Autoridad Cuando se utiliza el método basado en la autoridad, algo se considera verdadero por el solo hecho de que una tradición o alguna persona eminente afirma que lo es. Así, podemos creer en la teoría de la evolución porque nuestros ilustres profesores nos dicen que es verdadera, o bien, creer que Dios existe porque así lo afirman nuestros padres. Aunque este método de conocimiento tiene poca aceptación en la actualidad y a veces conduce a errores, es indispensable en la vida cotidiana. Simplemente aceptamos una gran cantidad de información como cierta con base en una autoridad, aunque sólo sea porque no tenemos ni el tiempo ni la pericia que se requieren para comprobarlo de primera mano. Por ejemplo, yo creo, tomando como base lo que dicen los expertos en física, que los electrones existen, pero nunca he visto ninguno; o, para hablar de algo más cercano, si el jefe de cirugía me explica que el cigarrillo provoca cáncer, dejaré de fumar porque tengo fe en lo que el médico me dice y porque no dispongo ni de los medios ni del tiempo necesarios para investigar el tema por mí mismo.

Racionalismo Este método utiliza exclusivamente el razonamiento para llegar al conocimiento. Supone que si las premisas son válidas y el razonamiento se realiza de manera correcta, según las reglas de la lógica, entonces las conclusiones tienen que ser la verdad. Estamos muy familiarizados con el proceso de razonar porque lo utilizamos mucho. Por ejemplo, consideremos el siguiente silogismo:

· Todos los profesores de estadística son personas interesantes.
 · El señor X es profesor de estadística.
 Por lo tanto, el señor X es una persona interesante.

Si la primera afirmación es cierta (¿eso quién lo dudaría?) y si la segunda afirmación es verdadera, entonces la conclusión también debe ser verdadera. Hablando en serio, es difícil que alguien cuestione la importancia del proceso de razonamiento para encontrar la verdad. Sin embargo, existe un gran número de situaciones en que, por sí solo, el razonamiento es inadecuado para encontrarla.

Por ejemplo, supongamos que usted observa que Juan, un amigo suyo, ha estado deprimido desde hace algunos meses. Como estudiante de psicología, usted sabe que los problemas psicológicos pueden causar depresión. Por lo tanto, es razonable pensar que Juan pueda tener problemas psicológicos que le ocasionan esa depresión. Por otro lado, también sabe que ésta puede ser provocada por una dieta inadecuada. Ésta es otra causa posible del problema. En este caso, existen dos explicaciones lógicas razonables para el fenómeno. Por lo tanto, por sí sola, la razón resulta inadecuada para distinguir entre ellas. Entonces tenemos que recurrir a la experiencia. ¿En verdad es deficiente la dieta de Juan? ¿El mejoramiento de sus hábitos alimenticios corregiría la situación? O bien, ¿tiene en realidad Juan problemas psicológicos serios, y el tratamiento adecuado podría aliviar su depresión? De esta manera comprobamos que la razón pura puede ser suficiente para arrojar luz en algunos casos, pero que es claramente inadecuada en otros.

Como veremos, el método científico también utiliza la razón para llegar a la verdad, pero el razonamiento puro es sólo una parte del proceso. Así, el método científico incorpora la razón, pero no es un sinónimo de ésta.

Intuición El conocimiento se adquiere también a través de la intuición. Con ésta nos referimos a esa inspiración súbita, esa idea que aclara y salta a la conciencia como un todo. No se llega a ella por medio de la razón. Por el contrario, con frecuencia la idea parece surgir después que ha fallado el razonamiento consciente. Beveridge⁶ proporciona numerosos ejemplos tomados de personajes notables. A continuación presentaremos un par de ellos:

He aquí el relato del propio Metchnikoff acerca de cuál fue el origen de la idea de la fagocitosis: "Un día, cuando toda mi familia había ido al circo para ver la notable actuación de unos simios, me quedé solo en casa, con mi microscopio, y me puse a observar la vida en las células móviles de una larva transparente de un pez, cuando una nueva idea brilló súbitamente en mi cerebro. Me pareció de pronto que unas células similares podrían intervenir en la defensa del organismo contra los intrusos. Al sentir que había algo muy interesante en este asunto, me sentí tan excitado que comencé a andar a grandes zancadas por todo el cuarto, de un lado a otro, e incluso fui a la playa para ordenar mis ideas".

Hadamard cita una experiencia del matemático Gauss, quien escribió acerca de un problema que había tratado infructuosamente de resolver a lo largo de varios años: "Finalmente, hace dos días, tuve éxito . . . Como un relámpago, el acertijo empezó a resolverse. Yo no podría decir cuál fue el hilo conductor que me llevó desde lo que sabía con anterioridad hasta aquello que hizo posible el éxito".

Es interesante observar que con frecuencia la idea intuitiva surge después que el razonamiento consciente ha fracasado y cuando el individuo ha dejado de lado el problema durante un tiempo. Así, Beveridge⁷ cita a dos científicos en los siguientes términos:

Despejando de mi mente todas las ideas que tenía acerca del problema, caminé vigorosamente por la calle cuando de pronto, en cierto punto que aún podría localizar, como si viniera del cielo hacia mí, una idea surgió en mi cabeza como si una voz me la dijera a gritos.

Decidí abandonar el trabajo y todas las ideas acerca de él y entonces, al día siguiente, cuando estaba ocupado en una tarea de otra índole, una idea llegó a mi mente como un relámpago, y ésa era la solución . . . su sencillez misma me hizo cavilar por qué no se me había ocurrido pensar en ella antes.

A pesar del hecho de que la intuición ha sido utilizada como fuente de conocimiento durante toda la existencia de la humanidad, sigue siendo un proceso misterioso acerca del cual sólo tenemos una comprensión muy rudimentaria.

Método científico Aunque el método científico utiliza el razonamiento y la intuición para llegar a la verdad, se fundamenta en una evaluación objetiva que lo distingue de los otros métodos. El *experimento científico* se encuentra en el centro de la ciencia. El método de la ciencia es bastante directo. Por algún medio, generalmente un razonamiento deductivo a partir de la teoría existente o una inducción a partir de hechos reales o a través de la intuición, el científico llega a una hipótesis sobre cierta particularidad de la realidad. Entonces, el investigador diseña un experimento para verificar la hipótesis de manera objetiva. Posteriormente, los datos obtenidos como resultado del experimento se analizan en forma estadística y la hipótesis es aceptada o rechazada. La característica más importante de esta metodología es que no importa lo que el científico crea que es verdadero

⁶W. I. B. Beveridge, *The Art of Scientific Investigation*, Vintage Books-Random House, Nueva York, 1957, pp. 94-95.

⁷Ibid., p. 92.

con respecto a la hipótesis en cuestión, pues el experimento proporciona una base para realizar una evaluación *objetiva* de dicha hipótesis. Los datos del experimento obligan a extraer una conclusión acorde con la realidad. Así, la metodología científica trae un “salvavidas integrado” para garantizar que las afirmaciones de veracidad de cualquier tipo, sobre la realidad, se ajusten a la veracidad demostrada con respecto a los fenómenos estudiados, antes que las afirmaciones reciban la categoría de verdades científicas.

Un aspecto importante de esta metodología es que los datos demostrarán si el investigador ha mantenido ideas erróneas o no con respecto al objeto de estudio. Esas ideas pueden contrastarse entonces con los datos y son verificadas de nuevo. La metodología científica, aunque a veces es muy lenta, posee una característica de autocorrección que, a largo plazo, tiene una alta probabilidad de llegar a la verdad. Como en este texto se hace énfasis en el análisis estadístico y no en el diseño experimental, no podemos dedicar mucho tiempo al análisis del diseño de los experimentos. Sin embargo, estudiaremos algunos diseños experimentales, en virtud de que están íntimamente relacionados con el análisis estadístico.



DEFINICIONES

Al analizar éste y otros materiales del libro, utilizaremos ciertos términos técnicos. Presentamos a continuación estos términos y sus definiciones:

- ♦ **Población** *Conjunto completo de individuos, objetos o datos que el investigador está interesado en estudiar.* En un experimento, la población es el grupo más grande de individuos del cual se pueden tomar los sujetos que participarán en dicho experimento.
- ♦ **Muestra** *Subconjunto de la población.* En un experimento, por razones económicas, lo usual es que el investigador reúna los datos acerca de un grupo de sujetos de menor tamaño que la población total. Ese grupo menor es la muestra.
- ♦ **Variable** *Cualquier propiedad o característica de algún evento, objeto o persona, que puede tener diversos valores en diferentes instantes, según las condiciones.* La estatura, el peso, el tiempo de reacción y la dosis de un medicamento son ejemplos de variables. Una variable debe contrastarse con una *constante*, la cual, por supuesto, no asume diversos valores en diferentes instantes. Un ejemplo de una constante es el símbolo matemático π ; siempre tiene el mismo valor (3.14 si se expresa con una precisión de dos cifras decimales).
- ♦ **Variable independiente** *La variable independiente de un experimento es aquella que es controlada en forma sistemática por el investigador.* En la mayor parte de los experimentos, el científico está interesado en determinar el efecto que produce una variable, digamos, la variable *A*, sobre una o más variables. Para esto, el experimentador controla los niveles de la variable *A* y mide el efecto que produce sobre las demás variables. A la variable *A* se le llama *independiente* debido a que sus niveles son controlados por el investigador, sin importar los cambios que puedan ocurrir en las demás variables. Por ejemplo, un científico podría estar interesado en el efecto del alcohol sobre el comportamiento social. Para investigar esto, es probable que el experimentador vaya modificando la cantidad de alcohol que suministra a varias personas y mida sus consecuencias sobre la conducta social de éstas. La cantidad de alcohol suministrada es la variable independiente. En otro experimento se estudia el efecto de la privación del sueño con respecto al comportamiento agresivo. A los sujetos de dicho experimento se les priva de distintas horas de sueño y se observan las consecuencias resultantes en relación con la agresividad. En este caso, se controla la magnitud de la privación del sueño. Por lo tanto, ésta es la variable independiente.

- ◆ **Variable dependiente** *La variable dependiente en un experimento es la que el investigador mide para determinar el efecto de la variable independiente.* Por ejemplo, en el experimento que estudia los efectos del alcohol sobre el comportamiento social, la cantidad de alcohol es la variable independiente. El comportamiento social de los sujetos se mide para ver si resulta afectado por la cantidad de alcohol consumida. Así, el comportamiento social es la variable dependiente. Se le llama *dependiente* porque depende de la cantidad de alcohol consumida. En la investigación sobre la privación del sueño y el comportamiento agresivo, se controla la cantidad de privación del sueño, y se mide el comportamiento agresivo de los sujetos. La magnitud de la privación del sueño es la variable independiente y el comportamiento agresivo es la variable dependiente.
- ◆ **Datos** *Medidas que se realizan sobre los sujetos de un experimento.* Por lo general, los datos consisten en las medidas de la variable dependiente o de otras características del sujeto, como la edad, el sexo, el número de individuos, etcétera. Los datos, en la forma original en que son medidos, se conocen como *datos brutos*, en bruto u *originales*.
- ◆ **Estadística** *Número calculado a partir de los datos de la muestra, que cuantifica una característica de dicha muestra.* Así, el promedio de un conjunto de datos obtenidos a partir de la muestra sería una estadística.
- ◆ **Parámetro** *Número calculado sobre los datos de una población que cuantifica una característica de dicha población.* Por ejemplo, el valor promedio de un conjunto de datos poblacionales se llama parámetro. Conviene observar que una estadística y un parámetro son conceptos muy similares. La única diferencia es que la estadística se calcula sobre una muestra y el parámetro se calcula con respecto a una población.

Un ejemplo general: modo de presentación y retención

Consideraremos ahora un experimento ilustrativo en el cual aplicaremos los términos que hemos analizado anteriormente.

Un profesor lleva a cabo un experimento para determinar si el modo de presentación influye en el grado en que se recuerda un material. Para este experimento, el investigador utiliza varios pasajes en prosa que se presentan en forma visual u oral. Se elige a cincuenta estudiantes entre los asistentes a la universidad en la cual trabaja ese maestro. Los estudiantes se dividen en dos grupos de 25 individuos cada uno. El primer grupo recibe una presentación visual de los pasajes en prosa y el segundo grupo escucha los pasajes en una presentación oral. Al final de sus respectivas presentaciones, se pide a los sujetos que escriban la mayor parte del material que sean capaces de recordar. Se calcula el promedio de palabras recordadas por cada grupo y los promedios son comparados para ver si el modo de presentación tuvo algún efecto.

En este experimento, la variable independiente es el modo de presentación de los pasajes en prosa, es decir, si es oral o visual. La variable dependiente es el número de palabras recordadas. La muestra está conformada por los 50 estudiantes que participaron en el experimento. La población es el grupo total de los individuos de los cuales se tomó la muestra, es decir, todos los estudiantes que asisten a la universidad. Los datos son el número de palabras recordadas por cada estudiante incluido en la muestra. El promedio de palabras recordadas por cada grupo es una estadística, pues mide una característica de los datos de la muestra. Como no se hizo medición de ninguna característica de la población, no se calculó ningún parámetro en este experimento. Sin embargo, para ilustrar el concepto, supongamos que toda la población hubiera presenciado una presentación visual de los pasajes. Si entonces calculáramos el número promedio de pa-

labras recordadas por toda la población, ese resultado sería un parámetro, pues con él se mediría una característica de los datos correspondientes a toda la población.

Resolvamos ahora un problema para practicar la identificación de estos términos.



Especifique usted lo siguiente en el caso del experimento que describiremos a continuación: la variable independiente, la(s) variable(s) dependiente(s), la muestra, la población, los datos, la(s) estadística(s) y el(los) parámetro(s).

Un profesor de ginecología, miembro de una prestigiada escuela de medicina, quiere determinar si un implante de tipo experimental para el control de la natalidad tiene efectos colaterales sobre el peso corporal y la depresión en las mujeres. Un grupo de 5000 mujeres adultas, residentes de una ciudad cercana, se han ofrecido como voluntarias para el experimento. El ginecólogo elige a 100 de esas mujeres para que participen en el estudio. Cincuenta de las voluntarias son asignadas al grupo 1 y las otras 50 al grupo 2, de manera que los datos sobre el peso corporal promedio y la depresión promedio de cada grupo sean iguales al principio del experimento. Las condiciones del tratamiento son las mismas para ambos grupos, excepto porque a las mujeres del grupo 1 se les realiza el implante del dispositivo experimental para el control natal, mientras que las mujeres del grupo 2 reciben un implante placebo. El peso corporal y el nivel de depresión se miden al inicio y al final de la prueba. Para medir la depresión se utiliza un cuestionario estándar; mientras más alto sea el puntaje obtenido en este cuestionario, tanto más deprimida estará la persona. Al final del experimento, el peso corporal promedio y los niveles de depresión promedio de cada grupo se comparan entre sí para determinar si el implante experimental tuvo algún efecto sobre estas variables. Para proteger a las mujeres contra un embarazo no deseado, durante todo el experimento se utilizó también otro método de control natal que no interfería con el implante.

SOLUCIÓN

Variable independiente: el implante experimental para el control natal, versus el placebo.

Variables dependientes: el peso corporal y el nivel de depresión.

Muestra: las 100 mujeres que participaron en el experimento.

Población: las 5000 mujeres que se ofrecieron como voluntarias para el experimento.

Datos: los datos individuales del peso corporal y de los niveles de depresión de las 100 mujeres, al inicio y al final del experimento.

Estadísticas: la media de los pesos corporales del grupo 1 al inicio y al final del experimento, la media de los niveles de depresión del grupo al inicio y al final de la prueba, más las mismas cuatro estadísticas correspondientes al grupo 2.

Parámetro: no se dieron ni se calcularon parámetros para este experimento. Si el ginecólogo hubiera medido los pesos corporales de las 5000 voluntarias al inicio de la prueba, la media de esos 5000 pesos habría sido un parámetro.



INVESTIGACIÓN CIENTÍFICA Y ESTADÍSTICA

La investigación científica se puede clasificar en dos categorías: los estudios observacionales y los experimentos reales. Las técnicas estadísticas son importantes en ambos tipos de investigaciones.

Estudios observacionales En este tipo de investigación, ninguna variable es controlada en forma activa por el investigador. Dentro de esta categoría de investigación figuran: 1) la observación naturalista, 2) la estimación de parámetros y 3) los estudios correlacionales. En la investigación basada en la observación naturalista, uno de los principales objetivos consiste en obtener una descripción precisa de la situación que se estudia. Gran parte de la investigación antropológica y etiológica es de este tipo. La investigación de estimación de parámetros se realiza sobre las muestras para estimar el nivel de una o varias características poblacionales (por ejemplo, el promedio o porcentaje de una población). Las encuestas de opinión pública, los escrutinios y gran parte de la investigación de mercado pertenecen a esta categoría. En la investigación correlacional, el experimentador se centra en dos o más variables para precisar si están relacionadas unas con otras. Por ejemplo, para determinar si la obesidad y la alta presión sanguínea están relacionadas entre sí en los adultos mayores de 30 años, un investigador tendría que medir el nivel de grasa corporal y la presión sanguínea de una muestra de individuos adultos mayores de 30 años. El experimentador analizaría a continuación los resultados para verificar si existe alguna relación entre esas variables; es decir, tendría que comprobar si los individuos con bajos niveles de grasa tienen también baja presión sanguínea, si en los individuos con niveles moderados de grasa la presión sanguínea es moderada, y si aquellos con altos niveles de grasa registran una alta presión sanguínea.

Experimentos reales En este tipo de investigación se intenta determinar si los cambios registrados en una variable producen cambios en otra. Con este tipo de investigación se controla una variable independiente y se estudia el efecto que ésta causa sobre alguna variable dependiente. Sin embargo, puede haber más de una variable independiente y dependiente. En el caso más sencillo sólo existen una variable independiente y otra dependiente. Un ejemplo de este caso es el experimento ya mencionado en el que se intentó investigar el efecto del alcohol sobre el comportamiento social. En esa prueba, como el lector recordará, el nivel de alcohol fue controlado por el experimentador y se midió su efecto sobre el comportamiento social.



MUESTREO ALEATORIO

En todas las investigaciones antes descritas, los datos son recopilados, por lo general, a partir de una muestra de sujetos y no sobre toda la población en la cual se pretende aplicar los resultados. Lo ideal, por supuesto, sería que el experimento se realizara sobre toda la población, pero generalmente eso es demasiado costoso y por eso se considera sólo una muestra. Observe que no cualquier muestra sirve. Ésta deberá ser una muestra de tipo aleatorio. El muestreo *aleatorio* será analizado en el capítulo 8. Por ahora, basta saber que este muestreo permite aplicar las leyes de la probabilidad a los datos y, al mismo tiempo, ayuda a lograr que la muestra sea representativa de la población de la cual fue extraída. Así, los resultados obtenidos a partir de la muestra deberán aplicarse también a la población. Una vez reunidos los datos, éstos se analizan en forma estadística y de ellos se extraen las conclusiones adecuadas acerca de la población.



ESTADÍSTICA DESCRIPTIVA E INFERENCIAL

El análisis estadístico es el tema principal de este libro. Lo hemos dividido en dos áreas: 1) *estadística descriptiva* y 2) *estadística inferencial*. Ambas implican el análisis de datos. Si se realiza un análisis con el fin de describir o caracterizar un conjunto de datos que ha sido reunido, entonces estamos en el área de la estadística descriptiva. Por ejemplo, suponga que su profesor de biología acaba de registrar las calificaciones del grupo. Devuelve los exámenes y ahora quiere describir los resultados obtenidos. Él podría calcular el promedio de la distribución para describir su *tendencia central*. Es probable que también quiera determinar su rango para caracterizar su *variabilidad*. También podría representar las calificaciones en una gráfica para mostrar la *forma* de la distribución. Como todos estos procedimientos tienen la finalidad de describir o caracterizar los datos ya reunidos, corresponden al campo de la estadística descriptiva.

Por otro lado, la estadística inferencial no se refiere a la simple descripción de los datos obtenidos, sino que abarca las técnicas que nos permiten utilizar los datos muestrales para inferir o extraer conclusiones sobre las poblaciones de las cuales fueron obtenidos dichos datos. Ésta es la parte más compleja del análisis estadístico. Implica la probabilidad y varias pruebas de inferencia, como la *prueba t* de Student y el *análisis de varianza*.

Para ilustrar la diferencia entre la estadística descriptiva y la inferencial, suponga que estamos interesados en determinar el CI promedio de todos los estudiantes del primer año de su universidad. Como sería demasiado costoso y tardado medir el CI de cada uno de los alumnos que forman parte de la población, consideraríamos el empleo de una muestra aleatoria de, digamos, 200 estudiantes y les daríamos a cada uno una prueba de CI. Tendríamos entonces 200 calificaciones de CI como *muestra*, y queremos que éstas nos sirvan para determinar el CI promedio de la *población*. Aunque no podemos determinar el valor exacto del promedio de la población, podemos estimarlo utilizando los datos de la muestra junto con una prueba de inferencia llamada la prueba *t* de Student. Los resultados nos permitirían hacer afirmaciones tales como: "Tenemos 95% de confianza de que el intervalo 115-120 contiene al CI promedio de la población". En este caso, no sólo estamos describiendo las calificaciones obtenidas, como en el caso del examen de biología, sino estamos utilizando las calificaciones de la muestra para inferir un valor correspondiente a toda la población. Por lo tanto, estamos en el campo de la estadística inferencial. Así, la *estadística descriptiva* y la *inferencial* se pueden definir de la siguiente manera:

- definiciones**
- **La estadística descriptiva** estudia las técnicas que se utilizan para describir o caracterizar los datos obtenidos.
 - **La estadística inferencial** incluye las técnicas que emplean los datos obtenidos en la muestra para, a partir de ellos, hacer inferencias sobre sus respectivas poblaciones.



USO DE LA COMPUTADORA EN ESTADÍSTICA

El uso de las computadoras en la estadística se incrementó en gran medida durante la última década. De hecho, en la actualidad casi todos los datos de investigación en las ciencias del comportamiento se analizan mediante programas estadísticos de computadora, en lugar de hacerlo "a mano" con una calculadora. Éstas son buenas noticias para los estudiantes, quienes con frecuencia gustan de las ideas, conceptos y resultados de la

estadística, pero detestan los cálculos "manuales". El hecho es que los investigadores también odian los cálculos y, por lo tanto, casi siempre utilizan una computadora para analizar los conjuntos de datos de tamaño considerable. Las computadoras tienen la ventaja de ahorrar tiempo y esfuerzo, minimizar la posibilidad de errores de cálculo, facilitar la presentación gráfica de los datos y proporcionar un mejor manejo de grandes conjuntos de éstos. A pesar de su gran utilidad, es frecuente que un curso básico de estadística no las incluya. Por lo tanto, he escrito esta edición del libro de modo que usted pueda aprender el contenido estadístico con el material para computadora o sin él.

Existen varios programas de computadora para realizar el análisis estadístico. Los más populares son el Statistical Package for the Social Sciences (SPSS), Biomedical Computer Programs-P series (BMDP), Statistical Analysis System (SAS), SYSTAT, MINITAB y Excel. Existen versiones de SPSS, SAS, SYSTAT y MINITAB para mainframes y microcomputadoras. Vale la pena dedicar cierto tiempo para aprender uno o varios de estos programas. Para los estudiantes que desean usar el SPSS con este libro de texto, he incluido ejercicios para computadora en el material adicional, que se encuentra disponible en inglés, en: <http://www.thomsonlearning.com.mx>. Estos ejercicios son apropiados para usarse con la edición de SPSS para el estudiante.*

Cuando usted comience a resolver problemas con computadoras, creo que empezará a experimentar la diversión y el poder de cálculo que puede aportar esta máquina a su estudio de la estadística. De hecho, una vez que haya utilizado una computadora para analizar datos, es probable que se pregunte: "¿Por qué tengo que hacer todos estos cálculos complejos a mano?". Por desgracia, el uso de una computadora para calcular el valor de una estadística no le ayuda a comprenderla. La comprensión de la estadística y su mejor uso se logran a través de la realización de cálculos a mano. Por supuesto, una vez que aprenda todo lo que pueda a partir de los cálculos manuales, el uso de la computadora para obtener valores estadísticos parecerá bastante razonable.



LA ESTADÍSTICA Y EL "MUNDO REAL"

Como ya he mencionado, uno de los principales objetivos de la estadística es ayudar en la evaluación científica de ciertas afirmaciones. Aunque usted puede ver esto como un asunto esotérico y alejado de la vida cotidiana, creo que al terminar de leer este libro estará convencido de que la comprensión de la estadística tiene importantes aspectos prácticos para su éxito personal en la vida. A medida que lea este texto, espero que sea cada vez más consciente de la frecuencia con la que somos bombardeados en la vida diaria por "autoridades" que nos dicen, con base en "afirmaciones ciertas", lo que debemos hacer, cómo debemos vivir, lo que debemos comprar, lo que debemos valorar, etc. En áreas de verdadera importancia para usted, espero que comience a preguntarse cosas como: "¿Están respaldadas estas afirmaciones por datos?", "¿qué tan confiables son esos datos?", "¿es el azar una explicación razonable para los datos?" Si no existen datos, o si éstos se presentan con expresiones como "mi experiencia es que . . ." en lugar de haber sido obtenidos por medio de experimentos bien controlados, usted comenzará a cuestionarse si debe considerar en serio el consejo de esa autoridad.

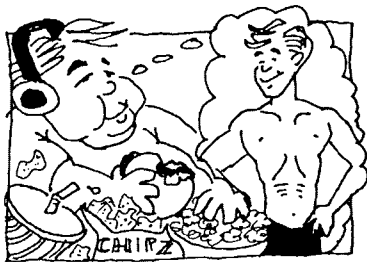
Para ayudarle a desarrollar el aspecto de la toma de decisiones mediante la estadística, al final de ciertos capítulos he incluido aplicaciones de la vida cotidiana con el título "¿Cuál es la verdad?". Para comenzar, consideremos el siguiente material.

*Para lo referente al disco que contiene esta versión de SPSS y un manual acompañante, véase: L. A. Kirkpatrick y B. C. Feeney, *A Simple Guide to SPSS for Windows: For Versions 8.0, 9.0, 10.0, & 11.0, Revised Edition*, Wadsworth, Belmont, CA, 2003.



Datos, datos, ¿dónde están los datos?

El anuncio que aquí reproducimos apareció en un número reciente de *Psychology Today*. Desde un punto de vista científico, ¿qué es lo que le falta?



Respuesta Este anuncio es similar a muchos otros que han proliferado en estos días. Prometen mucho, pero no ofrecen datos experimentales como respaldo de sus afirmaciones. El anuncio hace, de entrada, una afirmación de veracidad: **“Piense y adelgace”**. Después afirma: **“He aquí un programa grabado en cinta que en realidad funciona . . . ¡y de manera permanente!”**. El programa consiste en escuchar una cinta con un mensaje subliminal que supuestamente “programa” su

GRATIS Libro para aprender durmido
Valor \$3.95

HIPNOSIS

GRATIS Libro para aprender durmido
Valor \$3.95

Piense y adelgace

SUBLIMINAL

ESTÉ PRESENTE EN CADA PULSAMENTO DE LA CINTA PARA OBTENER EL MÁXIMO EFECTO. He aquí un programa grabado en cinta que en realidad funciona . . . ¡y de manera permanente! El programa de cintas Piense y Adelgace (programa H-1) ha sido diseñado por Robert Mazur, uno de los principales hipnotistas de Estados Unidos.

Más de 10 millones y 400 libras de personas han sido capaces de perder peso sin dieta, sin ejercicio físico y sin otros métodos de pérdida de peso. ¡Disfrute de un programa que le ayudará a adelgazar y a sentirse mejor! ¡Pídale nuestro folleto-catalogo de 64 páginas “Usted es lo que cree que es”. Adjunte el folleto “Piense y Adelgace” y el programa grabado en cinta. ¡Folleto-catalogo de 64 páginas GRATIS.

PARA OBTENER EL MÁXIMO EFECTO EN SU PROGRAMA PIENSE Y ADELGACE, PÍDELE EL PRIMER PÍLSULA.

Adjunto mis \$10.95. Por favor envíeme mi cinta “Piense y Adelgace”, e incluyan mi folleto-catalogo de 64 páginas GRATIS.

Escríbame al presidente de Health Travelers, 5017 So. Division Ave., Grand Rapids, MI 49508

Nombre _____
 Dirección _____
 Ciudad, Estado, Código Postal _____
 Teléfono (_____) _____
 Correo electrónico _____
 Fecha _____

*Advertencia: Este programa puede ser utilizado para otros propósitos que no sean los que se describen en este folleto.

¡Crea en un nuevo usted!

mente para que usted adelgace. Lo que falta de manera evidente son experimentos controlados y datos que apoyen tal afirmación. Éste es el tipo de afirmación que requiere desesperadamente de una verificación empírica. En apariencia, los autores de este anuncio publicitario no creen que

los lectores de *Psychology Today* sean muy sofisticados, desde el punto de vista de sus conocimientos de la estadística. Por cierto, espero que los lectores de este libro pidan esos datos antes de pasar seis meses de su tiempo oyendo esta cinta, ¡cuyo mensaje ni siquiera pueden escuchar!



Las autoridades están bien, pero . . .

Un anuncio que promueve el Anacin-3 apareció en un número reciente de *Cosmopolitan*. El título de la publicidad era "3 buenas razones para probar Anacin-3". El anuncio tenía la imagen de un doctor, una enfermera y un farmacéutico que afirmaban lo siguiente:

1. "Los *médicos* recomiendan acetaminofén, el analgésico sin aspirina de Anacin-3, más que

cualquier otro analgésico sin aspirina".

2. "Los *hospitales* utilizan acetaminofén, el analgésico sin aspirina de Anacin-3, más que cualquier otro analgésico sin aspirina".
3. "Los *farmacéuticos* recomiendan acetaminofén, el analgésico sin aspirina de Anacin-3, más que cualquier otro analgésico sin aspirina".

Desde un punto de vista científico, ¿falta algo?

Respuesta Este anuncio es un poco mejor que el anterior. Al menos se menciona cierta autoridad en

apoyo del producto. Sin embargo, el anuncio es incorrecto y tampoco presenta los datos adecuados. Una cuarta razón mucho mejor que las "3 buenas razones para probar Anacin-3" expuestas en el anuncio sería la presentación de datos de experimentos bien realizados donde se demostrara que: a) acetaminofén es un mejor analgésico que cualquier otro sin aspirina y b) que Anacin-3 alivia el dolor mejor que sus competidores. ¿Tiene alguna idea de la razón por la cual no se presentaron estos datos? Como estudiante de estadística, ¿está usted satisfecho con la argumentación presentada en el anuncio?

2 BUENAS RAZONES PARA PROBAR **Pyrd-Z**[®]



1. Los *médicos* recomiendan **Pyrd-Z** para aliviar el dolor, con preferencia sobre otros productos.



2. Los *hospitales* utilizan **Pyrd-Z**, el analgésico, más que ningún otro.



Datos, datos, ¿qué son los datos?-1

En las secciones anteriores de “¿Cuál es la verdad?” no se presentaron datos para justificar las declaraciones de los autores, lo cual es un error capital en ciencia. Un poco menos grave, pero también cuestionable, es el hecho de que si bien se presentan datos en un estudio o experimento, las conclusiones extraídas por los autores parecen estar muy alejadas de los datos reales. En un estudio publicado recientemente en un periódico se ilustra este argumento. Dicho artículo se presenta a continuación.

Los cavernícolas eran personas muy agradables

Por VIRGINIA FENTON

The Associated Press

PARÍS— Es probable que los neandertales no hayan sido tan salvajes como pensamos. El hueso de un maxilar de 200 000 años de antigüedad, descubierto en Francia, sugiere que los homínidos primitivos se atendían muy bien unos a otros, en este caso, por el hecho de que alimentaron a un camarada que carecía de dientes, afirmó el viernes un equipo de expertos internacionales.

Un maxilar estropeado, descubierto el año pasado en el sur de Francia, muestra que su poseedor sobrevivió sin dientes durante varios años, lo cual habría sido imposible de no haber contado con la ayuda de sus congéneres, afirmó el paleontólogo canadiense Serge Lebel.

Lebel dirigió un equipo internacional de expertos que descubrió el fósil en julio de 2000. En el equipo figuraba también el notable especialista del periodo Neandertal, Erik Trinkaus, de la Universidad de Washington en St. Louis, y varios colegas de Alemania, Portugal y Francia.

“Sin duda, ese individuo estaba bastante débil, necesitaba que alguien preparara su alimento, y es probable que el grupo social se haya hecho cargo de él o ella”, declaró Lebel en una conferencia de prensa.

“No debemos deshumanizar a esos seres. Ellos muestran una forma de comportamiento totalmente humana”, dijo Lebel, quien trabaja en la Universidad de Quebec.

Tal vez otros miembros del grupo llegaron al extremo de masticar el alimento antes que lo ingiriera su compañero enfermo, además de cortarlo y cocinarlo para él.

Con este descubrimiento, es posible que nuestras ideas sobre la fecha en que la sociedad empezó a atenderse en forma recíproca se atrasen 150 000 años más, explicó

Lebel. Una infección similar por la cual un homínido perdió los dientes sólo se había encontrado con anterioridad en fósiles que datan de hace 50 000 años, aseveró.

Los hallazgos del equipo fueron publicados el 25 de septiembre en el periódico estadounidense Proceedings of the National Academy of Sciences.

Sin embargo, el antropólogo de la Universidad de Pittsburgh, Jeffrey Schwartz, se mostró escéptico.

“La gente puede comer muchas cosas sin dientes. No hay razón para suponer que el individuo no pudo haber masticado alimentos blandos: caracoles, moluscos e incluso gusanos.”

Reproducido con autorización de The Associated Press.

¿Cree usted que el hallazgo de un maxilar dañado, de hace 200 000 años, realizado por ese equipo internacional de expertos sea una base de datos adecuada para extraer las conclusiones expuestas por el Dr. Lebel? O bien, ¿es usted escéptico como el Dr. Schwartz? Es probable que usted adivine a cuál de esos bandos pertenezco yo.





Datos, datos, ¿qué son los datos?-2

Como hemos dicho en este capítulo, el uso de los datos como base para hacer afirmaciones de veracidad es un aspecto importante del proceso científico. Sin embargo, a veces los datos están consignados en forma distorsionada y eso conduce a conclusiones falsas, no a la verdad. Este argumento se ilustra en el siguiente artículo que apareció recientemente en un periódico.

Los expertos retuercen los hechos y los convierten en leyendas urbanas

Por TIM RUTTEN

Las leyendas urbanas son anécdotas tan atractivas y entretenidas, tan afirmadoras de las ansiedades o confirmadoras de los prejuicios, que la gente las repite como verdaderas aunque no lo sean.

Su equivalente periodístico es el sustancioso "resultado de encuestas" o "hallazgo de un estudio" que circunscribe con tanta nitidez un problema social o un argumento de polémica partidista, que ningún presentador de programas de opinión u orador político es capaz de resistir.

Ronald Reagan, por ejemplo, se hizo tan notorio por su afición a las anécdotas de dudosa veracidad, que algunos de sus colaboradores cercanos especularon que sus años de actor hicieron que fuera incapaz de dejar pasar la ocasión de hacer un comentario ingenioso.

Esa actitud de hacer caso omiso de la veracidad de las fuentes, difícilmente concluyó con el retiro del Gran Comunicador.

El analista político de CNN William Schneider, una autoridad en

encuestas de opinión, considera que las cuestiones o las campañas políticas muy polarizadas son especialmente propensas a producir resultados de estudios basados en la ficción, que más tarde adquieren vida propia.

"La seguridad social estadounidense es un caso clásico", dijo Schneider. "Según una leyenda política, en una encuesta 'reciente' que no se especifica, se demostró que son más los jóvenes estadounidenses que creen que Elvis está vivo que los que esperan cobrar un cheque de Seguridad Social cuando se retiren. En los últimos años, todos los conservadores de la Tierra y todos los políticos republicanos de Washington, entre ellos el presidente Bush, se han referido a este hallazgo. Todos dicen sencillamente: 'Las encuestas o los estudios muestran esto', pero nada se ha demostrado. Los republicanos no dejan de citar este ficticio hallazgo porque —sea cierto o no— tiene un gran valor como moneda política."

Hace poco pudimos presenciar el nacimiento de otra leyenda periodística de ese tipo. La página de "ocio y arte" del Periódico Wall Street publicó un artículo extractado de un discurso que pronunció la semana anterior el presidente del National Endowment for the Humanities, Bruce Cole. En él,

lamentó el hecho de "que los estadounidenses no conocen su historia" y reflexionó acerca de los peligros específicos que implica esta amnesia colectiva en la secuela del 11 de septiembre. Además, expresó el temor de que hasta ese recuerdo se esté desvaneciendo.

Todos son argumentos válidos. Sin embargo, en apoyo de su declaración, Cole citó "una encuesta nacional comisionada recientemente por la Escuela de Derecho de Columbia", que, según dijo, "reveló que casi dos terceras partes de los estadounidenses piensan que el dogma socialista de Karl Marx, 'a cada cual según su capacidad, a cada cual según sus necesidades', fue escrito o pudo haber sido escrito por los Padres Fundadores de este país y que figuraba en la Constitución".

Claro; útil; alarmante. Incluso plausible, dado el gusto popular por la banalidad altruista. En otras palabras, todo lo que un buen argumento retórico debe ser.

La cuestión es que eso no fue lo que la encuesta de Columbia reveló.

Sin embargo, eso fue lo que se afirmó en el comunicado de prensa del 29 de mayo, donde se anunciaron los hallazgos de la encuesta. De hecho, su primer párrafo dice así: "Casi dos tercios de los estadounidenses creen que la máxima de Karl Marx, 'A cada cual



(continúa)



Datos, datos, ¿qué son los datos?-2

(continuación)

según su capacidad, a cada cual según sus necesidades', fue o pudo haber sido escrita por los fundadores del país y está incluida en la Constitución".

Observe el cambio sutil entre lo que dice el boletín, "pudo haber sido escrita" y la versión de Cole —o de los autores de su discurso— "tal vez fue escrita".

El segundo párrafo del anuncio cita al profesor de la Escuela de Derecho de Columbia, Michael Dorf, cuando menciona las problemáticas consecuencias del "hallazgo" sobre Marx y remite al lector a una columna escrita por el profesor en un sitio web de carácter jurídico, *Findlaw.com*. En esa columna, Dorf escribió que "según los hallazgos de la encuesta, el 69 por ciento de los encuestados dijeron que, o bien pensaban que la Constitución de los Estados Unidos contenía la máxima de Marx, o no sabían si la contenía o no".

Hay aún más en la descripción de Cole sobre los hallazgos.

¿Cuál fue la encuesta real, en la que a 1000 personas de una muestra nacional se les hicieron cinco preguntas sobre la Constitución, a las que debían responder "sí o no"?

En respuesta a la pregunta de si la máxima marxista está incluida en el documento, el 35 por ciento dijo que sí; el 31 por ciento dijo que no; y el 34 por ciento dijo "que no lo sabía".

¿Tenía derecho Dorf a unir a quienes respondieron "sí" y "no sé" en un solo y alarmante bloque de dos tercios? No. Schneider dijo: "Eso es una interpretación totalmente errónea de lo que la gente quiere

decir cuando le responde que no sabe a un encuestador. Cuando dice eso, la gente quiere decir 'No tengo la menor idea'. Optar por interpretar esto como: 'Creo que tal vez es cierto o que podría serlo' es un error de interpretación por parte del autor".

El efecto que esto produce es añadir mordacidad a un hallazgo que Schneider describió como "sin ningún rasgo notable. Nada tiene de sorprendente que cerca de un tercio de los encuestados considerara que [la máxima] formaba parte de la Constitución. Por ejemplo, la mayoría de la gente cree que el derecho a la privacidad figura en la Constitución, siendo que no hay en ella mención alguna a tal derecho".

De hecho, lo que resulta notable —y mucho más relevante por el momento— acerca de la encuesta de Columbia es que nos muestra lo que los estadounidenses saben acerca de su Constitución: 83 por ciento de los encuestados reconocieron la primera frase de la 14a Enmienda como parte de la carta magna nacional, mientras que el 60 por ciento de ellos entendieron correctamente que el presidente no puede suspender la Carta de Derechos en tiempo de guerra.

Más de dos tercios de los interrogados tenían conocimiento de que los cargos de los jueces de la Corte Suprema son vitalicios.

La otra pregunta que los entrevistados no supieron responder fue acerca de si el rechazo del veredicto del tribunal en el caso *Roe* contra *Wade* haría que el aborto fuera "ilegal en todo el territorio de Estados Unidos".

Menos de un tercio de los encuestados respondieron con acierto que no sería así, pues los estatutos de los estados que garantizan la libertad de elección seguirían estando vigentes.

A pesar de estos hechos, se puede decir con seguridad que los comentaristas de radio y televisión, los editorialistas y todo el resto de los que suelen desconfiar, pronto estarán suspirando, lamentándose, mofándose y enfureciéndose por la incapacidad de los estadounidenses para distinguir entre el pensamiento de James Madison y el de Karl Marx. ¿Cuánto tiempo pasará antes que se culpe de esto a la educación pública, a los medios liberales, al humanismo y a las familias de un solo progenitor?

Las páginas de comentarios son el vientre blando del periodismo estadounidense. Sus escritores, a pesar de que defienden sus propios intereses, son considerados bajo una norma de comprobación de veracidad diferente, lo cual significa más baja, en virtud de su supuesta pericia.

En realidad, ellos son los causantes de que se inyecte con regularidad información falsa de este tipo en el discurso público. En el mercado de las ideas, lo mismo que en el lote de automóviles usados, la mejor política sigue siendo tener precaución.

The Los Angeles Times, 14 de junio de 2002. Reproducido con autorización de Tribune Media Services.

¿Considera que William Schneider es demasiado escéptico o concuerda usted con la conclusión del artículo, es decir, que "en realidad, ellos son los causantes de que se inyecte con regularidad información falsa de este tipo en nuestro discurso público. En el mercado de las ideas, igual que en el lote de automóviles usados, la mejor política sigue siendo tener precaución"?

RESUMEN

En este capítulo analizamos la forma en que se establece la verdad. Tradicionalmente se han utilizado cuatro métodos: la autoridad, la razón, la intuición y la ciencia. El experimento científico es el elemento central de la ciencia. Mediante un razonamiento o por su intuición, el investigador crea una hipótesis acerca de cierta particularidad de la realidad. A continuación, el científico diseña un experimento a fin de probar de manera objetiva la hipótesis. Los datos del experimento se analizan estadísticamente y la hipótesis es confirmada o rechazada.

La mayor parte de la investigación científica cabe en alguna de estas dos categorías: los estudios observacionales y los experimentos reales. La observación natural, la estimación de parámetros y los estudios correlacionales se incluyen dentro de la categoría de observación. Su objetivo principal es brindar una descripción precisa de la situación, estimar los parámetros de la población, o determinar si dos o más de las variables están relacionadas entre sí. Como no existe un control sistemático de alguna variable por el experi-

mentador al realizar un estudio de observación, este tipo de investigación no permite determinar si los cambios en una variable producirán cambios en otra. Ese tipo de resultado sólo se puede determinar a partir de experimentos reales.

En los experimentos reales, el investigador controla en forma sistemática la variable independiente y observa su efecto sobre una o más variables dependientes. Debido a cuestiones prácticas, los datos se reúnen sólo sobre una muestra de sujetos y no sobre toda la población. Es importante que la muestra sea aleatoria. Los datos obtenidos se analizan después de manera estadística.

El análisis estadístico puede ser descriptivo o inferencial. Si el análisis sólo describe o caracteriza los datos obtenidos está en el campo de la estadística descriptiva. Si el análisis usa los datos obtenidos para hacer inferencias sobre las poblaciones, está en el campo de la estadística inferencial. El conocimiento del análisis estadístico tiene importantes consecuencias prácticas en la vida diaria.

TÉRMINOS IMPORTANTES

Constante (página 6)	Investigación observacional (página 9)	Muestra (página 6)
Datos (página 7)	Investigación para la estimación de parámetros (página 9)	Parámetro (página 7)
Estadística (página 7)	Investigación por observación natural (página 9)	Población (página 6)
Estadística descriptiva (página 10)	Método científico (página 5)	Racionalismo (página 4)
Estadística inferencial (página 10)	Método de autoridad (página 4)	SPSS (página 11)
Experimentos reales (página 9)		Variable (página 6)
Intuición (página 5)		Variable dependiente (página 7)
Investigación de correlación (página 9)		Variable independiente (página 6)

PREGUNTAS Y PROBLEMAS

Nota para el estudiante: usted observará que al final de algunos problemas de éste y otros capítulos, excepto el 2, identificará por medio de color un área específica de la psicología y otros campos afines en los que el problema tiene aplicación. Por ejemplo, el problema 6 parte **b**, página 18, es un problema del área de la psicología biológica. Agregué el letrero “biológica” al final del problema, omitiendo el término “psicología” en aras de la brevedad. Las áreas específicas identificadas son psicología cognitiva, psicología social, psicología del desarrollo, psicología biológica, psicología clínica, psicología industrial/organizacional (I/O), psicología de la salud, educación y otras. Como ya afirmé, en el letrero que utilicé omití el término “psicología” en aras de la brevedad. Espero que estos letreros sean útiles para que su profesor seleccione los problemas para realizar en casa y que le permitan a usted apre-

ciar la amplitud de las aplicaciones de este material, además de ayudarlo a identificar los problemas que tal vez le agradaría resolver, aparte de los que le hayan sido asignados.

- Defina cada uno de los siguientes términos:

Población	Variable dependiente
Muestra	Constante
Datos	Estadística
Variable	Parámetro
Variable independiente	
- ¿Cuáles son los cuatro métodos para adquirir el conocimiento? Escriba un párrafo breve que explique los aspectos esenciales de cada uno de ellos.
- ¿En qué difiere el método científico de cada uno de los siguientes métodos?
 - Método de autoridad.
 - Método de racionalismo.
 - Método de intuición.

4. Escriba un párrafo breve en el cual compare la observación naturalista con los experimentos reales.
5. Distinga entre la estadística descriptiva y la inferencial. Proponga ejemplos para ilustrar los puntos establecidos.
6. En cada uno de los siguientes experimentos, especifique: 1) la variable independiente, 2) la variable dependiente, 3) la muestra, 4) la población, 5) los datos y 6) la estadística:
 - a. Un psicólogo de la salud está interesado en saber si la motivación del miedo es eficaz para reducir la incidencia del consumo de cigarrillos. Se eligen 40 adultos fumadores entre los individuos que viven en la ciudad donde trabaja el psicólogo. A 20 de ellos se les pide que fumen un cigarrillo, después de lo cual presencian una película cruel acerca de la forma en que el hábito de fumar provoca cáncer. Se muestran imágenes cruentas de pulmones y otros órganos enfermos, de fumadores fallecidos, en un esfuerzo por inducir el miedo a fumar en esos sujetos. El otro grupo recibe el mismo tratamiento, excepto que ven una película neutral, no relacionada con el cigarrillo. Durante los dos meses posteriores a la exhibición de las cintas, el experimentador lleva un registro acerca de la cantidad de cigarrillos que fuman diariamente los participantes. Se calcula entonces, para cada grupo, la media de cigarrillos fumados diariamente a partir de la exhibición de la película, y estas medias son comparadas para determinar si el filme inductor del miedo tuvo algún efecto sobre los fumadores. *salud*
 - b. Un fisiólogo quiere saber si una región particular del cerebro (el hipotálamo) está implicada en la regulación del consumo de comida. Se realiza un experimento en el cual se eligen 30 ratas de un criadero universitario, las cuales se separan en dos grupos. A uno de estos grupos se le causan lesiones en el hipotálamo, mientras que el otro grupo es lesionado en un área neutral. Después de recuperarse de las operaciones, todos los animales tienen acceso libre al alimento durante dos semanas y se lleva un registro del consumo diario de cada animal. Al final del periodo de dos semanas se determina la media de consumo diario de alimento para cada grupo. Por último, estas medias son comparadas para ver si las lesiones en el hipotálamo han afectado la cantidad de alimento consumido. *biológica*
 - c. Un psicólogo clínico está interesado en evaluar tres métodos para tratar la depresión: la medicación, la reestructuración cognitiva y el ejercicio.

- Se incluye una cuarta condición de tratamiento, un grupo que sólo estará en espera de tratamiento, para tener un grupo de control de línea base. Se recluta a 60 personas deprimidas entre los estudiantes de licenciatura de una universidad estatal grande, y 15 de ellos son asignados a cada método de tratamiento. Los tratamientos se administran durante seis meses, después de lo cual cada estudiante recibe un cuestionario diseñado para medir el grado de depresión. El cuestionario tiene una escala de 0 a 100, donde los puntajes mayores indican un mayor grado de depresión. Luego se calculan las medias de los valores de depresión para los cuatro tratamientos y se comparan entre sí para determinar la eficacia relativa de cada tratamiento. *clínica, salud*
- d. Un psicólogo social está interesado en determinar si los individuos que terminan la enseñanza media superior y no reciben más educación ganan más que los desertores de dicho nivel de estudios. Entonces se lleva a cabo una encuesta nacional, en una ciudad grande del oeste medio, se obtiene una muestra de 100 individuos de cada categoría y se les pregunta cuál es el salario anual de cada uno de ellos. Los resultados se organizan en una tabla y se calcula la media de los valores de los salarios para cada grupo. *social*
 - e. Un psicólogo cognitivo se interesa en saber cómo resulta afectada la retención con el espaciamiento de las sesiones de práctica. Una muestra de 30 alumnos de séptimo grado es seleccionada, en una escuela de educación media inferior de la localidad, y se divide en tres grupos de 10 estudiantes cada uno. A todos los estudiantes se les pide que memoricen una lista de 15 palabras y se les imparten tres sesiones de práctica, de 5 minutos cada una, para que lo logren. Las sesiones de práctica de los sujetos del grupo 1 tienen un espaciamiento de 10 minutos; las del grupo 2 se espacian 20 minutos; y las del grupo 3 tienen 30 minutos de espaciamiento. A todos los grupos se les somete a una prueba de retención una hora después de la última sesión de práctica. Los resultados son registrados como el número de palabras recordadas correctamente en el periodo de prueba. Se calculan y comparan las medias de los valores de cada grupo. *cognitiva*
 - f. Una psicóloga deportiva utiliza la visualización para promover un más alto rendimiento en los atletas universitarios. Está interesada en evaluar la eficacia relativa de la visualización por sí sola, contra la visualización acompañada de una charla apropiada. Se realiza un experimento

con un equipo de baloncesto estudiantil. Diez miembros del equipo son seleccionados. Cinco de ellos son asignados a un grupo que sólo recibe la visualización, y otros cinco se asignan a un grupo de visualización más charla. Las dos técnicas han sido diseñadas para aumentar la precisión en los tiros libres. Cada grupo practica su técnica durante un mes. La precisión de cada jugador en los tiros libres es medida previamente y luego un mes después de que empiezan a practicar esa técnica. La diferencia en los puntos se calcula para cada jugador y las medias de las diferencias de puntaje de cada grupo son comparados para determinar la efectividad relativa de las dos técnicas. 1/0, otra

g. Un profesor de mecanografía piensa que el empleo de un orden diferente en las teclas de una máquina facilitará una escritura más rápida. Se elige a 20 estudiantes de secretariado en una escuela de comercio de gran tamaño para participar en un experimento diseñado para probar esta suposición. Diez de estas estudiantes aprenden a escribir con el teclado convencional. Las otras diez reciben su entrenamiento con el nuevo teclado. Al final del periodo de entrenamiento se mide la velocidad de escritura de cada estudiante, en palabras por minuto. Luego se calcula la media de velocidad de escritura para ambos grupos y se comparan las medias para determinar si el nuevo teclado ha tenido algún efecto. educación

7. Indique cuáles de las siguientes afirmaciones representan una variable y cuáles una constante:

- El número de letras del alfabeto.
- El número de horas que tiene un día.
- La hora en que usted come.
- El número de estudiantes que especializan en psicología cada año, en la universidad donde estudia usted.
- El número de centímetros en un metro.
- La cantidad de horas que duerme usted cada noche.
- Su peso.
- El volumen de un litro.

8. Indique cuáles de las siguientes situaciones corresponden a la estadística descriptiva y cuáles a la estadística inferencial:

- Un informe anual para accionistas que detalla los bienes de la corporación.
- Un profesor de historia que anuncia a su grupo el número de estudiantes que obtuvieron la máxima calificación en un examen reciente.

- El cálculo de la media de un conjunto de datos de una muestra para caracterizarla.
- El uso de los datos de una muestra en una encuesta para estimar la opinión de la población.
- Realizar un estudio de correlación sobre una muestra para determinar si el nivel educativo y el ingreso de la población están relacionados.
- Un artículo de un periódico informa sobre los salarios promedio de los empleados federales a partir de los datos reunidos de todos los interesados.

9. Para cada una de las siguientes situaciones, identifique los datos de la muestra y los datos de la población:

- Una psicóloga social interesada en el comportamiento de los bebedores investiga la cantidad de bebidas que se sirven en los bares de una ciudad en particular, cierto viernes, durante la "hora feliz". En la ciudad existen 213 bares. Como son demasiados para revisarlos, ella elige 20 y registra el número de bebidas que se sirven en ellos. Los datos son los siguientes:

50	82	47	65
40	76	61	72
35	43	65	76
63	66	83	82
57	72	71	58

social

- Para que un restaurante especializado en la venta de hamburguesas de 1/4 de libra de bajo costo pueda obtener ganancias, es necesario que cada hamburguesa servida tenga un peso cercano a 0.25 de libra. Sabiendo esto, el gerente del restaurante está interesado en la variabilidad de los pesos de las hamburguesas servidas cada día. Durante cierto día, se han servido 450 hamburguesas. Como tardaría mucho en pesarlas todas, el gerente decide pesar sólo 15. Así se obtuvieron los siguientes pesos, en libras:

0.25	0.27	0.25
0.26	0.35	0.27
0.22	0.32	0.38
0.29	0.22	0.28
0.27	0.40	0.31

otro

- Se sospecha que una máquina que corta piezas pequeñas de acero (utilizada para fabricar tor-

nillos) a una determinada longitud no es confiable. El capataz del taller decide entonces verificar la producción del artefacto. El día de la verificación, la máquina debe producir piezas de 2 centímetros. La tolerancia aceptable es de ± 0.05 centímetros. Como tardaría mucho en medir las 600 piezas producidas en un día, elige un grupo representativo de 25. Así obtiene las siguientes longitudes, en centímetros:

2.01	1.99	2.05	1.94	2.05
2.01	2.02	2.04	1.93	1.95
2.03	1.97	2.00	1.98	1.96
2.05	1.96	2.00	2.01	1.99
1.98	1.95	1.97	2.04	2.02

I/O

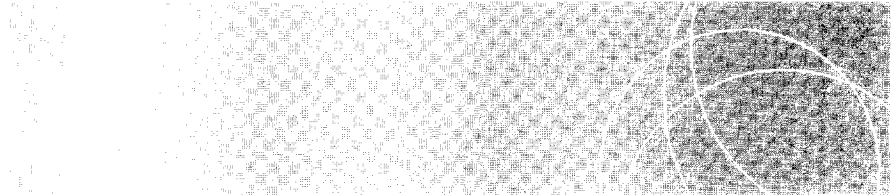
d. A una psicóloga fisiológica que trabaja en la Universidad de Tacoma le interesa conocer los ritmos cardíacos diastólicos, en reposo, de todas las estudiantes de dicha universidad. Para eso toma muestras aleatorias de 30 mujeres jóvenes del cuerpo estudiantil y registra los siguientes ritmos cardíacos diastólicos cuando las estudiantes están recostadas en una cama. Las mediciones se registran en latidos/minuto.

62	85	92	85	88	71
73	82	84	89	93	75
81	72	97	78	90	87
78	74	61	66	83	68
67	83	75	70	86	72

biológica

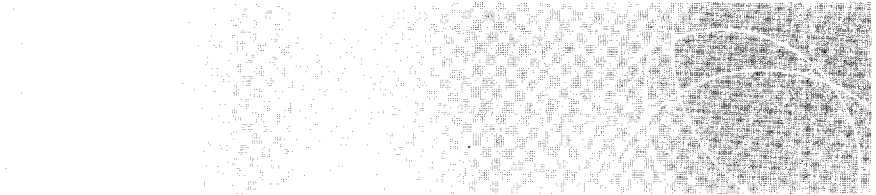
Si desea encontrar problemas de práctica adicionales totalmente resueltos sobre los temas que acaba de estudiar, busque Chapter 1 en el material adicional del libro, disponible en inglés en: <http://www.thomsonlearning.com.mx>.

ESTADÍSTICA DESCRIPTIVA



- 2** Conceptos básicos de matemáticas y de medición
- 3** Distribuciones de frecuencias
- 4** Medidas de tendencia central y variabilidad
- 5** La curva normal y los puntajes estándar
- 6** Correlación
- 7** Regresión lineal

Conceptos básicos de matemáticas y de medición



Sugerencias de estudio para el estudiante
Notación matemática
Sumatoria
Escala de medición: nominales, ordinales, de intervalos y de razón
Escala de medición en las ciencias del comportamiento
Variables continuas y discretas
Resumen
Términos importantes
Preguntas y problemas
Notas



SUGERENCIAS DE ESTUDIO PARA EL ESTUDIANTE

La estadística no es un tema sencillo; requiere el aprendizaje de conceptos difíciles e implica también la realización de cálculos matemáticos. Sin embargo, me gustaría darle a usted un consejo que, en mi opinión, le será de gran utilidad para comprender este material. Dicho consejo se basa en los muchos años que he enseñado esta materia: espero que lo tome muy en serio.

La mayoría de los estudiantes de ciencias del comportamiento sienten mucha ansiedad al emprender un curso de matemáticas o de estadística. Sin minimizar la dificultad del tema, no es necesario sentir tanta ansiedad. Para aprender el material contenido en este libro, usted no tiene que ser un “mago” del cálculo o de las ecuaciones diferenciales. He realizado un gran esfuerzo para presentar el material de modo que los alumnos que no tengan inclinación alguna por las matemáticas puedan comprenderlo. Sin embargo, no puedo apartarme del todo de las matemáticas. Para tener éxito, usted deberá dominar el álgebra elemental y unas cuantas operaciones matemáticas. Para ayudarle a repasar estos temas, he incluido el apéndice A en el que encontrará las matemáticas que necesita como previo requisito. Usted deberá estudiar ese material y asegurarse de que puede resolver

los problemas incluidos en él. Si tiene dificultades para resolver esos ejercicios, será muy conveniente que repase el tema en un libro de texto básico de álgebra elemental.

Otro factor que se debe tener presente es que en estadística se utilizan muchos símbolos. Por ejemplo, para referirnos a la media de una muestra de datos, utilizaremos el símbolo \bar{X} (que se lee “X barra”). Con frecuencia, los estudiantes creen que el material es más difícil de lo que es en realidad e intentan aprender lo que cada símbolo significa. Usted puede ahorrarse muchos dolores de cabeza si comprende los símbolos en su justa dimensión. Trátelos como si fueran el vocabulario de otro idioma. Memorícelos e intente trabajar con ellos desde un punto de vista conceptual. Por ejemplo, cuando vea que el texto dice \bar{X} , el concepto “la media de la muestra” deberá acudir de inmediato a su mente.

También es importante darse cuenta de que el material de la estadística es acumulativo. No se retrase en sus clases. Si lo hace, no podrá comprender tampoco los conceptos nuevos, en virtud de que no ha dominado los conceptos básicos. La situación puede convertirse en una bola de nieve y, antes que se dé cuenta, usted podría estar retrasado en forma irremediable. Recuerde que debe hacer todo lo posible por mantenerse al corriente del material presentado en clase, a medida que el curso vaya avanzando.

Por último, nuestra experiencia indica que gran parte de la comprensión de la estadística se obtiene con la resolución de muchos problemas. A menudo, un problema vale más que mil palabras. Con frecuencia, aunque el texto del libro sea claro, no entenderá el material hasta que usted haya resuelto los problemas asociados con el tema. Por lo tanto, resuelva muchos problemas y vuelva a leer las secciones para asegurarse de que lo ha comprendido.

En resumen, creo que si domina el álgebra elemental, trabaja de manera diligente, aprende los símbolos, estudia el texto, se mantiene al tanto del material que sea expuesto en clase y resuelve muchos problemas, tendrá éxito. Aunque usted no lo crea, cuando comience a percibir por sí mismo la elegancia y la diversión inherentes a la estadística, llegará incluso a disfrutarla.



NOTACIÓN MATEMÁTICA

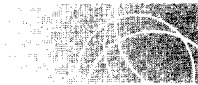
En estadística trabajamos por lo general con datos agrupados, resultantes de la medición de una o más variables. Con mucha frecuencia, los datos se obtienen a partir de muestras y en algunas ocasiones se obtienen a partir de poblaciones. Para fines matemáticos, es útil que ciertos símbolos representen a las variables medidas en el estudio. En todo este texto utilizaremos la letra mayúscula X , y a veces la Y , para representar la(s) variable(s) medida(s). Así, si estuviéramos midiendo la edad de los sujetos, emplearíamos la X para representar la variable “edad”. Si la variable puede adoptar muchos valores, es importante distinguir entre ellos. Esto se logra agregando un subíndice al símbolo X . Este proceso se ilustra en la tabla 2.1.

tabla 2.1 La edad de seis sujetos

Número del sujeto	Símbolo del dato	Valor del dato, edad (años)
1	X_1	8
2	X_2	10
3	X_3	7
4	X_4	6
5	X_5	10
6	X_6	12

En este ejemplo representamos la variable “edad” mediante el símbolo X . Además, N representa el número total de datos que hay en la distribución. En este ejemplo, $N = 6$. Cada uno de los seis datos representa un valor específico de X . Para distinguir entre los seis datos diferentes, se ha agregado un subíndice a X , que corresponde al número de sujeto que tiene el valor dado. Así, el símbolo X_1 corresponde al valor 8 del dato, X_2 al valor 10 del dato, X_3 al valor 7, X_4 a 6, X_5 a 10 y X_6 a 12. En general, podemos referirnos a uno cualquiera de los datos de la distribución X como X_i , donde i puede asumir cualquier valor de 1 a N , según el dato que deseamos designar. En resumen:

- ♦ X o Y representan la variable medida.
- ♦ N representa el número total de sujetos o de datos.
- ♦ X_i es el i -ésimo dato, donde i puede variar de 1 a N .



SUMATORIA

Una de las operaciones que se realizan con más frecuencia en estadística consiste en sumar todos los datos que pertenecen a una distribución o una parte de ellos. Como no resulta práctico escribir “suma de todos los datos” cada vez que se necesita indicar esta operación, particularmente en ecuaciones, se utiliza una abreviatura simbólica. La letra griega mayúscula sigma (Σ) indica la operación de sumatoria. La frase algebraica que se emplea para la sumatoria es:

$$\sum_{i=1}^N X_i$$

Esta expresión se lee como “la suma de la variable X de $i = 1$ a N ”. Las notaciones que aparecen arriba y abajo del signo de sumatoria indican los datos que deben incluirse en la operación. El término que aparece abajo del signo de la sumatoria nos indica el primer dato que interviene en esta operación, y el término que se encuentra arriba de dicho signo indica el último de esos datos. Así, esta frase señala que debemos sumar los datos de X , comenzando con el primero y concluyendo con el N -ésimo dato. Así,

$$\sum_{i=1}^N X_i = X_1 + X_2 + X_3 + \cdots + X_N \quad \text{ecuación de una sumatoria}$$

Al “aplicar la sumatoria” a los datos de las edades de la tabla anterior, tenemos que:

$$\begin{aligned} \sum_{i=1}^6 X_i &= X_1 + X_2 + X_3 + X_4 + X_5 + X_6 \\ &= 8 + 10 + 7 + 6 + 10 + 12 = 53 \end{aligned}$$

Cuando la sumatoria se realiza con todos los datos (de 1 a N), es frecuente que la propia frase con que se designa esta operación se abrevie, omitiendo las notaciones colocadas arriba y abajo del signo de la suma, al igual que el subíndice i . De esta manera,

$$\sum_{i=1}^N X_i \text{ se abrevia con frecuencia como } \Sigma X.$$

En el ejemplo anterior,

$$\Sigma X = 53$$

Esta expresión indica que la suma de todos los datos X es 53.

Observe que no es necesario que la sumatoria se realice desde 1 hasta N . Por ejemplo, podríamos querer sumar sólo el segundo, tercero, cuarto y quinto datos. Recuerde que la notación abajo del signo de sumatoria indica dónde debe comenzar la suma, y el término indicado arriba de dicho signo nos dice dónde terminarla. Así, para indicar la operación de suma del segundo, tercero, cuarto y quinto datos, utilizaríamos el símbolo

$\sum_{i=2}^5 X_i$. Para los datos anteriores sobre edades, tenemos que:

$$\sum_{i=2}^5 X_i = X_2 + X_3 + X_4 + X_5 = 10 + 7 + 6 + 10 = 33$$

Resolvamos algunos problemas de práctica para la sumatoria.



a. Para los siguientes datos, determine $\sum_{i=1}^N X_i$:

$$X: 6, 8, 13, 15 \quad \Sigma X = 6 + 8 + 13 + 15 = 42$$

$$X: 4, -10, -2, 20, 25, 8 \quad \Sigma X = 4 - 10 - 2 + 20 + 25 + 8 = 45$$

$$X: 1.2, 3.5, 0.8, 4.5, 6.1 \quad \Sigma X = 1.2 + 3.5 + 0.8 + 4.5 + 6.1 = 16.1$$

b. Para los siguientes datos, determine $\sum_{i=1}^3 X_i$:

$$X_1 = 10, X_2 = 12, X_3 = 13, X_4 = 18$$

$$\sum_{i=1}^3 X_i = 10 + 12 + 13 = 35$$

c. Para los siguientes datos, determine $\sum_{i=2}^4 X_i + 3$:

$$X_1 = 20, X_2 = 24, X_3 = 25, X_4 = 28, X_5 = 30, X_6 = 31$$

$$\sum_{i=2}^4 X_i + 3 = (24 + 25 + 28) + 3 = 80$$

d. Para los siguientes datos, determine $\sum_{i=2}^4 (X_i + 3)$:

$$X_1 = 20, X_2 = 24, X_3 = 25, X_4 = 28, X_5 = 30, X_6 = 31$$

$$\sum_{i=2}^4 (X_i + 3) = (24 + 3) + (25 + 3) + (28 + 3) = 86$$

Existen otros dos tipos de sumatorias que veremos con frecuencia en este libro. Son $\sum X^2$ y $(\sum X)^2$. Aunque se parecen, son distintos y, en general, proporcionan respuestas diferentes. El símbolo $\sum X^2$ (suma de los cuadrados de los datos X) indica que primero debemos elevar al cuadrado los datos X y luego sumarlos. Así,

$$\sum X^2 = X_1^2 + X_2^2 + X_3^2 + \cdots + X_N^2$$

Dados los valores $X_1 = 3$, $X_2 = 5$, $X_3 = 8$, y $X_4 = 9$,

$$\sum X^2 = 3^2 + 5^2 + 8^2 + 9^2 = 179$$

El símbolo $(\sum X)^2$ (el cuadrado de la suma de los datos X) indica que primero debemos sumar los datos X y luego elevar al cuadrado la suma resultante. Así,

$$(\sum X)^2 = (X_1 + X_2 + X_3 + \cdots + X_N)^2$$

Para los datos anteriores, es decir, $X_1 = 3$, $X_2 = 5$, $X_3 = 8$, y $X_4 = 9$,

$$(\sum X)^2 = (3 + 5 + 8 + 9)^2 = (25)^2 = 625$$

Observe que $\sum X^2 \neq (\sum X)^2$ ($179 \neq 625$). Confundir entre $\sum X^2$ y $(\sum X)^2$ es un error común que cometen los estudiantes, sobre todo cuando calculan la desviación estándar. Regresaremos a este punto en el capítulo 4.*



ESCALAS DE MEDICIÓN: NOMINALES, ORDINALES, DE INTERVALOS Y DE RAZÓN

En virtud de que la estadística analiza los datos y éstos son resultado de las mediciones, necesitamos dedicar cierto tiempo al estudio de las escalas de medición. Este tema es de suma importancia, pues el tipo de escala de medición utilizado para recolectar los datos ayuda a determinar el tipo de prueba de inferencia estadística que conviene emplear en el análisis de los datos. Desde un punto de vista teórico, una escala de medición puede tener uno o varios de los siguientes atributos matemáticos: magnitud, un intervalo igual entre unidades adyacentes y un cero absoluto. Existen cuatro clases de escalas que se utilizan de manera común en las ciencias del comportamiento: *nominal*, *ordinal*, de *intervalo* y de *razón*. Difieren en el número de atributos matemáticos que cada una posee.

Escalas nominales

Una *escala nominal* representa el nivel mínimo de medición y se utiliza con frecuencia para variables de naturaleza cualitativa y no cuantitativa. Algunos ejemplos de variables cualitativas son las marcas de zapatos deportivos, los distintos tipos de fruta o de música, los días de la semana, la nacionalidad, las creencias religiosas y el color de los ojos. Al utilizar una escala nominal, la variable se divide en sus diversas categorías. Estas categorías comprenden las “unidades” de la escala y los objetos se “miden” al determinar la categoría a la cual pertenecen. Así, la medición con una escala nominal equivale, en realidad, a clasificar los objetos y a darles el nombre (de ahí lo de escala nominal) de la categoría a la cual pertenecen.

*Véase la nota 2.1 al final de este capítulo para encontrar reglas adicionales sobre sumatoria, si así lo desea.

Por ejemplo, si a usted le gusta correr, es probable que esté interesado en las diversas marcas de zapatos deportivos que se ofrecen en el mercado, como Brooks, Nike, Adidas, Saucony y New Balance, para mencionar sólo algunas. Los zapatos deportivos son importantes porque, al correr, cada zapato entra en contacto con el suelo unas 800 veces por milla. En una carrera de 5 millas entrará en contacto 4000 veces. Si usted pesa 125 libras (57.5 kg), produciría un impacto total de 300 toneladas sobre cada pie durante una carrera de 5 millas. ¡Eso sí que es pesado! No es de sorprender que los corredores sean particularmente cuidadosos en la elección de su calzado.

La variable “marca de zapatos deportivos” es de tipo cualitativo. Se mide en una escala nominal. Las diversas marcas representan sólo algunas de las posibles categorías (unidades) de esta escala. Si tuviéramos un grupo de zapatos deportivos y quisiéramos medirlos con esta escala, consideraríamos cada uno y determinaríamos su marca. Es importante observar que las unidades de una escala nominal son categorías, por lo cual no existe una relación de magnitud entre ellas. Así, no existe una relación cuantitativa entre las categorías de Nike y Brooks. Nike no es “más marca” de zapatos deportivos que Brooks. Simplemente, son marcas distintas. Esto puede quedar más claro si llamamos a las categorías zapatos deportivos 1 y zapatos deportivos 2, en vez de Nike y Brooks. En este caso, los números 1 y 2 son sólo nombres y no poseen una relación de magnitud entre ellos.

Una propiedad fundamental de las escalas nominales es la *equivalencia*. Esto quiere decir que todos los miembros de una clase dada son iguales desde el punto de vista de la variable de clasificación. Así, a todos los pares de zapatos Nike se les considera iguales desde el punto de vista de “marca de zapatos deportivos”, a pesar de que pueda haber varios modelos distintos de zapatos Nike.

Una operación que se realiza con frecuencia junto con la medición nominal consiste en contar las menciones de cada tipo dentro de cada clase. Por ejemplo, si tuviéramos varios zapatos deportivos y determináramos la marca de cada uno, estaríamos realizando una medición nominal. Además, tal vez nos interesara contar el número de zapatos que pertenecen a cada categoría. Así, podríamos tener 20 zapatos Nike, 19 Saucony y 6 New Balance en total. Estas frecuencias nos permiten comparar el número de zapatos dentro de cada categoría. Esta comparación cuantitativa de las cifras dentro de cada categoría no debe confundirse con la afirmación hecha anteriormente de que no existe una relación de magnitud entre las unidades de una escala nominal. Podemos comparar de manera cuantitativa las cantidades de zapatos Nike con las de zapatos Saucony, pero Nike no es “más marca” de zapatos de lo que es Saucony. Así, una escala nominal no tiene los atributos matemáticos de magnitud, intervalos iguales o cero absoluto. Sólo permite la clasificación de los objetos en categorías mutuamente excluyentes.

Escalas ordinales

La *escala ordinal* representa el siguiente nivel de medición. Tiene un nivel relativamente bajo de la propiedad de magnitud. Con esta escala ordenamos los objetos medidos, tomando como base si poseen más, menos o la misma cantidad de la variable medida. Así, una escala ordinal permite determinar si $A > B$, $A = B$ o $A < B$.

Un ejemplo de una escala ordinal es el ordenamiento de los cinco mejores participantes en un concurso de oratoria, según su facilidad de palabra. Entre los oradores, la persona con el rango 1 fue considerada mejor que la persona que poseía el rango 2, quien a su vez fue mejor que la de rango 3. La persona de rango 3 fue estimada como mejor orador que la de rango 4, la cual fue considerada mejor que la persona del rango 5. Es importante observar que aunque esta escala permite hacer comparaciones del tipo “mejor que, igual a o menor que”, no dice nada sobre la magnitud de la diferencia entre las unidades adyacentes que pertenecen a dicha escala. En este ejemplo, la diferencia en

cuanto a la facilidad de palabra entre las personas de rango 1 y 2 podría ser grande y la que existe entre los individuos de rango 2 y 3 podría ser pequeña. Así, una escala ordinal no tiene la propiedad de mantener intervalos iguales entre las unidades adyacentes. Además, como lo único que tenemos son órdenes relativos, la escala no dice nada acerca del nivel absoluto de la variable. Así, los oradores de los cinco primeros lugares podrían tener un alto nivel de facilidad de palabra o un nivel bajo. Esta información no se puede obtener de una escala ordinal.

Otros ejemplos de escala ordinal son la clasificación de los corredores que participan en el Maratón de Boston según su orden de llegada, la clasificación de los equipos de fútbol americano universitario con base en las calificaciones de la Associated Press, el orden de los profesores según su capacidad de enseñanza y el orden de los estudiantes según su nivel de motivación.

Escalas de intervalos

La *escala de intervalo* representa un nivel superior de medición con respecto a la escala ordinal. Posee las propiedades de magnitud e igualdad de intervalos entre las unidades adyacentes, pero no tiene un cero absoluto. Así, la escala de intervalo posee las propiedades de la escala ordinal y tiene "intervalos iguales entre sus unidades adyacentes": estos intervalos indican que existen cantidades iguales de la variable medida entre las unidades adyacentes en la escala.

La escala Celsius para medir la temperatura es un buen ejemplo de una escala de intervalo. Tiene la propiedad de poseer intervalos iguales entre las unidades adyacentes, pero no tiene un punto correspondiente al cero absoluto. La propiedad de intervalos iguales se percibe por el hecho de que un cambio dado de calor producirá el mismo efecto en la lectura de temperatura en la escala, sin importar en qué parte de ésta ocurra dicho cambio. Así, la cantidad adicional de calor que provoca un cambio en la lectura de la temperatura de 2° a 3° Celsius producirá también un cambio en la lectura de 51° a 52° o de 105° a 106° Celsius. Esto ilustra el hecho de que cantidades iguales de calor están representadas entre las unidades adyacentes, en toda la escala.

Como en una escala de intervalo existen cantidades iguales de la variable entre las unidades adyacentes de la misma, las diferencias equivalentes entre los números de la escala representan diferencias de la misma magnitud en la variable. Así, podemos decir que la diferencia de calor es la misma entre 78° y 75° Celsius que entre 24° y 21° de la misma escala. Esto también implica, desde un punto de vista lógico, que diferencias grandes entre los números de la escala representan también diferencias grandes entre la magnitud de la variable medida, y que diferencias pequeñas entre los números de la escala representan pequeñas diferencias en la magnitud de la variable medida. Así, la diferencia de calor entre 80° y 65° es mayor que la diferencia entre 18° y 15° Celsius, y la diferencia de calor entre 93° y 91° es menor que la diferencia entre 48° y 40° Celsius. En vista del análisis anterior, podemos ver que además de determinar si $A = B$, $A > B$ o $A < B$, una escala de intervalo nos permite determinar si $A - B = C - D$, $A - B > C - D$ o $A - B < C - D$.

Escalas de razón

El siguiente y máximo nivel de medición es la *escala de razón*. Tiene todas las propiedades de una escala de intervalo y, además, posee un cero absoluto. Sin éste no se pueden calcular las razones con respecto a las lecturas de la escala. Como este tipo de escala tiene un cero absoluto, con ella se pueden utilizar razones (de aquí el nombre escala de *razón*).

Un ejemplo adecuado para ilustrar la diferencia entre la escala de intervalo y la de razón consiste en comparar la escala Celsius de temperatura con la escala Kelvin. El cero de la escala Kelvin es el cero absoluto (la ausencia completa de calor). El cero de la escala Celsius es la temperatura a la cual se congela el agua. Es un punto cero arbitrario que se presenta en realidad a los 273° Kelvin. La escala Celsius es de intervalo y la escala Kelvin es de razón. La diferencia de calor entre 8° y 9° es igual a la diferencia entre 99° y 100° sin importar que la escala sea Celsius o Kelvin. Sin embargo, no podemos calcular razones con la escala Celsius. Una lectura de 20° Celsius no es el doble de caliente que 10° Celsius. Esto se puede comprobar al convertir las lecturas Celsius al calor real que representan. En términos de calor real, 20° Celsius es en realidad 293°K (273° + 20°), y 10° Celsius es en realidad 283°K (273° + 10°). Es obvio que 293° no es el doble de 283°. Puesto que la escala Kelvin tiene un cero absoluto, una lectura de 20° en esta escala es el doble de caliente que una de 10°. De esta manera se pueden utilizar razones con la escala Kelvin.

Otros ejemplos de variables medidas con escalas de razones son: el tiempo de reacción, la longitud, el peso, la edad y la frecuencia de cualquier evento, como el número de zapatos Nike incluidos en el conjunto de zapatos deportivos que comentamos con anterioridad. Con una escala de este tipo, usted puede establecer razones y realizar todas las demás operaciones matemáticas asociadas por lo general a los números (es decir, suma, resta, multiplicación y división). Las cuatro escalas de medición y sus características se resumen en la figura 2.1.

Nominal

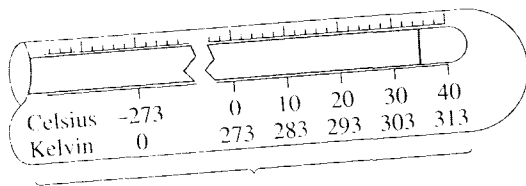
Las unidades de la escala son categorías. Los objetos se miden al determinar la categoría a la cual pertenecen. No existe una relación de magnitud entre las categorías.

Nike New Balance Saucony



Ordinal

Posee la propiedad de magnitud. Se puede ordenar a los objetos según que posean más, menos o igual cantidad de la variable medida. Así, se puede determinar si $A > B$, $A = B$, $A < B$. No se puede determinar cuánto mayor o menor es A que B en relación con el atributo medido.



Intervalo y razón

Intervalo: posee las propiedades de magnitud e intervalos iguales entre unidades adyacentes. Se pueden hacer las mismas determinaciones que en la escala ordinal y, además, determinar si $A - B = C - D$, $A - B > C - D$ o $A - B < C - D$.

Razón: posee las propiedades de magnitud, intervalos iguales entre unidades adyacentes y un punto cero absoluto. Se pueden realizar todas las operaciones matemáticas asociadas por lo general con los números, incluidas las razones.

figura 2.1 Escalas de medición y sus características.



ESCALAS DE MEDICIÓN EN LAS CIENCIAS DEL COMPORTAMIENTO

En las ciencias del comportamiento, muchas de las escalas utilizadas se consideran con frecuencia como si fuesen de intervalo sin establecer con claridad que la escala, en realidad, posee intervalos iguales entre unidades adyacentes. La medición del CI; las variables emocionales como la ansiedad y la depresión; las variables de la personalidad (como la autosuficiencia, la introversión, la extroversión y la dominancia); las variables de eficiencia o logros terminales; las variables de actitud, entre otras, corresponden a esta categoría. Con todas estas variables es claro que las escalas no son de razón. Por ejemplo, en el caso del CI, si una persona ha obtenido un cero en la escala de inteligencia para adultos de Wechsler (WAIS), no podríamos decir que tiene inteligencia nula. Se supone que se encontrarían algunas preguntas que podría contestar el individuo en cuestión, de modo que tendría un CI superior a cero. Así, la prueba WAIS no tiene un punto cero absoluto y las razones no son adecuadas. Por lo tanto, no es correcto decir que una persona con un CI de 160 es lo doble de inteligente que alguien con un CI de 80.

Por otro lado, parece que podemos hacer algo más que sólo especificar un ordenamiento de los individuos. Una persona que posee un CI de 100 tiene una inteligencia más cercana a alguien con un CI de 110 que a otra persona con un CI de 60. Esto parece ser una escala de intervalo, pero es difícil establecer que la escala en realidad posea intervalos iguales entre las unidades adyacentes. Muchos investigadores consideran tales variables como si fuesen medidas en escalas de intervalo, en particular cuando el instrumento de medición está bien normalizado, como en el caso del WAIS. Es mucho más discutible considerar como escalas de intervalos las escalas deficientemente normalizadas que miden variables psicológicas. Este tema surge en particular en la estadística inferencial, donde el nivel de las escalas puede influir en la selección de la prueba que se usará en el análisis de datos. Existen dos criterios al respecto. El primero afirma que el uso de ciertas pruebas, como la *t* de Student y el análisis de varianza, deben limitarse a los datos que pertenecen a escalas de intervalo y de razón. El segundo criterio no concuerda con lo anterior y afirma que esas pruebas también se pueden utilizar con datos nominales y ordinales. Sin embargo, este tema es demasiado complejo para ser considerado aquí.*



VARIABLES CONTINUAS Y DISCRETAS

En el capítulo 1 definimos la variable como una propiedad o característica de alguna cosa que puede asumir más de un valor. También distinguimos entre las variables independientes y dependientes. Además de esto, las variables pueden ser continuas o discretas:

*El lector interesado debe consultar N. H. Anderson, "Scales and Statistics: Parametric and Nonparametric", *Psychological Bulletin*, 58 (1961), 305-316; F. M. Lord, "On the Statistical Treatment of Football Numbers", *American Psychologist*, 8 (1953), 750-751; W. L. Hays, *Statistics for the Social Sciences*, segunda edición, Holt, Rinehart y Winston, Nueva York, 1973, pp. 87-90; S. Siegel, *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill, Nueva York, 1956, pp. 18-20; y S. S. Stevens, "Mathematics, Measurement, and Psychophysics", en *Handbook of Experimental Psychology*, S. S. Stevens, editor Wiley, Nueva York, 1951, pp. 23-30.

- definiciones**
- Una **variable continua** es aquella que teóricamente puede asumir un número infinito de valores entre las unidades adyacentes de una escala.
 - Una **variable discreta** es aquella para la cual no existen valores posibles localizados entre las unidades adyacentes de una escala.

El peso, la estatura y el tiempo son ejemplos de variables continuas. Para cada una de estas variables existe, de manera potencial, una infinidad de valores entre las unidades adyacentes. Si estamos midiendo el tiempo y la unidad mínima que utilizamos es un segundo, entonces entre 1 y 2 segundos existe una infinidad de valores posibles: 1.1 segundos, 1.01 segundos, 1.001 segundos, y así sucesivamente. Podemos aplicar el mismo argumento al peso y a la estatura.

Esto no ocurre con una variable discreta. El “número de hijos en una familia” es un ejemplo de variable discreta. En este caso, la menor unidad es un hijo y no existen valores posibles entre uno y dos hijos, dos y tres hijos, etc. La característica de una variable discreta es que cambia por cantidades fijas sin que haya valores intermedios. Otros ejemplos son: “el número de estudiantes en su grupo”, “el número de profesores en una universidad” y “el número de citas que tuvo el último mes”.

Límites reales de una variable continua

Como una variable continua puede tener una infinidad de valores entre las unidades adyacentes de una escala, todas las mediciones realizadas sobre esta variable son *aproximadas*. Utilizaremos el peso para ilustrar esto. Suponga que comenzó una dieta el día de ayer, que es primavera, casi verano, y que el clima ideal para ponerse el traje de baño está ya muy próximo. De cualquier modo, usted se pesó ayer en la mañana y su peso fue señalado por la aguja de la figura 2.2. Imagine que la escala de la figura tiene una precisión sólo a la libra (o kilo) más cercana. El peso registrado fue de 180 libras (83 k). Es-

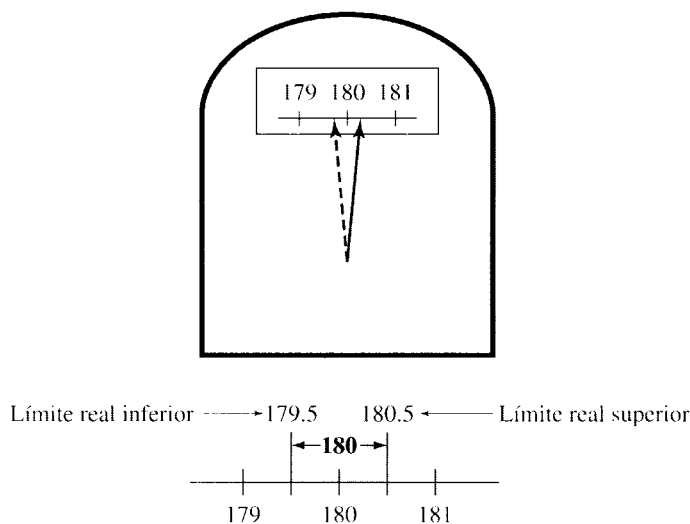


figura 2.2 Límites reales de una variable continua.

ta mañana, cuando usted se pesó después de un día de hacer dieta, la aguja llegó hasta el lugar indicado con línea interrumpida. ¿Qué peso registró esta vez? Por ser un rasgo muy humano, sabemos que a usted le agradaría registrar 179 libras; sin embargo, como futuro científico, usted busca la *verdad* a cualquier precio. Por eso registra de nuevo 180 libras. ¿Cuándo podría registrar 179 libras? Cuando la aguja esté debajo del punto medio entre 179 y 180 libras. De manera análoga, usted seguiría registrando 180 libras si la aguja estuviera arriba de 180, pero antes del punto medio entre 180 y 181 libras. Por consiguiente, cada vez que se registre el peso de 180 libras, eso no quiere decir que sean exactamente 180 libras, sino que el peso está entre 180 ± 0.5 libras. No conocemos el valor exacto del peso, pero estamos seguros de que se encuentra en el rango de 179.5 a 180.5. Este rango especifica los límites reales del peso de 180 libras. El valor de 179.5 es el límite real inferior y 180.5 es el límite real superior.

definiciones ■ *Los límites reales de una variable continua son aquellos valores que se encuentran por arriba y por debajo de los valores registrados, a una distancia igual a la mitad de la unidad de medida mínima de la escala utilizada.*

Por ejemplo, si la variable en cuestión es el peso, la menor unidad es 1 libra, y registramos 180 libras, los límites reales están arriba y abajo de 180 libras por $\frac{1}{2}$ libra. Así, los límites reales son 179.5 y 180.5 libras.* Si la unidad mínima fuese 0.1 libras y no 1 libra y registraríamos 180.0 libras, entonces los límites reales serían $180 \pm \frac{1}{2}(0.1)$, es decir, 179.95 y 180.05.

Cifras significativas

En estadística analizamos datos; este análisis implica numerosos cálculos matemáticos. Con mucha frecuencia obtenemos un residuo decimal, por ejemplo, después de realizar una división. Cuando esto ocurre, necesitamos decidir la cantidad de cifras decimales que utilizaremos para el residuo.

En las ciencias físicas, por lo general, se utiliza el mismo número de cifras significativas que tienen los datos en bruto. Por ejemplo, si medimos el peso de cinco sujetos hasta tres cifras significativas (173, 156, 162, 165 y 175 libras) y queremos calcular el promedio de estos pesos, nuestra respuesta deberá contener sólo tres cifras significativas. Así,

$$\bar{X} = \frac{\sum X}{N} = \frac{173 + 156 + 162 + 165 + 175}{5} = \frac{831}{5} = 166.2 = 166$$

La respuesta de 166.2 se redondea a tres cifras significativas, dando un resultado final de 166 libras. Por varias razones, este procedimiento no se ha seguido en las ciencias del comportamiento, sino que se ha establecido una tradición en la cual la mayor parte de los valores finales se expresan con dos o tres cifras decimales, sin importar el número de cifras significativas de los datos en bruto. Como éste es un texto para usarse en las ciencias del comportamiento, hemos optado por seguir esta tradición. Así, en este libro daremos la mayor parte de nuestras respuestas finales hasta con dos cifras decimales. En algunas ocasiones habrá excepciones. Por ejemplo, los coeficientes de correlación y de regresión tienen tres cifras decimales, y los valores de probabilidad se dan a menudo has-

*A decir verdad, los límites reales son 179.500000... y 180.499999..., pero no es necesario ser tan precisos.

ta con cuatro cifras decimales, lo cual es consistente con la tradición. *Es una práctica común realizar todos los cálculos intermedios con dos o más cifras decimales de las que tendrá la respuesta final.* Así, cuando se le pida que esta última tenga dos cifras decimales, usted deberá realizar los cálculos intermedios con cuatro cifras decimales por lo menos y redondeará la respuesta final a dos cifras.

Redondeo

Considerando que expresaremos nuestras respuestas finales con dos, y en algunas ocasiones hasta con tres o cuatro cifras decimales, necesitamos decidir cómo se determinará el valor del último dígito. Por fortuna, las siguientes reglas de redondeo son bastante sencillas y directas:

1. Divida en dos partes el número que desea redondear: la respuesta potencial y el residuo. La respuesta potencial es el número original y debe llegar solamente hasta el número deseado de cifras decimales. El residuo es el resto de ese número original.
2. Coloque un punto decimal antes del primer dígito del residuo, creando de esta manera un residuo decimal.
3. Si el residuo decimal es mayor que $\frac{1}{2}$, sume 1 al último dígito de la respuesta.
4. Si el residuo decimal es menor que $\frac{1}{2}$, deje el último dígito de la respuesta sin cambio alguno.
5. Si el residuo decimal es igual a $\frac{1}{2}$, sume 1 al último dígito de la respuesta si dicho dígito es impar, pero si es par, déjelo sin cambio alguno.

Probemos con algunos ejemplos. Redondearemos los números de la columna de la izquierda, en la tabla 2.2, hasta dos cifras decimales.

Para realizar este redondeo, el número se separa en dos partes: la respuesta potencial y el residuo. En vista de que estamos redondeando a dos cifras decimales, la respuesta potencial termina en la segunda cifra decimal. El resto del número constituye el residuo. Para el primer número, 34.01350, 34.01 constituye la respuesta potencial y .350 es el residuo. Como .350 es menor que $\frac{1}{2}$, el último dígito de la respuesta potencial permanece sin cambio alguno y la respuesta final es 34.01. Para el segundo número, 34.01761, el residuo decimal (.761) es mayor que $\frac{1}{2}$. Por lo tanto, tenemos que sumar 1 al último dígito, con lo cual la respuesta correcta es 34.02. Para los dos números siguientes, el residuo

tabla 2.2 Redondeo

Número	Respuesta; residuo	Residuo decimal	Respuesta final	Razón
34.01350	34.01;350	.350	34.01	El residuo decimal es menor que $\frac{1}{2}$.
34.01761	34.01;761	.761	34.02	El residuo decimal es mayor que $\frac{1}{2}$.
45.04500	45.04;500	.500	45.04	El residuo decimal es igual a $\frac{1}{2}$, y el último dígito es par.
45.05500	45.05;500	.500	45.06	El residuo decimal es igual a $\frac{1}{2}$, y el último dígito es impar.

decimal es igual a $\frac{1}{2}$. El número 45.04500 se convierte en 45.04 en virtud de que el último dígito de la respuesta potencial es par. El número 45.05500 se convierte en 45.06 debido a que el último dígito es impar.

RESUMEN

En este capítulo hemos analizado varios conceptos básicos de las matemáticas y de la medición. Los temas analizados fueron la notación, la suma, las escalas de medición, las variables discretas y continuas, y el redondeo. Además, señalamos que para tener éxito en

estadística, usted no tiene que ser un gran matemático. Si posee un conocimiento sólido del álgebra elemental, realiza un buen número de problemas, presta atención a los símbolos y se mantiene al corriente en sus clases, logrará comprender íntegramente el material.

TÉRMINOS IMPORTANTES

Escala de intervalo (página 29).

Escala de razón (página 29).

Escala nominal (página 27).

Escala ordinal (página 28).

Límites reales de una variable
(página 33).

Sumatoria (página 25).

Variable continua (página 32).

Variable discreta (página 32)

PREGUNTAS Y PROBLEMAS

- Defina los conceptos que aparecen en la sección "Términos importantes" y proponga un ejemplo de cada uno.
- Identifique cuáles de las siguientes variables son continuas y cuáles son discretas:
 - La hora del día.
 - Número de mujeres en su grupo.
 - Número de veces que una rata oprime la palanca en una caja de Skinner
 - Edad de los sujetos en un experimento.
 - Número de palabras recordadas.
 - Peso del alimento ingerido.
 - Porcentaje de estudiantes de su grupo que son mujeres.
 - Velocidad de los corredores en una carrera.
- Identifique la escala de cada una de las siguientes variables:
 - Número de bicicletas que utilizan los estudiantes de primer año.
 - Tipos de bicicletas utilizadas por los estudiantes de primer año.
 - El CI de sus profesores (suponga que utiliza una escala con intervalos iguales).
 - El dominio de las matemáticas, clasificado en las categorías de malo, regular y bueno.
 - Sensación de ansiedad al hablar en público, calificada según una escala de 0 a 100 (suponga que la diferencia en el grado de ansiedad entre las unidades adyacentes de toda la escala no siempre es la misma).
 - El peso de un grupo de personas sometidas a una dieta.
 - El tiempo de reacción al escuchar un sonido.
 - El dominio de las matemáticas, medido en una escala de 0 a 100. La escala está bien normalizada y se puede considerar que tiene intervalos iguales entre las unidades adyacentes.
 - Las calificaciones de los profesores asignadas por los estudiantes, en una escala de 50 puntos. No existe una base suficiente para suponer que los intervalos entre las unidades adyacentes son iguales.
- Un estudiante mide la asertividad con una escala de intervalos. ¿Es correcto decir que un puntaje de 30 en la escala representa la mitad de asertividad que un puntaje de 60? Explique.
- Para cada uno de los siguientes conjuntos de datos, determine

$$\sum_{i=1}^N X_i;$$
 - 2, 4, 5, 7
 - 2.1, 3.2, 3.6, 5.0, 7.2
 - 11, 14, 18, 22, 25, 28, 30
 - 110, 112, 115, 120, 133
- Redondee los siguientes números a dos cifras decimales:

a. 14.53670	b. 25.26231
c. 37.83500	d. 46.50499
e. 52.46500	f. 25.48501
- Determine los límites reales de los siguientes valores:
 - 10 libras (suponga que la unidad de medida mínima es de 1 libra).

- b. 2.5 segundos (suponga que la unidad de medida mínima es de 0.1 segundo).
 c. 100 gramos (suponga que la unidad de medida mínima es de 10 gramos).
 d. 2.01 centímetros (suponga que la unidad de medida mínima es de 0.01 centímetro).
 e. 5.232 segundos (suponga que la unidad de medida mínima es de 1 milisegundo).
8. Determine los valores de las siguientes expresiones:

a. Encuentre $\sum_{i=1}^4 X_i$ para los datos $X_1 = 3, X_2 = 5, X_3 = 7, X_4 = 10$.

b. Encuentre $\sum_{i=1}^6 X_i$ para los datos $X_1 = 2, X_2 = 3, X_3 = 4, X_4 = 6, X_5 = 9, X_6 = 11, X_7 = 14$.

c. Encuentre $\sum_{i=2}^N X_i$ para los datos $X_1 = 10, X_2 = 12, X_3 = 13, X_4 = 15, X_5 = 18$.

d. Encuentre $\sum_{i=3}^{N-1} X_i$ para los datos $X_1 = 22, X_2 = 24, X_3 = 28, X_4 = 35, X_5 = 38, X_6 = 40$.

9. En un experimento para medir los tiempos de reacción de ocho sujetos se obtuvieron los siguientes datos, en milisegundos:

Sujeto	Tiempo de reacción
1	250
2	378
3	451
4	275
5	225
6	430
7	325
8	334

- a. Si X representa la variable del tiempo de reacción, asigne a cada uno de los datos su símbolo X_i apropiado.
 b. Calcule $\sum X$ para estos datos.
10. Represente cada una de las siguientes expresiones por medio de la notación de sumatoria. Suponga que el número total de datos es 10.
- $X_1 + X_2 + X_3 + X_4 + \dots + X_{10}$
 - $X_1 + X_2 + X_3$
 - $X_2 + X_3 + X_4$
 - $X_2^2 + X_3^2 + X_4^2 + X_5^2$

11. Redondee los siguientes números a una cifra decimal:
- 1.423
 - 23.250
 - 100.750
 - 41.652
 - 35.348
12. Para cada uno de los conjuntos de datos correspondientes a los problemas 5b y 5c, demuestre que $\sum X^2 \neq (\sum X)^2$.

13. A partir de los datos $X_1 = 3, X_2 = 4, X_3 = 7$ y $X_4 = 12$, determine usted los valores de las siguientes expresiones (esta pregunta corresponde a la nota 2.1.)

a. $\sum_{i=1}^N (X_i + 2)$ b. $\sum_{i=1}^N (X_i - 3)$

c. $\sum_{i=1}^N (2X_i)$ d. $\sum_{i=1}^N (X_i/4)$

14. Redondee cada uno de los siguientes números a una cifra decimal y a dos cifras decimales.
- 4.1482
 - 4.1501
 - 4.1650
 - 4.1950

NOTAS

2.1 Muchos libros de texto incluyen un análisis de otras reglas para la sumatoria, como la sumatoria de una variable más una constante, la sumatoria de una variable por una constante, etc. Como estos conceptos no son necesarios para comprender el material de este libro, no los incluimos en el cuerpo principal del texto, pero hemos presentado aquí ese material. El conocimiento de las reglas de la sumatoria puede ser muy útil como antecedente para los cursos de estadística que se imparten en el nivel de posgrado.

Regla 1 La suma de los valores de una variable más una constante es igual a la suma de los valores de la variable más N veces la constante. En forma de ecuación, se representa así:

$$\sum_{i=1}^N (X_i + a) = \sum_{i=1}^N X_i + Na$$

La validez de esta ecuación se puede apreciar mediante la sencilla demostración algebraica que presentaremos a continuación:

$$\begin{aligned}\sum_{i=1}^N (X_i + a) &= (X_1 + a) + (X_2 + a) + (X_3 + a) \\ &\quad + \cdots + (X_N + a) \\ &= (X_1 + X_2 + X_3 + \cdots + X_N) \\ &\quad + (a + a + a + \cdots + a) \\ &= \sum_{i=1}^N X_i + Na\end{aligned}$$

Para ilustrar el uso de esta ecuación, suponga que queremos calcular la suma de los siguientes datos con la constante 3 sumada a cada uno de ellos:

$$X: 4, 6, 8, 9$$

$$\sum_{i=1}^N (X_i + 3) = \sum_{i=1}^N X_i + Na = 27 + 4(3) = 39$$

Regla 2 La suma de los valores de una variable menos una constante es igual a la suma de los valores de la variable menos N veces la constante. En forma de ecuación,

$$\sum_{i=1}^N (X_i - a) = \sum_{i=1}^N X_i - Na$$

La demostración algebraica de esta ecuación es la siguiente:

$$\begin{aligned}\sum_{i=1}^N (X_i - a) &= (X_1 - a) + (X_2 - a) + (X_3 - a) \\ &\quad + \cdots + (X_N - a) \\ &= (X_1 + X_2 + X_3 + \cdots + X_N) \\ &\quad + (-a - a - a - a - \cdots - a) \\ &= \sum_{i=1}^N X_i - Na\end{aligned}$$

Para ilustrar el uso de esta ecuación, suponga que queremos determinar la suma de los siguientes datos con la constante 2 restada de cada uno:

$$X: 3, 5, 6, 10$$

$$\sum_{i=1}^N (X_i - 2) = \sum_{i=1}^N X_i - Na = 24 - 4(2) = 16$$

Regla 3 La suma de una constante por el valor de una variable es igual a la constante multiplicada por la suma de los valores de la variable. En forma de ecuación,

$$\sum_{i=1}^N aX_i = a \sum_{i=1}^N X_i$$

A continuación demostraremos la validez de esta ecuación:

$$\begin{aligned}\sum_{i=1}^N aX_i &= aX_1 + aX_2 + aX_3 + \cdots + aX_N \\ &= a(X_1 + X_2 + X_3 + \cdots + X_N) \\ &= a \sum_{i=1}^N X_i\end{aligned}$$

Para ilustrar el uso de esta ecuación, suponga que queremos determinar la suma de 4 por cada uno de los siguientes datos:

$$X: 2, 5, 7, 8, 12$$

$$\sum_{i=1}^N 4X_i = 4 \sum_{i=1}^N X_i = 4(34) = 136$$

Regla 4 La suma de una constante dividida entre los valores de una variable es igual a la constante, dividida entre la suma de los valores de la variable. En forma de ecuación,

$$\sum_{i=1}^N \frac{X_i}{a} = \frac{\sum_{i=1}^N X_i}{a}$$

A continuación demostraremos la validez de esta ecuación:

$$\begin{aligned}\sum_{i=1}^N \frac{X_i}{a} &= \frac{X_1}{a} + \frac{X_2}{a} + \frac{X_3}{a} + \cdots + \frac{X_N}{a} \\ &= \frac{X_1 + X_2 + X_3 + \cdots + X_N}{a} \\ &= \frac{\sum_{i=1}^N X_i}{a}\end{aligned}$$

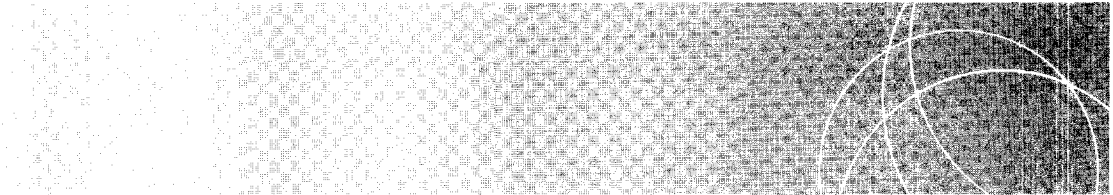
Veamos una vez más un ejemplo para ilustrar el uso de esta ecuación. Suponga que queremos determinar la suma de cada uno de los siguientes datos divididos entre 4:

$$X: 3, 4, 7, 10, 11$$

$$\sum_{i=1}^N \frac{X_i}{4} = \frac{\sum_{i=1}^N X_i}{4} = \frac{35}{4} = 8.75$$

Si desea encontrar problemas de práctica adicionales (en inglés) totalmente resueltos sobre el material que acaba de estudiar, busque Chapter 2 en el material adicional del libro, en: <http://www.thomsonlearning.com.mx>.

Distribuciones de frecuencias



- Introducción: distribuciones de frecuencias no agrupadas
- Datos agrupados
- Percentiles
- Rangos percentiles
- Graficación de las distribuciones de frecuencias
- Análisis exploratorio de los datos
- ¿CUÁL ES LA VERDAD? Al cambiar la escala, cambia el cuento
- Resumen
- Términos importantes
- Preguntas y problemas



INTRODUCCIÓN: DISTRIBUCIONES DE FRECUENCIAS NO AGRUPADAS

Supongamos que le acaban de devolver su primer examen de estadística y que en él obtuvo una calificación de 86. Es natural que sienta interés por saber qué tan bueno fue ese resultado con respecto a los demás estudiantes. Usted tiene muchas preguntas que hacer: ¿cuántos estudiantes obtuvieron 86?, ¿hubo muchas calificaciones más altas que la suya?, ¿cuántas calificaciones fueron más bajas que la de usted? Las calificaciones en bruto del examen se presentan desorganizadas en la tabla 3.1. Aunque en ella aparecen todas las calificaciones, resulta difícil entenderlas por la forma en la que han sido dispuestas en dicha tabla. Un arreglo más eficiente, y que tiene mayor significado, podría consistir en listar las calificaciones según su frecuencia de aparición. Esta lista se conoce como una *distribución de frecuencias*.

tabla 3.1 Calificaciones del examen de estadística ($N = 70$)

95	57	76	93	86	80	89
76	76	63	74	94	96	77
65	79	60	56	72	82	70
67	79	71	77	52	76	68
72	88	84	70	83	93	76
82	96	87	69	89	77	81
87	65	77	72	56	78	78
58	54	82	82	66	73	79
86	81	63	46	62	99	93
82	92	75	76	90	74	67

definición ■ Una **distribución de frecuencias** presenta los valores de los datos y la frecuencia con que se presentan. Al ser mostrados en una tabla, los valores de los datos se presentan en orden y, por lo general, el valor del dato más bajo aparece en la parte inferior de la tabla.

Las calificaciones de la tabla 3.1 han sido dispuestas según una distribución de frecuencias que se presenta en la tabla 3.2. Ahora los datos tienen mayor significado. En primer lugar, resulta fácil observar que hay 2 calificaciones de 86. Además, cuando sumamos las frecuencias apropiadas (f) podemos determinar cuál es el número de calificaciones mayores que y menores que 86. Se puede ver que hay 15 calificaciones más altas y 53 calificaciones más bajas que la de usted. También es fácil determinar el rango de los datos en una distribución de frecuencias. En el caso del examen de estadística, las calificaciones variaron en un rango de entre 46 y 99. A partir de este ejemplo podemos ver que el objetivo principal de una distribución de frecuencias consiste en presentar los datos de una manera que facilite su comprensión e interpretación.

tabla 3.2 Calificaciones de la tabla 3.1 organizadas en una distribución de frecuencias

Calificación	f	Calificación	f	Calificación	f	Calificación	f
99	1	85	0	71	1	57	1
98	0	84	1	70	2	56	2
97	0	83	1	69	1	55	0
96	2	82	5	68	1	54	1
95	1	81	2	67	2	53	0
94	1	80	1	66	1	52	1
93	3	79	3	65	2	51	0
92	1	78	2	64	0	50	0
91	0	77	4	63	2	49	0
90	1	76	6	62	1	48	0
89	2	75	1	61	0	47	0
88	1	74	2	60	1	46	1
87	2	73	1	59	0		
86	2	72	3	58	1		



DATOS AGRUPADOS

Cuando se dispone de muchos datos con un rango muy amplio, como en el caso del examen de estadística que estamos considerando, la enumeración de los datos individuales produce muchos valores cuya frecuencia es cero (nula) y el resultado es una presentación en la que resulta difícil visualizar la forma de la distribución y su tendencia central. En estas condiciones, los datos individuales se agrupan por lo general en intervalos de clase y se presentan como una *distribución de frecuencias de datos agrupados*. La tabla 3.3 muestra las calificaciones del examen de estadística agrupadas en dos distribuciones de frecuencias, una en la que cada intervalo tiene 2 unidades de amplitud y la otra con intervalos de 19 unidades de amplitud.

Al agrupar datos, uno de los puntos importantes consiste en determinar cuál será la amplitud de cada intervalo. Cada vez que los datos son agrupados se pierde un poco de información. Mientras más amplio sea el intervalo, más información se perderá. Por ejemplo, considere la distribución que aparece en la tabla 3.3 con intervalos de 19 unidades

tabla 3.3 Calificaciones de la tabla 3.1 agrupadas en intervalos de clase con amplitudes diferentes

Intervalo de clase (amplitud = 2)	f	Intervalo de clase (amplitud = 19)	f
98-99	1	95-113	4
96-97	2	76-94	38
94-95	2	57-75	23
92-93	4	38-56	5
90-91	1		$N = 70$
88-89	3		
86-87	4		
84-85	1		
82-83	6		
80-81	3		
78-79	5		
76-77	10		
74-75	3		
72-73	4		
70-71	3		
68-69	2		
66-67	3		
64-65	2		
62-63	3		
60-61	1		
58-59	1		
56-57	3		
54-55	1		
52-53	1		
50-51	0		
48-49	0		
46-47	1		
	$N = 70$		

des de amplitud. Aunque este amplitud de intervalo produce una presentación suave (no existen frecuencias nulas), se ha perdido mucha información. Por ejemplo, ¿cómo están distribuidas las 38 calificaciones en el intervalo de 76 a 94?, ¿están en 94? ¿o en 76? ¿o están distribuidas de manera uniforme en todo el intervalo? La cuestión es que no sabemos cómo están distribuidas en el intervalo. Hemos perdido esa información al agrupar los datos. Observe que mientras más amplio sea el intervalo, mayor será la ambigüedad.

Debe parecer obvio que mientras menor sea el intervalo, tanto más fielmente se preservarán los datos originales. El caso extremo se presenta cuando la amplitud del intervalo se reduce a una sola unidad, con lo cual regresamos a la situación de los datos individuales. Sin embargo, cuando el intervalo se vuelve demasiado estrecho, nos encontramos con los mismos problemas que en el caso de los datos individuales: a saber, valores cuya frecuencia es nula y una presentación poco clara en lo referente a la forma de la distribución y su tendencia central. La distribución de frecuencias con intervalos de 2 unidades de amplitud que aparece en la tabla 3.3 es un ejemplo en el cual los intervalos son demasiado pequeños.

Por la explicación anterior podemos ver que, al agrupar los datos, hay que resolver el dilema entre perder información y tener una presentación visual significativa. Para lograr lo mejor de ambas opciones debemos elegir una amplitud de intervalo que no sea ni muy amplia ni muy estrecha. En la práctica, determinamos de ordinario la amplitud del intervalo dividiendo la distribución en un total de 10 a 20 intervalos. A lo largo de los años se ha demostrado que este rango de intervalos funciona bien con la mayor parte de las distribuciones. Dentro de este rango, el número específico de intervalos que se utilice dependerá del número y el rango de los datos en bruto. Observe que mientras más intervalos utilice, tanto más estrecho será cada uno de éstos.

Elaboración de una distribución de frecuencias de datos agrupados

Los pasos para hacer una distribución de frecuencias para datos agrupados son los siguientes:

1. Encontrar el rango de los datos.
2. Determinar la amplitud de cada intervalo de clase (i).
3. Hacer una lista con los límites de cada intervalo de clase, colocando en la parte inferior de la misma al intervalo que contenga el dato más pequeño (mínimo).
Hacer una lista con los límites de cada intervalo de clase, colocando en la parte inferior de la misma al intervalo que contenga el dato más pequeño (mínimo).
4. Contar los datos en bruto contenidos en los intervalos de clase correspondientes.
5. Sumar las cuentas de cada intervalo para obtener la frecuencia del intervalo.

Apliquemos estos pasos a los datos de la tabla 3.1

1. *Determinación del rango.*

$$\text{Rango} = \text{Dato máximo} - \text{dato mínimo} = 99 - 46 = 53$$

2. *Determinación de la amplitud del intervalo (i).* Supongamos que nos interesa agrupar los datos en 10 intervalos de clase aproximadamente.

$$i = \frac{\text{Rango}}{\text{Número de intervalos de clase}} = \frac{53}{10} = 5.3 \quad (\text{se redondea a } 5)$$

Cuando i tenga un residuo decimal, deberemos aplicar la regla y redondear i al mismo número de cifras decimales que aparece en los datos en bruto. Así pues, i se redondea a 5.

3. *La lista de los intervalos.* Comenzaremos con el intervalo inferior. El primer paso consiste en determinar el límite inferior de dicho intervalo. Hay dos requisitos que deben cumplirse:

- a. El límite inferior de este intervalo debe ser tal que el intervalo contenga al dato mínimo.
- b. Se acostumbra hacer que el límite inferior de este intervalo sea exactamente divisible entre i .

En virtud de estos dos requisitos, al límite inferior se le asigna el valor del dato más pequeño (o mínimo) de la distribución si éste es exactamente divisible entre i . Si no lo es, entonces el límite inferior es asignado al siguiente valor más pequeño que sea exactamente divisible entre i . En nuestro ejemplo, el límite inferior del intervalo más bajo comienza en 45, ya que la calificación mínima (46) no es exactamente divisible entre 5.

Una vez que hemos determinado el límite inferior del primer intervalo, ya podemos hacer la lista de todos los intervalos. Como cada intervalo tiene 5 unidades de amplitud, el intervalo inferior abarca de 45 a 49. Aunque podría parecer que este intervalo tiene sólo 4 unidades de amplitud, en realidad tiene 5. Si tiene alguna duda, bastará que cuente las unidades (45, 46, 47, 48, 49). Al listar los demás intervalos, debemos cerciorarnos de que los intervalos son continuos y mutuamente excluyentes. Con esta expresión (mutuamente excluyentes) queremos decir que los intervalos deben ser tales, que ningún dato podrá quedar incluido de manera válida en más de un intervalo. Aplicando estas reglas obtenemos los intervalos que aparecen en la tabla 3.4. Observe que, según la exposición sobre el análisis de los límites realera presentada en el capítulo 2, los intervalos de clase que aparecen en la primera columna representan límites aparentes. Los límites reales se encuentran en la segunda columna. La forma usual de hacerlo consiste en listar sólo los límites aparentes de cada intervalo y omitir de la lista los límites reales. Seguiremos esta práctica en los ejemplos subsecuentes.

4. *Conteo de los datos.* A continuación, introducimos los datos en bruto en los intervalos de clase adecuados. Este conteo es un procedimiento por el cual se repasa toda la distribución de manera sistemática, de modo que por cada dato en bruto se escribe una raya vertical junto al intervalo que contiene a dicho dato. Así pues, para 95 (la primera calificación de la tabla 3.1) colocamos una raya en el intervalo 95-99. Hemos seguido este procedimiento con todas las calificaciones y de este modo obtuvimos los resultados que aparecen en la tabla 3.4.
5. *Suma de frecuencias.* Por último, las rayas del recuento se convierten en frecuencias, al sumar todas las que corresponden a cada intervalo. Estas frecuencias se muestran también en la tabla 3.4.

tabla 3.4 Construcción de una distribución de frecuencias para las calificaciones agrupadas

Intervalo de clase	Límites reales	Conteo	f
95-99	94.5-99.5	(calificación de 95) → <i> </i>	4
90-94	89.5-94.5	<i> </i> /	5
85-89	84.5-89.5	<i> </i> //	7
80-84	79.5-84.5	<i> </i> <i> </i>	10
75-79	74.5-79.5	<i> </i> <i> </i> <i> </i> /	16
70-74	69.5-74.5	<i> </i> <i> </i>	9
65-69	64.5-69.5	<i> </i> //	7
60-64	59.5-64.5	<i> </i>	4
55-59	54.5-59.5	<i> </i>	4
50-54	49.5-54.5	//	2
45-49	44.5-49.5	/	1
			$N = 70$

Tratemos de resolver un problema en plan de práctica. A partir de los siguientes 90 datos, haga una distribución de frecuencias de datos agrupados que tenga 12 intervalos aproximadamente.

112	68	55	33	72	80	35	55	62
102	65	104	51	100	74	45	60	58
92	44	122	73	65	78	49	61	65
83	76	95	55	50	82	51	138	73
83	72	89	37	63	95	109	93	65
75	24	60	43	130	107	72	86	71
128	90	48	22	67	76	57	86	114
33	54	64	82	47	81	28	79	85
42	62	86	94	52	106	30	117	98
58	32	68	77	28	69	46	53	38

SOLUCIÓN

1. Encuentre el rango. Rango = dato máximo – dato mínimo = 138 – 22 = 116.
2. Determine la amplitud del intervalo (*i*):

$$i = \frac{\text{Rango}}{\text{Número de intervalos}} = \frac{116}{12} = 9.7 \quad i \text{ se redondea a } 10$$

3. Haga una lista de los límites de cada intervalo de clase. En virtud de que el dato mínimo de la distribución (22) no es divisible de manera exacta entre *i*, el límite inferior del primer intervalo es 20. ¿Por qué 20? Porque es el siguiente valor inferior que es exactamente divisible entre 10. Los límites de cada intervalo de clase aparecen listados en la tabla 3.5.
4. Registre los datos en bruto en los intervalos de clase adecuados. Esto se llevó a cabo en la tabla 3.5.
5. Sume los registros de cada intervalo para obtener su frecuencia. Esto se llevó a cabo en la tabla 3.5.

tabla 3.5 Distribución de frecuencias de datos agrupados para el problema de práctica 3.1

Intervalo de clase	Registro	<i>f</i>	Intervalo de clase	Registro	<i>f</i>
130–139	//	2	70–79	/// /// ///	13
120–129	//	2	60–69	/// /// ///	15
110–119	///	3	50–59	/// /// //	12
100–109	/// /	6	40–49	/// ///	8
90–99	/// //	7	30–39	/// //	7
80–89	/// /// /	11	20–29	////	4
					N = 90



A partir de los 130 datos siguientes, haga una distribución de frecuencias para datos agrupados con 15 intervalos aproximadamente.

1.4	2.9	3.1	3.2	2.8	3.2	3.8	1.9	2.5	4.7
1.8	3.5	2.7	2.9	3.4	1.9	3.2	2.4	1.5	1.6
2.5	3.5	1.8	2.2	4.2	2.4	4.0	1.3	3.9	2.7
2.5	3.1	3.1	4.6	3.4	2.6	4.4	1.7	4.0	3.3
1.9	0.6	1.7	5.0	4.0	1.0	1.5	2.8	3.7	4.2
2.8	1.3	3.6	2.2	3.5	3.5	3.1	3.2	3.5	2.7
3.8	2.9	3.4	0.9	0.8	1.8	2.6	3.7	1.6	4.8
3.5	1.9	2.2	2.8	3.8	3.7	1.8	1.1	2.5	1.4
3.7	3.5	4.0	1.9	3.3	2.2	4.6	2.5	2.1	3.4
1.7	4.6	3.1	2.1	4.2	4.2	1.2	4.7	4.3	3.7
1.6	2.8	2.8	2.8	3.5	3.7	2.9	3.5	1.0	4.1
3.0	3.1	2.7	2.2	3.1	1.4	3.0	4.4	3.3	2.9
3.2	0.8	3.2	3.2	2.9	2.6	2.2	3.6	4.4	2.2

SOLUCIÓN

1. Encuentre el rango. Rango = dato máximo – dato mínimo = 5.0 – 0.6 = 4.4.
2. Determine la amplitud del intervalo (i):

$$i = \frac{\text{Rango}}{\text{Número de intervalos}} = \frac{4.4}{15} = 0.29 \quad i \text{ se redondea a } 0.3$$

3. Liste los límites de cada intervalo de clase. Como el dato mínimo en la distribución (0.6) se puede dividir de manera exacta entre i , ésta se convierte en el límite inferior del primer intervalo. Los límites de cada intervalo aparecen en la tabla 3.6.
4. Registre los datos en bruto en los intervalos de clase adecuados. Esto se llevó a cabo en la tabla 3.6.
5. Suma los registros de cada intervalo para obtener su frecuencia. Esto se llevó a cabo en la tabla 3.6. Observe que en virtud de que la unidad de medición más pequeña en los datos en bruto es de 0.1, los límites reales para cualquier dato están a ± 0.05 del dato mencionado. Así, los límites reales para el intervalo 4.8–5.0 son 4.75–5.05.

tabla 3.6 Distribución de frecuencias de datos agrupados para el problema de práctica 3.2

Intervalo de clase	Registro	<i>f</i>	Intervalo de clase	Registro	<i>f</i>
4.8–5.0	//	2	2.4–2.6	/// ///	10
4.5–4.7	///	5	2.1–2.3	/// ////	9
4.2–4.4	/// ///	8	1.8–2.0	/// ////	9
3.9–4.1	/// /	6	1.5–1.7	/// ///	8
3.6–3.8	/// /// /	11	1.2–1.4	/// /	6
3.3–3.5	/// /// /// /	16	0.9–1.1	////	4
3.0–3.2	/// /// /// /	16	0.6–0.8	///	3
2.7–2.9	/// /// /// //	17			<u>3</u>
					<i>N</i> = 130

Frecuencia relativa, frecuencia acumulada y distribuciones de porcentajes acumulados

A menudo se desea expresar los datos de una distribución de frecuencias como una frecuencia relativa, como una frecuencia acumulada o como una distribución de porcentajes acumulados.

- definiciones**
- Una **distribución de frecuencias relativas** indica la proporción del número total de datos que aparecen en cada intervalo.
 - Una **distribución de frecuencias acumuladas** indica el número de datos que están por debajo del límite real superior de cada intervalo.
 - Una **distribución de porcentajes acumulados** indica el porcentaje de datos que están por debajo del límite real superior de cada intervalo.

La tabla 3.7 muestra la distribución de frecuencias de las calificaciones del examen de estadística, expresadas como frecuencias relativas, frecuencias acumuladas y distribución de porcentajes acumulados. Para convertir una distribución de frecuencias en una distribución de frecuencias relativas, dividimos la frecuencia de cada intervalo entre el número total de calificaciones. Así,

$$\text{Relativa } f = \frac{f}{N}$$

Por ejemplo, para determinar la frecuencia relativa para el intervalo 45-49, dividimos su frecuencia (1) entre la cantidad total de calificaciones (70) disponibles. Por consiguiente, la frecuencia relativa que corresponde a este intervalo es de $\frac{1}{70} = 0.01$. La frecuencia relativa resulta útil porque nos indica la proporción de calificaciones que está contenida en el intervalo.

La frecuencia acumulada para cada intervalo se calcula sumando la frecuencia de ese intervalo a las frecuencias de todos los intervalos de clase que están debajo de él. Así, la frecuencia acumulada para el intervalo 60–64 = 4 + 4 + 2 + 1 = 11.

tabla 3.7 Distribuciones de frecuencias relativas, frecuencias acumuladas y porcentajes acumulados para los datos agrupados en la tabla 3.4

Intervalo de clase	f	Relativa f	Acumulada f	Acumulado %
95-99	4	0.06	70	100
90-94	6	0.09	66	94.29
85-89	7	0.10	60	85.71
80-84	10	0.14	53	75.71
75-79	16	0.23	43	61.43
70-74	9	0.13	27	38.57
65-69	7	0.10	18	25.71
60-64	4	0.06	11	15.71
55-59	4	0.06	7	10.00
50-54	2	0.03	3	4.29
45-49	1	0.01	1	1.43
	70	1.00		

El porcentaje acumulado de cada intervalo se determina convirtiendo las frecuencias acumuladas en porcentajes acumulados. La ecuación que se usa para esto es:

$$\% \text{ acumulado} = \frac{\text{Acumulada } f}{N} \times 100$$

Para el intervalo 60-64 tenemos que:

$$\% \text{ acumulado} = \frac{\text{Acumulada } f}{N} \times 100 = \frac{11}{70} \times 100 = 15.71\%$$

Las distribuciones de frecuencias acumuladas y de porcentajes acumulados son útiles para determinar percentiles y rangos percentiles.



PERCENTILES

Los percentiles son medidas de posición relativa. Se utilizan asiduamente en la educación con el fin de comparar el rendimiento de un individuo con el rendimiento de un grupo de referencia.

definición ■ *Un percentil o punto percentil es el valor sobre la escala de medición, debajo del cual se encuentra un porcentaje dado de los datos incluidos en la distribución.*

Así, el sexagésimo punto percentil es el valor sobre la escala de medida, debajo del cual está 60% de los datos de la distribución.

Cálculo de los puntos percentiles

Suponga usted que estamos interesados en calcular el quincuagésimo punto percentil para las calificaciones del examen de estadística. Estas calificaciones han sido presentadas en la tabla 3.8 como distribuciones de frecuencias acumuladas y de porcentajes acumulados. Emplearemos el símbolo P_{50} para representar al quincuagésimo punto percentil. ¿Qué entendemos por el quincuagésimo punto percentil? A partir de la definición de punto percentil, P_{50} es el valor de la escala debajo del cual se encuentra 50% de las calificaciones. Como quiera que hay en total 70 calificaciones en la distribución, P_{50} debe ser el valor debajo del cual se encuentran 35 calificaciones (50% de 70 es 35). Al observar la columna de las frecuencias acumuladas y desplazarnos de abajo hacia arriba, vemos que P_{50} se localiza en el intervalo 75-79. Sin embargo, en este momento no sabemos qué valor de la escala debemos asignar a P_{50} . Lo único que sabemos es que le corresponde algún valor dentro de los límites reales del intervalo 75-79, que son 74.5 y 79.5. Para determinar el punto en el cual se localiza P_{50} , suponemos que todas las calificaciones están distribuidas de manera uniforme en todo el intervalo.

En vista de que 27 de las calificaciones se encuentran debajo de 74.5, tenemos que desplazarnos dentro del intervalo hasta obtener 8 calificaciones adicionales (véase la figura 3.1). Puesto que existen 16 calificaciones en el intervalo y éste tiene 5 unidades de amplitud en la escala, cada calificación del intervalo ocupa $\frac{5}{16}$ de la unidad. Para obtener 8 calificaciones adicionales, tenemos que desplazarnos $\frac{5}{16} \times 8 = 2.5$ unidades dentro del intervalo. Al sumar 2.5 al límite inferior de 74.5, llegamos a P_{50} . Así,

$$P_{50} = 74.5 + 2.5 = 77.0$$

Para determinar cualquier punto percentil, siga estos pasos:

1. *Determine la frecuencia de los datos que están debajo del punto percentil. Simbolizaremos esta frecuencia como “ f_p acumulada”.*

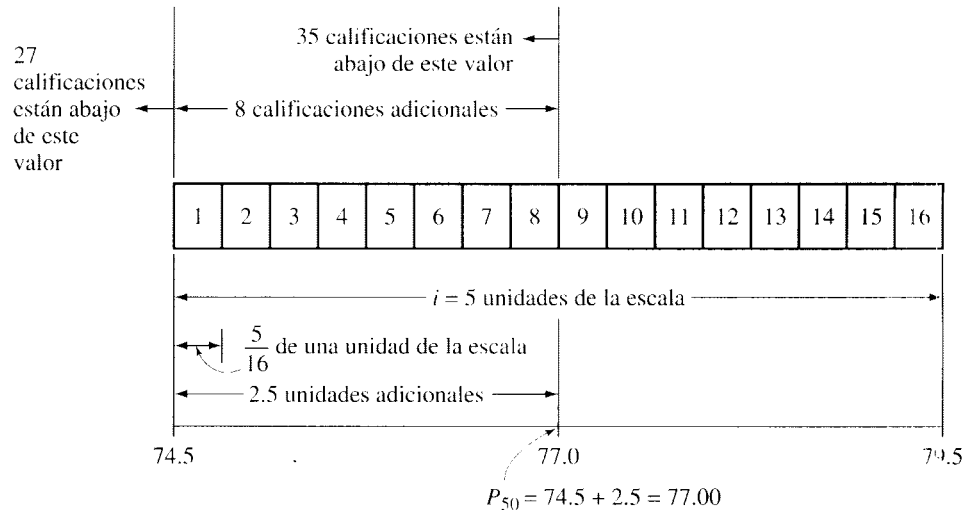
$$f_p \text{ acumulada} = (\% \text{ de los datos que están abajo del punto percentil}) \times N$$

$$f_p \text{ acumulada para } P_{50} = 50\% \times N = (0.50) \times 70 = 35$$

2. *Determine el límite real inferior del intervalo que contiene al punto percentil. Llamaremos a este límite real inferior X_L . Al conocer el número de datos que es-*

tabla 3.8 Cálculo de puntos percentiles a partir de las calificaciones de la tabla 3.1

Intervalo de clase	f	f acumulada	% acumulado	Cálculo del percentil
95-99	4	70	100	Punto percentil = $X_L + (if_i)(f_p \text{ acumulada} - f_l \text{ acumulada})$
90-94	6	66	94.29	
85-89	7	60	85.71	
80-84	10	53	75.71	
75-79	16	43	61.43	$P_{50} = 74.5 + \frac{5}{16} \cdot 235 = 77.00$
70-74	9	27	38.57	
65-69	7	18	25.71	$P_{20} = 64.5 + \frac{5}{16} \cdot 214 = 66.64$
60-64	4	11	15.71	
55-59	4	7	10.00	
50-54	2	3	4.29	
45-49	1	1	1.43	



Tomado de *Statistical Reasoning in Psychology and Education* por E. W. Minium. Derechos reservados © 1978 John Wiley & Sons, Inc. Adaptado con autorización.

figura 3.1 Determinación del valor de P_{50} en la escala, para las calificaciones del examen de estadística.

tán abajo del punto percentil podemos localizar el intervalo que contiene al punto percentil, comparando la f_p acumulada con la frecuencia acumulada de cada intervalo. Una vez localizado el intervalo que contiene al punto percentil, podemos encontrar de inmediato su límite real inferior, X_L . En este ejemplo, el intervalo que contiene a P_{50} es 75–79 y su límite real inferior, X_L , es 74.5

- Determine el número de datos adicionales que deben considerarse en el intervalo para alcanzar el punto percentil.

$$\text{Número de datos adicionales} = f_p \text{ acumulada} - f_L \text{ acumulada}$$

donde f_L acumulada = frecuencia de los datos que están abajo del límite real inferior del intervalo que contiene al punto percentil.

Para el ejemplo anterior,

$$\begin{aligned} \text{Número de datos adicionales} &= f_p \text{ acumulada} - f_L \text{ acumulada} \\ &= 35 - 27 \\ &= 8 \end{aligned}$$

- Determine el número de unidades adicionales que debemos desplazar en el intervalo para obtener el número de datos adicionales.

$$\begin{aligned} \text{Unidades adicionales} &= (\text{Número de unidades por dato}) \\ &\quad \times \text{Número de datos adicionales} \\ &= (i/f_i) \times \text{Número de datos adicionales} \\ &= \left(\frac{5}{16}\right) \times 8 \\ &= 2.5 \end{aligned}$$

Observe que:

f_i es el número de datos en el intervalo i

i/f_i nos proporciona el número de unidades por dato para dicho intervalo

5. *Determine el punto percentil.* Logramos esto sumando las unidades adicionales al límite real inferior del intervalo que contiene al punto percentil.

$$\begin{aligned} \text{Punto percentil} &= X_L + \text{Unidades adicionales} \\ P_{50} &= 74.5 + 2.5 = 77.00 \end{aligned}$$

Podemos escribir estos pasos en forma de ecuación. Por lo tanto,

$$\text{Punto percentil} = X_L + (i/f_i)(f_P \text{ acumulada} - f_L \text{ acumulada})^* \quad \begin{array}{l} \text{ecuación para} \\ \text{calcular un} \\ \text{punto percentil} \end{array}$$

donde X_L = Valor del límite real inferior del intervalo que contiene al punto percentil

f_P acumulada = frecuencia de los datos que están abajo del punto percentil

f_L acumulada = frecuencia de los datos que están abajo del límite real inferior del intervalo que contiene al punto percentil

f_i = frecuencia del intervalo que contiene al punto percentil

i = amplitud del intervalo

Si utilizamos esta ecuación para calcular P_{50} obtenemos:

$$\begin{aligned} \text{Punto percentil} &= X_L + (i/f_i)(f_P \text{ acumulada} - f_L \text{ acumulada}) \\ P_{50} &= 74.5 + (\frac{5}{16})(35 - 27) \\ &= 74.5 + 2.5 = 77.00 \end{aligned}$$



Intentemos resolver otro problema. Esta vez calcularemos P_{20} , es decir, el valor por debajo del cual se encuentra 20% de las calificaciones.

En términos de frecuencia acumulada, P_{20} es el valor por debajo del cual se encuentran 14 calificaciones (20% de 70 es 14). Podemos ver en la tabla 3.8 que P_{20} está en el intervalo 65-69. En vista de que 11 calificaciones se encuentran abajo del valor 64.5, necesitamos 3 calificaciones adicionales. Considerando que hay 7 calificaciones en el intervalo y que éste tiene 5 unidades de amplitud, debemos movernos $\frac{5}{7} \times 3 = 2.14$ unidades en el intervalo. Así,

$$P_{20} = 64.5 + 2.14 = 66.64$$

También podríamos haber hallado P_{20} en forma directa, utilizando la ecuación para el punto percentil. De esta manera,

$$\begin{aligned} \text{Punto percentil} &= X_L + (i/f_i)(f_P \text{ acumulada} - f_L \text{ acumulada}) \\ P_{20} &= 64.5 + (\frac{5}{7})(14 - 11) \\ &= 64.5 + 2.14 = 66.64 \end{aligned}$$

*Estoy en deuda con LeAnn Wilson, quien sugirió esta forma de la ecuación.

Intentemos resolver un problema más. Esta vez calcularemos P_{75} que es el valor de la escala abajo del cual se encuentra 75% de las calificaciones.

En términos de frecuencia acumulada, P_{75} es el valor de la escala por debajo del cual se localizan 52.5 calificaciones ($f_p = 75\%$ acumulado es igual a 75% de 70, que es 52.5). A partir de la tabla 3.8, vemos que P_{75} está en el intervalo 80-84. Como 43 calificaciones se encuentran abajo del límite inferior de este intervalo (79.5), necesitamos agregar a 79.5 el número adecuado de unidades de la escala para tener $52.5 - 43 = 9.5$ calificaciones adicionales. Como existen 10 calificaciones en el intervalo y éste tiene 5 unidades de amplitud, tendremos que desplazarlos $\frac{5}{10} \times 9.5 = 4.75$ unidades en el intervalo. Así pues,

$$P_{75} = 79.5 + 4.75 = 84.25$$

P_{75} también se puede determinar en forma directa, mediante la ecuación para el punto percentil. Así,

$$\begin{aligned} \text{Punto percentil} &= X_L + (i/f_i)(f_p \text{ acumulada} - f_L \text{ acumulada}) \\ P_{75} &= 79.5 + \left(\frac{5}{10}\right)(52.5 - 43) \\ &= 79.5 + 4.75 \\ &= 84.25 \end{aligned}$$



RANGOS PERCENTILES

A veces queremos conocer el rango percentil de un dato en bruto. Por ejemplo, como la calificación que usted obtuvo en el examen de estadística fue de 86, le podría ser útil conocer el rango percentil de este valor.

definición ■ *El rango percentil de un dato es el porcentaje de datos que tienen valores más bajos que el dato en cuestión.*

Cálculo de rangos percentiles

Esta situación es exactamente la opuesta a la anterior, en la que calculábamos un punto percentil. Ahora conocemos el dato y necesitamos calcular el porcentaje de datos que están por debajo de él. Una vez más, debemos suponer que los datos dentro de la totalidad de cualquier intervalo están distribuidos de manera uniforme. En la columna de intervalos de clase de la tabla 3.9 podemos ver que la calificación de 86 se encuentra en el

intervalo 85-89. Existen 53 calificaciones que están abajo de 84.5, el límite inferior de este intervalo. En virtud de que existen 7 calificaciones en el intervalo y éste tiene 5 unidades de amplitud, hay $\frac{7}{5}$ calificaciones por cada unidad de la escala. Entre una calificación de 86 y 84.5, existen $(\frac{7}{5})(86 - 84.5) = 2.1$ calificaciones adicionales. Por lo tanto, hay en total $53 + 2.1 = 55.1$ calificaciones que están por debajo de 86. Como quiera que la distribución tiene 70 calificaciones, el rango percentil de 86 = $(\frac{55.1}{70}) \times 100 = 78.71$.

Resumiremos estas operaciones en la siguiente ecuación:

$$\text{Rango percentil} = \frac{f_L \text{ acumulada} + (f_i/i)(X - X_L)}{N} \times 100 \quad \text{ecuación para calcular el rango percentil}$$

donde f_L acumulada = frecuencia de los datos que están abajo del límite real inferior del intervalo que contiene al dato X

X = dato cuyo rango percentil se desea determinar

X_L = valor de la escala que corresponde al límite real inferior del intervalo que contiene al dato X

i = amplitud del intervalo

f_i = frecuencia del intervalo que contiene al dato X

N = Número total de datos en bruto

Al utilizar esta ecuación para determinar el rango percentil de 86, obtenemos

$$\begin{aligned} \text{Rango percentil} &= \frac{f_L \text{ acumulada} + (f_i/i)(X - X_L)}{N} \times 100 \\ &= \frac{53 + (\frac{7}{5})(86 - 84.5)}{70} \times 100 \\ &= \frac{53 + 2.1}{70} \times 100 \\ &= \frac{55.1}{70} \times 100 \\ &= 78.71 \end{aligned}$$



Resolvamos otro problema para practicar. Determine el rango percentil de 59.

La calificación de 59 se encuentra en el intervalo 55-59. Existen 3 calificaciones que están por debajo de 54.5. En virtud de que existen 4 calificaciones dentro del intervalo, hay $(\frac{4}{3})(59 - 54.5) = 3.6$ calificaciones dentro del intervalo que se encuentran abajo de 59. En total, hay $3 + 3.6 = 6.6$ calificaciones por debajo de 59. Así pues, el rango percentil de 59 = $(\frac{6.6}{70}) \times 100 = 9.43$.

(continúa)

La solución se presenta en forma de ecuación en la tabla 3.9.

tabla 3.9 Cálculo del rango percentil para los datos de la tabla 3.1

Intervalo de clase	f	acumu- lada	% acumu- lado	Cálculo del rango percentil
95-99	4	70	100	Rango percentil = $\frac{f_L \text{ acumulada} + (f_i/i)(X - X_{L-})}{N} \times 100$
90-94	6	66	94.29	
85-89	7	60	85.71	
80-84	10	53	75.71	
75-79	16	43	61.43	
70-74	9	27	38.57	
65-69	7	18	25.71	
60-64	4	11	15.71	
55-59	4	7	10.00	Rango percentil de 59 = $\frac{3 + (\frac{3}{5})(59 - 54.5)}{70} \times 100$
50-54	2	3	4.29	= 9.43
45-49	1	1	1.43	



Vamos a resolver otro problema de práctica. Con la distribución de frecuencias de las calificaciones agrupadas de la tabla 3.5, determine el rango percentil de una calificación de 117.

La calificación de 117 se encuentra en el intervalo 110-119. El límite inferior de este intervalo es 109.5. Existen $6 + 7 + 11 + 13 + 15 + 12 + 8 + 7 + 4 = 83$ calificaciones por debajo de 109.5. Puesto que hay 3 calificaciones dentro del intervalo y éste tiene 10 unidades de amplitud, existen $(\frac{3}{10})(117 - 109.5) = 2.25$ calificaciones dentro del intervalo que se localizan abajo de la calificación de 117. En total, existen $83 + 2.25 = 85.25$ calificaciones que se encuentran por debajo de la calificación de 117. Así pues, el rango percentil de 117 = $(\frac{85.25}{90}) \times 100 = 94.72$.

También podríamos haber resuelto este problema mediante la ecuación del rango percentil. Así,

$$\begin{aligned} \text{Rango percentil} &= \frac{f_L \text{ acumulada} + (f_i/i)(X - X_{L-})}{N} \times 100 \\ &= \frac{83 + (\frac{3}{10})(117 - 109.5)}{90} \times 100 \\ &= 94.72 \end{aligned}$$



GRAFICACIÓN DE LAS DISTRIBUCIONES DE FRECUENCIAS

Las distribuciones de frecuencias se representan a menudo por medio de gráficas y no en forma de tablas. En virtud de que la gráfica se basa por completo en los datos de una tabla, no contiene información nueva. Sin embargo, una gráfica presenta los datos de una manera visual, lo cual frecuentemente facilita la tarea de observar las características importantes de los datos. Al escribir esta sección, hemos supuesto que usted está familiarizado con la elaboración de gráficas. Pues bien, aun en este caso, valdrá la pena repasar unos cuantos puntos importantes:

1. Una gráfica tiene dos ejes: uno vertical y el otro horizontal. El eje vertical es el eje de las *ordenadas* o eje Y , y el horizontal es el eje de las *abscisas* o eje X .
2. La variable independiente se coloca muy a menudo en el eje X y la variable dependiente en el eje Y . Al trazar la gráfica de una distribución de frecuencias, los valores de los datos se colocan de ordinario en el eje X y la frecuencia de los valores de los datos se registran en el eje Y .
3. Es necesario elegir las unidades apropiadas para registrar los datos en los ejes.
4. Para evitar la distorsión de los datos, se acostumbra colocar la intersección de los dos ejes en el punto cero y después se eligen sus escalas, de manera que la altura de los datos graficados sea más o menos de las tres cuartas partes de la anchura. La figura 3.2 muestra la forma en que la infracción de esta regla puede deformar la impresión que la gráfica produce. La figura muestra dos gráficas basadas en los mismos datos, a saber, el número de estudiantes inscritos en una universidad grande durante los años 1990-2002. En la parte (a) se ha respetado la regla mencionada con anterioridad. En la parte (b), la escala correspondiente al eje de las ordenadas no comienza en cero y se muestra considerablemente expandida, en comparación con la parte (a). Las impresiones que producen una y otra gráfica son muy diferentes. La parte (a) produce la impresión correcta de que la población estudiantil fue muy estable, mientras que la parte (b) distorsiona en gran medida los datos y produce la impresión de que hubo grandes fluctuaciones en la población universitaria.
5. Por lo general, la intersección de los dos ejes se encuentra en cero, en ambas escalas. Cuando no es así, esto se indica simbolizando una ruptura en el eje correspondiente, cerca de la intersección. Por ejemplo, en la figura 3.4 aparece una ruptura en el eje horizontal para indicar que una parte de la escala ha sido suprimida.
6. Cada eje debe tener un letrero y el título de la gráfica debe ser breve y explícito.

Al graficar distribuciones de frecuencias, se utilizan principalmente cuatro tipos de gráficas: la *gráfica de barras*, el *histograma*, el *polígono de frecuencias* y la *curva de porcentajes acumulados*.

La gráfica de barras

Se acostumbra graficar las distribuciones de frecuencias de los datos nominales u ordinales por medio de una gráfica de barras. Este tipo de gráfica se ilustra en la figura 3.3. Se traza una barra para cada categoría y la altura de la barra representa la frecuencia o el número de miembros de esa categoría. Como no existe una relación numérica entre las categorías de los datos nominales, podemos organizar los diversos grupos en cualquier orden a lo largo del eje horizontal. En la figura 3.3 están ordenados de izquierda a derecha, según la magnitud de la frecuencia de cada categoría. Ob-

serve que, en este tipo de gráfica, las barras de cada categoría no se tocan entre sí. Esto hace aún más patente la ausencia de una relación cuantitativa entre las distintas categorías.

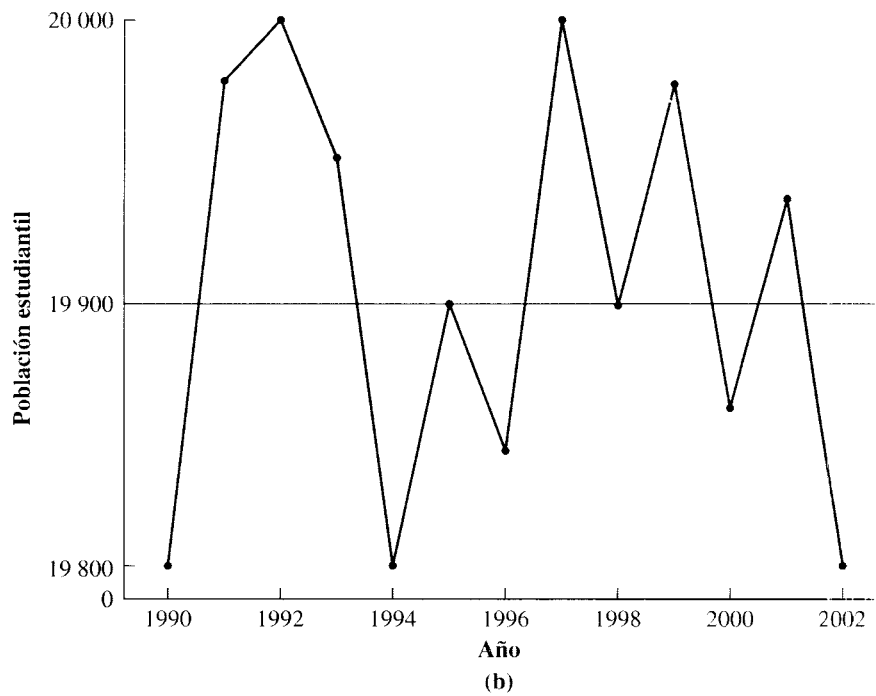
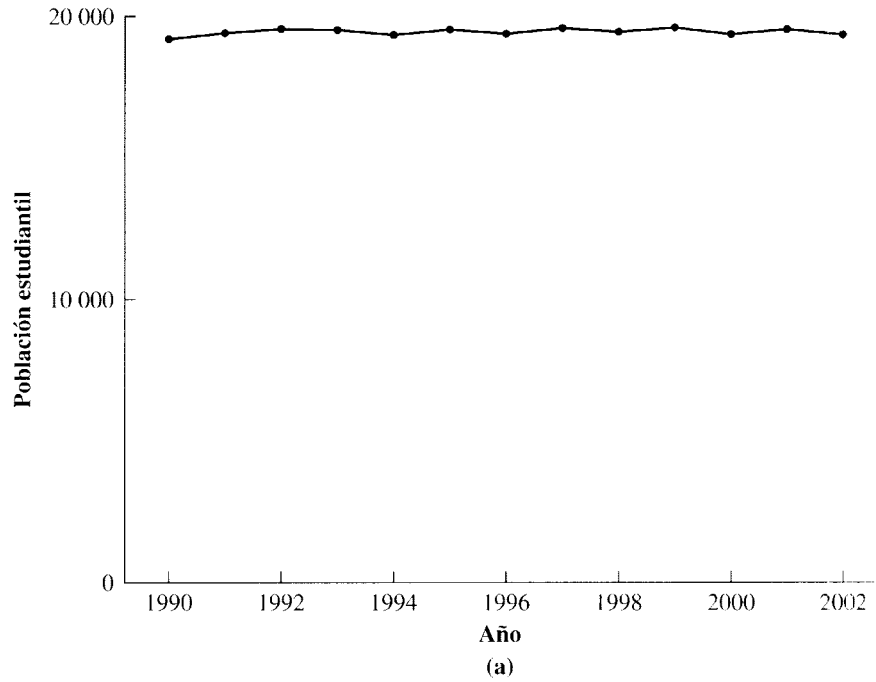


figura 3.2 Población estudiantil de una universidad grande entre 1990 y 2002.

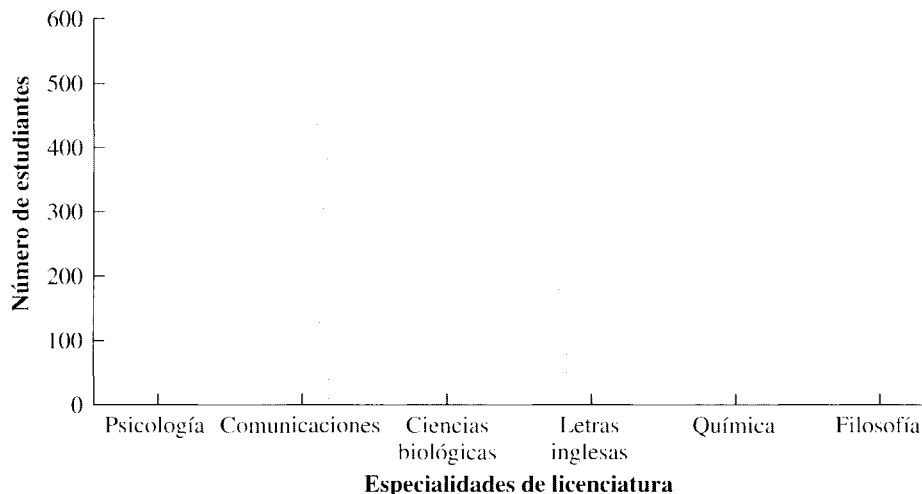


figura 3.3 Gráfica de barras: estudiantes inscritos en varias especialidades de licenciatura en una escuela superior de artes y ciencias.

El histograma

El histograma se utiliza para representar distribuciones de frecuencias constituidas por datos de intervalos o de razones (proporciones). Se asemeja a la gráfica de barras, salvo porque en el histograma se traza una barra para cada intervalo de clase. Los intervalos de clase se localizan sobre el eje horizontal, de modo que cada barra de clase comienza y termina en los límites reales del intervalo. La altura de la barra corresponde a la frecuencia del intervalo de clase representado. Como quiera que los intervalos son continuos, las barras verticales se deben tocar entre sí, en lugar de estar espaciadas como en las gráficas de barras. La figura 3.4 muestra las calificaciones del examen de

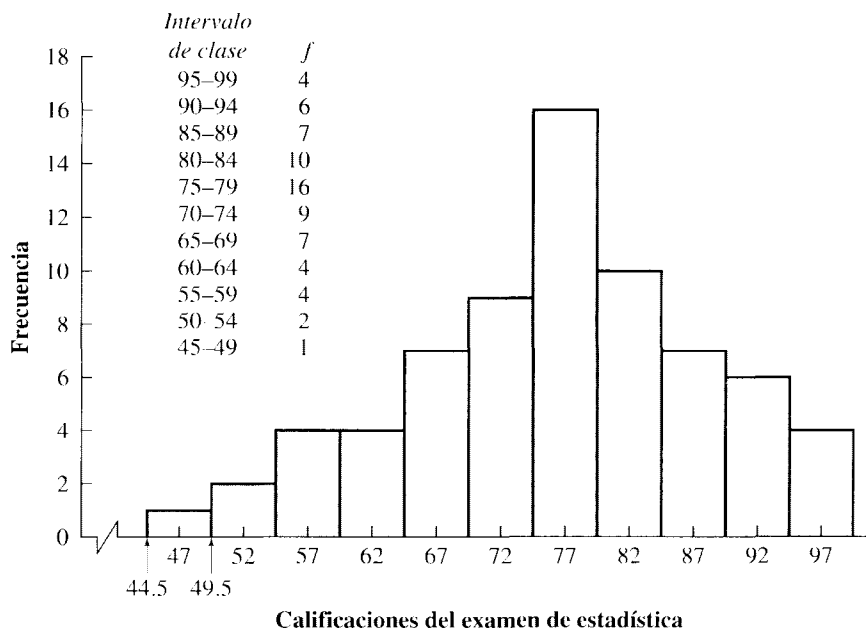


figura 3.4 Histograma: las calificaciones del examen de estadística de la tabla 3.4.

estadística (tabla 3.4) representadas mediante un histograma. Observe que se acostumbra marcar el punto medio de cada intervalo de clase sobre la abscisa. Las calificaciones agrupadas aparecen de nuevo en la figura para comodidad de usted.

El polígono de frecuencias

El polígono de frecuencias también se utiliza para representar datos de intervalo o de razón. El eje horizontal es idéntico al del histograma. Sin embargo, en este tipo de gráfica, en lugar de utilizar barras, se marca en la gráfica un punto que corresponde al punto medio de cada intervalo, a una altura correspondiente a la frecuencia del intervalo. A continuación se unen los puntos marcados, mediante líneas rectas. Por último, la línea que une los puntos se prolonga para que corte el eje horizontal en el punto medio de los dos intervalos de clase contiguos al primero y al último de los intervalos donde están contenidos los datos. Con esta manera de cerrar la figura con el eje horizontal, se forma un polígono del cual recibe su nombre esta gráfica. La figura 3.5 muestra las calificaciones de la tabla 3.4 como un polígono de frecuencias. La principal diferencia entre un histograma y un polígono de frecuencias es la siguiente: el histograma exhibe las calificaciones como si estuvieran distribuidas de manera uniforme en el intervalo, mientras que el polígono de frecuencias muestra las calificaciones como si todas estuviesen concentradas en el punto medio del intervalo. Algunos investigadores prefieren emplear el polígono de frecuencias para comparar la forma de dos o más distribuciones. Los polígonos de frecuencias también tienen el efecto de presentar los datos como si tuviesen una distribución continua, lo cual ocurre en realidad en muchos casos.

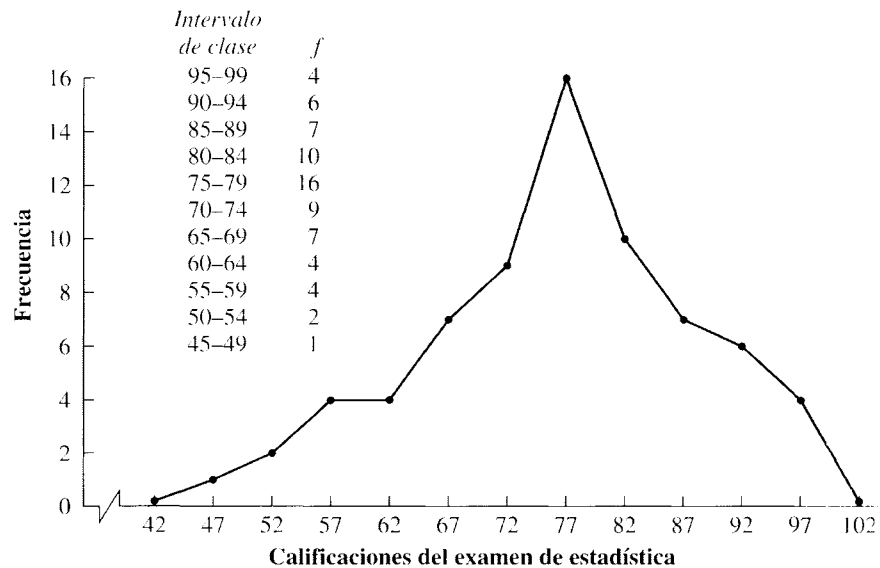


figura 3.5 Polígono de frecuencias: las calificaciones del examen de estadística de la tabla 3.4.

La curva de porcentaje acumulado

También las distribuciones de frecuencia acumulada y de porcentaje acumulado pueden presentarse en forma gráfica. Sólo ilustraremos la segunda de ellas, ya que las gráficas son esencialmente las mismas y las distribuciones de porcentajes acumulados se emplean más a menudo. Usted recordará que el porcentaje acumulado para un intervalo de clase indica el porcentaje de datos que están por debajo del límite real superior del intervalo. Así, el eje vertical utilizado para la curva de porcentajes acumulados se presenta en unidades

de porcentajes acumulados. En el eje horizontal, en lugar de dibujar los puntos en el punto medio de cada intervalo de clase, los marcamos en el límite real superior del intervalo. La figura 3.6 muestra las calificaciones de la tabla 3.7 representadas como una curva de porcentajes acumulados. Debe ser evidente que la curva de frecuencias acumuladas tendría la misma forma, siendo la única diferencia que esta última se grafica en unidades de frecuencias acumuladas y no de porcentajes acumulados. Los percentiles y los rangos percentiles se pueden leer directamente a partir de una curva de porcentajes acumulados. La curva de porcentajes acumulados recibe también el nombre de *ojiva*, con lo cual se indica que tiene la forma de una S.

Forma de las curvas de frecuencia

Las distribuciones de frecuencias pueden tener formas muy diversas. Algunas de las formas que se presentan con más frecuencia están en la figura 3.7. Las curvas se clasifican, en general, como *simétricas* o *asimétricas* (sesgadas).

definición ■ Una curva es **simétrica** si al doblarla a la mitad sus dos lados coinciden. Si una curva no es simétrica, entonces es **asimétrica**.

Las curvas que aparecen en la figura 3.7(a), (b) y (c) son simétricas. Las curvas (d), (e) y (f) son asimétricas. Si una curva es asimétrica, puede tener sesgo **positivo** o **negativo**.

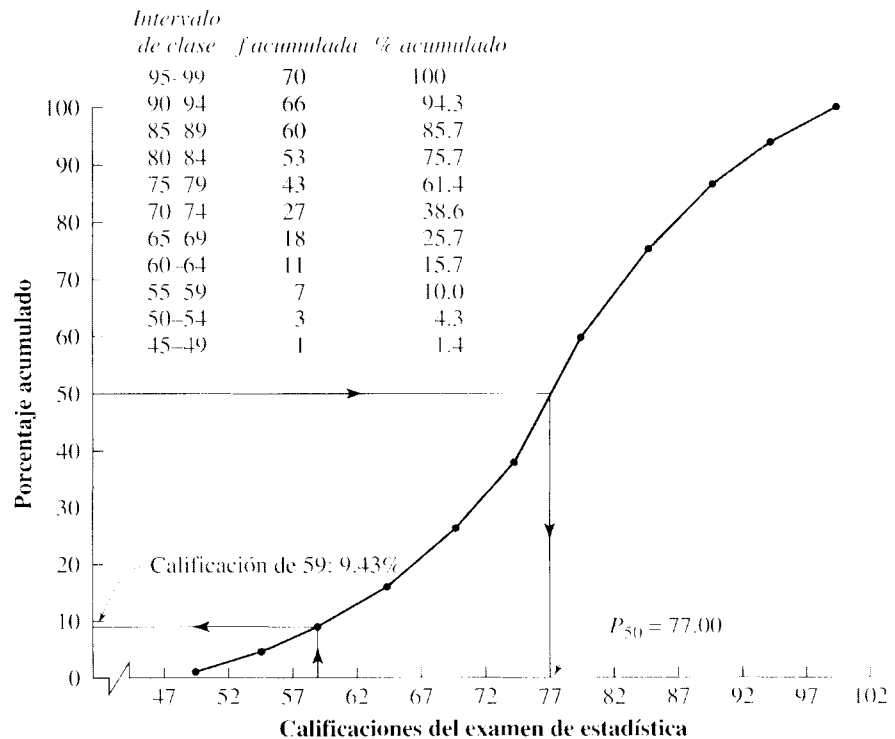


figura 3.6 Curva de porcentajes acumulativos: calificaciones del examen de estadística de la tabla 3.7.

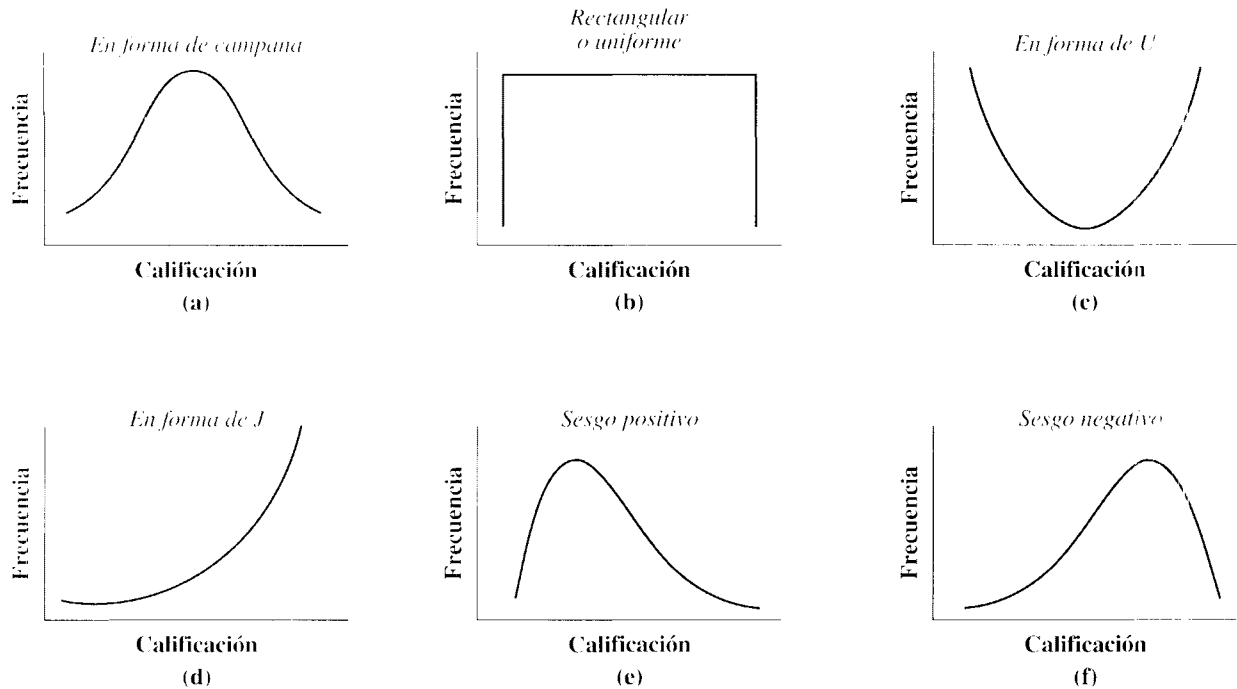


figura 3.7 Formas de las curvas de frecuencias.

definición ■ Cuando una curva está sesgada en **forma positiva**, la mayor parte de los datos se concentra en los valores bajos del eje horizontal y la curva disminuye hacia el extremo de los valores altos. Cuando una curva está sesgada en **forma negativa**, la mayor parte de los datos se concentra en los valores altos del eje horizontal y la curva se reduce hacia el extremo de los valores bajos.

La curva de la parte (e) está sesgada en forma positiva y la curva de la parte (f) está sesgada en forma negativa.

Las curvas de frecuencias son designadas a menudo según su forma. Así, a las curvas de las partes (a), (b), (c) y (d) se las denomina, respectivamente, curvas en forma de campana, rectangulares o uniformes, en forma de U y en forma de J.



ANÁLISIS EXPLORATORIO DE LOS DATOS

El análisis exploratorio de los datos es un procedimiento desarrollado en una época reciente. En él se utilizan diagramas fáciles de hacer que son muy útiles para resumir y describir los datos de una muestra. Uno de los diagramas más populares es el *diagrama de tallo y hojas*.

Diagramas de tallo y hojas

Los diagramas de tallo y hojas fueron desarrollados por primera vez en 1977 por John Tukey, cuando trabajaba en la Universidad de Princeton. Son una alternativa sencilla en lugar del histograma y resultan sumamente útiles para resumir y describir datos

cuando el número de éstos no es mayor de 100. Sin embargo, a diferencia de lo que ocurre con los histogramas, en los diagramas de tallo y hojas no se pierde ninguno de los datos originales. La figura 3.8 muestra un diagrama de tallo y hojas que muestra las calificaciones del examen de estadística de la tabla 3.1.

Al hacer un diagrama de tallo y hojas, cada dato es representado mediante un *tallo* y una *hoja*. El tallo se coloca a la izquierda de la línea vertical y la hoja a la derecha. Por ejemplo, los tallos y las hojas del primero y el último de los datos originales son:

tallo	hoja	tallo	hoja
9	5	6	7

En un diagrama de tallo y hojas, los tallos se colocan en orden vertical, de arriba abajo de la página, y las hojas se distribuyen en orden horizontal, de izquierda a derecha de la página. La hoja correspondiente a cada dato es, por lo general, el último dígito y el tallo está formado por los dígitos restantes. En ciertas ocasiones, la hoja contiene los dos últimos dígitos, según sea el rango de los datos.

Tome nota de que, en los diagramas de tallo y hojas, los valores del tallo pueden repetirse. En la figura 3.8, los valores del tallo se repiten dos veces. El resultado de es-

Calificaciones originales						
95	57	76	93	86	80	89
76	76	63	74	94	96	77
65	79	60	56	72	82	70
67	79	71	77	52	76	68
72	88	84	70	83	93	76
82	96	87	69	89	77	81
87	65	77	72	56	78	78
58	54	82	82	66	73	79
86	81	63	46	62	99	93
82	92	75	76	90	74	67

Diagrama de tallo y hojas	
4	6
5	2 4
5	6 6 7 8
6	0 2 3 3
6	5 5 6 7 7 8 9
7	0 0 1 2 2 2 3 4 4
7	5 6 6 6 6 6 6 7 7 7 7 8 8 9 9 9
8	0 1 1 2 2 2 2 3 4
8	6 6 7 7 8 9 9
9	0 2 3 3 3 4
9	5 6 6 9

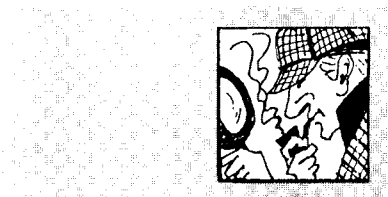
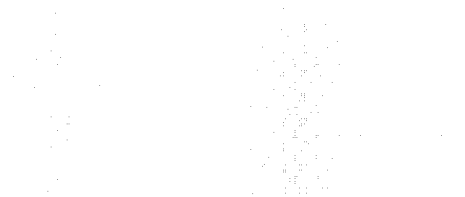
figura 3.8 Diagrama de tallo y hojas: calificaciones del examen de estadística de la tabla 3.1.

to consiste en que el tallo se alarga, es decir, se crean más intervalos y se dispersan los datos. A continuación presentaremos un diagrama de tallo y hojas para las calificaciones del examen de estadística, donde los valores del tallo se presentan sólo una vez.

4	6
5	2 4 6 6 7 8
6	0 2 3 3 5 5 6 7 7 8 9
7	0 0 1 2 2 2 3 4 4 5 6 6 6 6 6 6 7 7 7 7 8 8 9 9 9
8	0 1 1 2 2 2 2 2 3 4 6 6 7 7 8 9 9
9	0 2 3 3 3 4 5 6 6 9

Al registrar sólo una vez los valores del tallo, se obtiene como resultado menos intervalos, aunque más amplios, y cada uno generalmente con más datos. Esto hace que la presentación sea más “congestionada”. El hecho de que los valores del tallo deban aparecer una, dos o incluso más veces dependerá del rango de los datos.

Usted podrá observar que al hacer que el diagrama de tallo y hojas de la figura 3.8 gire 90° en dirección contraria a las manecillas del reloj, de manera que los tallos queden en la parte inferior, se obtiene un diagrama muy similar al histograma de la figura 3.4. Sin embargo, con el histograma perdemos los datos originales; en cambio, éstos se conservan con el diagrama de tallo y hojas.



Al cambiar la escala, cambia el cuento

En un artículo publicado en la sección de negocios de un periódico se analizó el alza de las tarifas de Puget Power & Light Company. La compañía estaba en malas condiciones financieras y

había propuesto un aumento más de tarifas en 1984 para tratar de resolver sus problemas. El tema era muy delicado porque las alzas de tarifas habían sido una plaga en la región, en fechas recientes, pues así se intentaba pagar las enormes pérdidas incurridas en la construcción de una planta de energía nuclear. La gráfica de la página siguiente fue publicada en

el artículo junto con la leyenda “Las tarifas de Puget Power han subido uniformemente durante los últimos 14 años”. ¿Nota usted algo peculiar en la gráfica?





Al cambiar la escala, cambia el cuento

(continuación)

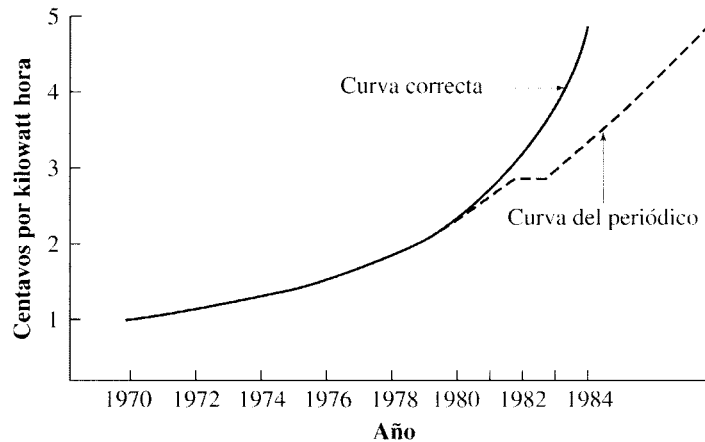
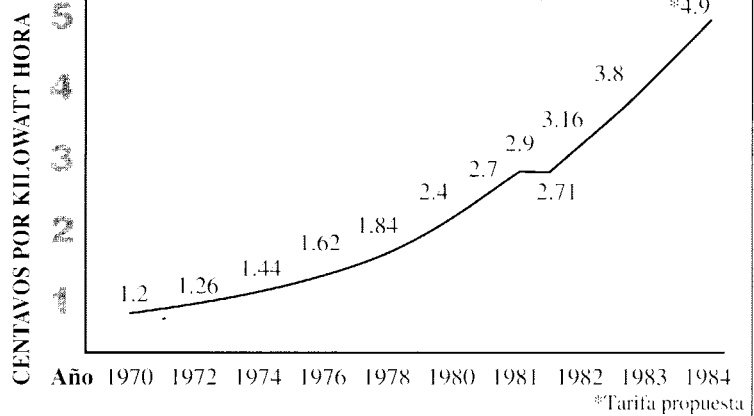
Respuesta Examine con atención el eje X. De 1970 a 1980, la escala se ha calibrado en intervalos de 2 años. A partir de 1980, la misma distancia en el eje X representa 1 año en lugar de 2 años.

Considerando los datos, al reducir esta parte de la escala se da la falsa impresión de que los costos han crecido “uniformemente”.

Pero al graficar esto en forma apropiada, como lo hemos hecho en gráfica inferior, se percibe que las tarifas no han aumentado de manera uniforme, sino que se han disparado en forma considerable durante los últimos 3 años (incluyendo el aumento de tarifas propuesto). El hecho de designar el crecimiento como “uniforme” y no como un aumento muy acelerado, favorece por supuesto los intereses de la compañía.

No se sabe con claridad si ésta proporcionó la gráfica o si el periódico la hizo por su cuenta. En cualquier caso, cuando los ejes de las gráficas no sean uniformes, ¡tenga cuidado, estimado lector!

Alzas en las tarifas de Puget Power



RESUMEN

En este capítulo hemos analizado las distribuciones de frecuencias y la forma de presentarlas en tablas y gráficas. En la estadística descriptiva nos interesa caracterizar un conjunto de datos de la manera más significativa. Si tenemos un gran número de datos, es más fácil comprenderlos, interpretarlos y analizarlos cuando se presentan por medio de una distribución de frecuencias. Una distribución de frecuencias es una lista de los valores de los datos, dispuestos en orden, junto con su frecuencia de aparición. Si existen muchos datos en un rango muy amplio, lo usual es que éstos se agrupen en intervalos iguales para obtener una interpretación más significativa. Los datos pueden presentarse como una distribución de frecuencias ordinaria, una distribución de frecuencias relativas, de frecuencias acumuladas o de porcentajes acumulados. Hemos analizado cada una de éstas y mostramos cómo hacerlas. También presentamos los

conceptos de punto percentil y rango percentil, e indicamos el modo de calcular cada uno.

Al trazar las gráficas de estas distribuciones, la frecuencia se localiza sobre el eje vertical y el valor de los datos sobre el eje horizontal. Se utilizan cuatro tipos principales de gráficas: la gráfica de barras, el histograma, el polígono de frecuencias y la curva de porcentajes acumulados. Estudiamos el uso de cada tipo de gráficas y la forma de hacerlas.

Las curvas de frecuencias pueden adoptar también varias formas. Hemos ilustrado algunas de las formas que encontramos más comúnmente en las gráficas (por ejemplo, en forma de campana, la forma de U y la forma de J); además, comentamos la diferencia entre las curvas simétricas y las asimétricas o sesgadas. Por último, estudiamos el empleo de una técnica para el análisis exploratorio de datos: los diagramas de tallo y hojas.

TÉRMINOS IMPORTANTES

Análisis exploratorio de los datos (página 58)	Curva simétrica (página 57)	Distribución de porcentajes acumulados (página 45)
Curva sesgada (página 58)	Diagramas de tallo y hojas (página 58)	Eje <i>X</i> (abscisas) (página 53)
Curva con sesgo negativo (página 58)	Distribución de frecuencias (página 39)	Eje <i>Y</i> (ordenadas) (página 53)
Curva con sesgo positivo (página 58)	Distribución de frecuencias acumuladas (página 45)	Gráfica de barras (página 53)
Curva en forma de campana (página 58)	Distribución de frecuencias de datos agrupados (página 40)	Histograma (página 55)
Curva en forma de J (página 58)	Distribución de frecuencias relativas (página 45)	Polígono de frecuencias (página 56)
Curva en forma de U (página 58)		Punto percentil (página 46)
		Rango percentil (página 50)

PREGUNTAS Y PROBLEMAS

1. Defina cada uno de los conceptos de la sección “Términos importantes”.
2. ¿En qué difiere la elaboración de las gráficas de barras, los histogramas y los polígonos de frecuencias? ¿Qué tipo de escala es adecuada para cada una?
3. La tabla siguiente muestra los salarios promedio anuales de varias categorías de científicos que ostentan doctorados, en Estados Unidos, en 2002. Haga una gráfica de barras para representar estos datos, colocando “Salario anual” sobre el eje *Y*, y “Categoría de científico” sobre el eje *X*. Ordene las categorías de manera que los salarios decrezcan de izquierda a derecha.

Categoría de científico	Salario anual (\$)
Ciencias biológicas y de la salud	70 100
Química	79 100
Ciencias matemáticas y de computación	75 000
Psicología	66 700
Sociología y antropología	63 100

4. Un estudiante de posgrado ha reunido datos acerca de 66 calificaciones. A partir de estos datos, ha hecho dos distribuciones de frecuencias de datos agrupados, las cuales presentamos a continuación.

¿Encuentra usted algún error en esas distribuciones? Explique.

a.

Intervalo de clase	f
48-63	17
29-47	28
10-28	21

b.

Intervalo de clase	f	Intervalo de clase	f
62-63	2	34-35	2
60-61	4	32-33	0
58-59	3	30-31	5
56-57	1	28-29	3
54-55	0	26-27	0
52-53	4	24-25	4
50-51	5	22-23	5
48-49	2	20-21	2
46-47	0	18-19	0
44-45	5	16-17	3
42-43	4	14-15	1
40-41	3	12-13	0
38-39	0	10-11	2
36-37	6		

5. Las siguientes calificaciones fueron obtenidas en un examen de letras inglesas por los miembros de un grupo de segundo grado de una universidad:

60	94	75	82	72	57	92	75	85	77	91
72	85	64	78	75	62	49	70	94	72	84
55	90	88	81	64	91	79	66	68	67	74
45	76	73	68	85	73	83	85	71	87	57
82	78	68	70	71	78	69	98	65	61	83
84	69	77	81	87	79	64	72	55	76	68
93	56	67	71	83	72	82	78	62	82	49
63	73	89	78	81	93	72	76	73	90	76

- Haga una distribución de frecuencias para los datos no agrupados ($i = 1$).
- Haga una distribución de frecuencias de datos agrupados que tenga 15 intervalos aproximadamente. Liste los límites aparentes y también los reales de cada intervalo.

- Elabore un histograma para la distribución de frecuencias construida en la parte **b**.
- ¿Es asimétrica o simétrica la distribución? Si está sesgada, ¿lo está en forma positiva o negativa?
- Haga un diagrama de tallo y hojas, en el que el último dígito sea una hoja y el primero sea un tallo. Incluya dos veces los valores en los tallos.
- ¿Cuál diagrama le gusta más, el histograma de la parte **c** o el diagrama de tallo y hojas de la parte **e**? Explique. educación

- Expresar la distribución de frecuencias agrupadas obtenida en la parte **b** del problema 5 como una distribución de frecuencias relativas, de frecuencias acumuladas y de porcentajes acumulados. educación
- Utilice la frecuencia acumulada obtenida en el problema 6 y determine:
 - P_{75}
 - P_{10} educación
- Use otra vez la distribución acumulada y los datos agrupados obtenidos en el problema 6 y determine:
 - El rango percentil de una calificación de 81
 - El rango percentil de una calificación de 66
 - El rango percentil de una calificación de 87 educación
- Elabore un histograma de la distribución de las calificaciones agrupadas del examen de letras inglesas, obtenida en el problema 5b. educación
- Los siguientes datos muestran la cantidad de peso (en libras) perdido por cada paciente de una clínica para el control del peso durante el último año:

10	13	22	26	16	23	35	53	17	32
41	35	24	23	27	16	20	60	48	43
52	31	17	20	33	18	23	8	24	15
26	46	30	19	22	13	22	14	21	39
28	43	37	15	20	11	25	9	15	21
21	25	34	10	23	29	28	18	17	24
16	26	7	12	28	20	36	16	14	
18	16	57	31	34	28	42	19	26	

- Haga una distribución de frecuencias de los datos agrupados con 10 intervalos aproximadamente.
- Elabore un histograma de la distribución de frecuencias obtenida en la parte **a**.

- c. ¿Es asimétrica o simétrica la distribución? Si está sesgada, ¿lo está en forma positiva o negativa?
- d. Realice un diagrama de tallo y hojas donde el último dígito sea una hoja y el primer dígito sea un tallo. Repita dos veces los valores del tallo.
- e. ¿Cuál diagrama le gusta más, el histograma de la parte **b** o el diagrama de tallo y hojas de la parte **d**? Explique. clínica, salud
11. Convierta la distribución de frecuencias agrupadas de las pérdidas de peso, determinada en el problema 10, en una distribución de frecuencias relativas y en una distribución de frecuencias acumuladas. clínica, salud
12. Utilice la distribución de frecuencias acumuladas obtenida en el problema 11 y determine:
- P_{50}
 - P_{25} clínica, salud
13. Utilice una vez más la distribución de frecuencias acumuladas obtenida en el problema 11 y determine:
- El rango percentil de un dato de 41
 - El rango percentil de un dato de 28 clínica, salud
14. Haga un polígono de frecuencias a partir de la distribución de frecuencias agrupadas determinada en el problema 10. ¿Es simétrica la curva? Si no lo es, ¿tiene sesgo positivo o negativo? clínica, salud
15. Una universidad pequeña utiliza el sistema de calificación de 0 a 4.0, donde 4.0 es la calificación más alta posible. Las calificaciones que aparecen a continuación son los promedios de las calificaciones de los estudiantes inscritos actualmente en las licenciaturas de psicología de esa universidad.
-
- | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|
| 2.7 | 1.9 | 1.0 | 3.3 | 1.3 | 1.8 | 2.6 | 3.7 |
| 3.1 | 2.2 | 3.0 | 3.4 | 3.1 | 2.2 | 1.9 | 3.1 |
| 3.4 | 3.0 | 3.5 | 3.0 | 2.4 | 3.0 | 3.4 | 2.4 |
| 2.4 | 3.2 | 3.3 | 2.7 | 3.5 | 3.2 | 3.1 | 3.3 |
| 2.1 | 1.5 | 2.7 | 2.4 | 3.4 | 3.3 | 3.0 | 3.8 |
| 1.4 | 2.6 | 2.9 | 2.1 | 2.6 | 1.5 | 2.8 | 2.3 |
| 3.3 | 3.1 | 1.6 | 2.8 | 2.3 | 2.8 | 3.2 | 2.8 |
| 2.8 | 3.8 | 1.4 | 1.9 | 3.3 | 2.9 | 2.0 | 3.2 |
-
- Haga una distribución de frecuencias para los datos agrupados, con 10 intervalos aproximadamente.
 - Elabore un histograma para la distribución de frecuencias obtenida en la parte **a**.
- c. ¿Es asimétrica o simétrica la distribución? Si está sesgada, ¿su sesgo es positivo o negativo?
- d. Realice un diagrama de tallo y hojas donde el último dígito sea una hoja y el primer dígito sea un tallo. Repita cinco veces los valores de los tallos.
- e. ¿Cuál diagrama le gusta más, el histograma de la parte **b** o el diagrama de tallo y hojas de la parte **d**? Explique. educación
16. Para las calificaciones agrupadas del problema 15, determine:
- P_{80}
 - P_{20} educación
17. El promedio de las calificaciones de Sarah es de 3.1. Tomando como base la distribución de frecuencias de las calificaciones agrupadas obtenidas en el parte **a**, ¿cuál es el rango percentil del promedio de las calificaciones de Sarah? educación
18. Según la política de la universidad del problema 15, el estudiante debe tener un promedio de calificaciones de 2.5 o mayor como requisito para que se le otorgue un título en psicología.
- A partir de los datos no agrupados que aparecen en el problema 15, ¿qué porcentaje de los estudiantes actuales de psicología necesita mejorar sus calificaciones?
 - Tomando como base la distribución de frecuencias obtenida para las calificaciones agrupadas, ¿qué porcentaje de los estudiantes necesita mejorar sus calificaciones?
 - Explique la diferencia entre las respuestas a las partes **a** y **b**. educación
19. Haga un polígono de frecuencias utilizando la distribución de calificaciones agrupadas obtenida en el problema 15. ¿La curva es simétrica, o tiene sesgo positivo o negativo?
20. El departamento de psicología de una universidad grande tiene su propio criadero de ratas con propósitos de investigación. Un muestreo reciente de 50 ratas tomadas del criadero reveló los siguientes pesos para dichas ratas:
-
- | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 320 | 282 | 341 | 324 | 340 | 302 | 336 | 265 | 313 | 317 |
| 310 | 335 | 353 | 318 | 296 | 309 | 308 | 310 | 277 | 288 |
| 314 | 298 | 315 | 360 | 275 | 315 | 297 | 330 | 296 | 274 |
| 250 | 274 | 318 | 287 | 284 | 267 | 292 | 348 | 302 | 297 |
| 270 | 263 | 269 | 292 | 298 | 343 | 284 | 352 | 345 | 325 |
-
- Haga una distribución de frecuencias de datos agrupados con 11 intervalos aproximadamente.
 - Elabore un histograma de la distribución de frecuencias obtenida en la parte **a**.

- c. ¿La distribución es simétrica o asimétrica?
- d. Construya un diagrama de tallo y hojas en el que el último dígito sea una hoja y los dos primeros dígitos sean un tallo. No repita los valores de los tallos.
- e. ¿Cuál diagrama le gusta más, el histograma o el diagrama de tallo y hojas? ¿Por qué? biológica
21. Convierta la distribución de frecuencias de datos agrupados de los pesos de las ratas obtenida en el problema 20, en una distribución de frecuencias relativas, de frecuencias acumuladas y de porcentajes acumulados. biológica
22. A partir de la distribución de frecuencias acumuladas que obtuvo en el problema 21, determine:
- P_{50}
 - P_{75} biológica
23. Utilizando otra vez la distribución de frecuencias acumuladas obtenida en el problema 21, determine:
- El rango percentil de un dato de 275
 - El rango percentil de un dato de 318 biológica
24. Una profesora está realizando investigaciones sobre las diferencias individuales de los estudiantes en cuanto a su susceptibilidad para ser hipnotizados. Como parte del experimento, la profesora decide administrar una parte de la Escala de Susceptibilidad Hipnótica de Stanford a 85 estudiantes

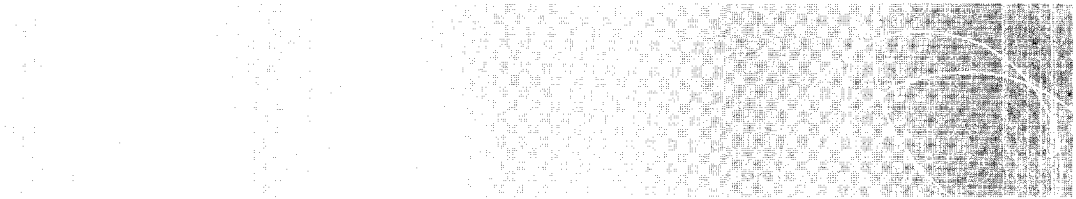
que se ofrecieron como voluntarios para el experimento. Los resultados obtenidos fueron calificados de 0 a 12, correspondiendo a 12 el grado más alto de susceptibilidad hipnótica y 0 al más bajo. Las calificaciones se presentan a continuación.

9	7	11	4	9	7	8	8	10	6
6	4	3	5	5	4	6	2	6	8
10	8	6	7	3	7	4	6	5	3
2	7	6	2	6	9	4	7	9	6
5	9	5	0	5	6	3	6	7	9
7	5	4	2	9	8	11	7	12	3
8	6	5	4	10	7	4	10	8	7
6	2	7	5	3	4	8	6	4	5
4	6	5	8	7					

- Haga una distribución de frecuencias de las calificaciones.
- Elabore un histograma de la distribución de frecuencias construida en la parte **a**.
- ¿La distribución es simétrica o sesgada?
- Determine el rango percentil de una calificación de 5 y de una calificación de 10. clínica, cognitiva, salud

Si desea encontrar problemas de práctica adicionales totalmente resueltos sobre los temas que acaba de estudiar, busque Chapter 3 en el material adicional del libro, disponible en inglés en: <http://www.thomsonlearning.com.mx>.

Medidas de tendencia central y variabilidad



Introducción

Medidas de tendencia central: media aritmética, mediana y moda

Medidas de la variabilidad: rango, desviación estándar y varianza

Resumen

Términos importantes

Preguntas y problemas

Notas



INTRODUCCIÓN

En el capítulo 3 analizamos la forma de organizar y presentar los datos de manera significativa. La distribución de frecuencias y sus muchos derivados son útiles a este respecto, pero, en sí mismas, no nos permiten hacer afirmaciones cuantitativas que caractericen a la distribución como un todo, ni tampoco permiten realizar comparaciones cuantitativas entre dos o más distribuciones. A menudo es conveniente describir las características de las distribuciones en forma cuantitativa. Por ejemplo, suponga que una psicóloga ha realizado un experimento para averiguar si los hombres y las mujeres tienen aptitudes matemáticas diferentes. Esta investigadora cuenta con dos conjuntos de calificaciones, uno para los hombres y otro para las mujeres del experimento. ¿Cómo puede comparar esas distribuciones? Para hacerlo, ella necesita cuantificarlas. La forma en que esto se realiza más a menudo consiste en calcular la calificación promedio de cada grupo y después comparar los resultados. La cantidad calculada es una medida de la *tendencia central* de cada distribución.

Una segunda característica de las distribuciones que resulta de gran utilidad para cuantificarlas es la *variabilidad* de la distribución. La variabilidad especifica la medida en la cual los datos son diferentes unos de otros debido a su dispersión. Esto es importante por dos razones. La primera es que en muchas de las pruebas de inferencia estadística

que realizaremos en el resto de este libro, tendremos que cuantificar la variabilidad de los datos. Además, la variabilidad de una distribución puede ser útil por derecho propio. Por ejemplo, supongamos que usted ha sido contratado para que diseñe y evalúe un programa educativo dirigido a jóvenes discapacitados. Al evaluar dicho programa, a usted no sólo le interesaría el valor promedio de las calificaciones al final del programa, sino también lo variable que éstos sean. La variabilidad de los datos es importante, ya que usted necesita saber si el efecto del programa es uniforme o ha variado de uno a otro de los jóvenes. Si varía, como casi siempre ocurre, ¿cuán grande es la variabilidad? ¿Ha funcionado el programa en forma satisfactoria con algunos estudiantes y mal con otros? En tal caso, es probable que sea necesario diseñar el programa de nuevo para que funcione mejor con los jóvenes que no se están beneficiando de manera adecuada con él en su estado actual.

La tendencia central y la variabilidad son las dos características de las distribuciones que son objeto de cuantificación más frecuentemente. En este capítulo vamos a analizar las medidas más importantes para estas dos características.



MEDIDAS DE TENDENCIA CENTRAL: MEDIA ARITMÉTICA, MEDIANA Y MODA

Las tres medidas de tendencia central que se usan más a menudo son la media aritmética, la mediana y la moda.

La media aritmética

Es probable que usted ya tenga conocimiento de la media aritmética, pues es justamente el valor que calcula al hallar el promedio de algo. Por ejemplo, si usted quiere saber el número promedio de horas al día que ha estudiado durante los últimos cinco días, tendrá que sumar las horas que estudió cada día y dividir el total entre 5. Al hacer esto, usted habrá calculado la media aritmética.

definición ■ La **media aritmética** se define como la suma de los datos dividida entre el número de los mismos. En forma de ecuación:

$$\bar{X} = \frac{\sum X_i}{N} = \frac{X_1 + X_2 + X_3 + \cdots + X_N}{N} \quad \text{media de una muestra}$$

o bien

$$\mu = \frac{\sum X_i}{N} = \frac{X_1 + X_2 + X_3 + \cdots + X_N}{N} \quad \text{media de un conjunto de datos de una población}$$

donde X_1, \dots, X_N = datos en bruto

\bar{X} (se lee "equis barra") = media de un conjunto de datos de una muestra

μ (se lee "mu") = media de un conjunto de datos de una población

Σ (se lee "sigma") = signo de sumatoria

N = número de datos

Observe que utilizamos dos símbolos para la media: \bar{X} cuando los datos provienen de una muestra y μ (la letra griega mu) cuando los datos provienen de una población. En cualquier caso, los cálculos son los mismos, independientemente de cuál sea la procedencia de los datos. Emplearemos μ sin subíndice alguno, para indicar que este símbolo se refiere a la media de una población de datos en bruto. En capítulos posteriores calcularemos las medias poblacionales de otros tipos de datos, y en esos casos les agregaremos los subíndices apropiados.

Resolvamos algunos problemas para practicar.



Calcule la media de cada uno de los siguientes conjuntos de datos de muestra:

$$\begin{aligned} \text{a. } X: 3, 5, 6, 8, 14 \quad \bar{X} &= \frac{\sum X_i}{N} = \frac{3+5+6+8+14}{5} \\ &= \frac{36}{5} = 7.20 \end{aligned}$$

$$\begin{aligned} \text{b. } X: 20, 22, 28, 30, 37, 38 \quad \bar{X} &= \frac{\sum X_i}{N} = \frac{20+22+28+30+37+38}{6} \\ &= \frac{175}{6} = 29.17 \end{aligned}$$

$$\begin{aligned} \text{c. } X: 2.2, 2.4, 3.1, 3.1 \quad \bar{X} &= \frac{\sum X_i}{N} = \frac{2.2+2.4+3.1+3.1}{4} \\ &= \frac{10.8}{4} = 2.70 \end{aligned}$$

Propiedades de la media La media tiene muchas propiedades o características importantes. La primera es que:

La media es sensible al valor exacto de todos los datos de la distribución.

En virtud de que para calcular la media es necesario sumar *todos* los datos, una modificación en cualquiera de éstos provocará un cambio en la media. No sucede así ni con la mediana ni con la moda.

La segunda propiedad de la media es la siguiente:

La suma de las desviaciones con respecto a la media es igual a cero. Expresada en forma algebraica, esta propiedad se convierte en $\sum (X_i - \bar{X}) = 0$.

Esta propiedad indica que si la media se resta de cada dato, la suma de las diferencias será igual a cero. La demostración algebraica de este hecho se presenta en la nota 4.1, que se encuentra al final de este capítulo. En la tabla 4.1 encontramos una demostración de su validez. Esta propiedad surge del hecho de que la media es el punto de equilibrio de la distribución. La media puede ser imaginada como el punto de apoyo de un sube y baja, si queremos emplear una analogía mecánica. Esta analogía se ilustra en la figura 4.1, la cual utiliza los datos de la tabla 4.1. Cuando los datos han sido distribuidos a lo largo del sube y baja según sus valores, la media de la distribución ocupa la posición en la que dichos datos se encuentran en equilibrio.

tabla 4.1 Demostración de que $\sum (X_i - \bar{X}) = 0$

X_i	$X_i - \bar{X}$	Cálculo de \bar{X}
2	-4	$\bar{X} = \frac{\sum X_i}{N} = \frac{30}{5}$ $= 6,00$
4	2	
6	0	
8	+2	
10	+4	
$\sum X_i = 30$	$\sum (X_i - \bar{X}) = 0$	

Una tercera propiedad de la media se deduce también del hecho de que la media es el punto de equilibrio de la distribución:

La media es muy sensible a los datos extremos.

Un vistazo a la figura 4.1 deberá convencerle de que si se agregara un dato extremo (muy lejano de la media), eso perturbaría considerablemente el equilibrio. La media tendría que desplazarse una gran distancia para restablecer el equilibrio. La media es más sensible a los valores extremos que la mediana o la moda. Comentaremos esto con mayor detalle cuando nos ocupemos de la mediana.

Una cuarta propiedad de la media se relaciona con la variabilidad de los datos que se encuentran en torno a ella. Esta propiedad establece lo siguiente:

La suma de los cuadrados de las desviaciones de todos los datos en torno a su media es la más pequeña posible. Expresando este concepto en forma algebraica, decimos que $\sum (X_i - \bar{X})^2$ es mínima.

Ésta es una característica importante y se utiliza en muchas áreas de la estadística, particularmente en la correspondiente a la regresión. Cuando se desarrolla con un poco más de detalle, esta propiedad establece que, aun cuando la suma de las desviaciones cuadradas en torno a la media no suele ser igual a cero, su magnitud es menor que si las desviaciones cuadradas hubieran sido calculadas en torno a cualquier otro valor. La validez de esta propiedad se demuestra en la tabla 4.2. Los datos de la distribución aparecen en la primera columna. Su media es igual a 5,00. La cuarta columna muestra los cuadrados de las desviaciones de los datos X_i en torno a su media $(X_i - 5)^2$. La suma de los cuadrados de estas desviaciones es igual a 20. Las otras columnas muestran los cuadrados de las desviaciones de los datos X_i en torno a valores diferentes de la media. En la tercera columna, el valor es $4(X_i - 4)^2$, en la segunda columna 3, en la quinta columna 6, y en la sexta columna 7. Observe que la suma de los cuadrados de las desviaciones en torno a cada uno de estos valores es mayor que la suma de los cuadrados de las desviaciones en torno a la

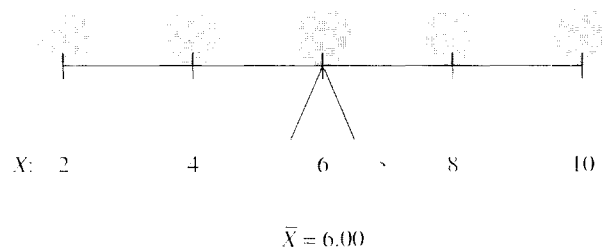
**figura 4.1** La media como el punto de equilibrio de la distribución.

tabla 4.2 Demostración de que $\sum (X_i - \bar{X})^2$ es mínima

X_i	$(X_i - 3)^2$	$(X_i - 4)^2$	$(X_i - 5)^2$	$(X_i - 6)^2$	$(X_i - 7)^2$	Cálculo de \bar{X}
2	1	4	9	16	25	$\bar{X} = \frac{\sum X_i}{N}$
4	1	0	1	4	9	
6	9	4	1	0	1	
8	25	16	9	4	1	
	36	24	20	24	36	= 5.00

media de la distribución. No sólo sucede que la suma es mayor, sino que cuando el valor se aleja de la media, la suma es cada vez más grande. Esto implica que a pesar de que sólo hayamos comparado cuatro cantidades distintas de la media, esta propiedad se cumple para cualquier valor. Por consiguiente, aun cuando la suma de los cuadrados de las desviaciones en torno a la media no se anula de ordinario, siempre es menor que la suma de los cuadrados de las desviaciones medidas en torno a cualquier otro valor.

Esta última propiedad tiene relación con el uso de la media para fines de inferencia estadística. Esta propiedad establece lo siguiente:

En la mayoría de los casos, de todas las medidas utilizadas para calcular la tendencia central, la media es la menos susceptible a las variaciones debidas al muestreo.

Si en repetidas ocasiones obtuviéramos muestras de una población, en forma aleatoria, la media variaría de una a otra de esas muestras. Lo mismo sucede con la mediana y la moda. Sin embargo, la media varía menos que estas otras medidas de tendencia central. Esto es muy importante en estadística inferencial y es una de las principales razones por las que la media se utiliza siempre que es posible en esta rama de la estadística.

La media global

En ciertas ocasiones se presenta el caso de que conocemos la media de varios grupos de datos y queremos calcular la media de todos ellos juntos. Por supuesto, podríamos comenzar otra vez desde el principio y sumar simplemente todos los datos en bruto, para luego dividir entre el número total de los mismos. Sin embargo, existe una forma de abreviar este proceso cuando ya conocemos la media y el número de datos que contiene cada grupo. La ecuación para este método se deduce de la definición básica de la media. Suponga que tenemos varios grupos de datos y los queremos combinar para calcular la media global. Sea k el número de grupos. Entonces,

$$\bar{X}_{\text{global}} = \frac{\text{Suma de todos los datos}}{N}$$

$$= \frac{\sum X_i (\text{primer grupo}) + \sum X_i (\text{segundo grupo}) + \dots + \sum X_i (\text{último grupo})}{n_1 + n_2 + \dots + n_k}$$

donde N = número total de datos
 n_1 = número de datos en el primer grupo
 n_2 = número de datos en el segundo grupo
 n_k = número de datos en el último grupo

Como $\bar{X}_1 = \sum X_i$ (primer grupo)/ n_1 , al multiplicar por n_1 , tenemos que $\sum X_i$ (primer grupo) = $n_1\bar{X}_1$. De manera análoga, $\sum X_i$ (segundo grupo) = $n_2\bar{X}_2$ y $\sum X_i$ (último grupo) = $n_k\bar{X}_k$, donde \bar{X}_k es la media del último grupo. Al sustituir estos valores en el numerador de la ecuación anterior, tenemos que:

$$\bar{X}_{\text{global}} = \frac{n_1\bar{X}_1 + n_2\bar{X}_2 + \dots + n_k\bar{X}_k}{n_1 + n_2 + \dots + n_k} \quad \text{media global de varios grupos}$$

En otras palabras, esta ecuación establece que la media global es igual a la suma de la media de cada grupo multiplicada por el número de datos del grupo, entre la suma del número de datos de cada grupo.

Con la finalidad de ilustrar el uso de esta ecuación, supondremos que un profesor de sociología ha administrado un examen final a dos grupos. La media de uno de esos grupos fue de 90 y el número de calificaciones registradas en el mismo fue de 20. La media del otro grupo fue de 70, y 40 estudiantes presentaron el examen. Calcule usted la media de los dos grupos combinados.

La solución es la siguiente: dado que $\bar{X}_1 = 90$ y $n_1 = 20$ y que $\bar{X}_2 = 70$ y $n_2 = 40$.

$$\bar{X}_{\text{global}} = \frac{n_1\bar{X}_1 + n_2\bar{X}_2}{n_1 + n_2} = \frac{20(90) + 40(70)}{20 + 40} = 76.67$$

La media global está mucho más próxima al promedio del grupo en el que hubo 40 calificaciones, que al promedio del grupo que tuvo 20 calificaciones. En este contexto, podemos apreciar que cada una de estas medias ha sido *ponderada* según el número de sus datos. Hemos contado la media de 70 cuarenta veces y la media de 90 fue contada únicamente veinte veces. Así pues, la media global es en realidad una media ponderada en la que los pesos (ponderaciones) son los números de datos empleados en la determinación de cada media. Resolvamos un problema más para practicar.



Un investigador realizó un experimento con tres grupos de sujetos. La media del primer grupo fue de 75 y había en él 50 sujetos. La media del segundo grupo fue de 80 y estaba integrado por 40 sujetos. En el tercer grupo había 25 sujetos y la media fue de 70. Calcule la media global de los tres grupos combinados.

SOLUCIÓN

La solución es la siguiente: dado que $\bar{X}_1 = 75$, $n_1 = 50$; $\bar{X}_2 = 80$, $n_2 = 40$; y $\bar{X}_3 = 70$, $n_3 = 25$,

$$\begin{aligned} \bar{X}_{\text{global}} &= \frac{n_1\bar{X}_1 + n_2\bar{X}_2 + n_3\bar{X}_3}{n_1 + n_2 + n_3} = \frac{50(75) + 40(80) + 25(70)}{50 + 40 + 25} \\ &= \frac{8700}{115} = 75.65 \end{aligned}$$

La mediana

La segunda de las medidas de tendencia central de uso más frecuente es la mediana.

definición ■ La **mediana** (cuyo símbolo es Mdn) se define como el valor de la escala debajo del cual se encuentra 50% de los datos. Por lo tanto, es lo mismo que P_{50} .

En el capítulo 3 analizamos la forma de calcular P_{50} , por lo cual usted ya sabe calcular la mediana para datos agrupados. Sin embargo, para practicar, el problema de práctica 4.3 presenta otro ejercicio y su solución. Es conveniente que usted trate de resolverlo antes de seguir adelante.



Calcular la mediana de los datos agrupados que muestra la tabla 4.3.

tabla 4.3 Cálculo de la mediana a partir de datos agrupados

Intervalo de clase	f	f acumulada	% acumulado	Cálculo de la mediana
3.6–4.0	4	52	100.00	
3.1–3.5	6	48	92.31	
2.6–3.0	8	42	80.77	
2.1–2.5	10	34	65.38	$Mdn = P_{50}$
1.6–2.0	9	24	46.15	$= X_L + (i/f_i)(f_p \text{ acumulada} - f_L \text{ acumulada})$
1.1–1.5	7	15	28.85	$= 2.05 + (0.5/10)(26 - 24)$
0.6–1.0	5	8	15.38	$= 2.05 + 0.10 = 2.15$
0.1–0.5	3	3	5.77	

SOLUCIÓN

La mediana es el valor por debajo del cual se encuentra 50% de los datos. En virtud de que $N = 52$, la mediana es el valor por debajo del cual se encuentran 26 de los datos (puesto que 50% de 52 es 26). La tabla 4.3 muestra que la mediana está localizada en el intervalo 2.1-2.5. Como 24 datos se encuentran por debajo de un valor de 2.05, necesitamos dos datos más para completar los 26 datos requeridos. En virtud de que existen 10 datos dentro del intervalo y éste tiene 0.5 unidades de amplitud, deberemos desplazarnos $0.5/10 \times 2 = 0.10$ unidad en él. De esta manera,

$$\text{Mediana} = 2.05 + 0.10 = 2.15$$

También podríamos haber obtenido la mediana aplicando la ecuación para el punto percentil. Esta solución se presenta en la tabla 4.3.

Cuando se trabaja con datos en bruto (no agrupados), es muy fácil determinar la mediana. En primer lugar, ordenamos por rango los datos.

La mediana es el dato que está ubicado en el centro cuando el número de datos es impar. Si el número de datos es par, se considera que la mediana es el promedio de los dos datos ubicados al centro.

Para ilustrar esto, suponga usted que tenemos los datos 5, 2, 3, 7 y 8 y queremos determinar la mediana. En primer lugar, ordenamos por rango los datos: 2, 3, 5, 7, 8. Como el número de datos es impar, la mediana es el dato central. En este ejemplo, la mediana es el 5. Podría parecer que 5 no es en realidad el P_{50} para este conjunto de datos. Sin embargo, consideremos que el dato 5 se encuentra distribuido de manera uniforme en todo el intervalo 4.5-5.5. Ahora sí nos resulta obvio que la mitad de los datos se encuentran por debajo de 5.0. Por consiguiente, 5.0 es P_{50} .

Probemos ahora con otro ejemplo, en esta ocasión con un número de datos par. Dados los datos 2, 8, 6, 4, 12 y 10, determine su mediana. En primer lugar, ordenamos por rango los datos: 2, 4, 6, 8, 10, 12. Como el número de datos es par, la mediana es el promedio de los dos datos centrales. En este ejemplo, la mediana es $(6 + 8)/2 = 7$. Como ejercicio adicional, el problema de práctica 4.4 presenta unos cuantos problemas que se refieren a datos en bruto.



Calcular la mediana de los siguientes conjuntos de datos:

a. 8, 10, 4, 3, 1, 15	En orden por rango: 1, 3, 4, 8, 10, 15	Mdn = $(4 + 8)/2 = 6$
b. 100, 102, 108, 104, 112	En orden por rango: 100, 102, 104, 108, 112	Mdn = 104
c. 2.5, 1.8, 1.2, 2.4, 2.0	En orden por rango: 1.2, 1.8, 2.0, 2.4, 2.5	Mdn = 2.0
d. 10, 11, 14, 14, 16, 14, 12	En orden por rango: 10, 11, 12, 14, 14, 14, 16	Mdn = 14

En el último conjunto de datos del problema de práctica 4.4, la mediana se presenta en 14, donde se ubican tres datos. Desde un punto de vista técnico, deberíamos considerar que los tres datos están igualmente dispersos sobre el intervalo 13.5-14.5. Entonces determinaríamos la mediana por medio de la ecuación que aparece en la tabla 4.3, con $i = 1$ ($Mdn = 13.67$). Sin embargo, al trabajar con datos en bruto no se acostumbra tanto refinamiento. En lugar de eso, se considera que la mediana está en 14. Nosotros seguiremos este procedimiento. Por consiguiente, si la mediana se presenta en un valor donde hay datos empatados, utilizaremos el dato empatado como la mediana.

Propiedades de la mediana Hay dos propiedades de la mediana que vale la pena mencionar. La primera es:

La mediana es menos sensible que la media a la presencia de datos extremos.

Para ilustrar esta propiedad, considere los datos de la primera columna de la tabla 4.4. Las tres distribuciones allí mostradas son iguales, excepto por el último dato. En la segunda distribución, el dato 100 es muy diferente, en valor, a los demás. En la tercera distribución, el dato 1000 es todavía más extremo. Observe lo que ocurre con la media en

tabla 4.4 Efecto de los datos extremos sobre la media y la mediana

Datos	Media	Mediana
3, 4, 6, 7, 10	6	6
3, 4, 6, 7, 100	24	6
3, 4, 6, 7, 1000	204	6

la segunda y la tercera de las distribuciones. Como quiera que la media es sensible a los datos extremos, su valor cambia de manera considerable por efecto de éstos. ¿Y qué se puede decir de la mediana? ¿Cambia también? Como podemos ver en la tercera columna, la respuesta es no. La mediana sigue siendo la misma; en virtud de que ella no depende de cada dato en forma individual, sino divide toda la distribución a la mitad, no resulta tan sensible como la media a la presencia de datos extremos. Por esta razón, cuando la distribución posee una marcada asimetría, tal vez sea mejor representar la tendencia central con la mediana y no con la media. No cabe duda de que, en la tercera distribución de la tabla 4.4, la mediana 6 constituye una representación mejor de la mayor parte de los datos, que la media de 204.

La segunda propiedad de la mediana se refiere a su estabilidad en el muestreo. Esta propiedad establece lo siguiente:

En circunstancias ordinarias, la mediana es más susceptible que la media a la variabilidad en el muestreo, pero menos susceptible que la moda a la variabilidad en éste.

Puesto que, por lo general, la mediana es menos estable que la media de una muestra a otra, no es tan útil como ésta en la estadística inferencial.

La moda

La tercera y última medida de tendencia central que comentaremos es la moda.

definición ■ *La moda se define como el dato más frecuente en la distribución.**

Resulta claro que, entre las tres medidas, ésta es la más sencilla de determinar. La moda se determina examinando los datos; no es necesario efectuar cálculo alguno. Por ejemplo, para determinar la moda de los datos de la tabla 3.2, lo único que debemos hacer es revisar la columna de frecuencias. La moda para estos datos es 76. Una vez agrupados los datos, la moda se define como el punto medio del intervalo que tenga la más alta frecuencia. La moda de los datos agrupados en la tabla 3.4 es 77.

Por lo general, las distribuciones son *unimodales*; es decir, sólo tienen una moda. Sin embargo, es posible que una distribución tenga muchas modas. Cuando una distribución tiene dos modas, como es el caso de los datos 1, 2, 3, 3, 3, 3, 4, 5, 7, 7, 7, 7, 8, 9, la distribución se llama *bimodal*. En la figura 4.2 se presentan los histogramas de una distribución

*Cuando todos los datos de una distribución tienen la misma frecuencia, se acostumbra decir que la distribución no tiene moda.

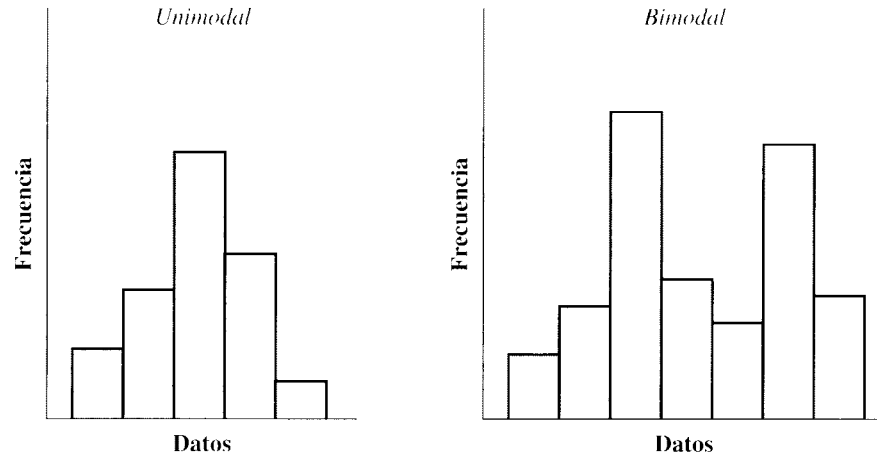
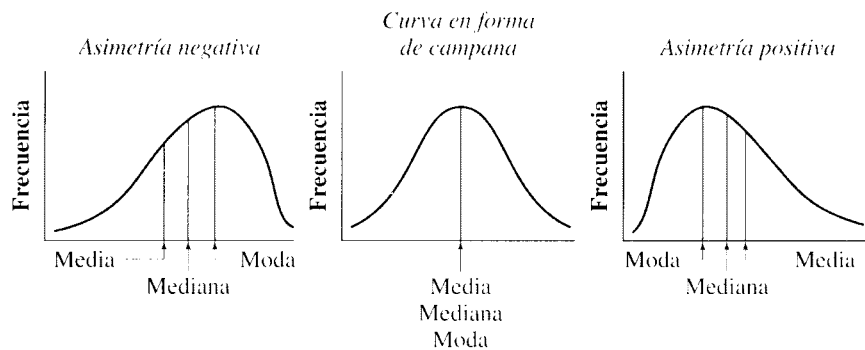


figura 4.2 Histogramas unimodal y bimodal.

unimodal y otra bimodal. Pese a ser la medida de tendencia central más fácil de determinar, la moda no se usa mucho en las ciencias del comportamiento porque no es muy estable de una muestra a otra y porque a menudo existe más de una moda para un conjunto determinado de datos.

Medidas de tendencia central y simetría

Si la distribución es unimodal y simétrica, entonces la media, la mediana y la moda serán iguales entre sí. Un ejemplo de esto es la curva en forma de campana que mencionamos en el capítulo 3 y que se ilustra en la figura 4.3. Cuando la distribución es asimétrica, los valores de la media y la mediana no son iguales. Como la media es la que resulta más afectada por los datos extremos, tendrá un valor más cercano a dichos datos que la mediana. Así, en el caso de una distribución asimétrica negativa, es decir, sesgada negativamente, la media será menor que la mediana. Con una curva asimétrica positiva o sesgada positivamente, la media será mayor que la mediana. La figura 4.3 muestra estas relaciones.



Tomado de *Statistical Reasoning in Psychology and Education* por E. W. Minium. Derechos reservados © 1978 John Wiley & Sons, Inc. Adaptado con autorización de John Wiley & Sons, Inc.

figura 4.3 Simetría y medidas de tendencia central.



MEDIDAS DE VARIABILIDAD: RANGO, DESVIACIÓN ESTÁNDAR Y VARIANZA

En una sección anterior de este capítulo dijimos que la variabilidad se refiere a qué tan alejados de la media están los datos. Así como las medidas de tendencia central son una cuantificación del valor promedio de la distribución, las medidas de variabilidad cuantifican la magnitud de la dispersión. Hay tres medidas de la variabilidad que se emplean comúnmente en las ciencias del comportamiento: el rango, la desviación estándar y la varianza.

El rango

Ya hemos empleado aquí el rango, al hacer distribuciones de frecuencias de datos agrupados.

definición ■ *El rango se define como la diferencia entre los datos máximo y mínimo de una distribución. Se expresa así, en forma de ecuación:*

$$\text{Rango} = \text{Dato máximo} - \text{Dato mínimo}$$

El rango es fácil de calcular, pero no nos proporciona más que una medida relativamente burda de la dispersión porque, en realidad, el rango mide únicamente la dispersión de los datos extremos y no la de los datos intermedios. A pesar de que el rango es fácil de calcular, hemos incluido algunos problemas al respecto para que usted lo practique. Siempre es mejor prevenir que lamentar.



Calcular el rango de las siguientes distribuciones:

- | | |
|----------------------------|---------------------------|
| a. 2, 3, 5, 8, 10 | Rango = $10 - 2 = 8$ |
| b. 18, 12, 28, 15, 20 | Rango = $28 - 12 = 16$ |
| c. 115, 107, 105, 109, 101 | Rango = $115 - 101 = 14$ |
| d. 1.2, 1.3, 1.5, 1.8, 2.3 | Rango = $2.3 - 1.2 = 1.1$ |

La desviación estándar

Antes de ocuparnos de la desviación estándar, es necesario presentar el concepto de desviación de un dato, es decir, el puntaje de desviación.

Puntajes de desviación Hasta ahora hemos trabajado principalmente con datos en bruto. Usted recordará que un dato en bruto es el dato tal como fue medido en forma original. Por ejemplo, si nos interesa conocer el CI y medimos un CI de 126, entonces 126 es un dato en bruto.

definición ■ *El puntaje de desviación nos indica qué tan lejos está el dato en bruto con respecto a la media de su distribución.*

En forma de ecuación, el puntaje de desviación se define como:

$$\begin{aligned} X - \bar{X} & \text{ puntaje de desviación para datos de muestra} \\ X - \mu & \text{ puntaje de desviación para datos de población} \end{aligned}$$

Para ilustrar esto, consideremos los datos muestrales de la tabla 4.5. Los datos en bruto figuran en la primera columna y sus puntajes de desviación transformados se presentan en la segunda columna. El puntaje de desviación nos indica a qué distancia se encuentra el dato en bruto, ya sea por encima o por debajo de la media. De esta manera, el dato en bruto de 2 ($X = 2$) se encuentra 4 unidades por debajo de la media ($X - \bar{X} = -4$). Los datos en bruto y los puntajes de desviación de los mismos se muestran también en forma gráfica en la figura 4.4.

Suponga que usted es un matemático con futuro (use la imaginación en caso necesario) y que se le ha asignado la tarea de deducir una medida de dispersión que proporcione la desviación promedio de los datos en torno a la media. Después de reflexionar un poco, usted dice: "Eso es fácil; solamente hay que calcular la desviación de cada dato con respecto a la media y promediar los puntajes de desviación de los datos". Su lógica es impecable. Sin embargo, hay un ligero problema. Considere los datos de la tabla 4.6. Para este ejemplo en especial, supondremos que se trata de un conjunto de datos poblacionales. La primera columna contiene los datos en bruto de la población; la segunda, la desviación de dichos datos. Queremos calcular desviación promedio de los datos en bruto en torno a su media. Según el método propuesto por usted, podríamos calcular primero los puntajes de desviación (segunda columna) y los promediaríamos, dividiendo la suma de la desviación de los datos [$\sum (X - \mu)$] entre N . La dificultad consiste en que $\sum (X - \mu) = 0$. Recuerde que ésta es una propiedad general de la media. La suma de las desviaciones en torno a la media siempre es igual a cero. Así pues, si siguiéramos su sugerencia, el promedio de las desviaciones siempre sería igual a cero, sin importar cuán dispersos estuvieran los datos [$\sum (X - \mu)/N = 0/N = 0$].

Por un momento, usted se queda desconcertado por este golpe inesperado. Sin embargo, no se da por vencido. Observa los puntajes de desviación y nota que los datos negativos se cancelan con los positivos. De pronto tiene una rápida inspiración: ¿por qué no elevar al cuadrado cada puntaje de desviación? Al hacer esto, todos los datos serían positivos y su suma ya no sería cero. ¡Eureka! Usted ha resuelto el problema. Ahora puede dividir la suma de los cuadrados de las desviaciones entre N para obtener el valor promedio [$\sum (X - \mu)^2/N$], y el resultado será distinto de cero. Notará usted que al numerador de esta fórmula [$\sum (X - \mu)^2$] se le llama la *suma de los cuadrados de las desviaciones*, lo cual se simboliza como SC_{pob} . El único problema, al llegar a este punto es que usted ha calculado la *desviación promedio al cuadrado*, no la desviación promedio. Lo que ahora debe hacer es "eliminar el cuadrado" de la respuesta. Esto se logra extrayendo la raíz cuadrada de SC_{pob}/N .

tabla 4.5 Cálculo de puntajes de desviación

X	$X - \bar{X}$	Cálculo de \bar{X}
2	2 - 6 = -4	$\bar{X} = \frac{\sum X}{N} = \frac{30}{6} = 6.00$
4	4 - 6 = -2	
6	6 - 6 = 0	
8	8 - 6 = +2	
10	10 - 6 = +4	

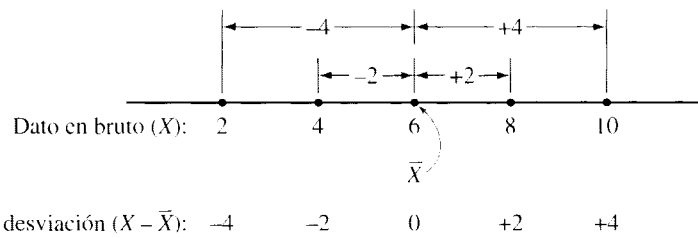


figura 4.4 Datos en bruto y sus correspondientes puntajes de desviación.

¡Fue salvada su reputación como matemático! Llegó usted a la ecuación que utilizan muchos estadísticos para la desviación estándar. El símbolo para la desviación estándar de datos poblacionales es σ (la letra griega sigma en minúscula) y para datos de muestras es s . La ecuación obtenida por usted para los datos de una población es la siguiente:

$$\sigma = \sqrt{\frac{SC_{\text{pob}}}{N}} = \sqrt{\frac{\sum (X - \mu)^2}{N}} \quad \text{desviación estándar de un conjunto poblacional de datos en bruto — método de la desviación}$$

donde $SC_{\text{pob}} = \sum (X - \mu)^2$ suma de cuadrados — datos poblacionales

El cálculo de la desviación estándar de un conjunto de datos poblacionales mediante el método de la desviación se presenta en la tabla 4.6.

Desde un punto de vista técnico, la ecuación es la misma para calcular la desviación estándar de los datos de una muestra. Sin embargo, cuando calculamos la desviación estándar de los datos de una muestra, generalmente queremos utilizar nuestro cálculo para estimar la desviación estándar de la población. Se puede demostrar en forma algebraica que la ecuación con la N en el denominador proporciona una estimación que, en promedio, es demasiado pequeña. Al dividir entre $N - 1$, en lugar de N , se obtiene una estimación más precisa de σ . Puesto que la estimación de la desviación estándar de la población es una aplicación importante de la desviación estándar de la muestra, y como más adelante nos ahorrará muchas confusiones, cuando analicemos la prueba t de Student y la prueba F , hemos optado por adoptar la ecuación con $N - 1$ en el denominador para calcular la desviación estándar de datos muestrales. Así,

$$s = \text{estimada } \sigma = \sqrt{\frac{SC}{N - 1}} = \sqrt{\frac{\sum (X - \bar{X})^2}{N - 1}} \quad \text{desviación estándar de un conjunto muestral de datos en bruto — método de la desviación}$$

donde $SC = \sum (X - \bar{X})^2$ suma de los cuadrados — datos muestrales

tabla 4.6 Cálculo de la desviación estándar de un conjunto de datos poblacionales mediante el método de la desviación

X	$X - \mu$	$(X - \mu)^2$	Cálculo de μ y de σ
3	-2	4	$\mu = \frac{\sum X}{N} = \frac{25}{5} = 5.00$ $\sigma = \sqrt{\frac{SC_{\text{pob}}}{N}} = \sqrt{\frac{\sum (X - \mu)^2}{N}} = \sqrt{\frac{10}{5}} = 1.41$
4	1	1	
5	0	0	
6	+1	1	
7	+2	4	
	$\sum (X - \mu) = 0$	$\sum (X - \mu)^2 = 10$	

tabla 4.7 Cálculo de la desviación estándar de datos muestrales mediante el método de la desviación

X	$X - \bar{X}$	$(X - \bar{X})^2$	Cálculo de \bar{X} y de s
2	-4	16	$\bar{X} = \frac{\sum X}{N} = \frac{30}{5} = 6.00$
4	-2	4	
6	0	0	$s = \sqrt{\frac{SC}{N-1}} = \sqrt{\frac{\sum (X - \bar{X})^2}{N-1}} = \sqrt{\frac{40}{5-1}}$
8	+2	4	
10	+4	16	
	$\sum (X - \bar{X}) = 0$	$SC = 40$	$= \sqrt{10} = 3.16$

En la mayor parte de las situaciones prácticas, los datos se obtienen más a menudo de muestras que de poblaciones. El cálculo de la desviación estándar de una muestra por medio de la ecuación anterior se ilustra en la tabla 4.7. A pesar de que esta ecuación proporciona la mejor comprensión conceptual de la desviación estándar y provee la respuesta correcta, su uso resulta muy engorroso en la práctica. Esto ocurre, en especial, cuando la media no es un número entero. La tabla 4.8 muestra un caso ilustrativo, utilizando la ecuación anterior con una media que tiene un residuo decimal. Observe que cada puntaje de desviación tiene un residuo decimal, el cual es necesario elevar al cuadrado para obtener $(X - \bar{X})^2$. Por consiguiente, es necesario redondear en varias ocasiones, y esto puede acentuar la imprecisión. Además, tenemos que efectuar sumas con números de 5 dígitos, lo cual aumenta la posibilidad de error. Usted puede apreciar lo engorroso que resulta el uso de esta ecuación cuando la media no es un entero; y en la mayor parte de los problemas prácticos, ¡la media no es un número entero!

Cálculo de la desviación estándar de una muestra por el método de los datos en bruto
Se puede demostrar algebraicamente que:

$$SC = \sum X^2 - \frac{(\sum X)^2}{N} \quad \text{suma de cuadrados}$$

tabla 4.8 Cálculo de la desviación estándar empleando puntajes de desviación, cuando la media no es un número entero

X	$X - \bar{X}$	$(X - \bar{X})^2$	Cálculo de \bar{X} y s
10	-6.875	47.2656	$\bar{X} = \frac{\sum X}{N} = \frac{135}{8} = 16.875$
12	-4.875	23.7656	
13	-3.875	15.0156	$s = \sqrt{\frac{SC}{N-1}} = \sqrt{\frac{\sum (X - \bar{X})^2}{N-1}}$
15	1.875	3.5156	
18	1.125	1.2656	
20	3.125	9.7656	
22	5.125	26.2656	
25	8.125	66.0156	$= \sqrt{\frac{192.8748}{7}}$
$\sum X = 135$	$\sum (X - \bar{X}) = 0.000$	$SC = 192.8748$	$= \sqrt{27.5535}$
$N = 8$			$= 5.25$

La deducción de este hecho se presenta en la nota 4.2. El empleo de esta ecuación para determinar SC nos permite utilizar los datos en bruto sin tener que calcular los puntajes de desviación. Esto, a su vez, evita los problemas con los residuos decimales descritos con anterioridad. A este procedimiento para calcular SC lo llamaremos “método de los datos en bruto” para distinguirlo del “método de la desviación”. Como quiera que el método de los datos en bruto generalmente es más sencillo y evita errores potenciales, es la opción preferida para calcular SC y lo utilizaremos en el resto de este texto. Al aplicar el método de los datos en bruto, *usted debe asegurarse de no confundir $\sum X^2$ y $(\sum X)^2$. $\sum X^2$ se lee “sumatoria de X cuadrada” o “suma de los cuadrados de X ”, y $(\sum X)^2$ se lee “el cuadrado de la sumatoria de X ” o “la suma de los datos X , al cuadrado”.* Para determinar $\sum X^2$, elevamos al cuadrado cada dato y luego sumamos los cuadrados. Para determinar $(\sum X)^2$, sumamos todos los datos y luego elevamos al cuadrado la suma. El resultado es distinto con ambos procedimientos. Además, SC debe ser positiva. Si en sus cálculos resulta negativa, es probable que usted haya confundido $\sum X^2$ y $(\sum X)^2$.

La tabla 4.9 muestra el cálculo de la desviación estándar presentada en la tabla 4.8, mediante el método de los datos en bruto. Al utilizar este método, calculamos primero SC a partir de la ecuación de los datos en bruto y a continuación sustituimos el valor obtenido en la ecuación para la desviación estándar.

Propiedades de la desviación estándar La desviación estándar tiene muchas características importantes. La primera de ellas es que nos proporciona una *medida de la dispersión con respecto a la media*. Esto es diferente del rango, el cual nos presenta una medida absoluta de la dispersión entre los dos datos más extremos. En segundo lugar, *la desviación estándar es sensible a cada uno de los datos de la distribución*. Si se cambiara un dato por un valor más cercano a la media, entonces la desviación estándar sería menor. Inversamente, si se le cambiara por otro más alejado de la media, entonces la desviación estándar aumentaría. En tercer lugar, *igual que la media, la desviación estándar es estable con respecto a las variaciones debidas al muestreo*. Si se tomaran varias veces muestras de poblaciones del tipo que suele encontrarse en las ciencias del comportamiento, la desviación estándar de las muestras variaría mucho menos que el rango, de una muestra a otra. Esta propiedad es una de las principales razones por las cuales se utiliza mucho más a menudo la desviación estándar, en lugar del rango, para analizar la variabilidad. Por último, tanto la media como la desviación estándar pueden manipularse algebraicamente. Esto permite realizar cálculos matemáticos con ellas para utilizarlas en estadística inferencial.

Pasemos ahora a los problemas de práctica 4.6 y 4.7.

tabla 4.9 Cálculo de la desviación estándar mediante el método de los datos en bruto

X	X^2	Cálculo de SC	Cálculo de s
10	100	$SC = \sum X^2 - \frac{(\sum X)^2}{N}$ $= 2471 - \frac{(135)^2}{8}$ $= 2471 - 2278.125$ $= 192.875$	$s = \sqrt{\frac{SC}{N-1}}$ $= \sqrt{\frac{192.875}{7}}$ $= \sqrt{27.5536}$ $= 5.25$
12	144		
13	169		
15	225		
18	324		
20	400		
22	484		
<u>25</u>	<u>625</u>		
$\sum X$ 135	$\sum X^2$ 2471		
N 8			

Si desea encontrar una solución en la que se use la ecuación conceptual (puntajes de desviación), busque Chapter 4 en el material adicional del libro. en: <http://www.thomsonlearning.com.mx>.



Calcule la desviación estándar de los datos contenidos en la primera columna de la siguiente tabla:

X	X^2	Cálculo de SC	Cálculo de s
25	625	$SC = \Sigma X^2 - \frac{(\Sigma X)^2}{N}$ $= 15\,545 - \frac{(387)^2}{10}$ $= 15\,545 - 14\,976.9$ $= 568.1$	$s = \sqrt{\frac{SC}{N-1}}$ $= \sqrt{\frac{568.1}{9}}$ $= \sqrt{63.1222}$ $= 7.94$
28	784		
35	1225		
37	1369		
38	1444		
40	1600		
42	1764		
45	2025		
47	2209		
<u>50</u>	<u>2500</u>		
$\Sigma X = 387$	$\Sigma X^2 = 15\,547$		
$N = 10$			

Si desea encontrar una solución en la que se use la ecuación conceptual (puntajes de desviación), busque Chapter 4 en el material adicional del libro. en: <http://www.thomsonlearning.com.mx>.



Calcule la desviación estándar de los datos contenidos en la primera columna de la siguiente tabla:

X	X^2	Cálculo de SC	Cálculo de s
1.2	1.44	$SC = \Sigma X^2 - \frac{(\Sigma X)^2}{N}$ $= 60.73 - \frac{(25.9)^2}{12}$ $= 60.73 - 55.9008$ $= 4.8292$	$s = \sqrt{\frac{SC}{N-1}}$ $= \sqrt{\frac{4.8292}{11}}$ $= \sqrt{0.4390}$ $= 0.66$
1.4	1.96		
1.5	2.25		
1.7	2.89		
1.9	3.61		
2.0	4.00		
2.2	4.84		
2.4	5.76		
2.5	6.25		
2.8	7.84		
3.0	9.00		
<u>3.3</u>	<u>10.89</u>		
$\Sigma X = 25.9$	$\Sigma X^2 = 60.73$		
$N = 12$			

La varianza

La varianza de un conjunto de datos es simplemente el cuadrado de la desviación estándar. Para los datos de una muestra, la varianza es igual a:

$$s^2 = \text{estimada } \sigma^2 = \frac{SC}{N-1} \quad \text{varianza de una muestra}$$

Para los datos de una población, la varianza es igual a:

$$\sigma^2 = \frac{SC_{\text{pob}}}{N} \quad \text{varianza de una población}$$

La varianza no se usa muy a menudo en estadística descriptiva porque proporciona unidades de medición elevadas al cuadrado. Sin embargo, se utiliza muy frecuentemente en estadística inferencial.

RESUMEN

En este capítulo hemos analizado la tendencia central y la variabilidad de las distribuciones. Las medidas de tendencia central más comunes son la media aritmética, la mediana y la moda. La media aritmética proporciona el promedio de los datos y se calcula sumando los datos y dividiendo el total entre N . La mediana divide la distribución a la mitad y, por lo tanto, es el valor de la escala que se encuentra en el quincuagésimo punto percentil de la distribución. La moda es el dato más frecuente dentro de una distribución. La media posee propiedades particulares que la hacen, por amplio margen, la medida de tendencia central de uso más común. Sin embargo, si la distribución está muy sesgada se deberá usar la mediana en lugar de la media, pues resulta menos afectada por los datos extremos. Además de presentar estas medidas, mostramos la forma de calcular cada una y revisamos sus propiedades más importantes.

También mostramos la manera de obtener la media global cuando se desea promediar varias medias. Por último, analizamos la relación entre la media, la mediana y la moda de una distribución y su simetría.

Las medidas de variabilidad más comunes son el rango, la desviación estándar y la varianza. El rango es una medida en bruto que indica únicamente la dispersión entre los dos datos más extremos. La desviación estándar es la medida de variabilidad que se utiliza más a menudo. Proporciona la dispersión promedio en torno a la media de la distribución. La varianza es simplemente el cuadrado de la desviación estándar. Igual que en el caso de las medidas de tendencia central, nuestro análisis de la variabilidad incluyó la forma de calcular cada medida. Por último, en vista de que la desviación estándar es la medida de variabilidad más importante, presentamos también sus propiedades.

TÉRMINOS IMPORTANTES

Desviación estándar (página 76)	Mediana (página 72)	Suma de cuadrados (página 78)
Dispersión (página 76)	Moda (página 74)	Tendencia central (página 67)
Media aritmética (página 67)	Puntaje de desviación (página 76)	Variabilidad (página 66)
Media global (página 70)	Rango (página 76)	Varianza (página 82)

PREGUNTAS Y PROBLEMAS

- Defina o identifique los conceptos de la sección “Términos importantes”.
- Enuncie cuatro propiedades de la media e ilustre cada una de ellas por medio de un ejemplo.
- ¿En qué condiciones preferiría usted utilizar la mediana, en lugar de la media, como la mejor medida de tendencia central? Explique por qué.
- ¿Por qué la moda no se utiliza mucho como medida de tendencia central?
- La media global (\bar{X}_{global}) es una media ponderada. ¿Es correcto este enunciado? Explique.
- Comente la relación entre la media y la mediana para distribuciones simétricas y asimétricas.
- ¿Por qué el rango no es una medida de dispersión tan útil como la desviación estándar?
- La desviación estándar es una medida relativa de la dispersión promedio. ¿Es correcto este enunciado? Explique.
- ¿Por qué usamos $N - 1$ en el denominador para calcular s , mientras que utilizamos N en el denominador para calcular σ ?
- ¿Cuál es la ecuación para SC en términos de los datos en bruto? ¿Cuándo es útil esta ecuación?
- Enuncie tres propiedades de la desviación estándar.
- ¿Cómo se relacionan entre sí la varianza y la desviación estándar?

13. Si $s = 0$, ¿qué debe ser verdad acerca de los datos de la distribución? Compruebe su respuesta por medio de un ejemplo.
14. ¿Es posible que el valor del rango, la desviación estándar o la varianza de un conjunto de datos sea negativo? Explique.
15. Proporcione el símbolo correspondiente a cada uno de los siguientes conceptos:
- Media de una muestra.
 - Media de una población.
 - Desviación estándar de una muestra.
 - Desviación estándar de una población.
 - Un dato en bruto.
 - Varianza de una muestra.
 - Varianza de una población.
16. Calcule la media, la mediana y la moda de los siguientes datos:
- 5, 2, 8, 2, 3, 2, 4, 0, 6
 - 30, 20, 17, 12, 30, 30, 14, 29
 - 1.5, 4.5, 3.2, 1.8, 5.0, 2.2
17. Calcule la media del siguiente conjunto de datos de una muestra: 1, 3, 4, 6, 6.
- Sume a cada dato una constante igual a 2. Calcule la media de los nuevos valores. Generalice los resultados para responder a la pregunta: “¿qué efecto produce sobre la media el hecho de sumar una constante a cada dato?”.
 - Reste de cada dato una constante igual a 2. Calcule la media de los nuevos valores. Generalice los resultados para responder a la pregunta: “¿qué efecto produce sobre la media el hecho de restar de cada dato una constante?”.
 - Multiplique cada dato por una constante igual a 2. Calcule la media de los nuevos valores. Generalice los resultados para responder a la pregunta: “¿qué efecto produce sobre la media el hecho de multiplicar cada dato por una constante?”.
 - Divida cada dato entre una constante igual a 2. Calcule la media de los nuevos valores. Generalice los resultados para responder a la pregunta: “¿qué efecto produce sobre la media el hecho de dividir cada dato entre una constante?”.
18. En un examen de biología se obtuvieron las siguientes calificaciones:

Calificaciones	f	Calificaciones	f
95–99	3	65–69	7
90–94	3	60–64	6
85–89	5	55–59	5
80–84	6	50–54	3
75–79	6	45–49	2
70–74	8		

- ¿Cuál es la mediana para este examen?
- ¿Cuál es la moda? educación

19. Utilice los datos que aparecen en la tabla 3.5 para:
- Determinar la mediana.
 - Determinar la moda.
20. Utilice los datos que aparecen en la tabla 3.6 para:
- Determinar la mediana.
 - Determinar la moda.
21. Para las siguientes distribuciones, indique si utilizaría la media o la mediana para representar la tendencia central de la distribución. Explique por qué.
- 2, 3, 8, 5, 7, 8
 - 10, 12, 15, 13, 19, 22
 - 1.2, 0.8, 1.1, 0.6, 25
22. Dados los siguientes valores de la tendencia central para cada distribución, determine si ésta es simétrica o si es asimétrica con sesgo positivo o negativo:
- Media = 14, mediana = 12, moda = 10
 - Media = 14, mediana = 16, moda = 18
 - Media = 14, mediana = 14, moda = 14
23. Una estudiante llevó un recuento del número de horas que estudió cada día en un periodo de dos semanas. Registró los siguientes datos diarios (medidos en horas): 2.5, 3.2, 3.8, 1.3, 1.4, 0, 0, 2.6, 5.2, 4.8, 0, 4.6, 2.8, 3.3. Calcule:
- La media del número de horas de estudio por día.
 - La mediana del número de horas de estudio por día.
 - La moda del número de horas de estudio por día. educación
24. Dos agentes de ventas que trabajan en la misma compañía tienen una discusión. Cada uno de ellos afirma que el número promedio de artículos que vendió, promediados durante el último mes, fue el más alto en la compañía. ¿Pueden tener ambos la razón? Explique. 1/0, otra
25. Un ornitólogo que estudia a la gaviota de alas glaucas, en Puget Sound, cuenta el número de interacciones agresivas por minuto, en un grupo de gaviotas, durante 9 minutos consecutivos. Así registró los siguientes datos: 24, 9, 12, 15, 10, 13, 22, 20, 14. Calcule:
- El número que representa la media de las interacciones agresivas por minuto.
 - El número que representa la mediana de las interacciones agresivas por minuto.
 - El número que representa la moda de las interacciones agresivas por minuto. biológica
26. Un especialista pone a prueba la velocidad de lectura de los niños de cuatro grupos en la asignatura de inglés de noveno grado. El grupo A tiene 42 estudiantes; el grupo B, 35; el grupo C, 33; y el grupo D, 39. La velocidad media de lectura, en palabras por minuto, para cada grupo fue la siguiente: grupo A, 220; grupo B, 185; grupo C, 212; y grupo

- D, 172. ¿Cuál es la velocidad media de lectura para todos los grupos combinados? educación
27. Para el siguiente conjunto de datos muestrales, calcule el rango, la desviación estándar y la varianza:
- 6, 2, 8, 5, 4, 4, 7
 - 24, 32, 27, 45, 28
 - 2.1, 2.5, 6.6, 0.2, 7.8, 9.3
28. En un curso de estadística en particular se llevaron a cabo tres exámenes. La calificación de cada estudiante estará basada en un promedio ponderado de las calificaciones que haya obtenido en esos exámenes. El primer examen tuvo un peso de 1, el segundo tuvo un peso de 2, y el tercero tuvo un peso de 2. Presentamos a continuación las calificaciones de uno de esos estudiantes. ¿Cuál será su promedio global?

Examen	1	2	3
Calificación	83	97	92

educación

29. El supervisor de una carrera utiliza un cronómetro para determinar los tiempos de llegada de los competidores. Cuando calcula el tiempo medio de los tres primeros finalistas, resulta que éste fue de 4.25 minutos. Después de revisar su cronómetro, descubre que, para su desgracia, lo puso en marcha cuando marcaba 15 segundos en lugar de 0, por lo cual cada uno de los tiempos fue medido con un excedente de 15 segundos. ¿Cuál fue entonces el tiempo medio correcto de los tres primeros finalistas? 1/0, otra
30. La fabricante de la marca A de zapatos para trotar desea averiguar cuánto duran éstos antes que sea necesario cambiarles las suelas. Con ese propósito toma sendas muestras, en forma aleatoria, entre usuarios de Chicago, Nueva York y Seattle. En Chicago, el tamaño de la muestra fue de 28 y la duración media antes del cambio de suelas fue de 7.2 meses. En Nueva York, el tamaño de la muestra fue de 35 y la duración media antes del cambio de suelas fue de 6.3 meses. En Seattle, el tamaño de la muestra fue de 22 y la duración media antes del cambio de suelas fue de 8.5 meses. ¿Cuál es la duración media global de los zapatos para trotar de la marca A antes que necesiten un cambio de suelas? 1/0, otra
31. Calcule la desviación estándar del siguiente conjunto de datos muestrales: 1, 3, 4, 6, 6.
- Sumé a cada dato una constante igual a 2. Calcule la desviación estándar para los nuevos valores. Generalice los resultados para responder a la pregunta: “¿Qué efecto produce sobre la desviación estándar el hecho de sumar una constante a cada dato?”.
 - Reste de cada dato una constante igual a 2. Calcule la desviación estándar para los nuevos valores. Generalice los resultados para responder a la pregunta: “¿Qué efecto produce sobre la desviación estándar el hecho de restar de cada dato una constante?”.
 - Multiplique cada dato por una constante igual a 2. Calcule la desviación estándar para los nuevos valores. Generalice los resultados para responder a la pregunta: “¿Qué efecto produce sobre la desviación estándar el hecho de multiplicar cada dato por una constante?”.
 - Divida cada dato entre una constante igual a 2. Calcule la desviación estándar para los nuevos valores. Generalice los resultados para responder a la pregunta: “¿Qué efecto produce sobre la desviación estándar el hecho de dividir cada dato entre una constante?”.
32. Una psicóloga industrial observó a ocho operadores de prensa taladradora durante 3 días de trabajo. Ella anotó el número de veces que cada operador oprimió el botón de “más rápido” en lugar del botón de “alto”, para determinar si el diseño del panel de control era un factor que contribuía a la alta tasa de accidentes en la planta. Dados los datos +, 7, 0, 2, 7, 3, 6, 7, calcule lo siguiente:
- La media.
 - La mediana.
 - La moda.
 - El rango.
 - La desviación estándar.
 - La varianza 1/0
33. Sin calcular en forma explícita la variabilidad, estudie las siguientes distribuciones muestrales:
- Distribución a: 21, 24, 28, 22, 20
 Distribución b: 21, 32, 38, 15, 11
 Distribución c: 22, 22, 22, 22, 22
- Ordene por rangos las distribuciones, según su propia estimación de la variabilidad relativa.
 - Calcule la desviación estándar de cada una para verificar el orden de rango asignado por usted.
34. Calcule la desviación estándar de los siguientes datos muestrales. ¿Por qué es tan grande en la parte **b**, en comparación con la parte **a**?
- 6, 8, 7, 3, 6, 4
 - 6, 8, 7, 3, 6, 35
35. Un psicólogo está interesado en los hábitos de los estudiantes de licenciatura en materia de citas románticas. Con ese propósito, elige una muestra de 10 estudiantes y determina el número de citas que tuvieron durante el mes pasado. A partir de los datos 1, 8, 12, 3, 8, 14, 4, 5, 8, 16, calcule lo siguiente:
- Media
 - Mediana
 - Moda
 - Rango
 - Desviación estándar
 - Varianza social
36. Un psicólogo cognitivo mide los tiempos de reacción de 6 sujetos ante palabras provistas de carga emocional. Los siguientes datos, en milisegundos, han sido registrados: 250, 310, 360, 470, 425, 270. Calcule lo siguiente:
- Media
 - Mediana
 - Moda
 - Rango
 - Desviación estándar
 - Varianza cognitiva

37. Un psicólogo biológico registra el número de células de una región particular del cerebro de gatos que responden a un estímulo táctil. En su estudio intervienen nueve gatos. Se han registrado los siguientes conteos de células por animal: 15, 28, 33, 19, 24, 17, 21, 34, 12. Calcule lo siguiente:
- Media
 - Mediana
 - Moda
 - Rango
 - Desviación estándar
 - Varianza biológica
38. ¿Qué le ocurre a la media de un conjunto de datos cuando:
- A cada dato del conjunto se le suma una constante a ?
 - Se resta de cada dato del conjunto una constante a ?
 - Se multiplica cada dato por una constante a ?
 - Se divide cada dato entre una constante a ?
- Ilustre cada uno de estos casos con un ejemplo numérico.
39. ¿Qué le ocurre a la desviación estándar de un conjunto de datos cuando:
- A cada dato del conjunto se le suma una constante a ?
 - Se resta de cada dato del conjunto una constante a ?
 - Se multiplica cada dato por una constante a ?
 - Se divide cada dato entre una constante a ?
- Ilustre cada uno de estos casos con un ejemplo numérico.
40. Suponga que, como se hace en algunas rifas, extraemos bolas al azar de un gran recipiente. Dicho recipiente contiene un considerable número de bolas, cada una de ellas marcada con un solo número, del 0 al 9. Existe la misma cantidad de bolas marcadas con cada número y todas ellas se revuelven incesantemente. Para este ejemplo extraigamos 10 muestras de tres bolas cada una. Cada muestra se forma al elegir las bolas, una a la vez, y devolviendo cada bola al recipiente antes de extraer la siguiente. El proceso de selección utilizado garantiza que cada una de las bolas del recipiente tenga la misma posibilidad de ser elegida en cada extracción. Supongamos que se han obtenido las siguientes muestras:
- | | | | | |
|---------|---------|---------|---------|---------|
| 1, 3, 4 | 2, 2, 6 | 3, 8, 8 | 1, 6, 7 | 5, 6, 9 |
| 3, 4, 7 | 1, 2, 6 | 2, 3, 7 | 6, 8, 9 | 4, 7, 9 |
- Calcule la media de cada muestra.
 - Calcule la mediana de cada muestra.
 - Tomando como base las propiedades de la media y de la mediana analizadas en este capítulo, ¿esperaría usted una mayor variabilidad en las medias o en las medianas? Verifique esto calculando la desviación estándar de las medias y de las medianas. otra

NOTAS

4.1 Para demostrar que $\sum (X_i - \bar{X}) = 0$,

$$\begin{aligned}
 \sum (X_i - \bar{X}) &= \sum X_i - \sum \bar{X} \\
 &= \sum X_i - N\bar{X} \\
 &= \sum X_i - N\left(\frac{\sum X_i}{N}\right) \\
 &= \sum X_i - \sum X_i \\
 &= 0
 \end{aligned}$$

4.2 Para demostrar que $SC = \sum X^2 - [(\sum X)^2/N]$,

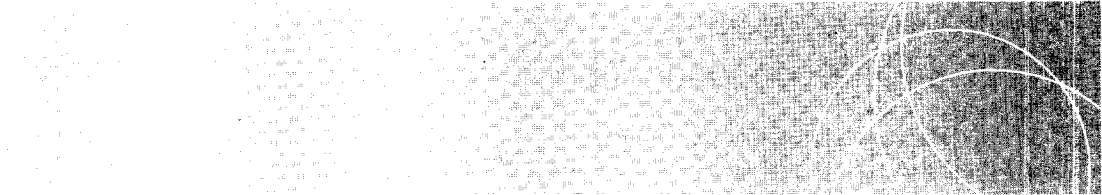
$$\begin{aligned}
 SC &= \sum (X - \bar{X})^2 \\
 &= \sum (X^2 - 2X\bar{X} + \bar{X}^2) \\
 &= \sum X^2 - \sum 2X\bar{X} + \sum \bar{X}^2 \\
 &= \sum X^2 - 2\bar{X} \sum X + N\bar{X}^2 \\
 &= \sum X^2 - 2\left(\frac{\sum X}{N}\right) \sum X + \frac{N(\sum X)^2}{N^2} \\
 &= \sum X^2 - \frac{2(\sum X)^2}{N} + \frac{(\sum X)^2}{N} \\
 &= \sum X^2 - \frac{(\sum X)^2}{N}
 \end{aligned}$$

Si desea encontrar problemas de práctica adicionales totalmente resueltos sobre los temas que acaba de estudiar, busque Chapter 4 en el material adicional del libro, disponible en inglés en: <http://www.thomsonlearning.com.mx>.

Si pretende resolver más problemas de práctica, relacionados con lo que se presentó en este capítulo y resueltos por medio de ecuaciones conceptuales, busque Chapter 4 en el material adicional del libro (en inglés), en: <http://www.thomsonlearning.com.mx>.

Si quiere resolver problemas para computadora sobre los temas que acaba de estudiar, busque Chapter 4 en el material adicional (en inglés) del libro, en: <http://www.thomsonlearning.com.mx>.

La curva normal y los puntajes estándar



Introducción

La curva normal

Área debajo de la curva normal

Puntajes estándar (puntajes z)

Características de los puntajes z

Determinación de las áreas correspondientes a cualquier dato en bruto

Determinación del dato en bruto correspondiente a un área dada

Resumen

Términos importantes

Preguntas y problemas



INTRODUCCIÓN

La curva normal es una distribución muy importante en las ciencias del comportamiento. Esto se explica por tres razones principales. La primera de ellas es que muchas de las variables que medimos en la investigación de las ciencias del comportamiento tienen distribuciones que se asemejan mucho a la curva normal. La estatura, el peso, la inteligencia y el rendimiento son unos cuantos ejemplos. En segundo lugar, muchas de las pruebas de inferencia que se utilizan en el análisis de los experimentos tienen distribuciones muestrales que adquieren una distribución normal cuando el tamaño de la muestra aumenta. La prueba de los signos y la prueba U de Mann-Whitney, que analizaremos posteriormente, son ejemplos de esto. Por último, muchas pruebas de inferencia requieren distribuciones muestrales que tengan semejanza con la curva normal (en el capítulo 12 nos ocuparemos de las distribuciones muestrales). La prueba z , la t de Student y la F son ejemplos de pruebas de inferencia que dependen de este aspecto. Así, gran parte de la importancia de la curva normal se percibe en conjunción con la estadística inferencial.



LA CURVA NORMAL

La curva normal es una distribución teórica de los datos de una población. Es una curva en forma de campana que puede ser descrita con la siguiente ecuación:

$$Y = \frac{N}{\sqrt{2\pi}\sigma} e^{-\frac{(X-\mu)^2}{2\sigma^2}} \quad \text{ecuación de la curva normal}$$

donde

- Y = frecuencia de un valor dado de X^*
- X = cualquier dato de la distribución
- μ = media de la distribución
- σ = desviación estándar de la distribución
- N = frecuencia total de la distribución
- π = constante cuyo valor aproximado es 3.1416
- e = constante cuyo valor aproximado es 2.7183

La mayoría de nosotros nunca necesitará conocer la ecuación exacta de la curva normal. La hemos presentado aquí para resaltar principalmente que es una curva teórica generada por procedimientos matemáticos. La figura 5.1 muestra un ejemplo de una curva normal.

Observe que la curva tiene dos puntos de inflexión, uno a cada lado de la media. Los puntos de inflexión se localizan allí donde la curvatura cambia de dirección. En la figura 5.1, los puntos de inflexión se localizan en el sitio donde la curva pasa de ser convexa hacia abajo a ser convexa hacia arriba. Si la curva en forma de campana es una curva normal, entonces los puntos de inflexión están ubicados a 1 desviación estándar de la media ($\mu + 1\sigma$ y $\mu - 1\sigma$). Tome note también de que, cuando la curva se aproxima al eje horizontal, su valor Y cambia muy lentamente. Desde el punto de vista teórico, la curva nunca llega a tocar al eje. Tiende hacia el eje horizontal y se acerca cada vez más a él, pero jamás lo toca. Se dice que la curva es *asintótica* al eje horizontal.

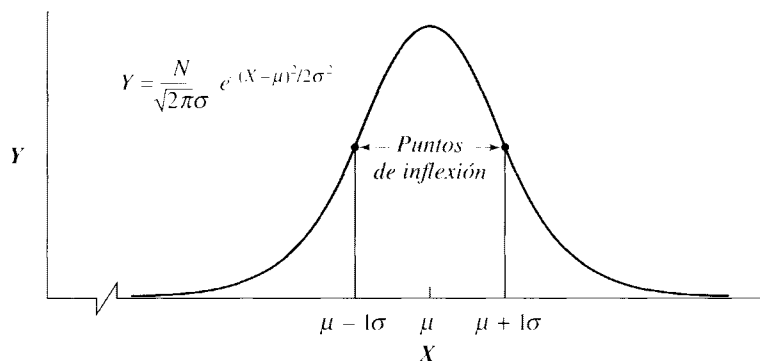


figura 5.1 Curva normal.

*Decir que Y es una "frecuencia" constituye una pequeña simplificación. Sin embargo, pienso que esta simplificación nos ayuda considerablemente a comprender y aplicar el material que veremos después. A lo que, en rigor, se le puede llamar con propiedad "frecuencia" es al área contenida debajo de la curva, entre dos valores X cualesquiera. Si desea encontrar una disertación acerca de este asunto, consulte *Statistical Reasoning in Psychology and Education*, 3a. ed., de E. Minium, B. King y G. Bear, John Wiley and Sons, Nueva York, 1993, pp. 123-124.

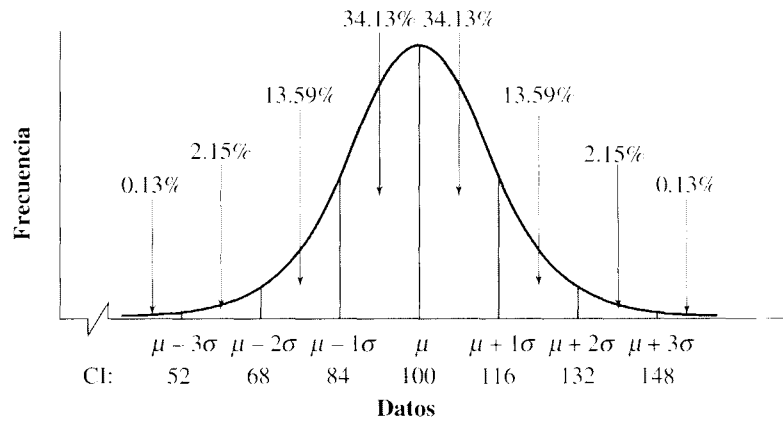


figura 5.2 Áreas debajo de la curva normal para datos seleccionados.

Área debajo de la curva normal

En las distribuciones cuya forma es normal, existe una relación especial entre la media y la desviación estándar con respecto al área contenida debajo de la curva. Cuando un conjunto de datos está distribuido en forma normal, 34.13% del área que se encuentra debajo de la curva está contenida entre la media (μ) y un dato que es igual a $\mu + 1\sigma$; 13.59% del área está contenida entre un dato igual a $\mu + 1\sigma$ y un dato de $\mu + 2\sigma$; 2.15% del área está contenida entre los datos de $\mu + 2\sigma$ y $\mu + 3\sigma$; y 0.13% del área se encuentra más allá $\mu + 3\sigma$. Esto representa 50% del área. En virtud de que la curva es simétrica, los mismos porcentajes son válidos para los datos que se localizan por debajo de la media. Estas relaciones aparecen en la figura 5.2. Como la frecuencia se localiza sobre el eje vertical, estos porcentajes representan el *porcentaje de los datos* contenidos dentro del área.

Por ejemplo, suponga usted que tenemos una población de 10 000 datos de CI. La distribución presenta una forma normal, con $\mu = 100$ y $\sigma = 16$. Puesto que los datos tienen una distribución normal, 34.13% de ellos se encuentra entre los valores de 100 y 116 ($\mu + 1\sigma = 100 + 16 = 116$), 13.59% se localiza entre 116 y 132 ($\mu + 2\sigma = 100 + 32 = 132$), 2.15% está entre 132 y 148; y 0.13% se encuentra por encima de 148. De manera análoga, 34.13% de los datos se ubica entre 84 y 100, 13.59% entre 68 y 84, 2.15% entre 52 y 68, y 0.13% está por debajo de 52. Estas relaciones se ilustran también en la figura 5.2.

Para calcular el número de datos que hay en cada área, lo único que debemos hacer es multiplicar el porcentaje correspondiente por la cantidad total de datos. Por lo tanto, existen $34.13\% \times 10\,000 = 3413$ datos entre 100 y 116, $13.59\% \times 10\,000 = 1359$ datos entre 116 y 132, y 215 datos entre 132 y 148; 13 datos son mayores que 148. Para la otra mitad de la distribución, existen 3413 datos entre 84 y 100, 1359 datos entre 68 y 84, y 215 datos entre 52 y 68; 13 datos se encuentran por debajo de 52. Observe que estas frecuencias sólo serían válidas si la distribución coincidiera exactamente con una curva normal. En la práctica, es posible que las frecuencias varíen un poco, según el grado en que la distribución se aproxime a este modelo teórico.



PUNTAJES ESTÁNDAR (PUNTAJES z)

Supongamos que alguien le dijo que el CI de usted es 132. ¿Estaría feliz o triste? Sin más información, es difícil decirlo. Un CI de 132 no tiene sentido, a menos que se cuente con un grupo de referencia con el cual podamos compararlo. Sin dicho grupo, usted no puede decir si el dato es alto, término medio o bajo. Para los fines de este ejemplo, suponga

que su puntaje es uno de los 10 000 que están representados en la distribución antes descrita. Ahora ya podemos comenzar a dar cierto significado a su CI de 132. Por ejemplo, podemos averiguar cuál es el porcentaje de datos de la distribución que son menores que 132. Sin duda reconocerá que esto equivale a calcular el rango percentil del dato 132. (Como usted seguramente recuerda, el rango percentil de un dato se define como el porcentaje de los datos que se encuentran por debajo del dato en cuestión.) En relación con la figura 5.2, podemos apreciar que el 132 se encuentra 2 desviaciones estándar más arriba de la media. En una curva normal, 34.13 + 13.59 = 47.72% de los datos se encuentran entre la media y un dato que se localiza 2 desviaciones estándar por encima de la media. Para obtener el rango percentil de 132, necesitamos sumar a este porcentaje el 50.00% que se encuentra más abajo de la media. En consecuencia, 97.72% (47.72 + 50.00) de los datos se localizan por debajo del CI de usted, que es de 132. Así pues, usted se debe sentir muy contento de ser tan inteligente. La solución se muestra en la figura 5.3.

Para resolver este problema, teníamos que determinar a cuántas desviaciones estándar estaba el dato bruto de 132, ya sea por arriba o por debajo de la media. Al hacer esto, transformamos el dato en bruto en un *puntaje estándar*, también conocido como *puntaje z* .

definición ■ Una **puntaje z** es un dato transformado que indica a cuántas unidades de desviación estándar, por encima o por debajo de la media, se encuentra un dato en bruto.

En forma de ecuación,

$$z = \frac{X - \mu}{\sigma} \quad \text{puntaje } z \text{ para los datos de una población}$$

$$z = \frac{X - \bar{X}}{s} \quad \text{puntaje } z \text{ para los datos de una muestra}$$

Para el ejemplo anterior,

$$z = \frac{X - \mu}{\sigma} = \frac{132 - 100}{16} = 2.00$$

El procedimiento por el cual un dato en bruto es modificado se llama *transformación de datos*. Veremos más adelante que la transformación z produce una distribución con media igual a 0 y una desviación estándar de 1. La razón por la cual a los puntajes z se los llama puntajes estándar es que éstos se expresan en relación con una distribución cuya media es 0 y cuya desviación estándar es 1.

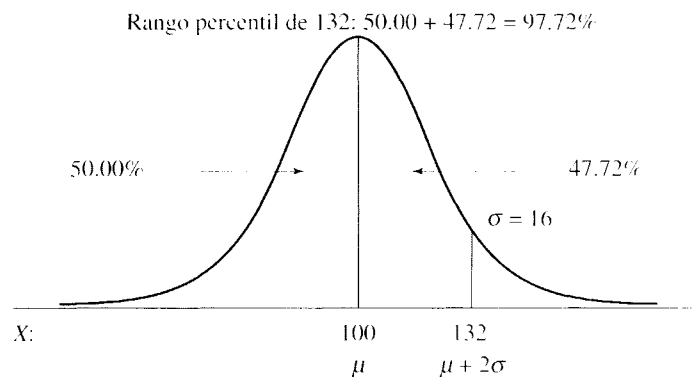


figura 5.3 Rango percentil de un CI de 132.

En combinación con una curva normal, los puntajes z nos permiten determinar el número o el porcentaje de los valores que están por encima o por debajo de cualquier dato en la distribución. Además, gracias a los puntajes z podemos hacer comparaciones entre datos de distribuciones distintas, aunque las unidades de estas distribuciones sean diferentes. Para ilustrar este punto, consideremos otro conjunto de datos poblacionales que posean una distribución normal. Supongamos que los pesos de todas las ratas que viven en un criadero universitario están distribuidos de manera normal, con $\mu = 300$ y $\sigma = 20$ gramos. ¿Cuál es el rango percentil de una rata que pesa 340 gramos?

La solución se muestra en la figura 5.4. En primer lugar, necesitamos convertir el dato en bruto de 340 gramos en el puntaje z correspondiente:

$$z = \frac{X - \mu}{\sigma} = \frac{340 - 300}{20} = 2.00$$

En virtud de que los datos están distribuidos de manera normal, $34.13 + 13.59 = 47.72\%$ de los datos se localizan entre el dato y la media. Al sumar el 50.00% restante que se encuentra por debajo de la media, obtenemos un rango percentil de $47.72 + 50.00 = 97.72\%$ para el peso de 340 gramos. Así pues, el dato de CI de 132 y el peso de la rata de 340 gramos tienen algo en común. Ambos ocupan la misma posición relativa en sus respectivas distribuciones. Tan pesada es la rata como inteligente es usted.

Este ejemplo, aunque un tanto jocoso, ilustra una aplicación importante de los datos z , a saber: nos permite comparar datos que no es posible cotejar directamente de ninguna otra manera. De ordinario no podríamos comparar la inteligencia con el peso. Los dos conceptos se miden con escalas distintas y tienen unidades diferentes. Sin embargo, al convertir los datos en sus datos z transformados, eliminamos las unidades originales y las reemplazamos por una unidad universal: la desviación estándar. Así, su dato de 132 unidades de CI se convierte en un dato de dos unidades de desviación estándar por encima de la media, y el peso de 340 gramos de la rata se transforma también en un dato de dos unidades de desviación estándar sobre la media. De esta manera es posible comparar “todo con todo”, mientras las escalas de medida permitan efectuar los cálculos de la media y la desviación estándar. La capacidad para comparar datos medidos con distintas escalas es de importancia fundamental para el tema de la correlación. Analizaremos este asunto con mayor detalle cuando nos ocupemos del tema en el capítulo 6.

Los ejemplos que hemos considerado hasta ahora se han referido a poblaciones. Sería útil practicar el cálculo de puntajes z utilizando datos muestrales. Esto lo haremos en el siguiente problema de práctica.

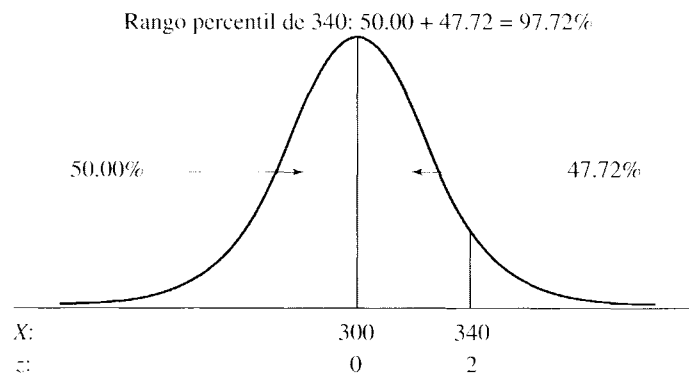


figura 5.4 Rango percentil de una rata que pesa 340 gramos.

Para cada uno de los datos en bruto de la muestra $X = 1, 4, 5, 7, 8$, determine el puntaje z correspondiente.

PASO 1: Determine la media de los datos en bruto.

$$\bar{X} = \frac{\sum X_i}{N} = \frac{25}{5} = 5.00$$

PASO 2: Determine la desviación estándar de los datos.

$$\begin{aligned} s &= \sqrt{\frac{SC}{N-1}} & SC &= \sum X^2 - \frac{(\sum X)^2}{N} \\ &= \sqrt{\frac{30}{4}} & &= 155 - \frac{(25)^2}{5} \\ &= 2.7386 & &= 30 \end{aligned}$$

PASO 3: Calcule el puntaje z para cada dato en bruto.

X	z
1	$z = \frac{X - \bar{X}}{s} = \frac{1 - 5}{2.7386} = -1.46$
4	$z = \frac{X - \bar{X}}{s} = \frac{4 - 5}{2.7386} = -0.37$
5	$z = \frac{X - \bar{X}}{s} = \frac{5 - 5}{2.7386} = 0.00$
7	$z = \frac{X - \bar{X}}{s} = \frac{7 - 5}{2.7386} = 0.73$
8	$z = \frac{X - \bar{X}}{s} = \frac{8 - 5}{2.7386} = 1.10$

Características de los puntajes z

Vale la pena destacar tres características de los puntajes z . En primer lugar, *los puntajes z tienen la misma forma que el conjunto de datos en bruto*. La transformación de los datos en bruto en sus puntajes z correspondientes no cambia la forma de la distribución. Tampoco las posiciones relativas de los datos se modifican. Los valores de los datos son los únicos que varían. La figura 5.5 ilustra este punto, mostrando los datos CI y sus correspondientes puntajes z . Usted puede observar que, aunque hemos utilizado puntajes z junto con la distribución normal, no todas las distribuciones z tienen la forma normal. Si usamos la ecuación z presentada anteriormente, los datos z pueden ser calculados para distribuciones de cualquier forma. Los puntajes z resultantes adoptarán la misma forma que los datos en bruto.

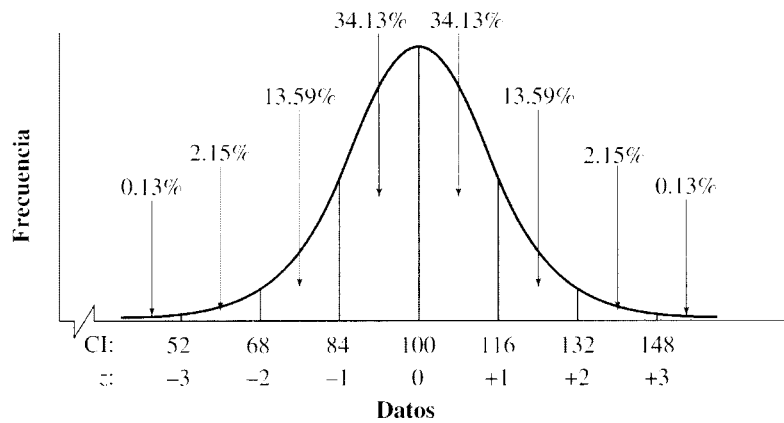


figura 5.5 Datos en bruto de CI con sus puntajes z correspondientes.

En segundo lugar, *la media de los puntajes z siempre es igual a cero* ($\mu_z = 0$). Esto es consecuencia de la observación según la cual los datos localizados en la media de los datos en bruto será también la media de los puntajes z (véase la figura 5.5). El valor z para los datos en bruto en la media es igual a cero. Por ejemplo, la transformación z para un dato ubicado en la media de la distribución de CI está dada por $z = (X - \mu) / \sigma = (100 - 100) / 16 = 0$. Por consiguiente, la media de la distribución z es igual a cero. La última característica de importancia es que *la desviación estándar de los puntajes z siempre es igual a 1* ($\sigma_z = 1$). Esto se debe a que un dato en bruto que se localiza 1 desviación estándar por arriba de la media tiene un puntaje z de +1:

$$z = \frac{(\mu + 1\sigma) - \mu}{\sigma} = 1$$

Determinación de las áreas correspondientes a cualquier dato en bruto

En los ejemplos anteriores acerca del CI y el peso, elegimos cuidadosamente el puntaje z para que la solución pudiera encontrarse a partir de la figura 5.2. Sin embargo, suponga que, en lugar de un CI de 132, queremos determinar el rango percentil de un CI de 142. Supongamos que persisten los mismos parámetros poblacionales. La solución aparece en la figura 5.6. En primer lugar, trace una curva que muestre la población; localice el área correspondiente ubicando el dato 142 sobre el eje horizontal. Luego sombree el área deseada. A continuación, calcule z:

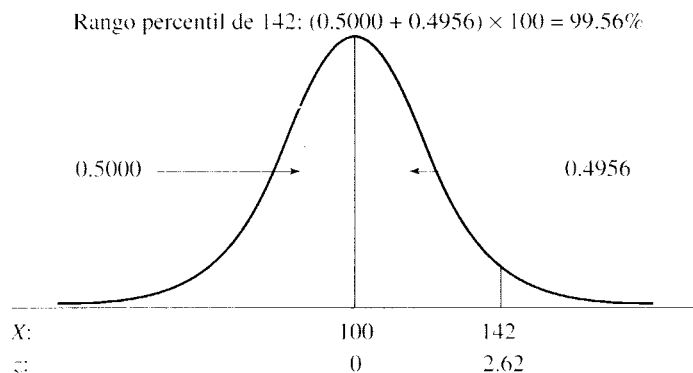


figura 5.6 Rango percentil de un CI de 142 en una distribución normal con $\mu = 100$ y $\sigma = 16$.

$$z = \frac{X - \mu}{\sigma} = \frac{142 - 100}{16} = \frac{42}{16} = 2.62$$

Como en las figuras 5.2 y 5.5 no se muestra el porcentaje correspondiente a un puntaje z de 2.62, no podemos utilizar ninguna de estas figuras para resolver el problema. Afortunadamente, las áreas que se encuentran debajo de la curva normal ya han sido calculadas para diversos puntajes z y los valores resultantes se presentan en la tabla A del apéndice D.

La primera columna de la tabla (columna A) contiene al puntaje z . La columna B enlista la proporción del área total que se encuentra entre un puntaje z dado y la media. La columna C enlista la proporción del área total que se localiza más allá del puntaje z .

Podemos utilizar la tabla A para determinar el rango percentil de 142. Primero localizamos el puntaje z de 2.62 en la columna A. Luego señalamos en la columna B la proporción del área total que se encuentra entre el puntaje z y la media. Para un puntaje z de 2.62, esta área es igual a 0.4956. A este valor le debemos sumar 0.5000 para tomar en cuenta los datos que están por debajo de la media (la gráfica nos ayuda a recordar que debemos hacer esto). Así, la proporción de los datos que están por debajo de un CI de 142 es $0.4956 + 0.5000 = 0.9956$. Para convertir esta proporción en un porcentaje, debemos multiplicarla por 100. De esta manera, el rango percentil de 142 es 99.56. Podemos utilizar la tabla A para determinar el área correspondiente a cualquier puntaje z , siempre que los datos estén distribuidos en forma normal. Cuando se utiliza la tabla A, generalmente basta redondear los valores de z hasta una precisión de dos cifras decimales. Resolvamos unos cuantos problemas ilustrativos adicionales para practicar.

Las calificaciones de un examen nacional de aptitudes matemáticas presentan una distribución normal, con $\mu = 80$ y $\sigma = 12$. ¿Cuál es el rango percentil de una calificación de 84?

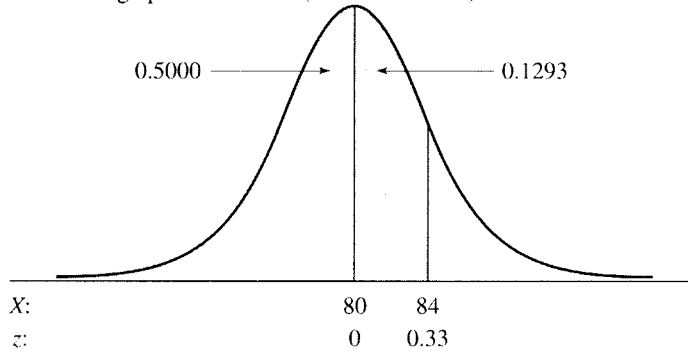
SOLUCIÓN

Al resolver problemas que corresponden a las áreas que se encuentran debajo de la curva normal, es conveniente trazar una gráfica de dicha curva desde un principio y localizar en ella las áreas correspondientes al problema. La figura siguiente muestra dicha gráfica. El área sombreada contiene todas las calificaciones menores que 84. Para determinar el rango percentil de 84, primero debemos convertir 84 en su puntaje z correspondiente:

$$z = \frac{X - \mu}{\sigma} = \frac{84 - 80}{12} = \frac{4}{12} = 0.33$$

Para calcular el área comprendida entre la media y un puntaje z de 0.33, consultamos la tabla A, localizamos el valor de z en la columna A y leemos el dato correspondiente en la columna B. Este valor es de 0.1293. Así pues, la proporción del área total comprendida entre la media y un puntaje z de 0.33 es de 0.1293. En la figura siguiente podemos ver que las calificaciones restantes debajo de la media ocupan una proporción de 0.5000 del área total. Si sumamos estas dos áreas, obtendremos la proporción de calificaciones que son menores que 84. Así, la proporción de calificaciones menores que 84 es $0.1293 + 0.5000 = 0.6293$. El rango percentil de 84 es, por consiguiente, $0.6293 \times 100 = 62.93$. *(continúa)*

Rango percentil de 84: $(0.5000 + 0.1293) \times 100 = 62.93\%$



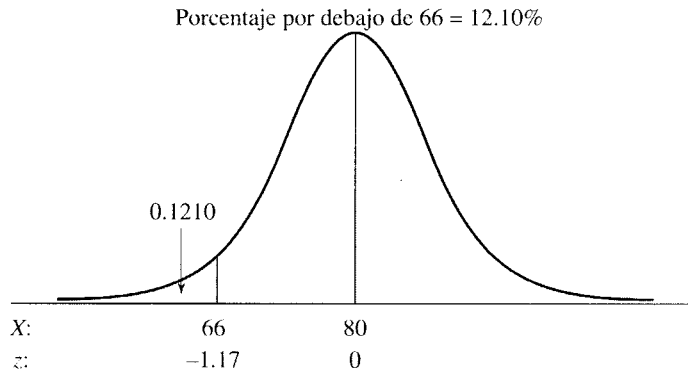
¿Qué porcentaje de las calificaciones de aptitud matemática se encuentra por debajo de una calificación de 66?

SOLUCIÓN

Una vez más, el primer paso consiste en trazar el diagrama apropiado. Éste se presenta en la figura adjunta. En este diagrama podemos apreciar que el área correspondiente (sombreada) se extiende más allá de la calificación 66. Para determinar el porcentaje de calificaciones contenidas en esta área, primero debemos convertir 66 en su puntaje z correspondiente. Así,

$$z = \frac{X - \mu}{\sigma} = \frac{66 - 80}{12} = \frac{-14}{12} = -1.17$$

Al leer la columna C de la tabla A, vemos que el área ubicada más allá de un puntaje z de 1.17 es 0.1210. Así pues, el porcentaje de las calificaciones que se encuentran por debajo de 66 es de $0.1210 \times 100 = 12.10\%$. La tabla A no muestra ningún puntaje z negativo. Sin embargo, esto no es un problema puesto que la curva normal es simétrica y a los puntajes z negativos les corresponde la misma proporción del área que a los puntajes z positivos de la misma magnitud. Por consiguiente, la proporción del área total que se encuentra “más allá” de un puntaje z de +1.17 es igual a la proporción que está más allá de un puntaje z de -1.17.



A partir de la misma población del problema de práctica 5.3, ¿qué porcentaje de las calificaciones se encuentran entre 64 y 90?

SOLUCIÓN

Al final del problema de práctica se presenta el diagrama correspondiente. Esta vez, las áreas sombreadas están a ambos lados de la media. Para resolver este problema, debemos determinar el área que se encuentra entre 64 y 80 y sumarla al área comprendida entre 80 y 90. Como antes, para hallar el área debemos calcular el puntaje z apropiado. Sin embargo, esta vez necesitamos calcular dos puntajes z. Para el área que se encuentra a la izquierda de la media:

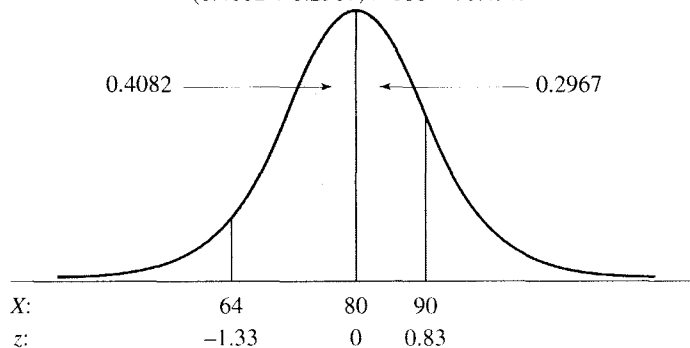
$$z = \frac{64 - 80}{12} = \frac{-16}{12} = -1.33$$

Para el área que se encuentra a la derecha de la media:

$$z = \frac{90 - 80}{12} = \frac{10}{12} = 0.83$$

En vista de que las áreas que queremos determinar están entre la media y el puntaje z, usaremos la columna B de la tabla A. El área que corresponde a un puntaje z de -1.33 es 0.4082 y la de un puntaje z de 0.83 es 0.2967. El área total es igual a la suma de estas dos áreas. Así pues, la proporción de las calificaciones comprendidas entre 64 y 90 es $0.4082 + 0.2967 = 0.7049$. El porcentaje de calificaciones ubicado entre 64 y 90 es $0.7049 \times 100 = 70.49\%$. Observe que en este problema no podemos restar simplemente 64 de 90 y luego dividir entre 12. Las áreas representadas en la tabla A han sido designadas con la media como punto de referencia. Por lo tanto, para resolver este problema tenemos que relacionar las calificaciones 64 y 90 con la media de la distribución. Notará usted también que no se puede restar sencillamente un valor z del otro, por el hecho de que la curva no es rectangular; en realidad, las magnitudes de las áreas que se encuentran debajo de dicha curva, correspondientes a los diversos puntos que la componen, son diferentes..

Porcentaje de calificaciones comprendidas entre 64 y 90:
 $(0.4082 + 0.2967) \times 100 = 70.49\%$



Otro tipo de problema se presenta cuando queremos determinar el área entre dos datos y los dos se localizan por debajo o por encima de la media. Tratemos de resolver un ejercicio de este tipo. Determinemos el porcentaje de las calificaciones del examen de aptitud que se encuentra entre las calificaciones 95 y 110.

SOLUCIÓN

La figura adjunta muestra la distribución y el área correspondiente. Igual que en el problema de práctica 5.4, no podemos restar sencillamente 95 de 110 y después dividir entre 12 para determinar el puntaje z adecuado, sino que es necesario utilizar la media como nuestro punto de referencia. En este problema tendremos que encontrar: 1) el área entre 110 y la media, y 2) el área entre 95 y la media. Una vez que hayamos restado estas dos áreas, obtendremos el área comprendida entre 95 y 110. Tal como lo hicimos antes, tendremos que calcular dos puntajes z :

$$z = \frac{110 - 80}{12} = \frac{30}{12} = 2.50 \quad \text{Transformación } z \text{ de } 110$$

$$z = \frac{95 - 80}{12} = \frac{15}{12} = 1.25 \quad \text{Transformación } z \text{ de } 95$$

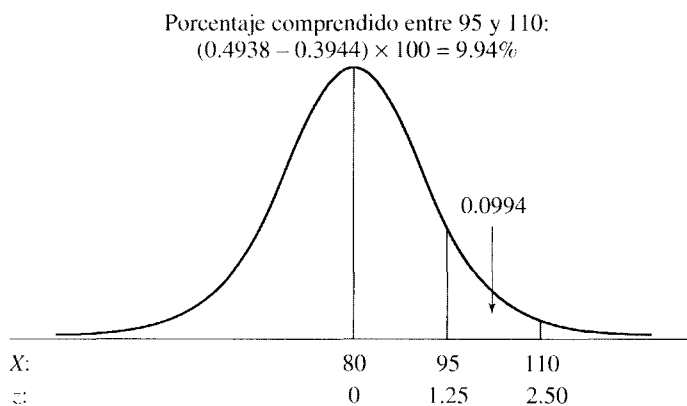
De la columna B de la tabla A,

$$\text{Área } (z = 2.50) = 0.4938$$

y también

$$\text{Área } (z = 1.25) = 0.3944$$

Así pues, la proporción de calificaciones que se encuentran entre 95 y 110 es $0.4938 - 0.3944 = 0.0994$. El porcentaje de calificaciones es $0.0994 \times 100 = 9.94\%$.



Determinación del dato en bruto correspondiente a un área dada

En algunas ocasiones ya conocemos el área y necesitamos saber cuál es el dato correspondiente. El siguiente problema es de este tipo. Encuentre usted el dato en bruto que divida la distribución de calificaciones del examen de aptitud matemática de tal manera que 70% de las calificaciones queden por debajo de él.

Este problema es exactamente el inverso del anterior. En este caso nos dan el área y nos piden determinar el dato. La figura 5.7 muestra el diagrama correspondiente. Aunque no conocemos el valor del dato en bruto, podemos determinar su puntaje z a partir de la tabla A. Una vez que conocemos el puntaje z , ya podemos despejar el dato en bruto mediante la ecuación de z . Si 70% de las calificaciones se encuentra por debajo del dato en bruto, entonces 30% debe estar por encima de él. Podemos determinar el puntaje z buscando en la columna C de la tabla A hasta localizar el área que más se aproxime a 0.3000 (30%). De esta manera veremos que el puntaje z correspondiente a esta área es 0.52. Para determinar el valor del dato en bruto, todo lo que necesitamos hacer es sustituir los valores correspondientes en la ecuación z y luego despejar X . Así.

$$z = \frac{X - \mu}{\sigma}$$

Al sustituir y despejar los términos en X , tenemos que:

$$0.52 = \frac{X - 80}{12}$$

$$X = 80 + 12(0.52) = 86.24$$

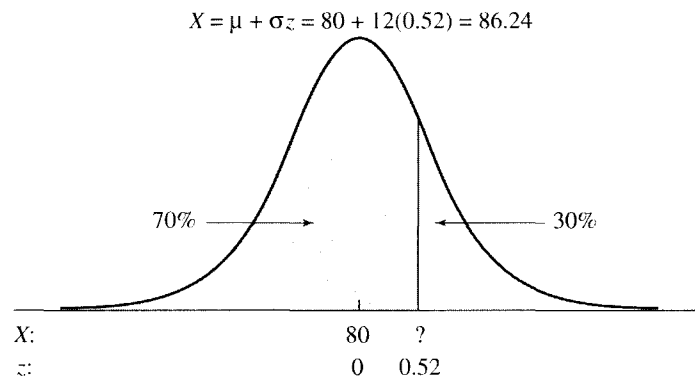


figura 5.7 Determinación de la calificación por debajo de la cual se encuentra 70% de la distribución, en una distribución normal con $\mu = 80$ y $\sigma = 12$.



Tratemos de resolver otro problema de este tipo. ¿Cuál es el dato que divide a la distribución de tal manera que 99% del área queda por debajo de él?

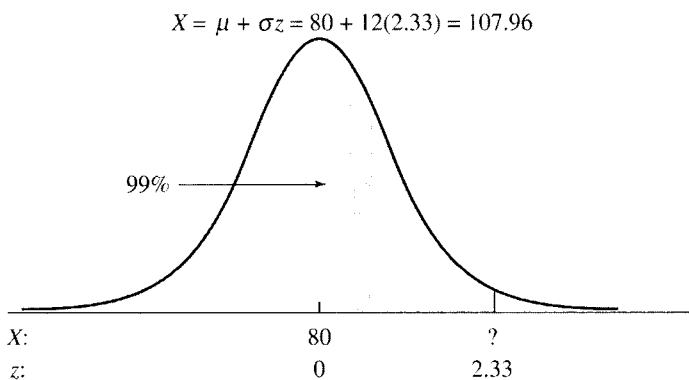
SOLUCIÓN

El diagrama correspondiente se presenta a continuación. Si 99% del área se encuentra por debajo del dato, entonces 1% deberá estar por encima de él. Para resolver este problema, localizamos el área en la columna C de la tabla A que más se aproxime a 0.0100 (1%) y allí observamos que $z = 2.33$. Para convertir el puntaje z en su dato en bruto correspondiente, sustituimos los valores apropiados en la ecuación de z y despejamos X . Así,

$$z = \frac{X - \mu}{\sigma}$$

$$2.33 = \frac{X - 80}{12}$$

$$X = 80 + 12(2.33) = 107.96$$





Resolvamos un problema más. ¿Cuáles son los datos que marcan los límites de 95% que se encuentra en la parte media de la distribución?

SOLUCIÓN

El diagrama se presenta a continuación. Existe un área de 2.5%, por arriba y por abajo de 95%, que se encuentra en la parte media. Para determinar los datos que marcan los límites de 95% de la distribución correspondiente a la porción media deberemos encontrar primero los valores z y los convertiremos a continuación en datos en bruto. Los puntajes z los encontramos en la tabla A, localizando en la columna C el área que más se aproxime a 0.0250 (2.5%) y leyendo el puntaje z asociado en la columna A. En este caso, $z = \pm 1.96$. Para encontrar los datos en bruto, sustituimos los valores apropiados en la ecuación z y despejamos X . Así,

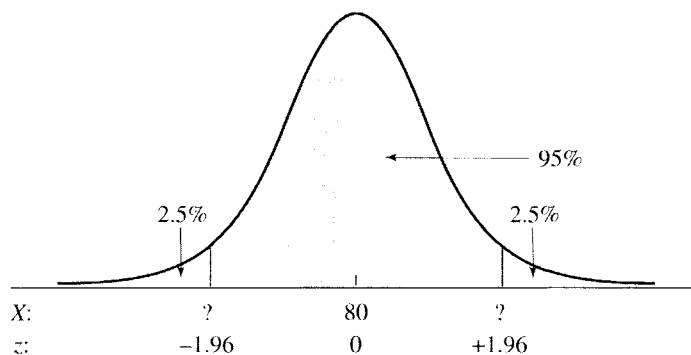
$$z = \frac{X - \mu}{\sigma}$$

$$-1.96 = \frac{X - 80}{12}$$

$$\begin{aligned} X &= 80 + 12(-1.96) \\ &= 56.48 \end{aligned}$$

$$+1.96 = \frac{X - 80}{12}$$

$$\begin{aligned} X &= 80 + 12(1.96) \\ &= 103.52 \end{aligned}$$



RESUMEN

En este capítulo hemos analizado la curva normal y los puntajes estándar. Comentamos que la curva normal tiene forma acampanada y presentamos la ecuación que la describe. A continuación, estudiamos el área que se encuentra debajo de la curva normal y su relación con los puntajes z . Un puntaje z es una transformación de un dato en bruto. Sirve para indicar a cuántas unidades de desviación estándar se encuentra el dato en bruto señalado, ya sea por encima o por debajo de la media. Una distribución z posee las siguientes características: 1) los puntajes z tienen la misma

forma que el conjunto de datos en bruto correspondientes; 2) la media de los puntajes z siempre es igual a cero y 3) la desviación estándar de los puntajes z siempre es igual a 1. Por último, hemos mostrado la manera de utilizar los puntajes z conjuntamente con una distribución normal, a fin de determinar: 1) el porcentaje o la frecuencia de datos correspondientes a cualquiera de los datos en bruto incluidos en la distribución y 2) el dato en bruto correspondiente a cualquiera de las frecuencias o los porcentajes de datos incluidos en la distribución.

TÉRMINOS IMPORTANTES

Asintótico (página 87)

Curva normal (página 87)

Puntajes estándar (puntajes z)
(página 88)

PREGUNTAS Y PROBLEMAS

1. Defina usted:
 - a. Asintótico.
 - b. La curva normal.
 - c. Los puntajes z .
 - d. Los puntajes estándar.
2. ¿Qué es una transformación de datos? Proponga un ejemplo.
3. ¿Cuáles son los valores de la media y de la desviación estándar de la distribución z ?
4. ¿Debe ser normal siempre la forma de una distribución z ? Explique.
5. ¿Son normales todas las distribuciones que tienen forma de campana? Explique.
6. Si los datos de un conjunto presentan una distribución normal, ¿qué información nos proporciona el área contenida debajo de la curva?
7. ¿Qué proporción de los datos de una distribución normal debe tener valores menores que $z = 0$? ¿Qué proporción de ellos tendrá valores mayores que $z = 0$?
8. Dado el conjunto de datos muestrales en bruto 10, 12, 16, 18, 19, 21,
 - a. Convierta cada dato en bruto en su valor z transformado.
 - b. Calcule la media y la desviación estándar de los puntajes z .
9. Suponga que los datos en bruto del problema 8 representan datos poblacionales; realice de nuevo los cálculos solicitados en los incisos **a** y **b**.
10. Una población de datos en bruto presenta una distribución normal con $\mu = 60$ y $\sigma = 14$. Determine los puntajes z para los siguientes datos en bruto tomados de esa población:

a. 76	b. 48
c. 86	d. 60
e. 74	f. 46
11. Para los siguientes puntajes z , determine el porcentaje de los datos que se encuentran más allá de z :

a. 0	b. 1
c. 1.54	d. -2.05
e. 3.21	f. -0.45
12. Para los siguientes puntajes z , determine el porcentaje de los datos que se encuentran entre la media y el puntaje z :

a. 1	b. -1
c. 2.34	d. -3.01
e. 0	f. 0.68
g. -0.73	
13. Para cada uno de los siguientes incisos, determine el puntaje z que divida la distribución, de tal manera que el porcentaje de puntajes dado se encuentre por encima del puntaje z (redondee a dos cifras decimales):

a. 50%	b. 2.50%
c. 5%	d. 30%
e. 80%	f. 90%
14. Considerando que una población de datos está normalmente distribuida, con $\mu = 110$ y $\sigma = 8$, determine lo siguiente:
 - a. El rango percentil de un dato de 120.
 - b. El porcentaje de datos que están por debajo de un dato de 99.

- c. El porcentaje de datos comprendidos entre un dato de 101 y 122.
- d. El porcentaje de datos comprendidos entre un dato de 114 y 124.
- e. El dato incluido en la población por encima del cual se encuentra 5% de los datos.
15. Al término de un trimestre en particular, Carol presentó cuatro exámenes finales. La media y la desviación estándar para cada examen, junto con la calificación obtenida por Carol en cada uno de ellos, se presentan a continuación. Supongamos que las calificaciones de cada examen presentan una distribución normal.

Examen	Media	Desviación estándar	Calificación de Carol
Francés	75.4	6.3	78.2
Historia	85.6	4.1	83.4
Psicología	88.2	3.5	89.2
Estadística	70.4	8.6	82.5

- a. ¿En cuál de los exámenes tuvo Carol mejor desempeño en relación con los demás estudiantes que lo presentaron?
- b. ¿Cuál fue su rango percentil en ese examen? educación
16. Uno de los hospitales de una gran ciudad registra el peso de todos los bebés que nacen en él. La distribución de pesos así obtenida presenta una forma normal, con una media $\mu = 2.9$ kilogramos y una desviación estándar $\sigma = 0.45$. Determine lo siguiente:
- a. El porcentaje de bebés que pesaron menos de 2.1 kilogramos.
- b. El rango percentil de un peso de 4.2 kilogramos.
- c. El porcentaje de bebés que pesaron entre 1.8 y 4.0 kilogramos.
- d. El porcentaje de bebés que pesaron entre 3.4 y 4.1 kilogramos.
- e. El peso que divida a la distribución de tal manera, que 1% de los pesos estén por encima de él.
- f. ¿Más allá de qué pesos se encuentra 5% de los datos más extremos?
- g. Si nacieron 15 000 bebés en el hospital, ¿cuántos pesaron menos de 3.5 kilogramos? salud, I/O
17. Una experta en estadística estudió los registros mensuales de lluvia en una región geográfica en particular. De esa manera determinó que la precipitación media mensual presentaba una distribución normal, con una media $\mu = 8.2$ centímetros

y una desviación estándar $\sigma = 2.4$. ¿Cuál es el rango percentil de los siguientes datos?

- a. 12.4 b. 14.3
- c. 5.8 d. 4.1
- e. 8.2 I/O, otra
18. Si se utilizan los mismos parámetros poblacionales del problema 17, ¿qué porcentaje de los datos están por encima de los siguientes valores?
- a. 10.5 b. 13.8
- c. 7.6 d. 3.5
- e. 8.2 I/O, otra
19. Utilizando los mismos parámetros poblacionales del ejercicio 17, ¿qué porcentaje de los datos se encuentra entre los siguientes valores?
- a. 6.8 y 10.2
- b. 5.4 y 8.0
- c. 8.8 y 10.5 I/O, otra
20. Un entusiasta del ejercicio lleva un registro del número de millas que trota cada semana. Los siguientes datos son una muestra de sus registros de 2002:

Semana	Distancia*	Semana	Distancia
5	32	30	36
8	35	32	38
10	30	38	35
14	38	43	31
15	37	48	33
19	36	49	34
24	38	52	37

*Los datos corresponden a las millas que corrió.

- a. Determine los puntajes z para las distancias que aparecen en la tabla. Observe que las distancias son datos muestrales.
- b. Grafique un polígono de frecuencias para los datos en bruto.
- c. Sobre la misma gráfica, trace un polígono de frecuencias para los puntajes z .
- d. ¿Tiene forma normal esta distribución z ? Si no la tiene, explique por qué.
- e. Calcule la media y la desviación estándar de la distribución z . I/O, otra
21. Un analista del mercado de valores bursátiles registró durante varios años el precio de venta diario de una acción selecta en particular. La distribución resultante de los datos tiene forma normal, con una media $\mu = \$84.10$ y una desviación estándar $\sigma = \$7.62$.
- a. Determine el porcentaje de los precios de venta que estuvieron por debajo de \$95.00.
- b. ¿Qué porcentaje de los precios de venta estuvieron entre \$76.00 y \$88.00?

- c. ¿Qué porcentaje de los precios de venta estuvieron por encima de \$70.00?
- d. ¿Qué precio de venta permite dividir a la distribución de tal manera que 2.5% de los datos quedan por encima de él? 1/0
22. Anthony tiene que decidir entre estudiar un posgrado en comercio o en derecho. Ya presentó los exámenes de aptitud, a nivel nacional, para ambas carreras. Las calificaciones de Anthony y también las normas nacionales se presentan a continuación. Tomando como base únicamente la posición relativa que ocupó Anthony en estos exámenes, ¿Cuál de esas especialidades deberá estudiar? Suponga que, para ambos exámenes, las distribuciones son normales.

Especialidad	Normas nacionales		Calificaciones de Anthony
	μ	σ	
Comercio	68	4.2	80.4
Derecho	85	3.6	89.8

educación

23. ¿En cuál de los siguientes exámenes tuvo Rebeca mejor desempeño? ¿Y Mauricio? Suponga que las calificaciones de los dos exámenes tuvieron una distribución normal.

	μ	σ	Calificaciones	Calificaciones
			de Rebeca	de Mauricio
Examen 1	120	6.8	130	132
Examen 2	50	2.4	56	52

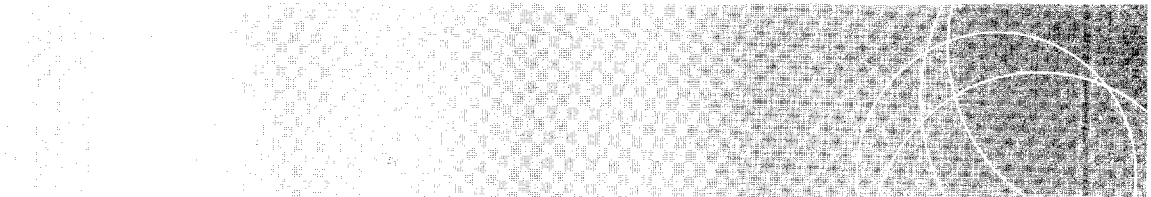
educación

24. Un psicólogo interesado en la inteligencia de los niños ha desarrollado una prueba estandarizada para seleccionar a los niños más "talentosos". Los puntajes de la prueba muestran una distribución normal, con $\mu = 75$ y $\sigma = 8$. Supongamos que a un niño talentoso se lo define como aquel que califica entre 1% superior de la distribución. ¿Cuál es el puntaje mínimo necesario para que se considere que un niño es talentoso? cognitiva, de desarrollo

Si desea encontrar problemas de práctica adicionales totalmente resueltos sobre los temas que acaba de estudiar, busque Chapter 5 en el material adicional del libro, disponible en inglés en: <http://www.thomsonlearning.com.mx>.

Si quiere resolver problemas para computadora sobre los temas que acaba de estudiar, busque Chapter 5 en el material adicional (en inglés) del libro, en: <http://www.thomsonlearning.com.mx>.

Correlación



Introducción

Relaciones

Relaciones lineales

Relaciones positivas y negativas

Relaciones perfectas e imperfectas

Correlación

El coeficiente de correlación lineal r de Pearson

Otros coeficientes de correlación

Efecto del rango sobre la correlación

Efecto del puntaje extremo

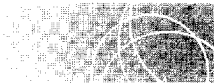
La correlación no implica causalidad

¿CUÁL ES LA VERDAD? "Un buen director = una buena escuela" ¿O no es así?

Resumen

Términos importantes

Preguntas y problemas



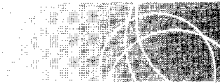
INTRODUCCIÓN

En los capítulos anteriores nos ocupamos sobre todo de las distribuciones individuales y la mejor forma de caracterizarlas. Además de describir estas distribuciones, con frecuencia es necesario determinar si los datos de una distribución se relacionan con los datos de otra. Por ejemplo, la persona encargada del departamento de recursos humanos de una gran empresa podría estar muy interesada en saber si existe alguna relación entre las calificaciones obtenidas en la universidad por los empleados y el éxito de éstos en la compañía. Si verdaderamente existe una relación fuerte entre estas dos variables, entonces las calificaciones obtenidas en la universidad se podrían usar para predecir el éxito en la empresa y, por lo tanto, serían muy útiles a la hora de seleccionar a los futuros empleados.

Además de la utilidad práctica que implica el uso de una relación para hacer predicciones, ¿por qué podría interesarnos saber si dos variables están relacionadas entre sí? Una razón importante de esto es que si las variables están relacionadas, es posible que una de ellas sea la causa de la otra.* Como veremos más adelante en este capítulo, el hecho de que dos variables estén relacionadas entre sí no es una base suficiente para demostrar la causalidad. Sin embargo, como los estudios de correlación son de los más fáciles de realizar, demostrar que existe una correlación entre las variables es con frecuencia el primer paso para demostrar que existe una relación causal entre ellas. Recíprocamente, si no existe correlación alguna entre las dos variables, entonces se puede descartar la idea de que exista una relación causal.

Otro uso muy importante de la correlación es la determinación de la “confiabilidad test-retest” de los instrumentos de prueba. La confiabilidad test-retest significa que existe consistencia en los puntajes obtenidos en aplicaciones repetidas de la prueba. Por ejemplo, suponiendo que el CI de un individuo permanezca estable de un mes a otro, esperaríamos que una buena prueba de CI mostrara una relación fuerte entre los puntajes obtenidos por el mismo individuo en dos administraciones de la prueba, realizadas con un mes de diferencia. Las técnicas correlacionales nos permiten cuantificar la relación entre los puntajes obtenidos en las dos aplicaciones y, de esta manera, medir la confiabilidad test-retest del instrumento.

La correlación y la regresión están muy relacionadas entre sí. Ambas implican la relación entre dos o más variables. La correlación sirve principalmente para averiguar si existe una relación y para determinar su magnitud y dirección, mientras que la regresión se refiere principalmente al hecho de utilizar la relación para hacer una predicción. En este capítulo analizaremos la correlación y en el capítulo 7 abordaremos el tema de la regresión lineal.



RELACIONES

La correlación se ocupa de establecer la magnitud y la dirección de las relaciones. Antes de profundizar en estos aspectos particulares de las relaciones, analizaremos algunas características generales de éstas, mediante las cuales podremos comprender mejor el material específico acerca de la correlación.

Relaciones lineales

Para iniciar nuestra exposición de las relaciones, veamos como ilustración una relación lineal entre dos variables. La tabla 6.1 muestra el salario mensual de cinco agentes de ventas y el valor en dólares de la mercancía vendida por cada uno de ellos en ese mes.

Podremos percibir mejor la relación entre estas variables si utilizamos los pares de valores X y Y que corresponden a cada agente de ventas, como los puntos para trazar una gráfica. La gráfica así obtenida se llama gráfica de *dispersión* o *dispersigrama*

definición ■ Una gráfica de **dispersión** o **dispersigrama** es una gráfica de pares de valores X y Y .

* Reconocemos que el tema de la causa y el efecto ha dado lugar a muchos debates filosóficos. Sin embargo, no podemos analizar todos los aspectos relacionados con ese tema. Al emplear el término causa, lo hacemos en la acepción de uso común que entienden quienes no son filósofos. Es decir, cuando decimos que A es la causa de B , queremos decir que una variación en A produjo una variación en B , estando controladas todas las demás variables en forma adecuada.

tabla 6.1 Salarios y mercancía vendida

Agente de ventas	Variable X mercancía vendida (\$)	Variable Y salario (\$)
1	0	500
2	1000	900
3	2000	1300
4	3000	1700
5	4000	2100

La gráfica de dispersión para los datos de los agentes de ventas aparece en la figura 6.1. En relación con esta figura, vemos que todos los puntos están sobre una línea recta. Cuando una línea recta describe la relación entre dos variables, se dice que esta relación es *lineal*.

definición ■ Una **relación lineal** entre dos variables es la que puede representarse con mayor exactitud por medio de una línea recta.

Observe que no todas las relaciones son lineales; algunas de ellas son *curvilíneas*. En esos casos, al trazar una gráfica de dispersión para las variables X y Y , una línea *curva* se ajusta mejor a los datos que una recta.

Obtención de la ecuación de la línea recta La relación entre “salario” y “mercancía vendida” que aparece en la figura 6.1 se puede describir por medio de una ecuación. Por

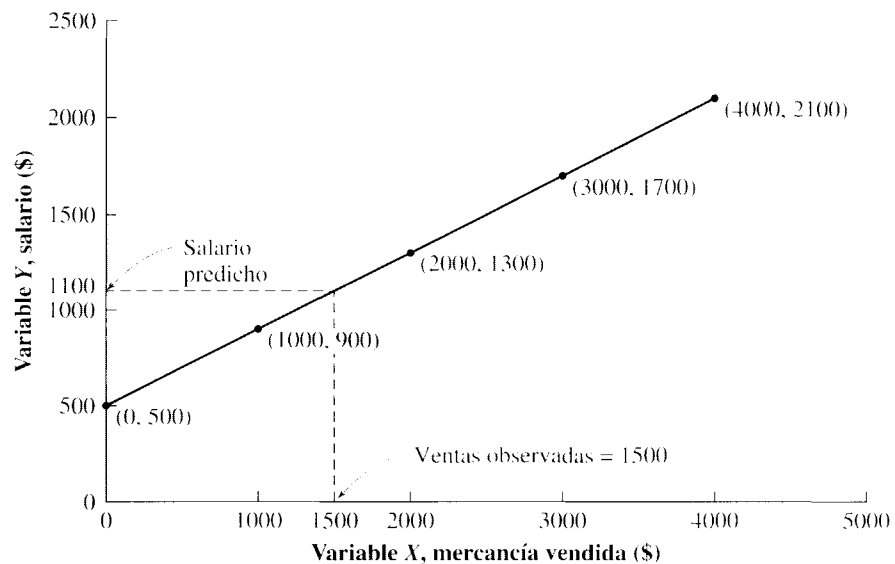


figura 6.1 Gráfica de la relación entre el salario y la mercancía vendida.

supuesto, esa ecuación es la que corresponde a la línea que une todos los puntos. La forma general de la ecuación está dada por

$$Y = bX + a \quad \text{ecuación de una línea recta}$$

donde $a = Y$ ordenada al origen (el valor de Y cuando $X = 0$)

$b =$ pendiente de la recta

Determinación de la Y ordenada al origen a La Y ordenada al origen es el valor de Y cuando la recta corta al eje vertical. Así pues, es el valor de Y cuando $X = 0$. Para este problema, podemos ver en la figura 6.1 que:

$$a = Y \text{ ordenada al origen} = 500$$

Determinación de la pendiente b La pendiente de una recta es una medida de su razón de cambio. Nos dice cuánto cambia un dato Y por cada cambio unitario en el dato X . Expresado como una ecuación:

$$b = \text{pendiente} = \frac{\Delta Y}{\Delta X} = \frac{Y_2 - Y_1}{X_2 - X_1} \quad \text{pendiente de una línea recta}$$

Como estamos trabajando con una línea recta, la pendiente es constante. Esto significa que no importa qué valores elijamos para X_2 y X_1 , los datos correspondientes de Y_2 y Y_1 , siempre resultarán en el mismo valor para la pendiente. Para calcular la pendiente hagamos variar X de 2000 a 3000. Si $X_1 = 2000$, entonces $Y_1 = 1300$. Si $X_2 = 3000$, entonces $Y_2 = 1700$. Sustituyendo estos valores en la ecuación de la pendiente:

$$b = \text{pendiente} = \frac{\Delta Y}{\Delta X} = \frac{Y_2 - Y_1}{X_2 - X_1} = \frac{1700 - 1300}{3000 - 2000} = \frac{400}{1000} = 0.40$$

Así, la pendiente es 0.40. Esto significa que el valor Y aumenta 0.40 de unidad por cada incremento unitario en X . Esta forma de determinar la pendiente y la ordenada al origen se muestra también en la figura 6.2. Observe que obtendríamos la misma pendiente

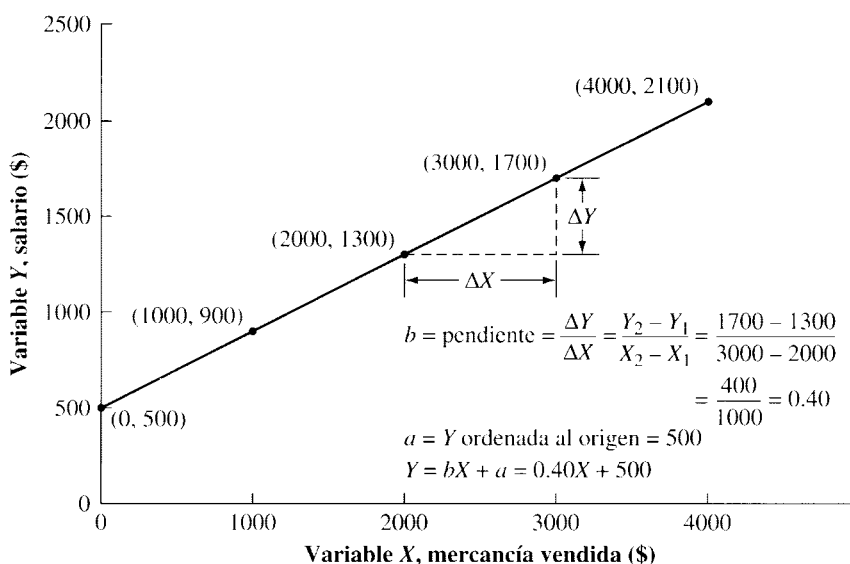


figura 6.2 Gráfica del salario y la cantidad de mercancía vendida.

si escogiéramos otros valores para X_1 y X_2 . Por ejemplo, si $X_1 = 1000$ y $X_2 = 4000$, entonces $Y_1 = 900$ y $Y_2 = 2100$. Al despejar la pendiente se obtiene:

$$b = \text{pendiente} = \frac{\Delta Y}{\Delta X} = \frac{Y_2 - Y_1}{X_2 - X_1} = \frac{2100 - 900}{4000 - 1000} = \frac{1200}{3000} = 0.40$$

Una vez más, la pendiente es 0.40.

Ahora podemos escribir la ecuación completa para la relación lineal que existe entre el salario y la mercancía vendida:

$$Y = bX + a$$

Al sustituir los valores de a y b , tenemos que:

$$Y = 0.40X + 500$$

La ecuación $Y = 0.40X + 500$ describe la relación entre la variable Y (salario) y la variable X (mercancía vendida). Nos dice que Y incrementa 0.40 por cada unidad de incremento en X . Además, mientras la relación se conserve, esta ecuación nos permite calcular un valor apropiado para Y , para cualquier valor de X dado. Esto hace que la ecuación sea muy útil para hacer predicciones.

Predicción de Y cuando X se conoce Cuando se utiliza para hacer una predicción, la ecuación se vuelve:

$$Y' = 0.40X + 500$$

donde Y' = a la predicción del valor de la variable Y

Con esta ecuación podemos predecir cualquier valor de Y si sólo conocemos el valor X . Por ejemplo, si $X = 1500$ como en nuestro problema anterior, entonces:

$$\begin{aligned} Y' &= 0.40X + 500 \\ &= 0.40(1500) + 500 \\ &= 600 + 500 \\ &= 1100 \end{aligned}$$

Así, si un agente vende \$1500 de mercancía, su salario será igual a \$1100.

Por supuesto, también podríamos haber hecho la predicción de manera gráfica, como se muestra en la figura 6.1. Al proyectar en forma vertical el valor X de \$1500 hasta que corte la línea recta, podemos leer el valor predicho de Y en el eje vertical. La predicción de ese valor es \$1100, que es el mismo valor que obtenemos utilizando la ecuación.

Relaciones positivas y negativas

Además de ser lineal o curvilínea, una relación entre dos variables puede ser *positiva* o *negativa*.

definición ■ Una **relación positiva** indica que existe una relación directa entre las variables. Una **relación negativa** existe cuando hay una relación inversa entre X y Y .

La pendiente de la recta indica si la relación es positiva o negativa. Cuando la relación es positiva, la pendiente es positiva. En el ejemplo anterior se presentó una pendiente positiva; es decir, los valores altos de X están asociados con los valores altos de Y , y los valores bajos de X están asociados con los valores bajos de Y . Cuando la pendiente es positiva, la recta asciende cuando uno la recorre de izquierda a derecha, lo cual indica que al aumentar X , aumenta también Y . En consecuencia, existe una relación directa entre las dos variables.

Cuando la relación es negativa, existe una relación inversa entre las variables, lo cual hace que la pendiente sea negativa. La figura 6.3 muestra un ejemplo de una relación negativa. Observe que cuando la pendiente es negativa, la curva “desciende” cuando es recorrida de izquierda a derecha. Los valores bajos de X están asociados con valores altos de Y , mientras que los valores altos de X se relacionan con valores bajos de Y . Esto se puede expresar también diciendo que: a medida que X crece, Y decrece.

Relaciones perfectas e imperfectas

En las relaciones que hemos graficado hasta este momento, todos los puntos están sobre la línea recta. Cuando esto ocurre, la relación es *perfecta*. Por desgracia, las relaciones perfectas son escasas en las ciencias del comportamiento. Lo más común es encontrar relaciones imperfectas.

definición ■ Una **relación perfecta** es aquella en la cual existe una relación positiva o negativa para la cual todos los puntos se localizan sobre la recta. Una **relación imperfecta** es aquella en la cual existe una relación, pero no todos los puntos se localizan sobre la recta.

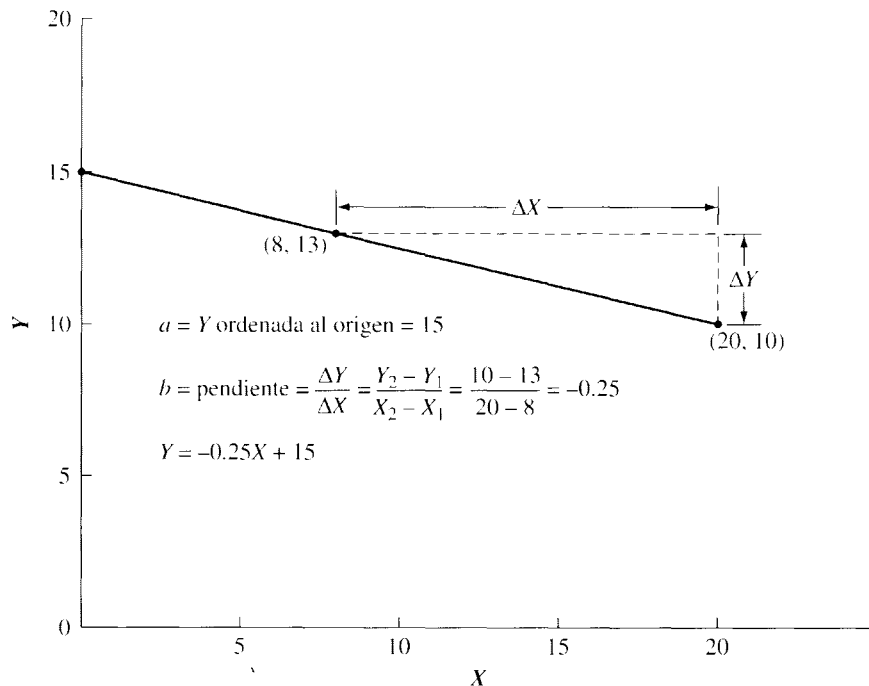


figura 6.3 Ejemplo de relación negativa.

tabla 6.2 CI y promedios de calificaciones de 12 estudiantes universitarios

Estudiante núm.	CI	Promedio de calificaciones
1	110	1.0
2	112	1.6
3	118	1.2
4	119	2.1
5	122	2.6
6	125	1.8
7	127	2.6
8	130	2.0
9	132	3.2
10	134	2.6
11	136	3.0
12	138	3.6

Por ejemplo, la tabla 6.2 muestra los puntos del CI y los promedios de calificaciones de una muestra de 12 estudiantes universitarios. Suponga que queríamos determinar si existe alguna relación entre estos datos hipotéticos. La gráfica de dispersión se presenta en la figura 6.4. Según la gráfica, es obvio que la relación entre el CI y las calificaciones escolares es imperfecta. La relación imperfecta es positiva, pues los valores bajos del CI

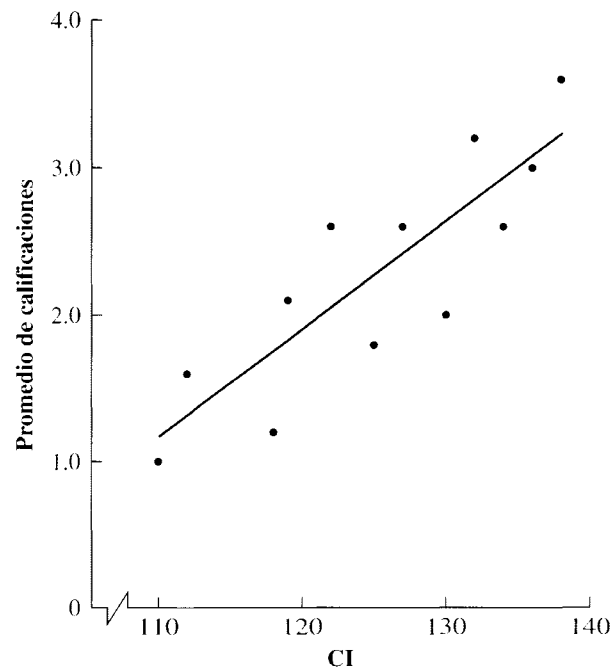


figura 6.4 Gráfica de dispersión del CI y del promedio de calificaciones.

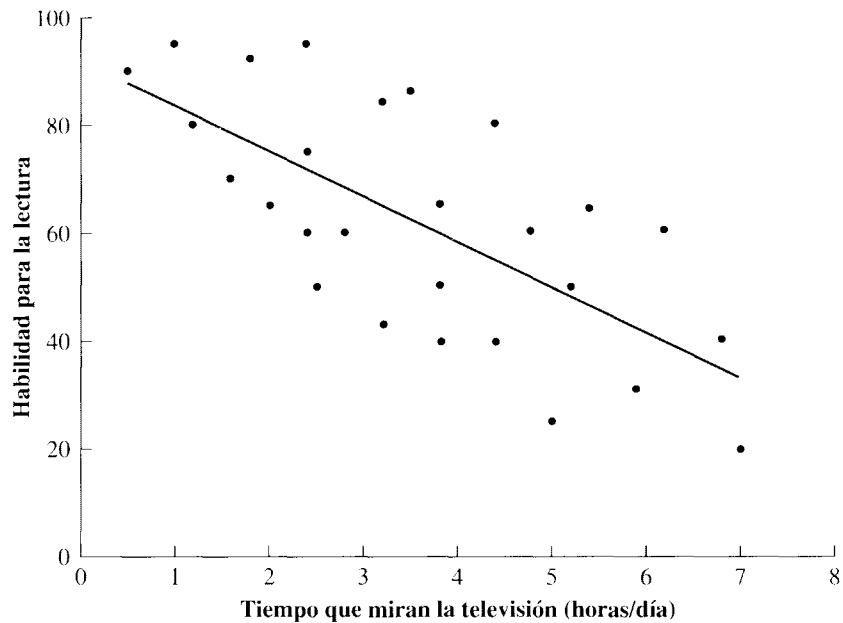


figura 6.5 Gráfica de dispersión de la habilidad para la lectura y la cantidad de tiempo que miran la televisión unos alumnos de sexto grado.

están asociados con valores bajos de las calificaciones y los valores altos del CI están asociados con valores altos de las calificaciones. Además, la relación parece ser lineal.

La mejor manera de describir esta relación por medio de una recta es trazar la línea recta que se ajuste mejor a los datos. En otras palabras, cuando la relación es imperfecta, no es posible trazar una sola línea recta que pase por todos los puntos. Sin embargo, sí podemos hacer la línea recta que mejor se ajuste a los datos.* Hemos trazado esta recta en la figura 6.4. Esta recta de mejor ajuste se usa frecuentemente con propósitos de predicción; cuando se utiliza de esta manera se llama línea de *regresión*.

Un artículo reciente de *USA Today* informó que existe una relación inversa entre la cantidad de tiempo que los alumnos de nivel primario dedican a mirar la televisión y su habilidad para la lectura. Suponga que los datos del artículo sobre los estudiantes de sexto grado aparecieron tal como se presentan en la figura 6.5. Éste es un ejemplo de una relación lineal negativa e imperfecta. La relación es negativa, ya que los valores altos de observación de la televisión están asociados con valores bajos de habilidad para la lectura, y los valores bajos de observación de la televisión están asociados con valores altos de habilidad para la lectura. La relación lineal es imperfecta porque no todos los puntos se localizan sobre una sola línea recta. La línea de regresión para estos datos se muestra también en la figura 6.5.

Una vez concluida nuestra exposición de los fundamentos de las relaciones, ya podemos pasar al tema de la correlación.

* Los detalles acerca de la realización de esta recta serán analizados en el capítulo 7.



CORRELACIÓN

La correlación es un tema que se centra en la *dirección* y el *grado* de la relación. La dirección de la relación se refiere a si ésta es positiva o negativa. El grado de la relación se refiere a su magnitud o fuerza. El grado de la relación puede variar entre inexistente y perfecto. Cuando la relación es perfecta, la correlación es máxima y podemos predecir con exactitud una variable a partir de la otra. En esta situación, cuando X cambia también Y presenta un cambio. Además, el mismo valor de X siempre conduce al mismo valor de Y . En forma alternativa, el mismo valor de Y siempre conduce al mismo valor de X . Todos los puntos están sobre una línea recta si la relación es lineal. Cuando la relación es inexistente, la correlación alcanza su mínimo, y el hecho de conocer el valor de una de las variables no ayuda en nada a predecir el valor de la otra. Las relaciones imperfectas tienen niveles intermedios de correlación y la predicción es aproximada. En este caso, el mismo valor de X no siempre conduce al mismo valor de Y . Sin embargo, en promedio, Y cambia de manera sistemática con X , y podemos lograr una predicción mejor de Y si conocemos el valor de X que si no lo conocemos.

Aunque para ciertos fines basta hablar, de manera vaga, de correlaciones “altas” o “bajas”, con bastante frecuencia es recomendable conocer la magnitud y dirección exactas de la correlación. Un *coeficiente de correlación* nos proporciona esta información.

definición ■ *Un coeficiente de correlación expresa de manera cuantitativa la magnitud y dirección de una relación.*

Un coeficiente de correlación puede variar de $+1$ a -1 . El *signo* del coeficiente nos indica si la relación es *positiva* o *negativa*. La parte numérica del coeficiente de correlación describe la magnitud de la correlación. Mientras mayor sea el número, mayor será la correlación. Como 1 es el mayor número posible, éste representa una correlación perfecta. Un coeficiente de correlación de $+1$ significa que la correlación es perfecta y que la relación es positiva. Un coeficiente de correlación de -1 significa que la correlación es perfecta y que la relación es negativa. Cuando la relación es inexistente, el coeficiente de correlación es igual a 0. Las relaciones imperfectas tienen coeficientes de correlación cuya magnitud varía entre 0 y 1. Tendrán el signo menos o más según sea la dirección de la relación.

La figura 6.6 muestra las gráficas de dispersión de varias relaciones lineales distintas y los coeficientes de correlación de cada una de ellas. Utilizamos el coeficiente de correlación r de Pearson, pues las relaciones son lineales. Analizaremos la r de Pearson en la siguiente sección. Cada gráfica de dispersión está formada por parejas de valores X y Y . Observe que si los puntos están más cerca de la línea de regresión, la magnitud del coeficiente de correlación será mayor y la predicción será más exacta. Además, cuando la correlación es 0, no existe relación entre X y Y . Esto significa que Y no crece ni decrece de manera sistemática cuando X aumenta o disminuye. Así, en el caso de la correlación cero, la recta de regresión para predecir Y es horizontal, y el conocimiento de X no nos ayuda a conocer el valor de Y .

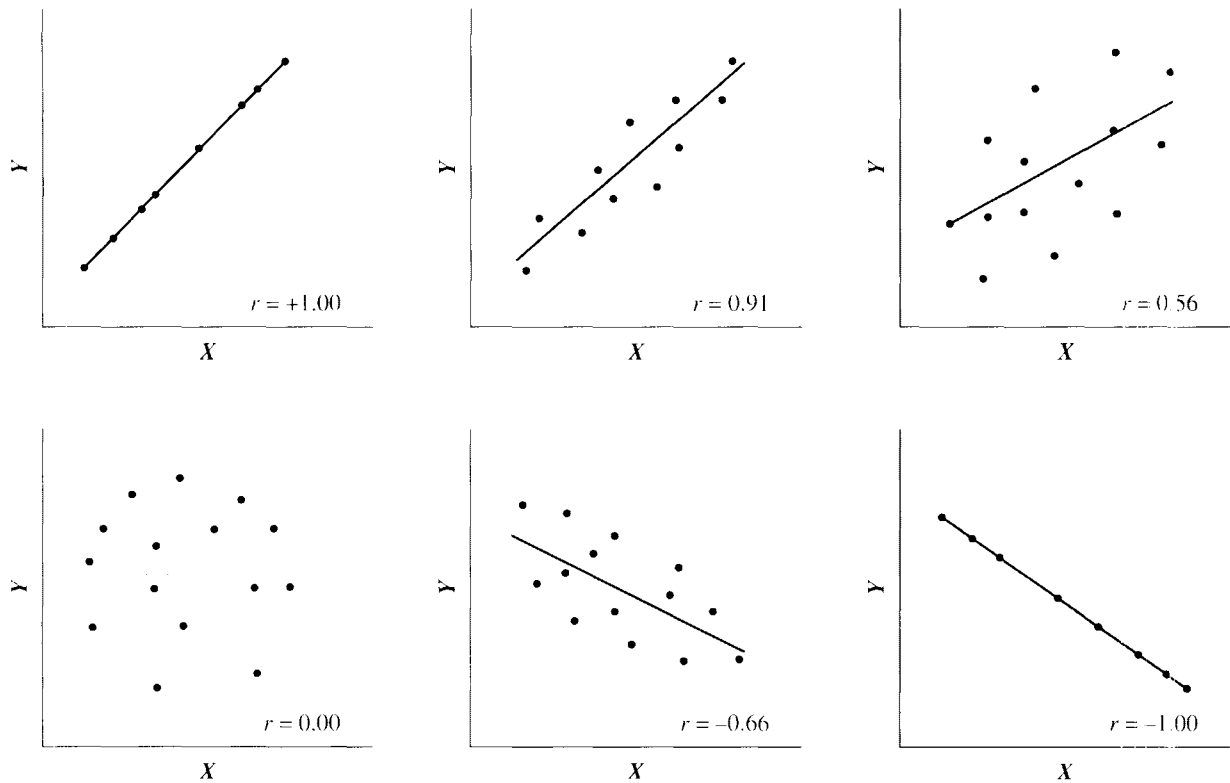


figura 6.6 Gráficas de dispersión de varias relaciones lineales.

El coeficiente de correlación lineal r de Pearson

Usted recordará, como se mencionó en el capítulo 5, que un problema básico para determinar la relación entre dos variables consiste en que, con mucha frecuencia, las variables han sido medidas en escalas y unidades diferentes. Por ejemplo, si nos interesa medir la correlación entre el CI y el promedio de calificaciones para los datos de la tabla 6.2, nos enfrentaríamos al problema de que ambos tienen escalas muy diferentes. Como mencionamos en el capítulo 5, podemos resolver esta dificultad al convertir cada calificación en su valor z transformado, lo cual colocaría a ambas variables en la misma escala, la escala z .

Para apreciar la utilidad de los puntajes z en la determinación de la correlación, consideremos el siguiente ejemplo. Suponga que el supermercado de su barrio tiene en venta naranjas empacadas; cada bolsa tiene marcado el precio total. Usted quiere saber si existe una relación entre el peso de las naranjas de cada bolsa y su costo. Como usted es un investigador nato, elige al azar seis bolsas y pesa cada una de ellas. El costo y el peso en libras de las seis bolsas aparecen en la tabla 6.3. La figura 6.7 muestra una gráfica de dispersión de los datos. ¿Están relacionadas estas dos variables? Sí; de hecho, todos los puntos están sobre una línea recta. Existe una correlación positiva perfecta entre el costo y el peso de las naranjas. Así, el coeficiente de correlación debe ser igual a $+1$.

Ahora veremos lo que ocurre al convertir estos datos en bruto en puntajes z . Los datos en bruto para el peso (X) y el costo (Y) están expresados como puntajes estándar en la cuarta y quinta columnas de la tabla 6.3. Ha ocurrido algo muy interesante: *las parejas de los datos en bruto para cada bolsa de naranjas tienen el mismo valor z* . Por ejemplo, las parejas de los datos en bruto para la bolsa A son 2.25 y 0.75. Sin embargo, sus

tabla 6.3 Costo y peso en libras de las seis bolsas de naranjas

Bolsa	Peso (libras) X	Costo (\$) Y	z_X	z_Y
A	2.25	0.75	-1.34	-1.34
B	3.00	1.00	-0.80	-0.80
C	3.75	1.25	-0.27	0.27
D	4.50	1.50	0.27	0.27
E	5.25	1.75	0.80	0.80
F	6.00	2.00	1.34	1.34

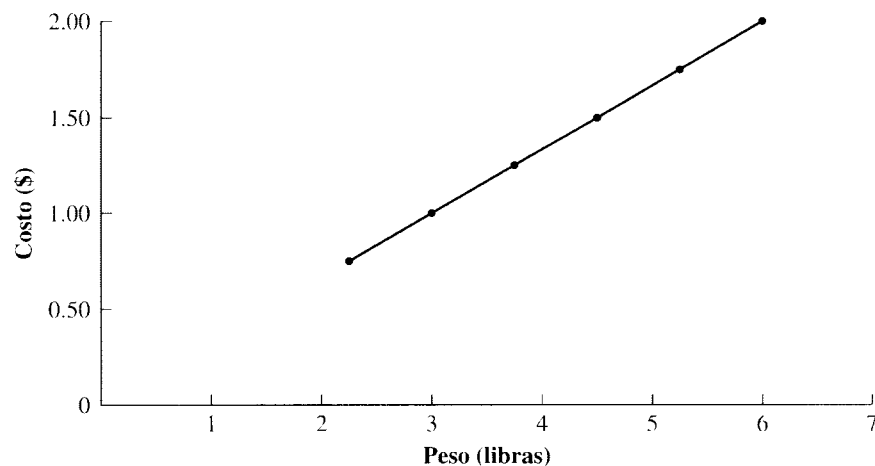
puntajes z respectivos son iguales a -1.34 . El dato en bruto de 2.25 está a tantas unidades de desviación estándar por debajo de la media de la distribución X como el dato en bruto de 0.75 está por debajo de la media de la distribución Y . Lo mismo es cierto para los otros pares de datos. Todos los pares de datos en bruto ocupan la misma posición relativa dentro de sus propias distribuciones. Es decir, tienen los mismos valores z . Al utilizar los datos en bruto, esta relación se oscurece debido a las diferencias de escala entre las dos variables. Si las parejas de datos ocupan la misma posición relativa dentro de sus propias distribuciones, entonces la correlación debe ser perfecta ($r = 1$) pues el hecho de conocer uno de los valores de la pareja de datos nos permite predecir de manera exacta al otro valor. Si la predicción es perfecta, la relación debe ser perfecta.

Esto nos lleva a la definición de la r de Pearson.

definición ■ *La r de Pearson es una medida del grado en el cual las parejas de datos ocupan posiciones iguales u opuestas dentro de sus propias distribuciones.*

Observe que esta definición también incluye a las parejas de datos que ocupan posiciones opuestas. Si las parejas de puntajes z tienen la misma magnitud, pero signos opuestos, la correlación sería perfecta de nuevo, y r sería igual a -1 .

Este ejemplo pone de relieve un aspecto muy importante. Como la correlación se refiere a la relación entre dos variables, y como es frecuente que las variables se midan en

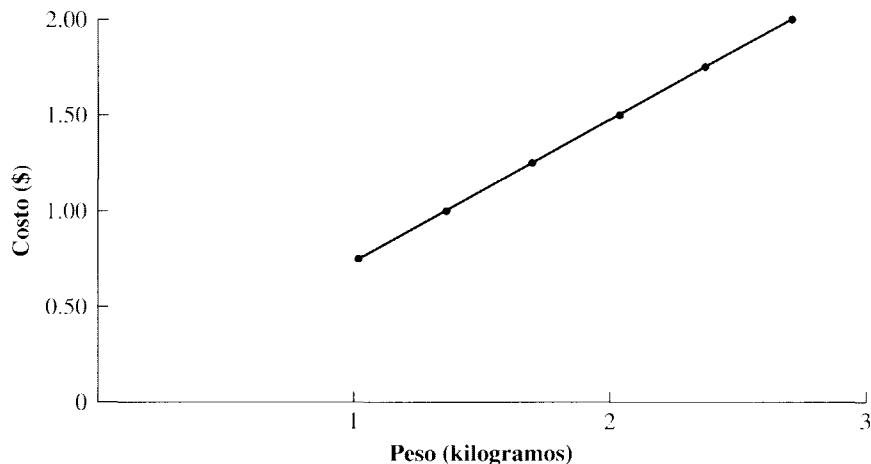
**figura 6.7** Costo de naranjas contra su peso en libras.

t a b l a 6.4 Costo y peso en kilogramos de las seis bolsas de naranjas

Bolsa	Peso (kilogramos) X	Costo (\$) Y	z_X	z_Y
A	1.02	0.75	-1.34	-1.34
B	1.36	1.00	-0.80	-0.80
C	1.70	1.25	-0.27	-0.27
D	2.04	1.50	0.27	0.27
E	2.38	1.75	0.80	0.80
F	2.72	2.00	1.34	1.34

diferentes escalas y unidades, la magnitud y la dirección del coeficiente de correlación debe ser independiente de las diferencias de escalas y de unidades que existan entre las dos variables. La r de Pearson logra esto mediante los puntajes z . Así, podemos correlacionar variables tan diversas como la hora del día y la posición del Sol, el porcentaje de grasa contenido en el cuerpo y el consumo de calorías, la ansiedad ante un examen y la calificación obtenida, etcétera.

Como el presente es un aspecto muy importante, queremos ilustrarlo una vez más, al llevar el ejemplo anterior un paso más adelante. En el ejemplo sobre la relación entre el costo de las naranjas y su peso, suponga que usted pesó las naranjas en kilogramos y no en libras. ¿Deberá cambiar el grado de relación entre el costo y el peso de las naranjas a causa de esto? En vista de lo que acabamos de exponer, la respuesta es un no rotundo. La correlación debe ser independiente de las unidades utilizadas para medir las dos variables. Si la correlación entre el costo de las naranjas y su peso en libras es igual a 1, entonces la correlación entre el costo de las naranjas y su peso en kilogramos también debe ser 1. Hemos convertido el peso de cada bolsa de naranjas de libras a kilogramos. Los datos correspondientes aparecen en la tabla 6.4 y la gráfica de los datos en bruto se presenta en la figura 6.8. De nuevo, todos los datos se localizan sobre una línea recta, de modo que la correlación es igual a 1.00. Observe los valores de las parejas de puntajes z en la cuarta y quinta columnas de la tabla 6.4. De nuevo se presentan los mismos valores, que son los mismos que cuando el peso de las naranjas fue expresado en libras. Así, el uso de los puntajes z permite medir la relación entre las dos variables, de manera independiente a las diferencias de escala y unidad empleadas para medir las variables.



f i g u r a 6.8 Costo de las naranjas contra su peso en kilogramos.

Cálculo de la r de Pearson La ecuación para calcular la r de Pearson mediante puntajes z es

$$r = \frac{\sum z_X z_Y}{N - 1}$$

donde $\sum z_X z_Y$ = es la suma de los productos de cada pareja de puntajes z

Para utilizar esta ecuación, primero hay que convertir cada dato en bruto en su valor z transformado. Esto puede ser muy tardado y da lugar a posibles errores de redondeo. Con un poco de álgebra, esta ecuación se puede transformar en una ecuación para hacer cálculos utilizando los datos en bruto:

$$r = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sqrt{\left[\sum X^2 - \frac{(\sum X)^2}{N}\right]\left[\sum Y^2 - \frac{(\sum Y)^2}{N}\right]}}$$

ecuación para el cálculo de la r de Pearson

donde $\sum XY$ = es la suma de los productos de cada pareja X y Y ($\sum XY$ se llama también suma de productos cruzados).

Si desea encontrar una solución en la que se use la ecuación conceptual (puntajes z), busque Chapter 6 en el material adicional del libro, en: <http://www.thomsonlearning.com.mx>.

La tabla 6.5 contiene algunos datos hipotéticos obtenidos de cinco sujetos, ahora vamos a utilizar estos datos para calcular la r de Pearson:

$$r = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sqrt{\left[\sum X^2 - \frac{(\sum X)^2}{N}\right]\left[\sum Y^2 - \frac{(\sum Y)^2}{N}\right]}}$$

tabla 6.5 Datos hipotéticos para el cálculo de la r de Pearson

Sujeto	X	Y	X^2	Y^2	XY
A	1	2	1	4	2
B	3	5	9	25	15
C	4	3	16	9	12
D	6	7	36	49	42
E	7	5	49	25	35
Total	21	22	111	112	106

$$r = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sqrt{\left[\sum X^2 - \frac{(\sum X)^2}{N}\right]\left[\sum Y^2 - \frac{(\sum Y)^2}{N}\right]}}$$

$$= \frac{106 - \frac{21(22)}{5}}{\sqrt{\left[111 - \frac{(21)^2}{5}\right]\left[112 - \frac{(22)^2}{5}\right]}}$$

$$= \frac{13.6}{18.616} = 0.731 = 0.73$$

ΣXY es la suma de los productos cruzados. Se determina multiplicando los datos X y Y para cada sujeto y sumando a continuación los productos resultantes. El cálculo de ΣXY y de los otros términos aparece ilustrado en la tabla 6.5. Al sustituir estos valores en la ecuación anterior, obtenemos:

$$r = \frac{106 - \frac{21(22)}{5}}{\sqrt{\left[111 - \frac{(21)^2}{5}\right]\left[112 - \frac{(22)^2}{5}\right]}} = \frac{13.6}{\sqrt{22.8(15.2)}} = \frac{13.6}{18.616} = 0.731 \approx 0.73$$

Si desea encontrar una solución en la que se use la ecuación conceptual (puntajes z), busque Chapter 6 en el material adicional del libro, en: <http://www.thomsonlearning.com.mx>.



Probemos con otro ejercicio. Esta vez utilizaremos los datos de la tabla 6.2. Para su comodidad, hemos reproducido estos datos en las primeras tres columnas de la tabla acompañante. En este ejemplo tenemos una relación lineal imperfecta y nos interesa calcular la magnitud y la dirección de la relación mediante la r de Pearson. La solución también aparece en la tabla siguiente.

SOLUCIÓN

Estudiante núm.	CI X	Promedio de calificaciones		X^2	Y^2	XY
		Y				
1	110	1.0		12 100	1.00	110.0
2	112	1.6		12 544	2.56	179.2
3	118	1.2		13 924	1.44	141.6
4	119	2.1		14 161	4.41	249.9
5	122	2.6		14 884	6.76	317.2
6	125	1.8		15 625	3.24	225.0
7	127	2.6		16 129	6.76	330.2
8	130	2.0		16 900	4.00	260.0
9	132	3.2		17 424	10.24	422.4
10	134	2.6		17 956	6.76	348.4
11	136	3.0		18 496	9.00	408.0
12	138	3.6		19 044	12.96	496.8
<i>Total</i>	1503	27.3		189 187	69.13	3488.7

$$r = \frac{\Sigma XY - \frac{(\Sigma X)(\Sigma Y)}{N}}{\sqrt{\left[\Sigma X^2 - \frac{(\Sigma X)^2}{N}\right]\left[\Sigma Y^2 - \frac{(\Sigma Y)^2}{N}\right]}}$$

$$= \frac{3488.7 - \frac{1503(27.3)}{12}}{\sqrt{\left[189187 - \frac{(1503)^2}{12}\right]\left[69.13 - \frac{(27.3)^2}{12}\right]}} = \frac{69.375}{81.088} = 0.856 \approx 0.86$$

Si desea encontrar una solución en la que se use la ecuación conceptual (puntajes z), busque Chapter 6 en el material adicional del libro, en: <http://www.thomsonlearning.com.mx>.

Tratemos de resolver otro problema. ¿Se ha puesto a reflexionar si es verdad que los opuestos se atraen? Todos hemos conocido alguna pareja formada por personas que parecen muy diferentes entre sí. Sin embargo, ¿es esto lo usual? ¿Qué es lo que fomenta la atracción: las diferencias o las similitudes?

Un psicólogo social que investigaba este problema pidió a 15 estudiantes que respondieran un cuestionario relacionado con sus actitudes hacia una amplia gama de temas. Tiempo después les mostró las “actitudes” de un desconocido hacia los mismos reactivos y les pidió que evaluaran su propio agrado o inclinación hacia dicho extraño y si consideraban probable que les resultara agradable trabajar con él. En realidad, las “actitudes” del desconocido habían sido inventadas por el experimentador y se variaron entre los individuos que participaron en el experimento, con respecto a la proporción de actitudes que cada uno compartía con el desconocido. De esta manera se obtuvieron datos para cada sujeto con respecto a sus actitudes y la atracción que sentía hacia el desconocido, con base en las actitudes de este último hacia los mismos reactivos. Si los iguales se atraen, entonces debe existir una relación directa entre la atracción que se siente hacia un desconocido y la proporción de actitudes que compartimos con él. Los datos se presentan en la tabla que aparece al final de este problema de práctica. Entre mayor sea la atracción, más alto será el puntaje. El máximo puntaje de atracción posible es 14. Calcule el coeficiente de correlación r de Pearson* para determinar si existe una relación directa entre la similitud de actitudes y el grado de atracción.

SOLUCIÓN

La solución aparece en la tabla siguiente.

Núm. de estudiante	Proporción de actitudes similares X	Atracción Y	X^2	Y^2	XY
1	0.30	8.9	0.090	79.21	2.670
2	0.44	9.3	0.194	86.49	4.092
3	0.67	9.6	0.449	92.16	6.432
4	0.00	6.2	0.000	38.44	0.000
5	0.50	8.8	0.250	77.44	4.400
6	0.15	8.1	0.022	65.61	1.215
7	0.58	9.5	0.336	90.25	5.510
8	0.32	7.1	0.102	50.41	2.272
9	0.72	11.0	0.518	121.00	7.920
10	1.00	11.7	1.000	136.89	11.700
11	0.87	11.5	0.757	132.25	10.005
12	0.09	7.3	0.008	53.29	0.657
13	0.82	10.0	0.672	100.00	8.200
14	0.64	10.0	0.410	100.00	6.400
15	0.24	7.5	0.058	56.25	1.800
<i>Total</i>	7.34	136.5	4.866	1279.69	73.273

* Como veremos en una sección posterior de este capítulo, el cálculo de la r de Pearson sólo es válido cuando los datos pertenecen a una escala de intervalo o de proporción. Por lo tanto, para calcular la r de Pearson en este problema, debemos suponer que los datos tienen al menos una escala de intervalo. (*continúa*)

$$\begin{aligned}
 r &= \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sqrt{\left[\sum X^2 - \frac{(\sum X)^2}{N}\right]\left[\sum Y^2 - \frac{(\sum Y)^2}{N}\right]}} \\
 &= \frac{73.273 - \frac{7.34(136.5)}{15}}{\sqrt{\left[4.866 - \frac{(7.34)^2}{15}\right]\left[1279.69 - \frac{(136.5)^2}{15}\right]}} \\
 &= \frac{6.479}{\sqrt{1.274(37.54)}} = \frac{6.479}{6.916} = 0.936 = 0.94
 \end{aligned}$$

Por lo tanto, tomando estos estudiantes como base, existe una relación muy fuerte entre las similitudes y la atracción.

Una segunda interpretación de la r de Pearson La r de Pearson se puede interpretar también en términos de la variabilidad de Y explicada por medio de X . Este punto de vista proporciona más información importante acerca de r y de la relación entre X y Y . Consideremos, por ejemplo, la figura 6.9, en la cual se muestra una relación imperfecta entre X y Y . En este ejemplo, la variable X representa el grado de destreza en ortografía, y la variable Y representa la habilidad en escritura de cada uno de seis estudiantes de tercer grado. Supongamos que se intenta predecir la calificación que obtendrá en escritura María, la estudiante cuya calificación en ortografía fue de 88. Si no existiera relación alguna entre la escritura y la ortografía, podríamos predecir una calificación de 50, que es la media general de todas las calificaciones de escritura. Si no existe relación en-

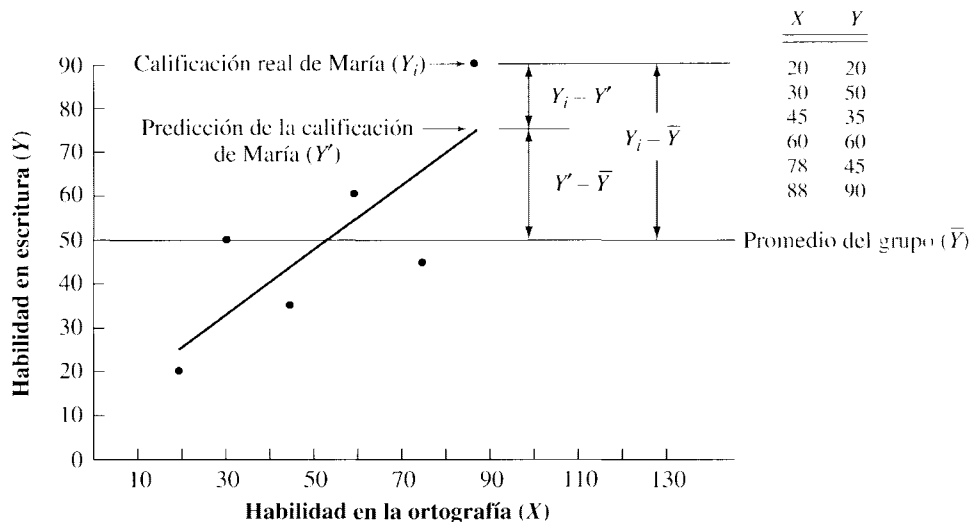


figura 6.9 Relación entre la ortografía y la escritura.

tre X y Y , la media general será la mejor predicción. Cuando no hay relación entre X y Y , el uso de la media minimiza los errores de predicción, puesto que la suma de los cuadrados de las desviaciones con respecto a ella es un mínimo. Usted reconocerá que ésta es la cuarta propiedad de la media, tal como fue analizada en el capítulo 4. Como quiera que la calificación real de María en escritura fue de 90, nuestra estimación de 50 tiene un error de 40 puntos. Así,

$$\text{Calificación real de María en escritura} - \text{Promedio del grupo} = Y_i - \bar{Y} = 90 - 50 = 40$$

Sin embargo, en este ejemplo la relación entre X y Y no es nula. Aunque no es perfecta, existe una relación mayor que cero entre X y Y . Por lo tanto, la media general de calificaciones en escritura no es la mejor predicción, sino, tal como lo hemos analizado en el capítulo, podemos utilizar la línea de regresión para estos datos como base de nuestra predicción. La línea de regresión para las calificaciones de escritura y ortografía se muestra en la figura 6.9. Utilizaremos esta línea para predecir una calificación en escritura de 75 para María. Ahora el error es de sólo 15 puntos. Así,

$$\left(\begin{array}{c} \text{Calificación real} \\ \text{de María} \end{array} \right) - \left(\begin{array}{c} \text{Predicción de la calificación} \\ \text{de María por medio de } X \end{array} \right) = Y_i - Y' = 90 - 75 = 15$$

En la figura 6.9 se puede observar que la distancia entre la calificación de María y la media de las calificaciones Y se puede separar en dos segmentos:

$$Y_i - \bar{Y} = (Y_i - Y') + (Y' - \bar{Y})$$

Desviación de Y_i = Error en predicción al usar la relación entre X y Y + La desviación de Y_i explicada mediante la relación entre X y Y

El segmento $Y_i - Y'$ indica el error en la predicción. El segmento restante $Y' - \bar{Y}$ representa la parte de la desviación de Y_i que se explica mediante la relación entre X y Y . Observe que “se explica mediante la relación entre X y Y ” se abrevia con frecuencia como “explicada por X ”.

Suponga que hemos determinado la predicción de la calificación Y (Y') para cada dato X mediante la línea de regresión. Entonces podríamos hacer $Y_i - Y$ para cada calificación. Si elevamos al cuadrado cada término $Y_i - Y$ y hacemos una sumatoria de todos estos datos, obtendríamos:

$$\Sigma (Y_i - \bar{Y})^2 = \Sigma (Y_i - Y')^2 + \Sigma (Y' - \bar{Y})^2$$

Variabilidad total de Y = Variabilidad de los errores de predicción + Variabilidad de Y explicada por X

Observe que $\Sigma (Y_i - \bar{Y})^2$ es la suma de los cuadrados de las calificaciones Y . Representa la variabilidad total de los datos Y . Así, esta ecuación establece que la variabilidad total de los datos Y se puede separar en dos partes: la variabilidad de los errores de predicción y la variabilidad de Y explicada por X .

Sabemos que cuando la relación es cada vez más fuerte, la predicción es cada vez más exacta. En la ecuación anterior, cuando la relación es más fuerte, los errores de predicción son menores, lo cual hace que la variabilidad de los errores de predicción

$\Sigma (Y_i - Y')^2$ también disminuya. Como la variabilidad total $\Sigma (Y_i - \bar{Y})^2$ no ha cambiado, la variabilidad de Y explicada por X , es decir, $\Sigma (Y' - \bar{Y})^2$, debe aumentar. Por lo tanto, la proporción de la variabilidad total de los datos Y que es explicada por X , a saber, $\Sigma (Y' - \bar{Y})^2 / \Sigma (Y_i - \bar{Y})^2$, es una medida de la fuerza de la relación. Se puede apreciar que si obtenemos la raíz cuadrada de esta razón y sustituimos en vez de Y' los valores adecuados, obtenemos la fórmula de cálculo para la r de Pearson. Ya antes habíamos definido a la r de Pearson como una medida del grado en que las parejas de datos ocupan posiciones idénticas u opuestas dentro de sus propias distribuciones. A partir de lo que hemos dicho, también se puede afirmar que la r de Pearson es igual a la raíz cuadrada de la proporción de la variabilidad de Y explicada por X . En forma de ecuación,

$$r = \sqrt{\frac{\Sigma(Y' - \bar{Y})^2}{\Sigma(Y_i - \bar{Y})^2}} = \sqrt{\frac{\text{Variabilidad de } Y \text{ explicada por } X}{\text{Variabilidad total de } Y}}$$

$$r = \sqrt{\text{Proporción de la variabilidad total de } Y \text{ explicada por } X}$$

Esta ecuación implica que si r crece, la proporción de la variabilidad de Y explicada por X aumenta también.

Relación de r^2 y la variabilidad explicada Si elevamos al cuadrado la ecuación anterior, obtenemos:

$$r^2 = \text{Proporción de la variabilidad total de } Y \text{ explicada por } X$$

Así, r^2 se conoce como el *coeficiente de determinación*. Como muestra la ecuación, r^2 es igual a la proporción de la variabilidad total de Y de la que es explicada por X . En el problema referente al promedio de calificaciones y el CI, la correlación era de 0.86. Si elevamos al cuadrado r , obtenemos:

$$r^2 = (0.86)^2 = 0.74$$

Esto significa que 74% de la variabilidad en Y puede ser explicado por el CI. Si resulta que el CI es un factor causal para determinar el promedio de las calificaciones, entonces r^2 nos señala que el CI explica 74% de la variabilidad observada en el promedio de calificaciones. ¿Y qué podemos decir del 26% restante? Deben existir otros factores relacionados con ese 26% restante que influyen sobre el promedio de las calificaciones. Lo importante aquí es que uno puede equivocarse al utilizar r pensando que X podría ser la causa principal de Y , cuando en realidad r^2 es la que nos indica cuál es la parte del cambio en Y que se puede atribuir a X . El error no es tan grave si se tiene un coeficiente de correlación tan alto como 0.86. Sin embargo, en las ciencias del comportamiento, esas correlaciones tan altas son poco comunes. Los coeficientes de correlación de $r = 0.50$ o 0.60 se consideran bastante altos; pero las correlaciones de esta magnitud sólo explican de 25 a 36% de la variabilidad en Y ($r^2 = 0.25$ a 0.36). La tabla 6.6 muestra la relación entre r y la variabilidad explicada, que se indica por medio de un porcentaje.

Otros coeficientes de correlación

Hasta ahora hemos analizado la correlación y describimos con cierto detalle el coeficiente de correlación lineal r de Pearson. Elegimos la r de Pearson por ser el coeficiente de correlación que aparece con más frecuencia en las investigaciones de las ciencias del comportamiento. Sin embargo, el lector debe observar que existen muchos coeficientes de correlación útiles, cada uno de los cuales es adecuado en diversas condiciones. Para decidir cuál coeficiente de correlación es más adecuado para el cálculo, es necesario considerar dos aspectos importantes, que son: la forma de la relación y la escala de medición en la que se expresan los datos.

tabla 6.6 Relación entre r y la variabilidad explicada

r	Variabilidad explicada
	r^2 (%)
0.10	1
0.20	4
0.30	9
0.40	16
0.50	25
0.60	36
0.70	49
0.80	64
0.90	81
1.00	100

Forma de la relación La elección del coeficiente de correlación apropiado para el cálculo depende de si la relación es lineal o curvilínea. Si los datos son curvilíneos, el uso de un coeficiente de correlación como la r de Pearson puede subestimar seriamente el grado de relación que existe entre X y Y . Según lo anterior, se utiliza otro coeficiente de correlación η (eta) para las relaciones curvilíneas. Como η no aparece a menudo en la investigación de las ciencias del comportamiento, no presentaremos un análisis de la misma.⁷ Sin embargo, esto pone de manifiesto la importancia de utilizar una gráfica de dispersión para determinar si la relación es lineal, antes de seguir adelante de manera rutinaria con el cálculo de un coeficiente de correlación lineal.

Escala de medición La elección del coeficiente de correlación depende también del tipo de escala de medición en que han sido expresados los datos. Ya hemos analizado el coeficiente de correlación lineal r de Pearson. En él se supone que los datos se miden en una escala de intervalo o de razón. Otros ejemplos de coeficientes de correlación lineal son: el coeficiente de correlación de rangos ordenados de Spearman rho (r_s), el coeficiente de correlación biserial (r_b), y el coeficiente phi (ϕ). En realidad, cada uno de estos coeficientes es la ecuación para la r de Pearson, simplificada para aplicarla a una escala de menor orden. Rho se utiliza cuando una o ambas variables tienen una escala ordinal, r_b se utiliza cuando una de las variables se mide al menos en intervalos y la otra es dicotómica, y phi se utiliza cuando cada una de las variables es dicotómica. Aunque la presentación detallada de cada coeficiente de correlación va más allá del alcance de este texto, el coeficiente de correlación de rangos ordenados de Spearman, rho, aparece con la frecuencia suficiente para que una exposición del mismo resulte justificada.

El coeficiente de correlación de rangos ordenados de Spearman rho (r_s) Tal como lo hemos mencionado, el coeficiente de correlación de rangos ordenados de Spearman, rho, se utiliza cuando una o ambas variables corresponden sólo a una escala ordinal. La rho de Spearman es en realidad el coeficiente de correlación lineal r de Pearson aplicado a datos que satisfacen los requisitos de una escala ordinal. La ecuación más sen-

⁷Una exposición sobre η y los demás coeficientes que aparecen en esta sección se encuentra en N. Downie y R. Heath, *Basic Statistical Methods*, cuarta edición, Harper & Row, Nueva York, 1974, pp. 102-114.

cilla para el cálculo de rho cuando no existen empates, o cuando existen muy pocos de ellos, con respecto al número de parejas de datos es:

$$r_s = 1 - \frac{6 \sum D_i^2}{N^3 - N} \quad \text{ecuación para el cálculo de rho}$$

donde D_i = diferencia entre el i -ésimo par de rangos = $R(X_i) - R(Y_i)$

$R(X_i)$ = rango del i -ésimo dato X

$R(Y_i)$ = rango del i -ésimo dato Y

N = número de parejas de rangos

Se puede demostrar que si entre los datos ordinales no existen empates, la r de Pearson se reduce algebraicamente a la ecuación anterior.

Para ilustrar el uso de rho, consideremos un ejemplo. Suponga que una gran corporación está interesada en calificar a un grupo de 12 aspirantes a gerentes, según su capacidad de liderazgo. Se ha contratado a dos psicólogos para realizar ese trabajo. Como resultado de sus exámenes y entrevistas, cada uno de los psicólogos, de manera independiente, ha ordenado por rangos a los aspirantes según su capacidad de liderazgo. Los rangos van de 1 a 12, donde 1 representa el nivel máximo de capacidad de liderazgo. Los datos se presentan en la tabla 6.7. ¿Cuál es la correlación entre los ordenamientos por rangos obtenidos por los dos psicólogos?

En virtud de que los datos tienen una escala ordinal, tenemos que calcular rho. La solución aparece en la tabla 6.7. Observe que los sujetos 5 y 6 quedaron empatados en el ordenamiento por rangos del psicólogo A. Cuando ocurren empates, la regla sagiere asignar a cada sujeto el promedio de los rangos empatados. Por ejemplo, los sujetos 5 y 6 quedaron empatados en los rangos 2 y 3. Por lo tanto, cada uno tuvo una clasificación de 2.5 [(2 + 3)/2 = 2.5]. Al asignar a los dos sujetos una clasificación de 2.5, utilizamos los rangos 2 y 3. El siguiente rango es 4. D_i es la diferencia entre las parejas de rangos para

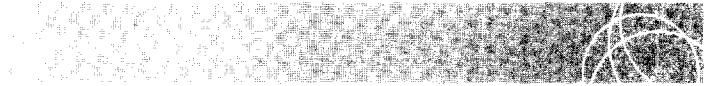
tabla 6.7 Cálculo de r_s para el ejemplo del liderazgo

Sujeto	Orden de rango del psicólogo A $R(X_i)$	Orden de rango del psicólogo B $R(Y_i)$	$D_i =$ $R(X_i) - R(Y_i)$	D_i^2
1	6	5	1	1
2	5	3	2	4
3	7	4	3	9
4	10	8	2	4
5	2.5	1	1.5	2.25
6	2.5	6	-3.5	12.25
7	9	10	-1	1
8	1	2	1	1
9	11	9	2	4
10	4	7	3	9
11	8	11	-3	9
12	12	12	0	0
$N=12$				$\sum D_i^2 = 56.5$

$$r_s = 1 - \frac{6 \sum D_i^2}{N^3 - N} = 1 - \frac{6(56.5)}{(12)^3 - 12} = 1 - \frac{339}{1716} = 0.80$$

el i -ésimo sujeto. De esta manera, $D_i = 1$ para el sujeto 1. No importa si se resta R de $R(Y_i)$ o $R(Y_i)$ de $R(X_i)$ para obtener D_i puesto que elevamos al cuadrado cada valor D_i . Luego se suman los cuadrados de las D_i ($\sum D_i^2 = 56.5$). Este valor se introduce después en la ecuación, junto con N ($N = 12$), para calcular r_s . Para este problema,

$$r_s = 1 - \frac{6 \sum D_i^2}{N^3 - N} = 1 - \frac{6(56.5)}{12^3 - 12} = 1 - \frac{339}{1716} = 0.80$$



Para ilustrar el cálculo de r_s , suponga que las calificaciones de las actitudes y el grado de atracción mutua planteadas en el problema de práctica 6.2 fuesen sólo de una escala ordinal. Bajo esta hipótesis, determine el valor del coeficiente de correlación lineal rho para estos datos y compárelo con la r de Pearson obtenida en el problema de práctica 6.2.

SOLUCIÓN

Los datos y la solución se presentan en la tabla siguiente.

Sujeto	Proporción de actitudes similares X_i	Atracción Y_i	Rango de X_i $R(X_i)$	Rango de Y_i $R(Y_i)$	$D_i = R(X_i) - R(Y_i)$	D_i^2
1	0.30	8.9	5	7	2	4
2	0.44	9.3	7	8	-1	1
3	0.67	9.6	11	10	1	1
4	0.00	6.2	1	1	0	0
5	0.50	8.8	8	6	2	4
6	0.15	8.1	3	5	-2	4
7	0.58	9.5	9	9	0	0
8	0.32	7.1	6	2	4	16
9	0.72	11.0	12	13	-1	1
10	1.00	11.7	15	15	0	0
11	0.87	11.5	14	14	0	0
12	0.09	7.3	2	3	-1	1
13	0.82	10.0	13	11.5	1.5	2.25
14	0.64	10.0	10	11.5	1.5	2.25
15	0.24	7.5	4	4	0	0
$N = 15$						$\sum D_i^2 = 36.5$

$$r_s = 1 - \frac{6 \sum D_i^2}{N^3 - N} = 1 - \frac{6(36.5)}{(15)^3 - 15} = 1 - \frac{219}{3360} = 0.93$$

Observe que $r_s = 0.93$ y que $r = 0.94$. Los valores no son idénticos, pero sí muy cercanos entre sí. En general, al calcular la r de Pearson utilizando las propiedades de intervalo o de razón de los datos, sus valores serán cercanos, aunque no exactamente iguales, a los que se obtendrían considerando únicamente las propiedades ordinales de aquellos datos.

Efecto del rango sobre la correlación

Si existe una correlación entre X y Y , la restricción del rango de cualquiera de las variables tendrá el efecto de *reducir* dicha correlación. Podemos ver esto en la figura 6.10, donde hemos trazado una gráfica de dispersión del promedio de las calificaciones de los estudiantes de primer año, y las calificaciones según el Comité para el Examen de Ingreso a la Universidad (CEEB, por sus siglas en inglés). Subdividimos la figura en calificaciones CEEB bajas, medias y altas. Al estimar la figura como un todo, es decir, al considerar todo el rango de las calificaciones del CEEB, hay una alta correlación entre las dos variables. Sin embargo, si consideramos las tres secciones por separado, la correlación de cada sección será mucho menor. Dentro de cada sección, los puntos muestran un cambio menos sistemático en Y con respecto a los cambios en X . Esto, por supuesto, indica una correlación menor entre X y Y . El efecto de la restricción del rango sobre la correlación se presenta con frecuencia en la educación o en la industria. Por ejemplo, suponga que sobre la base de la correlación alta entre las calificaciones de los estudiantes de primer año y las calificaciones del CEEB de la figura 6.10, una universidad decide admitir solamente a los graduados del bachillerato que hayan obtenido calificaciones dentro del rango alto de las calificaciones del CEEB. Si las calificaciones obtenidas posteriormente por estos estudiantes fueran correlacionadas con sus calificaciones del CEEB, esperaríamos ver una correlación mucho más baja, debido a la restricción de rango de las calificaciones del CEEB para estos estudiantes de primer grado. De manera similar, si alguien realiza un estudio de correlación y obtiene un coeficiente de correlación bajo, habrá que verificar si la restricción del rango no es la causa de este valor bajo.

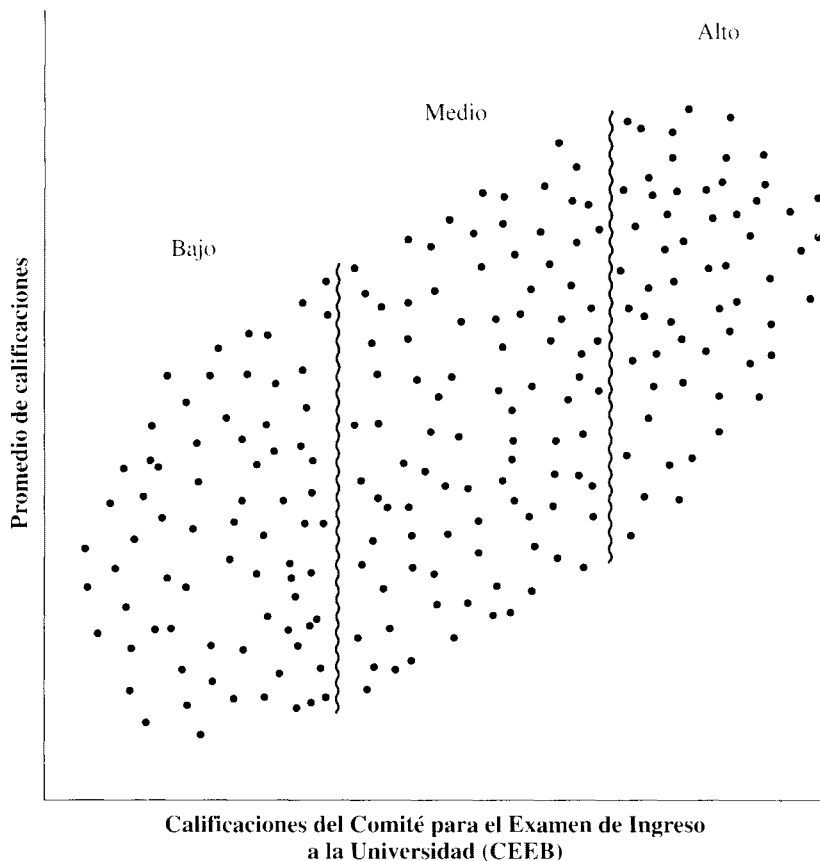


figura 6.10 Calificaciones de los estudiantes de primer grado y calificaciones CEEB.

Efecto de los datos extremos

Consideremos el efecto de un dato extremo sobre la magnitud del coeficiente de correlación. La figura 6.11(a) muestra un conjunto de datos en el que todos éstos se agrupan con un grado razonable de proximidad. El valor de la r de Pearson para este conjunto de datos es de 0.11. La figura 6.11(b) muestra el mismo conjunto de datos, pero se le ha agregado un dato extremo. El valor de la r de Pearson para este conjunto de datos es de 0.94. La magnitud de la r de Pearson ha cambiado de 0.11 a 0.94. Esto es una demostración del argumento según el cual *un dato extremo puede alterar drásticamente la magnitud del coeficiente de correlación* y, por lo tanto, modifica la interpretación de los datos. En consecuencia, antes de calcular el coeficiente de correlación, es conveniente examinar la gráfica de dispersión de los datos para comprobar si contiene datos extremos. Si tales datos extremos están presentes, se deberá tener cuidado al interpretar la relación. Si se trata de una muestra aleatoria grande, un valor extremo no alterará considerablemente, de ordinario, la magnitud de la correlación. En cambio, si la muestra es pequeña, como la de este ejemplo, un valor extremo puede producir un efecto grande.

La correlación no implica causalidad

Cuando existe una correlación entre dos variables (X y Y) es tentador concluir que una de ellas es la causa de la otra. Sin embargo, aceptar esto sin mayor investigación sería un grave error, pues siempre que dos variables estén correlacionadas puede haber cuatro explicaciones posibles de esa correlación: 1) la correlación entre X y Y degeneró en su origen, es espuria; 2) X es la causa de Y ; 3) Y es la causa de X , y 4) una tercera variable es la causa de la correlación entre X y Y . La primera posibilidad corresponde al caso en que la correlación exhibida por la muestra se debe a accidentes al realizar el muestreo de personas o comportamientos inusuales; es decir, que si el experimento se repitiese o si se tomaran más muestras, tal correlación desaparecería. Si la correlación es espuria en realidad, es obvio que sería incorrecto concluir que existe una relación causal entre X y Y .

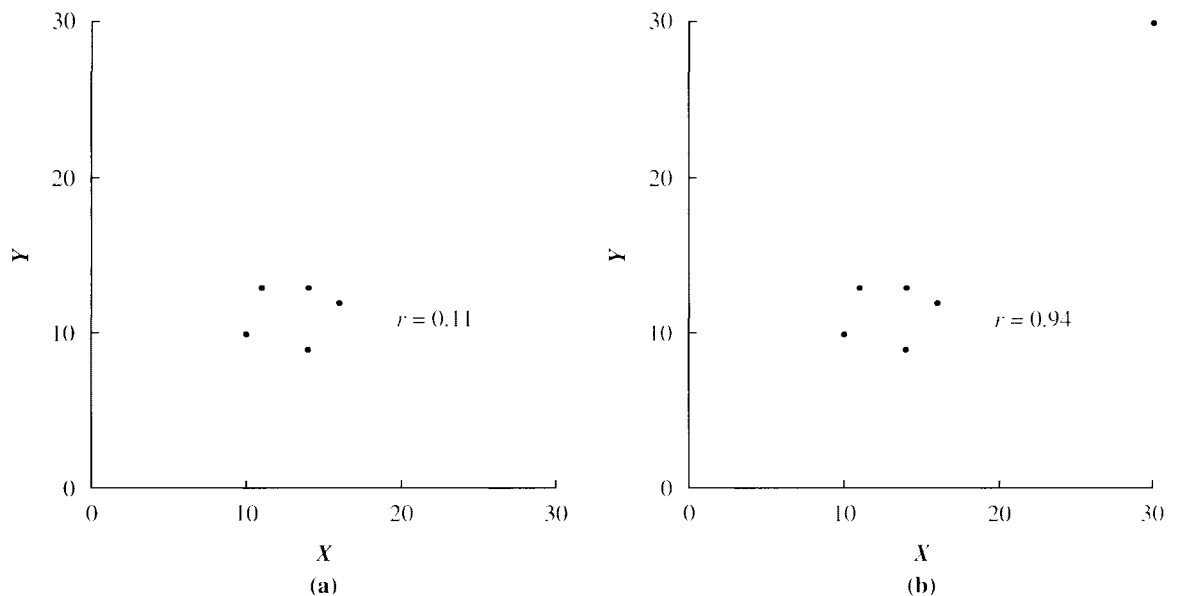


figura 6.11 Efecto de un dato extremo sobre la magnitud del coeficiente de correlación.

También es erróneo suponer una causalidad entre X y Y si la cuarta alternativa es correcta. Con mucha frecuencia, cuando X y Y están correlacionadas, no tienen una relación causal entre sí, sino que una tercera variable es responsable de esa correlación. Por ejemplo, ¿sabe usted que existe una relación estrecha entre los salarios de los profesores universitarios y el precio de una botella de whisky escocés? ¿Cuál es la causa y cuál el efecto de esto? ¿Acaso los salarios de los maestros universitarios dominan el mercado del whisky escocés, a tal grado que cuando los profesores reciben un aumento de sueldo y por ende pueden comprar más whisky, el precio de éste se incrementa en forma correspondiente? ¿O acaso a los profesores universitarios se les paga con las ganancias de las ventas de whisky escocés, de tal manera que cuando ellos necesitan un aumento, el precio de esta bebida se incrementa? En realidad, ninguna de estas explicaciones es correcta, pues la verdad es que existe un tercer factor responsable de esa correlación. ¿Cuál es ese factor? ¡La inflación! Hace poco, un artículo periodístico informó que había una correlación positiva entre la obesidad y la criminalidad femenina. ¿Significa esto que si una mujer aumenta 10 kilogramos se convertirá en criminal? O bien, ¿significa que si una mujer es criminal, está condenada a ser obesa? Ninguna de estas explicaciones parece ser satisfactoria. Francamente, no estamos seguros de cómo interpretar semejante correlación. Es posible que la correlación haya degenerado en su origen, es decir, sea espuria. También podría deberse a un tercer factor, como el nivel socioeconómico. Tanto la obesidad como la delincuencia están relacionadas con un nivel socioeconómico bajo.

El argumento es que la presencia de una correlación entre dos variables no basta para afirmar que existe una relación de causalidad entre ellas. Hay otras explicaciones posibles. Para establecer que una variable es la causa de otra, es necesario llevar a cabo un experimento en el cual se modifique de manera sistemática *sólo* la variable supuestamente causal y se mida el efecto que produce sobre la otra variable.



“Un buen director = una buena escuela”, ¿o no es así?

Uno de los principales periódicos de una gran ciudad publicó cierto día en sus titulares, con letra negrilla: “**Ecuación para el éxito: un buen director = una buena escuela elemental**”. En el artículo correspondiente se describía un estudio en el cual los directores de las escuelas de nivel básico fueron calificados por sus maestros, tomando como base una

serie de preguntas que indicaban si los directores tenían una capacidad de liderazgo alta, media o baja. Los estudiantes de esas escuelas fueron evaluados en lectura y matemáticas mediante los exámenes anuales de aprovechamiento de California. Según lo que podemos deducir del artículo del periódico, las clasificaciones y las calificaciones de los estudiantes fueron obtenidas a partir de las asignaciones habituales, sin que en dicho estudio se intentase asignar de manera aleatoria los directores a las escuelas. Los resultados mostraron que: 1) en 11 escuelas cuyos directores eran líderes fuertes, los estudiantes

estaban obteniendo magníficos resultados académicos; 2) en 11 escuelas donde los directores eran líderes débiles, los estudiantes mostraban avances inferiores al promedio o incluso se estaban retrasando, y 3) en 39 escuelas donde los directores fueron calificados como promedio, las calificaciones de los estudiantes sólo mostraron una mejoría de nivel promedio.

El reportero del periódico interpretó estos datos tal como lo anunciaban los titulares: “Ecuación para el éxito, un buen director = una buena escuela elemental”. En el artículo se citaban las palabras del



“Un buen director = una buena escuela”, ¿o no es así?

(continuación)

director de una escuela elemental: “Siempre lo he dicho, ‘muéstrenme una buena escuela y yo les mostraré un buen director’, pero ahora contamos con datos contundentes, indiscutibles, que corroboran esta afirmación”. En el artículo se citaba también al presidente de la asociación de directores, quien declaraba: “Esta información es emocionante y trae consigo una enorme responsabilidad. Muestra que nosotros podemos establecer una diferencia real en la vida de nuestros estudiantes”. ¿Cree usted que estas conclusiones están justificadas con los datos?

Respuesta Aun cuando creo personalmente que los directores de las escuelas son importantes para la calidad de la educación, el estudio parece ser estrictamente de

correlación: pares de mediciones de dos variables, sin asignación aleatoria alguna a los grupos. Por lo que dijimos anteriormente, es imposible determinar una causalidad a partir de ese tipo de estudios. Los individuos cuyas palabras fueron citadas consideraron un estudio que muestra la existencia de una correlación entre un liderazgo “fuerte” de los directores y los progresos de los alumnos en su educación, y concluyeron que los directores eran la causa de esos progresos. La conclusión resulta demasiado categórica. En realidad, tal

correlación podría haber degenerado en su origen, es decir, ser espuria o tener su causa en una tercera variable.

Es en realidad sorprendente la frecuencia con la que se presenta este error en la vida real. Manténgase alerta, y le sorprenderá ver la cantidad de veces que encontrará individuos dispuestos a concluir que existe causalidad, por el mero hecho de que los datos están correlacionados. Por supuesto, ahora que usted está tan bien informado sobre este punto, ¡jamás cometerá este error!



RESUMEN

En este capítulo hemos analizado el tema de la correlación. La correlación es una medida de la relación que existe entre dos variables. La magnitud y la dirección de la relación están expresadas por un coeficiente de correlación. El coeficiente de correlación puede variar de +1 a -1. El signo del coeficiente nos indica si la relación es positiva o negativa. La parte numérica especifica la magnitud de la correlación. Cuando la relación es perfecta, la magnitud es 1. Si no existe rela-

ción alguna, la magnitud es 0. Las magnitudes entre 0 y 1 corresponden a relaciones imperfectas.

Es posible calcular muchos coeficientes de correlación, según la escala de medición de los datos y la forma de la relación. En este capítulo destacamos la r de Pearson y la rho de Spearman. La r de Pearson se define como una medida del grado en que los pares de datos ocupan posiciones iguales u opuestas dentro de sus propias distribuciones. El uso de los puntajes estándar

permite medir la relación sin que importen las diferencias en la escala y las unidades utilizadas para medir las variables. La r de Pearson también es igual a la raíz cuadrada de la proporción de la variabilidad total en Y que es explicada por X . Además de estos conceptos, presentamos una ecuación para calcular r y practicamos dicho cálculo.

La rho de Spearman se usa para relaciones lineales cuando una o ambas variables se expresan solamente en la escala ordinal. Presentamos la ecuación para calcular rho y resolvimos varios problemas de práctica. A continuación, comentamos el efecto de un dato extremo sobre la magnitud de la correlación. Después, analizamos el efecto del rango sobre la correlación y señalamos que un rango truncado produce un coeficiente de correlación menor.

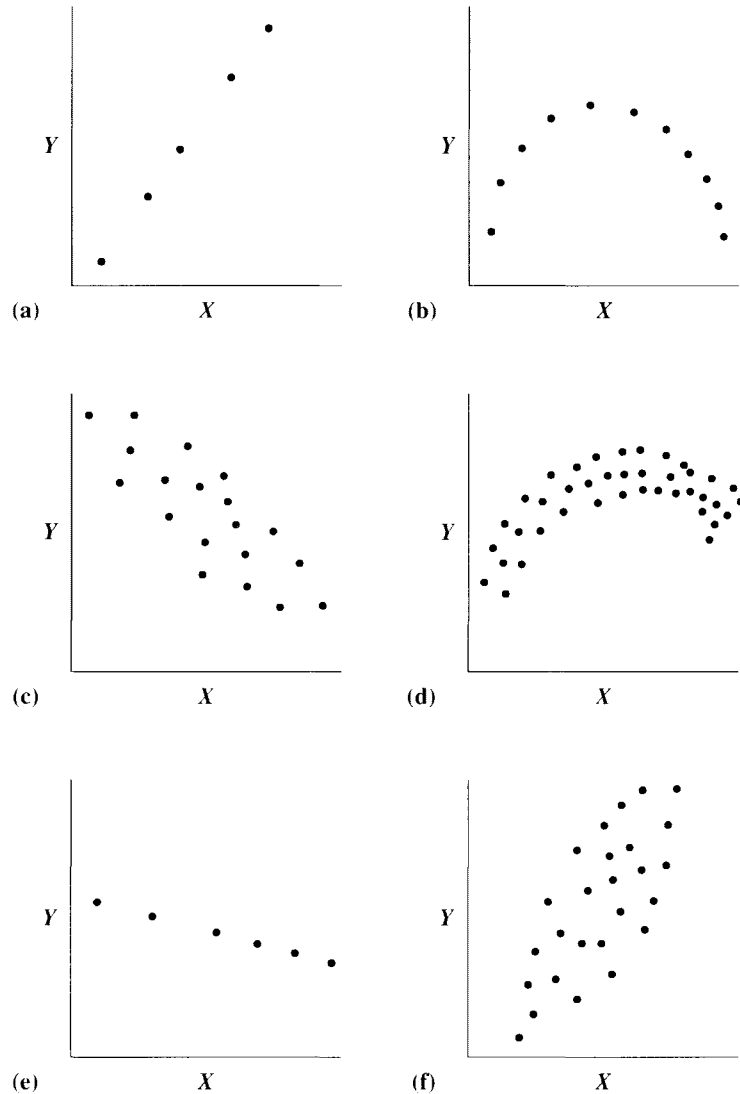
Como último punto en este tema, analizamos la correlación y la causalidad. Mostramos que aunque exista una correlación entre dos variables de un experimento, por ese simple hecho no se puede concluir que estén relacionadas de manera causal, ya que existen otras explicaciones posibles. La correlación puede ser espuria, o bien, una tercera variable puede ser causante de la correlación existente entre las dos primeras variables. Para establecer la causalidad, es necesario manipular una de las variables de manera independiente y luego medir el efecto resultante sobre la otra variable. Las demás variables deben mantenerse constantes o variar de manera no sistemática. Aunque las dos variables estén relacionadas de manera causal, es importante tener presente que r^2 , y no r , indica la magnitud del efecto de una variable sobre la otra.

TÉRMINOS IMPORTANTES

Coefficiente biserial (página 121)	Ordenada al origen (página 106)	Relación lineal (página 104)
Coefficiente de determinación (página 120)	Pendiente (página 106)	Relación negativa (página 107)
Coefficiente de correlación (página 111)	r de Pearson (página 112)	Relación perfecta (página 108)
Coefficiente phi (página 121)	Relación curvilínea (página 105)	Relación positiva (página 107)
Correlación (página 111)	Relación directa (página 107)	Rho de Spearman (página 121)
Gráfica de dispersión (página 104)	Relación imperfecta (página 108)	Variabilidad explicada por X (página 120)
	Relación inversa (página 107)	

PREGUNTAS Y PROBLEMAS

- Defina o identifique cada uno de los conceptos de la sección “Términos importantes”.
- Analice los distintos tipos de relación posibles entre dos variables.
- Para cada gráfica de dispersión de la figura anexa (partes **a–f** de la página 129) determine si la relación es:
 - Lineal o curvilínea. Si es lineal, determine si es positiva o negativa.
 - Perfecta o imperfecta.
- La profesora Taylor realiza un experimento y establece que existe una correlación entre las variables A y B . Tomando como base esta correlación, ella afirma que A es la causa de B . ¿Es correcta esta afirmación? Explique.
- Mencione dos significados de la r de Pearson.
- ¿Por qué se utilizan los puntajes z como base para determinar la r de Pearson?
- ¿Cuál es el rango de valores que puede asumir un coeficiente de correlación?
- Un estudio ha mostrado que la correlación entre la fatiga y la irritabilidad es de 0.53. Tomando como base esa correlación, el autor ha sacado la conclusión de que la fatiga es un factor importante para que se produzca la irritabilidad. ¿Considera justificada esta conclusión? Explique.
- ¿Qué factores influyen sobre la elección de un coeficiente de correlación particular? Proponga algunos ejemplos.
- Los coeficientes de correlación r de Pearson y rho de Spearman están relacionados entre sí. ¿Es ésta una afirmación correcta? Explique.
- Cuando dos variables están correlacionadas, existen cuatro explicaciones posibles para la correlación. ¿Cuáles son?
- ¿Qué efecto puede tener un dato extremo sobre la magnitud de la relación entre dos variables? Comente.
- ¿Qué efecto tiene la reducción del rango de las parejas de datos sobre el coeficiente de correlación?



Gráficas de dispersión para la pregunta 3

14. Dados los siguientes conjuntos de parejas de datos muestrales:

A		B		C	
X	Y	X	Y	X	Y
1	1	4	2	1	5
4	2	5	4	4	4
7	3	8	5	7	3
10	4	9	1	10	2
13	5	10	4	13	1

- a. Utilice la siguiente ecuación

$$r = \frac{\sum z_X z_Y}{(N - 1)}$$

para calcular el valor de la r de Pearson que corresponde a cada conjunto. Observe que en el conjunto B, en el cual la correlación es menor, algunos de los valores $z_X z_Y$ son positivos y otros son negativos. Estos valores tienden a cancelarse entre sí, lo cual hace que r tenga una menor magnitud. En cambio, en los conjuntos A y C todos los productos tienen el mismo signo, lo cual hace que la magnitud de r aumente. Cuando las parejas de datos ocupan las mismas posi-

ciones o posiciones opuestas dentro de sus propias distribuciones, los productos $z_X z_Y$ tienen el mismo signo, a causa de lo cual las magnitudes correspondientes a r son mayores.

- Calcule r para el conjunto B, utilizando la ecuación para los datos en bruto. ¿Qué prefiere, utilizar la ecuación de los datos en bruto o la de los puntajes z ?
 - Suma la constante 5 a los datos X en el conjunto A y calcule r de nuevo mediante la ecuación de los datos en bruto. ¿Se ha modificado el valor?
 - Multiplique los datos X del conjunto A por 5 y calcule r de nuevo. ¿Ha cambiado el valor?
 - Generalice los resultados obtenidos en las partes **c** y **d**, para restar y hacer la división de los datos entre una constante. ¿Qué le indica esto acerca de r ?
15. En un largo curso de introducción a la sociología, un profesor administra dos exámenes. El profesor quiere determinar si las calificaciones obtenidas por los estudiantes en el segundo examen están correlacionadas con las calificaciones del primero. Para facilitar los cálculos se elige una muestra de ocho estudiantes. Sus calificaciones se presentan en la siguiente tabla.

Estudiante	Examen 1	Examen 2
1	60	60
2	75	100
3	70	80
4	72	68
5	54	73
6	83	97
7	80	85
8	65	90

- Haga una gráfica de dispersión para estos datos, utilizando la calificación del primer examen como variable X . ¿Le parece a usted que la relación sea lineal?
 - Suponga que existe una relación lineal entre las calificaciones de los dos exámenes y calcule el valor de la r de Pearson.
 - ¿Qué tan bien permite explicar esta relación las calificaciones del segundo examen?
16. Una estudiante del posgrado en psicología del desarrollo tiene la impresión de que puede existir una relación entre el peso de los individuos al nacer y su futuro CI. Entonces muestrea al azar

a siete estudiantes de su universidad con especialidad en psicología y los somete a una prueba de CI. A continuación, investiga en los hospitales correspondientes cuál fue el peso de los siete estudiantes al nacer (por supuesto, con autorización de éstos). Los datos así obtenidos se presentan en la siguiente tabla.

Estudiante	Peso al nacer (libras)	CI
1	5.8	122
2	6.5	120
3	8.0	129
4	5.9	112
5	8.5	127
6	7.2	116
7	9.0	130

- Elabore una gráfica de dispersión de los datos, representando en el eje X el peso al nacer y el CI en el eje Y . ¿Parece ser lineal la relación?
 - Suponga que la relación es lineal y calcule el valor de la r de Pearson.
17. Un investigador realiza un estudio para averiguar la relación entre el consumo de cigarrillos y las enfermedades. El número de cigarrillos fumados diariamente y el número de días de ausencia al trabajo a causa de enfermedad durante el último año son calculados en el caso de 12 empleados de la compañía donde trabaja el investigador. Los datos se presentan en la tabla anexa.

Sujeto	Cigarrillos consumidos	Días de ausencia
1	0	1
2	0	3
3	0	8
4	10	10
5	13	4
6	20	14
7	27	5
8	35	6
9	35	12
10	44	16
11	53	10
12	60	16

- a. Haga una gráfica de dispersión para estos datos. ¿Le parece que la relación es lineal?
 - b. Determine el valor de la r de Pearson.
 - c. Elimine los datos de los sujetos 1, 2, 3, 10, 11 y 12. Esto reduce el rango de ambas variables. Vuelva a calcular r para los sujetos restantes. ¿Qué efecto tiene la disminución del rango sobre r ?
 - d. Al utilizar todo el conjunto de datos, ¿qué porcentaje de la variabilidad en el número de días de ausencia es explicado por la cantidad de cigarrillos fumados diariamente? ¿Para qué sirve ese valor? clínica, salud
18. Un educador ha elaborado una prueba para determinar las aptitudes mecánicas. Desea saber en qué medida es confiable dicha prueba, para lo cual la administra dos veces, con un lapso de un mes entre ellas. Lleva a cabo un estudio en el cual a 10 estudiantes se les administra en dos ocasiones la prueba, realizándose la segunda administración un mes después de la primera. Los datos correspondientes se presentan en la siguiente tabla.

Estudiante	Administración 1	Administración 2
1	10	10
2	12	15
3	20	17
4	25	25
5	27	32
6	35	37
7	43	40
8	40	38
9	32	30
10	47	49

- a. Elabore una gráfica de dispersión para las parejas de datos.
 - b. Determine el valor de r .
 - c. ¿Sería justo decir que ésta es una prueba confiable? Explíquelo utilizando r^2 . educación
19. Un grupo de investigadores ha diseñado un cuestionario sobre el tema del estrés, en el cual se mencionan 15 eventos de la vida. Están interesados en determinar si existe coincidencia entre dos culturas acerca de la cantidad relativa de ajustes que cada uno de esos eventos requiere. El cuestionario se aplica a 300 estadounidenses y 300 italianos. Cada individuo se debe basar en el evento "matrimonio" como estándar, calificando los demás eventos en relación con el grado de ajustes

necesario para adaptarse al matrimonio. Al matrimonio se le asigna un valor arbitrario de 50 puntos. Si se considera que un evento requiere más ajustes que el matrimonio, dicho evento deberá recibir más de 50 puntos. El número de puntos excedentes depende de la cantidad de ajustes que cada evento requiera. Una vez que todos los sujetos de ambas culturas han asignado puntos a todos los eventos, se calcula el promedio de los puntos correspondientes a cada evento. Los resultados obtenidos se presentan en la siguiente tabla.

Evento	Estadounidenses	Italianos
Muerte del cónyuge	100	80
Divorcio	73	95
Separación de la pareja	65	85
Temporada en prisión	63	52
Lesiones personales	53	72
Matrimonio	50	50
Despido del trabajo	47	40
Jubilación	45	30
Embarazo	40	28
Dificultades sexuales	39	42
Reajustes en la empresa	39	36
Problemas con la familia política	29	41
Problemas con el jefe	23	35
Vacaciones	13	16
Navidad	12	10

- a. Suponga que los datos tienen por lo menos una escala de intervalo y calcule la correlación entre los datos de los estadounidenses y de los italianos.
 - b. Suponga que los datos sólo tienen una escala ordinal y calcule la correlación entre los datos de ambas culturas. clínica, salud
20. Considerando el siguiente conjunto de pares de datos que fueron tomados de cinco sujetos:

Sujeto núm.	1	2	3	4	5
Y	5	6	9	9	11
X	6	8	4	8	7

- a. Haga una gráfica de dispersión de los datos.
- b. Calcule el valor de la r de Pearson.

- c. Agregue a los datos los siguientes pares de datos obtenidos de un sexto sujeto: $Y = 26$, $X = 25$.
- d. Elabore otro diagrama de dispersión, esta vez para los seis pares de datos.
- e. Calcule el valor de la r de Pearson para los seis pares de datos.
- f. ¿Hay mucha diferencia entre sus respuestas para las partes **b** y **e**? Explique esa diferencia.
21. La directora de una clínica para el control de peso, establecida en una ciudad grande del noroeste, cree que la ingestión de refrescos contribuye a la obesidad en los niños. Para determinar si existe una relación entre estas dos variables, ella realiza el siguiente estudio piloto. Ocho voluntarios de 12 años de edad han sido seleccionados aleatoriamente entre los niños que asisten a una escuela secundaria local. A los padres de los niños se les pide que lleven la cuenta del número de refrescos que hayan consumido sus hijos durante un periodo de una semana. Todos los niños son pesados al final de la semana y sus respectivos pesos son convertidos a valores correspondientes al índice de masa corporal (BMI por sus siglas en inglés). El BMI suele usarse para medir la obesidad; en él se toman en cuenta tanto la estatura como el peso de la persona. Se considera que un individuo es obeso si tiene un valor $BMI \geq 30$. Los datos que han sido recopilados con este procedimiento son los siguientes.

Niño	Número de refrescos que consumió	BMI
1	3	20
2	1	18
3	14	32
4	7	24
5	21	35
6	5	19
7	25	38
8	9	30

- a. Haga una gráfica de dispersión de los datos. ¿Le parece que esta relación sea lineal?
- b. Suponga que la relación es lineal y calcule la r de Pearson. salud
22. Un psicólogo social realiza un estudio para determinar la relación entre la religión y la autoestima. Diez alumnos de octavo grado han sido seleccionados al azar para el estudio. Cada uno de ellos es sometido a dos pruebas, una para medir su autoestima y la otra para medir su devoción religiosa. En la prueba de autoestima, cuanto más bajo sea el puntaje, tanto más alta será la autoestima; para la prueba que permite medir la devoción religiosa, cuanto más alto sea el puntaje, tanto más alta será la devoción religiosa. La prueba de autoestima tiene un rango de 1 a 10 y el rango de la prueba de la devoción religiosa es de 0 a 50. Para los propósitos de esta pregunta, supongamos que las dos pruebas están bien estandarizadas y tienen una escala de intervalo. Los datos recolectados son los siguientes:

nados al azar para el estudio. Cada uno de ellos es sometido a dos pruebas, una para medir su autoestima y la otra para medir su devoción religiosa. En la prueba de autoestima, cuanto más bajo sea el puntaje, tanto más alta será la autoestima; para la prueba que permite medir la devoción religiosa, cuanto más alto sea el puntaje, tanto más alta será la devoción religiosa. La prueba de autoestima tiene un rango de 1 a 10 y el rango de la prueba de la devoción religiosa es de 0 a 50. Para los propósitos de esta pregunta, supongamos que las dos pruebas están bien estandarizadas y tienen una escala de intervalo. Los datos recolectados son los siguientes:

Sujeto	Devoción religiosa	Autoestima
1	5	8
2	25	3
3	45	2
4	20	7
5	30	5
6	40	5
7	1	4
8	15	4
9	10	7
10	35	3

- a. Si efectivamente existe una relación por la cual cuanto más alta sea la devoción religiosa de una persona tanto más alta será su autoestima, ¿espera usted que la r calculada a partir de los valores proporcionados sea negativa o positiva? Explique.
- b. Calcule r . ¿Estaba usted en lo cierto cuando respondió a la parte **a**? social, de desarrollo
23. Un psicólogo elaboró una prueba “lápiz y papel” que tiene la finalidad de medir el grado de depresión de las personas. Para comparar los resultados de esa prueba con las calificaciones asignadas por expertos, 12 individuos “con perturbaciones emocionales” fueron sometidos a ella. Los mismos individuos fueron calificados también por dos psiquiatras, de manera independiente, según el grado de depresión observado por cada uno de ellos en el curso de entrevistas detalladas con esos individuos. Los datos así obtenidos se presentan a continuación. Los puntajes más altos corresponden a un grado mayor de depresión.

Individuo	Prueba "lápiz y papel"	Psiquiatra A	Psiquiatra B
1	48	12	9
2	37	11	12
3	30	4	5
4	45	7	8
5	31	10	11
6	24	8	7
7	28	3	4
8	18	1	1
9	35	9	6
10	15	2	2
11	42	6	10
12	22	5	3

- a. ¿Cuál es la correlación entre las calificaciones de los dos psiquiatras?
- b. ¿Cuál es la correlación entre las calificaciones de la prueba "lápiz y papel" y las calificaciones de cada psiquiatra? clínica, salud
24. Para resolver este problema, supongamos que usted es un psicólogo que trabaja en el departamento de recursos humanos de una gran corporación. El presidente de la compañía acaba de hablar con usted para manifestarle la importancia de contratar personal productivo para la sección de manufactura de la empresa, y le ha pedido que ayude a mejorar la capacidad de la institución para lograr ese propósito. En la actualidad hay 300 empleados en esa sección y todos ellos fabrican el mismo artículo. Hasta la fecha, la corporación sólo se ha ser-

vido del método de entrevistas para la selección de ese tipo de empleados. Usted busca en la bibliografía y descubre que existen dos pruebas "lápiz y papel" de desempeño bien estandarizadas, y estima que podrían ser apropiadas para determinar los requisitos de desempeño de la sección. Para averiguar si alguna de ellas puede ser utilizada como dispositivo de selección, usted elige a 10 empleados representativos de la sección de manufactura, cerciorándose de que en esa muestra quede representado un amplio rango de desempeño, y somete a cada uno de esos trabajadores a las dos pruebas anteriormente mencionadas. Los datos así obtenidos se presentan en la siguiente tabla. Mientras más alta sea la calificación, tanto mejor será el desempeño. Las calificaciones de desempeño en el trabajo son la cantidad real de artículos fabricados por cada empleado por semana, promediada durante los últimos 6 meses.

- a. Haga una gráfica de dispersión del desempeño en el trabajo y de la prueba 1, utilizando dicha prueba como variable X . ¿Cree usted que la relación sea lineal?
- b. Suponiendo que la relación anterior es lineal, calcule el valor de la r de Pearson.
- c. Elabore una gráfica de dispersión del desempeño en el trabajo y la prueba 2, utilizando dicha prueba 2 como variable X . ¿Cree usted que la relación sea lineal?
- d. Suponiendo que la relación anterior sea lineal, calcule el valor de la r de Pearson.
- e. Si solamente pudiera usted utilizar una de las pruebas para la selección de los empleados, ¿usaría alguna de ellas? En caso afirmativo, ¿cuál de ellas utilizaría? Explique. 1/0

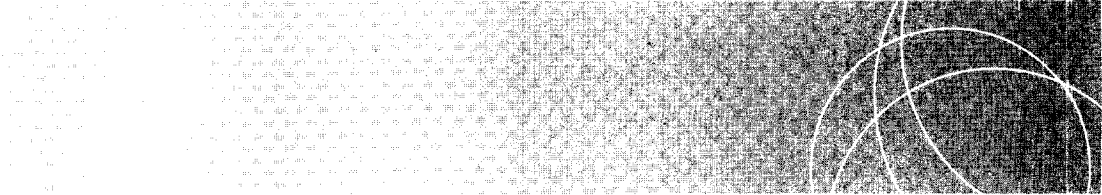
	Empleado									
	1	2	3	4	5	6	7	8	9	10
Desempeño en el trabajo	50	74	62	90	98	52	68	80	88	76
Prueba 1	10	19	20	20	21	14	10	24	16	14
Prueba 2	25	35	40	49	50	29	32	44	46	35

Si desea encontrar problemas de práctica adicionales totalmente resueltos sobre los temas que acaba de estudiar, busque Chapter 6 en el material adicional del libro, disponible en inglés en: <http://www.thomsonlearning.com.mx>.

Si pretende resolver más problemas de práctica, relacionados con lo que se presentó en este capítulo, resueltos por medio de ecuaciones conceptuales, busque Chapter 6 en el material adicional del libro (en inglés), en: <http://www.thomsonlearning.com.mx>.

Si quiere resolver problemas para computadora sobre los temas aquí presentados, busque Chapter 6 en el material adicional (en inglés) del libro, en: <http://www.thomsonlearning.com.mx>.

Regresión lineal



Introducción

Predicción y relaciones imperfectas

Construcción de la recta de regresión por mínimos cuadrados: Regresión de Y sobre X

Regresión de X sobre Y

Medición de errores de predicción: el error estándar de la estimación

Consideraciones al utilizar la regresión lineal para la predicción

Relación entre las constantes de regresión y la r de Pearson

Regresión múltiple y correlación múltiple

Resumen

Términos importantes

Preguntas y problemas



INTRODUCCIÓN

La regresión y la correlación están íntimamente relacionadas entre sí. En el nivel más elemental, ambas implican la relación entre dos variables y utilizan el mismo conjunto de datos básicos: las parejas de datos obtenidas de los mismos sujetos o de otros que forman parejas con ellos. Como estudiamos en el capítulo 6, la correlación tiene que ver con la magnitud y la dirección de la relación. La regresión se centra en el uso de la relación para elaborar una predicción. La predicción resulta bastante fácil de establecer cuando la relación es perfecta. Cuando esto sucede, todos los puntos están localizados sobre una línea recta y lo único que tenemos que hacer es obtener la ecuación de esa recta y utilizarla para llevar a cabo la predicción. Como puede suponer, cuando la relación es perfecta, también lo es la predicción. Todos los valores predichos son exactamente iguales a los valores observados y el error de predicción es igual a cero. La situación es más compleja cuando la relación es imperfecta.

definición ■ La **regresión** es un tema en el que se aplica la relación entre dos o más variables con propósitos de predicción.



PREDICCIÓN Y RELACIONES IMPERFECTAS

Consideremos de nuevo los datos referentes al promedio de las calificaciones y el CI, presentados en el capítulo 6. Por comodidad, hemos reproducido esos datos en la tabla 7.1. La figura 7.1 muestra una gráfica de dispersión o dispersigrama correspondiente a dichos datos. La relación es imperfecta, positiva y lineal. El problema de predicción que nos ocupa consiste en hallar una sola línea recta que describa esos datos mejor que cualquier otra. La solución que se usa más frecuentemente consiste en construir la recta que minimice los errores de predicción de acuerdo con el criterio de los *mínimos cuadrados*. Esa línea tiene el nombre, muy apropiado, de *recta de regresión por mínimos cuadrados*.

La recta de regresión por mínimos cuadrados para los datos de la tabla 7.1 aparece en la figura 7.2(a). La distancia vertical entre cada punto y la recta representa el error de la predicción. Si Y' es la predicción del valor Y y Y es el valor real, entonces $Y - Y'$ es el error para cada punto. Parecería que el error total de la predicción sería la suma algebraica simple de $Y - Y'$, sumado sobre todos los puntos. Si esto fuese cierto, y como estamos interesados en minimizar el error, construiríamos la recta que minimiza $\sum (Y - Y')$. Sin embargo, el error total de la predicción no es igual a $\sum (Y - Y')$ porque algunos de los valores Y' serán mayores que Y y otros serán menores. Así, habrá errores positivos y negativos, y las sumas algebraicas simples de estos errores se cancelarán entre sí. Ya nos encontramos con una situación similar cuando medimos el promedio de las dispersiones. Al deducir la ecuación de la desviación estándar, elevamos al cuadrado $X - \bar{X}$ para compensar el hecho de que había desviaciones positivas y negativas que se cancelaban entre sí. Aquí funciona también esta solución. En vez de sólo sumar $Y - Y'$, primero calcularemos $(Y - Y')^2$ para cada calificación. Esto elimina los valores negativos y el problema de cancelación queda resuelto. Ahora, si minimizamos $\sum (Y - Y')^2$, minimizamos el error total de la predicción.

tabla 7.1 CI y promedio de calificaciones de 12 estudiantes

Estudiante núm.	CI	Promedio de calificaciones
1	110	1.0
2	112	1.6
3	118	1.2
4	119	2.1
5	122	2.6
6	125	1.8
7	127	2.6
8	130	2.0
9	132	3.2
10	134	2.6
11	136	3.0
12	138	3.6

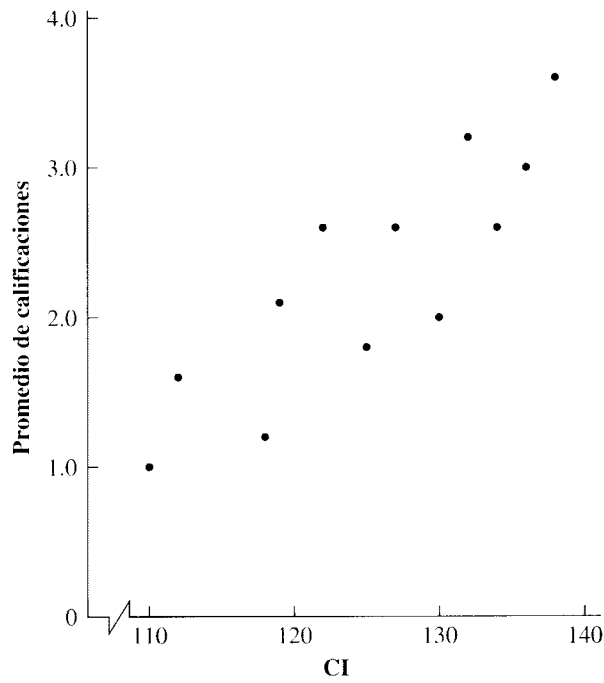


figura 7.1 Gráfica de dispersión del CI y el promedio de calificaciones.

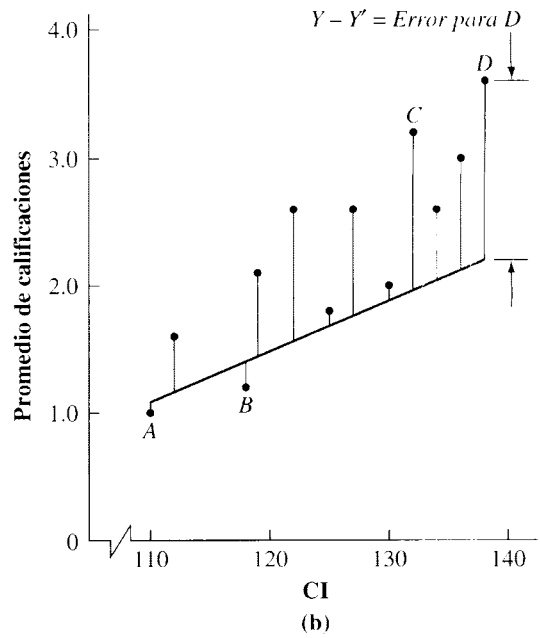
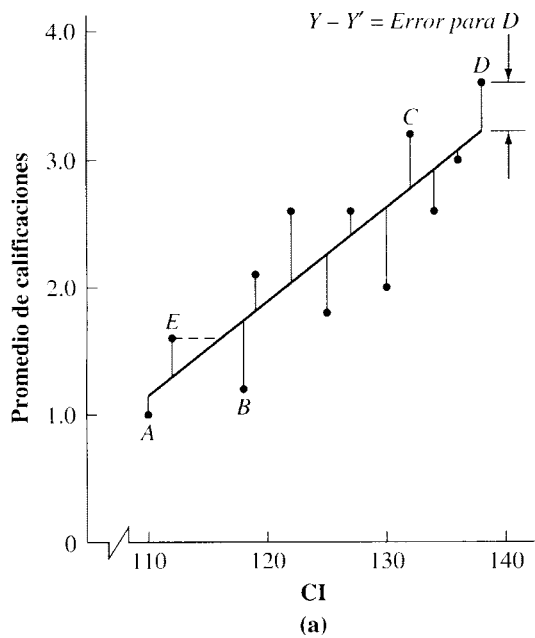


figura 7.2 Dos rectas de regresión y el error de predicción.

definición ■ *La línea de regresión por mínimos cuadrados es la línea de predicción que minimiza $\Sigma (Y - Y')^2$.*

Para cualquier relación lineal existe una única recta que minimiza $\Sigma (Y - Y')^2$. Por consiguiente, no existe más que una recta de regresión por mínimos cuadrados para cada relación lineal.

Habíamos dicho que es factible construir muchas líneas de predicción “posibles” cuando la relación es imperfecta. ¿Por qué deberíamos utilizar la línea de regresión por mínimos cuadrados? La empleamos porque es útil, ya que nos facilita la predicción con máxima precisión global. Para ilustrar este punto, en la figura 7.2(b) trazamos otra línea de predicción. Elegimos esta recta de manera arbitraria y es sólo una de la infinidad de líneas que podríamos haber trazado. ¿Cómo se compara la precisión de esta predicción con la de la recta de regresión por mínimos cuadrados? Podemos apreciar que en realidad se ajusta mejor a algunos puntos (por ejemplo, en los puntos A y B). Sin embargo, falla terriblemente en otros (por ejemplo, en los puntos C y D). Si consideramos todos los puntos, resulta claro que la recta de la figura 7.2(a) se ajusta mejor a los puntos que la recta de la figura 7.2(b). El error total de la predicción, representado por $\Sigma (Y - Y')^2$, es menor para la línea de regresión por mínimos cuadrados que para la recta de la figura 7.2(b). De hecho, el error total de predicción es menor con la línea de regresión por mínimos cuadrados que con cualquier otra recta de predicción posible. Así, utilizamos la recta de regresión por mínimos cuadrados porque proporciona predicciones con mayor precisión general que cualquier otra línea de regresión posible.



CONSTRUCCIÓN DE LA RECTA DE REGRESIÓN POR MÍNIMOS CUADRADOS: REGRESIÓN DE Y SOBRE X

La ecuación de la recta de regresión por mínimos cuadrados para predecir Y conociendo X es:

$$Y' = b_Y X + a_Y \quad \text{ecuación de regresión lineal para predecir } Y \text{ conociendo } X$$

donde Y' = el valor predicho o estimado de Y
 b_Y = la pendiente de la recta que minimiza los errores en la predicción de Y
 a_Y = Y la ordenada al origen de la recta que minimiza los errores en la predicción de Y

Esta es, por supuesto, la ecuación general de la recta que hemos utilizado todo el tiempo. Sin embargo, en este contexto, a_Y y b_Y se llaman *constantes de regresión*. Esta línea se llama recta de regresión de Y sobre X , o simplemente la regresión de Y sobre X , en virtud de que intentamos predecir Y a partir de una X conocida.

La constante de regresión b_Y es igual a:

$$b_Y = \frac{\Sigma XY - \frac{(\Sigma X)(\Sigma Y)}{N}}{SC_X}$$

donde $SC_X =$ la suma de los cuadrados de los datos $X = \Sigma X^2 - \frac{(\Sigma X)^2}{N}$
 $N =$ el número de parejas de datos

ΣXY = suma del producto de cada pareja X y Y (también conocida como la suma de los productos cruzados)

La ecuación para calcular b_Y a partir de los datos en bruto es:

$$b_Y = \frac{\Sigma XY - \frac{(\Sigma X)(\Sigma Y)}{N}}{\Sigma X^2 - \frac{(\Sigma X)^2}{N}} \quad \text{ecuación que permite calcular la constante de regresión } b \text{ para predecir } Y \text{ conociendo } X$$

La constante de regresión a_Y está dada por:

$$a_Y = \bar{Y} - b_Y \bar{X} \quad \text{ecuación para calcular la constante de regresión } a \text{ para predecir } Y \text{ conociendo } X$$

Como necesitamos la constante b_Y para determinar la constante a_Y el procedimiento consiste en hallar primero b_Y y después a_Y . Una vez determinadas ambas, éstas se sustituyen en la ecuación de regresión. Construyamos la línea de regresión por mínimos cuadrados para los datos del CI y las calificaciones que presentamos anteriormente. Por comodidad, los datos han sido presentados de nuevo en la tabla 7.2.

tabla 7.2 CI y promedio de las calificaciones de 12 estudiantes universitarios: predicción de Y a partir de X

Estudiante núm.	CI X	Promedio de las calificaciones		XY	X^2
		Y			
1	110	1.0		110.0	12 100
2	112	1.6		179.2	12 544
3	118	1.2		141.6	13 924
4	119	2.1		249.9	14 161
5	122	2.6		317.2	14 884
6	125	1.8		225.0	15 625
7	127	2.6		330.2	16 129
8	130	2.0		260.0	16 900
9	132	3.2		422.4	17 424
10	134	2.6		348.4	17 956
11	136	3.0		408.0	18 496
12	138	3.6		496.8	19 044
Total	1503	27.3		3488.7	189 187

$$b_Y = \frac{\Sigma XY - \frac{(\Sigma X)(\Sigma Y)}{N}}{\Sigma X^2 - \frac{(\Sigma X)^2}{N}} = \frac{3488.7 - \frac{1503(27.3)}{12}}{189\,187 - \frac{(1503)^2}{12}}$$

$$= \frac{69.375}{936.25} = 0.0741 = 0.074$$

$$a_Y = \bar{Y} - b_Y \bar{X} = 2.275 - 0.0741(125.25) = -7.006$$

$$Y' = b_Y X + a_Y = 0.074X - 7.006$$

$$\begin{aligned}
 b_Y &= \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sum X^2 - \frac{(\sum X)^2}{N}} \\
 &= \frac{3488.7 - \frac{1503(27.3)}{12}}{189.187 - \frac{(1503)^2}{12}} \\
 &= \frac{69.375}{936.25} = 0.0741 = 0.074 \\
 a_Y &= \bar{Y} - b_Y \bar{X} \\
 &= 2.275 - 0.0741(125.25) \\
 &= -7.006
 \end{aligned}$$

y también

$$Y' = 0.074X - 7.006$$

La solución completa aparece también en la tabla 7.2. La gráfica de la línea de regresión se presenta en la figura 7.3. Ahora podemos usar la ecuación para Y' para predecir el promedio de calificaciones, conociendo solamente el CI del estudiante. Por ejemplo, supongamos que el CI del estudiante es de 124; al emplear esta línea de regresión, ¿cuál es la predicción del promedio de las calificaciones del estudiante?

$$\begin{aligned}
 Y' &= 0.074X - 7.006 \\
 &= 0.074(124) - 7.006 \\
 &= 2.17
 \end{aligned}$$

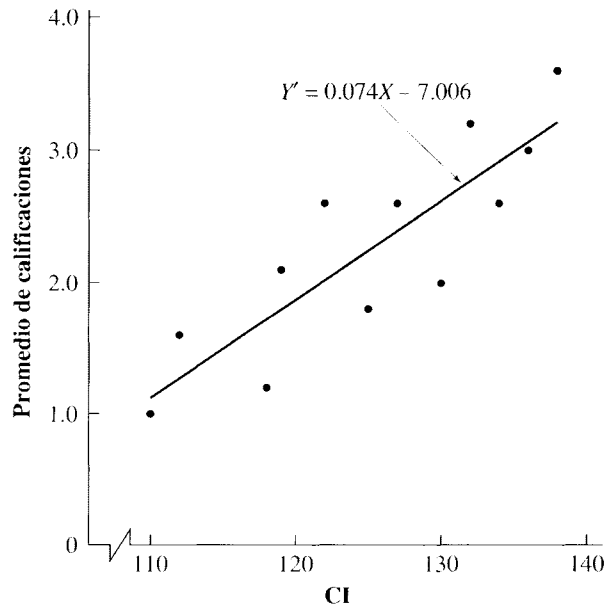


figura 7.3 Recta de regresión para el promedio de calificaciones y el CI.

Resolvamos un par de problemas de práctica.

Problema de práctica 7.1

Una psicóloga del desarrollo está interesada en determinar si es posible utilizar la estatura de los niños como base para predecir su posible estatura en la edad adulta. Para resolver esta cuestión, ha reunido los datos que aparecen en la tabla siguiente.

- Trace una gráfica de dispersión para los datos.
- Si los datos están relacionados de manera lineal, deduzca la recta de regresión por mínimos cuadrados.
- Con base en estos datos, ¿qué estatura podría predecir para una persona de 20 años si a los tres años de edad tenía una altura de 42 pulgadas?

Individuo número	Estatura a la edad de 3 años X (pulgadas)	Estatura a la edad de 20 años Y (pulgadas)	XY	X^2
1	30	59	1770	900
2	30	63	1890	900
3	32	62	1984	1024
4	33	67	2211	1089
5	34	65	2210	1156
6	35	61	2135	1225
7	36	69	2484	1296
8	38	66	2508	1444
9	40	68	2720	1600
10	41	65	2665	1681
11	41	73	2993	1681
12	43	68	2924	1849
13	45	71	3195	2025
14	45	74	3330	2025
15	47	71	3337	2209
16	48	75	3600	2304
<i>Total</i>	618	1077	41 956	24 408

$$b_Y = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sum X^2 - \frac{(\sum X)^2}{N}} = \frac{41\,956 - \frac{618(1077)}{16}}{24\,408 - \frac{(618)^2}{16}} = 0.6636 = 0.664$$

$$a_Y = \bar{Y} - b_Y \bar{X} = 67.3125 - 0.6636(38.625) = 41.681$$

$$Y' = b_Y X + a_Y = 0.664X + 41.681$$

SOLUCIÓN

- a. La gráfica de dispersión se presenta en la figura siguiente. Está claro que existe una relación imperfecta, lineal y positiva entre las estaturas a las edades de 3 y 20 años.
- b. Obtención de la recta de regresión por mínimos cuadrados.

$$Y' = b_Y X + a_Y$$

$$b_Y = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sum X^2 - \frac{(\sum X)^2}{N}} = \frac{41\,956 - \frac{618(1077)}{16}}{24\,408 - \frac{(618)^2}{16}}$$

$$= 0.6636 = 0.664$$

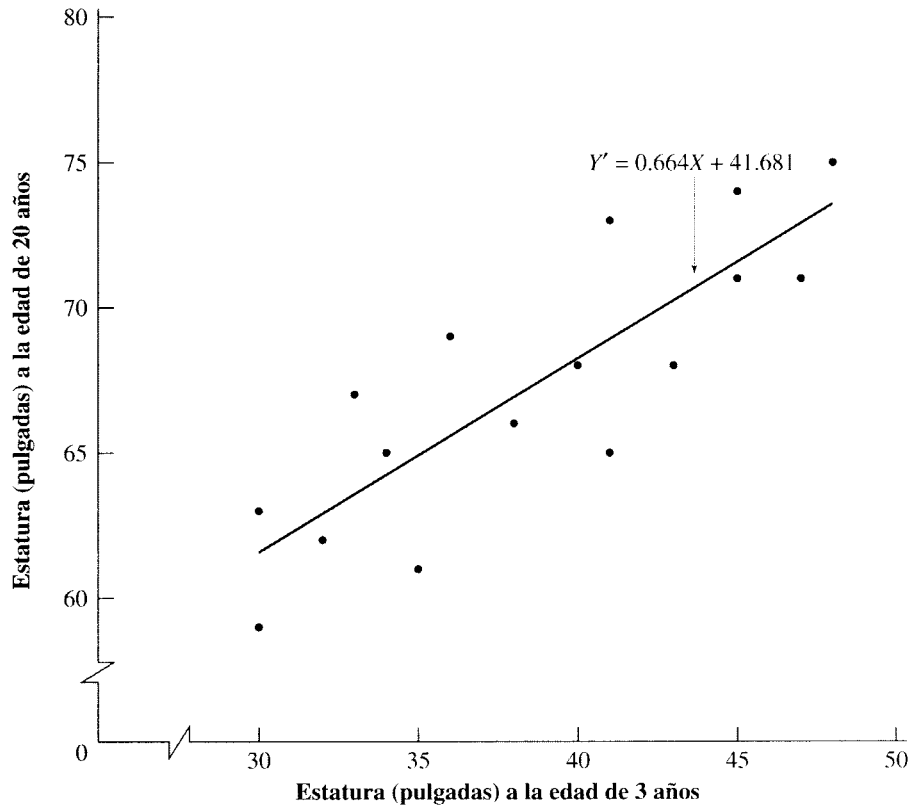
$$a_Y = \bar{Y} - b_Y \bar{X} = 67.3125 - 0.6636(38.625)$$

$$= 41.681$$

Por lo tanto,

$$Y' = b_Y X + a_Y = 0.664X + 41.681$$

Esta solución se presenta también en la tabla anterior. La recta de regresión por mínimos cuadrados se muestra sobre la gráfica de dispersión siguiente.



- c. Predicción de la estatura para el niño de 3 años de edad que mide 42 pulgadas:

$$Y' = 0.664X + 41.681$$

$$= 0.664(42) + 41.681$$

$$= 69.57 \text{ pulgadas}$$

Problema de práctica 7.2

Una neuróloga sospecha que los bajos niveles de serotonina, un neurotransmisor cerebral, pueden estar relacionados de manera causal con el comportamiento agresivo. Como primer paso en su investigación, esta científica decide realizar un estudio de correlación con nueve macacos de la India. Los monos son observados durante 6 meses y se registra el número de actos agresivos que realizan. También se miden los niveles de serotonina en el estriátum (una región del cerebro asociada con el comportamiento agresivo) una vez al día en cada animal. Los datos resultantes se presentan en la tabla siguiente. El número de actos agresivos por cada animal es el promedio de los 6 meses, expresado sobre una base diaria. Los niveles de serotonina también son los valores promedio durante el periodo de 6 meses.

- Trace una gráfica de dispersión para los datos.
- Si los datos están relacionados de manera lineal, obtenga la recta de regresión por mínimos cuadrados para predecir el número de actos agresivos a partir del nivel de serotonina.
- Con base en estos datos, ¿cuál es el número de actos agresivos por día que se podría predecir si en cierto macaco se midió un nivel de serotonina de 0.46 microgramos/gramo?

Sujeto número	Nivel de serotonina (microgramos/gramo) X	Número de actos agresivos/día Y	XY	X^2
1	0.32	6.0	1.920	0.1024
2	0.35	3.8	1.330	0.1225
3	0.38	3.0	1.140	0.1444
4	0.41	5.1	2.091	0.1681
5	0.43	3.0	1.290	0.1849
6	0.51	3.8	1.938	0.2601
7	0.53	2.4	1.272	0.2809
8	0.60	3.5	2.100	0.3600
9	0.63	2.2	1.386	0.3969
Total	4.16	32.8	14.467	2.0202

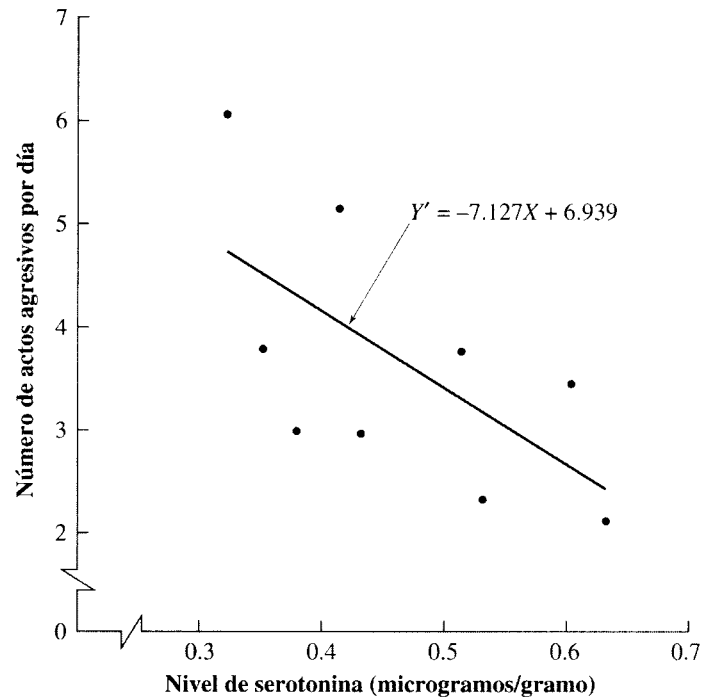
$$b_Y = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sum X^2 - \frac{(\sum X)^2}{N}} = \frac{14.467 - \frac{(4.16)(32.8)}{9}}{2.0202 - \frac{(4.16)^2}{9}} = -7.127$$

$$a_Y = \bar{Y} - b_Y \bar{X} = 3.6444 - (-7.1274)(0.4622) = 6.939$$

$$Y' = b_Y X + a_Y = -7.127X + 6.939$$

SOLUCIÓN

- a. A continuación se muestra la gráfica de dispersión. Está claro que existe una relación imperfecta, lineal y negativa entre las dos variables.



- b. Encuentre la recta de regresión por mínimos cuadrados. La solución se presenta al pie de la tabla precedente y la recta de regresión ha sido trazada sobre la gráfica de dispersión que aparece arriba.
- c. Predicción del número de actos agresivos:

$$\begin{aligned}
 Y' &= -7.127X + 6.939 \\
 &= -7.127(0.46) + 6.939 \\
 &= 3.7 \text{ actos agresivos por día}
 \end{aligned}$$

**REGRESIÓN DE X SOBRE Y**

Hasta ahora nos hemos ocupado de predecir los datos Y a partir de los datos X . Para lograrlo, obtuvimos una línea de regresión que nos permitió predecir Y conociendo X . Esta línea se conoce a veces como la *recta de regresión de Y sobre X* . También es posible predecir X conociendo Y . Sin embargo, para predecir X conociendo Y , debemos obtener una nueva recta de regresión. No podemos utilizar la ecuación de regresión para la predicción de Y cuando se conoce X . Por ejemplo, en el problema referente al CI (X) y el promedio de las calificaciones (Y), obtuvimos la siguiente recta de regresión:

$$Y' = 0.074X - 7.006$$

Esta recta resultó apropiada para predecir el promedio de calificaciones cuando el CI es conocido; es decir, para predecir Y conociendo X . Sin embargo, si queremos predecir el CI (X) cuando el promedio de calificaciones (Y) es conocido, no podemos utilizar esta recta de regresión. Ahora tendremos que obtener nuevas constantes de regresión, puesto que la recta de regresión anterior fue calculada con el propósito de minimizar los errores en la variable Y . Los errores que minimizamos han sido representados mediante rectas verticales paralelas al eje Y [véase la figura 7.2(a)]. Ahora queremos minimizar los errores en la variable X . Estos errores se representarían por medio de rectas horizontales paralelas al eje X . Un ejemplo ilustrativo de esto se representa en la figura 7.2(a), por medio de la línea punteada que une el punto E con la recta de regresión.

En términos generales, el proceso de minimizar los errores en Y' y el de minimizar los errores en X' no conducen a las mismas líneas de regresión. La excepción se presenta cuando la relación es perfecta. En este caso, ambas líneas de regresión coinciden, formando una recta única que pasa por todos los puntos. La línea de regresión para la predicción de X a partir de Y se conoce a veces como la *recta de regresión de X sobre Y* , o simplemente como la *regresión de X sobre Y* . Para ilustrar el cálculo de la línea emplearemos de nuevo los datos del CI y el promedio de calificaciones. Sin embargo, esta vez haremos la predicción del CI (X) a partir del promedio de las calificaciones (Y). Por comodidad, los datos se presentan ahora en la tabla 7.3. La ecuación de regresión lineal para predecir X dado Y es.

$$X' = b_X Y + a_X \quad \text{ecuación de regresión lineal para predecir } X \text{ dado } Y$$

tabla 7.3 CI y promedio de calificaciones de 12 estudiantes: Predicción de X a partir de Y

Estudiante número	CI X	Promedio de calificaciones		XY	Y^2
		Y			
1	110	1.0		110.0	1.00
2	112	1.6		179.2	2.56
3	118	1.2		141.6	1.44
4	119	2.1		249.9	4.41
5	122	2.6		317.2	6.76
6	125	1.8		225.0	3.24
7	127	2.6		330.2	6.76
8	130	2.0		260.0	4.00
9	132	3.2		422.4	10.24
10	134	2.6		348.4	6.76
11	136	3.0		408.0	9.00
12	138	3.6		496.8	12.96
<i>Total</i>	1503	27.3		3488.7	69.13

$$b_X = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sum Y^2 - \frac{(\sum Y)^2}{N}} = \frac{3488.7 - \frac{1503(27.3)}{12}}{69.13 - \frac{(27.3)^2}{12}}$$

$$= \frac{69.375}{7.0225} = 9.879$$

$$a_X = \bar{X} - b_X \bar{Y} = 125.25 - 9.879(2.275) = 102.775$$

$$X' = b_X Y + a_X = 9.879Y + 102.775$$

donde X' = predicción del valor de X
 b_X = pendiente de la recta que minimiza los errores en X'
 a_X = intersección con el eje X de la recta que minimiza los errores en X'

Las ecuaciones para b_X y a_X son

$$b_X = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{SC_Y} = \frac{(\sum XY) - \frac{(\sum X)(\sum Y)}{N}}{\sum Y^2 - \frac{(\sum Y)^2}{N}}$$

ecuación para el cálculo de la constante de regresión b para predecir X conociendo Y

donde SC_Y = suma de los cuadrados de los datos Y

$$= \sum Y^2 - \frac{(\sum Y)^2}{N}$$

$a_X = \bar{X} - b_X \bar{Y}$ constante de regresión a para predecir X conociendo Y

despejando b_X y a_X ,

$$b_X = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sum Y^2 - \frac{(\sum Y)^2}{N}}$$

$$b_X = \frac{3488.7 - \frac{1503(27.3)}{12}}{69.13 - \frac{(27.3)^2}{12}} = \frac{69.375}{7.0225} = 9.879$$

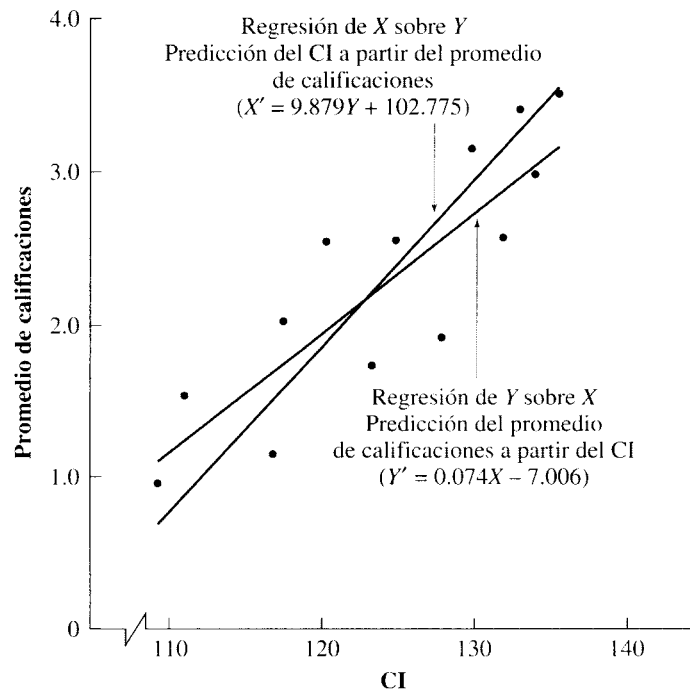


figura 7.4 Regresión de X sobre Y y regresión de Y sobre X.

$$\begin{aligned}
 a_X &= \bar{X} - b_X \bar{Y} \\
 &= 125.25 - 9.8790(2.275) \\
 &= 102.775
 \end{aligned}$$

La ecuación de regresión lineal para predecir X a partir de Y es

$$\begin{aligned}
 X' &= b_X Y + a_X \\
 &= 9.879Y + 102.775
 \end{aligned}$$

La figura 7.4 muestra esta recta, junto con la línea para la predicción de Y dado X . Observe que las dos rectas son distintas, tal como sería de esperar en virtud de que la relación es imperfecta. La solución se presenta resumida en la tabla 7.3.

Aunque existen diferentes ecuaciones para calcular la segunda recta de regresión, éstas se utilizan en muy pocas ocasiones. En lugar de eso, a la variable por predecir se la designa en la práctica como Y' y a la variable conocida se la designa como X . Así pues, si quisiéramos predecir el CI a partir del promedio de las calificaciones, designaríamos al CI como la variable Y' y al promedio de calificaciones como la variable X , para luego utilizar la ecuación de regresión para predecir Y conociendo X .



MEDICIÓN DE ERRORES DE PREDICCIÓN: EL ERROR ESTÁNDAR DE LA ESTIMACIÓN

La recta de regresión representa nuestra mejor estimación de los datos Y a partir de los valores X correspondientes. Sin embargo, a menos que la relación entre X y Y sea perfecta, la mayor parte de los valores reales de Y no estarán sobre la línea de regresión. Así pues, cuando la relación es imperfecta, necesariamente habrá errores en la predicción y será útil conocer la magnitud de éstos. Por ejemplo, tomando como base la relación entre el CI y el promedio de las calificaciones que vimos con anterioridad, parecería correcto afirmar que el promedio de calificaciones de John será de 3.2 cuando curse su último año de estudios de licenciatura. Sin embargo, considerando que la relación es imperfecta, es poco probable que nuestra predicción sea exacta. Bien, pero si no es perfecta, entonces ¿qué tan grande es el error? Si resulta probable que el error sea muy grande, entonces no podremos confiar mucho en la predicción. Sin embargo, si el error es pequeño, podremos tomar en serio la predicción y la utilizaremos como base para nuestras decisiones.

La medición de los errores de predicción requiere el cálculo del error estándar de la estimación. Este error es similar a la desviación estándar. Como usted recordará, la desviación estándar nos proporciona una medida de la desviación promedio en torno de la media. El error estándar de la estimación nos proporciona una medida de la desviación promedio de los errores de predicción en torno de la recta de regresión. En este contexto, la recta de regresión se puede considerar como una estimación de la media de los valores de Y para cada uno de los valores de X . Es algo así como una media “flotante” de los valores de Y , que cambia con valores de X . Con la desviación estándar, la suma de las desviaciones, $\sum (X - \bar{X})$, era igual a 0. Tuvimos que elevar al cuadrado las desviaciones para obtener un promedio significativo. La situación es la misma con el error estándar de la estimación. Como quiera que la suma de los errores de predicción, $\sum (Y - Y')$, es igual a 0, también debemos elevarlos al cuadrado. A continuación, obtenemos el promedio sumando los valores elevados al cuadrado, dividiendo entre $N - 2$, y extrayendo la raíz cuadrada del cociente (de una manera muy similar al caso de la desviación estándar). La ecuación para calcular el error estándar de la estimación para predecir Y conociendo X es

$$s_{Y|X} = \sqrt{\frac{\sum (Y - Y')^2}{N - 2}} \quad \text{error estándar de la estimación al predecir } Y \text{ dado } X$$

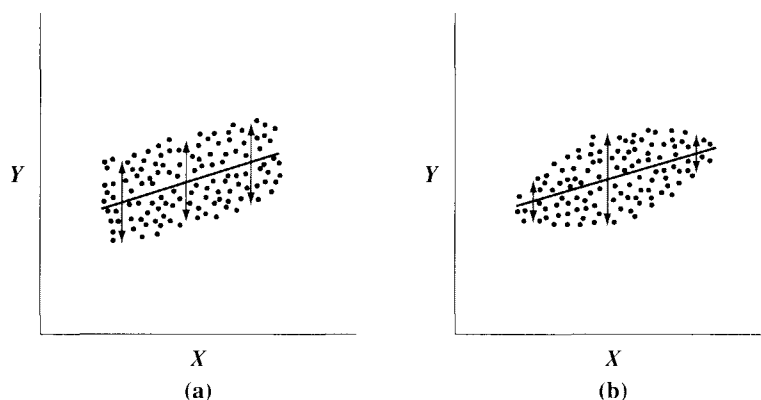
Observe que hemos dividido entre $N - 2$ y no entre $N - 1$, como en el caso de la desviación estándar de una muestra.* Los cálculos que intervienen en el uso de esta ecuación son bastante laboriosos. La ecuación que presentamos aquí para realizar dicho cálculo es mucho más sencilla de aplicar. Al determinar el coeficiente de regresión b_Y , ya habíamos calculado los valores de SC_X y SC_Y .

$$s_{Y|X} = \sqrt{\frac{SC_Y - \frac{[\sum XY - (\sum X)(\sum Y)/N]^2}{SC_X}}{N - 2}}$$

Para ilustrar el empleo de estas ecuaciones, calculemos el error estándar de estimación para los datos de los promedios de las calificaciones y el CI de las tablas 7.1 y 7.2. Como antes, la variable Y representará el promedio de las calificaciones y la variable X representará el CI, y calculemos el error estándar de estimación para la predicción del promedio de las calificaciones cuando se conoce el CI. Según lo calculado en las tablas, $SC_X = 936.25$, $SC_Y = 7.022$, $\sum XY - (\sum X)(\sum Y)/N = 69.375$, y $N = 12$. Al sustituir estos valores en la ecuación para el error estándar de estimación al predecir Y dado X , obtenemos:

$$\begin{aligned} s_{Y|X} &= \sqrt{\frac{SC_Y - \frac{[\sum XY - (\sum X)(\sum Y)/N]^2}{SC_X}}{N - 2}} \\ &= \sqrt{\frac{7.022 - \frac{(69.375)^2}{936.25}}{12 - 2}} \\ &= \sqrt{0.188} = 0.43 \end{aligned}$$

Por lo tanto, el error estándar de estimación es 0.43. Esta medida se calcula sobre todos los datos Y . Para que tenga sentido, debemos suponer que la variabilidad de Y permanece constante cuando pasamos de un dato X a otro. Esto se conoce como la hipótesis de *homosedasticidad*. La figura 7.5(a) muestra un caso supuesto de homosedasticidad. La figura 7.5(b) ilustra un caso en que la hipótesis es violada. La hipótesis de homosedasti-



Tomadas de E. W. Minium, *Statistical Reasoning in Psychology and Education*. Derechos reservados © 1978 por John Wiley & Sons, Inc. Adaptado con autorización de John Wiley & Sons, Inc.

figura 7.5 Gráficas de dispersión que muestran la variabilidad de Y como función de X .

*Dividimos entre $N - 2$ debido a que el cálculo del error estándar de la estimación implica el ajuste de los datos a una recta. Para esto debemos estimar dos parámetros, la pendiente y la ordenada al origen, dejando a las desviaciones en torno de la recta con $N - 2$ grados de libertad. En el capítulo 13 analizaremos los grados de libertad.

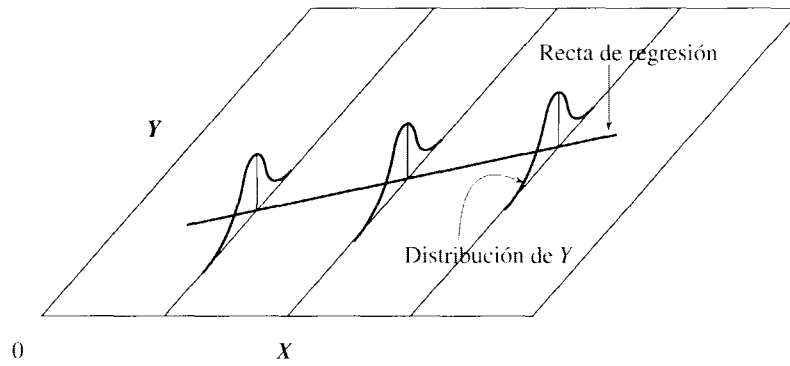


figura 7.6 Distribución normal de los datos Y en torno de la recta de regresión.

cidad implica que si separamos los datos X en columnas, la variabilidad de Y no cambia de una columna a otra. Podemos ver que esto es cierto para la figura 7.5(a), pero no para la figura 7.5(b).

¿Qué significado tiene el error estándar de estimación? Está claro que cuantifica los errores de predicción. Mientras mayor sea su valor, menor confianza tendremos en la predicción. Recíprocamente, mientras menor sea su valor, más probable será que la predicción sea precisa. Podemos ser aún más cuantitativos. Supongamos que los puntos se distribuyen de manera normal en torno a la recta de regresión (figura 7.6). Si la hipótesis fuese válida y quisiéramos construir dos rectas paralelas a la línea de regresión, a las distancias de $\pm 1s_{Y|X}$, $\pm 2s_{Y|X}$, y $\pm 3s_{Y|X}$, veríamos que cerca del 68% de los datos se localizaría entre las rectas en $\pm 1s_{Y|X}$, aproximadamente el 95% de los datos estaría entre $\pm 2s_{Y|X}$, y, en forma aproximada, el 99% de los datos se encontraría entre $\pm 3s_{Y|X}$. Pa-

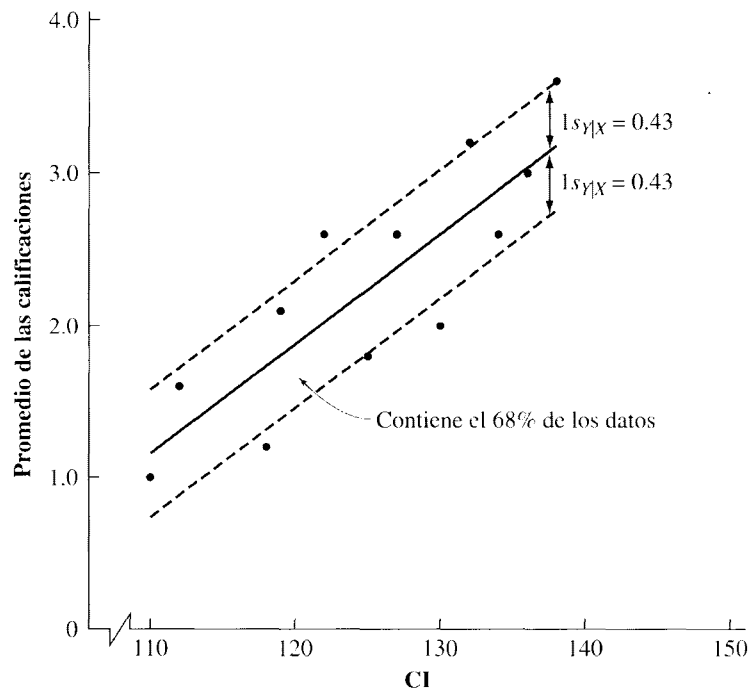


figura 7.7 Recta de regresión para los promedios de las calificaciones y el CI, con líneas paralelas $1s_{Y|X}$ arriba y abajo de la recta de regresión.

ra ilustrar este punto, en la figura 7.7 hemos trazado dos líneas punteadas, paralelas a la recta de regresión correspondiente a los promedios de las calificaciones y al CI, a una distancia de $\pm 1s_{y|X}$. También hemos incluido los datos en la figura. De acuerdo con lo dicho, cerca del 68% de los datos deberá estar entre estas líneas. Como existen 12 puntos, esperaríamos que $0.68(12) = 8$ datos estuviesen dentro de las líneas, y esto sucede efectivamente en realidad. Esta coincidencia no siempre es tan buena, en particular cuando sólo se tienen 12 datos en la muestra. Cuando N aumenta, generalmente también la coincidencia aumenta.



CONSIDERACIONES AL UTILIZAR LA REGRESIÓN LINEAL PARA LA PREDICCIÓN

Los procedimientos que hemos descrito son adecuados para predecir datos, suponiendo que exista una relación lineal entre las variables X y Y . Si la relación es no lineal, la predicción no será muy precisa. Entonces, la primera suposición que deberemos tomar en cuenta para el uso exitoso de esta técnica es que la *relación entre X y Y sea lineal*. En segundo lugar, no estamos interesados, por lo general, en utilizar la recta de regresión para predecir las calificaciones de los individuos que estaban en el grupo empleado para calcular dicha línea de regresión. Después de todo, ¿para qué predecir sus calificaciones cuando ya las conocemos? Generalmente, se determina una recta de regresión para utilizarla con sujetos en los que alguna de las variables es desconocida. Por ejemplo, en el problema del CI y el promedio de las calificaciones, un empleado del departamento de inscripciones de una universidad podría utilizar esa recta de regresión para predecir los promedios de las calificaciones de los aspirantes, al conocer sus puntajes de CI. No tiene sentido predecir los promedios de las calificaciones de los 12 estudiantes cuyos datos fueron utilizados en el problema. Dicho empleado ya conoce esos promedios. Si vamos a

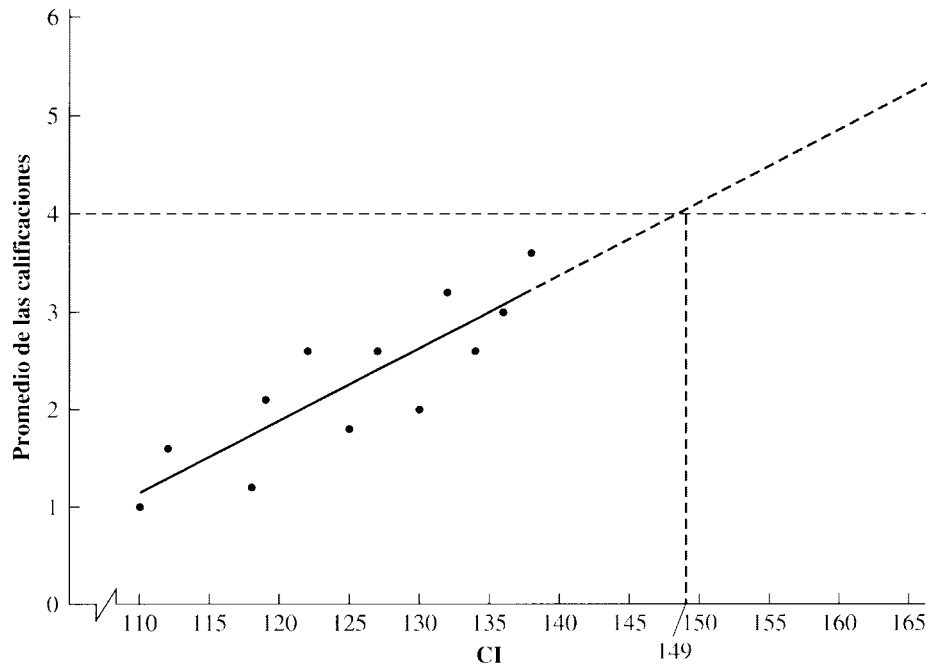


figura 7.8 Limitación de la predicción al rango de los datos básicos.

utilizar los datos de un grupo para predecir las calificaciones de otro, es importante que el grupo del cálculo básico sea representativo del grupo de predicción. Con frecuencia este requisito se controla mediante un muestreo aleatorio de la población objeto de la predicción, al utilizar la muestra para obtener la ecuación de regresión. El muestreo aleatorio se analiza en el capítulo 8. Por último, la ecuación de regresión lineal se utiliza de manera apropiada, sólo para el rango de la variable en la cual se basa. Por ejemplo, al predecir el promedio de las calificaciones mediante el CI, debemos limitar nuestras predicciones a datos de CI que varíen de 110 a 138. Como no tenemos datos fuera de este rango, no sabemos si la relación sigue siendo lineal para valores más extremos del CI.

Para ilustrar este punto, consideremos la figura 7.8, donde hemos ampliado la recta de regresión a fin de incluir los valores de CI hasta 165. En la universidad donde se tomó la muestra de los datos, el promedio máximo de calificaciones fue de 4.0. Si empleáramos la recta de regresión extendida para predecir el promedio de calificaciones que corresponde a un CI de 165, tendríamos un promedio de 5.2, un valor que obviamente es incorrecto. La predicción para un CI arriba de 165 sería todavía peor. Al observar la figura 7.8, podemos apreciar que cuando la relación se extiende más allá de un CI de 138, no puede extenderse también más allá de un CI de 149 aproximadamente (el valor del CI donde la recta de regresión alcanza un promedio de calificaciones de 4.0). Por supuesto, no hay razón alguna para creer que exista una relación más allá del punto básico de $CI = 138$, y por ese motivo no debemos hacer predicciones mediante esta relación para valores de CI mayores que 138.



RELACIÓN ENTRE LAS CONSTANTES DE REGRESIÓN Y LA r DE PEARSON

Aunque no hemos presentado este aspecto de la r de Pearson antes, se puede demostrar que la r de Pearson es la pendiente de la recta de regresión por mínimos cuadrados al graficar los datos como puntajes z . Por ejemplo, utilicemos los datos de la tabla 6.3 correspondientes al peso y el costo de seis bolsas de naranjas. Para comodidad del lector, hemos reproducido los datos en la tabla 7.4. La figura 7.9(a) muestra la gráfica de dispersión y la recta de regresión por mínimos cuadrados para estos datos en bruto. Como ésta es una relación lineal perfecta, $r = 1.00$. La figura 7.9(b) muestra la gráfica de dispersión para estas parejas de puntajes z y la recta de regresión por mínimos cuadrados para estos puntajes z . La pendiente de la recta de regresión para los datos en bruto es b y la pendiente de la recta de regresión para los puntajes z es r . Observe que la pendiente de esta última recta de regresión es 1.00, como debe ser, ya que $r = 1.00$.

tabla 7.4 Costo y peso en libras de seis bolsas de naranjas

Bolsa	Peso	Costo	z_x	z_y
	(libras) X	($\$$) Y		
A	2.25	0.75	-1.34	-1.34
B	3.00	1.00	-0.80	-0.80
C	3.75	1.25	-0.27	-0.27
D	4.50	1.50	0.27	0.27
E	5.25	1.75	0.80	0.80
F	6.00	2.00	1.34	1.34

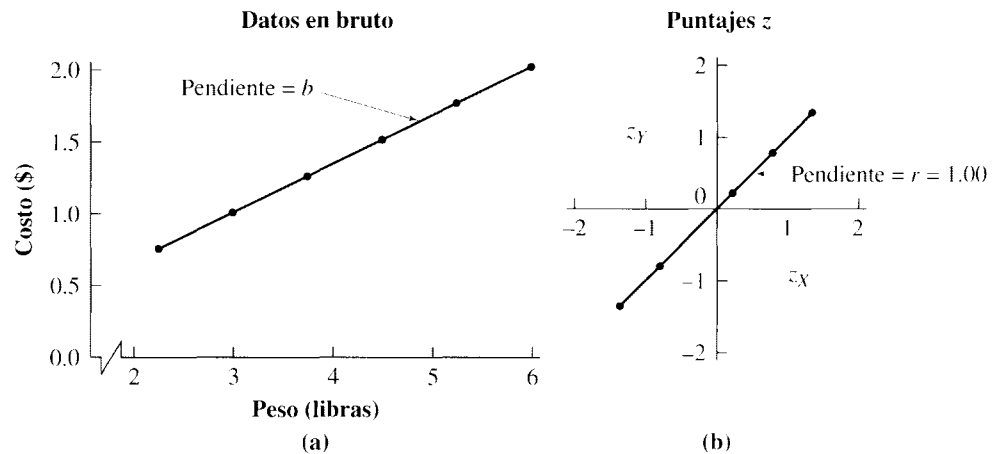


figura 7.9 Relación entre b y r .

Como la r de Pearson en una pendiente, se relaciona con b_Y y b_X . Se puede demostrar en forma algebraica que:

$$b_Y = r \frac{s_Y}{s_X}$$

y también

$$b_X = r \frac{s_X}{s_Y}$$

Estas ecuaciones son útiles si ya hemos calculado r , s_X y s_Y y queremos determinar la recta de regresión por mínimos cuadrados. Por ejemplo, en el problema del CI y el promedio de las calificaciones, $r = 0.8556$, $s_Y = 0.7990$ y $s_X = 9.2257$. Suponga que queremos determinar b_Y y a_Y cuando ya hemos calculado r , s_Y y s_X . La forma más sencilla de hacer consiste en emplear la ecuación

$$b_Y = r \frac{s_Y}{s_X} = 0.8556 \left(\frac{0.7990}{9.2257} \right) = 0.074$$

Observe que éste es el mismo valor al que llegamos anteriormente, en la página 139. Una vez determinado b_Y , calculamos a_Y de la manera usual.



REGRESIÓN MÚLTIPLE Y CORRELACIÓN MÚLTIPLE

Hasta ahora hemos analizado la regresión y la correlación mediante ejemplos con sólo dos variables. Cuando analizamos la relación entre el promedio de las calificaciones y el CI, determinamos que $r = 0.856$ y que la ecuación de la recta de regresión para predecir el promedio a partir del CI era:

$$Y' = 0.74X - 7.006$$

donde Y' = predicción del promedio de las calificaciones
 X = el dato del CI

Esta ecuación nos proporcionó una predicción bastante precisa. Aunque no lo calculamos, el error total de predicción al cuadrado [$\Sigma (Y - Y')^2$] era de 1.88 y la cantidad de varianza explicada correspondía al 73.2%. Por supuesto, existen otras variables, además del CI, que podrían afectar el promedio de las calificaciones. La cantidad de tiempo que los estudiantes dedican al estudio, la motivación para obtener mejores calificaciones y el interés que tienen por el curso son algunas de ellas. Aunque tengamos una precisión razonablemente buena de la predicción cuando se utiliza solamente el CI, podríamos obtener un mejor resultado si tuviéramos datos que relacionaran el promedio de las calificaciones con una o más de estas otras variables.

La regresión múltiple es una extensión de la regresión simple para incluir las situaciones en que hay dos o más variables de predicción. Por ejemplo, suponga que tenemos datos de los 12 estudiantes, que incluyen una segunda variable de predicción llamada "tiempo de estudio", así como los datos originales del promedio de las calificaciones y el CI. Los datos de estas tres variables aparecen en las columnas 2, 3 y 4 de la tabla 7.5. Ahora podemos obtener una ecuación de regresión para predecir el promedio de las calificaciones mediante las dos variables de predicción, el CI y el tiempo de estudio. La forma general de la ecuación de regresión múltiple para las dos variables de predicción es:

$$Y' = b_1X_1 + b_2X_2 + a$$

donde Y' = predicción del valor de Y
 b_1 = coeficiente de la primera variable de predicción
 X_1 = primera variable de predicción
 b = coeficiente de la segunda variable de predicción
 X_2 = segunda variable de predicción
 a = una constante de predicción

tabla 7.5 Comparación de la precisión de una predicción con una o dos variables de predicción

Estudiante núm.	CI (X_1)	Tiempo de estudio (horas/semana) (X_2)	Promedio de calificaciones (Y)	Predicción del promedio mediante el CI (Y')	Predicción del promedio mediante el CI y el tiempo de estudio (Y')	Error cuando se utiliza solamente el CI	Error cuando se utilizan el CI y el tiempo de estudio
1	110	8	1.0	1.14	1.13	-0.14	-0.13
2	112	10	1.6	1.29	1.46	0.31	0.13
3	118	6	1.2	1.74	1.29	-0.54	-0.09
4	119	13	2.1	1.81	2.16	0.29	-0.06
5	122	14	2.6	2.03	2.43	0.57	0.17
6	125	6	1.8	2.26	1.63	-0.46	0.17
7	127	13	2.6	2.40	2.56	0.20	0.04
8	130	12	2.0	2.63	2.59	-0.63	-0.59
9	132	13	3.2	2.77	2.81	0.42	0.39
10	134	11	2.6	2.92	2.67	-0.32	-0.07
11	136	12	3.0	3.07	2.88	-0.07	0.12
12	138	18	3.6	3.21	3.69	0.38	0.09
Error total al cuadrado = $\Sigma (Y - Y')^2 = 1.88$							= 0.63

Esta ecuación es muy similar a la que fue utilizada en el caso de la regresión simple, excepto que hemos agregado otra variable de predicción y su coeficiente. Como antes, los valores del coeficiente y la constante se determinan de acuerdo con el criterio de mínimos cuadrados, donde $\Sigma (Y - Y')^2$ alcanza un mínimo. Sin embargo, esta vez las matemáticas son complicadas y los cálculos reales casi siempre se realizan en una computadora dotada de un software estadístico. Para los datos de nuestro ejemplo, la ecuación de regresión múltiple que minimiza los errores en Y está dada por:

$$Y' = 0.049X_1 + 0.118X_2 - 5.249$$

donde Y' = predicción del valor del promedio de las calificaciones
 $b_1 = 0.049$
 X_1 = puntaje del CI
 $b = 0.118$
 X_2 = dato del tiempo de estudio
 $a = 5.249$

Para determinar si la precisión de la predicción aumenta al utilizar la ecuación de regresión múltiple, en la quinta columna de la tabla 7.5 colocamos las predicciones de los promedios de calificaciones al emplear sólo el CI para la predicción, y en la sexta columna aparecen las predicciones de dichos promedios al utilizar el CI y el tiempo de estudio como variable de predicción, y los errores de predicción respectivos se han presentado en las columnas 7 y 8 respectivamente. También graficamos en la figura 7.10(a) el valor Y real y las dos predicciones de los valores de Y' para cada estudiante. Ordenamos a los alumnos de izquierda a derecha sobre el eje X , de acuerdo con la mayor precisión de la predicción resultante para cada uno, al utilizar la ecuación de regresión múltiple. En la figura 7.10(b) graficamos el porcentaje de mejoría en la precisión de la predicción para cada estudiante, que resulta del hecho de utilizar el CI + el tiempo de estudio, en vez de usar sólo el CI. Al ver la tabla 7.5 y la figura 7.10, resulta claro que el uso de la ecuación de regresión múltiple ha mejorado en gran medida la precisión general de la predicción. Por ejemplo, la precisión de la predicción aumentó en todos los estudiantes, excepto en la del número 11, y para el estudiante número 3 la precisión aumentó en casi 40%. En la parte inferior de la tabla 7.5 mostramos también el valor de $\Sigma (Y - Y')^2$ para cada recta de regresión. Al agregar la segunda variable de predicción, el error de predicción total al cuadrado se redujo de 1.88 a 0.63, una mejora de más de 66 por ciento.

Como en este ejemplo aumentó la precisión de la predicción al utilizar dos variables en vez de una, esto implica que la proporción de la varianza de Y explicada ha aumentado también. Al tratar de determinar esta proporción, usted podría sentir la tentación de ampliar el concepto de r^2 de nuestro análisis anterior de la correlación, a calcular r^2 entre el promedio de calificaciones y cada variable, y luego sumar simplemente los valores resultantes. La tabla 7.6 muestra una matriz de correlación r de Pearson que corresponde al promedio de calificaciones, el CI y el tiempo de estudio. Si siguiéramos este procedimiento, la proporción de varianza explicada sería mayor que 1.00 $[(0.856)^2 + (0.829)^2 = 1.42]$, lo cual es imposible. No es factible explicar más del 100% de la variabilidad. El error se produce debido a que hay un traslape en la varianza explicada entre el CI y el tiempo de estudio. Los alumnos con mayor CI tienden a estudiar más tiempo. Por lo tanto, parte de la varianza en el promedio de las calificaciones explicada por el CI también puede explicarse por medio del tiempo de estudio. Para corregir esta situación, debemos tomar en cuenta la correlación entre el CI y el tiempo de estudio.

La ecuación correcta para el cálculo de la proporción de la varianza explicada, cuando existen dos variables de predicción está dada por:

$$R^2 = \frac{r_{YX_1}^2 + r_{YX_2}^2 - 2r_{YX_1}r_{YX_2}r_{X_1X_2}}{1 - r_{X_1X_2}^2}$$

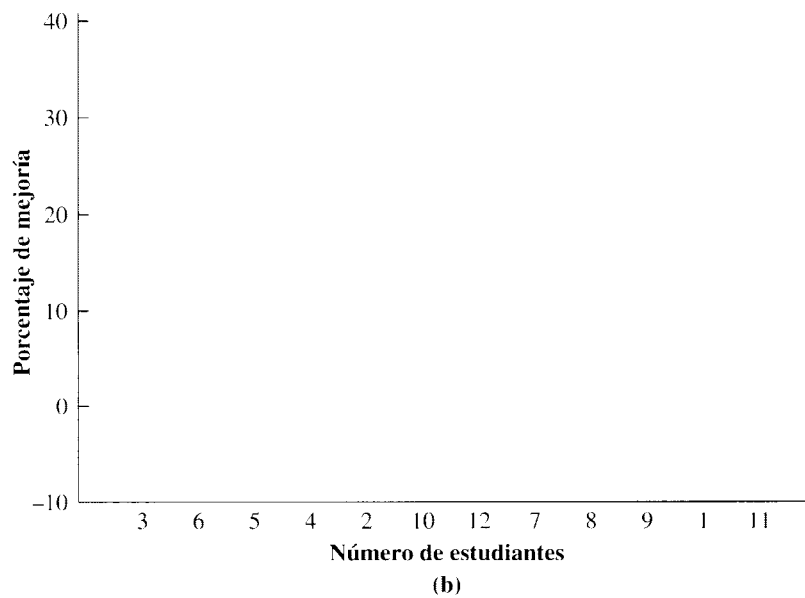
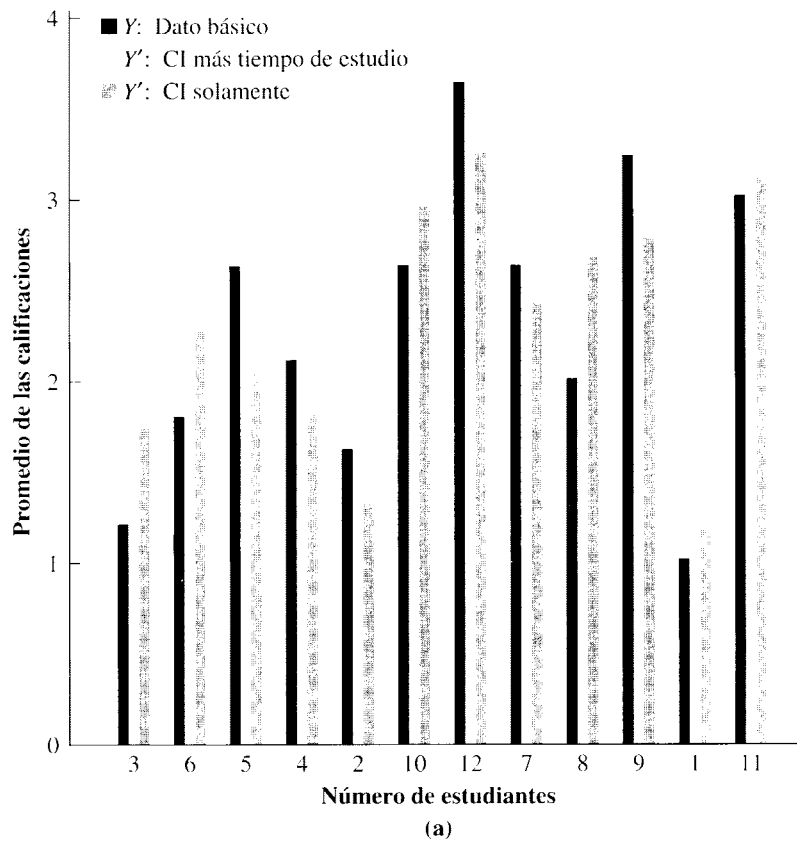


figura 7.10 Comparación de la precisión de las predicciones utilizando una o dos variables de predicción.

tabla 7.6 Matriz de correlación de Pearson entre el CI, el tiempo de estudio y el promedio de las calificaciones

	CI (X_1)	Tiempo de estudio (X_2)	Promedio de las calificaciones (Y)
CI (X_1)	1.000		
Tiempo de estudio (X_2)	0.560	1.000	
Promedio de las calificaciones (Y)	0.856	0.829	1.000

donde R^2 = coeficiente múltiple de determinación
 r_{YX_1} = correlación entre Y y la variable de predicción X_1
 r_{YX_2} = correlación entre Y y la variable de predicción X_2
 $r_{X_1X_2}$ = correlación entre las variables de predicción X_1 y X_2

A menudo R^2 recibe también el nombre de correlación múltiple cuadrada. Tomando como base los datos del presente estudio, r_{YX_1} = la correlación entre el promedio de las calificaciones y el CI = 0.856, r_{YX_2} = la correlación entre Y y el tiempo de estudio = 0.829, y $r_{X_1X_2}$ = la correlación entre el CI y el tiempo de estudio = 0.560. Para estos datos,

$$R^2 = \frac{(0.856)^2 + (0.829)^2 - 2(0.856)(0.829)(0.560)}{1 - (0.560)^2} = 0.910$$

Así pues, la proporción de varianza explicada ha aumentado de 0.73 a 0.91 por el hecho de utilizar el CI y el tiempo de estudio.

Por supuesto, el mero hecho de agregar otra variable de predicción no incrementa necesariamente la precisión de la predicción o la cantidad de varianza explicada. Esto depende de la fuerza de la relación entre la variable objeto de la predicción y la variable de predicción adicional, y también de la fuerza de la relación entre las propias variables de predicción. Por ejemplo, observemos lo que le sucede a R^2 cuando $r_{X_1X_2} = 0$. Este tema es analizado con mayor detalle en textos más avanzados.

RESUMEN

En este capítulo hemos analizado la forma de utilizar la relación entre dos variables para propósitos de predicción. Cuando en la predicción se utiliza la recta que mejor se ajusta a los puntos, dicha línea recibe el nombre de recta de regresión. La recta de regresión que se usa más comúnmente para las relaciones lineales imperfectas, se ajusta a los puntos de acuerdo con un criterio de mínimos cuadrados. A continuación presentamos las ecuaciones para determinar la recta de regresión por mínimos cuadrados cuando se desea predecir Y dado X , y la recta de regresión para predecir X dado Y . Las dos rectas no son la misma, a menos que la relación sea perfecta. Después aprovechamos estas ecuaciones para construir rectas de regresión para diversos conjuntos de datos y mostramos la forma de

utilizar esas rectas para hacer predicciones. Luego analizamos la forma de cuantificar los errores de predicción, calculando el error estándar de una estimación. Presentamos las condiciones bajo las cuales el uso de la recta de regresión lineal resulta apropiado: la relación tiene que ser lineal; la recta de regresión deberá haberse obtenido de datos representativos del grupo sobre el cual se desea hacer la predicción, y la predicción habrá de limitarse al rango de los datos básicos. A continuación, estudiamos la relación entre b y r . Por último, presentamos el tema de la regresión y la correlación múltiples; el coeficiente múltiple de determinación, R^2 , y mostramos cómo el hecho de usar dos variables puede incrementar la precisión de la predicción.

TÉRMINOS IMPORTANTES

Coefficiente múltiple de determinación (página 155)	Error estándar de estimación (página 146)	Regresión (página 134)
Constante de regresión (página 137)	Homosedasticidad (página 147)	Regresión de X sobre Y (página 143)
Correlación múltiple (página 151)	Recta de regresión (página 135)	Regresión de Y sobre X (página 137)
	Recta de regresión por mínimos cuadrados (página 137)	Regresión múltiple (página 151)

PREGUNTAS Y PROBLEMAS

- Defina o identifique cada uno de los conceptos contenidos en la sección “Términos importantes”.
- Mencione algunas situaciones en las que sería útil disponer de una predicción exacta.
- La recta de regresión por mínimos cuadrados minimiza $\sum (Y - Y')^2$, en lugar de $\sum (Y - Y')$. ¿Es correcta esta afirmación? Explique.
- La recta de regresión por mínimos cuadrados es la línea de predicción que produce los “éxitos” más directos. ¿Es correcta esta afirmación? Explique.
- En general, la recta de regresión de Y sobre X no es igual a la recta de regresión de X sobre Y . ¿Es correcta esta afirmación? Explique.
- ¿Qué utilidad tiene conocer el error estándar de estimación para un conjunto de parejas de datos X y Y ?
- ¿Por qué hay por lo general dos rectas de regresión, pero sólo un coeficiente de correlación para cualquier conjunto de parejas de datos?
- ¿Qué nombre recibe R^2 ? ¿Es cierto que, desde un punto de vista conceptual, R^2 es análogo a r^2 , excepto que R^2 se aplica a situaciones en las que hay dos o más variables de predicción? Explique. ¿Aumenta la exactitud de la predicción siempre que se usa una segunda variable de predicción? Explique.
- Dado el siguiente conjunto de parejas de datos X y Y .

X	7	10	9	13	7	11	13
Y	1	2	4	3	3	4	5

- Construya una gráfica de dispersión de las parejas de datos. ¿Le parece a usted que la relación sea lineal?
 - Determine la recta de regresión por mínimos cuadrados para predecir Y conociendo X .
 - Determine la recta de regresión por mínimos cuadrados para predecir X conociendo Y . ¿Son iguales estas rectas? Explique.
 - Trace las dos rectas de regresión sobre la gráfica de dispersión.
 - Utilice la relación entre X y Y , para predecir el valor de Y si $X = 12$ (redondeado a dos cifras decimales).
10. Un psicólogo clínico se interesa en estudiar la relación entre el nivel de testosterona de los hombres casados y la calidad de su relación marital. Entonces se realiza un estudio en el cual se miden

los niveles de testosterona de ocho hombres casados. Los ocho hombres responden también un cuestionario por escrito y estandarizado que permite evaluar la calidad de su relación marital. La escala del cuestionario es de 0 a 25, correspondiendo las cifras más altas a las mejores relaciones. Los datos acerca de la testosterona están expresados en nanomoles por litro de suero. Los datos se presentan a continuación.

Número de sujeto	1	2	3	4	5	6	7	8
Puntaje de la relación	24	15	15	10	19	11	20	19
Nivel de testosterona	12	13	19	25	9	16	15	21

- En una hoja de papel para graficar (“milimétrico”), elabore una gráfica de dispersión de los datos. Use el nivel de testosterona como variable X .
 - Describa la relación que se aprecia en la gráfica.
 - Calcule el valor de la r de Pearson.
 - Determine la recta de regresión por mínimos cuadrados para predecir el puntaje de la relación a partir del nivel de testosterona. ¿Cree usted que b_Y debe ser positivo o negativo? ¿Por qué?
 - Dibuje la recta de regresión por mínimos cuadrados obtenida en **d** sobre la gráfica de dispersión obtenida en **a**.
 - Tomando como base los datos de los ocho hombres, ¿qué puntaje podría usted predecir para la relación de un hombre cuyo nivel de testosterona es de 23 nanomoles por litro de suero? clínica, salud, biológica
11. Una de las atracciones de una feria recién llegada al pueblo es el puesto donde el señor Clarividente (un brillante estudiante de estadística con una moralidad un tanto cuestionable) afirma que puede adivinar el peso de las mujeres, con un error de 1 kilogramo, al estudiar las líneas de las manos y los dedos. Él apuesta que si se equivoca, la mujer podrá elegir cualquiera de los animales de peluche del puesto. Sin embargo, si adivina el peso, como recompensa por sus poderes la mujer le deberá pagar \$2. Lo que no perciben las mujeres que aceptan la apuesta es que el señor Clarividente mide la longitud del dedo índice de la mano iz-

quiera al “estudiarlas”. Las damas tampoco conocen, pero el Clarividente sí, la siguiente relación entre el peso de las mujeres y la longitud de los dedos índices de la mano izquierda:

Longitud del dedo índice izquierdo (cm)	5.6	6.2	6.0	5.4
Peso (kg)	79.0	83.5	82.0	77.5

- Si usted contara con toda esta información por adelantado, ¿apostarí con el señor Clarividente? Explique.
 - Con los datos de la tabla acompañante, ¿cuál es la recta de regresión por mínimos cuadrados que permite predecir el peso de una mujer, cuando la longitud de su índice izquierdo es conocida?
 - Con la recta de regresión por mínimos cuadrados obtenida en **b**, si el dedo índice de una mujer mide 5.7 centímetros, ¿cuál será la predicción acerca de su peso (redondeado a dos cifras decimales)? *cognitivo*
12. Un profesor de estadística realiza un estudio para investigar la relación que existe entre el rendimiento de sus estudiantes en los exámenes y su respectivo nivel de ansiedad. Elige a diez estudiantes de su clase para el experimento. Inmediatamente antes de presentar el examen final, los 10 estudiantes contestan un cuestionario sobre ansiedad. A continuación presentamos las calificaciones de la prueba final y los datos de ansiedad para los 10 estudiantes:

Ansiedad	28	41	35	39	31	42	50	46	45	37
Exam. final	82	58	63	89	92	64	55	70	51	72

- En una hoja de papel para graficar (“milimétrico”), elabore una gráfica de dispersión de las parejas de calificaciones. Utilice la ansiedad como variable X .
- Describa la relación que aparece en la gráfica.
- Suponga que la relación es lineal y calcule el valor de la r de Pearson.
- Determine la recta de regresión por mínimos cuadrados para predecir la calificación del examen final, dado el nivel de ansiedad. ¿Deberá b_y ser positiva o negativa? ¿Por qué?
- Trace la recta de regresión por mínimos cuadrados obtenida en **d** sobre la gráfica de dispersión obtenida en **a**.
- Tomando como base los datos de los 10 alumnos, si un estudiante tiene un nivel de ansiedad de 38, ¿qué valor se podría predecir para su calificación en el examen final (redondeado a dos cifras decimales)?
- Calcule el error estándar de estimación para la predicción de las calificaciones del examen fi-

nal, a partir de los niveles de ansiedad. *clínica, salud, educación*

13. El gerente de ventas de una gran tienda de artículos deportivos ha iniciado una campaña de publicidad a nivel nacional. Además, ha llevado un registro de los costos mensuales de la publicidad y las ganancias mensuales, los cuales se presentan aquí. Los datos están en miles de dólares.

Mes	Ene.	Feb.	Mar.	Abr.	May	Jun.	Jul.
Costo men. de la publicidad	10.0	14.0	11.4	15.6	16.8	11.2	13.2
Ganancia mensual	125	200	160	150	210	110	125

- Suponiendo que exista una relación lineal, obtenga la recta de regresión por mínimos cuadrados para predecir las ganancias mensuales a partir de los costos mensuales de publicidad.
 - En agosto, el gerente planea invertir \$17 000 en publicidad. Con base en estos datos, ¿cuál es la ganancia esperada ese mes (redondeada a miles de dólares)?
 - Dada la relación que presentan las parejas de datos, ¿podría pensar en alguna razón por la cual el gerente no quisiera gastar mucho más dinero en publicidad? *1/0*
14. En un artículo reciente de un periódico se informó que “entre los entrenadores de la NBA existe una fuerte correlación entre la continuidad y el éxito”. El artículo se basó en los siguientes datos:

Entrenador, equipo	Antigüedad como entrenador del mismo equipo (años)	Récord 1996–1997 (% de juegos ganados)
Jerry Sloan, Utah	9	79
Phil Jackson, Chicago	8	84
Rudy Tomjanovich, Houston	6	70
George Karl, Seattle	6	70
Lenny Wilkens, Atlanta	4	68
Mike Fratello, Cleveland	4	51
Larry Brown, Indiana	4	48

- ¿Es verdad lo que dice el artículo acerca de la fuerte correlación que existe entre la continuidad y el éxito en el caso de los entrenadores de la NBA?
- Obtenga la recta de regresión por mínimos cuadrados para predecir el éxito (% de juegos ganados) a partir de la antigüedad.
- Tomando como base su respuesta obtenida en **b**, ¿qué “% de juegos ganados” podría usted

predecir para un entrenador de la NBA que tiene 7 años de “antigüedad” con el mismo equipo? 1/0, otra

15. Al ver la alta inflación, el señor Chávez ha cuidado mucho de su presupuesto. Como su casa tiene calefacción eléctrica, llevó un registro de la cuenta del consumo mensual de energía eléctrica durante el año pasado, y del promedio mensual de la temperatura exterior. Los datos aparecen en la tabla siguiente. La temperatura está en grados Celsius y la cuenta de la energía, está en dólares.
- Suponga que existe una relación lineal entre la temperatura promedio mensual y la cuenta mensual de energía, y determine la recta de regresión por mínimos cuadrados para predecir la cuenta mensual de energía a partir de la temperatura promedio mensual.
 - Tomando como base la predicción del clima para este año, el señor Chávez espera un invierno más frío. Si en febrero hay 8 grados menos este año, ¿cuánto deberá reservar el señor Chávez en su presupuesto para pagar la cuenta de energía en febrero? Al pensar la respuesta, suponga que los costos de la electricidad aumentarán 10% con respecto a los costos del año anterior, a causa de la inflación.
 - Calcule el error estándar de estimación para predecir la cuenta mensual de energía eléctrica a partir de la temperatura promedio mensual.

Mes	Temperatura promedio	Cuenta de energía
Enero	10	120
Febrero	18	90
Marzo	35	118
Abril	39	60
Mayo	50	81
Junio	65	64
Julio	75	26
Agosto	84	38
Septiembre	52	50
Octubre	40	80
Noviembre	25	100
Diciembre	21	124

otra

16. Tome como referencia el problema 16 del capítulo 6. En ese problema se presentaron los datos so-

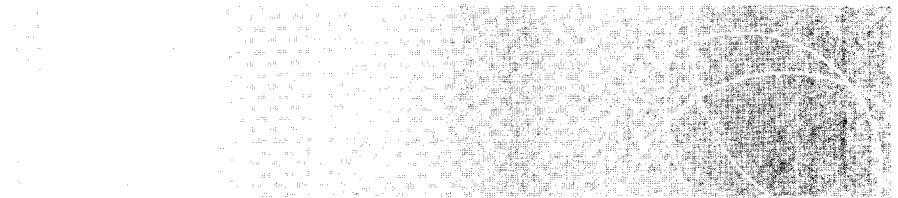
bre la relación entre el peso al nacer y el puntaje de CI futuro de siete estudiantes de psicología seleccionados de una universidad en particular.

- Suponiendo que exista una relación lineal, use estos datos para determinar la recta de regresión por mínimos cuadrados que nos permita predecir el CI a partir del peso al nacer.
 - Utilizando esta recta de regresión, ¿qué CI podría usted predecir si el peso al nacer fue de 7.5? de desarrollo
17. Tome como referencia el problema 21 del capítulo 6. En ese problema se presentaron datos sobre la relación entre el número de refrescos que ocho muchachos de 12 años de edad consumieron en una semana, y sus índices de masa corporal (BMI). Esos muchachos de 12 años fueron voluntarios seleccionados aleatoriamente en una escuela secundaria local de una ciudad grande del noroeste.
- Suponiendo que los datos presenten una relación lineal, obtenga la recta de regresión por mínimos cuadrados para predecir el BMI a partir del número de refrescos consumidos.
 - Utilizando esta recta de regresión, ¿qué BMI podría usted predecir para un muchacho de 12 años que estudia en esta escuela y consume un promedio de 17 refrescos a la semana? salud
18. Tome como referencia el problema 22 del capítulo 6. En ese problema se presentaron los datos obtenidos en un estudio realizado para determinar la relación entre la devoción religiosa y la autoestima.
- Suponiendo que exista una relación lineal, deduzca la recta de regresión por mínimos cuadrados que permita predecir la autoestima a partir de la devoción religiosa.
 - Utilizando esta recta de regresión, ¿qué valor de autoestima podría usted predecir en el caso de un estudiante de octavo grado cuya devoción religiosa fue valorada en 43? social, de desarrollo
19. Consulte el problema 24 del capítulo 6. En este problema se le pidió que recomendase un examen de selección para los aspirantes a empleados en la sección de manufactura de una corporación grande. Usted recomendó que se usara el examen 2. Ahora la pregunta es ésta: ¿sería mejor utilizar ambos exámenes, en lugar de usar solamente el 2? Explique su respuesta en términos de R^2 y r^2 . Utilice una computadora y software de estadística para resolver este problema, en caso de que disponga de ellos. 1/0

Si desea encontrar problemas de práctica adicionales totalmente resueltos sobre los temas que acaba de estudiar, busque Chapter 7 en el material adicional del libro, disponible en inglés en: <http://www.thomsonlearning.com.mx>.

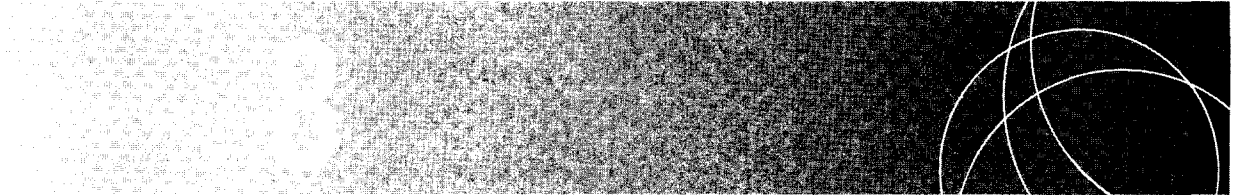
Si quiere resolver problemas para computadora sobre los temas que acaba de estudiar, busque Chapter 7 en el material adicional (en inglés) del libro, en: <http://www.thomsonlearning.com.mx>.

ESTADÍSTICA INFERENCIAL



- 8** Muestreo aleatorio y probabilidad
- 9** Distribución binomial
- 10** Introducción a la prueba de hipótesis mediante la prueba del signo
- 11** Potencia
- 12** Distribuciones muestrales, distribución muestral de la media, prueba (z) de la desviación normalizada
- 13** Prueba t de Student para muestras simples
- 14** Prueba t de Student para grupos independientes correlacionados
- 15** Introducción al análisis de varianza
- 16** Comparaciones múltiples
- 17** Introducción al análisis de varianza de dos factores o vías
- 18** Ji-cuadrada y otras pruebas no paramétricas
- 19** Repaso de estadística inferencial

Muestreo aleatorio y probabilidad



Introducción

Muestreo aleatorio

Técnicas para el muestreo aleatorio

Muestreo con o sin reemplazo

Probabilidad

Algunos aspectos básicos relativos a los valores de probabilidad

Cálculo de probabilidades

La regla de la suma

La regla del producto

Reglas de la suma y del producto

Probabilidad y variables continuas con distribución normal

¿CUÁL ES LA VERDAD? "No soy culpable, soy una víctima de las coincidencias":

¿excusa descartada o verdad?

¿CUÁL ES LA VERDAD? Declive del número de espermatozoides:

¿Falla del hombre o del muestreo?

¿CUÁL ES LA VERDAD? La muestra de una muestra

Resumen

Términos importantes

Preguntas y problemas

Notas



INTRODUCCIÓN

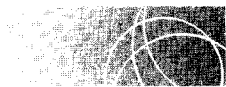
Hemos concluido nuestro análisis de la estadística descriptiva y ya estamos listos para comenzar a estudiar el área fascinante de la estadística inferencial. Con la estadística des-

criptiva, nuestra preocupación principal era lograr la presentación y descripción de los conjuntos de datos de la manera más significativa y eficaz. En la estadística inferencial vamos más allá de la mera descripción de los datos, ya que un objetivo básico es el uso de los datos provenientes de una muestra para hacer una afirmación acerca de una característica de la población. Se hacen dos tipos de afirmaciones: una tiene que ver con la actividad conocida como *prueba de hipótesis* y la otra con la *estimación de parámetros*.

En la prueba de hipótesis, el investigador recolecta datos de un experimento sobre el conjunto de sujetos de una muestra, en un intento de validar cierta hipótesis relativa a una población. Por ejemplo, suponga que una psicóloga dedicada a la educación cree que un nuevo método para enseñar la asignatura de matemáticas a los niños de tercer grado, en su distrito escolar (población), es mejor que la forma usual de enseñanza. En el experimento, ella toma dos muestras de los alumnos de tercer grado, a una de las cuales se le imparte el nuevo método de enseñanza y a la otra el método usual. Cada grupo presenta el mismo examen final. Al hacer esta prueba, la psicóloga no se conforma con sólo informar que la media del grupo que recibió el nuevo método fue mayor que la del otro grupo. Ella quiere hacer una afirmación como: “La mejoría observada en las calificaciones del examen final fue resultado del nuevo método de enseñanza y no del azar. Además, dicha mejoría no sólo se aplicaría a la muestra particular sometida a prueba, sino que aparecería en toda la población de estudiantes de tercer grado si se les enseñase con el nuevo método”. Las técnicas utilizadas en estadística inferencial permiten hacer afirmaciones de este tipo.

En los experimentos para la estimación de parámetros, el investigador está interesado en determinar la magnitud de una característica de la población. Por ejemplo, a un economista podría importarle la determinación de la cantidad mensual promedio de dinero gastada el año pasado en alimentos por los estudiantes de bachillerato. Con los datos de una muestra y las técnicas de la estadística inferencial, él puede estimar la cantidad media que gasta la población. El economista podría concluir con una afirmación como la siguiente: “Existe una probabilidad de 0.95 de que el intervalo \$250-\$300 contenga la media de la población”.

Los temas de *muestreo aleatorio* y de *probabilidad* son fundamentales para la metodología de la estadística inferencial. En la siguiente sección estudiaremos el muestreo aleatorio. El resto del capítulo tratará sobre los principios básicos de la probabilidad.



MUESTREO ALEATORIO

Para generalizar de manera válida, a partir de una muestra a toda la población, tanto en los experimentos de prueba de hipótesis como en los de estimación de parámetros, la muestra no puede ser *cualquier subconjunto* de la población. Es crucial que sea una muestra *aleatoria*.

definición ■ Una **muestra aleatoria** es elegida de la población mediante un proceso con el cual se asegura que 1) cada posible muestra de un tamaño dado tenga la misma probabilidad de ser elegida y 2) todos los miembros de la población tengan la misma probabilidad de ser seleccionados en la muestra.*

* Véase la nota 8.1 de la p. 196.

Para ilustrar esto, consideremos una situación en la que tenemos una población con los datos 2, 3, 4, 5 y 6, y queremos extraer de manera aleatoria una muestra de tamaño 2 a partir de dicha población. Observe que, por lo general, una población tendría más datos. La hemos restringido a cinco elementos para facilitar la comprensión de lo que queremos poner de manifiesto. Suponga usted que estamos realizando un muestreo de esta población y que tomamos un dato a la vez, para luego volver a colocarlo con el resto de dicha población antes de extraer otro. Esto se llama *muestreo con reemplazo* y será analizado posteriormente en este capítulo. A continuación presentamos todas las muestras de tamaño 2 que podemos obtener de la población aplicando este método de muestreo:

2, 2	3, 2	4, 2	5, 2	6, 2
2, 3	3, 3	4, 3	5, 3	6, 3
2, 4	3, 4	4, 4	5, 4	6, 4
2, 5	3, 5	4, 5	5, 5	6, 5
2, 6	3, 6	4, 6	5, 6	6, 6

Son 25 muestras de tamaño 2 las que podemos obtener al realizar el muestreo de un elemento a la vez con reemplazo. Para realizar un muestreo aleatorio, el proceso debe ser tal que 1) las 25 muestras posibles y 2) todos los datos de la población (2, 3, 4, 5 y 6) tengan la misma probabilidad de ser seleccionados en la muestra.

La muestra debe ser aleatoria por dos razones. La primera es que para que se pueda generalizar a toda la población una característica observada en una muestra, es necesario aplicar las leyes de la probabilidad. Si la muestra no ha sido obtenida mediante un procedimiento con el cual se garantice que cada muestra posible del tamaño elegido tenga la misma posibilidad de ser escogida, no podremos aplicar las leyes de probabilidad a esa muestra. La importancia que tiene para la inferencia estadística este aspecto de aleatoriedad y de probabilidad se podrá apreciar claramente cuando hayamos analizado los capítulos referentes a la prueba de hipótesis y a las distribuciones muestrales (véanse los capítulos 10 y 12 respectivamente).

La segunda razón para la utilización del muestreo aleatorio consiste en que, para asegurar la generalización de una característica de la muestra a toda la población, es necesario que dicha muestra sea representativa de la población en cuestión. Una forma de lograr la representatividad consiste en elegir la muestra por medio de un proceso que garantice que todos los miembros de la población tengan la misma posibilidad de ser escogidos. Por lo tanto, el hecho de requerir que una muestra sea aleatoria permite aplicar las leyes de probabilidad sobre la muestra misma y, por ende, que ésta sea representativa de la población.

Es tentador pensar que podemos lograr la representatividad de una muestra utilizando métodos distintos al del muestreo aleatorio. Sin embargo, con mucha frecuencia el procedimiento utilizado produce una muestra sesgada (no representativa). Un ejemplo de esto es la famosa encuesta presidencial del *Literary Digest* de 1936, la cual predijo una victoria aplastante de Landon (57 a 43%). De hecho, Roosevelt ganó con 62% de los votos. La predicción del *Literary Digest* resultó ser un error garrafal. ¿Por qué? Un análisis posterior reveló que el error se debió a que la muestra no era representativa de la población votante, sino que era una muestra *sesgada*. Los individuos fueron elegidos de fuentes tales como el directorio telefónico, registros de clubes y listas de propietarios de automóviles. Esos registros excluyeron de manera sistemática a las personas pobres, quienes tal vez no tenían teléfonos ni automóviles. Y ocurrió que los pobres votaron en forma aplastante a favor de Roosevelt. Aunque otros métodos de muestreo pudieran producir una muestra representativa, éstos no serían útiles para la inferencia, pues no podríamos aplicar las leyes de la probabilidad necesarias para generalizar de la muestra a toda la población.

Técnicas para el muestreo aleatorio

Un estudio profundo de las formas con las cuales se generan las muestras aleatorias está más allá del objetivo de este libro. Este tema puede ser complejo, en particular al trabajar con encuestas. Sin embargo, presentaremos algunas técnicas de uso común, junto con varias situaciones sencillas, de manera que usted pueda darse una idea de lo que esto implica. Suponga que tenemos una población de 100 personas y queremos obtener una muestra aleatoria de 20 personas para un experimento. Una forma de hacer esto sería enumerar a los individuos de la población del 1 al 100, luego tomar 100 pedazos de papel y numerarlos, colocar los papeles en un sombrero para revolverlos y después elegir uno. Volveríamos entonces a mezclar los papelitos para elegir otro. Luego, continuaríamos este proceso hasta haber elegido 20 pedazos de papel. Los números escritos en los pedazos de papel identificarían a los individuos que compondrían la muestra. En este método de muestreo aleatorio, es crucial que la población (los trozos de papel) esté bien “revuelta” para garantizar la aleatoriedad.

Un procedimiento común para producir muestras aleatorias consiste en utilizar una tabla de números aleatorios, como la que hemos incluido en la tabla J del apéndice D. Este tipo de tablas se hace, por lo regular, mediante un programa de computadora, el cual garantiza que todos los dígitos (0-9) tengan la misma posibilidad de ser elegidos cada vez que se imprima un dígito.

La tabla se puede emplear como una sucesión de dígitos simples consecutivos, números de dos dígitos, números de tres dígitos, etc. Por ejemplo, en la tabla J, página 546, si comenzamos en el primer renglón y nos movemos en dirección horizontal por la página, el orden aleatorio de los dígitos simples sería 3, 2, 9, 4, 2, Si quisiéramos utilizar números de dos dígitos, entonces el orden aleatorio sería 32, 94, 29, 54, 16,

Como los dígitos de la tabla son aleatorios, se pueden utilizar vertical u horizontalmente en ambas direcciones. La dirección que se va a usar debe especificarse antes de emplear la tabla; para manejarla de manera adecuada, hay que comenzar en forma aleatoria. Una posibilidad de lograr esto es hacer tarjetas con números de renglón y de columna, colocarlas en una caja, revolverlas y luego elegir un número para el renglón y otro para la columna. La intersección del renglón y la columna sería la posición del primer número aleatorio. Los demás números serían localizados al desplazarse desde el primer número, en la dirección convenida antes de entrar a la tabla. Por ejemplo, suponga que queremos formar una muestra aleatoria de 3 sujetos de una población de 10 sujetos.* Para este ejemplo hemos decidido movernos sobre la tabla en dirección horizontal, a la derecha. Para elegir la muestra, primero asignamos a cada individuo de la población un número del 0 al 9. Luego entramos en la tabla de manera aleatoria para localizar el primer número. Suponga que el dato es el primer número del séptimo renglón, que es igual a 3. Este número designa al primer sujeto de la muestra. Como hemos decidido movernos sobre la tabla hacia la derecha, los dos números siguientes son 5 y 6. De esta manera, los individuos que portan los números 5 y 6 completarían la muestra.

Resolvamos ahora un problema en el cual la población tiene más individuos. Con propósitos ilustrativos, supondremos que se desea extraer una muestra aleatoria de 15 sujetos, de una población de 100. Para variar las cosas un poco, para este ejercicio decidimos desplazarnos verticalmente sobre la tabla, hacia abajo, en lugar de hacerlo en dirección horizontal, hacia la derecha. Como antes, necesitamos asignar una cifra a cada miembro de la población. Esta vez, los números asignados serán del 00 al 99 en vez de 0

*Por supuesto, en los experimentos reales, el número de elementos de la población es mucho mayor que 10. Utilizamos 10 en el primer ejemplo para ayudar a comprender la forma correcta de emplear la tabla.

a 9. De nuevo, entramos a la tabla de manera aleatoria. Esta vez suponemos que el primer dato o primer número de dos dígitos aparece en la intersección de la columna 3 con el renglón 12. El guarismo de dos dígitos localizado en esa intersección es 70. Por lo tanto, el primer sujeto de la muestra es el individuo a quien le corresponde el número 70. Para localizar al siguiente sujeto de la muestra, bajamos verticalmente en la tabla a partir del 70. Así, el segundo sujeto de la muestra sería el individuo que tiene el número 33. Continuaríamos este procedimiento hasta que hayamos seleccionado 15 sujetos. El conjunto completo de números que representan a los sujetos de la muestra sería: 70, 33, 82, 22, 96, 35, 14, 12, 13, 59, 97, 37, 54, 42 y 89. Al obtener este conjunto de números, el 82 apareció dos veces en la tabla. Como el mismo individuo no puede estar en la muestra más de una vez, no incluimos el número repetido.

Muestreo con o sin reemplazo

Hasta ahora hemos definido una muestra aleatoria, analizamos la importancia del muestreo aleatorio y presentamos algunas técnicas para producir muestras aleatorias. Para completar nuestro estudio, debemos distinguir entre el “muestreo con reemplazo” y el “muestreo sin reemplazo”. Para ilustrar la diferencia entre estos dos métodos de muestreo, suponga que queremos formar una muestra de dos datos provenientes de una población compuesta por los números 4, 5, 8 y 10. Una manera de hacer esto consistiría en extraer un dato de la población, registrar su valor, y reintegrarlo a la población antes de extraer el segundo elemento. Así, el primer dato podría ser elegido de nuevo en la segunda extracción. Este método se llama *muestreo con reemplazo*. Una segunda alternativa consistiría en extraer al azar un dato de la población y no reemplazarlo antes de seleccionar el segundo. De esta manera, cualquier número proveniente de la población podría aparecer en la muestra solamente una vez. A este método se le llama *muestreo sin reemplazo*.

- definiciones**
- **El muestreo con reemplazo** se define como un método en el cual cada miembro de la población elegida para la muestra se reintegra a la población antes de seleccionar al siguiente miembro.
 - **El muestreo sin reemplazo** se define como un método en el cual los miembros de la muestra no son reintegrados a la población antes de seleccionar a los miembros subsecuentes.

Al elegir a los sujetos que participan en un experimento, hay que utilizar el muestreo sin reemplazo, pues el mismo individuo no puede estar en la muestra más de una vez. Es probable que usted reconozca éste como el método utilizado en la sección anterior. El muestreo con reemplazo constituye la base matemática para muchas de las pruebas de inferencia que analizaremos posteriormente. A pesar de que los dos métodos no producen resultados idénticos, cuando el tamaño de la muestra es pequeño en comparación con el tamaño de la población, las diferencias son insignificantes y las técnicas “con reemplazo” resultan mucho más fáciles de utilizar cuando se desea proporcionar la base matemática para la inferencia. Pasemos ahora al tema de la probabilidad.

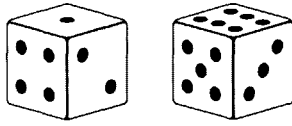


figura 8.1 Un par de dados.



PROBABILIDAD

La probabilidad se puede estudiar desde dos puntos de vista: 1) *a priori* o clásico y 2) *a posteriori* o empírico. *A priori* significa aquello que se puede deducir usando sólo la razón, sin recurrir a la experiencia. Desde el punto de vista *a priori* o clásico, la probabilidad se define como:

$$p(A) = \frac{\text{Número de eventos clasificables como } A}{\text{Número total de eventos posibles}} \quad \textit{probabilidad a priori}$$

El símbolo $p(A)$ se lee como “la probabilidad de ocurrencia del evento A ”. Así, la ecuación establece que la probabilidad de ocurrencia del evento A es igual al número de eventos clasificables como A dividida entre el número de eventos posibles. Para ilustrar el uso de esta ecuación, veamos un ejemplo relacionado con unos dados. La figura 8.1 muestra una ilustración con un par de dados. Cada dado tiene seis caras y ostenta un número distinto de puntos en cada cara. Los puntos varían de uno a seis. Estos cubos de apariencia inocente se utilizan para *juegos de azar*; son la base de muchas desgracias y alegrías, según la “suerte” del jugador.

Regresando a la probabilidad *a priori*, suponga que vamos a arrojar un dado una vez. ¿Cuál es la probabilidad de que el dado se detenga con un 2 (el lado que tiene dos puntos) hacia arriba? Como existen seis números posibles y sólo uno de ellos es un 2, la probabilidad de que salga un 2 en una tirada de un dado es:

$$p(A) = p(2) = \frac{\text{Número de eventos clasificables como } 2}{\text{Número total de eventos posibles}} = \frac{1}{6} = 0.1667$$

Tratemos de resolver otro problema mediante el punto de vista *a priori*. ¿Cuál es la probabilidad de obtener un número mayor que 4 en una tirada de un dado? Ahora existen dos eventos clasificables como A (obtener un 5 o un 6). Así,

$$p(A) = p(5 \text{ o } 6) = \frac{\text{Número de eventos clasificables como } 5 \text{ o } 6}{\text{Número total de eventos posibles}} = \frac{2}{6} = 0.3333$$

Observe que logramos resolver los dos problemas anteriores por medio de un razonamiento sin recurrir a ninguna colección de datos. Debemos comparar este punto de vista con el enfoque *a posteriori* o empírico de la probabilidad. *A posteriori* significa “después del hecho”; en el contexto de la probabilidad, significa después de haber reunido algunos datos. Desde el punto de vista *a posteriori* o empírico, la probabilidad se define como:

$$p(A) = \frac{\text{Número de veces que } A \text{ ha ocurrido}}{\text{Número total de ocurrencias}} \quad \textit{probabilidad a posteriori}$$

Para determinar la probabilidad de obtener un 2 al tirar un dado, desde el punto de vista empírico, tendríamos que considerar el dado real, arrojarlo en muchas ocasiones y

*En éste y en los demás problemas relativos a los dados, supondremos que éstos no pueden detenerse sobre ninguna de sus aristas.

contar el número de veces que se presenta un 2. Mientras más veces arrojemos el dado, tanto mejor será. Para este problema, suponga que lanzamos el dado 100 000 veces y que el 2 apareció en 16 000 ocasiones. La probabilidad de que ocurra un 2 en un lanzamiento del dado se determina mediante:

$$p(2) = \frac{\text{Número de veces que ha ocurrido el 2}}{\text{Número total de ocurrencias}} = \frac{16\,000}{100\,000} = 0.1600$$

Observe que, desde este punto de vista, antes de determinar la probabilidad es necesario tener el dado real y reunir algunos datos. El punto interesante es que si el dado no está cargado (entonces se dice que es un dado honesto), al lanzarlo muchas veces la probabilidad *a posteriori* se acerca a la probabilidad *a priori*. Si arrojamos el dado una infinidad de veces, las dos probabilidades serán iguales. Observe también que si el dado está cargado (con una distribución de peso por la cual uno de los lados quede arriba con mayor frecuencia que los demás), la probabilidad *a posteriori* será diferente que la probabilidad *a priori*. Por ejemplo, si el dado está “cargado” para que aparezca un 6, el 2 nunca aparecerá. Podemos ver ahora que la ecuación *a priori* supone que todos los resultados posibles tienen la misma probabilidad de ocurrir. Para la mayor parte de los problemas de este capítulo y el siguiente, aplicaremos el punto de vista *a priori*.

Algunos aspectos básicos relativos a los valores de probabilidad

Puesto que la probabilidad es esencialmente una proporción, su valor fluctúa entre 0.00 y 1.00. Si la probabilidad de que ocurra un evento es igual a 1.00, entonces es seguro que tal evento ocurrirá. Si la probabilidad es igual a 0.00, entonces es seguro que el evento *no* ocurrirá. Por ejemplo, un dado común no tiene 7 puntos en ninguno de sus lados. Por lo tanto, la probabilidad de obtener un 7 al arrojar un dado es igual a 0.00. Es seguro que no ocurrirá que salga un 7. Por otra parte, la probabilidad de que salga un número del 1 al 6 es igual a 1.00. Es seguro que se obtendrá alguno de los números 1, 2, 3, 4, 5 o 6.

La probabilidad de ocurrencia de un evento se expresa como una fracción o un número decimal. Por ejemplo, la probabilidad de elegir al azar el as de espadas en una baraja común es $\frac{1}{52}$ o bien 0.0192.* La respuesta se puede quedar como una fracción ($\frac{1}{52}$), pero generalmente se convierte a su equivalente decimal (0.0192).

Algunas veces, la probabilidad se expresa como “posibilidades con respecto a 100”. Por ejemplo, alguien podría comentar que la probabilidad de que ocurra el evento *A* es de 5 posibilidades en 100. Lo que esta persona quiere decir es que $p(A) = 0.05$. En ciertas ocasiones, la probabilidad se expresa también como las posibilidades a favor o en contra de que ocurra un evento. Por ejemplo, un apostador podría decir que las posibilidades de que Fred gane la carrera son de 3 a 1 a su favor. En términos de probabilidades, $p(\text{que gane Fred}) = \frac{3}{4} = 0.75$. Si las posibilidades fuesen de 3 a 1 en contra de que Fred triunfe, entonces $p(\text{que gane Fred}) = \frac{1}{4} = 0.25$.

Cálculo de probabilidades

La determinación de la probabilidad de los eventos puede ser compleja. De hecho, existen cursos completos dedicados a este tema y son bastante difíciles. Por fortuna, para nuestros propósitos sólo hay que aprender dos reglas fundamentales de la probabilidad: la de la suma y la del producto. Estas reglas proporcionarán las bases necesarias para comprender las pruebas de inferencia estadística que veremos más adelante en este libro de texto.

*Para los no iniciados, un mazo de barajas francesas ordinarias se compone de 52 cartas: 4 palos (espadas, corazones, diamantes y tréboles), y 13 cartas en cada palo (as, 2, 3, 4, 5, 6, 7, 8, 9, 10, jack, reina y rey).

La regla de la suma

La regla de la suma sirve para determinar la probabilidad de que ocurra alguno de varios eventos posibles. Para comenzar nuestro análisis, supongamos que sólo existen dos eventos posibles: A y B . Cuando existen dos eventos, la regla de la suma establece lo siguiente:

definición ■ *La probabilidad de ocurrencia de A o B es igual a la probabilidad de que ocurra A más la probabilidad de que ocurra B menos la probabilidad de que ocurran ambos, A y B .*

En forma de ecuación, la regla de la suma establece que:

$$p(A \text{ o } B) = p(A) + p(B) - p(A \text{ y } B) \quad \text{\textit{ecuación general de la regla de la suma para dos eventos}}$$

Vamos a ilustrar la forma en que se utiliza esta regla. Suponga que queremos determinar la probabilidad de sacar un as o un trébol de una baraja de cartas ordinarias. El problema ha sido resuelto de dos maneras en la figura 8.2. Por favor, consulte la figura al leer este párrafo. En la primera de esas maneras, enumeramos todos los eventos que pueden ser clasificados como un as o un trébol y utilizamos la ecuación básica para la probabilidad. Existen 16 formas de obtener un as o un trébol, por lo cual la probabilidad de sacar un as o un trébol es $= \frac{16}{52} = 0.3077$. El segundo método emplea la regla de la suma. La probabilidad de obtener un as es $= \frac{4}{52}$ y la probabilidad de sacar un trébol a la vez es $= \frac{13}{52}$. La probabilidad de lograr un as y un trébol es igual a $= \frac{1}{52}$. Según la regla de la suma, la probabilidad de conseguir un as o un trébol es $= \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52} = 0.3077$. ¿Por qué tenemos que restar la probabilidad de obtener un as y un trébol a la vez? Porque hemos contado el as de tréboles dos veces. Si no lo restáramos, pensaríamos de manera errónea que existen 17 eventos favorables siendo que sólo existen 16.

En este curso emplearemos la regla de la suma en casi todas las situaciones en las que los eventos sean *mutuamente excluyentes*.

definición ■ *Dos eventos son mutuamente excluyentes si no pueden ocurrir al mismo tiempo. Otra forma de decir esto es que **dos eventos son mutuamente excluyentes** si la ocurrencia de uno impide la ocurrencia del otro.*

Los eventos de obtener un 1 y un 2 al lanzar un dado son mutuamente excluyentes. Si el lanzamiento termina con un 1, entonces no puede haber terminado también con un 2. Los eventos de extraer a la vez un as y un rey al sacar una sola carta de un mazo de barajas ordinarias son mutuamente excluyentes. Si la carta es un as, eso impide que la misma sea también un rey. Esto se puede contrastar con los eventos de extraer a la vez un as y un trébol al sacar una sola carta del mazo. Estos dos eventos no son mutuamente excluyentes en virtud de que hay una carta que es simultáneamente un as y un trébol (el as de tréboles).

Cuando los eventos son mutuamente excluyentes, la probabilidad de que ambos eventos ocurran al mismo tiempo es igual a cero. Así, $p(A \text{ y } B) = 0$ cuando A y B son mutuamente excluyentes. En estas condiciones, la regla de la suma se simplifica así:

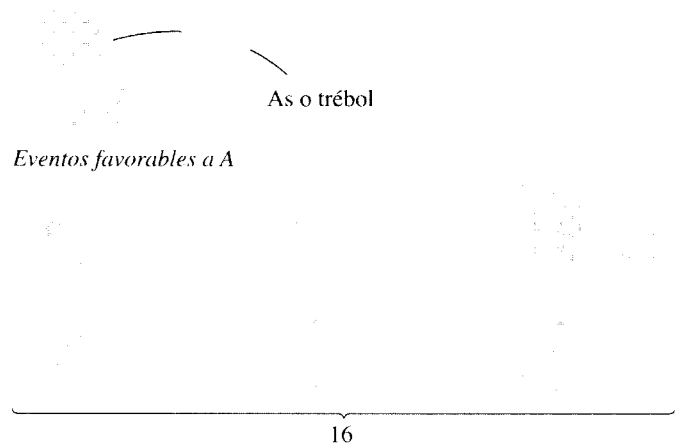
$$p(A \text{ o } B) = p(A) + p(B) \quad \text{regla de la suma cuando } A \text{ y } B \text{ son mutuamente excluyentes}$$

Practiquemos ahora resolviendo algunos problemas que incluyen situaciones en las que A y B son mutuamente excluyentes.

(a) Por enumeración, utilizando la definición básica de probabilidad

$$p(A) = \frac{\text{Número de eventos favorables a } A}{\text{Número total de eventos posibles}} = \frac{16}{52} = 0.3077$$

donde A = extracción de un as o un trébol



(b) Por la regla de la suma

$$p(A \text{ o } B) = p(A) + p(B) - p(A \text{ y } B) = \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52} = 0.3077$$

donde A = extracción de un as
 B = extracción de un trébol

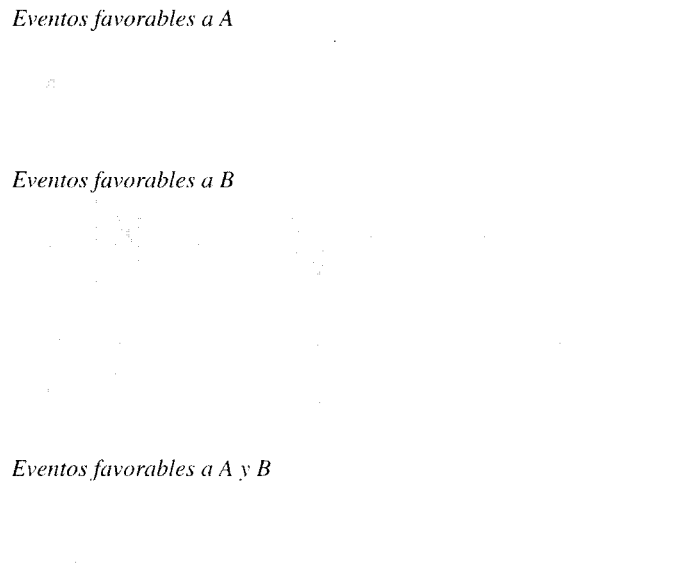


figura 8.2 Determinación de la probabilidad de elegir al azar un as o un trébol al extraer una carta de un mazo de barajas ordinarias.

¿Cuál es la probabilidad de elegir al azar un 10 o un 4 al extraer una carta de una baraja ordinaria?

SOLUCIÓN

La solución se ilustra en la figura siguiente. Como deseamos sacar un 10 o un 4 y estos dos eventos son mutuamente excluyentes, la regla de la suma con eventos mutuamente excluyentes resulta apropiada. Así, $p(10 \text{ o } 4) = p(10) + p(4)$.

Puesto que hay cuatro 10, cuatro 4 y 52 cartas, entonces $p(10) = \frac{4}{52}$ y $p(4) = \frac{4}{52}$. Por consiguiente, $p(10 \text{ o } 4) = \frac{4}{52} + \frac{4}{52} = \frac{8}{52} = 0.1538$.

$$\begin{aligned} p(A \text{ o } B) &= p(A) + p(B) \\ p(\text{un } 10 \text{ o un } 4) &= p(10) + p(4) \\ &= \frac{4}{52} + \frac{4}{52} \\ &= \frac{8}{52} = 0.1538 \end{aligned}$$

donde $A =$ extracción de un 10
 $B =$ extracción de un 4

10 o 4

Eventos favorables a A

Eventos favorables a B

Al lanzar una vez un dado que no está cargado, ¿cuál es la probabilidad de obtener un 1 o un número par?

SOLUCIÓN

La solución aparece en la figura acompañante. Como los eventos son mutuamente excluyentes y el problema pide obtener un 1 o un número par, podemos aplicar la regla de la suma para este tipo de eventos. Así, $p(1 \text{ o un número par}) = p(1) + p(\text{un número par})$. Existe una forma de conseguir un 1, tres formas de lograr un número par (2, 4, 6) y seis resultados posibles. Así, $p(1) = \frac{1}{6}$, $p(\text{un número par}) = \frac{3}{6}$ y $p(1 \text{ o un número par}) = \frac{1}{6} + \frac{3}{6} = \frac{4}{6} = 0.6667$

$$\begin{aligned}
 p(A \text{ o } B) &= p(A) + p(B) \\
 p(1 \text{ o un número par}) &= p(1) + p(\text{un número par}) \\
 &= \frac{1}{6} + \frac{3}{6} = \frac{4}{6} \\
 &= 0.6667
 \end{aligned}$$

Eventos favorables a A

donde A = representa obtener un 1
 B = representa obtener un número par

Eventos favorables a B

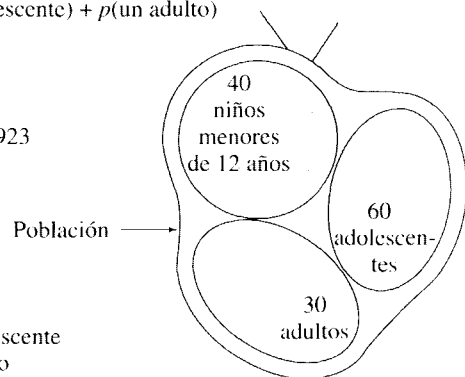
Suponga que va a elegir de manera aleatoria un individuo entre una población de 130 personas. En esa población hay 40 niños menores de 12 años, 60 adolescentes y 30 adultos. ¿Cuál es la probabilidad de que el individuo elegido sea un adolescente o un adulto?

SOLUCIÓN

La solución aparece en la figura acompañante. Como los eventos son mutuamente excluyentes y queremos obtener un adolescente o un adulto, podemos emplear la regla de la suma correspondiente. Así, $p(\text{adolescente o adulto}) = p(\text{adolescente}) + p(\text{adulto})$. Como hay 60 adolescentes, 30 adultos y 130 personas en la población, $p(\text{adolescente}) = \frac{60}{130}$ y $p(\text{adulto}) = \frac{30}{130}$. Por lo tanto, $p(\text{adolescente o adulto}) = \frac{60}{130} + \frac{30}{130} = \frac{90}{130} = 0.6923$.

$$\begin{aligned}
 p(A \text{ o } B) &= p(A) + p(B) \\
 p(\text{un adolescente o un adulto}) &= p(\text{un adolescente}) + p(\text{un adulto}) \\
 &= \frac{60}{130} + \frac{30}{130} \\
 &= \frac{90}{130} = 0.6923
 \end{aligned}$$

Adolescente o Adulto



donde A = representa obtener un adolescente
 B = representa obtener un adulto

La regla de la suma también se puede utilizar en situaciones donde hay más de dos eventos; sólo hacemos una sencilla extensión de la ecuación empleada para esta situación. Así, cuando existen más de dos eventos y todos son mutuamente excluyentes, la probabilidad de que ocurra cualquiera de ellos es igual a la suma de las probabilidades de cada evento. A continuación mostramos la ecuación:

$$p(A \text{ o } B \text{ o } C \text{ o } \dots \text{ o } Z) = p(A) + p(B) + p(C) + \dots + p(Z)$$

regla de la suma para más de dos eventos mutuamente excluyentes

donde Z = es el último evento

Con mucha frecuencia encontraremos situaciones en las cuales los eventos no sólo son mutuamente excluyentes, sino también exhaustivos. Ya hemos definido a los dos primeros, nos faltan los *exhaustivos*.

definición ■ *Un conjunto de eventos es **exhaustivo** si incluye a todos los eventos posibles.*

Por ejemplo, al lanzar un dado una sola vez, el conjunto de eventos que consisten en obtener 1, 2, 3, 4, 5 o 6 es exhaustivo, pues el conjunto incluye a todos los eventos posibles. Cuando un conjunto de eventos es exhaustivo y mutuamente excluyente, existe una relación muy útil. En estas condiciones, la suma de las probabilidades individuales de cada evento del conjunto debe ser igual a 1. Así,

$$p(A) + p(B) + p(C) + \dots + p(Z) = 1.00 \quad \text{cuando los eventos son exhaustivos y mutuamente excluyentes}$$

donde A, B, C, \dots, Z = son los eventos

Para ilustrar esta relación, consideremos el conjunto de eventos que consisten en obtener 1, 2, 3, 4, 5 o 6 al lanzar una sola vez un dado no cargado. Como los eventos son exhaustivos y mutuamente excluyentes, la suma de sus probabilidades debe ser igual a 1. Podemos ver que esto es cierto, pues $p(1) = \frac{1}{6}, p(2) = \frac{1}{6}, p(3) = \frac{1}{6}, p(4) = \frac{1}{6}, p(5) = \frac{1}{6}$ y $p(6) = \frac{1}{6}$. Así,

$$\begin{aligned} p(1) + p(2) + p(3) + p(4) + p(5) + p(6) &= 1.00 \\ \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} &= 1.00 \end{aligned}$$

Cuando sólo existen dos eventos y éstos son mutuamente excluyentes, se acostumbra asignar el símbolo P a la probabilidad de que ocurra uno de los eventos y Q a la probabilidad de que ocurra el otro evento. Por ejemplo, si estamos lanzando una moneda, ésta sólo puede caer de cara o de cruz y se presentará una situación en la que sólo hay dos eventos posibles (cara o cruz) en cada lanzamiento, y éstos son mutuamente excluyentes (si es cara, no puede ser cruz y viceversa). Se acostumbra asignar P a la probabilidad de ocurrencia de uno de los eventos, digamos cara, y Q a la probabilidad de ocurrencia del otro evento, cruz. En este caso, si la moneda no está cargada, entonces, $P = \frac{1}{2}$ y $Q = \frac{1}{2}$. Como los eventos son exhaustivos y mutuamente excluyentes, sus probabilidades deben ser iguales a 1. Por consiguiente,

$$P + Q = 1.00 \quad \text{cuando dos eventos son exhaustivos y mutuamente excluyentes}$$

En el capítulo 9 utilizaremos ampliamente los símbolos P y Q en relación con la distribución binomial.

La regla del producto

Así como la regla de la suma proporciona la probabilidad de que ocurra cualquiera de varios eventos, la regla del producto analiza la ocurrencia conjunta o sucesiva de varios eventos. Observe que la regla del producto analiza con frecuencia lo que ocurre en más de un lanzamiento o extracción, mientras que la regla de la suma estudia sólo uno de dichos lanzamientos o extracciones. Si nos interesa la ocurrencia conjunta o sucesiva de dos eventos A y B , la regla del producto establece lo siguiente:

definición ■ *La probabilidad de ocurrencia de A y B es igual a la probabilidad de ocurrencia de A por la probabilidad de ocurrencia de B , dado que A ha ocurrido.*

En forma de ecuación, la regla del producto es

$$p(A \text{ y } B) = p(A)p(B|A) \quad \textit{ecuación general de la regla del producto con dos eventos}$$

Observe que el símbolo $p(B|A)$ se lee como “la probabilidad de que ocurra B dado que A ha ocurrido”. Esto no significa B dividido entre A . Observe también que la regla del producto se refiere a la ocurrencia de A y B , mientras que la regla de la suma se aplica a la ocurrencia de A o B .

Al analizar la regla del producto, es útil distinguir tres condiciones: cuando los eventos son mutuamente excluyentes, cuando los eventos son independientes y cuando los eventos son dependientes.

Regla del producto: eventos mutuamente excluyentes Ya hemos analizado la ocurrencia conjunta de A y B cuando A y B son mutuamente excluyentes. Recordará que si A y B son mutuamente excluyentes,

$$p(A \text{ y } B) = 0 \quad \textit{regla del producto para eventos mutuamente excluyentes}$$

ya que cuando los eventos son mutuamente excluyentes, la ocurrencia de uno impide la ocurrencia del otro. La probabilidad de su ocurrencia conjunta es cero.

Regla del producto: eventos independientes Para comprender la forma apropiada de aplicar la regla del producto en esta situación, tenemos que definir primero qué son los eventos *independientes*.

definición ■ *Dos eventos son independientes si la ocurrencia de un evento no tiene efecto alguno sobre la probabilidad de ocurrencia del otro.*

El muestreo con reemplazo ilustra muy bien esta situación. Por ejemplo, suponga que vamos a extraer con reemplazo dos cartas, una a la vez, de una baraja ordinaria. Llamaremos A la carta extraída primero y B a la carta obtenida en segundo lugar. En vista de

que A es reintegrada al mazo de cartas antes de extraer B , la ocurrencia de A en la primera extracción no tiene efecto alguno sobre la probabilidad de ocurrencia de B . Por ejemplo, si A fue un as, por el hecho de que fue reintegrado a la baraja antes de extraer la segunda carta, la ocurrencia de un as en la primera extracción no tiene efecto alguno sobre la probabilidad de la ocurrencia de la carta elegida en la segunda extracción. Si A y B son independientes, entonces la probabilidad de B no resulta afectada por A . Por lo tanto, $p(B|A) = p(B)$. En este tipo de situaciones, la regla del producto se convierte en

$$p(A \text{ y } B) = p(A)p(B|A) = p(A)p(B) \quad \text{\textit{regla del producto para eventos independientes}}$$

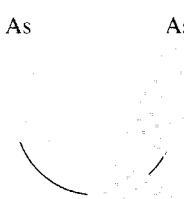
Veamos cómo se emplea esta ecuación. Suponga que vamos a extraer al azar dos cartas con reemplazo, una a la vez, de una baraja ordinaria. ¿Cuál es la probabilidad de que ambas cartas sean ases?

La solución se ilustra en la figura 8.3. Como quiera que el problema requiere la presencia de un as en la primera y la segunda extracciones, resulta apropiado aplicar la regla del producto. Supongamos que A es un as en la primera extracción y B es un as en la segunda. Como el muestreo es con reemplazo, A y B son independientes. Así, $p(\text{un as en la primera extracción y un as en la segunda extracción}) = p(\text{un as en la primera extracción})p(\text{un as en la segunda extracción})$. Puesto que existen cuatro ases posibles en la primera extracción, cuatro ases posibles en la segunda extracción (el muestreo se realiza *con* reemplazo) y hay 52 cartas en la baraja, entonces $p(\text{un as en la primera extracción}) = \frac{4}{52}$ y $p(\text{un as en la segunda extracción}) = \frac{4}{52}$. Por consiguiente, $p(\text{un as en la primera extracción y un as en la segunda extracción}) = \frac{4}{52} \left(\frac{4}{52} \right) = \frac{16}{2704} = 0.0059$.

Veamos algunos problemas en plan de práctica.

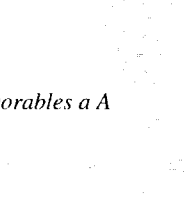
$$\begin{aligned}
 p(A \text{ y } B) &= p(A)p(B) \\
 \left[\begin{array}{l} p(\text{un as en la primera extracción}) \\ \text{y un as en la segunda extracción} \end{array} \right] &= \frac{p(\text{un as en la primera extracción})}{p(\text{un as en la segunda extracción})} \\
 &= \left(\frac{4}{52} \right) \left(\frac{4}{52} \right) \\
 &= \frac{16}{2704} = 0.0059
 \end{aligned}$$

As



Eventos favorables a A

As



Eventos favorables a B

donde A = representa obtener un as en la primera extracción
 B = representa obtener un as en la segunda extracción

figura 8.3 Determinación de la probabilidad de obtener dos ases en dos extracciones de un mazo de barajas ordinarias. El muestreo es de una carta a la vez, con reemplazo: regla del producto para eventos independientes.

Suponga que lanzamos una sola vez un par de dados no cargados. ¿Cuál es la probabilidad de obtener un 2 en el primer dado y un 4 en el segundo?

SOLUCIÓN

La solución aparece en la figura acompañante. Como quiera que los dados son independientes entre sí y en vista de que el problema requiere un 2 y un 4, podemos aplicar la regla del producto para eventos independientes. Así, $p(\text{un 2 en el primer dado y un 4 en el segundo}) = p(\text{un 2 en el dado 1})p(\text{un 4 en el dado 2})$. Existe una forma de obtener un 2 en el primer dado, una de conseguir un 4 en el segundo dado y seis resultados posibles para cada dado. Por lo tanto, $p(\text{un 2 en el primer dado}) = \frac{1}{6}$, $p(\text{un 4 en el dado 2}) = \frac{1}{6}$, y $p(\text{un 2 en el dado 1 y un 4 en el dado 2}) = \frac{1}{6}(\frac{1}{6}) = \frac{1}{36} = 0.0278$.

$$p(2 \text{ en el dado 1 y } 4 \text{ en dado 2}) = p(2 \text{ en el dado 1})p(4 \text{ en el dado 2}) \\ = \left(\frac{1}{6}\right)\left(\frac{1}{6}\right) = \frac{1}{36} = 0.0278$$

Eventos favorables a A
Dado 1

donde $A =$ representa obtener un 2
en el primer dado
 $B =$ representa obtener un 4
en el segundo dado

Eventos favorables a B
Dado 2

Si se lanzan dos monedas una sola vez, ¿cuál es la probabilidad de que ambas caigan con la cara hacia arriba? Suponga que las monedas no están cargadas y que una cara o una cruz es el único resultado posible para cada moneda.

SOLUCIÓN

La solución aparece en la figura acompañante. Como el resultado para la primera moneda no tiene efecto alguno sobre el resultado de la segunda, los eventos son independientes. En vista de que en el problema se requiere una cara con la primera moneda y una cara con la segunda moneda, podemos aplicar la regla del producto con eventos independientes. Así, $p(\text{una cara en la primera moneda y una cara en la$

(continúa)

segunda moneda) = $p(\text{una cara en la primera moneda})p(\text{una cara en la segunda moneda})$. Como sólo existe una forma de obtener una cara con cada moneda y hay dos posibilidades en cada moneda (cara o cruz), $p(\text{una cara en la primera moneda}) = \frac{1}{2}$ y $p(\text{una cara en la segunda moneda}) = \frac{1}{2}$. Por lo tanto, $p(\text{una cara en la primera moneda y una cara en la segunda moneda}) = \frac{1}{2}\left(\frac{1}{2}\right) = \frac{1}{4} = 0.2500$.

$$p(A \text{ y } B) = p(A)p(B)$$

$$\begin{aligned} \left[\begin{array}{l} p(\text{una cara en la primera moneda}) \\ \text{y una cara en la segunda moneda} \end{array} \right] &= \frac{p(\text{una cara en la primera moneda})}{p(\text{una cara en la segunda moneda})} \\ &= \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = 0.2500 \end{aligned}$$

donde A = representa obtener una cara en la primera moneda

B = representa obtener una cara en la segunda moneda

Eventos favorables a A
Primera moneda



Eventos favorables a B
Segunda moneda



Suponga que usted extrae una muestra aleatoria de una bolsa de frutas. La bolsa contiene cuatro manzanas, seis naranjas y cinco duraznos. Si selecciona dos frutas, una a la vez, con reemplazo, ¿cuál es la probabilidad de que obtenga una naranja y una manzana, en ese orden?

SOLUCIÓN

La solución aparece en la figura acompañante. Como las extracciones son independientes entre sí (el muestreo es con reemplazo) y como queremos obtener una naranja y una manzana, aplicamos la regla del producto con eventos independientes. Así, $p(\text{una naranja en la primera extracción y una manzana en la segunda extracción}) = p(\text{una naranja en la primera extracción})p(\text{una manzana en la segunda extracción})$. Como existen 6 naranjas y 15 frutas en la bolsa, $p(\text{una naranja en la primera extracción}) = \frac{6}{15}$. Puesto que la fruta obtenida en la primera selección se reemplaza antes de extraer la segunda, no tiene efecto alguno sobre la fruta elegi-

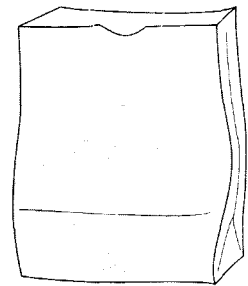
da en la segunda extracción. Como hay 4 manzanas y 15 frutas, $p(\text{una manzana en la segunda extracción}) = \frac{4}{15}$. Por lo tanto, $p(\text{una naranja en la primera extracción y una manzana en la segunda extracción}) = \frac{6}{15}(\frac{4}{15}) = 0.1067$.

$$p(A \text{ y } B) = p(A)p(B)$$

$$\begin{aligned} \left[p(\text{una naranja en la primera extracción y una manzana en la segunda extracción}) \right] &= \frac{p(\text{una naranja en la primera extracción})}{p(\text{una manzana en la segunda extracción})} \\ &= \left(\frac{6}{15} \right) \left(\frac{4}{15} \right) \\ &= \frac{24}{225} = 0.1067 \end{aligned}$$

donde A = representa una naranja en la primera extracción
 B = representa una manzana en la segunda extracción

Naranja Manzana



Eventos favorables a A
Naranjas

Eventos favorables a B
Manzanas

Suponga que usted extrae al azar 2 individuos de una población de 110 hombres y mujeres. Hay 50 varones y 60 mujeres en esa población. El muestreo es de un individuo a la vez, con reemplazo. ¿Cuál es la probabilidad de que la muestra contenga sólo mujeres?

SOLUCIÓN

La solución aparece en la figura acompañante. Como el problema requiere una mujer en la primera extracción y una mujer en la segunda extracción, y como estos dos eventos son independientes (el muestreo es con reemplazo), aplicamos la regla del producto con eventos independientes. Así, $p(\text{una mujer en la primera}$

(continúa)

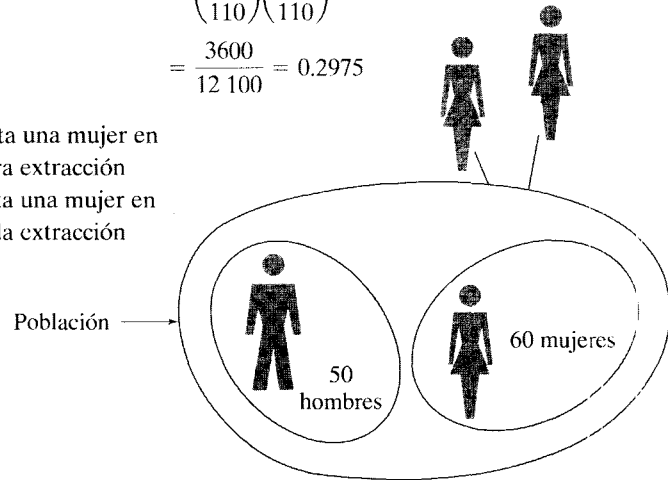
extracción y una mujer en la segunda extracción) = $p(\text{una mujer en la primera extracción})p(\text{una mujer en la segunda extracción})$. Como hay 60 mujeres y 110 personas en la población, $p(\text{una mujer en la primera extracción}) = \frac{60}{110}$, y $p(\text{una mujer en la segunda extracción}) = \frac{60}{110}$. Por lo tanto, $p(\text{una mujer en la primera extracción y una mujer en la segunda extracción}) = \frac{60}{110} \left(\frac{60}{110}\right) = \frac{3600}{12100} = 0.2975$.

$$\left[\begin{array}{l} p(\text{una mujer en la primera extracción}) \\ \text{y una mujer en la segunda extracción} \end{array} \right] = p(\text{una mujer en la primera extracción}) p(\text{una mujer en la segunda extracción})$$

$$= \left(\frac{60}{110}\right)\left(\frac{60}{110}\right)$$

$$= \frac{3600}{12100} = 0.2975$$

donde A = representa una mujer en la primera extracción
 B = representa una mujer en la segunda extracción



La regla del producto para eventos independientes también se aplica en situaciones con más de dos eventos. En tales casos, la probabilidad de la ocurrencia conjunta de los eventos es igual al producto de las probabilidades individuales de cada evento. En forma de ecuación:

$$p(A \text{ y } B \text{ y } C \text{ y } \dots \text{ y } Z) = p(A)p(B)p(C) \cdot \dots \cdot p(Z)$$

regla del producto para más de dos eventos independientes

Para mostrar el uso de esta ecuación, suponga que en vez de obtener una muestra de 2 individuos pertenecientes a la población descrita en el problema de práctica 8.7, se extrae una muestra de 4 personas. Salvo esta última condición, el problema es el mismo. La población está compuesta por 50 hombres y 60 mujeres. Como antes, el muestreo es de una persona a la vez, con reemplazo. ¿Cuál es la probabilidad de obtener 3 mujeres y 1 hombre, en ese mismo orden? La solución aparece en la figura 8.4. Como el problema requiere una mujer en la primera y segunda y tercera extracciones, y un hombre en la cuarta extracción, y como el muestreo es con reemplazo, aplicamos la regla del producto para más de dos eventos independientes. Esta regla es similar a la del producto para dos eventos independientes, sólo que tiene más factores. Así, $p(\text{una mujer en la primera extracción y una mujer en la segunda extracción, una mujer en la tercera extracción y un hombre en la cuarta extracción}) = p(\text{una mujer en la primera extracción}) p(\text{una mujer en la segunda extracción}) p(\text{una mujer en la tercera extracción}) p(\text{un hombre en la cuarta extracción})$. Hay 60 mujeres, 50 hombres y 110 personas en la población. Como el muestreo es con reemplazo, $p(\text{una mujer en la primera extracción}) = \frac{60}{110}$, $p(\text{una mujer en la segunda extracción}) = \frac{60}{110}$, $p(\text{una mujer en la tercera extracción}) = \frac{60}{110}$, y

$$p(A \text{ y } B \text{ y } C \text{ y } D) = p(A)p(B)p(C)p(D)$$

$$\begin{aligned} \left[p(\text{una mujer en la primera extracción y una mujer} \right. \\ \left. \text{en la segunda extracción y una mujer en la tercera} \right. \\ \left. \text{extracción y un hombre en la cuarta extracción)} \right] &= \left(\frac{60}{110} \right) \left(\frac{60}{110} \right) \left(\frac{60}{110} \right) \left(\frac{50}{110} \right) \\ &= \frac{1080}{14\,641} \\ &= 0.0738 \end{aligned}$$

- donde
- A = representa una mujer en la primera extracción
 - B = representa una mujer en la segunda extracción
 - C = representa una mujer en la tercera extracción
 - D = representa un hombre en la cuarta extracción

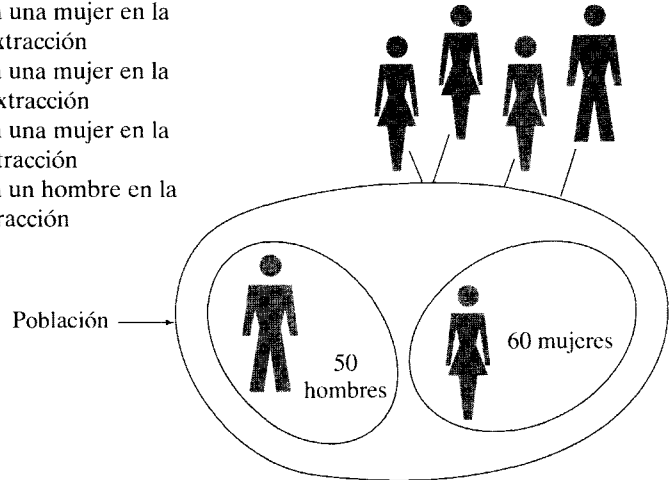


figura 8.4 Determinación de la probabilidad de obtener al azar 3 mujeres y 1 hombre, en ese orden, en cuatro extracciones de una población de 50 varones y 60 mujeres; el muestreo es de una persona a la vez, con reemplazo: regla del producto para más de dos eventos independientes.

$p(\text{un hombre en la cuarta extracción}) = \frac{50}{110}$. Así, $p(\text{una mujer en la primera extracción y una mujer en la segunda extracción y una mujer en la tercera extracción y un hombre en la cuarta extracción}) = \frac{60}{110} \left(\frac{60}{110} \right) \left(\frac{60}{110} \right) \left(\frac{50}{110} \right) = 1080/14\,641 = 0.0738$.

Regla del producto: Eventos dependientes Cuando A y B son dependientes, la probabilidad de que ocurra B resulta afectada por la ocurrencia de A . En este caso, no podemos simplificar la ecuación para la probabilidad de A y B . Tenemos que utilizarla en su forma original. Por consiguiente, si A y B son dependientes,

$$p(A \text{ y } B) = p(A)p(B|A) \quad \text{regla del producto para eventos dependientes}$$

El muestreo sin reemplazo proporciona una buena ilustración de los eventos dependientes. Supongamos que usted va a extraer dos cartas, una a la vez, *sin* reemplazo, de un mazo de barajas ordinarias. ¿Cuál es la probabilidad de que ambas cartas sean ases?

La solución se presenta en la figura 8.5. Convengamos en que A representa un as en la primera extracción y B representa un as en la segunda extracción. Como el muestreo se realiza sin reemplazo (la carta obtenida en la primera extracción es retirada de la baraja), la ocurrencia de A *realmente* afecta la probabilidad de B . A y B son dependientes entre sí. Como el problema requiere un as en la primera extracción y un as en la segunda, y como estos eventos son dependientes, aplicamos la regla del producto para eventos dependientes. Así, $p(\text{un as en la primera extracción y un as en la segunda extracción}) = p(\text{un as en la primera extracción})p(\text{un as en la segunda extracción, en virtud de que se$

$$p(A \text{ y } B) = p(A)p(B|A)$$

$$\left[p(\text{un as en la primera extracción}) \right. \\ \left. \text{y un as en la segunda extracción} \right] = \left[\begin{array}{l} p(\text{un as en la primera extracción}) \\ p(\text{un as en la segunda extracción después} \\ \text{de obtener un as en la primera extracción}) \end{array} \right]$$

$$= \left(\frac{4}{52} \right) \left(\frac{3}{51} \right)$$

$$= \frac{12}{2652} = 0.0045$$

donde $A =$ representa obtener un as en la primera extracción
 $B =$ representa obtener un as en la segunda extracción

Eventos favorables a A

Eventos favorables a B

figura 8.5 Determinación de la probabilidad de elegir de manera aleatoria dos ases en dos extracciones de una baraja ordinaria. El muestreo es de una carta a la vez y sin reemplazo: regla del producto para eventos dependientes.

obtuvo un as en la primera extracción). Para la primera extracción, hay 4 ases y 52 cartas. Por lo tanto, $p(\text{un as en la primera extracción}) = \frac{4}{52}$. En virtud de que el muestreo se realiza sin reemplazo, $p(\text{un as en la segunda extracción después de haber obtenido un as en la primera extracción}) = \frac{3}{51}$. Así, pues, $p(\text{un as en la primera extracción y un as en la segunda extracción}) = \frac{4}{52} \left(\frac{3}{51} \right) = \frac{12}{2652} = 0.0045$.

Suponga que extrae al azar dos frutas, una a la vez, de la misma bolsa descrita en el problema de práctica 8.6. Como antes, dicha bolsa contiene cuatro manzanas, seis naranjas y cinco duraznos. Sin embargo, esta vez obtendrá la muestra *sin* reemplazo. ¿Cuál es la probabilidad de obtener una naranja y una manzana, en ese orden?

SOLUCIÓN

La solución aparece en la figura acompañante. Como el problema requiere una naranja y una manzana, y como el muestreo se realiza sin reemplazo, aplicamos la regla del producto para eventos dependientes. Así, $p(\text{una naranja en la primera extracción y una manzana en la segunda extracción}) = p(\text{una naranja en la primera extracción})p(\text{una manzana en la segunda extracción después de haber obtenido una naranja en la primera extracción})$. En la primera extracción hay 6 naranjas y 15 frutas. Por lo tanto, $p(\text{una naranja en la primera extracción}) = \frac{6}{15}$. En virtud de

que el muestreo se realiza sin reemplazo, $p(\text{una manzana en la segunda extracción después de haber obtenido una naranja en la primera extracción}) = \frac{4}{14}$. Por lo tanto, $p(\text{una naranja en la primera extracción y una manzana en la segunda extracción}) = \frac{6}{15} \left(\frac{4}{14}\right) = \frac{24}{210} = 0.1143$.

$$p(A \text{ y } B) = p(A)p(B|A)$$

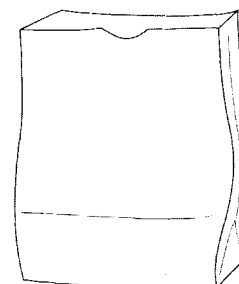
$$\left[\begin{array}{l} p(\text{una naranja en la primera extracción y un manzana en la segunda extracción}) \end{array} \right] = \left[\begin{array}{l} p(\text{una naranja en la primera extracción}) \\ p(\text{una manzana en la segunda extracción después de haber obtenido una naranja en la primera extracción}) \end{array} \right]$$

$$= \left(\frac{6}{15}\right)\left(\frac{4}{14}\right)$$

Naranja Manzana

$$= \frac{24}{210} = 0.1143$$

donde A = representa obtener una naranja en la primera extracción
 B = representa obtener una manzana en la segunda extracción



Eventos favorables a A
Naranjas

Eventos favorables a B
Manzanas



En una clase de una escuela superior hay 15 estudiantes de música, 24 de historia y 46 de psicología. Si usted extrac al azar 2 alumnos del grupo. ¿cuál es la probabilidad de que ambos sean estudiantes de historia? El muestreo es de una persona a la vez, sin reemplazo.

SOLUCIÓN

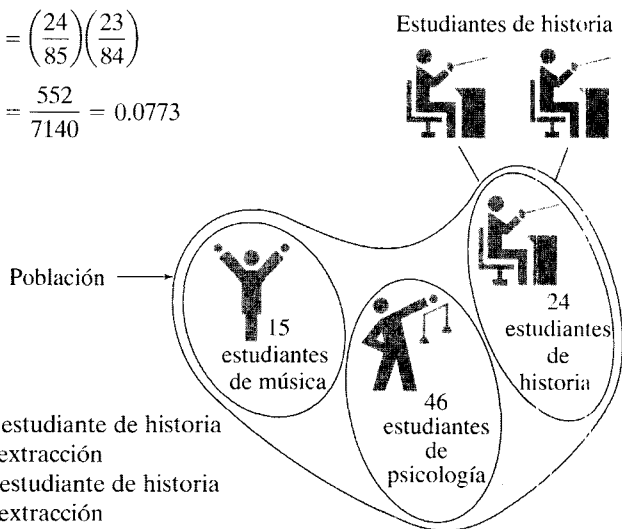
La solución se presenta en la figura acompañante. Como el problema requiere un estudiante de historia en la primera extracción y un estudiante de historia en la segunda extracción, y el muestreo se realiza sin reemplazo, es conveniente aplicar la regla del producto para eventos dependientes. Así, $p(\text{un estudiante de historia en la primera extracción y un estudiante de historia en la segunda extracción}) = p(\text{un estudiante de historia en la primera extracción})p(\text{un estudiante de historia en la segunda extracción después de haber obtenido un estudiante de historia en la$

(continúa)

primera extracción). Para la primera extracción, la población tenía 24 estudiantes de historia y 85 personas. Por lo tanto, $p(\text{un estudiante de historia en la primera extracción}) = \frac{24}{85}$. El muestreo se realiza sin reemplazo, $p(\text{un estudiante de historia en la segunda extracción después de haber obtenido un estudiante de historia en la primera extracción}) = \frac{23}{84}$. Por lo tanto, $p(\text{un estudiante de historia en la primera extracción y un estudiante de historia en la segunda extracción}) = \frac{24}{85} \left(\frac{23}{84} \right) = \frac{552}{7140} = 0.0773$.

$$p(A \text{ y } B) = p(A)p(B|A)$$

$$\begin{aligned} \left[p(\text{un estudiante de historia en la primera extracción y un estudiante de historia en la segunda extracción}) \right] &= \left[p(\text{un estudiante de historia en la primera extracción}) p(\text{un estudiante de historia en la segunda extracción después de haber obtenido un estudiante de historia en la primera extracción}) \right] \\ &= \left(\frac{24}{85} \right) \left(\frac{23}{84} \right) \\ &= \frac{552}{7140} = 0.0773 \end{aligned}$$



donde A = representa un estudiante de historia en la primera extracción
 B = representa un estudiante de historia en la segunda extracción

Tal como ocurre con la regla del producto para eventos independientes, la regla del producto para eventos dependientes se aplica también a situaciones en las que intervienen más de dos eventos. En esos casos, la ecuación adopta la forma siguiente:

$$p(A \text{ y } B \text{ y } C \text{ y } \dots \text{ y } Z) = p(A)p(B|A)p(C|AB) \cdots p(Z|ABC \dots)$$

regla del producto para más de dos eventos dependientes

donde $p(A)$ = probabilidad de A
 $p(B|A)$ = probabilidad de B dado que A ha ocurrido
 $p(C|AB)$ = probabilidad de C dado que A y B han ocurrido
 $p(Z|ABC \dots)$ = probabilidad de Z dado que A , B , C y todos los demás eventos han ocurrido

Para ilustrar el uso de esta ecuación, resolvamos un problema que incluye más de dos eventos dependientes. Suponga usted que va a extraer una muestra de 4 alumnos de la clase de una escuela superior presentada en el problema de práctica 8.9. En esa clase había 15 estudiantes de música, 24 de historia y 46 de psicología. Si el muestreo es de una persona a la vez, sin reemplazo, ¿cuál es la probabilidad de obtener en la muestra 4 estudiantes de historia?

La solución se presenta en la figura 8.6. Como el problema requiere que se obtenga un estudiante de historia en la primera y la segunda y la tercera y la cuarta extrac-

$$\begin{aligned}
 p(A \text{ y } B \text{ y } C \text{ y } D) &= p(A)p(B|A)p(C|AB)p(D|ABC) \\
 &= \left(\frac{24}{85}\right)\left(\frac{23}{84}\right)\left(\frac{22}{83}\right)\left(\frac{21}{82}\right) \\
 &= \frac{255.024}{48.594.840} = 0.0052
 \end{aligned}$$

donde A = representa obtener un estudiante de historia en la primera extracción
 B = representa obtener un estudiante de historia en la segunda extracción
 C = representa obtener un estudiante de historia en la tercera extracción
 D = representa obtener un estudiante de historia en la cuarta extracción

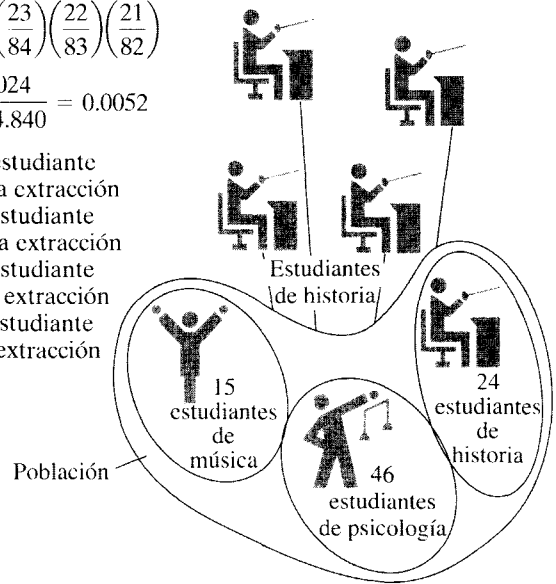


figura 8.6 Determinación de la probabilidad de extraer de manera aleatoria a 4 estudiantes de historia en cuatro extracciones de una población de 15 alumnos de música, 24 de historia y 46 de psicología. El muestreo es de una persona a la vez, sin reposición: regla del producto para varios eventos dependientes.

ciones, y como el muestreo se realiza sin reposición, es conveniente aplicar la regla del producto con más de dos eventos dependientes. Esta regla se parece mucho a la del producto para dos eventos dependientes, salvo que tiene más factores. Así, para este problema, $p(\text{un estudiante de historia en la primera extracción y un estudiante de historia en la segunda extracción y un estudiante de historia en la tercera extracción y un estudiante de historia en la cuarta extracción}) = p(\text{un estudiante de historia en la primera extracción})p(\text{un estudiante de historia en la segunda extracción después de haber obtenido un estudiante de historia en la primera extracción})p(\text{un estudiante de historia en la tercera extracción después de haber obtenido un estudiante de historia en la primera y la segunda extracciones})p(\text{un estudiante de historia en la cuarta extracción después de haber obtenido un estudiante de historia en la primera, la segunda y la tercera extracciones})$. Para la primera extracción, hay 24 estudiantes de historia y 85 individuos en la población. Así, $p(\text{un estudiante de historia en la primera extracción}) = \frac{24}{85}$. Como el muestreo se realiza sin reposición, $p(\text{un estudiante de historia en la segunda extracción después de haber obtenido un estudiante de historia en la primera extracción}) = \frac{23}{84}$, $p(\text{un estudiante de historia en la tercera extracción después de haber obtenido un estudiante de historia en la primera y la segunda extracciones}) = \frac{22}{83}$, y $p(\text{un estudiante de historia en la cuarta extracción después de haber obtenido un estudiante de historia en la primera, la segunda y la tercera extracciones}) = \frac{21}{82}$. Por lo tanto, $p(\text{un estudiante de historia en la primera extracción y un estudiante de historia en la segunda extracción y un estudiante de historia en la tercera extracción y un estudiante de historia en la cuarta extracción}) = \frac{24}{85} \left(\frac{23}{84}\right) \left(\frac{22}{83}\right) \left(\frac{21}{82}\right) = 255.024/48.594.840 = 0.0052$.

Reglas de la suma y del producto

En algunas situaciones, para hallar la solución es necesario aplicar las reglas de la suma y del producto. Por ejemplo, supongamos que voy a lanzar una vez dos dados no carga-

$$\begin{aligned}
 p(A) &= p(5 \text{ en el primer dado y } 6 \text{ en el segundo dado}) \\
 &= p(5 \text{ en el primer dado})p(6 \text{ en el segundo dado}) \\
 &= \left(\frac{1}{6}\right)\left(\frac{1}{6}\right) = \frac{1}{36}
 \end{aligned}$$

$$\begin{aligned}
 p(B) &= p(6 \text{ en el primer dado y } 5 \text{ en el segundo dado}) \\
 &= p(6 \text{ en el primer dado})p(5 \text{ en el segundo dado}) \\
 &= \left(\frac{1}{6}\right)\left(\frac{1}{6}\right) = \frac{1}{36}
 \end{aligned}$$

$$\begin{aligned}
 p(\text{suma de } 11) &= p(A \text{ o } B) = p(A) + p(B) \\
 &= \frac{1}{36} + \frac{1}{36} = \frac{2}{36} = 0.0556
 \end{aligned}$$

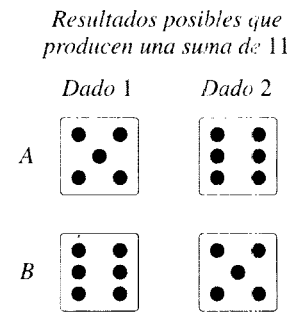


figura 8.7 Determinación de la probabilidad de obtener una suma de 11 puntos al lanzar dos dados no cargados: reglas del producto y de la suma.

dos. ¿Cuál es la probabilidad de que la suma de los números que aparezcan en los dados sea igual a 11? La solución se muestra en la figura 8.7. Existen dos resultados posibles cuya suma es 11 (5 en el primer dado y 6 en el segundo, que llamaremos resultado *A*; y 6 en el primer dado y 5 en el segundo, lo cual llamaremos resultado *B*). En virtud de que los dados son independientes entre sí, podemos aplicar la regla del producto para eventos independientes a fin de determinar la probabilidad de cada resultado. Al utilizar esta regla, $p(A) = \frac{1}{6}\left(\frac{1}{6}\right) = \frac{1}{36}$, y $p(B) = \frac{1}{6}\left(\frac{1}{6}\right) = \frac{1}{36}$. Como cualquiera de estos resultados produce una suma de 11 puntos, $p(\text{suma igual a } 11) = p(A \text{ o } B)$. Como estos resultados son mutuamente excluyentes, podemos aplicar la regla de la suma para eventos mutuamente excluyentes a fin de determinar $p(A \text{ o } B)$. Así, pues, $p(\text{suma de } 11) = p(A \text{ o } B) = p(A) + p(B) = \frac{1}{36} + \frac{1}{36} = \frac{2}{36} = 0.0556$.

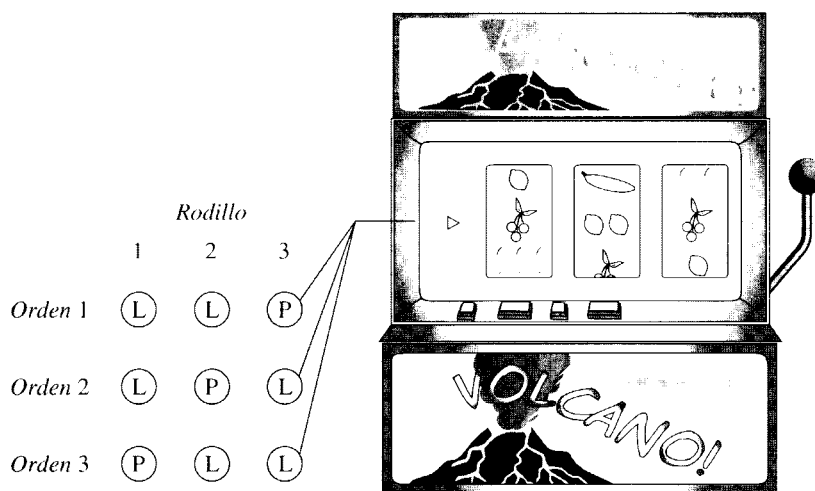
Tratemos de resolver otro problema en el que intervendrán las reglas del producto y de la suma.



Imagine usted que ha llegado a Las Vegas y decide probar su “suerte” con una máquina tragamonedas. Por si acaso no conoce ese tipo de máquinas, éstas consisten en tres rodillos que giran en forma independiente. Cada rodillo muestra las imágenes de diversos objetos. Suponga que la máquina con la que está jugando tiene siete frutas distintas en el rodillo 1: un limón, una ciruela, una manzana, una naranja, una pera, algunas cerezas y un plátano. Los rodillos 2 y 3 tienen las imágenes de las mismas frutas que el primero. Cuando se tira de la palanca hacia abajo, los tres rodillos giran en forma independiente y después se detienen. La máquina tiene una ventanilla al frente de cada rodillo. Las imágenes de las frutas pasan por la ventanilla durante la rotación. Cuando la rueda se detiene, una de las frutas de cada tambor se podrá ver inmóvil a través de la ventanilla. Supondremos que cada una de las frutas del rodillo tiene la misma probabilidad de aparecer en la ventanilla al final de la rotación. Usted inserta una moneda y tira de la palanca. ¿Cuál es la probabilidad de que aparezcan dos limones y una pera? El orden de éstos no es importante: lo único que debe importarle es obtener dos limones y una pera, en cualquier orden.

SOLUCIÓN

La solución se presenta en la figura acompañante. Hay tres órdenes posibles en los que pueden aparecer dos limones y una pera: limón, limón, pera; limón, pera, limón; y pera, limón, limón. Como los rodillos giran en forma independiente, podemos utilizar la regla del producto para eventos independientes a fin de determinar la probabilidad de obtener cada uno de esos órdenes. Puesto que la obtención de cada fruta es igualmente probable, $p(\text{limón y limón y pera}) = p(\text{limón})p(\text{limón})p(\text{pera}) = \frac{1}{7}(\frac{1}{7})(\frac{1}{7}) = \frac{1}{343}$. Esta misma probabilidad se aplica también a los otros dos órdenes. Como los tres órdenes producen dos limones y una pera, $p(\text{dos limones y una pera}) = p(\text{orden 1, 2 o 3})$. Al aplicar la regla de la suma para eventos independientes, $p(\text{orden 1, 2 o 3}) = \frac{3}{343} = 0.0087$. Así, la probabilidad de obtener dos limones y una pera, sin importar en qué orden, es igual a 0.0087.



$$\begin{aligned}
 p(\text{orden 1}) &= p(\text{limón en el rodillo 1 y limón en el rodillo 2 y pera en el rodillo 3}) \\
 &= p(\text{limón en el rodillo 1})p(\text{limón en el rodillo 2})p(\text{pera en el rodillo 3}) \\
 &= \left(\frac{1}{7}\right)\left(\frac{1}{7}\right)\left(\frac{1}{7}\right) = \frac{1}{343}
 \end{aligned}$$

$$\begin{aligned}
 p(\text{orden 2}) &= p(\text{limón en el rodillo 1 y pera en el rodillo 2 y limón en el rodillo 3}) \\
 &= p(\text{limón en el rodillo 1})p(\text{pera en el rodillo 2})p(\text{limón en el rodillo 3}) \\
 &= \left(\frac{1}{7}\right)\left(\frac{1}{7}\right)\left(\frac{1}{7}\right) = \frac{1}{343}
 \end{aligned}$$

$$\begin{aligned}
 p(\text{orden 3}) &= p(\text{pera en el rodillo 1 y limón en el rodillo 2 y limón en el rodillo 3}) \\
 &= p(\text{pera en el rodillo 1})p(\text{limón en el rodillo 2})p(\text{limón en el rodillo 3}) \\
 &= \left(\frac{1}{7}\right)\left(\frac{1}{7}\right)\left(\frac{1}{7}\right) = \frac{1}{343}
 \end{aligned}$$

(continúa)

$$\begin{aligned} \left(\begin{array}{l} p(2 \text{ limones}) \\ \text{y una pera} \end{array} \right) &= p(\text{orden 1 o 2 o 3}) = p(\text{orden 1}) + p(\text{orden 2}) + p(\text{orden 3}) \\ &= \frac{1}{343} + \frac{1}{343} + \frac{1}{343} = \frac{3}{343} = 0.0087 \end{aligned}$$

Probabilidad y variables continuas con distribución normal

Hasta este punto, en nuestro análisis de la probabilidad hemos considerado variables discretas, por ejemplo, el muestreo a partir de una baraja de cartas o el lanzamiento de un par de dados. Sin embargo, muchas de las variables dependientes que es necesario evaluar en los experimentos son continuas, no discretas. Cuando una variable es continua,

$$p(A) = \frac{\text{Área debajo de la curva, correspondiente a } A}{\text{Área total bajo la curva}} \quad \text{probabilidad de } A \text{ con una variable continua}$$

En virtud de que con frecuencia (aunque no siempre) estas variables están distribuidas de manera normal, concentraremos nuestro análisis en las variables continuas con distribución normal.

Para ilustrar el uso de la probabilidad con variables continuas que tienen una distribución normal, suponga que hemos determinado los pesos de todas las mujeres inscritas en el cuarto semestre de la universidad en la que usted estudia. Imaginemos que se trata de un conjunto de datos pertenecientes a una población cuya distribución es normal, con una media de 120 libras y una desviación estándar de 8 libras. Si elegimos al azar un dato de esta población, ¿cuál es la probabilidad de que el dato seleccionado sea igual o mayor que 134?

La población se muestra en la figura 8.8. La media de 120 y el dato 134 se localizan sobre el eje X . El área sombreada representa a todos los datos mayores o iguales a 134. Como el muestreo es aleatorio, todos los datos tienen la misma posibilidad de ser seleccionados. Así, la probabilidad de obtener un dato igual o mayor que 134 se determina calculando la proporción del total de datos contenidos en el área sombreada. Puesto que

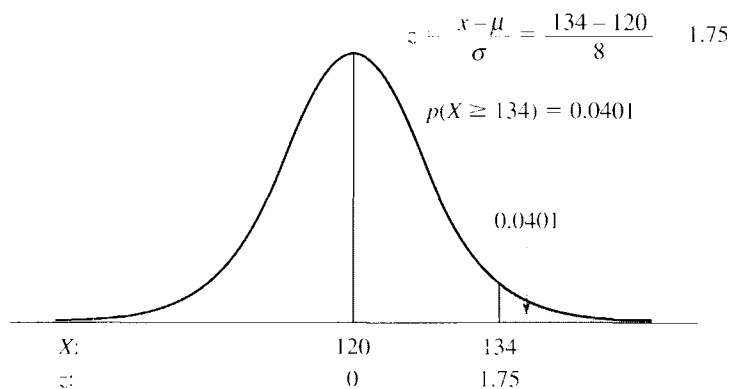


figura 8.8 Probabilidad de obtener $X \geq 134$ si se elige al azar un dato de una población normal con $\mu = 120$ y $\sigma = 8$.

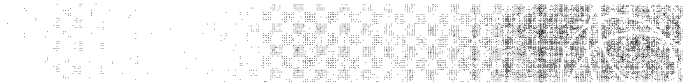
los datos están distribuidos de forma normal, podemos determinar esta proporción convirtiendo el dato en bruto en su valor z transformado, para luego buscar el área en la tabla A del apéndice D. Así, pues,

$$z = \frac{X - \mu}{\sigma} = \frac{134 - 120}{8} = \frac{14}{8} = 1.75$$

De la columna C de la tabla A,

$$p(X \geq 134) = 0.0401$$

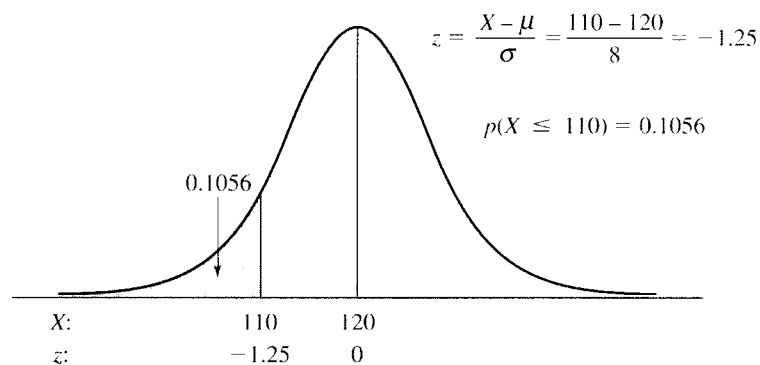
Estamos seguros de que el lector reconocerá que este tipo de problema es similar a los que presentamos en el capítulo 5, cuando estudiamos los puntajes estandarizados. La diferencia principal es que, en este capítulo, el problema se formula en términos de probabilidad y no en términos de la proporción o el porcentaje de los datos, como fue planteado en el capítulo 5. Como quiera que el lector ya se ha familiarizado con problemas de esta índole, no creemos que sea necesario presentar muchos problemas de práctica. Sin embargo, solucionaremos un par de ellos sólo para estar seguros.



Consideremos la misma población de mujeres universitarias que estudian el cuarto semestre. Si se elige un dato aleatoriamente de la población, ¿cuál es la probabilidad de que éste sea igual o menor que 110?

SOLUCIÓN

La solución se muestra en la figura acompañante. El área sombreada representa todos los datos que son iguales o menores que 110. Por el hecho de que el muestreo es aleatorio, todos los datos tienen la misma posibilidad de ser elegidos. Para determinar $p(X \leq 110)$, debemos transformar primero el dato en bruto de 110 en su puntaje z . A continuación, podemos calcular la proporción de datos contenidos en el área sombreada, con ayuda de la tabla A. Así,





Si consideramos de nuevo la población del ejercicio anterior, ¿cuál es la probabilidad de elegir al azar un dato que esté tan o más alejado de la media que un dato cuyo valor es de 138?

SOLUCIÓN

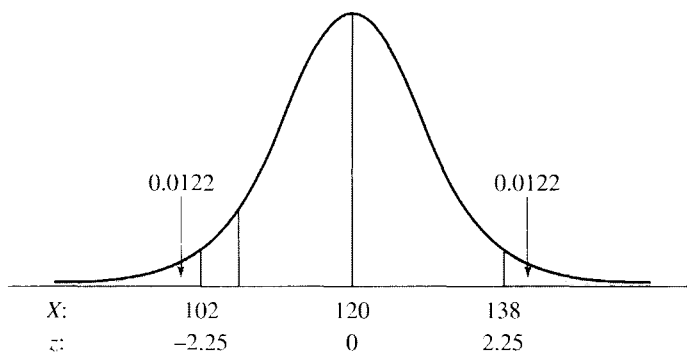
La solución se muestra en la figura acompañante. El dato de 138 está 18 unidades por arriba de la media. Ya que el problema pregunta por datos que estén tan o más alejados de la media, debemos considerar también los datos que estén a 18 unidades o más por debajo de la media. Las áreas sombreadas contienen todos los datos que se encuentran a 18 unidades o más de la media. Como el muestreo es aleatorio, cada dato tiene la misma posibilidad de ser elegido. Para determinar $p(X \leq 102$ o $X \geq 138)$, primero debemos tomar los datos en bruto 102 y 138 y transformarlos en sus puntajes z . Entonces podemos determinar la proporción de datos contenida en las áreas sombreadas, con ayuda de la tabla A. $p(X \leq 102$ o $X \geq 138)$ es igual a esta proporción. Así,

$$z = \frac{X - \mu}{\sigma} = \frac{102 - 120}{8} = -\frac{18}{8} \quad z = \frac{X - \mu}{\sigma} = \frac{138 - 120}{8} = \frac{18}{8}$$

$$= -2.25 \quad = 2.25$$

De la columna C de la tabla A,

$$p(X \leq 102 \text{ o } X \geq 138) = 0.0122 + 0.0122 = 0.0244$$





“No soy culpable, soy una víctima de las coincidencias”: ¿excusa descarada o verdad?

A pesar de una tradición que privilegia la toma de decisiones cualitativas sobre las cuantitativas, en el ámbito jurídico se utiliza cada vez más la estadística como una base para dictar sentencia. El siguiente caso, ocurrido en Suecia, es un ejemplo de esto.

En un juicio que se llevó a cabo en Suecia, al demandado se le impuso una multa porque dejó estacionado su automóvil en un lugar por más tiempo de lo permitido. Un oficial de policía anotó la posición de las válvulas de los neumáticos, delantero y trasero de uno de los lados del vehículo del acusado, según una representación basada en las manecillas del reloj, y resultó que la válvula delantera apuntaba a la una y la trasera a las seis, con redondeo a la hora más próxima en ambos casos (véase el diagrama). Después de transcurrido el tiempo permitido, el auto continuaba en el mismo lugar, con las dos válvulas apuntando igual que antes, es decir, a la una y a las seis, respectivamente. El acusado recibió la notificación de la multa.

Sin embargo, en el tribunal, él alegó ser inocente y afirmó que había dejado el lugar de estacionamiento a tiempo, pero que había regresado más tarde y simplemente sucedió que las válvulas quedaron en la misma posición que antes. El juez, que no había estudiado algún curso básico de estadística, llamó a un experto en la materia para que

evaluara la afirmación de coincidencia del demandado. ¿Es razonable su afirmación? Supongamos que usted es el experto. ¿Qué le diría al juez? Al formular su respuesta, suponga que las ruedas son independientes, tal como lo habría hecho el estadístico que asesoró al juez.

Respuesta Como estadístico, el trabajo de usted consiste en determinar en qué medida es razonable la afirmación de coincidencia. Si suponemos que la historia del demandado es verídica, es decir, que salió y luego regresó al mismo lugar de estacionamiento, ¿cuál es la probabilidad de que las dos válvulas volvieran exactamente a las mismas posiciones que ocupaban antes, según la representación de las manecillas del reloj, es decir, apuntando a la una y a las seis, respectivamente? Como existen 12 posiciones posibles para cada válvula, suponiendo que las ruedas sean independientes y utilizando la regla del producto,

$$p(\text{una y seis}) = \left(\frac{1}{12}\right)\left(\frac{1}{12}\right) = \left(\frac{1}{144}\right) = 0.0069$$

Pues bien, si fue una coincidencia (es decir, si sólo intervino el azar), la probabilidad de que las válvulas regresaran a sus posiciones originales es de aproximadamente 7 en 1000. ¿Qué cree usted que hizo el juez al recibir esta información? Por increíble que parezca, el juez exoneró al acusado, diciendo que si las cuatro ruedas hubieran sido verificadas y hubiera resultado que todas apuntaban en la misma dirección que antes ($p = \frac{1}{12} \times \frac{1}{12} \times \frac{1}{12} \times \frac{1}{12} = \frac{1}{20736} = 0.00005$), entonces habría rechazado la afirmación de coincidencia por ser altamente improbable y habría multado al acusado. Por lo tanto, el juez consideró que la explicación de coincidencia era demasiado probable para rechazarla, aunque esos resultados se habrían obtenido sólo 1 entre 144 veces si en verdad hubiera sido una coincidencia. En realidad, como las ruedas no giran en forma independiente, las fórmulas utilizadas subestiman un poco la probabilidad de que regresen a la posición original (¿cómo lo hizo usted? ¿Podremos llamarlo en el futuro como experto en estadística para asesorar a la justicia?)





Declive del número de espermatozoides: ¿falla del hombre o del muestreo?

El encabezado de un artículo publicado en 1995 en un importante periódico decía: "Un estudio de 20 años muestra una disminución del número de espermatozoides de los hombres fértiles". Reproducimos a continuación algunas partes de ese artículo.

Un nuevo estudio ha descubierto un notable declive en la cantidad de espermatozoides de los hombres fértiles durante los últimos 20 años. . . .

El artículo, publicado hoy en The New England Journal of Medicine, se basó en los datos reunidos durante un periodo de 20 años en un banco de esperma de París. En Estados

Unidos, algunos expertos presentaron sólidas objeciones a esos descubrimientos. . . .

En el nuevo estudio, realizado por el doctor Pierre Jouannet del Centro para el Estudio y la Conservación de Óvulos y Esperma de esa capital, se examinó el semen recolectado por un banco de esperma en París a partir de 1973. Se ha informado que la cantidad de espermatozoides encontrada se redujo en un promedio de 2.1% cada año, pasando de 89 millones de espermatozoides por mililitro en 1973 a un promedio de 60 millones por mililitro en 1992. Al mismo tiempo, se determinó que los porcentajes de espermatozoides dotados de movimiento normal y sin deformaciones se redujo entre 0.5 y 0.6 de 1% cada año.

El artículo está acompañado de un editorial incluido por invitación, escrito por un experto en esterilidad masculina, el doctor Richard Sherins, director de la división de andrología del Instituto de Genética

e IVF en Fairfax, Va., quien afirmó que ese estudio y otros anteriores adolecían de fallas metodológicas, por lo cual sus datos no podían ser interpretados correctamente.

Sherins afirmó que en los estudios no se analizó esperma de hombres elegidos al azar, que la cantidad de espermatozoides y su calidad variaba tanto de una semana a otra, que era muy azaroso basarse en una sola muestra para medir la calidad de los espermatozoides, como se hizo en esos estudios.

¿Qué opina usted? ¿Por qué pudo haber sido tan importante que se utilizaran muestras de hombres elegidos al azar y no de los hombres que depositan su semen en un banco de esperma? ¿Por qué se podría complicar la interpretación de los datos por la gran variabilidad observada en la cantidad y la calidad de los espermatozoides de una semana a otra?





La muestra de una muestra

El siguiente artículo sobre encuestas fue publicado en un número reciente de *The New York Times* en la red. Me pareció que es un artículo muy interesante y espero que ustedes también lo consideren así. Presentamos a continuación ese artículo en su totalidad.

Cómo se localiza al encuestado "típico"

Por **MICHAEL KAGAY**

"¿A qué clase de personas llama usted por teléfono para sus encuestas? Seguramente no está encuestando a ningún residente de Texas en este momento."

Los encuestadores reciben muchas llamadas como ésta, por lo general de personas que en verdad se sienten desconcertadas al ver que sus propios puntos de vista sobre algún tema de controversia, o incluso las opiniones de la mayoría de la gente de su respectiva localidad, no coinciden con la opinión registrada por las encuestas como la de la mayoría en todo el país.

Sin embargo, la mayor parte de los encuestadores tienen mucho cuidado y tratan de asegurarse de obtener una muestra representativa de entrevistados para sus encuestas. Es posible que la redacción correcta de las preguntas siga siendo un arte y que la conducción habilidosa de entrevistas sea todo un oficio, pero el muestreo apropiado es lo que introduce el elemento científico en las encuestas.

La marcación de dígitos al azar (o RDD por sus siglas en inglés) es

el procedimiento de muestreo estándar que todos los encuestadores profesionales usan en la actualidad. El objetivo es dar a todos los números telefónicos residenciales de los Estados Unidos la misma posibilidad de ser contactados para una entrevista. Casi todas las organizaciones de encuestas utilizan algún tipo de RDD.

Por ejemplo, la Encuesta de *The New York Times* y CBS News se basa en el sistema GENESYS, desarrollado y sostenido por el Marketing Systems Group de Filadelfia. Ese sistema consiste en una base de datos que contiene más de 42 000 conferencias telefónicas residenciales en todo el territorio del país, que se actualiza a intervalos de unos cuantos meses a fin de incluir los nuevos códigos de área y las conferencias realizadas en fechas más recientes.

Ese sistema incluye también software de computadora diseñado para extraer una muestra al azar de esas conferencias en cada nueva encuesta, y para formar al azar los cuatro últimos dígitos de cada uno de los números telefónicos que van a ser contactados.

Esa elección aleatoria del universo de todos los números telefónicos posibles es el factor que garantiza la representatividad de la muestra de las unidades familiares que son contactadas para una encuesta.

En el caso de la encuesta sobre la que preguntaba una persona al inicio de este artículo, los tejanos representan cerca de 7% del total de los estadounidenses, pero sólo tuvieron 6% de representación entre quienes respondieron a esa encuesta en particular. La cifra casi alcanzó la marca, pero quedó un poco baja.

La búsqueda insistente de entrevistados

Algunos tipos de personas son más difíciles de localizar y de entrevistar que otros. ¿Qué se hace para reducir o evitar la presencia de sesgos en los resultados?

El primer remedio consiste en la búsqueda insistente de las familias y luego del miembro designado de cada una de ellas que será encuestado.

En el curso de la mayoría de las encuestas hay que realizar múltiples llamadas a cada familia, lo cual incluye a menudo llamadas diurnas y nocturnas, entre semana y los fines de semana. Los entrevistadores dejan mensajes en máquinas contestadoras telefónicas sobre el propósito de la llamada, conciertan nuevas citas si la persona designada no está en la casa, y anuncian que volverán a llamar a una hora más conveniente para el entrevistado.

Cuando las personas se niegan a ser entrevistadas, en muchas encuestas se emplea también un escuadrón especial de entrevistadores dispuestos a hacer una llamada adicional para tratar de convertir una negativa en una entrevista completa. Esa tarea requiere destrezas especiales, habilidad para convencer a la persona que responde sobre la importancia de la entrevista, poseer la personalidad apropiada y ser capaz de tolerar un posible rechazo.

Ponderación o balanceo después de la entrevista

Sin embargo, casi todas las encuestas incluyen un número ligeramente excesivo de mujeres, personas jubiladas, egresados de la universidad e individuos blancos, en comparación con las proporciones de éstos que según la

(continúa)



La muestra de una muestra

(continuación)

Oficina del Censo existen en la población de Estados Unidos en un momento cualquiera.

Esto sucede así porque la frecuencia con la que las personas están en su casa es variable, así como su disposición para hablar por teléfono con desconocidos, y el grado de confianza que puedan sentir frente al hecho de compartir sus opiniones sobre temas políticos y sociales.

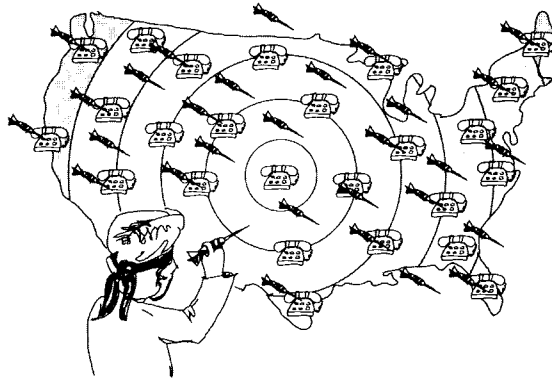
Por lo tanto, la mayoría de las encuestas ponderan ligeramente a sus encuestados —aumentando ligeramente la proporción de hombres, personas jóvenes, gente con menos instrucción y minorías raciales— para que sus proporciones tengan el equilibrio apropiado. Algunos encuestadores se refieren a este procedimiento como el balanceo.

Al mismo tiempo, la Encuesta del Times y CBS News aprovecha también la oportunidad y hace un ligero ajuste en cuanto al tamaño de la unidad familiar y el número de líneas telefónicas con que cuenta la vivienda. Con esto se equilibra la probabilidad de selección, en el caso de individuos en cuyas casas hay varias líneas telefónicas, y también el hecho de que en algunas viviendas familiares hay más residentes adultos que en otras.

Todos estos ajustes menores sirven para hacer que una buena muestra sea aún mejor y más representativa.

Una muestra que tenga la fisonomía del país

Entonces, ¿cómo son los entrevistados de una típica Encuesta del New



York Times y CBS News?

Después de todas las marcaciones telefónicas al azar, las segundas llamadas, la conversión de negativas en aceptaciones y la ponderación, la encuesta típica se compone de los siguientes elementos:

- ◆ En términos de geografía, 22% de los entrevistados viven en el nordeste, 33% en el sur, 24% en el medio oeste y 21% en el oeste.
- ◆ En términos de género, 47% de ellos son hombres y 53% son mujeres.
- ◆ En términos de raza, 80% son blancos, 11% negros, 1% asiáticos y 6% de otras razas.
- ◆ En cuanto al nivel de educación, hay 24% que terminó la universidad, otro 27% que hizo algunos estudios en escuelas superiores o vocacionales, 37% que no pasó de la preparatoria y 12% que no terminó la preparatoria.
- ◆ En lo que se refiere a política, el ánimo cambia de vez en cuando, pero a últimas fechas 27% se considera republicano, al tiempo que 36% se identifica como demócrata y otro 30% asegura ser independiente.
- ◆ En cuanto a la filosofía política, 20% suele considerarse liberal,

mientras que 42% se identifica como moderado y 32% se proclama conservador.

Ésa es una muestra que si refleja con bastante fidelidad a los Estados Unidos.

Bueno, cuando menos refleja a los Estados Unidos de habla inglesa, ya que la mayoría de las encuestas nacionales se realizan hoy únicamente en inglés. Ésta es una limitación de las encuestas que debe ser tomada en cuenta, ya que la población de habla española va en aumento en el país. Eso nos permite explicar también por qué la Encuesta del Times y CBS News se quedó un poco corta en cuanto a los tejanos incluidos en la encuesta a la que nos referíamos al principio del artículo; Texas es un estado en el cual la población de habla española es abundante.

Las encuestas tienen también otras limitaciones y problemas, y en futuras columnas de Poll Watch los iremos abordando también.

Fuente: Derechos de autor © 1999 por The New York Times Co. Reproducido con autorización.

RESUMEN

En este capítulo hemos analizado los temas del muestreo aleatorio y la probabilidad. Una muestra aleatoria se define como aquella que ha sido extraída de una población por medio de un proceso que garantice lo siguiente: 1) que cada muestra posible de un tamaño dado tenga la misma probabilidad de ser elegida y 2) que todos los miembros de la población tengan la misma posibilidad de ser elegidos en la muestra. Después de definir y analizar la importancia del muestreo aleatorio, describimos varios métodos para obtener una muestra aleatoria. A continuación, procedimos a analizar el muestreo con recemplazo y sin él.

Al presentar el tema de la probabilidad, señalamos que ésta se puede enfocar desde dos puntos de vista: *a priori* y *a posteriori*. Según el punto de vista *a priori*, $p(A)$ se define como:

$$p(A) = \frac{\text{Número de eventos clasificables como } A}{\text{Número total de eventos posibles}}$$

Desde un punto de vista *a posteriori*, $p(A)$ se define como:

$$p(A) = \frac{\text{Número de veces que ha ocurrido } A}{\text{Número total de ocurrencias}}$$

En virtud de que la probabilidad es fundamentalmente una proporción, ésta varía entre 0.00 a 1.00. A continuación, presentamos dos reglas de probabilidad necesarias para comprender la estadística inferencial: la regla de la suma y la del producto. Suponiendo que existen dos eventos (A y B), la regla de la suma determina la probabilidad de A o B , mientras que la del producto determina la probabilidad de A y B . La regla de la suma establece lo siguiente:

$$p(A \text{ o } B) = p(A) + p(B) - p(A \text{ y } B)$$

Si los eventos son mutuamente excluyentes:

$$p(A \text{ o } B) = p(A) + p(B)$$

Si los eventos son mutuamente excluyentes y exhaustivos:

$$p(A) + p(B) = 1.00$$

La regla del producto establece lo siguiente:

$$p(A \text{ y } B) = p(A)p(B|A)$$

Si los eventos son mutuamente excluyentes:

$$p(A \text{ y } B) = 0$$

Si los eventos son independientes:

$$p(A \text{ y } B) = p(A)p(B)$$

Si los eventos son dependientes, debemos utilizar la ecuación general

$$p(A \text{ y } B) = p(A)p(B|A)$$

Además, analizamos: 1) la generalización de esas ecuaciones a situaciones en las que hay más de dos eventos y 2) situaciones en las que para hallar la solución se requieren ambas reglas, la de la suma y la del producto. Por último, comentamos la probabilidad de A con variables continuas y describimos la forma de determinar $p(A)$ cuando la variable es continua y presenta una distribución normal. La ecuación para establecer la probabilidad de A cuando la variable es continua es la siguiente:

$$p(A) = \frac{\text{Área debajo de la curva correspondiente a } A}{\text{Área total debajo de la curva}}$$

TÉRMINOS IMPORTANTES

A posteriori (página 166)

A priori (página 166)

Conjunto de eventos exhaustivos (página 172)

Eventos mutuamente excluyentes (página 168)

Independencia entre dos eventos (página 173)

Muestra aleatoria (página 162)

Muestreo con recemplazo (página 165)

Muestreo sin recemplazo (página 165)

Probabilidad (página 166)

Probabilidad de ocurrencia de A o B (página 168)

Probabilidad de ocurrencia de A y B a la vez (página 173)

Regla de la suma (página 168)

Regla del producto (página 173)

PREGUNTAS Y PROBLEMAS

- Defina o identifique cada uno de los conceptos de la sección "Términos importantes".
- ¿Cuáles son los dos propósitos del muestreo aleatorio?
- Supongamos que se desea extraer una muestra aleatoria de 20 sujetos a partir de una población de 400 individuos. El muestreo se llevará a cabo sin recemplazo y usted ha decidido que utilizará la

- tabla J del apéndice D para conseguir la aleatoriedad. Explique de qué manera utilizaría la tabla para seleccionar la muestra.
4. Una psicóloga del desarrollo está interesada en evaluar la “inteligencia emocional” de estudiantes universitarios. El diseño experimental requiere que un cuestionario para medir la inteligencia emocional sea respondido por una muestra de voluntarios formada por 100 estudiantes de licenciatura que participan en un curso de introducción a la psicología impartido actualmente en su universidad. Suponga usted que ésta es la única muestra que se utilizará en este estudio y comente si le parece apropiada.
 5. ¿Cuál es la diferencia entre la probabilidad *a priori* y *a posteriori*?
 6. La regla de la suma permite obtener la probabilidad de ocurrencia de cualquiera de varios eventos, mientras que la regla del producto determina la probabilidad de la ocurrencia conjunta o sucesiva de varios eventos. ¿Es correcta esta afirmación? Explíquelo por medio de ejemplos que ilustren su respuesta.
 7. Al resolver problemas en los que interviene la regla del producto, es útil distinguir entre tres condiciones. ¿Cuáles son esas condiciones? ¿Por qué es útil diferenciarlas?
 8. ¿Cuál es la definición de la probabilidad cuando la variable es continua?
 9. ¿Cuáles de los siguientes ejemplos representan eventos independientes?
 - a. Obtener un 3 y un 4 al lanzar una vez un par de dados no cargados.
 - b. Obtener un as y un rey, en ese orden, al extraer sin reemplazo dos cartas de una baraja.
 - c. Obtener un as y un rey, en ese orden, al extraer con reemplazo dos cartas de una baraja.
 - d. Un cielo nublado antes de la lluvia.
 - e. La luna llena y comer una hamburguesa.
 10. ¿Cuáles de los siguientes ejemplos representan eventos mutuamente excluyentes?
 - a. Obtener un 4 y un 7 al extraer una sola carta de una baraja ordinaria.
 - b. Obtener un 3 y un 4 al lanzar una vez dos dados no cargados.
 - c. Ser hombre y estar embarazada.
 - d. Obtener un 1 y un número par al lanzar una vez un dado no cargado.
 - e. Casarse y seguir siendo soltero.
 11. ¿Cuáles de los siguientes ejemplos representan eventos exhaustivos?
 - a. Lanzar una moneda y obtener una cara o una cruz (la moneda no puede caer de canto).
 - b. Lanzar un dado y obtener un dos.
 - c. Presentar un examen y salir aprobado o no.
 - d. Acudir a una cita y pasar un buen rato.
 12. Al principio de la temporada de béisbol de un año en particular, las probabilidades de que los Yanquis de Nueva York ganen el gallardete de la Liga Americana son de 3 contra 2.
 - a. ¿Cuáles son las probabilidades de que los Yanquis pierdan el gallardete?
 - b. ¿Cuáles son las probabilidades de que los Yanquis ganen el gallardete? Expresé su respuesta en forma decimal.
 - c. ¿Cuáles son las probabilidades de que los Yanquis pierdan el gallardete? Expresé su respuesta en forma decimal. otra
 13. Si usted extrae una sola carta de una baraja ordinaria, ¿cuál es la probabilidad de que sea:
 - a. El as de diamantes?
 - b. Un 10?
 - c. Una reina o un corazón?
 - d. Un 3 o una carta negra? otra
 14. Si usted lanza una vez dos dados no cargados, ¿cuál es la probabilidad de que obtenga:
 - a. Un 2 en el primer dado y un 5 en el segundo?
 - b. Un 2 y un 5, sin importar cuál de los dados muestre el 2 y cuál el 5?
 - c. Por lo menos un 2 o un 5?
 - d. Una suma igual a 7? otra
 15. Si usted realiza un muestreo aleatorio, tomando un elemento a la vez, con reemplazo, de una bolsa que contiene ocho canicas azules, siete canicas rojas y cinco canicas verdes, ¿cuál es la probabilidad de que obtenga:
 - a. Una canica azul en una sola extracción de la bolsa?
 - b. Tres canicas azules en tres extracciones de la bolsa?
 - c. Una canica roja, una verde y una azul, en ese orden, en tres extracciones de la bolsa?
 - d. Por lo menos dos canicas rojas en tres extracciones de la bolsa? otra
 16. Responda las mismas preguntas del problema 13, excepto que ahora el muestreo es de una canica a la vez sin reemplazo. otra
 17. Suponga usted que está jugando con la máquina tragamonedas descrita en el problema de práctica 8.10 de la página 184. La máquina tiene tres rodillos, y en cada uno de ellos aparece una imagen de un limón, una ciruela, una manzana, una naranja, una pera, cerezas y un plátano (siete imágenes distintas). Usted inserta una moneda y tira de la palanca. ¿Cuál es la probabilidad de que aparezcan:
 - a. Tres naranjas?
 - b. Dos naranjas y un plátano, sin que importe el orden?

- c. Por lo menos dos naranjas? otra
18. Usted quiere llamar a una amiga por teléfono, pero sólo recuerda los tres primeros dígitos de su número telefónico y se ha olvidado de los últimos cuatro. ¿Cuál es la probabilidad de que marque al azar el número correcto si hace un solo intento? otra
19. Usted está planeando “hacer su agosto” en el hipódromo. En una determinada carrera participan siete caballos. Si los caballos están clasificados de la misma manera, ¿cuál es la probabilidad de que usted elija en forma correcta al ganador y al segundo lugar?
20. Una máquina vendedora de goma de mascar tiene 38 chicles de color anaranjado, 30 morados y 18 amarillos. La máquina opera de tal manera que cuando se introduce una moneda sale de ella un chicle.
- Si se usan tres monedas, ¿cuál es la probabilidad de obtener 3 chicles en este orden: anaranjado, morado, anaranjado?
 - Si se utiliza una moneda, ¿cuál es la probabilidad de extraer un chicle morado o amarillo?
 - Cuando se utilizan tres monedas, ¿cuál es la probabilidad de que, de los 3 chicles extraídos, exactamente 1 sea morado y 1 amarillo? otra
21. Si se extraen al azar dos cartas de una baraja ordinaria, una a la vez con reemplazo, ¿cuál es la probabilidad de obtener por lo menos un as? otra
22. En una lotería estatal se paga 1 000 000 de dólares al poseedor del billete que tenga el número correcto de 8 dígitos. Cada billete cuesta un dólar. Si usted ha comprado un billete, ¿cuál es la probabilidad de que gane la lotería? Supongamos que existe solamente un billete por cada número posible de 8 dígitos y que el número ganador es elegido por medio de un proceso aleatorio. otra
23. A partir de una población de 30 bats de béisbol, 15 guantes y 60 pelotas, si se realiza un muestreo aleatorio, tomando uno a la vez y sin reemplazo,
- ¿Cuál es la probabilidad de obtener un guante si se extrae un objeto de la población?
 - ¿Cuál es la probabilidad de obtener un bat y un guante, en ese orden, si se extraen dos objetos de la población?
 - ¿Cuál es la probabilidad de obtener un bat, un guante y un bat, en ese orden, si se extraen tres objetos de la población? otra
24. Una distribución de datos es normal, con una media $\mu = 85$ y una desviación estándar $\sigma = 4.6$. Si se extrae al azar un dato de esta distribución, ¿cuál es la probabilidad de que:
- Sea mayor que 96?
 - Esté entre 90 y 97?
 - Sea menor que 88? otra
25. Suponga que los datos de CI de los estudiantes de la universidad en la que usted estudia, tienen una distribución normal, con $\mu = 115$ y $\sigma = 8$. Si usted extrae al azar un dato de esta distribución, ¿cuál es la probabilidad de que:
- Sea mayor que 130?
 - Esté entre 110 y 125?
 - Sea menor que 100? cognitiva
26. Una prueba estandarizada para medir los conocimientos de matemáticas de los alumnos de sexto grado es administrada en toda la nación. Los resultados muestran una distribución normal de los datos, con $\mu = 50$ y $\sigma = 5.8$. Si un dato es extraído aleatoriamente de esta población, ¿cuál es la probabilidad de que:
- Sea mayor que 62?
 - Esté entre 40 y 65?
 - Sea menor que 45? educación
27. Suponga usted que todavía estamos tratando con la población de la pregunta 24. Si en lugar de tomar aleatoriamente una muestra de la población, el dato se hubiera extraído por medio de un procedimiento no aleatorio, ¿se afectaría por eso cualquiera de las respuestas del problema 24, parte **a**, **b** o **c**? Explique. otra
28. Un etólogo está interesado en averiguar cuánto tiempo tarda cierta especie de musaraña en atrapar a su presa. Para eso, 20 veces al día, él introduce una libélula en la jaula de la musaraña y mide el tiempo que transcurre hasta que este pequeño insectívoro captura a la libélula. Después de varios meses de investigación, el etólogo concluye que el tiempo promedio que este roedor tarda en atrapar a su presa es de 30 segundos, con una desviación estándar de 5.5 segundos, y los datos presentan una distribución normal. Con base en el registro anterior, ¿cuál es la probabilidad de que la musaraña:
- Atrape a una libélula en menos de 18 segundos?
 - Tarde en capturar una libélula entre 22 y 45 segundos?
 - Tarde más de 40 segundos en atrapar a una libélula? biológica
29. Un instructor de la escuela de demolición submarina, de la Marina de Guerra de los Estados Unidos, cree haber desarrollado una nueva técnica para permanecer debajo del agua durante más tiempo. El comandante de la escuela le ha otorgado permiso para que pruebe su técnica con un estudiante de su grupo actual, elegido al azar. Co-

mo parte de su examen de certificación, todos los alumnos son sometidos a una prueba para ver cuánto tiempo pueden estar debajo del agua sin un tanque de oxígeno. Los registros existentes muestran que los datos se distribuyen de forma normal, con una media $\mu = 130$ segundos y una desviación estándar $\sigma = 14$ segundos. Si la nueva técnica no produce efectos adicionales, ¿cuál es la probabilidad de que el estudiante elegido al azar permanezca debajo del agua:

- Más de 150 segundos?
- Entre 115 y 135 segundos?
- Menos de 90 segundos? educación

30. Si usted elige de manera aleatoria dos datos, uno a la vez, con reemplazo, de una población formada por los datos 2, 3, 4, 5 y 6, ¿cuál es la probabilidad de que:

- La media de la muestra (\bar{X}) sea igual a 6.0?
- $\bar{X} \geq 5.5$?
- $\bar{X} \leq 2.0$?

Sugerencia: Todas las muestras posibles de tamaño 2 se presentan en la página 163. otra

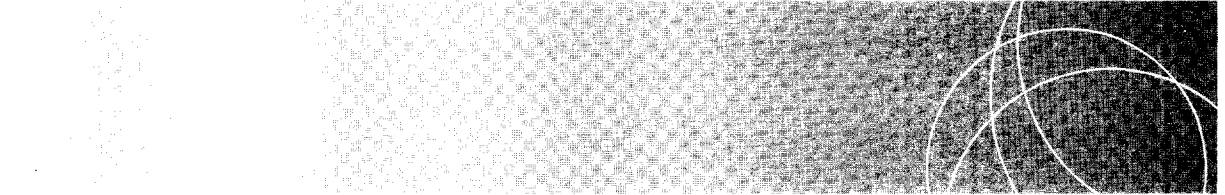
NOTAS

8.1 Observemos que si el proceso garantiza que cada muestra posible de un tamaño dado tenga la misma probabilidad de ser elegida, entonces también garantiza que todos los miembros de la población tengan la misma posibilidad de ser elegidos en la

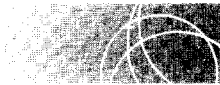
muestra. Si hemos incluido este último enunciado en la definición es porque consideramos que es lo suficientemente importante como para merecer este énfasis especial.

Si desea encontrar problemas de práctica adicionales totalmente resueltos sobre los temas que acaba de estudiar, busque *Chapter 8* en el material adicional del libro, disponible en inglés en: <http://www.thomsonlearning.com.mx>.

Distribución binomial



- Introducción
- Definición e ilustración de la distribución binomial
- Generación de la distribución binomial a partir del desarrollo binomial
- Uso de la tabla binomial
- Resumen
- Términos importantes
- Preguntas y problemas
- Notas



INTRODUCCIÓN

En el capítulo 10 analizaremos el tema de cómo realizar la prueba de hipótesis. Este tema es sumamente importante. Constituye la base de la mayor parte del material que abordaremos en el resto del libro. Por razones que serán explicadas en el capítulo 10, hemos optado por presentar los conceptos de la prueba de hipótesis mediante una prueba de inferencia sencilla llamada la prueba de los signos. Sin embargo, para comprender y utilizar esta última, primero tendremos que analizar una distribución de probabilidad que se conoce como distribución binomial.



DEFINICIÓN E ILUSTRACIÓN DE LA DISTRIBUCIÓN BINOMIAL

La distribución binomial se puede definir en la siguiente forma:

definición ■ *La **distribución binomial** es una distribución de probabilidad que se presenta cuando se cumplen estas cinco condiciones: 1) existe una serie de N ensayos; 2) en cada ensayo hay sólo dos resultados posibles; 3) en cada ensayo, los dos resultados posibles son mutuamente excluyentes; 4) los resultados de cada ensayo son independientes entre sí y 5) la probabilidad de obtener cada resultado posible en cualquier ensayo es la misma de un ensayo a otro. Cuando se cumplen estas condiciones, la distribución binomial proporciona todos los resultados posibles de los N ensayos y la probabilidad de obtener cada uno de estos resultados.*

Tomaremos el lanzamiento de una moneda como el medio para ilustrar la generación de una distribución binomial. Suponga que lanzamos una moneda no cargada una sola vez. Además, restringiremos los posibles resultados al final de cada lanzamiento a sólo dos: cara o cruz. Usted recordará, por el capítulo 8, que cuando una moneda no está cargada la probabilidad de obtener cara es igual que la probabilidad de obtener cruz. Como únicamente existen dos posibles resultados en cada lanzamiento,

$$\begin{aligned} p(\text{cara}) &= p(\text{cara}) = \frac{\text{Número de resultados clasificables como caras}}{\text{Cantidad total de resultados}} \\ &= \frac{1}{2} = 0.5000 \end{aligned}$$

$$\begin{aligned} p(\text{cruz}) &= p(\text{cruz}) = \frac{\text{Número de resultados clasificables como cruces}}{\text{Cantidad total de resultados}} \\ &= \frac{1}{2} = 0.5000 \end{aligned}$$

Suponga que ahora decidimos lanzar dos monedas que no están cargadas. El lanzamiento de cada moneda se considerará como un ensayo. Así, pues, con las dos monedas tendremos dos ensayos ($N = 2$). Los resultados posibles del lanzamiento de dos monedas se presentan en la tabla 9.1. Existen cuatro resultados posibles: uno en el cual se obtienen dos caras (renglón 1), dos en los que se obtienen una cara y una cruz (renglones 2 y 3) y uno en el cual se obtienen 2 cruces (renglón 4).

A continuación determinaremos la probabilidad de obtener cada uno de estos resultados, debido al azar. Si sólo interviene el azar, entonces cada uno de los resultados es igualmente probable. Así,

$$\begin{aligned} p(2 \text{ caras}) &= p(\text{cara cara}) = \frac{\text{Número de resultados clasificables como 2 caras}}{\text{Cantidad total de resultados}} \\ &= \frac{1}{4} = 0.2500 \end{aligned}$$

$$\begin{aligned} p(1 \text{ cara}) &= p\left(\begin{array}{l} \text{cara cruz o} \\ \text{cruz cara} \end{array}\right) = \frac{\text{Número de resultados clasificables como 1 cara}}{\text{Cantidad total de resultados}} \\ &= \frac{2}{4} = 0.5000 \end{aligned}$$

$$\begin{aligned} p(0 \text{ caras}) &= p(\text{cruz cruz}) = \frac{\text{Número de resultados clasificables como 0 caras}}{\text{Cantidad total de resultados}} \\ &= \frac{1}{4} = 0.2500 \end{aligned}$$

tabla 9.1 Todos los resultados posibles al lanzar dos monedas una sola vez

Renglón número	Moneda 1	Moneda 2	Número de resultados
1	<i>Cara</i>	<i>Cara</i>	1
2	<i>Cara</i>	<i>Cruz</i>	2
3	<i>Cruz</i>	<i>Cara</i>	
4	<i>Cruz</i>	<i>Cruz</i>	1
Total de resultados			4

Observe que también podríamos haber obtenido estas probabilidades a partir de las reglas de la suma y el producto. Por ejemplo, podríamos haber determinado $p(1 \text{ cara})$ a partir de una combinación de dichas reglas en la siguiente forma:

$$p(1 \text{ cara}) = p(\text{cara cruz o cruz cara})$$

Con la regla del producto obtenemos

$$\begin{aligned} p(\text{cara cruz}) &= p(\text{cara en la primera moneda y cruz en la segunda moneda}) \\ &= p(\text{cara en la primera moneda})p(\text{cruz en la segunda moneda}) \\ &= \frac{1}{2}\left(\frac{1}{2}\right) = \frac{1}{4} \end{aligned}$$

$$\begin{aligned} p(\text{cruz cara}) &= p(\text{cruz en la primera moneda y cara en la segunda moneda}) \\ &= p(\text{cruz en la primera moneda})p(\text{cara en la segunda moneda}) \\ &= \frac{1}{2}\left(\frac{1}{2}\right) = \frac{1}{4} \end{aligned}$$

Aplicando la regla de la suma obtenemos

$$\begin{aligned} P(1 \text{ cara}) &= p(\text{cara cruz o cruz cara}) \\ &= p(\text{cara cruz}) + p(\text{cruz cara}) \\ &= \frac{1}{4} + \frac{1}{4} = \frac{2}{4} = 0.5000 \end{aligned}$$

A continuación, suponga que aumentamos N de 2 a 3. Los resultados posibles al lanzar tres monedas no cargadas se presentan en la tabla 9.2. Esta vez tenemos ocho resultados posibles: una forma de obtener 3 caras (renglón 1), tres maneras de conseguir 2 caras y una cruz (renglones 2, 3 y 4), tres formas de obtener 1 cara y 2 cruces (renglones 5, 6 y 7) y una manera de conseguir 0 caras (renglón 8). En virtud de que todos los resultados son igualmente probables,

$$p(3 \text{ caras}) = \frac{1}{8} = 0.1250$$

$$p(2 \text{ caras}) = \frac{3}{8} = 0.3750$$

$$p(1 \text{ cara}) = \frac{3}{8} = 0.3750$$

$$p(0 \text{ caras}) = \frac{1}{8} = 0.1250$$

Las distribuciones producidas por el lanzamiento de una, dos o tres monedas se muestran en la tabla 9.3. Éstas son distribuciones binomiales, pues son distribuciones de probabilidad generadas por una situación donde hay una serie de ensayos ($N = 1, 2$ o 3), de manera que en cada ensayo sólo existen dos resultados posibles (cara o cruz); los resultados son mutuamente excluyentes (si es cara, no puede ser cruz); los ensayos inde-

t a b l a 9.2 Todos los resultados posibles al lanzar tres monedas una sola vez

Renglón número	Moneda 1	Moneda 2	Moneda 3	Número de resultados
1	<i>Cara</i>	<i>Cara</i>	<i>Cara</i>	1
2	<i>Cara</i>	<i>Cara</i>	<i>Cruz</i>	3
3	<i>Cara</i>	<i>Cruz</i>	<i>Cara</i>	
4	<i>Cruz</i>	<i>Cara</i>	<i>Cara</i>	
5	<i>Cruz</i>	<i>Cruz</i>	<i>Cara</i>	3
6	<i>Cruz</i>	<i>Cara</i>	<i>Cruz</i>	
7	<i>Cara</i>	<i>Cruz</i>	<i>Cruz</i>	
8	<i>Cruz</i>	<i>Cruz</i>	<i>Cruz</i>	1
Total de resultados				8

t a b l a 9.3 Distribución binomial para el lanzamiento de monedas, cuando el número de monedas es igual a 1, 2 o 3

<i>N</i>	Resultados posibles	Probabilidad
1	1 <i>Cara</i>	0.5000
	0 <i>Caras</i>	0.5000
2	2 <i>Caras</i>	0.2500
	1 <i>Cara</i>	0.5000
	0 <i>Caras</i>	0.2500
3	3 <i>Caras</i>	0.1250
	2 <i>Caras</i>	0.3750
	1 <i>Cara</i>	0.3750
	0 <i>Caras</i>	0.1250

pendientes (existe independencia entre los resultados de cada moneda) y la probabilidad de que se obtenga cara o cruz en cualquier ensayo es exactamente la misma de un ensayo a otro. Observe que cada distribución proporciona la siguiente información: 1) todos los resultados posibles de los N ensayos y 2) la probabilidad de obtener cada uno de los resultados..



GENERACIÓN DE LA DISTRIBUCIÓN BINOMIAL A PARTIR DEL DESARROLLO BINOMIAL

Podríamos continuar con este proceso de enumeración para valores mayores que N , pero la tarea resultaría demasiado laboriosa. Sería poco agradable tener que utilizar una enumeración como ésta para cada valor de N . Piense usted lo que pasaría si N llegara a ser 15. Con 15 monedas existen $(2)^{15} = 32\,768$ formas distintas en las que podrían caer. Por fortuna, contamos con una expresión matemática que nos permite generar de una manera sencilla todo lo que hemos considerado hasta ahora. Dicha expresión se conoce como el *desarrollo binomial*. El desarrollo binomial está determinado por:

$$(P + Q)^N \quad \text{desarrollo binomial}$$

donde P = es la probabilidad de uno de los dos resultados posibles en un ensayo
 Q = es la probabilidad del otro resultado posible
 N = es el número de ensayos

Para generar los resultados posibles y las probabilidades asociadas a las que hemos llegado en los anteriores experimentos con lanzamiento de monedas, lo único que necesitamos hacer es desarrollar la expresión $(P + Q)^N$ para el número de monedas que haya en el experimento y luego evaluar cada término del desarrollo. Por ejemplo, si hay dos monedas, $N = 2$ y

$$\begin{array}{ccc} \text{Eventos } 2P & \text{Evento } 1P \text{ y } 1Q & \text{Eventos } 2Q \\ \swarrow & \downarrow & \swarrow \\ (P + Q)^N = (P + Q)^2 = P^2 + 2P^1Q^1 + Q^2 \end{array}$$

Los términos P^2 , $2P^1Q^1$ y Q^2 representan todos los resultados posibles al lanzar dos monedas una sola vez.

Las letras de cada término (P o PQ o Q) nos indican los tipos de eventos que constituyen el resultado; el exponente de cada letra nos dice cuántos de ese tipo de evento hay en el resultado, y el coeficiente de cada término nos señala cuántas formas existen de obtener el resultado.

Así, pues,

1. P^2 indica que uno de los resultados posibles está formado por dos eventos P . Una P sola indica que el resultado en cuestión se compone únicamente de eventos P . El exponente 2 señala que existen dos eventos de este tipo. Si asociamos P con las caras, P^2 nos indica que dos caras son un resultado posible.
2. $2P^1Q^1$ indica que otro resultado posible es un evento P y un evento Q , o una cara y una cruz. El coeficiente 2 indica que existen dos formas de obtener un evento P y un evento Q .
3. Q^2 representa un resultado de dos eventos Q , o dos cruces (cero caras).

La probabilidad de obtener cada uno de estos resultados posibles se determina evaluando sus términos respectivos mediante los valores numéricos de P y Q . Si las monedas no están cargadas, entonces $P = Q = 0.50$. Así,

$$\begin{aligned} p(2 \text{ caras}) &= P^2 = (0.50)^2 = 0.2500 \\ p(1 \text{ cara}) &= 2P^1Q^1 = 2(0.50)(0.50) = 0.5000 \\ p(0 \text{ caras}) &= Q^2 = (0.50)^2 = 0.2500 \end{aligned}$$

Estos resultados son iguales a los obtenidos mediante la enumeración. Observe que al utilizar el desarrollo binomial para determinar la probabilidad de cada resultado posible no sumamos los términos, sino que los utilizamos por separado. En este momento parecería más fácil utilizar la enumeración que el desarrollo binomial. Sin embargo, la situación cambia rápidamente si N es cada vez mayor.

Presentemos otro ejemplo, esta vez con $N = 3$. Igual que antes, necesitamos desarrollar $(P + Q)^N$ y evaluar cada término del desarrollo utilizando $P = Q = 0.50$.⁶ Así,

$$(P + Q)^N = (P + Q)^3 = P^3 + 3P^2Q + 3PQ^2 + Q^3$$

Los términos P^3 , $3P^2Q$, $3PQ^2$ y Q^3 representan todos los resultados posibles cuando se lanzan tres monedas una sola vez. P^3 nos indica que hay tres eventos P o 3 caras. El término $3P^2Q$ indica que este resultado tiene dos eventos P y un evento Q , es decir, 2 caras y 1 cruz. El término $3PQ^2$ representa un evento P y dos eventos Q , es decir, 1 cara y

⁶ Véase en la nota 9.1 la ecuación general del desarrollo de $(P + Q)^N$.

2 cruces. Por último, el término Q^3 señala tres eventos Q y cero eventos P , es decir, 3 cruces y 0 caras. Podemos determinar la probabilidad de cada uno de estos resultados evaluando sus términos correspondientes. Puesto que ninguna de las monedas está cargada, $P = Q = 0.50$. En consecuencia,

$$p(3 \text{ caras}) = P^3 = (0.50)^3 = 0.1250$$

$$p(2 \text{ caras}) = 3P^2Q = 3(0.50)^2(0.50) = 0.3750$$

$$p(1 \text{ cara}) = 3PQ^2 = 3(0.50)(0.50)^2 = 0.3750$$

$$p(0 \text{ caras}) = Q^3 = (0.50)^3 = 0.1250$$

Éstos son los mismos resultados que obtuvimos previamente mediante la enumeración.

Podemos generar la distribución binomial para cualquier N , P y Q mediante el desarrollo binomial. En la figura 9.1 graficamos las distribuciones binomiales para $N = 3$, 8 y 15. $P = Q = 0.50$ para cada una de estas distribuciones.

Observe que 1) con $P = 0.50$, la distribución binomial es simétrica; 2) tiene dos colas (es decir, se va reduciendo a medida que nos desplazamos desde el centro hacia cualquier extremo); 3) implica una variable discreta (por ejemplo, no podemos tener $2\frac{1}{2}$ caras), y 4) cuando N aumenta, la distribución binomial se aproxima en su forma a la de una curva normal.



USO DE LA TABLA BINOMIAL

Aun cuando, en principio, cualquier problema relacionado con datos binomiales se puede resolver mediante la sustitución directa en el desarrollo binomial, los matemáticos nos han ahorrado todo ese trabajo. Ellos han resuelto el desarrollo binomial para muchos valores de N y los resultados han quedado registrados en tablas. Una de estas últimas es la tabla B del apéndice D. Esa tabla proporciona la distribución binomial para valores de N hasta 20. Al observar la tabla B, notará que N (el número de ensayos) se presenta en la primera columna y que los resultados posibles están en la segunda columna, bajo el encabezado "Número de eventos P o Q ". El resto de las columnas contienen datos de probabilidad para diferentes valores de P o Q . Los valores de P o Q se muestran en la parte superior de cada columna. Así, la segunda columna contiene valores de probabilidad para P o $Q = 0.10$ y la última columna muestra los valores correspondientes a P o $Q = 0.50$. En la práctica, cualquier problema relacionado con datos binomiales se puede resolver consultando la probabilidad adecuada en esta tabla. Esto, por supuesto, sólo se aplica a $N \leq 20$ y los valores de P o Q presentados en la tabla.

Será conveniente que el lector observe que la tabla B se puede emplear para resolver problemas en términos de P o de Q . Así, con excepción de la primera columna, los encabezados de las columnas están expresados en términos de P o Q . Para enfatizar cuál de ellos estamos utilizando (P o Q) en un problema dado, si consultamos la tabla bajo P y el número de eventos P , nos referiremos a la segunda columna como el "número de eventos P " y a los demás encabezados de columna como valores de probabilidad " P ". Si consultamos la tabla B bajo Q y el número de eventos Q , nos referiremos a la segunda columna como el "número de eventos Q " y a los demás encabezados de columna como valores de probabilidad " Q ". A continuación veremos la forma de utilizar esta tabla para resolver problemas en los que intervienen situaciones binomiales.

ejemplo : Si lanzo una sola vez tres monedas que no estén cargadas, ¿cuál será la probabilidad de obtener 2 caras y una cruz? Suponga que cada moneda sólo puede caer de cara o de cruz.

SOLUCIÓN

En este problema, N es el número de monedas, que es igual a 3. Denotamos por P la probabilidad de obtener una cara al lanzar una vez cualquier moneda. Las monedas no están cargadas,

por lo tanto, $P = 0.50$. Como queremos determinar la probabilidad de obtener 2 caras, el número de eventos P es igual a 2. Una vez determinado lo anterior, lo único que debemos hacer es consultar la parte de la tabla B bajo $N = 3$. A continuación, localizamos el 2 en la columna número de eventos P . La respuesta aparece donde se cruza el renglón que contiene al 2 con la columna cuyo encabezado es $P = 0.50$. Así,

$$p(2 \text{ caras y } 1 \text{ cruz}) = 0.3750$$

tabla 9.4 Datos de la tabla B

N	Número de eventos P	P 0.50
3	2	0.3750

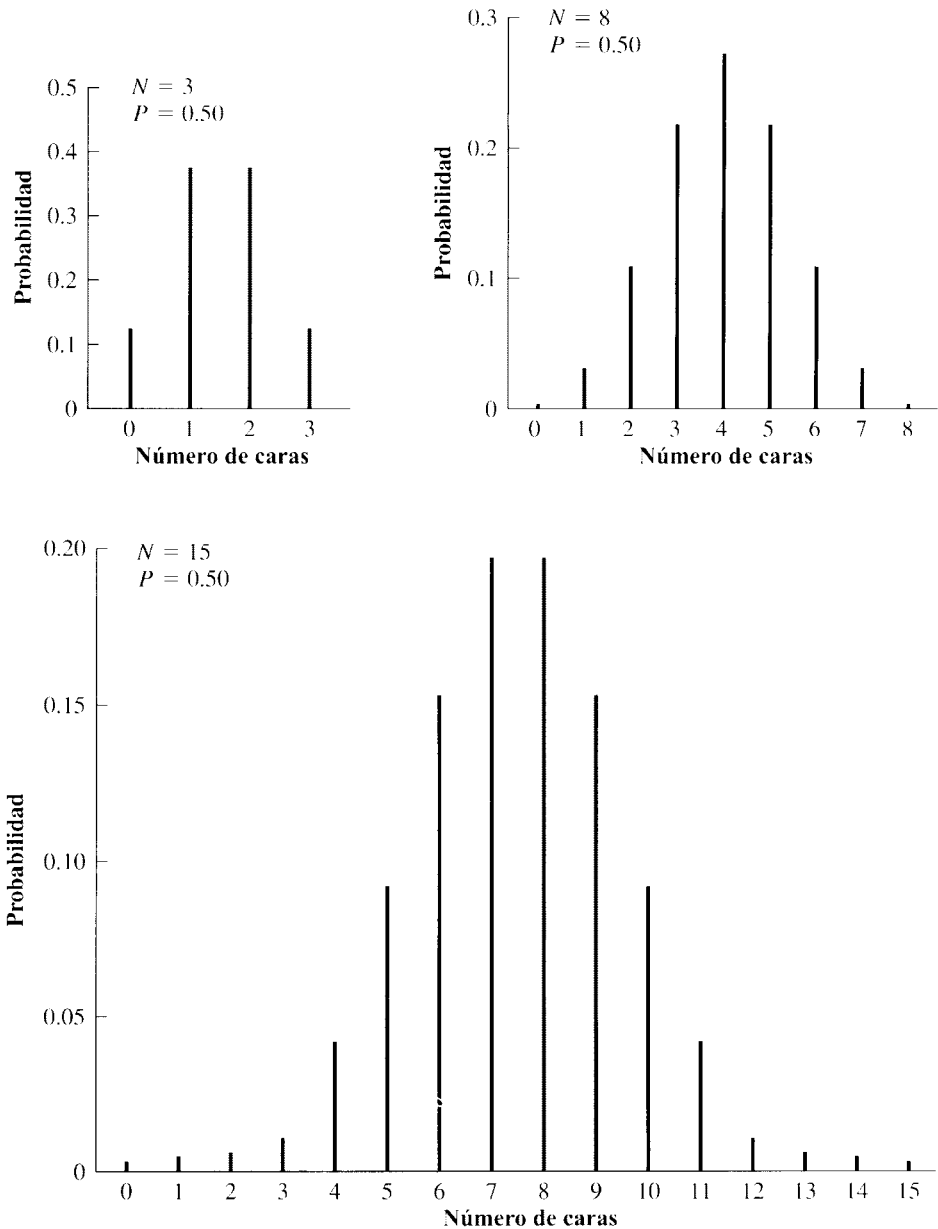


figura 9.1 Distribución binomial para $N = 3$, $N = 8$ y $N = 15$; $P = 0.50$.

Observe que ésta es la misma respuesta que habíamos obtenido antes. De hecho, si examina usted los demás datos de esa columna ($P = 0.50$) para $N = 2$ y 3 , comprobará que son los mismos que habíamos obtenido mediante el desarrollo binomial (y así debería ser, puesto que los datos de la tabla fueron calculados aplicando el desarrollo binomial). Realicemos ahora algunos problemas de práctica con ayuda de esta tabla.

Si lanzamos una vez seis monedas no cargadas, ¿cuál es la probabilidad de obtener:

- Exactamente 6 caras?
- 4, 5 o 6 caras?

SOLUCIÓN

- En vista de que existen seis monedas, $N = 6$. Una vez más, sea P la probabilidad de obtener una cara en el lanzamiento de cualquier moneda. Puesto que las monedas no están cargadas $P = 0.50$. Como queremos conocer la probabilidad de obtener exactamente 6 caras, el número de eventos P es igual a 6. Al consultar la tabla B bajo $N = 6$, número de eventos $P = 6$ y $P = 0.50$, tenemos que:

Datos de la tabla B			
$p(\text{exactamente } 6 \text{ caras}) = 0.0156$	N	Número de eventos P	P 0.50
	6	6	0.0156

- De nuevo, $N = 6$ y $P = 0.50$. Podemos determinar la probabilidad de 4, 5 y 6 caras consultando la tabla B bajo el número de eventos $P = 4, 5$ y 6 , respectivamente. De esta manera,

Datos de la tabla B			
	N	Número de eventos P	P 0.50
$p(4 \text{ caras}) = 0.2344$	6	4	0.2344
$p(5 \text{ caras}) = 0.0938$		5	0.0938
$p(6 \text{ caras}) = 0.0156$		6	0.0156

Según la regla de la suma para eventos mutuamente excluyentes,

$$\begin{aligned}
 p(4, 5 \text{ o } 6 \text{ caras}) &= p(4) + p(5) + p(6) \\
 &= 0.2344 + 0.0938 + 0.0156 \\
 &= 0.3438
 \end{aligned}$$



Si lanzamos una sola vez 10 monedas no cargadas, ¿cuál es la probabilidad de obtener un resultado tan extremo o más extremo que 9 caras?

SOLUCIÓN

Considerando que existen 10 monedas, $N = 10$. Igual que antes, sea P la probabilidad de obtener una cara en un solo lanzamiento de cualquier moneda. Como las monedas no están cargadas, $P = Q = 0.50$. La frase “tan extremo o más extremo que” significa “tan lejano o más lejano del centro de la distribución como”. Así, “tan extremo o más extremo que 9 caras” se refiere a los resultados que están tan lejanos o más del centro de la distribución como lo están 9 caras. Así, pues, el número de eventos P es 0, 1, 9 o 10. En la tabla B, bajo $N = 10$, número de eventos $P = 0, 1, 9$ o 10 y $P = 0.50$, tenemos

$$\begin{aligned}
 p \left(\begin{array}{l} \text{tan extremo o} \\ \text{más extremo} \\ \text{que 9 caras} \end{array} \right) &= p(0, 1, 9, \text{ o } 10) \\
 &= p(0) + p(1) + p(9) + p(10) \\
 &= 0.0010 + 0.0098 + 0.0098 \\
 &\quad + 0.0010 \\
 &= 0.0216
 \end{aligned}$$

Datos de la tabla B

N	Número de eventos P	P 0.50
10	0	0.0010
	1	0.0098
	9	0.0098
	10	0.0010

El desarrollo binomial es sumamente general. No se limita a los valores en los que $P = 0.50$. Así, pues, la tabla B enumera las probabilidades que corresponden a valores de P diferentes de 0.50. Tratemos de resolver algunos problemas en los que P no es igual a 0.50.



Suponga usted que tiene ocho monedas cargadas. Cada moneda tiene un peso tal que la probabilidad de obtener una cara con ella es de 0.30. Si las ocho monedas cargadas se lanzan una vez, ¿cuál es la probabilidad de obtener:

- a. 7 caras?
- b. 7 u 8 caras?
- c. ¿Cuál o cuáles de los términos del siguiente desarrollo binomial nos permite(n) evaluar la probabilidad determinada en la parte a?

$$P^8 + 8P^7Q^1 + 28P^6Q^2 + 56P^5Q^3 + 70P^4Q^4 + 56P^3Q^5 + 28P^2Q^6 + 8P^1Q^7 + Q^8$$

(continúa)

- d. Utilice su calculadora para evaluar el término o términos seleccionados en la parte c, con $P = 0.30$. Compare su respuesta con la de la parte a. Explique.

SOLUCIÓN

- a. En virtud de que existen ocho monedas, $N = 8$. Sea P la probabilidad de obtener una cara en un solo lanzamiento de cualquier moneda. Considerando que las monedas están cargadas de tal modo que la probabilidad de que resulte una cara cuando se lanza cualquier moneda es de 0.30, concluimos que $P = 0.30$. Puesto que queremos determinar la probabilidad de obtener exactamente 7 caras, el número de eventos P es igual a 7. En la tabla B, debajo de $N = 8$, número de eventos $P = 7$ y $P = 0.30$, encontraremos los siguientes datos:

$p(\text{exactamente 7 caras}) = 0.0012$

N	Número de eventos P	P 0.30
8	7	0.0012

- b. Una vez más, $N = 8$ y $P = 0.30$. Podemos determinar la probabilidad de obtener 7 y 8 caras, consultando la tabla B, debajo del número de eventos $P = 7$ y 8, respectivamente. Así,

$p(7 \text{ caras}) = 0.0012$
 $p(8 \text{ caras}) = 0.0001$

N	Número de eventos P	P 0.30
8	7	0.0012
	8	0.0001

Según la regla de la suma para eventos mutuamente excluyentes,

$$\begin{aligned} p(7 \text{ u } 8 \text{ caras}) &= p(7) + p(8) \\ &= 0.0012 + 0.0001 \\ &= 0.0013 \end{aligned}$$

- c. $8P^7Q^1$

- d. $8P^7Q^1 = 8(0.30)^7(0.7) = 0.0012$. Como era de esperarse, las respuestas son las mismas. Los datos de la tabla fueron calculados utilizando $8P^7Q^1$ con $P = 0.30$ y $Q = 0.70$.

Así, el uso de la tabla B en situaciones en las que P es menor que 0.50 es muy similar a los casos en los que $P = 0.50$. Solamente que ahora buscamos en la tabla el nuevo valor P , en lugar de consultar bajo $P = 0.50$.

La tabla B se puede utilizar también en situaciones en las que $P > 0.50$. Para ilustrar este punto, consideremos el siguiente ejemplo.

ejemplo

($P > 0.50$)

Si se lanzan cinco monedas cargadas una sola vez, ¿cuál es la probabilidad de obtener a) 5 caras y b) 4 o 5 caras? A cada moneda se le ha agregado un peso determinado que hace que la probabilidad de que se obtenga una cara al lanzar cualquier moneda sea de 0.75.

SOLUCIÓN

a. 5 caras. Dado que se trata de cinco monedas $N = 5$. Una vez más, sea P la probabilidad de obtener una cara en un solo lanzamiento de cualquier moneda. En vista de que las monedas están cargadas, la probabilidad de obtener una cara al lanzar una vez cualquier moneda es de 0.75, $P = 0.75$. Como queremos determinar la probabilidad de obtener 5 caras, el número de eventos P es igual a 5. Si aplicáramos nuestro procedimiento de costumbre, tendríamos que consultar la tabla B, bajo $N = 5$, número de eventos $P = 5$ y $P = 0.75$. Sin embargo, la tabla B no tiene ninguna columna encabezada por 0.75. Todos los encabezados de columna son iguales o menores que 0.50. Sin embargo, a pesar de todo, podemos utilizar la tabla B para resolver este problema.

Cuando $P > 0.50$, lo que tenemos que hacer es resolver el problema en términos de Q y el número de eventos Q , en lugar de hacerlo en términos de P y el número de eventos P . Como los valores de probabilidad presentados en la tabla B son válidos para P o para Q , una vez que hayamos planteado el problema en términos de Q , podremos consultar la tabla B utilizando Q en vez de P . La traducción del problema a términos de Q requiere dos pasos: la determinación de Q y el cálculo del número de eventos Q . Sigamos estos pasos con nuestro ejemplo actual:

1. *Determinación de Q .*

$$Q = 1 - P = 1 - 0.75 = 0.25$$

2. *Determinación del número de eventos Q .*

$$\text{Número de eventos } Q = N - \text{Número de eventos } P = 5 - 5 = 0$$

Por lo tanto, para resolver este ejemplo, consultamos la tabla B bajo $N = 5$, número de eventos $Q = 0$ y $Q = 0.25$. Así,

tabla 9.5 Datos de la tabla B

	Número de eventos Q	Q 0.25
$p(5 \text{ caras}) = p(0 \text{ cruces}) = 0.2373$	5	0
		0.2373

b. 4 o 5 caras. De nuevo, $N = 5$ y $Q = 0.25$. Esta vez, el número de eventos Q es 0 o 1. Así, pues,

tabla 9.6 Datos de la tabla B

	Número de eventos Q	Q 0.25
$p(4 \text{ o } 5 \text{ caras}) = p(0 \text{ o } 1 \text{ cruz})$	5	0
$= 0.2373 + 0.3955$		0.2373
$= 0.6328$		0.3955

Ahora ya estamos preparados para tratar de resolver un problema de práctica.



Si se lanzan una vez 12 monedas cargadas, ¿cuál es la probabilidad de obtener:

- 10 caras exactamente?
- 10 caras o más?

Las monedas están cargadas de tal manera que la probabilidad de que resulte una cara en cualquier moneda es igual a 0.65.

SOLUCIÓN

- En vista de que existen 12 monedas, $N = 12$. Sea P la probabilidad de obtener una cara en un lanzamiento de cualquier moneda. Como quiera que la probabilidad de que resulte una cara al lanzar cualquier moneda es igual a 0.65, $P = 0.65$. Como $P > 0.50$, consultamos la tabla B utilizando Q en lugar de P . Si existen 10 eventos P , deberán existir 2 eventos Q ($N = 12$). Si $P = 0.65$, entonces $Q = 0.35$. Al utilizar Q en la tabla B, obtenemos

Datos de la tabla B		
	Número de eventos Q	Q 0.35
$p(10 \text{ caras}) = 0.1088$	12	2 0.1088

- Nuevamente, $N = 12$ y $P = 0.65$. Esta vez, el número de eventos P es igual a 10, 11 o 12. Puesto que $P > 0.50$, deberemos emplear Q en vez de P en la tabla B. Con $N = 12$, el número de eventos Q es igual a 0, 1 o 2 y $Q = 0.35$. Al utilizar Q en la tabla B, obtenemos

Datos de la tabla B		
	Número de eventos Q	Q 0.35
$p(10, 11 \text{ o } 12 \text{ caras}) = 0.1088 + 0.0368$	12	0 0.0057
+ 0.0057		1 0.0368
= 0.1513		2 0.1088

Hasta ahora, hemos trabajado solamente con lanzamientos de monedas. Sin embargo, la distribución binomial no se limita a este tipo de eventos, sino que se aplica a todas las situaciones que implican una serie de ensayos, cuando en cada ensayo sólo existen dos resultados posibles: los resultados posibles en cada ensayo son mutuamente excluyentes, existe independencia entre los resultados de cada ensayo, y la probabilidad de cada uno de los resultados posibles en cualquier ensayo sigue siendo la misma de un ensayo a otro. Existen muchas situaciones que satisfacen estas condiciones. Para ilustrar esto, resolveremos un par de problemas de práctica.



Una estudiante presenta un examen de opción múltiple con 15 preguntas. Cada pregunta propone cinco opciones. Si la estudiante adivina la respuesta en cada pregunta, ¿cuál es la probabilidad de que apruebe el examen? La mínima calificación aprobatoria requiere que 60% de las preguntas sean contestadas en forma correcta. Suponga que todas las opciones de cada pregunta son igualmente probables.

SOLUCIÓN

Este problema se ajusta a las condiciones binomiales. Tenemos una serie de ensayos (preguntas). En cada ensayo, sólo hay dos resultados posibles. La estudiante está en lo correcto o está equivocada. Los resultados son mutuamente excluyentes. Si ella tiene una respuesta correcta, entonces no la puede tener incorrecta. Existe independencia entre los resultados de todos los ensayos. Si la estudiante acierta en la pregunta 1, esto no produce efecto alguno en el resultado de la pregunta 2. Por último, si suponemos que la estudiante adivina la respuesta en cada uno de los ensayos, entonces las probabilidades de que ésta sea correcta o incorrecta son las mismas de un ensayo a otro. En estas condiciones podemos aplicar la distribución binomial y la tabla B.

Podemos considerar cada pregunta como un ensayo. Por el hecho de que hay 15 preguntas, $N = 15$. Sea P la probabilidad de que la estudiante acierte al adivinar cualquier pregunta. Como existen cinco opciones igualmente probables para cada pregunta, $P = 0.20$. Una calificación aprobatoria implica 60% de las respuestas acertadas o más. Por lo tanto, la estudiante será aprobada si reúne 9 o más respuestas correctas (60% de 15 es 9). Así, pues, el número de eventos P es igual a 9, 10, 11, 12, 13, 14 y 15. Cuando buscamos en la tabla B debajo de $N = 15$, número de eventos $P = 9, 10, 11, 12, 13, 14$ y 15 , y $P = 0.20$, obtenemos lo siguiente:

Datos de la tabla B		
N	Número de eventos P	P 0.20
$p(9, 10, 11, 12, 13, 14 \text{ o } 15$ respuestas correctas) = 0.0007 + 0.0001 = 0.0008	15	9 0.0007 10 0.0001 11 0.0000 12 0.0000 13 0.0000 14 0.0000 15 0.0000

Una amiga de usted afirma que es una experta en materia de café. Ella siempre bebe Starbucks y afirma que ningún otro café se aproxima siquiera a su excelente sabor. Usted sospecha que ella es un poco presuntuosa. De hecho, usted se pregunta si en verdad puede distinguir la diferencia de sabor entre Starbucks y el café tostado en la localidad. Su amiga accede a realizar el siguiente experimento. Con los ojos vendados, ella tendrá seis oportunidades para probar el contenido de dos tazas de café y tratar de decir cuál de ellas contiene el de Starbucks. Las tazas son idénticas y contienen el mismo tipo de café, salvo que una de ellas contiene café de grano tostado y vendido por Starbucks y la otra tiene café de grano del proveedor local. Cada vez que ella prueba de las dos tazas, usted elimina cualquier posible pista y elige al azar la taza que le dará a probar primero en el siguiente ensayo. Por increíble que le parezca, su amiga consigue identificar correctamente el Starbucks en los seis ensayos. ¿Qué conclusión saca usted de esto? ¿Se le ocurre alguna manera de aumentar su nivel de confianza en esa conclusión?

SOLUCIÓN

La lógica de nuestro análisis es la siguiente. Supondremos que su amiga no es capaz de distinguir en realidad la diferencia entre las dos marcas de café. Entonces, con seguridad adivinó la respuesta en cada uno de los ensayos. Calcularemos la probabilidad de obtener seis respuestas correctas, suponiendo que la persona trata de adivinarla en cada ensayo. Si esta probabilidad resulta demasiado baja, rechazaremos la explicación de que haya adivinado y concluiremos que su amiga en verdad sabe apreciar la diferencia.

Este experimento se ajusta a las condiciones de la distribución binomial. Cada intento de comparación entre los dos cafés se puede considerar como un ensayo. En cada ensayo, sólo hay dos resultados posibles: o su amiga está en lo correcto o está equivocada. Los resultados son mutuamente excluyentes. Existe independencia entre los ensayos. Si su amiga está en lo correcto en el primer ensayo, eso no tendrá efecto alguno sobre el resultado del segundo ensayo. Por último, si suponemos que su amiga trata de adivinar la respuesta en todos los ensayos, entonces la probabilidad de que acierte o se equivoque en cada uno es la misma de un ensayo a otro.

Como quiera que cada comparación de los cafés es un ensayo, $N = 6$. Sea P la probabilidad de que su amiga acierte en cualquiera de los ensayos. Como sólo existen dos cafés, $P = 0.50$. Su amiga acertó en los seis ensayos. Por lo tanto, el número de eventos P es 6. Así, pues,

Datos de la tabla B		
N	Número de eventos P	P 0.50
6	6	0.0156

$$p(6 \text{ aciertos}) = 0.0156$$

Si su amiga hubiera tratado de adivinar las respuestas, la probabilidad de que hubiera obtenido seis respuestas correctas habría sido de 0.0156. Como éste es un valor demasiado bajo y representaría una explicación poco razonable, es probable que usted lo rechace y concluya que su amiga verdaderamente percibe la diferencia. Para aumentar su confianza en el rechazo a la idea de que su amiga acertó por adivinación, usted podría incluir más marcas de café en cada ensayo, o podría aumentar el número de ensayos. Por ejemplo, aunque sólo se trate de dos marcas de café, la probabilidad de adivinar correctamente 12 de los 12 ensayos es de 0.0002.

RESUMEN

En este capítulo hemos analizado la distribución binomial, la cual es una distribución de probabilidad que se presenta cuando las siguientes condiciones se cumplen: 1) existe una serie de N ensayos; 2) en cada ensayo hay sólo dos posibles resultados; 3) en cada ensayo, los dos resultados posibles son mutuamente excluyentes; 4) los resultados de cada ensayo son independientes entre sí, y 5) la probabilidad de cada resultado posible en cada uno de los ensayos es la misma de un ensayo a otro. Cuando se cumplen estas condiciones, la distribución binomial nos proporciona todos los resultados posibles de los N ensayos y la probabilidad de obtener cada uno de estos resultados.

Ilustramos la distribución binomial por medio de experimentos con lanzamiento de monedas y luego mostramos la forma de generar esta distribución mediante el desarrollo binomial. Este desarrollo está dado por $(P + Q)^N$, donde P es la probabilidad de que ocurra uno de los eventos y Q es la probabilidad de que se presente el otro. Por último, mostramos la manera de utilizar el valor binomial (tabla B del apéndice D) para resolver problemas relacionados con la distribución binomial. Esta distribución resulta adecuada siempre que se cumplan las cinco condiciones enumeradas al principio de este resumen.

TÉRMINOS IMPORTANTES

Desarrollo binomial (página 200)

Distribución binomial (página 197)

Monedas cargadas (página 205)

Monedas no cargadas (página 198)

Número de eventos P (página 201)

Número de eventos Q (página 201)

Tabla binomial (página 197)

PREGUNTAS Y PROBLEMAS

- Defina o explique brevemente cada uno de los conceptos mencionados en la sección "Términos importantes".
- ¿Cuáles son las cinco condiciones necesarias para que la distribución binomial resulte apropiada?
- En una situación binomial, si $P = 0.10$, $Q = \dots$.
- Utilizando la tabla B, si $N = 6$ y $P = 0.40$,
 - La probabilidad de obtener exactamente cinco eventos P es de $= \dots$.
 - ¿Cuáles son los términos de la siguiente ecuación que nos sirven para evaluar la probabilidad anterior?

$$P^6 + 6P^5Q + 15P^4Q^2 + 20P^3Q^3 + 15P^2Q^4 + 6PQ^5 + Q^6$$
 - Evalúe el o los términos de su respuesta para la parte **b** utilizando $P = 0.40$ y compare su respuesta con la parte **a**.
- Utilizando la tabla B, con $N = 12$ y $P = 0.50$,
 - ¿Cuál es la probabilidad de obtener exactamente 10 eventos P ?
 - ¿Cuál es la probabilidad de conseguir 11 o 12 eventos P ?
 - ¿Cuál es la probabilidad de lograr por lo menos 10 eventos P ?
 - ¿Cuál es la probabilidad de alcanzar un resultado tan extremo o más que 10 eventos P ?
- Utilizando la tabla B, con $N = 14$ y $P = 0.70$,
 - ¿Cuál es la probabilidad de obtener exactamente 13 eventos P ?
 - ¿Cuál es la probabilidad de conseguir por lo menos 13 eventos P ?
 - ¿Cuál es la probabilidad de lograr un resultado tan extremo o más que 13 eventos P ?
- Utilizando la tabla B, con $N = 20$ y $P = 0.20$,
 - ¿Cuál es la probabilidad de obtener exactamente dos eventos P ?
 - ¿Cuál es la probabilidad de conseguir dos eventos P o menos?
 - ¿Cuál es la probabilidad de lograr un resultado tan extremo o más que dos eventos P ?
- Una mujer lanza nueve monedas no cargadas. Si sólo permite una cara o una cruz en cada moneda,
 - ¿Cuál es la probabilidad de que todas caigan de cara?
 - ¿Cuál es la probabilidad de que obtenga siete caras o más?
 - ¿Cuál es la probabilidad de que obtenga un resultado tan extremo o más que siete caras?
- Una persona lanza una sola vez 15 monedas cargadas. Las monedas tienen un peso tal que la probabilidad de obtener una cara con cualquiera de ellas es de 0.85.
 - ¿Cuál es la probabilidad de que obtenga exactamente 14 caras?

- b. ¿Cuál es la probabilidad de que consiga por lo menos 14 caras?
- c. ¿Cuál es la probabilidad de que logre exactamente 3 cruces?
10. Una cerrajería anuncia que las llaves ahí fabricadas tienen una probabilidad $P = 0.90$ de funcionar bien. Si usted compró 10 llaves en esa cerrajería, ¿cuál es la probabilidad de que todas ellas funcionen correctamente?
11. Un estudiante presenta un examen de verdadero o falso con 15 preguntas. Si trata de adivinar todas las preguntas, ¿cuál es la probabilidad de que obtenga cuando menos 13 respuestas correctas? educación
12. Una estudiante presenta un examen de opción múltiple con 16 preguntas. Cada pregunta tiene cinco alternativas. Si ella trata de adivinar 12 de las 16 preguntas, ¿cuál es la probabilidad de que acierte por lo menos 8 preguntas? Suponga que todas las alternativas son igualmente probables en cada una de las preguntas en las que ella trató de adivinar la respuesta. educación
13. A usted le interesa averiguar si un niño en particular es capaz de distinguir entre los colores verde y azul. Con ese propósito, le muestra al pequeño cinco cubos de madera. Todos los cubos son idénticos, excepto que dos de ellos son verdes y tres son azules. Usted ordena de manera aleatoria los cubos en una fila y pide al niño que elija uno de color verde. Una vez que el pequeño lo ha elegido, usted reemplaza el cubo y modifica el orden de los cubos de manera aleatoria. A continuación, usted pide al niño que escoja un cubo verde. Este procedimiento se repite hasta que el niño ha realizado 14 selecciones. Si el niño no puede distinguir realmente entre el verde y el azul, ¿cuál es la probabilidad de que escoja un cubo verde por lo menos 11 veces? cognitiva
14. Supongamos que usted es un entusiasta aficionado a las carreras de caballos. Ahora se encuentra en el hipódromo y han anunciado ocho carreras. Este día, los caballos y los jockeys están tan parejos que sólo el azar determinará el orden de llegada en cada carrera. Participan 10 caballos en cada carrera. Si en cada carrera usted apuesta que un caballo en particular llegará en primero, segundo o tercer lugar, entonces,
- a. ¿Cuál es la probabilidad de que gane su apuesta en las ocho carreras?
- b. ¿Cuál es la probabilidad de que gane en seis de las carreras por lo menos? otra
15. Un fabricante de válvulas reconoce que su control de calidad ha decaído, de manera que actualmente su probabilidad de producir una válvula defectuosa es de 0.50. Si fabrica un millón de válvulas en un mes y usted elige al azar 10 000 muestras de esas válvulas, cada una formada por 15 válvulas, responda lo siguiente:
- a. ¿En cuántas muestras esperaría encontrar exactamente 13 válvulas en buen estado?
- b. ¿En cuántas muestras esperaría encontrar por lo menos 13 válvulas en buen estado? I/O
16. Imagine que 15% de la población es zurda y el resto utiliza la mano derecha (no hay ambidiestros). Si usted pregunta a las siguientes cinco personas que encuentre, ¿cuál es la probabilidad de que:
- a. Todas esas personas sean zurdas?
- b. Todas esas personas sean diestras?
- c. Exactamente dos de ellas sean zurdas?
- d. Por lo menos una de ellas sea zurda?
- Para resolver este problema, suponga que existe independencia en la selección de los cinco individuos. otra
17. En su distrito electoral, 25% de los votantes está en contra de cierta ley y el resto está en favor. Si usted realiza una encuesta y elige al azar cuatro votantes de su distrito, ¿cuál es la probabilidad de que:
- a. Ninguno de ellos esté en favor de la ley?
- b. Todos estén en favor de la ley?
- c. Por lo menos uno esté en contra de la ley? I/O
18. En la universidad en la que usted estudia, 30% de los estudiantes de licenciatura son originarios de otros estados. Si usted elige al azar ocho de esos alumnos, ¿cuál es la probabilidad de que:
- a. Todos sean del estado donde se encuentra la universidad?
- b. Todos sean de otros estados?
- c. Exactamente dos sean del estado donde se encuentra la universidad?
- d. Por lo menos cinco sean del estado donde se encuentra la universidad? educación
19. Veinte estudiantes, que están alojados en un dormitorio de la universidad, participaron en un concurso entre los dos principales refrescos de cola.
- a. Si en realidad no existe preferencia por ninguna, ¿cuál es la probabilidad de que los 20 estudiantes prefieran la marca X sobre la marca Y?
- b. Si en realidad no existe preferencia por ninguna, ¿cuál es la probabilidad de que por lo menos 17 de ellos prefieran la marca X sobre la marca Y?
- c. ¿Cuántos de los 20 estudiantes tendrían que preferir la marca X antes que usted estuviera

- dispuesto a concluir que verdaderamente existe una preferencia por dicha marca? otra
20. Un gran tazón contiene un millón de canicas. La mitad de éstas tiene pintado un signo más (+) y la otra mitad tiene un signo menos (-).
- a. Si usted elige al azar 10 canicas del tazón con reemplazo, una a la vez, ¿cuál será la probabili-

- dad de que escoja 9 canicas con signo más y 1 con signo menos?
- b. Si usted obtiene, una a la vez, 1000 muestras aleatorias de 10 canicas cada una, con reemplazo, ¿cuántas de esas muestras esperarías que tuviesen solamente canicas con el signo más? otra

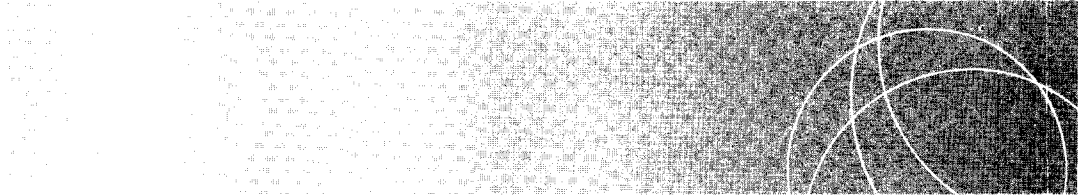
NOTAS

9.1 La ecuación para el desarrollo de $(P + Q)^N$ es

$$\begin{aligned} (P + Q)^N &= P^N + \frac{N}{1} P^{N-1} Q + \frac{N(N-1)}{1(2)} P^{N-2} Q^2 \\ &+ \frac{N(N-1)(N-2)}{1(2)(3)} P^{N-3} Q^3 \\ &+ \cdots + Q^N \end{aligned}$$

Si desea encontrar problemas de práctica adicionales totalmente resueltos sobre los temas que acaba de estudiar, busque Chapter 9 en el material adicional del libro, disponible en inglés en: <http://www.thomsonlearning.com.mx>.

Introducción a la prueba de hipótesis mediante la prueba del signo



Introducción

Un experimento: La marihuana y el tratamiento de pacientes con SIDA

Diseño de medidas repetidas

Hipótesis alternativa (H_1)

Hipótesis nula (H_0)

Regla de decisión (nivel α)

Evaluación del experimento de la marihuana mediante la distribución binomial

Errores de tipo I y de tipo II

El nivel alfa y el proceso de decisión

Evaluación de la cola de la distribución

Evaluaciones de probabilidad para una y dos colas

Magnitud del efecto: Significativo versus importante

¿CUÁL ES LA VERDAD? ¿Azar o efecto real? 1

¿CUÁL ES LA VERDAD? ¿Azar o efecto real? 2

¿CUÁL ES LA VERDAD? "Ningún producto es mejor que el nuestro"

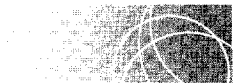
¿CUÁL ES LA VERDAD? Informes anecdóticos contra investigación sistemática

Resumen

Términos importantes

Preguntas y problemas

Notas



INTRODUCCIÓN

Ya hemos señalado que la estadística inferencial tiene dos propósitos principales: 1) la prueba de hipótesis y 2) la estimación de parámetros. Por lo general, la mayoría de las aplicaciones de la estadística inferencial pertenecen al área de la prueba de hipótesis. Como

estudiamos en el capítulo 1, la metodología científica se basa en esta aplicación de la estadística inferencial. Sin una verificación objetiva, la ciencia dejaría de existir y, por lo general, dicha verificación no se puede realizar sin la estadística inferencial. Usted recordará que el corazón de la metodología científica es un experimento. Por lo regular, el experimento ha sido diseñado para probar una hipótesis, por lo cual es preciso analizar los datos resultantes. A veces esos resultados son tan claros que no es necesario aplicar la inferencia estadística. Sin embargo, tales experimentos son raros. Debido a la variabilidad inherente de uno a otro sujeto con respecto a la variable que se mide, frecuentemente es difícil detectar el efecto de la variable independiente sin la ayuda de la estadística inferencial. En este capítulo iniciaremos un viaje fascinante, al estudiar cómo el diseño experimental, junto con el análisis matemático, permiten verificar diversas afirmaciones o hipótesis, como las hemos llamado. Le sugerimos poner especial atención a este capítulo. El material que contiene se aplica a todas las pruebas de inferencia que llevaremos a cabo (lo cual constituye la mayor parte del material que nos resta por estudiar en este texto).



UN EXPERIMENTO: LA MARIHUANA Y EL TRATAMIENTO DE PACIENTES CON SIDA

Supongamos que usted es un científico social y que actualmente trabaja en un hospital metropolitano que atiende a una numerosa población de pacientes con SIDA. A usted le preocupa mucho el dolor y el sufrimiento que afecta a esos pacientes. En particular, aunque no está del todo convencido, piensa que, desde una perspectiva ética, debería poderse utilizar la marihuana en el tratamiento de esos pacientes, especialmente en las etapas más avanzadas de la enfermedad. Por supuesto, antes de considerar con seriedad otros aspectos relacionados con el apoyo al uso de la marihuana con este propósito, usted se debe convencer de que esa sustancia tiene efectos positivos importantes. Hasta ahora, aunque existen muchos informes anecdóticos de pacientes con SIDA que al utilizar marihuana sienten menos náusea, aumenta su apetito y tienen más deseo de socializar, no existen experimentos científicos que respalden estos informes.

Como científico, usted observa que, aunque los informes personales son sugerentes, no son concluyentes. Es necesario realizar algunos experimentos antes que sea posible evaluar con propiedad las causas y los efectos; en este caso, los supuestos efectos de la marihuana. Dada la importancia de este tema, usted decide desarrollar un programa de investigación dirigido a ese fin. El primer experimento por realizar consiste en investigar el efecto de la marihuana sobre el apetito de los pacientes con SIDA. Por supuesto, si en realidad la marihuana redujera el apetito en lugar de aumentarlo, usted también debería detectar esto, debido a sus importantes consecuencias prácticas. Por lo tanto, éste será un experimento fundamental para determinar ciertos hechos, con los cuales intentará precisar si la marihuana tiene algún efecto, ya sea el de incrementar o reducir el apetito. El primer experimento será modesto. Usted planea obtener una muestra aleatoria con 10 individuos de la población de pacientes con SIDA tratados en su hospital. Usted observa que la generalización se limitará a esta población, pero por múltiples razones aceptará esa limitación en este experimento inicial. Después de haber obtenido el permiso apropiado de las autoridades correspondientes, usted realizará el siguiente experimento.

Se elige una muestra aleatoria de 10 pacientes con SIDA (los cuales están de acuerdo en participar en el experimento) a partir de una población bastante numerosa de pacientes, los cuales son tratados en el servicio de consultas externas del hospital donde usted trabaja. A ninguno de estos pacientes se le ha suministrado marihuana. Cada enfermo es internado en el hospital por una semana para que participe en el experimento. Durante los

2 primeros días, los pacientes se acostumbran al hospital. Al tercer día, la mitad de los enfermos ingieren antes de cada comida una píldora que contiene una forma sintética del ingrediente activo de la marihuana, THC. Al sexto día, esta píldora es sustituida por un placebo que ellos ingieren antes de cada comida. La otra mitad de los pacientes son tratados de igual manera que en la condición experimental, excepto que reciben las píldoras en orden inverso, es decir, el placebo al tercer día y las píldoras de THC en el sexto día. La variable dependiente es la cantidad de alimento ingerida por cada paciente en los días 3 y 6.

En este experimento, cada sujeto es estudiado bajo dos condiciones: una experimental y otra de control. La condición en la cual el sujeto recibe las píldoras de THC es la experimental, mientras que la condición en la que el individuo recibe las píldoras placebo es la condición de control. Por lo tanto, existen dos datos para cada sujeto: la cantidad de alimento ingerido (calorías) en la condición experimental y la cantidad de alimento ingerido en la condición de control. Si la marihuana influyera verdaderamente en el apetito, esperaríamos que los datos resultantes fueran diferentes para las dos condiciones. Por ejemplo, si la marihuana incrementa el apetito, entonces se deberá haber consumido más alimento en la condición experimental. Al restar el dato de control del dato experimental para cada sujeto, esperaríamos que predominaran las diferencias positivas. Los resultados del experimento se presentan en la tabla 10.1.

Podríamos analizar estos datos con diversas pruebas de inferencia estadística, como la prueba del signo, la de rangos con signo para pares igualados de Wilcoxon o la *t* de Student para grupos correlacionados. La elección de la prueba por utilizar en un experimento es muy importante. Dicha elección depende de la sensibilidad de la prueba y de que los datos del experimento cumplan los supuestos de la misma. En los siguientes capítulos estudiaremos cada uno de estos puntos. En este capítulo analizaremos los datos del experimento mediante la prueba del signo. Hemos elegido esta prueba porque 1) es fácil de comprender y 2) con ella podemos ilustrar de manera clara y sencilla todos los conceptos básicos relacionados con la prueba de hipótesis.

La prueba del signo “ignora” la magnitud de las diferencias entre los resultados y sólo toma en cuenta su dirección o signo. Con esto se omite mucha información, lo cual hace a la prueba bastante insensible (aunque mucho más fácil de entender). Si sólo consideramos los signos de la diferencia de los resultados, entonces el experimento produjo 9 signos positivos de 10 posibles. La cantidad de alimento consumida en la condición

tabla 10.1 Resultados del experimento sobre la marihuana

Paciente número	Condición experimental	Condición de control	Diferencias (calorías)
	Píldora THC Alimento ingerido (calorías)	Píldora placebo Alimento ingerido (calorías)	
1	1325	1012	+313
2	1350	1275	+ 75
3	1248	950	+298
4	1087	840	+247
5	1047	942	+105
6	943	860	+ 83
7	1118	1154	- 36
8	908	763	+145
9	1084	920	+164
10	1088	876	+212

experimental fue mayor después de haber ingerido la píldora THC en todos los pacientes, excepto uno. ¿Podemos entonces concluir que la marihuana produce un incremento en el apetito? No necesariamente.

Suponga que la marihuana no tiene efecto alguno sobre el apetito. ¿Sería posible obtener 9 signos positivos de 10 en el experimento? Sí, claro. Si la marihuana no tuviera efecto alguno sobre el apetito, entonces cada sujeto tendría dos condiciones idénticas, salvo por los factores debidos al azar. Tal vez, cuando el sujeto 1 estuvo bajo la condición THC, había dormido mejor la noche anterior y su apetito fue mayor que cuando fue estudiado bajo la condición de control, antes de tomar cualquier píldora. En tal caso, esperaríamos que comiera más en la condición THC, aunque el THC no tuviese efecto alguno sobre el apetito. Tal vez el sujeto 2 estaba resfriado cuando fue estudiado bajo la condición placebo, lo cual mitigó su apetito en comparación con el momento en que se encontraba bajo la condición experimental. También en ese caso esperaríamos que comiese más bajo la condición experimental, aunque el THC no produjera efecto alguno.

Podríamos seguir dando ejemplos para los demás sujetos. El punto es que estas explicaciones del mayor consumo de comida bajo la condición THC son factores aleatorios. Existen diferentes factores, independientes entre sí, que fácilmente podrían haberse presentado en cualquiera de los dos días de la prueba. Parecería improbable obtener 9 signos positivos de 10 posibles simplemente como resultado de los factores aleatorios. En realidad, la pregunta crucial es “¿qué tan improbable es esto?” Suponga que sabemos que, si el azar es el único responsable, obtendremos 9 signos positivos de 10 sólo una vez en mil millones. Como ésta es una ocurrencia demasiado rara, no dudaríamos en rechazar al azar y, con ello, la explicación de que la marihuana no tiene efecto alguno sobre el apetito. Concluiríamos entonces aceptando la hipótesis de que la marihuana afecta el apetito, porque ésa es la única explicación alternativa posible. Como la muestra fue aleatoria, podemos suponer que fue representativa de los pacientes con SIDA tratados en hospital donde usted labora y, por lo tanto, podríamos generalizar los resultados obtenidos a toda esa población.

Supongamos, sin embargo, que la probabilidad de obtener 9 signos positivos de 10 sólo por azar es de 1 en 3, no de 1 en mil millones. ¿Podemos rechazar al azar como causante de los resultados? Esta vez, la decisión no es tan clara. Necesitamos una regla para determinar cuándo la probabilidad obtenida es lo bastante pequeña como para rechazar al azar como la causa fundamental. Veremos que esto implica establecer un nivel crítico de probabilidad (llamado nivel alfa) como referencia para comparar los resultados.

Ahora formalizaremos algunos de los conceptos que hemos estado presentando.

Diseño de medidas repetidas

El diseño experimental que hemos utilizado se conoce como el diseño de medidas repetidas, de medidas replicadas o de grupos correlacionados. *Las características esenciales son la existencia de resultados pareados en las condiciones, y el hecho de que la diferencia entre éstos es analizada.* En el experimento de la marihuana, utilizamos los mismos sujetos en cada condición. Así, estos individuos hicieron las veces de su propio control. Formamos parejas con los resultados y, posteriormente, analizamos las diferencias entre ellas. En lugar de haber empleado a los mismos sujetos, podríamos haber utilizado parejas de gemelos idénticos o sujetos que se asemejaran entre sí de alguna otra manera. En los experimentos con animales, con frecuencia se utiliza a miembros de una misma camada para formar las parejas. En la forma más elemental de este diseño se utilizan únicamente dos condiciones: una experimental y una de control. Se procura que las dos condiciones sean lo más parecidas, dentro de lo posible, excepto por los valores de la variable independiente, los cuales, por supuesto, son diferentes de manera intencional. En nuestro ejemplo, la marihuana es la variable independiente.

Hipótesis alternativa (H_1)

En cualquier experimento existen dos hipótesis que compiten entre sí para explicar los resultados: la *hipótesis alternativa* y la *hipótesis nula*. La hipótesis alternativa es la que afirma que la variable independiente es la causa de la diferencia de los resultados entre las condiciones. En este caso, es la hipótesis de que “la marihuana afecta el apetito”. La hipótesis alternativa puede ser direccional o no direccional. La hipótesis “la marihuana afecta el apetito” pertenece a este último tipo, ya que no especifica la dirección del efecto. Si la hipótesis especifica la dirección del efecto, entonces es una hipótesis direccional. “La marihuana aumenta el apetito” es un ejemplo de hipótesis alternativa direccional.

Hipótesis nula (H_0)

La hipótesis nula es la contrapartida lógica de la hipótesis alternativa, de tal manera que si la hipótesis nula es falsa, la hipótesis alternativa debe ser verdadera. Por lo tanto, estas dos hipótesis tienen que ser mutuamente excluyentes y exhaustivas. Si la hipótesis alternativa es de tipo no direccional, entonces especifica que la variable independiente influye sobre la variable dependiente. Para esta clase de hipótesis (alternativa no direccional), la hipótesis nula indica que la variable independiente no influye sobre la variable dependiente. En el ejemplo que estamos presentando, como la hipótesis alternativa es de tipo no direccional, la hipótesis nula especifica que “la marihuana no tiene efecto sobre el apetito”. Señalamos anteriormente que la hipótesis alternativa nos dice que “la marihuana afecta el apetito”. Usted puede ver que estas dos hipótesis son mutuamente excluyentes y exhaustivas. Si la hipótesis nula es falsa, entonces la hipótesis alternativa debe ser verdadera. Como podrá ver, siempre evaluamos primero la hipótesis nula e intentamos demostrar que es falsa; si esto se logra, entonces la hipótesis alternativa debe ser verdadera.*

Si la hipótesis alternativa es direccional, la hipótesis nula afirma que la variable independiente no influye sobre la variable dependiente en la dirección especificada por la hipótesis alternativa.** Por ejemplo, para la hipótesis alternativa “la marihuana aumenta el apetito”, la hipótesis nula afirmarí que “la marihuana no aumenta el apetito”. De nuevo, observe que las dos hipótesis son mutuamente excluyentes y exhaustivas. Si la hipótesis nula es falsa, entonces la hipótesis alternativa debe ser verdadera.

Regla de decisión (nivel α)

Siempre evaluamos los resultados de un experimento mediante la evaluación de la hipótesis nula. La razón por la cual evaluamos directamente la hipótesis nula en vez de la hipótesis alternativa es que podemos calcular la probabilidad de eventos aleatorios, pero no hay manera de calcular la probabilidad de la hipótesis alternativa. Evaluamos la hipótesis nula suponiendo que es verdadera y verificando si dicho supuesto es razonable, al calcular la probabilidad de obtener el resultado *si sólo opera el azar*. Si la probabilidad resultante es menor o igual que un nivel de probabilidad crítico, llamado *nivel alfa* (α) rechazamos la hipótesis nula. Al rechazar la hipótesis nula, aceptamos de manera indirecta la hipótesis alternativa, ya que si el experimento se hace correctamente, ésta es la única explicación posible. Cuando rechazamos H_0 , decimos que los resultados son *significativos* o *confiables*. Si la probabilidad obtenida es mayor que el nivel alfa, concluimos que no podemos rechazar H_0 . Como el experimento no permite rechazar H_0 , conservamos H_0 como una explicación razonable de los datos. En este texto utilizaremos

* Véase la nota 10.1.

** Véase la nota 10.2.

las expresiones “no podemos rechazar H_0 ” y “conservamos H_0 ” de manera indistinta. Si conservamos H_0 , decimos que los resultados no son significativos o confiables. Por supuesto, cuando los resultados no son significativos, no podemos aceptar la hipótesis alternativa. De esta manera, la regla de decisión es:

Si la probabilidad obtenida $\leq \alpha$, rechazamos H_0 .

Si la probabilidad obtenida $> \alpha$, no podemos rechazar H_0 , es decir, conservamos H_0 .

El nivel alfa se establece al inicio del experimento. Por lo regular, los niveles alfa utilizados son $\alpha = 0.05$ y $\alpha = 0.01$. Más adelante, en el capítulo, analizaremos la razón por la cual se emplean estos niveles.

Por ahora, supongamos que $\alpha = 0.05$ para los datos referentes a la marihuana. Así pues, para evaluar los resultados del experimento de la marihuana, necesitamos: 1) determinar la probabilidad de obtener 9 signos positivos de 10 posibles cuando solamente interviene el azar y 2) comparar esta probabilidad con alfa.

Evaluación del experimento de la marihuana mediante la distribución binomial

Los datos de este experimento cumplen con los requisitos de la distribución binomial. El experimento consiste en una serie de ensayos (la exposición de cada uno de los pacientes a las condiciones experimental y de control constituye un ensayo). En cada ensayo hay solamente dos consecuencias posibles: un más o un menos. Observe que este modelo no permite empates. Si ocurre un empate debe ser descartado y, por consiguiente, habrá que reducir N . Los resultados son mutuamente excluyentes (un más y un menos no pueden ocurrir de manera simultánea), los ensayos son independientes (el resultado del paciente 1 no influye sobre el resultado del paciente 2, etcétera) y la probabilidad de un más y de un menos es la misma de un ensayo a otro. Como la distribución binomial es adecuada, utilizamos la tabla B del apéndice D (tabla 10.2) para determinar la probabilidad de obtener 9 signos positivos en 10 ensayos, cuando el único factor que influye es el azar. Resolvemos este problema de la misma manera que solucionamos los ejercicios de los lanzamientos de monedas en el capítulo 9.

Dado que hay 10 pacientes, $N = 10$. Sea P la probabilidad de obtener un signo positivo con cualquier paciente.* Si sólo opera el azar, la probabilidad de un signo positivo es igual a la probabilidad de un signo negativo. Como sólo hay dos alternativas igualmente probables, $P = 0.50$. Ya que deseamos determinar la probabilidad de obtener 9 signos positivos, el número de eventos P es igual a 9. En la tabla B, bajo $N = 10$, el número de eventos $P = 9$ y $P = 0.50$, obtenemos:

tabla 10.2 Datos de la tabla B

	Número de eventos P	P
$p(9 \text{ signos más}) = 0.0098$	9	0.50
	10	0.0098

* En este capítulo y el siguiente, cuando utilizemos la prueba del signo, P siempre denotará la probabilidad de obtener un signo positivo con cualquier sujeto. Este criterio es arbitrario. Por supuesto que podríamos haber elegido Q . Sin embargo, el uso de una misma letra (P o Q) para designar la probabilidad de un signo positivo en todos los problemas evita confusiones innecesarias.

Hemos fijado alfa en 0.05. El análisis muestra que sólo en 98 veces de 10 000 obtendríamos 9 signos positivos si el mero azar fuese la causa. Como 0.0098 es menor que alfa, rechazamos la hipótesis nula.⁸ Esto no parece ser una explicación razonable para esos datos. Por lo tanto, aceptamos la hipótesis alternativa: la marihuana afecta el apetito; parece que lo incrementa. Como quiera que la muestra es aleatoria, damos por cierto que es representativa de la población. Por lo tanto, es válido creer que esta conclusión se puede aplicar a la *población* de los pacientes con SIDA que son atendidos en el hospital donde usted trabaja.

Vale la pena observar que, en la práctica, suele ocurrir que los resultados de un experimento se generalizan a grupos que no forman parte de la población de donde fue tomada la muestra. Por ejemplo, en este experimento podríamos inclinarnos a declarar que la marihuana incrementa el apetito de los pacientes con SIDA que son atendidos en otros hospitales. Estrictamente, los resultados de un experimento sólo se aplican a la población de donde fue tomada la muestra aleatoria. Por lo tanto, la generalización a otros grupos debe hacerse con cuidado. Esta precaución es necesaria, ya que los demás grupos podrían diferir de los sujetos de la población original de alguna manera que cause un resultado diferente. Por supuesto, si el experimento se repite en varios hospitales con pacientes diferentes, la validez de la generalización es más amplia.



ERRORES DE TIPO I Y DE TIPO II

Al tomar decisiones con respecto a la hipótesis nula, podemos cometer errores de dos clases: Éstos reciben el nombre de *errores de tipo I* y de *tipo II*.

definiciones ■ *Un error de tipo I consiste en la decisión de rechazar la hipótesis nula cuando ésta es verdadera. Un error de tipo II consiste en la decisión de no rechazar la hipótesis nula cuando ésta es falsa.*

Para ilustrar estos conceptos, regresemos al ejemplo de la marihuana. Recuerde la lógica del proceso de decisión. Primero suponemos que H_0 es verdadera y evaluamos la probabilidad de obtener las mismas diferencias de puntajes observadas entre las condiciones si sólo interviene el azar. Si la probabilidad obtenida es $\leq \alpha$, entonces rechazamos H_0 . Si la probabilidad resultante es $> \alpha$, no rechazamos H_0 . En el experimento de la marihuana, la probabilidad resultante [$p(9 \text{ signos positivos})$] = 0.0098. En vista de que ésta es menor que alfa, rechazamos H_0 y concluimos que la marihuana fue la causante de los resultados observados. ¿Podemos tener la seguridad de que tomamos la decisión correcta? ¿Cómo sabemos que el azar no fue en realidad la causa? Tal vez la hipótesis nula sea verdadera. ¿Acaso no es posible que ésta haya sido una de esas 98 veces en 10 000 en las que obtendríamos 9 signos positivos y 1 negativo si sólo hubiera intervenido el azar? *La respuesta es que nunca sabemos con seguridad si el azar no fue el causante.* Es posible que los 9 signos positivos y 1 signo negativo se hayan debido al azar. En

⁸ En realidad, esto es una simplificación para dar mayor claridad a nuestra exposición. En la práctica, evaluamos la probabilidad de conseguir el mismo resultado obtenido o cualquier otro resultado más extremo. Analizaremos este punto con detalle, posteriormente, en este capítulo, en la sección "Evaluación de la cola de la distribución".

ese caso, habríamos cometido un error al rechazar H_0 . Esto sería un error de tipo I, es decir, rechazar la hipótesis nula cuando ésta es verdadera.

Un error de tipo II ocurre cuando conservamos H_0 y ésta es falsa. Supongamos que en el experimento de la marihuana, $p(9 \text{ signos positivos}) = 0.2300$ en lugar de 0.0098. En este caso, $0.2300 > \alpha$, por lo tanto, conservaríamos H_0 . Si H_0 es falsa habremos cometido un error de tipo II, es decir, habremos conservado H_0 siendo que no es verdadera.

Para aclarar la relación entre el proceso de decisión y los errores factibles, hemos resumido las posibilidades en la tabla 10.3. El encabezado de la columna es "Estado real". Esto indica el estado correcto de las situaciones referentes a la hipótesis nula. Sólo existen dos posibilidades: H_0 es verdadera o es falsa. El encabezado del renglón es la decisión tomada después de analizar los datos. De nuevo sólo hay dos posibilidades: que rechacemos H_0 o que la conservemos. Si conservamos H_0 y ésta es verdadera, habremos tomado la decisión correcta (véase la primera celda de la tabla). Si rechazamos H_0 y ésta es verdadera, entonces habremos cometido un error de tipo I. Esto se muestra en la celda 3. Si conservamos H_0 y ésta es falsa, habremos cometido un error de tipo II (celda 2). Por último, si rechazamos H_0 y ésta es falsa, habremos tomado la decisión correcta (celda 4). Observe que cuando rechazamos H_0 , el único error posible es el de tipo I. Si conservamos H_0 , el único error que podemos cometer es el de tipo II.

tabla 10.3 Conclusiones posibles y estado real

Decisión	Estado real	
	H_0 es verdadera	H_0 es falsa
Conservar H_0	¹ Decisión correcta	² Error de tipo II
Rechazar H_0	³ Error de tipo I	⁴ Decisión correcta

Usted preguntará quizá por qué nos hemos tomado la molestia de analizar todas las posibilidades lógicas. Lo hicimos porque consideramos que es muy importante conocer todos los posibles errores que podríamos cometer al deducir las conclusiones de un experimento. Por el análisis anterior sabemos que sólo existen dos posibilidades a ese respecto: un error de tipo I o un error de tipo II. Al saber que estos errores son posibles, podemos diseñar los experimentos, antes de realizarlos, con el fin de minimizar la probabilidad de cometer un error de tipo I o un error de tipo II. Cuando minimizamos la probabilidad de cometer estos errores, maximizamos la probabilidad de sacar la conclusión correcta, independientemente de que la hipótesis nula sea verdadera o falsa. En la siguiente sección veremos que alfa limita la probabilidad de incurrir en un error de tipo I. Por lo tanto, al controlar el nivel alfa, podemos minimizar la posibilidad de cometer un error de tipo I. Definiremos beta como la probabilidad de incurrir en un error de tipo II. En el siguiente capítulo analizaremos distintas formas en que se puede minimizar beta.



EL NIVEL ALFA Y EL PROCESO DE DECISIÓN

Debemos tomar en cuenta que cuando utilizamos datos muestrales para evaluar una hipótesis, nunca estamos seguros de que nuestra conclusión sea correcta. Cuando rechazamos H_0 , no sabemos con certeza que sea falsa. Por eso corremos el riesgo de cometer un error de tipo I. Por supuesto, mientras menos razonable sea la posibilidad de que el azar

sea la causa de los resultados, tanta más confianza tendremos de no haber cometido un error al rechazar la hipótesis nula. Por ejemplo, cuando la probabilidad de obtener los resultados bajo el supuesto del azar sea de 1 en un millón ($p = 0.000001$) tendremos más confianza de que la hipótesis nula sea falsa, que si la probabilidad es de 1 en 10 ($p = 0.10$).

definición ■ *El nivel alfa que los científicos establecen al principio de un experimento es el nivel al cual desean limitar la probabilidad de cometer un error de tipo I.*

Así pues, cuando un científico establece que $\alpha = 0.05$, está anunciando que al reunir los datos rechazará la hipótesis nula si, bajo el supuesto de que el azar es el único responsable, la probabilidad resultante es igual o menor que 5 veces en 100. Al hacer esto, señala que está dispuesto a limitar a 5 veces en 100 la probabilidad de rechazar la hipótesis nula cuando ésta sea verdadera. De esta manera, limita la probabilidad de cometer un error de tipo I a 0.05.

No existe fórmula mágica alguna que nos indique cuál debe ser el nivel alfa adecuado para alcanzar la verdad en cada experimento. A fin de determinar un nivel alfa razonable para una prueba, debemos considerar cuáles serían las consecuencias de cometer un error. En la ciencia, los efectos de rechazar la hipótesis nula cuando es verdadera (error de tipo I) son costosos. Cuando un científico publica un experimento donde rechaza la hipótesis nula, otros investigadores intentarán repetir los resultados o aceptar la conclusión como válida y diseñar otros experimentos, dando por hecho que el científico tomó la decisión correcta. Como se invierten muchas horas de trabajo y mucho dinero en esos experimentos posteriores, los científicos procuran minimizar la posibilidad de seguir avanzando por un camino erróneo. Por esa razón, los investigadores establecen niveles alfa bastante conservadores: $\alpha = 0.05$ y $\alpha = 0.01$ se utilizan con frecuencia. Usted podría preguntar “¿por qué no se instaure un criterio más riguroso, como $\alpha = 0.001$?”. Por desgracia, cuando el valor de alfa es más restrictivo se incrementa la probabilidad de cometer un error de tipo II.

Podemos apreciar esto por medio de un ejemplo. Dicho ejemplo será más comprensible si se plantea en combinación con la tabla 10.4. Suponga usted que realizaremos un experimento y establecemos $\alpha = 0.05$ (renglón superior de la tabla 10.4). Evaluamos la opción de que el resultado se deba al azar y obtenemos una probabilidad de 0.02. Entonces rechazamos H_0 . Sin embargo, si H_0 es verdadera habremos cometido un error de tipo I (celda 1). Ahora supongamos que $\alpha = 0.01$ en vez de 0.05 (renglón inferior de la tabla 10.4). En este caso, conservaríamos H_0 y ya no cometeríamos el error de tipo I (celda 3). Así, mientras más estricto sea el nivel alfa, menor será la probabilidad de cometer un error de tipo I.

Por otra parte, ¿qué sucede si en realidad H_0 es falsa (última columna de la tabla)? Con $\alpha = 0.05$ y la probabilidad obtenida igual a 0.02, rechazaríamos H_0 y al hacerlo to-

Tabla 10.4 Efecto sobre beta cuando alfa se hace más estricto

Nivel alfa	Probabilidad obtenida	Decisión	Estado real	
			H_0 es verdadera	H_0 es falsa
0.05	0.02	Rechazar H_0	¹ Error tipo I	² Decisión correcta
0.01	0.02	Aceptar H_0	³ Decisión correcta	⁴ Error tipo II

mariamos una decisión correcta (celda 2). Sin embargo, si cambiáramos el valor de alfa por $\alpha = 0.01$, conservaríamos H_0 y cometeríamos un error de tipo II (celda 4). De este modo, si restringimos el valor de alfa aún más, disminuimos la probabilidad de cometer un error de tipo I, pero aumentamos la probabilidad de incurrir en un error de tipo II. Debido a esta interacción entre alfa y beta, el nivel alfa empleado para un experimento depende del uso que se piense dar a los resultados del mismo. Como mencionamos anteriormente, si los resultados sirven para comunicar un nuevo hecho a la comunidad científica, entonces las consecuencias de un error de tipo I serán graves, por lo cual se utilizan niveles alfa muy estrictos (0.05 y 0.01). En cambio, si el experimento tiene una naturaleza exploratoria y los resultados servirán como "guía" para que el investigador decida si es conveniente realizar un experimento más completo, entonces sería una tontería utilizar niveles demasiado estrictos. En tales casos, los niveles alfa de que se usan comúnmente son tan altos como 0.10 o 0.20.

Consideremos un ejemplo más. Imagine que usted es presidente de una compañía farmacéutica. Una de sus principales investigadoras bioquímicas llega a su oficina y le dice que ha descubierto un medicamento que acrecienta la memoria. Por supuesto, usted se entusiasma, pero necesita ver los resultados experimentales. Supongamos que se requieren 30 millones de dólares para instalar los aparatos con los que se producirá el medicamento. Esto es un poco caro, pero si el medicamento realmente incrementa la memoria, los beneficios y las ganancias potenciales lo justificarán. En ese caso, usted deseará estar muy seguro de que los resultados obtenidos no sean producto del azar pues las consecuencias de cometer un error de tipo I serían graves. Se arriesgaría a perder 30 millones de dólares. Es probable que desee utilizar un nivel alfa extremadamente estricto antes de decidir si rechaza H_0 y arriesga la inversión de 30 millones de dólares.

Debemos apresurarnos a afirmar que la *verdad* no depende del nivel alfa que se use en un experimento. La marihuana afecta el apetito o no. El medicamento acrecienta la memoria o no. El establecimiento de un nivel alfa muy estricto sólo disminuye la posibilidad de que aceptemos la hipótesis alternativa cuando la hipótesis nula sea la verdadera.

En vista de que nunca sabemos con certeza cuál es la verdad si nos basamos en el resultado de un único experimento, la repetición es una parte necesaria y esencial del proceso científico. Para que un "presunto hecho" sea aceptado como parte del conocimiento científico, tendrá que ser comprobado de manera independiente en varios laboratorios. La probabilidad de cometer un error de tipo I disminuye en grado considerable por medio de la repetición independiente.



EVALUACIÓN DE LA COLA DE LA DISTRIBUCIÓN

En el análisis anterior determinamos la probabilidad utilizando solamente el resultado específico del experimento (es decir, 9 signos positivos y 1 negativo). Sin embargo, procedimos de esa manera para que las cosas resulten más sencillas y claras al presentar los demás conceptos principales. De hecho, es incorrecto emplear *solamente el resultado específico* al evaluar las conclusiones de un experimento. En vez de eso, debemos precisar la probabilidad de obtener el resultado conseguido o *cualquier otro resultado aún más extremo*. Ésta es la probabilidad que comparamos con alfa para evaluar lo razonable de la hipótesis nula. En otras palabras, evaluamos la *cola* de la distribución a partir del resultado obtenido y no solamente el resultado mismo. Si la hipótesis alternativa no es direccional, evaluamos el resultado obtenido, o cualquier otro resultado más extremo, en ambas direcciones (ambas colas). Si la hipótesis alternativa es direccional, evaluamos únicamente la cola de la distribución que está en la dirección especificada por H_1 .

Como ejemplo, evaluaremos de nuevo los datos del presente experimento, pero ahora examinaremos las colas y no sólo el resultado específico. La figura 10.1 muestra la distribución binomial para $N = 10$ y $P = 0.50$. La distribución tiene dos colas, una que contiene pocos signos positivos y otra que contiene muchos signos positivos. Como la hipótesis alternativa no es direccional, para calcular la probabilidad obtenida debemos determinar la probabilidad de conseguir el resultado deseado, o uno aún más extremo, en ambas direcciones. Como el resultado obtenido fue de 9 signos positivos, tenemos que incluir los resultados que son tan extremos o más que 9 signos positivos. En la figura 10.1, observamos que el resultado de 10 signos positivos es más extremo en una dirección, y que los resultados de un signo positivo y cero signos positivos son igualmente extremos o más en la otra dirección. Así, la probabilidad obtenida es:

$$\begin{aligned} p(0, 1, 9 \text{ o } 10 \text{ signos positivos}) &= p(0) + p(1) + p(9) + p(10) \\ &= 0.0010 + 0.0098 + 0.0098 + 0.0010 \\ &= 0.0216 \end{aligned}$$

Comparamos esta probabilidad (0.0216, y no 0.0098) con alfa para rechazar o conservar la hipótesis nula. Esta probabilidad es el valor de la *probabilidad de dos colas*, puesto que los resultados evaluados aparecen bajo ambas colas de la distribución. Así, las hipótesis alternativas no direccionales son evaluadas con valores de probabilidad de dos colas. Si la hipótesis alternativa es no direccional, el nivel alfa debe ser también de dos colas. Si $\alpha = 0.05_{2 \text{ colas}}$, esto significa que el valor obtenido de la probabilidad de dos colas debe ser menor o igual que 0.05 para que se pueda rechazar H_0 . En este ejemplo, como 0.0216 es menor que 0.05, rechazamos H_0 y concluimos, como antes, que la marihuana afecta el apetito.

Si la hipótesis alternativa es direccional, evaluamos la cola de la distribución que se encuentra en la dirección prevista por H_1 . Para ilustrar este punto, suponga que la hipótesis alternativa era “la marihuana aumenta el apetito” y que el resultado obtenido fue de 9 signos positivos y 1 negativo. Como H_1 especifica que la marihuana aumenta el apetito, evaluamos la cola que tenga mayor número de signos positivos. Recuerde que un signo positivo significa que se ingiere más comida bajo la condición de consumir marihuana. De este modo, si la marihuana aumenta el apetito, esperamos tener más signos

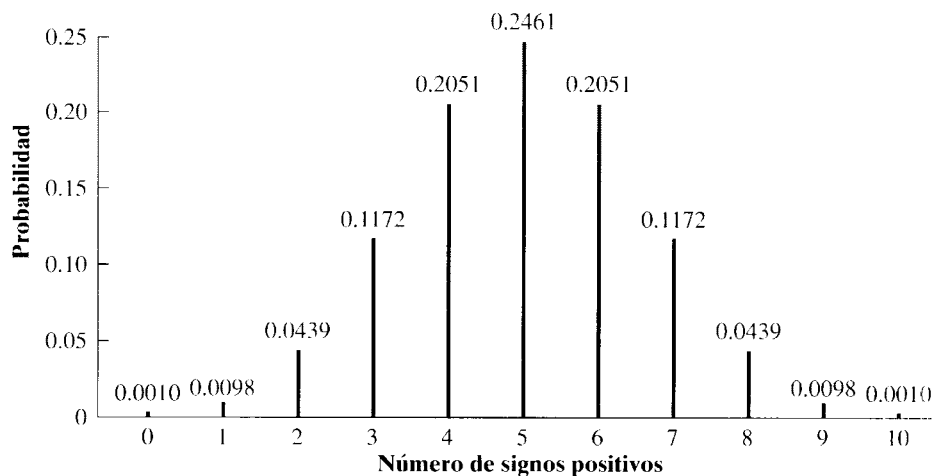


figura 10.1 Distribución binomial para $N = 10$ y $P = 0.50$.

positivos. El resultado de 10 signos positivos es el único posible en esta dirección, más extremo que 9 signos positivos. La probabilidad obtenida es:

$$\begin{aligned} p(9 \text{ o } 10 \text{ signos positivos}) &= 0.0098 + 0.0010 \\ &= 0.0108 \end{aligned}$$

Esta probabilidad se conoce como *probabilidad de una cola*, ya que todos los resultados evaluados están bajo una cola de la distribución. Así, las hipótesis alternativas que son direccionales se evalúan con probabilidades de una cola. Si la hipótesis alternativa es direccional, el nivel alfa debe ser de una cola. Por lo tanto, las hipótesis alternativas direccionales se evalúan con respecto a niveles alfa de una cola. En este ejemplo, si $\alpha = 0.05_{1 \text{ cola}}$, rechazaríamos H_0 , ya que 0.0108 es menor que 0.05.

La razón por la cual evaluamos la cola tiene que ver con el nivel alfa establecido al principio del experimento. En el ejemplo que hemos utilizado, suponga que la hipótesis es "la marihuana aumenta el apetito". Ésta es una hipótesis direccional, de modo que resulta adecuada la evaluación de una cola. Suponga que $N = 10$ y $\alpha = 0.05_{1 \text{ cola}}$. Si establecemos $\alpha = 0.05$ al principio del experimento, el investigador descarta limitar a 5 en 100 la probabilidad de cometer un error de tipo I. Suponga que los resultados de dicho experimento son 8 signos positivos y 2 negativos. ¿Es éste un resultado que permita rechazar H_0 de manera consistente con el nivel alfa? Su primer impulso sería responder "sí", ya que $p(8 \text{ signos positivos}) = 0.0439$. Sin embargo, si no admitimos H_0 con 8 signos positivos, entonces también debemos rechazarla cuando los resultados son 9 o 10 signos positivos. ¿Por qué? Porque estos resultados son aún más favorables a H_1 que 8 signos positivos y 2 negativos. En realidad, si la marihuana aumenta el apetito, la obtención de 10 signos positivos y 0 negativos es una evidencia mejor que 8 signos positivos y 2 negativos; de igual forma ocurre con 9 signos positivos y 1 negativo. Así, si rechazamos H_0 con 8 signos positivos, también deberemos hacerlo con 9 y 10 signos positivos. Pero, ¿cuál es la probabilidad de obtener 8, 9 o 10 signos positivos si sólo interviene el azar?

$$\begin{aligned} p(8, 9 \text{ o } 10 \text{ signos positivos}) &= p(8) + p(9) + p(10) \\ &= 0.0439 + 0.0098 + 0.0010 \\ &= 0.0547 \end{aligned}$$

La probabilidad es mayor que alfa. Por lo tanto, no podemos permitir que 8 signos positivos sean un resultado con el cual podamos rechazar H_0 ; la probabilidad de rechazar, de manera errónea, H_0 sería mayor que el nivel alfa. Observe que esto es cierto, aunque la propia probabilidad de obtener 8 signos positivos sea menor que alfa. Por lo tanto, no evaluamos el resultado exacto, sino la cola, para limitar la probabilidad de cometer un error de tipo I hasta el nivel alfa establecido al inicio del experimento. La razón por la cual se utiliza una evaluación de dos colas con una hipótesis alternativa no direccional es que los resultados en ambos extremos de la distribución son candidatos válidos para el rechazo de la hipótesis nula.



EVALUACIONES DE PROBABILIDAD PARA UNA Y DOS COLAS

Al establecer el nivel alfa, decidimos si la evaluación de la probabilidad debe ser de una o dos colas. Para tomar esta decisión, contamos con la siguiente regla:

La evaluación debe ser siempre de dos colas, a menos que el experimentador acepte H_0 cuando los resultados sean extremos en la dirección opuesta a la dirección predicha.

Al aplicar esta regla, encontramos por lo general dos situaciones que garantizan la hipótesis direccional. En primer lugar, cuando no existe diferencia práctica si los resultados aparecen en la dirección opuesta, es válido utilizar una hipótesis direccional y una evaluación de una cola. Por ejemplo, si un fabricante de neumáticos para automóvil está probando un nuevo tipo de llantas que supuestamente duran más, es válida una evaluación de una cola, ya que no hay gran diferencia práctica si los resultados experimentales se encuentran en la dirección opuesta. La conclusión será conservar H_0 y el fabricante continuará utilizando los viejos neumáticos. Otra situación que permite emplear una evaluación de una cola se presenta cuando existen buenas razones teóricas, así como datos o bases sólidas, para justificar la dirección predicha. En este caso, si los resultados experimentales aparecen en la dirección opuesta, el investigador conservará de nuevo H_0 (por lo menos hasta que el experimento se repita), ya que los resultados pierden validez frente a los datos y las teorías anteriores.

En situaciones en las que el experimentador rechaza H_0 si los resultados del experimento son extremos en la dirección opuesta a la dirección de la predicción, se utiliza una evaluación de dos colas. Para entender por qué, supongamos que el investigador inicia una predicción direccional, estableciendo $\alpha = 0.05_{1 \text{ cola}}$ y que los resultados son extremos en la dirección opuesta. Si él no quiere concluir que debe conservar H_0 , lo que probablemente hará, después de ver los datos, es usar una hipótesis no direccional con $\alpha = 0.05_{2 \text{ colas}}$ (0.025 bajo cada cola) para poder rechazar H_0 . A largo plazo, este procedimiento producirá una probabilidad de cometer un error de tipo I de 0.075 (0.05 bajo la cola en la dirección predicha y 0.025 bajo la otra cola). Así, el intercambio de las hipótesis alternativas después de ver los datos produce una probabilidad inflada de cometer un error de tipo I. Por supuesto, esto puede ser peor si, después de observar que los datos están en la dirección opuesta a la prevista, el experimentador cambia a $\alpha = 0.05_{1 \text{ cola}}$ en la dirección del resultado a fin de rechazar H_0 . En esta situación, la probabilidad de cometer un error de tipo I a largo plazo sería de 0.10 (0.05 bajo cada cola). Por ejemplo, se puede predecir que si un experimentador aplica este procedimiento en 100 experimentos y suponemos que en todos éstos la hipótesis nula es verdadera, rechazaría erróneamente en 10 ocasiones la hipótesis nula. Puesto que cada uno de esos rechazos sería un error de tipo I, el hecho de seguir este procedimiento implica una probabilidad de cometer un error de tipo I de 0.10 ($10/100 = 0.10$). Por lo tanto, para mantener en el nivel deseado la probabilidad de incurrir en un error de tipo I, es importante decidir al principio del experimento si H_1 tiene que ser direccional o no y establecer el nivel alfa según esto. Si se utiliza una H_1 direccional, hay que apegarse a la dirección prevista, aunque los resultados del experimento sean extremos en la dirección opuesta. Por consiguiente, H_0 debe ser conservada en ese tipo de casos.

Para resolver los problemas y los ejemplos de este libro, indicaremos si resulta apropiada la evaluación de una o de dos colas; nos gustaría que usted practicara con ambas. Tenga cuidado al resolver estos ejercicios. Cuando un científico realiza un experimento, por lo general sigue una corazonada para predecir un efecto direccional. Con frecuencia, los problemas de este libro se enuncian en términos de la corazonada direccional del experimentador. No obstante, *a menos que el investigador termine conservando H_0 si los resultados tienden a ser extremos en la dirección opuesta*, deberá utilizar una H_1 no direccional y una evaluación de dos colas, aunque su corazonada sea direccional. Cada uno de los problemas del libro le dirá si debe emplear una H_1 direccional o no direccional cuando se le pida una hipótesis alternativa. Si se le solicita una H_1 no direccional, debe suponer que no se ha cumplido el criterio adecuado para una hipótesis alternativa direccional, sin importar si la corazonada del científico en el problema es direccional. Si se le pide una H_1 direccional, suponga que se cumple el criterio apropiado y que es adecuado usar una H_1 direccional.

Ya estamos listos para resolver un problema completo, en la misma forma en que cualquier científico lo haría si usara la prueba del signo para evaluar los datos.

Suponga que realizamos un experimento para probar la hipótesis de que la marihuana afecta al apetito de los pacientes con SIDA. El procedimiento y la población son los mismos que describimos anteriormente, excepto que esta vez hemos tomado una muestra de 12 pacientes con SIDA. Los resultados se muestran a continuación (los datos están expresados en calorías):

Condición	Paciente											
	1	2	3	4	5	6	7	8	9	10	11	12
THC	1051	1066	963	1179	1144	912	1093	1113	985	1271	978	951
Placebo	872	943	912	1213	1034	854	1125	1042	922	1136	886	902

- ¿Cuál es la hipótesis alternativa no direccional?
- ¿Cuál es la hipótesis nula?
- Utilizando $\alpha = 0.05_{2 \text{ colas}}$ ¿qué se puede concluir?
- ¿Qué error podría cometer en la conclusión del inciso c'?
- ¿A qué población se aplica dicha conclusión?

A continuación se presenta la solución.

SOLUCIÓN

- Hipótesis alternativa no direccional: La marihuana afecta el apetito de los pacientes con SIDA, tratados en el hospital donde usted trabaja.
- Hipótesis nula: La marihuana no tiene efecto alguno sobre el apetito de los pacientes con SIDA, tratados en el hospital donde usted trabaja.
- Conclusión, al emplear $\alpha = 0.05_{2 \text{ colas}}$:

PASO 1: Cálculo del número de signos positivos y negativos. El primer paso consiste en calcular el número de signos positivos y negativos de la muestra. Hemos restado los resultados “placebo” de los resultados “THC” correspondientes. También podríamos restar al revés. Existen 10 signos positivos y 2 negativos.

PASO 2: Evaluación del número de signos positivos y negativos. Una vez calculado el número de signos positivos y negativos, determinamos la probabilidad de obtener este resultado o cualquier otro más extremo, en ambas direcciones, ya que se trata de una evaluación de dos colas.


(continúa)

La distribución binomial es adecuada en este caso. N es el número de diferencias (signos positivos y negativos) y es igual a 12. Sea P la probabilidad de un signo positivo con cualquier sujeto. Si la marihuana no tiene efecto sobre el apetito, sólo el azar cuenta cuando cualquier sujeto obtiene un signo positivo o uno negativo. Por lo tanto, $P = 0.50$. Como el resultado fue de 10 signos positivos y 2 negativos, el número de eventos P es 10. La probabilidad de conseguir un efecto tan extremo o más que 10 signos positivos (de dos colas) es igual a la probabilidad de obtener 0, 1, 2, 10, 11 o 12 signos positivos. Como la distribución es simétrica, $p(0, 1, 2, 10, 11$ o 12 signos positivos) es igual a $p(10, 11$ o 12 signos positivos) $\times 2$. Así, a partir de la tabla B:

$p(0, 1, 2, 10, 11$ o 12 signos positivos)	= $p(10, 11$ o 12 signos positivos) $\times 2$	= $[p(10) + p(11) + p(12)] \times 2$	= $(0.0161 + 0.0029 + 0.0002) \times 2$	= 0.0384	Datos de la tabla B												
					<table style="border-collapse: collapse; width: 100%;"> <thead> <tr> <th style="text-align: left; padding: 5px;">N</th> <th style="text-align: left; padding: 5px;">Número de eventos P</th> <th style="text-align: left; padding: 5px;">P 0.50</th> </tr> </thead> <tbody> <tr> <td style="padding: 5px;">12</td> <td style="padding: 5px;">10</td> <td style="padding: 5px;">0.0161</td> </tr> <tr> <td></td> <td style="padding: 5px;">11</td> <td style="padding: 5px;">0.0029</td> </tr> <tr> <td></td> <td style="padding: 5px;">12</td> <td style="padding: 5px;">0.0002</td> </tr> </tbody> </table>	N	Número de eventos P	P 0.50	12	10	0.0161		11	0.0029		12	0.0002
N	Número de eventos P	P 0.50															
12	10	0.0161															
	11	0.0029															
	12	0.0002															

Podríamos haber obtenido el mismo valor si hubiéramos sumado las seis probabilidades en vez de determinar la probabilidad de una cola y multiplicarla por 2. Como $0.0384 < 0.05$, rechazamos la hipótesis nula. No ofrece una explicación razonable de los resultados. Por lo tanto, concluimos que la marihuana sí afecta el apetito. Tal parece que lo incrementa.

- d. Error posible: Al rechazar la hipótesis nula, podría cometerse un error del tipo I. En realidad, la hipótesis nula puede ser verdadera y usted la ha rechazado.
- e. Población: Estos resultados se aplican a la población de pacientes con SIDA de la que fue tomada la muestra.



Usted tiene una buena razón para pensar que cierto programa de TV está causando un incremento en la conducta violenta de los adolescentes. Para probar esta hipótesis, usted realiza un experimento con una muestra aleatoria de 15 individuos, seleccionados entre los adolescentes del colegio más cercano de su localidad. Cada sujeto es sometido a una condición experimental y una de control. Bajo la condición experimental, los adolescentes observan el programa de TV durante tres meses y se registra el número de actos violentos que cometen. La condición de control

también dura tres meses, pero los adolescentes no pueden ver dicho programa. Al final del periodo de tres meses, se cuenta el número de actos violentos cometidos. Los resultados se presentan a continuación:

Condición	Sujeto														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Viendo el programa	25	35	10	8	24	40	44	18	16	25	32	27	33	28	26
Sin ver el programa	18	22	7	11	13	35	28	12	20	18	38	24	27	21	22

- ¿Cuál es la hipótesis alternativa direccional?
- ¿Cuál es la hipótesis nula?
- Al emplear $\alpha = 0.01_{1 \text{ cola}}$ ¿qué conclusión saca usted?
- ¿Qué error podría cometer a causa de la conclusión del inciso c?
- ¿A qué población se aplica dicha conclusión?

La solución se presenta a continuación.

SOLUCIÓN

- Hipótesis alternativa direccional: Observar el programa de TV causa un incremento en la conducta violenta de los adolescentes.
- Hipótesis nula: Observar el programa de TV no causa un incremento en la conducta violenta de los adolescentes.
- Conclusión, al emplear $\alpha = 0.01_{1 \text{ cola}}$:

PASO 1: Calcular el número de signos positivos y negativos. El primer paso es calcular el número de signos positivos y negativos contenidos en la muestra de datos. Hemos restado los datos de la condición “sin ver” de los datos bajo la condición “viendo”. El resultado obtenido es de 12 signos positivos y 3 negativos.

PASO 2: Evaluar el número de signos positivos y negativos. Ahora debemos determinar la probabilidad de obtener este resultado, o cualquier otro más extremo, en la dirección de la hipótesis alternativa. Ésta es una evaluación de una cola, ya que la hipótesis alternativa es direccional. La distribución binomial es la apropiada. N es el número de diferencias, igual a 15. Sea P la probabilidad de un signo positivo para cualquier sujeto. Podemos evaluar la hipótesis nula suponiendo que sólo interviene el azar cuando cualquier sujeto obtiene un signo positivo o negativo. Por lo tanto, $P = 0.50$. Como el resultado obtenido fue de 12 signos positivos y 3 negativos, el número de eventos $P = 12$. La probabilidad de obtener 12 signos positivos o más es igual a la probabilidad de obtener 12, 13, 14 o 15 signos positivos. Esto se presenta en la tabla B. Así,

(continúa)

$p(12, 13, 14 \text{ o } 15 \text{ signos positivos})$

$$= p(12) + p(13) + p(14) + p(15)$$

$$= 0.0139 + 0.0032 + 0.0005 + 0.0000$$

$$= 0.0176$$

Datos de la tabla B

N	Número de eventos P	P 0.50
15	12	0.0139
	13	0.0032
	14	0.0005
	15	0.0000

Como $0.0176 > 0.01$, no podemos rechazar la hipótesis nula. Por lo tanto, conservamos H_0 y no podemos concluir que el programa de TV cause un incremento en la conducta violenta de los adolescentes.

- d.** Error posible: Al conservar la hipótesis nula, podríamos cometer un error de tipo II. El programa de TV podría causar un incremento real en la conducta violenta de los adolescentes.
- e.** Población: Estos resultados se aplican a la población de adolescentes del colegio más cercano de su localidad.



Un psicólogo de una empresa piensa que el ejercicio influye en la imagen que cada persona tiene de sí misma. Para investigar esta posibilidad, fueron elegidos 14 empleados de la empresa, de manera aleatoria, para que participaran en un programa de carrera al trote. Antes de comenzar el programa, se les hizo responder un cuestionario para medir su imagen de sí mismos. Después, iniciaron el programa de carrera al trote. En dicho programa, las personas trotan a una velocidad moderada, durante 20 minutos diarios, 4 días a la semana. La imagen de los empleados se mide de nuevo, después de 2 meses de haber participado en el programa. A continuación, mostramos los resultados (entre más alto sea el resultado, más alta será la imagen de sí mismos); un puntaje de 20 es el resultado más alto posible.

Sujeto	Antes de trotar	Después de trotar	Sujeto	Antes de trotar	Después de trotar
1	14	20	8	16	13
2	13	16	9	10	16
3	8	15	10	14	18
4	14	12	11	6	14
5	12	15	12	15	17
6	7	13	13	12	18
7	10	12	14	9	15

- ¿Cuál es la hipótesis alternativa? Utilice una hipótesis no direccional.
- ¿Cuál es la hipótesis nula?
- Al utilizar $\alpha = 0.05_{2 \text{ colas}}$ ¿qué conclusión saca usted?
- ¿Qué error podría cometer a causa de su conclusión para el inciso c?
- ¿A qué población se aplica dicha conclusión?

La solución se presenta a continuación.

SOLUCIÓN

- Hipótesis alternativa no direccional: La carrera al trote afecta la imagen de sí mismo.
- Hipótesis nula: La carrera al trote no influye en la imagen de sí mismo.
- Conclusión, al utilizar $\alpha = 0.05_{2 \text{ colas}}$:

PASO 1: Calcular el número de signos positivos y negativos. Hemos restado los resultados o puntajes “antes de trotar”, de los resultados “después de trotar”. Existen 12 signos positivos y 2 negativos.

PASO 2: Evaluar el número de signos positivos y negativos. En virtud de que H_1 no es direccional, tendremos que determinar la probabilidad de obtener un resultado tan extremo o más que 12 signos positivos (con dos colas), suponiendo que el azar es la única razón que explica las diferencias. La distribución binomial es la apropiada. $N = 14$, $P = 0.50$, y el número de eventos $P = 0, 1, 2, 12, 13$ o 14 . Así pues, a partir de la tabla B, tenemos lo siguiente:

$$\begin{aligned}
 p(0, 1, 2, 12, 13 \text{ o } 14 \text{ signos positivos}) &= p(0) + p(1) + p(2) + p(12) \\
 &\quad + p(13) + p(14) \\
 &= 0.0001 + 0.0009 + 0.0056 \\
 &\quad + 0.0056 + 0.0009 + 0.0001 \\
 &= 0.0132
 \end{aligned}$$

Datos de la tabla B

N	Número de eventos P	P 0.50
14	0	0.0001
	1	0.0009
	2	0.0056
	12	0.0056
	13	0.0009
	14	0.0001

Podríamos haber obtenido el mismo valor si hubiéramos determinado la probabilidad de una cola y la hubiéramos multiplicado por 2. En virtud de que $0.0132 < 0.05$, rechazamos la hipótesis nula. Al parecer, la carrera al trote sí mejora la imagen de sí mismo.

- Error posible: Al rechazar la hipótesis nula, podríamos cometer un error de tipo I, ya que la hipótesis nula podría ser verdadera y haber sido rechazada.
- Población: Estos resultados se aplican a todos los empleados de la corporación que estaban contratados durante la realización del experimento.



MAGNITUD DEL EFECTO: SIGNIFICATIVO VERSUS IMPORTANTE

El procedimiento que hemos seguido para examinar los resultados de un experimento consiste en evaluar primero directamente la hipótesis nula y después concluir, en forma indirecta, con respecto a la hipótesis alternativa. Si rechazamos la hipótesis nula, eso indica que los resultados son significativos. Lo que queremos decir en realidad con la expresión “significativos” es “estadísticamente significativos”, es decir, que los resultados probablemente no se han debido al azar, que la variable independiente ha producido un efecto real, y que si repetimos el experimento obtendremos de nuevo resultados que nos permitirán rechazar la hipótesis nula. Tal vez habría sido mejor emplear el término confiable a este respecto, en vez de significativo. Sin embargo, el uso de la palabra significativo está plenamente establecido, por lo cual tendremos que adaptarnos a él. Observe usted que no debemos confundir estadísticamente significativo con “importante” en sentido práctico o teórico. El hecho de que un efecto sea significativo desde el punto de vista de la estadística no nos informa si dicho efecto es importante. Por ejemplo, supongamos que el efecto real de la marihuana consistiera en aumentar el apetito, pero sólo en el equivalente de 10 calorías. Un diseño experimental cuidadoso y una muestra suficientemente grande nos permitirían detectar este efecto, aunque fuera tan pequeño. En ese caso, concluiríamos que los resultados son significativos (confiables), pero aún así necesitamos preguntar: “¿Qué tan importante es este efecto real?” Para la mayoría de los propósitos, salvo tal vez para algunos fines teóricos, la importancia de un efecto aumenta en proporción directa a la magnitud del mismo. Para consultar un análisis más detallado de este punto, véase “¿Cuál es la verdad? Mucho ruido y pocas nueces”, en el capítulo 15.

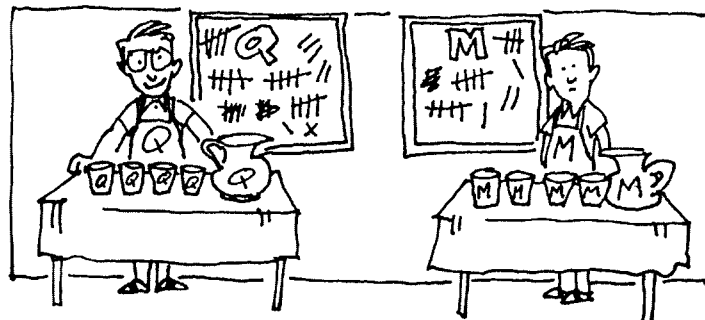


¿Azar o efecto real? 1

En la revista *Time* apareció un artículo acerca de la “Prueba de sabor del reto Pepsi”. Al otro lado del artículo apareció la publicidad de Pepsi que mostramos aquí. Los participantes en la prueba del sabor tomaban Coca-Cola en Michigan; se les pidió beber un vaso de Pepsi y otro de Coca-Cola para ver cuál de ellos preferían. Para evitar un sesgo obvio, los vasos no fueron marcados con los letreros “Coca” o “Pepsi”. En lugar de eso, para facilitar la administra-

ción “a ciegas” de las bebidas, los vasos de Coca-Cola fueron marcados con una “Q” y los de Pepsi con una “M”. Los resultados indicados en el anuncio fueron “más de la mitad de los consumidores de Coca-Cola en Michigan prefirieron

Pepsi”. Aparte de una posible preferencia real por Pepsi dentro de la población de consumidores de Coca-Cola que viven en Michigan, ¿cree usted que exista otra posible explicación para estos resultados muestrales?





¿Azar o efecto real? 1 (continuación)

Respuesta La explicación alternativa, más obvia, de estos resultados es que se deben solamente al azar: que en la población, las preferencias por Pepsi y por Coca-Cola son iguales ($P = 0.50$). Por supuesto, usted reconoce esto como la postulación de la hipótesis nula. Esta explicación podría y, desde nuestro punto de vista, debería ser descartada (dentro de los límites del error de tipo I) analizando los datos muestrales con la prueba de inferencia adecuada. Si los resultados realmente son significativos, no se necesita mucho espacio en un anuncio para comentarlo. Ésta es como muchas otras campañas publicitarias que establecen resultados muestrales que favorecen su producto, sin evaluar la posibilidad de que el azar sea una explicación razonable.

Por cierto, la Coca-Cola no sólo apeló "al azar", sino también afirmó que el estudio no era válido puesto que a mucha gente le gusta más la letra "M" que la "Q". La Coca-Cola realizó su propio estudio, colocando su bebida en ambos vasos "M" y "Q". La mayoría de la gente prefirió la bebida del vaso "M", aunque había Coca-Cola en ambos recipientes. Pepsi respondió con otro reto, sólo que esta vez utilizó las letras "S" y "L", de modo que Pepsi siempre estuviera en el vaso con la letra "L". Los resultados muestrales favorecieron de nuevo a Pepsi. Como era de esperar, los ejecutivos de Coca-Cola de nuevo reclamaron, argumentando

que había una preferencia por la letra "L". Entonces se consultó a una autoridad en el terreno de la motivación, quien dijo que no conocía ningún estudio que mostrara un sesgo a favor de la letra "L". Como un estadístico prometedor,

¿cómo diseñaría usted un experimento para determinar si existe una preferencia por Pepsi o por Coca-Cola en la población y, al mismo tiempo, eliminar la preferencia por los vasos como posible explicación?

Enfrenta el reto Pepsi.

Deja que tu gusto decida.

Prueba el sabor de Pepsi-Cola.

Tal vez hayas visto "El reto Pepsi" en la televisión. Es una prueba de gusto sencilla y directa, donde los consumidores de Coca-Cola prueban esta marca y Pepsi, sin saber cuál es cuál. Luego les preguntamos qué refresco prefieren.



Más de la mitad de los consumidores de Coca-Cola examinados en Michigan prefirieron Pepsi.

Cientos de consumidores de Coca-Cola fueron examinados en Michigan y se encontró que más de la mitad de las personas participantes en la prueba prefirieron el sabor de Pepsi.

Deja que tu gusto decida.

No te pedimos que creas en nuestra palabra, ni en la palabra de alguna otra persona. Sólo pruébala tú mismo. Enfrenta el reto Pepsi y deja que tu gusto decida.





¿Azar o efecto real? 2

En la investigación que mencionamos en la sección anterior de “¿Cuál es la verdad?” no se consignó ningún nivel de significación. A veces se informa que los experimentos contienen niveles de significación, pero aún así despiertan la sospecha de posibles errores tipo I. Considere el siguiente artículo publicado en un periódico acerca de ciertas investigaciones que se llevan a cabo en una importante universidad de Estados Unidos.

Estudio: Las mujeres levemente deprimidas viven más

Si usted desea una vida larga y feliz, piénselo una vez más.

Investigadores de la Universidad Duke observaron durante una década a más de 4000 personas de avanzada edad y descubrieron que las mujeres que viven más suelen estar ligeramente deprimidas. El estudio en cuestión fue publicado en el número de American Journal of Geriatric Psychiatry correspondiente a mayo-junio.

“No entendemos del todo este hallazgo”, dijo el Dr. Dan Blazer, profesor de psiquiatría y ciencia del comportamiento en Duke y coautor de dicho estudio. Este salto en materia de longevidad fue observado únicamente en mujeres. Las que estaban aquejadas de una depresión leve, certificada por medio de un cuestionario de diagnóstico, tuvieron 40% menos probabilidades de

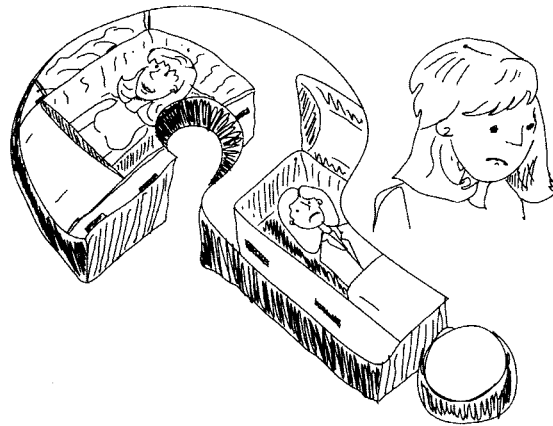
morir durante el periodo del estudio, que las mujeres que no sufrían depresión alguna o que las que padecían formas más severas de depresión. En el caso de los hombres, se observó que la depresión leve no producía efecto alguno sobre la mortalidad.

Blazer sospecha que las formas leves de la depresión podrían proteger contra la muerte, al hacer que las personas disminuyan su ritmo de vida, con lo cual pueden disponer de más tiempo para prestar atención a sus cuerpos y mentes.

Autor desconocido. Artículo tomado de *The Coloradoan*, 4 de mayo de 2002, p. A3.

A pesar de que en el artículo del periódico no se menciona ningún nivel de significación, supongamos que en el artículo publicado en el *American Journal of Geriatric Psychiatry* se dijo que, con un $\alpha = 0.05$, las mujeres que sufrían depresión leve fueron *significativamente* menos propensas a morir du-

rante el periodo de estudio, que las mujeres que no sufrían depresión alguna o que las que padecían otras formas más severas de depresión. Efectivamente, este resultado alcanzó la significación en el nivel de 0.05, pero todavía es un tanto cuestionable. En primer lugar, esto parece ser un resultado sorprendente e imprevisto que los autores no aciertan a entender, según lo han confesado. En segundo lugar, ni las mujeres sin depresión ni las que padecían depresiones más graves mostraron ese comportamiento. Por último, los hombres ligeramente deprimidos tampoco produjeron ese resultado. ¿A qué atribuye usted esto, al azar o a un efecto real? El argumento que se plantea aquí es que el simple hecho de que un resultado sea estadísticamente significativo no implica automáticamente que dicho resultado sea un efecto real. Por supuesto, una repetición habría ayudado a resolver la cuestión.





“Ningún producto es mejor que el nuestro”

Con frecuencia vemos anuncios donde se afirma algo sin presentar datos. “Ningún artículo es mejor para X que nuestro producto”. Un anuncio de Excedrin publicado recientemente en una revista de circulación nacional es un ejemplo de este tipo de anuncio. El anuncio en cuestión muestra una imagen de gran tamaño con un frasco de tabletas de Excedrin junto con las siguientes afirmaciones:

- “Nada de lo que usted pueda comprar es más fuerte.”
- “Nada de lo que usted pueda comprar funciona mejor.”
- “Nada le brinda mayor alivio.”

La pregunta es ¿cómo interpretamos estas afirmaciones? ¿Debemos correr a comprar Excedrin porque es más fuerte, funciona mejor y brinda mayor alivio que cualquier otro remedio disponible para el dolor de cabeza? Si existen datos experimentales que fundamenten estas afirmaciones, nos gustaría saber cuáles son los resultados. ¿Cuál es su opinión?

Respuesta Por supuesto, no sabemos qué ocurre en todos los casos; no pretendemos dirigir nuestros comentarios hacia algún anuncio específico. Sólo elegimos el anuncio de Excedrin como ejemplo de muchos otros. Sin embargo, no podemos menos que sospechar que en la mayor parte de los casos, si no es que en todos ellos, donde existen datos muestrales, los datos

reales revelan que no hay una diferencia significativa en el grado de eficacia del producto anunciado y el de los demás productos sometidos a prueba.

Para nuestro análisis, llamaremos “A” al producto del anuncio. Si los datos hubieran demostrado que “A” era mejor que los productos de la competencia, parecería razonable que el anunciante hubiera declarado de manera directa la superioridad de su producto, en vez de insinuarlo indirectamente haciendo una declaración más débil, al afirmar que ningún otro producto es mejor que el de ellos.

¿Por qué, entonces, se hace en el anuncio esa afirmación más débil? Tal vez porque los datos reales no demostraron que el producto “A” fuera superior a los demás. Lo más probable es que los datos muestrales hayan revelado que el producto “A” es igual o inferior a los otros, y que la prueba de inferencia no muestre una diferencia significativa entre los productos. Con estos da-

tos, en lugar de decir que la investigación ha demostrado que la eficacia de nuestro producto es inferior, o cuando mucho igual a la de los otros productos (lo cual naturalmente no ayudaría mucho a vender el producto “A”), enunciamos los resultados en esta forma más positiva, aunque engañosa desde nuestro punto de vista. Está claro que la declaración “Ningún otro producto es mejor que el nuestro en su eficacia” venderá más productos que “Todos los productos verificados fueron iguales en términos de eficacia”. Después de todo, si usted lee con cuidado la afirmación débil, advertirá que en realidad no dice que el producto “A” sea superior a los demás.

Así pues, a falta de datos en sentido contrario, creemos que la interpretación más precisa de la afirmación “Ningún otro producto es superior al nuestro en su eficacia” es que los productos son iguales en términos de eficacia.





Informes anecdóticos contra investigación sistemática

El siguiente artículo fue publicado en un número de *The New York Times* en la red. Resulta representativo de un suceso bastante común: el caso en que un testimonio individual acerca de un fenómeno entra en conflicto con la investigación sistemática sobre el mismo fenómeno. En este caso, el tema que se examina son los efectos de la hormona secretina sobre el autismo. Si después de leer el artículo alguien le preguntara a usted si la secretina produce un efecto benéfico sobre el autismo, ¿cuál sería su respuesta? ¿Se daría por satisfecho con los informes anecdóticos? ¿Le parecerían satisfactorios los resultados negativos de la investigación sistemática? ¿Cómo podría resolver esta cuestión?

Fármaco empleado contra el autismo es sometido a severos ataques en un nuevo estudio

Por SANDRA BLAKESLEE
Una hormona que fue aclamada en los medios informativos y en la Internet como una posible cura contra el autismo no produjo mejores resultados que el agua con sal, según han informado los científicos, cuando fue sometido a su primera prueba clínica controlada.

El estudio, descrito en el número del New England Journal of Medicine que se publicó el jueves pasado, es uno entre una docena de medidas patrocinadas por los Institutos Nacionales de Salud para poner a prueba la hormona secretina, la cual acaparó la atención a raíz de que un niño de tres años de edad, de Nueva Hampshire, mostró una rápida mejoría después de haber ingerido el fármaco en 1996.

“La secretina es un producto de uso generalizado en todo el país”, declaró la Dra. Marie Bristol-Power, asistente especial de programas contra el autismo en el Instituto de Salud Infantil y Desarrollo Humano, dependiente de los Institutos Nacionales de Salud. “Los padres necesitan saber si esto es algo prometedor que vale la pena continuar o si alguien se está aprovechando de ellos.”

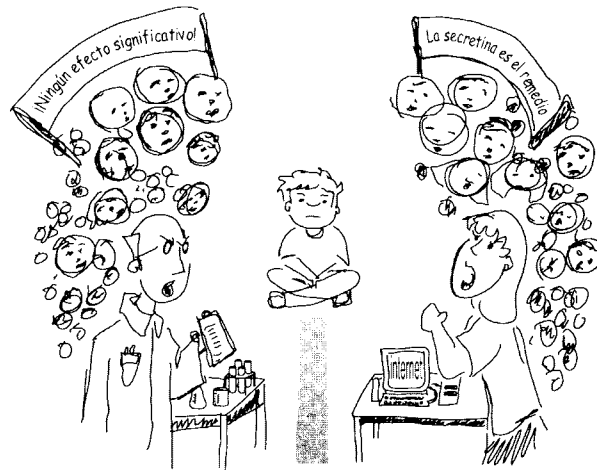
El estudio no fue definitivo, explicó la Dra. Bristol-Power, y los investigadores sabrán más a mediados del año próximo, cuando otros estudios se hayan llevado a cabo.

Los defensores de la secretina se apresuraron a desmentir el nuevo estudio, diciendo que en él no se había realizado una prueba imparcial de dicho fármaco.

“Nadie ha afirmado jamás que la secretina sea un remedio milagroso para el autismo”, dijo Victoria Beck, la madre del niño que mejoró después de tomar ese producto.

La Sra. Beck, que es autora de un libro sobre la secretina, dijo que la hormona representa “una pista que merece ser objeto de una cuidadosa investigación y no del rechazo de los expertos en autismo, quienes tienen intereses creados y desean conservar las cosas tal como las han mantenido en el pasado”.

El autismo es un grave trastorno cerebral que se inicia en la infancia e impide que los niños se desarrollen normalmente. Los síntomas, que pueden ser entre leves y severos, son: la incapacidad de comunicarse, la renuencia a establecer contacto visual, y diversas formas de comportamiento repetitivo, como golpearse la cabeza o dar palmadas, y una





Informes anecdóticos contra investigación sistemática (continuación)

obsesión por actividades o intereses insólitos.

Medio millón de estadounidenses padecen este trastorno, el cual es tratado de ordinario mediante terapia educacional y conductual; si bien es cierto que la terapia intensiva ha resultado eficaz con algunos pacientes, las dos terceras partes de los adultos que padecen autismo no pueden tener una vida normal independiente.

La secretina es una hormona que se elabora en el intestino y que estimula la liberación de fluidos digestivos. Al hijo de la Sra. Beck, Parker, se le aplicó una inyección para investigar la causa de su diarrea crónica. Al poco tiempo, su comportamiento mejoró: ya era capaz de quedarse quieto y alternaba mejor con los adultos.

Cuando su caso fue divulgado por televisión, en los periódicos y en la Internet, el año pasado, se inició una gran agitación entre los padres de familia que deseaban conseguir el fármaco.

La Dra. Bristol-Power declaró que los suministros de secretina no tardaron en agotarse y que entonces surgió un mercado negro que dio

lugar a la adulteración de la secretina, extorsiones en materia de precios y otras formas de acaparamiento; la Sra. Beck replicó que estas versiones eran exageradas y que, en la actualidad, para nadie sería un problema conseguir secretina al precio justo, que es de unos \$200 por inyección.

Las anécdotas sobre la secretina están proliferando en la Internet. Algunos padres consignan efectos espectaculares, otros no perciben ningún efecto y algunos informan que los síntomas de sus hijos han empeorado. En la controversia en torno a la secretina, ambos bandos estiman que entre 2000 y 3000 niños han tomado o están tomando el fármaco.

El estudio controlado de la secretina fue dirigido por el Dr. James Bodfish, del Centro del Oeste de Carolina y la Universidad de Carolina del Norte, en colaboración con el Dr. Adrian Sandler, del Centro para el Desarrollo Infantil dependiente del Hospital de Rehabilitación Thoms en Asheville, Carolina del Norte.

En total, 56 niños autistas, de 3 a 14 años de edad, participaron en el estudio. La mitad de ellos recibieron una sola inyección de secretina humana sintética y al resto se les inyectó agua salina. Los investigadores los sometieron a una batería de pruebas conductuales a intervalos de un día, una semana y cuatro semanas después del tratamiento. Los dos grupos no mostraron diferencias de carácter estadístico, según informó el Dr. Sandler.

Un tercio de los niños mejoraron en la misma medida, tanto en el grupo placebo como en el de tratamiento. A juicio del Dr. Sandler, la única explicación posible es que existe un "efecto placebo significativo".

"Los padres están sumamente interesados en la posibilidad de que surjan nuevos tratamientos y tienen grandes expectativas", explicó. "Ellos buscan mejoras sutiles.

Existe una gran variabilidad en el comportamiento que se observa de uno a otro día en los niños aquejados de autismo, y resulta fácil atribuir las variaciones normales al efecto de la secretina."

La Sra. Beck discute esta interpretación y asegura que, al concentrarse en pruebas conductuales estándar, los investigadores no han examinado la amplia gama de respuestas fisiológicas que presentan los niños cuando se les administra la secretina, y añadió que esas respuestas tienen que ser estudiadas.

"En verdad me molesta que se describa a los padres como un grupo de gente desesperada que tiene la cabeza hueca", afirmó. "Según se dice, no somos capaces de percibir la diferencia entre una mejoría y un espejismo. Sin embargo, cuando un niño deja de vomitar, duerme toda la noche y establece contacto visual, sabemos que algo está sucediendo."

Fuente: Derechos de autor © 1999 por la New York Times Co. Reproducido con autorización.

RESUMEN

En este capítulo abordamos el tema de la prueba de hipótesis, empleando la prueba del signo como vehículo. Esta prueba se utiliza junto con el diseño de medidas repetidas. Las características esenciales de este diseño son que existen parejas de datos entre dos condiciones y que se analizan las diferencias entre ellas.

En cualquier experimento cuyo propósito sea poner a prueba una hipótesis, siempre existen dos hipótesis que compiten por explicar los resultados obtenidos: la hipótesis alternativa y la hipótesis nula. La hipótesis alternativa especifica que la variable independiente es la causante de la diferencia observada entre los valores de los datos bajo una y otra de las condiciones. La hipótesis alternativa puede ser de tipo direccional o no direccional. Resulta legítimo utilizar una hipótesis direccional cuando se cuenta con una buena base teórica y con una buena evidencia de respaldo en la literatura sobre el tema. Si el propósito del experimento es determinar hechos básicos, generalmente se debe utilizar una hipótesis no direccional. Se evalúan también una hipótesis alternativa direccional, con un valor de probabilidad de una cola, y una hipótesis no direccional con un valor de probabilidad de dos colas.

La hipótesis nula es la contraparte lógica de la hipótesis alternativa, de manera que si la hipótesis nula es falsa, entonces la hipótesis alternativa debe ser verdadera. Si la hipótesis alternativa es no direccional, la hipótesis nula especifica que la variable independiente no produce efecto alguno sobre la variable dependiente. Si la hipótesis alternativa es direccional, entonces la hipótesis nula establece que la variable independiente no influye sobre la variable dependiente en la dirección especificada.

Al evaluar los datos de un experimento, nunca examinamos en forma directa la hipótesis alternativa. Siempre evaluamos primero la hipótesis nula, suponiendo que el azar es la única causa de las diferencias observadas en los datos entre las condiciones. Al hacer esta evaluación, calculamos la probabilidad de conseguir el resultado ya obtenido, o un resultado aún más extremo, si el azar es el único responsable. Si la probabilidad así obtenida es igual o menor que el nivel alfa, consideramos que la hipótesis nula es una explicación no razonable y, por lo tanto, rechazamos la hipótesis nula. Concluimos aceptando la hipótesis alternativa porque es la única explicación que nos queda. Si la probabilidad obtenida es mayor que el nivel

alfa, conservamos la hipótesis nula. Se la sigue considerando como una explicación razonable de los datos. Por supuesto, si la hipótesis nula no es rechazada, entonces la hipótesis alternativa no puede ser aceptada. La conclusión solamente se aplica en forma válida a la población de la cual fue extraída de manera aleatoria la muestra. Debemos tener cuidado para distinguir entre lo que es "significativo desde el punto de vista de la estadística" y lo que es práctica o teóricamente "importante".

El nivel alfa se establece, por lo general, como 0.05 o 0.01 para minimizar la probabilidad de incurrir en un error de tipo I. Este error se presenta cuando la hipótesis nula es rechazada, siendo que es verdadera. El nivel alfa limita la probabilidad de cometer un error de tipo I. También es posible incurrir en un error de tipo II. Esto sucede cuando conservamos la hipótesis nula y ésta es falsa. Beta se define como la probabilidad de cometer un error de tipo II. Cuando alfa se hace más restrictiva, beta aumenta. Al minimizar alfa y beta es posible tener una alta probabilidad de obtener una conclusión correcta de un experimento, independientemente de que H_0 o H_1 sean verdaderas. Un resultado significativo indica que se trata de un resultado confiable, pero proporciona poca información acerca de la magnitud del efecto. Cuanto mayor sea el efecto, tanto más probable será que dicho efecto sea importante.

Al analizar los resultados de un experimento por medio de la prueba del signo, pasamos por alto la magnitud de las diferencias entre los mismos y solamente consideramos su dirección. Existen únicamente dos resultados posibles para cada sujeto: un signo positivo o uno negativo. Sumamos los signos positivos y negativos de todos los sujetos, y el total así obtenido es el recuento de signos positivos y negativos. Para verificar la hipótesis nula, calculamos la probabilidad de obtener el número total de signos positivos, o un número de signos positivos aún más extremo, en caso de que sólo intervenga el azar. La distribución binomial, en la que P (la probabilidad de un signo positivo) = 0.50 y N = el número de diferencias entre los datos, es adecuada para tomar esta decisión. Presentamos un ejercicio ilustrativo y varios problemas de práctica para mostrar la forma de evaluar la hipótesis nula por medio de la distribución binomial.

TÉRMINOS IMPORTANTES

Beta (β) (página 221)	Error de tipo I (página 220)	Magnitud importante de un efecto (página 232)
Conservación de la hipótesis nula (página 220)	Error de tipo II (página 220)	Nivel alfa (α) (página 218, 222)
Decisión correcta (página 221)	Estado real (página 221)	Probabilidad de dos colas (página 224)
Diseño de grupos correlacionados (página 217)	Hipótesis alternativa (H_1) (página 218)	Probabilidad de una cola (página 225)
Diseño de medidas repetidas (página 217)	Hipótesis direccional (página 218)	Prueba del signo (página 216)
Diseño de medidas replicadas (página 217)	Hipótesis no direccional (página 218)	Rechazo de la hipótesis nula (página 220)
Distribución binomial (página 219)	Hipótesis nula (H_0) (página 218)	Significativo (página 219)
	Imposibilidad de rechazar la hipótesis nula (página 219)	
	Magnitud del efecto (página 232)	

PREGUNTAS Y PROBLEMAS

1. Describa o explique brevemente cada uno de los conceptos de la sección “Términos importantes”.
2. Describa en forma breve el procedimiento de la prueba de hipótesis. Asegúrese de incluir las hipótesis alternativa y nula, la regla de decisión, el tipo posible de error y la población a la cual pueden generalizarse los resultados.
3. Explique en sus propias palabras por qué es importante conocer los tipos de errores que podríamos cometer al rechazar o al no rechazar la hipótesis nula.
4. ¿Difiere la hipótesis nula para una H_1 no direccional, de la hipótesis nula para una H_1 direccional? Explique.
5. ¿Bajo qué condiciones es válido utilizar una H_1 direccional? ¿Por qué no es válido utilizar una H_1 direccional sólo porque el experimentador ha tenido una “corazonada” acerca de la dirección?
6. Si la probabilidad obtenida en un experimento es igual a 0.0200, ¿significa esto que la probabilidad de que H_0 sea verdadera es igual a 0.0200? Explique.
7. Analice la diferencia entre “significativo” e “importante”. Incluya en su exposición la “magnitud del efecto”.
8. ¿Qué consideraciones hay que hacer para determinar cuál es el mejor nivel alfa que se puede usar? Comente.
9. Una primatóloga tiene la impresión de que los monos rhesus son curiosos. Ella considera que, si está en lo cierto, esos primates preferirán una estimulación novedosa a una estimulación repetitiva. Entonces la investigadora realiza un experimento en el cual 12 monos rhesus son elegidos al azar entre la colonia de macacos que posee la universidad, y

luego se les enseña a oprimir dos barras. Cuando la barra 1 es oprimida se produce siempre el mismo sonido, mientras que al oprimir la barra 2 se produce un nuevo sonido en cada ocasión. Una vez que han aprendido a oprimir las barras, los monos son sometidos a una prueba de 15 minutos, durante los cuales tienen libre acceso a ambas barras. Se registra entonces el número de veces que oprimen cada barra durante esos 15 minutos. Los datos resultantes se presentan a continuación:

Sujeto	Barra 1	Barra 2
1	20	40
2	18	25
3	24	38
4	14	27
5	5	31
6	26	21
7	15	32
8	29	38
9	15	25
10	9	18
11	25	32
12	31	28

- a. ¿Cuál es la hipótesis alternativa? En este caso, supongamos que una hipótesis no direccional resulta apropiada, debido a que existe poca evidencia empírica que garantice una hipótesis direccional.
- b. ¿Cuál es la hipótesis nula?

- c. Utilice $\alpha = 0.05_{2 \text{ colas}}$ para obtener una conclusión.
- d. ¿Qué error se podría cometer con la conclusión del inciso c?
- e. ¿A qué población se aplica la conclusión? cognitiva, biológica
10. Un director de cierta escuela está interesado en un nuevo método de enseñanza para las ciencias sociales a nivel secundaria, pues piensa que con éste aumentará la cantidad de material aprendido. Para verificar ese método, el director realiza el siguiente experimento. Los estudiantes de secundaria del distrito escolar son agrupados en parejas según CI y sus calificaciones anteriores. Se eligen al azar 20 parejas para el experimento. Se elige al azar un miembro de cada pareja para formar un grupo que recibirá el nuevo método de enseñanza, mientras que el otro miembro de la pareja pasa a otro grupo que recibe la instrucción común. Al final del curso, todos los alumnos presentan un mismo examen final. Los resultados obtenidos son los siguientes:

Pareja número	Nuevo método	Instrucción estándar
1	95	83
2	75	68
3	73	80
4	85	82
5	78	84
6	86	78
7	93	85
8	88	82
9	75	84
10	84	68
11	72	81
12	84	91
13	75	72
14	87	81
15	94	83
16	82	87
17	70	65
18	84	76
19	72	63
20	83	80

- a. ¿Cuál es la hipótesis alternativa? Utilice una hipótesis direccional.
- b. ¿Cuál es la hipótesis nula?

- c. Utilice $\alpha = 0.05_{1 \text{ cola}}$ para sacar una conclusión.
- d. ¿Qué error se podría cometer con la conclusión del inciso c?
- e. ¿A qué población se aplica la conclusión? educación
11. Una fisióloga tiene la idea de que la hormona angiotensina II es un factor importante para la regulación de la sed. Para investigar esa idea, ella elige al azar una muestra de 16 ratas de la compañía de medicamentos en la cual trabaja y las encierra en jaulas individuales, dándoles libre acceso a la comida y al agua. Después que los roedores se aclimatan a sus nuevos "hogares", la experimentadora mide la cantidad de agua que cada rata bebe en un periodo de 20 minutos. Luego, la investigadora inyecta a cada animal, por vía intravenosa, una concentración conocida (100 microgramos por kilogramo) de angiotensina II. Posteriormente, las ratas regresan a su jaula y se mide la cantidad de agua que beben en otro periodo de 20 minutos. Los resultados se presentan en la siguiente tabla. Los datos se miden en mililitros consumidos durante 20 minutos.

Sujeto	Antes de la inyección	Después de la inyección
1	1.2	11.3
2	0.8	10.7
3	0.5	10.3
4	1.3	11.5
5	0.6	9.6
6	3.5	3.3
7	0.7	10.5
8	0.4	11.4
9	1.1	12.0
10	0.3	12.8
11	0.6	11.4
12	0.3	9.8
13	0.5	10.6
14	4.1	3.2
15	0.4	12.1
16	1.0	11.2

- a. ¿Cuál es la hipótesis alternativa no direccional?
- b. ¿Cuál es la hipótesis nula?
- c. Utilice $\alpha = 0.05_{2 \text{ colas}}$ para sacar una conclusión. Suponga que la inyección no influye en el comportamiento de consumo de agua.

- d. ¿Qué error se podría cometer con la conclusión del inciso c?
- e. ¿A qué población se aplica la conclusión? biológica
12. Un importante fabricante de dentífricos anuncia que, según se afirma en un estudio médico reciente, 70% de las personas que fueron sometidas a una prueba tenían los dientes más brillantes después de usar su producto (llamado Very Bright), en comparación con la gente que usó la marca de la competencia (Marca X). El anuncio dice a continuación: “por eso, use usted Very Bright y tendrá los dientes más brillantes”. En realidad, los datos sobre los cuales se basaron las afirmaciones anteriores fueron reunidos a partir de una muestra aleatoria de 10 empleados de la planta de Very Bright en Pasadena. En el experimento, cada trabajador usó las dos pastas de dientes. La mitad de los empleados usó la marca X durante 3 semanas, y después la marca Very Bright durante el mismo periodo. La otra mitad usó Very Bright primero y después la marca X. Se realizó una prueba de brillo dental al final de cada periodo de tres semanas. Por lo tanto, había dos datos por empleado: uno para la prueba de brillantez después de haber utilizado la marca X y otro para dicha prueba después de usar Very Bright. La siguiente tabla muestra los datos (cuanto más alto sea el número, tanto mayor fue el brillo dental):

Sujeto	Very Bright	Marca X
1	5	4
2	4	3
3	4	2
4	2	3
5	3	1
6	4	1
7	1	3
8	3	4
9	6	5
10	6	4

- a. ¿Cuál es la hipótesis alternativa? Utilice una hipótesis direccional.
- b. ¿Cuál es la hipótesis nula?
- c. Utilice $\alpha = 0.05$ para sacar una conclusión.
- d. ¿Qué error se podría cometer con la conclusión del inciso c?
- e. ¿A qué población se aplica la conclusión?

- f. ¿Es engañosa la publicidad? ¿/0
13. Un investigador está interesado en determinar si la acupuntura influye en el grado de tolerancia al dolor. Para averiguarlo, realiza un experimento en el cual elige al azar a 15 estudiantes de un numeroso grupo de voluntarios universitarios, estudiantes de licenciatura. Cada sujeto participa en dos situaciones. En las dos, a cada individuo se le aplica una descarga eléctrica de corta duración en la pulpa de un diente. La intensidad de la descarga se mantiene en un nivel que produce un dolor moderado al sujeto, el cual no se encuentra anestesiado. Después de cada descarga, cada individuo califica el nivel de dolor que ha percibido, en una escala de 0 a 10, donde 10 es el nivel máximo. En la situación experimental, cada voluntario recibe el tratamiento adecuado de acupuntura antes de someterse a la descarga. La situación de control es lo más parecida posible a la situación experimental, excepto que en ella se administra un tratamiento placebo en lugar de la acupuntura. Las dos situaciones se realizan en días separados, a la misma hora del día. La siguiente tabla muestra los datos obtenidos sobre el dolor.
- a. ¿Cuál es la hipótesis alternativa? Suponga que resulte apropiada una hipótesis no direccional.
- b. ¿Cuál es la hipótesis nula?
- c. Utilice $\alpha = 0.05$ para sacar una conclusión.
- d. ¿Qué error se podría cometer con la conclusión del inciso c?
- e. ¿A qué población se aplica la conclusión?

Sujeto	Acupuntura	Placebo
1	4	6
2	2	5
3	1	5
4	5	3
5	3	6
6	2	4
7	3	7
8	2	6
9	1	8
10	4	3
11	3	7
12	4	8
13	5	3
14	2	5
15	1	4

cognitivo, salud

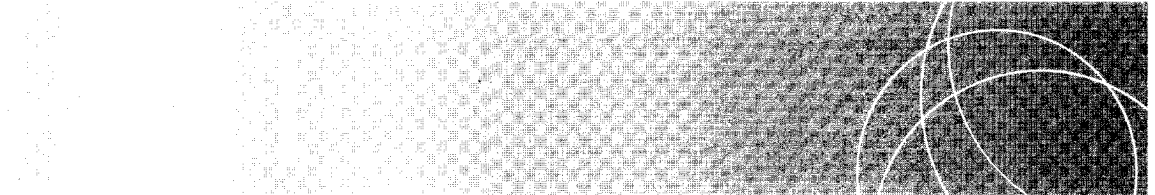
NOTAS

- 10.1 Si la hipótesis nula es falsa, entonces el azar no explica en los resultados. Estrictamente, esto significa que algo sistemático difiere en ambos grupos. Lo ideal es que la única diferencia sistemática se deba a la variable independiente. En ese caso, afirmamos que si la hipótesis nula es falsa, entonces la hipótesis alternativa debe ser verdadera. Sin embargo, en la práctica, el lector debe estar consciente de que es difícil realizar el experimento perfecto. En consecuencia, además de la hipótesis alternativa, con frecuencia existen otras explicaciones posibles de las diferencias sistemáticas. Por consiguiente, cuando decimos “aceptamos H_1 ”, debemos estar conscientes de que puede haber otras razones para explicar la diferencia sistemática.
- 10.2 Si la hipótesis alternativa es direccional, entonces la hipótesis nula afirma que la variable indepen-

diente no influye en la dirección especificada por la hipótesis alternativa. Esto es cierto en la enorme mayoría de los experimentos que se realizan. En ciertas ocasiones, se lleva a cabo un experimento en el cual la hipótesis alternativa no sólo especifica la dirección, sino también la magnitud del efecto. Por ejemplo, en el caso del experimento de la marihuana, una hipótesis alternativa de este tipo sostendría que “la marihuana aumenta el apetito de tal manera que el consumo alimenticio promedio diario se incrementa en más de 200 calorías”. La hipótesis nula para esta hipótesis alternativa sería: “la marihuana aumenta el apetito de tal manera que el consumo alimenticio promedio diario se incrementa en 200 calorías o menos”.

Si desea encontrar problemas de práctica adicionales totalmente resueltos sobre los temas que acaba de estudiar, busque Chapter 10 en el material adicional del libro, disponible en inglés en: <http://www.thomsonlearning.com.mx>.

Potencia



Introducción

¿Qué es potencia?

P_{nula} y P_{real}

P_{real} : Una medida de la magnitud y dirección del efecto real

Análisis de potencia y el experimento del SIDA

Efecto de N y magnitud del efecto real

Potencia y beta (β)

Potencia y alfa (α)

Alfa-beta y realidad

Interpretación de resultados no significativos

Cálculo de potencia

¿CUÁL ES LA VERDAD? Astrología y ciencia

Resumen

Términos importantes

Preguntas y problemas

Notas



INTRODUCCIÓN

En el capítulo 10 vimos que al someter a prueba una hipótesis podemos incurrir en dos clases de errores. A éstos los hemos llamado errores de tipo I y de tipo II. Comentamos también que el nivel alfa limita la probabilidad de cometer un error de tipo I. Cuando los experimentadores establecen alfa en 0.05 o 0.01, limitan la probabilidad de rechazar erróneamente la hipótesis nula a esos niveles bajos. Bien, pero ¿qué se puede decir de los errores tipo II? Hemos definido (β) como la probabilidad de cometer un error de tipo II. Más adelante, en este capítulo, veremos que $\beta = 1 -$ potencia. Al maximizar la

potencia minimizamos beta, lo cual significa que minimizamos la probabilidad de cometer un error de tipo II. Así pues, la potencia es un tema muy importante.



¿QUÉ ES POTENCIA?

Desde un punto de vista conceptual, la potencia de un experimento es una medida de la sensibilidad del experimento para detectar un *efecto real* de la variable independiente. Por “un efecto real de la variable independiente” nos referimos a un efecto que produzca un cambio en la variable dependiente. Si la variable independiente no genera un cambio en la variable dependiente, entonces carece de efecto sobre ella y por eso decimos entonces que la variable independiente no tiene un efecto real.

Cuando analizamos los datos obtenidos de un experimento, “detectamos” un efecto real de la variable independiente cuando rechazamos la hipótesis nula. Así, la potencia se define en términos del rechazo de H_0 .

definición ■ Desde un punto de vista matemático, **la potencia de un experimento** se define como la probabilidad de que los resultados del experimento permitan rechazar la hipótesis nula si la variable independiente tiene un efecto real.

Otra forma de enunciar esa definición consiste en decir que la potencia de un experimento es igual a la probabilidad de que los resultados del mismo permitan el rechazo de la hipótesis nula cuando la hipótesis nula es falsa.

Como la potencia es una probabilidad, su valor puede fluctuar entre 0.00 y 1.00. Mientras más alta sea la potencia, más sensible será el experimento para detectar un efecto real de la variable independiente. Los experimentos con potencias tan altas como 0.80 o más son muy deseables, pero rara vez los vemos en las ciencias del comportamiento. Los valores de 0.40 a 0.60 son mucho más comunes. Determinar la potencia de un experimento es especialmente útil, 1) cuando se elabora el diseño inicial del experimento y 2) cuando se interpretan los resultados de experimentos en los que no se logra detectar ningún efecto real de la variable independiente (es decir, experimentos en los que se conserva H_0).



P_{nula} Y P_{real}

Cuando se calcula la potencia de un experimento utilizando la prueba de los signos, resulta útil saber distinguir entre la P_{nula} y la P_{real} .

- definiciones**
- P_{nula} es la probabilidad de obtener un signo positivo con cualquiera de los sujetos de la muestra del experimento cuando la variable independiente no produce efecto alguno.
 - P_{real} es la probabilidad de obtener un signo positivo con cualquiera de los sujetos de la muestra del experimento cuando la variable independiente produce un efecto real.

P_{nula} equivale siempre a 0.50. En los experimentos donde H_1 es no direccional, P_{real} es igual a cualquiera de los otros valores posibles de P (es decir, cualquier valor de P que no sea igual a 0.50).

P_{real} : Una medida de la magnitud y dirección del efecto real

El valor real de P_{real} dependerá de la magnitud y dirección del efecto real. Para ilustrar este punto, consideraremos nuevamente el experimento de la marihuana presentado en el capítulo 10. (Use como referencia la figura 11.1 para el resto de esta exposición.) Supongamos, por el momento, que el experimento de la marihuana fue realizado con toda la población de los 10 000 pacientes de SIDA que son atendidos en el hospital donde usted trabaja, y no sólo con la muestra de 10. Si el efecto de la marihuana consiste en provocar un incremento en el apetito y si la magnitud de su efecto real es lo suficientemente grande para contrarrestar a todas las variables que podrían influir para reducir el apetito, entonces obtendríamos signos positivos con los 10 000 pacientes. Por consiguiente, en ese caso tendríamos 10 000 signos positivos y 0 signos negativos en la población. Así pues, para esta magnitud y dirección del efecto de la marihuana, $P_{real} = 1.00$. Esto se debe a que todos los signos de la población son positivos y los puntajes de los 10 sujetos del experimento real tendrían que ser una muestra aleatoria de esta población de puntajes. Ahora podemos profundizar más en la definición de P_{real} .

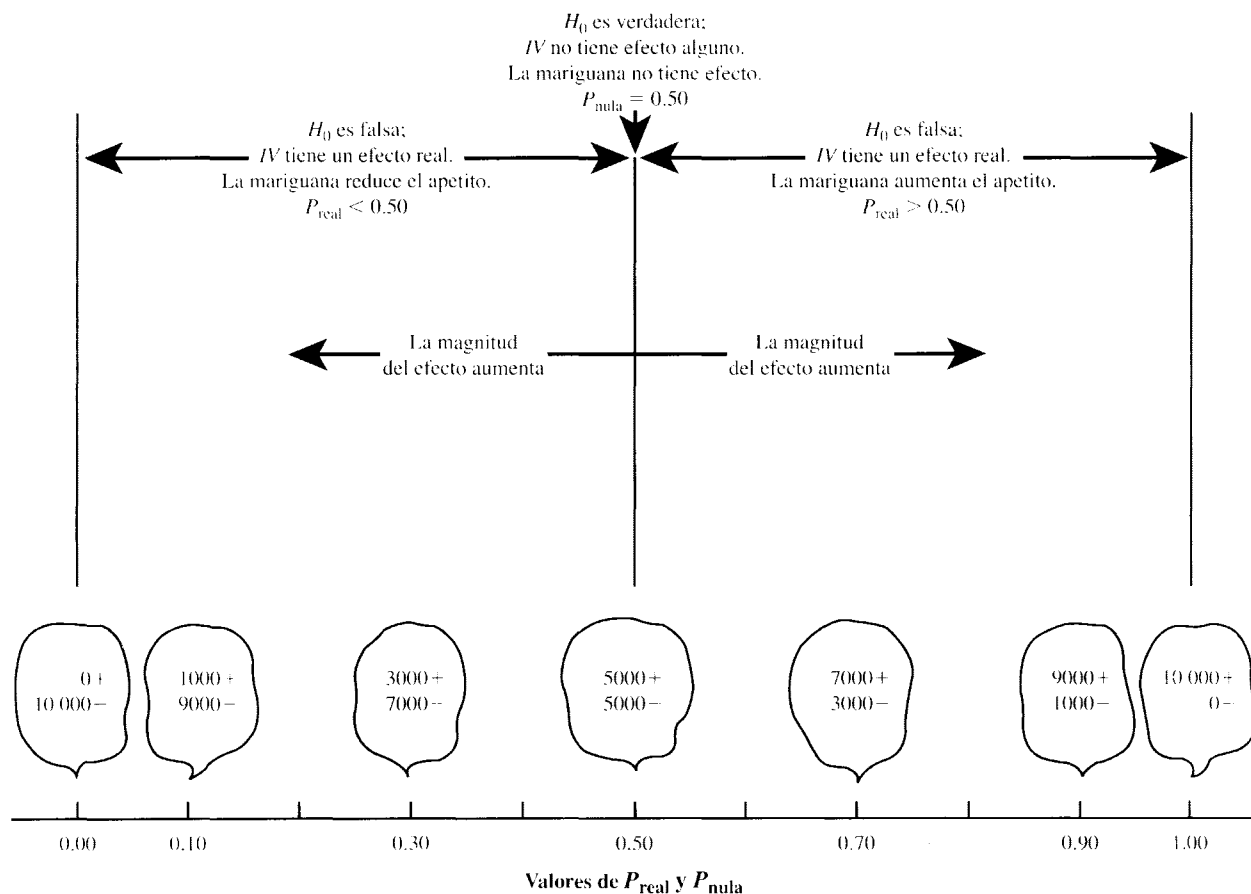


figura 11.1 Relación entre la hipótesis nula, la magnitud del efecto de la marihuana y los valores de P para una H_1 no direccional.

definición ■ *Tal como lo definimos anteriormente, P_{real} es la probabilidad de obtener un signo positivo con cualquiera de los sujetos de la muestra del experimento si la variable independiente tiene un efecto real. Sin embargo, significa también la proporción de signos positivos que habría en la población si el experimento se realizara con toda la población y si la variable independiente tiene un efecto real.*

Por supuesto, el valor de P_{real} es el mismo ya sea que se lo defina en términos de la proporción de signos positivos en la población o como la probabilidad de obtener un signo positivo con cualquier sujeto de la muestra. Regresemos a nuestra exposición de P_{real} y la magnitud del efecto de la variable independiente.

Si la marihuana aumenta el apetito, pero no con tanta intensidad como para producir solamente signos positivos —digamos, que sólo produzca 9 signos positivos por cada 1 negativo— en la población, habría 9000 signos positivos y 1000 signos negativos y $P_{\text{real}} = 0.90$.^{*} Si el efecto de aumento que la marihuana produce fuera todavía más pequeño —digamos, de 7 signos positivos por cada 3 signos negativos— la población tendría 7000 signos positivos y 3000 signos negativos. En este caso, $P_{\text{real}} = 0.70$. Finalmente, si la marihuana no tuviera efecto alguno sobre el apetito, entonces habría 5000 signos positivos y 5000 negativos, y $P_{\text{nula}} = 0.50$. Por supuesto, ésta es la predicción por el mero azar.

Por otra parte, si la marihuana reduce el apetito, podríamos esperar menos signos positivos que negativos. Por ejemplo, si el efecto de disminución del apetito es suficientemente grande, entonces todos los signos serán negativos (10 000 signos negativos y 0 positivos) en la población y $P_{\text{real}} = 0.00$. Un efecto reductor de menor magnitud, por ejemplo, un efecto que produjera 1000 signos positivos y 9000 negativos, daría lugar a una $P_{\text{real}} = 0.10$. Un efecto reductor del apetito aún más débil —por ejemplo, 3 signos positivos por cada 7 negativos— daría lugar a una $P_{\text{real}} = 0.30$. Si el efecto de disminución del apetito se siguiera debilitando todavía más, volveríamos finalmente a la especificación de la hipótesis nula de $P_{\text{nula}} = 0.50$ (la marihuana no tiene efecto alguno).

En la exposición anterior pudimos ver que P_{real} es una medida de la magnitud y la dirección del efecto real de la variable independiente. Cuanto más alejada esté P_{real} de 0.50, tanto mayor será la magnitud del efecto real. Resulta que la potencia del experimento varía en proporción a la magnitud de efecto real. Así pues, al realizar un análisis de potencia con la prueba de signos, debemos considerar todos los valores de P_{real} que puedan ser de interés.



ANÁLISIS DE POTENCIA Y EL EXPERIMENTO DEL SIDA

Suponga que usted está planeando un experimento para poner a prueba la hipótesis de que “la marihuana influye en el apetito de los pacientes de SIDA”. Su plan consiste en seleccionar al azar a cinco pacientes de SIDA, entre la población de enfermos de SIDA de su hospital, y va a realizar el experimento tal como lo hemos descrito antes. En virtud de que desea limitar a un nivel bajo la probabilidad de rechazar erróneamente la hipótesis nula, establece $\alpha = 0.05_{2\text{-colas}}$. Considerando este nivel tan riguroso de alfa, si usted rechaza H_0 tendrá un grado razonable de confianza en que sus resultados se deben a la marihuana y no al azar. Sin embargo, ¿cuál es la probabilidad de que usted rechace la hi-

^{*}El lector deberá observar que, a pesar de que hay signos negativos en la población, hemos supuesto que el efecto de la marihuana es el mismo en todos los sujetos (es decir, que provoca un aumento del apetito en todos ellos). Se supone que los signos negativos se presentan a causa de variables que intervienen en forma aleatoria para reducir el apetito.

pótesis nula como resultado de este experimento? Para responder esta pregunta, debemos determinar primero qué resultados de la muestra, si los hubiera, permitirán que H_0 sea rechazada. Los resultados más favorables para rechazar la hipótesis nula son que todos los signos sean positivos o que todos sean negativos. Supongamos que usted obtuvo el resultado más fuerte posible: solamente signos positivos en la muestra. ¿Podría entonces rechazar H_0 ? Como quiera que H_1 no es direccional, una evaluación de dos colas resulta apropiada en este caso. Con $N = 5$ y $P_{nula} = 0.50$, tomadas de la tabla B del apéndice D.

$$\begin{aligned} p(5 \text{ signos positivos o } 5 \text{ negativos}) &= \\ p(5 \text{ signos positivos o } 0 \text{ positivos}) &= 0.0312 + 0.0312 \\ &= 0.0624 \end{aligned}$$

Datos de la tabla B

N	Número de eventos P	
	0	0.50
5	0	0.0312
	5	0.0312

Como 0.0624 es mayor que alfa, si obtenemos estos resultados en el experimento, deberemos concluir que es necesario conservar H_0 . Así pues, aun cuando los resultados fueran los más favorables posibles para rechazar H_0 ¡ni aun en ese caso podríamos rechazarla!

Examinemos la situación un poco más de cerca. Supongamos que, en verdad, la marihuana produce un efecto muy intenso sobre el apetito ya que lo incrementa a tal grado que si el experimento se realizara con toda la población, se obtendrían únicamente signos positivos. Por ejemplo, si la población fuera de 10 000 pacientes, habría 10 000 signos positivos. Los cinco puntajes de la muestra serían una muestra aleatoria de esta población de puntajes, y la muestra sólo contendría signos positivos. Sin embargo, acabamos de determinar que, aun con cinco signos positivos en la muestra, no sería posible rechazar la hipótesis nula. En consecuencia, no importa cuán grande pudiera ser en realidad el efecto de la marihuana, no tendríamos posibilidad de rechazar H_0 . Con $N = 5$ y $\alpha = 0.05_{2\text{colas}}$, con ningún resultado de la muestra sería permisible rechazar H_0 . Éste es el experimento más insensible que puede existir. La potencia ha sido definida como la probabilidad de rechazar la hipótesis nula si la variable independiente produce un efecto real. En este experimento, la probabilidad de rechazar la hipótesis nula es cero, no importa cuán grande sea en realidad el efecto de la variable independiente. Así pues, la potencia de este experimento es cero para todos los valores de P_{real} . Podemos conceder muy poco valor a los resultados de un experimento tan insensible.

Efecto de N y magnitud del efecto real

Supongamos a continuación que N se incrementa a 10. ¿Existen ahora resultados muestrales que nos permitan rechazar la hipótesis nula? La solución se presenta en la tabla 11.1. Si el resultado de la muestra es 0 signos positivos, a partir de la tabla B, con $N = 10$ y usando $P_{nula} = 0.50$, $p(0 \text{ o } 10 \text{ signos positivos}) = 0.0020$. Observe que incluimos 10 signos positivos porque la hipótesis alternativa es no direccional, lo cual requiere una evaluación de dos colas. Puesto que 0.0020 es menor que alfa, tendríamos que rechazar H_0 si obtuviéramos este resultado muestral. En vista de que la probabilidad de dos colas para 10 signos positivos también es de 0.0020, con este resultado rechazaríamos igualmente H_0 . En la tabla B podemos ver que si el resultado de la muestra fuera de 1 o de 9 signos positivos, rechazaríamos también H_0 ($p = 0.0216$). Sin embargo, si el resultado de la

tabla 11.1 Determinación de los resultados muestrales que permitirán rechazar la hipótesis nula con $N = 10$, $P_{nula} = 0.50$ y $\alpha = 0.05$ colas

Resultado muestral	Probabilidad	Decisión
0 signos positivos	$p(0 \text{ o } 10 \text{ signos positivos}) = 2(0.0010) = 0.0020$	Rechazar H_0
10 signos positivos	$p(0 \text{ o } 10 \text{ signos positivos}) = 2(0.0010) = 0.0020$	Rechazar H_0
1 signo positivo	$p(0, 1, 9 \text{ o } 10 \text{ signos positivos}) = 2(0.0010 + 0.0098) = 0.0216$	Rechazar H_0
9 signos positivos	$p(0, 1, 9 \text{ o } 10 \text{ signos positivos}) = 2(0.0010 + 0.0098) = 0.0216$	Rechazar H_0
2 signos positivos	$p(0, 1, 2, 8, 9 \text{ o } 10 \text{ signos positivos}) = 2(0.0010 + 0.0098 + 0.0439) = 0.1094$	Conservar H_0
8 signos positivos	$p(0, 1, 2, 8, 9 \text{ o } 10 \text{ signos positivos}) = 2(0.0010 + 0.0098 + 0.0439) = 0.1094$	Conservar H_0

muestra fuera de 2 o de 8 signos positivos, el valor de la probabilidad de dos colas (0.1094) sería mayor que alfa. Por consiguiente, conservaríamos H_0 con 2 o con 8 signos positivos. Si no podemos rechazar H_0 con 2 u 8 signos positivos, sin duda no podemos rechazar H_0 si obtenemos un resultado menos extremo, como de 3, 4, 5, 6 o 7 signos positivos. Así, los únicos resultados que nos permitirán rechazar H_0 son 0, 1, 9 o 10 signos positivos. Advierta que, al tomar esta determinación, en virtud de que estábamos evaluando la hipótesis nula, usamos $P_{nula} = 0.50$ (que supone que no hay efecto alguno) y empezamos por los extremos, avanzando hacia el centro de la distribución hasta llegar al primer resultado para el cual la probabilidad de dos colas fue mayor que alfa. Los resultados que nos permiten rechazar H_0 son los que tienen valores más extremos que este primer resultado con el que conservamos H_0 .

¿Cómo podemos usar estos resultados para determinar la potencia? La potencia es igual a la probabilidad de rechazar H_0 si la variable independiente produce un efecto real. Acabamos de determinar que la única forma en que rechazaremos H_0 es si obtenemos un resultado muestral de 0, 1, 9 o 10 signos positivos. Por consiguiente, la potencia es igual a la probabilidad de obtener 0, 1, 9 o 10 signos positivos en nuestra muestra cuando la variable independiente produce un efecto real. Así pues,

Potencia = probabilidad de rechazar H_0 si la variable independiente (IV) produce un efecto real

$$= p(0, 1, 9 \text{ o } 10 \text{ signos positivos}) \text{ si } IV \text{ produce un efecto real}$$

Sin embargo, la probabilidad de obtener 0, 1, 9 o 10 signos positivos depende de la magnitud que tenga el efecto real de la marihuana sobre el apetito. Por consiguiente, la potencia será diferente con distintas magnitudes del efecto. Para ilustrar este punto calcularemos la potencia con varias magnitudes posibles del efecto real. Utilizando P_{real} como nuestra medida de la magnitud y la dirección del efecto de la marihuana, calcularemos la potencia para $P_{real} = 1.00, 0.90, 0.70, 0.30, 0.10$ y 0.00 . Estos valores han sido elegidos para cubrir todo el rango de los efectos reales posibles.

En primer lugar, supongamos que la marihuana tiene un efecto fuerte y creciente sobre el apetito y que, si la proporcionáramos a toda la población, produciría únicamente signos positivos. En este caso, $P_{\text{real}} = 1.00$. La determinación de la potencia para $P_{\text{real}} = 1.00$ es la siguiente:

$$\begin{aligned} \text{Potencia} &= \text{probabilidad de rechazar } H_0 \text{ si } IV \text{ produce un efecto real} \\ &= p(0, 1, 9 \text{ o } 10 \text{ signos positivos}) \text{ como resultado de la muestra si } P_{\text{real}} = 1.00 \\ &= p(0) + p(1) + p(9) + p(10) \quad \text{si } P_{\text{real}} = 1.00 \\ &= 0.0000 + 0.0000 + 0.0000 + 1.0000 \\ &= 1.0000 \end{aligned}$$

Si $P_{\text{real}} = 1.00$, entonces los únicos puntajes posibles son signos positivos. Por lo tanto, la muestra de 10 puntajes debe tener únicamente signos positivos. Por consiguiente, $p(0 \text{ signos positivos}) = p(1 \text{ signo positivo}) = p(9 \text{ signos positivos}) = 0.0000$ y $p(10 \text{ signos positivos}) = 1.0000$. De esta manera, según la regla de la adición, potencia = 1.0000. La probabilidad de rechazar la hipótesis nula cuando sea falsa, de tal modo que $P_{\text{real}} = 1.00$ es igual a 1.0000. Estamos seguros de que si el efecto de la marihuana es tan grande como se ha descrito, el experimento con 10 sujetos detectará su efecto. Entonces H_0 será rechazada con seguridad.

Sin embargo, supongamos que el efecto de la marihuana sobre el apetito no fuera tan grande como se ha descrito, es decir, que si fuera administrada a la población, todavía habría más signos positivos que negativos, pero en esta ocasión serían 9 signos positivos en promedio por cada 1 negativo. En este caso, $P_{\text{real}} = 0.90$. La potencia para esta magnitud un poco más baja del efecto real se localiza en la tabla B, utilizando $P = 0.90$ ($Q = 0.10$). Así pues,

$$\begin{aligned} \text{Potencia} &= \text{probabilidad de rechazar } H_0 \text{ si } IV \text{ produce un efecto real} \\ &= p(0, 1, 9 \text{ o } 10 \text{ signos positivos}) \text{ como resultado de la muestra si } P_{\text{real}} = 0.90 \\ &= p(0) + p(1) + p(9) + p(10) \quad \text{si } P_{\text{real}} = 0.90 \\ &= 0.0000 + 0.0000 + 0.3874 + 0.3487 \\ &= 0.7361 \end{aligned}$$

Datos de la tabla B

N	Número de eventos Q	Q 0.10
10	0	0.3487
	1	0.3874
	9	0.0000
	10	0.0000

La potencia de este experimento para detectar un efecto representado por $P_{\text{real}} = 0.90$ es de 0.7361. Entonces, la potencia del experimento ha disminuido. Observe que al determinar la potencia para $P_{\text{real}} = 0.90$, los resultados muestrales para rechazar H_0 no han cambiado. Igual que antes, son de 0, 1, 9 o 10 signos positivos. En virtud de que éstos son los resultados que nos permitirán rechazar H_0 , sólo son dependientes de N y α . Recuerde que calculamos estos resultados para la N y α dada y un nivel suponiendo que el azar

es el único factor que interviene ($P_{\text{nula}} = 0.50$) y determinando los resultados muestrales para los cuales la probabilidad obtenida es igual o menor que α utilizando P_{nula} .

¿Qué ocurre con la potencia del experimento si la marihuana tiene solamente un efecto moderado, de modo que $P_{\text{real}} = 0.70$?

Potencia = probabilidad de rechazar H_0 si IV produce un efecto real

$$\begin{aligned} &= p(0, 1, 9 \text{ o } 10 \text{ signos positivos}) \text{ como resultado de la muestra si } P_{\text{real}} = 0.70 \\ &= p(0) + p(1) + p(9) + p(10) \quad \text{si } P_{\text{real}} = 0.70 \\ &= 0.0000 + 0.0001 + 0.1211 + 0.0282 \\ &= 0.1494 \end{aligned}$$

Datos de la tabla B

N	Número de eventos Q	Q 0.30
10	0	0.0282
	1	0.1211
	9	0.0001
	10	0.0000

La potencia ha disminuido a 0.1494. También se han efectuado cálculos de potencia para las magnitudes del efecto representadas por $P_{\text{real}} = 0.30$, $P_{\text{real}} = 0.10$ y $P_{\text{real}} = 0.00$. Esos resultados se presentan resumidos en la tabla 11.2.

tabla 11.2 Cálculo de la potencia y de beta

N	H_0	α	Resultados de la muestra*	Magnitud del efecto de la marihuana	Potencia	β
5	$P_{\text{nula}} = 0.50$	$0.05_{2 \text{ colas}}$	Ninguno	Para todos los valores P_{real}	0	1.0000
10	$P_{\text{nula}} = 0.50$	$0.05_{2 \text{ colas}}$	0, 1, 9 o 10 signos positivos	$P_{\text{real}} = 1.00$	1.0000	0.0000
				$P_{\text{real}} = 0.90$	0.7361	0.2639
				$P_{\text{real}} = 0.70$	0.1494	0.8506
				$P_{\text{nula}} = 0.50$	†	
				$P_{\text{real}} = 0.30$	0.1494	0.8506
20	$P_{\text{nula}} = 0.50$	$0.05_{2 \text{ colas}}$	0-5 o 15-20 signos positivos	$P_{\text{real}} = 0.10$	0.7361	0.2639
				$P_{\text{real}} = 0.00$	1.0000	0.0000
				$P_{\text{real}} = 0.30$	0.4163	0.5837
				$P_{\text{real}} = 0.30$	0.1070	0.8930
20	$P_{\text{nula}} = 0.50$	$0.01_{2 \text{ colas}}$	0-3 o 17-20 signos positivos	$P_{\text{real}} = 0.30$	0.1070	0.8930

*Resultados muestrales con los cuales habría que rechazar H_0 .

† Véase la nota 11.1.

En este punto es posible hacer varias generalizaciones. Primera, cuando N aumenta, la potencia se incrementa. Segunda, para una N en particular, por ejemplo $N = 10$, la potencia del experimento varía en proporción directa a la magnitud del efecto real. Cuando la magnitud disminuye, la potencia del experimento disminuye también. Cuando la magnitud del efecto se aproxima al valor predicho por la hipótesis nula, la potencia llega a ser sumamente baja. Esta relación ha sido ilustrada en la figura 11.2.

Potencia y beta (β)

A medida que la potencia de un experimento aumenta, la probabilidad de cometer un error de tipo II disminuye. Esto se puede mostrar en la siguiente forma.

Cuando sacamos una conclusión a partir de un experimento, sólo hay dos posibilidades: o bien se rechaza H_0 o se conserva H_0 . Estas posibilidades también son mutuamente excluyentes. Por lo tanto, la suma de sus probabilidades debe ser igual a 1. Suponiendo que H_0 sea falsa,

$$p(\text{rechazar } H_0 \text{ si es falsa}) + p(\text{conservar } H_0 \text{ si es falsa}) = 1$$

sin embargo

$$\text{Potencia} = p(\text{rechazar } H_0 \text{ si es falsa})$$

$$\text{Beta} = p(\text{conservar } H_0 \text{ si es falsa})$$

Así pues,

$$\text{Potencia} + \text{beta} = 1$$

o bien

$$\text{Beta} = 1 - \text{potencia}$$

De esta manera, cuando la potencia aumenta, beta disminuye. Los valores apropiados de beta se presentan en la última columna de la tabla 11.2.

Notará usted que la tabla 11.2 tiene algunos datos adicionales. Cuando $N = 20$, la potencia de este experimento para detectar un efecto de $P_{\text{real}} = 0.30$ es igual a 0.4163.

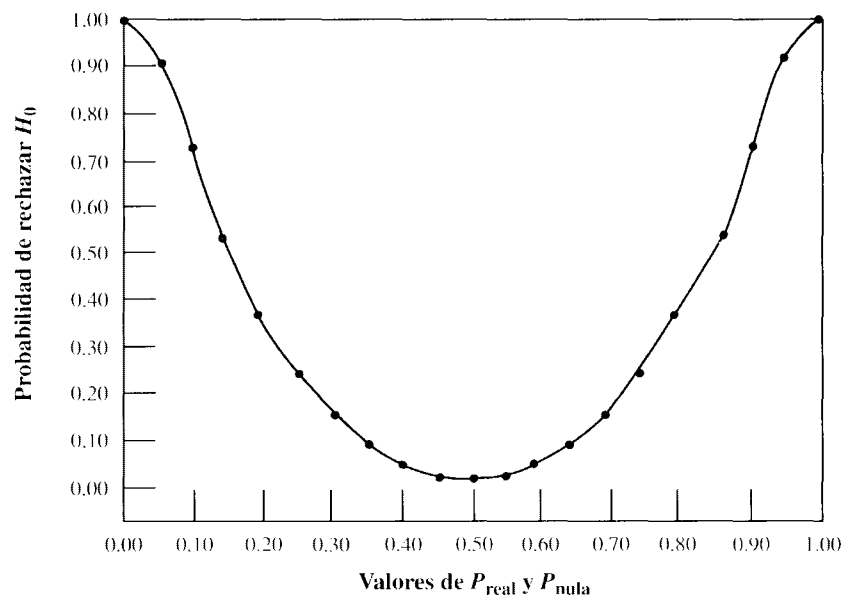


figura 11.2 Potencia de la prueba de signos con $N = 10$ y $\alpha = 0.05_{2 \text{ colas}}$.

Cuando $N = 10$, la potencia es de únicamente 0.1494. Ésta es una demostración más de que a medida que N aumenta, la potencia también aumenta.

Potencia y alfa (α)

El último renglón de la tabla 11.2 demuestra el hecho de que, a medida que alfa se hace más rigurosa, la potencia descende y beta aumenta. Con $N = 20$, $P_{\text{real}} = 0.30$, y $\alpha = 0.01_{2\text{colas}}$:

$$\begin{aligned} \text{Potencia} &= p(0-3 \text{ o } 17-20 \text{ signos positivos}) \\ &= 0.1070 \end{aligned}$$

$$\begin{aligned} \beta &= 1 - \text{potencia} \\ &= 1 - 0.1070 \\ &= 0.8930 \end{aligned}$$

Datos de la tabla B

N	Número de eventos P	P 0.30
20	0	0.0008
	1	0.0068
	2	0.0278
	3	0.0716
	17	0.0000
	18	0.0000
	19	0.0000
	20	0.0000

Al hacer que alfa sea más rigurosa, los posibles resultados muestrales para rechazar H_0 disminuyen. De esta manera, para $\alpha = 0.01_{2\text{colas}}$, sólo 0–3 o 17–20 signos positivos permitirán que H_0 sea rechazada, mientras que para $\alpha = 0.05_{2\text{colas}}$, 0–5 o 15–20 signos positivos darán como resultado el rechazo de H_0 . Esto reduce naturalmente la probabilidad de que la hipótesis nula sea rechazada. La disminución de la potencia se traduce en un incremento de beta.

Resumamos un poco estos conceptos.

1. La potencia de un experimento es la probabilidad de que el resultado del mismo consista en el rechazo de la hipótesis nula si la variable independiente tiene un efecto real.
2. Potencia = 1 – beta. Por lo tanto, cuanto más alta sea la potencia, tanto más baja será beta.
3. La potencia varía en proporción directa a N . Cuando N aumenta, la potencia aumenta también.
4. La potencia varía en proporción directa a la magnitud del efecto real de la variable independiente.
5. La potencia varía en proporción directa al nivel alfa. La potencia disminuye cuando los niveles alfa son más rigurosos.

El lector debe tener presente que, antes de realizar el experimento, el experimentador nunca sabe cuál es en realidad la fuerza del efecto de la variable independiente. Si no fuera así, ¿qué objeto tendría realizar el experimento? En la práctica, estimamos su fuerza a partir de un trabajo piloto o de otra investigación y después diseñamos un experimento que tenga una alta potencia para detectar esa fuerza. Por ejemplo, si se espera un efecto moderado ($P_{\text{real}} = 0.70$) mediante la selección de la N apropiada podemos obtener un grado de sensibilidad aceptable (por ejemplo, una potencia de 0.8000 o mayor). ¿Qué tan alta debe ser la potencia? ¿Qué fuerza del efecto se debe esperar? Estas preguntas deben ser respondidas por el investigador a partir de la experiencia y los recursos disponibles. Es conveniente señalar que al diseñar el experimento para que tenga una potencia de 0.8000 para $P_{\text{real}} = 0.70$, la potencia del experimento será todavía más

alta si el efecto de la variable independiente resulta ser más fuerte de lo esperado. *Por consiguiente, la estrategia consiste en diseñar el experimento para la máxima potencia que permita los recursos para la magnitud mínima del efecto real esperado.*



ALFA-BETA Y REALIDAD

Al realizar un experimento, sólo tenemos dos posibilidades: H_0 es realmente verdadera o es falsa. Al minimizar alfa y beta, maximizamos la posibilidad de que nuestras conclusiones sean correctas. Por ejemplo, si H_0 en realidad es verdadera, entonces la probabilidad de sacar una conclusión correcta del experimento es:

$$p(\text{conclusión correcta}) = p(\text{conservar } H_0) = 1 - \alpha$$

Si alfa se fija en un nivel muy restrictivo, como 0.05, entonces $p(\text{conclusión correcta})$ es:

$$p(\text{conclusión correcta}) = 1 - \alpha = 1 - 0.05 = 0.95$$

Por otra parte, si H_0 es verdaderamente falsa, la probabilidad de sacar una conclusión correcta es:

$$p(\text{conclusión correcta}) = p(\text{rechazar } H_0) = \text{potencia} = 1 - \beta$$

Si beta es baja (como por ejemplo, 0.10 para el efecto real mínimo de interés), entonces:

$$p(\text{conclusión correcta}) = 1 - \beta = 1 - 0.10 = 0.90$$

Así, sin importar el estado real en cuestión (H_0 es verdadera o H_0 es falsa), existe una alta probabilidad de concluir en forma correcta cuando α está en un nivel muy restrictivo y β es baja. Una forma de lograr un nivel bajo de beta cuando alfa se establece en un nivel restrictivo, consiste en tener una N grande. Otra manera consiste en utilizar la prueba de inferencia estadística más poderosa para los datos. Una tercera forma se basa en controlar las condiciones externas del experimento de manera que la variabilidad de los datos se reduzca. Analizaremos más adelante estos dos últimos métodos, cuando revisemos la prueba t de Student en los capítulos 13 y 14.



INTERPRETACIÓN DE RESULTADOS NO SIGNIFICATIVOS

Aunque la consideración de la potencia ayuda a diseñar un experimento, se utiliza mucho más a menudo para interpretar los resultados no significativos de una prueba ya realizada. El hecho de no rechazar H_0 puede deberse a que: 1) H_0 sea realmente verdadera o 2) H_0 sea falsa, pero el experimento poseía una potencia muy baja. Debido a esta segunda posibilidad, *nunca podremos aceptar H_0 como correcta si un experimento no produce resultados significativos. Tan sólo decimos que la prueba no nos ha permitido rechazar la hipótesis nula. Es posible que H_0 en realidad sea falsa, pero que el experimento no sea suficientemente sensible; es decir, que no dio suficientes posibilidades de rechazar H_0 . Un caso que está relacionado con esto es el ejemplo ya presentado con $N = 5$. En ese experimento, sin importar los resultados obtenidos, éstos no llegarían a ser significativos. No podríamos rechazar H_0 , independientemente de la fuerza que tuviera el efecto real. Sería un grave error aceptar H_0 como resultado de ese experimento. El experimento no dio ninguna oportunidad de rechazar H_0 . Desde este punto de vista, podemos ver que todo experimento existe para dar a H_0 la oportunidad de ser rechazada. Mientras mayor sea la potencia, la prueba nos dará mayores posibilidades de rechazar H_0 si ésta es falsa.*

Tal vez una analogía nos ayude a entender este punto. Podemos comparar la potencia de un experimento con el uso de un microscopio. Los neurocientíficos han estado interesados desde hace mucho tiempo en saber lo que sucede dentro del cerebro humano para que podamos memorizar un evento. Una hipótesis sostiene que un grupo de neu-

ronas se activan de manera conjunta como resultado de la presencia de un estímulo. Al presentarse las activaciones sucesivas (ensayos), existe un crecimiento de las sinapsis entre las células, de modo que, después de un tiempo, éstas llegan a excitarse juntas siempre que se presenta dicho estímulo. Este “ensamble celular” se convierte en el engrama fisiológico del estímulo (es decir, es la huella de la memoria).

Para probar esta hipótesis se realiza un experimento de reconocimiento visual. Después que algunos animales practiquen la tarea, las células apropiadas del cerebro de cada uno son preparadas en portaobjetos para ver si hay algún crecimiento a través de las sinapsis. H_0 predice que no habrá crecimiento; H_1 predice que habrá crecimiento. Primero se examinan los portaobjetos a simple vista; no se observa crecimiento alguno. ¿Podemos entonces aceptar H_0 ? No, porque el ojo humano no es tan poderoso como para notar el crecimiento aunque éste estuviese ahí. Lo mismo es válido para un experimento de baja potencia. Si los resultados no son significativos, no podemos aceptar H_0 ; pues incluso si H_0 es falsa, una potencia baja no permite rechazar la hipótesis nula. A continuación, se utiliza un microscopio ordinario y aún así no se distingue un crecimiento entre las sinapsis. Aunque éste es un experimento más poderoso, ¿podemos concluir que H_0 es verdadera? No, ya que un microscopio ordinario no tiene suficiente potencia como para ver las sinapsis con claridad. Por último, se emplea un microscopio electrónico, con el cual se tiene un experimento muy potente en el que todo se puede ver con claridad, salvo las estructuras más diminutas de las sinapsis. Si H_0 es falsa (es decir, si no hay crecimiento a través de las sinapsis), este poderoso experimento tiene una alta probabilidad de detectarlo. De este modo, mientras mayor sea la potencia del experimento, en mayor grado permitirá dicho experimento que rechacemos H_0 si ésta es falsa.

A la luz del análisis anterior, cuando un experimento no consigue producir resultados significativos, debemos tener cuidado en nuestra interpretación. Ciertamente, no podemos afirmar que la hipótesis nula sea correcta. Sin embargo, si la potencia de la prueba es alta, afirmamos que el experimento no nos ha permitido rechazar H_0 . Por ejemplo, si la potencia es 1.0000 para un efecto representado por $P_{\text{real}} = 1.00$ y no podemos rechazar H_0 , por lo menos concluimos que la variable independiente no es tan fuerte. Si la potencia es, digamos, 0.8000 para un efecto moderado ($P_{\text{real}} = 0.70$), podemos tener un grado razonable de confianza en que la variable independiente no es tan eficaz. Por otro lado, si la potencia es baja, los resultados no significativos nos dicen poco sobre el estado de la realidad. Así, un análisis de la potencia nos indica en qué grado debemos confiar en los experimentos que no permiten rechazar la hipótesis nula. Cuando no podemos desechar la hipótesis nula, mientras más alta sea la potencia para detectar un efecto real dado, tanta más confianza tendremos en que la variable independiente no es tan fuerte. Sin embargo, note que conforme el efecto real de la variable independiente se debilita, la potencia del experimento para detectarla también se debilita (véase figura 11.2). *Por lo tanto, es imposible demostrar que la hipótesis nula sea verdadera, ya que la potencia para detectar los efectos demasiado débiles, pero reales, de la variable independiente siempre es muy baja.*



CÁLCULO DE POTENCIA

El cálculo de la potencia requiere un proceso de dos pasos para cada nivel de P_{real} :

PASO 1: Supongamos que la hipótesis nula es verdadera. Mediante el uso de $P_{\text{crítica}} = 0.50$, determine los resultados muestrales posibles de un experimento que nos permitiría rechazar H_0 .

PASO 2: Para el nivel de P_{real} que se esté considerando (por ejemplo, $P_{\text{real}} = 0.30$), calcule la probabilidad de obtener cualquiera de los resultados muestrales anteriores. Esta probabilidad es la potencia del experimento para detectar dicho nivel de efecto real.

Usted tiene interés en averiguar con cuál de estos métodos se recuerda mejor una serie de palabras: 1) cuando éstas son memorizadas en forma directa o 2) cuando los sujetos inventan una historia en la que intervengan todas esas palabras. Al usar el segundo método será necesario recordar la historia a partir de la cual el sujeto logrará recordar las palabras. El plan de usted consiste en someter a 14 sujetos a un experimento de medidas repetidas y analizar posteriormente los datos con la prueba de los signos. Cada sujeto aplicará ambos métodos con conjuntos equivalentes de palabras. El número de palabras recordadas en cada situación será la variable dependiente; $\alpha = 0.05_{2\text{colas}}$.

- a. ¿Cuál es la potencia del experimento para detectar un efecto moderado* ($P_{\text{real}} = 0.80$ o 0.20)?
- b. ¿Cuál es la probabilidad de cometer un error de tipo II?

A continuación se presenta la solución. A partir de la solución, vemos que la potencia para detectar una diferencia moderada en el efecto, es decir, la capacidad de recordar palabras según que éstas sean memorizadas o que se invente una historia que las contenga, es de 0.4480. Esto significa que tenemos una probabilidad de 45% aproximadamente de rechazar H_0 si el efecto es moderado ($P_{\text{real}} = 0.80$ o 0.20) y 55% de probabilidad de incurrir en un error de tipo II. Si el efecto es más débil que moderado, entonces la probabilidad de cometer un error de tipo 2 será aún mayor. Por supuesto, si la N del experimento se incrementa, aumentará también la probabilidad de rechazar H_0 y disminuirá la probabilidad de cometer un error de tipo II.

SOLUCIÓN

- a. Cálculo de la potencia: El cálculo de la potencia se realiza en un proceso de dos pasos:

PASO 1: Suponga que la hipótesis nula es verdadera ($P_{\text{nula}} = 0.50$) y determine los posibles resultados muestrales del experimento que permitirían rechazar H_0 . $\alpha = 0.05_{2\text{colas}}$. Con $N = 14$ y $P = 0.50$, de la tabla B,

$p(0 \text{ signos positivos}) = 0.0001$	$p(0 \text{ signos positivos}) = 0.0001$
$p(1 \text{ signo positivo}) = 0.0009$	$p(1 \text{ signo positivo}) = 0.0009$
$p(2 \text{ signos positivos}) = 0.0056$	$p(2 \text{ signos positivos}) = 0.0056$
$p(12 \text{ signos positivos}) = 0.0056$	$p(3 \text{ signos positivos}) = 0.0222$
$p(13 \text{ signos positivos}) = 0.0009$	$p(11 \text{ signos positivos}) = 0.0222$
$p(14 \text{ signos positivos}) = 0.0001$	$p(12 \text{ signos positivos}) = 0.0056$
$p(0, 1, 2, 12, 13 \text{ o } 14) = 0.0132$	$p(13 \text{ signos positivos}) = 0.0009$
	$p(14 \text{ signos positivos}) = 0.0001$
	$p(0, 1, 2, 3, 11, 12, 13 \text{ o } 14) = 0.0576$

*Hemos dividido el rango de magnitudes en los tres intervalos siguientes: un efecto fuerte del orden de $P_{\text{real}} = 0.00-0.15$ o $0.85-1.00$, un efecto moderado del orden de $P_{\text{real}} = 0.15-0.35$ o $0.65-0.85$, y un efecto débil del orden de $P_{\text{real}} = 0.35-0.49$ o $0.51-0.65$. (continúa)

Comenzando por los extremos y avanzando hacia la parte media de la distribución, descubrimos que podremos rechazar H_0 si obtenemos 2 o 12 signos positivos ($p = 0.0132$), pero que no podremos rechazar H_0 si obtenemos 3 u 11 signos positivos ($p = 0.0576$) en la muestra. En consecuencia, los resultados que nos permitirán rechazar H_0 son 0, 1, 2, 12, 13 o 14 signos positivos.

PASO 2: Para $P_{\text{real}} = 0.20$, determine la probabilidad de obtener cualquiera de los resultados muestrales mencionados anteriormente. Esta probabilidad es la potencia del experimento para detectar ese efecto real hipotético. Con $N = 14$ y $P_{\text{real}} = 0.20$, tomados de la tabla B,

$$\begin{aligned} \text{Potencia} &= \text{probabilidad de rechazar } H_0 \text{ si } IV \text{ tiene un efecto real} \\ &= p(0, 1, 2, 12, 13 \text{ o } 14 \text{ signos positivos) como resultados muestrales si} \\ &\quad P_{\text{real}} = 0.20 \\ &= 0.0440 + 0.1539 + 0.2501 + 0.0000 + 0.0000 + 0.0000 \\ &= 0.4480 \end{aligned}$$

Datos de la tabla B

N	Número de eventos P	P 0.20
14	0	0.0440
	1	0.1539
	2	0.2501
	12	0.0000
	13	0.0000
	14	0.0000

Observe que obtendríamos la misma respuesta para $P_{\text{real}} = 0.80$.

b. Cálculo de beta:

$$\begin{aligned} \beta &= 1 - \text{potencia} \\ &= 1 - 0.4480 \\ &= 0.5520 \end{aligned}$$



Supongamos que usted está planeando un experimento para evaluar un medicamento. La hipótesis alternativa es direccional, en la dirección que produce principalmente signos positivos. Aplicará la prueba de los signos para analizar los datos; $\alpha = 0.05_{\text{cola}}$. Usted desea tener la posibilidad de detectar un efecto débil (digamos, $P_{\text{real}} = 0.60$) en la misma dirección que la hipótesis alternativa. Participarán 16 sujetos en el experimento.

- ¿Cuál es la potencia del experimento para detectar este efecto débil?
- ¿Cuál es la probabilidad de incurrir en un error de tipo II?

SOLUCIÓN

a. Cálculo de la potencia: El cálculo de la potencia se realiza en dos pasos:

PASO 1: Suponga que la hipótesis nula es verdadera ($P_{nula} = 0.50$) y determine los resultados muestrales posibles del experimento que nos permitiría rechazar H_0 . $\alpha = 0.05$, cola. Con $N = 16$ y $P = 0.50$, tomados de la tabla B,

$p(12 \text{ signos positivos}) = 0.0278$	$p(11 \text{ signos positivos}) = 0.0667$
$p(13 \text{ signos positivos}) = 0.0085$	$p(12 \text{ signos positivos}) = 0.0278$
$p(14 \text{ signos positivos}) = 0.0018$	$p(13 \text{ signos positivos}) = 0.0085$
$p(15 \text{ signos positivos}) = 0.0002$	$p(14 \text{ signos positivos}) = 0.0018$
$\underline{p(16 \text{ signos positivos}) = 0.0000}$	$\underline{p(15 \text{ signos positivos}) = 0.0002}$
$p(12, 13, 14, 15 \text{ o } 16) = 0.0383$	$\underline{p(16 \text{ signos positivos}) = 0.0000}$
	$p(11, 12, 13, 14, 15 \text{ o } 16) = 0.1050$

En vista de que la hipótesis alternativa está en la dirección en la cual la mayoría de los signos son positivos, los resultados que permiten rechazar H_0 se encuentran bajo la cola que tiene el número más alto de signos positivos. A partir de 16 signos positivos y avanzando hacia la parte media de la distribución, vemos que podremos rechazar H_0 si obtenemos 12 signos positivos ($p = 0.0383$), pero que no podremos rechazar H_0 si obtenemos 11 signos positivos ($p = 0.1050$) en la muestra. Por lo tanto, los resultados que nos permitirán rechazar H_0 son 12, 13, 14, 15 o 16 signos positivos.

PASO 2: Para $P_{real} = 0.60$, determine la probabilidad de obtener cualquiera de los resultados muestrales mencionados anteriormente. Esta probabilidad es la potencia del experimento para detectar este efecto real hipotético. Con $N = 16$ y $P_{real} = 0.60$ ($Q = 0.40$), tomados de la tabla B.

$$\begin{aligned}
 \text{Potencia} &= \text{probabilidad de rechazar } H_0 \text{ si } IV \text{ tiene un efecto real} \\
 &= p(12, 13, 14, 15 \text{ o } 16 \text{ signos positivos}) \text{ como resultados muestrales si} \\
 &\quad P_{real} = 0.60 \\
 &= 0.1014 + 0.0468 + 0.0150 + 0.0030 + 0.0003 \\
 &= 0.1665
 \end{aligned}$$

Datos de la tabla B

N	Número de eventos Q	$Q = 0.40$
16	0	0.0003
	1	0.0030
	2	0.0150
	3	0.0468
	4	0.1014

b. Cálculo de beta:

$$\begin{aligned}
 \beta &= 1 - \text{potencia} \\
 &= 1 - 0.1665 \\
 &= 0.8335
 \end{aligned}$$

(continúa)

Este experimento es muy insensible cuando el efecto del medicamento es débil. La probabilidad de cometer un error de tipo II es demasiado alta. La N deberá hacerse mayor a fin de aumentar la potencia del experimento para detectar el efecto débil del medicamento.



En el problema de práctica 10.2 (página 228), usted realizó el experimento para poner a prueba la hipótesis alternativa direccional según la cual el hecho de mirar un programa de televisión en particular provocaba un aumento de la violencia en los adolescentes. En el experimento participaron 15 sujetos y $\alpha = 0.01_{\text{cola}}$. Los datos fueron analizados con la prueba de los signos y conservamos H_0 .

- ¿Cuál fue, en ese experimento, la potencia para detectar un efecto moderado de $P_{\text{real}} = 0.70$ en la dirección de la hipótesis alternativa?
- ¿Cuál fue la probabilidad de cometer un error de tipo II?

SOLUCIÓN

- Cálculo de la potencia: El cálculo de la potencia se realiza en dos pasos:

PASO 1: Suponga que la hipótesis nula es verdadera ($P_{\text{nula}} = 0.50$) y determine los posibles resultados muestrales del experimento que nos permitirán rechazar H_0 . $\alpha = 0.01_{\text{cola}}$. Con $N = 15$ y $P = 0.50$, tomados de la tabla B,

$p(13 \text{ signos positivos})$	$= 0.0032$	$p(12 \text{ signos positivos})$	$= 0.0139$
$p(14 \text{ signos positivos})$	$= 0.0005$	$p(13 \text{ signos positivos})$	$= 0.0032$
$p(15 \text{ signos positivos})$	$= 0.0000$	$p(14 \text{ signos positivos})$	$= 0.0005$
$p(13, 14 \text{ o } 15)$	$= 0.0037$	$p(15 \text{ signos positivos})$	$= 0.0000$
		$p(12, 13, 14 \text{ o } 15)$	$= 0.0176$

En virtud de que la hipótesis alternativa está en la dirección en la cual la mayoría de los signos son positivos, los resultados que nos permiten rechazar H_0 se encuentran bajo la cola que tiene el mayor número de signos positivos. A partir de 15 signos positivos, nos desplazamos hacia la parte media de la distribución y encontramos que podremos rechazar H_0 si obtenemos 13 signos positivos ($p = 0.0037$), pero conservaremos H_0 si obtenemos 12 signos positivos ($p = 0.0176$) en la muestra. Por lo tanto, los resultados que nos permitirán rechazar H_0 son 13, 14 o 15 signos positivos.

PASO 2: Para $P_{\text{real}} = 0.70$, determine la probabilidad de obtener cualquiera de los resultados muestrales mencionados anteriormente. Esta probabilidad es la potencia del experimento para detectar este efecto real hipotético. Con $N = 15$ y $P_{\text{real}} = 0.70$, tomados de la tabla B,

$$\begin{aligned}
 \text{Potencia} &= \text{probabilidad de rechazar } H_0 \text{ si } IV \text{ tiene un efecto real} \\
 &= p(13, 14 \text{ o } 15 \text{ signos positivos}) \text{ como resultados muestrales si} \\
 &\quad P_{\text{real}} = 0.70 \\
 &= 0.0916 + 0.0305 + 0.0047 \\
 &= 0.1268
 \end{aligned}$$

Datos de la tabla B

N	Número de eventos Q	Q 0.30
15	0	0.0047
	1	0.0305
	2	0.0916

b. Cálculo de beta:

$$\begin{aligned}
 \beta &= 1 - \text{potencia} \\
 &= 1 - 0.1268 \\
 &= 0.8732
 \end{aligned}$$

Observe que como la potencia para detectar un efecto moderado es muy baja, aun cuando conservemos H_0 en el experimento, no podemos concluir que el programa no influye en la violencia. Habrá que realizar de nuevo el experimento con una potencia mayor para permitir una evaluación más satisfactoria del efecto de dicho programa sobre la violencia.



Astrología y ciencia

Un artículo publicado en un número reciente del *Post Gazette* de Pittsburgh tenía el encabezado: "Cuando los estudios clínicos mienten". A continuación reproducimos algunas partes de ese artículo:

Una gran conmoción surgió entre la comunidad médica hace dos semanas, cuando unos investigadores anunciaron que tres fármacos que se prescriben con frecuencia para ayu-

dar a un paciente después de un ataque al corazón han demostrado su inutilidad en nuevos estudios. . .

"Las personas quedan deslumbradas constantemente por la fuerza de los números, pero no saben qué mentiras pueden esconderse detrás de ellos", dijo Alvan R. Feinstein, un profesor de medicina y epidemiología de la Escuela de Medicina en Yale. "Incluso los científicos y los médicos han sufrido este lavado de cerebro, pensando que la frase mágica 'significativo desde el punto de vista de la estadística' es la respuesta para todas las cosas."

Los estudios recientes sobre me-

dicamentos para el corazón contradicen ese mito. Los ensayos clínicos realizados con miles de pacientes durante varios años ya habían demostrado, en el pasado, que los medicamentos que contienen nitrato, como la nitroglicerina, la enzima inhibidora captopril y el magnesio ayudaban a salvar vidas, cuando se administraban después de un ataque al corazón.

Al comparar la duración de la vida de los individuos que tomaron esos medicamentos con la de los que no lo hicieron, los investigadores en-

(continúa)



Astrología y ciencia (continuación)

contraron que la diferencia era estadísticamente significativa, y entonces dichos fármacos llegaron a ser parte de la práctica médica común. En Estados Unidos, a más del 80 por ciento de los pacientes que han sufrido ataques al corazón se les administran medicamentos que contienen nitrato.

Sin embargo en un nuevo estudio, con más de 50 000 pacientes, los investigadores no encontraron beneficio alguno en el uso de los nitratos o del magnesio. La utilidad del captopril fue marginal. Richard Peto, el epidemiólogo de Oxford que supervisó el último estudio, dijo que los resultados positivos de las pruebas anteriores deben haberse debido "al azar". . . Peto afirma que los errores numéricos pueden ser cuestión de vida o muerte.

Él y sus colegas abordaron este punto en 1988 cuando enviaron un artículo a la revista británica de medicina The Lancet. Su informe demostraba que las víctimas de ataques al corazón tenían una mayor oportunidad de sobrevivir si se les daba una aspirina durante las primeras horas después de sus ataques. Según describe los hechos Peto, los editores de la citada revista querían que los investigadores separasen los datos en varios subconjuntos para ver si ciertos tipos de pacientes, que difieren unos de otros por la edad o alguna otra característica, tenían más o menos probabilidades de beneficiarse con la aspirina.

Peto objetó esto aduciendo que la validez del estudio podría verse comprometida por el hecho de dividirlo en muchas piezas. Él afirma que si se comparara el número sufi-

ciente de subgrupos se podría obtener algún tipo de correlación sólo por efecto del azar. Cuando los editores insistieron Peto se dio por vencido, pero, entre otros criterios, decidió separar también a sus pacientes según su respectivo signo del zodiaco y exigió que sus descubrimientos fueran incluidos en el artículo publicado. Hoy en día, como una señal de advertencia a los no iniciados en estadística, se puede leer lo siguiente: La aspirina es inútil para las víctimas de ataques al corazón de Géminis y Libra, pero es un salvavidas para las personas nacidas bajo cualquier otro signo zodiacal. . .

Estudios como éste ejemplifican dos de los excesos estadísticos más comunes que cometen los científicos: hacer demasiadas comparaciones y prestar poca atención al hecho de que algo tenga sentido, dijo James L. Mills, jefe de la sección de epidemiología pediátrica del Instituto Nacional de Salud Infantil y Desarrollo Humano.

"Las personas esperan encontrar en los resultados las cosas más

emocionantes y positivas", afirmó. "Pero también hay que tomar en cuenta la plausibilidad biológica. Muchos hallazgos que finalmente no soportaron la prueba del tiempo en realidad no tenían sentido desde el principio." . . .

En los últimos años, muchos científicos han realizado ensayos clínicos cada vez más y más grandes con la esperanza de que así podrán minimizar la posibilidad de dejarse engañar por una casualidad.

¿Qué opina usted? Si fuese un médico, ¿continuaría recetando nitratos a los pacientes con ataques al corazón? ¿Será cierto que los primeros ensayos clínicos fueron un verdadero ejemplo del error de tipo I, tal como lo ha dicho el doctor Peto? ¿El uso de ensayos clínicos cada vez mayores minimizará la posibilidad de dejarse engañar por una casualidad? Por último, ¿es la aspirina inútil para las víctimas de ataques al corazón de Géminis y Libra, pero es un salvavidas para los pacientes nacidos bajo cualquier otro signo?



RESUMEN

En este capítulo nos hemos referido al tema de la potencia. La potencia se define como la probabilidad de rechazar la hipótesis nula cuando el efecto producido por la variable independiente es real. Como quiera que la potencia varía según la magnitud del efecto real, es preciso calcularla para el efecto real más pequeño que merezca nuestro interés. La potencia será todavía más alta cuando los efectos son mayores. El cálculo de la potencia incluye dos pasos. En el primer paso, se supone que la hipótesis nula es verdadera ($P_{nula} = 0.50$), y se determinan todos los resultados muestrales posibles del experimento que nos permitirían rechazar la hipótesis nula. A continuación se calcula la probabilidad de obtener cualquiera de esos resultados muestrales para el efecto real que estamos considerando (por ejemplo, el efecto representado por $P_{real} = 0.30$). Esta probabilidad es la potencia del experimento para detectar dicho efecto ($P_{real} = 0.30$).

Si todos los demás factores se mantienen constantes, la potencia aumenta cuando N aumenta y también lo hace cuando aumenta la magnitud del efecto real de la variable independiente. La potencia disminuye cuando se propone un nivel alfa más riguroso. La potencia es igual a $1 - \beta$, por lo cual cuando se maximiza la

potencia se minimiza la probabilidad de incurrir en un error de tipo II. Así, al minimizar alfa y beta, maximizamos la probabilidad de determinar correctamente el efecto verdadero de la variable independiente, no importa cuál sea el estado real.

Un análisis de potencia resulta útil: 1) durante el diseño inicial de un experimento y 2) cuando se interpretan los resultados de experimentos en los que la hipótesis nula se conserva. Cuando se realiza un experimento y los resultados no son significativos, esto puede ser porque la hipótesis nula es verdadera o porque la potencia del experimento es baja. Por esta razón, cuando los resultados no son significativos, nuestra conclusión no consiste en *aceptar* la hipótesis nula, sino más bien en *no rechazarla*. En realidad, es posible que la hipótesis nula sea falsa, pero el experimento no tiene la potencia suficiente para detectarlo. El propósito de todo experimento es brindar a la hipótesis nula la posibilidad de ser rechazada. Cuanto más potente sea el experimento, tanto mayor será la probabilidad de que la hipótesis nula sea rechazada cuando es falsa. En virtud de que la potencia se vuelve menor a medida que el efecto real de la variable independiente disminuye, es imposible demostrar que H_0 es verdadera.

TÉRMINOS IMPORTANTES

Alfa (página 252)

Beta (página 243)

Decisión correcta (página 253)

Efecto real (página 244)

Error de tipo I (página 243)

Error de tipo II (página 243)

Estado real (página 253)

Hipótesis nula (página 254)

No significativo (página 253)

P_{nula} (página 244)

Potencia (página 243)

P_{real} (página 244)

Resultados significativos (página 254)

PREGUNTAS Y PROBLEMAS

- ¿Qué es la potencia? ¿Cómo se define?
- ¿Cuáles son las dos situaciones en las que el análisis de potencia resulta especialmente útil? Explique.
- En los experimentos para una prueba de hipótesis, ¿por qué es preferible la conclusión "conservamos H_0 " a la conclusión "aceptamos H_0 como verdadera"?
- En los experimentos para una prueba de hipótesis, ¿es correcto concluir que la variable independiente no produce efecto alguno? Explique.
- En el cálculo de la potencia, ¿por qué siempre calculamos los resultados muestrales que nos permitirían rechazar H_0 ?
- Utilice α y β , para explicar la forma en que se puede maximizar la probabilidad de obtener la conclusión correcta de un experimento, independientemente de que H_0 sea verdadera o falsa. Como parte de su explicación, elija valores para α y β y determine la probabilidad de obtener la conclusión correcta cuando H_0 es verdadera y cuando H_0 es falsa.
- Usted está pensando en la forma de probar un nuevo fármaco que, al parecer, facilita el aprendizaje en los niños mentalmente retardados. Puesto que se sabe relativamente poco acerca de ese fármaco, usted planea usar una hipótesis alternativa no direccional. Como sus recursos son limitados, sólo podrá incluir en la prueba a 15 sujetos. Los sujetos serán observados según un diseño de medidas repetidas y los datos se analizarán por medio de la prueba de los signos, con $\alpha = 0.05_{2\text{colas}}$. Si el fármaco produce un efecto moderado sobre el aprendizaje, de tal manera que $P_{real} = 0.70$, ¿cuál es la probabilidad de que usted detecte dicho efecto al realizar su experimento? ¿Cuál es la probabilidad de que incurra en un error de tipo II? *cognitiva*
- En el problema 10 del capítulo 10 (página 240) se hizo la evaluación de un nuevo método de enseñanza. Veinte pares de sujetos fueron examinados según un diseño de medidas repetidas. Los resultados favorecieron al nuevo método, pero no alcanzaron un nivel significativo (H_0 no fue rechazada) al aplicar la prueba de los signos con $\alpha = 0.05_{1\text{cola}}$. Al tra-

tar de interpretar por qué no fueron significativos los resultados, usted razona que existen dos posibles explicaciones: 1) los dos métodos de enseñanza son en realidad igualmente eficaces (H_0 es verdadera), o bien, 2) el nuevo método es mejor, pero el experimento fue insensible. Con el fin de evaluar esta última posibilidad, usted realiza un análisis que le permite determinar la potencia del experimento para detectar una diferencia moderada a favor del nuevo método, tal que $P_{\text{real}} = 0.80$. ¿Cuál es la potencia necesaria del experimento para detectar este efecto? ¿Cuál es el valor de beta? educación

9. Un investigador se dispone a realizar un experimento para determinar si la pérdida de una noche de sueño influye sobre el desempeño. Suponga que se han cumplido los requisitos para establecer una hipótesis alternativa direccional. Catorce sujetos participarán en un diseño de medidas repetidas. Los datos serán analizados mediante la prueba de los signos, con $\alpha = 0.05_{2 \text{ colas}}$. Cada sujeto será observado en dos condiciones: condición 1, cuando el desempeño del sujeto se mide después de gozar de una buena noche de sueño y la condición 2, en la que el desempeño se mide después de una noche de privación del sueño. Cuanto mejor sea el desempeño, tanto más alto será el puntaje. Cuando los datos sean analizados, los puntajes de la condición 2 serán restados de los de la condición 1. Si la pérdida de una noche de sueño produce un efecto fuerte en detrimento del desempeño, de tal manera que $P_{\text{real}} = 0.90$, ¿qué potencia deberá tener el experimento para detectar este efecto? ¿Cuál es la probabilidad de cometer un error de tipo II? cognitiva
10. En el problema 12 del capítulo 10 (página 241), ¿qué potencia debe tener el experimento para detectar un efecto moderado, como el de $P_{\text{real}} = 0.70$? educación
11. Un psiquiatra planea un experimento para determinar si el aislamiento y la privación de todo es-

tímulo influye en la depresión. Dieciocho sujetos participarán en un diseño de medidas repetidas. Los datos serán analizados mediante la prueba de signos, con $\alpha = 0.05_{2 \text{ colas}}$. Cada sujeto será observado en dos condiciones: la condición 1, después de una semana de estar viviendo en un ambiente con una cantidad normal de estímulos externos y la condición 2, después de una semana de estar en un ambiente aislado donde los estímulos externos han sido suprimidos en forma radical. Después de someterlos a cada condición, se les aplicará un examen para medir la depresión. Cuanto más alto sea el puntaje que obtengan los sujetos en el cuestionario, tanto mayor será su grado de depresión. Al analizar los datos, los puntajes de la condición 1 serán restados de los puntajes de la condición 2. Si una semana de aislamiento de los estímulos produce algún efecto sobre la depresión, de manera que $P_{\text{real}} = 0.60$, ¿qué potencia deberá tener el experimento para detectar este efecto débil? ¿Cuál será su beta? Si los resultados del experimento no son significativos, ¿es legítimo que el psiquiatra concluya que la privación de estímulos no produce efecto alguno sobre la depresión? ¿Por qué? cognitiva, clínica, de salud

12. Tome como referencia el problema de práctica 10.2 (página 228). Supongamos que el programa de televisión afecta en realidad el fenómeno de la violencia en los adolescentes, incrementándolo en forma moderada ($P_{\text{real}} = 0.70$). ¿Cuál es la probabilidad de que el experimento realizado en el problema de práctica 10.2 nos permita detectarlo? ¿Cuál es la probabilidad de que se cometa un error de tipo II? Los datos recopilados en este experimento no permitieron rechazar H_0 . ¿Acaso por esta razón resulta justificado concluir que el programa de televisión no tiene efecto alguno sobre la violencia en los adolescentes? Explique. social

NOTAS

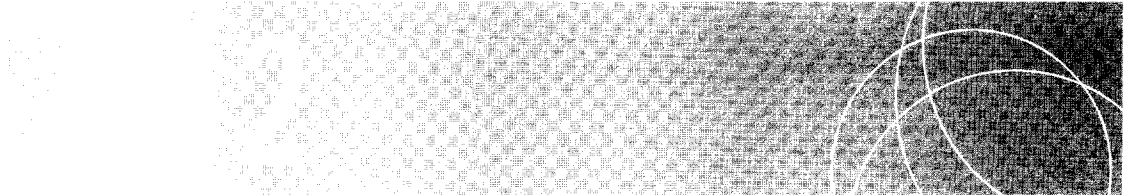
- 11.1 Esta probabilidad no es igual a la potencia porque cuando $P = 0.50$, H_0 es verdadera. La potencia se calcula cuando H_0 es falsa. La probabilidad de rechazar H_0 cuando ésta es verdadera se define como la probabilidad de incurrir en un error de tipo I. En el caso planteado en este ejemplo,

$$\begin{aligned}
 & p(\text{rechazar } H_0 \text{ cuando } P = 0.50) = \\
 & p(\text{error de tipo I}) \\
 & = p(0) + p(1) + p(9) + p(10) \\
 & = 0.0010 + 0.0098 + 0.0098 + 0.0010 \\
 & = 0.0216
 \end{aligned}$$

Observe que la probabilidad de cometer un error de tipo I (0.0216) no es igual al nivel α (0.05) porque el número de signos positivos es una variable discreta y no una variable continua. Para lograr que $p(\text{error de tipo I})$ fuera igual a alfa, necesitaríamos obtener un resultado de entre 8 y 9 signos positivos. Por supuesto esto es imposible porque el número de signos positivos sólo puede ser de 8 o 9 (valores discretos). La probabilidad de incurrir en un error de tipo I es igual a alfa cuando la variable es continua.

Si desea encontrar problemas de práctica adicionales totalmente resueltos sobre los temas que acaba de estudiar, busque Chapter 11 en el material adicional del libro, disponible en inglés en: <http://www.thomsonlearning.com.mx>.

Distribuciones muestrales, distribución muestral de la media, la prueba (z) de la desviación normalizada



Introducción

Distribuciones muestrales

Generación de las distribuciones muestrales

La prueba (z) de la desviación normalizada

Un experimento: Evaluación del programa de lectura en las escuelas públicas locales

Distribución muestral de la media

Revisión del experimento sobre la habilidad de lectura

Solución alternativa utilizando z_{obt} y la región crítica para el rechazo de H_0

Condiciones en las cuales resulta adecuada la prueba z

La potencia y la prueba z

Resumen

Términos importantes

Preguntas y problemas



INTRODUCCIÓN

En los capítulos 10 y 11 expusimos de qué manera se utiliza el método científico para investigar una hipótesis. Presentamos los diseños de medidas replicadas y de grupos independientes, y estudiamos la forma de analizar los datos resultantes. El elemento medular de este análisis consiste en responder la siguiente pregunta: *¿cuál es la probabilidad de conseguir el resultado obtenido, u otros resultados todavía más extremos, si el azar es en realidad la única causa de las diferencias que hemos observado entre los datos experimentales y los datos control?*

Aun cuando no hayamos insistido en este punto, la respuesta a esa pregunta requiere dos pasos: 1) el cálculo del estadístico apropiado y 2) la evaluación del mismo con base en su distribución muestral. En este capítulo vamos a analizar de manera más formal el tema de un estadístico y su distribución muestral. Después, daremos principio al estu-

dio de los experimentos con una sola muestra, utilizando la media de la muestra como estadístico. Para esto vamos a usar la distribución muestral de la media y la prueba z de la desviación normalizada.



DISTRIBUCIONES MUESTRALES

¿Qué es una distribución muestral?

definición ■ *La distribución muestral de un estadístico proporciona: 1) una lista de todos los valores que dicho estadístico puede asumir y 2) la probabilidad de obtener cada valor, suponiendo que éste sea producido únicamente por el azar.*

En el diseño de medidas replicadas utilizamos la prueba del signo para analizar los datos. El estadístico calculado fue el *número de signos positivos* contenidos en la muestra de N diferencias. En una versión del experimento “marihuana y apetito”, obtuvimos nueve signos positivos y uno negativo. Llevamos a cabo la evaluación de este resultado mediante la distribución binomial. La distribución binomial con $P = 0.50$ enumera todos los valores posibles del estadístico, el número de signos positivos, así como la probabilidad de obtener cada valor, bajo la hipótesis de que éste sólo depende del azar. *La distribución binomial con $P = 0.50$ es la distribución muestral del estadístico utilizado en la prueba del signo.* Observe que existe una distribución muestral diferente para cada tamaño de muestra (N).

A partir de este ejemplo podemos generalizar que el análisis de datos implica básicamente dos pasos:

1. El cálculo del estadístico apropiado; por ejemplo, el número de signos positivos y negativos para la prueba del signo.
2. La evaluación del estadístico con base en su distribución muestral.

Si la probabilidad de obtener el valor que se ha obtenido del estadístico, u otro valor más extremo, es menor o igual que el nivel alfa, rechazamos H_0 y aceptamos H_1 . En caso contrario, conservamos H_0 . Si rechazamos H_0 y ésta es verdadera, habremos cometido un error de tipo I. Si conservamos H_0 y ésta es falsa, habremos cometido un error de tipo II. Este proceso es aplicable a todos los experimentos que implican una prueba de hipótesis. *Lo que cambia de un experimento a otro es el estadístico utilizado y la distribución muestral correspondiente.* Una vez que usted comprenda este concepto, podrá apreciar que gran parte de los cursos de estadística inferencial se dedican a la presentación de los estadísticos de uso más frecuente, sus distribuciones muestrales y las condiciones en las cuales resulta apropiado utilizar cada tipo de estadístico.

Generación de las distribuciones muestrales

Hemos definido una distribución muestral como una distribución de probabilidad de todos los valores posibles que puede asumir un estadístico, suponiendo que sólo influye el azar. Una forma de deducir las distribuciones muestrales consiste en utilizar las consideraciones básicas de probabilidad. Adoptamos este enfoque para generar la distribución

binomial. También podemos deducir las distribuciones muestrales partiendo del punto de vista de un muestreo empírico. En este caso, tenemos un conjunto de datos poblacionales, real o teórico, que existe si la variable independiente no produce efecto alguno. La deducción de la distribución muestral del estadístico se hace de la siguiente forma:

1. Determinamos todas las diferentes muestras posibles de tamaño N que pueden formarse a partir de la población de datos.
2. Calculamos el estadístico para cada una de las muestras.
3. Calculamos la probabilidad de obtener cada uno de los valores del estadístico si sólo influye el azar.

Para ilustrar el enfoque de muestreo suponga que realizamos un experimento con un tamaño de muestra $N = 2$, y aplicamos la prueba de los signos para el análisis. Podemos imaginar un conjunto teórico de datos resultantes si el experimento se realiza sobre toda la población y la variable independiente no produce efecto alguno. Este conjunto de datos es la *población de la hipótesis nula*.

definición ■ *La población de la hipótesis nula es el conjunto real o teórico de datos provenientes de una población que resultaría si el experimento se realizara sobre toda la población en cuestión y la variable independiente no tuviese efecto alguno. Se la llama población de la hipótesis nula porque se utiliza para verificar la validez de la hipótesis nula.*

En el caso de la prueba del signo, si la variable independiente no hubiera producido efecto alguno, entonces la población de la hipótesis nula tendría el mismo número de signos positivos que de signos negativos ($P = Q = 0.50$).

A fin de facilitar los cálculos necesarios para generar la distribución muestral, suponga que hay solamente seis datos en la población: tres signos positivos y tres negativos. Para deducir la distribución muestral del “número de signos positivos” con $N = 2$, debemos determinar, en primer lugar, todas las distintas muestras de tamaño N que podemos formar con esta población. El muestreo es de un elemento a la vez, con reemplazo. La figura 12.1 ilustra la población y, en forma esquemática, todas las posibles muestras de tamaño 2 que es posible extraer de ella. Resulta que hay 36 muestras diferentes de tamaño 2. Éstas se presentan en la segunda columna de la tabla de la figura 12.1. A continuación, debemos calcular el valor del dato estadístico para cada muestra; esta información aparece en las columnas 3 y 4 de la tabla de la figura 12.1. Observe que de las 36 muestras distintas posibles, 9 tienen dos signos positivos, 18 tienen un signo positivo y 9 no presentan signos positivos. El último paso consiste en calcular la probabilidad de obtener cada uno de los valores del estadístico. Si sólo interviene el azar, cada muestra es igualmente probable que las demás. Así,

$$p(2 \text{ signos positivos}) = \frac{9}{36} = 0.2500$$

$$p(1 \text{ signo positivo}) = \frac{18}{36} = 0.5000$$

$$p(0 \text{ signos positivos}) = \frac{9}{36} = 0.2500$$

Hemos obtenido ahora la distribución muestral para $N = 2$ del estadístico “número de signos positivos”. La gráfica de la distribución se presenta en la figura 12.2. En este ejemplo utilizamos una población en la que sólo había seis datos. Habríamos obtenido la misma distribución muestral (aunque hubiese muchas más muestras “diferentes”) si hubiéramos empleado una población mayor, siempre que el número de signos positivos fuera

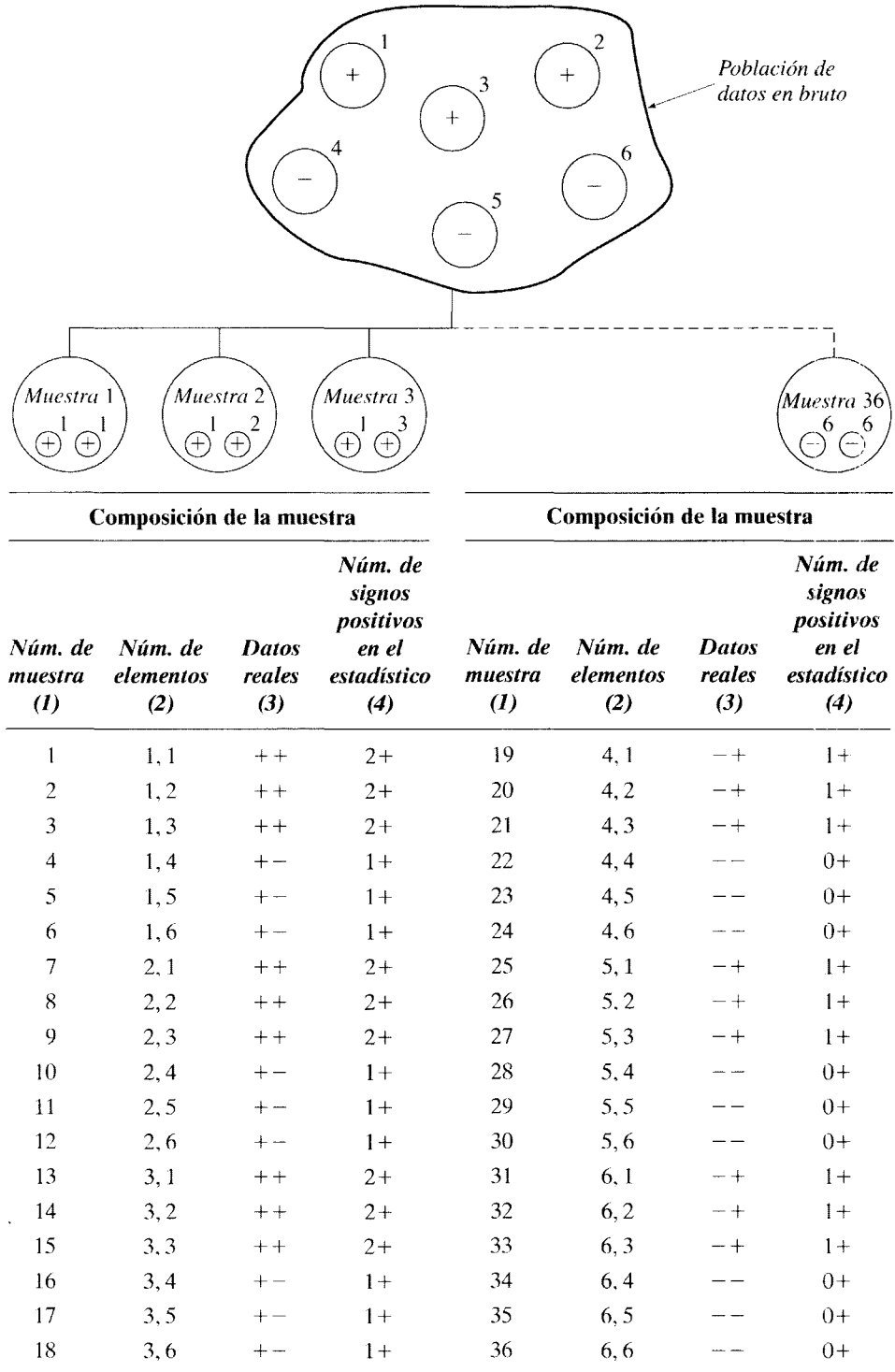


figura 12.1 Todas las posibles muestras de tamaño 2 que se pueden extraer de una población de tres signos positivos y tres negativos. El muestreo es de un elemento a la vez, con reemplazo.

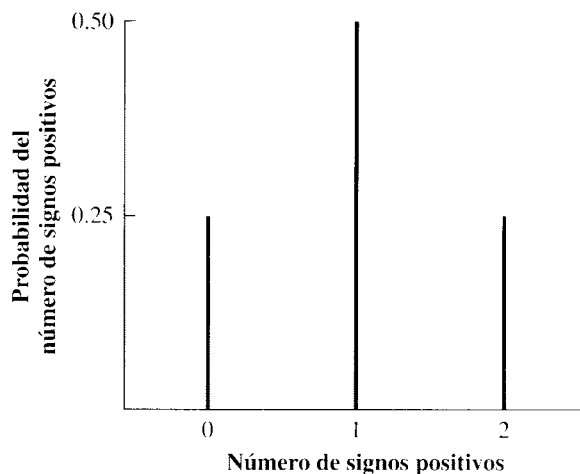


figura 12.2 Distribución muestral del "número de signos positivos" con $N = 2$ y $P = 0.50$.

igual al número de signos negativos y el tamaño de la muestra fuese igual a 2. Observe que ésta es la misma distribución muestral que obtuvimos mediante las consideraciones básicas de probabilidad, cuando analizamos la distribución binomial con $N = 2$ (véase la figura 12.3 para hacer una comparación). Sin embargo, esta vez generamos la distribución mediante un muestreo realizado en la población de la hipótesis nula. La distribución muestral de un estadístico se define con frecuencia en función de este proceso. Visto de esta forma, obtenemos la siguiente definición:

Enfoque empírico de muestreo			Enfoque <i>a priori</i>		
Primera extracción	Segunda extracción	Número de formas	Primera moneda	Segunda moneda	Número de formas
+	+	9	Cara	Cara	1
+	-	18	Cara	Cruz	2
-	+		Cruz	Cara	
-	-	9	Cruz	Cruz	1
		36			4

$$p(2+) = \frac{9}{36} = 0.2500$$

$$p(1+) = \frac{18}{36} = 0.5000$$

$$p(0+) = \frac{9}{36} = 0.2500$$

$$p(2H) = \frac{1}{4} = 0.2500$$

$$p(1H) = \frac{2}{4} = 0.5000$$

$$p(0H) = \frac{1}{4} = 0.2500$$

figura 12.3 Comparación del enfoque muestral empírico y de un enfoque *a priori* para generar distribuciones muestrales.

definición ■ Una **distribución muestral** proporciona todos los valores que puede asumir un estadístico, junto con la probabilidad de obtener cada valor si el muestreo es aleatorio a partir de la población de la hipótesis nula.



LA PRUEBA (z) DE LA DESVIACIÓN NORMALIZADA

Aunque gran parte de la exposición anterior fue abstracta y aparentemente poco práctica, es necesario conocerla para comprender las distribuciones muestrales subyacentes en muchas de las pruebas estadísticas que veremos a continuación. Una de esas pruebas, la prueba z de la desviación normalizada, se utiliza cuando conocemos los parámetros de la población de la hipótesis nula. En la prueba z la media muestral se utiliza como estadístico básico. Consideremos un experimento en el cual la prueba z resulta apropiada.

un experimento : Evaluación del programa de lectura en las escuelas públicas locales

Supongamos que usted es un inspector de las escuelas públicas de la ciudad donde vive. Recientemente, los ciudadanos locales han estado preocupados pues creen que el programa de lectura impartido en las escuelas públicas es de mala calidad. Como este asunto es grave, usted decide realizar un experimento para investigar la cuestión. Establece $\alpha = 0.05_{\text{cola}}$ para tomar su decisión. Para empezar, usted compara el nivel de lectura de los estudiantes del último año de bachillerato con las normas establecidas. Estas normas están basadas en las calificaciones de un examen que mide la habilidad de lectura y que fue aplicado a nivel nacional a un gran número de alumnos de este sector. Las calificaciones de esta población presentan una distribución normal, con $\mu = 75$ y $\sigma = 16$. Para su experimento, usted aplica el examen de lectura a 100 estudiantes de bachillerato, elegidos al azar, de su ciudad. La media muestral obtenida (\bar{X}_{obt}) = 72. ¿Cuál es su conclusión?

No cabe duda de que un valor de 72 como media muestral es menor que la media nacional de 75. Sin embargo, ¿es significativamente menor? Si solamente interviene el azar, entonces podemos considerar a los 100 datos muestrales como una muestra aleatoria de una población con $\mu = 75$. La hipótesis nula para este experimento establece que éste es efectivamente el caso. ¿Cuál es la probabilidad de obtener una media tan baja o más baja aún que 72, si los 100 datos provienen de una muestra aleatoria de una población con distribución normal, con una media igual a 75 y una desviación estándar de 16? Si la probabilidad resulta igual o menor que alfa, entonces rechazaremos H_0 y aceptaremos H_1 . En caso contrario, conservaremos H_0 . Resulta claro que el estadístico que estamos utilizando es la *media* de la muestra. Para determinar la probabilidad adecuada, *debemos conocer la distribución muestral de la media*.

En la siguiente sección analizaremos la distribución muestral de la media. Por el momento, dejaremos de lado al inspector y su problema, pero regresaremos a él posteriormente. Por ahora, basta con observar que utilizaremos la media muestral para evaluar H_0 , y para esto, debemos conocer la distribución muestral de la media.

Distribución muestral de la media

Si aplicamos la definición de la distribución muestral de un estadístico a la media, obtenemos lo siguiente:

definición ■ *La distribución muestral de la media proporciona todos los valores que puede asumir la media, junto con la probabilidad de obtener cada valor si se hace un muestreo aleatorio a partir de la población de la hipótesis nula.*

Podemos determinar la distribución muestral de la media empíricamente y en forma teórica; en este último caso se emplea el teorema del límite central. La deducción teórica es compleja y está más allá del nivel de este texto. En consecuencia, por razones pedagógicas, preferimos presentar aquí el enfoque empírico. Cuando aplicamos este enfoque, podemos determinar la distribución muestral de la media tomando una población específica de datos en bruto que tiene una media μ y una desviación estándar σ , 1) extrayendo todas las distintas muestras posibles de un tamaño fijo N , 2) calculando la media de cada muestra y 3) calculando la probabilidad de obtener cada valor de la media si sólo interviene el azar. Este proceso se muestra en la figura 12.4. Después de realizar estos tres pasos, habremos deducido la distribución muestral de la media para muestras de tamaño N extraídas de una población específica con media μ y desviación estándar σ . Esta distribu-

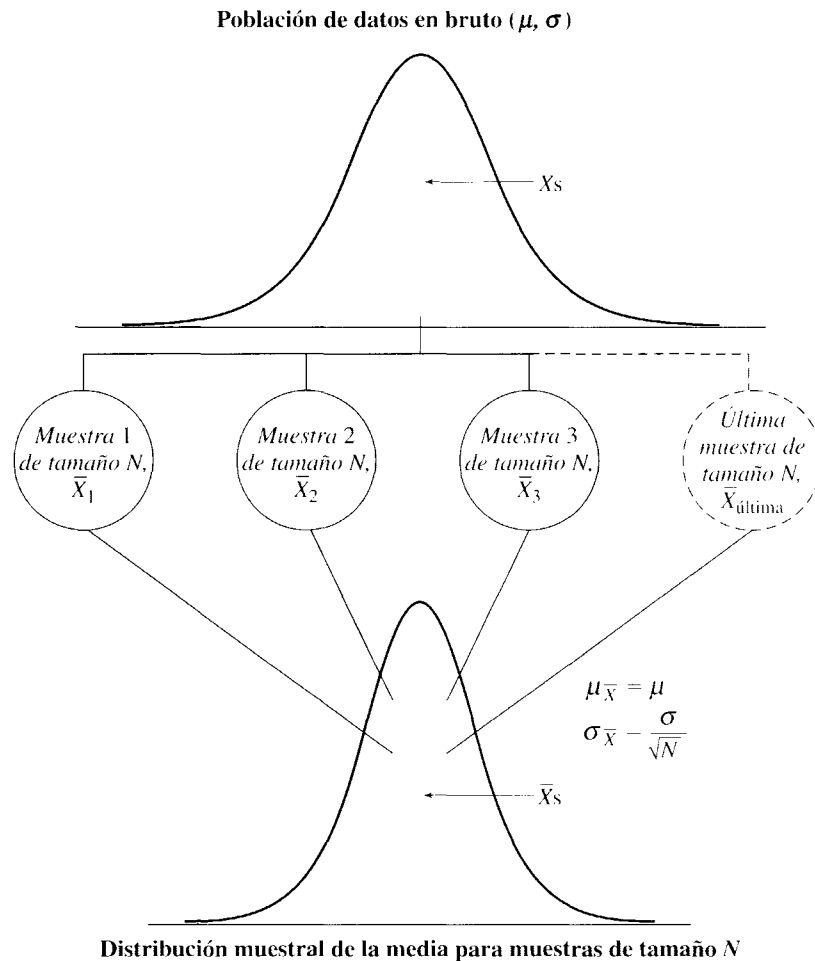


figura 12.4 Generación de la distribución muestral de la media para muestras de tamaño N extraídas de una población de datos en bruto.

ción muestral de la media nos proporcionará todos los valores que puede asumir la media para muestras de tamaño N , junto con la probabilidad de obtener cada valor si se realiza un muestreo aleatorio a partir de la población especificada. Al repetir este procedimiento en tres pasos para poblaciones con datos diferentes y variar de manera sistemática N , podemos establecer que la distribución muestral de la media tiene las siguientes características generales. Para muestras de un tamaño arbitrario N , la distribución muestral de la media posee las siguientes características:

1. Es una distribución de datos, cada uno de los cuales es una media muestral de N datos. Esta distribución tiene una media y una desviación estándar, como se indica en la parte inferior de la figura 12.4. Conviene observar que ésta es un conjunto de datos *poblacionales*, aunque los datos se basen en muestras, porque la distribución contiene al conjunto *completo* de medias muestrales. Simbolizaremos a la media de la distribución como $\mu_{\bar{x}}$ y a la desviación estándar como $\sigma_{\bar{x}}$. Por lo tanto,

$\mu_{\bar{x}}$ = media de la distribución muestral de la media

$\sigma_{\bar{x}}$ = desviación estándar de la distribución muestral de la media

$\sigma_{\bar{x}}$ se conoce también como el *error estándar de la media* porque cada una de las medias muestrales se puede considerar como una estimación de la media poblacional de los datos en bruto. La variabilidad entre las medias muestrales se presenta entonces debido a los errores de estimación, por eso se designa mediante la expresión *error estándar de la media* para $\sigma_{\bar{x}}$.

2. Tiene una media igual a la media poblacional de los datos en bruto. En forma de ecuación sería

$$\mu_{\bar{x}} = \mu$$

3. Tiene una desviación estándar igual a la desviación estándar poblacional de los datos en bruto, dividida entre \sqrt{N} . En forma de ecuación sería

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}}$$

4. Tiene la forma normal y depende de la forma de la población de datos en bruto y del tamaño de la muestra.

La primera característica es bastante obvia; sólo consiste en que la distribución muestral de la media está formada por datos que son medias muestrales. Como tal, debe tener también una media y una desviación estándar. La segunda característica consiste en que la media de la distribución muestral de la media es igual a la media de los datos en bruto ($\mu_{\bar{x}} = \mu$). Tendremos una idea más clara de esta relación si reconocemos que cada media muestral es una estimación de la media poblacional de los datos en bruto. Cada una mostrará diferencias con respecto a la media poblacional de los datos en bruto debido al azar. A veces, la media muestral será mayor que la media poblacional, y otras será menor, también debido al azar. Al considerar más medias muestrales, el promedio de éstas se acercará a la media poblacional de los datos en bruto, puesto que los factores aleatorios se cancelarán. Por último, cuando consideremos la totalidad de las distintas medias muestrales posibles, el promedio de todas ellas será igual a la media poblacional de los datos en bruto ($\mu_{\bar{x}} = \mu$).

La tercera característica consiste en que la desviación estándar de la distribución muestral de la media es igual a la desviación estándar poblacional de los datos en bruto dividida entre \sqrt{N} ($\sigma_{\bar{x}} = \sigma/\sqrt{N}$). Esto significa que la desviación estándar de la distribu-

ción muestral de la media varía en proporción directa con la desviación estándar poblacional de los datos en bruto y en proporción inversa con \sqrt{N} . Es bastante obvio el porqué $\sigma_{\bar{x}}$ debe variar en proporción directa con respecto a σ . Si los datos de la población son más variables, σ crece y lo mismo ocurre con la variabilidad entre las medias basadas en estos datos. La razón del porqué $\sigma_{\bar{x}}$ varía en proporción inversa con respecto a \sqrt{N} es un poco más difícil de entender. La clave consiste en reconocer que cada media muestral es una estimación de la media poblacional de los datos en bruto. A medida que N (el número de datos de cada muestra) aumenta, cada media muestral se convierte en una estimación más precisa de μ . Conforme las medias muestrales sean cada vez más precisas, variarán menos de una muestra a otra, haciendo que la varianza varíe en proporción inversa con respecto a $(\sigma_{\bar{x}}^2)$ de las medias muestrales disminuya. Así, $\sigma_{\bar{x}}^2$ varía en proporción inversa con respecto a N . Como $\sigma_{\bar{x}} = \sqrt{\sigma_{\bar{x}}^2}$, entonces $\sigma_{\bar{x}}$ varía en proporción inversa con \sqrt{N} . También queremos señalar que, como la desviación estándar de la distribución muestral de la media, $(\sigma_{\bar{x}})$ cambia con el tamaño de la muestra, existe una distribución muestral de la media que es diferente para cada tamaño de la muestra. Esto parece razonable, porque si el tamaño de la muestra se modifica, entonces los datos de cada muestra cambian y, en consecuencia, también las medias muestrales se modifican. Así, la distribución muestral de la media para muestras de tamaño 10 tiene que ser distinta de la distribución muestral de la media para muestras de tamaño 20, y así sucesivamente.

En relación con el cuarto punto, dos factores determinan la forma de la distribución muestral de la media: 1) la forma de la población de los datos en bruto y 2) el tamaño de la muestra (N). En lo que se refiere al primer factor, si la población de datos en bruto tiene una distribución normal, entonces la distribución muestral de la media también tendrá una distribución normal, sin importar cuál sea el tamaño de la muestra. Sin embargo, si la población de los datos en bruto no está distribuida de manera normal, la forma de la distribución muestral dependerá del tamaño de la muestra. *El teorema del límite central establece que, sin importar la forma que presente la población de los datos en bruto, la distribución muestral de la media tiende a distribuirse en forma normal cuando el tamaño de la muestra, N , aumenta.* Si N es bastante grande, la distribución muestral de la media es aproximadamente normal. ¿De qué tamaño tendrá que ser N para que la distribución muestral de la media se considere normal? Esto depende de la forma que tenga la población de los datos en bruto. Mientras más se alejen estos datos de la normalidad, mayor será el tamaño necesario de la muestra para que la distribución muestral de la media presente una forma normal. Cuando $N \geq 300$, la forma de la población de datos en bruto ya no es importante. Con este tamaño N , independientemente de la forma que tenga la población de los datos en bruto, la distribución muestral de la media se desviará tan poco de la normalidad que, para los cálculos estadísticos, podemos considerarla como una distribución normal. Como la mayor parte de las poblaciones que encontramos en las ciencias del comportamiento no difieren mucho de la normalidad, si $N \geq 30$, por lo general se supone que la distribución muestral de la media tiene forma normal.*

A pesar de que la demostración de estas características rebasa los alcances del presente texto, podemos ilustrarlas y, con ello, lograr una mejor comprensión de la distribución muestral de la media, al considerar una población dada y deducir la distribución muestral de la media para las muestras extraídas de ella. Para simplificar nuestros cálculos, utilizaremos una población con un número pequeño de datos. Para este ejemplo, supondremos que los datos en bruto de la población son 2, 3, 4, 5 y 6. La media (μ) de la población es igual a 4.00 y la desviación estándar (σ) es de 1.41. Queremos deducir la distribución muestral de la media para muestras de tamaño 2 extraídas de esta población. Una vez más, supondremos que el muestreo se realiza tomando un dato a la vez y con

*Existen algunas excepciones notables a esta regla; por ejemplo, los datos correspondientes a tiempos de reacción.

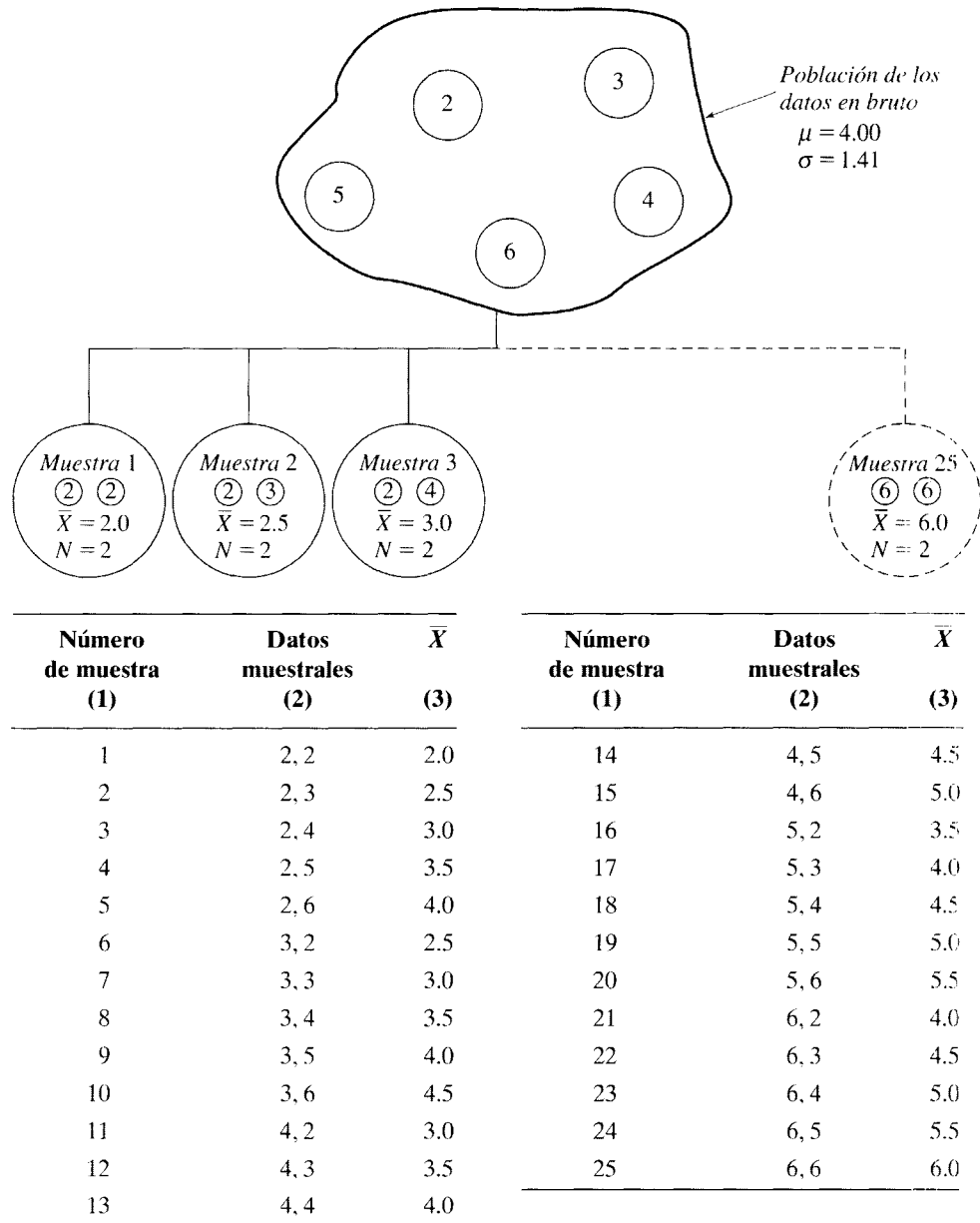


figura 12.5 Todas las posibles muestras de tamaño 2 que es posible extraer de una población formada por los datos en bruto 2, 3, 4, 5 y 6. El muestreo se realiza tomando un dato a la vez, con reemplazo.

reemplazo. El primer paso consiste en extraer de la población todas las distintas muestras posibles de tamaño 2. La figura 12.5 ilustra la población de los datos en bruto y, en forma esquemática, las diferentes muestras de tamaño 2 que es posible extraer de ella. Existen 25 muestras distintas posibles de tamaño 2. La lista de muestras se presenta en la segunda columna de la tabla de la figura 12.5. A continuación debemos calcular la media de cada muestra. Los resultados aparecen en la columna 3 de esta tabla. Ahora es fácil calcular la probabilidad de obtener cada uno de los valores de la media. Así,

$$\begin{aligned} p(\bar{X} = 2.0) &= \frac{\text{Número de } \bar{X} \text{ posibles de valor } 2.0}{\text{Número total de } \bar{X}} = \frac{1}{25} = 0.04 \\ p(\bar{X} = 2.5) &= \frac{2}{25} = 0.08 \\ p(\bar{X} = 3.0) &= \frac{3}{25} = 0.12 \\ p(\bar{X} = 3.5) &= \frac{4}{25} = 0.16 \\ p(\bar{X} = 4.0) &= \frac{5}{25} = 0.20 \\ p(\bar{X} = 4.5) &= \frac{4}{25} = 0.16 \\ p(\bar{X} = 5.0) &= \frac{3}{25} = 0.12 \\ p(\bar{X} = 5.5) &= \frac{2}{25} = 0.08 \\ p(\bar{X} = 6.0) &= \frac{1}{25} = 0.04 \end{aligned}$$

Ahora hemos deducido la distribución muestral de la media para muestras de tamaño $N = 2$ extraídas de una población constituida por los datos en bruto 2, 3, 4, 5 y 6. Determinamos todos los valores posibles de la media a partir de muestreos de dos datos de la población dada, junto con la probabilidad de obtener cada valor de la media si se realiza un muestreo aleatorio de la población. La distribución muestral completa se presenta en la tabla 12.1.

Suponga que por alguna razón queremos determinar la probabilidad de obtener $\bar{X} \geq 5.5$ como resultado de un muestreo aleatorio de dos datos, considerando uno a la vez, con reemplazo, tomados a partir de la población de datos en bruto. Podemos determinar la respuesta, consultando la distribución muestral de la media para $N = 2$. ¿Por qué? Porque esta distribución contiene todos los posibles valores de la media y su probabilidad, suponiendo que el muestreo es aleatorio. Así pues,

$$p(\bar{X} > 5.5) = 0.08 + 0.04 = 0.12$$

Ahora, consideremos las características de esta distribución; en primer lugar, su forma. La población original de datos en bruto y la distribución muestral se presentan en la figura 12.6 (a) y (b), respectivamente. En la parte (c) graficamos la distribución muestral de la media, con $N = 3$. Observe que la forma de las dos distribuciones muestrales difiere considerablemente de la población de datos en bruto. *Aun con una N tan pequeña como 3 y una población de datos en bruto con distribución distinta de la normal, la forma de la distribución muestral de la media se aproxima a la normalidad.* Esto ilustra lo que nos dice el teorema del límite central: que cuando N aumenta, la forma de la distribución muestral de la media se acerca a la de una distribución normal. Por supuesto, si la forma de la población de los datos en bruto fuera normal de por sí, la forma de la distribución muestral de la media también lo sería.

tabla 12.1
Distribución muestral de la media con $N = 2$ y datos de la población iguales a 2, 3, 4, 5 y 6

\bar{X}	$p(\bar{X})$
2.0	0.04
2.5	0.08
3.0	0.12
3.5	0.16
4.0	0.20
4.5	0.16
5.0	0.12
5.5	0.08
6.0	0.04

Ahora demostraremos que $\mu_{\bar{X}} = \mu$:

$$\begin{aligned} \mu &= \frac{\sum X}{\text{Número de datos en bruto}} & \mu_{\bar{X}} &= \frac{\sum \bar{X}}{\text{Número de datos de las medias}} \\ &= \frac{20}{5} = 4.00 & &= \frac{100}{25} = 4.00 \end{aligned}$$

Por lo tanto,

$$\mu_{\bar{X}} = \mu$$

Determinamos la media de los datos en bruto dividiendo la suma de estos datos entre el número de los mismos: $\mu = 4.00$. Calculamos la media de la distribución muestral de la

(a) Población de datos en bruto

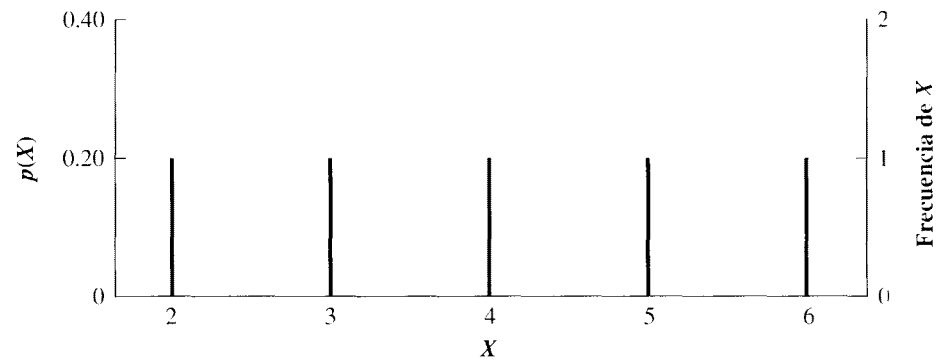
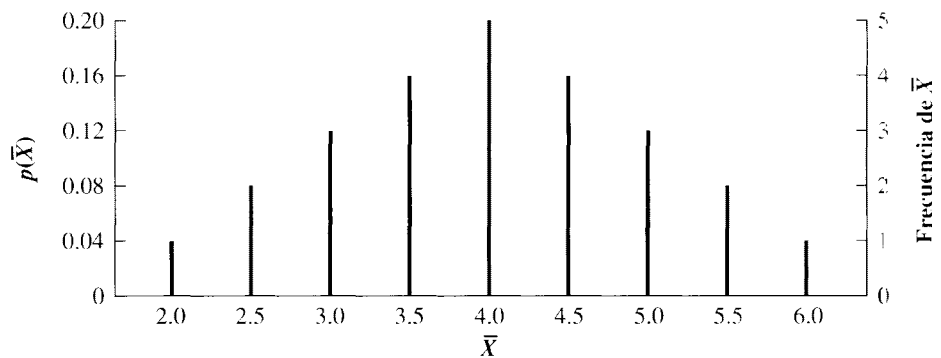
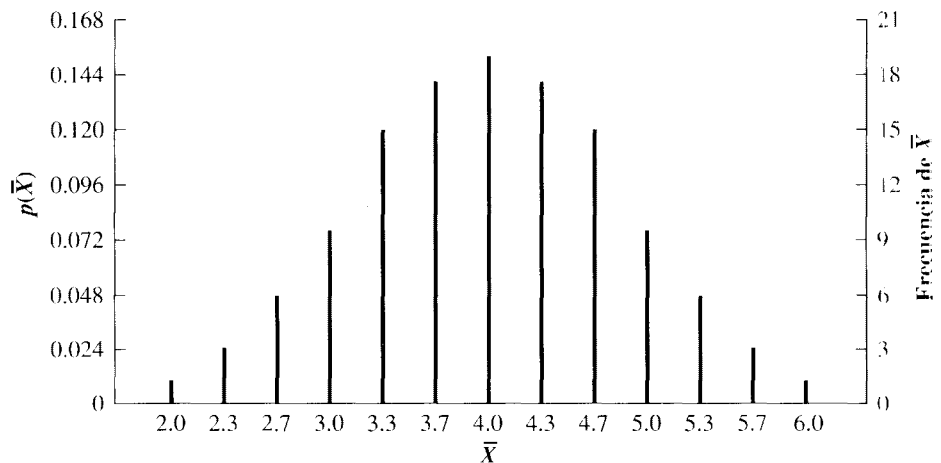
(b) Distribución muestral de \bar{X} con $N = 2$ (c) Distribución muestral de \bar{X} con $N = 3$ 

figura 12.6 Datos de la población y la distribución muestral de la media para muestras de tamaño $N = 2$ y $N = 3$.

media dividiendo la suma de las medias de las muestras entre el número de medias obtenidas: $\mu_{\bar{X}} = 4.00$. Así, $\mu_{\bar{X}} = \mu$.

Por último, es necesario señalar que $\sigma_{\bar{X}} = \sigma/\sqrt{N}$. El valor de $\sigma_{\bar{X}}$ se puede calcular de dos maneras: 1) a partir de la ecuación: $\sigma_{\bar{X}} = \sigma/\sqrt{N}$, o bien, 2) directamente a partir de los propios datos correspondientes a las medias muestrales. En nuestra demostra-

ción calculamos $\sigma_{\bar{X}}$ de las dos maneras y demostramos que conducen al mismo valor. Los cálculos aparecen en la tabla 12.2. Como ambos métodos proporcionan el mismo valor ($\sigma_{\bar{X}} = 1.00$), hemos demostrado que:

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}}$$

Observe que la N de la ecuación anterior es el número de datos que hay en cada muestra. Así pues, hemos demostrado que:

1. $\mu_{\bar{X}} = \mu$
2. $\sigma_{\bar{X}} = \sigma/\sqrt{N}$
3. La distribución muestral de la media asume una forma similar a la normal, aunque los datos en bruto no se distribuyan de manera normal.

Revisión del experimento sobre la habilidad de lectura

Ahora podemos regresar con el “inspector” y evaluar los datos del experimento donde se evalúa la habilidad para la lectura. Recordemos el experimento.

Usted es un inspector de escuelas públicas y realiza un experimento para investigar si la habilidad de lectura de los estudiantes de último año de bachillerato de la ciudad donde vive es deficiente. Una muestra aleatoria de 100 estudiantes de esa población mostró una habilidad media de lectura igual a 72 ($\bar{X}_{\text{obt}} = 72$). Las normas nacionales sobre la habilidad de lectura, para estudiantes de bachillerato, muestran una distribución normal de los datos con una media de 75 ($\mu = 75$) y una desviación estándar igual a 16 ($\sigma = 16$). ¿Es razonable considerar a los 100 datos como una muestra aleatoria de una población normal de calificaciones de lectura, donde $\mu = 75$ y $\sigma = 16$? Utilice $\alpha = 0.05$ cola.

Si tomamos todas las muestras posibles de tamaño 100, de la población de los datos de lectura que presentan una distribución normal, podemos determinar la distribución muestral de las medias muestrales con $N = 100$. De acuerdo a lo dicho anteriormente, esta distribución 1) tiene forma normal, 2) posee una media $\mu_{\bar{X}} = \mu = 75$, y 3) tiene una desviación estándar $\sigma_{\bar{X}} = \sigma/\sqrt{N} = 16/\sqrt{100} = 1.6$. Las dos distribuciones se presentan en la figura 12.7. Observe que la distribución muestral de la media contiene a todos los valores posibles de la media para muestras de tamaño 100, extraídas de la población de la hipótesis nula ($\mu = 75$, $\sigma = 16$). Para mayor claridad, en nuestra siguiente exposición hemos trazado nuevamente la distribución muestral de la media y la presentamos sola en la figura 12.8.

El área sombreada de la figura 12.8 contiene todos los valores muestrales de la media para $N = 100$ que son iguales o menores que $\bar{X}_{\text{obt}} = 72$. La proporción entre el área sombreada y el área total nos indica la probabilidad de obtener una media muestral me-

tabla 12.2 Demostración de que $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}}$

Utilizando $\sigma_{\bar{X}} = \sigma/\sqrt{N}$	Utilizando las medias muestrales
$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}}$	$\sigma_{\bar{X}} = \sqrt{\frac{\sum(\bar{X} - \mu_{\bar{X}})^2}{\text{Número de medias muestrales}}}$
$= \frac{1.41}{\sqrt{2}}$	$= \sqrt{\frac{(2.0 - 4.0)^2 + (2.5 - 4.0)^2 + \dots + (6.0 - 4.0)^2}{25}}$
$= 1.00$	$= \sqrt{\frac{25}{25}} = 1.00$

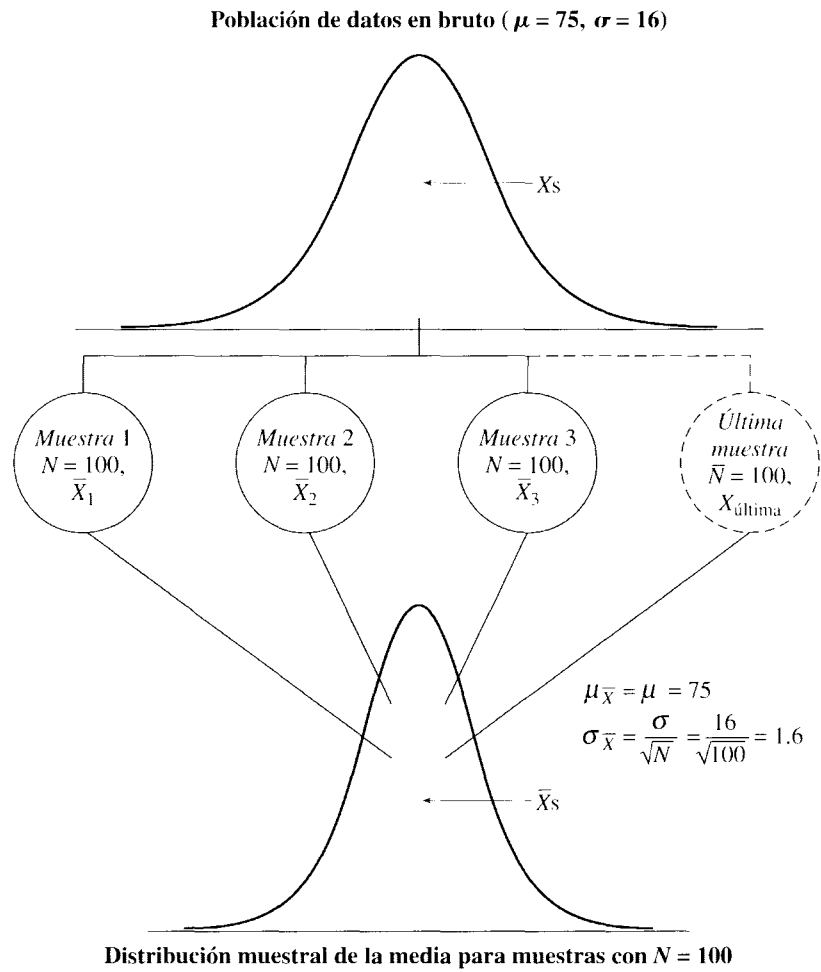


figura 12.7 Distribución muestral de la media para muestras de tamaño $N = 100$ extraídas de una población de datos en bruto con $\mu = 75$ y $\sigma = 16$.

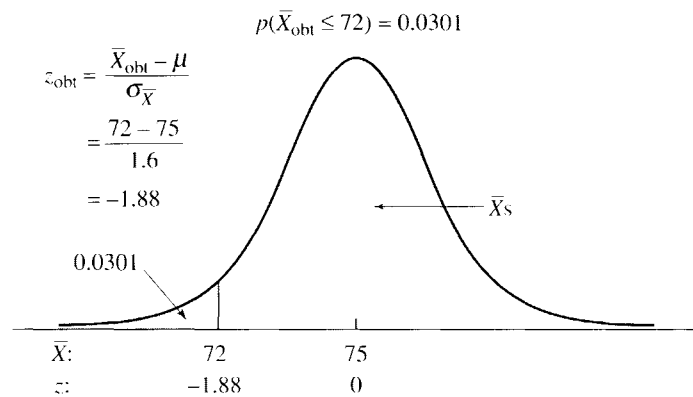


figura 12.8 Evaluación de los datos sobre la habilidad para la lectura, comparando la probabilidad obtenida con el nivel alfa.

nor o igual a 72 si solamente interviene el azar (otra forma de decir esto es: "si se trata de una muestra aleatoria proveniente de la población de la hipótesis nula"). Como la distribución muestral de la media tiene una forma normal, podemos determinar la proporción del área sombreada 1) al calcular la transformación z (z_{obt}) para $\bar{X}_{\text{obt}} = 72$ y 2) determinar el área apropiada en la tabla A, apéndice D, utilizando z_{obt} .

La ecuación para z_{obt} es muy similar a la ecuación para z del capítulo 5, excepto que en lugar de trabajar con datos en bruto, ahora lo hacemos con medias. En la tabla 12.3 se presentan las dos ecuaciones.

tabla 12.3 Ecuaciones z

Datos en bruto	Medias
$z = \frac{X - \mu}{\sigma}$	$z_{\text{obt}} = \frac{\bar{X}_{\text{obt}} - \mu_{\bar{X}}}{\sigma_{\bar{X}}}$

Puesto que $\mu_{\bar{X}} = \mu$, la ecuación z_{obt} para las medias muestrales se simplifica así:

$$z_{\text{obt}} = \frac{\bar{X}_{\text{obt}} - \mu}{\sigma_{\bar{X}}} \quad \text{transformación } z \text{ para } \bar{X}_{\text{obt}}$$

Si calculamos z_{obt} para este experimento, obtenemos:

$$\begin{aligned} z_{\text{obt}} &= \frac{\bar{X}_{\text{obt}} - \mu}{\sigma_{\bar{X}}} \\ &= \frac{72 - 75}{1.6} \\ &= -1.88 \quad \text{para } \bar{X}_{\text{obt}} = 72 \end{aligned}$$

De la columna C de la tabla A del apéndice D,

$$p(\bar{X}_{\text{obt}} \leq 72) = 0.0301$$

Como quiera que $0.0301 < 0.05$, rechazamos H_0 y concluimos que no es razonable suponer que los 100 datos sean una muestra aleatoria de una población donde $\mu = 75$. Tal parece que la habilidad de lectura que han demostrado los estudiantes del último año de bachillerato de la ciudad donde usted vive es deficiente.

Solución alternativa usando z_{obt} y la región crítica para el rechazo de H_0

Podemos analizar los resultados de este experimento de otra manera. En realidad, este método resulta preferible porque es más sencillo y al aplicarlo se establece un patrón para las pruebas de inferencia subsecuentes. Sin embargo, se basa en el método anterior y, por lo tanto, nos era imposible presentarlo antes. Para utilizar este método, definiremos primero algunos términos.

- definiciones**
- **La región crítica para el rechazo de la hipótesis nula** es el área debajo de la curva que contiene a todos los valores del estadístico que permiten el rechazo de la hipótesis nula.
 - **El valor crítico de un estadístico** es aquel que delimita la región crítica.

Para analizar los datos por medio de este método alternativo, basta calcular z_{obt} , determinar el valor crítico de z (z_{crit}), y verificar si z_{obt} queda localizada dentro de la región crítica para el rechazo de H_0 . Ya sabemos cuál es la forma de calcular z_{obt} .

La región crítica que permite rechazar H_0 queda determinada por el nivel alfa. Por ejemplo, si $\alpha = 0.05_{1\text{ cola}}$ en la dirección que predice un valor negativo de z_{obt} como en el ejemplo anterior, entonces la región crítica para el rechazo de H_0 es el área que se encuentra debajo de la cola izquierda de la curva, que es igual a 0.0500. Con el fin de determinar z_{crit} para esta área, hemos de utilizar la tabla A en sentido inverso. Así pues, consultamos la tabla A y nos desplazamos por la columna C hasta encontrar 0.0500, para determinar el valor z correspondiente a 0.0500. Observamos que 0.0500 se localiza a la mitad del camino entre los datos z de 1.64 y 1.65. Por consiguiente, el valor z que corresponde a 0.0500 es 1.645. En virtud de que hemos trabajado con la cola izquierda de la distribución.

$$z_{\text{crit}} = -1.645$$

Este dato define la región crítica para el rechazo de H_0 y, por lo tanto, se denomina z_{crit} . Si z_{obt} queda en la región crítica para el rechazo, rechazaremos H_0 . Estas relaciones se presentan en la figura 12.9(a). Si $\alpha = 0.05_{1\text{ cola}}$ en la dirección que predice un valor positivo z_{obt} entonces

$$z_{\text{crit}} = 1.645$$

Esto se muestra en la figura 12.9(b). Si $\alpha = 0.05_{2\text{ colas}}$, entonces el área conjunta bajo las dos colas de la curva deberá ser igual a 0.0500. De esta manera, el área debajo de cada cola deberá ser igual a 0.0250 como se aprecia en la figura 12.9(c). Para esta área:

$$z_{\text{crit}} = \pm 1.96$$

Para rechazar H_0 , la media muestral obtenida (\bar{X}_{obt}) deberá tener un valor z transformado (z_{obt}) que quede dentro de la región crítica para el rechazo.

Ahora utilizaremos estos conceptos para analizar los datos del examen de lectura. En primer lugar, calcularemos z_{obt} :

$$\begin{aligned} z_{\text{obt}} &= \frac{\bar{X}_{\text{obt}} - \mu}{\sigma_{\bar{X}}} \\ &= \frac{72 - 75}{1.6} = \frac{-3}{1.6} \\ &= -1.88 \end{aligned}$$

El siguiente paso consiste en determinar z_{crit} . Como $\alpha = 0.05_{1\text{ cola}}$, el área debajo de la cola izquierda es igual a 0.0500. Para esta área, a partir de la tabla A obtenemos:

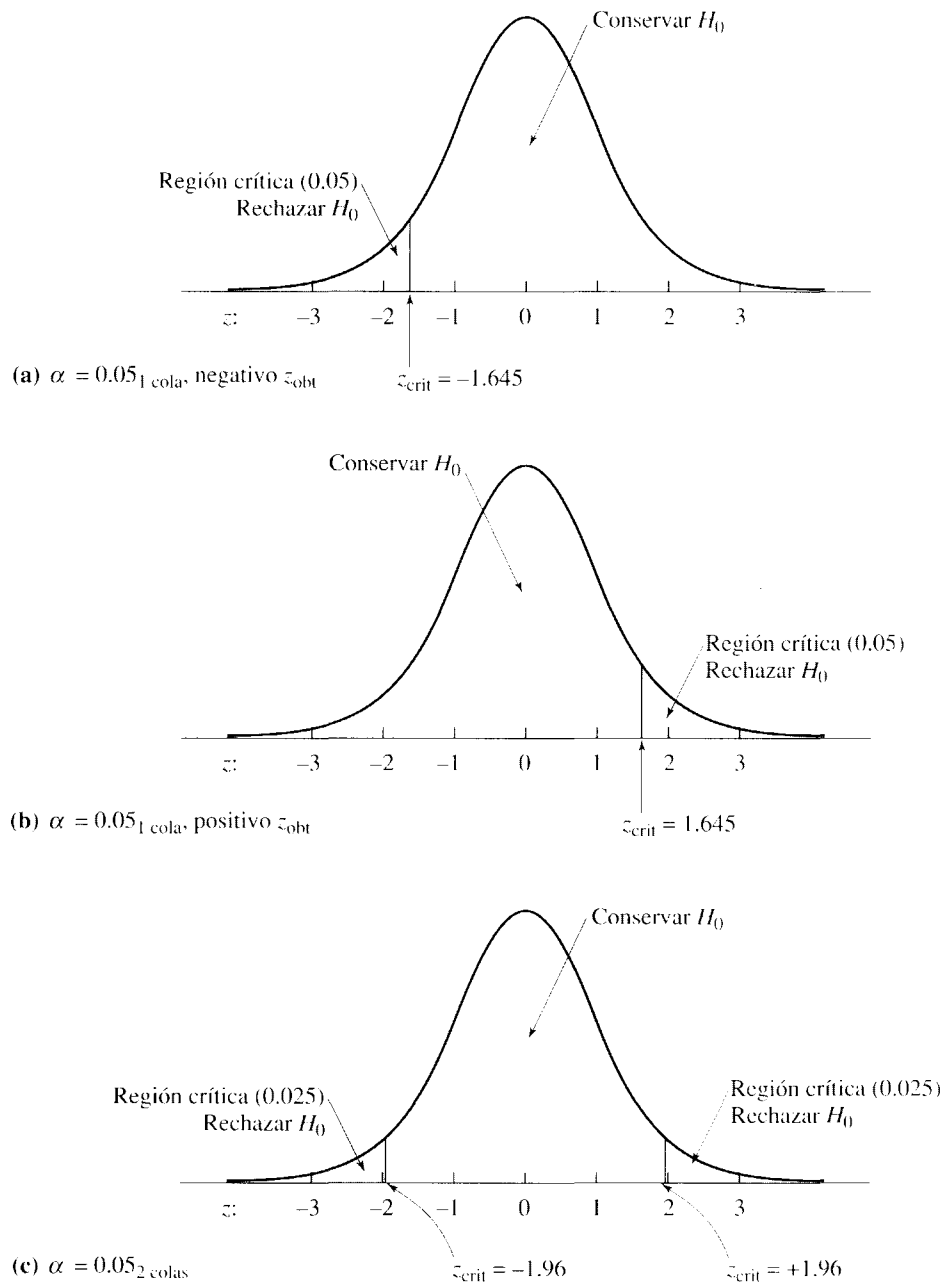
$$z_{\text{crit}} = -1.645$$

Por último, debemos determinar si z_{obt} queda dentro de la región crítica. En ese caso, rechazaremos la hipótesis nula. En caso contrario, la conservaremos. La regla de decisión establece lo siguiente:

Si $|z_{\text{obt}}| \geq |z_{\text{crit}}|$, se rechazará la hipótesis nula. En caso contrario, la hipótesis nula se conservará.

Observe que esta ecuación es sólo una manera abreviada de decir que si z_{obt} es positivo, entonces debe ser mayor o igual que $+z_{\text{crit}}$ para que quede dentro de la región crítica. Si z_{obt} es negativo, entonces debe ser igual o menor que $-z_{\text{crit}}$ para que quede dentro de la región crítica.

En el ejemplo presente, como $|z_{\text{obt}}| > 1.645$, rechazamos la hipótesis nula. La solución completa mediante este método se presenta en la figura 12.10. Queremos insistir en



Adaptado de *Fundamental Statistics for Behavioral Sciences* segunda edición, por Robert B. McCall, © 1998 por Brooks/Cole.

figura 12.9 Región crítica para el rechazo de H_0 para (a) $\alpha = 0.05$ $_1$ cola, z_{obt} negativo; (b) $\alpha = 0.05$ $_1$ cola, z_{obt} positivo, y (c) $\alpha = 0.05$ $_2$ colas.

que, al emplear este método, estamos siguiendo el procedimiento de dos pasos bosquejado en este capítulo para el análisis de los datos: 1) cálculo del estadístico apropiado y 2) evaluación del estadístico con base en su distribución muestral. En realidad, el experimentador calcula dos estadísticos, \bar{X}_{obt} y z_{obt} . Lo que se evalúa en último término es z_{obt} . Si la distribución muestral de \bar{X} tiene una forma de distribución normal, entonces la distribución z también será normal y las probabilidades apropiadas estarán dadas por la ta-

PASO 1: Cálculo del estadístico apropiado:

$$z_{\text{obt}} = \frac{\bar{X}_{\text{obt}} - \mu}{\sigma_{\bar{X}}} = \frac{72 - 75}{1.6} = -1.88$$

PASO 2: Evaluación del estadístico con base en su distribución muestral. La regla de decisión es la siguiente: si $|z_{\text{obt}}| \geq |z_{\text{crit}}|$, rechazamos H_0 . Como $\alpha = 0.05_{1 \text{ cola}}$, de acuerdo con la tabla A,

$$z_{\text{crit}} = -1.645$$

Vemos que $|z_{\text{obt}}| > 1.645$, queda dentro de la región crítica para el rechazo de H_0 . Por lo tanto, rechazamos H_0 .

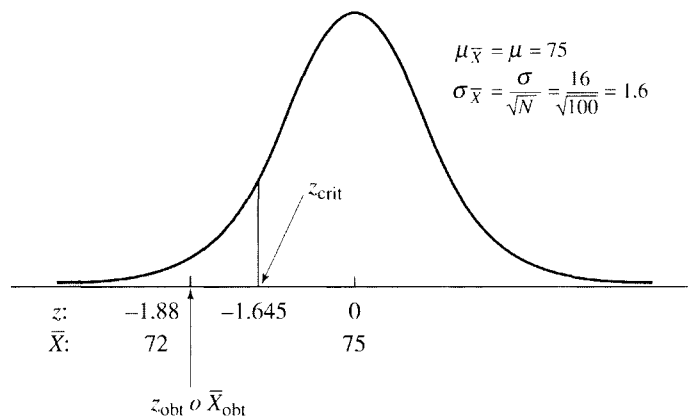


figura 12.10 Solución al experimento sobre la habilidad para la lectura, mediante z_{obt} y la región crítica.

bla A. Por supuesto, la distribución z tiene una media de 0 y una desviación estándar de 1, como lo analizamos en el capítulo 5.

Tratemos de resolver otro problema aplicando este enfoque.

El rector de una universidad tiene la impresión de que la edad promedio de los estudiantes que asisten a esa institución ha cambiado durante los últimos años. Para comprobar dicha hipótesis, se lleva a cabo un experimento en el cual se registra la edad de 150 alumnos elegidos al azar entre todos los estudiantes de este centro de educación superior. La edad promedio así registrada es de 23.5 años. Un censo realizado en la universidad unos cuantos años antes del experimento reveló que la edad promedio era de 22.4 años, con una desviación estándar de 7.6.

- ¿Cuál es la hipótesis alternativa no direccional?
- ¿Cuál es la hipótesis nula?
- Utilizando $\alpha = 0.05_{2 \text{ colas}}$ ¿cuál es la conclusión?

SOLUCIÓN

- a. Hipótesis alternativa no direccional: Durante los últimos años, la edad promedio de los estudiantes de la universidad ha cambiado. Por consiguiente, la muestra con $\bar{X}_{\text{obt}} = 23.5$ es una muestra aleatoria, extraída de una población en la cual $\mu \neq 22.4$.
- b. Hipótesis nula: De acuerdo con esta hipótesis, resulta razonable considerar a la muestra con $\bar{X}_{\text{obt}} = 23.5$ como una muestra aleatoria, extraída de una población con $\mu = 22.4$.
- c. Conclusión, utilizando $\alpha = 0.05_{2 \text{ colas}}$:

PASO 1: Cálculo del estadístico apropiado. Los datos se presentan con el problema.

$$\begin{aligned} z_{\text{obt}} &= \frac{\bar{X}_{\text{obt}} - \mu}{\sigma_{\bar{X}}} \\ &= \frac{\bar{X}_{\text{obt}} - \mu}{\sigma/\sqrt{N}} = \frac{23.5 - 22.4}{7.6/\sqrt{150}} \\ &= \frac{1.1}{0.6205} = 1.77 \end{aligned}$$

PASO 2: Evaluación del estadístico con base en su distribución muestral. La regla de decisión es la siguiente: si $|z_{\text{obt}}| \geq |z_{\text{crit}}|$, rechazamos H_0 . En caso contrario, conservamos H_0 . Puesto que $\alpha = 0.05_{2 \text{ colas}}$ de acuerdo con la tabla A,

$$z_{\text{crit}} = \pm 1.96$$

Como $|z_{\text{obt}}| < 1.96$, no queda dentro de la región crítica para el rechazo de H_0 . Por lo tanto, conservamos H_0 . No podemos concluir que la edad promedio de los estudiantes que asisten a la universidad haya cambiado.



El fabricante de cierta gasolina considera que un nuevo aditivo rendirá más millas por galón. La compañía ha realizado un gran número de mediciones de las millas recorridas con gasolina sin el aditivo, bajo condiciones controladas en forma rigurosa. Los resultados presentan una media de 24.7 millas por galón y una desviación estándar de 4.8. Se realizan pruebas con una muestra de 75 automóviles que utilizan la gasolina con el aditivo. La media de la muestra es igual a 26.5 millas por galón.

- a. Supongamos que existen suficientes fundamentos para realizar una prueba con una cola. ¿Cuál es la hipótesis alternativa direccional?
- b. ¿Cuál es la hipótesis nula?
- c. ¿Cuál es la conclusión? Utilice $\alpha = 0.05_{1 \text{ cola}}$.

(continúa)

SOLUCIÓN

- a. Hipótesis alternativa direccional: El nuevo aditivo incrementa el número de millas de rendimiento por galón. Por lo tanto, la muestra con $\bar{X}_{\text{obt}} = 26.5$ es una muestra aleatoria, extraída de una población donde $\mu > 24.7$.
- b. Hipótesis nula: H_0 : La muestra con $\bar{X}_{\text{obt}} = 26.5$ es una muestra aleatoria, extraída de una población en la cual $\mu \leq 24.7$.
- c. Conclusión, utilizando $\alpha = 0.05_{1 \text{ cola}}$:

PASO 1: Cálculo del estadístico adecuado. Los datos se presentan en el problema.

$$\begin{aligned} z_{\text{obt}} &= \frac{\bar{X}_{\text{obt}} - \mu}{\sigma/\sqrt{N}} \\ &= \frac{26.5 - 24.7}{4.8/\sqrt{75}} = \frac{1.8}{0.5543} \\ &= 3.25 \end{aligned}$$

PASO 2: Evaluación del estadístico con base en su distribución muestral. La regla de decisión establece que si $|z_{\text{obt}}| \geq |z_{\text{crit}}|$, rechazamos H_0 . En caso contrario, conservamos H_0 . Como $\alpha = 0.05_{1 \text{ cola}}$ de acuerdo con la tabla A,

$$z_{\text{crit}} = 1.645$$

Como $|z_{\text{obt}}| > 1.645$, entonces se encuentra dentro de la región crítica para el rechazo de H_0 . Por lo tanto, rechazamos la hipótesis nula y concluimos que el aditivo sí incrementa las millas por galón.

Condiciones en las cuales resulta adecuada la prueba z

El empleo de la prueba z resulta apropiado cuando en el experimento se considera una sola media muestral (\bar{X}_{obt}) y los parámetros de la población de la hipótesis nula son conocidos (es decir, si se sabe cuáles son los valores de μ y σ). Además, para utilizar esta prueba, la distribución muestral de la media debe tener la forma normal. Para esto, por supuesto, se requiere que $N \geq 30$ o bien, que la propia población de la hipótesis nula tenga una distribución normal.* Este requisito de normalidad se conoce como el "supuesto matemático que subyace a la prueba z ".

La potencia y la prueba z

En el capítulo 11 comentamos la potencia, juntamente con la prueba del signo. Vamos a repasar algunos de los puntos principales presentados en ese capítulo.

1. Desde un punto de vista conceptual, la potencia es la sensibilidad del experimento para detectar un efecto real de la variable independiente cuando dicho efecto existe.
2. La potencia se define en términos matemáticos como la probabilidad de que el experimento produzca como resultado el rechazo de la hipótesis nula cuando la variable independiente tiene un efecto real.

* Muchos autores limitarían el uso de la prueba z a datos que tienen una escala de intervalo o de razón. Por favor, consulte la nota al pie de la página 31, en el capítulo 2, para ver algunas referencias en torno de este punto.

3. Potencia + beta = 1. Por lo tanto, la potencia varía en proporción inversa con respecto al valor de beta.
4. La potencia varía en proporción directa con respecto al valor de N . Aumentar N incrementa la potencia.
5. La potencia varía en proporción directa con respecto a la magnitud del efecto real de la variable independiente. La potencia de un experimento es mayor para efectos grandes que para efectos pequeños.
6. La potencia varía en proporción de manera directa con respecto al nivel alfa. Si alfa se vuelve más restrictivo, la potencia disminuye.

Estos conceptos sobre la potencia son verdaderos, independientemente de la prueba de inferencia. En esta sección ilustraremos de nuevo estas conclusiones, sólo que en esta ocasión las presentaremos junto con la prueba de la desviación normalizada. Comenzaremos con un análisis de la potencia y el tamaño de la muestra.

ejemplo

Potencia y tamaño de la muestra (N)

Regresemos al ejemplo ilustrativo presentado al principio del capítulo. Supondremos una vez más que usted ha asumido el papel de inspector de escuelas públicas. Sin embargo, ahora usted solamente va a diseñar el experimento, por lo cual, éste todavía no se ha realizado. Le interesa averiguar si el programa de lectura para los estudiantes del último año de bachillerato de la ciudad donde usted habita es deficiente. Como mencionamos con anterioridad, las normas nacionales de habilidad de lectura para este sector de alumnos presentan datos poblacionales cuya distribución es normal, con $\mu = 75$ y $\sigma = 16$. Usted piensa aplicar un examen a una muestra aleatoria de estudiantes del último año de bachillerato de la ciudad donde vive, y está tratando de determinar cuál debe ser el tamaño de la muestra. Utilizará $\alpha = 0.05$ para evaluar los datos que van a ser recolectados. Desea estar en condiciones de detectar las deficiencias de las medias de desempeño de su programa que difieran de las normas nacionales en 3 puntos o más. Es decir, que si el promedio de habilidad para la lectura en la población de estudiantes del último año de bachillerato de la ciudad donde vive es inferior a las normas nacionales en 3 puntos o más, usted quiere que su experimento tenga una alta probabilidad de detectarlo.

- a. Si usted ha decidido que va a utilizar un tamaño de muestra de 25 ($N = 25$), ¿cuál será la potencia de su experimento para detectar una deficiencia de 3 unidades en la habilidad media para la lectura en esa población, con respecto de las normas nacionales?
- b. Si usted aumenta el tamaño de la muestra a $N = 100$, ¿cuál será ahora la potencia para detectar una deficiencia de 3 unidades en la habilidad media de lectura en la población?
- c. ¿Qué tamaño de N debe emplearse en la potencia para detectar una deficiencia de 3 unidades en la habilidad media de lectura que sea de aproximadamente 0.9000?

SOLUCIÓN

- a. Potencia con $N = 25$.

Como dijimos en el capítulo 11, la potencia es la probabilidad de rechazar H_0 si la variable independiente tiene un efecto real. Para calcular la potencia y ser capaz de detectar el efecto real hipotético, debemos determinar primero los resultados muestrales que permitirán el rechazo de H_0 . A continuación, tenemos que calcular la probabilidad de obtener cualquiera de esos resultados muestrales si la variable independiente produce el efecto real hipotético. La probabilidad resultante es la potencia necesaria para detectar ese efecto real hipotético. Así pues, el cálculo de la potencia se realiza en dos pasos:

PASO 1: Determinamos todas las medias muestrales posibles del experimento que permitirán el rechazo de H_0 . Con la prueba z , esto significa determinar la región crítica para el rechazo de H_0 , utilizando \bar{X} como el estadístico.

PASO 2: Suponiendo que el efecto real hipotético de la variable independiente es verdadero, determinamos la probabilidad de obtener una media muestral que se encuentre en la región crítica para el rechazo de H_0 .

Calculemos ahora cuál es la potencia necesaria para detectar una deficiencia media de 3 puntos en la habilidad de la población para la lectura, con respecto a las normas nacionales, cuando se usa $N = 25$.

PASO 1: Determinamos todas las medias muestrales posibles del experimento que permitan el rechazo de H_0 . Con la prueba z , esto significa determinar la región crítica para el rechazo de H_0 utilizando \bar{X} como el estadístico.

Cuando se evalúa H_0 mediante la prueba z , suponemos que la muestra es aleatoria y ha sido extraída de la población de la hipótesis nula. Designaremos la media de la población de la hipótesis nula como μ_{nula} . En nuestro ejemplo, la población de la hipótesis nula es el conjunto de calificaciones establecidas por medio del examen nacional, es decir, se trata de una población normal de datos con $\mu_{nula} = 75$, donde $\alpha = 0.05_1$ cola y, $z_{crit} = -1.645$. Para determinar el valor crítico de \bar{X} , podemos emplear la ecuación de z , de la cual hemos despejado \bar{X}_{crit} :

$$z_{crit} = \frac{\bar{X}_{crit} - \mu_{nula}}{\sigma_{\bar{X}}}$$

$$\bar{X}_{crit} = \mu_{nula} + \sigma_{\bar{X}}(z_{crit})$$

Al sustituir los datos con $N = 25$,

$$\begin{aligned}\bar{X}_{crit} &= 75 + 3.2(-1.645) & \sigma_{\bar{X}} &= \frac{\sigma}{\sqrt{N}} = \frac{16}{\sqrt{25}} = 3.2 \\ &= 75 - 5.264 \\ &= 69.74\end{aligned}$$

Así pues, con $N = 25$, rechazaremos H_0 si al realizar el experimento la media de la muestra (\bar{X}_{obt}) ≤ 69.74 . En la figura 12.11 se han representado estas relaciones de manera gráfica.

PASO 2: Suponiendo que el efecto real hipotético de la variable independiente es verdadero, determinamos la probabilidad de obtener una media muestral que se encuentre en la región crítica para el rechazo de H_0 .

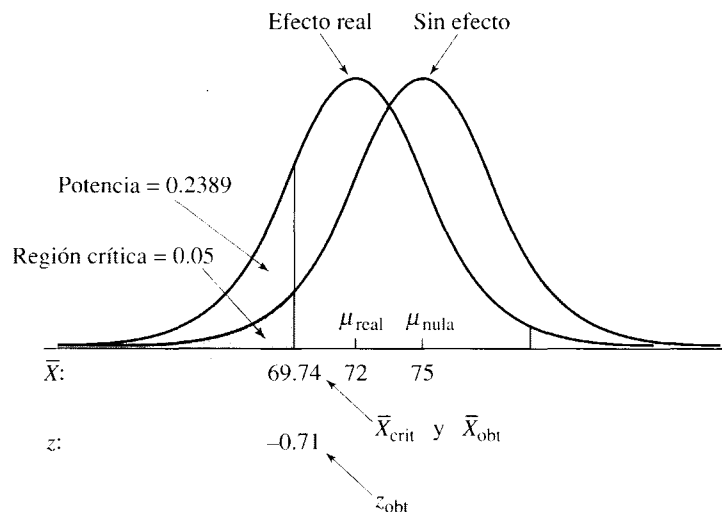


figura 12.11 Potencia para $N = 25$.

Si la variable independiente presenta el efecto real hipotético, entonces los datos muestrales del experimento no son una muestra aleatoria tomada de la población de la hipótesis nula. Más bien, son una muestra aleatoria de una población con una media especificada por el efecto real hipotético. Simbolizamos esta media como μ_{real} . De esta manera, si la habilidad de lectura de la población de estudiantes del último año de bachillerato de la ciudad donde usted vive está, en promedio, 3 puntos por debajo de las normas nacionales, entonces la muestra de su experimento es una muestra aleatoria extraída de una población en la que $\mu_{\text{real}} = 72$. La probabilidad de que la media muestral quede localizada en la región crítica, si la muestra es en realidad aleatoria y proviene de una población donde $\mu_{\text{real}} = 72$ se calcula obteniendo la transformación z para $\bar{X}_{\text{obt}} = 69.74$ y localizando el área correspondiente en la tabla A. De esta manera,

$$\begin{aligned} z_{\text{obt}} &= \frac{\bar{X}_{\text{obt}} - \mu_{\text{real}}}{\sigma_{\bar{X}}} & \sigma_{\bar{X}} &= \frac{\sigma}{\sqrt{N}} = \frac{16}{\sqrt{25}} = 3.2 \\ &= \frac{69.74 - 72}{3.2} \\ &= -0.71 \end{aligned}$$

Según la tabla A,

$$p(\bar{X}_{\text{obt}} \leq 69.74) = p(z_{\text{obt}} \leq -0.71) = 0.2389$$

Por lo tanto,

$$\text{Potencia} = 0.2389$$

$$\text{Beta} = 1 - \text{potencia} = 1 - 0.2389 = 0.7611$$

Así pues, la potencia para detectar una deficiencia media de 3 puntos promedio, con $N = 25$ es de 0.2389 y beta = 0.7611.

Al ver que la probabilidad de cometer un error de tipo II es demasiado alta, usted decide no continuar con el experimento para $N = 25$. Veamos ahora qué ocurre con la potencia y con beta si N es incrementado a 100.

- b. Si $N = 100$, ¿cuál es la potencia necesaria para detectar una diferencia poblacional media de 3 puntos en la habilidad para la lectura?

PASO 1: Determinamos todas las medias muestrales posibles del experimento que permitan el rechazo de H_0 . Con la prueba z, esto significa determinar la región crítica para el rechazo de H_0 , al utilizar \bar{X} como el estadístico:

$$\begin{aligned} \bar{X}_{\text{crit}} &= \mu_{\text{nula}} + \sigma_{\bar{X}}(z_{\text{crit}}) & \sigma_{\bar{X}} &= \frac{\sigma}{\sqrt{N}} = \frac{16}{\sqrt{100}} = 1.6 \\ &= 75 + 1.6(-1.645) \\ &= 75 - 2.632 \\ &= 72.37 \end{aligned}$$

PASO 2: Suponiendo que el efecto real hipotético de la variable independiente es verdadero, determinamos la probabilidad de obtener una media muestral que se encuentre en la región crítica para el rechazo de H_0 .

$$\begin{aligned} z_{\text{obt}} &= \frac{\bar{X}_{\text{obt}} - \mu_{\text{real}}}{\sigma_{\bar{X}}} & \sigma_{\bar{X}} &= \frac{\sigma}{\sqrt{N}} = \frac{16}{\sqrt{100}} = 1.6 \\ &= \frac{72.37 - 72}{1.6} \\ &= 0.23 \end{aligned}$$

Según la tabla A,

$$p(\bar{X}_{\text{obt}} \leq 72.37) = 0.5000 + 0.0910 = 0.5910$$

Por lo tanto,

$$\text{Potencia} = 0.5910$$

$$\text{Beta} = 1 - \text{potencia} = 1 - 0.5910 = 0.4090$$

Así pues, al incrementar N de 25 a 100, la potencia para detectar una deficiencia media de 3 puntos ha aumentado de 0.2389 a 0.5910. Beta ha disminuido de 0.7611 a 0.4090. Esto es una demostración de que la potencia varía en proporción directa con respecto a N y que beta varía en proporción inversa con respecto a N . De esta manera, al incrementarse N , la potencia aumenta y beta disminuye. La figura 12.12 resume las relaciones para este problema.

c. ¿Qué tamaño de N se debe utilizar para que la potencia sea aproximadamente igual a 0.9000?

Para que la potencia sea de 0.9000 a fin de detectar una deficiencia media de 3 puntos en la población, la probabilidad de que \bar{X}_{obt} quede en la región crítica debe ser igual a 0.9000. Como muestra la figura 12.13, esto implica que el área comprendida entre z_{obt} y $\mu_{\text{real}} = 0.4000$. Según la tabla A, $z_{\text{obt}} = 1.28$. (Observe que hemos considerado el valor más cercano en la tabla, en lugar de interpolarlo. Esto producirá una potencia próxima a 0.9000, pero no exactamente igual a 0.9000). Al despejar z_{obt} en la ecuación de \bar{X}_{obt} y hacer \bar{X}_{obt} igual a \bar{X}_{crit} , podemos determinar N . Por lo tanto,

$$\bar{X}_{\text{obt}} = \mu_{\text{real}} + \sigma_{\bar{X}}(z_{\text{obt}})$$

$$\bar{X}_{\text{crit}} = \mu_{\text{nula}} + \sigma_{\bar{X}}(z_{\text{crit}})$$

Estableciendo $\bar{X}_{\text{obt}} = \bar{X}_{\text{crit}}$, tenemos que,

$$\mu_{\text{real}} + \sigma_{\bar{X}}(z_{\text{obt}}) = \mu_{\text{nula}} + \sigma_{\bar{X}}(z_{\text{crit}})$$

Despejando N ,

$$\mu_{\text{real}} - \mu_{\text{nula}} = \sigma_{\bar{X}}(z_{\text{crit}}) - \sigma_{\bar{X}}(z_{\text{obt}})$$

$$\mu_{\text{real}} - \mu_{\text{nula}} = \sigma_{\bar{X}}(z_{\text{crit}} - z_{\text{obt}})$$

$$\mu_{\text{real}} - \mu_{\text{nula}} = \frac{\sigma}{\sqrt{N}}(z_{\text{crit}} - z_{\text{obt}})$$

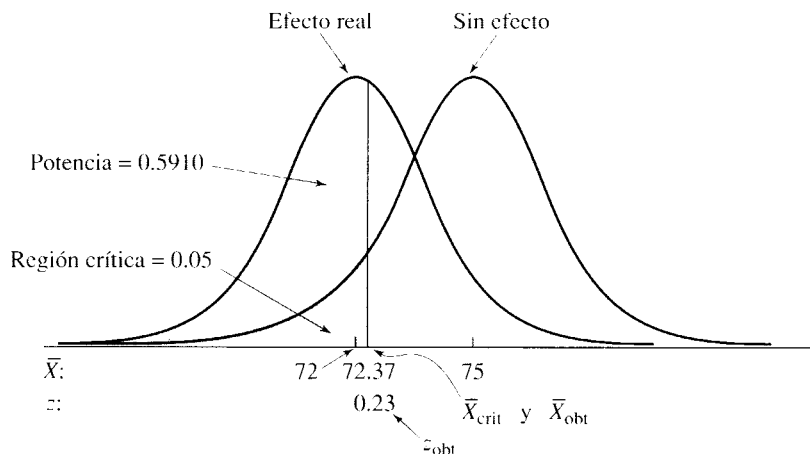


figura 12.12 Potencia para $N = 100$.

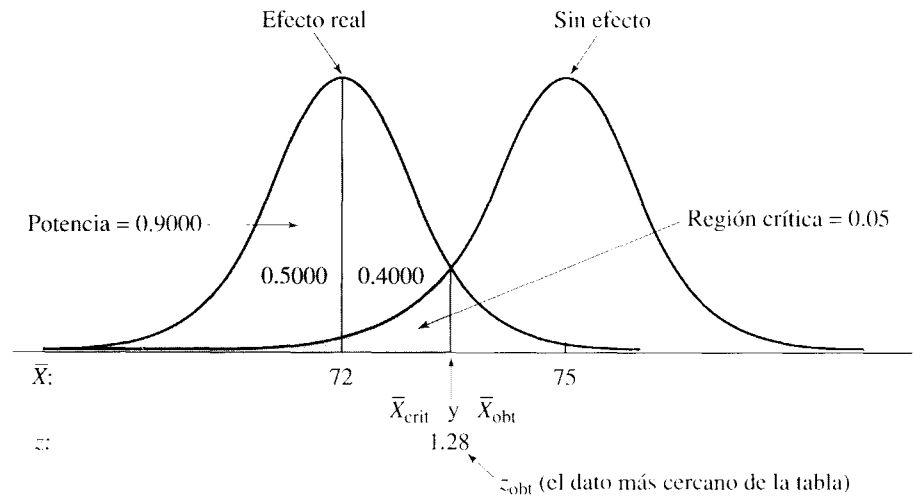


figura 12.13 Determinación de N para una potencia = 0.9000.

$$\sqrt{N}(\mu_{\text{real}} - \mu_{\text{nula}}) = \sigma(z_{\text{crit}} - z_{\text{obt}})$$

$$N = \left[\frac{\sigma(z_{\text{crit}} - z_{\text{obt}})}{\mu_{\text{real}} - \mu_{\text{nula}}} \right]^2$$

Así, para determinar N , la ecuación que usamos es

$$N = \left[\frac{\sigma(z_{\text{crit}} - z_{\text{obt}})}{\mu_{\text{real}} - \mu_{\text{nula}}} \right]^2 \quad \text{ecuación para determinar } N$$

Aplicando esta ecuación al problema que hemos estado considerando, tenemos que

$$\begin{aligned} N &= \left[\frac{\sigma(z_{\text{crit}} - z_{\text{obt}})}{\mu_{\text{real}} - \mu_{\text{nula}}} \right]^2 \\ &= \left[\frac{16(-1.645 - 1.28)}{72 - 75} \right]^2 \\ &= 243 \end{aligned}$$

Así vemos que, si N aumenta hasta 243 sujetos, la potencia apropiada para detectar una deficiencia media de 3 puntos en la capacidad de lectura en la población será aproximadamente igual a 0.9000 (potencia = igual a 0.8997). Le sugiero que confirme estos cálculos de potencia con $N = 243$ como ejercicio de práctica.

Potencia y nivel alfa Para continuar, demos un vistazo a la relación entre la potencia y alfa. Suponga que se hubiera establecido $\alpha = 0.01_{1 \text{ cola}}$ en vez de $0.05_{1 \text{ cola}}$. ¿Qué habría ocurrido con la potencia resultante? (Supondremos que $N = 100$ para esta pregunta.)

SOLUCIÓN

PASO 1: Determine todas las medias muestrales posibles del experimento que permitan el rechazo de H_0 . Con la prueba z , esto significa determinar la región crítica para el rechazo de H_0 , al utilizar \bar{X} como el estadístico:

$$\begin{aligned}\bar{X}_{\text{crit}} &= \mu_{\text{nula}} + \sigma_{\bar{X}}(z_{\text{crit}}) & \sigma_{\bar{X}} &= \frac{\sigma}{\sqrt{N}} = \frac{16}{100} = 1.6 \\ &= 75 + 1.6(-2.33) & z_{\text{crit}} &= -2.33 \\ &= 71.27\end{aligned}$$

PASO 2: Suponiendo que el efecto real hipotético de la variable independiente es verdadero, determinamos la probabilidad de obtener una media muestral que se encuentre en la región crítica para el rechazo de H_0 .

$$\begin{aligned}z_{\text{obt}} &= \frac{\bar{X}_{\text{obt}} - \mu_{\text{real}}}{\sigma_{\bar{X}}} & \sigma_{\bar{X}} &= \frac{\sigma}{\sqrt{N}} = \frac{16}{100} = 1.6 \\ &= \frac{71.27 - 72}{1.6} \\ &= -0.46\end{aligned}$$

Según la tabla A.

$$p(\bar{X}_{\text{obt}} \leq 71.27) = 0.3228$$

Así pues,

$$\text{Potencia} = 0.3228$$

$$\text{Beta} = 1 - \text{potencia} = 0.6772$$

En consecuencia, cuando se restringió alfa para hacerla más rigurosa (al cambiarla de 0.05_{cola} a 0.01_{cola}), la potencia ha disminuido de 0.5910 a 0.3228. Beta ha aumentado de 0.4090 a 0.6772. Esto demuestra que existe una relación directa entre alfa y la potencia, y una relación inversa entre alfa y beta. La figura 12.14 muestra las relaciones que intervienen en este problema.

Relación entre la magnitud del efecto real y la potencia A continuación examinaremos la relación que existe entre la magnitud del efecto real y la potencia. Con este propósito, calcularemos la potencia para detectar una deficiencia media de 5 puntos en la habilidad de la población para la lectura, con respecto a las normas nacionales. Supondremos que $N = 100$ y $\alpha = 0.05_{\text{cola}}$. La figura 12.15 muestra las relaciones que intervienen en este problema.

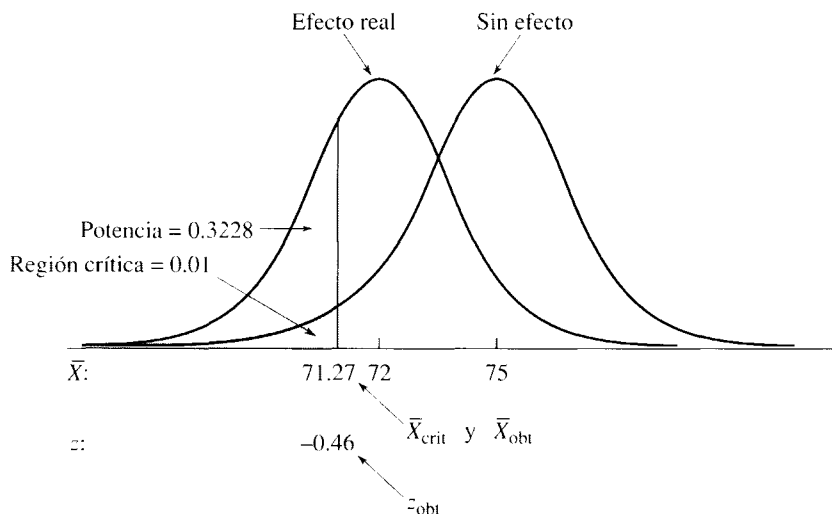


figura 12.14 Potencia para $N = 100$ y $\alpha = 0.01_{\text{cola}}$.

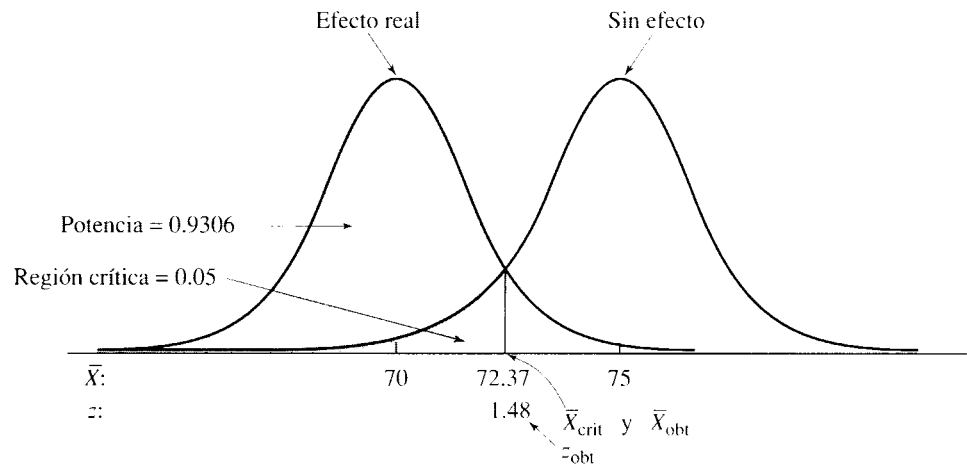


figura 12.15 Potencia para $N = 100$ y $\mu_{\text{real}} = 70$.

SOLUCIÓN

PASO 1: Determine todas las medias muestrales posibles del experimento que permitan el rechazo de H_0 . Con la prueba z, esto significa determinar la región crítica para el rechazo de H_0 , utilizando \bar{X} como el estadístico:

$$\begin{aligned}\bar{X}_{\text{crit}} &= \mu_{\text{nula}} + \sigma_{\bar{X}}(z_{\text{crit}}) & \sigma_{\bar{X}} &= \frac{\sigma}{\sqrt{N}} = \frac{16}{100} = 1.6 \\ &= 75 + 1.6(-1.645) \\ &= 72.37\end{aligned}$$

PASO 2: Suponiendo que el efecto real hipotético de la variable independiente es verdadero, determine la probabilidad de obtener una media muestral que se encuentre en la región crítica para el rechazo de H_0 :

$$\begin{aligned}z_{\text{obt}} &= \frac{\bar{X}_{\text{obt}} - \mu_{\text{real}}}{\sigma_{\bar{X}}} & \sigma_{\bar{X}} &= \frac{\sigma}{\sqrt{N}} = \frac{16}{100} = 1.6 \\ &= \frac{72.37 - 70}{1.6} \\ &= 1.48\end{aligned}$$

Según la tabla A,

$$p(\bar{X}_{\text{obt}} \leq 72.37) = 0.5000 + 0.4306 = 0.9306$$

Así pues,

$$\text{Potencia} = 0.9306$$

$$\text{Beta} = 1 - \text{potencia} = 1 - 0.9306 = 0.0694$$

Vemos que al aumentar la magnitud del efecto real de 3 a 5 puntos de la media, la potencia se ha incrementado de 0.5910 a 0.9306. Beta ha disminuido de 0.4090 a 0.0694. Esto demuestra que existe una relación directa entre la magnitud del efecto real y la potencia necesaria para detectarlo.

RESUMEN

En este capítulo analizamos los temas referentes a la distribución muestral de un estadístico, la forma de generar distribuciones muestrales cuando se parte de un enfoque de muestreo empírico, la distribución muestral de la media y el análisis de experimentos con una sola muestra por medio de la prueba z . Señalamos que el procedimiento para el análisis de los datos, en la mayor parte de los experimentos para la prueba de hipótesis, consiste en calcular el estadístico adecuado y, a continuación, evaluar ese estadístico con base en su distribución muestral. La distribución muestral de un estadístico proporciona todos los valores que éste puede asumir, junto con la probabilidad de obtener cada valor si se realiza un muestreo aleatorio en la población de la hipótesis nula. La distribución muestral se puede generar de manera teórica, por medio del teorema del límite central, o bien en forma empírica: 1) determinando todas las distintas muestras posibles de tamaño N que se puedan formar a partir de la población de datos en bruto, 2) calculando el estadístico para cada una de esas muestras y 3) calculando la probabilidad de obtener cada uno de los valores del estadístico si se realiza un muestreo aleatorio a partir de la población de la hipótesis nula.

La distribución muestral de la media es una distribución de valores de medias muestrales que tiene una media ($\mu_{\bar{x}}$) igual a μ y una desviación estándar ($\sigma_{\bar{x}}$) igual a σ/\sqrt{N} . Presenta una distribución normal si la pobla-

ción de los datos en bruto tiene una distribución normal o si $N \geq 30$, suponiendo que la población de datos en bruto no sea radicalmente distinta de una normal. La prueba z es apropiada para el análisis de experimentos con una sola muestra, cuando μ y σ son conocidas y la media muestral se utiliza como el estadístico básico. Al aplicar esta prueba, se calcula z_{obt} y luego se evalúa para determinar si queda en la región crítica para rechazar la hipótesis nula. Para que se pueda utilizar la prueba z , la distribución muestral de la media debe mostrar una distribución normal. Esto, a su vez, requiere que la población de la hipótesis nula se distribuya de manera normal o que $N \geq 30$.

Por último, examinamos la potencia en conjunción con la prueba z . La potencia es la probabilidad de rechazar H_0 cuando la variable independiente produce efectivamente un efecto real. Para realizar el cálculo de la potencia, aplicamos un procedimiento que incluye dos pasos: la determinación de las medias muestrales posibles que permitan rechazar H_0 , y el cálculo de la probabilidad de obtener cualquiera de esas medias muestrales, suponiendo que el efecto real hipotético de la variable independiente sea verdadero. La potencia varía en proporción directa con respecto a N , alfa, y la magnitud del efecto real de la variable independiente. La potencia varía en proporción inversa con respecto a beta.

TÉRMINOS IMPORTANTES

Distribución muestral de la media (página 268)	Media de la distribución muestral de la media (página 270)	Valor(es) crítico(s) de un estadístico (página 277)
Distribución muestral de un estadístico (página 264)	Población de la hipótesis nula (página 265)	Valor(es) crítico(s) de \bar{X} (página 284)
Error estándar de la media (página 270)	Potencia (página 282)	Valor(es) crítico(s) de z (página 278)
	Región crítica (página 277)	μ_{nula} (página 284)
		σ_{real} (página 285)

PREGUNTAS Y PROBLEMAS

- Defina usted cada uno de los conceptos de la sección "términos importantes".
- ¿Por qué es importante la distribución muestral de un estadístico cuando se realiza una prueba de hipótesis? Explíquelo en unas cuantas líneas.
- Cómo se generan las distribuciones muestrales por medio del enfoque de muestreo empírico?
- ¿Cuáles son los dos pasos básicos que se utilizan para el análisis de los datos?
- ¿Cuáles son los supuestos para la aplicación de la prueba z ?
- ¿Cuáles son las características de la distribución muestral de la media?
- Explique por qué la desviación estándar de la distribución muestral de la media se conoce algunas veces como el "error estándar de la media".
- ¿En qué difiere cada uno de los siguientes conceptos?
 - s y $s_{\bar{x}}$
 - s^2 y σ^2
 - μ y $\mu_{\bar{x}}$
 - σ y $\sigma_{\bar{x}}$
- Explique por qué $\sigma_{\bar{x}}$ debe variar en proporción directa con respecto a σ y en proporción inversa respecto a N .
- ¿Por qué $\mu_{\bar{x}} = \mu$?

11. ¿La forma de la distribución muestral de la media siempre es igual a la forma de la población de la hipótesis nula? Explique.
 12. Al utilizar la prueba z , ¿por qué es importante que la distribución muestral de la media esté distribuida de manera normal?
 13. Si se cumplen los supuestos de la prueba z , ¿cuáles son las características de la distribución muestral de z ?
 14. Defina potencia, tanto conceptual como matemáticamente.
 15. Explique lo que ocurre con la potencia de la prueba z cuando cada una de las siguientes variables crece.
 - a. N
 - b. Nivel alfa
 - c. Magnitud del efecto real de la variable independiente
 - d. σ
 16. Cuando la N de un experimento aumenta, ¿cómo son afectados los siguientes aspectos?
 - a. Potencia
 - b. Beta
 - c. Alfa
 - d. Magnitud del efecto real
 17. Dado el conjunto de datos poblacionales 3, 4, 5, 6, 7,
 - a. Determine la distribución muestral de la media para muestras de tamaño 2. Suponga que el muestreo es de uno a la vez, con reemplazo.
 - b. Demuestre que $\mu_{\bar{X}} = \mu$.
 - c. Demuestre que $\sigma_{\bar{X}} = \sigma/\sqrt{N}$.
 18. Si una población de datos en bruto tiene una distribución normal, con una media $\mu = 80$ y una desviación estándar $\sigma = 8$, determine los parámetros ($\mu_{\bar{X}}$ y $\sigma_{\bar{X}}$) de la distribución muestral de la media para los siguientes tamaños de muestra:
 - a. $N = 16$
 - b. $N = 35$
 - c. $N = 50$
 - d. Explique lo que ocurre cuando N crece. otra
 19. ¿Es razonable considerar que una muestra de 40 datos con $\bar{X}_{\text{obt}} = 65$ es una muestra aleatoria de una población de datos que presenta una distribución normal con $\mu = 60$ y $\sigma = 10$? Utilice $\alpha = 0.05_{2 \text{ colas}}$ para tomar su decisión. otra
 20. Un conjunto de datos muestrales obtenidos en un experimento tiene una $N = 30$ y una $\bar{X}_{\text{obt}} = 19$.
 - a. ¿Podemos rechazar la hipótesis nula en el sentido de que la muestra es una muestra aleatoria que proviene de una población normal con $\mu = 22$ y $\sigma = 8$? Utilice $\alpha = 0.01_{1 \text{ cola}}$. Suponga que la media muestral está en la dirección correcta.
 - b. ¿Cuál es la potencia del experimento para detectar un efecto real, como que $\mu_{\text{real}} = 20$?
 - c. ¿Cuál es la potencia necesaria para detectar una $\mu_{\text{real}} = 20$ si N se incrementa hasta 100?
 - d. ¿Qué valor debe tener N para lograr una potencia igual a 0.8000 a fin de detectar una $\mu_{\text{real}} = 20$? Utilice el valor más cercano de la tabla para z_{obt} . otra
21. De acuerdo con una técnica que ella misma ha desarrollado, una estudiante cree que puede reducir el tiempo que los esquizofrénicos permanecen recluidos en una institución de salud. Como director de capacitación de una institución cercana, usted ha permitido que esa estudiante pruebe su método con 20 esquizofrénicos, elegidos al azar, de la institución que usted dirige. El tiempo medio de estancia de los esquizofrénicos en esta institución es de 85 semanas, con una desviación estándar de 15 semanas. Los datos tienen una distribución normal. Los resultados del experimento muestran que los pacientes que han sido tratados con la nueva técnica desarrollada por la estudiante tienen una estancia media de 78 semanas, con una desviación estándar de 20 semanas.
 - a. ¿Cuál es la hipótesis alternativa? En este caso, supongamos que resulta apropiada una hipótesis no direccional, pues no hay bases teóricas ni empíricas suficientes que justifiquen una hipótesis direccional.
 - b. ¿Cuál es la hipótesis nula?
 - c. ¿Qué puede usted concluir acerca de la técnica de la estudiante? Utilice $\alpha = 0.05_{2 \text{ colas}}$. clínica, salud
 22. Un profesor ha enseñado estadística durante muchos años. Sus archivos muestran que el promedio general de sus exámenes finales es de 82, con una desviación estándar de 10. El profesor cree que su grupo de este año es superior a los anteriores. El promedio de las calificaciones del examen final de este grupo de 65 estudiantes es de 87. ¿Qué podría usted concluir? Utilice $\alpha = 0.05_{1 \text{ cola}}$. educación
 23. Una ingeniera automotriz cree que su nuevo motor ahorrará mucha gasolina. Numerosas pruebas realizadas en motores de diseño antiguo produjeron un consumo medio de gasolina de 27.5 millas por galón, con una desviación estándar de 5.2. Ahora hace sus pruebas en 15 motores nuevos. El consumo medio de gasolina es de 29.6 millas por galón. ¿Qué conclusión saca usted? Utilice $\alpha = 0.05_{1 \text{ cola}}$. otra

24. Consulte el problema de práctica 12.2 de la página 281 para responder esta pregunta. Suponga que antes de realizar el experimento, el fabricante quiere averiguar cuál es la probabilidad de detectar un incremento promedio real de 2.0 millas por galón con el aditivo, si éste tiene por lo menos esa efectividad.
- Si el fabricante prueba 20 automóviles, ¿cuál es la potencia necesaria para detectar un incremento promedio de 2.0 millas por galón?
 - Si el fabricante incrementa N a 75 automóviles, ¿cuál es la potencia necesaria para detectar un aumento promedio de 2.0 millas por galón?
 - ¿Cuántos automóviles deberá utilizar el fabricante si quiere tener el 99% de posibilidades de detectar un incremento promedio de 2.0 millas por galón? /10
25. Un profesor de educación física considera que el ejercicio físico puede retardar el proceso de envejecimiento. Durante los últimos 10 años, él mismo ha impartido una clase de ejercicios físicos a 14 individuos que actualmente tienen 50 años de edad. Por regla general, cuando una persona envejece, su consumo máximo de oxígeno disminuye. La norma nacional para el consumo máximo de oxígeno por los individuos de 50 años de edad es de 30 mililitros por kilogramo y por minuto, con una desviación estándar de 8.6. La media observada en los 14 individuos de la clase es de 40 mililitros por kilogramo por minuto. ¿Qué conclusión podría usted sacar? Utilice $\alpha = 0.05$ 1 cola, biológica, salud

Si desea encontrar problemas de práctica adicionales totalmente resueltos sobre el material del capítulo 12, busque Chapter 12 en el material adicional del libro, disponible en inglés en: <http://www.thomsonlearning.com.mx>.

Prueba t de Student para muestras simples

Introducción

Comparación de las pruebas z y t

Un experimento: Técnica para incrementar el habla temprana en los niños

La distribución muestral de t

Grados de libertad

Comparación de las distribuciones t y z

Revisión del experimento de la emisión de las primeras palabras

Cálculo de t_{obt} a partir de los datos originales

Condiciones bajo las cuales la prueba t es adecuada

Distribución t : Intervalos de confianza para la media poblacional

Construcción del intervalo de confianza del 95 por ciento

Ejemplo: Estimación de la Imedia del CI de los profesores de una universidad

Ecuaciones generales para cualquier intervalo de confianza

Prueba de la significancia de la r de Pearson

Resumen

Términos importantes

Preguntas y problemas

Notas



INTRODUCCIÓN

En el capítulo 12 analizamos la prueba z y determinamos que resulta apropiada para utilizarse en situaciones en las que se conoce tanto la media como la desviación estándar poblacionales de la hipótesis nula. Sin embargo, estas situaciones son relativamente raras. Es más frecuente encontrar situaciones en las cuales se conoce la media poblacional de la hipótesis nula y la desviación estándar es *desconocida*. En estos casos no se puede emplear la prueba z . En lugar de esta prueba, aplicamos otra que se llama la *prueba t de*

tabla 13.1 Comparación de las ecuaciones de las pruebas *z* y *t*

Prueba <i>z</i>	Prueba <i>t</i>
$\bar{z}_{\text{obt}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{N}}$ $= \frac{\bar{X}_{\text{obt}} - \mu}{\sigma_{\bar{X}}}$	$t_{\text{obt}} = \frac{\bar{X} - \mu}{s/\sqrt{N}}$ $= \frac{\bar{X}_{\text{obt}} - \mu}{s_{\bar{X}}}$
<p>donde s = estimación de σ $s_{\bar{X}}$ = estimación de $\sigma_{\bar{X}}$</p>	

Student. Esta última es muy similar a la prueba *z*. Fue desarrollada por W. S. Gosset, quien escribía bajo el seudónimo de “Student” (estudiante). La *t* de Student es una prueba práctica y muy poderosa que se usa ampliamente en las ciencias del comportamiento. En este capítulo vamos a estudiar la prueba *t* junto con los experimentos en los que interviene una sola muestra. En el capítulo 14 analizaremos la prueba *t* aplicada a experimentos con dos muestras o condiciones.



COMPARACIÓN DE LAS PRUEBAS *z* Y *t*

Las pruebas *z* y *t* que se aplican en experimentos con una sola muestra son muy parecidas entre sí. Las ecuaciones que corresponden a cada una de ellas se presentan en la tabla 13.1.

Al comparar estas ecuaciones, vemos que la única diferencia entre ellas es que la prueba *z* utiliza la desviación estándar poblacional de la hipótesis nula (σ), mientras que la prueba *t* emplea la desviación estándar de la muestra (s). Si σ es desconocida, podemos estimarla a partir de la estimación que s proporciona, y el estadístico resultante se llama *t*. De este modo, el denominador de la prueba *t* es s/\sqrt{N} en lugar de σ/\sqrt{N} . El símbolo $s_{\bar{X}}$ reemplaza a $\sigma_{\bar{X}}$ donde:

$$s_{\bar{X}} = \frac{s}{\sqrt{N}} \quad \text{error estándar estimado de la media}$$

Ahora ya estamos listos para estudiar un experimento donde se utiliza la prueba *t* para analizar los datos.

un experimento

Técnica para incrementar el habla temprana en los niños

Supongamos que usted tiene una técnica que, según cree, puede modificar la edad en la que los niños comienzan a hablar. En su localidad, el promedio de edad en que un niño pronuncia su primera palabra es de 13.0 meses. No se conoce la desviación estándar poblacional. Usted aplica su técnica a una muestra aleatoria de 15 niños. Los resultados obtenidos demuestran que la media de la edad en que los niños de la muestra pronunciaron la primera palabra fue de 11.0 meses, con una desviación estándar de 3.34.

1. ¿Cuál es la hipótesis alternativa no direccional?
2. ¿Cuál es la hipótesis nula?
3. ¿Dio resultado la técnica? Utilice $\alpha = 0.05_{2 \text{ colas}}$.

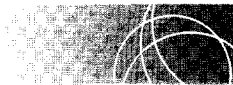
SOLUCIÓN

1. Hipótesis alternativa: La técnica afecta la edad en la que los niños comienzan a hablar. Por lo tanto, la muestra con $\bar{X}_{\text{obt}} = 11.0$ es una muestra aleatoria, extraída de una población donde $\mu \neq 13.0$.
2. Hipótesis nula: H_0 : La muestra con $\bar{X}_{\text{obt}} = 11.0$ es una muestra aleatoria, extraída de una población en la cual $\mu = 13.0$.
3. Conclusión utilizando $\alpha = 0.05$ colas:

PASO 1: Cálculo de los estadísticos apropiados. Como σ es desconocida, resulta imposible determinar z_{obt} . Sin embargo, conocemos s y por eso podemos calcular t_{obt} . Así pues,

$$\begin{aligned} t_{\text{obt}} &= \frac{\bar{X}_{\text{obt}} - \mu}{s/\sqrt{N}} \\ &= \frac{11.0 - 13.0}{3.34/\sqrt{15}} \\ &= \frac{-2}{0.862} \\ &= -2.32 \end{aligned}$$

El siguiente paso consiste de ordinario en evaluar t_{obt} utilizando la distribución muestral de t . Sin embargo, como aún no estamos familiarizados con esta distribución, la analizaremos antes de continuar con la evaluación.

**LA DISTRIBUCIÓN MUESTRAL DE t**

Empleamos la definición de la distribución muestral desarrollada en el capítulo 12 para observar lo siguiente:

definición ■ *La distribución muestral de t es una distribución de probabilidad de los valores t que se pueden presentar si se consideran todas las posibles muestras de un tamaño fijo N , extraídas de la población de la hipótesis nula. Esta distribución proporciona: 1) todos los diferentes valores t para muestras de tamaño N y 2) la probabilidad de obtener cada uno de los valores si la muestra extraída de la población de la hipótesis nula es aleatoria.*

Como en el caso de la distribución muestral de la media, podemos determinar la distribución muestral de t de manera teórica o empírica. De nuevo, por razones pedagógicas, preferimos el punto de vista empírico. Deducimos la distribución muestral de t en forma empírica, considerando una población específica de datos en bruto, extrayendo todas las diferentes muestras posibles de un tamaño N fijo y luego calculando el valor t para cada una de estas muestras. Una vez obtenidos todos los valores t , es sencillo calcular la probabilidad de cada uno de ellos, bajo el supuesto de un muestreo aleatorio de la población. Modificando N y los datos poblacionales, podemos determinar las distribuciones muestrales para diversas poblaciones y tamaños de muestra. Ya sea en forma empírica o teórica, se puede ver que si la población de la hipótesis nula tiene forma normal, o bien, si $N \geq 30$, la distribución t es similar a la z , excepto que existe una familia de curvas t que varían junto con el tamaño de la muestra. Usted recordará que la distribución z sólo tiene una curva para todos los tamaños de muestra (los valores representados en la tabla A

del apéndice D). Por otro lado, la distribución *t*, al igual que la distribución muestral de la media, tiene muchas curvas que dependen del tamaño de la muestra. Como tratamos de estimar σ utilizando s en la ecuación *t* y en vista de que el tamaño de la muestra influye en el grado de precisión de la estimación, tiene sentido que exista una distribución muestral *t* diferente para cada tamaño de la muestra.

Grados de libertad

Aunque la distribución *t* varía junto con el tamaño de la muestra, Gosset encontró que cambia de manera *única* con respecto a los *grados de libertad* asociados con *t*, en lugar de variar solamente con el tamaño de la muestra. La razón de esto no será evidente sino hasta que revisemos el capítulo 14. Por ahora, sólo presentaremos el concepto de grados de libertad.

definición ■ *Los grados de libertad (gl) para cualquier estadístico es el número de datos que pueden variar libremente al calcular dicho estadístico.*

Por ejemplo, la media tiene asociados N grados de libertad. ¿Cómo lo sabemos? Para cualquier conjunto de datos, N está dado. Si existen tres datos y conocemos los dos primeros, el último puede asumir cualquier valor; no tiene restricciones. No podemos decir cuál es su valor si sólo conocemos los otros dos datos. Lo mismo ocurre con los dos primeros datos. Así, los tres datos pueden variar libremente al calcular la media, por lo cual existen N grados de libertad.

Contrasta esto con el cálculo de la desviación estándar:

$$s = \sqrt{\frac{\sum (X - \bar{X})^2}{N - 1}}$$

Como la suma de las desviaciones con respecto a la media debe ser igual a cero, sólo $N - 1$ de las desviaciones son libres de asumir cualquier valor. Así, existen $N - 1$ grados de libertad asociados con s . ¿Por qué? Considere los datos en bruto 4, 8 y 12. La media es 8. La tabla 13.2 muestra lo que sucede cuando calculamos s .

Como la media es 8, la desviación del dato en bruto 4 es -4 , y para el dato en bruto 8 la desviación es 0. Como $\sum (X - \bar{X}) = 0$, la última desviación queda determinada por las otras desviaciones. Tiene que ser $+4$ (véase la “?” en la tabla 13.2). No puede asumir ningún otro valor pues está fijo en $+4$ a causa de los otros dos datos de desviación. Por lo tanto, solamente dos de las tres desviaciones pueden cambiar con libertad. Independientemente del valor que asuman, el tercer dato estará fijo. Al calcular s , solamente $N - 1$ desviaciones pueden cambiar libremente. Así, existen $N - 1$ grados de libertad asociados con la desviación estándar.

tabla 13.2 Número de datos de desviación que pueden variar

X	\bar{X}	$X - \bar{X}$
4	8	-4
8	8	0
12	8	?

Al calcular t para una muestra, debemos determinar primero el valor de s . Como al calcular s perdemos un grado de libertad, existen $N - 1$ grados de libertad asociados con t . De esta manera, para la prueba t ,

$$gl = N - 1 \quad \text{grados de libertad para la prueba } t \text{ (una muestra)}$$



COMPARACIÓN DE LAS DISTRIBUCIONES t Y z

La figura 13.1 presenta la distribución t para diversos grados de libertad. La distribución t es simétrica con respecto a cero y, al aumentar los grados de libertad, se aproxima a la distribución normal z . Observe la rapidez con la cual se acerca a la curva normal. Aun con un número de grados de libertad tan pequeño como 20, la distribución t se aproxima cada vez más a la curva normal. Desde el punto de vista teórico, cuando el número de grados de libertad es infinito, $gl = \infty$,⁶ la distribución t es idéntica a la distribución z . Esto resulta lógico, pues al aumentar los grados de libertad, el tamaño de la muestra aumenta y la estimación s se acerca a σ . Para un número cualquiera de grados de libertad no infinito, la distribución t tiene más valores extremos que la distribución z , ya que existe más variabilidad en t debido a que utilizamos s para estimar σ . Otra forma de decir esto es que las colas de la distribución t son más elevadas con respecto a la distribución z . Así, para cierto nivel alfa, el valor crítico de t es mayor que el de z , lo que hace a la prueba t menos sensible que la prueba z . Es decir, para cualquier nivel alfa, t_{obt} debe ser mayor que

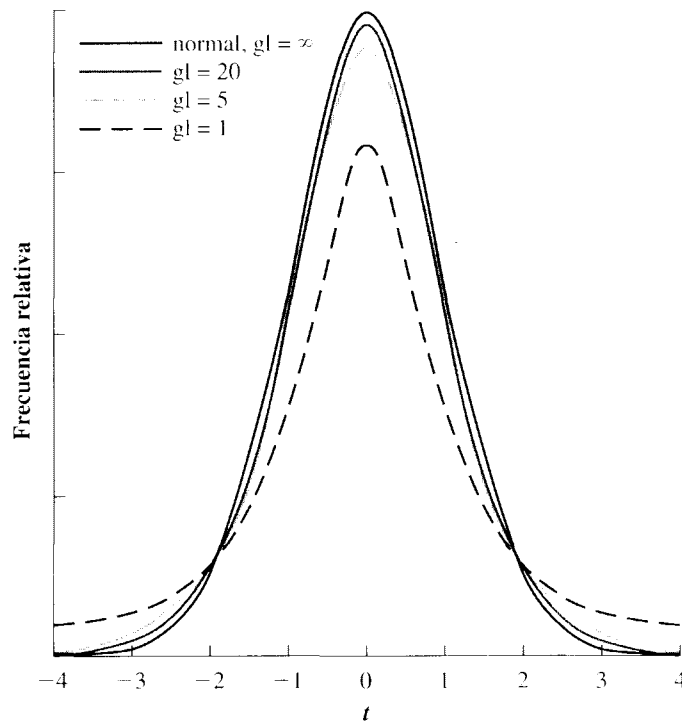


figura 13.1 Distribución t para diversos grados de libertad.

⁶ Cuando el número de grados de libertad se acerca a infinito, la distribución t se aproxima a la curva normal.

tabla 13.3 Valores críticos de *z* y *t* para los niveles alfa 0.05 y 0.01, de una cola.

<i>gl</i>	$z_{0.05}$	$t_{0.05}$	$z_{0.01}$	$t_{0.01}$
5	1.645	2.015	2.326	3.365
30	1.645	1.697	2.326	2.457
60	1.645	1.671	2.326	2.390
∞	1.645	1.645	2.326	2.326

z_{obt} a fin de rechazar la hipótesis nula. La tabla 13.3 muestra los valores críticos de *z* y *t* para los niveles alfa de 0.05 y 0.01. Al aumentar los grados de libertad, el valor crítico de *t* tiende al de *z*. Por supuesto, el valor crítico de *z* no cambia con el tamaño de la muestra. Los valores críticos de *t* para varios niveles de alfa y grados de libertad aparecen en la tabla D del apéndice D. Estos valores han sido obtenidos a partir de la distribución muestral de *t* para cada grado de libertad. Emplearemos la tabla para evaluar t_{obt} en cualquier experimento. Ahora ya estamos listos para regresar al ejemplo ilustrativo.



REVISIÓN DEL EXPERIMENTO DE LA EMISIÓN DE LAS PRIMERAS PALABRAS

Usted está investigando una técnica que, supuestamente, modifica la edad en la que los niños empiezan a hablar $\mu = 13.0$ meses; σ es desconocida; la muestra de 15 niños, que utilizan su técnica, tiene una media de 11.0 meses para las primeras palabras pronunciadas y una desviación estándar de 3.34.

1. ¿Cuál es la hipótesis alternativa no direccional?
2. ¿Cuál es la hipótesis nula?
3. ¿Funcionó la técnica? Utilice $\alpha = 0.05_{2 \text{ colas}}$.

SOLUCIÓN

1. Hipótesis alternativa: La técnica modifica la edad en la que los niños comienzan a hablar. Por lo tanto, la muestra con $\bar{X}_{\text{obt}} = 11.0$ es una muestra aleatoria, extraída de la población donde $\mu \neq 13.0$.
2. Hipótesis nula: H_0 : La muestra con $\bar{X}_{\text{obt}} = 11.0$ es una muestra aleatoria, extraída de la población con $\mu = 13.0$.
3. Conclusión al utilizar $\alpha = 0.05_{2 \text{ colas}}$:

PASO 1: Cálculo de los estadísticos apropiados. Como éste es un experimento con una sola muestra y σ es desconocida, el uso de t_{obt} resulta apropiado:

$$\begin{aligned}
 t_{\text{obt}} &= \frac{\bar{X}_{\text{obt}} - \mu}{s/\sqrt{N}} \\
 &= \frac{11.0 - 13.0}{3.34/\sqrt{15}} \\
 &= -2.32
 \end{aligned}$$

PASO 1: Cálculo de los estadísticos apropiados. Como α es desconocida, t_{obt} resulta apropiada.

$$t_{\text{obt}} = \frac{\bar{X}_{\text{obt}} - \mu}{s/\sqrt{N}} = \frac{11.0 - 13.00}{3.34/\sqrt{15}} = -2.32$$

PASO 2: Evaluación de los estadísticos. Si $|t_{\text{obt}}| \geq |t_{\text{crit}}|$, rechace H_0 . Como $\alpha = 0.05_{2\text{colas}}$ y $gl = N - 1 = 15 - 1 = 14$, según la tabla D,

$$t_{\text{crit}} = \pm 2.145$$

Como $|t_{\text{obt}}| > 2.145$, queda dentro de la región crítica. Por lo tanto, rechazamos H_0

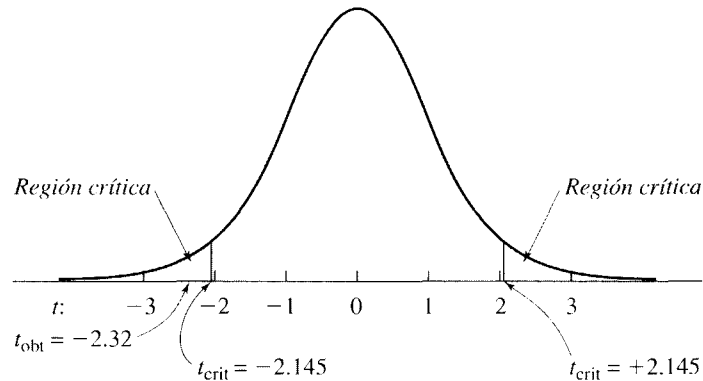


Figura 13.2 Solución al experimento de la primera palabra pronunciada utilizando la prueba t de Student.

PASO 2: Evaluación de los estadísticos con base en su distribución muestral. De la misma manera, como en la prueba z , si

$$|t_{\text{obt}}| \geq |t_{\text{crit}}|$$

entonces queda dentro de la región crítica para el rechazo de la hipótesis nula. Localizamos t_{crit} en la tabla D, bajo el nivel alfa y los grados de libertad apropiados. Para este ejemplo, con $\alpha = 0.05_{2\text{colas}}$ y $gl = N - 1 = 15 - 1 = 14$, según la tabla D,

$$t_{\text{crit}} = \pm 2.145$$

Puesto que $|t_{\text{obt}}| > 2.145$, rechazamos H_0 y concluimos que la técnica empleada influye en la edad en la cual comienzan a hablar los niños de la localidad escogida. La solución se presenta en la figura 13.2.



CÁLCULO DE t_{obt} A PARTIR DE LOS DATOS ORIGINALES

Si en una situación dada los datos originales están disponibles, entonces podemos calcular t de manera directa sin tener que calcular primero s . La ecuación apropiada se presenta aquí:⁸

⁸ La deducción se presenta en la nota 13.1.

$$t_{\text{obt}} = \frac{\bar{X}_{\text{obt}} - \mu}{\sqrt{\frac{SC}{N(N-1)}}} \quad \text{ecuación para calcular } t_{\text{obt}} \text{ a partir de los datos en bruto}$$

$$\text{donde } SC = \sum (X - \bar{X})^2 = \sum X^2 - \frac{(\sum X)^2}{N}$$

ejemplo

Supongamos que los datos originales del problema anterior son los que aparecen en la tabla 13.4. Ahora calcularemos t_{obt} directamente a partir de esos datos en bruto.

Si desea encontrar una solución basada en la ecuación conceptual, busque Chapter 13 en el material de apoyo de este libro que está disponible en: www.thomsonlearning.com.mx

SOLUCIÓN

$$\begin{aligned} t_{\text{obt}} &= \frac{\bar{X}_{\text{obt}} - \mu}{\sqrt{\frac{SC}{N(N-1)}}} & SC &= \sum X^2 - \frac{(\sum X)^2}{N} \\ &= \frac{11.0 - 13.0}{\sqrt{\frac{156}{15(14)}}} & &= 1971 - \frac{(165)^2}{15} \\ &= \frac{-2}{0.862} & &= 156 \\ & & &= -2.32 \end{aligned}$$

tabla 13.4 Datos en bruto para el ejemplo de la emisión de las primeras palabras

Edad (meses) <i>X</i>	Cálculo de <i>X</i> ²
8	64
9	81
10	100
15	225
18	324
17	289
12	144
11	121
7	49
8	64
10	100
11	121
8	64
9	81
<u>12</u>	<u>144</u>
165	1971
<i>N</i> = 15	$\bar{X}_{\text{obt}} = \frac{165}{15} = 11.0$

- : Éste es el mismo valor que obtuvimos anteriormente. Observe usted que es correcto calcular
 : primero s y después utilizar la ecuación original t_{obt} . Sin embargo, la respuesta está más expues-
 : ta a los errores de redondeo.
 : Intentemos resolver otro problema.



Una investigadora cree que la estatura de las mujeres ha aumentado en años recientes. Ella sabe que hace 10 años, en la ciudad donde vive, el promedio de estatura de una mujer adulta joven era de 63 pulgadas. No se conoce la desviación estándar. La investigadora toma una muestra aleatoria de ocho mujeres jóvenes adultas que residen en dicha ciudad y mide sus estaturas. Así obtiene los siguientes datos:

Estatura (en pulgadas) X	Cálculo de X^2
64	4096
66	4356
68	4624
60	3600
62	3844
65	4225
66	4356
63	3969
514	33 070
$N = 8$	$\bar{X}_{\text{obt}} = \frac{514}{8} = 64.25$

- ¿Cuál es la hipótesis alternativa? Al evaluar este experimento, suponga que una hipótesis no direccional resulta apropiada porque no existen suficientes bases teóricas y empíricas para garantizar una hipótesis direccional.
- ¿Cuál es la hipótesis nula?
- ¿Cuál es la conclusión de usted? Use $\alpha = 0.01_{2 \text{ colas}}$.

SOLUCIÓN

- Hipótesis alternativa no direccional: En años recientes, la estatura de las mujeres ha ido cambiando. Por lo tanto, la muestra con $\bar{X}_{\text{obt}} = 64.25$ es una muestra aleatoria extraída de una población donde $\mu \neq 63$.
- Hipótesis nula: La hipótesis nula asegura que es razonable considerar la muestra con $\bar{X}_{\text{obt}} = 64.25$ como una muestra aleatoria extraída de una población con $\mu = 63$.
- Conclusión, utilizando $\alpha = 0.01_{2 \text{ colas}}$:

(continúa)

Si desea encontrar una solución basada en la ecuación conceptual, busque Chapter 13 en el material de apoyo de este libro que está disponible en: www.thomsonlearning.com.mx

PASO 1: Cálculo de los estadísticos apropiados. Los datos fueron dados anteriormente. Como σ es desconocida, t_{obt} resulta apropiada. Existen dos formas de determinar t_{obt} : 1) calcular primero s y después t_{obt} y 2) evaluar t_{obt} directamente, a partir de los datos en bruto. Aquí mostramos ambos métodos:

s primero y después t_{obt} :

$$s = \sqrt{\frac{SC}{N-1}} = \sqrt{\frac{45.5}{7}} \quad SC = \sum X^2 - \frac{(\sum X)^2}{N} = 33\,070 - \frac{(514)^2}{8}$$

$$= \sqrt{6.5} = 2.550 \quad = 33\,070 - 33\,024.5 = 45.5$$

$$t_{\text{obt}} = \frac{\bar{X}_{\text{obt}} - \mu}{s/\sqrt{N}} = \frac{64.25 - 63}{2.550/\sqrt{8}}$$

$$= \frac{1.25}{0.902} = 1.39$$

directamente de los datos en bruto:

$$t_{\text{obt}} = \frac{\bar{X}_{\text{obt}} - \mu}{\sqrt{\frac{SC}{N(N-1)}}} = \frac{64.25 - 63}{\sqrt{\frac{45.5}{8(7)}}} \quad SC = \sum X^2 - \frac{(\sum X)^2}{N}$$


$$= \frac{1.25}{\sqrt{0.812}} = 1.39 \quad = 33\,070 - \frac{(514)^2}{8}$$

$$= 33\,070 - 33\,024.5 = 45.5$$

PASO 2: Evaluación de los estadísticos. Si $|t_{\text{obt}}| \geq |t_{\text{crit}}|$, rechazamos H_0 . Si no es así, conservamos H_0 . Con $\alpha = 0.01$ 2 colas y $gl = N - 1 = 8 - 1 = 7$, según la tabla D,

$$t_{\text{crit}} = \pm 3.499$$

Puesto que $|t_{\text{obt}}| < 3.499$, no queda dentro de la región crítica. Por lo tanto conservamos H_0 . No podemos concluir que la estatura de las mujeres adultas jóvenes de la ciudad en cuestión se haya modificado en los últimos años.



Un amigo suyo ha estado “jugando” en el mercado accionario. Él afirma que ha dedicado años enteros a hacer una investigación en este terreno y que ideó un exitoso método empírico para invertir. Como usted no se niega a ser un poco más rico, está pensando en darle algún dinero a este amigo para que lo invierta por usted. Sin embargo, antes de esto, decide evaluar su método. Usted lo convence de realizar una “prueba simulada” en la cual él usará su método y le recomendará qué valores debe adquirir, pero usted, en vez de comprar y vender acciones, sólo comprobará el rendimiento de las acciones recomendadas para ver si el método funciona. Duran-

te ese periodo de prueba, las acciones recomendadas mostraron los siguientes cambios de precios (un signo positivo significa un incremento de precio, y uno negativo indica una disminución):

Acción	Cambio en precio (\$) X	Cálculo de X^2
A	+4.52	20.430
B	+5.15	26.522
C	+3.28	10.758
D	+4.75	22.562
E	+6.03	36.361
F	+4.09	16.728
G	+3.82	14.592
	<u>31.64</u>	<u>147.953</u>
$N = 7$		$\bar{X}_{\text{obt}} = 4.52$

Durante el mismo lapso, el cambio promedio de precios de todo el mercado de acciones fue de +\$3.25. Como usted desea saber si el método es mejor o peor que el azar, opta por usar una evaluación de dos colas.

- ¿Cuál es la hipótesis alternativa no direccional?
- ¿Cuál es la hipótesis nula?
- ¿Cuál es su conclusión? Use $\alpha = 0.05_{2 \text{ colas}}$.

SOLUCIÓN

- Hipótesis alternativa no direccional: El método de su amigo produce una selección de acciones cuyo cambio de precio es diferente del que se esperaría de un muestreo aleatorio del mercado bursátil en general. Por lo tanto, la muestra con $\bar{X}_{\text{obt}} = \$4.52$ no puede ser considerada como una muestra aleatoria, extraída de una población donde $\mu = \$3.25$.
- Hipótesis nula: El método de su amigo produce una selección de acciones cuyo cambio de precio no es diferente del que se esperaría de un muestreo aleatorio del mercado bursátil en general. Por lo tanto, la muestra con $\bar{X}_{\text{obt}} = \$4.52$ puede ser considerada como una muestra aleatoria, extraída de una población donde $\mu = \$3.25$.
- Conclusión, utilizando una $\alpha = 0.05_{2 \text{ colas}}$:

PASO 1: Cálculo de los estadísticos apropiados. Los datos fueron presentados en la tabla anterior. Como no conocemos σ , resulta apropiada t_{obt}

$$\begin{aligned}
 t_{\text{obt}} &= \frac{\bar{X}_{\text{obt}} - \mu}{\sqrt{\frac{SC}{N(N-1)}}} \\
 &= \frac{4.52 - 3.25}{\sqrt{\frac{4.940}{7(6)}}} \\
 &= \frac{1.27}{0.343} = 3.70
 \end{aligned}$$

$$\begin{aligned}
 SC &= \sum X^2 - \frac{(\sum X)^2}{N} \\
 &= 147.953 - \frac{(31.64)^2}{7} \\
 &= 4.940
 \end{aligned}$$

Si desea encontrar una solución basada en la ecuación conceptual, busque Chapter 13 en el material de apoyo de este libro que está disponible en: www.thomsonlearning.com.mx

(continúa)

PASO 2: Evaluación de los estadísticos. Con $\alpha = 0.05_{2 \text{ colas}}$ y el $gl = N - 1 = 7 - 1 = 6$, según la tabla D,

$$t_{\text{crit}} = \pm 2.447$$

En virtud de que $|t_{\text{obt}}| > 2.447$, rechazamos H_0 . Tal parece que este amigo suyo es un ganador. ¡Su método parece funcionar! Sin embargo, antes que usted invierta mucho, le sugerimos que realice el experimento por lo menos una vez más para asegurarse de no cometer un error de Tipo I. Recuerde que es esencial repetir el experimento antes de aceptar un resultado como real. Es mejor que peque de precavido.



CONDICIONES BAJO LAS CUALES LA PRUEBA t ES ADECUADA

La prueba t (para una muestra) resulta apropiada cuando el experimento tiene una sola muestra, se conoce μ , se desconoce σ y se utiliza la media de la muestra como estadístico básico. Al igual que la prueba z , la prueba t requiere que la distribución muestral de \bar{X} sea normal. Para que la distribución muestral de \bar{X} sea normal, N debe ser igual o mayor que 30 o la población de datos en bruto debe ser normal.*



DISTRIBUCIÓN t : INTERVALOS DE CONFIANZA PARA LA MEDIA POBLACIONAL

A veces se desea conocer el valor de la media de una población. Como no es muy económico medir a cada uno de los miembros de dicha población, se extrae de ella una muestra aleatoria y la media muestral se utiliza como estimación de la media poblacional. Por ejemplo, supongamos que la administradora de una universidad quiere conocer el CI promedio de los profesores que laboran en ese centro educativo. La administradora extrae una muestra aleatoria y $\bar{X} = 135$. La estimación sería entonces de 135. Este número se conoce como una *estimación puntual*, ya que se utiliza solamente un valor para dicha estimación. Sin embargo, si preguntamos a la administradora su opinión con respecto a que la media de la población fue exactamente igual a 135, seguramente responderá que "no" lo creería. Bien, pero entonces, ¿qué tan cerca está 135 de la media poblacional?

La forma usual de responder esta pregunta consiste en seleccionar un rango de valores que con certeza incluyan a la media de la población. Ésta es una estimación de intervalo. Por ejemplo, la administradora podría tener cierta confianza en que la media poblacional se encuentre en el rango 130-140. Ciertamente, tendría más confianza en el rango de 130-140 que en el valor único de 135. ¿Y qué pasaría con el rango 110-160? Está claro que tendría más confianza en este rango que en el de 130-140. Por lo tanto, mientras más amplio sea el rango, mayor confianza se tendrá de que contiene a la media de la población.

*Muchos autores se limitan a utilizar la prueba t con datos que poseen escalas de intervalo o de razón. Consulte por favor la nota al pie de la página 31 del capítulo 2, para ver referencias en las que se comenta este punto.

- definiciones** ■ *Un intervalo de confianza es un rango de valores que probablemente contiene el valor poblacional.*
- *Los límites de confiabilidad son los valores que establecen las fronteras del intervalo de confianza.*

Es posible cuantificar el grado de confianza que se tiene de que el intervalo en cuestión contenga la media poblacional. De hecho, podemos construir intervalos de confianza con grados específicos; por ejemplo, un intervalo de confianza de 95 por ciento:

El intervalo de confianza del 95% es aquel en el que la probabilidad de que dicho intervalo contenga el valor poblacional es de 0.95.

Aunque podríamos construir muchos intervalos diferentes, los más utilizados en la práctica son los de 95 y 99 por ciento. Consideraremos ahora cómo podemos construir dichos intervalos.

Construcción del intervalo de confianza del 95 por ciento

Suponga que obtenemos un conjunto de N datos mediante un muestreo aleatorio en una población de datos en bruto, con media $= \mu$ y que hemos calculado t_{obt} . Si se cumplen los supuestos de t , vemos que la probabilidad de que la siguiente desigualdad sea cierta es de 0.95:

$$-t_{0,025} \leq t_{\text{obt}} \leq +t_{0,025}$$

$t_{0,025}$ es el valor crítico de t para $\alpha = 0.025_{1 \text{ cola}}$ y $gl = N - 1$. Esta desigualdad indica que si obtenemos N datos mediante un muestreo aleatorio en una población de datos en bruto con media μ y calculamos t_{obt} , se encuentre entre t_{obt} la probabilidad de que $-t_{0,025}$ y $+t_{0,025}$ es de 0.95. La verdad de esta afirmación puede entenderse mejor si nos referimos a la figura 13.3. Ésta muestra la distribución t para $N - 1$ grados de libertad. Hemos localizado $+t_{0,025}$ y $-t_{0,025}$ en la distribución. Recuerde que éstos son los valores críticos de t para $\alpha = 0.025_{1 \text{ cola}}$. Por definición, el 2.5% de los valores de t debe quedar debajo de cada cola, y el 95% de los valores debe estar entre $-t_{0,025}$ y $+t_{0,025}$. De lo anterior se desprende que la probabilidad de que t_{obt} se encuentre entre $-t_{0,025}$ y $+t_{0,025}$.

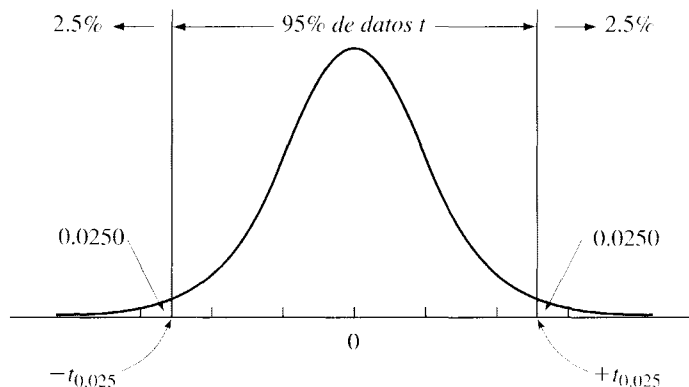


figura 13.3 Porcentaje de datos t entre $\pm t_{\text{crit}}$ para $\alpha = 0.05_{2 \text{ colas}}$ y $gl = N - 1$.

Podemos utilizar la desigualdad anterior para deducir una ecuación a fin de estimar el valor de una μ desconocida. Por lo tanto,

$$-t_{0.025} \leq t_{\text{obt}} \leq t_{0.025}$$

pero

$$t_{\text{obt}} = \frac{\bar{X}_{\text{obt}} - \mu}{s_{\bar{X}}}$$

Por consiguiente,

$$-t_{0.025} \leq \frac{\bar{X}_{\text{obt}} - \mu}{s_{\bar{X}}} \leq t_{0.025}$$

Al despejar μ en esta desigualdad, obtenemos*

$$\bar{X}_{\text{obt}} - s_{\bar{X}} t_{0.025} \leq \mu \leq \bar{X}_{\text{obt}} + s_{\bar{X}} t_{0.025}$$

Esto afirma que las posibilidades de que el intervalo $\bar{X}_{\text{obt}} \pm s_{\bar{X}} t_{0.025}$ contenga a la media poblacional son 95 en 100. Así $\bar{X}_{\text{obt}} \pm s_{\bar{X}} t_{0.025}$ es el intervalo de confianza del 95%. Los límites de confianza inferior y superior son

$$\mu_{\text{inferior}} = \bar{X}_{\text{obt}} - s_{\bar{X}} t_{0.025} \quad \text{límite inferior para el intervalo de confianza de 95\%}$$

$$\mu_{\text{superior}} = \bar{X}_{\text{obt}} + s_{\bar{X}} t_{0.025} \quad \text{límite superior para el intervalo de confianza de 95\%}$$

Ahora estamos listos para presentar un ejemplo al respecto. Regresemos con la administradora de la universidad.

ejemplo

Estimación de la media del CI de los profesores de una universidad

Supongamos que la administradora de una universidad quiere determinar el CI promedio de los profesores que laboran en esa institución. Como es muy costoso hacer una prueba a todos los maestros, se extrae una muestra aleatoria de 20 individuos de toda la población. Cada uno de los profesores resuelve una prueba diseñada para medir el CI; los resultados proporcionan una media muestral de 135 y una desviación estándar de 8. Construya el intervalo de confianza de 95% para la media poblacional.

SOLUCIÓN

A fin de determinar el intervalo de confianza de 95% para la media de la población, despejamos los límites de confianza superior e inferior de las respectivas ecuaciones. Así,

$$\mu_{\text{inferior}} = \bar{X}_{\text{obt}} - s_{\bar{X}} t_{0.025} \quad \text{y} \quad \mu_{\text{superior}} = \bar{X}_{\text{obt}} + s_{\bar{X}} t_{0.025}$$

Despejando $s_{\bar{X}}$,

$$s_{\bar{X}} = \frac{s}{\sqrt{N}} = \frac{8}{\sqrt{20}} = 1.789$$

Según la tabla D, con $\alpha = 0.025$ en cola y $gl = N - 1 = 20 - 1 = 19$,

$$t_{0.025} = 2.093$$

*Para los pasos intermedios de esta deducción, véase la nota 13.2.

Al sustituir los valores de $s_{\bar{x}}$ y $t_{0,025}$ en las ecuaciones de los límites de confianza, obtenemos:

$$\begin{aligned}\mu_{\text{inferior}} &= \bar{X}_{\text{obt}} - s_{\bar{x}} t_{0,025} & \text{y} & & \mu_{\text{superior}} &= \bar{X}_{\text{obt}} + s_{\bar{x}} t_{0,025} \\ &= 135 - 1.789(2.093) & & & &= 135 + 1.789(2.093) \\ &= 135 - 3.744 & & & &= 135 + 3.744 \\ &= 131.26 & \text{límite inferior} & & &= 138.74 & \text{límite superior}\end{aligned}$$

Así, el intervalo de confianza de 95% = 131.26-138.74.

¿Qué significa exactamente decir que el intervalo de confianza de 95% es igual a cierto rango? En el caso del ejemplo anterior, el rango es de 131.26-138.74. Una segunda muestra podría producir un \bar{X}_{obt} distinto y un rango diferente, tal vez $\bar{X}_{\text{obt}} = 138$ y un rango de 133.80-142.20. Si consideráramos todas las muestras posibles de $N = 20$ a partir de la población, habríamos determinado la distribución muestral del intervalo de confianza de 95% para muestras de tamaño 20. *Lo importante en este caso es que el 95% de estos intervalos contendrán la media poblacional y el 5% de ellos no la contendrán.* Por consiguiente, cuando decimos que “el intervalo de confianza de 95% es de 131.26-138.74”, esto significa que la probabilidad de que el intervalo contenga a la media poblacional es de 0.95. Observe que el valor de la probabilidad se aplica al intervalo y no a la media poblacional, puesto que ésta es constante. Lo que varía de una muestra a otra es el intervalo. Por eso, desde el punto de vista técnico, no es correcto decir que “la probabilidad de que la media poblacional quede dentro del intervalo es de 0.95”. La afirmación correcta es que “la probabilidad de que el intervalo contenga la media poblacional es de 0.95”.

Ecuaciones generales para cualquier intervalo de confianza

Las ecuaciones que hemos presentado hasta ahora se refieren únicamente al intervalo de confianza de 95%. Sin embargo, podemos extenderlas fácilmente con el fin de obtener ecuaciones generales para cualquier intervalo de confianza. Así,

$$\begin{aligned}\mu_{\text{inferior}} &= \bar{X}_{\text{obt}} - s_{\bar{x}} t_{\text{crit}} & \text{ecuación general para el límite de confianza inferior} \\ \mu_{\text{superior}} &= \bar{X}_{\text{obt}} + s_{\bar{x}} t_{\text{crit}} & \text{ecuación general para el límite de confianza superior}\end{aligned}$$

donde t_{crit} = valor crítico de t , de una cola, que corresponde al intervalo de confianza deseado.

Así pues, si nos interesara el intervalo de confianza del 99%, $t_{\text{crit}} = t_{0,005}$ = valor crítico de t para $\alpha = 0.005$ _{1 cola}. Para ilustrar este punto, resolvamos el problema anterior para el intervalo de confianza de 99 por ciento.

SOLUCIÓN

Según la tabla D, con $gl = 19$ y $\alpha = 0.005$ _{1 cola},

$$t_{0,005} = 2.861$$

De acuerdo con la solución anterior, $s_{\bar{x}} = 1.789$. Al sustituir estos valores en las ecuaciones para calcular los límites de confianza, obtenemos:

$$\begin{aligned}\mu_{\text{inferior}} &= \bar{X}_{\text{obt}} - s_{\bar{x}} t_{\text{crit}} & \text{y} & & \mu_{\text{superior}} &= \bar{X}_{\text{obt}} + s_{\bar{x}} t_{\text{crit}} \\ &= \bar{X}_{\text{obt}} - s_{\bar{x}} t_{0,005} & & & &= \bar{X}_{\text{obt}} + s_{\bar{x}} t_{0,005} \\ &= 135 - 1.789(2.861) & & & &= 135 + 1.789(2.861)\end{aligned}$$

$$\begin{aligned} &= 135 - 5.118 && = 135 + 5.118 \\ &= 129.88 \text{ límite inferior} && = 140.12 \text{ límite superior} \end{aligned}$$

Por lo tanto, el intervalo de confianza de 99% es = 129.88-140.12.

Observe que este intervalo es mayor que el intervalo de confianza de 95% (131.26-138.74). Como lo hemos comentado con anterioridad, mientras más grande sea el intervalo, tanta más confianza tendremos de que éste contenga la media poblacional.

Tratemos de resolver un problema de práctica.



Un etólogo desea determinar el peso promedio de las marmotas olímpicas adultas (que sólo han sido encontradas en la Península Olímpica de Washington). Como sería muy caro y poco práctico atrapar y medir a toda la población de marmotas, solamente se captura y se pesa una muestra aleatoria de 15 ejemplares adultos. La muestra tiene una media de 7.2 kilogramos y una desviación estándar de 0.48. Construya el intervalo de confianza del 95% para la media poblacional.

SOLUCIÓN

Los datos se presentan en el problema. Para determinar el intervalo de confianza de 95% para la media de la población, calculamos los límites de confianza superior e inferior. Así,

$$\mu_{\text{inferior}} = \bar{X}_{\text{obt}} - s_{\bar{X}}t_{0.025} \quad \text{y} \quad \mu_{\text{superior}} = \bar{X}_{\text{obt}} + s_{\bar{X}}t_{0.025}$$

Despejando $s_{\bar{X}}$,

$$s_{\bar{X}} = \frac{s}{\sqrt{N}} = \frac{0.48}{\sqrt{15}} = 0.124$$

Según la tabla D, con $\alpha = 0.025_{1\text{ cola}}$ y $\text{gl} = N - 1 = 15 - 1 = 14$,

$$t_{0.025} = 2.145$$

Sustituyendo los valores de $s_{\bar{X}}$ y $t_{0.025}$ en las ecuaciones de los límites de confianza, obtenemos:

$$\begin{aligned} \mu_{\text{inferior}} &= \bar{X}_{\text{obt}} - s_{\bar{X}}t_{0.025} && \text{y} && \mu_{\text{superior}} &= \bar{X}_{\text{obt}} + s_{\bar{X}}t_{0.025} \\ &= 7.2 - 0.124(2.145) && && &= 7.2 + 0.124(2.145) \\ &= 7.2 - 0.266 && && &= 7.2 + 0.266 \\ &= 6.93 \text{ límite inferior} && && &= 7.47 \text{ límite superior} \end{aligned}$$

Así, el intervalo de confianza de 95% es 6.93-7.47 kilogramos.

Para estimar el promedio de vida útil de sus bombillas de 100 watts, un fabricante extrae una muestra aleatoria de 200 bombillas y las mantiene encendidas hasta que se funden. La muestra tiene una vida media de 215 horas y una desviación estándar de 8 horas. Construya los límites de confianza de 99% para la media poblacional. Para resolver este problema, utilice el valor de grados de libertad más cercano que aparezca en la tabla.

SOLUCIÓN

Los datos se presentan en el mismo problema.

$$\mu_{\text{inferior}} = \bar{X}_{\text{obt}} - s_{\bar{X}}t_{0,005} \quad \text{y} \quad \mu_{\text{superior}} = \bar{X}_{\text{obt}} + s_{\bar{X}}t_{0,005}$$

$$s_{\bar{X}} = \frac{s}{\sqrt{N}} = \frac{8}{\sqrt{200}} = 0.567$$

Según la tabla D, con $\alpha = 0.005$ cola y $gl = N - 1 = 200 - 1 = 199$,

$$t_{0,005} = 2.617$$

Observe que éste es el valor más cercano que se puede encontrar en la tabla D. Sustituyendo los valores para $s_{\bar{X}}$ y $t_{0,005}$ en las ecuaciones de los límites de confianza, obtenemos

$$\begin{aligned} \mu_{\text{inferior}} &= \bar{X}_{\text{obt}} - s_{\bar{X}}t_{0,005} & \text{y} & & \mu_{\text{superior}} &= \bar{X}_{\text{obt}} + s_{\bar{X}}t_{0,005} \\ &= 215 - 0.567(2.617) & & & &= 215 + 0.567(2.617) \\ &= 213.52 \quad \text{límite inferior} & & & &= 213.52 \quad \text{límite superior} \end{aligned}$$

Así, el intervalo de confianza de 99% es 213.52-216.48 horas.

PRUEBA DE LA SIGNIFICANCIA DE LA r DE PEARSON

Cuando se realiza un estudio de correlación, rara vez se incluye a toda la población. Por lo general, el procedimiento que se aplica consiste en extraer una muestra aleatoria de la población y calcular el coeficiente de correlación para los datos de la muestra. Para determinar si existe una correlación en la población, debemos verificar si la r obtenida (r_{obt}) es significativa. Por supuesto, éste es el mismo procedimiento que hemos utilizado siempre para pruebas de hipótesis. Representamos el coeficiente de correlación poblacional con la letra griega ρ (ro). Una hipótesis alternativa no direccional asegura que $\rho \neq 0$. Una hipótesis alternativa direccional afirma que ρ es positiva o negativa, según la dirección prevista de la relación. Para verificar la hipótesis nula, suponemos que el conjunto muestra de datos X y Y con una correlación de r_{obt} es una muestra aleatoria de una población en la que $\rho = 0$. Podemos generar la distribución muestral de r de manera empírica extrayendo todas las muestras de tamaño N de una población en la cual $\rho = 0$ y calculando r para cada muestra. Al modificar de manera sistemática los datos poblacionales y N se genera la distribución muestral de r .

La significancia de r puede ser evaluada mediante la prueba t . De esta manera,

$$t_{\text{obt}} = \frac{r_{\text{obt}} - \rho}{s_r} \quad \text{prueba } t \text{ para verificar la significancia de } r$$

donde r_{obt} = correlación obtenida en una muestra de N sujetos
 ρ = coeficiente de correlación poblacional
 s_r = estimación de la desviación estándar de la distribución muestral de r

Observe que esto es muy similar a la ecuación t utilizada al analizar la media de una sola muestra. La única diferencia es que ahora el estadístico es r en lugar de \bar{X} .

$$t_{\text{obt}} = \frac{r_{\text{obt}} - \rho}{s_r} = \frac{r_{\text{obt}}}{\sqrt{\frac{1 - r_{\text{obt}}^2}{N - 2}}}$$

donde $\rho = 0$
 $s_r = \sqrt{(1 - r_{\text{obt}}^2)/(N - 2)}$
 $gl = N - 2$

Utilicemos esta ecuación para evaluar la significancia de la correlación obtenida en el problema sobre “el CI y el promedio de calificaciones” del capítulo 6, página 116. Supongamos que los 12 estudiantes forman una muestra aleatoria extraída de una población de alumnos universitarios de licenciatura y que nos interesa averiguar si existe una correlación en la población. Utilizaremos $\alpha = 0.05_{2 \text{ colas}}$ para tomar nuestra decisión.

Por lo general, el primer paso en la resolución de un problema de este tipo consiste en calcular r_{obt} . Sin embargo, esto ya lo hemos hecho y sabemos que $r_{\text{obt}} = 0.856$. Al sustituir este valor en la ecuación t , obtenemos:

$$t_{\text{obt}} = \frac{r_{\text{obt}}}{\sqrt{\frac{1 - r_{\text{obt}}^2}{N - 2}}} = \frac{0.856}{\sqrt{\frac{1 - (0.856)^2}{10}}} = \frac{0.856}{0.163} = 5.252 = 5.25$$

Según la tabla D, con $gl = N - 2 = 10$ y $\alpha = 0.05_{2 \text{ colas}}$,

$$t_{\text{crit}} = \pm 2.228$$

Como $|t_{\text{obt}}| > 2.228$, rechazamos H_0 y concluimos que existe una correlación positiva significativa en la población.

Si bien es cierto que el método anterior funciona, hay una manera aún más sencilla de resolver este problema. Al sustituir t_{crit} en la ecuación t , podemos determinar r_{crit} para cualquier gl y para cualquier nivel de α . Una vez que conocemos r_{crit} lo único que falta es comparar r_{obt} con r_{crit} . La regla de decisión es:

$$\text{Si } |r_{\text{obt}}| \geq |r_{\text{crit}}|, \text{ rechazamos } H_0.$$

Los estadísticos ya han calculado r_{crit} para diversos gl y varios niveles de α . Los resultados se presentan en la tabla E del apéndice D. Esta tabla se utiliza igual que la tabla t (tabla D), excepto que los datos mencionan r_{crit} en lugar de t_{crit} .

Aplicando el método r_{crit} al presente problema, podríamos calcular primero r_{obt} y determinar después r_{crit} con la ayuda de la tabla E. Por último, compararíamos r_{obt} con r_{crit} mediante la regla de decisión. En este ejemplo, ya hemos determinado que $r_{\text{obt}} = 0.856$. Según la tabla E, con $gl = 10$ y $\alpha = 0.05_{2 \text{ colas}}$,

$$r_{\text{crit}} = \pm 0.5760$$

En virtud de que $|r_{\text{obt}}| > 0.5760$, rechazamos H_0 igual que antes. Preferimos esta solución porque es más breve y sencilla que la obtenida mediante la comparación de t_{obt} con t_{crit} .

Intentemos resolver algunos problemas para practicar.



De acuerdo con una idea popular, existe una correlación inversa entre la capacidad matemática y la artística. Una psicóloga ha decidido averiguar si existe algo de verdad en esa creencia. Con ese propósito, extrae al azar una muestra de 15 estudiantes de licenciatura y les pide que resuelvan exámenes para medir los dos tipos de capacidades. Los datos resultantes se presentan aquí. ¿Existe alguna correlación en la población entre la capacidad matemática y la capacidad artística? Utilice una $\alpha = 0.01_{2 \text{ colas}}$

Sujeto número	Habilidad matemática X	Habilidad artística Y	X^2	Y^2	XY
1	15	19	225	361	285
2	30	22	900	484	660
3	35	17	1225	289	595
4	10	25	100	625	250
5	28	23	784	529	644
6	40	21	1600	441	840
7	45	14	2025	196	630
8	24	10	576	100	240
9	21	18	441	324	378
10	25	19	625	361	475
11	18	30	324	900	540
12	13	32	169	1024	416
13	9	16	81	256	144
14	30	28	900	784	840
15	<u>23</u>	<u>24</u>	<u>529</u>	<u>576</u>	<u>552</u>
Total	366	318	10 504	7250	7489

SOLUCIÓN

PASO 1: Cálculo del estadístico adecuado:

$$r_{\text{obt}} = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sqrt{\left[\sum X^2 - \frac{(\sum X)^2}{N}\right] \left[\sum Y^2 - \frac{(\sum Y)^2}{N}\right]}}$$


(continúa)

$$\begin{aligned}
&= \frac{7489 - \frac{366(318)}{15}}{\sqrt{\left[10,504 - \frac{(366)^2}{15}\right]\left[7,250 - \frac{(318)^2}{15}\right]}} \\
&= \frac{-270.2}{894.437} \\
&= -0.302 \\
&= -0.30
\end{aligned}$$

PASO 2: Evaluación del estadístico. Según la tabla E, con $gl = N - 2 = 15 - 2 = 13$ y $\alpha = 0.01_{2\text{ colas}}$,

$$r_{\text{crit}} = \pm 0.6411$$

Como $|r_{\text{obt}}| < 0.6411$, decidimos conservar H_0 .



En el problema de práctica 6.2 del capítulo 6, calculamos la r de Pearson para la relación entre similitud en actitudes y la atracción en una muestra de 15 estudiantes. En ese ejemplo, $r_{\text{obt}} = 0.94$. Usando una $\alpha = 0.05_{2\text{ colas}}$, ahora determinaremos si esto es un valor significativo para r_{obt} .

SOLUCIÓN

Según la tabla E, con $gl = N - 2 = 13$ y $\alpha = 0.05_{2\text{ colas}}$,

$$r_{\text{crit}} = \pm 0.5139$$

Puesto que $|r_{\text{obt}}| > 0.5139$, rechazamos H_0 y concluimos que existe una correlación significativa en la población.

RESUMEN

En este capítulo hemos examinado el uso de la prueba *t* de Student para: 1) poner a prueba hipótesis en experimentos con una sola muestra; 2) estimar la media de una población construyendo intervalos de confianza, y 3) evaluar la significancia de la r de Pearson.

Al poner a prueba hipótesis en experimentos con una sola muestra, la prueba *t* resulta apropiada cuando se conoce la media poblacional de la hipótesis nula y se desconoce la desviación estándar. En este caso, estimamos σ mediante la desviación estándar de la

muestra. La ecuación para calcular t_{obt} es muy similar a la de z_{obt} , excepto que ahora utilizamos s en lugar de σ . La distribución muestral de t es una familia de curvas que varían con los grados de libertad asociados al cálculo de t . Existen $N - 1$ grados de libertad asociados con la prueba *t* para una sola muestra. Las curvas de la distribución muestral son simétricas, con forma de campana y media igual a 0. Sin embargo, éstas se elevan en las colas, en comparación con la distribución normal. Cuando se usa la prueba *t*, calculamos, t_{obt} y

después la evaluamos para determinar si queda dentro de la región crítica. La prueba t resulta apropiada cuando la distribución muestral de \bar{X} es normal. Para que la distribución muestral de \bar{X} sea normal, la población de datos en bruto debe tener una distribución normal, o bien $N \geq 30$.

Después, estudiamos la construcción de los intervalos de confianza para la media poblacional. Definimos un intervalo de confianza como un rango de valores que, probablemente, contiene al valor poblacional. Los límites de confianza son aquellos valores que delimitan el intervalo de confianza. Al analizar este tema, mostramos la manera de construir intervalos de confianza acerca de los cuales tenemos un grado de confianza conocido de que en ellos esté contenida la

media poblacional. Resolvimos algunos problemas, tanto ilustrativos como de práctica, para construir los intervalos de confianza de 95 y 99 por ciento.

El último tema fue evaluar la significancia de la r de Pearson. Señalamos que, como la mayor parte de los datos correlativos se recolectan en muestras, debemos examinar el valor muestral de r (r_{obt}) para ver si existe una correlación en la población. En la evaluación interviene la prueba t . Sin embargo, al sustituir t_{crit} en la ecuación t , podemos determinar r_{crit} para cualquier grado de libertad y cualquier nivel alfa. Evaluamos el valor de r_{obt} comparándolo con r_{crit} para el grado de libertad y nivel alfa dados. Resolvimos varios problemas para practicar la evaluación de r_{obt} .

TÉRMINOS IMPORTANTES

Distribución muestral de t (página 295)	Intervalo de confianza (página 305)	Valor(es) crítico(s) de r (página 310)
Grados de libertad (página 296)	Límites de confianza (página 305)	Valor(es) crítico(s) de t (página 297)

PREGUNTAS Y PROBLEMAS

- Defina cada uno de los conceptos de la sección "términos importantes".
- Supongamos que se cumplen los supuestos fundamentales de la prueba t . ¿Cuáles son las características de la distribución muestral de t ?
- Explique lo que significa *grados de libertad*. Proponga un ejemplo.
- ¿Cuáles son los supuestos fundamentales para el uso apropiado de la prueba t ?
- Analice las semejanzas y las diferencias entre las pruebas z y t .
- Explique brevemente por qué la prueba z es más poderosa que la prueba t .
- ¿Cuál de las siguientes afirmaciones es más correcta desde el punto de vista técnico? 1) Tenemos el 95% de confianza de que la media de la población se encuentra en el intervalo 80-90, o bien, 2) tenemos el 95% de confianza de que el intervalo 80-90 contiene la media de la población. Explique.
- Explique por qué $gl = N - 1$ cuando se utiliza la prueba t con una sola muestra.
- Si el coeficiente de correlación de una muestra tiene un valor diferente de cero (por ejemplo, $r = 0.45$), esto significa automáticamente que la correlación en la población también es diferente de cero. ¿Es correcta esta afirmación? Explique.
- Con el mismo conjunto de datos muestrales, ¿el intervalo de confianza de 99% para la media poblacional es mayor o menor que el intervalo de confianza de 95%? ¿Le parece a usted lógico? Explique.
- Un conjunto muestral de 30 datos tiene una media de 82 y una desviación estándar de 12. ¿Podemos rechazar la hipótesis de que se trata de una muestra aleatoria, extraída de una población normal con $\mu = 85$? Utilice $\alpha = 0.01_{2 \text{ colas}}$ para tomar su decisión. otra
- Un conjunto muestral de 29 datos tiene una media de 76 y una desviación estándar de 7. ¿Podemos aceptar la hipótesis de que se trata de una muestra aleatoria, extraída de una población normal con una media mayor que 72? Utilice $\alpha = 0.01_{1 \text{ cola}}$ para tomar su decisión. otra
- ¿Es razonable considerar una muestra con $N = 22$, $\bar{X}_{\text{obt}} = 42$ y $s = 9$ como una muestra aleatoria extraída de una población normal con $\mu = 38$? Utilice $\alpha = 0.05_{1 \text{ cola}}$ para tomar su decisión. Suponga que \bar{X}_{obt} está en la dirección correcta. otra
- En cada una de las siguientes muestras aleatorias, determine los intervalos de confianza de 95 y el 99% para la media poblacional:
 - $\bar{X}_{\text{obt}} = 25$, $s = 6$, $N = 15$
 - $\bar{X}_{\text{obt}} = 120$, $s = 8$, $N = 30$
 - $\bar{X}_{\text{obt}} = 30.6$, $s = 5.5$, $N = 24$
 - Vuelva a resolver la parte **a** con $N = 30$. ¿Qué ocurre con el intervalo de confianza cuando N crece? otra

15. Supongamos que se desconoce la desviación estándar poblacional del problema 21 del capítulo 12, página 291. Utilice de nuevo $\alpha = 0.05_{2 \text{ colas}}$. ¿Qué podría concluir con respecto a la técnica de la estudiante? Explique la diferencia entre la conclusión del problema 21 y la de este problema. clínica, salud
16. Como director de una escuela preparatoria privada, usted quiere comparar la enseñanza de matemáticas en su escuela con la que se imparte en las escuelas públicas que se encuentran en la misma área. Durante los últimos 5 años, las escuelas públicas han aplicado una prueba de habilidades matemáticas a todos sus estudiantes del último año. La distribución tiene una media de 78. Usted aplica a todos sus alumnos de último año la misma prueba. Los resultados muestran una distribución de 41 datos, con una media de 83 y una desviación estándar de 12.2.
- ¿Cuál es la hipótesis alternativa? Utilice una hipótesis no direccional.
 - ¿Cuál es la hipótesis nula?
 - Si $\alpha = 0.05_{2 \text{ colas}}$, ¿qué conclusión saca usted? educación
17. Un consejero universitario quiere determinar el tiempo promedio de estudio que dedican los estudiantes de primer año a sus materias. Para eso extrae una muestra aleatoria de 61 alumnos de primer año y les pregunta cuántas horas a la semana estudian. La media de los datos resultantes es de 20 horas y la desviación estándar es de 6.5 horas.
- Construya un intervalo de confianza de 95% para la media poblacional.
 - Construya el intervalo de confianza de 99% para la media poblacional. educación
18. Una profesora del programa de estudios sobre la mujer cree que la cantidad de cigarrillos que fuman las mujeres se ha incrementado en años recientes. Un censo exhaustivo realizado hace 2 años con mujeres de una ciudad vecina mostró que el número promedio de cigarrillos que fuman diariamente las mujeres era de 5.4, con una desviación estándar de 2.5. Para evaluar su hipótesis, la profesora calculó el número de cigarrillos fumados diariamente por una muestra aleatoria de 200 mujeres que viven actualmente en esa ciudad. Los datos mostraron que el número de cigarrillos fumados diariamente por las 200 mujeres tiene una media de 6.1 y una desviación estándar de 2.7.
- ¿Es correcta la hipótesis de la profesora? Supongamos que una H_1 direccional resulta apropiada y utilice $\alpha = 0.05_{1 \text{ cola}}$ para tomar su decisión. Asegúrese de aplicar la prueba más sensible para analizar los datos.
- b. Supongamos que se desconoce la media poblacional y analice de nuevo los datos con el mismo nivel alfa. ¿Cuál sería ahora su conclusión? Explique las diferencias entre la parte **a** y la parte **b**. social
19. Un psicólogo cognitivo cree que un fármaco en particular mejora la memoria a corto plazo. El fármaco es seguro y no produce efectos secundarios. Así, realiza un experimento en el cual 8 sujetos son elegidos al azar. Entonces se les administra el fármaco y luego se les concede un tiempo breve para que memoricen una lista de diez palabras. Después de 15 minutos de concluido el periodo de memorización, los sujetos son sometidos a un examen de retención. El número de palabras que cada uno de los sujetos logró recordar correctamente fue: 8, 9, 10, 6, 8, 7, 9, 7. En los últimos años, los psicólogos han recolectado muchos datos por medio de este procedimiento con sujetos similares. Aunque el psicólogo ya extravió los datos originales, recuerda que el promedio fue de 6 palabras correctamente recordadas y que los datos presentaron una distribución normal. Tomando como base estos datos, ¿qué conclusión puede usted sacar acerca del efecto de ese fármaco sobre la memoria a corto plazo? Utilice $\alpha = 0.05_{2 \text{ colas}}$ para tomar su decisión. cognitiva
20. Una doctora en medicina, empleada de una gran empresa, cree que debido a un incremento en la vida sedentaria durante la década pasada, los hombres de mediana edad han engordado. En 1995, la empresa midió el porcentaje de grasa de sus empleados. Para los hombres de mediana edad, los datos tenían una distribución normal, con una media de 22%. Para probar su hipótesis, la médica mide el porcentaje de grasa en una muestra aleatoria de 12 hombres de mediana edad que trabajan actualmente en la empresa. Los porcentajes de grasa registrados fueron los siguientes: 24, 40, 29, 32, 33, 25, 15, 22, 18, 25, 16, 27. Con base en estos datos, ¿podemos concluir que los hombres de edad intermedia que laboran en la empresa han engordado? Suponga que es válida una H_1 direccional y utilice $\alpha = 0.05_{1 \text{ cola}}$ para tomar su decisión. salud
21. Una escuela de comercio local afirma que sus estudiantes graduados consiguen empleos mejor remunerados que el promedio nacional para ese tipo de escuelas. Los salarios del año pasado, pagados a todos los graduados de las escuelas de comercio en su primer empleo, mostraron una media de \$10.20

- por hora. Una muestra aleatoria de 10 graduados de la mencionada escuela en el último año, mostró los siguientes salarios por hora en su primer empleo: \$9.40, \$10.30, \$11.20, \$10.80, \$10.40, \$9.70, \$9.80, \$10.60, \$10.70, \$10.90. Como usted no cree en lo declarado por dicha escuela de comercio, decide evaluar el salario de los graduados de esta escuela de comercio con $\alpha = 0.05_{\text{colas}}$, educación
22. Usted desea estimar el número promedio de vehículos que cruzaron cada mañana un puente congestionado, cercano a su casa, en las horas pico, durante el último año. Para averiguarlo, usted y algunos ayudantes se colocaron en uno de los extremos del puente durante 18 mañanas de ese año, elegidas de manera aleatoria, y contaron el número de vehículos que cruzaban el puente en un periodo de 10 minutos durante la hora pico. Usted encontró que la media fue de 125 vehículos por minuto, con una desviación estándar de 32.
 - a. Construya los límites de confianza de 95% para la media poblacional (vehículos por minuto).
 - b. Construya los límites de confianza de 99% para la media poblacional (vehículos por minuto). otra
 23. En relación con los datos del problema 17 del capítulo 6, página 130, responda las siguientes preguntas:
 - a. Construya una gráfica de dispersión para estos datos.
 - b. Calcule el valor de la r de Pearson.
 - c. ¿Es significativa la correlación entre los cigarrillos fumados y los días de ausencia? Utilice $\alpha = 0.05_{\text{colas}}$ para decidir. salud
 24. En relación con los datos del problema 18 del capítulo 6, página 131, responda las siguientes preguntas.
 - a. Calcule el valor de la r de Pearson para las dos aplicaciones de la prueba de aptitud mecánica.
 - b. ¿Es significativa la correlación? Utilice $\alpha = 0.05_{\text{colas}}$. I/O
 25. En relación con los datos del problema 15 del capítulo 6, página 130, responda las siguientes preguntas.
 - a. Calcule el valor de la r de Pearson para las dos pruebas.
 - b. Utilice $\alpha = 0.05_{\text{colas}}$ y determine si la correlación es significativa. En caso contrario, ¿significa esto que $\rho = 0$? Explique.
 - c. Suponga que incrementa el número de estudiantes a 20 y que $r = 0.653$, utilizando el mismo nivel alfa que en la parte **b**. ¿qué conclusión puede sacar esta vez? Explique. educación
 26. Un psicólogo del desarrollo quiere saber si los padres que sufren tensiones tienden a tener hijos tensos. Se realiza un estudio que incluye a un padre por cada 15 familias y al hijo mayor de cada familia y se mide la tensión de cada par de ellos. La r de Pearson es igual 0.582. Utilizando $\alpha = 0.05_{\text{colas}}$, ¿es significativa la correlación? desarrollo, clínica

NOTAS

13.1

$$t_{\text{obt}} = \frac{\bar{X}_{\text{obt}} - \mu}{s/\sqrt{N}}$$

$$= \frac{\bar{X}_{\text{obt}} - \mu}{\sqrt{\frac{SC}{N-1}}/\sqrt{N}}$$

$$(t_{\text{obt}})^2 = \frac{(\bar{X}_{\text{obt}} - \mu)^2}{\left(\frac{SC}{N-1}\right)\left(\frac{1}{N}\right)} = \frac{(\bar{X}_{\text{obt}} - \mu)^2}{\frac{SC}{N(N-1)}}$$

$$t_{\text{obt}} = \frac{\bar{X}_{\text{obt}} - \mu}{\sqrt{\frac{SC}{N(N-1)}}}$$

13.2

$$-t_{0.025} \leq \frac{\bar{X}_{\text{obt}} - \mu}{s_{\bar{X}}} \leq t_{0.025}$$

Dado

Sustituyendo $\sqrt{\frac{SC}{N-1}}$ en lugar de s

Elevando al cuadrado ambos lados de la ecuación y agrupando los términos.

Aplicando la raíz cuadrada en ambos lados de la ecuación

Dado

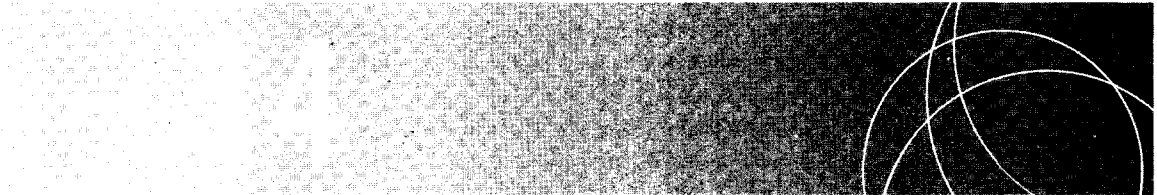
$$\begin{aligned}
 -s_{\bar{X}}t_{0.025} &\leq \bar{X}_{\text{obt}} - \mu \leq s_{\bar{X}}t_{0.025} && \textit{Multiplicando por } s_{\bar{X}} \\
 -\bar{X}_{\text{obt}} - s_{\bar{X}}t_{0.025} &\leq -\mu \leq -\bar{X}_{\text{obt}} + s_{\bar{X}}t_{0.025} && \textit{Restando } \bar{X}_{\text{obt}} \\
 \bar{X}_{\text{obt}} + s_{\bar{X}}t_{0.025} &\geq \mu \geq \bar{X}_{\text{obt}} - s_{\bar{X}}t_{0.025} && \textit{Multiplicando por } -1 \\
 \bar{X}_{\text{obt}} - s_{\bar{X}}t_{0.025} &\leq \mu \leq \bar{X}_{\text{obt}} + s_{\bar{X}}t_{0.025} && \textit{Reagrupando términos}
 \end{aligned}$$

Si desea encontrar problemas de práctica adicionales totalmente resueltos sobre los temas que acaba de estudiar, busque Chapter 13 en el material adicional del libro, disponible en inglés en: <http://www.thomsonlearning.com.mx>.

Si pretende resolver más problemas de práctica, relacionados con lo que se presentó en este capítulo y resueltos por medio de ecuaciones conceptuales, busque Chapter 13 en el material adicional del libro (en inglés), en: <http://www.thomsonlearning.com.mx>.

Si quiere resolver problemas para computadora sobre los temas aquí presentados, busque Chapter 13 en el material adicional (en inglés) del libro, en: <http://www.thomsonlearning.com.mx>.

Prueba t de Student para grupos independientes correlacionados



Introducción

Prueba t de Student para grupos correlacionados

Un experimento: Estimulación cerebral y alimentación

Comparación de la prueba t para grupos correlacionados y la prueba t para muestras únicas

Revisión y análisis del experimento de estimulación cerebral

Comparación de la prueba t para grupos correlacionados y la prueba de los signos

Supuestos subyacentes a la prueba t para grupos correlacionados

Pruebas z y t para grupos independientes

Diseño de grupos independientes

Prueba z para grupos independientes

Un experimento: Hormona X y comportamiento sexual

Distribución muestral de la diferencia entre las medias muestrales $(\bar{X}_1 - \bar{X}_2)$

Un experimento: Revisión del experimento de la hormona X

Prueba t de Student para grupos independientes

Comparación de las ecuaciones para z_{obt} y t_{obt}

Análisis del experimento de la hormona X con la prueba t para grupos independientes

Cálculo de t_{obt} cuando $n_1 = n_2$

Supuestos subyacentes a la prueba t para grupos independientes

Violación de los supuestos de la prueba t para grupos independientes

Magnitud del efecto

Potencia de la prueba t

Comparación de los diseños para grupos correlacionados y grupos independientes

Resumen

Términos importantes

Preguntas y problemas

Notas



INTRODUCCIÓN

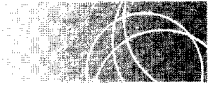
En los capítulos 12 y 13 vimos que la prueba de hipótesis consiste fundamentalmente en dos pasos: 1) el cálculo del estadístico apropiado y 2) la evaluación de dicho estadístico por medio de su distribución muestral. Además, analizamos la forma en que podemos utilizar las pruebas *z* y *t* para evaluar las hipótesis que han sido investigadas en experimentos con una sola muestra. En este capítulo presentamos la prueba *t* en conjunción con experimentos que implican dos condiciones o dos muestras.

Ya cuando aplicamos las pruebas del signo nos encontramos con experimentos que presentaban dos condiciones. El experimento con dos condiciones, ya sea con el diseño de grupos correlacionados o con el de grupos independientes, tiene grandes ventajas sobre el experimento de una sola muestra que analizamos anteriormente. Una de las principales limitaciones del experimento con una sola muestra es que en él es necesario especificar un parámetro de la población cuando menos (μ). En la inmensa mayoría de los casos no se dispone de esta información. Como veremos más adelante en este capítulo, con el experimento con dos condiciones se suprime por completo la necesidad de medir los parámetros de la población al realizar la prueba de una hipótesis. Obviamente, esto tiene una utilidad práctica muy amplia.

Una segunda ventaja muy importante del experimento con dos condiciones guarda relación con la interpretación de los resultados del estudio. En la mayoría de los casos, la metodología científica aplicada de la manera correcta no permite que un investigador, al realizar un experimento, use datos poblacionales adquiridos con anterioridad. Por ejemplo, en el problema ilustrativo acerca de la emisión de las primeras palabras (página 298) empleamos una media poblacional igual a 13.0 meses. ¿Cómo sabemos que la media es realmente de 13.0 meses? Supongamos que estas cifras fueron recopiladas entre 3 y 5 años antes de la realización del experimento. ¿Cómo podemos saber si los nuevos infantes no han cambiado en estos años? ¿Y qué podemos decir de las condiciones en las cuales fueron recopilados los datos poblacionales? ¿Son las mismas cuando se realiza el experimento? ¿No es posible que las personas que recopilaron los datos de la población no hayan estado tan motivadas como el experimentador y que, por lo tanto, no compilaron los datos con cuidado? ¿Cómo fueron recopilados los datos? ¿Estando disponibles en el momento en que el niño pronunciaba su primera palabra? Esto es poco probable. Tal vez, los datos fueron recopilados preguntando a los padres cuándo hablaron sus hijos por primera vez. ¿Qué tan precisa es, entonces, la media de la población?

Aunque no existieran los problemas anteriores, se presentan otros que están relacionados con el método mismo con el cual se realiza el experimento. Por ejemplo, suponiendo que el valor de 13.0 meses sea exacto y que se aplica en forma apropiada a la muestra de 15 niños, ¿cómo podemos estar seguros de que la técnica aplicada por el experimentador fue lo que propició la emisión de las primeras palabras? ¿No podría deberse este hecho a la atención, el manejo o la estimulación adicionales que se prodigaron a los niños junto con el método, en lugar de ser producto exclusivo de este último?

Podemos superar muchos de estos problemas realizando un experimento con dos condiciones. Si utilizamos dos grupos de niños (ya sea mediante la formación de pares igualados [diseño de grupos correlacionados] o por asignaciones aleatorias [diseño de grupos independientes]); si damos a cada grupo el mismo tratamiento, salvo por la técnica particular diseñada por el experimentador (la misma atención, el mismo manejo, etcétera); si trabajamos con estos grupos de manera concurrente; si utilizamos a las mismas personas para reunir los datos de ambos grupos, y así sucesivamente, podremos eliminar muchas de las explicaciones alternativas en torno de los resultados. En el siguiente análisis consideraremos en primer lugar la prueba *t* para el diseño de grupos correlacionados, y después para el diseño de grupos independientes.



PRUEBA *t* DE STUDENT PARA GRUPOS CORRELACIONADOS*

Usted recordará que, en el diseño de medidas repetidas o de grupos correlacionados, cada sujeto recibe dos o más tratamientos: Se calcula la diferencia entre los puntajes de cada condición para cada sujeto y se analizan los puntajes de diferencia resultantes. En el experimento más sencillo de este tipo se utilizan dos condiciones, llamadas con frecuencia *control* y *experimental*, o *antes* y *después*. En una variante de este diseño, en lugar de someter al mismo sujeto a las dos condiciones, se usan parejas de sujetos que coincidan en una o varias características y cada sujeto es asignado aleatoriamente a una condición. De esta manera, las parejas pueden igualarse por CI, edad, género, etcétera. A continuación, los puntajes de diferencia entre las parejas son analizados en la misma forma que cuando se somete el mismo sujeto a ambas condiciones. Este diseño se conoce también como *diseño de grupos correlacionados*, porque los sujetos no son asignados a los grupos de modo independiente: es decir, las parejas comparten características comunes igualadas de manera específica. En el diseño de grupos independientes, que estudiaremos más adelante en este capítulo, no se recurre a la formación de parejas.

La primera vez que nos encontramos con el diseño de grupos correlacionados fue cuando utilizamos la prueba del signo. Sin embargo, la prueba del signo tenía baja potencia porque pasaba por alto la magnitud de los puntajes de diferencia. A pesar de eso usamos la prueba del signo por su sencillez. En el análisis de experimentos reales, es probable que se utilice otro instrumento, como la prueba *t*. Esta prueba para grupos correlacionados permite usar *ambas* cosas: la magnitud y la dirección de los puntajes de diferencia. En esencia, se considera en ella a los puntajes de diferencia como si fueran datos en bruto y se pone a prueba la hipótesis de que dichos puntajes constituyen una muestra aleatoria extraída de una población de puntajes de diferencias cuya media es cero. Esto se puede apreciar mejor por medio de un ejemplo.

un experimento

Estimulación cerebral y alimentación

Como caso ilustrativo, supongamos que un psicólogo tiene la creencia de que una región del cerebro llamada hipotálamo lateral interviene en el comportamiento de alimentación. Una forma de probar esto consiste en servirse de un grupo de animales (por ejemplo, ratas) al cual se le implanta permanentemente, en el hipotálamo lateral, un electrodo que descarga estímulos eléctricos. Si el hipotálamo lateral influye en el comportamiento alimentario, es posible que los estímulos eléctricos aplicados en esa parte del cerebro alteren la cantidad de alimento ingerida. Para establecer el control, determinando el efecto de la estimulación cerebral por sí misma, se implanta otro electrodo en un área neutral del cerebro de cada animal. Cada área se estimula durante un periodo fijo y se registra la cantidad de alimento ingerida. Después, se calcula un puntaje de diferencia para cada animal.

Supongamos que no existen evidencias suficientes para plantear una hipótesis alternativa direccional. Por lo tanto, se planea una evaluación de dos cosas. La hipótesis alternativa establece que el estímulo eléctrico al hipotálamo lateral influye en la cantidad de alimento ingerida. Por el contrario, la hipótesis nula especifica que el estímulo eléctrico del hipotálamo lateral no afecta la cantidad de alimento ingerida. Si H_0 es verdadera, el puntaje de diferencia para cada rata se deberá a factores aleatorios. A veces será positivo y en otras ocasiones será negativo; en algunas ocasiones tendrá una magnitud grande y en otras pequeña. Si el experimento se realiza con un gran número de ratas, por ejemplo con toda la población, la media de los puntajes de diferencia será igual a cero.[‡] La figura 14.1

*Véase la nota 14.1.

‡Véase la nota 14.2.

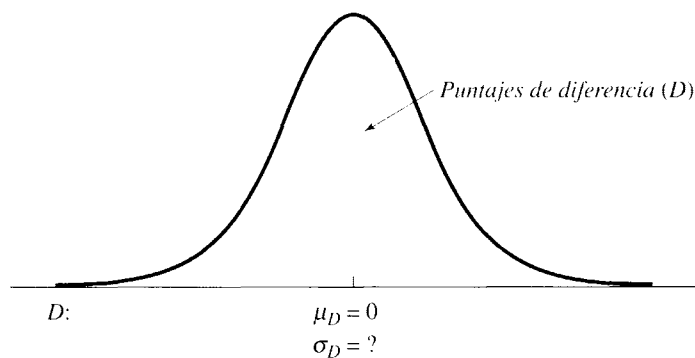


figura 14.1 Población de puntajes de diferencia de la hipótesis nula.

muestra esa distribución. Observe con atención y verá que la media de esta población es conocida ($\mu_D = 0$) y que la desviación estándar ($\sigma_D = ?$) no se conoce. La explicación con base en factores aleatorios supone que los puntajes de diferencia de la muestra de este experimento constituyen una muestra aleatoria de esa población de puntajes de diferencia. Por lo tanto, tenemos un caso en el que existe un conjunto de datos (por ejemplo, los puntajes de diferencia de la muestra) y deseamos determinar si resulta razonable que consideremos estos datos como una muestra aleatoria extraída de una población de puntajes de diferencia, con una media conocida ($\mu_D = 0$) y una desviación estándar desconocida.

Comparación de la prueba *t* para grupos correlacionados y la prueba *t* para muestras únicas

La situación que acabamos de describir es casi idéntica a la que consideramos anteriormente con la prueba *t* de una sola muestra. La única diferencia radica en que, en el experimento de grupos correlacionados, trabajamos con puntajes de *diferencia* y no con datos en *bruto*. En consecuencia, las ecuaciones para cada caso tienen que ser bastante similares. Estas ecuaciones se presentan en la tabla 14.1.

tabla 14.1 Prueba *t* para muestras simples y grupos correlacionados

Prueba <i>t</i> para muestras simples	Prueba <i>t</i> para grupos correlacionados
$t_{\text{obt}} = \frac{\bar{X}_{\text{obt}} - \mu}{s/\sqrt{N}}$	$t_{\text{obt}} = \frac{\bar{D}_{\text{obt}} - \mu_D}{s_D/\sqrt{N}}$
$t_{\text{obt}} = \frac{\bar{X}_{\text{obt}} - \mu}{\sqrt{\frac{SC}{N(N-1)}}}$	$t_{\text{obt}} = \frac{\bar{D}_{\text{obt}} - \mu_D}{\sqrt{\frac{SC_D}{N(N-1)}}}$
$SC = \sum X^2 - \frac{(\sum X)^2}{N}$	$SC_D = \sum D^2 - \frac{(\sum D)^2}{N}$
donde	<p>D = puntaje de diferencia</p> <p>\bar{D}_{obt} = media de los puntajes de diferencia de la muestra</p> <p>μ_D = media de la población de puntajes de diferencia</p> <p>s_D = desviación estándar de los puntajes de diferencia de la muestra</p> <p>N = número de puntajes de diferencia</p> <p>$SC_D = \sum (D - \bar{D})^2$ = suma de los cuadrados de los puntajes de diferencia de la muestra</p>

Es obvio que los dos conjuntos de ecuaciones son idénticos, excepto por el hecho de que, en el caso de una sola muestra, trabajamos con datos en bruto, mientras que en los experimentos de grupos correlacionados analizamos puntajes de diferencia. Ahora vamos a agregar algunos números al experimento de estimulación cerebral y veremos cómo se utiliza la prueba *t* para grupos correlacionados.

Revisión y análisis del experimento de estimulación cerebral

Un psicólogo piensa que el hipotálamo lateral está relacionado con el comportamiento de alimentación. De ser así, entonces con un estímulo eléctrico aplicado a esta área podría modificarse la cantidad de alimento ingerida. Para probar esta posibilidad, a 10 ratas se les implantan electrodos permanentemente. Cada roedor tiene insertados dos electrodos, uno implantado en el hipotálamo lateral y el otro en un área donde se sabe que el estímulo eléctrico no produce efecto alguno. Después que los animales se recuperan de la operación quirúrgica, cada uno de ellos recibe 30 minutos de estimulación eléctrica en las áreas cerebrales correspondientes y se mide la cantidad de comida que ingieren en ese periodo de estimulación. La cantidad de alimento, en gramos, ingerida por los animales durante la estimulación se presenta en la tabla 14.2.

1. ¿Cuál es la hipótesis alternativa? Supongamos que una hipótesis no direccional resulta apropiada.
2. ¿Cuál es la hipótesis nula?
3. ¿Cuál es la conclusión? Utilice una $\alpha = 0.05_{2 \text{ colias}}$.

SOLUCIÓN

1. Hipótesis alternativa: Esta hipótesis específica que la estimulación eléctrica en el hipotálamo lateral afecta la cantidad de alimento ingerida. Los puntajes de dife-

tabla 14.2 Datos obtenidos en el experimento sobre la estimulación cerebral

Sujeto	Alimento ingerido		Diferencia	
	Hipotálamo lateral (g)	Área neutral (g)	<i>D</i>	<i>D</i> ²
1	10	6	4	16
2	18	8	10	100
3	16	11	5	25
4	22	14	8	64
5	14	10	4	16
6	25	20	5	25
7	17	10	7	49
8	22	18	4	16
9	12	14	-2	4
10	21	13	8	64
			53	379

$$N = 10 \quad \bar{D}_{\text{obt}} = \frac{\sum D}{N} = \frac{53}{10} = 5.3$$

- rencia de la muestra con una media $\bar{D}_{\text{obt}} = 5.3$ son una muestra aleatoria extraída de una población de puntajes de diferencia cuya media $\mu_D \neq 0$.
- Hipótesis nula: Esta hipótesis establece que la estimulación eléctrica en el hipotálamo lateral no produce efecto alguno sobre la cantidad de alimento ingerida. Los puntajes de diferencia de la muestra con una media $\bar{D}_{\text{obt}} = 5.3$ son una muestra aleatoria extraída de una población de diferencias cuya media $\mu_D = 0$.
 - Conclusión, utilizando $\alpha = 0.05_{2 \text{ colas}}$:

PASO 1: Cálculo del estadístico apropiado. Como se trata de un diseño de grupos correlacionados, lo que nos interesa conocer es la diferencia que se presenta entre las parejas de datos, y no los datos de cada condición en sí misma. Las diferencias entre los datos se presentan en la tabla 14.2. De las pruebas que hemos analizado hasta ahora, podemos elegir tanto la prueba del signo como la prueba t . Queremos usar la prueba más sensible, por lo cual elegimos la prueba t . Según la tabla de datos, $N = 10$ y $\bar{D}_{\text{obt}} = 5.3$. El cálculo de t_{obt} se realiza de la siguiente manera:

$$\begin{aligned}
 t_{\text{obt}} &= \frac{\bar{D}_{\text{obt}} - \mu_D}{\sqrt{\frac{SC_D}{N(N-1)}}} & SC_D &= \sum D^2 - \frac{(\sum D)^2}{N} \\
 &= \frac{5.3 - 0}{\sqrt{\frac{98.1}{10(9)}}} & &= 379 - \frac{(53)^2}{10} \\
 &= \frac{5.3}{\sqrt{1.09}} & &= 98.1 \\
 &= 5.08
 \end{aligned}$$

PASO 2: Evaluación del estadístico. Como en el caso de la prueba t para una sola muestra, si t_{obt} queda dentro de la región crítica, la conclusión consiste en rechazar H_0 . Por lo tanto, aplicamos la misma regla de decisión, es decir:

$$\text{Si } |t_{\text{obt}}| \geq |t_{\text{crit}}| \text{ rechazamos } H_0.$$

Los grados de libertad son iguales al número de puntajes de diferencia menos uno. Así pues, $gl = N - 1 = 10 - 1 = 9$. Según la tabla D del apéndice D, con $\alpha = 0.05_{2 \text{ colas}}$ y $gl = 9$,

$$t_{\text{crit}} = \pm 2.262$$

Como quiera que $|t_{\text{obt}}| > 2.262$, rechazamos H_0 y concluimos que la estimulación eléctrica del hipotálamo lateral afecta el consumo de alimento. Tal parece que incrementa la cantidad ingerida.



A fin de motivar a los ciudadanos para que ahorren gasolina, el gobierno estudia la idea de emprender una campaña nacional a favor del ahorro de ese energético. Sin embargo, antes de emprenderla a nivel nacional, el gobierno decide realizar un experimento que le permita evaluar la eficacia de la campaña. Para el experimento, realiza la campaña de ahorro de gasolina en un área geográfica pequeña, pero

representativa. Entonces, 12 familias de esa área son elegidas al azar y se mide la cantidad de gasolina que consumieron durante el mes anterior al lanzamiento de la campaña de publicidad y durante el mes siguiente a la campaña. Así se obtienen los siguientes datos:

Familia	Antes de la campaña (galones/mes)	Después de la campaña (galones/mes)	Diferencia	
			D	D^2
A	55	48	7	49
B	43	38	5	25
C	51	53	-2	4
D	62	58	4	16
E	35	36	-1	1
F	48	42	6	36
G	58	55	3	9
H	45	40	5	25
I	48	49	-1	1
J	54	50	4	16
K	56	58	-2	4
L	32	25	7	49
			<u>35</u>	<u>235</u>

$$N = 12 \quad \bar{D}_{\text{obt}} = \frac{\sum D}{N} = \frac{35}{12} = 2.917$$

- ¿Cuál es la hipótesis alternativa? Utilice una hipótesis no direccional.
- ¿Cuál es la hipótesis nula?
- ¿Cuál es la conclusión? Utilizando $\alpha = 0.05_{2 \text{ colas}}$.

SOLUCIÓN

- Hipótesis alternativa: La campaña de ahorro afecta la cantidad de gasolina consumida. La muestra con $\bar{D}_{\text{obt}} = 2.917$ es una muestra aleatoria extraída de una población de puntajes de diferencia en la cual $\mu_D \neq 0$.
- Hipótesis nula: La campaña de ahorro no influye en la cantidad de gasolina consumida. La muestra con $\bar{D}_{\text{obt}} = 2.917$ es una muestra aleatoria extraída de una población de puntajes de diferencia en la cual $\mu_D = 0$.
- Conclusión, utilizando $\alpha = 0.05_{2 \text{ colas}}$:

PASO 1: Cálculo del estadístico apropiado. Los puntajes de diferencia se presentan en la tabla anterior. Hemos restado los datos de la columna “después” de los datos de la columna “antes”. Si se cumplen los supuestos de t , entonces el estadístico apropiado es t_{obt} . Según la tabla de datos, $N = 12$ y $\bar{D}_{\text{obt}} = 2.917$.

(continúa)

$$\begin{aligned}
 t_{\text{obt}} &= \frac{\bar{D}_{\text{obt}} - \mu_D}{\sqrt{\frac{SC_D}{N(N-1)}}} & SC_D &= \sum D^2 - \frac{(\sum D)^2}{N} \\
 &= \frac{2.917 - 0}{\sqrt{\frac{132.917}{12(11)}}} & &= 235 - \frac{(35)^2}{12} \\
 &= 2.91 & &= 132.917
 \end{aligned}$$

PASO 2: Evaluación del estadístico. Grados de libertad = $N - 1 = 12 - 1 = 11$. Según la tabla D, con $\alpha = 0.05_{2 \text{ colas}}$ y 11 grados de libertad.

$$t_{\text{crit}} = \pm 2.201$$

En virtud de que $|t_{\text{obt}}| > 2.201$, rechazamos H_0 . La campaña de ahorro influye en la cantidad de gasolina consumida. Según parece, disminuye el consumo de ese combustible.

Resolvamos otro problema de práctica.



En fecha reciente fue descubierto un neurotransmisor cerebral endógeno llamado galanina. Según parece, éste afecta de manera directa el deseo de ingerir alimentos con un alto contenido de grasa. Mientras más alta sea la cantidad de este neurotransmisor de origen natural en un individuo, mayor será el apetito que éste sienta por la comida con alto contenido de grasa. Recientemente, una compañía farmacéutica desarrolló una sustancia experimental que bloquea la galanina sin alterar el apetito por otros alimentos más saludables (es decir, con menos grasa). Un neurocientífico de la compañía piensa que esa sustancia experimental será muy útil para controlar la obesidad. En forma específica, él cree que la administración diaria de este medicamento hará que la persona ingiera alimentos con menos grasa y, por ende, promoverá la pérdida de peso. En estas condiciones, se realiza un experimento, para el cual se elige de manera aleatoria a 15 mujeres obesas, todas ellas voluntarias, y se les administra el medicamento experimental durante 6 meses. Se registra el peso inicial y el peso final (después de 6 meses) de cada persona. Los pesos registrados se presentan en la siguiente tabla.

Número de sujeto	Peso inicial (libras)	Peso final (libras)	<i>D</i>	<i>D</i> ²
1	165	145	20	400
2	143	137	6	36
3	175	170	5	25

Número de sujeto	Peso inicial (libras)	Peso final (libras)	<i>D</i>	<i>D</i> ²
4	135	136	-1	1
5	148	141	7	49
6	155	138	17	289
7	158	137	21	441
8	140	125	15	225
9	172	161	11	121
10	164	156	8	64
11	178	165	13	169
12	182	170	12	144
13	190	176	14	196
14	169	154	15	225
15	157	143	14	196
			<u>14</u>	<u>196</u>
			177	2581

$$N = 15 \quad \bar{D}_{\text{obt}} = \frac{\sum D}{N} = \frac{177}{15} = 11.80$$

- ¿Cuál es la hipótesis alternativa direccional?
- ¿Cuál es la hipótesis nula?
- ¿Cuál es la conclusión? utilice $\alpha = 0.05$ cola.

SOLUCIÓN

- Hipótesis alternativa: El uso del medicamento experimental produce pérdida de peso. La muestra con $\bar{D}_{\text{obt}} = 11.80$ es una muestra aleatoria, extraída de una población de puntajes de diferencia donde $\mu_D > 0$.
- Hipótesis nula: El uso del medicamento experimental no influye ni en la pérdida ni en el aumento de peso. La muestra con $\bar{D}_{\text{obt}} = 11.80$ es aleatoria y fue extraída de una población de puntajes de diferencia con $\mu_D \leq 0$.
- Conclusión, utilizando $\alpha = 0.05$ cola.

PASO 1: Cálculo del estadístico apropiado. Los puntajes de diferencia se presentan en la tabla. Hemos restado los “pesos finales” de los “pesos iniciales”. Si se cumplen los supuestos acerca de *t*, el estadístico apropiado es t_{obt} . Según la tabla de datos, $N = 15$, $\bar{D}_{\text{obt}} = 11.80$.

$$t_{\text{obt}} = \frac{\bar{D}_{\text{obt}} - \mu_D}{\sqrt{\frac{SC_D}{N(N-1)}}} \quad SC_D = \sum D^2 - \frac{(\sum D)^2}{N}$$

(continúa)

$$\begin{aligned}
 &= \frac{11.80 - 0}{\sqrt{\frac{492.40}{15(14)}}} &&= 2581 - \frac{(177)^2}{15} \\
 &= 7.71 &&= 492.40
 \end{aligned}$$

PASO 2: Evaluación del estadístico. Grados de libertad = $N - 1 = 14$. Según la tabla D, con $\alpha = 0.05$ cola y 14 gl,

$$t_{\text{crit}} = 1.761$$

En virtud de que $|t_{\text{obt}}| > 1.761$, rechazamos H_0 . Tal parece que la administración del medicamento experimental produce una pérdida de peso.

Comparación de la prueba t para grupos correlacionados y la prueba de los signos

Podríamos haber resuelto los dos problemas anteriores por medio de la prueba de los signos; sin embargo, en lugar de eso, elegimos la prueba t porque es más sensible. A fin de ilustrar este punto, aplicaremos la prueba de los signos para resolver el problema referente al ahorro de gasolina.

SOLUCIÓN

PASO 1: Cálculo del estadístico. Existen 8 signos positivos en la muestra.

PASO 2: Evaluación del estadístico. Con $N = 12$, $P = 0.50$ y $\alpha = 0.05$ colas,

$$p(8 \text{ o más signos positivos}) = p(8) + p(9) + p(10) + p(11) + p(12)$$

Según la tabla B del apéndice D.

$$\begin{aligned}
 p(8 \text{ o más signos positivos}) &= 0.1208 + 0.0537 + 0.0161 + 0.0029 + 0.0002 \\
 &= 0.1937
 \end{aligned}$$

Como el nivel alfa tiene dos colas,

$$\begin{aligned}
 p(\text{resultado tan extremo como 8 signos positivos}) &= 2[p(8 \text{ o más signos positivos})] \\
 &= 2(0.1937) \\
 &= 0.3874
 \end{aligned}$$

Puesto que $0.3874 > 0.05$, la conclusión es conservar H_0 .

No podemos rechazar H_0 con la prueba de los signos, pero lo pudimos rechazar con la prueba t . ¿Significa esto que la campaña resulta eficaz si analizamos los datos con la prueba t y que no es eficaz si los analizamos con la prueba de los signos? Desde luego que no. A causa de la baja potencia que posee de la prueba de los signos, en ella hay una alta probabilidad de cometer un error de tipo II (es decir, conservar H_0 cuando ésta es falsa). Por lo general, la prueba t es más potente que la prueba de los signos. La potencia adicional hace que H_0 tenga una mayor probabilidad de ser rechazada cuando es falsa. En este caso, la potencia adicional hizo posible el rechazo de H_0 . Cuando varias pruebas son apropiadas para analizar los datos, una regla general del análisis estadístico aconseja utilizar la más potente, pues ésta es la que ofrece la más alta probabilidad de rechazar H_0 cuando sea falsa.

Supuestos subyacentes a la prueba t para grupos correlacionados

Los supuestos son muy similares a los que subyacen a la prueba t para una sola muestra. La prueba t para grupos correlacionados requiere que la distribución muestral de \bar{D} sea una distribución normal. Esto significa que N debe ser igual o mayor que 30; suponiendo que la forma de la población no difiere mucho de la normal, o bien que los propios datos poblacionales están distribuidos de manera normal.*



PRUEBAS z Y t PARA GRUPOS INDEPENDIENTES

Diseño de grupos independientes

Dos diseños básicos son los que se usan con más frecuencia en el estudio del comportamiento. El primero fue presentado cuando analizamos la prueba de los signos y la prueba t para grupos correlacionados. Su nombre es diseño de medidas repetidas o replicadas. La forma más sencilla de este diseño utiliza dos condiciones: una experimental y una de control. La característica esencial del diseño es que usa datos apareados entre ambas condiciones y los puntajes de *diferencia* de cada par de datos son analizados para determinar si el azar, por sí solo, puede explicarlos razonablemente.

El otro tipo de diseño se conoce como *diseño de grupos independientes*. Igual que el diseño de grupos correlacionados, el diseño de grupos independientes incluye experimentos en los cuales se usan dos o más condiciones. En cada condición se emplea un nivel diferente de la variable independiente. El más fundamental de esos experimentos tiene únicamente dos condiciones: la condición experimental y la condición control. En este capítulo vamos a considerar el experimento básico, en el cual intervienen sólo dos condiciones. Otros experimentos más complicados serán considerados en el capítulo 15.

En el diseño de grupos independientes, los sujetos son seleccionados aleatoriamente a partir de la población correspondiente, y luego se les asigna, también en forma aleatoria, ya sea a la condición experimental o a la control. No hay base alguna para formar las parejas de sujetos y cada uno de éstos es sometido a prueba sólo una vez. Todos los sujetos de la condición experimental reciben el nivel apropiado de la variable independiente para la condición experimental, y al conjunto de esos sujetos se le conoce como el "grupo experimental". Todos los sujetos de la condición control reciben el nivel de la variable independiente que se considere apropiado para la condición control y se dice que pertenecen al "grupo control".

En virtud de que los sujetos son asignados aleatoriamente a una u otra condición, cuando se analizan los datos no existe base alguna para formar las parejas de datos entre las dos condiciones. En lugar de eso, se calcula un estadístico para cada grupo por *separado* y luego los estadísticos de los dos grupos son comparados para determinar si el azar, por sí solo, es una explicación razonable de las diferencias observadas entre los estadísticos de los dos grupos. El estadístico que se calcula para cada muestra depende de la prueba de inferencia que se utilice. En la prueba t para grupos independientes se calcula la media de cada muestra y luego se analiza la diferencia entre esas dos medias muestrales para determinar si el azar, por sí solo, es una explicación razonable de la diferencia observada entre ambas medias.

* Muchos autores limitan el uso de la prueba t a los datos con escala de intervalo o razón. Véase la nota al pie de la página 31 del capítulo 2, para consultar otras referencias al analizar este punto.

H_1 y H_0 Los datos muestrales, en una de las condiciones (digamos, la condición 1), pueden considerarse como una muestra aleatoria extraída de una población formada por datos que se distribuyen de manera normal; esta población sería resultado de que todos los individuos de la población estuvieran en esa condición (la condición 1). Llamemos μ_1 a la media de esta población hipotética y σ_1 a su desviación estándar. De igual manera, los datos muestrales de la condición 2 pueden considerarse como una muestra aleatoria extraída de una población formada por datos distribuidos normalmente, la cual es resultado de que todos los individuos estuvieran bajo la condición 2. Podemos llamar μ_2 a la media de esta segunda población y σ_2 a su desviación estándar. Así,

μ_1 = media de la población que está en la condición 1

σ_1 = desviación estándar de la población que está en la condición 1

μ_2 = media de la población que está en la condición 2

σ_2 = desviación estándar de la población que está en la condición 2

Se supone que el hecho de modificar el nivel de la variable independiente afecta a la media de la distribución (μ_2), pero no a la desviación estándar (σ_2) ni a la varianza (σ_2^2). Por lo tanto, según esta suposición, si la variable independiente tiene un efecto real, entonces las medias de las poblaciones serán distintas, pero sus varianzas serán las mismas. Así, suponemos que σ_1^2 es igual a σ_2^2 . Un caso en el que se cumple este supuesto es cuando la variable independiente tiene el mismo efecto sobre cada individuo. Una hipótesis alternativa direccional predice que las muestras son muestras aleatorias seleccionadas de poblaciones en las cuales $\mu_1 > \mu_2$ o $\mu_1 < \mu_2$, según la dirección del efecto. Una hipótesis alternativa no direccional predice que $\mu_1 \neq \mu_2$. Si la variable independiente no tiene efecto alguno, las muestras serían aleatorias extraídas de poblaciones en las cuales $\mu_1 = \mu_2^*$ y el azar sería la única causa de la diferencia entre las medias muestrales.



PRUEBA z PARA GRUPOS INDEPENDIENTES

Antes de analizar la prueba *t* para grupos independientes, presentaremos la prueba *z*. En el caso de dos grupos, casi nunca se utiliza la prueba *z*, pues para usarla hay que conocer σ_1^2 o σ_2^2 . Sin embargo, esta prueba proporciona una importante base conceptual para comprender la prueba *t*. Después de presentar la prueba *z*, regresaremos a la prueba *t*.

Comencemos con un experimento.

un experimento

Hormona X y comportamiento sexual

Un fisiólogo piensa que la hormona X es importante para generar comportamiento sexual. Para probar esta hipótesis, el investigador elige una muestra aleatoria de 20 ratas macho y las asigna al azar a dos grupos. Los animales del grupo 1 son inyectados con la hormona X y luego se les coloca en un espacio individual junto con una hembra sexualmente receptiva. Los animales del grupo 2 reciben un tratamiento similar, salvo que se les inyecta una solución placebo. Se

*En este caso habría dos poblaciones de la hipótesis nula; una con una media μ_1 y una desviación estándar σ_1 y la otra con una media μ_2 y una desviación estándar σ_2 . Sin embargo, como $\mu_1 = \mu_2$ y $\sigma_1 = \sigma_2$, las poblaciones serían idénticas.

$$\begin{aligned}
 t_{\text{obt}} &= \frac{\bar{D}_{\text{obt}} - \mu_D}{\sqrt{\frac{SC_D}{N(N-1)}}} & SC_D &= \Sigma D^2 - \frac{(\Sigma D)^2}{N} \\
 &= \frac{2.917 - 0}{\sqrt{\frac{132.917}{12(11)}}} & &= 235 - \frac{(35)^2}{12} \\
 &= 2.91 & &= 132.917
 \end{aligned}$$

PASO 2: Evaluación del estadístico. Grados de libertad = $N - 1 = 12 - 1 = 11$. Según la tabla D, con $\alpha = 0.05_{2 \text{ colas}}$ y 11 grados de libertad,

$$t_{\text{crit}} = \pm 2.201$$

En virtud de que $|t_{\text{obt}}| > 2.201$, rechazamos H_0 . La campaña de ahorro influye en la cantidad de gasolina consumida. Según parece, disminuye el consumo de ese combustible.

Resolvamos otro problema de práctica.



En fecha reciente fue descubierto un neurotransmisor cerebral endógeno llamado galanina. Según parece, éste afecta de manera directa el deseo de ingerir alimentos con un alto contenido de grasa. Mientras más alta sea la cantidad de este neurotransmisor de origen natural en un individuo, mayor será el apetito que éste sienta por la comida con alto contenido de grasa. Recientemente, una compañía farmacéutica desarrolló una sustancia experimental que bloquea la galanina sin alterar el apetito por otros alimentos más saludables (es decir, con menos grasa). Un neurocientífico de la compañía piensa que esa sustancia experimental será muy útil para controlar la obesidad. En forma específica, él cree que la administración diaria de este medicamento hará que la persona ingiera alimentos con menos grasa y, por ende, promoverá la pérdida de peso. En estas condiciones, se realiza un experimento, para el cual se elige de manera aleatoria a 15 mujeres obesas, todas ellas voluntarias, y se les administra el medicamento experimental durante 6 meses. Se registra el peso inicial y el peso final (después de 6 meses) de cada persona. Los pesos registrados se presentan en la siguiente tabla.

Número de sujeto	Peso inicial (libras)	Peso final (libras)	<i>D</i>	<i>D</i> ²
1	165	145	20	400
2	143	137	6	36
3	175	170	5	25

Número de sujeto	Peso inicial (libras)	Peso final (libras)	D	D ²
4	135	136	-1	1
5	148	141	7	49
6	155	138	17	289
7	158	137	21	441
8	140	125	15	225
9	172	161	11	121
10	164	156	8	64
11	178	165	13	169
12	182	170	12	144
13	190	176	14	196
14	169	154	15	225
15	157	143	14	196
			<u>14</u>	<u>196</u>
			177	2581

$$N = 15 \quad \bar{D}_{\text{obt}} = \frac{\sum D}{N} = \frac{177}{15} = 11.80$$

- ¿Cuál es la hipótesis alternativa direccional?
- ¿Cuál es la hipótesis nula?
- ¿Cuál es la conclusión? utilice $\alpha = 0.05_{1 \text{ cola}}$.

SOLUCIÓN

- Hipótesis alternativa: El uso del medicamento experimental produce pérdida de peso. La muestra con $\bar{D}_{\text{obt}} = 11.80$ es una muestra aleatoria, extraída de una población de puntajes de diferencia donde $\mu_D > 0$.
- Hipótesis nula: El uso del medicamento experimental no influye ni en la pérdida ni en el aumento de peso. La muestra con $\bar{D}_{\text{obt}} = 11.80$ es aleatoria y fue extraída de una población de puntajes de diferencia con $\mu_D \leq 0$.
- Conclusión, utilizando $\alpha = 0.05_{1 \text{ cola}}$.

PASO 1: Cálculo del estadístico apropiado. Los puntajes de diferencia se presentan en la tabla. Hemos restado los “pesos finales” de los “pesos iniciales”. Si se cumplen los supuestos acerca de t , el estadístico apropiado es t_{obt} . Según la tabla de datos, $N = 15$, $\bar{D}_{\text{obt}} = 11.80$.

$$t_{\text{obt}} = \frac{\bar{D}_{\text{obt}} - \mu_D}{\sqrt{\frac{SC_D}{N(N-1)}}} \quad SC_D = \sum D^2 - \frac{(\sum D)^2}{N}$$

(continúa)

$$\begin{aligned}
 &= \frac{11.80 - 0}{\sqrt{\frac{492.40}{15(14)}}} &= 2581 - \frac{(177)^2}{15} \\
 &= 7.71 &= 492.40
 \end{aligned}$$

PASO 2: Evaluación del estadístico. Grados de libertad = $N - 1 = 14$. Según la tabla D, con $\alpha = 0.05_{1 \text{ cola}}$ y 14 gl,

$$t_{\text{crit}} = 1.761$$

En virtud de que $|t_{\text{obt}}| > 1.761$, rechazamos H_0 . Tal parece que la administración del medicamento experimental produce una pérdida de peso.

Comparación de la prueba t para grupos correlacionados y la prueba de los signos

Podríamos haber resuelto los dos problemas anteriores por medio de la prueba de los signos: sin embargo, en lugar de eso, elegimos la prueba t porque es más sensible. A fin de ilustrar este punto, aplicaremos la prueba de los signos para resolver el problema referente al ahorro de gasolina.

SOLUCIÓN

PASO 1: Cálculo del estadístico. Existen 8 signos positivos en la muestra.

PASO 2: Evaluación del estadístico. Con $N = 12$, $P = 0.50$ y $\alpha = 0.05_{2 \text{ colas}}$,

$$p(8 \text{ o más signos positivos}) = p(8) + p(9) + p(10) + p(11) + p(12)$$

Según la tabla B del apéndice D,

$$\begin{aligned}
 p(8 \text{ o más signos positivos}) &= 0.1208 + 0.0537 + 0.0161 + 0.0029 + 0.0002 \\
 &= 0.1937
 \end{aligned}$$

Como el nivel alfa tiene dos colas,

$$\begin{aligned}
 p(\text{resultado tan extremo como 8 signos positivos}) &= 2[p(8 \text{ o más signos positivos})] \\
 &= 2(0.1937) \\
 &= 0.3874
 \end{aligned}$$

Puesto que $0.3874 > 0.05$, la conclusión es conservar H_0 .

No podemos rechazar H_0 con la prueba de los signos, pero lo pudimos rechazar con la prueba t . ¿Significa esto que la campaña resulta eficaz si analizamos los datos con la prueba t y que no es eficaz si los analizamos con la prueba de los signos? Desde luego que no. A causa de la baja potencia que posee de la prueba de los signos, en ella hay una alta probabilidad de cometer un error de tipo II (es decir, conservar H_0 cuando ésta es falsa). Por lo general, la prueba t es más potente que la prueba de los signos. La potencia adicional hace que H_0 tenga una mayor probabilidad de ser rechazada cuando es falsa. En este caso, la potencia adicional hizo posible el rechazo de H_0 . Cuando varias pruebas son apropiadas para analizar los datos, *una regla general del análisis estadístico aconseja utilizar la más potente*, pues ésta es la que ofrece la más alta probabilidad de rechazar H_0 cuando sea falsa.

Supuestos subyacentes a la prueba t para grupos correlacionados

Los supuestos son muy similares a los que subyacen a la prueba t para una sola muestra. La prueba t para grupos correlacionados requiere que la distribución muestral de \bar{D} sea una distribución normal. Esto significa que N debe ser igual o mayor que 30; suponiendo que la forma de la población no difiere mucho de la normal, o bien que los propios datos poblacionales están distribuidos de manera normal.*



PRUEBAS z Y t PARA GRUPOS INDEPENDIENTES

Diseño de grupos independientes

Dos diseños básicos son los que se usan con más frecuencia en el estudio del comportamiento. El primero fue presentado cuando analizamos la prueba de los signos y la prueba t para grupos correlacionados. Su nombre es diseño de medidas repetidas o replicadas. La forma más sencilla de este diseño utiliza dos condiciones: una experimental y una de control. La característica esencial del diseño es que usa datos apareados entre ambas condiciones y los puntajes de *diferencia* de cada par de datos son analizados para determinar si el azar, por sí solo, puede explicarlos razonablemente.

El otro tipo de diseño se conoce como *diseño de grupos independientes*. Igual que el diseño de grupos correlacionados, el diseño de grupos independientes incluye experimentos en los cuales se usan dos o más condiciones. En cada condición se emplea un nivel diferente de la variable independiente. El más fundamental de esos experimentos tiene únicamente dos condiciones: la condición experimental y la condición control. En este capítulo vamos a considerar el experimento básico, en el cual intervienen sólo dos condiciones. Otros experimentos más complicados serán considerados en el capítulo 15.

En el diseño de grupos independientes, los sujetos son seleccionados aleatoriamente a partir de la población correspondiente, y luego se les asigna, también en forma aleatoria, ya sea a la condición experimental o a la control. No hay base alguna para formar las parejas de sujetos y cada uno de éstos es sometido a prueba sólo una vez. Todos los sujetos de la condición experimental reciben el nivel apropiado de la variable independiente para la condición experimental, y al conjunto de esos sujetos se le conoce como el "grupo experimental". Todos los sujetos de la condición control reciben el nivel de la variable independiente que se considere apropiado para la condición control y se dice que pertenecen al "grupo control".

En virtud de que los sujetos son asignados aleatoriamente a una u otra condición, cuando se analizan los datos no existe base alguna para formar las parejas de datos entre las dos condiciones. En lugar de eso, se calcula un estadístico para cada grupo por *separado* y luego los estadísticos de los dos grupos son comparados para determinar si el azar, por sí solo, es una explicación razonable de las diferencias observadas entre los estadísticos de los dos grupos. El estadístico que se calcula para cada muestra depende de la prueba de inferencia que se utilice. En la prueba t para grupos independientes se calcula la media de cada muestra y luego se analiza la diferencia entre esas dos medias muestrales para determinar si el azar, por sí solo, es una explicación razonable de la diferencia observada entre ambas medias.

* Muchos autores limitan el uso de la prueba t a los datos con escala de intervalo o razón. Véase la nota al pie de la página 31 del capítulo 2, para consultar otras referencias al analizar este punto.

H_1 y H_0 Los datos muestrales, en una de las condiciones (digamos, la condición 1), pueden considerarse como una muestra aleatoria extraída de una población formada por datos que se distribuyen de manera normal: esta población sería resultado de que todos los individuos de la población estuvieran en esa condición (la condición 1). Llame-mos μ_1 a la media de esta población hipotética y σ_1 a su desviación estándar. De igual manera, los datos muestrales de la condición 2 pueden considerarse como una muestra aleatoria extraída de una población formada por datos distribuidos normalmente, la cual es resultado de que todos los individuos estuvieran bajo la condición 2. Podemos llamar μ_2 a la media de esta segunda población y σ_2 a su desviación estándar. Así,

μ_1 = media de la población que está en la condición 1

σ_1 = desviación estándar de la población que está en la condición 1

μ_2 = media de la población que está en la condición 2

σ_2 = desviación estándar de la población que está en la condición 2

Se supone que el hecho de modificar el nivel de la variable independiente afecta a la media de la distribución (μ_2), pero no a la desviación estándar (σ_2) ni a la varianza (σ_2^2). Por lo tanto, según esta suposición, si la variable independiente tiene un efecto real, entonces las medias de las poblaciones serán distintas, pero sus varianzas serán las mismas. Así, suponemos que σ_1^2 es igual a σ_2^2 . Un caso en el que se cumple este supuesto es cuando la variable independiente tiene el mismo efecto sobre cada individuo. Una hipótesis alternativa direccional predice que las muestras son muestras aleatorias seleccionadas de poblaciones en las cuales $\mu_1 > \mu_2$ o $\mu_1 < \mu_2$, según la dirección del efecto. Una hipótesis alternativa no direccional predice que $\mu_1 \neq \mu_2$. Si la variable independiente no tiene efecto alguno, las muestras serían aleatorias extraídas de poblaciones en las cuales $\mu_1 = \mu_2^*$ y el azar sería la única causa de la diferencia entre las medias muestrales.



PRUEBA z PARA GRUPOS INDEPENDIENTES

Antes de analizar la prueba *t* para grupos independientes, presentaremos la prueba *z*. En el caso de dos grupos, casi nunca se utiliza la prueba *z*, pues para usarla hay que conocer σ_1^2 o σ_2^2 . Sin embargo, esta prueba proporciona una importante base conceptual para comprender la prueba *t*. Después de presentar la prueba *z*, regresaremos a la prueba *t*.

Comencemos con un experimento.

un experimento

Hormona X y comportamiento sexual

Un fisiólogo piensa que la hormona X es importante para generar comportamiento sexual. Para probar esta hipótesis, el investigador elige una muestra aleatoria de 20 ratas macho y las asigna al azar a dos grupos. Los animales del grupo 1 son inyectados con la hormona X y luego se les coloca en un espacio individual junto con una hembra sexualmente receptiva. Los animales del grupo 2 reciben un tratamiento similar, salvo que se les inyecta una solución placebo. Se

*En este caso habría dos poblaciones de la hipótesis nula: una con una media μ_1 y una desviación estándar σ_1 y la otra con una media μ_2 y una desviación estándar σ_2 . Sin embargo, como $\mu_1 = \mu_2$ y $\sigma_1 = \sigma_2$, las poblaciones serían idénticas.

tabla 14.3 Datos del experimento de la hormona X y el comportamiento sexual

Hormona X, Grupo 1	Placebo, Grupo 2
8	5
10	6
12	3
6	4
6	7
7	8
9	6
8	5
7	4
<u>11</u>	<u>8</u>
84	56
$n_1 = 10$	$n_2 = 10$
$\bar{X}_1 = 8.4$	$\bar{X}_2 = 5.6$
$\bar{X}_1 - \bar{X}_2 = 2.8$	

cuenta entonces el número de apareamientos producidos durante un periodo de 20 minutos y se obtienen los siguientes resultados.

Como muestra la tabla 14.3, la media del grupo 1 es mayor que la del grupo 2. La diferencia entre las medias de las dos muestras es 2.8 y está en la dirección que indica un efecto positivo. ¿Es válido concluir que la hormona X fue la causa de la diferencia registrada entre las medias? Por supuesto, la respuesta es no. Antes de sacar esta conclusión, debemos evaluar la explicación de la hipótesis nula. El estadístico que utilizamos para esta evaluación es la diferencia entre las medias de las dos muestras. Como ocurre con los demás estadísticos, debemos conocer su distribución muestral antes de poder evaluar la hipótesis nula.

Distribución muestral de la diferencia entre las medias muestrales ($\bar{X}_1 - \bar{X}_2$)

Al igual que la distribución muestral de la media, esta distribución muestral se puede determinar en forma teórica o empírica. Una vez más, para fines pedagógicos, preferimos el punto de vista empírico. Para deducir de manera empírica la distribución muestral de $\bar{X}_1 - \bar{X}_2$, extraeríamos todas las diferentes muestras posibles de tamaño n_1 , de una población con media μ_1 y varianza σ_1^2 . De la misma manera, extraeríamos todas las diferentes muestras posibles de tamaño n_2 de otra población con media μ_2 y varianza σ_2^2 . Luego calcularíamos los valores de \bar{X}_1 y \bar{X}_2 para cada muestra. A continuación, determinaríamos $\bar{X}_1 - \bar{X}_2$ para todos los posibles pares de muestras de tamaños n_1 y n_2 . La distribución resultante contendría todos los distintos datos $\bar{X}_1 - \bar{X}_2$ que podrían obtenerse a partir de las poblaciones, cuando los tamaños de muestra son n_1 y n_2 . Una vez conseguida esta distribución, es sencillo calcular la probabilidad de obtener cada diferencia entre las medias ($\bar{X}_1 - \bar{X}_2$), suponiendo que sea un muestreo aleatorio de los datos n_1 y n_2 extraídos de sus respectivas poblaciones. Ésta sería entonces la distribución muestral de la diferencia entre las medias muestrales para las muestras de n_1 y n_2 tomadas de sus poblaciones específicas. Este proceso se repite para los diferentes tamaños de muestra y datos poblacionales. Ya sea que se determine en forma teórica o empírica, la distribución muestral de la diferencia entre las medias muestrales tiene las siguientes características:

1. Si las poblaciones de donde se extraen las muestras son normales, entonces la distribución de la diferencia entre las medias muestrales también es normal.

$$2. \mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2$$

donde $\mu_{\bar{X}_1 - \bar{X}_2}$ = media de la distribución de la diferencia entre las medias muestrales

$$3. \sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2}$$

donde $\sigma_{\bar{X}_1 - \bar{X}_2}$ = desviación estándar de la distribución muestral de la diferencia entre las medias muestrales; o bien, error estándar de la diferencia entre las medias muestrales

$\sigma_{\bar{X}_1}^2$ = varianza de la distribución muestral de la media para muestras de tamaño n_1 extraídas de la primera población

$\sigma_{\bar{X}_2}^2$ = varianza de la distribución muestral de la media para muestras de tamaño n_2 extraídas de la segunda población

Si, como mencionamos antes, suponemos que las varianzas de las dos poblaciones son iguales ($\sigma_1^2 = \sigma_2^2$), entonces la ecuación para $\sigma_{\bar{X}_1 - \bar{X}_2}$ se puede simplificar como sigue:

$$\begin{aligned}\sigma_{\bar{X}_1 - \bar{X}_2} &= \sqrt{\sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2} \\ &= \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \\ &= \sqrt{\sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}\end{aligned}$$

donde $\sigma^2 = \sigma_1^2 = \sigma_2^2 =$ varianza de cada población.

La distribución se presenta en la figura 14.2. Ahora regresemos a nuestro ejemplo.

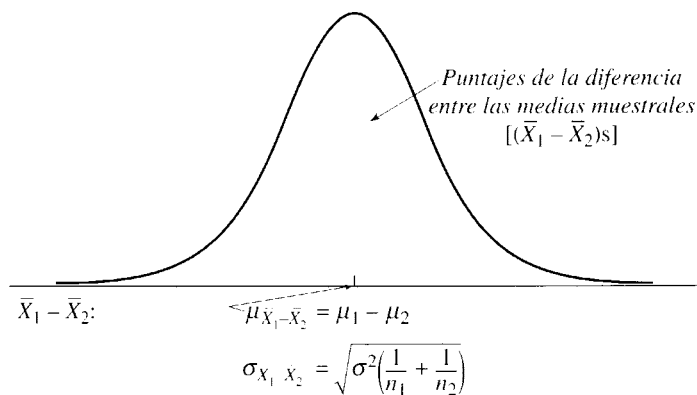


figura 14.2 Distribución muestral de la diferencia entre las medias muestrales.

un experimento

Revisión del experimento de la hormona X

Los resultados del experimento mostraron que 10 ratas inyectadas con la hormona X tuvieron una media de 8.4 apareamientos, mientras que la media de las 10 ratas inyectadas con un placebo fue de 5.6. ¿Es significativa la diferencia ($\bar{X}_1 - \bar{X}_2 = 2.8$) entre las medias? Utilice $\alpha = 0.05$, colas.

SOLUCIÓN

La distribución muestral de $\bar{X}_1 - \bar{X}_2$ se presenta en la figura 14.3. El área sombreada contiene todos los puntajes de diferencia entre las medias que son mayores o iguales que +2.8. Si la distribución muestral de $\bar{X}_1 - \bar{X}_2$ es normal y si la diferencia entre las medias muestrales (2.8) se puede convertir a un valor z , entonces podemos utilizar la prueba z para resolver el problema. La ecuación para z_{obt} es similar a las otras ecuaciones z que ya hemos considerado. Sin embargo, en este caso el valor que estamos convirtiendo es $\bar{X}_1 - \bar{X}_2$. Por lo tanto,

$$z_{\text{obt}} = \frac{(\bar{X}_1 - \bar{X}_2) - \mu_{\bar{X}_1 - \bar{X}_2}}{\sigma_{\bar{X}_1 - \bar{X}_2}} \quad \text{ecuación para } z_{\text{obt}} \text{ en el caso del diseño de grupos independientes}$$

Si la hormona X no tiene efecto alguno sobre el comportamiento de apareamiento, entonces ambas muestras son aleatorias extraídas de poblaciones donde $\mu_1 = \mu_2$ y $\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2 = 0$. Por consiguiente,

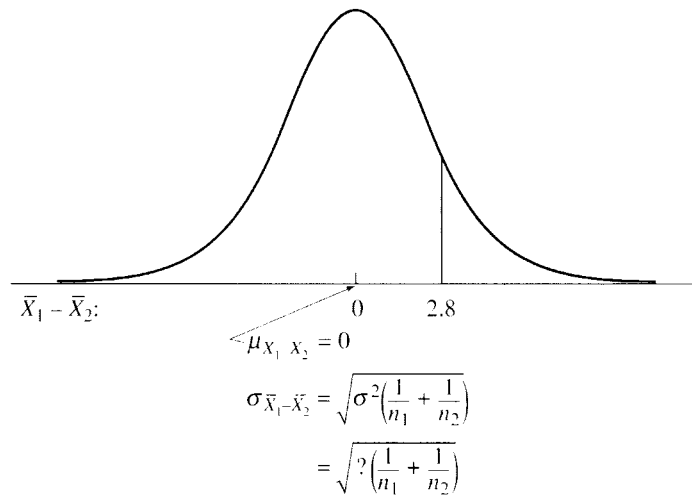


figura 14.3 Distribución muestral de la diferencia entre las medias muestrales para el problema de la hormona.

$$z_{\text{obt}} = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{\sqrt{\sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{2.8}{\sqrt{\sigma^2\left(\frac{1}{10} + \frac{1}{10}\right)}}$$

Observe que para calcular z_{obt} , primero debemos conocer la varianza de las poblaciones (σ^2). Como casi nunca se conoce σ^2 esta limitación restringe severamente la utilidad práctica de la prueba z para este diseño. Sin embargo, como podrá usted suponer, σ^2 se puede estimar a partir de los datos muestrales. Cuando se hace esto, tenemos la prueba t para grupos independientes.



PRUEBA t DE STUDENT PARA GRUPOS INDEPENDIENTES

Comparación de las ecuaciones para z_{obt} y t_{obt}

Las ecuaciones para las pruebas z y t se presentan en la tabla 14.4. Las ecuaciones z y t son idénticas, salvo por el hecho de que la ecuación t usa s_w^2 para estimar la varianza de la población (σ^2). Esta situación es similar al caso de la prueba t para una sola muestra.

tabla 14.4 Comparación entre las ecuaciones z y t

Prueba z	Prueba t
$z_{\text{obt}} = \frac{(\bar{X}_1 - \bar{X}_2) - \mu_{\bar{X}_1 - \bar{X}_2}}{\sigma_{\bar{X}_1 - \bar{X}_2}}$	$t_{\text{obt}} = \frac{(\bar{X}_1 - \bar{X}_2) - \mu_{\bar{X}_1 - \bar{X}_2}}{s_{\bar{X}_1 - \bar{X}_2}}$
$= \frac{(\bar{X}_1 - \bar{X}_2) - \mu_{\bar{X}_1 - \bar{X}_2}}{\sqrt{\sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$	$= \frac{(\bar{X}_1 - \bar{X}_2) - \mu_{\bar{X}_1 - \bar{X}_2}}{\sqrt{s_w^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

donde s_w^2 = estimación ponderada de σ^2
 $s_{\bar{X}_1 - \bar{X}_2}$ = estimación de $\sigma_{\bar{X}_1 - \bar{X}_2}$ = error estándar estimado de la diferencia entre las medias muestrales

Usted recordará que en esa ocasión empleamos la(s) desviación(es) estándar muestral(es) para estimar σ . Sin embargo, en la prueba *t* para grupos independientes existen dos muestras y nos interesa estimar σ^2 . Como s es una estimación precisa de σ , s^2 también es una estimación precisa de σ^2 . Existen dos muestras y podríamos utilizar cualquiera de ellas para estimar σ^2 , sin embargo, podemos obtener una estimación más precisa si utilizamos ambas muestras. La estimación más precisa de σ^2 se obtiene utilizando un promedio ponderado de las varianzas muestrales s_1^2 y s_2^2 . Utilizamos los grados de libertad como pesos en la ponderación. De esta manera,

$$s_w^2 = \frac{gl_1 s_1^2 + gl_2 s_2^2}{gl_1 + gl_2} = \frac{(n_1 - 1) \left(\frac{SC_1}{n_1 - 1} \right) + (n_2 - 1) \left(\frac{SC_2}{n_2 - 1} \right)}{(n_1 - 1) + (n_2 - 1)} = \frac{SC_1 + SC_2}{n_1 + n_2 - 2}$$

donde s_w^2 = estimación ponderada de σ^2
 s_1^2 = varianza de la primera muestra
 s_2^2 = varianza de la segunda muestra
 SC_1 = suma de cuadrados de la primera muestra
 SC_2 = suma de cuadrados de la segunda muestra

Al sustituir s_w^2 en la ecuación *t*, obtenemos la ecuación que nos permite calcular t_{obt} . Así pues,

$$t_{\text{obt}} = \frac{(\bar{X}_1 - \bar{X}_2) - \mu_{\bar{X}_1 - \bar{X}_2}}{\sqrt{s_w^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{(\bar{X}_1 - \bar{X}_2) - \mu_{\bar{X}_1 - \bar{X}_2}}{\sqrt{\left(\frac{SC_1 + SC_2}{n_1 + n_2 - 2} \right) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

*ecuación para el cálculo de t_{obt} ,
diseño de grupos independientes*

Para evaluar la hipótesis nula, suponemos que ambas son muestras aleatorias extraídas de poblaciones que tienen la misma media ($\mu_1 = \mu_2$). Por lo tanto, $\mu_{\bar{X}_1 - \bar{X}_2} = 0$.^{*} La ecuación anterior se reduce a:

$$t_{\text{obt}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{SC_1 + SC_2}{n_1 + n_2 - 2} \right) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

*ecuación para el cálculo de t_{obt} ,
suponiendo que $\mu_{\bar{X}_1 - \bar{X}_2} = 0$*

Podríamos continuar y calcular t_{obt} para el ejemplo de la hormona X, pero para evaluarla debemos conocer la distribución muestral de *t*. Resulta que al deducir la distribución muestral de *t* para grupos independientes, obtenemos la misma familia de curvas que se obtiene con la distribución muestral de *t* para una sola muestra, salvo que el número de grados de libertad es distinto. Como usted recordará, se pierde un grado de libertad por cada vez que se calcula una desviación estándar. Como calculamos s_1^2 y s_2^2 para el caso de dos muestras, perdemos 2 grados de libertad, uno de cada muestra. Así,

$$gl = (n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2 = N - 2$$

donde $N = n_1 + n_2$

^{*}Véase la nota 14.3.

Entonces empleamos la tabla D tal como lo hicimos en el caso de la prueba *t* para una sola muestra, excepto que en la situación de dos muestras utilizamos la tabla con $N = 2$ gl. Así pues, la distribución *t* varía tanto con N como con los grados de libertad, pero varía de manera única sólo con los grados de libertad. Es decir, que la distribución *t* correspondiente a 13 gl es la misma, ya sea que provenga del caso de una sola muestra con $N = 14$ o bien, del caso de dos muestras con $N = 15$.

Análisis del experimento de la hormona X con la prueba *t* para grupos independientes

Ahora podremos evaluar por fin los datos en el caso de la hormona. Para comodidad del lector, volveremos a enunciar el problema y los datos.

Un fisiólogo ha realizado un experimento para evaluar el efecto de la hormona X sobre el comportamiento sexual. Diez ratas fueron inyectadas con la hormona X y otras tantas recibieron una inyección placebo. Se contó el número de aparcamientos registrados durante un periodo de 20 minutos.

Los resultados se presentan en la tabla 14.5.

1. ¿Cuál es la hipótesis alternativa? Utilice una hipótesis no direccional.
2. ¿Cuál es la hipótesis nula?
3. ¿Cuál es su conclusión? Utilice $\alpha = 0.05_{2 \text{ colas}}$.

tabla 14.5 Datos del experimento de la hormona X

Grupo 1, hormona X		Grupo 2, placebo	
X_1	X_1^2	X_2	X_2^2
8	64	5	25
10	100	6	36
12	144	3	9
6	36	4	16
6	36	7	49
7	49	8	64
9	81	6	36
8	64	5	25
7	49	4	16
<u>11</u>	<u>121</u>	<u>8</u>	<u>64</u>
84	744	56	340
	$n_1 = 10$	$n_2 = 10$	
	$\bar{X}_1 = 8.4$	$\bar{X}_2 = 5.6$	
	$\bar{X}_1 - \bar{X}_2 = 2.8$		

SOLUCIÓN

1. Hipótesis alternativa: La hipótesis alternativa especifica que la hormona X afecta el comportamiento sexual. La diferencia de 2.8 entre las medias muestrales se debe al muestreo aleatorio de poblaciones en las que $\mu_1 \neq \mu_2$.

2. Hipótesis nula: La hipótesis nula establece que la hormona X no está relacionada con el comportamiento sexual. La diferencia de 2.8 entre las medias muestrales se debe a un muestreo aleatorio de las poblaciones en las que $\mu_1 = \mu_2$.
3. Conclusión. al utilizar $\alpha = 0.05_{2 \text{ colas}}$:

PASO 1: Cálculo del estadístico apropiado. Por el momento, supongamos que t es apropiada. En una sección posterior comentaremos los supuestos de t . Según la tabla 14.5, $n_1 = 10$, $n_2 = 10$, $\bar{X}_1 = 8.4$ y $\bar{X}_2 = 5.6$. Despejando SC_1 y SC_2 ,

$$\begin{aligned} SC_1 &= \sum X_1^2 - \frac{(\sum X_1)^2}{n_1} & SC_2 &= \sum X_2^2 - \frac{(\sum X_2)^2}{n_2} \\ &= 744 - \frac{(84)^2}{10} & &= 340 - \frac{(56)^2}{10} \\ &= 38.4 & &= 26.4 \end{aligned}$$

Sustituyendo estos valores en la ecuación de t_{obt} , tenemos:

$$t_{\text{obt}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{SC_1 + SC_2}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{8.4 - 5.6}{\sqrt{\left(\frac{38.4 + 26.4}{10 + 10 - 2}\right)\left(\frac{1}{10} + \frac{1}{10}\right)}} = 3.30$$

PASO 2: Evaluación del estadístico. Igual que en las pruebas t anteriores, si t_{obt} queda en la región crítica tenemos que rechazar H_0 . Así pues,

Si $|t_{\text{obt}}| \geq |t_{\text{crit}}|$, rechazamos H_0 .

Si no es así, conservamos H_0 .

El número de grados de libertad es $gl = N - 2 = 20 - 2 = 18$. Según la tabla D, con $\alpha = 0.05_{2 \text{ colas}}$ y $gl = 18$,

$$t_{\text{crit}} = \pm 2.101$$

Como $|t_{\text{obt}}| > 2.101$, la conclusión consiste en rechazar H_0 .

Cálculo de t_{obt} cuando $n_1 = n_2$

Cuando los tamaños de las muestras son iguales, podemos simplificar la ecuación de t_{obt} . Por lo tanto,

$$t_{\text{obt}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{SC_1 + SC_2}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

pero $n_1 = n_2 = n$. Sustituyendo n por n_1 y n_2 en la ecuación de t_{obt} ,

$$t_{\text{obt}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{SC_1 + SC_2}{n + n - 2}\right)\left(\frac{1}{n} + \frac{1}{n}\right)}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{SC_1 + SC_2}{2(n - 1)}\right)\left(\frac{2}{n}\right)}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{SC_1 + SC_2}{n(n - 1)}\right)}}$$

Así pues,

$$t_{\text{obt}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{SC_1 + SC_2}{n(n - 1)}}} \quad \text{ecuación para calcular } t_{\text{obt}} \text{ cuando } n_1 = n_2$$

En virtud de que $n_1 = n_2$ en el problema anterior, podemos utilizar la ecuación simplificada para calcular t_{obt} . Así,

$$t_{\text{obt}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{SC_1 + SC_2}{n(n-1)}}} = \frac{8.4 - 5.6}{\sqrt{\frac{38.4 + 26.4}{10(9)}}} = 3.30$$

Éste es el mismo valor de t_{obt} que obtuvimos al utilizar la ecuación más complicada. Siempre que $n_1 = n_2$, resulta más fácil usar la ecuación simplificada. Cuando $n_1 \neq n_2$, se debe utilizar la ecuación más complicada.

Resolvamos otro problema en el que las dos n sean iguales.



Un psicólogo desea determinar si la capacidad de la memoria inmediata resulta afectada por la falta de sueño. La memoria inmediata se define como la cantidad de material que la persona puede recordar poco después de que éste se le presenta. Doce estudiantes del curso de introducción a la psicología son elegidos al azar y se les asigna, de manera aleatoria, a dos grupos formados por seis individuos cada uno. A uno de los grupos se le priva del sueño durante las 24 horas anteriores a la presentación del material para recordar. Todos los sujetos del otro grupo duermen el tiempo normal (de 7 a 8 horas). El material consiste en una serie de diapositivas, cada una de las cuales contiene nueve números. Cada diapositiva se les presenta durante un periodo muy breve (50 milisegundos), después de lo cual el sujeto debe recordar la mayor cantidad posible de los números que acaban de ser presentados. Los resultados se muestran a continuación. Los datos representan el porcentaje de números correctamente recordados.

Sueño normal Grupo 1		Privado del sueño Grupo 2	
X_1	X_1^2	X_2	X_2^2
68	4624	70	4900
73	5329	62	3844
72	5184	68	4624
65	4225	63	3969
70	4900	69	4761
73	5329	60	3600
421	29 591	392	25 698
$n_1 = 6$		$n_2 = 6$	
$\bar{X}_1 = 70.167$		$\bar{X}_2 = 65.333$	
$\bar{X}_1 - \bar{X}_2 = 4.834$			

- ¿Cuál es la hipótesis alternativa? Utilice una hipótesis no direccional.
- ¿Cuál es la hipótesis nula?
- ¿Cuál es la conclusión de usted? Utilice $\alpha = 0.01_{2 \text{ colas}}$.

(continúa)

SOLUCIÓN

- a. Hipótesis alternativa: La hipótesis alternativa sostiene que la privación del sueño influye en la memoria inmediata. La diferencia de 4.834 registrada entre las medias muestrales se debe al muestreo aleatorio de poblaciones en las cuales $\mu_1 \neq \mu_2$.
- b. Hipótesis nula: La hipótesis nula establece que la privación del sueño no influye en la memoria inmediata. La diferencia de 4.834 registrada entre las medias muestrales se debe al muestreo aleatorio de poblaciones en las cuales $\mu_1 = \mu_2$.
- c. Conclusión, utilizando $\alpha = 0.01_{2 \text{ colas}}$:

PASO 1: Cálculo del estadístico apropiado. Si se cumplen los supuestos de t , entonces t_{obt} es el estadístico apropiado. Según la tabla de datos, $n_1 = n_2 = n = 6$, $\bar{X}_1 = 70.167$ y $\bar{X}_2 = 65.333$. Despejando SC_1 , y SC_2 , obtenemos:

$$\begin{aligned} SC_1 &= \sum X_1^2 - \frac{(\sum X_1)^2}{n_1} & SC_2 &= \sum X_2^2 - \frac{(\sum X_2)^2}{n_2} \\ &= 29\,591 - \frac{(421)^2}{6} & &= 25\,698 - \frac{(392)^2}{6} \\ &= 50.833 & &= 87.333 \end{aligned}$$

Sustituyendo estos valores en la ecuación para t_{obt} con valores de n iguales, tenemos:

$$t_{\text{obt}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{SC_1 + SC_2}{n(n-1)}}} = \frac{70.167 - 65.333}{\sqrt{\frac{50.833 + 87.333}{6(5)}}} = 2.25$$

PASO 2: Evaluación del estadístico. Grados de libertad = $N - 2 = 12 - 2 = 10$. Según la tabla D, con $\alpha = 0.01_{2 \text{ colas}}$ y $gl = 10$,

$$t_{\text{crit}} = \pm 3.169$$

Como $|t_{\text{obt}}| \leq 3.169$, conservamos H_0 . No podemos concluir que la privación del sueño influya en la memoria inmediata.

Resolvamos otro problema, esta vez con valores de n distintos.

Un neurocirujano tiene la idea de que las lesiones producidas en cierta área del cerebro, conocida como tálamo, hacen que disminuya la percepción del dolor. Si está en lo cierto, eso podría ser importante para el tratamiento de enfermedades terminales acompañadas de dolor intenso. Como un primer intento de poner a prueba esta hipótesis, el investigador realiza un experimento en el cual 16 ratas son divididas, al azar, en dos grupos de 8 especímenes cada uno. A los animales del grupo experimental se les provoca una pequeña lesión en la parte del tálamo que, según el investigador, está implicada en la percepción del dolor. A los animales del grupo control se les provoca una lesión comparable en un área del cerebro que, según se cree, no está relacionada con el dolor. Dos semanas después de las operaciones quirúrgicas, a cada roedor se le aplica una serie de breves descargas eléctricas en las patas. Esas descargas van creciendo en intensidad: tienen un nivel de intensidad muy bajo y éste aumenta hasta que el animal retrocede. De esta manera se determina el umbral del dolor provocado por las descargas eléctricas en cada rata. Al final se obtienen los siguientes datos. Cada dato representa el nivel más bajo (en miliamperes) en que se produce el retroceso. Mientras más alto sea el nivel de la corriente, tanto más alto será el umbral de dolor. Observe que un animal murió durante la operación quirúrgica y no fue reemplazado.

Lesiones en el área neutral grupo control Grupo 1		Lesiones en el tálamo grupo experimental Grupo 2	
X_1	X_1^2	X_2	X_2^2
0.8	0.64	1.9	3.61
0.7	0.49	1.8	3.24
1.2	1.44	1.6	2.56
0.5	0.25	1.2	1.44
0.4	0.16	1.0	1.00
0.9	0.81	0.9	0.81
1.4	1.96	1.7	2.89
1.1	1.21	10.1	15.55
7.0	6.96		
$n_1 = 8$		$n_2 = 7$	
$\bar{X}_1 = 0.875$		$\bar{X}_2 = 1.443$	
		$\bar{X}_1 - \bar{X}_2 = 0.568$	

- ¿Cuál es la hipótesis alternativa? En este problema, suponga que existen suficientes fundamentos teóricos y experimentales para utilizar una hipótesis direccional.
- ¿Cuál es la hipótesis nula?
- ¿Cuál es la conclusión de usted? Utilice $\alpha = 0.05_{1\text{ cola}}$.

(continúa)

SOLUCIÓN

- a. Hipótesis alternativa: La hipótesis alternativa sostiene que las lesiones del tálamo hacen que disminuya la percepción del dolor. La diferencia de -0.568 registrada entre las medias muestrales se debe al muestreo aleatorio de poblaciones en las cuales $\mu_1 < \mu_2$.
- b. Hipótesis nula: La hipótesis nula establece que las lesiones del tálamo no tienen efecto alguno o hacen que aumente la percepción del dolor. La diferencia -0.568 registrada entre las medias muestrales se debe al muestreo aleatorio de poblaciones en las cuales $\mu_1 \geq \mu_2$.
- c. Conclusión, utilizando $\alpha = 0.05_{1\text{ cola}}$:

PASO 1: Cálculo del estadístico apropiado. Suponiendo que los supuestos de t se hayan cumplido, entonces t_{obt} será el estadístico apropiado. Según la tabla de datos, $n_1 = 8$, $n_2 = 7$, $\bar{X}_1 = 0.875$ y $\bar{X}_2 = 1.443$. Despejando SC_1 y SC_2 obtenemos:

$$\begin{aligned} SC_1 &= \sum X_1^2 - \frac{(\sum X_1)^2}{n_1} & SC_2 &= \sum X_2^2 - \frac{(\sum X_2)^2}{n_2} \\ &= 6.960 - \frac{(7)^2}{8} & &= 15.550 - \frac{(10.1)^2}{7} \\ &= 0.835 & &= 0.977 \end{aligned}$$

Sustituyendo estos valores en la ecuación general para t_{obt} , tenemos:

$$\begin{aligned} t_{\text{obt}} &= \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{SC_1 + SC_2}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \\ &= \frac{0.875 - 1.443}{\sqrt{\left(\frac{0.835 + 0.977}{8 + 7 - 2}\right)\left(\frac{1}{8} + \frac{1}{7}\right)}} = -2.94 \end{aligned}$$

PASO 2: Evaluación del estadístico. Grados de libertad = $N - 2 = 15 - 2 = 13$. De la tabla D, con $\alpha = 0.05_{1\text{ cola}}$ y $gl = 13$,

$$t_{\text{crit}} = -1.771$$

Como $|t_{\text{obt}}| > 1.771$, rechazamos H_0 y concluimos que las lesiones del tálamo disminuyen la percepción del dolor.

Supuestos subyacentes a la prueba t para grupos independientes

Los supuestos subyacentes a la prueba t para grupos independientes son los siguientes:

1. La distribución muestral de $\bar{X}_1 - \bar{X}_2$ es normal. Esto significa que las poblaciones de donde fueron extraídas las muestras deben tener una distribución normal.
2. Existe homogeneidad de la varianza. Usted recordará que, al principio de nuestro análisis de la prueba t para grupos independientes, señalamos que la prueba t supone que la variable independiente afecta a las medias de las poblaciones, pero no a sus desviaciones estándar ($\sigma_1 = \sigma_2$). Como la varianza no es más que el

cuadrado de la desviación estándar, la prueba *t* para grupos independientes supone también que las varianzas de las dos poblaciones son iguales, es decir, $\sigma_1^2 = \sigma_2^2$. Esto se conoce como el *supuesto de la homogeneidad de la varianza*. Si las varianzas de las muestras del experimento (s_1^2 y s_2^2) son muy diferentes (por ejemplo, si una de las varianzas es 4 o más veces mayor que la otra), es probable que las dos muestras no sean muestras aleatorias extraídas de poblaciones en las cuales $\sigma_1^2 = \sigma_2^2$. Si esto es cierto, entonces no se cumple el supuesto de homogeneidad de la varianza ($\sigma_1^2 \neq \sigma_2^2$).*

Violación de los supuestos de la prueba *t* para grupos independientes

Se han realizado experimentos para determinar el efecto que la violación de los supuestos de normalidad en las poblaciones de datos en bruto y de homogeneidad de la varianza, produce sobre la prueba *t* para grupos independientes. Por fortuna, la prueba *t* es *robusta*. Decimos que *una prueba es robusta si es relativamente insensible a la violación de los supuestos matemáticos subyacentes*. Por lo tanto, la prueba *t* es relativamente insensible a la violación de la normalidad y de la homogeneidad de la varianza, dependiendo del tamaño de la muestra, y del tipo y magnitud de la violación de estos supuestos.[†] Si $n_1 = n_2$ y si el tamaño de cada muestra es igual o mayor que 30, entonces podemos emplear la prueba *t* para grupos independientes, sin un error apreciable, a pesar de que sean infringidos de manera moderada los supuestos de normalidad y/o la homogeneidad de la varianza. Si existe una violación excesiva de estos supuestos, entonces debe usarse una prueba alternativa, como la prueba *U* de Mann-Whitney. Esta prueba será analizada en el capítulo 18.

Antes de concluir este tema, vale la pena observar que cuando las dos muestras exhiben una diferencia muy grande en sus varianzas, esto puede indicar que la variable independiente no produce el mismo efecto en todos los sujetos que están en la misma condición. Este hecho puede representar un hallazgo importante por derecho propio, porque nos conduce a nuevos experimentos que permiten observar la forma en la cual los efectos de la variable independiente se modifican en los diferentes tipos de sujetos.

Magnitud del efecto

Hasta ahora hemos comentado la prueba *t* para grupos independientes y mostramos la forma de usarla para determinar si la variable independiente que nos interesa en un experimento básico con grupos independientes produce un efecto real sobre la variable dependiente que se va a medir. Para determinar si los datos presentan un efecto real, calculamos t_{obt} ; si t_{obt} es significativa, concluimos que la variable independiente produce un efecto real. Por supuesto, esto nos aporta información muy importante. Nos permite sustentar y delinear mejor las teorías en las que intervienen las dos variables, y también nos proporciona información que puede tener consecuencias prácticas importantes.

*Existen muchas pruebas de inferencia para determinar si los datos cumplen los supuestos de homogeneidad de la varianza. Sin embargo, este tema queda fuera del alcance del presente libro. Véase R. F. Kirk, *Experimental Design*, tercera edición, Brooks/Cole, Pacific Grove, CA, 1995, pp. 100-103.

†Algunos estadísticos exigen, además, que los datos estén en una escala de intervalo o de razón para utilizar la prueba *z*, la prueba *t* de Student y el análisis de varianza (que consideraremos en el capítulo 15). Para una exposición sobre este punto, véanse las referencias en la nota al pie de la página 31 del capítulo 2.

‡Para un repaso de este tema, véase C. A. Boneau, "The Effects of Violations of Assumptions Underlying the *t* Test". *Psychological Bulletin*, 57 (1960), pp. 49-64.

Además de determinar si existe un efecto real, a menudo es conveniente calcular la magnitud de dicho efecto. Por ejemplo, en el experimento encaminado a investigar la función del tálamo en la percepción del dolor (página 337), el valor de t_{obt} fue significativo y pudimos sacar la conclusión de que las lesiones del tálamo reducen la percepción del dolor. Así pues, dentro de los límites de un posible error de tipo I, el resultado de este experimento nos permitió concluir que el tálamo desempeña un papel en la percepción del dolor. Sin embargo, sin duda sería deseable saber también qué tan grande es el papel que le corresponde al tálamo. ¿Acaso él controla por completo la percepción del dolor, al grado que si los núcleos relevantes del tálamo fueran destruidos por completo, el sujeto ya no podría percibir el dolor? Por otra parte, ¿es su efecto tan pequeño que pueda ser pasado por alto desde cualquier punto de vista práctico? Después de todo, hasta los efectos pequeños tienen probabilidades de ser significativos si se elige N suficientemente grande. La determinación de la magnitud del efecto del tálamo sería importante, en particular, para el neurocirujano que realiza esta investigación con la esperanza de crear un tratamiento que permita reducir los intensos dolores que sufren algunos pacientes terminales.

Ya hablamos de la magnitud del efecto de la variable X sobre la variable Y en conjunción con la investigación de correlación, cuando analizamos el coeficiente de determinación (r^2) en el capítulo 6, página 120. Usted recordará que si la variable X es de índole causal con respecto a la variable Y , el coeficiente de determinación es una medida de la magnitud del efecto de X sobre Y . Llegamos a esta reflexión porque demostramos que r^2 es una medida de la variabilidad de Y que puede ser atribuida a X .

La situación es muy similar en el caso del experimento de grupos independientes. En el experimento de grupos independientes, la variable independiente es X y la variable dependiente es Y . En la figura 14.4 se presentan las dos situaciones. En la figura 14.4(a), hemos mostrado la gráfica de dispersión que se utilizó para desarrollar la explicación en torno a r^2 en el capítulo 6. Observe que la variable X , “competencia en deletreo” se encuentra dispersa sobre muchos valores de X . En la figura 14.4(b) se presenta una gráfica de los datos correspondientes al experimento sobre el tálamo y la percepción del dolor. En este caso, la variable X , “lesiones”, tiene solamente dos valores: “Lesiones en un área

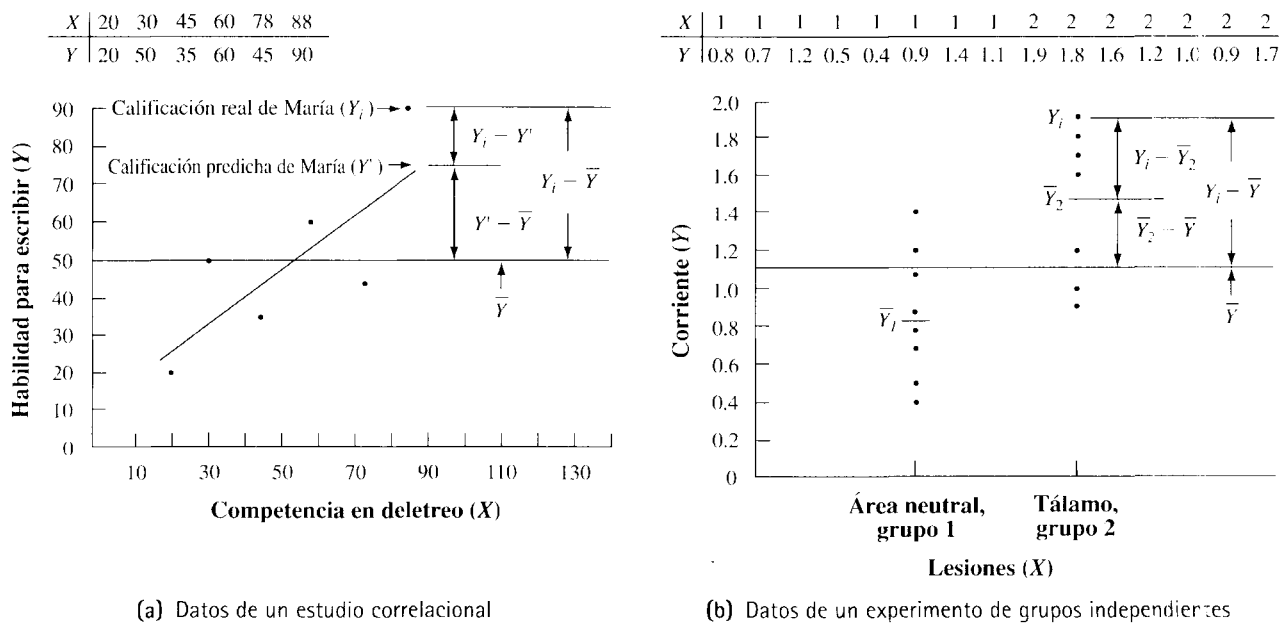


figura 14.4 Gráficas XY de datos tomados de un estudio correlacional y un experimento de grupos independientes.

neutral, grupo 1" y "Lesiones en el tálamo, grupo 2". Por lo tanto, todos los datos del diseño básico de grupos independientes están acumulados en dos grupos.

En los datos del estudio correlacional ilustrados en la figura 14.4(a), la variabilidad total de Y está representada por $Y_i - \bar{Y}$, la variabilidad de Y atribuible a X está representada por $Y' - \bar{Y}$, y el error de predicción está representado por $Y_i - Y'$. \bar{Y} es el promedio de los valores Y de todos los datos, y Y' (el valor que se predijo para Y) está determinado por la línea de regresión que relaciona a X con Y . La situación es esencialmente la misma con los datos de los grupos independientes que se presentan en la figura 14.4(b), salvo que la recta de regresión ha sido sustituida por la media de cada grupo. Esto se debe a que, sin saber a cuál de los grupos pertenece un punto, habríamos predicho \bar{Y} para cada punto, donde \bar{Y} es el promedio de los valores Y de todos los puntos en ambos grupos. Sin embargo, si supiéramos a cuál de los grupos pertenece un punto (su valor X), habríamos predicho la media de su grupo (\bar{Y}_1 o \bar{Y}_2) para el valor Y de dicho punto. De esta manera, la variabilidad total de Y está representada por $Y_i - \bar{Y}$, la variabilidad de Y atribuible a la pertenencia a un grupo determinado (por ejemplo, a la variable independiente) está representada por $\bar{Y}_2 - \bar{Y}$, y la variabilidad de Y que aún no puede atribuirse después de conocer cuál es la pertenencia a los grupos, está representada por $Y_i - \bar{Y}_2$.

En ambas situaciones, la proporción de la variabilidad total de Y que es atribuible a X es una medida de la magnitud del efecto de X sobre Y . Recuerde que en el experimento de grupos independientes, X es la variable independiente. Para diseños de grupos independientes, *omega al cuadrado* (ω^2) es un estadístico que representa un estimado de la proporción de la variabilidad total de Y que es atribuible a X , la variable independiente. Por lo tanto, podemos usar ω^2 como medida de la magnitud del efecto en los experimentos donde se usa el diseño de grupos independientes. El valor de omega al cuadrado se estima mediante la siguiente ecuación:

$$\omega^2 = \frac{t_{\text{obt}}^2 - 1}{t_{\text{obt}}^2 + N - 1} \quad \text{ecuación de cálculo para estimar omega al cuadrado}$$

donde ω^2 = valor estimado de omega al cuadrado

t_{obt} = valor t obtenido para los datos en el experimento de grupos independientes

N = número total de sujetos en ambos grupos = $n_1 + n_2$

Para los datos del experimento sobre el tálamo y la percepción del dolor, $t_{\text{obt}} = -2.94$, $n_1 = 8$, $n_2 = 7$. Por lo tanto,

$$\omega^2 = \frac{(-2.94)^2 - 1}{(-2.94)^2 + 15 - 1} = 0.208$$

Así pues, a las lesiones del tálamo pudo atribuirse cerca de 20.8% de la variabilidad observada en el umbral del dolor frente a las descargas eléctricas. Esto se podría considerar como un efecto entre pequeño y moderado. Si se deseaba encontrar un efecto más grande, una opción podría haber sido acrecentar el tamaño de las lesiones en el tálamo.



POTENCIA DE LA PRUEBA t

Las tres ecuaciones para t_{obt} son las siguientes:

<i>Una sola muestra</i>	<i>Grupos correlacionados</i>	<i>Grupos independientes</i>
$t_{\text{obt}} = \frac{\bar{X}_{\text{obt}} - \mu}{\sqrt{\frac{SC}{N(N-1)}}}$	$t_{\text{obt}} = \frac{\bar{D}_{\text{obt}} - 0}{\sqrt{\frac{SC_D}{N(N-1)}}}$	$t_{\text{obt}} = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{\sqrt{\frac{SC_1 + SC_2}{n(n-1)}}}$

Es obvio que mientras más grande sea t_{obt} tanto más probable será que H_0 sea rechazada. Por lo tanto, cualquier cosa que aumente la posibilidad de obtener valores grandes de t_{obt} producirá una prueba *t* más potente. Esta situación puede presentarse en varias formas. En primer lugar, si el efecto real de la variable independiente es más grande, es más probable que $\bar{X}_{\text{obt}} - \mu$, $\bar{D}_{\text{obt}} - 0$ o $(\bar{X}_1 - \bar{X}_2) - 0$ sea grande. Como las diferencias se presentan en el numerador de la ecuación *t*, esto implica que *mientras más grande sea el efecto de la variable independiente, tanto mayor será la potencia de la prueba t* (manteniendo constantes los demás factores). Por supuesto, antes de realizar el experimento no sabemos cuál es el efecto real de la variable independiente. Si lo supiéramos, ¿qué sentido tendría el experimento? Sin embargo, este análisis es útil porque nos muestra que cuando se diseña un experimento, es recomendable emplear el nivel de la variable independiente que el experimentador considere más eficaz para maximizar la posibilidad de detectar el efecto. Además, este análisis sugiere que aunque se cuente con pocos recursos para llevar a cabo un experimento, éste puede ser lo bastante potente como para detectar el efecto, si la variable independiente produce en realidad un efecto grande.

El denominador de la ecuación *t* cambia en función del tamaño y la variabilidad de la muestra. Cuando el tamaño de la muestra aumenta, el denominador disminuye. Por lo tanto,

$$\sqrt{\frac{SC}{N(N-1)}}, \quad \sqrt{\frac{SC_D}{N(N-1)}}, \quad \text{y} \quad \sqrt{\frac{SC_1 + SC_2}{n(n-1)}}$$

decrecen y, por lo tanto t_{obt} aumenta. *Así pues, al incrementarse el tamaño de la muestra, se incrementa también la potencia de la prueba t.*

El denominador también cambia en función de la variabilidad de la muestra. En el caso de una sola muestra, *SC* es la medida de variabilidad. *SC_D* en los experimentos correspondientes a grupos correlacionados y *SC₁ + SC₂* en los experimentos a base de grupos independientes, reflejan la variabilidad. Cuando aumenta la variabilidad, también el denominador aumenta en cada caso, haciendo que t_{obt} disminuya. *Así pues, una alta variabilidad en la muestra disminuye la potencia.* Por lo tanto, es deseable disminuir la variabilidad lo más posible. Una forma de lograr esto es a través de un control cuidadoso de las condiciones experimentales. Por ejemplo, en un experimento de tiempo de reacción, el investigador podría emplear una señal de advertencia que preceda de manera inmediata al estímulo al cual debe responder el sujeto examinado. De esta forma, podríamos eliminar la variabilidad provocada por los momentos en que decae de atención. Otro método para controlar las condiciones experimentales consiste en usar un diseño experimental adecuado. Por ejemplo, en ciertas situaciones, el empleo de un diseño para grupos correlacionados y no el diseño para grupos independientes disminuirá la variabilidad.



COMPARACIÓN DE LOS DISEÑOS PARA GRUPOS CORRELACIONADOS Y GRUPOS INDEPENDIENTES

Es probable que usted haya observado que muchas de las hipótesis de los ejemplos ilustrativos se habrían podido analizar, ya sea con el diseño de grupos correlacionados o con el de grupos independientes. Por ejemplo, en el problema de práctica 14.1 presentamos un experimento donde se evaluaba el efecto de una campaña para promover el ahorro de gasolina. En el experimento se utilizó un diseño de grupos correlacionados, y los datos fueron analizados con la prueba *t* para grupos correlacionados. Para comodidad del lector, los datos y el análisis se presentan de nuevo en la tabla 14.6.

t a b l a 14.6 Datos y análisis del experimento sobre la campaña para el ahorro de gasolina

Familia	Antes de la campaña (galones) (1)	Después de la campaña (galones) (2)	Diferencia	
			<i>D</i>	<i>D</i> ²
A	55	48	7	49
B	43	38	5	25
C	51	53	-2	4
D	62	58	4	16
E	35	36	-1	1
F	48	42	6	36
G	58	55	3	9
H	45	40	5	25
I	48	49	-1	1
J	54	50	4	16
K	56	58	-2	4
L	32	25	7	49
			<u>35</u>	<u>235</u>

$$N = 12 \quad \bar{D}_{\text{obt}} = \frac{\sum D}{N} = \frac{35}{12} = 2.917$$

$$\begin{aligned}
 t_{\text{obt}} &= \frac{\bar{D}_{\text{obt}} - \mu_D}{\sqrt{\frac{SC_D}{N(N-1)}}} & SC_D &= \sum D^2 - \frac{(\sum D)^2}{N} \\
 &= \frac{2.917 - 0}{\sqrt{\frac{132.917}{12(11)}}} & &= 235 - \frac{(35)^2}{12} \\
 &= 2.907 & &= 132.917 \\
 &= 2.91
 \end{aligned}$$

A partir de la tabla D, con $gl = 11$ y $\alpha = 0.05_{2 \text{ colas}}$,

$$t_{\text{crit}} = \pm 2.201$$

Puesto que $|t_{\text{obt}}| > 2.201$, rechazamos H_0 y concluimos que la campaña de ahorro afectó realmente el consumo de gasolina. Lo redujo de manera significativa.

También habríamos podido evaluar la campaña del experimento anterior mediante el diseño de grupos independientes. En lugar de utilizar los mismos sujetos en cada condición, tendríamos dos grupos de individuos. Un grupo sería analizado antes de la campaña y el otro después de ésta. Para evaluar la hipótesis nula, consideraríamos a cada una de las muestras como independientes extraídas de manera aleatoria de poblaciones en las cuales $\mu_1 = \mu_2$. El estadístico fundamental del cálculo sería $\bar{X}_1 - \bar{X}_2$. A manera de comparación, analicemos los datos de la campaña para la promoción del ahorro de gasolina como si se hubiesen obtenido mediante un diseño de grupos independientes. Su-

ponga que existen dos grupos diferentes. Las familias del grupo 1 (antes) son observadas durante un mes para medir la cantidad de gasolina que consumen antes de la campaña de ahorro, mientras que las familias del grupo 2 son observadas a este respecto durante un mes, después de haberse realizado la campaña.

Como quiera que $n_1 = n_2$, podemos aplicar la ecuación t_{obt} para n iguales. En este experimento, $n_1 = n_2 = 12$. Al despejar \bar{X}_1 , \bar{X}_2 , SC_1 y SC_2 obtenemos:

$$\begin{aligned}\bar{X}_1 &= \frac{\sum X_1}{n_1} = \frac{587}{12} = 48.917 & \bar{X}_2 &= \frac{\sum X_2}{n_2} = \frac{552}{12} = 46.000 \\ SC_1 &= \sum X_1^2 - \frac{(\sum X_1)^2}{n_1} & SC_2 &= \sum X_2^2 - \frac{(\sum X_2)^2}{n_2} \\ &= 29\,617 - \frac{(587)^2}{12} & &= 26\,496 - \frac{(552)^2}{12} \\ &= 902.917 & &= 1104\end{aligned}$$

Al sustituir estos valores en la ecuación para t_{obt} con n iguales, obtenemos:

$$t_{\text{obt}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{SC_1 + SC_2}{n(n-1)}}} = \frac{48.917 - 46.000}{\sqrt{\frac{902.917 + 1104}{12(11)}}} = 0.748 = 0.75$$

Según la tabla D, con $gl = N - 2 = 24 - 2 = 22$ y $\alpha = 0.05_{2 \text{ colas}}$,

$$t_{\text{crit}} = \pm 2.074$$

Como $|t_{\text{obt}}| < 2.074$, conservamos H_0 .

Parece que aquí hay algo raro. Cuando los datos fueron recopilados con un diseño de grupos correlacionados rechazamos H_0 . Sin embargo, en el caso de un diseño de grupos independientes, no conseguimos rechazar H_0 , aunque los datos eran idénticos. ¿Por qué? El diseño de grupos correlacionados nos permite utilizar a los sujetos como su propio control. Esto maximiza la posibilidad de que exista una alta correlación entre los datos que se encuentran en las dos condiciones. En el ejemplo en cuestión, la r de Pearson para la correlación entre las parejas de datos, antes y después, es igual a 0.938. Cuando la correlación es alta,* los puntajes de diferencia serán mucho menos variables que los datos originales. Por ejemplo, consideremos los datos de las familias A y L. La familia A consume mucha gasolina (55 galones), mientras que la familia L consume mucho menos (32 galones). Como resultado de la campaña para el ahorro de gasolina, los datos de ambas familias decrecen en 7 galones. Sus diferencias son idénticas (7). No existe variabilidad entre las diferencias correspondientes a estas familias, aunque sí hay una gran variabilidad entre sus datos en bruto. El potencial de una alta correlación y, por ende, de una disminución de la variabilidad, es la causa de que el diseño de grupos correlacionados sea potencialmente más poderoso que el diseño de grupos independientes. La disminución de la variabilidad en el presente ejemplo se puede ver con más claridad al comparar las dos soluciones. Esto se presenta en la tabla 14.7.

* Véase la nota 14.4 para una comparación directa entre las dos ecuaciones t en las que interviene la r de Pearson.

tabla 14.7 Soluciones para diseños con grupos correlacionados e independientes

Grupos correlacionados	Grupos independientes
$t_{\text{obt}} = \frac{D}{\sqrt{SC_D}}$ $= \frac{2.917}{\sqrt{132.917}}$ $= 2.91$	$t_{\text{obt}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{SC_1 + SC_2}}$ $= \frac{2.917}{\sqrt{902.917 + 1104}}$ $= 0.75$

Las dos ecuaciones producen los mismos valores, excepto SC_D en el diseño de grupos correlacionados y $SC_1 + SC_2$ en el diseño de grupos independientes. SC_D es una medida de la variabilidad de las diferencias. $SC_1 + SC_2$ son medidas de la variabilidad de los datos en bruto. $SC_D = 132.917$, mientras que $SC_1 + SC_2 = 2006.917$; SC_D es mucho más pequeña que $SC_1 + SC_2$. Esta disminución de la variabilidad hace que t_{obt} sea mayor en el análisis de grupos correlacionados.

Si el diseño de grupos correlacionados es potencialmente más poderoso, ¿por qué no lo utilizamos siempre? En primer lugar, el diseño de grupos independientes es mucho más eficaz desde el punto de vista del análisis de grados de libertad por medición. Los grados de libertad son importantes, debido a que mientras mayor sea el número de éstos, t_{crit} será menor. En nuestro ejemplo, para el diseño de grupos correlacionados se realizaron 24 mediciones, pero sólo se obtuvieron 11 grados de libertad. En cambio, para el diseño de grupos independientes se llevaron a cabo 24 mediciones y se consiguieron 22 grados de libertad. De esta manera, el diseño de grupos independientes produjo el doble de grados de libertad para la misma cantidad de mediciones.

En segundo lugar, muchos experimentos excluyen el uso del mismo sujeto bajo ambas condiciones. Por ejemplo, suponga que queremos investigar si los hombres y las mujeres son diferentes con respecto a la agresividad. Está claro que no podríamos utilizar al mismo sujeto bajo ambas condiciones. A veces, el efecto de la primera condición persiste durante mucho tiempo. Si el experimento requiere que las dos condiciones se administren en un periodo breve, tal vez no será posible emplear al mismo sujeto en ambas condiciones sin que la primera condición afecte su desempeño en la segunda condición. Con frecuencia, cuando se utiliza al sujeto en la primera condición, éste se "gasta" y no es posible utilizarlo en la segunda condición. Esto es particularmente válido en los experimentos para medir el aprendizaje. Por ejemplo, si nos interesa estudiar los efectos que el ejercicio produce sobre el aprendizaje del esquí, sabemos que una vez que los sujetos hayan aprendido a esquiar, ya no podrán ser utilizados de nuevo, puesto que ya sabrán esquiar. Cuando el mismo sujeto no puede ser utilizado en las dos condiciones, es posible formar parejas de sujetos. Sin embargo, la formación de parejas consume tiempo y es costosa. Además, con frecuencia ocurre que el experimentador no sabe cuáles son las variables importantes, en la formación de parejas, para producir una mayor correlación. Por todas esas razones, el diseño de grupos independientes se utiliza con más frecuencia que el diseño de grupos correlacionados.

RESUMEN

En este capítulo hemos examinado la prueba t para grupos correlacionados e independientes. Señalamos que, en realidad, este tipo de prueba para grupos co-

relacionados es un caso particular de la prueba t para una sola muestra. En el diseño de grupos correlacionados, se analizaron las diferencias entre parejas de da-

tos. Si la variable independiente no producía efecto alguno y el azar era el único responsable de las diferencias, entonces éstas se podían considerar como una muestra aleatoria extraída de una población de diferencias en la cual $\mu_D = 0$ y σ_D se desconoce. Sin embargo, éstas son precisamente las condiciones en que se aplica la prueba *t* para una sola muestra. La única modificación consistió en que, en el diseño de grupos correlacionados, analizamos puntajes de diferencia, mientras que en el diseño de una sola muestra examinamos datos en bruto. Después de presentar algunos problemas ilustrativos y de práctica, concluimos nuestro análisis de la prueba *t* para grupos correlacionados comparándola con la prueba de los signos. Mostramos que aunque ambas pruebas son adecuadas para el diseño de grupos correlacionados, siempre que sus supuestos se cumplan, debe emplearse la prueba *t* porque es más sensible.

La prueba *t* para grupos independientes se utiliza cuando el experimento posee dos grupos independientes. El estadístico que se analiza es la diferencia entre las medias de las dos muestras ($\bar{X}_1 - \bar{X}_2$). Los datos de la muestra 1 se consideran como una muestra aleatoria extraída de una población que tiene una media μ_1 y una desviación estándar σ_1 . Los datos de la muestra 2 son una muestra aleatoria extraída de una población que tiene una media μ_2 y una desviación estándar σ_2 .

Si la variable independiente tiene un efecto real, entonces la diferencia entre las medias muestrales se debe a un muestreo aleatorio de poblaciones en las que $\mu_1 \neq \mu_2$. Se supone que el cambio de nivel de la variable independiente afecta las medias de las poblaciones, pero no las desviaciones estándar ni las varianzas. Si la variable independiente no produce efecto alguno y el azar es el único responsable de las diferencias entre las dos muestras, entonces la diferencia entre las medias muestrales se debe a un muestreo aleatorio de poblaciones en las cuales $\mu_1 = \mu_2$. En estas condiciones, la distribución muestral de $\bar{X}_1 - \bar{X}_2$ tiene media cero y una desviación estándar cuyo valor se determina conociendo la varianza de las poblaciones a partir

de las cuales se extrajeron las muestras. Como este valor nunca se conoce, no podemos emplear la prueba *z*. Sin embargo, podemos estimar la varianza mediante una estimación ponderada a partir de ambas muestras. Al hacer esto, el estadístico resultante es t_{ob} .

Entonces, se usa el estadístico *t* para el análisis de los datos del experimento que presenta grupos independientes con dos muestras. La distribución muestral de *t* para este diseño es la misma que para el diseño de una sola muestra, salvo que los grados de libertad son diferentes. En el diseño de grupos independientes, $gl = N - 2$. Después de presentar varios problemas ilustrativos y de práctica, estudiamos los supuestos subyacentes de la prueba *t* para grupos independientes. Señalamos que esta prueba requiere: 1) que las poblaciones de datos en bruto tengan una distribución normal, y 2) que exista homogeneidad de la varianza. También señalamos que la prueba *t* es robusta con respecto a la violación de ambos supuestos: el de la normalidad de la población y el de la homogeneidad de la varianza. Además de determinar si existe un efecto significativo, también es importante determinar la magnitud de dicho efecto. En un experimento con grupos independientes, la magnitud del efecto de la variable independiente se puede estimar calculando $\hat{\omega}^2$.

A continuación analizamos la potencia de la prueba *t*. Mostramos que dicha potencia varía en proporción directa con la magnitud del efecto real de la variable independiente y la *N* del experimento, pero que varía en proporción inversa con la variabilidad de los datos de la muestra.

Por último, comparamos los diseños de grupos correlacionados y los de grupos independientes. Cuando la correlación entre las parejas de datos es alta, el diseño de grupos correlacionados es más sensible que el diseño de grupos independientes. Sin embargo, en lo que se refiere a los grados de libertad, resulta más sencillo y eficaz realizar un experimento con grupos independientes. Además, existen muchas situaciones en las cuales el diseño de grupos correlacionados no resultaría apropiado.

TÉRMINOS IMPORTANTES

Diferencia entre medias muestrales (página 329)	Grados de libertad (página 332)	Puntajes de diferencia (página 319)
Diseño de grupos correlacionados (página 319)	Homogeneidad de la varianza (página 338)	Prueba <i>t</i> para grupos correlacionados (página 319)
Diseño de grupos independientes (página 327)	Magnitud del efecto (página 339)	Prueba <i>t</i> para grupos independientes (página 331)
Distribución muestral de la diferencia entre medias muestrales (página 329)	Media de los puntajes de diferencia (página 320)	Variabilidad (página 342)

PREGUNTAS Y PROBLEMAS

1. Identifique o defina los conceptos de la sección "términos importantes".
2. Analice las ventajas del experimento con dos condiciones, en comparación con las ventajas del experimento que presenta una sola muestra.
3. La prueba t para grupos correlacionados se puede considerar como un caso particular de la prueba t para una sola muestra, analizada en el capítulo anterior. Explique.
4. ¿Cuál es la principal ventaja del uso de la prueba t para grupos correlacionados, sobre el uso de la prueba de los signos, para analizar los datos procedentes de un experimento de grupos correlacionados?
5. ¿Cuáles son las características de la distribución muestral de la diferencia entre medias muestrales?
6. ¿Por qué nunca se utiliza la prueba z para grupos independientes?
7. ¿Qué es lo que se estima en la prueba t para grupos independientes? ¿Cómo se obtiene la estimación?
8. Se dice que la varianza de los datos de la muestra tiene una influencia importante en la potencia de la prueba t . ¿Es cierta esta afirmación? Explique.
9. ¿Cuáles son las ventajas y las desventajas del uso de un diseño de grupos correlacionados en lugar de un diseño de grupos independientes?
10. ¿Cuáles son los supuestos subyacentes de la prueba t para grupos independientes?
11. Después de cometer lo que cree es un error de tipo II al usar un diseño de grupos independientes y un análisis de prueba t , indique todo lo que haría en el próximo experimento para reducir la probabilidad de cometer ese error.
12. ¿Es importante la magnitud del efecto de la variable independiente? Explique.
13. Si el efecto de la variable independiente es significativo, ¿quiere decir esto necesariamente que dicho efecto es grande? Explique.

Para cada uno de los siguientes problemas, suponga que la población tiene una distribución normal, a menos que se indique otra cosa.

14. Usted quiere determinar si una píldora experimental para el control de la natalidad produce el efecto colateral de modificar la presión sanguínea. Usted elige al azar a 10 mujeres de la ciudad en la que vive. A 5 de ellas les administra un placebo durante un mes y después mide su presión sanguínea diastólica. Luego, sustituye el placebo por la píldora y mide de nuevo la presión sanguínea. Las otras 5 mujeres reciben el mismo tratamiento, sólo que se les administra primero la píldora durante un mes y después el placebo, también durante un mes. Las lecturas de la presión sanguínea se presentan en la tabla siguiente. Observe que para evitar que las mujeres resultaran con embarazos no deseados, durante todo el experimento se utilizaron otros métodos anticonceptivos que no interfirieron con la píldora.

Sujeto No.	Presión sanguínea diastólica	
	Píldora para el control de la natalidad	Placebo
1	108	102
2	76	68
3	69	66
4	78	71
5	74	76
6	85	80
7	79	82
8	78	79
9	80	78
10	81	85

- a. ¿Cuál es la hipótesis alternativa? Supongamos que resulta apropiada una hipótesis no direccional.
 - b. ¿Cuál es la hipótesis nula?
 - c. ¿Cuál es la conclusión de usted? Utilice $\alpha = 0.01$ en las pruebas social, biológica, salud
15. A partir de investigaciones anteriores y consideraciones teóricas sólidas, un psicólogo cognoscitivo sospecha que el recuerdo de imágenes es superior al recuerdo de sustantivos. Para verificar esta hipótesis, el psicólogo realiza un experimento en el cual los estudiantes de una clase de introducción a la psicología son utilizados como sujetos del estudio. Ocho estudiantes elegidos al azar miran 30 diapositivas con sustantivos impresos en ellas, y otro grupo de ocho alumnos seleccionados al azar miran 30 diapositivas con imágenes que corresponden a esos mismos sustantivos. Cada diapositiva contiene un sustantivo o una imagen y es observada durante 4 segundos. Después de haber visto las diapositivas, los sujetos reciben una prueba de recuerdo y se cuenta la cantidad de ítems recordados en forma correcta. Los datos son los siguientes:

Número de imágenes recordadas	Número de sustantivos recordados
18	12
21	9
14	21
25	17
23	16
19	10
26	19
15	22

- a. ¿Cuál es la hipótesis alternativa? Supongamos que resulta apropiada una hipótesis direccional.
- b. ¿Cuál es la hipótesis nula?
- c. ¿Cuál es la conclusión de usted? Utilice $\alpha = 0.05$ cola?
- d. Estime la magnitud del efecto real. ¿Es grande el efecto? cognitiva
16. Una enfermera fue contratada por una oficina de gobierno que se encarga de temas ecológicos, para que investigue el efecto causado por una fundición de plomo sobre el nivel de éste en la sangre de los niños que viven cerca de esa fábrica. Se eligen al azar 10 niños, de entre todos los que viven cerca de la fundición. Se escoge al azar un grupo de comparación de 7 niños, extraídos de una población que vive en un área relativamente libre de contaminación por plomo. Se extrajeron muestras de sangre de los niños que conformaban estos grupos experimentales y se midieron los niveles de plomo. Los resultados son los siguientes (los datos se miden en microgramos de plomo por 100 mililitros de sangre):

Niveles de plomo	
Niños que viven cerca de la fundición	Niños que viven en un área contaminada
18	9
16	13
21	8
14	15
17	17
19	12
22	11
24	
15	
18	

- a. ¿Cuál es la conclusión de usted? Utilice $\alpha = 0.01$ cola.
- b. Estime la magnitud del efecto real. ¿Es pequeño el efecto? salud
17. La gerente de la sección de cosméticos de una gran tienda de departamentos quiere determinar si la publicidad en los periódicos afecta verdaderamente el nivel de ventas. Para este experimento elige, de manera aleatoria, 15 artículos disponibles en el almacén y establece su línea base. Los precios de los 15 artículos se registran a su valor competitivo ordinario. Se registra la cantidad vendida de cada producto durante un periodo de una semana. Después, sin modificar el precio, la gerente publica un gran anuncio en el periódico, en el cual promueve los 15 artículos. De nuevo, ella registra la cantidad vendida en un periodo de una semana. A continuación presentamos los resultados.

Artículo	Cantidad vendida antes del anuncio	Cantidad vendida después del anuncio
1	25	32
2	18	24
3	3	7
4	42	40
5	16	19
6	20	25
7	23	23
8	32	35
9	60	65
10	40	43
11	27	28
12	7	11
13	13	12
14	23	32
15	16	28

- ¿Qué conclusión saca usted? Utilice $\alpha = 0.05$ 2 colas. /0
18. Como la tensión muscular en la región de la cabeza se ha asociado con los dolores de cabeza, usted razona que si la tensión muscular disminuye, es probable que los dolores de cabeza se reduzcan o desaparezcan. Usted diseña un experimento en el cual participan nueve sujetos que padecen dolores de cabeza. Estos individuos llevan un registro diario del número de dolores de cabeza que sufren durante un periodo de línea base de 2 semanas. Posteriormente usted los entrena para que aprendan a reducir la tensión muscular en la región de la cabeza, utilizando un dispositivo de biorretroa-

limentación. Para este experimento, el dispositivo mencionado se conecta al músculo frontal, que se encuentra en la región del frente de la cara. El dispositivo indica al sujeto la cantidad de tensión que existe en el músculo al cual está unido (en este caso, el frontal) y le ayuda a reducir los niveles de tensión. Después de 6 semanas de entrenamiento, los sujetos han logrado mantener una baja tensión en el músculo frontal; entonces se lleva nuevamente un registro del número de dolores de cabeza que sufren durante otras dos semanas. A continuación presentamos el número de dolores de cabeza registrados durante cada uno de esos periodos de 2 semanas.

Sujeto No.	Número de dolores de cabeza	
	Línea base	Después del entrenamiento
1	17	3
2	13	7
3	6	2
4	5	3
5	5	6
6	10	2
7	8	1
8	6	0
9	7	2

- a. ¿Qué conclusión saca usted si $\alpha = 0.05_{2 \text{ colas}}$? Supongamos que la distribución muestral de la media de las diferencias (\bar{D}) presenta una distribución normal. Suponga que una hipótesis no direccional resulta apropiada, porque no existen suficientes bases empíricas para garantizar una hipótesis direccional.
- b. Si la distribución muestral de \bar{D} no es normal, ¿qué otra prueba utilizaría usted para analizar los datos? ¿Cuál sería la conclusión? clínica, salud

19. Existe un problema de interpretación en el problema 18. Está claro que los dolores de cabeza disminuyeron de manera significativa. Sin embargo, es posible que la disminución no se deba al entrenamiento con biorretroalimentación, sino a algún otro factor también presente en la situación, como la atención que se proporcionó a los sujetos. Lo que en realidad se necesita es tener un grupo que controle esta posibilidad. Suponga que se utilizó otro grupo de nueve pacientes con dolor de cabeza, al mismo tiempo que el grupo del problema 18. Este grupo fue tratado en la misma forma, excep-

to que los sujetos no recibieron ningún entrenamiento relacionado con la biorretroalimentación. Cada semana, ellos sólo hablaron con usted acerca de sus dolores de cabeza, durante 6 semanas, y usted les mostró mucho afecto, amabilidad y atención. El número de dolores de cabeza durante la línea base y en el periodo de seguimiento de dos semanas para el grupo de control se presenta en la siguiente tabla:

Sujeto No.	Número de dolores de cabeza	
	Línea base	Periodo de seguimiento
1	5	4
2	8	9
3	14	12
4	16	15
5	6	4
6	5	3
7	8	7
8	10	6
9	9	7

Evalúe el efecto de estos otros factores, como la atención, sobre la incidencia de los dolores de cabeza. Utilice $\alpha = 0.05_{2 \text{ colas}}$, clínica, salud

20. Como el grupo control del problema 19 mostró también una reducción significativa de los dolores de cabeza, la interpretación de los resultados del problema 18 queda en duda. ¿Contribuyó el entrenamiento de relajación a la disminución de los dolores de cabeza, o bien esta disminución se debió solamente a factores tales como la atención? Para responder esta pregunta, podemos comparar los cambios registrados en ambos grupos. Estos datos se presentan en la siguiente tabla:

Modificación de los dolores de cabeza	
Grupo de entrenamiento de relajación	Grupo control
14	1
6	-1
4	2
2	1
-1	2
8	2
7	1
6	4
5	2

¿Qué conclusión saca usted? Utilice $\alpha = 0.05_{2 \text{ colas}}$, clínica, salud

21. El director de recursos humanos de una compañía grande piensa contratar empleados de tiempo parcial para cubrir los puestos que antes estaban asignados a personal de tiempo completo. Sin embargo, él se pregunta si esto afectará la productividad. Por lo tanto, lleva a cabo un experimento para evaluar la idea antes de implantarla en toda la fábrica. Cada una de las seis plazas de tiempo completo en la división de fabricación de partes, son ocupadas por dos empleados contratados para trabajar medio tiempo. La producción de estas seis parejas de medio tiempo es comparada con el resultado de una muestra elegida al azar de seis empleados de tiempo completo de la misma división. Observe que todos los trabajadores del experimento se dedican a fabricar las mismas partes. El número promedio de partes producidas al día por las parejas de medio tiempo y por los obreros de tiempo completo se presentan a continuación:

Partes producidas por día	
Parejas de medio tiempo	Obreros de tiempo completo
24	20
26	28
46	40
32	36
30	24
36	30

¿Afecta a la productividad la contratación de obreros de medio tiempo? Utilice $\alpha = 0.05_{2 \text{ colas}}$ para tomar su decisión. 1/0

22. A partir de la experiencia que ha adquirido mediante el trato con sus clientes, una psicóloga clínica piensa que la depresión puede afectar el sueño. Entonces decide verificar esta idea. Para eso, el sueño de nueve pacientes con depresión y de ocho controles que no padecen este problema, es observado durante tres noches sucesivas. El número promedio de horas que durmió cada sujeto durante las dos últimas noches se presenta en la siguiente tabla:

Horas de sueño	
Pacientes deprimidos	Controles normales
7.1	8.2
6.8	7.5
6.7	7.7
7.3	7.8
7.5	8.0
6.2	7.4
6.9	7.3
6.5	6.5
7.2	

- a. ¿Está en lo correcto la psicóloga? Utilice $\alpha = 0.05_{2 \text{ colas}}$ para tomar una decisión.
 b. Si el efecto es significativo, estime la magnitud del efecto. ¿Es grande el efecto? clínica, salud
23. Un maestro quiere determinar si el ingreso temprano a la escuela podría afectar el CI. Para eso, consigue la ayuda de los padres de 12 parejas de gemelos idénticos en edad preescolar; los padres están de acuerdo en que sus hijos participen en el experimento. Un miembro de cada pareja de gemelos entra a preescolar y allí pasa 2 años, mientras que el otro miembro de la pareja permanece en su casa. Al final de los 2 años, se mide el CI de todos los niños y así se obtiene la siguiente tabla.

Pareja	CI	
	Gemelos en la escuela	Gemelos en la casa
1	110	114
2	121	118
3	107	103
4	117	112
5	115	117
6	112	106
7	130	125
8	116	113
9	111	109
10	120	122
11	117	116
12	106	104

¿Afecta al CI el ingreso temprano a la escuela? Utilice $\alpha = 0.05_{2 \text{ colas}}$, cognitiva, de desarrollo, educación

24. Los investigadores de una prestigiosa universidad estaban interesados en averiguar el efecto del sueño sobre la consolidación de la memoria. Veinticuatro estudiantes de un curso de introducción a la filosofía se ofrecieron voluntariamente para un experimento y fueron asignados al azar, ya sea a un grupo que "sí duerme" o a un grupo que "no duerme", de tal manera que en cada grupo había 12 estudiantes. En el primer día, todos los estudiantes presenciaron una exhibición de 15 objetos diferentes, cada uno de los cuales fue visible durante 200 milisegundos en una pantalla de computadora, y se les pidió que recordaran el mayor número posible de esos objetos. Esa noche, el grupo que "sí duerme" gozó de una noche de sueño normal. El grupo que "no duerme" se mantuvo despierto hasta la segunda noche. Todos los sujetos durmieron normalmente la segunda noche y también la tercera. El cuarto día, todos los sujetos fueron sometidos a una prueba para ver a cuántos de los 15 objetos originales podían recordar. A continuación se presenta el número de objetos que recordó cada uno de los sujetos en esta prueba:

Grupo que sí duerme	Grupo que no duerme
14	8
13	9
8	6
9	13
11	7
10	9
9	10
13	12
12	8
11	11
14	9
13	12

Utilizando $\alpha = 0.05_{2\text{ colas}}$, ¿qué conclusión saca usted? *cognitiva*

25. Algunos psicólogos del desarrollo que trabajan en una importante universidad de California realizaron un estudio longitudinal para investigar el efecto de los niveles altos de curiosidad a edad temprana sobre la inteligencia. La población de los niños de 3 años de edad de la localidad fue estudiada mediante una batería de pruebas para evaluar la curiosidad. A 12 de los niños de 3 años que calificaron en el 90% superior en esta variable, se les sometió a una prueba de CI en ese momento y nue-

vamente a la edad de 11 años. Así se obtuvieron los siguientes puntajes de CI.

Número de sujeto	CI (3 años de edad)	CI (11 años de edad)
1	100	114
2	105	116
3	125	139
4	140	151
5	108	106
6	122	119
7	117	131
8	112	136
9	135	148
10	128	139
11	104	122
12	98	113

Utilizando $\alpha = 0.01_{2\text{ colas}}$, ¿qué conclusión saca usted? Al extraer su conclusión, suponga que ha quedado bien establecido que el CI permanece relativamente constante durante esos años, en el caso de individuos con niveles de curiosidad promedio o inferiores al promedio. *cognitiva, de desarrollo*

26. Al observar que las mujeres parecen estar más interesadas por las emociones que los hombres, una investigadora de la especialidad de estudios sobre la mujer se planteó la pregunta de si las mujeres recuerdan mejor que los hombres los eventos emocionales. Entonces decidió reunir algunos datos sobre esa cuestión. Llevó a cabo un experimento para el cual fueron elegidos al azar ocho hombres y mujeres, a quienes se les mostraron 20 fotografías con alto contenido emocional, después de lo cual se les pidió que las recordaran al cabo de una semana de habérselas mostrado. Así se obtuvieron los siguientes datos sobre su capacidad de recuerdo. Los puntajes representan el porcentaje de respuestas correctas; un hombre no se presentó a la prueba de memoria.

Hombres	Mujeres
75	85
85	92
67	78
77	80
83	88
88	94
86	90
	89

Utilizando $\alpha = 0.05_{2 \text{ colas}}$, ¿qué conclusión saca usted? cognitiva, social

27. En virtud de que los resultados del experimento del problema 26 estuvieron muy cerca de ser significativos, la investigadora decide repetir el experimento, sólo que en esta ocasión aumentará la potencia del mismo incrementando N . Este estudio incluyó a diez hombres y diez mujeres. Así se obtuvieron los siguientes resultados.

Hombres	Mujeres
74	87
87	90
64	80
76	77
85	91
86	95
84	89
78	92
77	90
80	94

Utilizando $\alpha = 0.05_{2 \text{ colas}}$, ¿cuál es la conclusión de usted en esta ocasión? cognitiva, social

28. Un profesor de física cree que la iluminación natural en las aulas puede mejorar el aprendizaje de los estudiantes. Entonces realiza un experimento

en el cual imparte la misma unidad de física a dos grupos de siete estudiantes cada uno, asigna los al azar a cada grupo. Todos los detalles son similares en uno y otro grupo, salvo que a uno de ellos se le imparte la clase en un aula que recibe mucha luz natural además del alumbrado eléctrico, mientras que el otro ocupa un aula que sólo está iluminada con luz eléctrica. Al final de la unidad, ambos grupos presentan el mismo examen sobre dicho material. La calificación máxima en el examen es de 20 puntos; cuanto más alta sea la calificación, tanto mejor habrá sido el desempeño. Así se obtuvieron las siguientes calificaciones.

Luz natural además del alumbrado eléctrico	Alumbrado eléctrico únicamente
16	17
18	13
14	12
17	14
16	13
19	15
17	14

Utilizando $\alpha = 0.05_{2 \text{ colas}}$, ¿qué conclusión saca usted? educación

NOTAS

- 14.1 La mayor parte de los libros de texto presentan dos métodos para determinar t_{obt} en el caso del diseño de grupos correlacionados: 1) el método de la diferencia directa y 2) un método que requiere el cálculo del grado de relación (el coeficiente de correlación) existente entre los dos conjuntos de datos en bruto. Hemos omitido el segundo método, porque rara vez se utiliza en la práctica y, desde nuestro punto de vista, confunde a los estudiantes. El método de la diferencia directa se desarrolla de manera más natural y lógica, a partir del análisis de la prueba *t* para una sola muestra. Es mucho más fácil de emplear y, de hecho, es utilizado con mucha frecuencia en la práctica.
- 14.2 En ciertas ocasiones, en un experimento con medidas repetidas donde la hipótesis alternativa es

direccional, el investigador podría desear verificar si la variable independiente tiene un efecto mayor que algún valor dado distinto de cero. Por ejemplo, si en el presente experimento se justificara una hipótesis direccional, el experimentador podría investigar si el valor de recompensa promedio del área A era cinco veces por minuto mayor que el del área B, tomando como unidad el número de veces que se presiona la palanca. En este caso, la hipótesis nula sería que el valor de recompensa del área A no es cinco veces por minuto mayor que el del área B, tomando como unidad el número de veces que se presiona la palanca. En este caso, $\mu_D = 5$ en lugar de 0.

- 14.3 A veces, en un experimento con grupos independientes, la hipótesis alternativa es direccional y específica que la variable independiente tiene un

efecto mayor que algún valor específico distinto de 0. Por ejemplo, en el experimento "hormona X", si una hipótesis direccional es válida, el fisiólogo podría verificar si la hormona X posee un efecto promedio de tres apareamientos o más sobre el efecto del placebo. En este caso, la hipótesis nula es que el efecto promedio de la hormona X es \leq tres apareamientos más que el efecto producido por el placebo. Quisiéramos verificar esta hipótesis, al suponer que la muestra que recibió la hormona X fue una muestra aleatoria, extraída de una población con una media 3 unidades mayor que la población de la cual se extrajo la muestra placebo. En este caso, $\mu_{\bar{X}_1 - \bar{X}_2} = 3$, y

$$t_{\text{obt}} = \frac{(\bar{X}_1 - \bar{X}_2) - 3}{\sqrt{\left(\frac{SS_1 + SS_2}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

14.4 Aunque no hemos presentado antes las ecuaciones t en esta forma, podemos escribir lo siguiente:

Prueba t para grupos independientes	Prueba t para grupos correlacionados
$t_{\text{obt}} = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}}$	$t_{\text{obt}} = \frac{\bar{D}}{s_D}$
$= \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_{\bar{X}_1}^2 + s_{\bar{X}_2}^2}}$	$= \frac{\bar{D}}{\sqrt{s_{\bar{X}_1}^2 + s_{\bar{X}_2}^2} - 2rs_{\bar{X}_1}s_{\bar{X}_2}}$

En virtud de que $\bar{X}_1 - \bar{X}_2$ es igual a \bar{D} , las ecuaciones t_{obt} para los grupos independientes y los grupos correlacionados son idénticas, excepto por el término $-2rs_{\bar{X}_1}s_{\bar{X}_2}$ en el denominador de la ecuación para grupos correlacionados. Así pues, la mayor sensibilidad del diseño de grupos correlacionados depende de la magnitud de r . Mientras mayor sea el valor de r , más sensible será el diseño de grupos correlacionados con respecto al diseño de grupos independientes. Si utilizamos los datos del experimento de la película de conservación para ilustrar el uso de estas ecuaciones, obtenemos lo siguiente:

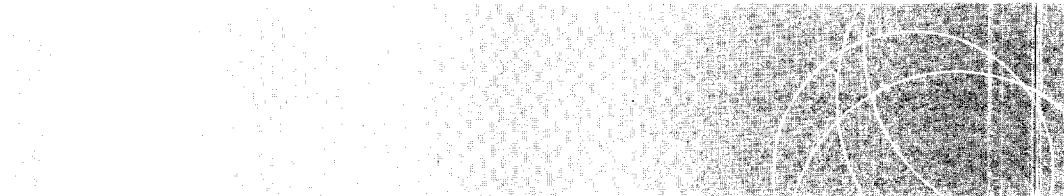
Grupos independientes	Grupos correlacionados
$t_{\text{obt}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_{\bar{X}_1}^2 + s_{\bar{X}_2}^2}}$	$t_{\text{obt}} = \frac{\bar{D}}{\sqrt{s_{\bar{X}_1}^2 + s_{\bar{X}_2}^2} - 2rs_{\bar{X}_1}s_{\bar{X}_2}}$
donde $s_{\bar{X}_1}^2 = \frac{s_1^2}{n_1} = \frac{82.083}{12} = 6.840$	$= \frac{2.917}{\sqrt{6.840 + 8.364} - 2(0.938)(2.615)(2.892)}$
y $s_{\bar{X}_2}^2 = \frac{s_2^2}{n_2} = \frac{100.364}{12} = 8.364$	$= \frac{2.917}{\sqrt{15.204}}$
Por lo tanto,	$= 2.91$
$t_{\text{obt}} = \frac{48.917 - 46.000}{\sqrt{6.840 + 8.364}} = \frac{2.917}{\sqrt{15.204}} = 0.748$	
$= 0.75$	
Observe que éstos son los mismos valores que obtuvimos anteriormente.	

Si desea encontrar problemas de práctica adicionales totalmente resueltos sobre los temas que acaba de estudiar, busque Chapter 14 en el material adicional del libro, disponible en inglés en: <http://www.thomsonlearning.com.mx>.

Si pretende resolver más problemas de práctica, relacionados con lo que se presentó en este capítulo y resueltos por medio de ecuaciones conceptuales, busque Chapter 14 en el material adicional del libro (en inglés), en: <http://www.thomsonlearning.com.mx>.

Si quiere resolver problemas para computadora sobre los temas aquí presentados, busque Chapter 14 en el material adicional (en inglés) del libro, en: <http://www.thomsonlearning.com.mx>.

Introducción al análisis de varianza



- Introducción: La distribución F
- La prueba F y el análisis de varianza (ANOVA)
- Panorama de la técnica del análisis de varianza de un factor o vía
 - Estimación de la varianza dentro de grupos, s_W^2*
 - Estimación de la varianza entre grupos, s_B^2*
 - La razón F*
- Análisis de datos con la técnica del análisis de varianza
 - Un experimento: Situaciones diversas y tensión*
- Lógica subyacente al análisis de varianza de un factor
- La relación entre el análisis de varianza y la prueba t
 - Ejemplo: Citas. Las jóvenes de clubes femeninos estudiantiles versus las de dormitorios*
- Supuestos subyacentes al análisis de varianza
- Magnitud del efecto
- Potencia del análisis de varianza
 - Potencia y N*
 - Potencia y el efecto real de la variable independiente*
 - Potencia y la variabilidad de la muestra*
- ¿CUÁL ES LA VERDAD? Mucho ruido y pocas nueces
- Resumen
- Términos importantes
- Preguntas y problemas



INTRODUCCIÓN: LA DISTRIBUCIÓN F

En los capítulos 12, 13 y 14 utilizamos la *media* como el estadístico básico para evaluar la hipótesis nula. También es posible emplear la *varianza* de los datos para llevar a cabo la prueba de hipótesis. Una de las pruebas más importantes en las que se usa la varianza es la prueba F , llamada así en honor de R. A. Fisher, el estadístico que la desarrolló. Cuando

se utiliza esta prueba, calculamos el estadístico F_{obt} , que básicamente es la razón de dos estimaciones independientes de la misma varianza poblacional σ^2 . En forma de ecuación,

$$F_{\text{obt}} = \frac{\text{Primera estimación de la varianza } \sigma^2}{\text{Segunda estimación de la varianza } \sigma^2}$$

La distribución muestral de F se puede generar de manera empírica, 1) considerando todas las posibles muestras de tamaño n_1 y n_2 de la misma población, 2) estimando la varianza σ^2 de la población de cada una de las muestras mediante s_1^2 y s_2^2 , 3) calculando F_{obt} para cada una de las posibles combinaciones de s_1^2 y s_2^2 , y luego, 4) determinando $p(F)$ para cada valor diferente de F_{obt} . El resultado es la *distribución muestral de F* . Así, la definición de esta distribución es:

definición ■ La **distribución muestral de F** proporciona todos los posibles valores F junto con la $p(F)$ de cada uno de ellos, suponiendo que se realiza un muestreo aleatorio en la población.

Igual que la distribución t , la distribución F varía según los *grados de libertad*. Sin embargo, la distribución F tiene dos valores para los grados de libertad: uno para el numerador y otro para el denominador. Como ya habrá imaginado, perdemos un grado de libertad por cada cálculo de la varianza. Por lo tanto,

$$\text{Grados de libertad para el numerador} = \text{gl}_1 = n_1 - 1$$

$$\text{Grados de libertad para el denominador} = \text{gl}_2 = n_2 - 1$$

La figura 15.1 muestra una distribución F con 3 gl en el numerador y 16 gl en el denominador. Podemos apreciar varias características de esta distribución. En primer lugar, como F es la razón de las estimaciones de la varianza, nunca tiene un valor negativo (s_1^2 y s_2^2 siempre serán positivas). En segundo lugar, la distribución F está sesgada positivamente. Por último, el valor mediano de F es aproximadamente igual a 1.

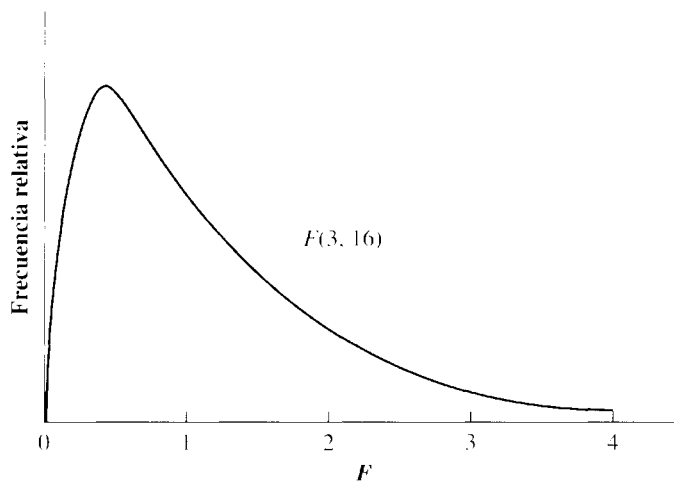


figura 15.1 Distribución F con 3 grados de libertad en el numerador y 16 grados de libertad en el denominador.

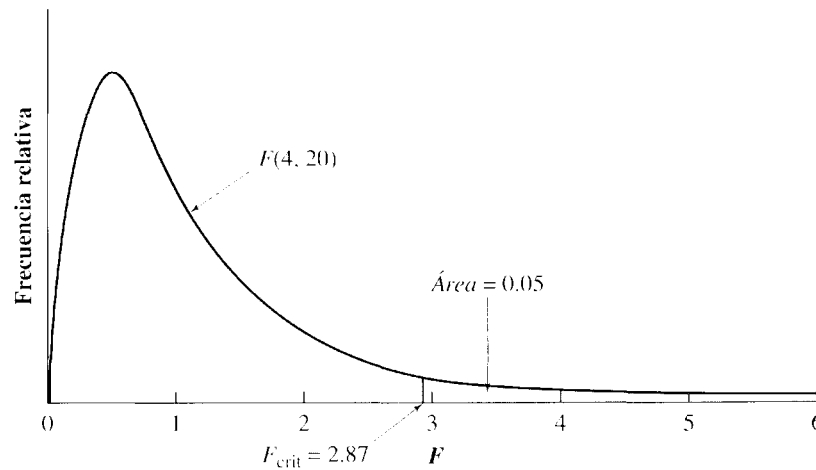


figura 15.2 Ejemplo donde se muestra que F_{crit} en la tabla F tiene una cola para el lado derecho.

Igual que en el caso de la prueba t , existe una familia de curvas F . Sin embargo, para la prueba F hay una curva distinta para cada combinación de los grados de libertad gl_1 y gl_2 . La tabla F del apéndice D proporciona los valores críticos de F para diferentes combinaciones de gl_1 y gl_2 . Existen dos datos por cada celda. El dato en tipo normal provee el valor crítico de F para el nivel 0.05. El dato en negrilla proporciona el valor crítico de F para el nivel 0.01. Observe que se trata de valores de una sola cola para el lado derecho de la distribución F . Por ejemplo, la figura 15.2 muestra la distribución F para 4 gl en el numerador y 20 gl en el denominador. De la tabla F , F_{crit} en el nivel 0.05. Esto significa que 5% de los valores F son iguales o mayores que 2.87. El área que contiene estos valores aparece sombreada en la figura 15.2.



LA PRUEBA F Y EL ANÁLISIS DE VARIANZA (ANOVA)

La prueba F resulta apropiada para cualquier experimento en el cual sea posible utilizar los datos para obtener dos estimaciones independientes de la varianza poblacional. Una situación bastante frecuente en las ciencias del comportamiento, para la cual es adecuada la prueba F , se presenta al analizar los datos de los experimentos en los que se utilizan más de dos grupos o condiciones.

Hasta este punto, en el texto hemos considerado el experimento más fundamental: el estudio de dos grupos, uno control y otro experimental. Aunque este diseño se sigue utilizando con frecuencia, es más común encontrar experimentos que involucren a tres o más grupos. Una de las principales limitantes que presenta el estudio con dos grupos es que a menudo esos dos grupos no bastan para interpretar con claridad los hallazgos experimentales. Por ejemplo, en el experimento sobre “el tálamo y la percepción del dolor” (página 337) fueron incluidos dos grupos. Uno recibió lesiones en el tálamo y el otro en un área que según se “cree” no tiene relación con el dolor. Los resultados mostraron un umbral de dolor significativamente mayor para las ratas que tenían lesiones en el tálamo. Nuestra conclusión fue que las lesiones en el tálamo elevaron el umbral del dolor. Sin embargo, la diferencia entre los dos grupos bien podría deberse a una reducción del umbral del dolor como resultado de la otra lesión, y no como una elevación del umbral ocasionado por el daño sufrido en el tálamo. Esta ambigüedad se podría haber resuelto

si se hubieran utilizado tres grupos en lugar de dos. El tercer grupo podría haber sido un grupo control sin lesión alguna. Una comparación del umbral de dolor de los dos grupos lesionados con el del grupo sin lesiones, habría ayudado a resolver la cuestión.

Otro tipo de experimentos en los que se requieren más de dos grupos son aquellos en los que la variable independiente se varía como factor; es decir, que se elige un rango predeterminado para esta variable independiente y se utilizan distintos valores de dicho rango en el experimento. Por ejemplo, en el problema de “la hormona X y el comportamiento sexual”, en lugar de seleccionar al azar un valor de la hormona, el experimentador podría elegir diversos niveles en el rango de valores eficaces posibles. Cada nivel se aplicaría a un grupo diferente de sujetos, elegidos al azar entre la población. En el experimento habría tantos grupos como niveles de la hormona. Los experimentos de este tipo ofrecen la ventaja de permitir que el investigador determine cómo cambia la variable dependiente con respecto a los diferentes niveles de la variable independiente. En este ejemplo, el investigador encontraría la forma en que el comportamiento presentado en el apareamiento varía con respecto a los diferentes niveles de la hormona X. El uso de diversos niveles no sólo muestra una relación válida, si existe alguna, sino que, en situaciones en las que el experimentador no está seguro de cuál nivel puede resultar eficaz, el empleo de varios niveles aumenta la posibilidad de obtener un resultado positivo en el experimento.

En vista de que a menudo es recomendable realizar experimentos con más de dos grupos, usted se preguntará por qué estos experimentos no se analizan de la manera usual. Por ejemplo, si el experimento empleó cuatro grupos independientes, ¿por qué no basta comparar las medias de los grupos por parejas, al utilizar la prueba t para grupos independientes? Es decir, ¿por qué no basta calcular los valores t comparando el grupo 1 con el 2, 3 y 4; el grupo 2 con el 3 y 4, y el grupo 3 con el 4?

La respuesta implica consideraciones acerca del error de tipo I. Usted recordará que cuando establecemos el nivel alfa en 0.05 decimos que estamos dispuestos a arriesgar estar equivocados 5% de las ocasiones cuando rechazamos H_0 . En un experimento con dos grupos, sólo habría un cálculo de t , y compararíamos t_{obt} con t_{crit} para ver si t_{obt} queda en la región crítica para rechazar H_0 . Supongamos que alfa = 0.05. El valor crítico de t en el nivel 0.05 se determinó inicialmente considerando la distribución muestral de t para el número adecuado de grados de libertad y localizando el valor de t , de manera que la proporción del número total de valores t iguales o más extremos que él fuera igual a 0.05. Es decir, si realizamos un muestreo aleatorio de un puntaje t según la distribución t , la probabilidad de que sea $\geq t_{\text{crit}}$ es 0.05.

Ahora bien, ¿qué ocurre cuando realizamos un experimento con muchas comparaciones t , por ejemplo 20? Ya no podemos llevar a cabo el muestreo de un solo valor t de la distribución t , sino el de 20 valores. Está claro que la probabilidad de obtener valores t iguales o mayores que t_{crit} aumenta. Ya no es igual a 0.05. La probabilidad de cometer un error de tipo I aumenta como resultado de la realización de un experimento con muchos grupos y el análisis de los datos con más de una comparación.



PANORAMA DE LA TÉCNICA DEL ANÁLISIS DE VARIANZA DE UN FACTOR O VÍA

El análisis de varianza es una técnica estadística utilizada para analizar los experimentos en que intervienen varios grupos. El uso de la prueba F nos permite hacer una comparación global que nos indica si existe una diferencia significativa entre las medias de los grupos. Así se evita el problema de aumentar la probabilidad de cometer un error de tipo I, que se presenta cuando se evalúan muchos valores t . El análisis de varianza, o ANOVA

como se le suele llamar, se emplea tanto con los diseños de grupos independientes como con los de medidas repetidas. También se usa cuando en el mismo experimento son investigados uno o más factores (variables). En esta sección consideraremos al más sencillo de estos diseños: el diseño simple de grupos formados aleatoriamente. Este diseño se conoce también como análisis de varianza de un factor para grupos independientes. Una tercera denominación que se usa con frecuencia es la del diseño de experimentos con un solo factor, de grupos independientes.⁸ De acuerdo con este diseño, los sujetos son elegidos de la población en forma aleatoria y luego se les asigna a las distintas condiciones de manera igualmente aleatoria, procurando que, de preferencia, haya el mismo número de sujetos en cada condición. Existen tantos grupos independientes como condiciones haya. Si el estudio investiga el efecto de una variable independiente como factor, entonces las condiciones serán los diferentes niveles de la variable independiente utilizada. Cada grupo recibirá un nivel distinto de la variable independiente (por ejemplo, una concentración distinta de la hormona X). Así, en este diseño se analizan los datos de varios grupos independientes.

La hipótesis alternativa que se utiliza en el análisis de varianza es de tipo no direccional. Esta hipótesis establece que los efectos de la variable independiente en una o más condiciones difieren de los que tiene en al menos una de las otras condiciones. La hipótesis nula nos indica que las distintas condiciones son igualmente eficaces, por lo cual los datos de cada grupo son muestras aleatorias extraídas de poblaciones cuyas medias tienen el mismo valor. Si existen k grupos, entonces la hipótesis nula especifica que:

$$\mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

donde μ_1 = media de la población de la cual se ha extraído el grupo 1
 μ_2 = media de la población de la cual se ha extraído el grupo 2
 μ_3 = media de la población de la cual se ha extraído el grupo 3
 μ_k = media de la población de la cual se ha extraído el grupo k

Al igual que la prueba t , el análisis de varianza supone que sólo la media de los datos, y no la varianza, es afectada por la variable independiente. Por lo tanto, el análisis de varianza supone que:

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \dots = \sigma_k^2$$

donde σ_1^2 = varianza de la población de la cual se extrae el grupo 1
 σ_2^2 = varianza de la población de la cual se extrae el grupo 2
 σ_3^2 = varianza de la población de la cual se extrae el grupo 3
 σ_k^2 = varianza de la población de la cual se extrae el grupo k

En esencia, el análisis de varianza divide la variabilidad total de los datos (SC_T) en dos fuentes: a la variabilidad existente dentro de cada grupo se le conoce como la *suma de cuadrados dentro de los grupos* (SC_W), y la variabilidad que hay entre los grupos, llamada la *suma de cuadrados entre los grupos* (SC_B) (véase la figura 15.3). Cada suma de cuadrados se utiliza para calcular una estimación independiente de la varianza de la población H_0 . La estimación basada en la variabilidad dentro de los grupos se conoce como la *estimación de la varianza dentro de los grupos* (s_W^2), y la estimación que se basa en la variabilidad entre los grupos se llama *estimación de la varianza entre grupos* (s_B^2).⁹ Por último, se calcula la razón F , donde:

⁸El uso del análisis de varianza con los diseños de medidas repetidas se estudia en B. J. Winer, D. R. Brown y K. M. Michels, *Statistical Principles in Experimental Design*, tercera edición, McGraw-Hill, Nueva York, 1991.
⁹ s_B^2 se simboliza con frecuencia como MC_B y se denomina "media de cuadrados entre". De manera análoga, s_W^2 se simboliza con frecuencia como MC_W y se denomina "media de cuadrados dentro" o bien "error cuadrado medio".

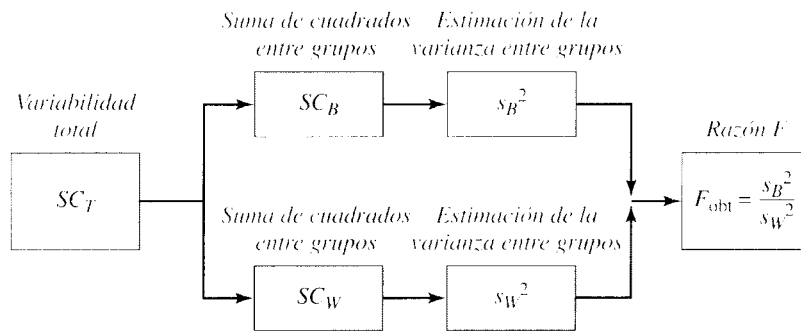


figura 15.3 Panorama de la técnica del análisis de varianza con el diseño simple de grupos aleatorizados.

$$F_{\text{obt}} = \frac{\text{Estimación de la varianza entre grupos } (s_B^2)}{\text{Estimación de la varianza intra grupos } (s_W^2)}$$

Este proceso se presenta en la figura 15.3. La estimación de la varianza entre grupos aumenta con la magnitud del efecto de la variable independiente, mientras que la estimación de la varianza intra grupos no resulta afectada. Por consiguiente, mientras mayor sea la razón F , menos razonable será que se cumpla la hipótesis nula. Igual que en el caso de los demás índices estadísticos, evaluamos F_{obt} comparándola con F_{crit} . Si F_{obt} es igual o mayor que F_{crit} , rechazamos H_0 . Así pues, la regla de decisión establece lo siguiente:

Si $F_{\text{obt}} \geq F_{\text{crit}}$, rechazamos H_0 .

Si $F_{\text{obt}} < F_{\text{crit}}$, conservamos H_0 .

Estimación de la varianza intra grupos, s_W^2

Una estimación de la varianza de la población H_0 , σ^2 , se basa en la variabilidad dentro de cada grupo. Se simboliza como s_W^2 y se determina exactamente en la misma forma que la estimación ponderada s_W^2 que se utiliza en la prueba t para grupos independientes. En el análisis de varianza la llamamos estimación de la varianza *intra grupos*, para distinguirla de la estimación de la varianza *entre grupos* que estudiaremos en la siguiente sección. Usted recordará que en la prueba t para grupos independientes,

$$\begin{aligned} s_W^2 &= \text{estimación ponderada de la varianza, } \sigma^2, \text{ de la población } H_0 \\ &= \text{promedio ponderado de } s_1^2 \text{ y } s_2^2 \\ &= \frac{SC_1 + SC_2}{(n_1 - 1) + (n_2 - 1)} \\ &= \frac{SC_1 + SC_2}{N - 2} \end{aligned}$$

En el análisis de varianza se emplea la misma estimación, excepto que, por lo general, trabajamos con tres o más grupos. Así, para el análisis de varianza,

$$\begin{aligned} s_W^2 &= \text{estimación de la varianza intra grupos} \\ &= \text{estimación ponderada de la varianza de la población } H_0, \sigma^2 \\ &= \text{promedio ponderado de } s_1^2, s_2^2, s_3^2, \dots, \text{ y } s_k^2 \\ &= \frac{SC_1 + SC_2 + SC_3 + \dots + SC_k}{(n_1 - 1) + (n_2 - 1) + (n_3 - 1) + \dots + (n_k - 1)} \end{aligned}$$

donde k = número de grupos
 n_k = número de sujetos en el grupo k
 SC_k = suma de cuadrados del grupo k

Podemos simplificar esta ecuación como:

$$s_w^2 = \frac{SC_1 + SC_2 + SC_3 + \cdots + SC_k}{N - k} \quad \text{ecuación conceptual para la estimación de la varianza intra grupos}$$

donde $N = n_1 + n_2 + n_3 + \cdots + n_k$

El numerador de esta ecuación se llama *suma de cuadrados dentro de grupos*. Se simboliza con SC_W . El denominador es igual al número de grados de libertad para la estimación de la varianza intra grupos. En virtud de que perdemos 1 grado de libertad por cada varianza muestral calculada, y como tenemos k varianzas, existen $N - k$ grados de libertad. Así,

$$s_w^2 = \frac{SC_W}{gl_W} \quad \text{estimación de la varianza intra grupos}$$

donde $SC_W = SC_1 + SC_2 + SC_3 + \cdots + SC_k$ *suma de cuadrados intra grupos*
 $gl_W = N - k$ *grados de libertad intra grupos*

Esta ecuación para SC_W está bien desde el punto de vista conceptual, pero en un cálculo real es mejor utilizar otra ecuación que algebraicamente es equivalente a la ecuación conceptual, pero que resulta más fácil de usar y nos permite reducir los errores de redondeo. Aquí presentamos la ecuación para dicho cálculo y la analizaremos más adelante, cuando analicemos los datos de un experimento:

$$SC_W = \sum_{\text{todos los datos}} X^2 - \left[\frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} + \frac{(\sum X_3)^2}{n_3} + \cdots + \frac{(\sum X_k)^2}{n_k} \right]$$

ecuación para el cálculo de SC_W

Estimación de la varianza entre grupos, s_B^2

La segunda estimación de la varianza para las poblaciones de la hipótesis nula, σ^2 , está basada en la variabilidad *entre* los grupos. Se simboliza con s_B^2 . La hipótesis nula establece que cada grupo es una muestra aleatoria extraída de poblaciones en las cuales $\mu_1 = \mu_2 = \mu_3 = \cdots = \mu_k$. Si la hipótesis nula es correcta, entonces podemos utilizar la variabilidad entre las medias de las muestras para estimar la varianza de estas poblaciones, σ^2 .

Por los resultados obtenidos en el capítulo 12 sabemos que si tomamos todas las muestras posibles de tamaño n de una población y calculamos sus valores medios, la distribución muestral resultante de las medias tendrá una varianza de $\sigma_X^2 = \sigma^2/n$. Al despejar σ^2 , tenemos que $\sigma^2 = n\sigma_X^2$. Si podemos estimar σ_X^2 entonces podemos sustituir esa estimación en la ecuación anterior para obtener una estimación independiente de σ^2 . En virtud de que en un experimento real tenemos varias medias muestrales, podemos emplear la varianza de estas medias para estimar la varianza de todo el conjunto de medias muestrales, $\sigma_{\bar{X}}^2$. Como tenemos k medias muestrales, dividimos entre $k - 1$, de la misma manera que cuando tenemos N datos en bruto, dividimos entre $N - 1$. Así pues,

$$s_{\bar{X}}^2 = \frac{\sum (\bar{X} - \bar{X}_i)^2}{k - 1} \quad \text{estimación de } \sigma_X^2$$

donde \bar{X}_G = gran media (media general de todos los datos combinados)
 k = número de grupos

Al utilizar $s_{\bar{X}}^2$ para nuestra estimación de $\sigma_{\bar{X}}^2$, llegamos a la segunda estimación independiente de σ^2 . Ésta recibe el nombre de *estimación de la varianza entre grupos* y se simboliza con s_B^2 . Como quiera que $\sigma^2 = n\sigma_{\bar{X}}^2$, entonces:

$$s_B^2 = \text{estimación de } \sigma^2 = ns_{\bar{X}}^2$$

Al sustituir el valor de $s_{\bar{X}}^2$,

$$s_B^2 = \frac{n \sum (\bar{X} - \bar{X}_G)^2}{k - 1}$$

Desarrollando la suma, obtenemos:

$$s_B^2 = \frac{n [(\bar{X}_1 - \bar{X}_G)^2 + (\bar{X}_2 - \bar{X}_G)^2 + (\bar{X}_3 - \bar{X}_G)^2 + \dots + (\bar{X}_k - \bar{X}_G)^2]^*}{k - 1}$$

*ecuación conceptual para la
estimación de la varianza entre grupos*

El numerador de esta ecuación se conoce como la *suma de cuadrados entre grupos* y se simboliza como SC_B . El denominador es el número de grados de libertad para la estimación de la varianza entre grupos y se simboliza como gl_B . Así,

$$s_B^2 = \frac{SC_B}{gl_B} \quad \text{estimación de la varianza entre grupos}$$

donde $SC_B = n[(\bar{X}_1 - \bar{X}_G)^2 + (\bar{X}_2 - \bar{X}_G)^2 + (\bar{X}_3 - \bar{X}_G)^2 + \dots + (\bar{X}_k - \bar{X}_G)^2]$ *suma de cuadrados entre grupos*
 $gl_B = k - 1$ *grados de libertad entre grupos*

Debe quedar claro que cuando el efecto de la variable independiente se incrementa, las diferencias entre las medias muestrales aumentan también. Esto hace que $(\bar{X}_1 - \bar{X}_G)^2$, $(\bar{X}_2 - \bar{X}_G)^2$, \dots , $(\bar{X}_k - \bar{X}_G)^2$ incremente, y esto, a su vez, provoca un incremento en SC_B . Por el hecho de que SC_B se encuentra en el numerador, cualquier aumento en éste produce un incremento en s_B^2 . Por lo tanto, la estimación de la varianza entre grupos (s_B^2) aumenta con el efecto de la variable independiente.

Esta ecuación para SC_B está bien desde el punto de vista conceptual, pero para realizar un cálculo real es mejor utilizar otra ecuación. Igual que en el caso de SC_W , existe una ecuación para el cálculo de SC_B que algebraicamente es equivalente a la ecuación conceptual, pero que resulta más fácil de emplear y nos permite reducir los errores de redondeo. Aquí presentamos la ecuación para dicho cálculo y la consideraremos en breve, al analizar los datos de un experimento:

$$SC_B = \left[\frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} + \frac{(\sum X_3)^2}{n_3} + \dots + \frac{(\sum X_k)^2}{n_k} \right] - \frac{\left(\sum_{\text{os datos}} X \right)^2}{N}$$

ecuación para el cálculo de SC_B

*Esto supone que hay n sujetos en cada grupo.

La razón F

Ya observamos anteriormente que s_B^2 aumenta por el efecto de la variable independiente. Sin embargo, como uno de los supuestos del análisis de varianza establece que la variable independiente sólo afecta a la media de cada grupo, y no a la varianza de cada grupo, por lo tanto, la estimación de la varianza dentro de los grupos no cambia a causa del efecto de la variable independiente. Como $F = s_B^2/s_W^2$, F crece por el efecto de la variable independiente. En consecuencia, mientras más grande sea la razón F , tanto más razonable será que la variable independiente produzca un efecto real. En otras palabras s_B^2 es en realidad una estimación de σ^2 más los efectos de la variable independiente, mientras que s_W^2 es tan sólo una estimación de σ^2 . Así pues,

$$F_{\text{obt}} = \frac{s_B^2}{s_W^2} = \frac{\sigma^2 + \text{Efectos de la variable independiente}}{\sigma^2}$$

Mientras más grande sea F_{obt} tanto más razonable será que la variable independiente tenga un efecto real. Por supuesto, F_{obt} debe ser igual o mayor que F_{crit} para que sea posible rechazar H_0 . Si F_{obt} es menor que 1, entonces ni siquiera tenemos que compararla con F_{crit} . En este caso, es obvio que el tratamiento no produjo un efecto significativo y podemos concluir, de inmediato, que conservaremos H_0 .



ANÁLISIS DE DATOS CON LA TÉCNICA DEL ANÁLISIS DE VARIANZA

Hasta ahora hemos sido un tanto teóricos. Resolvamos un problema para ilustrar la técnica del análisis de varianza.

un experimento

Situaciones diversas y tensión

Supongamos que usted quiere determinar si ciertas situaciones producen distintos niveles de tensión. Usted sabe que la cantidad de la hormona corticosterona que circula en la sangre es una buena medida de la tensión que presenta una persona. Usted dividió de manera aleatoria a 15 estudiantes en tres grupos de 5 individuos cada uno. Se midió el nivel de corticosterona de los alumnos del grupo 1, inmediatamente después de que éstos regresaron de vacaciones (tensión baja). Se midió el nivel de corticosterona de los estudiantes del grupo 2, después que éstos asistieron a sus clases durante una semana (tensión moderada). Por último, se midió el nivel de corticosterona de los alumnos del grupo 3 inmediatamente antes de la semana de los exámenes finales (tensión alta). Todas las mediciones se llevaron a cabo a la misma hora del día. Usted registró los datos tal como se presentan en la tabla 15.1. Los datos están expresados en miligramos de corticosterona por cada 100 mililitros de sangre.

1. ¿Cuál es la hipótesis alternativa?
2. ¿Cuál es la hipótesis nula?
3. ¿Cuál es la conclusión? Utilice $\alpha = 0.05$.

SOLUCIÓN

1. Hipótesis alternativa: La hipótesis alternativa sostiene que por lo menos una de las situaciones afecta la tensión de una manera diferente que por lo menos una de las otras situaciones. Por lo tanto, cuando menos una de las medias (μ_1 , μ_2 , o μ_3) es diferente de una de las otras.
2. Hipótesis nula: La hipótesis nula sostiene que las distintas situaciones afectan a la tensión de igual manera. Por lo tanto, los tres conjuntos muestrales de datos son muestras aleatorias extraídas de poblaciones donde $\mu_1 = \mu_2 = \mu_3$.

tabla 15.1 Datos del experimento sobre la tensión

Grupo 1, vacaciones		Grupo 2, clases		Grupo 3, examen final	
X_1	X_1^2	X_2	X_2^2	X_3	X_3^2
2	4	10	100	10	100
3	9	8	64	13	169
7	49	7	49	14	196
2	4	5	25	13	169
6	36	10	100	15	225
20	102	40	338	65	859
$n_1 = 5$		$n_2 = 5$		$n_3 = 5$	
$\bar{X}_1 = 4.00$		$\bar{X}_2 = 8.00$		$\bar{X}_3 = 13.00$	
$\sum_{\text{todos los datos}} X = 125$		$\sum_{\text{todos los datos}} X^2 = 1299$		$\bar{X}_G = \frac{\sum_{\text{todos los datos}} X}{N} = 8.333$	
$N = 15$					

3. Conclusión, con $\alpha = 0.05$: La conclusión se alcanza de la misma manera general que en las otras pruebas de inferencia. En primer lugar, calculamos el valor estadístico apropiado, en este caso, F_{obt} , y luego evaluamos F_{obt} tomando como base su distribución muestral.

A. Cálculo de F_{obt} .

PASO 1: Cálculo de la suma de cuadrados entre grupos, SC_B . Para calcular SC_B , utilizaremos la siguiente ecuación:

$$SC_B = \left[\frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} + \frac{(\sum X_3)^2}{n_3} + \dots + \frac{(\sum X_k)^2}{n_k} \right] - \frac{\left(\sum_{\text{todos los datos}} X \right)^2}{N}$$

ecuación para el cálculo de SC_B

En este problema, como $k = 3$, esta ecuación se reduce a:

$$SC_B = \left[\frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} + \frac{(\sum X_3)^2}{n_3} \right] - \frac{\left(\sum_{\text{todos los datos}} X \right)^2}{N}$$

donde $\sum_{\text{todos los datos}} X =$ suma de todos los datos

Sustituyendo los valores apropiados de la tabla 15.1 en esta ecuación, obtenemos:

$$\begin{aligned} SC_B &= \left[\frac{(20)^2}{5} + \frac{(40)^2}{5} + \frac{(65)^2}{5} \right] - \frac{(125)^2}{15} = 80 + 320 + 845 - 1041.667 \\ &= 203.333 \end{aligned}$$

PASO 2: Cálculo de la suma de cuadrados intra grupos, SC_W . La ecuación para calcular SC_W es la siguiente:

$$SC_W = \sum_{\text{los datos}} X^2 - \left[\frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} + \frac{(\sum X_3)^2}{n_3} + \dots + \frac{(\sum X_k)^2}{n_k} \right]$$

ecuación para el cálculo de SC_W

donde $\sum_{\text{los datos}} X^2 =$ suma de los cuadrados de todos los datos

Como $k = 3$, la ecuación de este problema se reduce a:

$$SC_W = \sum_{\text{los datos}} X^2 - \left[\frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} + \frac{(\sum X_3)^2}{n_3} \right]$$

Sustituyendo los valores apropiados en esta ecuación, obtenemos:

$$\begin{aligned} SC_W &= \sum_{\text{los datos}} X^2 - \left[\frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} + \frac{(\sum X_3)^2}{n_3} \right] \\ &= 1299 - \left[\frac{(20)^2}{5} + \frac{(40)^2}{5} + \frac{(65)^2}{5} \right] \\ &= 54 \end{aligned}$$

PASO 3: Cálculo de la suma total de cuadrados, SC_T . Este paso es sólo una verificación para comprobar que los cálculos realizados en los pasos 1 y 2 son correctos. Como usted recordará, al inicio de la sección sobre el análisis de varianza, página 358, mencionamos que esta técnica divide la variabilidad total en dos partes: la variabilidad intra grupos y entre grupos. La medida de la variabilidad total es SC_T , la medida de la variabilidad intra grupos es SC_W , y la medida de la variabilidad entre grupos es SC_B . Así pues,

$$SC_T = SC_W + SC_B$$

Por medio del cálculo independiente de SC_T , podemos verificar si esta relación sigue siendo válida para los cálculos realizados en los pasos 1 y 2:

$$SC_T = \sum_{\text{los datos}} X^2 - \frac{\left(\sum_{\text{los datos}} X \right)^2}{N}$$

El lector podrá reconocer que esta ecuación es sumamente parecida a la suma de cuadrados en cada muestra, excepto que ahora utilizaremos los datos de todas las muestras como si fuera un solo grupo. Al calcular SC_T , obtenemos:

$$\begin{aligned} SC_T &= \sum_{\text{los datos}} X^2 - \frac{\left(\sum_{\text{los datos}} X \right)^2}{N} \\ &= 1299 - \frac{(125)^2}{15} \\ &= 257.333 \end{aligned}$$

Sustituyendo los valores de SC_T , SC_W y SC_B en la ecuación, obtenemos lo siguiente:

$$\begin{aligned}SC_T &= SC_W + SC_B \\257.333 &= 54 + 203.333 \\257.333 &= 257.333\end{aligned}$$

Observe que si la suma de cuadrados intra grupos más la suma de cuadrados entre grupos no es igual a la suma total de cuadrados, eso indicará que usted ha cometido un error de cálculo. Regrese al problema y verifique los pasos 1, 2 y 3 hasta que la ecuación quede balanceada (considerando los posibles errores a causa del redondeo).

PASO 4: Cálculo de los grados de libertad para cada estimación:

$$\begin{aligned}gl_B &= k - 1 = 3 - 1 = 2 \\gl_W &= N - k = 15 - 3 = 12 \\gl_T &= N - 1 = 15 - 1 = 14\end{aligned}$$

PASO 5: Cálculo de la estimación de la varianza entre grupos, s_B^2 . Las estimaciones de la varianza son simplemente las sumas de cuadrados divididas entre sus grados de libertad. Así,

$$s_B^2 = \frac{SC_W}{gl_B} = \frac{203.333}{2} = 101.667$$

PASO 6: Cálculo de la estimación de la varianza intra grupos, s_W^2 :

$$s_W^2 = \frac{SC_W}{gl_W} = \frac{54}{12} = 4.5$$

PASO 7: Cálculo de F_{obt} . Hemos calculado dos estimaciones independientes de σ^2 : la estimación de la varianza entre grupos y la estimación de la varianza intra grupos. El valor F es la razón de s_B^2 sobre s_W^2 . De esta manera,

$$F_{\text{obt}} = \frac{s_B^2}{s_W^2} = \frac{101.667}{4.5} = 22.59$$

Observe que s_B^2 siempre se escribe en el numerador y s_W^2 se escribe en el denominador.

B. Evaluación de F_{obt} . Como quiera que s_B^2 es una medida del efecto de la variable independiente y también una estimación de σ^2 , tiene que ser mayor que s_W^2 , a menos que el único elemento que intervenga sea el azar. Si $F_{\text{obt}} \leq 1$, está claro que la variable independiente no ha tenido un efecto significativo y concluimos conservando H_0 sin molestarnos siquiera en comparar F_{obt} con F_{crit} . Si $F_{\text{obt}} > 1$, debemos compararla con F_{crit} . Si $F_{\text{obt}} \geq F_{\text{crit}}$, rechazamos H_0 . Según la tabla F del apéndice D, con $\alpha = 0.05$, $gl_{\text{numerador}} = 2$ y $gl_{\text{denominador}} = 12$,

$$F_{\text{crit}} = 3.88$$

Observe que al buscar F_{crit} en la tabla F, es importante fijarse bien en los grados de libertad que corresponden al numerador y al denominador. Si por error hubiera buscado 2 gl para el denominador y 12 gl para el numerador, F_{crit} habría resultado igual a 19.41, una cifra muy distinta de 3.88. Como $F_{\text{obt}} > 3.88$, rechazamos H_0 . Las tres situaciones no son iguales en lo que se refiere a los niveles de tensión que producen. En la tabla 15.2 se presenta un resumen de la solución.

Tabla 15.2 Resumen del problema de análisis de varianza referente a la tensión

Fuente	SC	gl	s ²	F _{obt}
Entre grupos	203.333	2	101.667	22.59*
Intra grupos	54.000	12	4.500	
Total	257.333	14		

*Con $\alpha = 0.05$, $F_{\text{crit}} = 3.88$. Por lo tanto, se rechaza H_0 .

LÓGICA SUBYACENTE AL ANÁLISIS DE VARIANZA DE UN FACTOR

Ahora que hemos realizado los cálculos de un ejemplo ilustrativo, nos gustaría comentar con más detalle la lógica que sirve de fundamento al análisis de varianza de un factor. Ya hemos señalado que este análisis divide la variabilidad total (SC_T) en dos partes: la suma de cuadrados intra grupos (SC_W) y la suma de cuadrados entre grupos (SC_B). Podemos formarnos una idea más clara de esta división si reconocemos que se basa en la sencilla idea de que la desviación de cada dato con respecto a la gran media está formada por dos partes: la desviación del dato con respecto a la media de su propio grupo y la desviación de la media de ese grupo con respecto a la gran media. Si aplicamos esta idea al primer dato del grupo 1, obtenemos:

$$\begin{array}{rcccl}
 \text{Desviación de cada dato} & & \text{Desviación del dato} & & \text{Desviación de la media} \\
 \text{con respecto a la} & = & \text{con respecto a la} & + & \text{de ese grupo con} \\
 \text{gran media} & & \text{media de su grupo} & & \text{respecto a la gran media} \\
 \\
 2 - 8.33 & = & 2 - 4.00 & + & 4.00 - 8.33 \\
 X - \bar{X}_G & = & X - \bar{X}_1 & + & \bar{X}_1 - \bar{X}_G \\
 \downarrow & & \downarrow & & \downarrow \\
 SC_T & = & SC_W & + & SC_B
 \end{array}$$

Observe que al elevar al cuadrado el término de la izquierda ($X - \bar{X}_G$) y sumar, para todos los datos, obtenemos SC_T . Así,

$$SC_T = \sum_{\text{todos los datos}} (X - \bar{X}_G)^2$$

Al elevar al cuadrado el término intermedio ($X - \bar{X}_1$) y sumar, para todos los datos (por supuesto, tenemos que restar la media de grupo apropiada de cada dato) obtenemos SC_W . Así,

$$\begin{aligned}
 SC_W &= SC_1 + SC_2 + SC_3 \\
 &= \sum (X - \bar{X}_1)^2 + \sum (X - \bar{X}_2)^2 + \sum (X - \bar{X}_3)^2
 \end{aligned}$$

Es importante observar que, como los sujetos dentro de cada grupo reciben el mismo nivel de la variable independiente, la variabilidad entre los datos dentro de cada grupo no

se puede deber a diferencias en el efecto de la variable independiente. Así, la suma de cuadrados intra grupos (SC_W) no es una medida del efecto de la variable independiente. Como $SC_W/gl_W = s_W^2$, esto significa que la estimación de la varianza intra grupos, (s_W^2) tampoco es una medida del efecto real de la variable independiente, sino que proporciona una estimación de la variabilidad inherente de los propios datos. Así, s_W^2 es una estimación de σ^2 que no resulta afectada por diferencias de tratamiento.

El último término de la ecuación que divide la variabilidad del dato 2 con respecto a la gran media es $\bar{X}_1 - \bar{X}_G$. Cuando este término se eleva al cuadrado y se suma para todos los datos, se convierte en SC_B . Así,


$$SC_B = n(\bar{X}_1 - \bar{X}_G)^2 + n(\bar{X}_2 - \bar{X}_G)^2 + n(\bar{X}_3 - \bar{X}_G)^2$$

Como ya hemos analizado, SC_B es sensible al efecto de la variable independiente, pues mientras mayor sea este efecto, tanto más diferirán las medias de cada grupo entre sí y, por lo tanto, serán diferentes de \bar{X}_G . Como $SC_B/gl_B = s_B^2$, esto significa que la estimación de la varianza entre grupos (s_B^2) también es sensible al efecto real de la variable independiente. Así, s_B^2 nos proporciona una estimación de σ^2 más los efectos de la variable independiente. Puesto que

$$F_{\text{obt}} = \frac{s_B^2}{s_W^2} = \frac{\sigma^2 + \text{efectos de la variable independiente}}{\sigma^2}$$

mientras mayor sea F_{obt} menos razonable será la explicación de la hipótesis nula. Si la variable independiente no tiene efecto alguno, entonces tanto s_B^2 como s_W^2 son estimaciones independientes de σ^2 y su razón se distribuye como F con $gl = gl_B$ (numerador) y gl_W (denominador). Evaluamos la hipótesis nula al comparar F_{obt} con F_{crit} . Si $F_{\text{obt}} \geq F_{\text{crit}}$, rechazamos H_0 .

Tratemos de resolver otro problema para practicar.



Un profesor quiere encontrar la mejor forma de presentar un importante tema frente al grupo que tiene a su cargo. Para eso puede escoger una de las tres opciones siguientes: 1) dar clase, 2) dar clase y asignar una lectura complementaria o 3) proyectar una película y asignar una lectura complementaria. Él decide hacer un experimento para evaluar las tres opciones. Entonces solicita 27 voluntarios de su grupo y asigna aleatoriamente 9 a cada una de las tres condiciones. Bajo la condición 1, él da clase a los estudiantes. Bajo la condición 2, él da la clase y asigna una lectura complementaria. Bajo la condición 3, los alumnos presencian una película acerca del tema y realizan la misma lectura complementaria que los estudiantes de la condición 2. Después se aplica a los alumnos un examen sobre el material. Se obtuvieron los siguientes datos (porcentaje de respuestas correctas):

(continúa)

Clase, condición 1		Clase + lectura, condición 2		Película + lectura, condición 3	
X_1	X_1^2	X_2	X_2^2	X_3	X_3^2
92	8464	86	7396	81	6561
86	7396	93	8649	80	6400
87	7569	97	9409	72	5184
76	5776	81	6561	82	6724
80	6400	94	8836	83	6889
87	7569	89	7921	89	7921
92	8464	98	9604	76	5776
83	6889	90	8100	88	7744
84	7056	91	8281	83	6889
767	65 583	819	74 757	734	60 088
$n_1 = 9$		$n_2 = 9$		$n_3 = 9$	
$\bar{X}_1 = 85.222$		$\bar{X}_2 = 91$		$\bar{X}_3 = 81.556$	
$\sum_{\text{todos los datos}} X = 2320$		$\sum_{\text{todos los datos}} X^2 = 200 428$		$\bar{X}_G = \frac{\sum_{\text{todos los datos}} X}{N} = 85.926$	
$N = 27$					

- a. ¿Cuál es la hipótesis nula global?
b. ¿Cuál es la conclusión? Utilice $\alpha = 0.05$.

SOLUCIÓN

- a. Hipótesis nula: La hipótesis nula establece que los distintos métodos para presentar el material tienen la misma eficacia. Por lo tanto, $\mu_1 = \mu_2 = \mu_3$.
b. Conclusión, utilizando $\alpha = 0.05$: Para evaluar H_0 debemos calcular F_{obt} y luego evaluarla mediante su distribución muestral.

A. Cálculo de F_{obt} .

PASO 1: Cálculo de SC_B :

$$\begin{aligned}
 SC_B &= \left[\frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} + \frac{(\sum X_3)^2}{n_3} \right] - \frac{\left(\sum_{\text{todos los datos}} X \right)^2}{N} \\
 &= \left[\frac{(767)^2}{9} + \frac{(819)^2}{9} + \frac{(734)^2}{9} \right] - \frac{(2320)^2}{27} \\
 &= 408.074
 \end{aligned}$$

PASO 2: Cálculo de SC_W :

$$\begin{aligned}
 SC_W &= \sum_{\text{todos los datos}} X^2 - \left[\frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} + \frac{(\sum X_3)^2}{n_3} \right] \\
 &= 200\,428 - \left[\frac{(767)^2}{9} + \frac{(819)^2}{9} + \frac{(734)^2}{9} \right] \\
 &= 671.778
 \end{aligned}$$

PASO 3: Cálculo de SC_T :

$$\begin{aligned}
 SC_T &= \sum_{\text{todos los datos}} X^2 - \frac{\left(\sum_{\text{todos los datos}} X \right)^2}{N} \\
 &= 200\,428 - \frac{(2320)^2}{27} \\
 &= 1079.852
 \end{aligned}$$

Este paso es una verificación para ver si SC_B y SC_W fueron calculadas correctamente. En caso afirmativo, $SC_T = SC_B + SC_W$. Esta verificación se ilustra aquí:

$$\begin{aligned}
 SC_T &= SC_B + SC_W \\
 1079.852 &= 408.074 + 671.778 \\
 1079.852 &= 1079.852
 \end{aligned}$$

PASO 4: Cálculo de los grados de libertad:

$$\begin{aligned}
 gl_B &= k - 1 = 3 - 1 = 2 \\
 gl_W &= N - k = 27 - 3 = 24 \\
 gl_T &= N - 1 = 27 - 1 = 26
 \end{aligned}$$

PASO 5: Cálculo de s_B^2 :

$$s_B^2 = \frac{SC_B}{gl_B} = \frac{408.074}{2} = 204.037$$

PASO 6: Cálculo de s_W^2 :

$$s_W^2 = \frac{SC_W}{gl_W} = \frac{671.778}{24} = 27.991$$

PASO 7: Cálculo de F_{obt} :

$$F_{\text{obt}} = \frac{s_B^2}{s_W^2} = \frac{204.037}{27.991} = 7.29$$

- B.** Evaluar F_{obt} como $\alpha = 0.05$, $gl_{\text{numerador}} = 2$ y $gl_{\text{denominador}} = 24$, a partir de la tabla F,

$$F_{\text{crit}} = 3.40$$

(continúa)

Como $F_{\text{obt}} > 3.40$, rechazamos H_0 . No todos los métodos de presentación tienen la misma eficacia. La solución se resume en la tabla 15.3.

tabla 15.3 Resumen del análisis de varianza para el experimento sobre los métodos de presentación

Fuente	SC	gl	s^2	F_{obt}
Entre grupos	408.074	2	204.037	7.29*
Intra grupos	671.778	24	27.991	
Total	1079.852	26		

*Con $\alpha = 0.05$, $F_{\text{crit}} = 3.40$. Por lo tanto, se rechaza H_0 .



LA RELACIÓN ENTRE EL ANÁLISIS DE VARIANZA Y LA PRUEBA t

Cuando el estudio implica sólo dos grupos independientes y estamos poniendo a prueba la hipótesis nula $\mu_1 = \mu_2$, podemos utilizar la prueba t para grupos independientes o el análisis de varianza. En esos casos resulta que $t^2 = F$. Para ilustrar este punto, examinaremos una situación adecuada para la prueba t y el análisis de varianza.

ejemplo

Citas: Las jóvenes de clubes femeninos estudiantiles versus las de dormitorios

Un sociólogo quiere determinar quiénes tienen citas con más frecuencia en una universidad, si las estudiantes que pertenecen a los clubes femeninos o las de los dormitorios. Él elige al azar 12 mujeres de los clubes femeninos y 12 de los dormitorios, y registra el número de citas que tiene cada una durante el siguiente mes. Los resultados se presentan en la tabla 15.4.

tabla 15.4 Resultados del experimento de las citas

Estudiantes de los clubes femeninos		Estudiantes de los dormitorios	
X_1	X_1^2	X_2	X_2^2
8	64	9	81
5	25	7	49
6	36	3	9
4	16	4	16
12	144	4	16
7	49	8	64
9	81	7	49
10	100	5	25
5	25	8	64
3	9	6	36
7	49	3	9
Σ	Σ	Σ	Σ
81	623	69	443
$n_1 = 12$		$n_2 = 12$	
$\bar{X}_1 = 6.750$		$\bar{X}_2 = 5.750$	

A partir de los datos de la tabla 15.4, ¿cuál es la conclusión del sociólogo? Utilice $\alpha = 0.05_{2 \text{ colas}}$.

SOLUCIÓN

La hipótesis nula es que $\mu_1 = \mu_2$. Como éste es un diseño de grupos independientes, para evaluar esta hipótesis podemos utilizar la prueba t para grupos independientes o el análisis de varianza. Primero examinaremos los datos con la prueba t , de la siguiente manera:

$$\begin{aligned}
 SC_1 &= \sum X_1^2 - \frac{(\sum X_1)^2}{N} & SC_2 &= \sum X_2^2 - \frac{(\sum X_2)^2}{N} \\
 &= 623 - \frac{(81)^2}{12} & &= 443 - \frac{(69)^2}{12} \\
 &= 76.250 & &= 46.250 \\
 t_{\text{obt}} &= \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{SC_1 + SC_2}{n(n-1)}}} = \frac{6.750 - 5.750}{\sqrt{\frac{76.250 + 46.250}{12(11)}}} \\
 &= \frac{1.000}{0.963} = 1.038 = 1.04
 \end{aligned}$$

Con $\alpha = 0.05_{2 \text{ colas}}$ y $gl = N - 2 = 22$, según la tabla D del apéndice D,

$$t_{\text{crit}} = +2.074$$

Puesto que $|t_{\text{obt}}| < 2.074$, conservamos H_0 . No podemos concluir quiénes tienen citas con mayor frecuencia, las jóvenes de los clubes femeninos o las de los dormitorios.

También podemos resolver este problema con el análisis de varianza. A continuación presentamos la solución:

PASO 1: Cálculo de SC_B :

$$\begin{aligned}
 SC_B &= \left[\frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} \right] - \frac{\left(\sum_{\text{todos los datos}} X \right)^2}{N} \\
 &= \left[\frac{(81)^2}{12} + \frac{(69)^2}{12} \right] - \frac{(150)^2}{24} \\
 &= 6.000
 \end{aligned}$$

PASO 2: Cálculo de SC_W :

$$\begin{aligned}
 SC_W &= \sum_{\text{todos los datos}} X^2 - \left[\frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} \right] \\
 &= 1066 - \left[\frac{(81)^2}{12} + \frac{(69)^2}{12} \right] \\
 &= 122.500
 \end{aligned}$$

PASO 3: Cálculo de SC_T :

$$\begin{aligned}
 SC_T &= \sum_{\text{todos los datos}} X^2 - \frac{\left(\sum_{\text{todos los datos}} X \right)^2}{N} \\
 &= 1066 - \frac{(150)^2}{24} \\
 &= 128.500
 \end{aligned}$$

Como verificación de lo anterior:

$$\begin{aligned}SC_T &= SC_B + SC_W \\128.500 &= 6.000 + 122.500 \\128.500 &= 128.500\end{aligned}$$

PASO 4: Cálculo de los grados de libertad:

$$\begin{aligned}gl_B &= k - 1 = 2 - 1 = 1 \\gl_W &= N - k = 24 - 2 = 22 \\gl_T &= N - 1 = 24 - 1 = 23\end{aligned}$$

PASO 5: Cálculo de s_B^2 :

$$s_B^2 = \frac{SC_B}{gl_B} = \frac{6.000}{1} = 6.000$$

PASO 6: Cálculo de s_W^2 :

$$s_W^2 = \frac{SC_W}{gl_W} = \frac{122.500}{22} = 5.568$$

PASO 7: Cálculo de F_{obt} :

$$F_{\text{obt}} = \frac{s_B^2}{s_W^2} = \frac{6.000}{5.568} = 1.078 = 1.08$$

Con $\alpha = 0.05$, $gl_{\text{numerador}} = 1$ y $gl_{\text{denominador}} = 22$, a partir de la tabla F,

$$F_{\text{crit}} = 4.30$$

Como $F_{\text{obt}} < 4.30$, conservamos H_0 . No podemos concluir quiénes tienen citas con mayor frecuencia, si las jóvenes de los clubes femeninos o las de los dormitorios.

Observe que hemos llegado a la misma conclusión al emplear la prueba t y la prueba F . Podemos comparar las dos estimaciones estadísticas obtenidas y los valores críticos asociados a ellos como se muestra aquí:

$$\begin{aligned}t_{\text{obt}}^2 &= F_{\text{obt}} & t_{\text{crit}}^2 &= F_{\text{crit}} \\(1.04)^2 &= 1.08 & (2.074)^2 &= 4.30\end{aligned}$$

Así, en el diseño de grupos independientes, cuando tenemos dos grupos y estamos poniendo a prueba la hipótesis nula, donde $\mu_1 = \mu_2$, $t_{\text{obt}}^2 = F$ y $t_{\text{crit}}^2 = F_{\text{crit}}$. Por lo tanto, las dos pruebas proporcionan los mismos resultados.



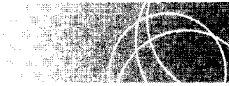
SUPUESTOS SUBYACENTES AL ANÁLISIS DE VARIANZA

Los supuestos subyacentes al análisis de varianza son similares a los de la prueba t para grupos independientes:

1. Las poblaciones de las cuales se extraen las muestras tienen una distribución normal.

2. Las muestras se extraen de poblaciones que tienen la misma varianza. Como señalamos en el capítulo 14, en relación con la prueba t para grupos independientes, este supuesto es el de la homogeneidad de la varianza. También el análisis de varianza supone la homogeneidad de la varianza.*

Al igual que la prueba t , el análisis de varianza es una prueba robusta. Ésta se afecta en forma mínima por las violaciones a la normalidad poblacional. También es relativamente insensible a las violaciones a la homogeneidad de la varianza, siempre que las muestras sean del mismo tamaño.†



MAGNITUD DEL EFECTO

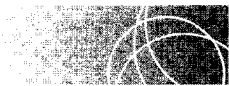
En el capítulo 14 describimos un método para determinar la magnitud del efecto en el caso del experimento básico de dos condiciones, para grupos independientes. El método incluía el cálculo de una estimación de omega al cuadrado ($\hat{\omega}^2$), lo cual nos proporcionaba una estimación del porcentaje de la variabilidad total de la variable dependiente que puede ser atribuido a la variable independiente. Cuanto más grande sea este porcentaje, tanto mayor será el efecto de la variable independiente. El método que emplearemos para determinar la magnitud del efecto con experimentos en los que interviene el análisis de varianza de un factor, para grupos independientes, es muy similar. Una vez más, calcularemos una estimación del porcentaje de la variabilidad total de la variable dependiente que es imputable a la variable independiente, y de nuevo el concepto estadístico que utilizaremos será $\hat{\omega}^2$. Sin embargo, la ecuación para estimar omega al cuadrado con experimentos en los que se usa el ANOVA de un factor, para variables independientes, es diferente del que aplicamos en el capítulo 14. Esta ecuación para experimentos en los que participan k grupos es la siguiente:

$$\hat{\omega}^2 = \frac{SC_B - (k - 1)s_W^2}{SC_T + s_W^2}$$

Para los datos del experimento sobre la tensión, página 362, $SC_B = 203.333$, $SC_T = 257.333$, $s_W^2 = 4.500$ y $k = 3$. Por lo tanto,

$$\hat{\omega}^2 = \frac{203.333 - (3 - 1)4.500}{257.333 + 4.500} = 0.7422$$

Las situaciones de tensión representan 74.2% de la variabilidad en los niveles de corticosterona. Éste es un efecto de gran magnitud.



POTENCIA DEL ANÁLISIS DE VARIANZA

La potencia del análisis de varianza resulta afectada por las mismas variables y en la misma forma que sucedía en el caso de la prueba t para grupos independientes. Como usted recordará, en el caso de la prueba t para grupos independientes, la potencia resulta afectada por lo siguiente:

*Véase la nota al pie de la página 339, capítulo 14. Algunos estadísticos limitarían también el uso del ANOVA a los datos con escala de intervalo o de razón. Para un estudio de este punto, véanse las referencias en la nota al pie de la página 31 en el capítulo 2.

†Para un análisis amplio de estos puntos, véanse G. V. Glass, P. D. Peckham y J. R. Sanders, "Consequences of Failure to Meet the Assumptions Underlying the Use of Analysis of Variance and Covariance", *Review of Educational Research*, 42 (1972), pp. 237-288.

1. La potencia varía en proporción directa con respecto a N . Cuando N se incrementa, la potencia aumenta.
2. La potencia varía en proporción directa con la magnitud del efecto real de la variable independiente. La potencia de la prueba t para detectar un efecto real es mayor con efectos grandes que con efectos pequeños.
3. La potencia varía en proporción inversa con la variabilidad de la muestra. Cuanto más grande sea la variabilidad de la muestra, tanto más baja será la potencia para detectar un efecto real.

Examinemos ahora cada una de estas variables y la forma en que afectan el análisis de varianza. Esta exposición se puede comprender con más facilidad si se usa como referencia la siguiente ecuación de F_{obt} , para un experimento en el que participan tres grupos.

$$F_{\text{obt}} = \sqrt{\frac{s_B^2}{s_w^2}} = \sqrt{\frac{n[(\bar{X}_1 - \bar{X}_G)^2 + (\bar{X}_2 - \bar{X}_G)^2 + (\bar{X}_3 - \bar{X}_G)^2]/2}{(SC_1 + SC_2 + SC_3)/(N - 3)}}$$

Potencia y N

Como es obvio, cualquier cosa que incremente F_{obt} aumentará también la potencia. A medida que se incrementa N , el número total de los sujetos que participan en el experimento, lo mismo sucede con n , el número de sujetos incluidos en cada grupo. Los incrementos en cualquiera de estas variables se traducen en un aumento de F_{obt} . Esto puede visualizarse en la siguiente forma. Tomando como referencia la ecuación F_{obt} cuando N aumenta, en virtud de que está en el denominador de la estimación de la varianza s_w^2 , intra grupos, entonces s_w^2 disminuye. Puesto que s_w^2 está en el denominador de la ecuación de F_{obt} entonces, F_{obt} aumenta. En cuanto a n , por el hecho de que se encuentra en el numerador de la ecuación de F_{obt} y es un multiplicador de valores positivos, los incrementos de n producen como resultado un aumento en la estimación s_B^2 de la varianza entre grupos. Como s_B^2 está en el numerador de la ecuación de F_{obt} los incrementos en s_B^2 producen un aumento en F_{obt} . Como se dijo con anterioridad, cualquier cosa que incremente F_{obt} hace que la potencia aumente también. *Así pues, los incrementos de N y n se traducen en un aumento de la potencia.*

Potencia y el efecto real de la variable independiente

Cuanto más grande sea el efecto real de la variable independiente, tanto mayores serán los valores de $(\bar{X}_1 - \bar{X}_G)^2$, $(\bar{X}_2 - \bar{X}_G)^2$ y $(\bar{X}_3 - \bar{X}_G)^2$. Los aumentos de esos valores producen un incremento en s_B^2 . Como s_B^2 está en el numerador de la ecuación de F_{obt} los incrementos de s_B^2 producen como resultado un aumento en F_{obt} . *Así pues, cuanto más grande sea el efecto real de la variable independiente, tanto más alta será la potencia.*

Potencia y la variabilidad de la muestra

SC_1 (la suma de los cuadrados del grupo 1), SC_2 (la suma de los cuadrados del grupo 2), y SC_3 (la suma de los cuadrados del grupo 3) son medidas de la variabilidad dentro de cada grupo. Los incrementos de SC_1 , SC_2 y SC_3 dan lugar a un aumento en la estimación de la varianza dentro de grupos, s_w^2 . En virtud de que s_w^2 está en el denominador de la ecuación de F_{obt} los incrementos en s_w^2 se traducen en una disminución de F_{obt} . *Así pues, los incrementos de la variabilidad se traducen en disminuciones de la potencia.*



Mucho ruido y pocas nueces

En un anuncio publicado recientemente por la revista *Rawlings Golf Company*, Rawlings afirma que ha desarrollado una nueva pelota de golf que recorre mayores distancias. La pelota lleva el nombre de Tony Penna DB (siglas en inglés de pelota de distancia). En el anuncio, Rawlings no sólo menciona términos como *núcleo de alto rebote, cubierta Surlyn, acción centrífuga*, etcétera, para explicar por qué es razonable pensar que esta pelota recorre mayores distancias; también ha contratado a un instituto de pruebas del consumidor para que realice un experimento y determine si en realidad la pelota Tony Penna DB “llega más lejos”. En este experimento se evaluán seis marcas diferentes de pelotas. Cada uno de 51 golfistas golpea 18 pelotas nuevas (3 de cada marca) con un palo de golf. La distancia media recorrida por cada pelota se presenta en la siguiente tabla:

1. Tony Penna DB	254.57 yardas
2. Titleist Pro Trajectory	252.50 yardas

3. Wilson Pro Staff	249.24 yardas
4. Titleist DT	249.16 yardas
5. Spalding Top-Flite	247.12 yardas
6. Dunlop Blue Max	244.22 yardas

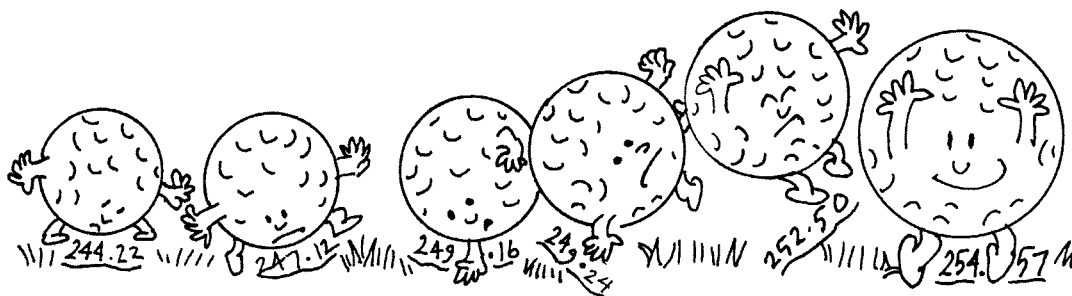
Aunque no se informó acerca de una prueba de inferencia, el anuncio proclama esta conclusión: “Como es evidente, aunque no le podamos prometer que *usted* enviará la pelota a 250 yardas, sí podemos ofrecerle un margen competitivo, de una yarda o dos; pero un margen es un margen”. Como ahora usted dispone de una amplia base de estadística inferencial, ¿cuál sería su respuesta a este anuncio?

Respuesta En primer lugar, creo que usted debería recomendar a la compañía que realice una investigación de evaluación en la cual compare su producto con los de la competencia, basándose en una variable dependiente muy importante. También se recomienda encargar a una organización imparcial la realización de dicha investigación. Por último, se debe informar de los resultados y llamar la atención de los lectores hacia el hecho de que las diferencias entre las pelotas son

muy pequeñas (aunque, desde luego, en el texto del anuncio se pretende lograr un resultado un tanto diferente).

Una de las principales críticas a este anuncio (que ya es muy conocida para nosotros) es que no se nos ha aclarado si estos resultados son significativos desde el punto de vista estadístico. Sin establecer este punto, la explicación más razonable de las diferencias sería el “azar”. Por supuesto, si el azar es la explicación correcta, ¡entonces el uso de la pelota Tony Penna DB no le dará ni siquiera una o dos yardas de ventaja! Antes de considerar seriamente la afirmación de una supuesta superioridad, el fabricante debe indicar que las diferencias son significativas desde el punto de vista estadístico. Sin esta afirmación, por regla general creo que debemos suponer que las diferencias evaluadas no fueron significativas, en cuyo caso el azar queda como la única explicación razonable de los datos. (Por cierto, ¿qué prueba de inferencia habría utilizado usted? ¿Respondió que el análisis de varianza? ¡Muy bien!)

(continúa)





Mucho ruido y pocas nueces (continuación)

Ahora bien, digamos que se llevó a cabo la prueba de inferencia adecuada y que los datos son significativos desde el punto de vista estadístico. Todavía tendríamos que preguntarnos: “¿Y eso qué? Aunque los resultados sean significativos, ¿vale la pena preocuparse por la magnitud del efecto?” En lo que se refiere a la diferencia entre las yardas medidas en las dos primeras marcas, creo que la respuesta es “no”. A pesar de que fuera un apa-

sionado del golf, no veo cómo una o dos yardas podrían establecer una diferencia práctica en mi nivel de juego. Con toda probabilidad, mi promedio en los 18 hoyos no cambiaría ni siquiera por un golpe, independientemente de qué pelota haya utilizado. Por otra parte, si acostumbrara usar pelotas Blue Max de Dunlop, estos resultados me inducirían a probar alguna de las dos primeras marcas. En relación con las marcas que ocuparon los lugares tercero, cuarto y quinto, una persona sensata podría hacer cualquier cosa. Si no hubiera diferencia de precios, creo que podría utilizar una de las dos primeras marcas en plan de prueba.

En resumen, he intentado destacar dos puntos. El primero es que *al afirmar la superioridad del producto con base en los datos de la muestra, se debería informar si los resultados son significativos desde el punto de vista estadístico*. El segundo es que la “significación estadística” y la “importancia” son conceptos distintos. Una vez establecida la significación estadística, debemos considerar la magnitud del efecto para averiguar si es lo suficientemente grande para justificar que modifiquemos nuestro comportamiento.

RESUMEN

En este capítulo hemos estudiado la prueba F y el análisis de varianza. La prueba F es, en esencia, la razón entre dos estimaciones independientes de la varianza de una misma población, σ^2 . La distribución F es una familia de curvas que varían según los grados de libertad. Como quiera que F_{obt} es un cociente, existen dos valores para los grados de libertad, uno para el numerador y otro para el denominador. La distribución F 1) está sesgada positivamente, 2) no tiene valores negativos y 3) posee una mediana aproximadamente igual a 1, dependiendo de la n de las estimaciones.

La técnica del análisis de varianza se utiliza en conjunción con experimentos en los que hay más de dos grupos independientes. Básicamente, esta técnica permite comparar las medias de los distintos grupos en una sola evaluación global, con lo cual se evita el aumento de la probabilidad de cometer un error de tipo I al realizar muchas pruebas t . En el análisis de varianza de un factor, la variabilidad total de los datos (SC_T) se divide en dos partes: la variabilidad que existe dentro de cada grupo, llamada suma de cuadrados in-

tra grupos (SC_W), y la variabilidad que existe entre los grupos, llamada suma de cuadrados entre grupos (SC_B). Cada suma de cuadrados se emplea para formar una estimación independiente de la varianza de la población de la hipótesis nula. Por último, se calcula una razón F , donde la estimación de la varianza entre grupos (s_B^2) aparece en el numerador y la estimación de la varianza intra grupos (s_W^2) se muestra en el denominador. Como la estimación de la varianza entre grupos crece con el efecto de la variable independiente, y la estimación de la varianza intra grupos se mantiene constante, mientras más grande sea la razón F , menos razonable será aceptar la hipótesis nula. Para evaluar F_{obt} la comparamos con F_{crit} . Si $F_{\text{obt}} \geq F_{\text{crit}}$, rechazamos la hipótesis nula y concluimos que por lo menos una de las condiciones difiere, de manera significativa, de una de las demás conclusiones por lo menos.

A continuación, estudiamos los supuestos subyacentes al análisis de varianza. Existen dos supuestos: 1) las poblaciones de las cuales se extraen las muestras deben ser normales y 2) debe haber homogeneidad de

la varianza. La prueba F es robusta con respecto a las violaciones de la normalidad y la homogeneidad de la varianza.

Después de analizar los supuestos, presentamos una sección sobre la forma de estimar la magnitud del efecto de la variable independiente. Observamos que esto es muy similar a la estimación de la magnitud del efecto mediante la prueba t para grupos independientes. Una vez más calculamos $\hat{\omega}^2$ aun cuando, con el análisis de varianza, la ecuación para $\hat{\omega}^2$ es diferente de la que se usa en la prueba t . En ambos casos, lo que

se calcula es el porcentaje de la variabilidad total de la variable dependiente que se puede atribuir a la variable independiente. Cuanto más grande sea el porcentaje, tanto mayor será el tamaño del efecto.

Finalmente, presentamos una sección sobre la potencia del análisis de varianza. Igual que en el caso de la prueba t , la potencia del ANOVA cambia en proporción directa con respecto a N y la magnitud del efecto real, y en proporción inversa con la variabilidad de la muestra.

TÉRMINOS IMPORTANTES

Análisis de varianza (página 356)	Estimación de la varianza intra grupos (s_w^2) (página 358)	Gran media (\bar{X}_G) (página 361)
Análisis de varianza de un factor con diseño de grupos independientes (página 357)	Estimación de la varianza entre grupos (s_B^2) (página 358)	Prueba F (página 355)
Diseño simple con grupos aleatorizados (página 358)	Experimento de un solo factor con diseño de grupos independientes (página 358)	Suma de cuadrados intra grupos (SC_W) (página 358)
Distribución muestral de F (página 355)	F_{crit} (página 356)	Suma de cuadrados entre grupos (SC_B) (página 358)
		Variabilidad total (SC_T) (página 358)

PREGUNTAS Y PROBLEMAS

- Identifique y defina los conceptos de la sección “términos importantes”.
- ¿Cuáles son las características de la distribución F ?
- ¿Cuáles son las ventajas de realizar experimentos con más de dos grupos o condiciones?
- Al realizar un experimento con muchos grupos, ¿cuál es el problema de llevar a cabo pruebas t entre todos los grupos posibles sin correcciones? ¿Por qué el uso del análisis de varianza evita ese problema?
- La técnica del análisis de varianza estudia la variabilidad de los datos. Sin embargo, un valor F significativo indica que existe por lo menos una diferencia significativa entre las medias de las condiciones. ¿En qué forma el estudio de la variabilidad de los datos permite extraer conclusiones acerca de las medias de las condiciones?
- ¿Cuáles son los pasos para formar una razón F al utilizar la técnica del análisis de varianza de un factor?
- En el análisis de varianza, si F_{obt} es menor que 1, ni siquiera necesitamos compararla con F_{crit} . Es obvio que la variable independiente no produjo un efecto significativo. ¿Por qué?
- ¿Cuáles son los supuestos fundamentales o subyacentes al análisis de varianza?
- El análisis de varianza es una técnica no direccional, y sin embargo, utiliza una evaluación con una cola. ¿Es correcta esta afirmación? Explique.
- Determine F_{crit} para las siguientes situaciones:
 - $\text{gl}(\text{numerador}) = 2$, $\text{gl}(\text{denominador}) = 16$, $\alpha = 0.05$
 - $\text{gl}(\text{numerador}) = 3$, $\text{gl}(\text{denominador}) = 36$, $\alpha = 0.05$
 - $\text{gl}(\text{numerador}) = 3$, $\text{gl}(\text{denominador}) = 36$, $\alpha = 0.01$
 ¿Qué ocurre con F_{crit} cuando los grados de libertad aumentan y alfa se mantiene constante? ¿Qué ocurre con F_{crit} cuando los grados de libertad se mantienen constantes y alfa se vuelve más restrictiva?
- Verifique que $F = t^2$ cuando sólo hay dos grupos independientes, al calcular F_{obt} para los datos del problema de práctica 14.4 de la página 337.
- ¿Cuáles son las variables que influyen en la potencia de la técnica del análisis de varianza de un factor?
- Para cada una de las variables identificadas en la pregunta 12, indique cómo resulta afectada la potencia si la variable se incrementa. Aplique la ecuación de F_{obt} página 374, para justificar su respuesta.

14. La siguiente tabla es el resumen de un análisis de varianza de un factor, con grupos independientes, en la cual falta una parte de la información.

Fuente	SC	gl	s ²	F _{obt}
Entre grupos	1253.68	3		
Intra grupos				
Total	5016.40	39		

- Proporcione usted los valores faltantes.
 - ¿Cuántos grupos tiene el experimento?
 - Si existe el mismo número de sujetos en cada grupo, ¿cuántos sujetos tiene cada grupo?
 - ¿Cuál es el valor de F_{crit} , si $\alpha = 0.05$?
 - ¿Existe un efecto significativo?
15. Supongamos que usted es un nutriólogo a quien se le ha pedido que determine si existe una diferencia en el contenido de azúcar de las tres principales marcas de cereal para el desayuno (A, B y C). Para evaluar la cantidad de azúcar que contienen los cereales, usted selecciona una muestra aleatoria de seis paquetes de cada marca y lleva a cabo un examen químico del contenido de azúcar de cada uno. Así midió los siguientes gramos de azúcar:

Cereal para el desayuno		
A	B	C
1	7	5
4	5	4
3	3	4
3	6	5
2	4	7
5	7	8

- Utilice las ecuaciones conceptuales del análisis de varianza de un factor y determine si alguna de las marcas tiene un contenido de azúcar diferente. Utilice $\alpha = 0.05$.
 - La misma pregunta de la parte **a**, pero ahora utilice las ecuaciones para el cálculo. ¿Cuáles prefiere usted? ¿Por qué? salud
16. Un investigador realiza un experimento para determinar si la pérdida del sueño afecta la capacidad de mantener constante la atención. Quince individuos son divididos de manera aleatoria entre

los siguientes tres grupos de 5 sujetos cada uno: el grupo 1, que duerme de manera normal (de 7 a 8 horas); el grupo 2, que no duerme durante 24 horas, y el grupo 3, que no duerme durante 48 horas. A los tres grupos se les somete a la misma prueba de percepción auditiva, la cual consiste en hacerlos oír tonos de medio segundo a intervalos irregulares durante una hora. En ciertas ocasiones, un tono es ligeramente más corto que los demás. La tarea del sujeto consiste en detectar los tonos más cortos. Así se observaron los siguientes porcentajes de percepciones correctas:

Sueño normal	Sin dormir durante 24 horas	Sin dormir durante 48 horas
85	60	60
83	58	48
76	76	38
64	52	47
75	63	50

- Determine si existe un efecto global por privación del sueño, utilizando las ecuaciones conceptuales del análisis de varianza de un factor. Utilice $\alpha = 0.05$.
 - La misma pregunta de la parte **a**, pero ahora utilice las ecuaciones para el cálculo.
 - ¿Cuál prefiere usted? ¿Por qué?
 - A partir de los valores obtenidos en **b**, estime usted la magnitud del efecto. cognitiva
17. Para verificar si la memoria cambia con la edad, un investigador realiza un experimento en el cual participan cuatro grupos de seis sujetos cada uno. Los grupos difieren en cuanto a la edad de los sujetos. En el grupo 1, los sujetos tienen 30 años; en el grupo 2 tienen 40 años; en el grupo 3 tienen 50 años; y en el grupo 4 tienen 60 años. Suponga que todos los sujetos gozan de buena salud y que los grupos concuerdan en otras variables importantes, tales como la escolaridad, el CI, el género, la motivación, y así sucesivamente. Se muestra a cada sujeto una serie de sílabas sin sentido (una combinación de tres letras sin significado, como DAF o FUM) a razón de una sílaba cada 4 segundos. La serie se muestra dos veces, después de lo cual se pide a los sujetos que escriban el mayor número posible de sílabas que sean capaces de re-

cordar. El número de sílabas que logró recordar cada sujeto se presenta en la siguiente tabla:

30 años de edad	40 años de edad	50 años de edad	60 años de edad
14	12	17	13
13	15	14	10
15	16	14	7
17	11	9	8
12	12	13	6
10	18	15	9

- Utilice el análisis de varianza con $\alpha = 0.05$ para determinar si la edad tiene algún efecto sobre la memoria.
 - Si existe un efecto significativo en **a.** estime la magnitud de dicho efecto. *cognitiva*
18. Supongamos que usted es contratado por un servicio de evaluación de productos y que debe evaluar acumuladores para automóvil. En esta parte del estudio, usted quiere determinar si existe una diferencia entre la vida útil de los mejores acumuladores de tres fabricantes (A, B y C). Como base de su evaluación, usted selecciona una muestra aleatoria de cuatro acumuladores de cada fabricante y les aplica pruebas de laboratorio, las cuales le permitirán determinar la vida útil de cada uno. En la siguiente tabla presentan los resultados, en términos de meses de vida útil de los acumuladores:

Fabricante del acumulador		
A	B	C
52	46	44
57	52	53
55	51	50
59	50	51

Utilice el análisis de varianza con $\alpha = 0.05$ para determinar si existe alguna diferencia entre estas tres marcas de acumuladores. *l/O*

19. Un experimento del capítulo 14 trató acerca del estudio del efecto de la hormona X sobre el com-

portamiento sexual. Aunque sólo presentamos dos concentraciones en aquel problema, supongamos que el experimento hubiera implicado, en realidad, cuatro concentraciones diferentes de la hormona. Aquí mostramos los datos completos, donde las concentraciones están ordenadas en forma creciente; es decir, la concentración 0 es aquella que no presenta la hormona X (éste es el grupo placebo) y la concentración 3 representa la máxima cantidad de la hormona:

Concentración de la hormona X			
0	1	2	3
5	4	8	13
6	5	10	10
3	6	12	9
4	4	6	12
7	5	6	12
8	7	7	14
6	7	9	9
5	8	8	13
4	4	7	10
8	8	11	12

- Utilice el análisis de varianza con $\alpha = 0.05$, para determinar si la hormona X afecta el comportamiento sexual.
 - Si existe un efecto real, estime usted la magnitud de dicho efecto. *biológica*
20. Un psicólogo clínico está interesado en evaluar la eficacia de las tres técnicas siguientes para el tratamiento de la depresión leve: reestructuración cognitiva, entrenamiento de asertividad, y un programa de ejercicios y nutrición. A partir de la lista de espera del centro de asesoría de la universidad, se extrae una muestra de 40 estudiantes que sufren depresión leve, y se asigna a 10 de ellos, de manera aleatoria, a las tres técnicas ya mencionadas; los 10 restantes se integran a un grupo de control placebo. El tratamiento se lleva a cabo durante 10 semanas, después de lo cual se mide la depresión mediante la escala de depresión de Beck. A continuación se presentan los niveles de depresión observados después del tratamiento. Una mayor puntuación indica mayor depresión.

Tratamiento			
	Reestructuración	Entrenamiento de	Ejercicio y
Placebo	cognitiva	asertividad	nutrición
27	10	16	26
16	8	18	24
18	14	12	17
26	16	15	23
18	18	9	25
28	8	13	22
25	12	17	16
20	14	20	15
24	9	21	18
26	7	19	23

- a. ¿Cuál es la hipótesis nula global?
- b. Utilizando $\alpha = 0.05$, ¿cuál es la conclusión de usted? clínica, salud
21. Un investigador universitario con conocimientos de medicina china realizó un estudio para averiguar si la acupuntura puede ayudar a reducir la adicción a la cocaína. En este experimento, 18 adictos a la cocaína fueron asignados aleatoriamente a tres grupos, incluyendo a 6 adictos en cada grupo. Un grupo recibió durante 10 semanas un tratamiento de acupuntura, en el cual a los sujetos se les insertaron agujas en ciertos puntos del pabellón auricular donde se cree que la estimulación es eficaz. A otro grupo, el grupo placebo, se le insertaron agujas en puntos de la oreja que, según se cree, no son eficaces. El tercer grupo no recibió tratamiento alguno de acupuntura y, en su lugar, los adictos que lo integraban recibieron terapia de relajación. A todos los grupos se les impartió también asesoría durante las 10 semanas del periodo de tratamiento. La variable dependiente fue el deseo de consumir cocaína, medida por el número de accesos de avidez de esa sustancia que experimentó cada uno de los adictos en la última semana del tratamiento. Los resultados se presentan a continuación.

Acupuntura + asesoría	Placebo + asesoría	Terapia de relajación + asesoría
4	8	12
7	12	7
6	11	9
5	8	6
2	10	11
3	7	6

- Utilizando $\alpha = 0.05$, ¿cuál es la conclusión de usted? clínica, salud
22. Una profesora imparte tres secciones de un curso de introducción a la psicología y en cada sección cubre el mismo material. Ha preparado un examen final diferente para cada sección, pero sospecha que una de las versiones es más difícil que las otras dos. Entonces decide realizar un experimento para evaluar la dificultad del examen. Durante el periodo de revisión, inmediatamente antes de los exámenes finales, ella selecciona al azar a cinco voluntarios de cada clase. A los voluntarios de la clase 1 les entrega la versión 1 del examen; los voluntarios de la clase 2 reciben la versión 2, y a los voluntarios de la clase 3 les proporciona la versión 3. Por supuesto, todos los voluntarios prometen no revelar ninguna de las preguntas de ese examen y además, naturalmente, todos los voluntarios resolverán un examen final diferente del que presentaron durante el experimento. A continuación se presentan los resultados.

Versión 1 del examen	Versión 2 del examen	Versión 3 del examen
70	95	88
92	75	76
85	81	84
83	83	93
78	72	77

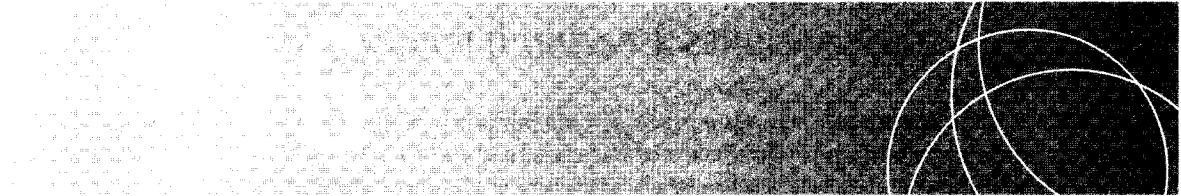
Utilizando $\alpha = 0.05$, ¿cuál es la conclusión de usted? educación

Si desea encontrar problemas de práctica adicionales totalmente resueltos sobre los temas que acaba de estudiar, busque Chapter 15 en el material adicional del libro, disponible en inglés en: <http://www.thomsonlearning.com.mx>.

Si pretende resolver más problemas de práctica, relacionados con lo que se presentó en este capítulo y resueltos por medio de ecuaciones conceptuales, busque Chapter 15 en el material adicional del libro (en inglés), en: <http://www.thomsonlearning.com.mx>.

Si quiere resolver problemas para computadora sobre los temas aquí presentados, busque Chapter 15 en el material adicional (en inglés) del libro, en: <http://www.thomsonlearning.com.mx>.

Comparaciones múltiples



Introducción

Comparaciones *a priori* o planeadas

Comparaciones *a posteriori* o *post hoc*

La prueba de diferencia honestamente significativa (DHS) de Tukey

La prueba de Newman-Keuls

Pruebas DHS y Newman-Keuls con n distinta

Contrastación entre las comparaciones planeadas, la prueba DHS de Tukey y la prueba de Newman-Keuls

Resumen

Términos importantes

Preguntas y problemas

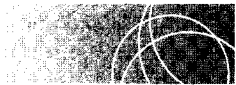
Notas



INTRODUCCIÓN

En el capítulo 15 estudiamos el análisis de varianza de un factor cuando se utiliza con un diseño de grupos independientes. Allí señalamos que el análisis de varianza es una técnica estadística que se utiliza en los experimentos con varios grupos y que nos permite determinar, si se emplea junto con una prueba global (una prueba F), si la variable independiente ha producido un efecto significativo.

En el análisis de varianza de un factor, un valor F significativo indica que no todas las condiciones producen el mismo efecto sobre la variable dependiente. Por ejemplo, en el experimento ilustrativo presentado en el capítulo 15, en el cual se investigaba la cantidad de estrés producido por tres situaciones distintas, se obtuvo un valor F significativo y concluimos que las tres situaciones no eran iguales en cuanto a los niveles de estrés que producían. Por razones de carácter pedagógico decidimos terminar el análisis con es-



COMPARACIONES *A POSTERIORI* O *POST HOC*

Cuando las comparaciones no se planean de antemano, debemos utilizar una prueba *a posteriori*. Estas comparaciones surgen, por lo general, después que el experimentador revisa los datos y elige a grupos cuyas medias están muy apartadas, o bien, se producen cuando se realizan todas las comparaciones posibles sin una base teórica *a priori*. Como estas comparaciones no fueron planeadas antes del experimento, debemos corregir los valores inflados de probabilidad que se presentan cuando se llevan a cabo comparaciones múltiples, como se comentó en la sección anterior.

Existen muchos métodos para efectuar esta corrección.* El tema es bastante complejo y la presentación de todos los métodos está fuera del objetivo de este texto. Sin embargo, mostraremos dos de los métodos que gozan de mayor aceptación: un método ideado por Tukey, que se llama prueba DHS (diferencia honestamente significativa) y la prueba de Newman-Keuls. Ambas pruebas son de comparación múltiple *post hoc*, y en las dos se mantiene la tasa de error de Tipo I en α y se realizan todas las comparaciones posibles entre las parejas de las medias muestrales.

Como usted recordará, el problema que se presenta al realizar comparaciones múltiples con las pruebas *t* es que los valores críticos de *t* se obtienen bajo el supuesto de que sólo hay dos muestras cuyas medias van a ser comparadas. Esto podría hacerse mediante una prueba *t*. Cuando hay muchas muestras, y por lo tanto más de una comparación, la distribución muestral de *t* ya no resulta adecuada. De hecho, si ésta se utilizara, la probabilidad real de cometer un error de tipo I rebasaría en mucho a alfa, en particular si se llevara a cabo una gran cantidad de comparaciones. Tanto el método de Tukey como el de Newman-Keuls permiten evitar esta dificultad, utilizando distribuciones muestrales basadas en la comparación de las medias de muchas muestras y no sólo de dos. Estas distribuciones, llamadas *Q* o de rango Student, se desarrollaron al extraer, de manera aleatoria *k* muestras con igual *n* de una misma población (en lugar de extraer sólo 2, como en la prueba *t*) y determinar la diferencia entre las medias muestrales máxima y mínima. Las diferencias se dividieron después entre $\sqrt{s_w^2/n}$, esto produjo distribuciones similares a las distribuciones *t*, excepto que esto proporcionó la base para establecer comparaciones múltiples y no sólo una, como en la prueba *t*. Los puntos percentiles números 95 y 99 de la distribución *Q* se presentan en la tabla G del apéndice D. Estos percentiles son los valores críticos de *Q* para los niveles alfa de 0.05 y 0.01. Como podrá imaginar, los valores críticos dependen del número de medias muestrales y de los grados de libertad asociados a s_w^2 .

Al analizar las pruebas DHS y de Newman-Keuls, es útil distinguir entre dos aspectos de los errores de tipo I: las tasas de error *por experimento* y *por comparación*.

definiciones ■ *La tasa de error por experimento es la probabilidad de cometer uno o más errores de Tipo I para todo el conjunto de las comparaciones posibles en un experimento. La tasa de error por comparación es la probabilidad de cometer un error de tipo I para cualquiera de las comparaciones posibles.*

*Para un análisis detallado de estos métodos, véase R. F. Kirk, *Experimental Design*, tercera edición, Brooks/Cole, Pacific Grove, CA, 1995, pp. 144–159.

Como veremos en las siguientes secciones, las pruebas DHS y de Newman-Keuls difieren en cuanto a la tasa que se mantiene igual a alfa en cada una de ellas.

La prueba de diferencia honestamente significativa (DHS) de Tukey

Esta prueba está diseñada para comparar todas las parejas posibles de medias, al tiempo que el error de Tipo I para el conjunto completo de comparaciones se mantiene en α . Por consiguiente, la prueba DHS mantiene la tasa de error de tipo I por experimento en α . El estadístico que se calcula en el caso de esta prueba es Q y esto se define mediante la siguiente ecuación:

$$Q_{\text{obt}} = \frac{\bar{X}_i - \bar{X}_j}{\sqrt{s_w^2/n}}$$

donde \bar{X}_i = la mayor de las dos medias que se comparan

\bar{X}_j = la menor de las dos medias que se comparan

s_w^2 = estimación de la varianza dentro de los grupos

n = número de sujetos en cada grupo

Observe que, al calcular Q_{obt} , siempre se resta la media menor de la mayor, por lo cual Q_{obt} siempre es positiva. Por lo demás, el estadístico Q es muy similar al estadístico t , excepto por el hecho de que utiliza las distribuciones Q en lugar de las distribuciones t . Para emplear el estadístico, calculamos Q_{obt} para las comparaciones deseadas y comparamos Q_{obt} con Q_{crit} , determinadas a partir de la tabla G. La regla de decisión establece que si $Q_{\text{obt}} \geq Q_{\text{crit}}$, entonces rechazamos H_0 . En caso contrario, conservamos H_0 .

Para ilustrar el uso de la prueba DHS, emplearemos los datos del experimento sobre el estrés. Para este ejemplo, supondremos que se desea hacer las tres comparaciones. Para utilizar la prueba DHS necesitamos dos pasos. En primer lugar, tenemos que calcular el valor de Q_{obt} para cada comparación y luego comparar cada valor con Q_{crit} . Los cálculos para Q_{obt} son los siguientes:

Grupo 2 y grupo 1:

$$Q_{\text{obt}} = \frac{\bar{X}_2 - \bar{X}_1}{\sqrt{s_w^2/n}} = \frac{8.00 - 4.00}{\sqrt{4.50/5}} = \frac{4.00}{0.949} = 4.21$$

Grupo 3 y grupo 1:

$$Q_{\text{obt}} = \frac{\bar{X}_3 - \bar{X}_1}{\sqrt{s_w^2/n}} = \frac{13.00 - 4.00}{\sqrt{4.50/5}} = \frac{9.00}{0.949} = 9.48$$

Grupo 3 y grupo 2:

$$Q_{\text{obt}} = \frac{\bar{X}_3 - \bar{X}_2}{\sqrt{s_w^2/n}} = \frac{13.00 - 8.00}{\sqrt{4.50/5}} = \frac{5.00}{0.949} = 5.27$$

El siguiente paso consiste en comparar los valores de Q_{obt} con los de Q_{crit} . El valor de Q_{crit} se determina mediante la tabla G. Para localizar el valor adecuado, debemos conocer el número de grados de libertad, el nivel alfa y k . Los grados de libertad son los asociados con s_w^2 . En este experimento, $gl = 12$. Como ya hemos mencionado, k representa el número de grupos en el experimento. En nuestro experimento, $k = 3$. Para este caso, establecemos alfa en 0.05. Según la tabla G, con $gl = 12$, $k = 3$ y $\alpha = 0.05$, obtenemos:

$$Q_{\text{crit}} = 3.77$$

Como quiera que $Q_{\text{obt}} > 3.77$ para cada comparación, rechazamos H_0 en cada caso y concluimos que $\mu_1 \neq \mu_2 \neq \mu_3$. Las tres condiciones difieren en cuanto a su valor como productores de estrés. La solución se resume en la tabla 16.2.

La prueba de Newman-Keuls

La prueba de Newman-Keuls es de tipo *post hoc* y nos permite realizar todas las comparaciones posibles entre las parejas de medias muestrales. Se parece a la prueba DHS por el hecho de que Q_{obt} se calcula para cada comparación y se compara con Q_{crit} a fin de evaluar la hipótesis nula. Difiere de la prueba DHS en que la tasa de error de tipo I se mantiene en α para cada comparación, en lugar del conjunto completo de comparaciones. Por lo consiguiente:

La prueba de Newman-Keuls mantiene la tasa de error por comparación en α , mientras que la prueba DHS mantiene la tasa de error por experimento en α .

Para mantener la tasa de error por comparación en α , el método de Newman-Keuls modifica el valor de Q_{crit} para cada comparación. El valor de Q_{crit} empleado para cada comparación está dado por la distribución muestral de Q para el número de grupos que tienen medias entre \bar{X}_i y \bar{X}_j después de ordenar todas las medias. Este número se designa con r para distinguirlo de k , que indica el número total de grupos que participan en el experimento.

A fin de ejemplificar el método de Newman-Keuls, lo utilizaremos ahora para analizar los datos del experimento sobre el estrés. El primer paso consiste en ordenar las medias. Esto ya se ha hecho en la tabla 16.3. A continuación, calculamos Q_{obt} para cada comparación. Los cálculos y los valores de Q_{obt} se presentan también en la tabla 16.3. El siguiente paso consiste en determinar los valores de Q_{crit} . Observe que a fin de mantener la tasa de error de tipo I en α para cada comparación, existe un valor crítico distinto para cada comparación, el cual depende de r . Recuerde que para cualquier comparación, r es igual al número de grupos con medias entre \bar{X}_i y \bar{X}_j , después de haber ordenado todas las medias. Así, para la comparación entre los grupos 3 y 1, $\bar{X}_i = \bar{X}_3$ y $\bar{X}_j = \bar{X}_1$. Al ordenar las medias, existen tres grupos (los grupos 1, 2 y 3) cuyas medias

tabla 16.2 Análisis de las comparaciones individuales *post hoc* del experimento sobre el estrés, mediante la prueba DHS de Tukey

	Grupo			Cálculo
	1	2	3	
\bar{X}	4.00	8.00	13.00	
$\bar{X}_i - \bar{X}_j$		4.00	9.00	Grupos 2 y 1: $Q_{\text{obt}} = \frac{\bar{X}_2 - \bar{X}_1}{\sqrt{s_w^2/n}} = \frac{8.00 - 4.00}{\sqrt{4.50/5}} = 4.21$
Q_{obt}		4.21*	9.48*	Grupos 3 y 1: $Q_{\text{obt}} = \frac{\bar{X}_3 - \bar{X}_1}{\sqrt{s_w^2/n}} = \frac{13.00 - 4.00}{\sqrt{4.50/5}} = 9.48$
Q_{crit}		3.77	3.77	Grupos 3 y 2: $Q_{\text{obt}} = \frac{\bar{X}_3 - \bar{X}_2}{\sqrt{s_w^2/n}} = \frac{13.00 - 8.00}{\sqrt{4.50/5}} = 5.27$
gl = 12			3.77	
k = 3				
$\alpha = 0.05$				

*Rechazar H_{0i} .

tabla 16.3 Análisis de comparaciones individuales *post hoc* del experimento sobre el estrés, mediante la prueba de Newman-Keuls

	Grupo			Cálculo
	1	2	3	
\bar{X}	4.00	8.00	13.00	Grupos 3 y 1:
$\bar{X}_i - \bar{X}_j$		4.00	9.00	$Q_{\text{obt}} = \frac{\bar{X}_3 - \bar{X}_1}{\sqrt{s_w^2/n}} = \frac{13.00 - 4.00}{\sqrt{4.50/5}} = 9.48$
Q_{obt}		4.21*	9.48*	Grupos 3 y 2:
Q_{crit}		3.08	3.77	$Q_{\text{obt}} = \frac{\bar{X}_3 - \bar{X}_2}{\sqrt{s_w^2/n}} = \frac{13.00 - 8.00}{\sqrt{4.50/5}} = 5.27$
gl = 12			3.08	Grupos 2 y 1:
r = 2, 3				$Q_{\text{obt}} = \frac{\bar{X}_2 - \bar{X}_1}{\sqrt{s_w^2/n}} = \frac{8.00 - 4.00}{\sqrt{4.50/5}} = 4.21$
$\alpha = 0.05$				

*Rechazar H_0 .

están entre \bar{X}_3 y \bar{X}_1 . Por lo tanto, $r = 3$ para esta comparación. Para la comparación entre los grupos 3 y 2, $\bar{X}_i = \bar{X}_3$ y $\bar{X}_j = \bar{X}_2$. Después de ordenar las medias, \bar{X}_3 y \bar{X}_2 son adyacentes, de modo que sólo existen dos medias entre \bar{X}_3 y \bar{X}_2 , a saber, \bar{X}_2 y \bar{X}_3 . Así, $r = 2$ para esta comparación. Lo mismo ocurre con la comparación entre los grupos 2 y 1. Para esta comparación, $\bar{X}_i = \bar{X}_2$, $\bar{X}_j = \bar{X}_1$ y $r = 2$ (las dos medias entre ellas son \bar{X}_2 y \bar{X}_1).

Ahora podemos determinar Q_{crit} para cada comparación. El valor de Q_{crit} se presenta en la tabla G, empleando los valores apropiados para gl, r y α . Los grados de libertad son los gl correspondientes a s_w^2 , que son 12, y utilizaremos el mismo nivel α . Así pues, $\alpha = 0.05$. Para la comparación entre los grupos 1 y 3, con gl = 12, $r = 3$, y $\alpha = 0.05$, $Q_{\text{crit}} = 3.77$. Para la comparación entre los grupos 1 y 2, y los grupos 2 y 3 con gl = 12, $r = 2$ y $\alpha = 0.05$, $Q_{\text{crit}} = 3.08$. Estos valores también se presentan en la tabla 16.3.

El último paso consiste en comparar Q_{obt} con Q_{crit} . Al hacer estas comparaciones seguimos la regla que consiste en comenzar con el valor máximo de Q_{obt} que contiene la tabla (el que está en la esquina superior derecha de los valores de Q_{obt}) y compararlo con el Q_{crit} apropiado. Si el valor encontrado es significativo, nos desplazamos un paso hacia la izquierda y comparamos el siguiente Q_{obt} con el correspondiente Q_{crit} . Continuamos de la misma manera hasta el final del renglón o hasta llegar a un valor de Q_{obt} no significativo. En el último caso, los valores restantes de Q_{obt} que se encuentran a la izquierda de este primer Q_{obt} no significativo se consideran también como valores no significativos. Al finalizar el renglón o al llegar a un Q_{obt} no significativo, pasamos al siguiente renglón y empezamos de nuevo con el valor Q_{obt} del extremo derecho. Continuamos en esta forma hasta que todos los valores de Q_{obt} han sido evaluados.

En nuestro ejemplo, comenzamos con 9.48 y lo comparamos con 3.77. Éste es significativo. Luego comparamos 4.21 con 3.08, el cual también es significativo. Como terminamos el renglón con esta comparación, pasamos al siguiente y comparamos 5.27 con 3.08. De nuevo, el valor de Q_{obt} es significativo. Así pues, podemos rechazar H_0 en cada comparación. Por lo tanto, concluimos el análisis al establecer que $\mu_1 \neq \mu_2 \neq \mu_3$. La solución se resume en la tabla 16.3.

Ahora resolveremos un problema de práctica.



- Con los datos del problema de práctica 15.1 (página 367), lleve a cabo lo siguiente:
- Ponga a prueba las comparaciones planeadas según las cuales 1) clase + lectura y clase tienen distintos efectos y 2) clase + lectura y película + lectura poseen diferentes efectos. Utilice $\alpha = 0.05_{2 \text{ colas}}$.
 - Realice todas las comparaciones *post hoc* posibles mediante la prueba DHS. Utilice $\alpha = 0.05$.
 - Realice todas las comparaciones *post hoc* posibles mediante la prueba de Newman-Keuls. Utilice $\alpha = 0.05$.

SOLUCIÓN

Para comodidad del lector, presentamos de nuevo los datos y la solución mediante el análisis de varianza.

Condición 1 Clase		Condición 2 Clase + lectura		Condición 3 Película + lectura	
X_1	X_1^2	X_2	X_2^3	X_3	X_3^2
92	8464	86	7396	81	6561
86	7396	93	8649	80	6400
87	7569	97	9409	72	5184
76	5776	81	6561	82	6724
80	6400	94	8836	83	6889
87	7569	89	7921	89	7921
92	8464	98	9604	76	5776
83	6889	90	8100	88	7744
84	7056	91	8281	83	6889
767	65 583	819	74 757	734	60 088

$$n_1 = 9$$

$$\bar{X}_1 = 85.222$$

$$n_2 = 9$$

$$\bar{X}_2 = 91$$

$$n_3 = 9$$

$$\bar{X}_3 = 81.556$$

$$\sum_{\text{todos los datos}} X = 2320$$

$$N = 27$$

$$\sum_{\text{todos los datos}} X^2 = 200\,428$$

$$\bar{X}_G = \frac{\sum_{\text{todos los datos}} X}{N}$$

$$= 85.926$$

Fuente	SC	gl	s^2	F_{obt}
Entre grupos	408.074	2	204.037	7.29*
Dentro de los grupos	671.778	24	27.991	
Total	1079.852	26		

*Con $\alpha = 0.05$, $F_{\text{crit}} = 3.40$. Por lo tanto, H_0 se rechaza.

- a. Comparaciones planeadas: Las comparaciones en cuestión son las siguientes:
Clase + lectura y clase (condición 2 y condición 1):

$$t_{\text{obt}} = \frac{\bar{X}_2 - \bar{X}_1}{\sqrt{2s_w^2/n}} = \frac{91 - 85.222}{\sqrt{2(27.991)/9}} = 2.32$$

Clase + lectura y película + lectura (condición 2 y condición 3):

$$t_{\text{obt}} = \frac{\bar{X}_2 - \bar{X}_3}{\sqrt{2s_w^2/n}} = \frac{91 - 81.556}{\sqrt{2(27.991)/9}} = 3.79$$

Para evaluar estos valores de t_{obt} , debemos determinar t_{crit} . Según la tabla D, con $\alpha = 0.05_{2 \text{ colas}}$ y $gl = 24$,

$$t_{\text{crit}} = \pm 2.064$$

En vista de que $|t_{\text{obt}}| > 2.064$ en ambas comparaciones, rechazamos H_0 en cada caso y concluimos que $\mu_1 \neq \mu_2$ y $\mu_2 \neq \mu_3$. Al utilizar las pruebas *a priori*, tal parece que el método más eficaz es el de clase + lectura.

- b. Comparaciones *post hoc* mediante la prueba DHS: Con la prueba DHS determinamos Q_{obt} para cada comparación y luego la evaluamos en comparación con Q_{crit} . El valor de Q_{crit} es el mismo para cada comparación, y es apropiado para que la tasa de error por experimento se mantenga en α . Aunque en esta prueba no es necesario hacerlo, primero ordenamos las medias para poder hacer comparaciones con la prueba de Newman-Keuls. Los resultados se presentan en la siguiente tabla. Los cálculos para obtener Q_{obt} son los siguientes:

Clase (1) y Película + lectura (3):

$$Q_{\text{obt}} = \frac{\bar{X}_1 - \bar{X}_3}{\sqrt{s_w^2/n}} = \frac{85.222 - 81.556}{\sqrt{27.991/9}} = \frac{3.666}{1.764} = 2.08$$

Clase + lectura (2) y Película + lectura (3):

$$Q_{\text{obt}} = \frac{\bar{X}_2 - \bar{X}_3}{\sqrt{s_w^2/n}} = \frac{91 - 81.556}{\sqrt{27.991/9}} = \frac{9.444}{1.764} = 5.35$$

Clase + lectura (2) y Clase (1):

$$Q_{\text{obt}} = \frac{\bar{X}_2 - \bar{X}_1}{\sqrt{s_w^2/n}} = \frac{91 - 85.222}{\sqrt{27.991/9}} = \frac{5.778}{1.764} = 3.28$$

Ahora debemos determinar Q_{crit} . Según la tabla G, con $gl = 24$, $k = 3$ y $\alpha = 0.05$, obtenemos:

$$Q_{\text{crit}} = 3.53$$

Al comparar los tres valores de Q_{obt} con Q_{crit} , observamos que sólo la comparación entre película + lectura y clase + lectura es significativa. (Para esta comparación, $Q_{\text{obt}} > 3.53$, mientras que para las demás, $Q_{\text{obt}} < 3.53$.) Por lo tanto, con base en la prueba DHS, podemos rechazar H_0 para la comparación clase + lectura y película + lectura (condiciones 2 y 3). Clase + lectura parece más eficaz

(continúa)

que película + lectura. Sin embargo, no podemos rechazar H_0 con respecto a las demás comparaciones. Los resultados se resumen en la siguiente tabla.

	Condición			Cálculo
	3	1	2	
\bar{X}	81.556	85.222	91	
$\bar{X}_i - \bar{X}_j$		3.666	9.444	Condiciones 1 y 3: $Q_{\text{obt}} = \frac{\bar{X}_1 - \bar{X}_3}{\sqrt{s_w^2/n}} = \frac{85.222 - 81.556}{\sqrt{27.991/9}} = 2.08$
Q_{obt}		2.08	5.778	
$Q_{\text{crit}} = 3.53$		2.08	5.35*	Condiciones 2 y 3: $Q_{\text{obt}} = \frac{\bar{X}_2 - \bar{X}_3}{\sqrt{s_w^2/n}} = \frac{91 - 81.556}{\sqrt{27.991/9}} = 5.35$
gl = 24		3.53	3.28	
k = 3		3.53	3.53	Condiciones 2 y 1: $Q_{\text{obt}} = \frac{\bar{X}_2 - \bar{X}_1}{\sqrt{s_w^2/n}} = \frac{91 - 85.222}{\sqrt{27.991/9}} = 3.28$
$\alpha = 0.05$				

*Rechazar H_0 .

c. Comparaciones *post hoc* mediante la prueba de Newman-Keuls: Como en el caso de la prueba DHS, calculamos, Q_{obt} para cada comparación y luego la evaluamos con respecto a Q_{crit} . Sin embargo, con la prueba de Newman-Keuls, el valor de Q_{crit} cambia con cada comparación, a fin de mantener la tasa de error por comparación en α . En primer lugar se ordenan las medias, de menor a mayor. Esto se muestra en la tabla siguiente. Luego, calculamos Q_{obt} para cada comparación. Estos cálculos, y los valores de Q_{obt} se presentan en la tabla. A continuación, determinamos Q_{crit} para cada comparación, a partir de la tabla G. Los valores de Q_{crit} dependen de α , de gl para s_w^2 , y de r , donde r es el número de grupos cuyas medias se encuentran entre \bar{X}_i y \bar{X}_j después de ordenar las medias. Así pues, para la comparación entre \bar{X}_2 y \bar{X}_1 , $r = 2$; entre \bar{X}_2 y \bar{X}_3 , $r = 3$; y entre \bar{X}_1 y \bar{X}_3 , $r = 2$. Para este experimento, gl = 24 y $\alpha = 0.05$. Los valores de Q_{crit} para cada comparación se presentan en la tabla. Al comparar Q_{obt} con Q_{crit} , comenzamos con el valor máximo de Q_{obt} en el primer renglón y, continuando hacia la izquierda, observamos que podemos rechazar H_0 en la comparación entre las condiciones 2 y 3, pero no en la comparación entre las condiciones 1 y 3. Al pasar al siguiente renglón, como $3.28 > 2.92$, también podemos rechazar H_0 en la comparación entre las condiciones 2 y 1. Por consiguiente, con base en la prueba de Newman-Keuls, parece que clase + lectura es superior tanto al método de clase solamente como al método de película + lectura.

(continúa)

	Condición			Cálculo
	3	1	2	
\bar{X}	81.556	85.222	91	
$\bar{X}_i - \bar{X}_j$		3.666	9.444	
Q_{obt}		2.08	5.35*	Condiciones 1 y 3: $Q_{obt} = \frac{\bar{X}_1 - \bar{X}_3}{\sqrt{s_w^2/n}} = \frac{85.222 - 81.556}{\sqrt{27.991/9}} = 2.08$
Q_{crit}		2.92	3.28*	Condiciones 2 y 3: $Q_{obt} = \frac{\bar{X}_2 - \bar{X}_3}{\sqrt{s_w^2/n}} = \frac{91 - 81.556}{\sqrt{27.991/9}} = 5.35$
gl = 24 r = 2, 3 $\alpha = 0.05$		2.92	3.53	Condiciones 2 y 1: $Q_{obt} = \frac{\bar{X}_2 - \bar{X}_1}{\sqrt{s_w^2/n}} = \frac{91 - 85.222}{\sqrt{27.991/9}} = 3.28$

*Rechazar H_0 .

Pruebas DHS y Newman-Keuls con n distinta

Como ya hemos señalado, las pruebas DHS y de Newman-Keuls resultan apropiadas cuando existe el mismo número de sujetos en cada grupo. Cuando las n son distintas, podemos seguir utilizando estas pruebas siempre que las n no difieran mucho entre sí. Para utilizar estas pruebas con n distintas, calculamos la media armónica (\tilde{n}) de las diversas n y la escribimos en el denominador de la ecuación de Q . La ecuación de \tilde{n} es

$$\tilde{n} = \frac{k}{(1/n_1) + (1/n_2) + (1/n_3) + \dots + (1/n_k)} \quad \text{media armónica}$$

donde k = número de grupos

n_k = es el número de sujetos incluidos en el k -ésimo grupo

Supongamos que en el experimento sobre el estrés, $n_1 = 5$, $n_2 = 7$ y $n_3 = 8$. Entonces,

$$\tilde{n} = \frac{k}{(1/n_1) + (1/n_2) + (1/n_3)} = \frac{3}{\frac{1}{5} + \frac{1}{7} + \frac{1}{8}} = 6.41$$



CONTRASTACIÓN ENTRE LAS COMPARACIONES PLANEADAS, LA PRUEBA DHS DE TUKEY Y LA PRUEBA DE NEWMAN-KEULS

En virtud de que las comparaciones planeadas no incluyen una corrección para compensar el incremento de la probabilidad de cometer un error de tipo I, son más potentes que cualquiera de las pruebas *post hoc* que hemos analizado hasta ahora. Éste es el método que debemos elegir siempre que sea posible. Sin embargo, es importante observar que las comparaciones planeadas deben ser relativamente escasas y haber surgido en forma significativa y lógica del diseño experimental.

La decisión de si se debe utilizar la prueba DHS de Tukey o la de Newman-Keuls depende en realidad de la filosofía de cada persona. Como quiera que la prueba DHS mantiene la tasa de error de tipo I en α para todo el conjunto de comparaciones, mientras que la de Newman-Keuls mantiene dicha tasa en α para cada comparación, la prueba de Newman-Keuls tiene una tasa de error de tipo I por experimento algo mayor que la prueba DHS (aunque aún mucho menor que cuando no se lleva a cabo ajuste alguno). Debido al uso de un nivel α por experimento menos estricto que la prueba DHS, la prueba de Newman-Keuls es más potente. Un ejemplo de lo anterior es el último experimento que hemos analizado. Con Newman-Keuls, conseguimos rechazar H_0 para las comparaciones entre las condiciones 2 y 3, y entre las condiciones 2 y 1. Con la prueba DHS rechazamos H_0 sólo en la comparación entre las condiciones 2 y 3. Por lo tanto, debemos decidir cuál es preferible. Newman-Keuls tiene una tasa de error de tipo I mayor por experimento, pero una tasa de error de tipo II menor. Entre las dos pruebas, un experimentador conservador elegiría probablemente la prueba DHS de Tukey, mientras que un investigador más liberal podría escoger la prueba de Newman-Keuls. Por supuesto, si las consecuencias de cometer un error de tipo I son mayores que las de cometer un error de tipo II o viceversa, el investigador elegirá la prueba que minimice la tasa de error apropiada.

RESUMEN

En los experimentos que utilizan la técnica de análisis de varianza, un valor F significativo indica que los efectos no son iguales en todas las condiciones. Para determinar qué condiciones difieren entre sí, por lo general se realizan comparaciones múltiples entre pares de medias de grupo. Existen dos enfoques para realizar las comparaciones múltiples: comparaciones *a priori* o planeadas, y comparaciones *a posteriori* o *post hoc*.

En el enfoque *a priori* existen comparaciones entre los grupos que han sido planeadas antes de reunir los datos. Éstas se pueden realizar de la manera usual, sin importar que el valor F obtenido sea significativo o no, al calcular t_{obt} para los dos grupos y evaluar t_{obt} para compararlo con t_{crit} . Al realizar el análisis, utilizamos la estimación de la varianza dentro de grupos calculada al hacer el análisis de varianza. Como esta estimación se basa en más grupos que la estimación de dos grupos utilizada en la prueba t ordinaria, es más precisa. No se necesita corrección alguna para realizar comparaciones múltiples. Sin embargo, los estadísticos no se ponen de acuerdo en cuanto a si las comparaciones deben ser ortogonales. Hemos adoptado la posición de que las comparaciones *a priori* no forzosamente tienen que ser ortogonales, siempre que surjan en for-

ma significativa y lógica del diseño experimental y no sean numerosas.

Las comparaciones *a posteriori* o *post hoc*, no se planean antes de realizar el experimento, sino que surgen después de observar los datos. Por consiguiente, debemos tener mucho cuidado con las consideraciones referentes a los errores de tipo I. Las comparaciones *post hoc* se deben llevar a cabo con un método que permita corregir el incremento de la probabilidad de cometer un error de tipo I. Existen muchos métodos para este propósito.

En este capítulo hemos descrito las pruebas DHS de Tukey y de Newman-Keuls. Ambas técnicas mantienen la tasa de error de tipo I en α , mientras realizan todas las comparaciones posibles entre las parejas de medias muestrales. La prueba DHS mantiene la tasa de error por experimento en α , mientras que la prueba de Newman-Keuls mantiene la tasa de error por comparación en α . Ambas pruebas utilizan el estadístico Q o de rango Student. Igual que en el caso de la prueba t , Q_{obt} se calcula para cada comparación y se evalúa con respecto a Q_{crit} la cual se determina a partir de la distribución muestral de Q . Si $Q_{\text{obt}} \geq Q_{\text{crit}}$, entonces rechazamos la hipótesis nula.

TÉRMINOS IMPORTANTES

Comparaciones <i>a posteriori</i> (página 384)	Comparaciones <i>post hoc</i> (página 384)	Q_{obt} (página 385)
Comparaciones <i>a priori</i> (página 382)	Prueba DHS de Tukey (página 385)	Tasa de error por comparación (página 384)
Comparaciones planeadas (página 382)	Prueba de Newman-Keuls (página 386)	Tasa de error por experimento (página 384)
	Q_{crit} (página 385)	

PREGUNTAS Y PROBLEMAS

- Identifique o defina cada uno de los conceptos de la sección “términos importantes”.
- Explique por qué debemos hacer correcciones al llevar a cabo comparaciones múltiples, cuando realizamos comparaciones *post hoc*.
- ¿Cuál es la diferencia entre las comparaciones planeadas, las comparaciones *post hoc* mediante la prueba DHS, y las comparaciones *post hoc* al utilizar la prueba de Newman-Keuls con respecto a:
 - La potencia? Explique.
 - La probabilidad de cometer un error de tipo I? Explique.
- ¿Qué son las distribuciones Q o de rango Student? ¿Cómo evitan el problema del incremento de la probabilidad de incurrir en errores de tipo I, que resulta al realizar varias comparaciones con la distribución t ?
- Al realizar comparaciones planeadas, es mejor utilizar s_w^2 del análisis de varianza, en lugar de la estimación ponderada de la varianza de los dos grupos bajo comparación. ¿Es correcta esta afirmación? ¿Por qué?
- Consulte el problema 15 del capítulo 15, página 378. En dicho problema, el análisis de varianza produjo una F significativa, lo cual indica que por lo menos uno de los cereales tiene un contenido de azúcar diferente.
 - Realice un análisis *post hoc* de cada pareja de medias con ayuda de la prueba DHS de Tukey, con $\alpha = 0.05$ para determinar qué cereales tienen un contenido de azúcar distinto.
 - Lo mismo que en la parte **a**, pero en esta ocasión utilice la prueba de Newman-Keuls. salud
- Consulte el problema 16 del capítulo 15, página 378. En dicho problema, F_{obt} fue significativa, lo cual indica que por lo menos una de las condiciones produjo un efecto diferente sobre la capacidad de mantener la atención.
 - Realice una comparación planeada entre las medias del grupo que no durmió durante 48 horas y el grupo normal, para ver si estas condiciones difieren en su efecto sobre la capacidad de mantener la atención. Utilice $\alpha = 0.05_{2 \text{ colas}}$. ¿Cuál es la conclusión de usted?
 - Realice comparaciones *post hoc* para contrastar cada pareja de medias utilizando la prueba de Newman-Keuls y $\alpha = 0.05_{2 \text{ colas}}$. ¿Cuál es la conclusión de usted?
 - Lo mismo que en la parte **b**, pero utilice la prueba DHS. Compare sus respuestas para las partes **b** y **c**. Explique cualquier diferencia entre ellas. cognitiva
- Consulte el problema 17 del capítulo 15, página 378. El análisis de varianza para este ejercicio produjo un valor F significativo, lo cual indica que por lo menos una de las condiciones de edad producía un efecto diferente sobre la memoria.
 - Utilice las comparaciones planeadas con $\alpha = 0.05_{2 \text{ colas}}$ para comparar las medias de los grupos de 60 y de 30 años de edad. ¿Cuál es la conclusión de usted?
 - Aplique la prueba de Newman-Keuls con $\alpha = 0.05_{2 \text{ colas}}$ para comparar todas las parejas de medias posibles. ¿Cuál es la conclusión de usted? cognitiva
- Consulte el problema 18 del capítulo 15, página 379. Suponga que se le pide que haga una recomendación sobre los acumuladores, basándose en la vida útil observada en éstos. Utilice la prueba DHS con $\alpha = 0.05_{2 \text{ colas}}$ como ayuda para tomar una decisión. I/O
- Utilice como referencia el problema 19 del capítulo 15, página 379.
 - Por medio de comparaciones planeadas, con $\alpha = 0.05_{2 \text{ colas}}$, compare la media de la concentración 3 con la de la concentración 0. ¿Cuál es la conclusión de usted?
 - Utilizando la prueba de Newman-Keuls con $\alpha = 0.05_{2 \text{ colas}}$, compare todas las parejas de medias posibles. ¿Cuál es la conclusión de usted?
 - Igual que en la parte **b**, pero en esta ocasión utilice la prueba DHS. biológica
- Utilice como referencia el problema 20 del capítulo 15, página 379. Compare *post hoc* utilizando la prueba DHS de Tukey, con $\alpha = 0.05_{2 \text{ colas}}$. ¿Cuál es la conclusión de usted? clínica, salud

NOTAS

16.1 *Comparaciones ortogonales.* Cuando hacemos comparaciones entre las medias de los grupos, podemos representar cualquier comparación como una suma ponderada de dichas medias. Por ejemplo, si tenemos cuatro medias, podemos representar la comparación entre cualesquiera de los grupos como la suma ponderada de \bar{X}_1 , \bar{X}_2 , \bar{X}_3 y \bar{X}_4 . De esta manera, al evaluar $\bar{X}_1 - \bar{X}_2$, la suma ponderada sería $(1)\bar{X}_1 + (-1)\bar{X}_2 + (0)\bar{X}_3 + (0)\bar{X}_4$, donde 1, -1, 0 y 0 son los pesos. Si evalua-

mos $\bar{X}_3 - \bar{X}_4$, la suma ponderada será $(0)\bar{X}_1 + (0)\bar{X}_2 + (1)\bar{X}_3 + (-1)\bar{X}_4$, donde 0, 0, 1 y -1 son los pesos. *En general, dos comparaciones son ortogonales si la suma de los productos de los dos pesos para cada media es igual a cero.* Con esta información, determinaremos si las dos comparaciones anteriores son ortogonales. Multiplicamos y sumamos las parejas de pesos apropiados de la siguiente manera:

Comparación	Suma ponderada	Pesos
$\bar{X}_1 - \bar{X}_2$	$(1)\bar{X}_1 + (-1)\bar{X}_2 + (0)\bar{X}_3 + (0)\bar{X}_4$	1, -1, 0, 0
$\bar{X}_3 - \bar{X}_4$	$(0)\bar{X}_1 + (0)\bar{X}_2 + (1)\bar{X}_3 + (-1)\bar{X}_4$	0, 0, 1, -1
	$(1)(0) + (-1)(0) + (0)(1) + (0)(-1) = 0$	

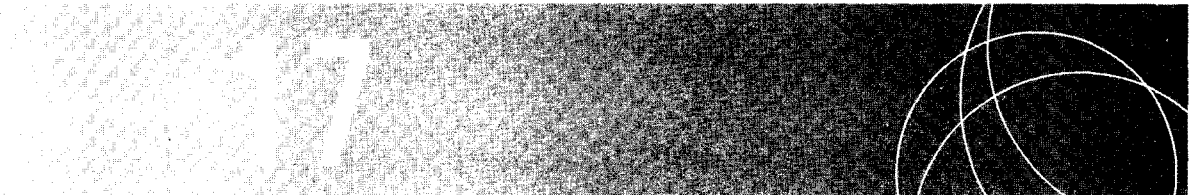
En virtud de que la suma de los productos de los dos pesos para cada media es igual a cero, estas dos comparaciones son ortogonales (es decir, in-

dependientes). En general, si existen k grupos en un experimento, entonces hay $k - 1$ comparaciones independientes posibles.*

Si desea encontrar problemas de práctica adicionales totalmente resueltos sobre los temas que acaba de estudiar, busque Chapter 16 en el material adicional del libro, disponible en inglés en: <http://www.thomsonlearning.com.mx>.

*Para un análisis más detallado de las comparaciones ortogonales, véase R. E. Kirk, *Experimental Design*, tercera edición, Brooks/Cole, Pacific Grove, CA, 1995, pp. 115-118.

Introducción al análisis de varianza de dos factores o vías



Introducción al análisis de varianza de dos factores–Presentación cualitativa

Presentación cuantitativa del análisis de varianza de dos factores

Estimación de la varianza dentro de las celdas (s_W^2)

Estimación de la varianza por renglón (s_R^2)

Estimación de la varianza por columna (s_C^2)

Estimación de la varianza renglón \times columna (s_{RC}^2)

Cálculo de las razones F

Análisis de un experimento con ANOVA de dos factores

Un experimento: Efecto del ejercicio sobre el sueño

Interpretación de los resultados

Comparaciones múltiples

Supuestos subyacentes al análisis de varianza de dos factores

Resumen

Términos importantes

Preguntas y problemas



INTRODUCCIÓN AL ANÁLISIS DE VARIANZA DE DOS FACTORES–PRESENTACIÓN CUALITATIVA

En el capítulo 15 expusimos el análisis más elemental del diseño de varianza. Le dimos los nombres de diseño simple de grupos aleatorios, análisis de varianza de un factor con diseño de grupos independientes o experimento de un solo factor con un diseño de grupos independientes. Las características de este diseño son las siguientes: que existe sólo

una variable independiente (un factor) por investigar, que en él aparecen representados varios niveles de la variable independiente (varias condiciones) y que los sujetos son asignados en forma aleatoria a cada condición.

En realidad, el diseño del análisis de varianza no se limita a los experimentos con un solo factor. De hecho, el efecto de muchos factores diferentes puede ser investigado al mismo tiempo en un experimento. En esos casos se dice que los experimentos son factoriales.

definición ■ *Un experimento factorial es aquel en el cual los efectos de dos o más factores son evaluados en un mismo experimento. En un experimento factorial, los tratamientos utilizados son combinaciones de los distintos niveles de los factores.*

El análisis de varianza de dos factores es un poco más complicado que el diseño de un solo factor. Sin embargo, con el diseño de dos factores obtenemos mucha más información.

Básicamente, el análisis de varianza de dos factores permite evaluar en un mismo experimento el efecto de dos variables independientes y la interacción entre ellas.

Para ilustrar este diseño, supongamos que un profesor de educación física realiza un experimento con la idea de comparar los efectos que producen sobre el sueño nocturno diferentes intensidades de ejercicio y la hora del día en que dicho ejercicio se ejecuta. Para este ejemplo, supongamos que existen dos niveles de ejercicio (ligero e intenso) y dos horas del día (por la mañana y por la tarde). El experimento se presenta descrito en forma de diagrama en la figura 17.1. En esta figura podemos apreciar que existen dos factores (o variables independientes): el factor A, que es la *hora del día*, y el factor B que es la *intensidad del ejercicio*. Cada factor tiene dos niveles. Por eso, esta disposición se conoce como un diseño de 2×2 (léase “dos por dos”), en el cual cada número representa un factor y la magnitud del número corresponde al número de niveles que existen dentro del factor. Por ejemplo, si en el factor A existen tres niveles, se dirá que el expe-

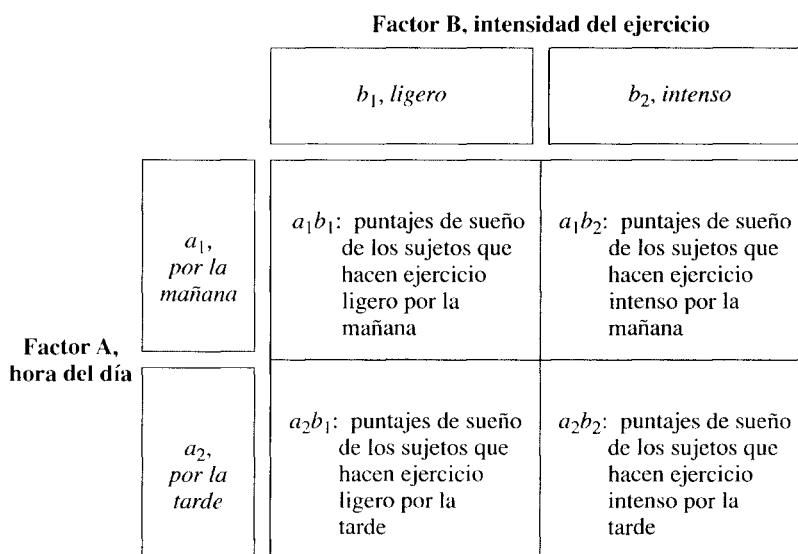


figura 17.1 Diagrama esquemático del ejemplo del análisis de varianza de dos factores, sobre el tema de la intensidad del ejercicio y la hora del día.

rimento tiene un diseño de 3×2 . En un diseño de $2 \times 4 \times 3$, habrá tres factores, los cuales tendrán dos, cuatro y tres niveles respectivamente. En el ejemplo presente hay dos factores y cada uno de ellos tiene dos niveles. De esto resultan cuatro celdas o condiciones: a_1b_1 (ejercicio ligero por la mañana), a_1b_2 (ejercicio intenso por la mañana), a_2b_1 (ejercicio ligero por la tarde) y a_2b_2 (ejercicio intenso por la tarde). En virtud de que se trata de un diseño de grupos independientes, los sujetos deben ser asignados en forma aleatoria a cada una de las celdas, de manera que un grupo diferente de sujetos ocupe cada celda. Como quiera que los niveles de cada factor fueron seleccionados sistemáticamente por el experimentador, en lugar de ser elegidos de manera aleatoria, decimos que se trata de un diseño de *efectos fijos*.

Se han llevado a cabo tres análisis en este diseño. En primer lugar, queremos averiguar si el factor A tiene un efecto significativo, haciendo caso omiso del efecto del factor B. En este ejemplo nos interesa determinar si “la hora del día” produce alguna diferencia en el efecto del ejercicio sobre el sueño, sin tomar en cuenta el efecto de “la intensidad del ejercicio”. En segundo lugar, deseamos determinar si el factor B tiene un efecto significativo, sin tomar en cuenta el efecto del factor A. En este experimento, nos interesa averiguar si la intensidad del ejercicio produce algunas diferencias en la actividad del sueño, haciendo caso omiso del efecto de la hora del día. Por último, queremos determinar si existe alguna interacción entre los factores A y B. En el presente experimento, nos proponemos averiguar si hay alguna interacción entre la hora del día y la intensidad del ejercicio, en cuanto a su efecto sobre el sueño.

La figura 17.2 muestra algunos resultados posibles de este experimento. En la parte (a) no se perciben efectos significativos. En la parte (b) hay un efecto principal significativo para la hora del día, pero no hay efecto alguno para la intensidad del ejercicio y tampoco hay interacción. Por lo tanto, los sujetos disfrutaron de un grado significativamente

definición ■ *El efecto del factor A (promediado sobre los niveles del factor B) y el efecto del factor B (promediado sobre los niveles del factor A) reciben el nombre de **efectos principales**. Un **efecto de interacción** se presenta cuando el efecto de un factor no es igual en todos los niveles del otro factor.*

mayor de sueño cuando hacen ejercicio por la mañana, en lugar de hacerlo por la tarde. Sin embargo, parece que no tiene importancia que dicho ejercicio sea ligero o intenso. En la parte (c) se presenta un efecto principal significativo en cuanto a la intensidad del ejercicio, pero ningún efecto en relación con la hora del día y tampoco hay interacción. En este ejemplo, el ejercicio intenso produce un grado significativamente mayor de sueño que el ejercicio ligero, y no importa si el ejercicio se realiza en la mañana o por la tarde, ya que en ambos casos el efecto parece ser el mismo. La parte (d) muestra un efecto principal significativo en cuanto a la intensidad del ejercicio y la hora del día, y no hay efecto alguno de interacción.

Tanto la parte (e) como la (f) muestran efectos de interacción significativos. Como ya señalamos, la esencia de una interacción consiste en que el efecto de un factor no es igual en todos los niveles del otro factor. Esto significa que, cuando se presenta una interacción entre los factores A y B, las diferencias observadas en la variable dependiente a causa de los cambios introducidos en un factor no son iguales en cada uno de los niveles del otro factor. En la parte (e) existe un efecto de interacción significativo entre la intensidad del ejercicio y la hora del día. El efecto de diferentes intensidades del ejercicio no es igual en todos los niveles de la hora del día. De esta manera, si el ejercicio se realiza por la tarde, entonces el ejercicio ligero propicia un grado significativamente mayor de sueño que el ejercicio intenso. Por otra parte, si el ejercicio se realiza por la mañana,

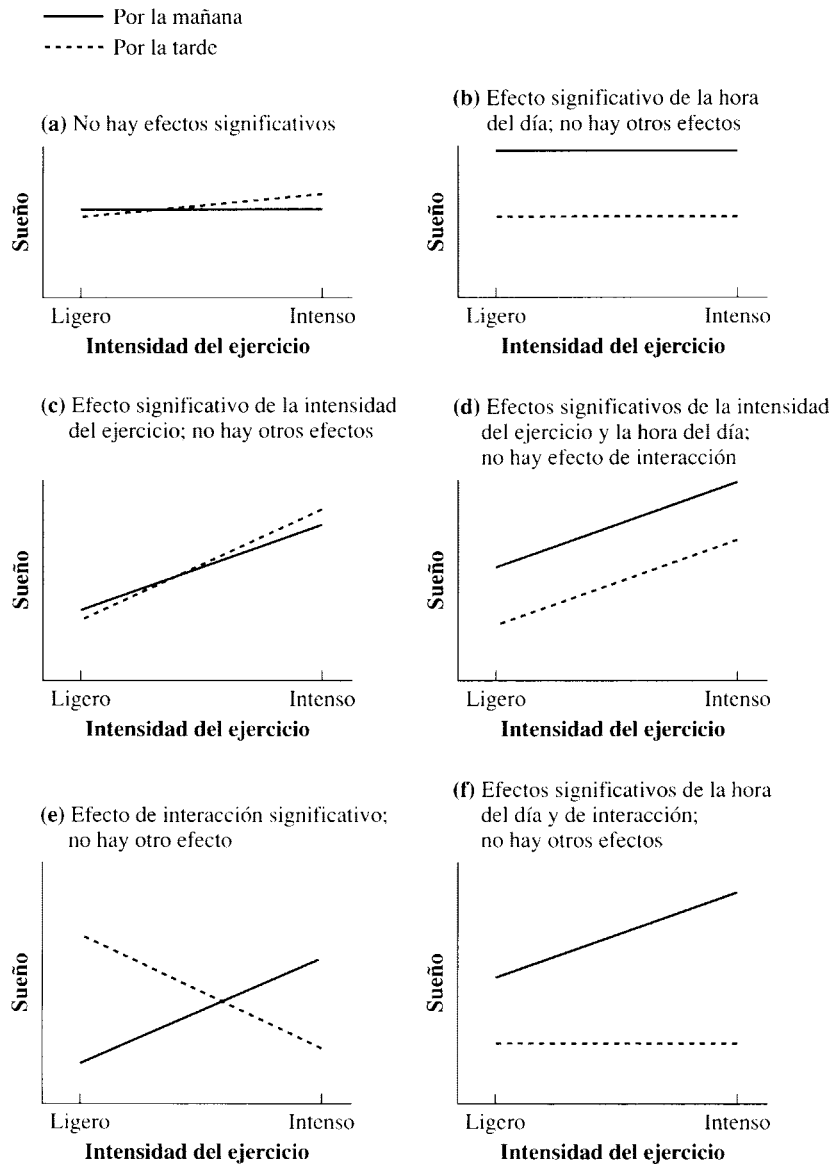


figura 17.2 Algunos resultados posibles del experimento para investigar los efectos de la intensidad del ejercicio y la hora del día.

entonces el ejercicio ligero propicia un grado significativamente menor de sueño que el ejercicio intenso. En la parte (f), la hora del día produce un efecto principal significativo y el efecto de interacción también es significativo. Por lo tanto, cuando el ejercicio se realiza por la mañana, el resultado es un grado significativamente mayor de sueño que cuando se realiza por la tarde, independientemente de que el ejercicio sea ligero o intenso. Además de este efecto principal, existe una interacción entre la intensidad del ejercicio y la hora del día. De esta manera, no hay diferencia alguna entre el efecto de las dos intensidades cuando el ejercicio se realiza por la tarde; en cambio, cuando se realiza por la mañana, el ejercicio intenso propicia mayor cantidad de sueño que el ejercicio ligero.

Cuando examinamos los datos de un diseño de análisis de varianza de dos factores, determinamos cuatro estimaciones de varianza: s_W^2 , s_R^2 , s_C^2 y s_{RC}^2 . La estimación de s_W^2 es la *estimación de la varianza dentro de las celdas* y corresponde a la estimación de la

varianza intra grupos que se utiliza en el análisis de varianza de un factor. Ésta se convierte en la norma con la cual se compara cada una de las demás estimaciones. Las otras estimaciones son sensibles a los efectos de las variables independientes. La expresión s_R^2 se conoce como la *estimación de la varianza por renglones*. Se basa en la variabilidad de las medias de los renglones (véase la figura 17.1) y, por lo tanto, es sensible a los efectos de la variable A. La expresión s_C^2 se conoce como la *estimación de la varianza por columnas*. Se basa en la variabilidad de las medias de las columnas y, por lo tanto, es sensible a los efectos de la variable B. La estimación s_{RC}^2 recibe el nombre de *estimación de la varianza renglón \times columna* (léase "renglón por columna"). Se basa en la variabilidad de las medias de las celdas y, por lo tanto, es sensible a los efectos de interacción de las variables A y B. Si la variable A no produce efecto alguno s_R^2 es una estimación independiente de σ^2 . Si la variable B no produce efecto alguno, entonces s_C^2 es una estimación independiente de σ^2 . Por último, si no existe interacción entre las variables A y B, entonces s_{RC}^2 también es una estimación independiente de σ^2 . Así pues, las estimaciones s_R^2 , s_C^2 y s_{RC}^2 son análogas a la estimación de la varianza entre grupos en el diseño de un factor. Para poner a prueba la significación, se formulan tres razones F :

$$\begin{array}{ll} \text{Para la variable A,} & F_{\text{obt}} = \frac{s_R^2}{s_W^2} \\ \text{Para la variable B,} & F_{\text{obt}} = \frac{s_C^2}{s_W^2} \\ \text{Para la interacción} & F_{\text{obt}} = \frac{s_{RC}^2}{s_W^2} \\ \text{entre A y B.} & \end{array}$$

Cada valor de F_{obt} se evalúa contra F_{crit} , tal como se hace en el análisis de un factor. Para realizar la comparación por renglones, si $F_{\text{obt}} \geq F_{\text{crit}}$, entonces hay un efecto principal significativo para el factor A. Si $F_{\text{obt}} \geq F_{\text{crit}}$ para la comparación por columnas, entonces hay un efecto principal significativo para el factor B. Finalmente, si $F_{\text{obt}} \geq F_{\text{crit}}$ para la comparación renglón \times columna, entonces hay un efecto de interacción significativo. Así pues, existen muchas semejanzas entre los diseños de un factor y los de dos factores. La principal diferencia entre ellos consiste en que, con el diseño de dos factores, podemos realizar esencialmente dos experimentos de un factor y, además, podemos evaluar la interacción entre las dos variables independientes.

Hasta este punto hemos estudiado de manera cualitativa el análisis de varianza de dos factores, con un diseño de efectos fijos para grupos independientes. En el resto de este capítulo presentaremos un estudio cuantitativo más detallado sobre el análisis de los datos para este diseño.



PRESENTACIÓN CUANTITATIVA DEL ANÁLISIS DE VARIANZA DE DOS FACTORES

En el análisis de varianza de un factor, la suma total de cuadrados se divide en dos componentes: la suma de cuadrados intra grupos y la suma de cuadrados entre grupos. Estos dos componentes se dividen entre los grados de libertad apropiados para formar dos estimaciones de la varianza: la estimación de la varianza intra grupos (s_W^2) y la estimación de la varianza entre grupos (s_B^2). Si la hipótesis nula es correcta, entonces ambas son estimaciones de la varianza poblacional de la hipótesis nula (σ^2) y la razón s_B^2/s_W^2 se distribuirá como F . Si la variable independiente tiene un efecto real, entonces s_B^2 tenderá a ser mayor que en el caso contrario y lo mismo ocurrirá con la razón F . De esta manera,

mientras mayor sea la razón F , tanto menos razonable será la hipótesis nula. Cuando $F_{\text{obt}} \geq F_{\text{crit}}$, rechazaremos H_0 porque será una explicación poco razonable de los datos.

La situación es muy similar en el análisis de varianza de dos factores. Sin embargo, en este último caso dividimos la suma total de cuadrados (SC_T) en cuatro componentes: la suma de cuadrados dentro de las celdas (SC_W), la suma de cuadrados por renglón (SC_R), la suma de cuadrados por columna (SC_C), y la suma de cuadrados renglón \times columna (SC_{RC}). Esta división se ilustra en la figura 17.3. Cuando estas sumas de cuadrados se dividen entre los grados de libertad apropiados, forman cuatro estimaciones de la varianza que son las siguientes: la estimación de la varianza dentro de las celdas (s_W^2), la estimación de la varianza por renglón (s_R^2), la estimación de la varianza por columna (s_C^2), y la estimación de la varianza renglón \times columna (s_{RC}^2). Al analizar cada una de estas estimaciones de la varianza será conveniente usar como referencia la figura 17.4, la cual muestra la notación y la distribución general de los datos para un análisis de varianza en dos factores, con un diseño de grupos independientes. En la siguiente exposición hemos supuesto que el número de sujetos contenidos en cada celda es el mismo.

Estimación de la varianza dentro de las celdas (s_W^2)

Esta estimación se deduce de la variabilidad que presentan los datos dentro de cada celda. En virtud de que todos los sujetos reciben los mismos niveles de las variables A y B, dentro de cada celda, la variabilidad que muestran los datos de estos sujetos no puede deberse a diferencias en el tratamiento. Esta estimación es análoga a la de la varianza intra grupos utilizada en el análisis de varianza de un factor; es una medida de la variabilidad inherente a los datos y, por tanto, proporciona una estimación de la varianza

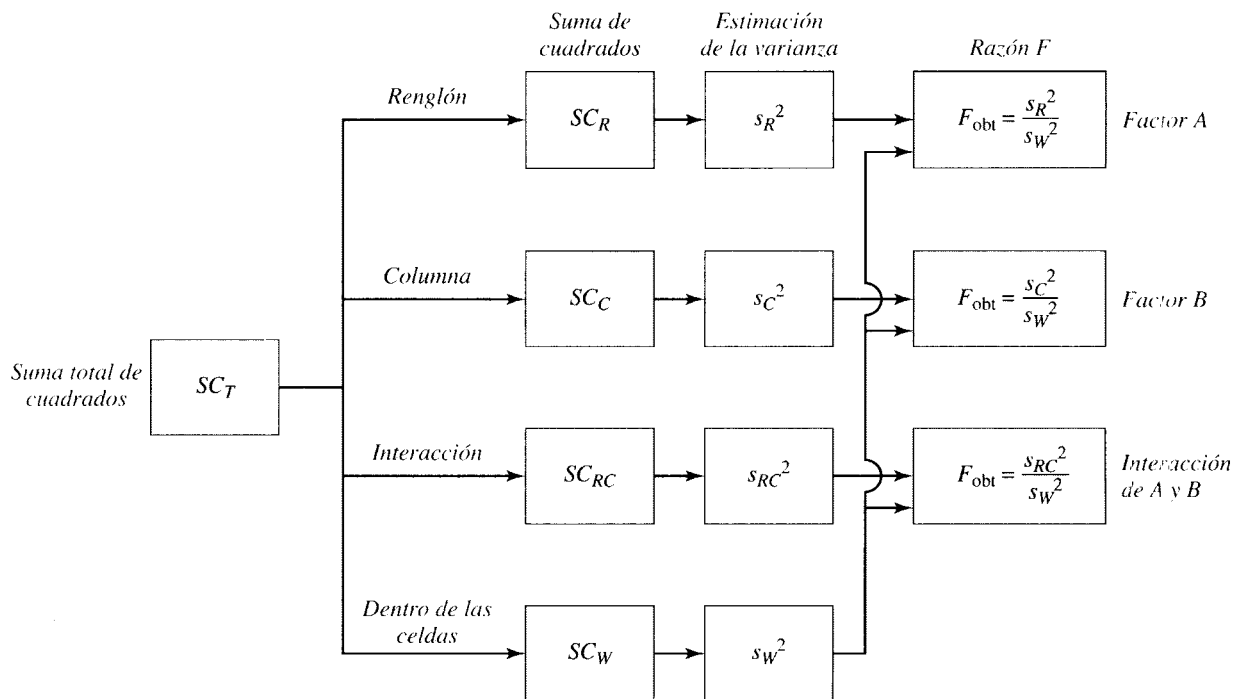


figura 17.3 Panorama de la técnica del análisis de varianza de dos factores, con un diseño de grupos independientes.

		Niveles del factor B				
		b_1	b_2		b_c	
<i>Niveles del factor A</i>	a_1	Celda 11 — — — —	Celda 12 — — — —	— — — —	Celda 1c — — — —	Renglón 1 $\sum X$
	a_2	Celda 21 — — — —	— — — —	— — — —	— — — —	Renglón 2 $\sum X$
		— — — —	— — — —	— — — —	— — — —	
	a_r	Celda r1 — — — —	— — — —	— — — —	Celda rc — — — —	Renglón r $\sum X$
		col. 1 $\sum X$	col. 2 $\sum X$		col. c $\sum X$	

donde a_1 = primer nivel del factor A
 a_r = último nivel del factor A
 b_1 = primer nivel del factor B
 b_c = último nivel del factor B

renglón 1
 $\sum X$ = suma de los datos del renglón 1

renglón r
 $\sum X$ = suma de los datos del renglón r

col. 1
 $\sum X$ = suma de los datos de la columna 1

col. c
 $\sum X$ = suma de los datos de la columna c

r = número de renglones
 c = número de columnas
 N = número total de datos
 n_{celda} = número de datos en cada celda
 $n_{\text{renglón}}$ = número de datos en cada renglón
 n_{col} = número de datos en cada columna

figura 17.4 Notación y distribución general de los datos en un análisis de varianza de dos factores.

poblacional (σ^2) de la hipótesis nula. Es el patrón contra el cual se compara cada una de las demás estimaciones de la varianza. En forma de ecuación, tenemos que,

$$s_w^2 = \frac{SC_w}{gl_w} \quad \text{ecuación para la estimación de la varianza dentro de las celdas}$$

donde SC_w = suma de cuadrados dentro de las celdas
 gl_w = grados de libertad dentro de las celdas

La suma de cuadrados dentro de las celdas (SC_w) es precisamente el total de las sumas de cuadrados dentro de cada celda. En forma conceptual, tenemos,

$$SC_w = SC_{11} + SC_{12} + \dots + SC_{rc} \quad \text{ecuación conceptual para la suma de cuadrados dentro de las celdas}$$

donde SC_{11} = suma de cuadrados para los datos de la celda definida por la intersección del renglón 1 y la columna 1
 SC_{12} = suma de cuadrados para los datos de la celda definida por la intersección del renglón 1 y la columna 2
 SC_{rc} = suma de cuadrados para los datos de la celda definida por la intersección del renglón r y la columna c ; ésta es la última celda de la matriz

Como suele suceder, la ecuación conceptual no es la mejor para propósitos de cálculo. La ecuación más apropiada para el cálculo es la siguiente:

$$SC_w = \sum_{\text{todos los datos}} X^2 - \left[\frac{\left(\sum_{\text{celda 11}} X \right)^2 + \left(\sum_{\text{celda 12}} X \right)^2 + \dots + \left(\sum_{\text{celda rc}} X \right)^2}{n_{\text{celda}}} \right]$$

ecuación para el cálculo de la suma de cuadrados dentro de las celdas

Observe la semejanza de estas ecuaciones con las que se utilizaron para la estimación de la varianza intra grupos, empleada en el análisis de varianza de un factor. La única diferencia es que, en el análisis de varianza de dos factores, la suma se realiza con respecto a las celdas, mientras que en el análisis de varianza de un factor, la suma se realiza conforme a los grupos.

Al calcular SC_w , existen n puntajes de desviación para cada celda. Por lo tanto, existen $n - 1$ grados de libertad por cada celda. En vista de que sumamos en todas las celdas para calcular SC_w , el número de grados de libertad dentro de las celdas es igual a $n - 1$ por el número de celdas. Si r es igual al número de renglones y c el número de columnas, entonces rc es igual al número de celdas. Por lo tanto, el número de grados de libertad dentro de las celdas es igual a $rc(n - 1)$. De esta manera, obtendremos,

$$gl_w = rc(n - 1) \quad \text{grados de libertad dentro de las celdas}$$

donde r = número de renglones
 c = número de columnas

Estimación de la varianza por renglón (s_R^2)

Esta estimación se basa en las diferencias que existen entre las medias por renglón; es análoga a la estimación de la varianza entre grupos (s_B^2) que se usa en el análisis de varianza de un factor. Usted recordará que s_B^2 es una estimación de σ^2 más el efecto de la variable independiente. De manera similar, el cálculo de la varianza por renglón (s_R^2) en el análisis de varianza de dos factores es una estimación de σ^2 más el efecto del factor A. Si el factor A no tiene efecto alguno, entonces las medias de la población por renglón son

iguales ($\mu_{a_1} = \mu_{a_2} = \dots = \mu_{a_r}$), y las diferencias entre las medias muestrales por renglón se deberán sólo a un muestreo aleatorio de poblaciones que son idénticas. En este caso, s_R^2 es una estimación de σ^2 . Si el factor A produce algún efecto, entonces las diferencias entre las medias por renglón, y por lo tanto s_R^2 , tenderán a ser mayores que en el caso anterior. Expresado en forma de ecuación, sería así:

$$s_R^2 = \frac{SC_R}{gl_R} \quad \text{ecuación para la estimación de la varianza por renglón}$$

donde SC_R = suma de cuadrados por renglón
 gl_R = grados de libertad por renglón

La suma de cuadrados por renglón es muy similar a la suma de cuadrados entre grupos que vimos en el análisis de varianza de un factor. La única diferencia es que con la suma de cuadrados por renglón empleamos las medias por renglón, mientras que en la suma de cuadrados entre grupos utilizamos las medias de grupo. De ahí obtenemos la ecuación conceptual para SC_R . Observe que al calcular las medias por renglón combinamos y promediamos todos los datos de un renglón dado. Esto se conoce como el cálculo de las medias de renglón “promediado sobre las columnas” (véase la figura 17.4). Así, obtenemos las medias de renglón promediando las columnas:

$$SC_R = n_{\text{renglón}} [(\bar{X}_{\text{renglón 1}} - \bar{X}_G)^2 + (\bar{X}_{\text{renglón 2}} - \bar{X}_G)^2 + \dots + (\bar{X}_{\text{renglón } r} - \bar{X}_G)^2] \quad \text{ecuación conceptual para la suma de cuadrados por renglón}$$

$$\text{donde } \bar{X}_{\text{renglón 1}} = \frac{\sum_{\text{renglón 1}} X}{n_{\text{renglón 1}}}$$

$$\bar{X}_{\text{renglón } r} = \frac{\sum_{\text{renglón } r} X}{n_{\text{renglón } r}}$$

$$\bar{X}_G = \text{gran media}$$

La ecuación conceptual nos permite apreciar que SC_R aumenta con el efecto de la variable A. Cuando este efecto se incrementa, las medias del renglón se separan cada vez más, lo cual, a su vez, hace que $(\bar{X}_{\text{renglón 1}} - \bar{X}_G)^2, (\bar{X}_{\text{renglón 2}} - \bar{X}_G)^2, \dots, (\bar{X}_{\text{renglón } r} - \bar{X}_G)^2$ aumente. Como estos términos están en el numerador, SC_R se incrementa. Por supuesto, SC_R se incrementa, lo mismo sucede con s_R^2 .

Al calcular SC_R , tenemos r puntajes de desviación. Por lo tanto, el número de grados de libertad por renglón es $r - 1$. En forma de ecuación, tenemos:

$$gl_R = r - 1 \quad \text{grados de libertad por renglón}$$

Recuerde que los grados de libertad entre grupos (gl_B) = $k - 1$ para el análisis de varianza de un factor. Los grados de libertad por renglón son muy similares, excepto que ahora utilizamos renglones en lugar de grupos.

De nuevo, la ecuación conceptual no es la mejor para calcular SC_R . La ecuación para el cálculo es la siguiente:

$$SC_R = \left[\frac{\left(\sum_{\text{renglón 1}} X \right)^2 + \left(\sum_{\text{renglón 2}} X \right)^2 + \dots + \left(\sum_{\text{renglón } r} X \right)^2}{n_{\text{renglón}}} \right] - \frac{\left(\sum_{\text{todos los datos}} X \right)^2}{N}$$

ecuación para el cálculo de la suma de cuadrados por renglón

Estimación de la varianza por columna (s_C^2)

Esta estimación se basa en las diferencias que se presentan entre las medias por columna. Es exactamente igual a s_R^2 , excepto que utiliza las medias por columna en lugar de las medias por renglón. Como el factor B afecta a las medias por columna, el cálculo de la varianza por columna (s_C^2) es una estimación de σ^2 más los efectos del factor B. Si los niveles del factor B no poseen un efecto diferencial, entonces las medias poblacionales por columna son iguales ($\mu_{b_1} = \mu_{b_2} = \mu_{b_3} = \dots = \mu_{b_c}$) y las diferencias entre las medias muestrales por columna se deben a un muestreo aleatorio de poblaciones que son idénticas. En este caso, s_C^2 será sólo una estimación de σ^2 . Si el factor B produce algún efecto, entonces las diferencias entre las medias por columna, y por lo tanto s_C^2 , tenderán a ser más grandes de lo que serían en caso contrario.

La ecuación para s_C^2 es

$$s_C^2 = \frac{SC_C}{gl_C} \quad \text{estimación de la varianza por columna}$$

donde SC_C = suma de cuadrados por columna
 gl_C = grados de libertad por columna

La suma de cuadrados por columna también es muy parecida a la suma de cuadrados por renglón. La única diferencia es que para calcular la suma de los cuadrados empleamos las medias por columna, en lugar de las medias por renglón. Observe que al calcular las medias por columna, combinamos y promediamos todos los datos de una columna determinada. De esta manera, obtenemos las medias por columna promediando los renglones.

$$SC_C = n_{\text{col.1}}(\bar{X}_{\text{col.1}} - \bar{X}_G)^2 + (\bar{X}_{\text{col.2}} - \bar{X}_G)^2 + \dots + (\bar{X}_{\text{col.c}} - \bar{X}_G)^2$$

ecuación conceptual para la suma de cuadrados por columna

$$\text{donde } \bar{X}_{\text{col.1}} = \frac{\sum_{\text{col.1}}^{\text{col.}} X}{n_{\text{col.1}} W}$$

$$\bar{X}_{\text{col.c}} = \frac{\sum_{\text{col.c}}^{\text{col.}} X}{n_{\text{col.c}}}$$

De nuevo, la ecuación conceptual muestra que SC_C aumenta con el efecto de la variable B. Cuando este efecto se incrementa, las medias por columna quedan más separadas, lo cual, a su vez, hace que $(\bar{X}_{\text{col.1}} - \bar{X}_G)^2$, $(\bar{X}_{\text{col.2}} - \bar{X}_G)^2$, \dots , $(\bar{X}_{\text{col.c}} - \bar{X}_G)^2$ aumenten. Como estos términos están en el numerador de la ecuación de SC_C , esto implica que SC_C también se incrementa. Por supuesto, un aumento en SC_C produce un incremento en s_C^2 .

Como utilizamos c datos de desviación para calcular SC_C , los grados de libertad por columna son $c - 1$. Por lo tanto,

$$gl_C = c - 1 \quad \text{grados de libertad por columna}$$

La ecuación para el cálculo de SC_C es

$$SC_C = \left[\frac{\left(\sum_{\text{col.1}}^{\text{col.}} X \right)^2 + \left(\sum_{\text{col.2}}^{\text{col.}} X \right)^2 + \dots + \left(\sum_{\text{col.c}}^{\text{col.}} X \right)^2}{n_{\text{col.}}} \right] - \frac{\left(\quad \right)^2}{N}$$

ecuación para el cálculo de la suma de cuadrados por columna

Estimación de la varianza renglón \times columna (s_{RC}^2)

En un pasaje anterior de este capítulo comentamos que existe una interacción cuando el efecto de una de las variables no es el mismo en todos los niveles de la otra variable. En otras palabras, se presenta una interacción cuando el efecto de la acción conjunta de las variables es distinto del que se podría predecir con los efectos individuales de las variables. Por ejemplo, consideremos la figura 17.2(f) de la página 398, donde hay una interacción entre la hora del día y la intensidad del ejercicio. Dicha interacción se debe a que el nivel de sueño que corresponde al ejercicio intenso, realizado en la mañana, es mayor del que podría predecirse con base en los efectos individuales de las variables hora del día e intensidad del ejercicio. Si no hubiera interacción, entonces esperaríamos que las rectas fueran paralelas. La variable intensidad del ejercicio tendría el mismo efecto si el ejercicio se realizara por la mañana o por la tarde.

Utilizamos la estimación de la varianza renglón \times columna (s_{RC}^2) para evaluar la interacción de las variables A y B. Como tal, ésta se basa en las diferencias entre las medias de las celdas, más allá de lo que se pueda predecir mediante los efectos individuales de las dos variables. El cálculo de la varianza renglón \times columna es una estimación de σ^2 más la interacción de A y B. Si no existe interacción alguna, y se eliminan los efectos principales, entonces las medias de las celdas son iguales ($\mu_{a_1b_1} = \mu_{a_1b_2} = \dots = \mu_{a_r b_c}$) y las diferencias entre las medias de las celdas se deben a un muestreo aleatorio de poblaciones que son idénticas. En este caso, s_{RC}^2 será tan sólo una estimación de σ^2 . Si se presenta una interacción entre los factores A y B, entonces las diferencias entre las medias de las celdas, y por lo tanto, s_{RC}^2 , tenderán a ser mayores de lo que serían en caso contrario.

La ecuación para s_{RC}^2 es

$$s_C^2 = \frac{SC_C}{gl_C} \quad \text{estimación de la varianza renglón} \times \text{columna}$$

donde SC_{RC} = suma de cuadrados renglón \times columna
 gl_{RC} = grados de libertad renglón \times columna

La suma de cuadrados renglón \times columna es igual a la variabilidad de las medias de las celdas, cuando se elimina la variabilidad ocasionada por los efectos individuales de los factores A y B. A continuación presentamos las ecuaciones conceptual y de cálculo:

$$SC_{RC} = n_{\text{celda}} [(\bar{X}_{\text{celda } 11} - \bar{X}_G)^2 + (\bar{X}_{\text{celda } 12} - \bar{X}_G)^2 + \dots + (\bar{X}_{\text{celda } rc} - \bar{X}_G)^2] - SC_R - SC_C$$

ecuación conceptual para la suma de cuadrados renglón \times columna

$$SC_{RC} = \left[\frac{\left(\sum_{\text{celda } 11}^{\text{celda } 11} X \right)^2 + \left(\sum_{\text{celda } 12}^{\text{celda } 12} X \right)^2 + \dots + \left(\sum_{\text{celda } rc}^{\text{celda } rc} X \right)^2 }{n_{\text{celda}}} \right] - \frac{\left(\sum_{\text{todos los datos}} X \right)^2}{N} - SC_R - SC_C$$

ecuación para el cálculo de la suma de cuadrados renglón \times columna

El número de grados de libertad para la estimación de la varianza renglón \times columna es igual a $(r - 1)(c - 1)$. Así,

$$gl_{RC} = (r - 1)(c - 1) \quad \text{grados de libertad renglón} \times \text{columna}$$

Cálculo de las razones F

Una vez determinadas las estimaciones de la varianza, éstas se utilizan junto con s_W^2 para formar razones F (véase la figura 17.3) y para verificar los efectos principales de las variables y su interacción. Calculemos las tres razones F siguientes:

Para verificar el efecto principal de la variable A (efecto por renglón):

$$F_{\text{obt}} = \frac{s_R^2}{s_W^2} = \frac{\sigma^2 + \text{efectos de la variable A}}{\sigma^2}$$

Para verificar el efecto principal de la variable B (efecto por columna):

$$F_{\text{obt}} = \frac{s_C^2}{s_W^2} = \frac{\sigma^2 + \text{efectos de la variable B}}{\sigma^2}$$

Para verificar la interacción de las variables A y B (efecto renglón \times columna):

$$F_{\text{obt}} = \frac{s_{RC}^2}{s_W^2} = \frac{\sigma^2 + \text{efectos de la interacción de A y B}}{\sigma^2}$$

La razón s_R^2/s_W^2 se emplea para verificar el efecto principal de la variable A. Si la variable A no tiene efecto principal, entonces, s_R^2 es una estimación independiente de σ^2 y s_R^2/s_W^2 se distribuye como F con grados de libertad iguales a gl_R y gl_W . Si la variable A tiene un efecto principal, s_R^2 será mayor de lo que sería en caso contrario, y el valor F_{obt} para los renglones aumentará.

La razón s_C^2/s_W^2 se utiliza para verificar el efecto principal de la variable B. Si la variable B no tiene efecto principal, entonces s_C^2 es una estimación independiente de σ^2 y s_C^2/s_W^2 se distribuye como F con grados de libertad iguales a gl_C y gl_W . Si la variable B tiene un efecto principal, s_C^2 será mayor de lo que sería en caso contrario, y el valor F_{obt} para las columnas aumentará.

Confirmamos la interacción entre A y B con la razón s_{RC}^2/s_W^2 . Si no existe interacción, s_{RC}^2 es una estimación independiente de σ^2 y s_{RC}^2/s_W^2 distribuye como F con grados de libertad iguales a gl_{RC} y gl_W . Si hay una interacción, s_{RC}^2 será mayor de lo que sería en caso contrario, causando un incremento en el valor de F_{obt} para la interacción.

Verificamos el efecto principal de cada variable y su interacción, comparando el valor F_{obt} apropiado con F_{crit} . F_{crit} se encuentra en la tabla F del apéndice D, al utilizar α y los grados de libertad del valor F que se está evaluando. La regla de decisión es la misma que en cualquier análisis de varianza de un factor, a saber:

Si $F_{\text{obt}} \geq F_{\text{crit}}$, rechazamos H_0 . *regla de decisión para evaluar H_0 en un análisis de varianza de dos factores*



ANÁLISIS DE UN EXPERIMENTO CON ANOVA DE DOS FACTORES

Ahora ya estamos en condiciones de analizar los datos de un ejemplo ilustrativo.

un experimento

Efecto del ejercicio sobre el sueño

Supongamos que un profesor de educación física realiza un experimento para comparar los efectos que producen sobre el sueño nocturno diferentes cantidades de ejercicio y la hora del día en que se lleva a cabo dicho ejercicio. Para este experimento se utiliza un diseño factorial de 3×2 de efectos fijos con grupos independientes. Hay tres niveles de ejercicio (ligero, moderado e intenso) y dos horarios al día (por la mañana y por la tarde). Treinta y seis estudiantes universitarios en buena condición física son repartidos de manera aleatoria entre las seis

celdas, de manera que haya 6 sujetos en cada celda. Los individuos que realizan un ejercicio intenso corren al trote 3 millas, los que hacen un ejercicio moderado trotan 1 milla y los sujetos sometidos a la condición del ejercicio ligero corren al trote $\frac{1}{4}$ de milla solamente. El ejercicio matutino se realiza a las 7:30 horas, mientras que el vespertino se lleva a cabo a las 19:00 horas. Cada sujeto realiza su ejercicio una vez y se lleva un registro del número de horas que duerme esa noche. Así se obtienen los datos presentados en la tabla 17.1.

1. ¿Cuáles son las hipótesis nulas para este experimento?
2. Utilizando $\alpha = 0.05$, ¿cuál es la conclusión de usted?

SOLUCIÓN

1. Hipótesis nulas:

- a. *Para la variable A (efecto principal):* La hora del día en que se realiza el ejercicio no afecta el sueño por la noche. Las medias poblacionales por renglón para el ejercicio matutino y vespertino, promediadas sobre los distintos niveles de ejercicio, son iguales ($\mu_{a_1} = \mu_{a_2}$).
- b. *Para la variable B (efecto principal):* Los distintos niveles de ejercicio tienen el mismo efecto sobre el sueño por la noche. Las medias poblacionales por columna para el ejercicio ligero, moderado e intenso, promediadas sobre las distintas horas del día, son iguales ($\mu_{b_1} = \mu_{b_2} = \mu_{b_3}$).
- c. *Para la interacción entre A y B:* No hay interacción alguna entre la hora del día y el nivel del ejercicio. Si se eliminan los efectos principales, entonces las medias poblacionales por celda son iguales ($\mu_{a_1b_1} = \mu_{a_1b_2} = \mu_{a_1b_3} = \mu_{a_2b_1} = \mu_{a_2b_2} = \mu_{a_2b_3}$).

2. Conclusión, utilizando $\alpha = 0.05$:

- a. Calculamos F_{obt} para cada hipótesis:

PASO 1: Cálculo de la suma de cuadrados por renglón, SC_R :

$$\begin{aligned}
 SC_R &= \left[\frac{\left(\sum_1^{\text{renglón}} X \right)^2 + \left(\sum_2^{\text{renglón}} X \right)^2}{n_{\text{renglón}}} \right] - \frac{\left(\sum_{\text{todos los datos}} X \right)^2}{N} \\
 &= \left[\frac{(129.2)^2 + (147.2)^2}{18} \right] - \frac{(276.4)^2}{36} \\
 &= 9.000
 \end{aligned}$$

tabla 17.1 Datos del experimento sobre el ejercicio

Hora del día	Ejercicio						
	Ligero		Moderado		Intenso		
Mañana	6.5	7.4	7.4	7.3	8.0	7.6	$\sum X = 129.20$
	7.3	7.2	6.8	7.6	7.7	6.6	$\sum X^2 = 930.50$
	6.6	6.8	6.7	7.4	7.1	7.2	$n = 18$
	$\bar{X} = 6.97$		$\bar{X} = 7.20$		$\bar{X} = 7.37$		$\bar{X} = 7.18$
Tarde	7.1	7.7	7.4	8.0	8.2	8.7	$\sum X = 147.20$
	7.9	7.5	8.1	7.6	8.5	9.6	$\sum X^2 = 1212.68$
	8.2	7.6	8.2	8.0	9.5	9.4	$n = 18$
	$\bar{X} = 7.67$		$\bar{X} = 7.88$		$\bar{X} = 8.98$		$\bar{X} = 8.18$
	$\sum X = 87.80$	$\sum X = 90.50$	$\sum X = 98.10$	$\sum X = 276.40$			
	$\sum X^2 = 645.30$	$\sum X^2 = 685.07$	$\sum X^2 = 812.81$	$\sum X^2 = 2143.18$			
	$n = 12$	$n = 12$	$n = 12$	$N = 36$			
	$\bar{X} = 7.32$	$\bar{X} = 7.54$	$\bar{X} = 8.18$				

PASO 2: Cálculo de la suma de cuadrados por columna, SC_C :

$$\begin{aligned}
 SC_C &= \left[\frac{\left(\sum_{\text{col. 1}} X \right)^2 + \left(\sum_{\text{col. 2}} X \right)^2 + \left(\sum_{\text{col. 3}} X \right)^2}{n_{\text{col.}}} \right] - \frac{\left(\sum_{\text{los datos}} X \right)^2}{N} \\
 &= \left[\frac{(87.8)^2 + (90.5)^2 + (98.1)^2}{12} \right] - \frac{(276.4)^2}{36} \\
 &= 4.754
 \end{aligned}$$

PASO 3: Cálculo de la suma de cuadrados renglón \times columna, SC_{RC} :

$$\begin{aligned}
 SC_{RC} &= \left[\frac{\left(\sum_{\text{celda 11}} X \right)^2 + \left(\sum_{\text{celda 12}} X \right)^2 + \left(\sum_{\text{celda 13}} X \right)^2 + \left(\sum_{\text{celda 21}} X \right)^2 + \left(\sum_{\text{celda 22}} X \right)^2 + \left(\sum_{\text{celda 23}} X \right)^2}{n_{\text{celda}}} \right] \\
 &\quad - \frac{\left(\sum_{\text{los datos}} X \right)^2}{N} - SC_R - SC_C - SC_C \\
 &= \left[\frac{(41.8)^2 + (43.2)^2 + (44.2)^2 + (46.0)^2 + (47.3)^2 + (53.9)^2}{6} \right] - \frac{(276.4)^2}{36} \\
 &\quad - 9.000 - 4.754 \\
 &= 1.712
 \end{aligned}$$

PASO 4: Cálculo de la suma de cuadrados dentro de las celdas, SC_W :

$$\begin{aligned}
 SC_W &= \sum_{\text{los datos}} X^2 \\
 &\quad - \left[\frac{\left(\sum_{\text{celda 11}} X \right)^2 + \left(\sum_{\text{celda 12}} X \right)^2 + \left(\sum_{\text{celda 13}} X \right)^2 + \left(\sum_{\text{celda 21}} X \right)^2 + \left(\sum_{\text{celda 22}} X \right)^2 + \left(\sum_{\text{celda 23}} X \right)^2}{n_{\text{celda}}} \right] \\
 &= 2143.18 - \left[\frac{(41.8)^2 + (43.2)^2 + (44.2)^2 + (46.0)^2 + (47.3)^2 + (53.9)^2}{6} \right] \\
 &= 5.577
 \end{aligned}$$

PASO 5: Cálculo de la suma total de cuadrados, SC_T : Este paso es una verificación para garantizar que los cálculos anteriores sean correctos. Una vez calculada SC_T , podemos utilizar la siguiente ecuación para comprobar los demás cálculos:

$$SC_T = SC_R + SC_C + SC_{RC} + SC_W$$

Primero debemos calcular SC_T :

$$\begin{aligned}
 SC_T &= \sum_{\text{los datos}} X^2 - \frac{\left(\sum_{\text{los datos}} X \right)^2}{N} \\
 &= 2143.18 - \frac{(276.4)^2}{36} \\
 &= 21.042
 \end{aligned}$$

Sustituyendo los valores obtenidos de SC_T , SC_R , SC_C , SC_{RC} y SC_W en la ecuación de componentes de la SC_T , obtenemos:

$$\begin{aligned}SC_T &= SC_R + SC_C + SC_{RC} + SC_W \\21.042 &= 9.000 + 4.754 + 1.712 + 5.577 \\21.042 &= 21.043\end{aligned}$$

La ecuación es correcta, salvo por un pequeño error de redondeo. Por lo tanto, podemos suponer que, hasta ahora, nuestros cálculos son correctos.

PASO 6: Cálculo de los grados de libertad para cada estimación de la varianza:

$$\begin{aligned}gl_R &= r - 1 = 2 - 1 = 1 \\gl_C &= c - 1 = 3 - 1 = 2 \\gl_{RC} &= (r - 1)(c - 1) = 2(1) = 2 \\gl_W &= rc(n_{celda} - 1) = 2(3)(5) = 30 \\gl_T &= N - 1 = 35\end{aligned}$$

Observe que:

$$\begin{aligned}gl_T &= gl_R + gl_C + gl_{RC} + gl_W \\35 &= 1 + 2 + 2 + 30 \\35 &= 35\end{aligned}$$

PASO 7: Cálculo de las estimaciones de la varianza s_R^2 , s_C^2 , s_{RC}^2 y s_W^2 : Cada estimación de la varianza es igual a la suma de cuadrados dividida entre los grados de libertad correspondientes. Así,

$$\begin{aligned}\text{Estimación de la varianza por renglón} &= s_R^2 = \frac{SC_R}{gl_R} = \frac{9.000}{1} = 9.000 \\ \text{Estimación de la varianza por columna} &= s_C^2 = \frac{SC_C}{gl_C} = \frac{4.754}{2} = 2.377 \\ \text{Estimación de la varianza renglón} \times \text{columna} &= s_{RC}^2 = \frac{SC_{RC}}{gl_{RC}} = \frac{1.712}{2} = 0.856 \\ \text{Estimación de la varianza dentro de las celdas} &= s_W^2 = \frac{SC_W}{gl_W} = \frac{5.577}{30} = 0.186\end{aligned}$$

PASO 8: Cálculo de razones F : Para el efecto por renglón,

$$F_{\text{obt}} = \frac{s_R^2}{s_W^2} = \frac{9.000}{0.186} = 48.42$$

Para el efecto por columna:

$$F_{\text{obt}} = \frac{s_C^2}{s_W^2} = \frac{2.377}{0.186} = 12.78$$

Para el efecto de interacción renglón \times columna:

$$F_{\text{obt}} = \frac{s_{RC}^2}{s_W^2} = \frac{0.856}{0.186} = 4.60$$

b. Evaluar los valores de F_{obt} :

Para el efecto por renglón: Según la tabla F, con $\alpha = 0.05$, $gl_{\text{numerador}} = gl_R = 1$ y $gl_{\text{denominador}} = gl_W = 30$, $F_{\text{crit}} = 4.17$. Como quiera que $F_{\text{obt}} (48.39) > 4.17$, rechazamos H_0 con respecto a la variable A, la cual es en este experimento la hora del día. Hay un efecto significativo para la hora del día.

Para el efecto por columna: Según la tabla F, con $\alpha = 0.05$, $gl_{\text{numerador}} = gl_C = 2$ y $gl_{\text{denominador}} = gl_W = 30$, $F_{\text{crit}} = 3.32$. Como quiera que $F_{\text{obt}} (12.78) > 3.32$, rechazamos H_0 con respecto a la variable B, la cual es en este experimento la cantidad de ejercicio. Hay un efecto principal significativo para la cantidad de ejercicio.

Para el efecto de interacción renglón \times columna: Según la tabla F, con $\alpha = 0.05$, $gl_{\text{numerador}} = gl_{RC} = 2$ y $gl_{\text{denominador}} = gl_W = 30$, $F_{\text{crit}} = 3.32$. Como quiera que $F_{\text{obt}} (4.60) > 3.32$, rechazamos H_0 con respecto a la interacción de las variables A y B. Hay una interacción significativa entre la cantidad de ejercicio y la hora del día a la cual se realiza.

El análisis se resume en la tabla 17.2.

tabla 17.2 Tabla resumida del análisis de varianza para el experimento sobre el ejercicio y la hora del día

Fuente	SC	gl	s^2	F_{obt}	F_{crit}
Renglones (hora del día)	9.000	1	9.000	48.42*	4.17
Columnas (ejercicio)	4.754	2	2.377	12.79*	3.32
Renglones \times columnas	1.712	2	0.856	4.60*	3.32
Dentro de las celdas	5.577	30	0.186		
Total	21.042	35			

*En vista de que $F_{\text{obt}} > F_{\text{crit}}$, se rechaza H_0 .

Interpretación de los resultados

En el análisis anterior rechazamos la hipótesis nula tanto para efectos de renglón como de columna. Un efecto significativo de renglón indica que la variable A tuvo un efecto principal significativo. Las diferencias entre las medias de renglón, promediadas sobre las columnas, fueron demasiado grandes como para ser atribuidas a un muestreo aleatorio de poblaciones en las que $\mu_{a1} = \mu_{a2}$. En el presente experimento, el efecto significativo de renglón indica que hubo un efecto principal significativo del factor hora del día. Las diferencias entre las medias para las condiciones de hora del día promediadas sobre las condiciones de cantidad de ejercicio fueron demasiado grandes como para atribuir las al azar. En la figura 17.5 graficamos la media de cada celda. En esta figura se puede ver que el ejercicio realizado por la tarde condujo a mayor sueño que el realizado por la mañana.

Un efecto significativo de columna indica que la variable B tuvo un efecto principal significativo: que las diferencias entre las medias de columnas, calculadas con los promedios de los renglones, fueron demasiado grandes como para ser atribuidas a muestreo aleatorio de la población correspondiente a la hipótesis nula. En el presente experimento, el efecto significativo de columnas nos señala que las diferencias entre las medias de las tres condiciones de ejercicio calculadas sobre las condiciones de hora del día fueron demasiado grandes como para ser atribuidas a fluctuaciones del muestreo aleatorio. En la figura 17.5 se puede ver que el efecto de aumentar la cantidad de ejercicio, promediado sobre las condiciones de hora del día, fue aumentar la cantidad de sueño.

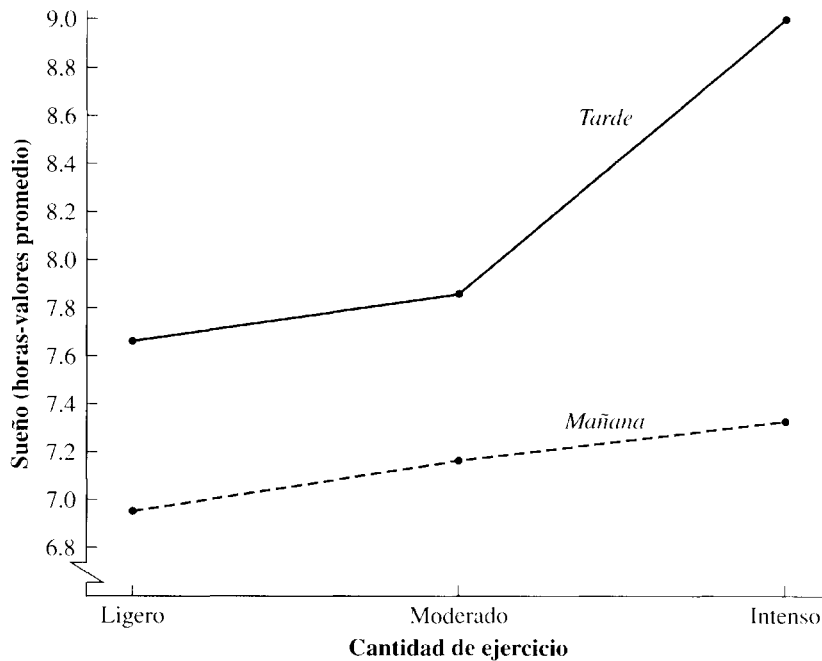
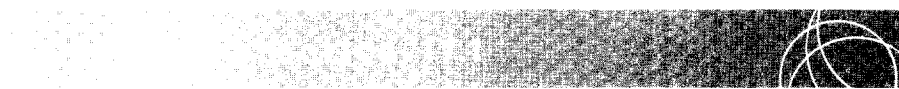


figura 17.5 Medias por celda para el experimento sobre ejercicio y sueño.

Los resultados de este experimento también muestran un efecto de interacción renglón por columna significativo. Como antes se dijo, un efecto de interacción significativo indica que los efectos de uno de los factores sobre la variable dependiente no son iguales para todos los niveles del otro factor. Para interpretar un efecto de interacción resulta de mucha ayuda graficar la media de cada celda. En la figura 17.5 podemos ver que la cantidad de sueño aumenta en la misma medida al pasar de ejercicio ligero a moderado, independientemente de que el ejercicio se realice por la mañana o por la tarde. Sin embargo, al pasar de ejercicio moderado a intenso, la diferencia en cantidad de sueño es distinta dependiendo de si el ejercicio se hace por la mañana o por la tarde. El ejercicio intenso da lugar a un aumento mucho mayor en la cantidad de sueño cuando el ejercicio se hace por la tarde que cuando se hace por la mañana.

Para practicar, resolvamos otro problema.



Un profesor de estadística realiza un experimento para comparar la efectividad de dos métodos para enseñar su materia. El método I consiste en la forma usual en que imparte su materia: clases, asignar tareas y un examen final. El método II es igual al método I, excepto que los estudiantes que reciben el método II toman una hora adicional por semana en la que resuelven problemas ilustrativos bajo la orientación del profesor. El profesor también está interesado en la manera en que los

(continúa)

métodos afectan a los estudiantes con diferentes habilidades para las matemáticas. Por consiguiente, los participantes voluntarios en el experimento se subdividen según su habilidad matemática en tres grupos: superior, promedio y pobre. Cinco estudiantes de cada grupo son asignados aleatoriamente al método I y cinco estudiantes de cada grupo al método II. Al terminar el curso, los 30 estudiantes reciben el mismo examen final. Se obtuvieron las siguientes calificaciones en el examen final:

Habilidad matemática	Método de enseñanza			
	Método I (1)		Método II (2)	
Superior	39*	41	49	47
(1)	48	42	47	48
	44		43	
Promedio	43	36	38	46
(2)	40	35	45	44
	42		42	
Pobre	30	33	37	41
(3)	29	36	34	33
	37		40	

*Cada dato es el total de puntos obtenidos de los 50 puntos posibles.

- ¿Cuál es la hipótesis nula para este experimento?
- Considerando un $\alpha = 0.05$, ¿cuál es su conclusión?

SOLUCIÓN

- Hipótesis nulas:

- Para la variable A (efecto principal):* Los tres niveles de habilidad matemática no afectaron de manera diferenciada las calificaciones en el examen final de esta materia. Las medias poblacionales por renglón para los tres niveles de habilidad matemática, promediadas respecto a los métodos de enseñanza, son iguales ($\mu_{a_1} = \mu_{a_2} = \mu_{a_3}$).
- Para la variable B (efecto principal):* Los métodos de enseñanza I y II tienen los mismos efectos sobre las calificaciones en el examen final de esta materia ($\mu_{b_1} = \mu_{b_2}$).
- Para la interacción entre las variables A y B:* No existe efecto de interacción entre las variables A y B. Removiendo cualquier efecto principal, las medias poblacionales de las celdas son iguales ($\mu_{a_1b_1} = \mu_{a_1b_2} = \mu_{a_2b_1} = \mu_{a_2b_2} = \mu_{a_3b_1} = \mu_{a_3b_2}$).

- Conclusión, utilice $\alpha = 0.05$:

- Cálculo de F_{obt} :

PASO 1: Cálculo de SC_R :

$$SC_R = \left[\frac{\left(\sum_1 X \right)^2 + \left(\sum_2 X \right)^2 + \left(\sum_3 X \right)^2}{n_{\text{renglón}}} \right] - \frac{\left(\sum_{\text{todos los datos}} X \right)^2}{N}$$

$$= \frac{(448)^2 + (411)^2 + (350)^2}{10} - \frac{(1209)^2}{30} = 489.800$$

PASO 2: Cálculo de SC_C :

$$SC_C = \left[\frac{\left(\sum_{\text{col. 1}} X \right)^2 + \left(\sum_{\text{col. 2}} X \right)^2}{n_{\text{col.}}} \right] - \frac{\left(\sum_{\text{todos los datos}} X \right)^2}{N}$$

$$= \frac{(575)^2 + (634)^2}{15} - \frac{(1209)^2}{30} = 116.033$$

PASO 3: Cálculo de SC_{RC} :

$$SC_{RC} = \left[\frac{\left(\sum_{\text{celda 11}} X \right)^2 + \left(\sum_{\text{celda 12}} X \right)^2 + \left(\sum_{\text{celda 21}} X \right)^2 + \left(\sum_{\text{celda 22}} X \right)^2 + \left(\sum_{\text{celda 31}} X \right)^2 + \left(\sum_{\text{celda 32}} X \right)^2}{n_{\text{celda}}} \right]$$

$$- \frac{\left(\sum_{\text{todos los datos}} X \right)^2}{N} - SC_R - SC_C$$

$$= \left[\frac{(214)^2 + (234)^2 + (196)^2 + (215)^2 + (165)^2 + (185)^2}{5} \right] - \frac{(1209)^2}{30}$$

$$- 498.8 - 116.083 = 49.328.6 - 48.722.7 - 489.8 - 116.033 = 0.067$$

PASO 4: Cálculo de SC_W :

$$SC_W = \sum_{\text{todos los datos}} X^2$$

$$- \left[\frac{\left(\sum_{\text{celda 11}} X \right)^2 + \left(\sum_{\text{celda 12}} X \right)^2 + \left(\sum_{\text{celda 21}} X \right)^2 + \left(\sum_{\text{celda 22}} X \right)^2 + \left(\sum_{\text{celda 31}} X \right)^2 + \left(\sum_{\text{celda 32}} X \right)^2}{n_{\text{celda}}} \right]$$

$$= 49.587 - \left[\frac{(214)^2 + (234)^2 + (196)^2 + (215)^2 + (165)^2 + (185)^2}{5} \right]$$

$$= 49.587 - 49.328.6 = 258.4$$

PASO 5: Cálculo de SC_T : Este paso permite verificar los cálculos anteriores:

$$SC_T = \sum_{\text{todos los datos}} X^2 - \frac{\left(\sum_{\text{todos los datos}} X \right)^2}{N}$$

$$= 49.587 - \frac{(1209)^2}{30}$$

$$= 864.3$$

$$SC_T = SC_R + SC_C + SC_{RC} + SC_W$$

$$864.3 = 489.8 + 116.033 + 0.067 + 258.4$$

$$864.3 = 864.3$$

(continúa)

En vista de que esta ecuación de partición es correcta, podemos suponer que nuestros cálculos hasta este punto son correctos.

PASO 6: Cálculo de los grados de libertad:

$$\begin{aligned}gl_R &= r - 1 = 3 - 1 = 2 \\gl_C &= c - 1 = 2 - 1 = 1 \\gl_{RC} &= (r - 1)(c - 1) = (3 - 1)(2 - 1) = 2 \\gl_W &= rc(n_{\text{celda}} - 1) = 6(4) = 24 \\gl_T &= N - 1 = 29\end{aligned}$$

PASO 7: Cálculo de s_R^2 , s_C^2 , s_{RC}^2 y s_W^2 :

$$\begin{aligned}s_R^2 &= \frac{SC_R}{gl_R} = \frac{489.8}{2} = 244.9 \\s_C^2 &= \frac{SC_C}{gl_C} = \frac{116.033}{1} = 116.033 \\s_{RC}^2 &= \frac{SC_{RC}}{gl_{RC}} = \frac{0.067}{2} = 0.034 \\s_W^2 &= \frac{SC_W}{gl_W} = \frac{258.4}{24} = 10.767\end{aligned}$$

PASO 8: Cálculo de F_{obt} : Para el efecto por renglón:

$$F_{\text{obt}} = \frac{s_R^2}{s_W^2} = \frac{244.9}{10.767} = 22.75$$

Para el efecto por columna:

$$F_{\text{obt}} = \frac{s_C^2}{s_W^2} = \frac{116.033}{10.767} = 10.78$$

Para el efecto de interacción renglón \times columna:

$$F_{\text{obt}} = \frac{s_{RC}^2}{s_W^2} = \frac{0.034}{10.767} = 0.00$$

2. Evaluación de los valores F_{obt} :

Para el efecto por renglón: Considerando la tabla F, con $\alpha = 0.05$, $gl_{\text{numerador}} = gl_R = 2$, y $gl_{\text{denominador}} = gl_W = 24$, $F_{\text{crit}} = 3.40$. Como $F_{\text{obt}} (22.75) > 3.40$, rechazamos H_0 con respecto a la variable A. Existe un efecto significativo de la habilidad matemática.

Para el efecto por columna: Con $\alpha = 0.05$, $gl_{\text{numerador}} = gl_C = 1$, y $gl_{\text{denominador}} = gl_W = 24$, según la tabla F, $F_{\text{crit}} = 4.26$. Puesto que $F_{\text{obt}} (10.78) > 4.26$, rechazamos H_0 con respecto a la variable B. Existe un efecto principal significativo para el método de enseñanza.

Para el efecto de interacción renglón \times columna: Puesto que $F_{\text{obt}} (0.00) < 1$, conservamos H_0 y concluimos que los datos no apoyan la hipótesis de que existe una interacción entre habilidad matemática y métodos de enseñanza.

La solución de este problema se resume en la tabla 17.3

tabla 17.3 Tabla resumida del análisis de varianza para el experimento sobre método de enseñanza y habilidad matemática

Fuente	SC	gl	s^2	F_{obt}	F_{crit}
Renglones (habilidad matemática)	489.800	2	244.900	22.75*	3.40
Columnas (método de enseñanza)	116.033	1	116.033	10.78*	4.26
Renglones \times columnas	0.067	2	0.034	0.00	3.40
Dentro de las celdas	258.400	24	10.767		
Total	864.300	29			

*En virtud de que $F_{\text{obt}} > F_{\text{crit}}$, se rechaza H_0 .

Interpretación de los resultados del problema de práctica 17.1 En el análisis anterior, rechazamos la hipótesis nula para los efectos por renglón y también por columna. El rechazo de H_0 para renglones nos indica que hubo un efecto principal significativo para la variable A, la habilidad matemática. Las diferencias entre las medias para los distintos niveles de habilidad matemática, promediada sobre los métodos de enseñanza, eran demasiado grandes como para ser atribuidas al azar. La media de cada celda se presenta en la figura 17.6. Esta figura muestra que a medida que se incrementa el nivel de habilidad matemática se obtienen mejores calificaciones en el examen final.

El rechazo de H_0 para las columnas indica que hubo un efecto principal significativo para la variable B, es decir, para los métodos de enseñanza. La diferencia registrada entre las medias, para los métodos I y II, promediadas sobre la habilidad matemática, resultó demasiado grande como para atribuirse a las fluctuaciones provocadas por el muestreo aleatorio. En la figura 17.6 se muestra que el método II fue superior al método I.

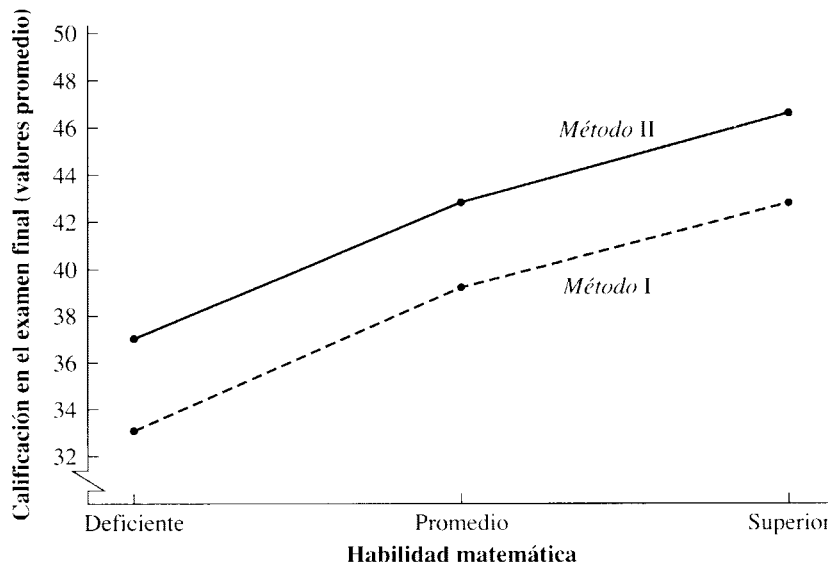


figura 17.6 Medias por celda para el experimento sobre los métodos de enseñanza y la habilidad matemática.

En este experimento no se presentó un efecto significativo de interacción. Lo anterior nos indica que, dentro de los límites de sensibilidad del experimento, el efecto de cada variable fue el mismo sobre todos los niveles de la otra variable. Esto lo observamos más claramente al analizar la figura 17.6 con respecto a la variable A. La carencia de un efecto significativo de interacción indica que el efecto de los diversos niveles de habilidad matemática sobre las calificaciones del examen final fue el mismo para los métodos de enseñanza I y II. Esto produce rectas paralelas al graficar las medias por celda (véase la figura 17.6). *De hecho, una regla general establece que cuando se presentan rectas paralelas en una gráfica de medias individuales por celda, usted puede estar seguro de que no existe un efecto de interacción.* Esto se debe a que, para que haya un efecto de interacción, las líneas deben desviarse significativamente de las paralelas. A este respecto, será útil ver la figura 17.2 y analizar qué gráficas muestran efectos de interacción.*



A una psicóloga clínica le interesa analizar el efecto de la ansiedad sobre la capacidad que poseen las personas para aprender nuevos materiales. Esta investigadora pretende averiguar también si el efecto de la ansiedad depende de la dificultad que presenta el nuevo material. Entonces realiza un experimento con tres niveles de ansiedad (alta, media y baja) y tres niveles de dificultad (alta, media y baja) del material por aprender. De un grupo de voluntarios, elige a 15 sujetos que muestran una ansiedad baja, 15 con ansiedad media y otros 15 con ansiedad alta, y asigna 5 individuos de cada grupo, de manera aleatoria, a cada uno de los tres niveles de dificultad. Cada sujeto dispone de 30 minutos para aprender el nuevo material, después de lo cual presenta una prueba para determinar la cantidad de material que ha aprendido.

Así se reunieron los siguientes datos:

Dificultad del material	Ansiedad*					
	Baja		Media		Alta	
Baja	18	17	18	18	18	17
	20	16	19	15	16	18
	17		17		19	
Media	18	14	18	17	14	15
	17	16	18	15	17	12
	14		14		16	
Alta	11	6	15	12	9	8
	10	10	13	11	7	8
	8		12		5	

*Cada dato es el total de puntos obtenidos de los 20 puntos posibles.

- a. ¿Cuáles son las hipótesis nulas?
- b. Utilice $\alpha = 0.05$ ¿Cuál es la conclusión de usted?

*Por supuesto, no se puede saber si la interacción es significativa, a menos que haga un análisis estadístico.

SOLUCIÓN**a. Hipótesis nulas:**

1. *Para la variable A (efectos principales):* La hipótesis nula establece que la dificultad del material presentado a los sujetos del experimento no tiene efecto alguno sobre la cantidad de material aprendido. Las medias poblacionales por renglón para los niveles de dificultad baja, media y alta, promediadas sobre los niveles de ansiedad, son iguales ($\mu_{a_1} = \mu_{a_2} = \mu_{a_3}$).
2. *Para la variable B (efecto principal):* La hipótesis nula establece que el nivel de ansiedad no influye sobre la cantidad de material aprendido. Las medias poblacionales por columna para los niveles de ansiedad baja, media y alta, promediadas sobre los tres niveles de dificultad, son iguales ($\mu_{b_1} = \mu_{b_2} = \mu_{b_3}$).
3. *Para la interacción entre las variables A y B:* La hipótesis nula establece que no existe efecto alguno de interacción entre la dificultad del material y la ansiedad. Al eliminar cualquiera de los efectos principales, las medias poblacionales por celda son iguales ($\mu_{a_1b_1} = \mu_{a_1b_2} = \dots = \mu_{a_3b_3}$).

b. Conclusión, utilice $\alpha = 0.05$:**1. Cálculo de F_{obt} :****PASO 1: Cálculo de SC_R :**

$$SC_R = \left[\frac{\left(\overset{\text{renglón}}{\sum_1 X} \right)^2 + \left(\overset{\text{renglón}}{\sum_2 X} \right)^2 + \left(\overset{\text{renglón}}{\sum_3 X} \right)^2}{n_{\text{renglón}}} \right] - \frac{\left(\overset{\text{todos los datos}}{\sum X} \right)^2}{N}$$

$$= \frac{(263)^2 + (235)^2 + (145)^2}{15} - \frac{(643)^2}{45} = 506.844$$

PASO 2: Cálculo de SC_C :

$$SC_C = \left[\frac{\left(\overset{\text{col.}}{\sum_1 X} \right)^2 + \left(\overset{\text{col.}}{\sum_2 X} \right)^2 + \left(\overset{\text{col.}}{\sum_3 X} \right)^2}{n_{\text{col.}}} \right] - \frac{\left(\overset{\text{todos los datos}}{\sum X} \right)^2}{N}$$

$$= \frac{(212)^2 + (232)^2 + (199)^2}{15} - \frac{(643)^2}{45}$$

$$= 36.844$$

PASO 3: Cálculo de SC_{RC} :

$$SC_{RC} = \left[\frac{\left(\overset{\text{celda}}{\sum_{11} X} \right)^2 + \left(\overset{\text{celda}}{\sum_{12} X} \right)^2 + \dots + \left(\overset{\text{celda}}{\sum_{33} X} \right)^2}{n_{\text{celda}}} \right] - \frac{\left(\overset{\text{todos los datos}}{\sum X} \right)^2}{N} - SC_R - SC_C$$

$$= \left[\frac{(88)^2 + (87)^2 + (88)^2 + (79)^2 + (82)^2 + (74)^2 + (45)^2 + (63)^2 + (37)^2}{5} \right]$$

$$- \frac{(643)^2}{45} - 506.844 - 36.844$$

$$= 40.756$$

(continúa)

PASO 4: Cálculo de SC_W :

$$SC_W = \sum_{\text{todos los datos}} X^2 - \left[\frac{\left(\sum_{\text{celda 11}} X \right)^2 + \left(\sum_{\text{celda 12}} X \right)^2 + \cdots + \left(\sum_{\text{celda 33}} X \right)^2}{n_{\text{celda}}} \right] = 9871$$

$$- \left[\frac{(88)^2 + (87)^2 + (88)^2 + (79)^2 + (82)^2 + (74)^2 + (45)^2 + (63)^2 + (37)^2}{5} \right]$$

$$= 98.800$$

PASO 5: Cálculo de SC_T : Este paso permite verificar los cálculos anteriores:

$$SC_T = \sum_{\text{todos los datos}} X^2 - \frac{\left(\sum_{\text{los datos}} X \right)^2}{N} = 9871 - \frac{(643)^2}{45}$$

$$= 683.244$$

$$SC_T = SC_R + SC_C + SC_{RC} + SC_W$$

$$683.224 = 506.844 + 36.844 + 40.756 + 98.800$$

$$683.224 = 683.224$$

En vista de que esta ecuación de partición es correcta, podemos suponer que nuestros cálculos hasta este punto son correctos.

PASO 6: Cálculo de los grados de libertad:

$$gl_R = r - 1 = 3 - 1 = 2$$

$$gl_C = c - 1 = 3 - 1 = 2$$

$$gl_{RC} = (r - 1)(c - 1) = 2(2) = 4$$

$$gl_W = rc(n_{\text{celda}} - 1) = 3(3)(5 - 1) = 36$$

$$gl_T = N - 1 = 45 - 1 = 44$$

PASO 7: Cálculo de s_R^2 , s_C^2 , s_{RC}^2 y s_W^2 :

$$s_R^2 = \frac{SC_R}{gl_R} = \frac{506.844}{2} = 253.422$$

$$s_C^2 = \frac{SC_C}{gl_C} = \frac{36.844}{2} = 18.442$$

$$s_{RC}^2 = \frac{SC_{RC}}{gl_{RC}} = \frac{40.756}{4} = 10.189$$

$$s_W^2 = \frac{SC_W}{gl_W} = \frac{98.800}{36} = 2.744$$

PASO 8: Cálculo de F_{obt} : Para el efecto por renglón:

$$F_{\text{obt}} = \frac{s_R^2}{s_W^2} = \frac{253.422}{2.744} = 92.35$$

Para el efecto por columna:

$$F_{\text{obt}} = \frac{s_C^2}{s_W^2} = \frac{18.442}{2.744} = 6.72$$

Para el efecto de interacción renglón \times columna:

$$F_{\text{obt}} = \frac{s_{RC}^2}{s_W^2} = \frac{10.189}{2.744} = 3.71$$

2. Evaluación de F_{obt} :

Para el efecto por renglón: Con $\alpha = 0.05$, $gl_{\text{numerador}} = gl_R = 2$, y $gl_{\text{denominador}} = gl_W = 36$, según la tabla F, $F_{\text{crit}} = 3.26$. Como $F_{\text{obt}} (92.35) > 3.26$, rechazamos H_0 con respecto a la variable A. Existe un efecto significativo principal para la dificultad del material.

Para el efecto por columna: Con $\alpha = 0.05$, $gl_{\text{numerador}} = gl_C = 2$, y $gl_{\text{denominador}} = gl_W = 36$, según la tabla F, $F_{\text{crit}} = 3.26$. Puesto que $F_{\text{obt}} (6.72) > 3.26$, rechazamos H_0 con respecto a la variable B. Existe un efecto principal significativo para el nivel de ansiedad.

Para el efecto de renglón \times columna: Con $\alpha = 0.05$, $gl_{\text{numerador}} = gl_{RC} = 4$ y $gl_{\text{denominador}} = gl_W = 36$, según la tabla F, $F_{\text{crit}} = 2.63$. Puesto que $F_{\text{obt}} (3.71) > 2.63$, rechazamos H_0 . Existe una interacción significativa entre la dificultad del material y el nivel de ansiedad.

La solución de este problema se resume en la tabla 17.4.

tabla 17.4 Tabla resumida del análisis de varianza para el experimento sobre la dificultad del material y el nivel de ansiedad

Fuente	SC	gl	s^2	F_{obt}	F_{crit}
Renglones (dificultad del material)	506.844	2	253.422	92.35*	3.26
Columnas (nivel de ansiedad)	36.844	2	18.442	6.72*	3.26
Renglones \times columnas	40.756	4	10.189	3.71*	2.63
Dentro de las celdas	98.800	36	2.744		
Total	683.244	44			

*En virtud de que $F_{\text{obt}} > F_{\text{crit}}$, se rechaza H_0 .

Interpretación de los resultados del problema de práctica 17.2 En el análisis anterior se registró un efecto principal significativo, tanto para la dificultad del material como para el nivel de ansiedad. El efecto principal significativo de la dificultad del material indica que las diferencias entre las medias para los tres niveles de dificultad, promediadas sobre los niveles de ansiedad, fueron demasiado grandes como para ser atribuidos al azar. Las medias por celda se presenta como gráfica en la figura 17.7. Esta figura muestra que al incrementar la dificultad del material se obtienen medias menores, cuando las calificaciones se promedian sobre los niveles de ansiedad.

El efecto principal significativo del nivel de ansiedad es más difícil de interpretar. Por supuesto, en el nivel operativo, este efecto principal nos indica que las diferencias entre las medias para los tres niveles de ansiedad, cuando se promediaron sobre los nive-

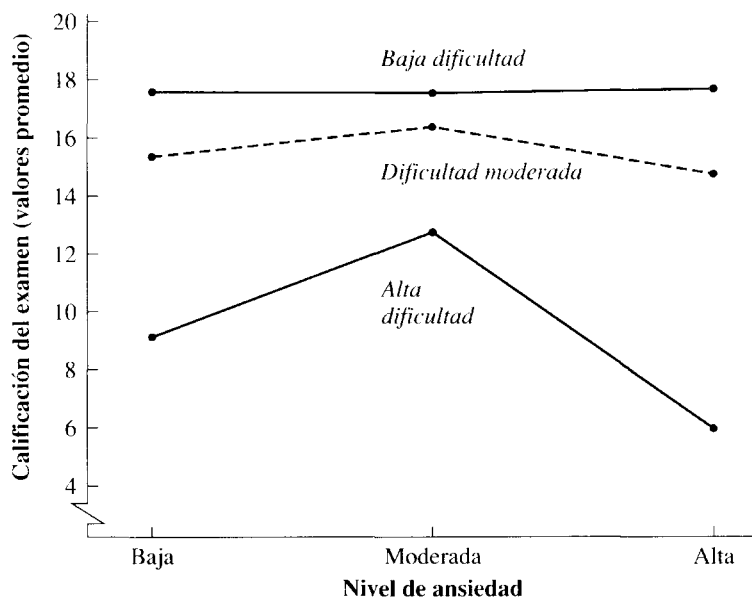


figura 17.7 Medias por celda del experimento sobre la dificultad del material y el nivel de ansiedad.

les de dificultad, fueron demasiado grandes como para ser atribuidas al azar. Sin embargo, más allá de esto, la interpretación no está clara, debido a la interacción entre las dos variables. En la figura 17.7 podemos apreciar que el efecto de los distintos niveles de ansiedad depende de la dificultad del material. En el nivel más bajo de dificultad, las diferencias en el nivel de ansiedad parecen no influir sobre las calificaciones. Sin embargo, en los otros dos niveles de dificultad, las diferencias en el nivel de ansiedad realmente afectan el rendimiento. La interacción es compleja porque tal parece que los niveles alto y bajo de ansiedad interfieren más con el rendimiento que la ansiedad moderada. Éste es un ejemplo de la curva en forma de U invertida que se presenta con frecuencia en psicología, en relación con los niveles de rendimiento y excitación.



COMPARACIONES MÚLTIPLES

En los tres ejemplos que hemos analizado, concluimos el análisis evaluando los valores F_{obt} . En la práctica, dicho análisis se extiende un poco más, pues se realizan múltiples comparaciones con las parejas de medias apropiadas. Por ejemplo, en el problema de práctica 17.2, había un valor F_{obt} significativo para la dificultad del material. Por lo general, el siguiente paso consiste en determinar qué niveles de dificultad son significativamente diferentes unos de otros. Desde el punto de vista conceptual, este tema es muy similar al que presentamos en el capítulo 16, cuando estudiamos las comparaciones múltiples junto con el análisis de varianza de un factor. Una de las principales diferencias consiste en que, en el análisis de varianza de dos factores, evaluamos con frecuencia parejas de medias por *renglón* o por *columna*, en lugar de parejas de medias por grupo. Un análisis más detallado de este tema está fuera de los objetivos de este libro.*

*El lector interesado en este tema debe consultar B. J. Winer *et al.*, *Statistical Principles in Experimental Design*, tercera edición, McGraw-Hill, Nueva York, 1991.



SUPUESTOS SUBYACENTES AL ANÁLISIS DE VARIANZA DE DOS FACTORES

Los supuestos fundamentales o subyacentes al análisis de varianza de dos factores son los mismos que en el análisis de varianza de un factor:

1. Las poblaciones de las cuales se extraen las muestras tienen una distribución normal.
2. Las varianzas de la población para cada celda son iguales. Ésta es la suposición de *homogeneidad de la varianza*.

Igual que el análisis de varianza de un factor, el análisis de varianza de dos factores es robusto con respecto a las violaciones de estos supuestos, siempre que las muestras sean del mismo tamaño.*

RESUMEN

En primer lugar, presentamos una exposición cualitativa del diseño del análisis de varianza de dos factores para grupos independientes. Igual que en el diseño de un factor, en el diseño de dos factores los sujetos son asignados aleatoriamente a las condiciones. Sin embargo, el diseño de dos factores nos permite investigar en un experimento dos variables independientes y la interacción entre ellas. El efecto de cualquiera de las variables independientes (promediadas sobre los niveles de la otra variable) se conoce como efecto principal. Se presenta una interacción cuando el efecto de una de las variables no es igual en todos los niveles de la otra variable.

El análisis de varianza de dos factores es muy similar al ANOVA de un factor. Sin embargo, en el ANOVA de dos factores, la suma total de cuadrados (SC_T) está dividida en cuatro componentes: la suma de cuadrados dentro de las celdas (SC_W), la suma de cuadrados por renglón (SC_R), la suma de cuadrados por columna (SC_C), y la suma de cuadrados renglón \times columna (SC_{RC}). Cuando estas sumas de cuadrados se dividen entre los grados de libertad correspondientes, forman las siguientes cuatro estimaciones de la varianza: la estimación de la varianza dentro de las celdas (s_W^2), la estimación de la varianza por renglón (s_R^2), la estimación de la varianza por columna (s_C^2) y la estimación de la varianza renglón \times columna (s_{RC}^2).

La estimación de la varianza dentro de las celdas (s_W^2) es el patrón con el cual se comparan las demás estimaciones de la varianza. Como quiera que todos los sujetos, dentro de cada celda, reciben el mismo nivel

de las variables A y B, la variabilidad dentro de las celdas no se puede deber a diferencias en el tratamiento. Más bien representa una medida de la variabilidad inherente de los datos y, por lo tanto, nos proporciona una estimación de la varianza poblacional de la hipótesis nula (σ^2). La estimación de la varianza por renglón (s_R^2) se basa en las diferencias entre las medias por renglón. Es una estimación de σ^2 más el efecto del factor A. Se utiliza para evaluar el efecto principal de la variable A. La estimación de la varianza por columna (s_C^2) se fundamenta en las diferencias que se presentan entre las medias por columna. Es una estimación de σ^2 más el efecto del factor B y se emplea para evaluar el efecto principal de la variable B. La estimación de la varianza renglón \times columna (s_{RC}^2) se basa en las diferencias existentes entre las medias por celda, más allá de lo que puede predecirse por los efectos individuales de las dos variables. Es una estimación de σ^2 más la interacción de A y B. Como tal, se utiliza para evaluar la interacción de las variables A y B.

Además de presentar la base conceptual del análisis de varianza de dos factores, desarrollamos ecuaciones para calcular cada una de las cuatro estimaciones de la varianza; presentamos varios ejemplos ilustrativos con el propósito de practicar la técnica del ANOVA de dos factores. Además, señalamos que las técnicas de comparación múltiple, similares a las utilizadas en el análisis de varianza de un factor, se emplean también en el análisis de varianza de dos factores. Por último, presentamos los supuestos subyacentes al ANOVA de dos factores.

*Algunos estadísticos requieren también que los datos tengan una escala de intervalo o de razón. Para un análisis de este tema, véanse las referencias que se encuentran en la nota al pie de la página 31, en el capítulo 2.

TÉRMINOS IMPORTANTES

Análisis de varianza de dos factores (página 395)	Estimación de la varianza renglón \times columna (s_{RC}^2) (página 405)	Grados de libertad renglón \times columna (gl_{RC}) (página 405)
Efecto de interacción (página 397)	Experimento factorial (página 396)	Suma de cuadrados dentro de las celdas (SC_W) (página 402)
Efecto principal (página 397)	Grados de libertad dentro de las celdas (gl_W) (página 402)	Suma de cuadrados por columna (SC_C) (página 404)
Estimación de la varianza dentro de las celdas (s_W^2) (página 400)	Grados de libertad por columna (gl_C) (página 404)	Suma de cuadrados por renglón (SC_R) (página 403)
Estimación de la varianza por columna (s_C^2) (página 404)	Grados de libertad por renglón (gl_R) (página 403)	Suma de cuadrados renglón \times columna (SC_{RC}) (página 405)
Estimación de la varianza por renglón (s_R^2) (página 402)		

PREGUNTAS Y PROBLEMAS

- Defina o identifique cada uno de los conceptos de la sección “términos importantes”.
- ¿Cuáles son las ventajas del ANOVA de dos factores, comparado con el ANOVA de un factor?
- ¿Qué es un experimento factorial?
- ¿Cuál es el efecto principal en el análisis de varianza de dos factores? ¿Qué es una interacción? ¿Es posible tener un efecto principal sin una interacción? ¿Y tener una interacción sin un efecto principal? Explique.
- En el ANOVA de dos factores, la suma total de cuadrados se divide en cuatro componentes. ¿Cuáles son éstos?
- ¿Por qué se utiliza la estimación de la varianza dentro de las celdas como el patrón con el cual se comparan las otras estimaciones de la varianza?
- Las cuatro estimaciones de la varianza (s_R^2 , s_C^2 , s_{RC}^2 y s_W^2) se conocen también como cuadrados medios. ¿Puede usted explicar por qué?
- Si el efecto de la variable A aumenta, ¿qué cree que pasará con las diferencias entre las medias por renglón? ¿Qué le pasará a s_R^2 ? Explique. Si no hay interacción, ¿qué sucederá con las diferencias entre las medias por columna?
- Si el efecto de la variable B aumenta, ¿qué pasará con las diferencias entre las medias por columna? ¿Qué le ocurrirá a s_C^2 ? Explique. Si no hay interacción, ¿qué sucederá con las diferencias entre las medias por renglón?
- ¿Cuáles son los supuestos subyacentes al diseño del ANOVA de dos factores con grupos independientes?
- Según cierta teoría, la repetición es útil para recordar y el aprendizaje de un material adicional puede interferir con los intentos de recordar el material ya aprendido. Un profesor que está interesado en el aprendizaje y la memoria humana realiza un experimento factorial de 2×3 a fin de investigar los efectos de estas dos variables sobre la memoria. El material por recordar consiste en una lista de parejas de sílabas carentes de sentido. Se presentan todas las parejas, una a la vez, durante 4 segundos, recorriendo en forma cíclica toda la lista antes de mostrar de nuevo la primera pareja. Existen tres niveles de repetición: el nivel 1, en el que cada pareja se muestra 4 veces; el nivel 2, en el que cada pareja se presenta 8 veces, y el nivel 3, en el cual cada pareja se muestra 12 veces. Después de presentar la lista el número correspondiente de veces y antes de verificar si se puede recordar, cada sujeto debe aprender cierto material. Este material de interferencia puede ser de dos tipos: el tipo 1 consiste en parejas de números y el tipo 2 consiste en parejas de sílabas sin sentido. Una vez que el material de interferencia ha sido mostrado, los sujetos presentan una prueba para ver si recuerdan la lista original de 16 parejas de sílabas sin sentido. Treinta y seis estudiantes participan como sujetos. Se les divide de manera aleatoria, de modo que haya seis por celda. Al fin se registraron las siguientes calificaciones; cada una corresponde a la cantidad de parejas de sílabas de la lista original que fueron correctamente recordadas.

Material de interferencia	Número de repeticiones					
	4 veces		8 veces		12 veces	
Parejas de números	10	11	16	12	16	14
	12	15	11	15	16	13
	14	10	13	14	15	16
Parejas de sílabas sin sentido	8	7	11	13	14	12
	4	5	9	10	16	15
	5	6	8	9	12	13

Fertilizante	Número de hojas de pasto por pulgada cuadrada					
	Red fescue		Kentucky blue		Green velvet	
Tipo 1	14	15	15	17	20	19
	16	17	12	18	15	22
	10		11		25	
Tipo 2	11	7	10	6	15	11
	11	8	8	13	18	10
	14		12		19	

- a. ¿Cuáles son las hipótesis nulas para este experimento?
- b. Al utilizar $\alpha = 0.05$, ¿qué puede usted concluir? Trace una gráfica de las medias por celda como ayuda para interpretar los resultados. cognitiva
12. Supongamos que usted ha aceptado un empleo como científico en jefe de una compañía agrícola líder en el mercado. Su primera tarea será recomendar el mejor tipo de pasto que puede cultivarse en la zona Pacífico Noroeste (de Estados Unidos) y el mejor fertilizante para él. A fin de proveer la base de datos que justifique la recomendación que se dispone a hacer, después de graduarse *summa cum laude* en estadística, usted decide llevar a cabo un experimento utilizando el diseño factorial con grupos independientes. Como existen tres tipos de pasto y hay dos fertilizantes en consideración, el experimento es factorial de 2×3 , donde la variable A es el tipo de fertilizante y la variable B es el tipo de pasto. En una estación de campo, usted reproduce el suelo y el clima del Pacífico Noroeste. A continuación, divide el suelo en 30 áreas iguales y elige al azar 5 de ellas para cada combinación de tratamientos. Posteriormente aplica el fertilizante y planta la semilla de pasto más adecuada para cada área. A partir de ese momento, las 30 áreas reciben el mismo tratamiento. Cuando el pasto ha crecido lo suficiente, usted determina el número de hojas de pasto por pulgada cuadrada que brotaron en cada área. La recomendación que usted está a punto de hacer estará basada en esta variable dependiente. Mientras más "denso" sea el pasto, tanto mejor será el resultado. De esta manera ha obtenido los siguientes datos:
- a. ¿Cuáles son las hipótesis nulas para este experimento?
- b. Utilizando $\alpha = 0.05$, ¿cuáles son las conclusiones de usted? Trace una gráfica de las medias por celda como ayuda para interpretar los resultados. I/O
13. Un investigador realiza un experimento para saber si un medicamento hipnótico llamado Drowson, que se anuncia como remedio contra el insomnio, provoca verdaderamente el sueño. Además, este investigador está interesado en determinar si se desarrolla alguna tolerancia al medicamento con el uso crónico del mismo. El diseño del experimento es factorial de 2×2 con grupos independientes. Una de las variables es la concentración empleada del medicamento Drowson. Existen dos niveles: 1) concentración cero (placebo) y 2) la dosis mínima recomendada por el fabricante. La otra variable se refiere al uso previo del Drowson. De nuevo, aquí se presentan dos niveles: 1) los sujetos que no lo han usado previamente y 2) los usuarios crónicos. Contamos con 16 individuos que padecen insomnio y que no han utilizado Drowson, los cuales son asignados de manera aleatoria a las dos condiciones de concentración, de tal manera que haya 8 sujetos en cada condición. Dieciséis usuarios crónicos de Drowson también son asignados al azar a las dos condiciones, a razón de 8 individuos en cada condición. Todos los sujetos ingieren su "dosis" prescrita durante tres noches consecutivas y se registra el tiempo que cada uno de ellos tarda en conciliar el sueño. Los datos que aparecen en la siguiente tabla son los tiempos promedio, en minutos, que cada sujeto tardó en conciliar el sueño, promediados durante los tres días:

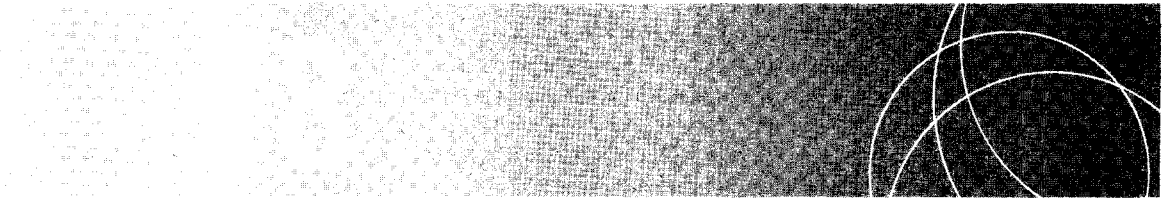
Uso previo	Concentración de Drowson			
	Placebo		Dosis mínima recomendada	
Sin uso previo	45	53	30	47
	48	58	33	35
	62	55	40	31
	70	64	50	39
Usuarios crónicos	47	68	52	46
	52	64	60	49
	55	58	58	50
	62	59	68	55

- ¿Cuáles son las hipótesis nulas para este experimento?
- Utilizando $\alpha = 0.05$, ¿qué puede usted concluir? Trace una gráfica de las medias por celda como ayuda para interpretar los resultados. clínica, salud

Si pretende resolver más problemas de práctica, relacionados con lo que se presentó en este capítulo y resueltos por medio de ecuaciones conceptuales, busque Chapter 17 en el material adicional del libro (en inglés), en: <http://www.thomsonlearning.com.mx>.

Si quiere resolver problemas para computadora sobre los temas aquí presentados, busque Chapter 17 en el material adicional (en inglés) del libro, en: <http://www.thomsonlearning.com.mx>.

Ji-cuadrada y otras pruebas no paramétricas



Introducción: Distinción entre las pruebas paramétricas y las no paramétricas
Ji-cuadrada (χ^2)

Experimentos con una variable

Un experimento: Preferencia por las diversas marcas de cerveza ligera

Prueba de independencia entre dos variables

Un experimento: Afiliación política y actitud

Supuestos subyacentes a χ^2

La prueba de rangos con signo y pares igualados de Wilcoxon

Un experimento: Promoción de actitudes más favorables hacia la conservación de la vida silvestre

Supuestos de la prueba de rangos con signo de Wilcoxon

La prueba *U* de Mann-Whitney

Un experimento: Efecto de una dieta alta en proteínas sobre el desarrollo intelectual

Rangos empatados

*Supuestos subyacentes de la prueba *U* de Mann-Whitney*

La prueba de Kruskal-Wallis

Un experimento: Evaluación de dos programas para adelgazar

Supuestos subyacentes de la prueba de Kruskal-Wallis

¿CUÁL ES LA VERDAD? Estadística e investigación social aplicada: ¿Útiles o "abusivas"?

Resumen

Términos importantes

Preguntas y problemas

Notas



INTRODUCCIÓN: DISTINCIÓN ENTRE LAS PRUEBAS PARAMÉTRICAS Y LAS NO PARAMÉTRICAS

Las pruebas de inferencia estadística se clasifican con frecuencia como paramétricas o no paramétricas. Como usted recordará, en nuestro análisis del capítulo 1 vimos que un parámetro es una característica de una población. Una prueba de inferencia paramétrica es aquella cuyo uso depende de manera considerable de las características de la población o parámetros. Las pruebas z , t y F son ejemplos de pruebas paramétricas. Por ejemplo, en la prueba z es necesario especificar la media y la desviación estándar de la población de la hipótesis nula y, además, los datos de la población deben tener una distribución normal para una N pequeña. La prueba t para una muestra debe cumplir los mismos requisitos, excepto que en este caso no especificamos σ . Las pruebas t para dos muestras o condiciones (t correlacionada o t independiente) requieren que los datos de la población tengan una distribución normal cuando las muestras sean pequeñas. La prueba t independiente también exige que las varianzas de la población sean iguales. El análisis de varianza tiene requisitos similares a los de la prueba t independiente.

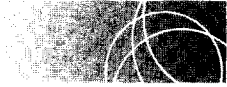
Aunque todas las pruebas de inferencia dependen en cierta medida de las características de la población, los requisitos de las pruebas no paramétricas son mínimos. Por ejemplo la prueba de los signos es una prueba no paramétrica. Para usar la prueba de los signos no es necesario conocer la media, la varianza o la forma de los datos poblacionales. En virtud de que las pruebas no paramétricas dependen poco del conocimiento de la distribución de la población, con frecuencia son conocidas como *pruebas de distribución libre*.

Como las pruebas de inferencia no paramétrica tienen menos requisitos o supuestos relacionados con las características de la población, surge la duda de si no deberíamos utilizarlas todo el tiempo, olvidándonos de las pruebas paramétricas. La respuesta tiene dos partes. En primer lugar, muchas de las pruebas de inferencia paramétrica son robustas con respecto a la violación de los supuestos subyacentes. Usted recordará que una prueba es robusta si la violación de los supuestos no perturba de manera grave la distribución muestral del estadístico que se está empleando. Así, la prueba t es robusta en cuanto a la violación de la normalidad en la población. Aunque, desde un punto de vista teórico, se exige la normalidad en la población para muestras pequeñas, desde un punto de vista empírico se tiene que, a menos que la desviación con respecto a la normalidad sea sustancial, la distribución muestral de t sigue siendo esencialmente la misma. Por lo tanto, la prueba t se puede emplear aun con datos que violen los supuestos de normalidad.

Las principales razones para preferir las pruebas paramétricas sobre las no paramétricas son que, en general, las primeras son más potentes y versátiles que las segundas. Vimos un ejemplo de esta superioridad de las pruebas paramétricas, en términos de potencia, al comparar la prueba t con la de los signos para grupos correlacionados. El diseño factorial analizado en el capítulo 17 proporciona un buen ejemplo de la versatilidad de la estadística paramétrica. Con este diseño podemos poner a prueba dos, tres, cuatro o más variables y las interacciones entre ellas. No existe una técnica estadística no paramétrica que pueda compararse.

Por regla general, los investigadores emplean las pruebas paramétricas siempre que sea posible. Sin embargo, cuando existe una violación grave de algún supuesto relacionado con una prueba paramétrica, o si el investigador piensa que la escala de los datos hace inadecuado el uso de la prueba paramétrica en cuestión, se utiliza una prueba de inferencia no paramétrica. Ya hemos presentado una prueba no paramétrica: la del signo. En las secciones posteriores de este capítulo presentaremos otras cuatro: la ji-cuadrada, la prueba de rangos con signo y pares igualados de Wilcoxon, la prueba U de Mann-Whitney y la de Kruskal-Wallis.*

*Si bien hemos analizado sólo algunas pruebas no paramétricas, existen muchas más. El lector interesado en este punto debe consultar S. Siegel y N. Castellan, Jr., *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill, Nueva York, 1988, o W. Daniel, *Applied Nonparametric Statistics*, segunda edición, PWS-Kent, Boston, 1990.



Ji-CUADRADA (χ^2)

Experimentos con una variable

Hasta este momento, hemos presentado pruebas de inferencia que han sido utilizadas principalmente con datos ordinales, de intervalo o de razón. ¿Pero qué se puede decir de los datos nominales? Los experimentos relacionados con datos nominales son muy frecuentes, en particular en la psicología social. Usted recordará que con este tipo de datos, las observaciones se agrupan en varias categorías discretas, mutuamente excluyentes, y se puede contar la frecuencia con la cual se presentan en cada categoría. La prueba de inferencia que se emplea más a menudo con los datos nominales es una prueba no paramétrica que se conoce como *ji-cuadrada* (χ^2). De acuerdo con el procedimiento que hemos venido aplicando a lo largo de todo este texto, iniciaremos nuestro análisis de la ji-cuadrada con un experimento.

un experimento

Preferencia por las diversas marcas de cerveza ligera

Supongamos que usted quiere determinar si los consumidores de cerveza que viven en Puget Sound difieren con respecto a sus preferencias por las diferentes marcas de cerveza ligera. Para eso decide realizar un experimento en el cual elige al azar a 150 consumidores de cerveza y los invita a probar las tres marcas principales de esa bebida. Suponga que se han observado todas las precauciones de un buen diseño experimental, como no mostrar los nombres de las marcas a los sujetos del experimento y otros cuidados similares. Los datos resultantes se presentan en la tabla 18.1.

tabla 18.1 Preferencia por tres marcas de cerveza ligera

	Marca A	Marca B	Marca C	Total
1	45	40	65	150

SOLUCIÓN

Los datos de cada celda son el número o la frecuencia de los sujetos que corresponde a esa celda. Así, 45 sujetos prefirieron la marca A (celda 1), 40 prefirieron la marca B (celda 2) y 65 prefirieron la marca C (celda 3). ¿Podemos concluir, a partir de estos datos, que existe alguna diferencia en las preferencias de la población? La hipótesis nula para este experimento establece que no existe diferencia alguna en la preferencia de la población por las diversas marcas de cerveza ligera. Más específicamente, establece que, en la población, la proporción de individuos que prefiere la marca A es igual a la de quienes prefieren la marca B y ésta es igual a la de los que prefieren la marca C. Tomando la tabla como referencia, resulta claro que el número de individuos de la muestra que prefirió cada marca es diferente. Sin embargo, esto no implica necesariamente que exista una diferencia en la población. ¿No sería posible que estos datos se debieran sólo al muestreo aleatorio realizado en una población de consumidores de cerveza en la cual la proporción de individuos que prefiere cada marca es la misma? Por supuesto, la respuesta es "sí". La ji-cuadrada nos permite evaluar esta posibilidad.

Cálculo de χ_{obt}^2 Para calcular χ_{obt}^2 , debemos determinar primero la frecuencia que esperaríamos obtener en cada celda si se realizara un muestreo aleatorio de la población de la hipótesis nula. Estos datos reciben el nombre de frecuencias esperadas y se simbolizan como f_e . Las frecuencias que en realidad se obtienen en el experimento se conocen como *frecuencias observadas* y se simbolizan como f_o . Así,

f_e = frecuencia esperada bajo el supuesto de que se realice un muestreo aleatorio de la población de la hipótesis nula

f_o = frecuencia observada en la muestra

Debe quedar claro que mientras más se acerque la frecuencia observada de cada celda a la frecuencia esperada de dicha celda, tanto más razonable será H_0 . Por otra parte, mientras mayor sea la diferencia entre f_o y f_e tanto más razonable se volverá H_1 .

Después de determinar f_e para cada celda, obtenemos la diferencia entre f_o y f_e , la elevamos al cuadrado y la dividimos entre f_e . En forma simbólica, calculamos $(f_o - f_e)^2/f_e$ para cada celda. Por último, sumamos los valores resultantes para cada celda. En forma de ecuación,

$$\chi_{\text{obt}}^2 = \sum \frac{(f_o - f_e)^2}{f_e} \quad \text{ecuación para calcular } \chi^2$$

donde f_o = frecuencia observada en la celda
 f_e = frecuencia esperada en la celda, y la Σ se realiza sobre todas las celdas

Esta ecuación muestra que χ^2 es básicamente una medida de la diferencia entre las frecuencias observadas y las frecuencias esperadas.

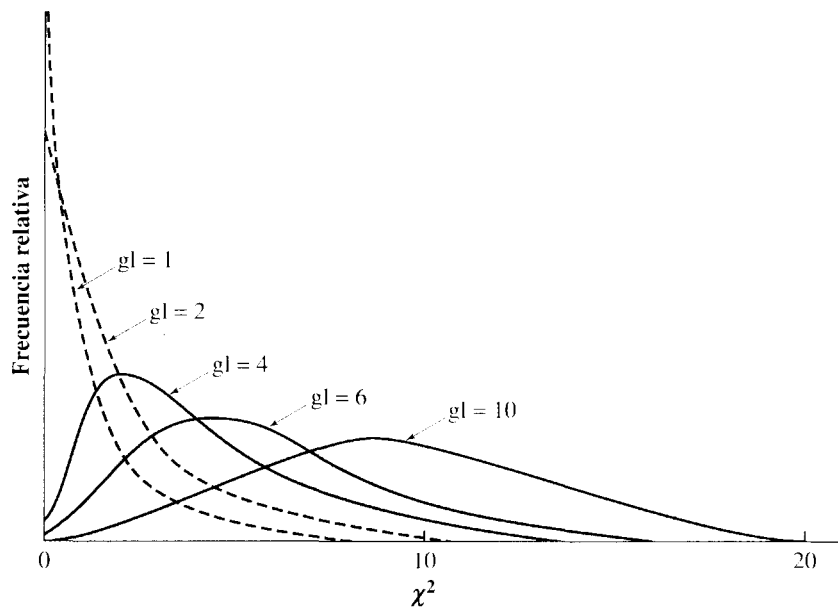
Para calcular el valor de χ_{obt}^2 en este experimento, debemos determinar f_e para cada celda. Los valores de f_o se presentan en la tabla. Si la hipótesis nula es verdadera, entonces, dentro de la población, la proporción de consumidores de cerveza que prefiere la marca A debe ser igual a la proporción que prefiere la marca B, la cual, a su vez, tendrá que ser igual a la proporción que prefiere la marca C. Esto significa que una tercera parte de la población deberá preferir la marca A, otra tercera parte la marca B y la tercera parte restante preferirá la marca C. Por lo tanto, si la hipótesis nula es cierta, esperamos que un tercio de los individuos de la población, y por lo tanto de la muestra, prefiera la marca A, un tercio prefiera la marca B y un tercio prefiera la marca C. Como la muestra tiene 150 sujetos, f_e para cada celda = $\frac{1}{3}(150) = 50$. Reproducimos aquí la tabla de datos y escribimos los valores de f_e entre paréntesis:

Marca A	Marca B	Marca C	Total
45	40	65	150
(50)	(50)	(50)	(150)

Ahora que hemos determinado el valor de f_e para cada celda, podemos calcular χ_{obt}^2 . Basta sumar los valores de $(f_o - f_e)^2/f_e$ para cada celda. Así pues,

$$\begin{aligned} \chi_{\text{obt}}^2 &= \sum \frac{(f_o - f_e)^2}{f_e} \\ &= \frac{(45 - 50)^2}{50} + \frac{(40 - 50)^2}{50} + \frac{(65 - 50)^2}{50} \\ &= 0.50 + 2.00 + 4.50 = 7.00 \end{aligned}$$

Evaluación de χ_{obt}^2 La distribución muestral teórica de χ^2 se presenta en la figura 18.1. Esta distribución consta de una familia de curvas que, como en el caso de la distribución t , varían según los grados de libertad. Para los grados de libertad menores, las curvas están sesgadas positivamente. Los grados de libertad quedan determinados por el número de datos f_o que varían libremente. En nuestro experimento, dos de los datos de f_o pueden variar con libertad. Una vez que conocemos dos de los datos f_o , el tercer dato f_o queda fijo, en virtud de que la suma de los tres datos f_o tiene que ser igual a N . Por lo tanto, $gl = 2$. En general, para los experimentos con una sola variable existen $k - 1$ gra-



Tomado de *Design and Analysis of Experiments in Psychology and Education*, por E. F. Lindquist. Derechos reservados © 1953 Houghton Mifflin Company. Reproducido con autorización.

figura 18.1 Distribución de χ^2 para varios grados de libertad.

dos de libertad, donde k es igual al número de grupos o categorías. Cuando consideremos el uso de χ^2 con tablas de contingencia, estudiaremos otra ecuación para establecer los grados de libertad. Esto lo estudiaremos cuando se presente el tema.

La tabla H del apéndice D proporciona los valores críticos de χ^2 para diversos niveles alfa. Como quiera que χ^2 es básicamente una medida de la discrepancia total entre f_o y f_e , de esto se deduce que mientras más grande sea la discrepancia entre las frecuencias observadas y las esperadas, tanto más grande será el valor de χ_{obt}^2 . Por lo tanto, mientras mayor sea el valor de χ_{obt}^2 , tanto menos razonable será la hipótesis nula. Como en el caso de las pruebas t y F si χ_{obt}^2 queda dentro de la región crítica, entonces rechazaremos la hipótesis nula. La regla de decisión establece lo siguiente:

$$\text{Si } \chi_{\text{obt}}^2 \geq \chi_{\text{crit}}^2 \text{ rechazamos } H_0.$$

Debemos observar que en el cálculo de χ_{obt}^2 no importa si f_o es mayor o menor que f_e . La diferencia se eleva al cuadrado, se divide entre f_e y se suma a las demás celdas para obtener χ_{obt}^2 . Como la dirección de la diferencia no es importante, χ^2 es una prueba no direccional.* Además, como cada diferencia agrega algo al valor de χ^2 , la región crítica para el rechazo está siempre en la cola derecha de la distribución χ^2 .

En este experimento hemos determinado que $\chi_{\text{obt}}^2 = 7.00$. Para evaluarla, debemos determinar χ_{crit}^2 para ver si χ_{obt}^2 queda dentro de la región crítica para el rechazo de H_0 . Según la tabla H, con $gl = 2$ y $\alpha = 0.05$,

$$\chi_{\text{crit}}^2 = 5.991$$

*Por favor, véase la nota 18.1 para conocer un análisis más detallado de este punto.

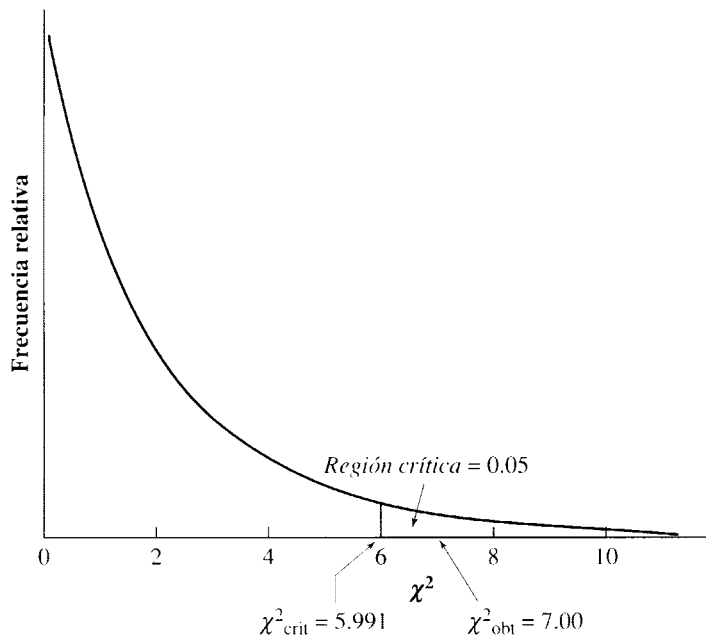


figura 18.2 Evaluación de χ^2_{obt} para el problema sobre los consumidores de cerveza ligera, con $gl = 2$ y $\alpha = 0.05$.

La figura 18.2 muestra la distribución de χ^2 con $gl = 2$ y la región crítica para $\alpha = 0.05$. Debido a que $\chi^2_{\text{obt}} > 5.991$ queda dentro de la región crítica, rechazamos H_0 . Existe una diferencia en la población, en cuanto a la preferencia demostrada por las tres marcas de cerveza puestas a prueba. Al parecer, la marca C fue la favorita.

Resolvamos otro problema para practicar.



Un politólogo cree que la composición étnica de la ciudad donde él vive ha cambiado durante los últimos años. Las cifras más recientes (recopiladas hace unos cuantos años) muestran que los habitantes de dicha ciudad presentaban la siguiente composición étnica: 53% noruegos, 32% suecos, 8% irlandeses, 5% hispanos y 2% italianos (observe que no hemos incluido las nacionalidades cuyos porcentajes son menores que 2%). Para poner a prueba su idea, el científico social obtiene una muestra aleatoria de 750 habitantes. Los resultados que obtuvo se presentan en la siguiente tabla:

	Noruegos	Suecos	Irlandeses	Hispanos	Italianos	Total
¹	399	² 193	³ 63	⁴ 82	⁵ 13	750

- a. ¿Cuál es la hipótesis nula?
- b. ¿Cuál es la conclusión de usted? Utilice $\alpha = 0.05$.

SOLUCIÓN

- a. Hipótesis nula: La composición étnica de la ciudad no ha cambiado. Por lo tanto, la muestra de 750 individuos es una muestra aleatoria extraída de una población compuesta por 53% de noruegos, 32% de suecos, 8% de irlandeses, 5% de hispanos y 2% de italianos.
- b. Conclusión, utilizando $\alpha = 0.05$:

PASO 1: Cálculo del estadístico apropiado. El estadístico apropiado es χ^2_{obt} . Los cálculos se presentan a continuación:

Celda número	f_o	f_e	$\frac{(f_o - f_e)^2}{f_e}$
1	399	$0.53(750) = 397.5$	$\frac{(399 - 397.5)^2}{397.5} = 0.006$
2	193	$0.32(750) = 240.0$	$\frac{(193 - 240)^2}{240} = 9.204$
3	63	$0.08(750) = 60.0$	$\frac{(63 - 60)^2}{60} = 0.150$
4	82	$0.05(750) = 37.5$	$\frac{(82 - 37.5)^2}{37.5} = 52.807$
5	13	$0.02(750) = 15.0$	$\frac{(13 - 15)^2}{15} = 0.267$
			$\chi^2_{\text{obt}} = \sum \frac{(f_o - f_e)^2}{f_e} = 62.434$
			$= 62.43$

PASO 2: Evaluación del estadístico. Grados de libertad = $5 - 1 = 4$. Con $gl = 4$ y $\alpha = 0.05$, según la tabla H,

$$\chi^2_{\text{crit}} = 9.488$$

En virtud de que $\chi^2_{\text{obt}} > 9.488$, rechazamos H_0 . La composición étnica de la ciudad parece haber cambiado efectivamente. Ha habido un incremento en la población de hispanos y un decremento en la de suecos.

Prueba de independencia entre dos variables

Una de las principales aplicaciones de χ^2 consiste en determinar si dos variables categóricas son independientes o están relacionadas entre sí. Para ilustrar este punto, consideremos el siguiente ejemplo.

un experimento

Afiliación política y actitud

Supongamos que en la legislatura local está en estudio un proyecto de ley cuyo propósito es reducir la edad legal a fin de permitir el consumo de bebidas alcohólicas a las personas de 18 años. Un politólogo que vive en esa ciudad quiere determinar si existe alguna relación entre la afiliación política y la actitud frente al proyecto de ley. Para averiguarlo, este científico envía cartas a dos muestras aleatorias, una de ellas formada por 200 miembros del Partido Republicano y la otra por 200 miembros del Partido Demócrata. En la misiva, el politólogo explica su interés por el proyecto de ley y pregunta a los destinatarios si están a favor, indecisos, o en contra de esa iniciativa de ley. Les asegura la plena confidencialidad de sus respuestas. En la carta incluye un sobre con porte pagado para facilitar la respuesta. El investigador recibe respuesta de los 400, republicanos y demócratas. Los resultados obtenidos se presentan en la tabla 18.2.

tabla 18.2 Datos sobre afiliación política y actitud

	Actitud			Marginal por renglón
	<i>A favor</i>	<i>Indecisos</i>	<i>En contra</i>	
Republicanos	68	22	110	200
Demócratas	92	18	90	200
Marginal por columna	160	40	200	400

Los datos de cada celda son la frecuencia de los sujetos correspondientes. Por ejemplo, en el caso de los republicanos, 68 están a favor del proyecto, 22 están indecisos y 110 están en contra. En el caso de los demócratas, 92 están a favor del proyecto, 18 están indecisos y 90 están en contra. Este tipo de arreglo se conoce como *tabla de contingencia*.

definición ■ Una **tabla de contingencia** se compone de dos vías o entradas y muestra la relación contingente entre dos variables, cuando éstas han sido clasificadas en categorías mutuamente excluyentes y cuando los datos de cada celda son frecuencias.

Observe que al construir una tabla de contingencia, es esencial que las categorías sean mutuamente excluyentes. De esta manera, si un dato es apropiado para una de las celdas, las categorías deben ser de tal naturaleza que no resulte apropiado asentar ese dato en ninguna otra celda.

Esta tabla de contingencia incluye los datos referentes a la relación contingente entre la afiliación política y la actitud hacia el proyecto de ley. La hipótesis nula establece que no existe una relación contingente entre las variables en la población. Para este ejemplo, H_0 establece que, en la población, la actitud hacia el proyecto y la afiliación política son independientes. Si esto es cierto, entonces los republicanos y los demócratas en la población deberán tener la misma proporción de personas “a favor”, “indecisos” y “en contra” del proyecto. Resulta claro que las frecuencias asentadas en estas tres columnas en la tabla de contingencia son diferentes para los republicanos y para los demócratas. La hipótesis nula establece que estas frecuencias son simplemente el resultado de un muestreo aleatorio realizado en una población en la cual la proporción de republicanos es igual a

la proporción de demócratas en cada una de las categorías consideradas. La hipótesis alternativa establece que los republicanos y los demócratas tienen actitudes diferentes en cuanto al proyecto de ley. Si en verdad es así, entonces las proporciones deberán ser diferentes en la población.

Cálculo de χ_{obt}^2 Para poner a prueba la hipótesis nula, debemos calcular χ_{obt}^2 y compararla con χ_{crit}^2 . En los experimentos con dos variables, la parte más difícil del proceso consiste en determinar f_e para cada celda. Como ya hemos visto, la hipótesis nula establece que, dentro de la población, la proporción de republicanos en cada categoría es igual a la proporción de demócratas. Si conociéramos sus respectivas proporciones, bastaría multiplicarlas por el número de republicanos o de demócratas incluidos en la muestra y así podríamos determinar f_e para cada celda. Por ejemplo, supongamos que si H_0 es cierta, la proporción de republicanos que está en contra del proyecto en la población es igual a 0.50. Para determinar f_e para cada celda, basta multiplicar 0.50 por el número de republicanos incluidos en la muestra. Así, para la celda “republicanos en contra”, f_e sería igual a $0.50(200) = 100$.

En virtud de que no conocemos las proporciones reales en la población, las estimamos a partir de la muestra. En nuestro experimento, 160 de los 400 republicanos y demócratas estuvieron a favor del proyecto, 40 de 400 estuvieron indecisos, y 200 de los 400 se pronunciaron en contra. Puesto que la hipótesis nula supone que la afiliación política y la actitud son independientes entre sí, utilizamos estas proporciones muestrales como estimaciones de las proporciones poblacionales de la hipótesis nula. Entonces, podemos utilizar estas estimaciones para calcular las frecuencias esperadas. Nuestras estimaciones de las proporciones poblacionales de la hipótesis nula son las siguientes:

$$\text{Proporción estimada de la población de la } H_0 \text{ a favor del proyecto} = \frac{\text{Número de sujetos a favor del proyecto}}{\text{Número total de sujetos}} = \frac{160}{400}$$

$$\text{Proporción estimada de la población de la } H_0 \text{ indecisa} = \frac{\text{Número de sujetos a favor del proyecto}}{\text{Número total de sujetos}} = \frac{40}{400}$$

$$\text{Proporción estimada de la población de la } H_0 \text{ en contra del proyecto} = \frac{\text{Número de sujetos en contra del proyecto}}{\text{Número de sujetos indecisos}} = \frac{200}{400}$$

Utilizando estas estimaciones para calcular las frecuencias esperadas, obtenemos los siguientes valores para f_e :

Para la celda republicanos a favor (celda 1 en la tabla de la página 434):

$$f_e = \left(\begin{array}{l} \text{Proporción estimada de} \\ \text{la población de la } H_0 \\ \text{a favor del proyecto} \end{array} \right) \left(\begin{array}{l} \text{Número total} \\ \text{de republicanos} \end{array} \right) = \frac{160}{400}(200) = 80$$

Para la celda republicanos indecisos (celda 2):

$$f_e = \left(\begin{array}{l} \text{Proporción estimada de la} \\ \text{población de la } H_0 \text{ indecisa} \end{array} \right) \left(\begin{array}{l} \text{Número total} \\ \text{de republicanos} \end{array} \right) = \frac{40}{400}(200) = 20$$

Para la celda republicanos en contra (celda 3):

$$f_e = \left(\begin{array}{l} \text{Proporción estimada de} \\ \text{la población de la } H_0 \\ \text{en contra del proyecto} \end{array} \right) \left(\begin{array}{l} \text{Número total} \\ \text{de republicanos} \end{array} \right) = \frac{200}{400}(200) = 100$$

Para la celda de demócratas a favor (celda 4):

$$f_e = \left(\begin{array}{l} \text{Proporción estimada de} \\ \text{la población de la } H_0 \\ \text{a favor del proyecto} \end{array} \right) \left(\begin{array}{l} \text{Número total} \\ \text{de demócratas} \end{array} \right) = \frac{160}{400}(200) = 80$$

Para la celda de demócratas indecisos (celda 5):

$$f_e = \left(\begin{array}{l} \text{Proporción estimada de la} \\ \text{población de la } H_0 \text{ indecisa} \end{array} \right) \left(\begin{array}{l} \text{Número total} \\ \text{de demócratas} \end{array} \right) = \frac{40}{400}(200) = 20$$

Para la celda de demócratas en contra (celda 6):

$$f_e = \left(\begin{array}{l} \text{Proporción estimada de} \\ \text{la población de la } H_0 \\ \text{en contra del proyecto} \end{array} \right) \left(\begin{array}{l} \text{Número total} \\ \text{de demócratas} \end{array} \right) = \frac{200}{400}(200) = 100$$

Para comodidad del lector, hemos reproducido de nuevo la tabla de contingencia 2×3 y agregamos los valores de f_e entre paréntesis, en las celdas apropiadas:

	Actitud			Marginal por renglón
	A favor	Indecisos	En contra	
Republicanos	1 68 (80)	2 22 (20)	3 110 (100)	200
Demócratas	4 92 (80)	5 18 (20)	6 90 (100)	200
Marginal por columna	160	40	200	400

Podemos hallar los mismos valores de f_e en forma directa, multiplicando los “marginales” para esa celda y dividiendo entre N . Los marginales son los totales por renglón y por columna que aparecen fuera de la tabla. Por ejemplo, los marginales para la celda 1 son 160 (total por columna) y 200 (total por renglón). Utilizaremos este método a fin de establecer f_e para cada celda. Si multiplicamos los marginales y los dividimos entre N , obtenemos:

$$f_e(\text{celda 1}) = \frac{160(200)}{400} = 80$$

$$f_e(\text{celda 2}) = \frac{40(200)}{400} = 20$$

$$f_e(\text{celda 3}) = \frac{200(200)}{400} = 100$$

$$f_e(\text{celda 4}) = \frac{160(200)}{400} = 80$$

$$f_e(\text{celda 5}) = \frac{40(200)}{400} = 20$$

$$f_e(\text{celda 6}) = \frac{200(200)}{400} = 100$$

Por supuesto, éstos son los mismos valores que obtuvimos anteriormente. Aunque el uso de los marginales no proporciona mucha información del porqué f_c debe tener ese valor, desde un punto de vista práctico es la mejor forma de calcular f_c para las distintas celdas. Conviene observar que una buena forma de verificar que los cálculos de f_c son correctos consiste en ver si los totales por renglón y por columna de f_c son iguales a los marginales por renglón y por columna.

Una vez que se ha determinado f_c para cada celda, el siguiente paso consiste en calcular χ_{obt}^2 . Como antes, para hacer esto sumamos $(f_o - f_c)^2/f_c$ para cada celda. Así pues, para nuestro experimento,

$$\begin{aligned} \chi_{\text{obt}}^2 &= \sum \frac{(f_o - f_c)^2}{f_c} \\ &= \frac{(68 - 80)^2}{80} + \frac{(22 - 20)^2}{20} + \frac{(110 - 100)^2}{100} + \frac{(92 - 80)^2}{80} \\ &\quad + \frac{(18 - 20)^2}{20} + \frac{(90 - 100)^2}{100} \\ &= 1.80 + 0.20 + 1.00 + 1.80 + 0.20 + 1.00 = 6.00 \end{aligned}$$

Evaluación de χ_{obt}^2 Para realizar la evaluación de χ_{obt}^2 , tenemos que compararla con χ_{crit}^2 para los gl apropiados. Como dijimos antes, el número de grados de libertad es igual al número de datos f_o que están en condiciones de variar libremente, mientras los totales se mantienen constantes. En el experimento con dos variables, debemos mantener los marginales por columna y por renglón en los mismos valores. Así pues, el número de grados de libertad para los experimentos que implican una contingencia entre dos variables es igual al número de datos f_o que pueden variar libremente, a la vez que se mantienen constantes los marginales por columna y por renglón. En el caso de una tabla de 2×3 , tenemos dos grados de libertad únicamente. Sólo pueden variar con libertad dos datos f_o , mientras que los otros datos f_o y f_c se mantienen fijos.

Como ilustración de esto, consideremos la siguiente tabla de 2×3 :

1	68	2	22	3		200
4		5		6		200
	160		40		200	400

Si escribimos en la tabla dos datos f_o , todos los demás datos f_o quedan determinados en forma completa, siempre que los marginales se mantengan con los mismos valores. Por ejemplo, en la tabla hemos llenado los datos f_o para las celdas 1 y 2. Observe que el valor de todos los demás datos queda fijo en cuanto dos datos f_o son conocidos; por ejemplo, el dato f_o para la celda 3 debe ser 110 [$200 - (68 + 22)$].

También existe una ecuación a fin de calcular los gl para las tablas de contingencia. Dicha ecuación es:

$$\text{gl} = (r - 1)(c - 1)$$

donde r = número de renglones
 c = número de columnas

Al aplicar la ecuación a nuestro experimento, obtenemos:

$$\text{gl} = (r - 1)(c - 1) = (2 - 1)(3 - 1) = 2$$

La prueba χ^2 no se limita a las tablas de 2×3 . Se puede utilizar con tablas de contingencia que tengan cualquier número de renglones y columnas. Esta ecuación es totalmente

general y se aplica a todas las tablas de contingencia. Así, si realizáramos un experimento con dos variables cuya tabla tuviera cuatro renglones y seis columnas, entonces $gl = (r - 1)(c - 1) = (4 - 1)(6 - 1) = 15$.

Considerando nuevamente la evaluación de nuestro experimento, supongamos que $\alpha = 0.05$. Con $gl = 2$ y $\alpha = 0.05$, según la tabla H,

$$\chi_{\text{crit}}^2 = 5.991$$

En vista de que $\chi_{\text{obt}}^2 > 5.991$, rechazamos H_0 . La afiliación política está relacionada con la actitud hacia el proyecto de ley. Los demócratas parecen tener una disposición más favorable hacia el proyecto que los republicanos. La solución completa se presenta en la tabla 18.3.

Tratemos de resolver un problema para practicar.

tabla 18.3 Solución al problema sobre la afiliación política y la actitud

- a. Hipótesis nula: La afiliación política y la actitud hacia el proyecto son independientes entre sí. La frecuencia obtenida en cada celda se debe a un muestreo aleatorio realizado en una población donde la proporción de republicanos y demócratas que están a favor, indecisos, o en contra de dicho proyecto, son iguales.
- b. Conclusión, utilizando $\alpha = 0.05$:

PASO 1: Cálculo del estadístico apropiado. El estadístico apropiado es χ_{obt}^2 . Los datos se presentan en la página 432. Los cálculos se muestran aquí.

PASO 2: Evaluación del estadístico. Grados de libertad = $(r - 1)(c - 1) = (2 - 1)(3 - 1) = 2$. Con $gl = 2$ y $\alpha = 0.05$, según la tabla H,

$$\chi_{\text{crit}}^2 = 5.991$$

En vista de que $\chi_{\text{obt}}^2 > 5.991$, rechazamos H_0 . La afiliación política y la actitud hacia el proyecto de ley están relacionadas entre sí. Tal parece que los demócratas están más a favor del decreto que los republicanos.

Celda número	f_o	f_e	$\frac{(f_o - f_e)^2}{f_e}$
1	68	$\frac{160(200)}{400} = 80$	$\frac{(68 - 80)^2}{80} = 1.80$
2	22	$\frac{40(200)}{400} = 20$	$\frac{(22 - 20)^2}{20} = 0.20$
3	110	$\frac{200(200)}{400} = 100$	$\frac{(110 - 100)^2}{100} = 1.00$
4	92	$\frac{160(200)}{400} = 80$	$\frac{(92 - 80)^2}{80} = 1.80$
5	18	$\frac{40(200)}{400} = 20$	$\frac{(18 - 20)^2}{20} = 0.20$
6	90	$\frac{200(200)}{400} = 100$	$\frac{(90 - 100)^2}{100} = 1.00$
			$\chi_{\text{obt}}^2 = \sum \frac{(f_o - f_e)^2}{f_e} = 6.00$

Una universidad estudia la posibilidad de implantar uno de los tres sistemas de calificaciones siguientes: 1) todas las calificaciones son de aprobado-reprobado; 2) todas las calificaciones están en el sistema 4.0, y 3) 90% de las calificaciones están en el sistema 4.0 y 10% son de aprobado-reprobado. Se realiza una encuesta para determinar si existe alguna relación entre el área de la licenciatura que estudia cada alumno y la preferencia que manifiesta por algún sistema de calificación en particular. Entonces se elige una muestra aleatoria constituida por 200 estudiantes del área de ingeniería, 200 de artes y ciencias y 100 de bellas artes. Se pregunta a cada alumno cuál de los tres sistemas de calificación prefiere. Los resultados se presentan en la siguiente tabla de contingencia de 3×3 :

	Sistema de calificación			Marginal por renglón
	Aprobado-reprobado	4.0 y aprobado-reprobado	4.0	
Bellas artes	¹ 26	² 55	³ 19	100
Artes y ciencias	⁴ 24	⁵ 118	⁶ 58	200
Ingeniería	⁷ 20	⁸ 112	⁹ 68	200
Marginal por columna	70	285	145	500

- a. ¿Cuál es la hipótesis nula?
- b. ¿Cuál es la conclusión de usted? Utilice $\alpha = 0.05$.

SOLUCIÓN

- a. Hipótesis nula: El área de estudio de los estudiantes y su preferencia por algún sistema de calificación son independientes entre sí. La frecuencia obtenida en cada celda se debe al muestreo aleatorio realizado en una población donde las proporciones de los estudiantes de bellas artes, artes y ciencias, e ingeniería, que prefieren cada sistema de calificación son iguales.
- b. Conclusión, utilizando $\alpha = 0.05$:

PASO 1: Cálculo del estadístico apropiado. Los datos se presentan en la siguiente tabla. El estadístico apropiado es χ^2_{obt} . Antes de calcular χ^2_{obt} , debemos determinar primero f_e , para cada celda. Los valores de f_e fueron encontrados a partir de los marginales.

Celda número	f_o	f_e	$\frac{(f_o - f_e)^2}{f_e}$
1	26	$\frac{70(100)}{500} = 14$	$\frac{(26 - 14)^2}{14} = 10.286$
2	55	$\frac{285(100)}{500} = 57$	$\frac{(55 - 57)^2}{57} = 0.070$

(continúa)

Celda número	f_o	f_e	$\frac{(f_o - f_e)^2}{f_e}$
3	19	$\frac{145(100)}{500} = 29$	$\frac{(19 - 29)^2}{29} = 3.448$
4	24	$\frac{70(200)}{500} = 28$	$\frac{(24 - 28)^2}{28} = 0.571$
5	118	$\frac{285(200)}{500} = 114$	$\frac{(118 - 114)^2}{114} = 0.140$
6	58	$\frac{145(200)}{500} = 58$	$\frac{(58 - 58)^2}{58} = 0.000$
7	20	$\frac{70(200)}{500} = 28$	$\frac{(20 - 28)^2}{28} = 2.286$
8	112	$\frac{285(200)}{500} = 114$	$\frac{(112 - 114)^2}{114} = 0.035$
9	68	$\frac{145(200)}{500} = 58$	$\frac{(68 - 58)^2}{58} = 1.724$
			$\chi^2_{\text{obt}} = \sum \frac{(f_o - f_e)^2}{f_e} = 18.560$
			= 18.56

PASO 2: Evaluación del estadístico. Grados de libertad = $(r - 1)(c - 1) = (3 - 1)(3 - 1) = 4$. Con $gl = 4$ y $\alpha = 0.05$, según la tabla H:

$$\chi^2_{\text{crit}} = 9.488$$

Como $\chi^2_{\text{obt}} > 9.488$, rechazamos H_0 . El área de estudio elegida y la preferencia por algún sistema de calificación están relacionadas entre sí.

Para tratar de determinar cuáles fueron las diferencias entre las preferencias de los grupos (considerando que el número de sujetos difiere notablemente en el caso de los estudiantes de bellas artes), es necesario convertir los datos de las frecuencias en proporciones. Estas proporciones se presentan en la tabla 18.4.

tabla 18.4 Preferencias por los sistemas de calificación, expresadas como proporciones

	Aprobado-reprobado	4.0 y aprobado-reprobado	4.0
Bellas artes	0.26	0.55	0.19
Artes y ciencias	0.12	0.59	0.29
Ingeniería	0.10	0.56	0.34

Al examinar esta tabla, tal parece que las diferencias entre los grupos de estudiantes están divididas entre el sistema de calificación aprobado-reprobado y el sistema de 4.0 exclusivamente. Entre los alumnos de bellas artes existe una mayor proporción a favor del sistema aprobado-reprobado que por el sistema de 4.0 exclusivamente, mientras que entre los estudiantes de artes y ciencias y los de ingeniería se presenta el patrón inverso. Todos los grupos muestran casi la misma proporción a favor de un sistema de calificación que consiste en una combinación del sistema de 4.0 y el sistema aprobado-reprobado.

Intentemos realizar otro problema para practicar.



Un psicólogo social está interesado en determinar si existe alguna relación entre el nivel de educación de los padres y el número de hijos que tienen. Para averiguarlo, realiza una encuesta y así obtiene los siguientes resultados:

	Número de hijos		Marginal por renglón
	<i>Dos o menos</i>	<i>Más de dos</i>	
Educación universitaria	1 53	2 22	75
Solamente bachillerato	3 37	4 38	75
Marginal por columna	90	60	150

- a. ¿Cuál es la hipótesis nula?
- b. ¿Cuál es la conclusión? Utilice $\alpha = 0.05$.

SOLUCIÓN

- a. Hipótesis nula: El nivel educativo de los padres y el número de hijos que tienen son independientes. La frecuencia obtenida en cada celda se debe a un muestreo aleatorio realizado en una población donde las proporciones de los padres que poseen educación universitaria y los que sólo terminaron el bachillerato, que tienen 1) dos hijos o menos y 2) más de dos hijos son iguales.
- b. Conclusión, utilizando $\alpha = 0.05$:

PASO 1: Calcule el estadístico apropiado. Los datos se presentan en la siguiente tabla. El estadístico apropiado es χ^2_{obt} . Los cálculos se muestran a continuación.

(continúa)

Celda número	f_o	f_e	$\frac{(f_o - f_e)^2}{f_e}$
1	53	$\frac{90(75)}{150} = 45$	$\frac{(53 - 45)^2}{45} = 1.422$
2	22	$\frac{60(75)}{150} = 30$	$\frac{(22 - 30)^2}{30} = 2.133$
3	37	$\frac{90(75)}{150} = 45$	$\frac{(37 - 45)^2}{45} = 1.422$
4	38	$\frac{60(75)}{150} = 30$	$\frac{(38 - 30)^2}{30} = 2.133$
			$\chi^2_{\text{obt}} = \sum \frac{(f_o - f_e)^2}{f_e} = 7.110$

PASO 2: Evaluación del estadístico. Grados de libertad = $(r - 1)(c - 1) = (2 - 1)(2 - 1) = 1$. Con $gl = 1$, y $\alpha = 0.05$, según la tabla H:

$$\chi^2_{\text{crit}} = 3.841$$

Como $\chi^2_{\text{obt}} > 3.841$, rechazamos H_0 . El nivel educacional de los padres y el número de hijos que tienen están relacionados entre sí.

Supuestos subyacentes a χ^2

Un supuesto básico para utilizar χ^2 consiste en que cada una de las observaciones registradas en la tabla de contingencia es independiente de las demás. Esto significa que cada sujeto puede tener sólo un dato en la tabla. No se permite realizar varias mediciones con el mismo sujeto e introducirlas como frecuencias separadas en la misma celda o en otra distinta. Este error produciría una N mayor que el número de observaciones independientes realizadas.

Un segundo supuesto establece que el tamaño de la muestra deberá ser lo suficientemente grande para que la frecuencia esperada en cada celda sea de 5, por lo menos, para las tablas en las que r o c es mayor que 2. Si la tabla es de 1×2 o de 2×2 , entonces cada frecuencia esperada tendrá que ser de 10 cuando menos. Si el tamaño de la muestra es lo bastante pequeño para que se produzcan frecuencias esperadas que violen estos requisitos, entonces la distribución muestral real de χ^2 se desviará considerablemente de la teórica y no podremos aplicar los valores de probabilidad contenidos en la tabla H. Si el experimento utiliza una tabla de contingencia de 2×2 y los datos violan este supuesto, debe usarse la prueba de probabilidad exacta de Fisher.*

Aunque χ^2 se utiliza con frecuencia cuando los datos sólo tienen una escala nominal, sus aplicaciones no se limitan a este tipo de datos. La ji-cuadrada se puede emplear con datos en escalas ordinales, de intervalo o de razón. A pesar de esto e independientemente de cuál sea la escala, los datos se deben reducir a categorías mutuamente excluyentes y a frecuencias adecuadas antes de poder usar χ^2 .

*Esta prueba es analizada en S. Siegel y N. Castellan, Jr., *Nonparametric Statistics for the Behavioral Sciences*, 2a. ed., McGraw-Hill, Nueva York, 1988, pp. 103-111. También es examinada en W. Daniel, *Applied Nonparametric Statistics*, 2a. ed., PWS-Kent, Boston, 1990, pp. 150-162.



LA PRUEBA DE RANGOS CON SIGNO Y PARES IGUALADOS DE WILCOXON

La prueba de rangos con signo de Wilcoxon se utiliza en combinación con el diseño de grupos correlacionados, el cual debe contener datos que tengan por lo menos una escala ordinal. Se trata de una prueba relativamente potente que a veces se emplea en lugar de la prueba t para grupos correlacionados, cuando hay una violación extrema del supuesto de normalidad o cuando los datos no tienen la escala adecuada. En la prueba de rangos con signo de Wilcoxon se toman en cuenta tanto la magnitud como la dirección de los puntajes de diferencia, lo cual la hace más potente que la prueba de los signos. Sin embargo, es menos potente que la prueba t para grupos correlacionados. Con el fin de ilustrar esta prueba, consideremos ahora el siguiente experimento.

un experimento

Promoción de actitudes más favorables hacia la conservación de la vida silvestre

Un destacado grupo ecológico planea realizar una campaña activa para fomentar la conservación de la flora y la fauna silvestres en su país. Como parte de la campaña, piensan exhibir una película producida con el fin de promover actitudes más favorables hacia la conservación de la vida silvestre. Antes de exhibirla a todo el público, los promotores de la campaña desean evaluar los efectos de la película. Para eso se elige al azar a un grupo de 10 sujetos y se les proporciona un cuestionario que mide la actitud individual hacia la conservación de la vida silvestre. A continuación, se les proyecta la película y después de la exhibición contestan de nuevo el cuestionario de actitudes. Éste tiene 50 puntos posibles y cuanto más alta sea la puntuación, tanto más favorable será la actitud hacia la conservación de la vida silvestre. Los resultados se presentan en la tabla 18.5.

1. ¿Cuál es la hipótesis alternativa? Utilice una hipótesis no direccional.
2. ¿Cuál es la hipótesis nula?
3. ¿Cuál es la conclusión de usted? Utilice $\alpha = 0.05_{2\text{ colas}}$.

SOLUCIÓN

1. La hipótesis alternativa se enuncia por lo general sin especificar los parámetros de la población. Para este ejemplo, establece que la película sí influye en las actitudes hacia la conservación de la vida silvestre.
2. La hipótesis nula también se enuncia sin especificar los parámetros de la población. Para este ejemplo, establece que la película no influye en las actitudes hacia la conservación de la vida silvestre.
3. Conclusión, utilizando $\alpha = 0.05_{2\text{ colas}}$: Igual que en todas las demás pruebas de inferencia, el primer paso consiste en calcular el estadístico adecuado. Como los datos fueron obtenidos por medio de cuestionarios, tienen por lo menos una escala ordinal. Para ilustrar el uso de la prueba de rangos con signo de Wilcoxon, supondremos que los datos satisfacen los supuestos de esta prueba (los cuales analizaremos un poco más adelante). El estadístico calculado por la prueba de rangos con signo de Wilcoxon es T_{obt} . Para determinar T_{obt} se requieren cuatro pasos:
 - a. Calcular la diferencia entre cada pareja de datos.
 - b. Ordenar los valores absolutos de los puntajes de diferencia, de menor a mayor.
 - c. Asignar a los rangos resultantes el signo del puntaje de diferencia cuyo valor absoluto produjo ese rango.
 - d. Calcular, por separado, la suma de los rangos para los rangos con signos positivos y negativos. La suma más baja es T_{obt} .

Hemos realizado estos cuatro pasos con los datos obtenidos mediante el cuestionario de actitudes y los valores resultantes se presentan en la tabla 18.5. Así pues, calculamos los puntajes de diferencia que se muestran en la cuarta columna de dicha tabla. Los rangos de los valores absolutos

t a b l a 18.5 Datos y solución del problema sobre la conservación de la vida silvestre

Sujeto	Actitud		Diferencia	Rango de la Diferencia	Rango con signo de la diferencia	Suma de los rangos positivos	Suma de los rangos negativos
	Antes	Después					
1	40	44	4	4	4	4	
2	33	40	7	6	6	6	
3	36	49	13	10	10	10	
4	34	36	2	2	2	2	
5	40	39	-1	1	-1		1
6	31	40	9	8	8	8	
7	30	27	-3	3	-3		3
8	36	42	6	5	5	5	
9	24	35	11	9	9	9	
10	20	28	8	7	7	7	
				55		51	4

$$\frac{n(n+1)}{2} = \frac{10(11)}{2} = 55$$

$$T_{\text{obt}} = 4$$

Según la tabla I, con $N = 10$ y $\alpha = 0.05_{2 \text{ colas}}$,

$$T_{\text{crit}} = 8$$

En vista de que $T_{\text{obt}} < 8$, rechazamos H_0 . Según parece, la película sí promueve actitudes más favorables hacia la conservación de la vida silvestre.

de los puntajes de diferencia se presentan en la quinta columna. Observe que, como comprobación de que el orden quedó correctamente establecido, la suma de los rangos sin signo tendrá que ser igual a $n(n+1)/2$. En nuestro ejemplo, esta suma debe ser igual a 55 [$10(11)/2 = 55$], lo cual es cierto. En el paso c debemos asignar a cada rango el signo del puntaje de diferencia cuyo valor absoluto fue el origen de ese rango. Esto es lo que hemos hecho en la sexta columna. De esta manera, los rangos 1 y 3 reciben signos negativos, y el resto son positivos. Los rangos 1 y 3 reciben signos negativos porque sus puntajes de diferencia asociados son negativos. Determinamos T_{obt} calculando la suma de los rangos positivos y la suma de los rangos negativos. T_{obt} es la menor de esas sumas. En este ejemplo, la suma de los rangos positivos es igual a 51, mientras que la suma de los rangos negativos es igual a 4. Por consiguiente,

$$T_{\text{obt}} = 4$$

Observe que con frecuencia no es necesario calcular las dos sumas. Por lo general, una ligera inspección muestra cuál será la menor. El paso final consiste en evaluar T_{obt} . La tabla I del apéndice D contiene los valores críticos de T para diversos valores de N . Con $N = 10$ y $\alpha = 0.05_{2 \text{ colas}}$, según la tabla I,

$$T_{\text{crit}} = 8$$

Con la prueba de rangos con signo de Wilcoxon, la regla de decisión es:

$$\text{Si } T_{\text{obt}} \leq T_{\text{crit}}, \text{ deberá rechazarse } H_0.$$

Observe que esto es lo opuesto a la regla que hemos utilizado para la mayor parte de las otras pruebas. Como $T_{\text{obt}} < 8$, rechazamos H_0 y concluimos que la película en verdad influye en las actitudes hacia la conservación de la vida silvestre. Tal parece que promueve actitudes más favorables.

Es fácil ver por qué la prueba de rangos con signo de Wilcoxon es más potente que la prueba de los signos, pero no tan potente como la prueba t para grupos correlacionados. La prueba de rangos con signo de Wilcoxon toma en cuenta la magnitud de los puntajes de diferencia, lo cual la hace más potente que la prueba de los signos. Sin embargo, sólo considera el orden de los puntajes de diferencia y no su magnitud real, como la prueba t . Por lo tanto, la prueba de rangos con signo de Wilcoxon no es tan potente como la prueba t .

Tratemos de resolver otro problema para practicar.



Un investigador quiere determinar si la dificultad del material que han de aprender afecta el nivel de ansiedad de los estudiantes universitarios. A cada uno de los miembros de una muestra aleatoria de 12 alumnos se le asignan ciertas tareas de aprendizaje que se clasifican como fáciles y difíciles. Antes que los estudiantes inicien cada tarea, se les presentan algunos ejemplos como muestra del material que van a aprender. A continuación, se mide el nivel de ansiedad que mostraron los alumnos, mediante un cuestionario adecuado. De esta manera, se mide el nivel de ansiedad antes de cada tarea de aprendizaje. Los datos se presentan en la tabla siguiente. Mientras más alta sea la calificación, mayor será el nivel de ansiedad. ¿Cuál es la conclusión, utilizando la prueba de rangos con signo de Wilcoxon y $\alpha = 0.05_{2 \text{ colas}}$?

SOLUCIÓN

La solución se presenta en la tabla siguiente. Observe que existen empates en algunos de los puntajes de diferencia. En general, hay dos tipos posibles de empate. En primer término, los datos en bruto pueden ser iguales, lo cual produce un puntaje de diferencia igual a 0. Si esto ocurre, entonces se eliminan estos datos y la N general se reduce en 1 por cada puntaje de diferencia 0. Los empates también pueden producirse en los puntajes de diferencia, como en este ejemplo. Cuando esto sucede, los rangos de estos datos reciben un valor igual a la media de los rangos empatados. Éste es el mismo procedimiento que aplicamos para el coeficiente de correlación rho de Spearman. Así pues, en este ejemplo, a los dos puntajes de diferencia 3 que están empatados se les asignan rangos de 2.5 $[(2 + 3)/2 = 2.5]$ y los puntajes de diferencia 10 empatados tendrán el rango 9.5. Por lo demás, la solución es muy similar a la del ejemplo anterior.

Estudiante número	Ansiedad		Diferencia	Rango de la diferencia	Rango con signo de la diferencia	Suma de los rangos positivos	Suma de los rangos negativos
	Tareas difíciles	Tareas fáciles					
1	48	40	8	7	7	7	
2	33	27	6	5	5	5	
3	46	34	12	11	11	11	
4	42	28	14	12	12	12	
5	40	30	10	9.5	9.5	9.5	
6	27	24	3	2.5	2.5	2.5	

(continúa)

Estudiante número	Ansiedad		Diferencia	Rango de la diferencia	Rango con signo de la diferencia	Suma de los rangos positivos	Suma de los rangos negativos
	Tareas difíciles	Tareas fáciles					
7	31	33	-2	1	-1		1
8	42	39	3	2.5	2.5	2.5	
9	38	31	7	6	6	6	
10	34	39	-5	4	-4		4
11	38	29	9	8	8	8	
12	44	34	10	9.5	9.5	9.5	
				<u>78</u>		<u>73</u>	5

$$\frac{n(n+1)}{2} = \frac{12(13)}{2} = 78$$


TÉRMINOS Y CONCEPTOS

Hipótesis alternativa (H_1) Según la hipótesis alternativa, la diferencia en los resultados obtenidos de una a otra de las condiciones se debe a la influencia de la variable independiente. La hipótesis alternativa puede ser de dos tipos: direccional o no direccional. Una hipótesis no direccional establece que la variable independiente produce realmente un efecto sobre la variable dependiente, pero no especifica cuál es la dirección de dicho efecto. Una hipótesis alternativa direccional permite precisar cuál es la dirección del efecto esperado.

Hipótesis nula (H_0) La hipótesis nula se formula como la contrapartida lógica de la hipótesis alternativa, de tal manera que si la hipótesis nula es falsa, la hipótesis alternativa debe ser verdadera. Inversamente, si la hipótesis nula es verdadera, entonces la hipótesis alternativa debe ser falsa. En el caso de una hipótesis alternativa no direccional, la hipótesis nula especifica que la variable independiente no produce efecto alguno sobre la variable dependiente. La hipótesis nula para una hipótesis alternativa direccional establece que la variable independiente no produce efecto alguno en la dirección especificada.

Población(es) de la hipótesis nula La población o poblaciones de la hipótesis nula es el conjunto o los conjuntos de datos que se obtendrían en el caso de que el experimento se realizara sobre toda la población y la variable independiente no tuviera efecto alguno. En el diseño de una muestra, es la población correspondiente a la μ conocida. En el caso de un diseño de medidas replicadas, es la población de puntajes de diferencia cuya $\mu_D = 0$ o cuya $P = 0.50$. En un diseño de grupos independientes existen tantas poblaciones como grupos y las muestras son de carácter aleatorio, extraídas de poblaciones donde $\mu_1 = \mu_2 = \mu_3 = \mu_k$.

Distribución muestral La distribución muestral de un estadístico proporciona todos los valores que puede asumir, y muestra la probabilidad de obtener cada valor suponiendo que tal distribución se debe en forma exclusiva al azar o si el muestreo fue realizado en forma aleatoria en la población o las poblaciones de la hipótesis nula. Esta distribución puede ser deducida teóricamente a partir de la probabilidad básica, como lo hicimos con la prueba de los signos, o de manera empírica, como en el caso de las pruebas



EXPERIMENTOS CON VARIOS GRUPOS

Aun cuando el diseño de dos grupos se utiliza con bastante frecuencia en las ciencias del comportamiento, es más común encontrar experimentos con tres o más grupos. El hecho de tener más de dos grupos reporta dos ventajas principales: 1) los grupos adicionales aclaran, generalmente, la interpretación de los resultados y 2) permiten examinar varios niveles de la variable independiente en un experimento. Sin embargo, existe un problema al realizar experimentos con varios grupos. Como hay que hacer muchas comparaciones, aumenta el riesgo de cometer un error de tipo I en el análisis de los datos. La técnica del análisis de varianza nos permite analizar los datos sin correr ese riesgo.

Análisis de varianza de un factor, prueba F

Prueba	Estadístico calculado	Regla de decisión
Análisis de varianza paramétrico de una función, prueba F	$F_{\text{obt}} = \frac{s_B^2}{s_W^2}$	Si $F_{\text{obt}} \geq F_{\text{crit}}$, rechazamos H_0 .

Comentarios generales El análisis de varianza paramétrico utiliza la prueba F para evaluar los datos. Cuando se emplea esta prueba, calculamos F_{obt} , que es esencialmente la razón entre dos estimaciones independientes de la varianza σ^2 de una población. La distribución muestral de F está compuesta por una familia de curvas sesgadas positivamente que varían según el número de grados de libertad. Existen dos valores para los grados de libertad: uno para el numerador y otro para el denominador. La distribución F : 1) está sesgada positivamente, 2) no tiene valores negativos y 3) tiene una mediana aproximadamente igual a 1.

La técnica del análisis de varianza paramétrico se puede utilizar tanto con el diseño de grupos independientes como con el diseño de grupos correlacionados. En este texto hemos considerado únicamente el análisis de varianza de un factor con el diseño de grupos independientes. Dicha técnica permite comparar las medias de todos los grupos en una sola evaluación; esto evita que se incremente la probabilidad de cometer un error de tipo I, como suele suceder cuando se realizan muchas comparaciones individuales. En esencia, el análisis de varianza divide la variabilidad total de los datos en dos partes: la variabilidad que existe dentro de cada grupo (la suma de cuadrados intra grupos) y la variabilidad que se presenta entre los grupos (la suma de cuadrados entre grupos). Cada suma de cuadrados la utilizamos para hacer una estimación independiente de la varianza σ^2 de las poblaciones de la hipótesis nula. Por último, calculamos una razón F en la cual la estimación de la varianza entre grupos aparece en el numerador y la estimación de la varianza intra grupos aparece en el denominador.

Los pasos y las ecuaciones para el cálculo de F_{obt} son los siguientes:

PASO 1: Cálculo de la suma de cuadrados entre grupos, SC_B :

$$SC_B = \left[\frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} + \frac{(\sum X_3)^2}{n_3} + \dots + \frac{(\sum X_k)^2}{n_k} \right] - \frac{\left(\sum_{\text{todos los datos}} X \right)^2}{N}$$

PASO 2: Cálculo de la suma de cuadrados intra grupos, SC_W :

$$SC_W = \sum_{\text{todos los datos}} X^2 - \left[\frac{(\sum X_1)^2}{n_1} + \frac{(\sum X_2)^2}{n_2} + \frac{(\sum X_3)^2}{n_3} + \dots + \frac{(\sum X_k)^2}{n_k} \right]$$

PASO 3: Cálculo de la suma total de cuadrados, SC_T , verificar que $SC_T = SC_W + SC_B$:

$$SC_T = \sum_{\text{todos los datos}} X^2 - \frac{\left(\sum_{\text{todos los datos}} X \right)^2}{N}$$

PASO 4: Cálculo de los grados de libertad para cada estimación:

$$gl_B = k - 1$$

$$gl_W = N - k$$

$$gl_I = N - 1$$

PASO 5: Cálculo de la estimación de la varianza entre grupos, s_B^2 :

$$s_B^2 = \frac{SC_B}{gl_B}$$

PASO 6: Cálculo de la estimación de la varianza intra grupos, s_W^2 :

$$s_W^2 = \frac{SC_W}{gl_W}$$

PASO 7: Cálculo de F_{obt} :

$$F_{\text{obt}} = \frac{s_B^2}{s_W^2}$$

En la hipótesis nula para el análisis de varianza se supone que la variable independiente no produce efecto alguno y que las muestras son muestras aleatorias, extraídas de poblaciones en las que $\mu_1 = \mu_2 = \mu_3 = \mu_k$. En vista de que la estimación de la varianza entre grupos se incrementa a causa del efecto de la variable independiente y la estimación de la varianza intra grupos permanece constante, mientras más grande sea la razón F , tanto menos razonable resultará la hipótesis nula. Evaluamos F_{obt} comparándola con F_{crit} . Si $F_{\text{obt}} \geq F_{\text{crit}}$, we reject H_0 rechazamos H_0 y concluimos que por lo menos una de las condiciones es diferente de una de las demás condiciones, también cuando menos. Observe que la técnica de análisis de varianza no es del tipo direccional.

Comparaciones múltiples Para determinar cuáles son las condiciones que difieren unas de otras, se realizan comparaciones *a priori* o *a posteriori* entre pares de grupos. Las comparaciones *a priori* (llamadas también comparaciones planeadas) resultan apropiadas cuando las comparaciones se han planeado de antemano. No se realiza ajuste alguno para comparaciones múltiples. Las comparaciones planeadas deben ser relativamente pocas y haber surgido de la lógica y el significado del experimento. Al llevar a cabo las comparaciones planeadas, en general comparamos las medias de los grupos especificados, mediante la prueba t para grupos independientes. Determinamos el valor de t_{obt} de la manera usual, excepto que ahora utilizamos la s_W^2 del análisis de varianza en el denominador de la ecuación de t . Comparamos el valor de t_{obt} con el de t_{crit} , utilizando gl_W

con el nivel alfa y la tabla D para determinar t_{crit} . Las ecuaciones para el cálculo de t_{obt} se presentan a continuación:

$$t_{\text{obt}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_W^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

Si $n_1 = n_2$,

$$t_{\text{obt}} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{2s_W^2/n}}$$

Las comparaciones *a posteriori* o *post hoc* no se planean antes de llevar a cabo el experimento; surgen después de examinar los datos o al utilizar la técnica de “escopeta”, la cual consiste en realizar todas las comparaciones factibles de las medias en un intento de extraer la mayor cantidad posible de información del experimento. Por estas razones, en las comparaciones *post hoc* es necesario corregir la probabilidad de cometer un error de tipo I, que se incrementa a causa de las múltiples comparaciones. Existen muchas técnicas para lograr esto. En este texto hemos descrito la prueba de la diferencia honestamente significativa (DHS) de Tukey y la prueba de Newman–Keuls.

Prueba DHS de Tukey

Prueba	Estadístico calculado	Regla de decisión
Prueba DHS de Tukey	Q_{obt} , donde $Q_{\text{obt}} = \frac{\bar{X}_i - \bar{X}_j}{\sqrt{s_W^2/n}}$	Si $Q_{\text{obt}} \geq Q_{\text{crit}}$, rechazamos H_0 .

La prueba DHS ha sido diseñada para comparar todas las parejas posibles de medias, manteniendo en α la tasa de error de tipo I para el conjunto completo de comparaciones. El estadístico Q es muy similar al estadístico t , excepto que siempre es positivo y utiliza las distribuciones Q en lugar de las distribuciones t . Las distribuciones Q (con rango tipo Student) se deducen tomando k muestras aleatorias con la misma n , de la misma población, en lugar de extraer sólo dos muestras, como en el caso de las distribuciones t , y determinando la diferencia entre las medias muestrales máxima y mínima. Para utilizar esta prueba, calculamos Q_{obt} para las comparaciones deseadas y comparamos Q_{obt} con Q_{crit} . Los valores de Q_{crit} se localizan en la tabla G del apéndice D, utilizando k , el número de grados de libertad (gl) para s_W^2 y α . La regla de decisión se presenta en la tabla anterior.

Prueba de Newman–Keuls

Prueba	Estadístico calculado	Regla de decisión
Prueba de Newman–Keuls	Q_{obt} , donde $Q_{\text{obt}} = \frac{\bar{X}_i - \bar{X}_j}{\sqrt{s_W^2/n}}$	Si $Q_{\text{obt}} \geq Q_{\text{crit}}$, rechazamos H_0 .

Comentarios generales La prueba de Newman–Keuls es también una prueba *post hoc* que nos permite realizar todas las comparaciones posibles, por parejas, entre las medias de las muestras. La prueba de Newman–Keuls se parece a la DHS por el hecho de que

Q_{obt} se calcula y se compara con Q_{crit} para evaluar H_0 . Sin embargo, mantiene en α la tasa de error de tipo I, pero no para el conjunto completo de comparaciones, sino para cada comparación. Esto lo realiza modificando el valor de Q_{crit} para cada comparación. El valor de Q_{crit} para una comparación cualquiera está dado por la distribución muestral de Q para el número de medias que se encuentran entre \bar{X}_i y \bar{X}_j después de haber ordenado por rango todas las medias. Este número se simboliza como r . Los valores específicos de Q_{crit} para cualquier análisis se presentan en la tabla G, utilizando r , gl para s_w^2 y α .

Los supuestos subyacentes en el análisis de varianza son los mismos que en el caso de la prueba t para grupos independientes. Existen dos supuestos: 1) las poblaciones de las cuales serán extraídas las muestras deben tener una distribución normal y 2) tiene que haber homogeneidad de varianza entre los grupos. La prueba F es robusta con respecto a las violaciones de la normalidad y la homogeneidad de la varianza, siempre que exista el mismo número de sujetos en cada grupo y $n \geq 30$.

Análisis de varianza de un factor, prueba de Kruskal-Wallis

Prueba	Estadístico calculado	Regla de decisión
Análisis de varianza no paramétrico de un factor, prueba de Kruskal-Wallis	$H_{\text{obt}} = \left[\frac{12}{N(N+1)} \right] \left[\sum_{i=1}^k \frac{(R_i)^2}{n_i} \right] - 3(N+1)$	Si $H_{\text{obt}} > H_{\text{crit}}$, rechazamos H_0 .

Comentarios generales La prueba de Kruskal-Wallis es una técnica no paramétrica, adecuada para un diseño de grupos independientes con k grupos. Se utiliza como opción alternativa al análisis de varianza paramétrico de un factor cuando los supuestos de dicha prueba han sido violados severamente. La prueba de Kruskal-Wallis no supone la normalidad de la población y sólo exige una escala ordinal de la variable dependiente. Todos los datos se agrupan y se ordenan, asignando el rango 1 al menor de los datos, el 2 al siguiente y N al mayor de todos. Luego, sumamos los rangos de cada condición. La prueba de Kruskal-Wallis evalúa si estas sumas de rangos son lo suficientemente diferentes como para que no sea razonable considerar que provienen de muestras elegidas al azar entre la misma población.



ANÁLISIS DE VARIANZA DE DOS FACTORES, PRUEBA *F*

El análisis de varianza paramétrico de dos factores nos permite evaluar los efectos de dos variables y su interacción, en un experimento. En el análisis de varianza paramétrico de dos factores, dividimos la suma total de cuadrados (SC_T) en cuatro componentes: la suma de cuadrados dentro de las celdas (SC_w), la suma de cuadrados por renglón (SC_R), la suma de cuadrados por columna (SC_C) y la suma de cuadrados renglón \times columna (SC_{RC}). Cuando estas sumas de cuadrados se dividen entre los grados de libertad adecuados, forman cuatro estimaciones de la varianza: la estimación de la varianza dentro de las celdas (s_w^2), la estimación de la varianza por renglón (s_R^2), la estimación de la varianza por columna (s_C^2) y la estimación de la varianza renglón \times columna (s_{RC}^2). El efecto de cada variable se determina calculando el valor F_{obt} apropiado y comparándolo con F_{crit} .

Análisis de varianza de dos factores, prueba F

Prueba	Estadístico calculado	Regla de decisión
Análisis de varianza de dos factores, prueba F	$F_{\text{obt}} = \frac{s_R^2}{s_W^2}$ $F_{\text{obt}} = \frac{s_C^2}{s_W^2}$ $F_{\text{obt}} = \frac{s_{RC}^2}{s_W^2}$	Si $F_{\text{obt}} > F_{\text{crit}}$, rechazamos H_0 .

Los pasos y las ecuaciones para calcular los diferentes valores F_{obt} son los siguientes: Calculamos F_{obt} para los efectos principales y la interacción:

PASO 1: Cálculo de la suma de cuadrados por renglón, SC_R :

$$SC_R = \left[\frac{\left(\sum_{\text{renglón } 1} X \right)^2 + \left(\sum_{\text{renglón } 2} X \right)^2 + \cdots + \left(\sum_{\text{renglón } r} X \right)^2}{n_{\text{renglón}}} \right] - \frac{\left(\sum_{\text{todos los datos}} X \right)^2}{N}$$

PASO 2: Cálculo de la suma de cuadrados por columna, SC_C :

$$SC_C = \left[\frac{\left(\sum_{\text{col. } 1} X \right)^2 + \left(\sum_{\text{col. } 2} X \right)^2 + \cdots + \left(\sum_{\text{col. } c} X \right)^2}{n_{\text{col.}}} \right] - \frac{\left(\sum_{\text{todos los datos}} X \right)^2}{N}$$

PASO 3: Cálculo de la suma de cuadrados renglón \times columna, SC_{RC} :

$$SC_{RC} = \left[\frac{\left(\sum_{\text{celda } 11} X \right)^2 + \left(\sum_{\text{celda } 12} X \right)^2 + \cdots + \left(\sum_{\text{celda } rc} X \right)^2}{n_{\text{celda}}} \right] - \frac{\left(\sum_{\text{todos los datos}} X \right)^2}{N} - SC_R - SC_C$$

PASO 4: Cálculo de la suma de cuadrados dentro de las celdas, SC_W :

$$SC_W = \sum_{\text{todos los datos}} X^2 - \left[\frac{\left(\sum_{\text{celda } 11} X \right)^2 + \left(\sum_{\text{celda } 12} X \right)^2 + \cdots + \left(\sum_{\text{celda } rc} X \right)^2}{n_{\text{celda}}} \right]$$

PASO 5: Cálculo del total de la suma de cuadrados, SC_T y verificación de que $SC_T = SC_R + SC_C + SC_{RC} + SC_W$:

$$SC_T = \sum_{\text{todos los datos}} X^2 - \frac{\left(\sum_{\text{todos los datos}} X \right)^2}{N}$$

PASO 6: Cálculo de los grados de libertad para cada estimación de la varianza:

$$gl_R = r - 1$$

$$gl_C = c - 1$$

$$gl_{RC} = (r - 1)(c - 1)$$

$$gl_W = rc(n_{\text{celda}} - 1)$$

$$gl_I = N - 1$$

PASO 7: Cálculo de las estimaciones de la varianza:

$$\text{Estimación de la varianza por renglón} = s_R^2 = \frac{SC_R}{gl_R}$$

$$\text{Estimación de la varianza por columna} = s_C^2 = \frac{SC_C}{gl_C}$$

$$\text{Estimación de la varianza renglón} \times \text{columna} = s_{RC}^2 = \frac{SC_{RC}}{gl_{RC}}$$

$$\text{Estimación de la varianza dentro de las celdas} = s_W^2 = \frac{SC_W}{gl_W}$$

PASO 8: Cálculo de las razones F :

Para el efecto por renglón,

$$F_{\text{obt}} = \frac{s_R^2}{s_W^2}$$

Para el efecto por columna,

$$F_{\text{obt}} = \frac{s_C^2}{s_W^2}$$

Para el efecto de interacción renglón \times columna,

$$F_{\text{obt}} = \frac{s_{RC}^2}{s_W^2}$$

Se comparan los valores F_{obt} con F_{crit} y se sacan conclusiones.



ANÁLISIS DE DATOS NOMINALES

Usted recordará que en el caso de los datos nominales, las observaciones se agrupan en varias categorías discretas, mutuamente excluyentes, y contamos la frecuencia de su ocurrencia en cada categoría. La prueba de inferencia que se utiliza más frecuentemente con datos nominales es la prueba de ji-cuadrada.

Prueba Ji-cuadrada

Prueba	Estadístico calculado	Regla de decisión
Ji-cuadrada	$\chi_{\text{obt}}^2 = \sum \frac{(f_o - f_e)^2}{f_e}$	Si $\chi_{\text{obt}}^2 \geq \chi_{\text{crit}}^2$, rechazamos H_0 .

Comentarios generales Esta prueba resulta apropiada para el análisis de datos de frecuencias cuando intervienen una o dos variables. En el caso de dos variables, los datos de frecuencia se presentan en una tabla de contingencia y hacemos pruebas para saber

si existe alguna relación entre las dos variables. La hipótesis nula establece que no existe relación entre ellas; es decir, que las variables son independientes. La hipótesis alternativa señala que las dos variables están relacionadas entre sí.

La prueba ji-cuadrada mide la discrepancia entre la frecuencia observada (f_o) y la frecuencia esperada (f_e) para cada una de las celdas de la tabla, y luego suma a través de las celdas. La ecuación para χ^2_{obt} se presenta en la tabla. Cuando en los datos intervienen dos variables, la frecuencia esperada para cada celda es la frecuencia que se podría esperar como resultado de un muestreo aleatorio realizado en una población donde las dos variables se presentan en las mismas proporciones en cada categoría. Como las proporciones de la población son desconocidas, estimamos sus valores esperados bajo H_0 a partir de los datos de la muestra, y calculamos las frecuencias esperadas utilizando estas estimaciones. La forma más sencilla de determinar f_e para cada celda consiste en multiplicar los marginales para esa celda y dividir entre N . Si en los datos sólo interviene una variable, las proporciones de la población se determinan sobre alguna base *a priori* (por ejemplo, las mismas proporciones de la población para cada categoría).

Evaluamos el valor obtenido de χ^2 comparándolo con χ^2_{crit} de acuerdo con la regla de decisión presentada en la página 479. El valor crítico de χ^2 se determina mediante la distribución muestral de χ^2 y el nivel alfa. La distribución muestral de χ^2 es una familia de curvas que varía según los grados de libertad. En el experimento con una variable, $gl = k - 1$. En el caso de dos variables, $gl = (r - 1)(c - 1)$. Los valores de χ^2_{crit} se localizan en la tabla H del apéndice D, utilizando gl y α .

Para el uso apropiado de esta prueba, se supone que: 1) cada sujeto tiene sólo una entrada (dato) en la tabla (no hay medidas repetidas de los mismos sujetos), 2) si r o c es mayor que 2, f_e en cada celda debe ser 5 cuando menos y 3) si la tabla es de 1×2 o de 2×2 entonces cada f_e debe ser 10 por lo menos.

La prueba ji-cuadrada se puede utilizar también con datos ordinales, de intervalo o de razón. Sin embargo, no importa cuál sea la escala real, los datos se deben reducir a categorías mutuamente excluyentes y a frecuencias apropiadas para que pueda usarse la prueba χ^2 .



ELECCIÓN DE LA PRUEBA ADECUADA

Uno de los aspectos importantes de la inferencia estadística es la elección de la prueba que se debe utilizar en un experimento o problema cualquiera. Hasta el momento, esto ha sido sencillo. Sólo hemos empleado la prueba que se estudia en cada sección del texto. Sin embargo, la situación representa un reto mayor en este capítulo de repaso. Después de haber analizado muchas pruebas de inferencia, ahora podemos elegir de entre ellas a la que deseamos emplear. Esto, por supuesto, se parece más a la situación que enfrentamos al realizar una investigación.

Al seleccionar una prueba de inferencia, la regla fundamental que debemos aplicar es:

Utilizar la prueba más potente posible.

Para determinar cuáles son las pruebas que podemos utilizar en un experimento o problema determinado, es necesario tomar en cuenta dos factores: la escala de medición de la variable dependiente y el diseño del experimento. En relación con el diagrama de flujo de la figura 19.1, la primera pregunta es: “¿cuál es el nivel de medición utilizado para la variable dependiente?”. Si es nominal, la prueba χ^2 es la única prueba de inferencia que hemos estudiado en este texto y que resulta apropiada para los datos nominales. Por consiguiente, si los datos son de una escala nominal y se cumplen los requisitos de χ^2 (datos de frecuencia, N suficientemente grande, categorías mutuamente excluyentes y ob-

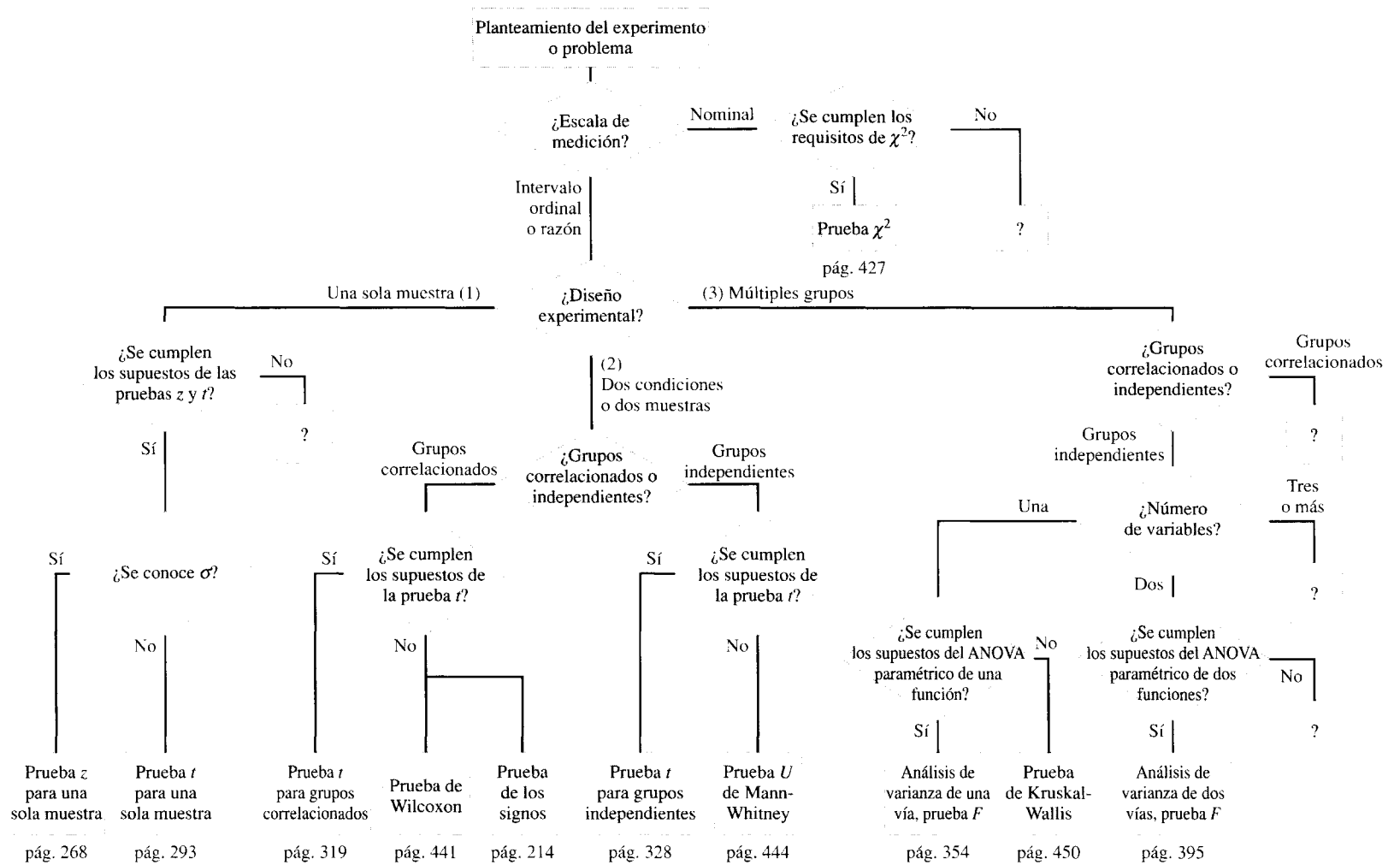


figura 19.1 Diagrama de flujo de decisiones para la selección de la prueba de inferencia apropiada

servaciones independientes), entonces debemos elegir la prueba χ^2 . Pero en caso de que no se cumplan estos supuestos, entonces no sabremos cuál de las pruebas debemos usar, porque no hemos analizado esa cuestión en este texto introductorio. En el diagrama de flujo, esta lamentable situación se indicada por medio de un signo de interrogación (“?”). Sin embargo, me apresuro a garantizarle que las pruebas de inferencia que hemos estudiado son las más comunes, con la posible excepción de experimentos muy complejos con tres o más variables.

Si los datos no son nominales, entonces deben ser de una escala ordinal, de intervalo o de razón. Después de descartar los datos nominales, debemos preguntarnos: “¿cuál es el diseño experimental?”. El diseño utilizado en el experimento limita las pruebas de inferencia que podemos emplear para analizar los datos. Hemos estudiado tres diseños básicos: los de una muestra, los de dos muestras o dos condiciones y los experimentos que incluyen múltiples grupos. Si el diseño utilizado es de una sola muestra (ruta 1 en la figura 19.1), las dos pruebas que hemos estudiado para este diseño son la prueba z y la prueba t para muestras simples. Si los datos cumplen con los supuestos de estas pruebas, para decidir cuál de ellas conviene utilizar debemos preguntarnos: “¿conocemos σ ?” Si la respuesta es “sí”, entonces la prueba apropiada es la z para muestras simples. Si la respuesta es “no”, entonces debemos estimar σ y emplear la prueba t para muestras simples.

Si el diseño experimental es de dos muestras o condiciones (ruta 2), debemos determinar si se trata de grupos correlacionados o independientes. Si se trata de grupos correlacionados y se cumplen los supuestos de t , entonces la prueba adecuada es la prueba t para grupos correlacionados. ¿Por qué? Porque, si los supuestos se cumplen, ésta es la prueba más potente que podemos utilizar para ese diseño. Si los supuestos son seriamente violados, tendremos que emplear una técnica alternativa, como la prueba de Wilcoxon (si se cumplen los supuestos correspondientes) o la prueba de los signos. Si se trata de un diseño de grupos independientes y se cumplen los supuestos de t , deberemos usar la prueba t para grupos independientes. Si los supuestos de t son seriamente violados, entonces tendremos que emplear una técnica alternativa, como la prueba U de Mann-Whitney.

Si el diseño experimental incluye varios grupos (ruta 3), habrá que determinar si es un diseño de grupos independientes o de grupos correlacionados. En este texto analizamos experimentos con varios grupos, utilizando el diseño de grupos independientes. Si el experimento es de grupos múltiples, emplea un diseño de grupos independientes, incluye una variable y se cumplen los supuestos de un análisis de varianza paramétrico, la técnica apropiada es el ANOVA de un factor (prueba F). Si los supuestos son violados gravemente, deberemos usar una técnica alternativa: la prueba de Kruskal-Wallis. Si el diseño es de varios grupos independientes, incluye dos variables y se cumplen los supuestos de un análisis de varianza paramétrico de dos factores, deberemos emplear un ANOVA de dos factores (prueba F) para analizar los datos. No hemos considerado los diseños más complejos con tres o más variables.

PREGUNTAS Y PROBLEMAS

Nota al estudiante: En los capítulos anteriores referentes a la estadística inferencial, cuando a usted se le pedía que resolviera un problema de fin de capítulo, no había duda de qué prueba de inferencia debía utilizar: usted empleaba la prueba estudiada en el capítulo. Por ejemplo, si se trataba de un problema del capítulo 13, usted sabía que debía utilizar la prueba t para una sola muestra, pues era la técnica analizada en ese capítulo. Pero ahora, usted ha avanzado y ahora sabe tanto

de estadística que podría usar más de una prueba de inferencia para resolver un problema. Con frecuencia, podrá utilizar una prueba paramétrica o no paramétrica. Éste es un nuevo reto. La regla a seguir consiste en emplear la técnica más potente que nos permitan los datos. Para los problemas de este capítulo, suponga siempre que se satisfacen los supuestos subyacentes a la prueba paramétrica, a menos que el problema indique de modo explícito lo contrario.

1. Describa brevemente los siguientes términos:
 Hipótesis alternativa
 Hipótesis nula
 Población de la hipótesis nula
 Distribución muestral
 Región crítica para el rechazo de H_0
 Nivel alfa
 Error de tipo I
 Error de tipo II
 Potencia
2. Describa en pocas palabras el proceso de prueba de una hipótesis. Asegúrese de incluir en el análisis los términos enumerados en la pregunta 1.
3. ¿Por qué son tan importantes las distribuciones muestrales en la prueba de hipótesis?
4. Un maestro realiza un experimento mediante un diseño de grupos independientes que ha preparado a fin de evaluar dos métodos para enseñanza de ortografía de tercer grado. Los resultados no son significativos, por lo cual el maestro concluye que los dos métodos son iguales. ¿Está bien fundamentada esta conclusión? Suponga que el estudio fue diseñado y realizado de manera adecuada; por ejemplo, que se tenían los controles adecuados, que el tamaño de la muestra fue suficientemente grande, que se utilizó el estadístico apropiado, etcétera.
5. ¿Por qué son preferibles las pruebas paramétricas sobre las no paramétricas?
6. Enumere los factores que afectan la potencia de un experimento y explique cómo se pueden utilizar para aumentar dicha potencia.
7. ¿Qué factores determinan el tipo de prueba de inferencia a utilizar en el análisis de los datos de un experimento?
8. Enumere los distintos diseños experimentales analizados en este texto. Además, liste las pruebas de inferencia apropiadas para cada diseño, por orden de sensibilidad.
9. ¿Cuáles son los supuestos subyacentes a cada prueba de inferencia?
10. ¿Cuáles son los dos pasos que se siguen al analizar los datos de cualquier estudio que implique una prueba de hipótesis?
11. Un nuevo competidor, en la industria del whisky escocés, lleva a cabo un experimento para comparar el whisky que elabora (llamado McPherson's Joy) con las otras tres marcas principales en el mercado. Para eso elige una muestra aleatoria de 200 bebedores de whisky, entre los habitantes de Nueva York. Se le pide a cada individuo que pruebe los cuatro whiskys y que elija el que le parezca mejor. Por supuesto, los whiskys no tienen etique-

tas y el orden en que se prueban está bien equilibrado. El número de sujetos que prefirieron cada marca se presenta en la siguiente tabla:

<i>McPherson's Joy</i>	<i>Marca X</i>	<i>Marca Y</i>	<i>Marca Z</i>	
58	52	48	42	200

- a. ¿Cuál es la hipótesis alternativa? Utilice una hipótesis no direccional.
 - b. ¿Cuál es la hipótesis nula?
 - c. Utilizando $\alpha = 0.05$, ¿cuál es la conclusión de usted? I/O
12. Un psicólogo que está interesado en el aprendizaje animal, realiza un experimento para determinar el efecto de la hormona ACT (adrenocorticotrópica) sobre el aprendizaje de evitación. Para este experimento, se elige de manera aleatoria a 20 ratas macho de 100 días de edad, de entre los ejemplares del bioterio de la universidad. Diez de las 20 ratas seleccionadas al azar reciben inyecciones de ACT, 30 minutos antes de ser colocadas en la situación de evitación. Las otras 10 reciben inyecciones de placebo. El número de ensayos necesarios para que cada animal aprendiera la tarea se presenta a continuación:
- | ACT | Placebo |
|------------|----------------|
| 58 | 74 |
| 73 | 92 |
| 80 | 87 |
| 78 | 84 |
| 75 | 72 |
| 74 | 82 |
| 79 | 76 |
| 72 | 90 |
| 66 | 95 |
| 77 | 85 |
- a. ¿Cuál es la hipótesis alternativa no direccional?
 - b. ¿Cuál es la hipótesis nula?
 - c. Utilizando $\alpha = 0.01$ 2 colas, ¿cuál es la conclusión de usted?
 - d. ¿Qué error podría cometer si saca su conclusión como lo hizo en la parte c?
 - e. ¿A qué población se aplican estos resultados?
 - f. ¿Cuál es la magnitud del efecto? biológica
13. Una nutrióloga de una universidad se pregunta si el reciente énfasis en las dietas saludables ha afectado a los estudiantes de primer grado de su uni-

versidad. Para responder la interrogante, ella realiza un estudio con el cual intenta determinar si la dieta de los alumnos actuales contiene menos grasa que la dieta de los estudiantes anteriores de primer grado. Para determinar el consumo diario de grasa, 15 estudiantes de la clase actual de primer grado llevan un registro de todo lo que comen durante 7 días. Los resultados muestran que para estos 15 individuos, el porcentaje promedio del consumo diario de grasa es de 37%, con una desviación estándar de 12%. Los registros correspondientes a un gran número de alumnos de años anteriores muestran un porcentaje promedio de consumo de grasa del 40%, una desviación estándar del 10.5% y una distribución normal de los datos.

- Con base en estos datos, ¿es menor el consumo diario de grasa de los estudiantes actuales que el de los estudiantes de años anteriores? Utilice $\alpha = 0.05$ cola.
 - Si el consumo promedio diario de grasa de los estudiantes actuales es de 35%, ¿cuál es la potencia del experimento para detectar este nivel del efecto real?
 - Si N aumenta a 30, ¿cuál es la potencia para detectar un consumo promedio diario de grasa de 35%?
 - Si la nutrióloga quiere alcanzar una potencia de 0.9000 para detectar un efecto real de, cuando menos, cinco puntos medios por debajo de las normas establecidas de la población, ¿qué valor de N debe usar? salud, I/O
14. Un fisiólogo realiza un experimento diseñado para determinar el efecto de la tiroxina exógena (hormona producida por la glándula tiroides) sobre la actividad. Cuarenta ratas macho son asignadas de manera aleatoria a cuatro grupos, de manera que haya 10 roedores en cada grupo. Cada uno de los grupos es inyectado con una cantidad diferente de tiroxina. El grupo 1 no recibe tiroxina y sólo se le administra una solución salina. El grupo 2 recibe una pequeña cantidad, el grupo 3 una cantidad moderada y el grupo 4 una alta cantidad de tiroxina. Después de las inyecciones, cada animal pasa a un aparato de campo abierto para medir su nivel de actividad. Este aparato está formado por una plataforma relativamente grande, cuyos lados evitan que el animal salga de la plataforma. Se traza una retícula sobre la superficie de la plataforma, de modo que toda la superficie quede cubierta por cuadrados. Para

medir la actividad, el experimentador cuenta simplemente el número de cuadrados recorridos por el animal durante cierto periodo fijo. En nuestro experimento, cada rata permanece en el aparato durante 10 minutos. Los resultados obtenidos se presentan en la tabla; los datos son el número de cuadrados recorridos por minuto.

- ¿Cuál es la hipótesis nula general?
- Utilizando $\alpha = 0.05$, ¿cuál es la conclusión de usted?
- ¿Cuál es la magnitud del efecto?
- Evalúe la hipótesis a priori en el sentido de que una alta cantidad de tiroxina exógena produce un efecto sobre la actividad, diferente al del efecto de la solución salina. Utilice $\alpha = 0.05$ 2 colas.
- Utilice la prueba DHS de Tukey con $\alpha = 0.05$ 2 colas para comparar todas las parejas posibles de las medias. ¿Cuál es la conclusión de usted? biológica

Cantidad de tiroxina

<i>Nula,</i> 1	<i>Baja,</i> 2	<i>Moderado,</i> 3	<i>Alta,</i> 4
2	4	8	12
3	3	7	10
3	5	9	8
2	5	6	7
5	3	5	9
2	2	8	13
1	4	9	11
3	3	7	8
4	6	8	7
5	4	4	9

- Se realiza un estudio para determinar si una dieta adecuada, aunada al ejercicio físico, es más eficaz para propiciar la pérdida de peso que la dieta por sí sola. Se utilizan 12 parejas de sujetos. Los sujetos son dispuestos en parejas que concuerdan en el peso y en el nivel de ejercicio iniciales, la edad y el sexo. Un miembro de cada pareja se pone a dieta durante 3 meses. El otro miembro recibe la misma dieta, pero adicionalmente emprende un régimen de ejercicio moderado. Los siguientes datos indican la pérdida de peso de cada sujeto, en libras, al terminar el periodo de 3 meses:

Pareja	Dieta más ejercicio	Solamente dieta
1	24	16
2	20	18
3	22	19
4	15	16
5	23	18
6	21	18
7	16	17
8	17	19
9	19	13
10	25	18
11	24	19
12	13	14

Al responder las siguientes preguntas, suponga que los datos son tan poco normales que no permiten utilizar una prueba paramétrica apropiada.

- ¿Cuál es la hipótesis alternativa? Utilice una hipótesis direccional.
 - ¿Cuál es la hipótesis nula?
 - Utilizando $\alpha = 0.05_{1 \text{ cola}}$, ¿cuál es la conclusión de usted? salud
16. a. ¿Qué otra prueba no paramétrica habría empleado usted para analizar los datos del problema 15?
- Use esta prueba para analizar los datos. ¿Cuál es la conclusión de usted con $\alpha = 0.05_{1 \text{ cola}}$?
 - Explique la diferencia entre sus conclusiones para los problemas 16b y 15c.
 - Sea P la probabilidad, para cada sujeto, de que la dieta más el ejercicio dé por resultado una mayor pérdida de peso. Si $P_{\text{real}} = 0.75$, utilizando la prueba de los signos con $\alpha = 0.05_{1 \text{ cola}}$, ¿cuál es la potencia del experimento para detectar este nivel del efecto? ¿Cuál es la probabilidad de cometer un error de tipo II? salud
17. Un investigador de la sexualidad humana quiere determinar si existe una relación entre el género y la hora del día preferida para tener relaciones sexuales. Se realiza una encuesta cuyos resultados aparecen en la siguiente tabla; los datos de las entradas son la cantidad de individuos que prefieren la mañana o la tarde:

Género	Relaciones		
	Mañana	Tarde	
Masculino	36	24	60
Femenino	28	32	60
	64	56	120

- ¿Cuál es la hipótesis nula?
 - Utilizando $\alpha = 0.05$, ¿cuál es la conclusión de usted? social
18. Una psicóloga quiere saber si el estado interno de los individuos afecta las percepciones de éstos. Específicamente, quiere determinar si el hambre influye sobre la percepción. Para verificar esta hipótesis, ella divide al azar a 24 sujetos en tres grupos de 8 individuos cada uno. Se pide a los voluntarios que describan las "imágenes" que observan en una pantalla. En realidad no son imágenes, sino formas ambiguas. El hambre se controla privando de alimento a los sujetos. A un grupo se le muestran las imágenes una hora después de comer; al otro, a las 4 horas de haber comido y al último grupo, 12 horas después de comer. Se lleva un registro del número de objetos relacionados con la comida que fueron mencionados por cada sujeto: Así se obtuvieron los siguientes datos:

Privación del alimento

1 hora	4 horas	12 horas
1	2	3
2	6	8
5	7	10
7	6	15
2	10	19
1	15	9
8	12	14
7	7	15
6	6	12

- ¿Cuál es la hipótesis nula general?
 - ¿Cuál es la conclusión de usted? Utilice $\alpha = 0.05$.
 - Utilizando la prueba de Newman-Keuls con $\alpha = 0.05_{2 \text{ colas}}$, realice todas las comparaciones posibles *post hoc* entre las parejas de medias. ¿Cuál es la conclusión de usted? cognitiva
19. Una ingeniera de una empresa líder en electrónica, afirma que ha inventado un procedimiento para que los cinescopios de TV tengan mayor duración. Entonces se realizan pruebas con 24 cinescopios fabricados con el nuevo procedimiento; éstos muestran un promedio de vida de 1725 horas y una desviación estándar de 85 horas. Las pruebas realizadas durante los últimos tres años con un gran número de cinescopios elaborados por el procedimiento antiguo, muestran un promedio de vida de 1538 horas. ¿Es correcta la afirmación de la ingeniera? Utilice $\alpha = 0.01_{1 \text{ cola}}$ para tomar una decisión. I/O.

20. En un estudio realizado para determinar el efecto del alcohol sobre la agresividad, 17 voluntarios adultos fueron asignados, al azar, a dos grupos: uno experimental y otro control. Los sujetos del grupo experimental bebieron vodka diluido en jugo de naranja, mientras que los sujetos del grupo de control bebieron solamente jugo. Después de beber, se les aplicó una prueba de agresividad; así se obtuvieron los datos siguientes. Los puntajes más altos indican una mayor agresividad:

Jugo de naranja	Vodka más jugo de naranja
11	14
9	13
14	19
15	16
7	15
10	17
8	11
10	18
8	

- a. ¿Cuál es la hipótesis alternativa? Utilice una hipótesis no direccional.
 - b. ¿Cuál es la hipótesis nula?
 - c. Utilizando $\alpha = 0.05_{2 \text{ colas}}$, ¿cuál es la conclusión de usted? social, clínica
21. El encargado de la oficina de admisiones de una gran universidad se pregunta cuán fuerte es la relación entre las calificaciones obtenidas en el bachillerato y las que se obtienen en la universidad. Durante los 2 años que ha permanecido en el puesto, él ha ponderado fuertemente las calificaciones del bachillerato al decidir qué alumnos deben ser admitidos en la universidad, aunque aún no ha visto datos que relacionen ambas variables. Como posee buenos fundamentos experimentales, decide realizar un estudio por su cuenta. Para eso elige de manera aleatoria a 15 estudiantes del último grado de su universidad y solicita las calificaciones que ellos obtuvieron en el bachillerato y en la universidad. Así se consiguieron los siguientes datos:

Sujeto	Calificaciones	
	Bachillerato	Universidad
1	2.2	1.5
2	2.6	1.7
3	2.5	2.0
4	2.2	2.4
5	3.0	1.7
6	3.0	2.3
7	3.1	3.0
8	2.6	2.7
9	2.8	3.2
10	3.2	3.6
11	3.4	2.5
12	3.5	2.8
13	4.0	3.2
14	3.6	3.9
15	3.8	4.0

- a. Calcule r_{obt} para estos datos.
 - b. ¿Es significativa la correlación? Utilice $\alpha = 0.05_{2 \text{ colas}}$.
 - c. ¿Qué proporción de la variabilidad en las calificaciones de la universidad es explicada por las calificaciones del bachillerato?
 - d. ¿Tiene razón el funcionario en ponderar fuertemente las calificaciones del bachillerato al determinar a qué estudiantes debe admitir en la universidad?
22. Se realiza un experimento para evaluar el efecto del cigarrillo sobre el ritmo cardiaco. Para ello, se elige de manera aleatoria a 10 personas que fuman cigarrillos. Cada sujeto participa en dos condiciones. En la condición 1, el sujeto descansa durante una hora, después de lo cual se mide su ritmo cardiaco. En la condición 2, el sujeto descansa durante una hora y luego fuma dos cigarrillos. En la condición 2, el ritmo cardiaco se mide después que el sujeto ha terminado de fumar los cigarrillos. A continuación presentamos los datos obtenidos.

Sujeto	Latidos por minuto	
	<i>Sin fumar,</i> 1	<i>Fumando,</i> 2
1	72	76
2	80	84
3	68	75
4	74	73
5	80	86
6	85	88
7	86	84
8	78	80
9	68	72
10	67	70

- a. ¿Cuál es la hipótesis alternativa no direccional?
- b. ¿Cuál es la hipótesis nula?
- c. Utilizando $\alpha = 0.05_{2 \text{ colas}}$, ¿cuál es la conclusión de usted? biológica, clínica
23. Para enfrentar la actual crisis petrolera, el gobierno debe decidirse por una línea de acción. Hay dos opciones a este respecto: 1) permitir que suba el precio del petróleo o 2) imponer un racionamiento de la gasolina. Entonces se realiza una encuesta entre individuos de diversas ocupaciones para comprobar si existe alguna relación entre éstas y la línea de acción con la cual concuerda cada uno de ellos. Los resultados se presentan en la siguiente tabla de 3×2 ; los datos de las celdas son el número de individuos que están a favor de la línea de acción que encabeza la celda:

Ocupación	Línea de acción		
	<i>Aumento del precio del petróleo</i>	<i>Racionamiento de la gasolina</i>	
Empresario	180	120	300
Ama de casa	135	165	300
Obrero	152	148	300
	467	433	900

- a. ¿Cuál es la hipótesis nula?
- b. Utilizando $\alpha = 0.05$, ¿cuál es la conclusión de usted? I/O
24. Usted quiere verificar la hipótesis de que los hombres y las mujeres adultos tienen diferente habilidad para el razonamiento lógico. Con este

propósito, elige al azar a 16 adultos de la ciudad donde vive y les aplica una prueba de razonamiento lógico. Una calificación más alta indica una mayor habilidad para el razonamiento lógico. De esta manera obtuvo los siguientes datos:

Hombres	Mujeres
70	80
60	50
82	81
65	75
83	95
92	85
85	93
	75
	90

Al responder las siguientes preguntas, suponga que los datos violan los supuestos subyacentes al uso de la prueba paramétrica apropiada y que usted debe analizar los datos con una prueba no paramétrica.

- a. ¿Cuál es la hipótesis nula?
- b. Utilizando $\alpha = 0.05_{2 \text{ colas}}$, ¿cuál es la conclusión de usted?
25. Para elaborar su tesis doctoral, una estudiante del posgrado en estudios sobre la mujer investigó los efectos de la tensión sobre el ciclo menstrual. Obtuvo un muestreo aleatorio de 42 mujeres y utilizó un diseño de medidas repetidas con dos condiciones. Sin embargo, una de las voluntarias abandonó el estudio. En la condición de tensión, la longitud promedio del ciclo menstrual de las 41 mujeres restantes fue de 29 días, con una desviación estándar de 14 días. Con base en estos datos, determine usted la longitud promedio del ciclo menstrual bajo tensión, con un intervalo de confianza de 95%, para toda la población. salud, social
26. Una investigadora interesada en la justicia social piensa que los hispanos no están representados suficientemente entre los maestros de bachillerato, en la región del país donde ella vive. Para comprobarlo, extrae una muestra aleatoria de 150 maestros de bachillerato de esa localidad geográfica. Los resultados muestran que había 15 maestros hispanos en la muestra. El porcentaje de hispanos residentes en esa localidad es de 22%.
- a. ¿Cuál es la hipótesis nula?
- b. Utilizando $\alpha = 0.05$, ¿cuál es la conclusión de usted? social

27. Una estudiante sospecha que los profesores de ciencias físicas son más autoritarios que los de ciencias sociales. Para comprobarlo, realiza un estudio en el que elige al azar a 6 profesores de física, 6 de psicología y 6 de sociología y los hace responder un cuestionario para medir el autoritarismo. Los resultados se presentan a continuación. Mientras más alta sea la calificación, tanto más autoritario será el individuo. Suponga que los datos violan de manera severa los supuestos de normalidad. ¿Cuál es la conclusión de usted, utilizando $\alpha = 0.05$.

Profesores		
<i>Física</i>	<i>Psicología</i>	<i>Sociología</i>
75	73	71
82	80	80
80	85	90
97	92	78
94	70	94
76	69	68

social

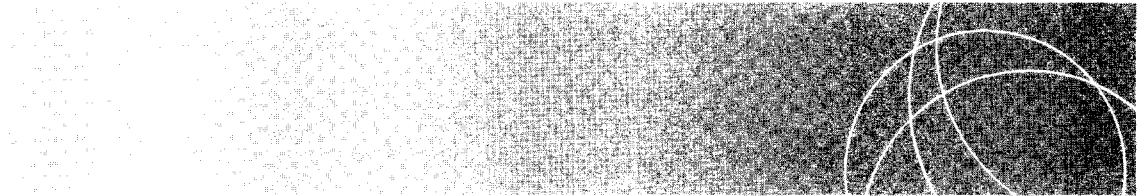
28. Un investigador del sueño quiere determinar si el hecho de dormir la siesta mejora el rendimiento y, en tal caso, si es importante el hecho de que la siesta se tome en las primeras horas de la tarde o casi al anochecer. Para averiguarlo, toma

una muestra aleatoria de 30 estudiantes y los asigna a una de seis condiciones: dormir una siesta por la tarde o al anochecer, descansar sin dormir por la tarde o al anochecer, o una condición de control con actividad normal, también por la tarde o al anochecer. Hay 5 sujetos bajo cada condición. Cada individuo realiza la actividad correspondiente, según la condición en la cual se encuentra, después de lo cual se somete a una prueba de rendimiento. Una calificación más alta indica mejor rendimiento. De esta manera se han obtenido los siguientes resultados. ¿Cuál es la conclusión de usted? Utilice $\alpha = 0.05$ y suponga que los datos provienen de poblaciones que presentan una distribución normal.

Hora del día	Actividad					
	Siesta		Descanso sin dormir		Normal	
Tarde	8	9	7	8	3	4
	7	5	6	4	5	5
	6		5		6	
Anochecer	6	5	5	3	4	2
	7	4	4	5	3	4
	6		4		3	

cognitiva

APÉNDICES



- A** Repaso de los prerrequisitos de matemáticas 491
- B** Ecuaciones 501
- C** Respuestas a las preguntas y problemas de final de capítulo 509
- D** Tablas 524

Repaso de los prerrequisitos de matemáticas

Introducción
Resolución de ecuaciones con una incógnita
Interpolación lineal



INTRODUCCIÓN

En este apéndice presentaremos un repaso de ciertas habilidades matemáticas básicas que creemos importantes como fundamento de un curso introductorio de estadística. Este apéndice pretende ser un repaso del material que ha aprendido con anterioridad, pero que podría estar "empolvado" por falta de uso. Para los estudiantes que han estado alejados de las matemáticas durante muchos años y que no tengan mucha seguridad acerca de sus bases matemáticas, recomendamos los siguientes libros: H. M. Walker, *Mathematics Essential for Elementary Statistics* (edición revisada, Holt, Nueva York, 1951) o bien A. J. Washington, *Arithmetic and Beginning Algebra* (Addison-Wesley, Reading, MA, 1984).

Símbolos algebraicos

<i>Símbolo</i>	<i>Explicación</i>
$>$	Es mayor que
$5 > 4$	5 es mayor que 4.
$X > 10$	X es mayor que 10.
$a > b$	a es mayor que b .
$<$	Es menor que
$7 < 9$	7 es menor que 9.
$X < 12$	X es menor que 12.
$a < b$	a es menor que b .
$2 < X < 20$	X es mayor que 2 y menor que 20, o bien, el valor de X está entre 2 y 20.
$X < 2$ o $X > 20$	X es menor que 2 o mayor que 20, o bien, el valor de X está fuera del intervalo de 2 a 20.
\geq	Es igual o mayor que
$X \geq 3$	X es igual o mayor que 3.
$a \geq b$	a es igual o mayor que b .
\leq	Es igual o menor que
$X \leq 5$	X es igual o menor que 5.
$a \leq b$	a es igual o menor que b .
\neq	Es diferente de
$3 \neq 5$	3 es diferente de 5.
$X \neq 8$	X es diferente de 8.
$a \neq b$	a es diferente de b .
$ X $	El valor absoluto de X ; el valor absoluto de X es igual a la magnitud de X , independientemente de cuál sea su signo.
$ +7 $	El valor absoluto de 7; $ +7 = 7$.
$ -5 $	El valor absoluto de -5 ; $ -5 = 5$.

Operaciones aritméticas

<i>Operación</i>	<i>Ejemplo</i>
1. <i>Suma de dos números positivos:</i> Para sumar dos números positivos, se suman sus valores absolutos y el resultado tendrá signo positivo.	$2 + 8 = 10$
2. <i>Suma de dos números negativos:</i> Para sumar dos números negativos, se suman sus valores absolutos y el resultado tendrá signo negativo.	$-3 + (-4) = -7$
3. <i>Suma de dos números con signos opuestos:</i> Para sumar dos números con signos opuestos, se calcula la diferencia entre sus valores absolutos y el resultado tendrá el signo del número cuyo valor absoluto sea mayor.	$16 + (-10) = 6$ $3 + (-14) = -11$
4. <i>Resta de un número de otro:</i> Para restar un número de otro, se cambia el signo del número por restar y se procede igual que en la suma (operaciones 1, 2 o 3).	$16 - 4 = 16 + (-4) = 12$ $5 - 8 = 5 + (-8) = -3$ $9 - (-6) = 9 + (+6) = 15$ $-3 - 5 = -3 + (-5) = -8$
5. <i>Multiplicación de una serie de números:</i>	$2(-5)(-6)(3) = 180$
a. Al multiplicar una serie de números, el resultado es positivo cuando el número de valores negativos incluidos en la serie es par.	$-3(-7)(-2)(-1) = 42$ $-a(-b) = ab$
b. Al multiplicar una serie de números, el resultado es negativo cuando el número de valores negativos incluidos en la serie es impar.	$4(-5)(2) = -40$ $-8(-2)(-5)(3) = -240$ $a(-b)(-c) = -abc$
6. <i>División de una serie de números:</i>	
a. Al dividir una serie de números, el resultado es positivo cuando el número de valores negativos incluidos en la serie es par.	$\frac{-4}{-8} = \frac{1}{2}$ $\frac{-3(-4)(2)}{6} = 4$ $\frac{-a}{-b} = \frac{a}{b}$
b. Al dividir una serie de números, el resultado es negativo cuando el número de valores negativos incluidos en la serie es impar.	$\frac{-2}{5} = -0.40$ $\frac{3(-2)}{-4} = -1.5$ $\frac{-a}{b} = -\frac{a}{b}$

Reglas que rigen el orden de las operaciones aritméticas

<i>Regla</i>	<i>Ejemplo</i>
1. El orden en que se suman los números no modifica el resultado.	$6 + 4 + 11 = 4 + 6 + 11 = 11 + 6 + 4 = 21$ $6 + (-3) + 2 = -3 + 6 + 2 = 2 + 6 + (-3) = 5$
2. El orden en que se multiplican los números no modifica el resultado.	$3 \times 5 \times 8 = 8 \times 5 \times 3 = 5 \times 8 \times 3 = 120$
3. Si aparecen una multiplicación y una suma o resta, la multiplicación debe realizarse en primer lugar, a menos que los paréntesis o corchetes indiquen lo contrario.	$4 \times 5 + 2 = 20 + 2 = 22$ $6 \times (14 + 12) \times 3 = 6 \times 2 \times 3 = 36$ $6 \times (4 + 3) \times 2 = 6 \times 7 \times 2 = 84$
4. Si aparecen una división y una suma o resta, la división debe realizarse en primer lugar, a menos que los paréntesis o corchetes indiquen lo contrario.	$12 \div 4 + 2 = 3 + 2 = 5$ $12 \div (4 + 2) = 12 \div 6 = 2$ $12 \div 4 + 2 \div 3 + 2 = 1$ $12 \div (4 + 2) = 12 \div 2 = 6$

Reglas sobre los paréntesis y corchetes

<i>Regla</i>	<i>Ejemplo</i>
1. Los paréntesis y corchetes indican que lo que está encerrado dentro de ellos debe considerarse como un solo número.	$(2 + 8)(6 - 3 + 2) = 10(5) = 50$
2. Cuando hay paréntesis dentro de unos corchetes, se realizan primero las operaciones que están indicadas dentro de los paréntesis.	$[(4)(6 - 3 + 2) + 6][2] = [(4)(5) + 6][2]$ $= [20 + 6][2] = [26][2] = 52$
3. Cuando no sea conveniente reducir a un solo número lo que está encerrado entre paréntesis, los paréntesis pueden eliminarse de la siguiente manera:	
a. Si hay un signo positivo antes de los paréntesis, éstos se eliminan sin modificar los signos de los números contenidos dentro de ellos.	$1 + (3 + 5 - 2) = 1 + 3 + 5 - 2 = 7$
b. Si hay un signo negativo antes de los paréntesis, éstos se eliminan, y se cambian los signos de los números contenidos dentro de ellos.	$4 - (6 - 2 - 1) = 4 - 6 + 2 + 1 = 1$
c. Si afuera de los paréntesis hay un número multiplicando, entonces todos los términos que están dentro de los paréntesis deberán multiplicarse por dicho número.	$3(2 + 3 - 4) = 6 + 9 - 12 = 3$ $a(b + c + d) = ab + ac + ad$
d. El producto de dos sumas se obtiene multiplicando cada elemento de una suma por los elementos de la otra.	$(a + b)(c + d) = ac + ad + bc + bd$
e. Si dentro de los paréntesis hay operaciones indicadas, siempre se realizan primero dichas operaciones antes de combinar el contenido de los paréntesis con otros términos.	$5 + 4(3 + 1) - 6 = 5 + 4(4) - 6 = 5 + 16 - 6 = 27$ $4 + (3 + 1)/2 = 4 + 4/2 = 4 + 2 = 6$ $1 + (3 + 2)^2 = 1 + (5)^2 = 1 + 25 = 26$

Fracciones

Operaciones**Ejemplo**1. *Suma de fracciones:*

Para sumar dos fracciones, 1) se determina el mínimo común denominador. 2) se expresa cada fracción en términos del mínimo común denominador y 3) se suman los numeradores y se divide la suma entre el común denominador.

$$\frac{1}{3} + \frac{1}{2} = \frac{2}{6} + \frac{3}{6} = \frac{5}{6}$$

$$\frac{a}{b} + \frac{c}{d} = \frac{ad}{bd} + \frac{cb}{bd} = \frac{ad + cb}{bd}$$

2. *Multipliación de fracciones:*

Para multiplicar dos fracciones, se multiplican entre sí los numeradores y se divide este resultado entre el producto de los denominadores.

$$\frac{2}{5} \left(\frac{3}{7} \right) = \frac{6}{35}$$

$$\frac{a}{b} \left(\frac{c}{d} \right) = \frac{ac}{bd}$$

3. *Conversión de una fracción en su equivalente decimal:*

Para convertir una fracción en decimal, se realiza la división indicada, redondeando al número necesario de cifras (dos dígitos en el ejemplo presentado).

$$\frac{3}{7} = 0.429 = 0.43$$

4. *Conversión de un decimal en un porcentaje:*

Para convertir una fracción decimal en un porcentaje, se multiplica la fracción decimal por 100.

$$0.43 \times 100 = 43\%$$

5. *Multipliación de un entero por una fracción:*

Para multiplicar un entero por una fracción, se multiplica el entero por el numerador de la fracción y se divide este producto entre el denominador.

$$\frac{2}{5}(4) = \frac{8}{5} = 1.60$$

6. *Cancelación:*

Al multiplicar varias fracciones entre sí, podemos cancelar los factores idénticos en el numerador y el denominador.

$$\frac{1}{12} \left(\frac{4}{9} \right) \left(\frac{3}{10} \right) = \frac{1}{18}$$

Exponentes

<i>Operaciones</i>	<i>Ejemplo</i>
1. <i>Multiplicación de un número por sí mismo 2 veces</i>	$(4)^2 = 4(4) = 16$ $a^2 = aa$
2. <i>Multiplicación de un número por sí mismo 3 veces</i>	$(4)^3 = 4(4)(4) = 64$ $a^3 = aaa$
3. <i>Multiplicación de un número por sí mismo N veces</i>	$4^N = \underbrace{4(4)(4) \cdots (4)}_N$
4. <i>Multiplicación de dos cantidades exponenciales con la misma base:</i> El producto de dos cantidades exponenciales con la misma base es dicha base elevada a la suma de los exponentes.	$(2)^2(2)^4 = (2)^{2+4} = (2)^6$ $a^N a^P = a^{N+P}$
5. <i>División de dos cantidades exponenciales con la misma base:</i> El cociente de dos cantidades exponenciales con la misma base es dicha base elevada a un exponente igual a la resta del exponente que está en el numerador menos el exponente que está en el denominador.	$\frac{(2)^4}{(2)^2} = (2)^2$ $\frac{a^N}{a^P} = a^{N-P}$
6. <i>Elevar una base a un exponente negativo:</i> Una base elevada a un exponente negativo es igual a 1 entre la base elevada al valor positivo del exponente.	$(2)^{-3} = \frac{1}{(2)^3}$ $a^{-N} = \frac{1}{a^N}$

Factorización Al factorizar una expresión algebraica, intentamos reducir la expresión a los componentes más sencillos, tales que al ser multiplicados entre sí den como resultado la expresión original.

<i>Ejemplo</i>	<i>Explicación</i>
$ab + ac + ad = a(b + c + d)$	Factorizamos a de cada elemento.
$abc - 2ab = ab(c - 2)$	Factorizamos ab de ambos términos.
$a^2 + 2ab + b^2 = (a + b)^2$	Esta expresión se puede reducir $a + b$ multiplicado por sí mismo.



SOLUCIÓN DE ECUACIONES CON UNA INCÓGNITA

Al resolver ecuaciones con una incógnita, la idea básica es dejar la incógnita de un lado de la ecuación y reducir el otro lado a su menor valor posible. Para esto utilizamos el principio de que la ecuación seguirá siendo una igualdad si todo lo que hagamos en uno de los lados de la ecuación lo hacemos también en el otro lado. Así, por ejemplo, la ecuación sigue siendo una igualdad si sumamos el mismo número a ambos lados. Para resolver una ecuación, modificamos ésta sumando, restando, multiplicando, dividiendo, elevando al cuadrado, etcétera, de modo que la incógnita quede sola en un lado de la ecuación. Esto es válido siempre que realicemos la misma operación en ambos lados de la ecuación, con lo cual se mantiene la igualdad. Los siguientes ejemplos ilustran muchas

de las operaciones que se utilizan comúnmente para resolver ecuaciones con una incógnita. En cada uno de los ejemplos, despejaremos Y en la ecuación.

Ejemplo	Explicación
$Y + 5 = 2$ $Y = 2 - 5$ $= -3$	Para despejar Y , restamos 5 a ambos lados de la ecuación.
$Y - 4 = 6$ $Y = 6 + 4$ $= 10$	Para despejar Y , sumamos 4 a ambos lados de la ecuación.
$\frac{Y}{2} = 8$ $Y = 8(2)$ $= 16$	Para despejar Y , multiplicamos ambos lados de la ecuación por 2.
$3Y = 7$ $Y = \frac{7}{3}$ $= 2.33$	Para despejar Y , dividimos ambos lados de la ecuación entre 3.
$6 = \frac{Y - 3}{2}$ $12 = Y - 3$ $3 + 12 = Y$ $15 = Y$ $Y = 15$	Para despejar Y , 1) multiplicamos ambos lados de la ecuación por 2, y 2) sumamos 3 a ambos lados de la ecuación.
$3 = \frac{2}{Y}$ $3Y = 2$ $Y = \frac{2}{3}$	Para despejar Y , 1) multiplicamos ambos lados de la ecuación por Y , y 2) dividimos ambos lados de la ecuación entre 3.
$4(Y + 1) = 3$ $Y + 1 = \frac{3}{4}$ $Y = \frac{3}{4} - 1$ $= -\frac{1}{4}$	Para despejar Y , 1) dividimos ambos lados de la ecuación entre 4, y 2) restamos 1 de ambos lados de la ecuación.
$\frac{4}{Y + 2} = 8$ $Y + 2 = \frac{1}{8}$ $Y = 2 - \frac{1}{8}$ $= 1\frac{1}{8}$	Para despejar Y , 1) obtenemos el recíproco de ambos lados, 2) multiplicamos ambos lados por 4, y 3) restamos 2 de ambos lados.
$2Y \cdot 4 = 10$ $2Y = 10 \div 4$ $Y = \frac{10 \div 4}{2}$ $= 3$	Para despejar Y , 1) restamos 4 de ambos lados, y 2) dividimos ambos lados entre 2.



INTERPOLACIÓN LINEAL

Al buscar valores en una tabla, con frecuencia es necesario realizar una interpolación lineal. Por ejemplo, suponga que queremos determinar la raíz cuadrada de 96.5 utilizando una tabla que sólo tiene la raíz cuadrada de 96 y 97, pero no la de 96.5, como aquí se muestra:

<i>Número</i>	<i>Raíz cuadrada</i>
96	9.7980
97	9.8489

Al buscar en la columna Número, observamos que no hay un valor correspondiente a 96.5. Los valores más cercanos son 96 y 97. En la tabla vemos que la raíz cuadrada de 96 es 9.7980 y que la raíz cuadrada de 97 es 9.8489. Está claro que la raíz cuadrada de 96.5 debe estar entre 9.7980 y 9.8489. Al utilizar la interpolación *lineal*, suponemos que existe una relación lineal entre el número y su raíz cuadrada, y utilizamos esta relación lineal para aproximar la raíz cuadrada de los números que no aparecen en la tabla. Como quiera que 96.5 está a la mitad del camino entre 96 y 97, cuando se utiliza la interpolación lineal es previsible que la raíz cuadrada de 96.5 esté a la mitad del camino entre 9.7980 y 9.8489. Si X es igual a la raíz cuadrada de 96.5, entonces:

$$X = 9.7980 + 0.5(9.8489 - 9.7980) = 9.8234$$

Aunque no se explicó en forma explícita, el valor calculado para X fue el resultado de plantear la siguiente proporción y despejar X :

$$\frac{96.5 - 96}{97 - 96} = \frac{X - 9.7980}{9.8489 - 9.7980}$$

Número	Raíz cuadrada
96	9.7980
96.5	X
97	9.8489

Esta relación se presenta de manera gráfica en la figura A.1.

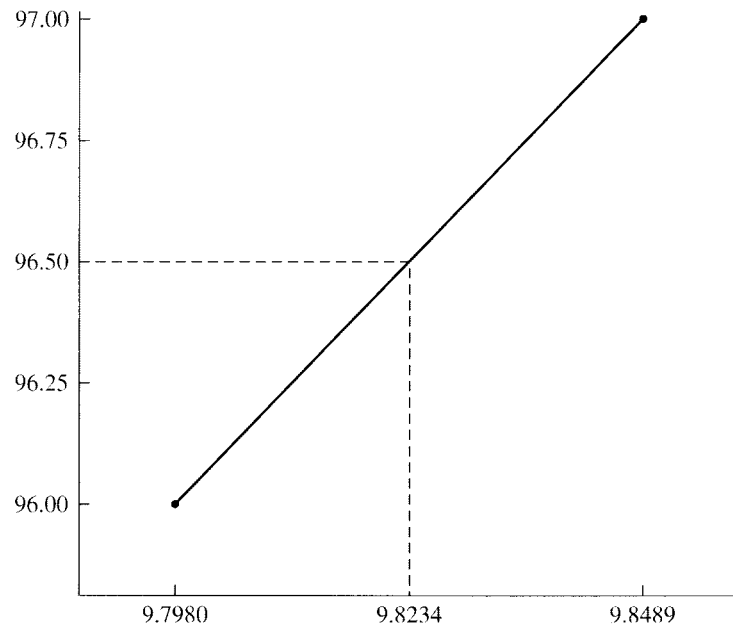


figura A.1 Interpolación lineal para $\sqrt{96.5}$.

Ecuaciones

A continuación presentamos la lista de las ecuaciones utilizadas en este libro.

Ecuación	Descripción	La ecuación aparece por primera vez en la página:
$\sum_{i=1}^N X_i = X_1 + X_2 + X_3 + \dots + X_N$	<i>sumatoria</i>	25
Porcentaje acumulado = $\frac{f \text{ acumulada}}{N} \times 100$	<i>porcentaje acumulativo</i>	46
Punto percentil = $X_{i.} + (i/f_i)(f_r \text{ acumulada} - f_{i.} \text{ acumulada})$	<i>ecuación para el cálculo del punto percentil</i>	49
Rango percentil = $\frac{f_{i.} \text{ acumulada} + (f_i/i)(X - X_{i.})}{N} \times 100$	<i>ecuación para el cálculo del rango percentil</i>	51
$\bar{X} = \frac{\sum X_i}{N}$	<i>media de una muestra</i>	67
$\mu = \frac{\sum X_i}{N}$	<i>media de una población de datos en bruto</i>	67

Ecuación	Descripción	La ecuación aparece por primera vez en la página:
$\bar{X}_{\text{global}} = \frac{n_1\bar{X}_1 + n_2\bar{X}_2 + \dots + n_k\bar{X}_k}{n_1 + n_2 + \dots + n_k}$	media global de varios grupos	71
Mediana = $P_{50} = X_L + (i/f_i)(f_p \text{ acumulada} - f_L \text{ acumulada})$	mediana de una distribución	72
Rango = Dato máximo - Dato mínimo	rango de una distribución	76
$X - \bar{X}$	puntaje de desviación para datos muestrales	77
$X - \mu$	puntaje de desviación para datos poblacionales	77
$s = \sqrt{\frac{SC}{N-1}} = \sqrt{\frac{\sum(X - \bar{X})^2}{N-1}}$	desviación estándar de una muestra de datos en bruto	78
$\sigma = \sqrt{\frac{SC_{\text{pob}}}{N}} = \sqrt{\frac{\sum(X - \mu)^2}{N}}$	desviación estándar de una población de datos en bruto	78
$SC = \sum X^2 - \frac{(\sum X)^2}{N}$	suma de cuadrados	79
$s^2 = \frac{SC}{N-1}$	varianza de una muestra de datos en bruto	82
$\sigma^2 = \frac{SC_{\text{pob}}}{N}$	varianza de una población de datos en bruto	82
$z = \frac{X - \mu}{\sigma}$	puntaje z para datos poblacionales	89
$z = \frac{X - \bar{X}}{s}$	puntaje z para datos muestrales	89
$X = \mu + \sigma z$	ecuación para determinar un dato en bruto de la población a partir de su valor z	97
$Y = bX + a$	ecuación de una línea recta	106
$b = \text{Pendiente} = \frac{\Delta Y}{\Delta X} = \frac{Y_2 - Y_1}{X_2 - X_1}$	pendiente de una línea recta	106
$r = \frac{\sum z_A z_Y}{N-1}$	ecuación para la r de Pearson mediante los valores z	105
$r = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sqrt{\left[\sum X^2 - \frac{(\sum X)^2}{N}\right] \left[\sum Y^2 - \frac{(\sum Y)^2}{N}\right]}}$	ecuación para el cálculo de la r de Pearson	105
$r_s = 1 - \frac{6 \sum D_i^2}{N^3 - N}$	ecuación para el cálculo de la rho de Spearman	102
$Y' = b_Y X + a_Y$	ecuación de regresión lineal para predecir Y dada X	107

La ecuación
aparece
por primera
vez en la
página:

Ecuación	Descripción	
$b_Y = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sum X^2 - \frac{(\sum X)^2}{N}}$	constante de regresión b para predecir Y dada X, ecuación para calcularla con datos en bruto	138
$a_Y = Y - b_Y \bar{X}$	constante de regresión a para predecir Y dada X	138
$X' = b_X Y + a_X$	ecuación de regresión lineal para predecir X dada Y	144
$b_X = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sum Y^2 - \frac{(\sum Y)^2}{N}}$	constante de regresión b para predecir X dada Y, ecuación para calcularla con datos en bruto	145
$a_X = \bar{X} - b_X Y$	constante de regresión a para predecir X dada Y	145
$s_{Y X} = \sqrt{\frac{SC_Y - \frac{[\sum XY - (\sum X)(\sum Y)/N]^2}{SC_X}}{N - 2}}$	ecuación para el cálculo del error estándar de la estimación al predecir Y	147
$b_Y = r \frac{s_Y}{s_X}$	ecuación que relaciona r con la constante de regresión b_Y	151
$b_X = r \frac{s_X}{s_Y}$	ecuación que relaciona r con la constante de regresión b_X	151
$R^2 = \frac{r_{YX_1}^2 + r_{YX_2}^2 - 2r_{YX_1}r_{YX_2}r_{X_1X_2}}{1 - r_{X_1X_2}^2}$	ecuación para el cálculo del cuadrado de la correlación múltiple	153
$p(A) = \frac{\text{Número de eventos clasificables como } A}{\text{Número total de eventos posibles}}$	probabilidad a priori	166
$p(A) = \frac{\text{Número de veces que } A \text{ ha ocurrido}}{\text{Cantidad total de ocurrencias}}$	probabilidad a posteriori	166
$p(A \text{ o } B) = p(A) + p(B) - p(A \text{ y } B)$	ecuación general de la regla de la suma para dos eventos	168
$p(A \text{ o } B) = p(A) + p(B)$	regla de la suma cuando A y B son mutuamente excluyentes	169
$p(A \text{ o } B \text{ o } C \text{ o } \dots \text{ o } Z) = p(A) + p(B) + p(C) + \dots + p(Z)$	regla de la suma para más de dos eventos mutuamente excluyentes	172
$p(A) + p(B) + p(C) + \dots + p(Z) = 1.00$	cuando los eventos son exhaustivos y mutuamente excluyentes	172

Tablas

- Tabla A Áreas bajo la curva normal
 - Tabla B Distribución binomial
 - Tabla C Valores críticos de U y U'
 - Tabla D Valores críticos de la distribución t de Student
 - Tabla E Valores críticos de la r de Pearson
 - Tabla F Valores críticos de la distribución F
 - Tabla G Valores críticos de la distribución con rango de Student (Q)
 - Tabla H Distribución ji-cuadrada (χ^2)
 - Tabla I Valores críticos de T para la prueba de rangos con signo de Wilcoxon
 - Tabla J Números aleatorios
- Agradecimientos

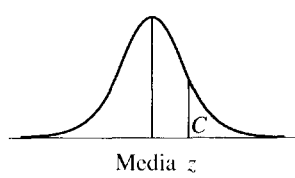
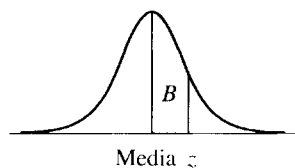


tabla A Áreas bajo la curva normal

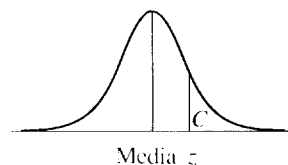
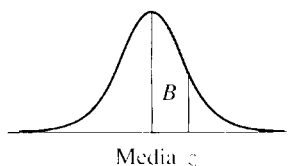
z	Área entre la media y z	Área más allá de z	z	Área entre la media y z	Área más allá de z
A	B	C	A	B	C
0.00	.0000	.5000	0.45	.1736	.3264
0.01	.0040	.4960	0.46	.1772	.3228
0.02	.0080	.4920	0.47	.1808	.3192
0.03	.0120	.4880	0.48	.1844	.3156
0.04	.0160	.4840	0.49	.1879	.3121
0.05	.0199	.4801	0.50	.1915	.3085
0.06	.0239	.4761	0.51	.1950	.3050
0.07	.0279	.4721	0.52	.1985	.3015
0.08	.0319	.4681	0.53	.2019	.2981
0.09	.0359	.4641	0.54	.2054	.2946
0.10	.0398	.4602	0.55	.2088	.2912
0.11	.0438	.4562	0.56	.2123	.2877
0.12	.0478	.4522	0.57	.2157	.2843
0.13	.0517	.4483	0.58	.2190	.2810
0.14	.0557	.4443	0.59	.2224	.2776
0.15	.0596	.4404	0.60	.2257	.2743
0.16	.0636	.4364	0.61	.2291	.2709
0.17	.0675	.4325	0.62	.2324	.2676
0.18	.0714	.4286	0.63	.2357	.2643
0.19	.0753	.4247	0.64	.2389	.2611
0.20	.0793	.4207	0.65	.2422	.2578
0.21	.0832	.4168	0.66	.2454	.2546
0.22	.0871	.4129	0.67	.2486	.2514
0.23	.0910	.4090	0.68	.2517	.2483
0.24	.0948	.4052	0.69	.2549	.2451
0.25	.0987	.4013	0.70	.2580	.2420
0.26	.1026	.3974	0.71	.2611	.2389
0.27	.1064	.3936	0.72	.2642	.2358
0.28	.1103	.3897	0.73	.2673	.2327
0.29	.1141	.3859	0.74	.2704	.2296
0.30	.1179	.3821	0.75	.2734	.2266
0.31	.1217	.3783	0.76	.2764	.2236
0.32	.1255	.3745	0.77	.2794	.2206
0.33	.1293	.3707	0.78	.2823	.2177
0.34	.1331	.3669	0.79	.2852	.2148
0.35	.1368	.3632	0.80	.2881	.2119
0.36	.1406	.3594	0.81	.2910	.2090
0.37	.1443	.3557	0.82	.2939	.2061
0.38	.1480	.3520	0.83	.2967	.2033
0.39	.1517	.3483	0.84	.2995	.2005
0.40	.1554	.3446	0.85	.3023	.1977
0.41	.1591	.3409	0.86	.3051	.1949
0.42	.1628	.3372	0.87	.3078	.1922
0.43	.1664	.3336	0.88	.3106	.1894
0.44	.1700	.3300	0.89	.3133	.1867

La columna A proporcione el valor z positivo.

La columna B proporciona el área entre la media y z . Como quiera que la curva es simétrica, las áreas para los valores z negativos son iguales a las áreas para los valores z positivos.

La columna C proporciona el área que está más allá de z .

(continúa)



La columna A proporcione el valor z positivo.

La columna B proporciona el área entre la media y z . Como quiera que la curva es simétrica, las áreas para los valores z negativos son iguales a las áreas para los valores z positivos.

La columna C proporciona el área que está más allá de z .

tabla A (continuación)

z	Área entre la media y z	Área más allá de z	z	Área entre la media y z	Área más allá de z
A	B	C	A	B	C
0.90	.3159	.1841	1.35	.4115	.0885
0.91	.3186	.1814	1.36	.4131	.0869
0.92	.3212	.1788	1.37	.4147	.0853
0.93	.3238	.1762	1.38	.4162	.0838
0.94	.3264	.1736	1.39	.4177	.0823
0.95	.3289	.1711	1.40	.4192	.0808
0.96	.3315	.1685	1.41	.4207	.0793
0.97	.3340	.1660	1.42	.4222	.0778
0.98	.3365	.1635	1.43	.4236	.0764
0.99	.3389	.1611	1.44	.4251	.0749
1.00	.3413	.1587	1.45	.4265	.0735
1.01	.3438	.1562	1.46	.4279	.0721
1.02	.3461	.1539	1.47	.4292	.0708
1.03	.3485	.1515	1.48	.4306	.0694
1.04	.3508	.1492	1.49	.4319	.0681
1.05	.3531	.1469	1.50	.4332	.0668
1.06	.3554	.1446	1.51	.4345	.0655
1.07	.3577	.1423	1.52	.4357	.0643
1.08	.3599	.1401	1.53	.4370	.0630
1.09	.3621	.1379	1.54	.4382	.0618
1.10	.3643	.1357	1.55	.4394	.0606
1.11	.3665	.1335	1.56	.4406	.0594
1.12	.3686	.1314	1.57	.4418	.0582
1.13	.3708	.1292	1.58	.4429	.0571
1.14	.3729	.1271	1.59	.4441	.0559
1.15	.3749	.1251	1.60	.4452	.0548
1.16	.3770	.1230	1.61	.4463	.0537
1.17	.3790	.1210	1.62	.4474	.0526
1.18	.3810	.1190	1.63	.4484	.0516
1.19	.3830	.1170	1.64	.4495	.0505
1.20	.3849	.1151	1.65	.4505	.0495
1.21	.3869	.1131	1.66	.4515	.0485
1.22	.3888	.1112	1.67	.4525	.0475
1.23	.3907	.1093	1.68	.4535	.0465
1.24	.3925	.1075	1.69	.4545	.0455
1.25	.3944	.1056	1.70	.4554	.0446
1.26	.3962	.1038	1.71	.4564	.0436
1.27	.3980	.1020	1.72	.4573	.0427
1.28	.3997	.1003	1.73	.4582	.0418
1.29	.4015	.0985	1.74	.4591	.0409
1.30	.4032	.0968	1.75	.4599	.0401
1.31	.4049	.0951	1.76	.4608	.0392
1.32	.4066	.0934	1.77	.4616	.0384
1.33	.4082	.0918	1.78	.4625	.0375
1.34	.4099	.0901	1.79	.4633	.0367

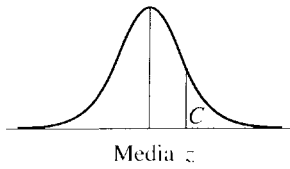
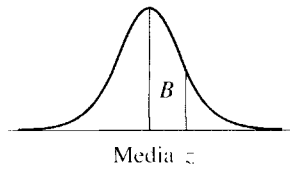


tabla A Áreas bajo la curva normal (continuación) (continuación)

z	Área entre la media y z	Área más allá de z	z	Área entre la media y z	Área más allá de z
A	B	C	A	B	C
1.80	.4641	.0359	2.25	.4878	.0122
1.81	.4649	.0351	2.26	.4881	.0119
1.82	.4656	.0344	2.27	.4884	.0116
1.83	.4664	.0336	2.28	.4887	.0113
1.84	.4671	.0329	2.29	.4890	.0110
1.85	.4678	.0322	2.30	.4893	.0107
1.86	.4686	.0314	2.31	.4896	.0104
1.87	.4693	.0307	2.32	.4898	.0102
1.88	.4699	.0301	2.33	.4901	.0099
1.89	.4706	.0294	2.34	.4904	.0096
1.90	.4713	.0287	2.35	.4906	.0094
1.91	.4719	.0281	2.36	.4909	.0091
1.92	.4726	.0274	2.37	.4911	.0089
1.93	.4732	.0268	2.38	.4913	.0087
1.94	.4738	.0262	2.39	.4916	.0084
1.95	.4744	.0256	2.40	.4918	.0082
1.96	.4750	.0250	2.41	.4920	.0080
1.97	.4756	.0244	2.42	.4922	.0078
1.98	.4761	.0239	2.43	.4925	.0075
1.99	.4767	.0233	2.44	.4927	.0073
2.00	.4772	.0228	2.45	.4929	.0071
2.01	.4778	.0222	2.46	.4931	.0069
2.02	.4783	.0217	2.47	.4932	.0068
2.03	.4788	.0212	2.48	.4934	.0066
2.04	.4793	.0207	2.49	.4936	.0064
2.05	.4798	.0202	2.50	.4938	.0062
2.06	.4803	.0197	2.51	.4940	.0060
2.07	.4808	.0192	2.52	.4941	.0059
2.08	.4812	.0188	2.53	.4943	.0057
2.09	.4817	.0183	2.54	.4945	.0055
2.10	.4821	.0179	2.55	.4946	.0054
2.11	.4826	.0174	2.56	.4948	.0052
2.12	.4830	.0170	2.57	.4949	.0051
2.13	.4834	.0166	2.58	.4951	.0049
2.14	.4838	.0162	2.59	.4952	.0048
2.15	.4842	.0158	2.60	.4953	.0047
2.16	.4846	.0154	2.61	.4955	.0045
2.17	.4850	.0150	2.62	.4956	.0044
2.18	.4854	.0146	2.63	.4957	.0043
2.19	.4857	.0143	2.64	.4959	.0041
2.20	.4861	.0139	2.65	.4960	.0040
2.21	.4864	.0136	2.66	.4961	.0039
2.22	.4868	.0132	2.67	.4962	.0038
2.23	.4871	.0129	2.68	.4963	.0037
2.24	.4875	.0125	2.69	.4964	.0036

La columna A proporcione el valor z positivo.

La columna B proporciona el área entre la media y z . Como quiera que la curva es simétrica, las áreas para los valores z negativos son iguales a las áreas para los valores z positivos.

La columna C proporciona el área que está más allá de z .

(continued)

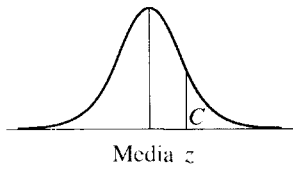
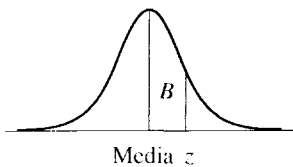


tabla A (continuación)

<i>z</i>	Área entre la media y <i>z</i>	Área más allá de <i>z</i>	<i>z</i>	Área entre la media y <i>z</i>	Área más allá de <i>z</i>
A	B	C	A	B	C
2.70	.4965	.0035	3.00	.4987	.0013
2.71	.4966	.0034	3.01	.4987	.0013
2.72	.4967	.0033	3.02	.4987	.0013
2.73	.4968	.0032	3.03	.4988	.0012
2.74	.4969	.0031	3.04	.4988	.0012
2.75	.4970	.0030	3.05	.4989	.0011
2.76	.4971	.0029	3.06	.4989	.0011
2.77	.4972	.0028	3.07	.4989	.0011
2.78	.4973	.0027	3.08	.4990	.0010
2.79	.4974	.0026	3.09	.4990	.0010
2.80	.4974	.0026	3.10	.4990	.0010
2.81	.4975	.0025	3.11	.4991	.0009
2.82	.4976	.0024	3.12	.4991	.0009
2.83	.4977	.0023	3.13	.4991	.0009
2.84	.4977	.0023	3.14	.4992	.0008
2.85	.4978	.0022	3.15	.4992	.0008
2.86	.4979	.0021	3.16	.4992	.0008
2.87	.4979	.0021	3.17	.4992	.0008
2.88	.4980	.0020	3.18	.4993	.0007
2.89	.4981	.0019	3.19	.4993	.0007
2.90	.4981	.0019	3.20	.4993	.0007
2.91	.4982	.0018	3.21	.4993	.0007
2.92	.4982	.0018	3.22	.4994	.0006
2.93	.4983	.0017	3.23	.4994	.0006
2.94	.4984	.0016	3.24	.4994	.0006
2.95	.4984	.0016	3.30	.4995	.0005
2.96	.4985	.0015	3.40	.4997	.0003
2.97	.4985	.0015	3.50	.4998	.0002
2.98	.4986	.0014	3.60	.4998	.0002
2.99	.4986	.0014	3.70	.4999	.0001

La columna A proporcione el valor *z* positivo.

La columna B proporciona el área entre la media y *z*. Como quiera que la curva es simétrica, las áreas para los valores *z* negativos son iguales a las áreas para los valores *z* positivos.

La columna C proporciona el área que está más allá de *z*.

tabla B Distribución binomial

N	Número de eventos P o Q	P o Q									
		.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
1	0	.9500	.9000	.8500	.8000	.7500	.7000	.6500	.6000	.5500	.5000
	1	.0500	.1000	.1500	.2000	.2500	.3000	.3500	.4000	.4500	.5000
2	0	.9025	.8100	.7225	.6400	.5625	.4900	.4225	.3600	.3025	.2500
	1	.0950	.1800	.2550	.3200	.3750	.4200	.4550	.4800	.4950	.5000
	2	.0025	.0100	.0225	.0400	.0625	.0900	.1225	.1600	.2025	.2500
3	0	.8574	.7290	.6141	.5120	.4219	.3430	.2746	.2160	.1664	.1250
	1	.1354	.2430	.3251	.3840	.4219	.4410	.4436	.4320	.4084	.3750
	2	.0071	.0270	.0574	.0960	.1406	.1890	.2389	.2880	.3341	.3750
	3	.0001	.0010	.0034	.0080	.0156	.0270	.0429	.0640	.0911	.1250
4	0	.8145	.6561	.5220	.4096	.3164	.2401	.1785	.1296	.0915	.0625
	1	.1715	.2916	.3685	.4096	.4219	.4116	.3845	.3456	.2995	.2500
	2	.0135	.0486	.0975	.1536	.2109	.2646	.3105	.3456	.3675	.3750
	3	.0005	.0036	.0115	.0256	.0469	.0756	.1115	.1536	.2005	.2500
	4	.0000	.0001	.0005	.0016	.0039	.0081	.0150	.0256	.0410	.0625
5	0	.7738	.5905	.4437	.3277	.2373	.1681	.1160	.0778	.0503	.0312
	1	.2036	.3280	.3915	.4096	.3955	.3602	.3124	.2592	.2059	.1562
	2	.0214	.0729	.1382	.2048	.2637	.3087	.3364	.3456	.3369	.3125
	3	.0011	.0081	.0244	.0512	.0879	.1323	.1811	.2304	.2757	.3125
	4	.0000	.0004	.0022	.0064	.0146	.0284	.0488	.0768	.1128	.1562
	5	.0000	.0000	.0001	.0003	.0010	.0024	.0053	.0102	.0185	.0312
6	0	.7351	.5314	.3771	.2621	.1780	.1176	.0754	.0467	.0277	.0156
	1	.2321	.3543	.3993	.3932	.3560	.3025	.2437	.1866	.1359	.0938
	2	.0305	.0984	.1762	.2458	.2966	.3241	.3280	.3110	.2780	.2344
	3	.0021	.0146	.0415	.0819	.1318	.1852	.2355	.2765	.3032	.3125
	4	.0001	.0012	.0055	.0154	.0330	.0595	.0951	.1382	.1861	.2344
	5	.0000	.0001	.0004	.0015	.0044	.0102	.0205	.0369	.0609	.0938
	6	.0000	.0000	.0000	.0001	.0002	.0007	.0018	.0041	.0083	.0156
7	0	.6983	.4783	.3206	.2097	.1335	.0824	.0490	.0280	.0152	.0078
	1	.2573	.3720	.3960	.3670	.3115	.2471	.1848	.1306	.0872	.0547
	2	.0406	.1240	.2097	.2753	.3115	.3177	.2985	.2613	.2140	.1641
	3	.0036	.0230	.0617	.1147	.1730	.2269	.2679	.2903	.2918	.2734
	4	.0002	.0026	.0109	.0287	.0577	.0972	.1442	.1935	.2388	.2734
	5	.0000	.0002	.0012	.0043	.0115	.0250	.0466	.0774	.1172	.1641
	6	.0000	.0000	.0001	.0004	.0013	.0036	.0084	.0172	.0320	.0547
	7	.0000	.0000	.0000	.0000	.0001	.0002	.0006	.0016	.0037	.0078
8	0	.6634	.4305	.2725	.1678	.1001	.0576	.0319	.0168	.0084	.0039
	1	.2793	.3826	.3847	.3355	.2670	.1977	.1373	.0896	.0548	.0312
	2	.0515	.1488	.2376	.2936	.3115	.2965	.2587	.2090	.1569	.1094
	3	.0054	.0331	.0839	.1468	.2076	.2541	.2786	.2787	.2568	.2188
	4	.0004	.0046	.0185	.0459	.0865	.1361	.1875	.2322	.2627	.2734
	5	.0000	.0004	.0026	.0092	.0231	.0467	.0808	.1239	.1719	.2188
	6	.0000	.0000	.0002	.0011	.0038	.0100	.0217	.0413	.0703	.1094
	7	.0000	.0000	.0000	.0001	.0004	.0012	.0033	.0079	.0164	.0312
	8	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0007	.0017	.0039

(continúa)

tabla B Distribución binomial (continuación)

N	Número de eventos P o Q	P o Q									
		.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
13	0	.5133	.2542	.1209	.0550	.0238	.0097	.0037	.0013	.0004	.0001
	1	.3512	.3672	.2774	.1787	.1029	.0540	.0259	.0113	.0045	.0016
	2	.1109	.2448	.2937	.2680	.2059	.1388	.0836	.0453	.0220	.0095
	3	.0214	.0997	.1900	.2457	.2517	.2181	.1651	.1107	.0660	.0349
	4	.0028	.0277	.0838	.1535	.2097	.2337	.2222	.1845	.1350	.0873
	5	.0003	.0055	.0266	.0691	.1258	.1803	.2154	.2214	.1989	.1571
	6	.0000	.0008	.0063	.0230	.0559	.1030	.1546	.1968	.2169	.2095
	7	.0000	.0001	.0011	.0058	.0186	.0442	.0833	.1312	.1775	.2095
	8	.0000	.0000	.0001	.0011	.0047	.0142	.0336	.0656	.1089	.1571
	9	.0000	.0000	.0000	.0001	.0009	.0034	.0101	.0243	.0495	.0873
	10	.0000	.0000	.0000	.0000	.0001	.0006	.0022	.0065	.0162	.0349
	11	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0012	.0036	.0095
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0016
13	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	
14	0	.4877	.2288	.1028	.0440	.0178	.0068	.0024	.0008	.0002	.0001
	1	.3593	.3559	.2539	.1539	.0832	.0407	.0181	.0073	.0027	.0009
	2	.1229	.2570	.2912	.2501	.1802	.1134	.0634	.0317	.0141	.0056
	3	.0259	.1142	.2056	.2501	.2402	.1943	.1366	.0845	.0462	.0222
	4	.0037	.0349	.0998	.1720	.2202	.2290	.2022	.1549	.1040	.0611
	5	.0004	.0078	.0352	.0860	.1468	.1963	.2178	.2066	.1701	.1222
	6	.0000	.0013	.0093	.0322	.0734	.1262	.1759	.2066	.2088	.1833
	7	.0000	.0002	.0019	.0092	.0280	.0618	.1082	.1574	.1952	.2095
	8	.0000	.0000	.0003	.0020	.0082	.0232	.0510	.0918	.1398	.1833
	9	.0000	.0000	.0000	.0003	.0018	.0066	.0183	.0408	.0762	.1222
	10	.0000	.0000	.0000	.0000	.0003	.0014	.0049	.0136	.0312	.0611
	11	.0000	.0000	.0000	.0000	.0000	.0002	.0010	.0033	.0093	.0222
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0019	.0056
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0009
14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	
15	0	.4633	.2059	.0874	.0352	.0134	.0047	.0016	.0005	.0001	.0000
	1	.3658	.3432	.2312	.1319	.0668	.0305	.0126	.0047	.0016	.0005
	2	.1348	.2669	.2856	.2309	.1559	.0916	.0476	.0219	.0090	.0032
	3	.0307	.1285	.2184	.2501	.2252	.1700	.1110	.0634	.0318	.0139
	4	.0049	.0428	.1156	.1876	.2252	.2186	.1792	.1268	.0780	.0417
	5	.0006	.0105	.0449	.1032	.1651	.2061	.2123	.1859	.1404	.0916
	6	.0000	.0019	.0132	.0430	.0917	.1472	.1906	.2066	.1914	.1527
	7	.0000	.0003	.0030	.0138	.0393	.0811	.1319	.1771	.2013	.1964
	8	.0000	.0000	.0005	.0035	.0131	.0348	.0710	.1181	.1647	.1964
	9	.0000	.0000	.0001	.0007	.0034	.0116	.0298	.0612	.1048	.1527
	10	.0000	.0000	.0000	.0001	.0007	.0030	.0096	.0245	.0515	.0916
	11	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0074	.0191	.0417
	12	.0000	.0000	.0000	.0000	.0000	.0001	.0004	.0016	.0052	.0139
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010	.0032
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005
15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	

(continúa)

tabla B (continuación)

N	Número de eventos <i>P o Q</i>	<i>P o Q</i>									
		.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
16	0	.4401	.1853	.0743	.0281	.0100	.0033	.0010	.0003	.0001	.0000
	1	.3706	.3294	.2097	.1126	.0535	.0228	.0087	.0030	.0009	.0002
	2	.1463	.2745	.2775	.2111	.1336	.0732	.0353	.0150	.0056	.0018
	3	.0359	.1423	.2285	.2463	.2079	.1465	.0888	.0468	.0215	.0085
	4	.0061	.0514	.1311	.2001	.2252	.2040	.1553	.1014	.0572	.0278
	5	.0008	.0137	.0555	.1201	.1802	.2099	.2008	.1623	.1123	.0667
	6	.0001	.0028	.0180	.0550	.1101	.1649	.1982	.1983	.1684	.1222
	7	.0000	.0004	.0045	.0197	.0524	.1010	.1524	.1889	.1969	.1746
	8	.0000	.0001	.0009	.0055	.0197	.0487	.0923	.1417	.1812	.1964
	9	.0000	.0000	.0001	.0012	.0058	.0185	.0442	.0840	.1318	.1746
	10	.0000	.0000	.0000	.0002	.0014	.0056	.0167	.0392	.0755	.1222
	11	.0000	.0000	.0000	.0000	.0002	.0013	.0049	.0142	.0337	.0667
	12	.0000	.0000	.0000	.0000	.0000	.0002	.0011	.0040	.0115	.0278
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0008	.0029	.0085
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0018
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002
16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	
17	0	.4181	.1668	.0631	.0225	.0075	.0023	.0007	.0002	.0000	.0000
	1	.3741	.3150	.1893	.0957	.0426	.0169	.0060	.0019	.0005	.0001
	2	.1575	.2800	.2673	.1914	.1136	.0581	.0260	.0102	.0035	.0010
	3	.0415	.1556	.2359	.2393	.1893	.1245	.0701	.0341	.0144	.0052
	4	.0076	.0605	.1457	.2093	.2209	.1868	.1320	.0796	.0411	.0182
	5	.0010	.0175	.0668	.1361	.1914	.2081	.1849	.1379	.0875	.0472
	6	.0001	.0039	.0236	.0680	.1276	.1784	.1991	.1839	.1432	.0944
	7	.0000	.0007	.0065	.0267	.0668	.1201	.1685	.1927	.1841	.1484
	8	.0000	.0001	.0014	.0084	.0279	.0644	.1143	.1606	.1883	.1855
	9	.0000	.0000	.0003	.0021	.0093	.0276	.0611	.1070	.1540	.1855
	10	.0000	.0000	.0000	.0004	.0025	.0095	.0263	.0571	.1008	.1484
	11	.0000	.0000	.0000	.0001	.0005	.0026	.0090	.0242	.0525	.0944
	12	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0081	.0215	.0472
	13	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0021	.0068	.0182
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0004	.0016	.0052
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	
18	0	.3972	.1501	.0536	.0180	.0056	.0016	.0004	.0001	.0000	.0000
	1	.3763	.3002	.1704	.0811	.0338	.0126	.0042	.0012	.0003	.0001
	2	.1683	.2835	.2556	.1723	.0958	.0458	.0190	.0069	.0022	.0006
	3	.0473	.1680	.2406	.2297	.1704	.1046	.0547	.0246	.0095	.0031
	4	.0093	.0070	.1592	.2153	.2130	.1681	.1104	.0614	.0291	.0117
	5	.0014	.0218	.0787	.1507	.1988	.2017	.1664	.1146	.0666	.0327
	6	.0002	.0052	.0310	.0816	.1436	.1873	.1941	.1655	.1181	.0708
	7	.0000	.0010	.0091	.0350	.0820	.1376	.1792	.1892	.1657	.1214
	8	.0000	.0002	.0022	.0120	.0376	.0811	.1327	.1734	.1864	.1669
	9	.0000	.0000	.0004	.0033	.0139	.0386	.0794	.1284	.1694	.1855
	10	.0000	.0000	.0001	.0008	.0042	.0149	.0385	.0771	.1248	.1669
11	.0000	.0000	.0000	.0001	.0010	.0046	.0151	.0374	.0742	.1214	

tabla C.1 Valores críticos de U y U' para una prueba de una cola en $\alpha = 0.005$ o para una prueba de dos colas en $\alpha = 0.01$

Para que sea significativo, dados n_1 y n_2 : U_{obs} deberá ser igual o **menor que** el valor que se presenta en la tabla. U'_{obs} deberá ser igual o **mayor que** el valor que se presenta en la tabla.

$n_1 \backslash n_2$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0
3	—	—	—	—	—	—	—	—	0	0	0	1	1	1	2	2	2	2	3	3
4	—	—	—	—	—	0	0	1	1	2	2	3	3	4	5	5	6	6	7	8
5	—	—	—	—	0	1	1	2	3	4	5	6	7	7	8	9	10	11	12	13
6	—	—	—	0	1	2	3	4	5	6	7	9	10	11	12	13	15	16	17	18
7	—	—	—	0	1	3	4	6	7	9	10	12	13	15	16	18	19	21	22	24
8	—	—	—	1	2	4	6	7	9	11	13	15	17	18	20	22	24	26	28	30
9	—	—	0	1	3	5	7	9	11	13	16	18	20	22	24	27	29	31	33	36
10	—	—	0	2	4	6	9	11	13	16	18	21	24	26	29	31	34	37	39	42
11	—	—	0	2	5	7	10	13	16	18	21	24	27	30	33	36	39	42	45	48
12	—	—	1	3	6	9	12	15	18	21	24	27	31	34	37	41	44	47	51	54
13	—	—	1	3	7	10	13	17	20	24	27	31	34	38	42	45	49	53	56	60
14	—	—	1	4	7	11	15	18	22	26	30	34	38	42	46	50	54	58	63	67
15	—	—	2	5	8	12	16	20	24	29	33	37	42	46	51	55	60	64	69	73
16	—	—	2	5	9	13	18	22	27	31	36	41	45	50	55	60	65	70	74	79
17	—	—	2	6	10	15	19	24	29	34	39	44	49	54	60	65	70	75	81	86
18	—	—	2	6	11	16	21	26	31	37	42	47	53	58	64	70	75	81	87	92
19	—	0	3	7	12	17	22	28	33	39	45	51	56	63	69	74	81	87	93	99
20	—	0	3	8	13	18	24	30	36	42	48	54	60	67	73	79	86	92	99	105

Los guiones en la tabla indican que no se puede tomar una decisión en el nivel de significación indicado.

tabla C.2 Valores críticos de U y U' para una prueba de una cola en $\alpha = 0.001$ o para una prueba de dos colas en $\alpha = 0.02$

Para que sea significativo, dados n_1 y n_2 : U_{obt} deberá ser igual o **menor que** el valor que se presenta en la tabla. U'_{obt} deberá ser igual o **mayor que** el valor que se presenta en la tabla.

$n_1 \backslash n_2$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
2	—	—	—	—	—	—	—	—	—	—	—	—	0	0	0	0	0	0	1	1
3	—	—	—	—	—	—	0	0	1	1	1	2	2	2	3	3	4	4	4	5
4	—	—	—	—	0	1	1	2	3	3	4	5	5	6	7	7	8	9	9	10
5	—	—	—	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
6	—	—	—	1	2	3	4	6	7	8	9	11	12	13	15	16	18	19	20	22
7	—	—	0	1	3	4	6	7	9	11	12	14	16	17	19	21	23	24	26	28
8	—	—	0	2	4	6	7	9	11	13	15	17	20	22	24	26	28	30	32	34
9	—	—	1	3	5	7	9	11	14	16	18	21	23	26	28	31	33	36	38	40
10	—	—	1	3	6	8	11	13	16	19	22	24	27	30	33	36	38	41	44	47
11	—	—	1	4	7	9	12	15	18	22	25	28	31	34	37	41	44	47	50	53
12	—	—	2	5	8	11	14	17	21	24	28	31	35	38	42	46	49	53	56	60
13	—	0	2	5	9	12	16	20	23	27	31	35	39	43	47	51	55	59	63	67
14	—	0	2	6	10	13	17	22	26	30	34	38	43	47	51	56	60	65	69	73
15	—	0	3	7	11	15	19	24	28	33	37	42	47	51	56	61	66	70	75	80
16	—	0	3	7	12	16	21	26	31	36	41	46	51	56	61	66	71	76	82	87
17	—	0	4	8	13	18	23	28	33	38	44	49	55	60	66	71	77	82	88	93
18	—	0	4	9	14	19	24	30	36	41	47	53	59	65	70	76	82	88	94	100
19	—	1	4	9	15	20	26	32	38	44	50	56	63	69	75	82	88	94	101	107
20	—	1	5	10	16	22	28	34	40	47	53	60	67	73	80	87	93	100	107	114

Los guiones en la tabla indican que no se puede tomar una decisión en el nivel de significación indicado.

tabla C.3 Valores críticos de U y U' para una prueba de una cola en $\alpha = 0.025$ o para una prueba de dos colas en $\alpha = 0.05$

Para que sea significativo, dados n_1 y n_2 : U_{obt} deberá ser igual o **menor que** el valor que se presenta en la tabla. U'_{obt} deberá ser igual o **mayor que** el valor que se presenta en la tabla.

$n_1 \backslash n_2$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
2	—	—	—	—	—	—	—	0	0	0	0	1	1	1	1	1	2	2	2	2
3	—	—	—	—	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
4	—	—	—	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	13
5	—	—	0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6	—	—	1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7	—	—	1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
8	—	0	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
9	—	0	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
10	—	0	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	55
11	—	0	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	62
12	—	1	4	7	11	14	18	22	26	29	33	37	41	45	49	53	57	61	65	69
13	—	1	4	8	12	16	20	24	28	33	37	41	45	50	54	59	63	67	72	76
14	—	1	5	9	13	17	22	26	31	36	40	45	50	55	59	64	67	74	78	83
15	—	1	5	10	14	19	24	29	34	39	44	49	54	59	64	70	75	80	85	90
16	—	1	6	11	15	21	26	31	37	42	47	53	59	64	70	75	81	86	92	98
17	—	2	6	11	17	22	28	34	39	45	51	57	63	67	75	81	87	93	99	105
18	—	2	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99	106	112
19	—	2	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	119
20	—	2	8	13	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127

Los guiones en la tabla indican que no se puede tomar una decisión en el nivel de significación indicado.

tabla C.4 Valores críticos de U y U' para una prueba de una cola en $\alpha = 0.05$ o para una prueba de dos colas en $\alpha = 0.10$

Para que sea significativo, dados n_1 y n_2 : U_{obt} deberá ser igual o **menor que** el valor que se presenta en la tabla. U'_{obt} deberá ser igual o **mayor que** el valor que se presenta en la tabla.

$n_1 \backslash n_2$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0	0
																			19	20
2	—	—	—	—	0	0	0	1	1	1	1	2	2	2	3	3	3	4	4	4
					10	12	14	15	17	19	21	22	24	26	27	29	31	32	34	36
3	—	—	0	0	1	2	2	3	3	4	5	5	6	7	7	8	9	9	10	11
			9	12	14	16	19	21	24	26	28	31	33	35	38	40	42	45	47	49
4	—	—	0	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17	18
			12	15	18	21	24	27	30	33	36	39	42	45	48	50	53	56	59	62
5	—	0	1	2	4	5	6	8	9	11	12	13	15	16	18	19	20	22	23	25
		10	14	18	21	25	29	32	36	39	43	47	50	54	57	61	65	68	72	75
6	—	0	2	3	5	7	8	10	12	14	16	17	19	21	23	25	26	28	30	32
		12	16	21	25	29	34	38	42	46	50	55	59	63	67	71	76	80	84	88
7	—	0	2	4	6	8	11	13	15	17	19	21	24	26	28	30	33	35	37	39
		14	19	24	29	34	38	43	48	53	58	63	67	72	77	82	86	91	96	101
8	—	1	3	5	8	10	13	15	18	20	23	26	28	31	33	36	39	41	44	47
		15	21	27	32	38	43	49	54	60	65	70	76	81	87	92	97	103	108	113
9	—	1	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54
		17	24	30	36	42	48	54	60	66	72	78	84	90	96	102	108	114	120	126
10	—	1	4	7	11	14	17	20	24	27	31	34	37	41	44	48	51	55	58	62
		19	26	33	39	46	53	60	66	73	79	86	93	99	106	112	119	125	132	138
11	—	1	5	8	12	16	19	23	27	31	34	38	42	46	50	54	57	61	65	69
		21	28	36	43	50	58	65	72	79	87	94	101	108	115	122	130	137	144	151
12	—	2	5	9	13	17	21	26	30	34	38	42	47	51	55	60	64	68	72	77
		22	31	39	47	55	63	70	78	86	94	102	109	117	125	132	140	148	156	163
13	—	2	6	10	15	19	24	28	33	37	42	47	51	56	61	65	70	75	80	84
		24	33	42	50	59	67	76	84	93	101	109	118	126	134	143	151	159	167	176
14	—	2	7	11	16	21	26	31	36	41	46	51	56	61	66	71	77	82	87	92
		26	35	45	54	63	72	81	90	99	108	117	126	135	144	153	161	170	179	188
15	—	3	7	12	18	23	28	33	39	44	50	55	61	66	72	77	83	88	94	100
		27	38	48	57	67	77	87	96	106	115	125	134	144	153	163	172	182	191	200
16	—	3	8	14	19	25	30	36	42	48	54	60	65	71	77	83	89	95	101	107
		29	40	50	61	71	82	92	102	112	122	132	143	153	163	173	183	193	203	213
17	—	3	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96	102	109	115
		31	42	53	65	76	86	97	108	119	130	140	151	161	172	183	193	204	214	225
18	—	4	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109	116	123
		32	45	56	68	80	91	103	114	123	137	148	159	170	182	193	204	215	226	237
19	0	4	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	130
	19	34	47	59	72	84	96	108	120	132	144	156	167	179	191	203	214	226	238	250
20	0	4	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138
	20	36	49	62	75	88	101	113	126	138	151	163	176	188	200	213	225	237	250	262

Los guiones en la tabla indican que no se puede tomar una decisión en el nivel de significación indicado.

tabla D Valores críticos de la distribución *t* de Student

Los valores que se presentan en la tabla son los valores críticos de *t* para los grados de libertad dados (columna de la izquierda) y el nivel alfa (encabezado de la columna). Para los niveles alfa con dos colas, t_{crit} es + y -. Para que sea significativo, $|t_{obt}| \geq |t_{crit}|$.

df	Nivel de significación para la prueba con una cola					
	.10	.05	.025	.01	.005	.0005
	Nivel de significación para la prueba con dos colas					
	.20	.10	.05	.02	.01	.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.598
3	1.638	2.353	3.182	4.541	5.841	12.941
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.859
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.405
8	1.397	1.860	2.306	2.986	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587
11	1.363	1.796	2.201	2.718	3.106	4.437
12	1.356	1.782	2.179	2.681	3.055	4.318
13	1.350	1.771	2.160	2.650	3.012	4.221
14	1.345	1.761	2.145	2.624	2.977	4.140
15	1.341	1.753	2.131	2.602	2.947	4.073
16	1.337	1.746	2.120	2.583	2.921	4.015
17	1.333	1.740	2.110	2.567	2.898	3.965
18	1.330	1.734	2.101	2.552	2.878	3.922
19	1.328	1.729	2.093	2.539	2.861	3.883
20	1.325	1.725	2.086	2.528	2.845	3.850
21	1.323	1.721	2.080	2.518	2.831	3.819
22	1.321	1.717	2.074	2.508	2.819	3.792
23	1.319	1.714	2.069	2.500	2.807	3.767
24	1.318	1.711	2.064	2.492	2.797	3.745
25	1.316	1.708	2.060	2.485	2.787	3.725
26	1.315	1.706	2.056	2.479	2.779	3.707
27	1.314	1.703	2.052	2.473	2.771	3.690
28	1.313	1.701	2.048	2.467	2.763	3.674
29	1.311	1.699	2.045	2.462	2.756	3.659
30	1.310	1.697	2.042	2.457	2.750	3.646
40	1.303	1.684	2.021	2.423	2.704	3.551
60	1.296	1.671	2.000	2.390	2.660	3.460
120	1.289	1.658	1.980	2.358	2.617	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.291

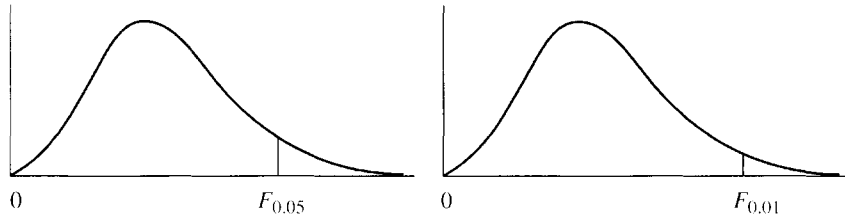
table E Valores críticos de la r de Pearson

Los valores que se presentan en la tabla son los valores críticos de r para los grados de libertad dados (columna de la izquierda) y el nivel alfa (encabezado de la columna). Para los niveles alfa con dos colas, t_{crit} es $+ y -$. Para que sea significativo, $|r_{\text{obt}}| \geq |r_{\text{crit}}|$.

$gl = N - 2$	Nivel de significación para la prueba con una cola				
	.05	.025	.01	.005	.0005
	Nivel de significación para la prueba con dos colas				
	.10	.05	.02	.01	.001
1	.9877	.9969	.9995	.9999	1.0000
2	.9000	.9500	.9800	.9900	.9990
3	.8054	.8783	.9343	.9587	.9912
4	.7293	.8114	.8822	.9172	.9741
5	.6694	.7545	.8329	.8745	.9507
6	.6215	.7067	.7887	.8343	.9249
7	.5822	.6664	.7498	.7977	.8982
8	.5494	.6319	.7155	.7646	.8721
9	.5214	.6021	.6851	.7348	.8471
10	.4973	.5760	.6581	.7079	.8233
11	.4762	.5529	.6339	.6835	.8010
12	.4575	.5324	.6120	.6614	.7800
13	.4409	.5139	.5923	.6411	.7603
14	.4259	.4973	.5742	.6226	.7420
15	.4124	.4821	.5577	.6055	.7246
16	.4000	.4683	.5425	.5897	.7084
17	.3887	.4555	.5285	.5751	.6932
18	.3783	.4438	.5155	.5614	.6787
19	.3687	.4329	.5034	.5487	.6652
20	.3598	.4227	.4921	.5368	.6524
25	.3233	.3809	.4451	.4869	.5974
30	.2960	.3494	.4093	.4487	.5541
35	.2746	.3246	.3810	.4182	.5189
40	.2573	.3044	.3578	.3932	.4896
45	.2428	.2875	.3384	.3721	.4648
50	.2306	.2732	.3218	.3541	.4433
60	.2108	.2500	.2948	.3248	.4078
70	.1954	.2319	.2737	.3017	.3799
80	.1829	.2172	.2565	.2830	.3568
90	.1726	.2050	.2422	.2673	.3375
100	.1638	.1946	.2301	.2540	.3211

tabla F Valores críticos de la distribución F para $\alpha = 0.05$ (en tipo normal) y $\alpha = 0.01$ (en negritas)

Los valores que se presentan en la tabla son los valores críticos de F para los grados de libertad del numerador del cociente F (encabezados de columna) y los grados de libertad del denominador del cociente F (encabezados de renglón). Para que sea significativo, $F_{\text{obt}} \geq F_{\text{crit}}$.



Grados de libertad: denominador	Grados de libertad: numerador																							
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞
1	161 4052	200 4999	216 5403	225 5625	230 5764	234 5859	237 5928	239 5981	241 6022	242 6056	243 6082	244 6106	245 6142	246 6169	248 6208	249 6234	250 6258	251 6286	252 6302	253 6323	253 6334	254 6352	254 6361	254 6366
2	18.51 98.49	19.00 99.00	19.16 99.17	19.25 99.25	19.30 99.30	19.33 99.33	19.36 99.34	19.37 99.36	19.38 99.38	19.39 99.40	19.40 99.41	19.41 99.42	19.42 99.43	19.43 99.44	19.44 99.45	19.45 99.46	19.46 99.47	19.47 99.48	19.47 99.48	19.48 99.49	19.49 99.49	19.49 99.49	19.50 99.50	19.50 99.50
3	10.13 34.12	9.55 30.82	9.28 29.46	9.12 28.71	9.01 28.24	8.94 27.91	8.88 27.67	8.84 27.49	8.81 27.34	8.78 27.23	8.76 27.13	8.74 27.05	8.71 26.92	8.69 26.83	8.66 26.69	8.64 26.60	8.62 26.50	8.60 26.41	8.58 26.35	8.57 26.27	8.56 26.23	8.54 26.18	8.54 26.14	8.53 26.12
4	7.71 21.20	6.94 18.00	6.59 16.69	6.39 15.98	6.26 15.52	6.16 15.21	6.09 14.98	6.04 14.80	6.00 14.66	5.96 14.54	5.93 14.45	5.91 14.37	5.87 14.24	5.84 14.15	5.80 14.02	5.77 13.93	5.74 13.83	5.71 13.74	5.70 13.69	5.68 13.61	5.66 13.57	5.65 13.52	5.64 13.48	5.63 13.46
5	6.61 16.26	5.79 13.27	5.41 12.06	5.19 11.39	5.05 10.97	4.95 10.67	4.88 10.45	4.82 10.27	4.78 10.15	4.74 10.05	4.70 9.96	4.68 9.89	4.64 9.77	4.60 9.68	4.56 9.55	4.53 9.47	4.50 9.38	4.46 9.29	4.44 9.24	4.42 9.17	4.40 9.13	4.38 9.07	4.37 9.04	4.36 9.02
6	5.99 13.74	5.14 10.92	4.76 9.78	4.53 9.15	4.39 8.75	4.28 8.47	4.21 8.26	4.15 8.10	4.10 7.98	4.06 7.87	4.03 7.79	4.00 7.72	3.96 7.60	3.92 7.52	3.87 7.39	3.84 7.31	3.81 7.23	3.77 7.14	3.75 7.09	3.72 7.02	3.71 6.99	3.69 6.94	3.68 6.90	3.67 6.88
7	5.59 12.25	4.47 9.55	4.35 8.45	4.12 7.85	3.97 7.46	3.87 7.19	3.79 7.00	3.73 6.84	3.68 6.71	3.63 6.62	3.60 6.54	3.57 6.47	3.52 6.35	3.49 6.27	3.44 6.15	3.41 6.07	3.38 5.98	3.34 5.90	3.32 5.85	3.29 5.78	3.28 5.75	3.25 5.70	3.24 5.67	3.23 5.65
8	5.32 11.26	4.46 8.65	4.07 7.59	3.84 7.01	3.69 6.63	3.58 6.37	3.50 6.19	3.44 6.03	3.39 5.91	3.34 5.82	3.31 5.74	3.28 5.67	3.23 5.56	3.20 5.48	3.15 5.36	3.12 5.28	3.08 5.20	3.05 5.11	3.03 5.06	3.00 5.00	2.98 4.96	2.96 4.91	2.94 4.88	2.93 4.86
9	5.12 10.56	4.26 8.02	3.86 6.99	3.63 6.42	3.48 6.06	3.37 5.80	3.29 5.62	3.23 5.47	3.18 5.35	3.13 5.26	3.10 5.18	3.07 5.11	3.02 5.00	2.98 4.92	2.93 4.80	2.90 4.73	2.86 4.64	2.82 4.56	2.80 4.51	2.77 4.45	2.76 4.41	2.73 4.36	2.72 4.33	2.71 4.31
10	4.96 10.04	4.10 7.56	3.71 6.55	3.48 5.99	3.33 5.64	3.22 5.39	3.14 5.21	3.07 5.06	2.97 4.95	2.92 4.85	2.91 4.78	2.86 4.71	2.82 4.60	2.77 4.52	2.72 4.41	2.70 4.33	2.67 4.25	2.67 4.17	2.64 4.12	2.61 4.05	2.59 4.01	2.56 3.96	2.55 3.93	2.54 3.91
11	4.84 9.65	3.98 7.20	3.59 6.22	3.36 5.67	3.20 5.32	3.09 5.07	3.01 4.88	2.95 4.74	2.90 4.63	2.86 4.54	2.82 4.46	2.79 4.40	2.74 4.29	2.70 4.21	2.65 4.10	2.61 4.02	2.57 3.94	2.53 3.86	2.50 3.80	2.47 3.74	2.45 3.70	2.42 3.66	2.41 3.62	2.40 3.60
12	4.75 9.33	3.88 6.93	3.49 5.95	3.26 5.41	3.11 5.06	3.00 4.82	2.92 4.65	2.85 4.50	2.80 4.39	2.76 4.30	2.72 4.22	2.69 4.16	2.64 4.05	2.60 3.98	2.54 3.86	2.50 3.78	2.46 3.70	2.42 3.61	2.40 3.56	2.36 3.49	2.35 3.46	2.32 3.41	2.31 3.38	2.30 3.36
13	4.67 9.07	3.80 6.70	3.41 5.74	3.18 5.20	3.02 4.86	2.92 4.62	2.84 4.44	2.77 4.30	2.72 4.19	2.67 4.10	2.63 4.02	2.60 3.96	2.55 3.85	2.51 3.78	2.46 3.67	2.42 3.59	2.38 3.51	2.34 3.42	2.32 3.37	2.28 3.30	2.26 3.27	2.24 3.21	2.22 3.18	2.21 3.16
14	4.60 8.86	3.74 6.51	3.34 5.56	3.11 5.03	2.96 4.69	2.85 4.46	2.77 4.28	2.70 4.14	2.65 4.03	2.60 3.94	2.56 3.86	2.53 3.80	2.48 3.70	2.44 3.62	2.39 3.51	2.35 3.43	2.31 3.34	2.27 3.26	2.24 3.21	2.21 3.14	2.19 3.11	2.16 3.06	2.14 3.02	2.13 3.00

Grados de libertad:
denominador

Grados de libertad: numerador

	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞	
15	4.54 8.68	3.68 6.36	3.29 5.42	3.06 4.89	2.90 4.56	2.79 4.32	2.70 4.14	2.64 4.00	2.59 3.89	2.55 3.80	2.51 3.73	2.48 3.67	2.43 3.56	2.39 3.48	2.33 3.36	2.29 3.29	2.25 3.20	2.21 3.12	2.18 3.07	2.15 3.00	2.12 2.97	2.10 2.92	2.08 2.89	2.08 2.89	2.07 2.87
16	4.49 8.53	3.63 6.23	3.24 5.29	3.01 4.77	2.85 4.44	2.74 4.20	2.66 4.03	2.59 3.89	2.54 3.78	2.49 3.69	2.45 3.61	2.42 3.55	2.37 3.45	2.33 3.37	2.28 3.25	2.24 3.18	2.20 3.10	2.16 3.01	2.13 2.96	2.09 2.89	2.07 2.86	2.04 2.82	2.04 2.80	2.02 2.77	2.02 2.75
17	4.45 8.40	3.59 6.11	3.20 5.18	2.96 4.67	2.81 4.34	2.70 4.10	2.62 3.93	2.55 3.79	2.50 3.68	2.45 3.59	2.41 3.52	2.38 3.45	2.33 3.35	2.29 3.27	2.23 3.16	2.19 3.08	2.15 3.00	2.11 2.92	2.08 2.86	2.04 2.79	2.02 2.76	1.99 2.70	1.99 2.67	1.97 2.65	1.96 2.65
18	4.41 8.28	3.55 6.01	3.16 5.09	2.93 4.58	2.77 4.25	2.66 4.01	2.58 3.85	2.51 3.71	2.46 3.60	2.41 3.51	2.37 3.44	2.34 3.37	2.29 3.27	2.25 3.19	2.19 3.07	2.15 3.00	2.11 2.91	2.07 2.83	2.04 2.78	2.00 2.71	1.98 2.68	1.95 2.62	1.95 2.59	1.93 2.57	1.92 2.57
19	4.38 8.18	3.52 5.93	3.13 5.01	2.90 4.50	2.74 4.17	2.63 3.94	2.55 3.77	2.48 3.63	2.43 3.52	2.38 3.43	2.34 3.36	2.31 3.30	2.26 3.19	2.21 3.12	2.15 3.00	2.11 2.92	2.07 2.84	2.02 2.76	2.00 2.70	1.96 2.63	1.94 2.60	1.91 2.54	1.91 2.51	1.90 2.51	1.88 2.49
20	4.35 8.10	3.49 5.85	3.10 4.94	2.87 4.43	2.71 4.10	2.60 3.87	2.52 3.71	2.45 3.56	2.40 3.45	2.35 3.37	2.31 3.30	2.28 3.23	2.23 3.13	2.18 3.05	2.12 2.94	2.08 2.86	2.04 2.77	1.99 2.69	1.96 2.63	1.92 2.56	1.90 2.53	1.87 2.47	1.87 2.44	1.85 2.42	1.84 2.42
21	4.32 8.02	3.47 5.78	3.07 4.87	2.84 4.37	2.68 4.04	2.57 3.81	2.49 3.65	2.42 3.51	2.37 3.40	2.32 3.31	2.28 3.24	2.25 3.17	2.20 3.07	2.15 2.99	2.09 2.88	2.05 2.80	2.00 2.72	1.96 2.63	1.93 2.58	1.89 2.51	1.87 2.47	1.84 2.42	1.84 2.38	1.82 2.36	1.81 2.36
22	4.30 7.94	3.44 5.72	3.05 4.82	2.82 4.31	2.66 3.99	2.55 3.76	2.47 3.59	2.40 3.45	2.35 3.35	2.30 3.26	2.26 3.18	2.23 3.12	2.18 3.02	2.13 2.94	2.07 2.83	2.03 2.75	1.98 2.67	1.93 2.58	1.91 2.53	1.87 2.46	1.84 2.42	1.81 2.37	1.81 2.33	1.80 2.31	1.78 2.31
23	4.28 7.88	3.42 5.66	3.03 4.76	2.80 4.26	2.64 3.94	2.53 3.71	2.45 3.54	2.38 3.41	2.32 3.30	2.28 3.21	2.24 3.14	2.20 3.07	2.14 2.97	2.10 2.89	2.04 2.78	2.00 2.70	1.96 2.62	1.91 2.53	1.88 2.48	1.84 2.41	1.82 2.37	1.79 2.32	1.79 2.28	1.77 2.28	1.76 2.26
24	4.26 7.82	3.40 5.61	3.01 4.72	2.78 4.22	2.62 3.90	2.51 3.67	2.43 3.50	2.36 3.36	2.30 3.25	2.26 3.17	2.22 3.09	2.18 3.03	2.13 2.93	2.09 2.85	2.02 2.74	1.98 2.66	1.94 2.58	1.89 2.49	1.86 2.44	1.82 2.36	1.80 2.33	1.76 2.27	1.76 2.23	1.74 2.23	1.73 2.21
25	4.24 7.77	3.38 5.57	2.99 4.68	2.76 4.18	2.60 3.86	2.49 3.63	2.41 3.46	2.34 3.32	2.28 3.21	2.24 3.13	2.20 3.05	2.16 2.99	2.11 2.89	2.06 2.81	2.00 2.70	1.96 2.62	1.92 2.54	1.87 2.45	1.84 2.40	1.80 2.32	1.77 2.29	1.74 2.23	1.74 2.19	1.72 2.19	1.71 2.17
26	4.22 7.72	3.37 5.53	2.98 4.64	2.74 4.14	2.59 3.82	2.47 3.59	2.39 3.42	2.32 3.29	2.27 3.17	2.22 3.09	2.18 3.02	2.15 2.96	2.10 2.86	2.05 2.77	1.99 2.66	1.95 2.58	1.90 2.50	1.85 2.41	1.82 2.36	1.78 2.28	1.76 2.25	1.72 2.19	1.72 2.15	1.70 2.15	1.69 2.13
27	4.21 7.68	3.35 5.49	2.96 4.60	2.73 4.11	2.57 3.79	2.46 3.56	2.37 3.39	2.30 3.26	2.25 3.14	2.20 3.06	2.16 2.98	2.13 2.93	2.08 2.83	2.03 2.74	1.97 2.63	1.93 2.55	1.88 2.47	1.84 2.38	1.80 2.33	1.76 2.25	1.74 2.21	1.71 2.16	1.71 2.12	1.68 2.12	1.67 2.10
28	4.20 7.64	3.34 5.45	2.95 4.57	2.71 4.07	2.56 3.76	2.44 3.53	2.36 3.36	2.29 3.23	2.24 3.11	2.19 3.03	2.15 2.95	2.12 2.90	2.06 2.80	2.02 2.71	1.96 2.60	1.91 2.52	1.87 2.44	1.81 2.35	1.78 2.30	1.75 2.22	1.72 2.18	1.69 2.13	1.69 2.09	1.67 2.09	1.65 2.06
29	4.18 7.60	3.33 5.42	2.93 4.54	2.70 4.04	2.54 3.73	2.43 3.50	2.35 3.33	2.28 3.20	2.22 3.08	2.18 3.00	2.14 2.92	2.10 2.87	2.05 2.77	2.00 2.68	1.94 2.57	1.90 2.49	1.85 2.41	1.80 2.32	1.77 2.27	1.73 2.19	1.71 2.15	1.68 2.10	1.68 2.06	1.65 2.03	1.64 2.03
30	4.17 7.56	3.32 5.39	2.92 4.51	2.69 4.02	2.53 3.70	2.42 3.47	2.34 3.30	2.27 3.17	2.21 3.06	2.16 2.98	2.12 2.90	2.09 2.84	2.04 2.74	1.99 2.66	1.93 2.55	1.89 2.47	1.84 2.38	1.79 2.29	1.76 2.24	1.72 2.16	1.69 2.13	1.66 2.07	1.66 2.03	1.64 2.01	1.62 2.01
32	4.15 7.50	3.30 5.34	2.90 4.46	2.67 3.97	2.51 3.66	2.40 3.42	2.32 3.25	2.25 3.12	2.19 3.01	2.14 2.94	2.10 2.86	2.07 2.80	2.02 2.70	1.97 2.62	1.91 2.51	1.86 2.42	1.82 2.34	1.76 2.25	1.74 2.20	1.69 2.12	1.67 2.08	1.66 2.02	1.64 1.98	1.61 1.98	1.59 1.96
34	4.13 7.44	3.28 5.29	2.88 4.42	2.65 3.93	2.49 3.61	2.38 3.38	2.30 3.21	2.23 3.08	2.17 2.97	2.12 2.89	2.08 2.82	2.05 2.76	2.00 2.66	1.95 2.58	1.89 2.47	1.84 2.38	1.80 2.30	1.74 2.21	1.71 2.15	1.67 2.08	1.64 2.04	1.64 1.98	1.61 1.94	1.59 1.91	1.57 1.91
36	4.11 7.39	3.26 5.25	2.86 4.38	2.63 3.89	2.48 3.58	2.36 3.35	2.28 3.18	2.21 3.04	2.15 2.94	2.10 2.86	2.06 2.78	2.03 2.72	1.98 2.62	1.93 2.54	1.87 2.43	1.82 2.35	1.78 2.26	1.72 2.17	1.69 2.12	1.65 2.04	1.62 2.00	1.59 1.94	1.56 1.90	1.55 1.87	1.55 1.87
38	4.10 7.35	3.25 5.21	2.85 4.34	2.62 3.86	2.46 3.54	2.35 3.32	2.26 3.15	2.19 3.02	2.14 2.91	2.09 2.82	2.05 2.75	2.02 2.69	1.96 2.59	1.92 2.51	1.85 2.40	1.80 2.32	1.76 2.22	1.71 2.14	1.67 2.08	1.63 2.00	1.60 1.97	1.57 1.90	1.54 1.86	1.53 1.84	1.53 1.84

(continúa)

tabla F (continuación)

Los valores que se presentan en la tabla son los valores críticos de F para los grados de libertad del numerador del cociente F (encabezados de columna) y los grados de libertad del denominador del cociente F (encabezados de renglón). Para que sea significativo, $F_{\text{obt}} \geq F_{\text{crit}}$.

Grados de libertad: denominador	Grados de libertad: numerador																							
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞
40	4.08 7.31	3.23 5.18	2.84 4.31	2.61 3.83	2.45 3.51	2.34 3.29	2.25 3.12	2.18 2.99	2.12 2.88	2.07 2.80	2.04 2.73	2.00 2.66	1.95 2.56	1.90 2.49	1.84 2.37	1.79 2.29	1.74 2.20	1.69 2.11	1.66 2.05	1.61 1.97	1.59 1.94	1.55 1.88	1.53 1.84	1.51 1.81
42	4.07 7.27	3.22 5.15	2.83 4.29	2.59 3.80	2.44 3.49	2.32 3.26	2.24 3.10	2.17 2.96	2.11 2.86	2.06 2.77	2.02 2.70	1.99 2.64	1.94 2.54	1.89 2.46	1.82 2.35	1.78 2.26	1.73 2.17	1.68 2.08	1.64 2.02	1.60 1.94	1.57 1.91	1.54 1.85	1.51 1.80	1.49 1.78
44	4.06 7.24	3.21 5.12	2.82 4.26	2.58 3.78	2.43 3.46	2.31 3.24	2.23 3.07	2.16 2.94	2.10 2.84	2.05 2.75	2.01 2.68	1.98 2.62	1.92 2.52	1.88 2.44	1.81 2.32	1.76 2.24	1.72 2.15	1.66 2.06	1.63 2.00	1.58 1.92	1.56 1.88	1.52 1.82	1.50 1.78	1.48 1.75
46	4.05 7.21	3.20 5.10	2.81 4.24	2.57 3.76	2.42 3.44	2.30 3.22	2.22 3.05	2.14 2.92	2.09 2.82	2.04 2.73	2.00 2.66	1.97 2.60	1.91 2.50	1.87 2.42	1.80 2.30	1.75 2.22	1.71 2.13	1.65 2.04	1.62 1.98	1.57 1.90	1.54 1.86	1.51 1.80	1.48 1.76	1.46 1.72
48	4.04 7.19	3.19 5.08	2.80 4.22	2.56 3.74	2.41 3.42	2.30 3.20	2.21 3.04	2.14 2.90	2.08 2.80	2.03 2.71	1.99 2.64	1.96 2.58	1.90 2.48	1.86 2.40	1.79 2.28	1.74 2.20	1.70 2.11	1.64 2.02	1.61 1.96	1.56 1.88	1.53 1.84	1.50 1.78	1.47 1.73	1.45 1.70
50	4.03 7.17	3.18 5.06	2.79 4.20	2.56 3.72	2.40 3.41	2.29 3.18	2.20 3.02	2.13 2.88	2.07 2.78	2.02 2.70	1.98 2.62	1.95 2.56	1.90 2.46	1.85 2.39	1.78 2.26	1.74 2.18	1.69 2.10	1.63 2.00	1.60 1.94	1.55 1.86	1.52 1.82	1.48 1.76	1.46 1.71	1.44 1.68
55	4.02 7.12	3.17 5.01	2.78 4.16	2.54 3.68	2.38 3.37	2.27 3.15	2.18 2.98	2.11 2.85	2.05 2.75	2.00 2.66	1.97 2.59	1.93 2.53	1.88 2.43	1.83 2.35	1.76 2.23	1.72 2.15	1.67 2.06	1.61 1.96	1.58 1.90	1.52 1.82	1.50 1.78	1.46 1.71	1.43 1.66	1.41 1.64
60	4.00 7.08	3.15 4.98	2.76 4.13	2.52 3.65	2.37 3.34	2.25 3.12	2.17 2.95	2.10 2.82	2.04 2.72	1.99 2.63	1.95 2.56	1.92 2.50	1.86 2.40	1.81 2.32	1.75 2.20	1.70 2.12	1.65 2.03	1.59 1.93	1.56 1.87	1.50 1.79	1.48 1.74	1.44 1.68	1.41 1.63	1.39 1.60
65	3.99 7.04	3.14 4.95	2.75 4.10	2.51 3.62	2.36 3.31	2.24 3.09	2.15 2.93	2.08 2.79	2.02 2.70	1.98 2.61	1.94 2.54	1.90 2.47	1.85 2.37	1.80 2.30	1.73 2.18	1.68 2.09	1.63 2.00	1.57 1.90	1.54 1.84	1.49 1.76	1.46 1.71	1.42 1.64	1.39 1.60	1.37 1.56
70	3.98 7.01	3.13 4.92	2.74 4.08	2.50 3.60	2.35 3.29	2.23 3.07	2.14 2.91	2.07 2.77	2.01 2.67	1.97 2.59	1.93 2.51	1.89 2.45	1.84 2.35	1.79 2.28	1.72 2.15	1.67 2.07	1.62 1.98	1.56 1.88	1.53 1.82	1.47 1.74	1.45 1.69	1.40 1.62	1.37 1.56	1.35 1.53
80	3.96 6.96	3.11 4.88	2.72 4.04	2.48 3.56	2.33 3.25	2.21 3.04	2.12 2.87	2.05 2.74	1.99 2.64	1.95 2.55	1.91 2.48	1.88 2.41	1.82 2.32	1.77 2.24	1.70 2.11	1.65 2.03	1.60 1.94	1.54 1.84	1.51 1.78	1.45 1.70	1.42 1.65	1.38 1.57	1.35 1.52	1.32 1.49
100	3.94 6.90	3.09 4.82	2.70 3.98	2.46 3.51	2.30 3.20	2.19 2.99	2.10 2.82	2.03 2.69	1.97 2.59	1.92 2.51	1.88 2.43	1.85 2.36	1.79 2.26	1.75 2.19	1.68 2.06	1.63 1.98	1.57 1.89	1.51 1.79	1.48 1.73	1.42 1.64	1.39 1.59	1.34 1.51	1.30 1.46	1.28 1.43
125	3.92 6.84	3.07 4.78	2.68 3.94	2.44 3.47	2.29 3.17	2.17 2.95	2.08 2.79	2.01 2.65	1.95 2.56	1.90 2.47	1.86 2.40	1.83 2.33	1.77 2.23	1.72 2.15	1.65 2.03	1.60 1.94	1.55 1.85	1.49 1.75	1.45 1.70	1.39 1.68	1.36 1.59	1.31 1.54	1.27 1.46	1.25 1.40
150	3.91 6.81	3.06 4.75	2.67 3.91	2.43 3.44	2.27 3.14	2.16 2.92	2.07 2.76	2.00 2.62	1.94 2.53	1.89 2.44	1.85 2.37	1.82 2.30	1.76 2.20	1.71 2.12	1.64 2.00	1.59 1.91	1.54 1.83	1.47 1.72	1.44 1.66	1.37 1.56	1.34 1.51	1.29 1.43	1.25 1.37	1.22 1.33
200	3.89 6.76	3.04 4.71	2.65 3.88	2.41 3.41	2.26 3.11	2.14 2.90	2.05 2.73	1.98 2.60	1.92 2.50	1.87 2.41	1.83 2.34	1.80 2.28	1.74 2.17	1.69 2.09	1.62 1.97	1.57 1.88	1.52 1.79	1.45 1.69	1.42 1.62	1.35 1.53	1.32 1.48	1.26 1.39	1.22 1.33	1.19 1.28
400	3.86 6.70	3.02 4.66	2.62 3.83	2.39 3.36	2.23 3.06	2.12 2.85	2.03 2.69	1.96 2.55	1.90 2.46	1.85 2.37	1.81 2.29	1.78 2.23	1.72 2.12	1.67 2.04	1.60 1.92	1.54 1.84	1.49 1.74	1.42 1.64	1.38 1.57	1.32 1.47	1.28 1.42	1.22 1.32	1.16 1.24	1.13 1.19
1000	3.85 6.66	3.00 4.62	2.61 3.80	2.38 3.34	2.22 3.04	2.10 2.82	2.02 2.66	1.95 2.53	1.89 2.43	1.84 2.34	1.80 2.26	1.76 2.20	1.70 2.09	1.65 2.01	1.58 1.89	1.53 1.81	1.47 1.71	1.41 1.61	1.36 1.54	1.30 1.44	1.26 1.38	1.19 1.28	1.13 1.19	1.08 1.11
	3.81 6.64	2.99 4.60	2.60 3.78	2.37 3.32	2.21 3.02	2.09 2.80	2.01 2.64	1.94 2.51	1.88 2.41	1.83 2.32	1.79 2.24	1.73 2.18	1.69 2.07	1.64 1.99	1.57 1.87	1.52 1.79	1.46 1.69	1.40 1.59	1.35 1.52	1.28 1.41	1.24 1.36	1.17 1.25	1.11 1.15	1.00 1.00

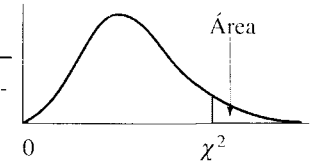
tabla G Valores críticos de la distribución con rango de Student (Q)

Los valores que se presentan en la tabla son los valores críticos de Q para $\alpha = 0.05$ y 0.01 , como función de los grados de libertad de s_w^2 y k (el número de medias) o r (el número de medias que se encuentran entre \bar{X}_i y \bar{X}_j , después de ordenarlas). Para que sea significativo, $Q_{\text{obt}} \geq Q_{\text{crit}}$.

s_w^2 gl	α	k (el número de medias) o r (el número de medias que se encuentran entre \bar{X}_i y \bar{X}_j)									
		2	3	4	5	6	7	8	9	10	11
5	.05	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17
	.01	5.70	6.98	7.80	8.42	8.91	9.32	9.67	9.97	10.24	10.48
6	.05	3.46	4.34	4.90	5.30	5.63	5.90	6.12	6.32	6.49	6.65
	.01	5.24	6.33	7.03	7.56	7.97	8.32	8.61	8.87	9.10	9.30
7	.05	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16	6.30
	.01	4.95	5.92	6.54	7.01	7.37	7.68	7.94	8.17	8.37	8.55
8	.05	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05
	.01	4.75	5.64	6.20	6.62	6.96	7.24	7.47	7.68	7.86	8.03
9	.05	3.20	3.95	4.41	4.76	5.02	5.24	5.43	5.59	5.74	5.87
	.01	4.60	5.43	5.96	6.35	6.66	6.91	7.13	7.33	7.49	7.65
10	.05	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60	5.72
	.01	4.48	5.27	5.77	6.14	6.43	6.67	6.87	7.05	7.21	7.36
11	.05	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49	5.61
	.01	4.39	5.15	5.62	5.97	6.25	6.48	6.67	6.84	6.99	7.13
12	.05	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.39	5.51
	.01	4.32	5.05	5.50	5.84	6.10	6.32	6.51	6.67	6.81	6.94
13	.05	3.06	3.73	4.15	4.45	4.69	4.88	5.05	5.19	5.32	5.43
	.01	4.26	4.96	5.40	5.73	5.98	6.19	6.37	6.53	6.67	6.79
14	.05	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25	5.36
	.01	4.21	4.89	5.32	5.63	5.88	6.08	6.26	6.41	6.54	6.66
15	.05	3.01	3.67	4.08	4.37	4.59	4.78	4.94	5.08	5.20	5.31
	.01	4.17	4.84	5.25	5.56	5.80	5.99	6.16	6.31	6.44	6.55
16	.05	3.00	3.65	4.05	4.33	4.56	4.74	4.90	5.03	5.15	5.26
	.01	4.13	4.79	5.19	5.49	5.72	5.92	6.08	6.22	6.35	6.46
17	.05	2.98	3.63	4.02	4.30	4.52	4.70	4.86	4.99	5.11	5.21
	.01	4.10	4.74	5.14	5.43	5.66	5.85	6.01	6.15	6.27	6.38
18	.05	2.97	3.61	4.00	4.28	4.49	4.67	4.82	4.96	5.07	5.17
	.01	4.07	4.70	5.09	5.38	5.60	5.79	5.94	6.08	6.20	6.31
19	.05	2.96	3.59	3.98	4.25	4.47	4.65	4.79	4.92	5.04	5.14
	.01	4.05	4.67	5.05	5.33	5.55	5.73	4.89	6.02	6.14	6.25
20	.05	2.95	3.58	3.96	4.23	4.45	4.62	4.77	4.90	5.01	5.11
	.01	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09	6.19
24	.05	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81	4.92	5.01
	.01	3.96	4.55	4.91	5.17	5.37	5.54	5.69	5.81	5.92	6.02
30	.05	2.89	3.49	3.85	4.10	4.30	4.46	4.60	4.72	4.82	4.92
	.01	3.89	4.45	4.80	5.05	5.24	5.40	5.54	5.65	5.76	5.85
40	.05	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.73	4.82
	.01	3.82	4.37	4.70	4.93	5.11	5.26	5.39	5.50	5.60	5.69
60	.05	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65	4.73
	.01	3.76	4.28	4.59	4.82	4.99	5.13	5.25	5.36	5.45	5.53
120	.05	2.80	3.36	3.68	3.92	4.10	4.24	4.36	4.47	4.56	4.64
	.01	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30	5.37
∞	.05	2.77	3.31	3.63	3.86	4.03	4.17	4.29	4.39	4.47	4.55
	.01	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16	5.23

tabla H Distribución ji-cuadrada (χ^2)

La primera columna (*gl*) permite localizar cada distribución χ^2 . Las demás columnas indican la proporción del área bajo la distribución χ^2 que se encuentra por encima del valor de χ^2 que se ha localizado en la tabla. Los valores de χ^2 que aparecen bajo los encabezados de las columnas correspondientes a .05 y .01 son los valores críticos de χ^2 para $\alpha = 0.05$ and 0.01. Para que sea significativo, $\chi_{\text{obt}}^2 \geq \chi_{\text{crit}}^2$.



Grados de libertad <i>gl</i>	<i>P</i> = .99	.98	.95	.90	.80	.70	.50	.30	.20	.10	.05	.02	.01
1	.000157	.000628	.00393	.0158	.0642	.148	.455	1.074	1.642	2.706	3.841	5.412	6.635
2	.0201	.0404	.103	.211	.446	.713	1.386	2.408	3.219	4.605	5.991	7.824	9.210
3	.115	.185	.352	.584	1.005	1.424	2.366	3.665	4.642	6.251	7.815	9.837	11.341
4	.297	.429	.711	1.064	1.649	2.195	3.357	4.878	5.989	7.779	9.488	11.668	13.277
5	.554	.752	1.145	1.610	2.343	3.000	4.351	6.064	7.289	9.236	11.070	13.388	15.086
6	.872	1.134	1.635	2.204	3.070	3.828	5.348	7.231	8.558	10.645	12.592	15.033	16.812
7	1.239	1.564	2.167	2.833	3.822	4.671	6.346	8.383	9.803	12.017	14.067	16.622	18.475
8	1.646	2.032	2.733	3.490	4.594	5.527	7.344	9.524	11.030	13.362	15.507	18.168	20.090
9	2.088	2.532	3.325	4.168	5.380	6.393	8.343	10.656	12.242	14.684	16.919	19.679	21.666
10	2.558	3.059	3.940	4.865	6.179	7.267	9.342	11.781	13.442	15.987	18.307	21.161	23.209
11	3.053	3.609	4.575	5.578	6.989	8.148	10.341	12.899	14.631	17.275	19.675	22.618	24.725
12	3.571	4.178	5.226	6.304	7.807	9.034	11.340	14.011	15.812	18.549	21.026	24.054	26.217
13	4.107	4.765	5.892	7.042	8.634	9.926	12.340	15.119	16.985	19.812	22.362	25.472	27.688
14	4.660	5.368	6.571	7.790	9.467	10.821	13.339	16.222	18.151	21.064	23.685	26.873	29.141
15	5.229	5.985	7.261	8.547	10.307	11.721	14.339	17.322	19.311	22.307	24.996	28.259	30.578
16	5.812	6.614	7.962	9.312	11.152	12.624	15.338	18.418	20.465	23.542	26.296	29.633	32.000
17	6.408	7.255	8.672	10.085	12.002	13.531	16.338	19.511	21.615	24.769	27.587	30.995	33.409
18	7.015	7.906	9.390	10.865	12.857	14.440	17.338	20.601	22.760	25.989	28.869	32.346	34.805
19	7.633	8.567	10.117	11.651	13.716	15.352	18.338	21.689	23.900	27.204	30.144	33.687	36.191
20	8.260	9.237	10.851	12.443	14.578	16.266	19.337	22.775	25.038	28.412	31.410	35.020	37.566
21	8.897	9.915	11.591	13.240	15.445	17.182	20.337	23.858	26.171	29.615	32.671	36.343	38.932
22	9.542	10.600	12.338	14.041	16.314	18.101	21.337	24.939	27.301	30.813	33.924	37.659	40.289
23	10.196	11.293	13.091	14.848	17.187	19.021	22.337	26.018	28.429	32.007	35.172	38.968	41.638
24	10.856	11.992	13.848	15.659	18.062	19.943	23.337	27.096	29.553	33.196	36.415	40.270	42.980
25	11.524	12.697	14.611	16.473	18.940	20.867	24.337	28.172	30.675	34.382	37.652	41.566	44.314
26	12.198	13.409	15.379	17.292	19.820	21.792	25.336	29.246	31.795	35.563	38.885	42.856	45.642
27	12.879	14.125	16.151	18.114	20.703	22.719	26.336	30.319	32.912	36.741	40.113	44.140	46.963
28	13.565	14.847	16.928	18.939	21.588	23.647	27.336	31.391	34.027	37.916	41.337	45.419	48.278
29	14.256	15.574	17.708	19.768	22.475	24.577	28.336	32.461	35.139	39.087	42.557	46.693	49.588
30	14.953	16.306	18.493	20.599	23.364	25.508	29.336	33.530	36.250	40.256	43.773	47.962	50.892

tabla 1 Valores críticos de T para la prueba de rangos con signo de Wilcoxon

Los valores que se presentan en la tabla son los valores críticos de T para la N (columna de la izquierda) y el nivel alfa (encabezado de la columna) especificados. Para que sea significativo, $T_{\text{obt}} \leq T_{\text{crit}}$.

N	Nivel de significancia para la prueba con una cola				N	Nivel de significancia para la prueba con una cola			
	.05	.025	.01	.005		.05	.025	.01	.005
	Nivel de significancia para la prueba con dos colas					Nivel de significancia para la prueba con dos colas			
	.10	.05	.02	.01		.10	.05	.02	.01
5	0	—	—	—	28	130	116	101	91
6	2	0	—	—	29	140	126	110	100
7	3	2	0	—	30	151	137	120	109
8	5	3	1	0	31	163	147	130	118
9	8	5	3	1	32	175	159	140	128
10	10	8	5	3	33	187	170	151	138
11	13	10	7	5	34	200	182	162	148
12	17	13	9	7	35	213	195	173	159
13	21	17	12	9	36	227	208	185	171
14	25	21	15	12	37	241	221	198	182
15	30	25	19	15	38	256	235	211	194
16	35	29	23	19	39	271	249	224	207
17	41	34	27	23	40	286	264	238	220
18	47	40	32	27	41	302	279	252	233
19	53	46	37	32	42	319	294	266	247
20	60	52	43	37	43	336	310	281	261
21	67	58	49	42	44	353	327	296	276
22	75	65	55	48	45	371	343	312	291
23	83	73	62	54	46	389	361	328	307
24	91	81	69	61	47	407	378	345	322
25	100	89	76	68	48	426	396	362	339
26	110	98	84	75	49	446	415	379	355
27	119	107	92	83	50	466	434	397	373

table J Números aleatorios

	1	2	3	4	5	6	7	8	9
1	32942	95416	42339	59045	26693	49057	87496	20624	14819
2	07410	99859	83828	21409	29094	65114	36701	25762	12827
3	59981	68155	45673	76210	58219	45738	29550	24736	09574
4	46251	25437	69654	99716	11563	08803	86027	51867	12116
5	65558	51904	93123	27887	53138	21488	09095	78777	71240
6	99187	19258	86421	16401	19397	83297	40111	49326	81686
7	35641	00301	16096	34775	21562	97983	45040	19200	16383
8	14031	00936	81518	48440	02218	04756	19506	60695	88494
9	60677	15076	92554	26042	23472	69869	62877	19584	39576
10	66314	05212	67859	89356	20056	30648	87349	20389	53805
11	20416	87410	75646	64176	82752	63606	37011	57346	69512
12	28701	56992	70423	62415	40807	98086	58850	28968	45297
13	74579	33844	33426	07570	00728	07079	19322	56325	84819
14	62615	52342	82968	75540	80045	53069	20665	21282	07768
15	93945	06293	22879	08161	01442	75071	21427	94842	26210
16	75689	76131	96837	67450	44511	50424	82848	41975	71663
17	02921	16919	35424	93209	52133	87327	95897	65171	20376
18	14295	34969	14216	03191	61647	30296	66667	10101	63203
19	05303	91109	82403	40312	62191	67023	90073	83205	71344
20	57071	90357	12901	08899	91039	67251	28701	03846	94589
21	78471	57741	13599	84390	32146	00871	09354	22745	65806
22	89242	79337	59293	47481	07740	43345	25716	70020	54005
23	14955	59592	97035	80430	87220	06392	79028	57123	52872
24	42446	41880	37415	47472	04513	49494	08860	08038	43624
25	18534	22346	54556	17558	73689	14894	05030	19561	56517
26	39284	33737	42512	86411	23753	29690	26096	81361	93099
27	33922	37329	89911	55876	28379	81031	22058	21487	54613
28	78355	54013	50774	30666	61205	42574	47773	36027	27174
29	08845	99145	94316	88974	29828	97069	90327	61842	29604
30	01769	71825	55957	98271	02784	66731	40311	88495	18821
31	17639	38284	59478	90409	21997	56199	30068	82800	69692
32	05851	58653	99949	63505	40409	85551	90729	64938	52403
33	42396	40112	11469	03476	03328	84238	26570	51790	42122
34	13318	14192	98167	75631	74141	22369	36757	89117	54998
35	60571	54786	26281	01855	30706	66578	32019	65884	58485
36	09531	81853	59334	70929	03544	18510	89541	13555	21168
37	72865	16829	86542	00396	20363	13010	69645	49608	54738
38	56324	31093	77924	28622	83543	28912	15059	80192	83964
39	78192	21626	91399	07235	07104	73652	64425	85149	75409
40	64666	34767	97298	92708	01994	53188	78476	07804	62404
41	82201	75694	02808	65983	74373	66693	13094	74183	73020
42	15360	73776	40914	85190	54278	99054	62944	47351	89098
43	68142	67957	70896	37983	20487	95350	16371	03426	13895
44	19138	31200	30616	14639	44406	44236	57360	81644	94761
45	28155	03521	36415	78452	92359	81091	56513	88321	97910
46	87971	29031	51780	27376	81056	86155	55488	50590	74514
47	58147	68841	53625	02059	75223	16783	19272	61994	71090
48	18875	52809	70594	41649	32935	26430	82096	01605	65846
49	75109	56474	74111	31966	29969	70093	98901	84550	25769
50	35983	03742	76822	12073	59463	84420	15868	99505	11426

tabla J Números aleatorios (continuación)

	1	2	3	4	5	6	7	8	9
51	12651	61646	11769	75109	86996	97669	25757	32535	07122
52	81769	74436	02630	72310	45049	18029	07469	42341	98173
53	36737	98863	77240	76251	00654	64688	09343	70278	67331
54	82861	54371	76610	94934	72748	44124	05610	53750	95938
55	21325	15732	24127	37431	09723	63529	73977	95218	96074
56	74146	47887	62463	23045	41490	07954	22597	60012	98866
57	90759	64410	54179	66075	61051	75385	51378	08360	95946
58	55683	98078	02238	91540	21219	17720	87817	41705	95785
59	79686	17969	76061	83748	55920	83612	41540	86492	06447
60	70333	00201	86201	69716	78185	62154	77930	67663	29529
61	14042	53536	07779	04157	41172	36473	42123	43929	50533
62	59911	08256	06596	48416	69770	68797	56080	14223	59199
63	62368	62623	62742	14891	39247	52242	98832	69533	91174
64	57529	97751	54976	48957	74599	08759	78494	52785	68526
65	15469	90574	78033	66885	13936	42117	71831	22961	94225
66	18625	23674	53850	32827	81647	80820	00420	63555	74489
67	74626	68394	88562	70745	23701	45630	65891	58220	35442
68	11119	16519	27384	90199	79210	76965	99546	30323	31664
69	41101	17336	48951	53674	17880	45260	08575	49321	36191
70	32123	91576	84221	78902	82010	30847	62329	63898	23268
71	26091	68409	69704	82267	14751	13151	93115	01437	56945
72	67680	79790	48462	59278	44185	29616	76531	19589	83139
73	15184	19260	14073	07026	25264	08388	27182	22557	61501
74	58010	45039	57181	10238	36874	28546	37444	80824	63981
75	56425	53996	86245	32623	78858	08143	60377	42925	42815
76	82630	84066	13592	60642	17904	99718	63432	88642	37858
77	14927	40909	23900	48761	44860	92467	31742	87142	03607
78	23740	22505	07489	85986	74420	21744	97711	36648	35620
79	32990	97446	03711	63824	07953	85965	87089	11687	92414
80	05310	24058	91946	78437	34365	82469	12430	84754	19354
81	21839	39937	27534	88913	49055	19218	47712	67677	51889
82	08833	42549	93981	94051	28382	83725	72643	64233	97252
83	58336	11139	47479	00931	91560	95372	97642	33856	54825
84	62032	91144	75478	47431	52726	30289	42411	91886	51818
85	45171	30557	53116	04118	58301	24375	65609	85810	18620
86	91611	62656	60128	35609	63698	78356	50682	22505	01692
87	55472	63819	86314	49174	93582	73604	78614	78849	23096
88	18573	09729	74091	53994	10970	86557	65661	41854	26037
89	60866	02955	90288	82136	83644	94455	06560	78029	98768
90	45043	55608	82767	60890	74646	79485	13619	98868	40857
91	17831	09737	79473	75945	28394	79334	70577	38048	03607
92	40137	03981	07585	18128	11178	32601	27994	05641	22600
93	77776	31343	14576	97706	16039	47517	43300	59080	80392
94	69605	44104	40103	95635	05635	81673	68657	09559	23510
95	19916	52934	26499	09821	87331	80993	61299	36979	73599
96	02606	58552	07678	56619	65325	30705	99582	53390	46357
97	65183	73160	87131	35530	47946	09854	18080	02321	05809
98	10740	98914	44916	11322	89717	88189	30143	52687	19420
99	98642	89822	71691	51573	83666	61642	46683	33761	47542
100	60139	25601	93663	25547	02654	94829	48672	28736	84994



AGRADECIMIENTOS

Las tablas que aparecen en este apéndice fueron adaptadas con el debido permiso de las fuentes siguientes:

- Tabla A** R. Clarke, A. Coladarch y J. Caffrey, *Statistical Reasoning and Procedures*, Charles E. Merrill Publishers, Columbus, Ohio, 1965, Apéndice 2.
- Tabla B** R. S. Burington y D. C. Mayo, *Handbook of Probability and Statistics with Tables*, 2a. ed., McGraw-Hill Book Company, Nueva York, 1970.
- Tabla C** H. B. Mann y D. R. Whitney, "On a Test of Whether One of Two Random Variables Is Stochastically Larger Than the Other," *Annals of Mathematical Statistics*, 18 (1947), 50-60, y D. Auble, "Extended Tables for the Mann-Whitney Statistic," *Bulletin of the Institute of Educational Research at Indiana University*, 1, núm. 2 (1953), as used in Runyon and Haber, *Fundamentals of Behavioral Statistics*, 3a. ed., Addison-Wesley Publishing Company, Inc., Reading, Mass., 1976.
- Tabla D** Fisher y Yates, *Statistical Tables for Biological, Agricultural, y Medical Research*, Longman Group Ltd., Londres (previamente publicado por Oliver & Boyd Ltd., Edinburgh), 1974, Tabla III.
- Tabla E** Fisher y Yates, *Statistical Tables for Biological, Agricultural, and Medical Research*, Longman Group Ltd., Londres (previamente publicado por Oliver & Boyd Ltd., Edinburgh), 1974, Tabla VII.
- Tabla F** G. W. Snedecor, *Statistical Methods*, 5a. ed., Iowa State University Press, Ames, 1956.
- Tabla G** E. S. Pearson y H. O. Hartley, eds., *Biometrika Tables for Statisticians*, Vol. 1, 3a. ed., Cambridge University Press, Nueva York, 1966, Tabla 29.
- Tabla H** Fisher y Yates, *Statistical Tables for Biological, Agricultural, and Medical Research*, Longman Group Ltd., Londres (previamente publicado por Oliver & Boyd Ltd., Edinburgh), 1974, Tabla IV.
- Tabla I** F. Wilcoxon, S. Katte, y R. A. Wilcox, *Critical Values and Probability Levels for the Wilcoxon Rank Sum Test y the Wilcoxon Signed Ranks Test*, American Cyanamid Co., Nueva York, 1963, y F. Wilcoxon y R. A. Wilcox, *Some Rapid Approximate Statistical Procedures*, Lederle Laboratories, Nueva York, 1964, previamente publicado por, *Fundamentals of Behavioral Statistics*, 3a. ed., Addison-Wesley Publishing Company, Inc., Reading, Mass., 1976.
- Tabla J** RAND Corporation, *A Million Random Digits*, Free Press of Glencoe, Glencoe, Ill., 1955.



GLOSARIO

- Beta** Probabilidad de cometer un error de tipo II. (página 221)
- Coefficiente de correlación** Expresión cuantitativa de la magnitud y la dirección de una relación. (página 111)
- Conjunto exhaustivo de eventos** Conjunto que incluye todos los eventos posibles. (página 172)
- Curva asimétrica** Curva cuyos dos lados no coinciden si ésta se dobla por la mitad; es decir, una curva que no es simétrica. (página 57)
- Curva asimétrica en forma negativa** Curva en la cual la mayor parte de los datos se presentan en los valores más altos, de modo que la curva se reduce hacia el extremo inferior del eje horizontal. (página 58)
- Curva asimétrica en forma positiva** Curva en la cual la mayor parte de los datos se presentan en los valores más bajos, de modo que la curva se reduce hacia el extremo superior del eje horizontal. (página 58)
- Curva simétrica** Curva cuyos dos lados coinciden si la curva se dobla a la mitad. (página 57)
- Dato de desviación** Distancia del dato en bruto con respecto a la media de su distribución. (página 77)
- Dato z** Dato transformado que indica a cuántas unidades de desviaciones estándar se encuentra un dato, por arriba o por abajo de la media. (página 89)
- Datos** Medidas que se realizan sobre los sujetos de un experimento. (página 7)
- Distribución binomial** Distribución de probabilidad que resulta cuando se cumplen cinco condiciones: 1) existe una serie de N ensayos; 2) en cada ensayo sólo hay dos resultados posibles; 3) en cada ensayo, los dos resultados posibles son mutuamente excluyentes; 4) los resultados de cada ensayo son independientes entre sí, y 5) la probabilidad de cada resultado posible en cualquier ensayo es la misma de un ensayo a otro. La distribución binomial proporciona cada resultado posible de los N ensayos y la probabilidad de obtener cada uno de dichos resultados. (página 198)
- Distribución de frecuencias** Lista de valores de datos y su frecuencia de aparición. (página 39)
- Distribución de frecuencias acumuladas** Número de datos que quedan por debajo del límite superior real de cada intervalo. (página 45)
- Distribución de frecuencias relativas** Proporción del número total de datos que aparecen en cada intervalo. (página 45)
- Distribución de porcentajes acumulados** Porcentaje de datos que quedan por debajo del límite superior de cada intervalo. (página 45)
- Distribución muestral de F** Todos los valores posibles de F junto con la $p(F)$ de cada valor, suponiendo que la muestra extraída de la población es aleatoria. (página 355)
- Distribución muestral de la media** Lista de todos los valores que puede asumir la muestra, junto con la probabilidad de obtener cada valor si el muestreo es aleatorio, a partir de la población de la hipótesis nula. (página 269)
- Distribución muestral de t** Distribución de probabilidad de los valores de t que aparecerían si todas las diversas muestras posibles de un tamaño fijo N se extrajeran de la población de la hipótesis nula. Proporciona 1) todos los distintos valores de t para las muestras de tamaño N y 2) la probabilidad de obtener cada valor si el muestreo es aleatorio, a partir de la población de la hipótesis nula. (página 295)
- Distribución muestral de un estadístico** Lista de 1) todos los valores que puede asumir un estadístico y 2) la probabilidad de obtener cada valor, suponiendo que éste se debe solamente al azar, o si el muestreo es aleatorio a partir de la población de la hipótesis nula. (páginas 264, 268)

Efecto de interacción Resultado que se observa cuando el efecto de un factor no es igual en todos los niveles del otro factor. (página 397)

Efectos principales El efecto del factor *A* (promediado sobre los niveles del factor *B*) y el efecto del factor *B* (promediado sobre los niveles del factor *A*). (página 397)

Error de tipo I Efecto de la decisión de rechazar la hipótesis nula cuando ésta es verdadera. (página 220)

Error de tipo II Efecto de la decisión de conservar la hipótesis nula cuando ésta es falsa. (página 220)

Estadística Número calculado a partir de datos muestrales, que cuantifica una característica de la muestra. (página 7)

Estadística descriptiva Técnicas que se utilizan para describir o caracterizar los datos muestrales obtenidos. (página 10)

Estadística inferencial Técnicas que utilizan los datos muestrales obtenidos para hacer inferencias sobre poblaciones. (página 10)

Eventos mutuamente excluyentes Dos eventos que no pueden ocurrir al mismo tiempo; es decir, que la ocurrencia de uno impide la ocurrencia del otro. (página 168)

Experimento factorial Experimento en el cual se evalúan los efectos de dos o más factores, y los tratamientos utilizados son combinaciones de los niveles de los factores. (página 396)

Grados de libertad (gl) Número de datos que pueden variar libremente al calcular un estadístico. (página 296)

Gráfica de dispersión Gráfica de parejas de valores *X* y *Y*. (página 104)

Independencia de dos eventos Ocurrencia de un evento que no tiene efecto alguno sobre la probabilidad de que ocurra el otro evento. (página 173)

Intervalo de confianza Rango de valores que probablemente contiene el valor de la población. (página 305)

Límites de confianza Valores que establecen la frontera del intervalo de confianza. (página 305)

Límites reales de una variable continua Valores que están por arriba y por debajo del valor registrado, a una distancia igual a la mitad de la menor unidad de medida de la escala. (página 33)

Línea de regresión por mínimos cuadrados Línea de predicción que minimiza $\sum (Y - Y')^2$. (página 137)

Media aritmética Suma de los datos dividida entre el número de los mismos. En forma de ecuación,

$$\bar{X} = \frac{\sum X_i}{N} = \frac{X_1 + X_2 + X_3 + \dots + X_N}{N}$$

media de una muestra

$$\text{o bien } \mu = \frac{\sum X_i}{N} = \frac{X_1 + X_2 + X_3 + \dots + X_N}{N}$$

media de un conjunto poblacional de datos

donde X_1, \dots, X_N = los datos en bruto

\bar{X} (se lee "X barra") = la media de un conjunto muestral de datos

μ (se lee "mu") = la media de un conjunto poblacional de datos

Σ (se lee "sigma") = el signo de sumatoria

N = el número de datos

(página 67)

Mediana Valor de la escala debajo del cual se encuentra el 50% de los datos. (página 72)

Moda El dato más frecuente en la distribución. (página 74)

Muestra Subconjunto de la población. (página 6)

Muestra aleatoria Muestra elegida a partir de una población por medio de un procedimiento con el cual se tiene la seguridad de que 1) todas las muestras aleatorias de un tamaño determinado tengan la misma probabilidad de ser elegidas y 2) todos los miembros de la población tengan la misma probabilidad de ser incluidos en la muestra. (página 162)

Muestreo con reemplazo Método de muestreo en el cual cada miembro de la población que es elegido para la muestra es reintegrado a la población antes de elegir al siguiente miembro. (página 165)

Muestreo sin reemplazo Método de muestreo en el cual los miembros de la muestra no se reintegran a la población antes de elegir a los miembros subsiguientes. (página 165)

Nivel alfa Nivel de probabilidad que un investigador establece al inicio de un experimento, para limitar la probabilidad de cometer un error de tipo I. (páginas 218, 222)

Parámetro Número que se calcula a partir de los datos de una población, con el cual se cuantifica una característica de dicha población. (página 7)

Percentil Valor sobre la escala de medida, por debajo del cual queda un porcentaje dado de los datos, dentro de la distribución. (página 46)

P_{nula} Probabilidad de obtener un signo positivo con cualquier sujeto de la muestra del experimento

cuando la variable independiente no produce efecto alguno (es apropiada para la prueba de los signos). (*página 244*)

P_{real} Probabilidad de obtener un signo positivo con cualquier sujeto de la muestra del experimento cuando la variable independiente produce un efec-

to real; proporción de signos positivos dentro de la población si el experimento se realizó con toda la población y si la variable independiente produce un efecto real (es apropiada para la prueba de los signos). (*página 244*)



ÍNDICE

Los números de página seguidos de “n” se refieren a notas al pie de la página o al final del capítulo.

- Abscisa (eje X), 53
- Álgebra
- factorización de expresiones algebraicas, 497
 - resolución de ecuaciones con una incógnita, 497-498
 - símbolos utilizados en, 492
- Análisis de varianza de dos factores
- análisis de experimentos con, 406-420
 - comparaciones múltiples y, 420
 - efectos de interacción y, 397-399, 416
 - efectos principales y, 397
 - estimación de la varianza por columnas (s_C^2), 398-399, 400, 404, 409
 - estimación de la varianza dentro de las celdas (s_W^2) en, 398-399, 400, 402, 409
 - estimación de la varianza por renglón (s_R^2) en, 398-399, 400, 402-403, 409
 - estimación de la varianza por renglón (columna) (s_{RC}^2) en, 398-399, 400, 405, 409
 - experimento factorial y, 396-399
 - experimento sobre el ejercicio y el sueño, 406-411
 - notación y disposición general de los datos para, 399-400
 - panorama del, 398-399, 400
 - razones F para, 398-399, 400, 406, 409-410, 478-479
 - resumen del, 377, 421, 477-479
 - supuesto de la homogeneidad de la varianza y, 421
 - supuestos subyacentes, 421
 - variabilidad total/suma total de cuadrados (SC_T), 364-365, 400, 408-409, 477-478
- Análisis de varianza de un factor
- análisis de datos con, 362-366
 - comparaciones múltiples y, 381-392, 475-476
 - lógica subyacente, 366-367
 - magnitud del efecto y, 373
 - panorama de la técnica, 357-362, 474-475
 - relación entre la prueba t y, 370-372
 - resumen de, 376-377, 474-476
 - supuestos subyacentes, 372-373
- Análisis de varianza. Véase también
- Análisis de varianza de un factor; Análisis de varianza de dos factores
 - análisis de datos con, 362-366
 - supuestos subyacentes, 372-373
- Análisis exploratorio de los datos, 58-60
- Ansiedad a causa de las matemáticas y la estadística, 23-24
- Anuncio de Anacin-3, 13
- Anuncio de Excedrin, 235
- Anuncio sobre la hipnosis, 12
- Aplicaciones de la estadística en el mundo real, 11-16, 60-61, 126-127, 189-192, 232-237, 259-260, 375-376, 455-457
- Área bajo una curva normal, tabla sobre el tema, 525-528
- Astrología y ciencia, 259-260
- Autoridad
- como método de conocimiento, 4
 - uso de la misma en publicidad, 12
- Beta (β), 221, 223, 243-244, 251-252, 253
- Beveridge, W.I.B., 5
- Cálculos matemáticos
- exponentes, 497
 - factorización de expresiones algebraicas, 497
 - fracciones, 496
 - interpolación lineal, 499-500
 - operaciones aritméticas, 493-494
 - paréntesis y corchetes, 495
 - redondeo de, 34-35
 - repaso de, 491-500
 - residuo decimal en, 33-35
 - resolución de ecuaciones con una incógnita, 497-498
- Causalidad
- correlación versus, 125-126
 - relaciones entre causa y efecto, 104, 104n
- Ciencia y astrología, 259-260
- Coefficiente de correlación bi-erial (r_b), 121
- Coefficiente de correlación et al., 121

- Coeficiente de correlación ϕ , 121
- Coeficiente de correlación rho (r_s), 121-123, 309
- Coeficiente de determinación (r^2), 120, 340
- Coeficientes de correlación
 - coeficiente de correlación biserial (r_b), 121
 - coeficiente de correlación eta, 121
 - datos extremos y, 125
 - definición de, 111
 - escala de medición y, 121
 - signo negativo de, 111
 - r de Pearson, 111, 112-120
 - coeficiente de correlación ϕ , 121
 - signo positivo de, 111
 - gráficas de dispersión de, 111-112
 - forma de la relación y, 121
- Cola de la distribución, evaluación de una, 223-231
- Comparaciones *a posteriori*, 384-392, 476
- Comparaciones *a priori*, 382-383, 391-392, 475-476
- Comparaciones múltiples
 - comparaciones *a posteriori*, 384-392, 476
 - comparaciones *a priori*, 382-383, 475-476
 - introducción a las, 381-382
 - prueba de Newman-Keuls, 386-392, 476-477
 - análisis de varianza de un factor, 381-392, 475-476
 - análisis de varianza de dos factores y, 420
 - comparaciones ortogonales, 382, 393-394n
 - resumen de, 392
 - prueba DHS de Tukey, 385-386, 398-392, 476
- Comparaciones ortogonales, 382, 393-394n
- Comparaciones planeadas, 382-383, 391-392, 475-476
- Comparaciones *post hoc*, 384-392, 476
 - con eventos exhaustivos y mutuamente excluyentes, 172-173
 - con eventos mutuamente excluyentes, 168-172
 - con más de dos eventos mutuamente excluyentes, 172
 - utilizada conjuntamente con la regla del producto, 183-186
- Condición control, 216, 319, 327
- Condición experimental, 216, 319, 327
- Confiabilidad del método de test-retest, 104
- Confiabilidad
 - resultados significativos y, 219, 232
 - confiabilidad de la test-retest, 104
- Conocimiento
 - autoridad y, 4
 - intuición y, 5
 - método científico y, 5-6
 - racionalismo, 4-5
- Conservación de la hipótesis nula, 219
- Constante, 6
- Constantes de regresión
 - para la regresión de X sobre Y , 144, 145-146
 - para la regresión de Y sobre X , 137-138
 - r de Pearson y, 150-151
- Correlación múltiple al cuadrado (R^2), 153, 155
- Correlación múltiple de determinación (R^2), 153, 155
- Correlación múltiple, 151-155
 - aplicaciones en el mundo real, 126-127
 - coeficiente de correlación r de Pearson, 111, 112-120
 - correlación múltiple, 151-155
 - datos (puntajes) z y, 90, 112-115
 - gráficas de dispersión de coeficientes de correlación, 111-112
 - rango y, 124
 - rho de Spearman (r_s), 121-123
 - resumen de, 127-128
- Correlación
 - causalidad versus, 125-126
 - coeficiente de correlación eta, 121
 - comparación con la regresión, 104, 134
 - datos extremos y, 125
 - definición de coeficiente de correlación, 111
 - escala de medición y, 121
 - introducción a la, 103-104
 - y la dirección y el grado de una relación, 111
- Curva asimétrica
 - de una distribución de frecuencias, 57-58
 - tendencia central y, 75
- Curva asintótica, 87
- Curva de porcentaje acumulativo, 56-57
- Curva en forma de campana
 - curva normal, 86-88
 - distribución de frecuencias y, 58
 - tendencia central y, 75
- Curva en forma de J, 58
- Curva en forma de U, 58
- Curva normal
 - área contenida bajo la, 88
 - cálculo del área que corresponde a cualquier dato en bruto, 92-95
 - cálculo del dato en bruto que corresponde a un área dada, 97-99
 - características de los datos z , 91-92
 - como asíntota del eje horizontal, 87
 - datos estándar (puntajes z) y, 88-99
 - ecuación de, 87
 - importancia de la, 86
 - probabilidad y variables continuas normalmente distribuidas, 186-188
 - punto de inflexión de la, 87
 - resumen de, 100
 - tabla de áreas bajo la, 525-528
- Curva rectangular (uniforme), 57
- Curva simétrica de distribución de frecuencias, 57, 58
- Curva uniforme (rectangular), 58
- Curvas asimétricas negativamente sesgadas
 - tendencia central, 75
 - distribuciones de frecuencias, 57-58
- Curvas positivamente sesgadas
 - tendencia central, 75
 - distribuciones de frecuencias, 57-58
- Curvas, véase Gráficas
- Datos (puntajes) de desviación
 - cálculo de, 77-78
 - definición de, 76-77
 - para datos poblacionales, 77

- para datos muestrales, 77
- Datos (puntajes) z
 - cálculo de las áreas
 - correspondientes a cualquier dato en bruto, 92-95
 - cálculo del dato en bruto correspondiente a un área dada, 97-99
 - características de, 91-92
 - correlación y, 90, 112-115
 - definición de, 89
 - desviación estándar de, 92
 - ecuación de, 89
 - forma de los mismos en
 - comparación con la forma de los datos en bruto, 91
 - introducción a los, 88-89
 - para datos de población, 89-90
 - para datos muestrales, 89, 90-91
 - transformación de datos, 89
 - resumen de, 100
 - uso de, 90
- Datos agrupados, distribución de frecuencias de, 40-46
- Datos en bruto, 7. *Véase también* Datos
- Datos estándar (datos z)
 - cálculo del dato en bruto que corresponde a un área dada, 97-99
 - cálculo de las áreas correspondientes a cualquier dato en bruto, 92-95
 - características de los, 91-92
 - correlación y, 90, 112-115
 - definición de, 89
 - desviación estándar de, 92
 - ecuación para, 89
 - introducción a los, 88-89
 - media de, 92
 - para datos de población, 89-90
 - para datos muestrales, 89, 90-91
 - resumen de, 100
 - su forma en comparación con la forma de los datos en bruto, 91
 - transformación de datos, 89
 - uso de, 90
- Datos extremos, véase Datos
- Datos originales, 7. *Véase también* Datos
- Datos
 - definición de, 7
 - falta de datos en anuncios, 12
 - generalización indebida a partir de, 14
 - inexactos, 15
- Datos. *Véase también* Puntajes (datos) z
 - correlación y dato extremo, 125
 - datos (puntajes) de desviación y, 76-78
 - distribución de frecuencias de datos agrupados, 40-46
 - en una tabla de distribución de frecuencias, 38-39
 - sensibilidad de las medidas de tendencia central ante los datos extremos, 69, 73-74
- Decisión correcta sobre la hipótesis nula, 221, 253
- Depresión en las mujeres, 234
- Desarrollo binomial
 - cálculo de, 201
 - definición de, 200
 - ecuación para, 213n
 - generación de la distribución binomial a partir de él, 200-203
- Desviación al cuadrado, 77
- Desviación estándar
 - cálculo de la misma utilizando datos de desviación cuando la media no es un número entero, 79
 - cálculo de la misma utilizando el método de datos en bruto, 79-81, 85n
 - cálculo de la misma utilizando el método de la desviación, 78-79
 - datos de desviación y, 76-78
 - como medida de la dispersión en relación con la media, 80
 - de datos de población utilizando el método de desviación, 78
 - de la distribución muestral de la media, 270-271
 - de datos muestrales utilizando el método de la desviación, 78-79
 - de puntajes (datos) z , 92
 - de una muestra utilizando el método de datos en bruto, 79-81, 85n
 - propiedades de la, 80
 - sensibilidad de la misma con respecto a cada dato de la distribución, 80
 - símbolos de, 78
 - variabilidad de una muestra y, 80
- Diagrama de flujo para elegir la prueba apropiada, 480-482
- Diagramas de tallo y hojas, 58-60
- Diferencia entre medias muestrales, 329-331
- Directores de escuela, 126-127
- Directores de escuelas elementales, 126-127
- Directores de escuelas elementales, 126-127
- Diseño de efectos fijos, 397
- Diseño de grupos aleatorios simples, 358, 395-396
- Diseño de grupos correlacionados
 - en comparación con el diseño de grupos independientes, 342-345, 352n4
 - descripción de, 217
 - prueba de los signos y, 217, 319
 - prueba t para, 319-327, 341-345, 352n
- Diseño de grupos independientes. *Véase también* Prueba U de Mann-Whitney; Análisis de varianza de una función; Prueba t (grupos independientes)
 - comparación con el diseño de grupos correlacionados, 342-345
 - descripción de, 327-328, 472
 - prueba t para, 327-328, 331-345, 352-353n
 - prueba z para, 328-332
- Diseño de medidas repetidas (replicadas), 217, 327, 352n2, 470. *Véase también* Prueba t (grupos correlacionados); Prueba t (grupos independientes)
- Disminución del recuento de esperma, 190
- Dispersión, la desviación estándar como medida de, 80
- Distribución binomial
 - condiciones apropiadas para la, 471
 - definición de, 198
 - ejemplo de, 198-200
 - examen de opción múltiple como ejemplo ilustrativo, 209
 - evaluación del experimento de la marihuana mediante, 219-220

- generación de la distribución binomial a partir del desarrollo binomial, 200-203
- lanzamiento de una moneda como ejemplo ilustrativo, 198-208
- prueba de la degustación de café como ejemplo ilustrativo, 210
- resumen de, 211, 471
- tabla binomial empleada para, 203-210
- tabla de, 529-533
- Distribución de frecuencias relativas, 45-46
- Distribución de la frecuencia acumulativa, 45, 46
- Distribución del porcentaje acumulativo, 45, 46
- Distribución F , 354-356
- tabla de, 540-542
- Distribución muestral de F , 355-356
- Distribución muestral de la diferencia entre medias muestrales, 329-331
- Distribución muestral de la media características de la, 270-275
- definición de, 269
- deducción empírica de la, 269-275
- deducción teórica de la, 269
- desviación estándar de la, 270-271
- forma de la, 270, 271, 273, 275
- media de la, 270
- para el experimento de la habilidad para la lectura, 275-277
- resumen de, 290
- Distribución muestral de t , 295-297
- Distribución muestral de un estadístico, 264
- Distribución Q , tabla de, 543
- Distribución t
- comparada con la distribución z , 297-298
- construcción de un intervalo de confianza de 95%, 305-307
- estimación del CI medio de profesores universitarios, 306-307
- intervalos de confianza para la media poblacional, 304-309
- tabla de, 538
- Distribución z comparada con la distribución t , 297-298
- Distribuciones de frecuencias
- análisis de datos exploratorio, 58-60
- aplicaciones en el mundo real, 60-61
- construcción de gráficas de, 53-58
- definición de, 38, 39
- de datos agrupados, 40-46
- diagramas de tallo y hojas, 58-60
- distribución de frecuencias
- acumuladas, 45, 46
- distribuciones de frecuencias no agrupadas, 38-39
- distribución de porcentajes acumulados, 45, 46
- distribución de frecuencias relativas, 45-46
- forma de las curvas de frecuencias, 57-58
- percentiles, puntos percentiles, 46-50
- rango percentil, 50-52
- resumen de, 62
- su construcción para datos agrupados, 41-45
- Distribuciones muestrales. Véase también Prueba z
- definición de, 268, 466
- definición de la distribución muestral de la media, 269
- definición de la distribución muestral de un estadístico, 264
- de la diferencia entre medias muestrales, 329-331
- distribución muestral de f , 355-356
- distribución muestral de la media, 268-275
- distribución muestral de t , 295-297
- ecuación para calcularla, 168, 169
- generación de, 264-268
- introducción a las, 263-264
- población de la hipótesis nula, 265
- resumen de, 280
- Ecuaciones
- con una incógnita, 497-498
- de la recta, 105-107
- lista de, 501-508
- Efecto real, 244-251, 254-259, 288-289
- potencia y, 374
- Efectos de interacción, 397-399, 416
- Efectos principales, 397
- Eje X (abscisas), 53
- Eje Y (ordenadas), 53
- Encuesta presidencial de 1936 del Literary Digest, 163
- Encuesta presidencial de 1936, 163
- Encuestas, 163, 191-192
- Equivalencia y escalas nominales, 28
- Error de la media cuadrada, 358n
- Error de tipo I
- nivel alfa (α) y, 222-223, 225, 243, 262n
- análisis de varianza y, 357
- tasa de error por comparación y, 384, 386
- definición de, 220, 264, 467
- tasa de error por experimento y, 384, 386, 392
- prueba de hipótesis utilizando la prueba de los signos, 220-223, 226
- prueba de Newman-Keuls y, 391-392
- probabilidad de cometerlo, 262n
- prueba DHS de Tukey y, 385, 386, 391-392
- Error de tipo II
- beta (β) y, 221-223, 243-244
- definición de, 220, 264, 467
- Error estándar de la estimación, 146-149, 147n
- Error estándar de la media
- definición de, 270
- error estándar estimado de la media, 294
- Error estándar estimado de la media, 270
- Errores de predicción, error estándar de la estimación, 146-149, 147n
- Errores
- error de tipo I, 220-223, 226, 243, 262n, 264, 357, 384, 385, 467
- error de tipo II, 220-221, 223, 243-244, 264, 326, 392, 467
- error cuadrado medio, 358n
- errores de predicción, 146-149
- error estándar del estimado, 146-149, 147n
- error estándar de la media, 270

- error estándar estimado de la media, 270
- tasa de errores por comparación, 384-385
- tasa de errores por experimento, 384-385
- Escala Celsius, 29-30
- Escala de intervalos
definición y descripción de, 29
 r de Pearson para, 121
- Escala de razón (proporción)
punto cero absoluto en la, 29-30
definición y descripción de, 29-30
 r de Pearson y, 121
- Escala Kelvin, 30
- Escala nominal, 27-28
- Escala ordinal
definición y descripción de, 28-29
 ρ de Spearman (r_s), 121-123
- Escala Wechsler para la inteligencia de adultos (WAIS), 31
- Escalas de medición
coeficiente de correlación y, 121
en las ciencias del comportamiento, 31
escala de intervalo, 29
escala nominal, 27-28
escala ordinal, 28-29
escala de proporción (razón), 29-30
- Escalas. *Véase* Escalas de medición
- Estadística (singular), definición de, 7
- Estadística descriptiva
correlación, 103-126
curva normal, 86-88
datos estándar (valores z), 88-99
definición de, 10
distribuciones de frecuencias, 38-62
escalas de medición, 27-31
mediciones de la variabilidad, 76-82
medidas de la tendencia central, 67-75
notación matemática, 24-25
regresión lineal, 134-155
sumatoria, 25-27
variables continuas y discretas, 31-35
- Estadística inferencial. *Véase también* Análisis de varianza de un factor; Análisis de varianza de dos factores
comparaciones múltiples, 381-420
- definición de, 10
- distribución binomial, 197-211
- distribuciones muestrales, 263-290
- elección de la prueba apropiada, 480-482
- ji-cuadrada, 427-440, 479-480
- muestreo aleatorio, 162-165
- potencia, 243-261
- probabilidad, 166-188
- prueba de hipótesis mediante la prueba de los signos, 214-238, 466-468
- prueba de Kruskal-Wallis, 450-454, 477
- prueba de rangos con signo y pares igualados de Wilcoxon, 441-444, 471
- prueba t para grupos correlacionados e independientes, 317-346, 470, 472-473
- prueba t para muestras simples, 293-313, 469
- prueba U de Mann-Whitney, 444-450, 473
- pruebas de la desviación normal (z), 268-290, 468
- pruebas no paramétricas, 425-458
- repaso de, 465-482
- Estadística. *Véanse también* Técnicas estadísticas específicas
abusos de la 455-457
ansiedad frente a la, 23-24
aplicaciones en el mundo real, 11-16, 60-61, 126-127, 189-192, 232-237, 259-260, 375-376, 455-457
estadística descriptiva, 10
estadística inferencial, 10
sugerencias para el estudio de la, 23-24
uso de la computadora en, 10-11
- Estado de realidad, 221, 253
- Estimación de intervalos, 304
- Estimación de la variación por columnas (S_C^2), 398-400, 404, 409, 479
- Estimación de la varianza dentro de las celdas (s_W^2), 398-399, 400, 402, 409, 479
- Estimación de la varianza entre grupos (s_B^2), 360-361, 365, 475
- Estimación de la varianza intra grupos (s_W^2), 358-360, 365, 475
- Estimación de la varianza por renglón (columna) (S_{RC}^2), 398-400, 405, 409, 479
- Estimación de la varianza por renglón (s_R^2), 398-399, 400, 402-403, 479
- Estimación puntual, 304
- Estudio de correlación, 9
- Estudios de observación, 9
- Eventos dependientes
definición de, 179
regla del producto con, 179-183
regla del producto con más de dos eventos dependientes, 182-183
- Eventos exhaustivos
regla de la suma con, 172-173
definición de, 172
- Eventos independientes
definición de, 173
regla del producto con, 173-179
regla del producto con más de dos eventos independientes, 178-179
- Eventos mutuamente excluyentes
definición de, 168
regla de la suma con, 168-172
regla del producto con, 173
- Eventos Q . número de, 201-203, 207-208
- Experimento de un solo factor.
diseño de grupos independientes, 358, 395-396
- Experimento factorial, 396-398
- Experimento sobre el desarrollo intelectual y la dieta, 444-447
- Experimento sobre el ejercicio y el sueño, 406-411
- Experimento sobre el sueño y el ejercicio, 406-411
- Experimento sobre la afiliación política y la actitud, 432-436
- Experimento sobre la dieta y el desarrollo intelectual, 444-447
- Experimento sobre la emisión de la primera palabra, 294-295, 298-299
- Experimento sobre la estimulación del cerebro y la alimentación, 319-322
- Experimento sobre la habilidad para la lectura, 268, 275-280, 283-289

- Experimento sobre la hormona X y el comportamiento sexual, 328-331, 333-334
- Experimento sobre la marihuana con pacientes con SIDA, 215-220, 246-253
- Experimento sobre la primera palabra que emiten los niños, 294-295, 298-299
- Experimento sobre la reducción de peso, 450-452
- Experimento sobre la tensión, 362-366
- Experimento sobre las actitudes ante la conservación de la vida silvestre, 441-443
- Experimento sobre las citas, 370-372
- Experimento sobre los pacientes con SIDA y la marihuana, 215-220, 246-253
- Experimento sobre marcas de cerveza clara, 427-430
- Experimento sobre marcas de cerveza, 427-430
 - efectos colaterales del implante en el control de la natalidad, 8
 - experimento sobre la estimulación cerebral y la alimentación, 319-322
 - experimento sobre la emisión de la primera palabra por los niños, 294-295, 298-299
 - experimento sobre las citas, 370-372
 - experimento sobre la dieta y el desarrollo intelectual, 444-447
 - experimento sobre el ejercicio y el sueño, 406-411
 - experimento sobre la hormona X y el comportamiento sexual, 328-331, 333-334
 - experimento sobre la marihuana con pacientes de SIDA, 215-220, 246-253
 - experimento sobre marcas de cerveza, 427-430
- Experimentos científicos (ejemplos)
- Experimentos científicos. *Véase también* Prueba de hipótesis; Experimentos científicos (ejemplos)
 - condición control en, 216, 319, 327
 - condición experimental en, 216, 319, 327
 - definiciones en, 6-7
 - diseño de medidas repetidas, 217
 - distribución binomial para la evaluación de, 219-220
 - errores de tipo I y tipo II en, 220-223, 264, 326, 357, 384, 391-392, 467
 - hipótesis alternativa en, 218, 242n, 466
 - hipótesis nula en, 218-219, 232, 242n, 466
 - informes anecdóticos versus, 236-237
 - regla de decisión y nivel alfa, 218-219, 221-223, 466
 - repaso de, 5-6
 - resultados significativos de, 219, 232, 253-254
 - resumen de, 17
 - ventajas de los experimentos con dos condiciones, 318
- Experimentos con dos condiciones, 217, 318, 327, 352n2, 470. *Véase también* Prueba t (grupos correlacionados); Prueba t (grupos independientes)
- Experimentos con una sola muestra. *Véanse también* Prueba t (una sola muestra); Prueba z
- Experimentos con una sola variable, ji-cuadrada (χ^2) para, 427-431
- Experimentos. *Véanse también* Experimentos científicos; Experimentos científicos (ejemplos)
- Exponentes, 497
- Factorización de expresiones algebraicas, 497
- F_{crit} , 359, 362, 365, 399, 406, 475
- Fisher, R. A., 354
- Fraciones, 496
- Frecuencias esperadas, 427-428
- Frecuencias observadas, 427-428
- gl_C (grados de libertad por columna), 405, 409, 478
- gl_R (grados de libertad por renglón), 402-403, 409, 478
- gl_{RC} (grados de libertad por renglón (columna)), 405, 409, 479
- gl_W (grados de libertad dentro de las celdas), 400-402, 409, 480
- Gosset, W.S., 294, 296
- Grado de separación, 447
- Grados de libertad dentro de las celdas (gl_W), 400, 402, 409, 479
- Grados de libertad por columnas (gl_C), 405, 409, 478
- Grados de libertad por renglón (columna (gl_{RC})), 405, 409, 479
- Grados de libertad por renglón (gl_R), 402-403, 478
- Grados de libertad
 - grados de libertad dentro de las celdas (gl_W), 400-402, 409
 - grados de libertad por columnas (gl_C), 405, 409, 478
 - grados de libertad por renglón (gl_R), 402-403, 409, 478
 - grados de libertad por renglón (columna (gl_{RC})), 405, 478
 - prueba t (grupos independientes), 345
 - prueba t (una sola muestra), 296-297
 - ji-cuadrada y, 428-430, 464n
- Gráfica de barras, 53, 55
- Gráficas de dispersión
 - coeficientes de correlación y, 111-112
 - estudio de correlación, 340-341
 - experimento con grupos independientes, 340-341
 - predicción y relaciones imperfectas, 135, 136
 - relaciones imperfectas y, 109-110
 - relaciones lineales y, 104-105
- Gráficas. *Véase también* Curva normal; Gráficas de dispersión
 - construcción de gráficas en general, 53
 - curva de porcentaje, 56-57
 - curvas F , 356
 - de distribuciones de frecuencias, 53-58
 - eje X (abscisa) de, 53
 - eje Y (ordenada) de, 53
 - gráfica de barras, 53, 55
 - graficación de datos en general, 53
 - gráfica de dispersión, 104-105
 - histograma, 55-56, 60, 74-75
 - líneas paralelas en las, 416

- polígono de frecuencias, 56
tendencia central, 75
- Gran media, 360-361, 366-367
- Grupo control, 327
- Grupo experimental, 327
- $H_{crit.}$, 452
- Hipótesis alternativa
definición de, 218, 466
hipótesis direccional, 218, 223-230, 242n
hipótesis no direccional, 218, 223-224, 226, 227-228, 230-231
y poder de cálculo, 256-259
- Hipótesis direccional, 218, 223-230, 242n
- Hipótesis no direccional, 218, 223-224, 226, 227-228, 230-231
- Hipótesis nula (H_0)
análisis de varianza de dos factores y , 406
análisis de varianza de un factor y , 359
decisión correcta sobre la, 221, 253
definición y descripción de, 218-219, 466
estado de realidad y la, 221
error de tipo I y, 220-223, 226, 264, 467
error de tipo II y, 220-221, 223, 264, 326, 467
media de la población de la hipótesis nula (μ_{nula}), 284-289
incapacidad de rechazarla (o conservarla), 219, 253-254, 264
población de la hipótesis nula, 265, 466
prueba t (grupos independientes), 328
rechazo de la, 218-219, 232, 242n, 264
región crítica para el rechazo de la hipótesis nula, 277-282, 466
resultados no significativos y, 253-254
- Histograma bimodal, 74-75
- Histograma unimodal, 74-75
- Histograma
comparación con el diagrama de tallo y hojas, 60
descripción de, 55-56
- histograma bimodal, 74-75
- histograma unimodal, 74-75
- Hobt, 451-454
- Homoscedasticidad, 147-148
- Imposibilidad de rechazar la hipótesis nula, 219, 253-254, 264
- Interpolación lineal, 499-500
- Intervalo de confianza de 95%, 305-307
- Intervalos de confianza
construcción de un intervalo de confianza de 95%, 305-307
definición de, 305
estimación de la media del CI de profesores universitarios, 306-307
ecuaciones generales para cualquier intervalo de confianza, 307-309
para la media poblacional y la prueba t (una sola muestra), 304-309
- Intervalos en distribuciones de frecuencias, 40-44
- Intuición, 5
- Investigación científica. *Véase también* Experimentos científicos
informes anecdóticos versus, 236-237
investigación social aplicada, 455-457
estudios de observación, 9
resumen de, 17
experimentos verdaderos, 9
- Investigación de observación de la naturaleza, 9
- Investigación para la estimación de parámetros, 9, 162
- Investigación social aplicada, 455-457
- Investigación. *Véanse también* Experimentos científicos; Experimentos científicos (ejemplos); Investigación científica
- CI de profesores universitarios, 306-307
- Ji-cuadrada (χ^2)
cálculo de $\chi^2_{obt.}$, 427-428, 433-435, 479
- experimento sobre la afiliación política y la actitud, 432-436
- experimento sobre marcas de cerveza, 427-430
- evaluación de $\chi^2_{obt.}$, 428-430, 435-436
- experimentos con una sola variable, 427-431
- prueba de la independencia entre dos variables, 431-440
- resumen de, 457-458, 479-480
- supuestos subyacentes, 440
- tabla de, 544
- Juegos de dados, 166
- Keppel, G., 382
- Límites de confianza, 305
- Lanzamiento de monedas
monedas sin cargar (honestas), 175-176, 198-200, 203-205
para ilustrar la distribución binomial, 198-208
- Límite de confianza inferior, 307-308
- Límite de confianza superior, 307-308
- Límite inferior que corresponde a un intervalo de confianza de 95%, 306-307
- Límite inferior real de una variable continua, 33
- Límite real superior de una variable continua, 33
- Límite superior para un intervalo de confianza de 95%, 306-307
- Límites reales de una variable continua, 32-33
- Línea recta
ecuación de la, 105-107
ordenada al origen de una, 106
predicción de Y dada X , 107
pendiente (b) de una, 106-107
relaciones perfectas e imperfectas y, 108-110
- Línea. *Véase también* Relaciones lineales; Línea recta
- Magnitud del efecto
análisis de varianza y, 373
potencia y magnitud del efecto real, 244-251, 288-289
resultados significativos versus importantes, 232

- prueba t (grupos independientes) y, 339-341
- Máquina tragamonedas. 184-186
- Matemáticas, ansiedad ante las. 23-24
- Media aritmética
 - cálculo de, 67
 - de datos (puntajes) z , 92
 - definición de, 67
 - de un conjunto poblacional de datos, 67-68
 - de una muestra, 67-68
 - media global, 70-71
 - propiedades de, 68-70
 - símbolos de, 68
 - su sensibilidad frente a datos extremos, 69
 - su sensibilidad para el valor exacto de los datos incluidos en la distribución, 68
 - suma de cuadrados de las desviaciones de todos los datos en torno de su media, 69-70
 - suma de desviaciones en torno a la media, 68-69
 - variación muestral y, 70
- Media cuadrada dentro, 358n
- Media cuadrada entre, 358n
- Media de la distribución muestral de la media, 270
- Media de la población de la hipótesis nula (μ_{nula}), 284-289
- Media de puntajes de diferencia, 320
- Media del CI de profesores universitarios, 306-307
- Media global, 70-71
- Media poblacional, intervalos de confianza para la, 304-309
- Media
 - de datos z , 92
 - de un conjunto poblacional de datos, 67-68
 - de una muestra, 67-68
 - desviación estándar como medida de la dispersión, en relación con la, 80
 - gran media, 360-361, 366-367
 - media aritmética, 67-70
 - media global, 70-71
 - propiedades de la, 68-70
 - símbolos de, 68
 - simetría y, 75
- símbolos de, 68
- simetría y, 75
- su sensibilidad al valor exacto de los datos incluidos en la distribución, 68
- su sensibilidad a los datos extremos, 69
- suma de desviaciones cuadradas de todos los datos en torno a sus medias, 69-70
- suma de desviaciones en torno a la media, 68-69
- variación de muestreo y, 70
- Mediana
 - cálculo de la, 72-73
 - definición de, 72
 - propiedades de la, 73-74
 - símbolo de la, 72
 - simetría y, 75
 - su sensibilidad a los datos extremos, 73-74
 - variabilidad muestral y, 74
- Mediana, 72-75
- Medicamento contra el autismo, 236-237
- Medición del IQ, 31
- Memoria y forma de presentar un pasaje en prosa, 7-8
 - experimento sobre la afiliación política y la actitud, 432-436
 - experimento sobre la habilidad para la lectura, 268, 275-280, 283-289
 - experimento sobre la reducción de peso, 450-452
 - experimento sobre la tensión, 362-366
 - experimento sobre las actitudes ante la conservación de la vida silvestre, 441-443
 - obesidad y alta presión arterial, 9
- Método científico. *Véase también* Experimentos científicos
 - definición y descripción del, 5-6
 - resumen del, 17
- Método de desviación para la desviación estándar, 78-79
- Método de la autoridad, 4
- Método de la diferencia directa para la prueba t , 319-327, 352n1
- Método de los datos en bruto para la desviación estándar, 79-81, 85n
- MINITAB, 11
- Moda
 - cálculo de la, 74
 - definición de, 74
 - variabilidad muestral y, 74
 - simetría y, 75
- Monedas cargadas, 205-208
- Monedas no cargadas (honestas), 175-176, 198-200, 203-205
- Monedas sin cargar, 175-176, 198-200, 203-205
- Mu (μ), para la media del conjunto poblacional de datos, 67-68
- Muestra asimétrica (sesgada), 163
- Muestra
 - datos (puntajes) z para una, 89, 90-91
 - definición de, 6
 - media de una, 67-68
 - puntajes de desviación para, 77
 - su desviación estándar y el uso del método de datos en bruto, 79-81, 85n
 - su desviación estándar y el uso del método de desviación, 78-79
 - su tamaño y la potencia, 283-287
 - varianza de una, 82
- Muestreo aleatorio. *Véase también* Probabilidad
 - aplicaciones en el mundo real, 189-192
 - definición de, 9, 162
 - encuestas y, 163, 191-192
 - muestreo con sustitución, 163, 165, 173-174
 - muestreo sin sustitución, 165
 - tabla de números aleatorios para, 164
 - técnicas de, 164-165
 - resumen de, 193
 - razones para realizarlo, 163
- Muestreo con reemplazo, 163, 165, 173-174
- Muestreo sin reemplazo, 165
- Muestreo. *Véase también* Probabilidad; Muestreo aleatorio
- N (tamaño de la muestra)
 - potencia y, 247-251, 283-287, 374
 - símbolo de, 25
- Nivel alfa (α)
 - definición de, 222, 466
 - error de tipo I y, 222-223, 225, 243, 262n
 - potencia y, 252-253, 287-288

- regla de decisión y, 218-219, 221-223
y probabilidad con una o dos colas, 225-226
- Notación matemática, 24-25, 492
- Número de eventos P , 201-207
- Número de eventos P , 201-207
- Número de eventos Q , 201-202, 207-208
- Omega cuadrada, 341, 373
- Operaciones aritméticas. *Véase también* Cálculos matemáticos
- Ordenada (eje Y), 53
- Ordenada al origen, 106
- Paquete de estadística para las ciencias sociales (SPSS), 11
- Parámetro, definición de, 7, 426
- Pendiente (b)
de una línea recta, 106-107
relaciones positivas y negativas y, 107-108
- Percentiles, puntos percentiles
cálculo de, 47-50
definición de, 46
ecuación para calcular los, 49
 P_{nula} , 244-245
- Población de la hipótesis nula.
Véase también Prueba de desviación normal (z)
definición de, 265, 466
media de (μ_{nula}), 284-289
- Población
datos de desviación para, 77
definición de, 6
desviación estándar poblacional, uso del método de desviación, 78
población de la hipótesis nula, 265, 466
puntajes (datos) z para, 89-90
media del conjunto poblacional de datos, 67-68
varianza de, 82
- Polígono de frecuencias, 56
- Potencia
alfa (α) y, 252-253, 287-288
análisis del experimento sobre el SIDA, 246-253
aplicación en el mundo real, 259-260
beta (β) y, 251-252, 253
cálculo de la, 254-259
características de, 283
de la prueba t , 341-342
definición de, 244, 282-283, 467
del análisis de varianza, 373-374
e interpretación de resultados no significativos, 253-254
efecto de N y magnitud del efecto real, 247-251, 374
efecto real y, 244-251, 254-259, 288-289, 374
medición de la magnitud y dirección del efecto real, 245-246
 P_{nula} , 244-245
 P_{real} , 244-246
prueba z y, 282-289
tamaño de la muestra y, 283-287, 374
- Predicción
de Y dada X para una línea recta, 107
error estándar de estimación, 146-149, 147n
recta de regresión para, 110, 119
recta de regresión por mínimos cuadrados, 135-143
regresión lineal para, 135-150
regresión múltiple y, 152-155
relaciones imperfectas y, 135-137
relaciones perfectas y, 134
Preal, 244-246, 254-259
- Probabilidad *a posteriori*, 166-167
- Probabilidad *a priori*, 166, 167
- Probabilidad con una cola, 223-227, 228-230
- Probabilidad de dos colas, 224-228, 230-231
- Probabilidad. *Véase también*
Distribución binomial;
Muestreo aleatorio
aplicaciones en el mundo real, 189
cálculo de, 167-186
como fracción de un número decimal, 167
probabilidad *a posteriori*, 166-167
probabilidad *a priori*, 166-167
probabilidad con dos colas, 224-228, 230-231
probabilidad con una cola, 223-227, 228-230
puntos básicos sobre valores de probabilidad, 167
regla de la suma para, 168-173, 199
regla del producto para, 173-183, 199
reglas del producto y de la suma empleadas ambas para la, 183-186
resumen de, 193
y variables continuas normalmente distribuidas, 186-188
- Programas de computación
MINITAB, 11
paquete de estadística para las ciencias sociales (SPSS), 11
programas de computación biomédicos serie P , 11
sistema de análisis estadístico (SAS), 11
SYSTAT, 11
- Programas de la serie P de computación biomédica, 11
- Prueba de degustación de Coca versus Pepsi, 232-233
- Prueba de diferencia honestamente significativa (DHS) de Tukey, 385-386, 389-392, 476
prueba de hipótesis utilizando la prueba de los signos, 220-221, 326
prueba de Newman-Keuls y, 392
- Prueba de hipótesis. *Véanse también* Prueba t ; Prueba z
aplicaciones en el mundo real, 232-237
diseño de medidas repetidas, 217
distribución binomial para su evaluación, 219-220
errores de tipo I y tipo II en, 220-223, 264, 326, 467
experimento sobre marihuana con pacientes con SIDA como ejemplo de, 215-220
hipótesis alternativa en la, 218, 242n, 466
hipótesis nula en, 218-219, 232, 242n
introducción a la, 162, 214-215
magnitud del efecto, 232
probabilidad con dos colas y, 224-228, 230-231
probabilidad con una cola y, 224-226, 228-230
procedimiento para, 467-468
prueba con signos, 216-237, 264, 471-472
regla de decisión y nivel α fa, 218-219, 221-223, 466
resultados significativos y, 219, 232, 253-254
resumen de, 227-238, 467-468
términos y conceptos en, 466-467

- Prueba de Kruskal-Wallis, 450-454, 458, 477
- Prueba de la desviación normal (z)
 comparación con la prueba t , 294
 condiciones apropiadas para, 282-289, 293-294
 distribución muestral de la media y, 268-275
 ecuación para la, 294
 experimento sobre la habilidad para la lectura, 268, 275-281, 283-289
 magnitud del efecto real y potencia, 288-289
 nivel alfa y potencia, 287-288
 potencia y, 282-289
 región crítica para el rechazo de la hipótesis nula, 277-282
 resumen de, 290
 supuesto matemático subyacente, 282
 tamaño de la muestra y potencia, 283-287
 valor crítico de un estadístico, 277-282
- Prueba de la probabilidad exacta de Fisher, 440
- Prueba de los signos
 comparación con la prueba t (grupos correlacionados), 326
 descripción de la, 216-217, 471-472
 diseño de grupos correlacionados y, 217, 319
 para la prueba de hipótesis, 216-237, 264, 471-472
 requisitos de la, 426
 sumatoria en la, 227-238, 471-472
- Prueba de Newman-Keuls, 386-392, 476-477
- Prueba de rangos con signo y pares igualados de Wilcoxon, 441-444, 458, 471
 tabla de la, 545
- Prueba del desafío de degustación de Pepsi, 232-233
- Prueba DHS. *Véase también* Prueba de las diferencias honestamente significativas de Tukey (DHS)
- Prueba F , 356-357, 474-475, 478-479
- Prueba k , 450-451, 477
- Prueba robusta, 339, 426
- Prueba t (grupos correlacionados)
 supuestos subyacentes, 327
 comparación con la prueba de los signos, 326
 comparación con la prueba t (grupos independientes), 342-345, 352n4
 comparación con la prueba t (una sola muestra), 320-321
 ecuación de t_{obt} , 341-342, 470
 experimento con medidas repetidas, 352n2
 experimento sobre el estímulo cerebral y la alimentación, 319-322
 método de la diferencia directa para, 319-327, 352n1
 potencia de la, 341-342
 requisitos de la, 426
 resumen de la, 345-346, 470
- Prueba t (grupos independientes)
 para comparaciones *a priori* o planeadas, 382-383, 391-392, 475-476
 supuestos subyacentes, 338-339
 cálculo de t_{obt} cuando $n_1 + n_2$, 334-338, 472
 como una prueba robusta, 339
 comparación con la prueba t (grupos correlacionados), 342-345
 comparación con la prueba z , 331-333
 diseño de grupos independientes y, 327-328
 grados de libertad y, 345
 ecuación de la, 331-333, 472-473
 ecuación de t_{obt} , 341-342, 472
 y homogeneidad del supuesto de la varianza, 338-339
 experimento sobre la hormona X y el comportamiento sexual, 333-334
 potencia de la, 341-342
 relación entre el análisis de varianza y la, 370-372
 requisitos de la, 426
 magnitud del efecto y, 339-341
 resumen de la, 345-346, 472-473
 violación de los supuestos de la, 339
- Prueba t (una sola muestra)
 cálculo de t_{obt} a partir de datos originales, 299-304, 315n
 comparación con la prueba t (grupos correlacionados), 320-321
 comparación con la prueba z , 294
 condiciones apropiadas para la, 304, 426
 experimento sobre la emisión de la primera palabra por los niños, 294-295, 298-299
 distribución muestral de t , 295-297
 distribución z comparada con la distribución t , 297-298
 ecuación de la, 294-469
 ecuación de t_{obt} , 341-342, 469
 error estándar estimado de la media, 294
 grados de libertad y, 296-297
 potencia de la, 341-342
 para comprobar la significación de la r de Pearson, 309-312, 469
 resumen de la, 312-313, 469
- Prueba t de Student. *Véase también* Prueba t
- Prueba U de Mann-Whitney como opción alternativa a la prueba t para grupos independientes, 339
 experimento sobre la dieta y el desarrollo intelectual, 444-447
 supuestos subyacentes, 450
 rangos empatados y, 448-450
 resumen de, 458, 473
- Prueba U . *Véase* Prueba U de Mann-Whitney
- Prueba z
 condiciones apropiadas para la, 282-289, 293-294, 426
 distribución muestral de la media y, 268-275
 ecuación de la, 294, 468
 en comparación con la prueba t , 294, 331-332
 experimento sobre la habilidad para la lectura, 268, 275-280, 283-289
 para grupos independientes, 328-332
 potencia y, 282-289
 magnitud del efecto real y potencia, 288-289
 nivel alfa y potencia, 287-288
 región crítica para el rechazo de la hipótesis nula, 277-282
 resumen de, 290, 468
 supuestos matemáticos subyacentes, 282

- tamaño de la muestra y potencia, 283-287
valor crítico de un estadístico y, 277-282
- Pruebas no paramétricas, 426, 457-458. *Véase también* Ji-cuadrada (χ^2); Prueba de Kruskal-Wallis; Prueba *U* de Mann-Whitney; Prueba de los signos; Prueba de rangos con signo y pares igualados de Wilcoxon
- Pruebas paramétricas, 426, 457. *Véase también* Prueba *F*; Prueba *t*; Prueba *z*
- Pruebas que prescinden de la distribución, 426
- Publicidad, 12-13, 233-235, 375-376
- Puget Power & Light Company, 60-61
- Puntajes (datos) de diferencia, 320-321
- Punto cero absoluto en una escala de proporción (razón), 29-30
- Puntos de inflexión de una curva normal, 87
- Q_{crit} , 385-391, 476
 Q_{obt} , 385-391, 476
- r* de Pearson
cálculo de, 115-116, 117n
constantes de regresión y, 150-151
definición de, 113
relación de r^2 y variabilidad explicada, 120
prueba *t* para comprobar la significación de, 309-312, 469
puntajes (datos) *z* y, 112-115
tabla de valores críticos de la, 539
y variabilidad de *Y* explicada por medio de *X*, 118-120
- R^2 (correlación múltiple cuadrada), 153, 155
- Racionalismo, 4-5
- Rango percentil
cálculo del, 50-52
definición de, 50
ecuación para calcularlo, 51
- Rango
cálculo del, 76
como medida de la variabilidad, 76
correlación y, 124
definición de, 76
en distribuciones de frecuencias, 41, 43, 44
- Rangos empatados, 448-450
- Razón *F*
para el análisis de varianza de dos factores, 399-400, 406, 409-410, 478-479
para el análisis de varianza de un factor, 358-359, 362, 365, 474-475
- Razonamiento. *Véase también* Racionalismo
- r_b (coeficiente de correlación biserial), 121
- Recta de regresión de *X* sobre *Y*, 143-146
- Recta de regresión de *Y* sobre *X*, 137-143
- Recta de regresión por mínimos cuadrados
construcción de, 137-143
definición de, 137
ecuación para la, 137-138
predicción y relaciones imperfectas, 135-137
- Recta de regresión
definición de, 110
de *X* sobre *Y*, 143-146
de *Y* sobre *X*, 137-143
error estándar de una estimación, 146-149
para la predicción, 110, 119
recta de regresión por mínimos cuadrados, 135-143
- Rechazo de la hipótesis nula, 218-219, 232, 242n, 264
región crítica para el, 277-282, 466
- Redondeo, 34-35
- Región crítica para el rechazo de la hipótesis nula, 277-282, 466
- Regla de decisión
análisis de varianza de dos factores y, 406
análisis de varianza de un factor y, 359
nivel alfa y, 218-219, 221-223, 466
- Regla de la suma para probabilidades
definición de, 168, 169
distribución binomial y, 199
- Regla del producto para la probabilidad
con eventos dependientes, 179-183
con eventos independientes, 173-179
con eventos mutuamente excluyentes, 173
con más de dos eventos dependientes, 182-183
con más de dos eventos independientes, 178-179
definición de, 173
distribución binomial y, 199
ecuación para, 173
regla de la suma utilizada con la, 183-186
- Reglas sobre corchetes, 495
Reglas sobre paréntesis, 495
- Regresión de *X* sobre *Y*, 143-146
Regresión de *Y* sobre *X*, 137-143
- Regresión lineal
consideraciones sobre su uso con propósitos de predicción, 149-150
construcción de una recta de regresión por mínimos cuadrados, 137-143
ecuación para la recta de regresión por mínimos cuadrados, 137-138
errores de predicción y, 146-149
introducción a la, 134
predicciones y relaciones imperfectas, 135-137
recta de regresión por mínimos cuadrados, 135-143
regresión de *X* sobre *Y*, 143-146
regresión de *Y* sobre *X*, 137-143
relación entre las constantes de regresión y la *r* de Pearson, 150-151
error estándar de estimación, 146-149, 147n
resumen de, 155
- Regresión múltiple, 151-155
- Regresión
comparación con la correlación, 104, 134
consideraciones sobre el uso de la regresión lineal para hacer predicciones, 149-150
construcción de la recta de regresión por mínimos cuadrados, 137-143
definición de, 135
ecuación de la recta de regresión por mínimos cuadrados, 137-138



SÍMBOLOS

A continuación enumeramos los símbolos utilizados en el texto. La definición de cada símbolo aparece a su derecha. En la última columna se proporciona el número de la página en la que aparece el símbolo por primera vez.

Letras griegas	Significado	El símbolo aparece por primera vez en la página:
α	umbral de probabilidad para el rechazo de H_0 ; en el caso de una variable continua, la probabilidad de cometer un error de tipo I	218
β	probabilidad de cometer un error de tipo I	239
χ^2	ji-cuadrada	427
ϕ	coeficiente de correlación para variables dicotómicas	121
η	coeficiente de correlación curvilínea	121
μ	media de una población	67
μ_D	media de una población de puntajes de diferencia	320
μ_{nula}	media de la población de la hipótesis nula	284
μ_{real}	media de la población cuando existe un efecto real	284
$\mu_{\bar{X}}$	media de la distribución muestral de la media	270
$\mu_{\bar{X}_1 - \bar{X}_2}$	media de la distribución muestral de la diferencia entre las medias muestrales	329
ρ	coeficiente de correlación lineal de la población	310
Σ	la suma de	25
σ	desviación estándar de una población	78
σ^2	varianza de una población	82
$\sigma_{\bar{X}}$	error estándar de la media	270
$\sigma_{\bar{X}}^2$	varianza de la distribución muestral de la media	271
$\sigma_{\bar{X}_1 - \bar{X}_2}$	error estándar de la diferencia entre medias muestrales	329
ω^2	estimación de la magnitud del efecto	341
a_X	intersección con el eje X de la recta de regresión por mínimos cuadrados que predice X conociendo Y	144
a_Y	ordenada al origen de la recta de regresión por mínimos cuadrados que predice Y conociendo X	137

Letras griegas	Significado	El símbolo aparece por primera vez en la página:
b_X	pendiente de la recta de regresión por mínimos cuadrados que predice X conociendo Y	144
b_Y	pendiente de la recta de regresión por mínimos cuadrados que predice Y conociendo X	137
c	número de columnas en una tabla de contingencia	435
	número de columnas en una tabla de datos para análisis de varianza de dos factores	401
$\text{cum } f$	frecuencia acumulativa	46
$\text{cum } f_L$	frecuencia de datos por debajo del límite real inferior del intervalo que contiene al punto percentil	48
$\text{cum } f_P$	frecuencia de datos por debajo del punto percentil	47
$\text{cum } \%$	porcentaje acumulativo	46
D	diferencia entre las parejas de datos	320
\bar{D}_{obt}	media de las diferencias entre parejas de datos	320
gl	grados de libertad	296
gl_B	grados de libertad entre grupos	361
gl_C	grados de libertad por columna	404
gl_R	grados de libertad por renglón	403
gl_{RC}	grados de libertad por renglón \times columna	405
gl_W	grados de libertad dentro de las celdas	402
	grados de libertad intra grupos	360
F	razón de dos estimaciones de la varianza	355
f	frecuencia	39
f_e	frecuencia esperada	428
f_i	frecuencia del intervalo que contiene al punto percentil	49
f_o	frecuencia observada	428
H_0	hipótesis nula	218
H_1	hipótesis alternativa	218
H_{obt}	estadística calculada con la prueba de Kruskal–Wallis	450
i	ancho del intervalo	44
k	número de grupos o medias	361
Mdn	mediano	72
N	numero total de datos	25
	número de pares de datos	137

(c ontinúa)

- error estándar de estimación, 146-149, 147n
- errores de predicción y, 146-149
- predicción y relaciones imperfectas, 135-137
- recta de regresión por mínimos cuadrados, 135-143
- regresión múltiple, 151-155
- regresión de X sobre Y , 143-146
- regresión de Y sobre X , 137-143
- relación entre las constantes de regresión y la r de Pearson, 150-151
- resumen de, 155
- Relación directa, 107-108
- Relación indirecta, 107-108
- Relaciones curvilíneas, 105, 121
- Relaciones imperfectas
 - definición y descripción de, 108-110
 - predicción y, 135-137
- Relaciones lineales
 - definición de, 105
 - ecuación de una línea recta, 105-107
 - gráfica de dispersión, 104-105
 - ordenada al origen de una recta, 106
 - pendiente (b) de una línea recta, 106-107
- Relaciones negativas, 107-108
- Relaciones perfectas
 - definición y descripción de, 108-110
 - predicción y, 134
- Relaciones positivas, 107-108
- Relaciones. *Véase también*
 - Correlación; Regresión
 - introducción a las, 103-104
 - relaciones curvilíneas, 105, 121
 - relación directa, 107-108
 - relaciones imperfectas, 108-110, 135-137
 - relación indirecta, 107-108
 - relaciones lineales, 104-107
 - relaciones negativas, 107-108
 - relaciones perfectas, 108-110, 134
 - relaciones positivas, 107-108
- Reportajes anecdóticos versus investigación científica, 236-237
- Residuo decimal, 33-35
- Residuo
 - residuo decimal, 33-35
 - redondeo del, 34-35
- Resultados importantes versus significativos, 232
- Resultados no significativos, 253-254
- Resultados significativos de experimentos, 219, 232, 253-254
- Rho de Spearman (r_s), 121-123
- Roosevelt, Franklin D., 163
- r_s (rho) coeficiente de correlación, 121-123, 309
- s , para la desviación estándar de datos muestrales, 78
- SAS (sistema de análisis estadístico), 11
- s_B^2 (estimación de la varianza entre grupos), 360-361, 475
- s_C^2 (estimación de la varianza por columna), 398-400, 404, 409, 479
- SC_B (suma de cuadrados entre grupos), 358-359, 361, 363
- SC_C (suma de cuadrados por columna), 404, 408, 477-478
- SC_{pob} (suma de desviaciones al cuadrado), 77-78
- SC_{RC} (suma de cuadrados por renglón (columna), 399-400, 408, 477-478
- SC_T (variabilidad total/suma total de cuadrados), 364-365, 399-400, 408-409, 475, 477-478
- SC_W (suma de cuadrados dentro de las celdas), 400-402, 408, 477-478
- SC_W (suma de cuadrados intra grupos), 358-360
- Sigma (Σ), 25-26
- Sigma (σ), para la desviación estándar de datos poblacionales, 78
- Sigma $\Sigma (X_i - \bar{X})^2$, 68-69, 85n
- Sigma $\Sigma (X_i - \bar{X})^2$, 69
- $(\Sigma X)^2$, 27
- Significación de la r de Pearson, comprobación de la, 309-312
- Símbolos
 - para sujetos o para datos, 25
 - para variables, 24-25
 - para el estudio de, 24
- Simetría
 - curvas de frecuencias y, 57-58
 - tendencia central y, 75
- Sistema de análisis estadístico (SAS), 11
- Sistema de bienestar social, 455-457
- Sistema legal, 189
- SPSS (paquete de estadística para las ciencias sociales), 11
- s_R^2 (estimación de la varianza por renglón), 398-399, 400, 402-403, 409, 479
- s_{RC}^2 (estimación de la varianza por renglón (columna), 398-399, 400, 405, 409, 479
- SC_R (suma de cuadrados por renglón), 399-400, 402-403, 407, 477-478
- Suecia, 189
- Sugerencias para el estudio de la estadística, 23-24
- Suma de cuadrados dentro de las celdas (SCW), 400-402, 408, 477-478
- Suma de cuadrados entre grupos (SC_B), 358-359, 361, 363
 - aplicaciones en el mundo real, 375-376
 - comparaciones múltiples y, 381-392
 - distribución F y, 354-356
 - estimación de la varianza entre grupos (S_B^2), 360-361, 365
 - estimación de la varianza intra grupos (s_W^2), 358-360, 365
 - experimento sobre citas, 370-372
 - experimento sobre la tensión, 362-366
 - lógica suoyacente en el análisis de varianza de un factor, 366-367
 - técnica del análisis de varianza de una función, 357-370, 399-400
 - potencia de, 373-374
 - prueba F y, 356-357, 474-475, 478-479
 - relación entre la prueba t y, 370-372
 - magnitud del efecto y, 373
 - razón F y, 358-359, 362, 365, 474-475, 478-479
 - resumen de, 376-377, 474-479
 - suma de cuadrados entre grupos (SC_B), 358-359, 361, 363
 - suma de cuadrados intra grupos (SC_W), 358-360, 364
 - variabilidad total (SC_T), 364-365
- Suma de cuadrados intra grupos (SC_W), 358-360, 364
- Suma de cuadrados por columna (SC_C), 404, 408, 477-478
- Suma de cuadrados por renglón (columna (SC_{RC}), 399-400, 408, 477-478

- Suma de cuadrados por renglón (SC_R), 399-400, 402-403, 407, 477-478
- Suma de cuadrados, 77-78
- Suma de datos X al cuadrado, ΣX^2 , 27
- Suma de desviaciones al cuadrado, 77-78
- Suma de la cantidad de datos X al cuadrado, $(\Sigma X)^2$, 27
- Suma de las desviaciones al cuadrado de todos los datos en torno a su media, 69-70
- Suma de las desviaciones en torno a la media, 68-69
- Suma de los valores de la variable más una constante, 36-37n
- Suma de los valores de la variable menos una constante, 37n
- Suma de una constante dividida entre el valor de una variable, 37n
- Suma de una constante multiplicada por el valor de una variable, 37n
- Suma total de cuadrados (SC_T), 364-365, 399-400, 408-409, 477-478
- Sumatoria
ecuación de la, 25-27
regla de la, 36-37n
- Supuesto sobre la homogeneidad de la varianza, 338-339, 421
- s_w^2 (estimación de la varianza dentro de las celdas), 398-399, 400, 402, 409, 479
- s_w^2 (estimación de la varianza intra grupos), 358-359, 475
- SYSTAT, 11
- Tabla binomial, 203-210, 219
- Tabla de contingencias, 432-433
- Tabla de números aleatorios, 164, 546-547
- Tablas
área bajo una curva normal, 525-528
distribución binomial, 529-533
distribución F , 540-542
distribución Q , 543
distribución t , 538
interpolación lineal y, 499-500
ji-cuadrada, 544
números aleatorios, 546-547
prueba de rangos con signo de Wilcoxon, 545
- valores críticos de la r de Pearson, 539
- valores críticos U y U' , 534-537
- Tallo y hojas, diagramas de, véase Diagramas de tallo y hojas
- Tasa de error por comparación, 384-385
- Tasa de error por experimento, 384-385, 392
- Tendencia central
introducción a la, 66
media aritmética, 67-71
media global, 70-71
mediana, 72-74
moda, 74-75
resumen de, 82
simetría y, 75
- Teorema del límite central, 269, 271
- t_{obt}
cálculo de t_{obt} para grupos correlacionados, 341-342, 470
cálculo de t_{obt} para una sola muestra, 299-304, 341-342, 469
cálculo de t_{obt} a partir de datos originales, 299-304, 315n
cálculo de t_{obt} en la prueba t para grupos independientes, 332-333, 341-342
cálculo de t_{obt} cuando $n_1 + n_2$, 334-338
ecuación de, 341-342
- Totales marginales, 434-435
- Transformación de datos, 89
- Tukey, John, 58
- U y U' , tabla de valores críticos de, 534-537
- Uso de la computadora en estadística, 10-11
- Valor crítico de \bar{X} , 284-289
- Valor crítico de r , 310-312
- Valor crítico de t , 298
- Valor crítico de un estadístico, 277-282
- Valor crítico de z , 278-280, 298
- Variabilidad explicada por X , 118-120
- Variabilidad muestral
desviación estándar y, 80
media y, 70
mediana y, 74
moda y, 74
potencia y, 374
- Variabilidad total/suma total de cuadrados (SC_T), 364-365, 399-400, 408-409, 477-478
- Variabilidad
de Y explicada por X , 118-120
desviación estándar, 76-81
introducción a la, 66-67
prueba t y, 342
 r de Pearson y, 118-120
rango, 76
resumen de, 82
y prueba t para grupos independientes, 339-341
varianza, 82
- Variable dependiente, 7
- Variable independiente, 6
- Variables continuas
mediciones aproximadas de, 32-33
definición de, 32
probabilidad y variables continuas normalmente distribuidas, 186-188
límites reales de una variable continua, 32-33
- Variables discretas, 32
- Variables
definición de, 6, 31
símbolos para las, 24-25
variable continua, 32-33, 186-188
variable dependiente, 7
variable discreta, 32
variable independiente, 6
- Varianza. Véase también Análisis de varianza
cálculo de la, 82
de datos de población, 82
de datos muestrales, 82
definición de, 82
supuesto de la homogeneidad de la varianza, 338-339, 421
- WAIS (escala Wechsler para la inteligencia de adultos), 31
- Winer, 382
- X , como símbolo de variables, 24-25
- \bar{X} , para la media de una muestra, 67-68
- X_i , 25
- Y como símbolo de variables, 24-25
 μ_{mua} , 284-289
 μ_{real} , 285-289
 μ , X^2 , 27
 μ^2 . Véase también Ji-cuadrada (χ^2)

Letras griegas	Significado	El símbolo aparece por primera vez en la página:
n_k	numero de datos en el k -ésimo grupo	70
P	en una situación con dos eventos, la probabilidad de uno de ellos	172
p	probabilidad	166
$p(A)$	probabilidad del evento A	166
$p(B A)$	probabilidad de B sabiendo que A ha ocurrido	173
P_{nula}	proporción de signos positivos en la población si la variable independiente no tiene efecto alguno	244
P_{real}	proporción de signos positivos en la población si la variable independiente tiene un efecto real	244
Q	en una situación con dos eventos, la probabilidad de uno de ellos estadística con rango de tipo Student	172 385
r	coeficiente de correlación producto momento de Pearson número de medias abarcadas por \bar{X}_i y \bar{X}_j numero de renglones en una tabla de contingencias número de renglones en una tabla de datos para análisis de varianza de dos factores	112 386 435 401
r^2	coeficiente de determinación	120
R^2	coeficiente de determinación múltiple correlación cuadrática múltiple	153 153
r_b	coeficiente de correlación biserial	121
r_s	coeficiente de correlación de orden de rangos de Spearman, rho	121
s	desviación estándar de una muestra estimación de la desviación estándar de una población	78 78
s_D	desviación estándar de los puntajes de diferencias muestrales	320
s_X	desviación estándar de la variable X	151
s_Y	desviación estándar de la variable Y	151
$s_{Y.X}$	error estándar de la estimación al predecir Y conociendo X	147
$s_{\bar{X}}$	error estándar estimado de la media	294
$s_{\bar{X}_1 - \bar{X}_2}$	error estándar estimado de la diferencia entre medias muestrales	331
s^2	varianza de una muestra	82
s_B^2	estimación de la varianza entre grupos	357
s_C^2	column variance estimate	399
s_R^2	estimación de la varianza por columna	399
s_{RC}^2	estimación de la varianza por renglón \times columna	399
s_W^2	estimación ponderada de la varianza de una población estimación de la varianza intra grupos estimación de la varianza entre las celdas	331 357 399

Letras griegas	Significado	El símbolo aparece por primera vez en la página:
SC	suma de cuadrados de una muestra	78
SC_B	suma de cuadrados entre grupos	361
SC_C	suma de cuadrados por columna	404
SC_D	suma de cuadrados de los puntajes de diferencia de la muestra	320
SC_{pop}	suma de cuadrados de una población	78
SC_R	suma de cuadrados por renglón	403
SC_{RC}	suma de cuadrados por renglón \times columna	405
SC_T	suma total de cuadrados	364
SC_W	suma de cuadrados intra grupos	360
	suma de cuadrados dentro de las celdas	402
SC_X	suma de cuadrados de la variable X	137
SC_Y	suma de cuadrados de la variable Y	145
T	suma inferior de los rangos	441
t	estadística de Student	294
U, U'	estadística calculada en la prueba U de Mann–Whitney	445
X	datos en bruto	24
	una variable	24
X'	predicción del valor de X	144
X_i	i -ésimo dato en bruto	25
X_L	valor del límite inferior real del intervalo que contiene al dato X	49
\bar{X}	media de un conjunto muestral de datos en bruto	67
\bar{X}_{global}	media general de varios grupos	70
Y	datos en bruto	25
	una variable	25
Y'	predicción del valor	137
Y_i	i -ésimo dato en bruto	118
\bar{Y}	media de un conjunto muestral de datos en bruto	138
z	número de unidades de desviación estándar que un dato se desvía de la media	89
	estadística calculada para la prueba z	268
	dato estándar	88