10

Pruebas de hipótesis de una muestra

Dole Pineapple, Inc., está preocupada porque supone que las latas de 16 onzas de piña rebanada contienen un exceso de producto. Suponga que la desviación estándar del proceso es de 0.03 onzas. El departamento de control de calidad tomó una muestra aleatoria de 50 latas y comprobó que la media aritmética del peso era de 16.05 onzas. ¿Puede concluir que el peso medio es mayor que 16 onzas con un nivel de significancia de 5%? Determine el valor *p*. (Vea el ejercicio 32, objetivo 6.)

Objetivos de aprendizaje

Al concluir el capítulo, será capaz de:

OA1 Definir una *hipótesis.*

OA2 Describir el procedimiento de prueba de cinco pasos de una hipótesis.

OA3 Definir los errores tipo I y tipo II

OA4 Definir el término *prueba estadística* y explicar la forma de utilizarla.

OA5 Distinguir entre las pruebas de hipótesis de una y dos colas.

OA6 Llevar a cabo una prueba de hipótesis de una media poblacional.

OA7 Calcular e interpretar el valor *p.*

OA8 Llevar a cabo una prueba de hipótesis de una proporción poblacional.

OA9 Calcular la probabilidad de un error tipo II.

10.1 Introducción

En el capítulo 8 dio inicio el estudio de la inferencia estadística. Se describió la forma de seleccionar una muestra aleatoria y, a partir de ella, calcular el valor de un parámetro poblacional. Por ejemplo, se seleccionó una muestra de 5 empleados de Spence Sprockets para determinar la cantidad de años de servicio de cada empleado entrevistado, se calculó la media de los años de servicio y se utilizó la media de la muestra para estimar la media de los años de servicio de todo el personal. En otras palabras, se estimó un parámetro poblacional a partir de un estadístico de la muestra.

En el capítulo 9 se prosiguió con el estudio de la inferencia estadística mediante la construcción de un intervalo de confianza. Un intervalo de confianza es un conjunto de valores en el que se encuentra el parámetro de la población. En este capítulo, en lugar de crear un conjunto de valores en el que se espera que se presente el parámetro poblacional, se expone un procedimiento para *probar* la validez de un enunciado relativo a un parámetro poblacional. Algunos ejemplos de enunciados por probar son los siguientes:

- La velocidad media de los automóviles que pasan por la señal de 150 millas de la carretera West Virginia Turnpike es de 68 millas por hora.
- La cantidad media de millas que recorre una Chevy TrailBlazer rentada durante tres años es de 32 000 millas.
- El tiempo medio que una familia estadounidense vive en una vivienda en particular es de 11.8 años.
- En 2010, el salario inicial medio en ventas de un graduado de universidad fue de $47 673.
- Treinta y cinco por ciento de los jubilados de la región norte de Estados Unidos vende su hogar y se muda a un clima más cálido después de un año de haberse retirado.
- Ochenta por ciento de los jugadores asiduos a la lotería estadounidense jamás gana más de $100 en un juego.

Este capítulo y algunos de los siguientes se relacionan con pruebas de hipótesis estadísticas. Primero hay que definir los términos de hipótesis estadística y pruebas de hipótesis estadísticas. Después se muestran los pasos para llevar a cabo una prueba de hipótesis estadística. A continuación se aplican pruebas de hipótesis para medias y proporciones. En la última sección del capítulo se describen los posibles errores que se deben al muestreo en las pruebas de hipótesis.

10.2 ¿Qué es una hipótesis?

Una hipótesis es un enunciado acerca de un parámetro poblacional.

Una hipótesis es una declaración relativa a una población. A continuación se utilizan los datos para verificar lo razonable del enunciado. Para comenzar, es necesario definir la palabra *hipótesis.* En el sistema legal estadounidense, una persona es inocente hasta que se prueba su culpabilidad. Un jurado plantea como hipótesis que una persona a la que se le imputa un crimen es inocente, y someten esta hipótesis a verificación, para lo cual revisan la evidencia y escuchan el testimonio antes de llegar a un veredicto. De forma similar, un paciente visita al médico y acusa varios síntomas. Con base en ellos, el médico indicará ciertos exámenes de diagnóstico; en seguida, de acuerdo con los síntomas y los resultados de los exámenes, determina el tratamiento.

En el análisis estadístico se establece una afirmación, una hipótesis, se recogen datos que posteriormente se utilizan para probar la aserción. Entonces, una hipótesis estadística es:

OA1 Definir una *hipótesis.*

HIPÓTESIS Afirmación relativa a un parámetro de la población sujeta a verificación.

En la mayoría de los casos, la población es tan grande que no es viable estudiarla por completo. Por ejemplo, no sería posible contactar a todos los analistas de sistemas de Estados Unidos para preguntarles su ingreso mensual. Del mismo modo, la calidad del departamento de control de calidad de Cooper Tire no puede verificar todos los neumáticos que la empresa produce para ver si duran más de 60 000 millas.

Como se observó en el capítulo 8, una opción para medir o entrevistar a toda la población es tomar una muestra de ella. Por lo tanto, así se pone a prueba una declaración para determinar si la muestra apoya o no la declaración en lo concerniente a la población.

Estadística en acción

LASIK es un procedimiento quirúrgico de 15 minutos de duración con un rayo láser para modificar la forma de la córnea con el fin de mejorar la visión. Las investigaciones demuestran que alrededor de 5% de las cirugías presenta complicaciones, como deslumbramientos, visión borrosa, corrección excesiva o insuficiente de la visión, y su pérdida. Desde una perspectiva estadística, las investigaciones someten a prueba una hipótesis nula acerca de que la cirugía no mejorará la visión frente a la hipótesis alternativa de que la cirugía la mejorará. Los datos de la muestra de la cirugía LASIK indican que 5% de los casos presenta complicaciones. Este término de 5% representa un índice de error tipo I. Cuando una persona decide someterse a la cirugía, espera rechazar la hipótesis nula. En 5% de los casos futuros, esta expectativa no se cumplirá. (Fuente: *American Academy of Ophthalmology Journal*, San Francisco, vol. 16, núm. 43.)

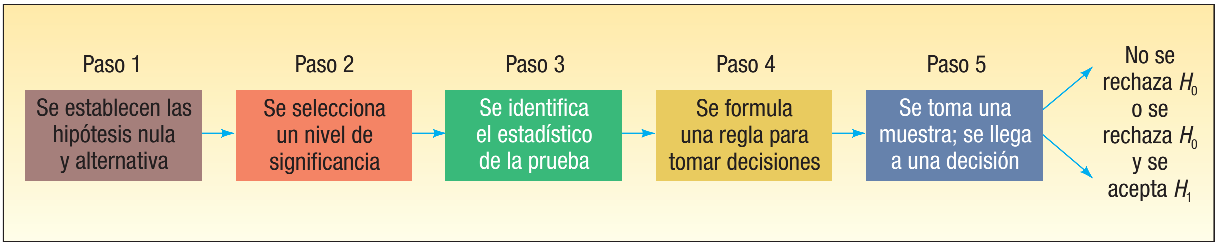
10.3 ¿Qué es la prueba de hipótesis?

Los términos *prueba de hipótesis* y *probar una hipótesis* se utilizan indistintamente. La prueba de hipótesis comienza con una afirmación, o suposición, sobre un parámetro de la población, como la media poblacional. Como ya se indicó, esta afirmación recibe el nombre de *hipótesis.* Una hipótesis puede ser que la comisión mensual media de las comisiones de los vendedores de tiendas al menudeo de aparatos electrónicos, como Circuit City, es de $2 000. No es posible entrar en contacto con todos los vendedores para asegurarnos de que la media en realidad sea de $2 000. El costo de localizar a y entrevistarse con todos los vendedores de aparatos electrónicos en Estados Unidos sería exorbitante. Para probar la validez de la afirmación (μ = $2 000) se debe seleccionar una muestra de la población de vendedores de aparatos electrónicos, calcular el estadístico muestral y, con base en ciertas reglas de decisión, aceptar o rechazar la hipótesis. Una media muestral de $1 000 de los vendedores de aparatos electrónicos provocaría con certeza el rechazo de la hipótesis. Sin embargo, suponga que la media de la muestra es de $1 995. ¿Está lo bastante cerca de $2 000 para aceptar la suposición de que la media de la población es de $2 000? ¿La diferencia de $5 entre las dos medias se puede atribuir al error de muestreo, o dicha diferencia resulta estadísticamente significativa?

PRUEBA DE HIPÓTESIS Procedimiento basado en evidencia de la muestra y la teoría de la probabilidad para determinar si la hipótesis es una afirmación razonable.

10.4 Procedimiento de cinco pasos para probar una hipótesis

Existe un procedimiento de cinco pasos que sistematiza la prueba de una hipótesis; al llegar al paso 5, se está en posibilidades de rechazar o no la hipótesis. Sin embargo, la prueba de hipótesis, como la emplean los especialistas en estadística, no prueba que algo es verdadero de la forma en que un matemático *demuestra* un enunciado. Más bien, proporciona un tipo de *prueba más allá de toda duda razonable,* como en el sistema judicial. De ahí que existan reglas específicas de evidencia, o procedimientos. En el siguiente diagrama aparecen los pasos. Analizaremos con detalle cada uno de ellos.



Paso 1: Se establece la hipótesis nula (H_0) y la hipótesis alternativa (H_1)

OA2 Describir el procedimiento de prueba de una hipótesis en cinco pasos.

El primer paso consiste en establecer la hipótesis que se debe probar. Ésta recibe el nombre de **hipótesis nula**, la cual se designa H_0, y se lee "*H* subíndice cero". La letra mayúscula *H* representa la hipótesis, y el subíndice cero implica que "no hay diferencia". Por lo general se incluye un término *no* en la hipótesis nula, que significa que "no hay cambio". Por ejemplo, la hipótesis nula que se refiere a la cantidad media de millas que recorre cada neumático con cinturón de acero no es diferente de 60 000. La hipótesis nula se escribiría H_0: $\mu = 60\ 000$. En términos generales, la hipótesis nula se formula para realizar una prueba. O se rechaza o no se rechaza. Es una afirmación que no se rechaza a menos que la información de la muestra ofrezca evidencia convincente de que es falsa.

Cabe hacer hincapié en que, si la hipótesis nula no se rechaza con base en los datos de la muestra, no es posible decir que la hipótesis nula sea verdadera. En otras palabras, el hecho de no rechazar una hipótesis no prueba que H_0 sea verdadera, sino que *no rechazamos* H_0. Para probar sin lugar a dudas que la hipótesis nula es verdadera, sería necesario conocer el parámetro poblacional. Para determinarlo, habría que probar, entrevistar o contar cada elemento de la población. Esto no resulta factible. La alternativa consiste en tomar una muestra de la población.

Se establecen la hipótesis nula y la hipótesis alternativa.

También debe destacarse que con frecuencia la hipótesis nula inicia con las expresiones: "No existe diferencia *significativa* entre…" o "La resistencia media del vidrio a los impactos no es *significativamente* diferente de…" Al seleccionar una muestra de una población, el estadístico de la muestra es numéricamente distinto del parámetro poblacional hipotético. Como ejemplo, suponga que la hipótesis de la resistencia de un platón de vidrio a los impactos es de 70 psi, y que la resistencia media de una muestra de 12 platones de vidrio es de 69.5 psi. Se debe tomar la decisión con la diferencia de 0.5 psi. ¿Se trata de una diferencia real, es decir, una diferencia significativa, o la diferencia entre el estadístico de la muestra (69.5) y el parámetro poblacional hipotético (70.0) es aleatorio y se debe al error de muestreo? Según se dijo, la respuesta a esta pregunta implica una prueba de significancia, que recibe el nombre de *prueba de hipótesis.* Una hipótesis nula es:

> **HIPÓTESIS NULA** Enunciado relativo al valor de un parámetro poblacional que se formula con el fin de probar evidencia numérica.

La **hipótesis alternativa** describe lo que se concluirá si se rechaza la hipótesis nula. Se representa H_1 y se lee "*H* subíndice uno". También se le conoce como *hipótesis de investigación.* La hipótesis alternativa se acepta si la información de la muestra ofrece suficiente evidencia estadística para rechazar la hipótesis nula.

> **HIPÓTESIS ALTERNATIVA** Enunciado que se acepta si los datos de la muestra ofrecen suficiente evidencia para rechazar la hipótesis nula.

El siguiente ejemplo aclara los términos hipótesis nula y alternativa. Un artículo reciente indicó que el tiempo de uso medio de los aviones comerciales estadounidenses es de 15 años. Para llevar a cabo una prueba estadística relacionada con esta afirmación, el primer paso consiste en determinar las hipótesis nula y alternativa. La hipótesis nula representa el estado actual o reportado. Se escribe: H_0: $\mu = 15$. La hipótesis alternativa se refiere al hecho de que la afirmación no es verdadera, es decir, H_1: $\mu \neq 15$. Es necesario recordar que, sin que importe la manera de plantear el problema, la *hipótesis nula siempre incluirá el signo de igual.* Este signo (=) nunca aparecerá en la hipótesis alternativa. ¿Por qué? Porque es la afirmación que se va a probar, y es necesario un valor específico para incluir en los cálculos. Se recurre a la hipótesis alternativa sólo si la información sugiere que la hipótesis nula es falsa.

Paso 2: Se selecciona un nivel de significancia

Después de establecer las hipótesis nula y alternativa, el siguiente paso consiste en determinar el nivel de significancia.

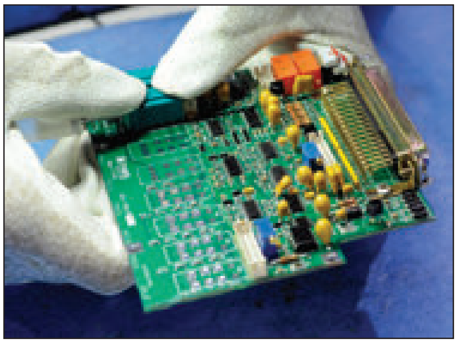
NIVEL DE SIGNFICANCIA Probabilidad de rechazar la hipótesis nula cuando es verdadera.

Se selecciona un nivel de significancia o riesgo.

El nivel de significancia se expresa con la letra griega alfa, α. En ocasiones también se conoce como *nivel de riesgo*. Éste quizá sea un término más adecuado porque se trata del riesgo que se corre al rechazar la hipótesis nula cuando es verdadera.

No existe ningún nivel de significancia que se aplique a todas las pruebas. Se toma la decisión de utilizar el nivel de 0.05 (expresado con frecuencia como nivel de 5%), nivel de 0.01, nivel de 0.10 o cualquier otro nivel entre 0 y 1. Se acostumbra elegir el nivel de 0.05 en el caso de los proyectos de investigación relacionados con los consumidores; el nivel de 0.01 en relación con el del control de calidad, y el de 0.10 en el de las encuestas políticas. Usted, como investigador, debe elegir el nivel de significancia *antes* de formular una regla de decisión y recopilar los datos de la muestra.

Para ilustrar cómo es posible rechazar una hipótesis verdadera, suponga que una empresa fabricante de computadoras personales utiliza una gran cantidad de tarjetas con circuitos impresos. Los proveedores participan en una licitación y el que presenta la cotización más baja obtiene el contrato. Suponga que éste especifica que el departamento de control de calidad del fabricante de computadoras tomará una muestra de los envíos que llegan. Si más de 6% de las tarjetas de la muestra no cumple con las normas, el envío se rechaza. La hipótesis nula consiste en que el envío de tarjetas contiene 6% o menos tarjetas que no satisfacen las normas. La hipótesis alternativa consiste en que más de 6% de las tarjetas están defectuosas.



Una muestra de 50 tarjetas de circuitos de Allied Electronics, que se recibieron el 21 de julio, reveló que 4, es decir, 8%, no cumplían con las normas. El envío se rechazó en virtud de que excedía el máximo de 6% de tarjetas que no cumplían con las normas. Si en realidad el envío no cumplía con las normas, fue acertada la decisión de devolver las tarjetas al proveedor. No obstante, suponga que las 4 tarjetas elegidas de la muestra de 50 eran las únicas que no cumplían con las normas en un envío de 4 000 tarjetas. Entonces, sólo 0.1% se encontraba defectuoso (4/4 000 = 0.001). En este caso, menos de 6% de todo el envío no satisfacía las normas, y rechazarlo fue un error. En términos de la prueba de hipótesis, rechazamos la hipótesis nula de que el envío cumplía con las normas cuando se debió aceptar. Al rechazar la hipótesis nula, se incurrió en un error tipo I. La probabilidad de cometer este tipo de error es α.

ERROR TIPO I Rechazar la hipótesis nula, H_0, cuando es verdadera.

La probabilidad de cometer otro tipo de error, conocido como error tipo II, se expresa con la letra griega beta (β).

ERROR TIPO II Aceptar la hipótesis nula cuando es falsa.

La empresa que fabrica computadoras personales cometería un error del tipo II si, sin que lo sepa el fabricante, un envío de tarjetas de Allied Electronics contiene 15% de tarjetas que no cumplen con las normas, y aún así lo aceptara. ¿Cómo puede suceder esto? Suponga que

2 de las 50 tarjetas (4%) no son aceptables, mientras que 48 de 50 lo son. De acuerdo con el procedimiento mencionado, como la muestra contiene menos de 6% de tarjetas que no cumplen con las normas, el envío se acepta. ¡Puede suceder que, *por azar*, las 48 tarjetas que contiene la muestra sean las únicas aceptables en todo el envío, que consta de miles de tarjetas!

En retrospectiva, el investigador no puede estudiar cada elemento o individuo de la población. Por lo tanto, existe la posibilidad de que se presenten dos clases de error: un error tipo I, en el que se rechaza la hipótesis nula cuando en realidad debe aceptarse, y un error tipo II, en el que se acepta la hipótesis nula cuando en realidad debe rechazarse.

OA3 Definir los errores tipo I y tipo II.

Con frecuencia se hace referencia a la probabilidad de cometer estos dos posibles errores como *alfa,* α, y *beta,* β. Alfa (α) es la probabilidad de cometer un error tipo I, y beta (β), la probabilidad de cometer un error tipo II.

La siguiente tabla resume las decisiones que el investigador puede tomar y sus posibles consecuencias.

	Investigador	
Hipótesis nula	**No rechaza H_0**	**Rechaza H_0**
H_0 es verdadera	Decisión correcta	Error tipo I
H_0 es falsa	Error tipo II	Decisión correcta

Paso 3: Se selecciona el estadístico de prueba

Hay muchos estadísticos de prueba. En este capítulo se utilizan z y t como estadísticos de prueba. En otros capítulos aparecen estadísticos de prueba como F y χ^2, conocida como *ji-cuadrada.*

OA4 Definir el término *prueba estadística* y explicar la forma de utilizarlo.

ESTADÍSTICO DE PRUEBA Valor, determinado a partir de la información de la muestra, para determinar si se rechaza la hipótesis nula.

La prueba de hipótesis de la media (μ), cuando se conoce σ o el tamaño de la muestra es grande, es el estadístico de prueba z que se calcula de la siguiente manera:

PRUEBA DE LA MEDIA CUANDO SE CONOCE σ

$$z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \qquad \textbf{(10-1)}$$

El valor z se basa en la distribución muestral de $\overline{X}$, que sigue la distribución normal cuando la muestra es razonablemente grande, con una media ($\mu_{\overline{X}}$) igual a μ y una desviación estándar $\sigma_{\overline{X}}$ igual a $\sigma/\sqrt{n}$. Por consiguiente, puede determinar si la diferencia entre $\overline{X}$ y μ es significativa desde una perspectiva estadística al determinar el número de desviaciones estándares a las que se encuentra $\overline{X}$ de μ, con la fórmula (10.1).

Paso 4: Se formula la regla de decisión

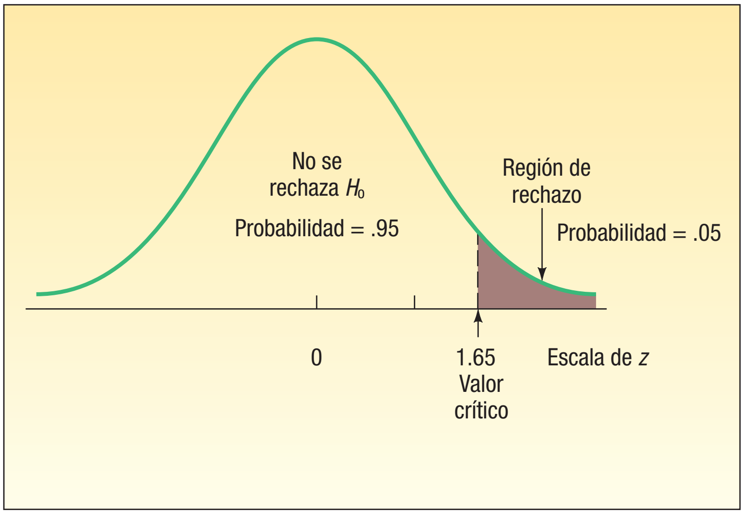
La regla de decisión establece las condiciones cuando se rechaza H_0.

Una regla de decisión es un enunciado sobre las condiciones específicas en que se rechaza la hipótesis nula y aquellas en las que no se rechaza. La región o área de rechazo define la ubicación de todos esos valores que son tan grandes o tan pequeños que la probabilidad de que ocurran en una hipótesis nula verdadera es muy remota.

En la gráfica 10-1 se presenta la región de rechazo de una prueba de significancia que se efectuará más adelante en este capítulo.

Estadística en acción

Durante la Segunda Guerra Mundial, los encargados de la planeación militar de los aliados necesitaban cálculos aproximados de la cantidad de tanques alemanes. No era confiable la información que proporcionaban los métodos de espionaje tradicionales, y, en cambio, los métodos estadísticos probaron ser muy valiosos. Por ejemplo, el espionaje y el reconocimiento llevaron a los analistas a calcular que durante junio de 1941 se produjeron 1 550 tanques. Sin embargo, por medio de la utilización de los números de serie de los tanques capturados y el análisis estadístico, los encargados de la planeación militar calcularon 244. La cantidad real de tanques producidos, de acuerdo con los registros de producción alemanes, fue de 271. El cálculo a través del análisis estadístico resultó ser mucho más preciso. Un tipo de análisis similar se empleó para calcular la cantidad de tanques iraquíes que fueron destruidos en la Tormenta del Desierto.



GRÁFICA 10-1 Distribución muestral del estadístico z; prueba de una cola a la derecha; nivel de significancia de 0.05

Observe lo siguiente en la gráfica:

- El área en que se acepta la hipótesis nula se localiza a la izquierda de 1.65. En breve se explicará la forma de obtener el valor de 1.65.
- El área de rechazo se encuentra a la derecha de 1.65.
- Se aplica una prueba de una sola cola (este hecho también se explicará más adelante).
- Se eligió el nivel de significancia de 0.05.
- La distribución muestral del estadístico z tiene una distribución normal.
- El valor 1.65 separa las regiones en que se rechaza la hipótesis nula y en la que se acepta.
- El valor de 1.65 es el **valor crítico**.

VALOR CRÍTICO Punto de división entre la región en que se rechaza la hipótesis nula y aquella en la que se acepta.

Paso 5: Se toma una decisión

El quinto y último paso en la prueba de hipótesis consiste en calcular el estadístico de la prueba, comparándola con el valor crítico, y tomar la decisión de rechazar o no la hipótesis nula. De acuerdo con la gráfica 10-1, si, a partir de la información de la muestra, se calcula que z tiene un valor de 2.34, se rechaza la hipótesis nula con un nivel de significancia de 0.05. La decisión de rechazar H_0 se tomó porque 2.34 se localiza en la región de rechazo; es decir, más allá de 1.65. Se rechaza la hipótesis nula porque es poco probable que un valor z tan alto se deba al error de muestreo (azar).

Si el valor calculado hubiera sido de 1.65 o menos, supongamos 0.71, la hipótesis nula no se habría rechazado. Un valor calculado tan bajo no se atribuye al azar, es decir, al error de muestreo.

Como se indicó, en la prueba de hipótesis sólo es posible una de las dos decisiones: la hipótesis nula se acepta o se rechaza. En lugar de *aceptar* la hipótesis nula, H_0, algunos investigadores prefieren expresar la decisión como “no se rechaza H_0”, “se decide no rechazar H_0” o “los resultados de la muestra no permiten rechazar H_0”.

Es necesario subrayar de nuevo que siempre existe la posibilidad de que la hipótesis nula se rechace cuando en realidad no se debe rechazar (error tipo I). Asimismo, existe una posibilidad definible de que la hipótesis nula se acepte cuando debiera rechazarse (error tipo II).

RESUMEN DE LOS PASOS DE LA PRUEBA DE HIPÓTESIS

1. Se establecen la hipótesis nula (H_0) y la hipótesis alternativa (H_1).
2. Se selecciona el nivel de significancia, es decir, α.
3. Se selecciona un estadístico de prueba adecuado.
4. Se formula una regla de decisión con base en los pasos 1, 2 y 3 anteriores.
5. Se toma una decisión en lo que se refiere a la hipótesis nula con base en la información de la muestra. Se interpretan los resultados de la prueba.

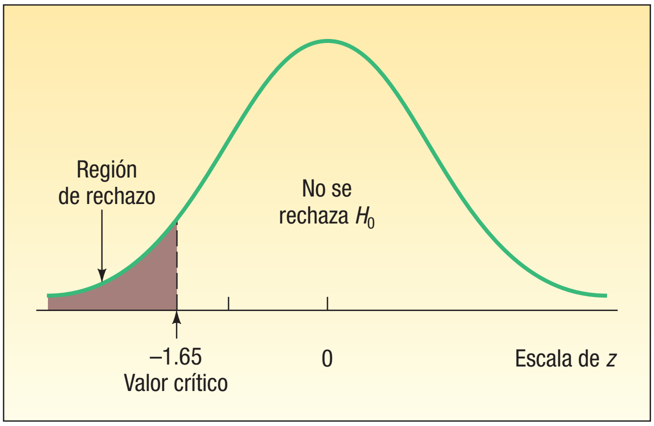
Antes de llevar a cabo una prueba de hipótesis, es importante diferenciar entre una prueba de significancia de una cola y una prueba de dos colas.

10.5 Pruebas de significancia de una y dos colas

OA5 Distinguir entre las pruebas de hipótesis de una y dos colas.

Consulte la gráfica 10-1. En ella se describe una prueba de una cola. La región de rechazo se localiza sólo en la cola derecha (superior) de la curva. Para ilustrarlo, suponga que el departamento de empaque de General Foods Corporation se preocupa porque algunas cajas de Grape Nuts exceden considerablemente el peso. El cereal se empaca en cajas de 453 gramos, por lo que la hipótesis nula es H_0: $\mu \leq 453$, que se lee: “la media poblacional (μ) es igual o menor que 453”. Por consiguiente, la hipótesis alternativa es H_0: $\mu > 453$, que se lee: “μ es mayor que 453”. Observe que el signo de desigualdad en la hipótesis alternativa ($>$) señala hacia la región de rechazo ubicada en la cola superior. (Vea la gráfica 10-1.) También observe que la hipótesis nula incluye el signo igual. Es decir, H_0: $\mu \leq 453$. La condición de igualdad siempre aparece en H_0 y jamás en H_1.

La gráfica 10-2 representa un caso en el que la región de rechazo se encuentra en la cola izquierda (inferior) de la distribución normal. Como ejemplo, considere el problema de los fabricantes de automóviles. Por ejemplo, las grandes compañías de renta de autos y otras empresas que compran grandes cantidades de neumáticos desean que duren un promedio de 60 000 millas en condiciones normales. Por consiguiente, rechazarán un envío de neumáticos si las pruebas revelan que la vida de éstas es mucho menor a 60 000 millas en promedio. Con gusto aceptarán el envío si la vida media es mayor a 60 000 millas. Sin embargo, esta posibilidad no les preocupa. Sólo les interesa si cuentan con evidencias suficientes para concluir que los neumáticos tendrán un promedio de vida útil inferior a 60 000 millas. Por lo tanto, la prueba se plantea de manera que satisfaga la preocupación de los fabricantes de automóviles respecto de que la *vida media de los neumáticos sea menor a 60 000 millas.* Este enunciado



GRÁFICA 10-2 Distribución muestral del estadístico z, prueba de cola izquierda, nivel de significancia de 0.05

La prueba es de una cola si H_1 afirma que $\mu >$ o $\mu <$.

aparece en la hipótesis alternativa. En este caso, las hipótesis nula y alternativa se escriben H_0: $\mu \geq 60\,000$ y H_1: $\mu < 60\,000$.

Una manera para determinar la ubicación de la región de rechazo consiste en mirar en la dirección en la que señala el signo de desigualdad en la hipótesis alternativa ($<$ o $>$). En este problema, señala a la izquierda, y, por consiguiente, la región de rechazo se localiza en la cola izquierda.

Si H_1 indica una dirección, la prueba es de una cola.

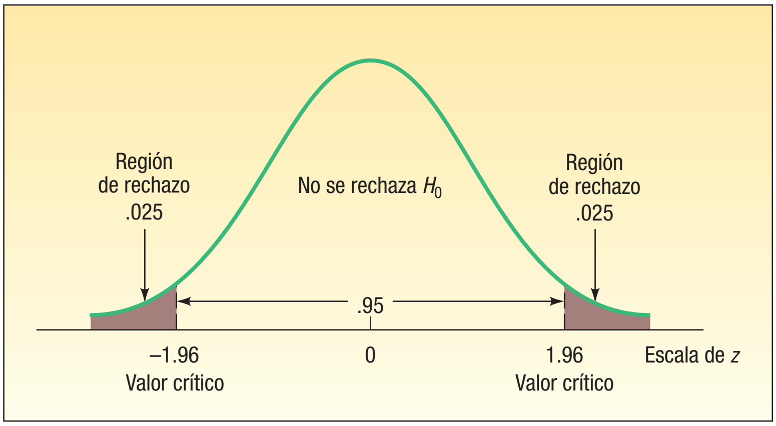
En resumen, una prueba es de *una cola* cuando la hipótesis alternativa, H_1, indica una dirección, como:

H_0: el ingreso medio anual de las corredoras de bolsa es *menor o igual a* \$65 000.
H_1: el ingreso medio anual de las corredoras de bolsa es *mayor a* \$65 000 anuales.

Si no se especifica dirección alguna en la hipótesis alternativa, utilice una prueba de *dos colas.* Si cambia el problema anterior con fines de ilustración, puede decir lo siguiente:

H_0: el ingreso medio anual de las corredoras de bolsa es de \$65 000 anuales.
H_1: el ingreso medio anual de las corredoras de bolsa *no es igual a* \$65 000 anuales.

Si se rechaza la hipótesis nula y se acepta H_1 en el caso de las dos colas, el ingreso medio puede ser significativamente mayor o inferior a \$65 000 anuales. Para dar cabida a estas dos posibilidades, el área de 5% de rechazo se divide con equidad en las dos colas de la distribución muestral (2.5% cada una). La gráfica 10-3 presenta las dos áreas y los valores críticos. Observe que el área total en la distribución normal es de 1.0000, que se calcula por medio de 0.9500 + 0.0250 + 0.0250.



GRÁFICA 10-3 Regiones de aceptación y rechazo de una prueba de dos colas con un nivel de significancia de 0.05

10.6 Pruebas de la media de una población: se conoce la desviación estándar poblacional

Prueba de dos colas

Un ejemplo mostrará los detalles del procedimiento para probar una hipótesis en cinco pasos. También se desea usar una prueba de dos colas. Es decir, *no interesa* si los resultados de la muestra son más grandes o más pequeños que la media poblacional propuesta. Lo que interesa es si ésta es *diferente del* valor propuesto para la media poblacional. Como en el capítulo anterior, conviene iniciar con un caso del cual se cuente con un historial de datos sobre la población y, de hecho, se conozca la desviación estándar.

Ejemplo

OA6 Llevar a cabo una prueba de hipótesis de una media poblacional.

Jamestown Steel Company fabrica y arma escritorios y otros muebles para oficina en diferentes plantas en el oeste del estado de Nueva York. La producción semanal del escritorio modelo A325 en la planta de Fredonia tiene una distribución normal, con una media de 200 y una desviación estándar de 16. Hace poco, con motivo de la expansión del mercado, se introdujeron nuevos métodos de producción y se contrató a más empleados. El vicepresidente de fabricación pretende investigar si hubo algún *cambio* en la producción semanal del escritorio modelo A325. En otras palabras, ¿la cantidad media de escritorios que se produjeron en la planta de Fredonia es *diferente de* 200 escritorios semanales con un nivel de significancia de 0.01?

Solución

En este ejemplo, tenemos dos datos importantes: 1) la población de la producción semanal sigue una distribución normal, y 2) la desviación estándar de esta distribución normal es de 16 escritorios por semana. Por ello, es apropiado utilizar el estadístico z para resolver este problema. Aplique el procedimiento de prueba de hipótesis estadística para investigar si cambió el índice de producción de 200 escritorios semanales.

Paso 1: Se establecen las hipótesis nula y alternativa. La hipótesis nula es: "la media de la población es de 200". La hipótesis alternativa es: "la media es diferente de 200" o "la media no es de 200". Estas dos hipótesis se expresan de la siguiente manera:

$$H_0\colon \mu = 200$$
$$H_1\colon \mu \neq 200$$

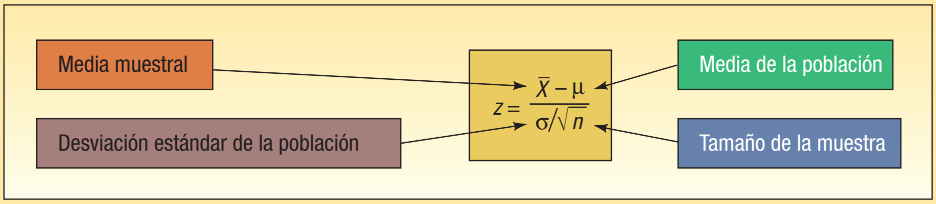
Ésta es una *prueba de dos colas,* pues la hipótesis alternativa no indica dirección alguna. En otras palabras, no establece si la producción media es mayor o menor a 200. El vicepresidente sólo desea saber si la tasa de producción es distinta de 200.

Paso 2: Se selecciona el nivel de significancia. Como ya se indicó, se utiliza el nivel de significancia de 0.01. Éste es α, la probabilidad de cometer un error tipo I, que es la probabilidad de rechazar una hipótesis nula verdadera.

Paso 3: Se selecciona el estadístico de prueba. El estadístico de prueba de una muestra grande es z. Este tema se estudió lo suficiente en el capítulo 7. La transformación de los datos de producción en unidades estándares (valores z) permite que se les utilice no sólo en este problema, sino en otros relacionados con la prueba de hipótesis. A continuación se repite la fórmula (10-1) para z y se identifican las diferentes letras.

(10-1)

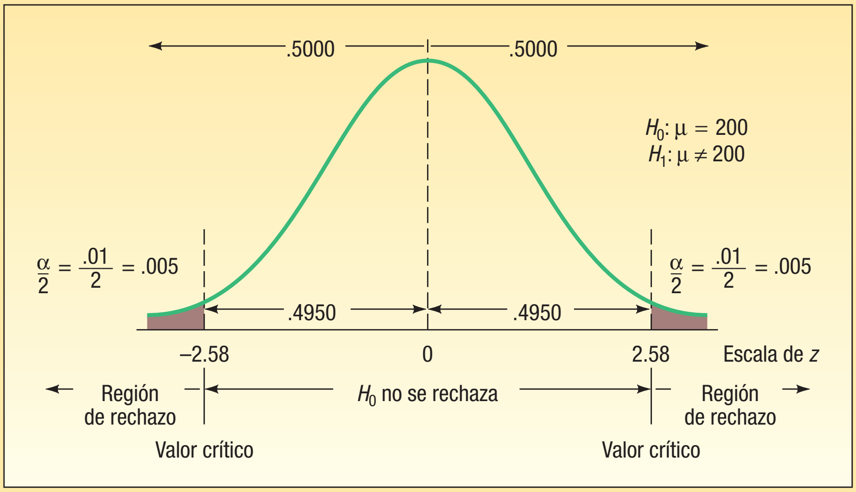
Fórmula del estadístico de la prueba.



Paso 4: Se formula la regla de decisión. La regla de decisión se formula al encontrar los valores críticos de z con ayuda del apéndice B.1. Como se trata de una prueba de dos colas, la mitad de 0.01, o 0.005, se localiza en cada cola. Por consiguiente, el área en la que no se rechaza H_0, que se ubica entre las dos colas, es 0.99. El apéndice B.1 se basa en la mitad del área bajo la curva, o 0.5000. Entonces, 0.5000 − 0.0050 es 0.4950, por lo que 0.4950 es el área entre 0 y el valor crítico.

Se localiza 0.4950 en el cuerpo de la tabla. El valor más cercano a 0.4950 es 0.4951. En seguida se lee el valor crítico en el renglón y columna correspondientes a 0.4951. Éste es de 2.58. Por conveniencia, se repite el apéndice B.1, Áreas bajo la curva normal, en la tercera de forros.

Todas las facetas de este problema aparecen en el diagrama de la gráfica 10-4.



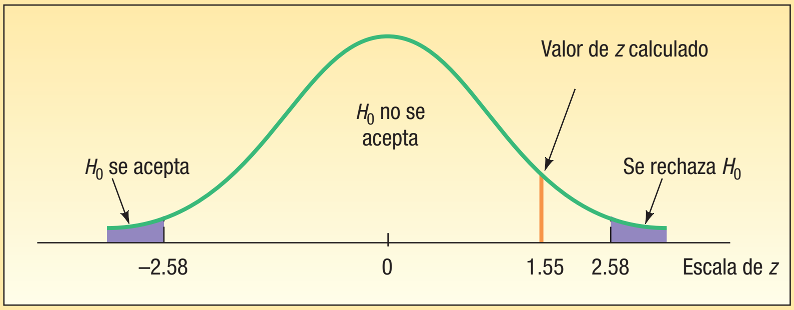
GRÁFICA 10-4 Regla de decisión del nivel de significancia de 0.01

Por lo tanto, la regla de decisión es: rechazar la hipótesis nula y aceptar la hipótesis alternativa (que indica que la media de la población no es 200) si el valor z calculado no se encuentra entre -2.58 y $+2.58$. La hipótesis nula no se rechaza si z se ubica entre -2.58 y $+2.58$.

Paso 5: Se toma una decisión y se interpreta el resultado. Se toma una muestra de la población (producción semanal), se calcula z, se aplica la regla de decisión y se llega a la decisión de rechazar o no H_0. La cantidad media de escritorios que se produjeron el año pasado (50 semanas, pues la planta cerró 2 semanas por vacaciones) es de 203.5. La desviación estándar de la población es de 16 escritorios semanales. Al calcular el valor z a partir de la fórmula (10-1), se obtiene:

$$z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{203.5 - 200}{16/\sqrt{50}} = 1.55$$

Como 1.55 no cae en la región de rechazo, H_0 no se rechaza. La conclusión es: la media de la población *no* es distinta de 200. Por lo tanto, se informa al vicepresidente de fabricación que la evidencia de la muestra no indica que la tasa de producción en la planta de Fredonia haya cambiado de 200 semanales. La diferencia de 3.5 unidades entre la producción semanal histórica y la del año pasado puede atribuirse razonablemente al error de muestreo. Esta información se resume en el siguiente diagrama:



¿Se demostró que el ritmo de montaje aún es de 200 a la semana? No. Lo que se hizo, desde un punto de vista técnico, fue *no desaprobar la hipótesis nula.* No refutar la hipótesis de que la media poblacional es de 200 no es lo mismo que probar que necesariamente es verdadera. Como se sugiere en la introducción del capítulo, la conclusión es análoga a la del sistema jurídico estadounidense. Para explicarlo, suponga que se acusa a una persona de un crimen, pero un jurado la absuelve. Si la persona es absuelta, se concluye que no había suficiente evidencia para probar su culpabilidad. El juicio no probó que el individuo era necesariamente inocente, sino que no había suficiente evidencia para probar su culpabilidad. Eso evidencia las pruebas de hipótesis estadísticas cuando no se rechaza la hipótesis nula. La interpretación correcta consiste en que no se probó la falsedad de la hipótesis nula.

En este caso se eligió el nivel de significancia de 0.01 antes de establecer la regla de decisión y tomar una muestra de la población. Ésta es la estrategia adecuada. El investigador debe establecer el nivel de significancia, pero debe determinarlo *antes* de reunir la evidencia de la muestra y no realizar cambios con base en la evidencia de ella.

Comparación de intervalos de confianza y pruebas de hipótesis.

¿Cómo se confronta el procedimiento de prueba de hipótesis, recién descrito, con el procedimiento de los intervalos de confianza que se estudió en el capítulo anterior? Al realizar la prueba de hipótesis en la producción de escritorios, se cambiaron las unidades de escritorios semanales a un valor *z.* Después se comparó el valor calculado del estadístico de la prueba (1.55) con el de los valores críticos (-2.58 y 2.58). Como el valor calculado se localizó en la región de no rechazo de la hipótesis nula, se concluyó que la media poblacional podía ser de 200. Por otro lado, para aplicar el enfoque del intervalo de confianza, se debía construir un intervalo de confianza con la fórmula (9-1) (p. 302). El intervalo iría de 197.66 a 209.34, el cual se calcula de la siguiente manera: $203.5 \pm 2.58(16/\sqrt{50})$. Observe que el valor poblacional propuesto, 200, se encuentra en este intervalo. De ahí que la media poblacional podría ser, razonablemente, 200.

En general, H_0 se rechaza si el intervalo de confianza no incluye el valor hipotético. Si el intervalo de confianza incluye el valor hipotético, no se rechaza H_0. Así, la *región de no rechazo* en una prueba de hipótesis equivale al valor poblacional propuesto en el intervalo de confianza. La diferencia fundamental entre un intervalo de confianza y la *región de no rechazo* en una prueba de hipótesis depende de que el intervalo se centre en torno al estadístico de la muestra, como $\overline{X}$, al intervalo de confianza o alrededor de 0, como en la prueba de hipótesis.

Autoevaluación 10-1

Heinz, un fabricante de catsup, utiliza una máquina para vaciar 16 onzas de su salsa en botellas. A partir de su experiencia de varios años con la máquina despachadora, la empresa sabe que la cantidad del producto en cada botella tiene una distribución normal con una media de 16 onzas y una desviación estándar de 0.15 onzas. Una muestra de 50 botellas llenadas durante la hora pasada reveló que la cantidad media por botella era de 16.017 onzas. ¿Sugiere la evidencia que la cantidad media despachada es diferente de 16 onzas? Utilice un nivel de significancia de 0.05.

a) Establezca la hipótesis nula y la hipótesis alternativa.
b) ¿Cuál es la probabilidad de cometer un error tipo I?
c) Proporcione la fórmula del estadístico de la prueba.
d) Enuncie la regla de decisión.
e) Determine el valor del estadístico de la prueba.
f) ¿Cuál es su decisión respecto de la hipótesis nula?
g) Interprete, en un enunciado, el resultado de la prueba estadística.

Prueba de una cola

En el ejemplo anterior sólo se destacó el interés por informar al vicepresidente si ocurrió un cambio en la cantidad media de escritorios armados en la planta de Fredonia. No importaba si el cambio era un incremento o una disminución de la producción.

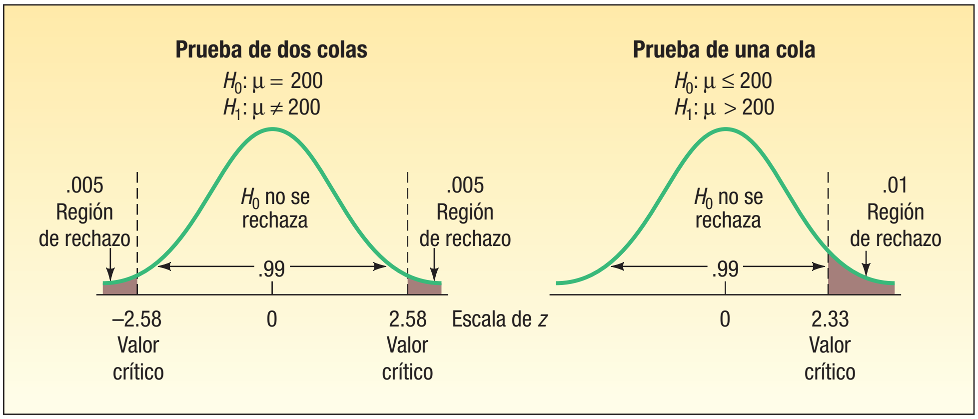
OA5 Distinguir entre las pruebas de hipótesis de una y dos colas.

Para ilustrar la prueba de una cola, vea otro problema. Suponga que el vicepresidente desea saber si hubo un *incremento* de la cantidad de unidades que se armaron. ¿Puede concluir, debido al mejoramiento de los métodos de producción, que la cantidad media de escritorios que se ensamblaron en las pasadas 50 semanas fue superior a 200? Observe la diferencia al formular el problema. En el primer caso deseaba conocer si había una *diferencia* en la cantidad media armada; en cambio, ahora desea saber si hubo un *incremento.* Como se investigan diferentes cuestiones, se plantea la hipótesis de otra manera. La diferencia más importante se presenta en la hipótesis alternativa. Antes se enunció la hipótesis alternativa como "diferente de"; ahora se enuncia como "mayor que". En símbolos:

Prueba de dos colas:	Prueba de una cola:
H_0: $\mu = 200$	H_0: $\mu \leq 200$
H_1: $\mu \neq 200$	H_1: $\mu > 200$

Los valores críticos en una prueba de una cola son diferentes a los de una prueba de dos colas en el mismo nivel de significancia. En el ejemplo anterior, se dividió el nivel de significancia a la mitad y se colocó una mitad en la cola inferior y la otra en la cola superior. En una prueba de una cola, toda la región de rechazo se coloca en una cola. Vea la gráfica 10-5.

En el caso de la prueba de una cola, el valor crítico es de 2.33, que se calcula: 1) se resta 0.01 de 0.5000 y 2) se determina el valor *z* correspondiente a 0.4900.



GRÁFICA 10-5 Regiones de rechazo de las pruebas de una y dos colas; $\alpha = 0.01$

10.7 Valor *p* en la prueba de hipótesis

Cuando se desea probar una hipótesis, se compara el estadístico de la prueba con un valor crítico. Se toma la decisión de rechazar la hipótesis nula o de no hacerlo. Así, por ejemplo, si el valor crítico es de 1.96 y el valor calculado del estadístico de prueba es de 2.19, la decisión consiste en rechazar la hipótesis nula.

OA7 Calcular e interpretar el valor *p*.

En años recientes, debido a la disponibilidad del software de computadora, con frecuencia se da información relacionada con la seguridad del rechazo o aceptación. Es decir, ¿cuánta confianza hay en el rechazo de la hipótesis nula? Este enfoque indica la probabilidad (en el

Estadística en acción

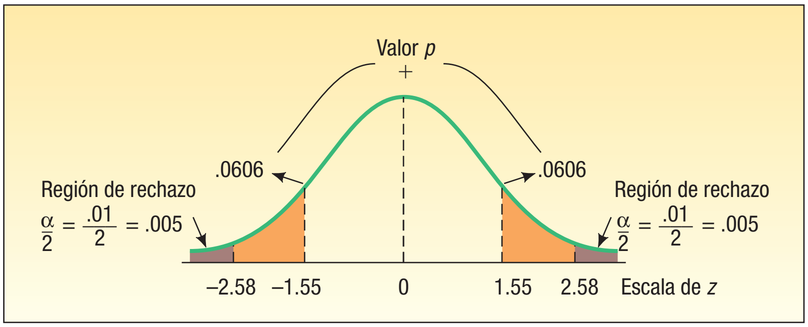
Existe una diferencia entre *estadísticamente significativo* y *prácticamente significativo.* Para explicarlo, suponga que crea una nueva píldora para adelgazar y la prueba en 100 000 personas. Concluye que la persona común que toma la píldora durante dos años pierde una libra. ¿Cree usted que mucha gente se interesaría en tomar la píldora para perder una libra? Los resultados de ingerir la nueva píldora fueron estadísticamente significativos, pero no prácticamente significativos.

supuesto de que la hipótesis nula sea verdadera) de obtener un valor del estadístico de la prueba por lo menos tan extremo como el valor real que se obtuvo. Este proceso compara la probabilidad, denominada **valor *p***, con el nivel de significancia. Si el valor *p* es menor que el nivel de significancia, H_0 se rechaza. Si es mayor que el nivel de significancia, H_0 no se rechaza.

VALOR *p* Probabilidad de observar un valor muestral tan extremo o más que el valor observado, si la hipótesis nula es verdadera.

La determinación del valor *p* no sólo da como resultado una decisión respecto de H_0, sino que brinda la oportunidad de observar la fuerza de la decisión. Un valor *p* muy pequeño, como 0.0001, indica que existe poca probabilidad de que H_0 sea verdadera. Por otra parte, un valor *p* de 0.2033 significa que H_0 no se rechaza y que existe poca probabilidad de que sea falsa.

¿Cómo calcular el valor *p*? Para ilustrarlo se recurre al ejemplo en el que se probó la hipótesis nula relativa a que la cantidad de escritorios producidos a la semana en Fredonia fue de 200. No se rechazó la hipótesis nula, pues el valor *z* de 1.55 cayó en la región comprendida entre -2.58 y 2.58. Se decidió no rechazar la hipótesis nula si el valor calculado de *z* caía en esta región. La probabilidad de hallar un valor *z* de 1.55 o más es de 0.0606, que se calcula mediante la diferencia de $0.5000 - 0.4394$. En otras palabras, la probabilidad de obtener una $\overline{X}$ mayor de 203.5 si $\mu = 200$ es de 0.0606. Para calcular el valor *p,* es necesario concentrarse en la región menor a -1.55, así como en los valores superiores a 1.55 (pues la región de rechazo se localiza en ambas colas). El valor *p* de dos colas es de 0.1212, que se calcula así: 2(0.0606). El valor *p* de 0.1212 es mayor que el nivel de significancia de 0.01 que se estableció al inicio, así que no se rechaza H_0. En la siguiente gráfica se muestran los detalles. En general, el área se duplica en una prueba de dos colas. Entonces, el valor *p* se compara con facilidad con el nivel de significancia. Se aplica la misma regla de decisión en el caso de una prueba de una cola.



Un valor *p* es una manera de expresar la probabilidad de que H_0 sea falsa. Pero, ¿cómo interpretar un valor *p*? Ya se mencionó que si el valor *p* es menor que el nivel de significancia, se rechaza H_0; si es mayor que el nivel de significancia, no se la rechaza. Asimismo, si el valor *p* es muy grande, es probable que H_0 sea verdadera. Si el valor *p* es pequeño, es probable que H_0 no lo sea. El siguiente recuadro permite interpretar los valores *p.*

INTERPRETACIÓN DE LA IMPORTANCIA DE LA EVIDENCIA EN CONTRA DE H_0 Si el valor *p* es menor que

a) 0.10, hay *cierta* evidencia de que H_0 no es verdadera.
b) 0.05, hay evidencia *fuerte* de que H_0 no es verdadera.
c) 0.01, hay evidencia *muy fuerte* de que H_0 no es verdadera.
d) 0.001, hay evidencia *extremadamente fuerte* de que H_0 no es verdadera.

Autoevaluación 10-2

Consulte la autoevaluación 10-1.

a) Suponga que se modifica el penúltimo enunciado para que diga: ¿La evidencia sugiere que la cantidad media despachada es *mayor a* 16 onzas? Establezca la hipótesis nula y la hipótesis alternativa en estas condiciones.

b) ¿Cuál es la regla de decisión en las nuevas condiciones definidas en el inciso *a*)?

c) Una segunda muestra de 50 contenedores llenos reveló que la media es de 16.040 onzas. ¿Cuál es el valor del estadístico de la prueba en esta muestra?

d) ¿Cuál es su decisión respecto de la hipótesis nula?

e) Interprete, en un solo enunciado, el resultado de la prueba estadística.

f) ¿Cuál es el valor *p*? ¿Cuál es su decisión respecto de la hipótesis nula con base en el valor *p*? ¿Es la misma conclusión a la que se llegó en el inciso *d*)?

Ejercicios

connect™

Responda las siguientes preguntas en los ejercicios 1 a 4: *a*) ¿es una prueba de una o de dos colas?; *b*) ¿cuál es la regla de decisión?; *c*) ¿cuál es el valor del estadístico de la prueba?; *d*) ¿cuál es su decisión respecto de H_0?; *e*) ¿cuál es el valor *p*? Interprete este valor.

1. Se selecciona una muestra de 36 observaciones de una población normal. La media muestral es de 49, y el tamaño de la muestra, de 36. La desviación estándar de la población es 5. Utilice el nivel de significancia de 0.05.

$$H_0: \mu = 50$$
$$H_1: \mu \neq 50$$

2. Se selecciona una muestra de 36 observaciones de una población normal. La media muestral es de 12, y el tamaño de la muestra, 36. La desviación estándar de la población es 3. Utilice el nivel de significancia 0.02.

$$H_0: \mu \leq 10$$
$$H_1: \mu > 10$$

3. Se selecciona una muestra de 36 observaciones de una población normal. La media de la muestra es 21, y la desviación estándar de la población, 5. Lleve a cabo la prueba de hipótesis con el nivel de significancia de 0.05.

$$H_0: \mu \leq 20$$
$$H_1: \mu > 20$$

4. Se selecciona una muestra de 64 observaciones de una población normal. La media de la muestra es 215, y la desviación estándar de la población, 15. Lleve a cabo la prueba de hipótesis, utilice el nivel de significancia 0.03.

$$H_0: \mu \geq 220$$
$$H_1: \mu < 220$$

En el caso de los ejercicios 5 a 8: *a*) establezca la hipótesis nula y la hipótesis alternativa; *b*) defina la regla de decisión; *c*) calcule el valor del estadístico de la prueba; *d*) ¿cuál es su decisión respecto de H_0?; *e*) ¿cuál es el valor *p*? Interprételo.

5. El fabricante de neumáticos radiales con cinturón de acero X-15 para camiones señala que el millaje medio que cada uno recorre antes de que se desgasten las cuerdas es de 60 000 millas. La desviación estándar del millaje es de 5 000 millas. La Crosset Truck Company compró 48 neumáticos y comprobó que el millaje medio para sus camiones es de 59 500 millas. ¿La experiencia de Crosset es diferente de lo que afirma el fabricante en el nivel de significancia de 0.05?

6. La cadena de restaurantes MacBurger afirma que el tiempo de espera de los clientes es de 8 minutos con una desviación estándar poblacional de 1 minuto. El departamento de control de calidad halló en una muestra de 50 clientes en Warren Road MacBurger que el tiempo medio de espera era de 2.75 minutos. Con el nivel de significancia de 0.05, ¿puede concluir que el tiempo medio de espera sea menor a 3 minutos?

7. Una encuesta nacional reciente determinó que los estudiantes de secundaria veían en promedio (media) 6.8 películas en DVD al mes, con una desviación estándar poblacional de 0.5 horas. Una muestra aleatoria de 36 estudiantes universitarios reveló que la cantidad media de películas en DVD que vieron el mes pasado fue de 6.2. Con un nivel de significancia de 0.05, ¿puede concluir que los estudiantes universitarios ven menos películas en DVD que los estudiantes de secundaria?

8. En el momento en que fue contratada como mesera en el Grumney Family Restaurant, a Beth Brigden le dijeron: “Puedes ganar en promedio más de \$80 al día en propinas.” Suponga que la desviación estándar de la distribución de población es de \$3.24. Los primeros 35 días de trabajar en el restaurante, la suma media de sus propinas fue de \$84.85. Con el nivel de significancia de 0.01, ¿la señorita Brigden puede concluir que gana un promedio de más de \$80 en propinas?

10.8 Prueba de la media poblacional: desviación estándar de la población desconocida

En el ejemplo anterior se conocía σ, la desviación estándar de la población. No obstante, en la mayoría de los casos, la desviación estándar de la población es desconocida. Por consiguiente, σ debe basarse en estudios previos o calcularse por medio de la desviación estándar de la muestra, *s.* La desviación estándar poblacional en el siguiente ejemplo no se conoce, por lo que se emplea la desviación estándar muestral para estimar σ.

Para determinar el valor del estadístico de la prueba utilice la distribución *t* y modifique la fórmula (10.1) de la siguiente manera:

PRUEBA DE UNA MEDIA; σ DESCONOCIDA

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} \qquad \textbf{(10-2)}$$

con $n - 1$ grados de libertad, donde:

$\bar{X}$ representa la media de la muestra.
μ, la media poblacional hipotética.
s, la desviación estándar de la muestra.
n, el número de observaciones incluidas en la muestra.

Es una situación similar a cuando construyó intervalos de confianza en el capítulo anterior. Vea las páginas 306-312, capítulo 9. En la gráfica 9-3 de la página 309 se resumió el problema. En estas condiciones, el procedimiento estadístico correcto consiste en sustituir la distribución normal estándar con la distribución *t.* Para repasar las principales características de la distribución *t*:

- Es una distribución continua.
- Tiene forma de campana y es simétrica.
- Existe una familia de distribuciones *t*; cada vez que se cambia de grados de libertad, se crea una nueva distribución.
- Conforme se incrementa el número de grados de libertad, la forma de la distribución *t* se aproxima a la de la distribución normal estándar.
- La distribución *t* es plana, o más dispersa, que la distribución normal estándar.

El siguiente ejemplo muestra los detalles.

Ejemplo

El departamento de quejas de McFarland Insurance Company informa que el costo medio para tramitar una queja es de \$60. Una comparación en la industria demostró que esta cantidad es mayor que en las demás compañías de seguros, así que la compañía tomó medidas para reducir gastos. Para evaluar el efecto de las medidas de reducción de gastos, el supervisor del departamento de quejas seleccionó una muestra aleatoria de 26 quejas atendidas el mes pasado. La información de la muestra aparece a continuación.

$45	$49	$62	$40	$43	$61
48	53	67	63	78	64
48	54	51	56	63	69
58	51	58	59	56	57
38	76				

¿Es razonable concluir que el costo medio de atención de una queja ahora es menor a $60 con un nivel de significancia de 0.01?

Solución

OA6 Llevar a cabo una prueba de hipótesis de una media poblacional.

Aplique la prueba de hipótesis con el procedimiento de los cinco pasos.

Paso 1: Se establecen las hipótesis nula y alternativa. La hipótesis nula consiste en que la media poblacional es de por lo menos $60. La hipótesis alternativa consiste en que la media poblacional es menor a $60. Se expresan las hipótesis nula y alternativa de la siguiente manera:

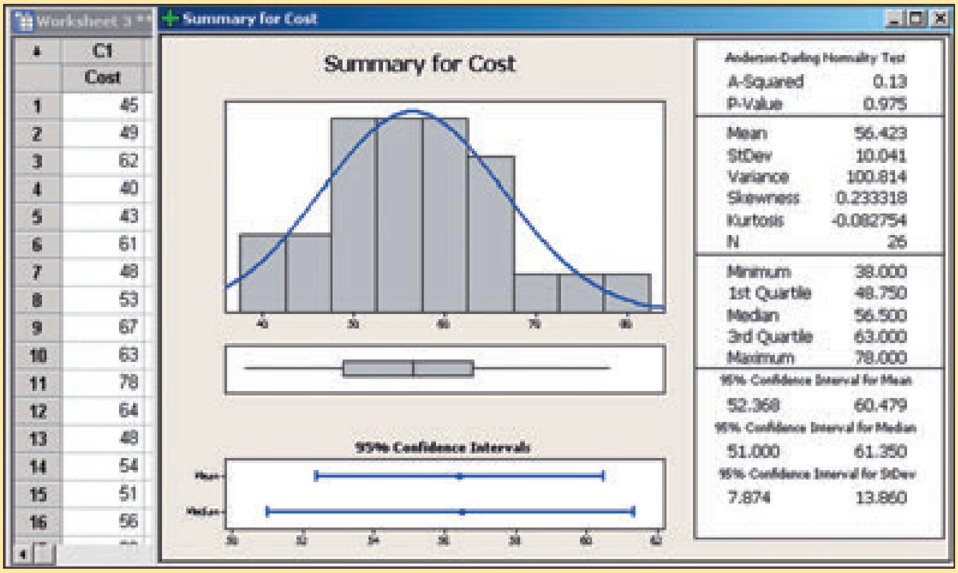
$$H_0: \mu \geq \$60$$

$$H_1: \mu < \$60$$

La prueba es de *una* cola, pues desea determinar si hubo una *reducción* en el costo. La desigualdad en la hipótesis alternativa señala la región de rechazo en la cola izquierda de la distribución.

Paso 2: Se selecciona un nivel de significancia. El nivel de significancia es 0.01.

Paso 3: Se identifica el estadístico de la prueba. En este caso, el estadístico de la prueba es la distribución *t.* ¿Por qué? Primero, porque resulta razonable concluir que la distribución del costo por queja sigue la distribución normal. Puede confirmarlo a partir del histograma a la derecha de la siguiente captura de pantalla de Minitab. Observe la distribución normal superpuesta en la distribución de frecuencias.



No se conoce la desviación estándar de la población, por lo que ésta se sustituye por la desviación estándar de la muestra. El valor del estadístico de la prueba se calcula por medio de la fórmula (10-2):

$$t = \frac{\overline{X} - \mu}{s/\sqrt{n}}$$

Paso 4: Se formula una regla para tomar decisiones. Los valores críticos de t aparecen en el apéndice B.2, una parte del cual se reproduce en la tabla 10-1. La columna extrema izquierda de la tabla está rotulada como *gl,* que representa los grados de libertad. El número de grados de libertad es el total de observaciones incluidas en la muestra menos el número de poblaciones muestreadas, lo cual se escribe $n - 1$. En este caso, el número de observaciones de la muestra es de 26, y se muestrea una población, así que hay $26 - 1 = 25$ grados de libertad. Para determinar el valor crítico, primero localice el renglón con los grados de libertad adecuados. Este renglón se encuentra sombreado en la tabla 10-1. Luego determine si la prueba es de una o de dos colas. En este caso, es una prueba de una cola, así que busque la sección de la tabla rotulada *una cola.* Localice la columna con el nivel de significancia elegido. En este ejemplo, el nivel de significancia es de 0.01. Desplácese hacia abajo por la columna rotulada *0.01* hasta intersecar el renglón con 25 grados de libertad. El valor es de 2.485. Como se trata de una prueba de una cola y la región de rechazo se localiza en la cola izquierda, el valor crítico es negativo. La regla de decisión consiste en rechazar H_0 si el valor de t es menor a -2.485.

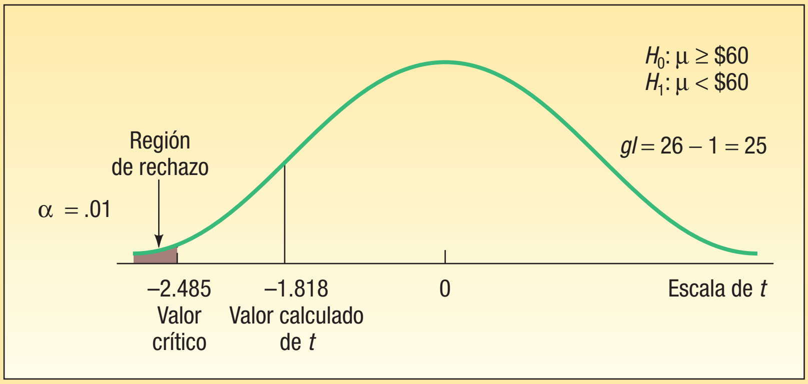
TABLA 10-1 Parte de la tabla de la distribución t

	Intervalos de confianza					
	80%	**90%**	**95%**	**98%**	**99%**	**99.9%**
	Nivel de significancia de una prueba de una cola, α					
gl	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.0005**
	Nivel de significancia de una prueba de dos colas, α					
	0.20	**0.10**	**0.05**	**0.02**	**0.01**	**0.001**
⋮	⋮	⋮	⋮	⋮	⋮	⋮
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	3.768
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

Paso 5: Se toma una decisión y se interpreta el resultado. De acuerdo con la pantalla de Minitab, a la derecha del histograma, el costo medio por queja de la muestra de 26 observaciones es de \$56.42. La desviación estándar de esta muestra es de \$10.04. Al sustituir estos valores en la fórmula (10-2) y calcular el valor de t:

$$t = \frac{\overline{X} - \mu}{s/\sqrt{n}} = \frac{\$56.42 - \$60}{\$10.04/\sqrt{26}} = -1.818$$

Como -1.818 se localiza en la región ubicada a la derecha del valor crítico de -2.485, la hipótesis nula no se rechaza con el nivel de significancia de 0.01. No se demostró que las medidas de reducción de costos hayan bajado el costo medio por queja a menos de \$60. En otras palabras, la diferencia de \$3.58 (\$56.52 − \$60) entre la media muestral y la media poblacional puede deberse al error de muestreo. El valor calculado de t aparece en la gráfica 10-6. Éste se encuentra en la región en que la hipótesis nula *no* se rechaza.



GRÁFICA 10-6 Región de rechazo, distribución t, nivel de significancia 0.01

En el ejemplo anterior, la media y la desviación estándar se calcularon con Minitab. El siguiente ejemplo muestra los detalles cuando se calculan la media y la desviación estándar a partir de los datos de la muestra.

Ejemplo

La longitud media de una pequeña barra de contrapeso es de 43 milímetros. Al supervisor de producción le preocupa que hayan cambiado los ajustes de la máquina de producción de barras. Solicita una investigación al departamento de ingeniería, que selecciona una muestra aleatoria de 12 barras y las mide. Los resultados aparecen en seguida, expresados en milímetros.

42	39	42	45	43	40	39	41	40	42	43	42

¿Es razonable concluir que cambió la longitud media de las barras? Utilice el nivel de significancia 0.02.

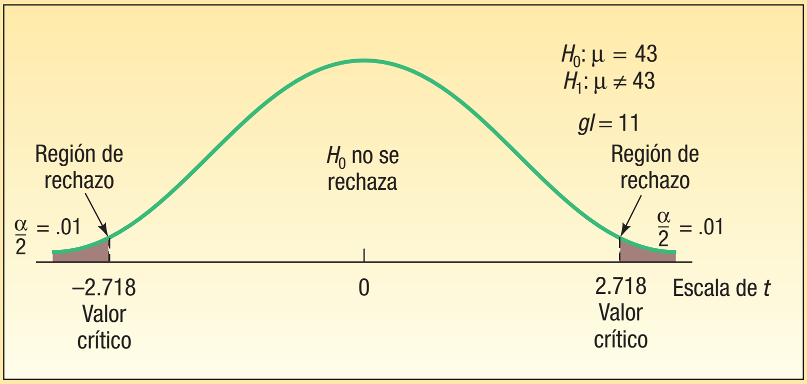
Solución

OA6 Llevar a cabo una prueba de hipótesis de una media poblacional.

Primero formule la hipótesis nula y la hipótesis alternativa.

$$H_0: \mu = 43$$
$$H_1: \mu \neq 43$$

La hipótesis alternativa no señala una dirección, así que se trata de una prueba de dos colas. Hay 11 grados de libertad, que se calculan por medio de $n - 1 = 12 - 1 = 11$. El valor t es de 2.718, que se determina con el apéndice B.2 en el caso de una prueba de dos colas con un nivel de significancia de 0.02 y 11 grados de libertad. La regla de decisión es: se rechaza la hipótesis nula si el valor calculado de t se localiza a la izquierda de −2.718 o a la derecha de 2.718. Esta información se resume en la gráfica 10-7.



GRÁFICA 10-7 Regiones de rechazo, prueba de dos colas, distribución t de Student, α = 0.02

Se calcula la desviación estándar de la muestra con la fórmula (3-11). La media, $\bar{X}$, es de 41.5 milímetros, y la desviación estándar, s, 1.784 milímetros. Los detalles aparecen en la tabla 10-2.

TABLA 10-2 Cálculos de la desviación estándar de la muestra

X (mm)	$X - \bar{X}$	$(X - \bar{X})^2$
42	0.5	0.25
39	−2.5	6.25
42	0.5	0.25
45	3.5	12.25
43	1.5	2.25
40	−1.5	2.25
39	−2.5	6.25
41	−0.5	0.25
40	−1.5	2.25
42	0.5	0.25
43	1.5	2.25
42	0.5	0.25
498	0	35.00

$$\bar{X} = \frac{498}{12} = 41.5 \text{ mm}$$

$$s = \sqrt{\frac{\Sigma(X - \bar{X})^2}{n - 1}} = \sqrt{\frac{35}{12 - 1}} = 1.784$$

Ahora puede calcular el valor de t con la fórmula (10-2).

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} = \frac{41.5 - 43.0}{1.784/\sqrt{12}} = -2.913$$

La hipótesis nula que afirma que la media poblacional es de 43 milímetros se rechaza porque el valor calculado de t de −2.913 se encuentra en el área a la izquierda de −2.718. Se acepta la hipótesis alternativa y se concluye que la media poblacional no es de 43 milímetros. La máquina está fuera de control y necesita algunos ajustes.

Autoevaluación 10-3

La vida media de una batería de un reloj digital es de 305 días. Las vidas medias de las baterías se rigen por la distribución normal. Hace poco se modificó la batería para que tuviera mayor duración. Una muestra de 20 baterías modificadas exhibió una vida media de 311 días con una desviación estándar de 12 días. ¿La modificación incrementó la vida media de la batería?

a) Formule la hipótesis nula y la hipótesis alternativa.
b) Muestre la gráfica de la regla de decisión. Utilice el nivel de significancia 0.05.
c) Calcule el valor de *t.* ¿Cuál es su decisión respecto de la hipótesis nula? Resuma sus resultados.

Ejercicios

9. Sean las siguientes hipótesis:

$$H_0: \mu \leq 10$$
$$H_1: \mu > 10$$

En el caso de una muestra aleatoria de 10 observaciones seleccionada de una población normal, la media muestral fue de 12, y la desviación estándar de la muestra, de 3. Utilice el nivel de significancia 0.05:

a) Formule la regla de decisión.
b) Calcule el valor del estadístico de prueba.
c) ¿Cuál es su decisión respecto de la hipótesis nula?

10. Sean las siguientes hipótesis:

$$H_0: \mu = 400$$
$$H_1: \mu \neq 400$$

En el caso de una muestra aleatoria de 12 observaciones seleccionada de una población normal, la media muestral fue de 407, y la desviación estándar de la muestra, de 6. Utilice el nivel de significancia 0.01:

a) Formule la regla de decisión.
b) Calcule el valor del estadístico de prueba.
c) ¿Cuál es su decisión respecto de la hipótesis nula?

11. El gerente de ventas del distrito de las Montañas Rocallosas de Rath Publishing, Inc., editorial de textos universitarios, afirma que los representantes de ventas realizan en promedio 40 llamadas de ventas a la semana a profesores. Varios representantes señalan que el cálculo es muy bajo. Una muestra aleatoria de 28 representantes de ventas revela que la cantidad media de llamadas que se realizó la semana pasada fue de 42. La desviación estándar de la muestra es de 2.1 llamadas. Con el nivel de significancia de 0.05, ¿puede concluir que la cantidad media de llamadas semanales por vendedor es de más de 40?

12. La administración de White Industries analiza una nueva técnica para armar un carro de golf; la técnica actual requiere 42.3 minutos de trabajo en promedio. El tiempo medio de montaje de una muestra aleatoria de 24 carros, con la nueva técnica, fue de 40.6 minutos, y la desviación estándar, de 2.7 minutos. Con un nivel de significancia de 0.10, ¿puede concluir que el tiempo de montaje con la nueva técnica es más breve?

13. El ingreso promedio por persona en Estados Unidos es de $40 000, y la distribución de ingresos sigue una distribución normal. Una muestra aleatoria de 10 residentes de Wilmington, Delaware, presentó una media de $50 000, con una desviación estándar de $10 000. A un nivel de significancia de 0.05, ¿existe suficiente evidencia para concluir que los residentes de Wilmington, Delaware, ganan más que el promedio nacional?

14. En la actualidad, la mayoría de quienes viajan por avión compra sus boletos por internet. De esta forma, los pasajeros evitan la preocupación de cuidar un boleto de papel, además de que las aerolíneas ahorran. No obstante, en fechas recientes, las aerolíneas han recibido quejas relacionadas con los boletos, en particular cuando se requiere hacer un enlace para cambiar de línea. Para analizar el problema, una agencia de investigación independiente tomó una muestra aleatoria de 20 aeropuertos y recogió información relacionada con la cantidad de quejas que hubo sobre los boletos durante marzo. A continuación se presenta la información.

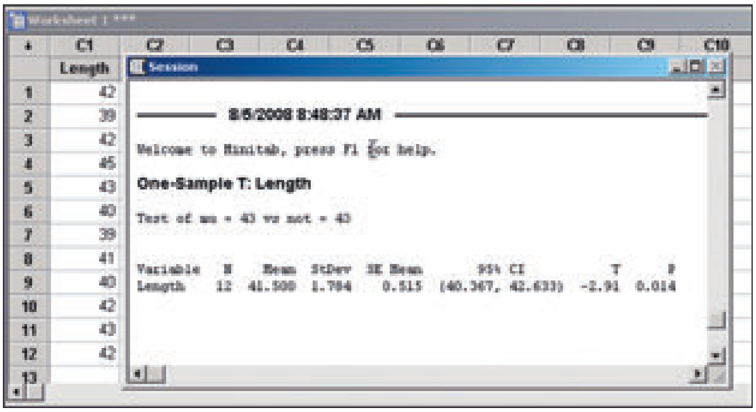
14	14	16	12	12	14	13	16	15	14
12	15	15	14	13	13	12	13	10	13

Con un nivel de significancia de 0.05, ¿la agencia de investigación puede concluir que la cantidad media de quejas por aeropuerto es menor de 15 al mes?

a) ¿Qué suposición se requiere antes de llevar a cabo una prueba de hipótesis?
b) Ilustre la cantidad de quejas por aeropuerto en una distribución de frecuencias o en un diagrama de dispersión. ¿Es razonable concluir que la población se rige por una distribución normal?
c) Realice una prueba de hipótesis e interprete los resultados.

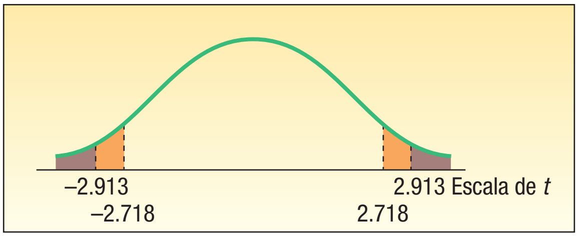
Solución con software

El sistema de software de estadística Minitab, que se utilizó en los capítulos precedentes y en la sección anterior, proporciona una forma eficaz de llevar a cabo una prueba de hipótesis de una cola para la media de la población. Los pasos para generar la siguiente captura de pantalla aparecen en la sección de **Comandos de software**, al final del capítulo.



Una característica adicional de la mayoría de los paquetes de software consiste en que calculan el valor *p*, el cual proporciona más información sobre la hipótesis nula. El valor *p* es la probabilidad de un valor *t* tan extremo como el que se calculó, en caso de que la hipótesis nula sea verdadera. De acuerdo con los datos del ejemplo anterior, de la barra de contrapeso, el valor *p* de 0.014 es la probabilidad de un valor *t* de −2.91 o menor más la probabilidad de un valor *t* de 2.91 o mayor, con una media poblacional de 43. Así, la comparación del valor *p* con el nivel de significancia indica si la hipótesis nula se encontraba cerca de ser rechazada, si apenas se rechazó, etcétera.

El siguiente diagrama contiene una explicación más detallada. El valor *p* de 0.014 es el área más oscura o sombreada, y el nivel de significancia es la totalidad del área sombreada. Como el valor *p* de 0.014 es menor que el nivel de significancia de 0.02, la hipótesis nula se rechaza. Si el valor *p* hubiera sido mayor que el nivel de significancia, 0.06, 0.19 o 0.57, la hipótesis nula no se habría rechazado. Si se hubiera elegido un valor de 0.01 para el nivel de significancia, la hipótesis nula no se habría rechazado.



En el ejemplo anterior, la hipótesis alternativa era de dos colas, así que había áreas de rechazo tanto en la cola inferior (izquierda) como en la superior (derecha). Para calcular el valor *p* fue necesario determinar el área a la izquierda de −2.913 de una distribución *t* con 11 grados de libertad y sumarla al valor del área a la derecha de 2.913, también con 11 grados de libertad.

¿Y si se tratara de una prueba de una cola, de forma que toda la región de rechazo se localizara ya en la cola superior, ya en la cola inferior? En dicho caso, se indicaría un área a partir de una sola cola. En el ejemplo de la barra de contrapeso, si H_1 se definiera como $\mu < 43$, la desigualdad apuntaría hacia la izquierda. Por consiguiente, se señalaría el valor *p* como el área a la izquierda de −2.913. Este valor es 0.007, que se calcula al dividir 0.014/2. Por lo tanto, el valor *p* de una prueba de una cola sería 0.007.

¿Cómo calcular un valor *p* sin una computadora? Para ilustrarlo, recuerde que, en el ejemplo relativo a la longitud de la barra de contrapeso, se rechazó la hipótesis nula que indicaba

TABLA 10-3 Parte de la distribución *t* de Student

	Intervalos de confianza					
	80%	**90%**	**95%**	**98%**	**99%**	**99.9%**
	Nivel de significancia de una prueba de una cola, α					
gl	**0.10**	**0.05**	**.0025**	**0.01**	**0.005**	**0.0005**
	Nivel de significancia de una prueba de dos colas, α					
	0.20	**0.10**	**0.05**	**0.02**	**0.01**	**0.001**
⋮	⋮	⋮	⋮	⋮	⋮	⋮
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

que $\mu = 43$ y se aceptó la hipótesis alternativa que indicaba que $\mu \neq 43$. El nivel de significancia era de 0.02, así que, por lógica, el valor *p* es menor que 0.02. Para calcular el valor *p* con mayor precisión, vea el apéndice B.2 y localice el renglón con 11 grados de libertad. El valor calculado de *t,* 2.913, se localiza entre 2.718 y 3.106 (parte del apéndice B.2 se reproduce en la tabla 10-3). El nivel de significancia de dos colas correspondiente a 2.718 es 0.02, y en el caso de 3.106, es 0.01. Por lo tanto, el valor *p* se encuentra entre 0.01 y 0.02. Se acostumbra indicar que el valor *p* es *menor* que el mayor de los dos niveles de significancia. Así: "el valor *p* es menor que 0.02".

Autoevaluación 10-4

Se programa una máquina para llenar un frasco pequeño con 9.0 gramos de medicamento. Una muestra de ocho frascos arrojó las siguientes cantidades (en gramos) por botella.

9.2	8.7	8.9	8.6	8.8	8.5	8.7	9.0

¿Puede concluir que el peso medio es inferior a 9.0 gramos si el nivel de significancia es de 0.01?
a) Formule la hipótesis nula y la hipótesis alternativa.
b) ¿Cuántos grados de libertad existen?
c) Establezca la regla de decisión.
d) Calcule el valor de *t.* ¿Qué decide respecto de la hipótesis nula?
e) Estime el valor *p.*

Ejercicios

connect

15. Sean las siguientes hipótesis:

$$H_0: \mu \geq 20$$
$$H_1: \mu < 20$$

Una muestra aleatoria de cinco elementos dio como resultado los siguientes valores: 18, 15, 12, 19 y 21. ¿Puede concluir que la media poblacional es menor que 20 con un nivel de significancia de 0.01?
a) Establezca la regla de decisión.
b) Calcule el valor del estadístico de prueba.
c) ¿Cuál es su decisión en lo que se refiere a la hipótesis nula?
d) Calcule el valor de *p.*

16. Sean las siguientes hipótesis:

$$H_0: \mu = 100$$
$$H_1: \mu \neq 100$$

Una muestra aleatoria de seis elementos dio como resultado los siguientes valores: 118, 105, 112, 119, 105 y 111. ¿Puede concluir que la media poblacional es diferente de 100 con un nivel de significancia de 0.05?
a) Establezca la regla de decisión.
b) Calcule el valor del estadístico de prueba.
c) ¿Cuál es su decisión en lo que se refiere a la hipótesis nula?
d) Calcule el valor de *p.*

17. La cantidad de agua consumida al día por un adulto sano sigue una distribución normal, con una media de 1.4 litros. Una campaña de salud promueve el consumo de cuando menos 2.0 litros diarios. Después de la campaña, una muestra de 10 adultos muestra el siguiente consumo en litros:

1.5	1.6	1.5	1.4	1.9	1.4	1.3	1.9	1.8	1.7

A un nivel de significancia de 0.01, ¿se puede concluir que se ha elevado el consumo de agua? Calcule e interprete el valor *p.*

18. El cloro líquido que se agrega a las albercas para combatir las algas tiene una duración relativamente corta en las tiendas antes de que pierda su eficacia. Los registros indican que la duración media de un frasco de cloro es de 2 160 horas (90 días). Como experimento, se agregó Holdlonger al cloro para saber si éste incrementaba la duración del cloro. Una muestra de nueve frascos de cloro arrojó los siguientes tiempos de duración (en horas) en las tiendas:

2 159	2 170	2 180	2 179	2 160	2 167	2 171	2 181	2 185

¿Con el nivel de significancia de 0.025, ¿incrementó el Holdlonger la duración del cloro en las tiendas? Calcule el valor *p*.

19. Un grupo de expertos en Washington, D.C. anuncia que el adolescente típico envió 50 mensajes de texto por día durante 2009. Para actualizar la estimación, usted contacta por teléfono a una muestra de adolescentes y les pregunta cuántos mensajes enviaron el día anterior. Sus respuestas fueron:

51	175	47	49	44	54	145	203	21	59	42	100

A un nivel de significancia de 0.05, ¿puede concluir que el número medio es mayor a 50? Estime el valor *p* y describa qué le revela.

20. Hugger Polls afirma que un agente realiza una media de 53 entrevistas extensas a domicilio a la semana. Se introdujo un nuevo formulario para las entrevistas, y Hugger desea evaluar su eficacia. La cantidad de entrevistas extensas por semana de una muestra aleatoria de agentes es:

53	57	50	55	58	54	60	52	59	62	60	60	51	59	56

Con un nivel de significancia de 0.05, ¿puede concluir que la cantidad media de entrevistas de los agentes es más de 53 a la semana? Calcule el valor de *p*.

10.9 Pruebas relacionadas con proporciones

En el capítulo anterior se analizaron los intervalos de confianza para proporciones. Vea la sección 9.4 en las páginas 313-316. También puede llevar a cabo una prueba de hipótesis de una proporción. Recuerde que una proporción es la razón entre el número de éxitos y el número de observaciones. Si X se refiere al número de éxitos y n al de observaciones, la proporción de éxitos en una cantidad fija de pruebas es X/n. Por consiguiente, la fórmula para calcular una proporción muestral, p, es $p = X/n$. Considere los siguientes casos de posibles pruebas de hipótesis.

- Según sus registros, General Motors informa que 70% de los vehículos rentados se devuelve con menos de 36 000 millas. Una muestra reciente de 200 vehículos devueltos al final de su periodo de renta mostró que 158 tenían menos de 36 000 millas. ¿Se incrementó la proporción?
- La American Association of Retired Persons (AARP) informa que 60% de los retirados de menos de 65 años de edad regresaría a trabajar de tiempo completo si hubiera disponible un trabajo adecuado. Una muestra de 500 retirados de menos de 65 años reveló que 315 volverían a trabajar. ¿Puede concluir que más de 60% volvería a trabajar?
- Able Moving and Storage, Inc., anuncia a sus clientes que el traslado a largas distancias de los bienes familiares se entregarán de 3 a 5 días a partir del momento de recogerlos. Los registros de Able muestran que han tenido éxito 90% de las veces. Una auditoría reciente mostró que de 200 veces, 190 tuvieron éxito. ¿La compañía puede concluir que aumentó este registro de éxitos?

OA8 Llevar a cabo una prueba de hipótesis de una proporción poblacional.

Se deben hacer algunas suposiciones antes de probar una proporción de población. Para probar una hipótesis relativa a una proporción de población, se elige una muestra aleatoria de ésta. Se supone que se satisfacen los supuestos binomiales del capítulo 6: 1) los datos de la muestra que se recogen son resultado de conteos; 2) el resultado de un experimento se clasifica en una de dos categorías mutuamente excluyentes —"éxito" o "fracaso"—; 3) la probabilidad de un éxito es la misma para cada prueba; 4) las pruebas son independientes, lo cual significa que el resultado de una prueba no influye en el resultado de las demás. La prueba que realizará en breve es adecuada cuando $n\pi$ y $n(1 - \pi)$ son de al menos 5. El tamaño de la muestra es n, y π, la proporción poblacional. Se tiene la ventaja de que una distribución binomial puede aproximarse por medio de la distribución normal.

Ejemplo

Suponga que a partir de las elecciones anteriores en un estado, para que sea electo un candidato a gobernador, es necesario que gane por lo menos 80% de los votos de la zona norte. El gobernador de turno está interesado en evaluar sus posibilidades de volver al cargo y hace planes para llevar a cabo una encuesta de 2 000 votantes registrados en esa región.

Aplique el procedimiento para probar hipótesis y evalúe las posibilidades de que el gobernador se reelija.

Solución

Este caso de la reelección del gobernador satisface las condiciones binomiales.

- Sólo hay dos posibles resultados. Es decir, un votante entrevistado votará o no por el gobernador.
- La probabilidad de un éxito es la misma para cada prueba. En este caso, la probabilidad de que cualquier votante entrevistado apoye la reelección es de 0.80.
- Las pruebas son independientes. Esto significa, por ejemplo, que la probabilidad de que el votante 23 entrevistado apoye la reelección no resulta afectada por lo que hagan los votantes 24 y 52.
- Los datos de la muestra son el resultado de conteos. Vamos a contar el número de votantes que apoya la reelección en la muestra de 2 000.

Se puede utilizar la aproximación normal de la distribución binomial que se analizó en el capítulo 7, pues $n\pi$ y $n(1 - \pi)$ exceden de 5. En este caso, $n = 2\,000$ y $\pi = 0.80$ (π es la proporción de votos en la parte norte del estado, u 80%, necesarios). Por lo tanto, $n\pi = 2\,000(0.80) = 1\,600$ y $n(1 - \pi) = 2\,000(1 - 0.80) = 400$. Ambos, 1 600 y 400, son mayores que 5.

Paso 1: Se establecen las hipótesis nula y alternativa. La hipótesis nula, H_0, consiste en que la proporción de la población π es 0.80 o mayor. La hipótesis alternativa, H_1, es que la proporción es menor a 0.80. Desde un punto de vista práctico, al gobernador de turno sólo le interesa cuando la proporción es menor de 0.80. Si es igual o mayor que 0.80, no pondrá objeciones; es decir, los datos de la muestra indicarían que probablemente se le reelija. Estas hipótesis se escriben simbólicamente de la siguiente manera:

$$H_0: \pi \geq .80$$
$$H_1: \pi < .80$$

H_1 establece una dirección. Por consiguiente, como se hizo notar antes, la prueba es de una cola, en la que el signo de desigualdad apunta a la cola de la distribución que contiene la región de rechazo.

Paso 2: Se selecciona el nivel de significancia. El nivel de significancia es de 0.05. Ésta es la probabilidad de rechazar una hipótesis verdadera.

Paso 3: Seleccione el estadístico de prueba. El estadístico adecuado es z, que se determina de la siguiente manera:

PRUEBA DE HIPÓTESIS DE UNA PROPORCIÓN

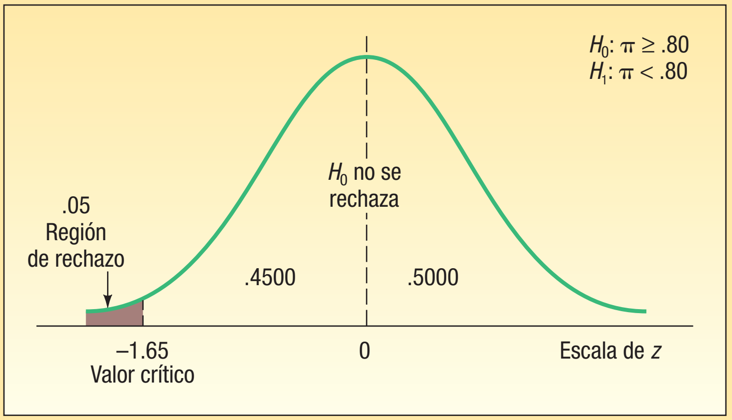
$$z = \frac{p - \pi}{\sqrt{\dfrac{\pi(1 - \pi)}{n}}} \qquad \textbf{(10-3)}$$

donde:

π es la proporción poblacional.
p es la proporción de la muestra.
n es el tamaño de la muestra.

Determinación del valor crítico.

Paso 4: Se formula la regla de decisión. El valor o los valores críticos de z forman el punto o puntos de división entre las regiones en las que se rechaza H_0 y en la que no se rechaza. Como la hipótesis alternativa indica una dirección, se trata de una prueba de una cola. El signo de la desigualdad apunta hacia la izquierda, así que sólo se utiliza el lado izquierdo de la curva. (Vea la gráfica 10-8.) El nivel de significancia del paso 2 fue de 0.05. Esta probabilidad se encuentra en la cola



GRÁFICA 10-8 Región de rechazo del nivel de significancia de 0.05, prueba de una cola

izquierda y determina la región de rechazo. El área entre cero y el valor crítico es de 0.4500, que se calcula mediante 0.5000 − 0.0500. Y cuando se busca 0.4500 en el apéndice B.1, se halla que el valor crítico de z es 1.65. Por lo tanto, la regla de decisión es: se rechaza la hipótesis nula y se acepta la hipótesis alternativa si el valor calculado de z cae a la izquierda de −1.65; de otra forma no se rechaza H_0.

Se selecciona una muestra y se toma una decisión respecto de H_0.

Paso 5: Se toma una decisión y se interpreta el resultado. Se selecciona una muestra y se toma una decisión respecto de H_0. Un sondeo de muestra de 2 000 posibles electores en la parte norte del estado reveló que 1 550 pensaban votar por el gobernador de turno. ¿Se encuentra la proporción de la muestra de 0.775 (calculada con la operación 1 550/2 000) lo bastante cerca de 0.80 para concluir que la diferencia se debe al error de muestreo? En este caso:

p tiene un valor de 0.775 y representa la proporción en la muestra que planea votar por el gobernador.
n tiene un valor de 2 000 y representa el número de votantes entrevistados.
π tiene un valor de 0.80 y representa la proporción de población hipotética.
z es un estadístico de prueba con una distribución normal cuando la hipótesis es verdadera y los demás supuestos son verdaderos.

Con la fórmula (10-3) se calcula el valor de z:

$$z = \frac{p - \pi}{\sqrt{\dfrac{\pi(1 - \pi)}{n}}} = \frac{\dfrac{1\,550}{2\,000} - .80}{\sqrt{\dfrac{.80(1 - .80)}{2\,000}}} = \frac{.775 - .80}{\sqrt{.00008}} = -2.80$$

El valor calculado de z (−2.80) se encuentra en la región de rechazo, por lo que la hipótesis nula se rechaza en el nivel 0.05. La diferencia de 2.5 puntos porcentuales entre el porcentaje de la muestra (77.5%) y el porcentaje de la población hipotética en la parte norte del estado que se requiere para ganar las elecciones estatales (80%) resulta estadísticamente significativa. Quizá no se deba a la variación muestral. En otras palabras, le evidencia no apoya la afirmación de que el gobernador de turno vuelva a su mansión otros cuatro años.

El valor p es la probabilidad de hallar un valor z inferior a −2.80. De acuerdo con el apéndice B.1, la probabilidad de un valor de z entre cero y −2.80 es de 0.4974. Así, el valor p es 0.0026, que se determina con el cálculo de 0.5000 − 0.4974. El gobernador no puede confiar en la reelección porque el valor p es inferior al nivel de significancia.

Autoevaluación 10-5

Un informe reciente de la industria de seguros indicó que 40% de las personas implicadas en accidentes de tránsito menores había tenido por lo menos un accidente los pasados cinco años. Un grupo de asesoría decidió investigar dicha afirmación, pues creía que la cantidad era muy grande. Una muestra de 200 accidentes de tránsito de este año mostró que 74 personas también estuvieron involucradas en otro accidente los pasados cinco años. Utilice el nivel de significancia 0.01.

a) ¿Se puede emplear z como estadístico de la prueba? Indique la razón.
b) Formule la hipótesis nula y la hipótesis alternativa.
c) Muestre gráficamente la regla de decisión.
d) Calcule el valor z y plantee su decisión respecto de la hipótesis nula.
e) Determine e interprete el valor p.

Ejercicios

21. Sean las siguientes hipótesis:

$$H_0: \pi \leq .70$$
$$H_1: \pi > .70$$

Una muestra de 100 observaciones reveló que $p = 0.75$. ¿Puede rechazar la hipótesis nula en el nivel de significancia de 0.05?
a) Formule la regla de decisión.
b) Calcule el valor del estadístico de prueba.
c) ¿Cuál es su decisión respecto de la hipótesis nula?

22. Sean las siguientes hipótesis:

$$H_0: \pi = .40$$
$$H_1: \pi \neq .40$$

Una muestra de 120 observaciones reveló que $p = 0.30$. ¿Puede rechazar la hipótesis nula en el nivel de significancia de 0.05?
a) Formule la regla de decisión.
b) Calcule el valor del estadístico de prueba.
c) ¿Cuál es su decisión respecto de la hipótesis nula?

Nota: se recomienda utilizar el procedimiento de los cinco pasos de la prueba de hipótesis y resolver los siguientes problemas.

23. El National Safety Council informó que 52% de los conductores estadounidenses que viajan por autopista de cuota es de género masculino. Una muestra de 300 automóviles que viajaron el día de ayer por la autopista de Nueva Jersey reveló que a 170 los manejaban hombres. Con un nivel de significancia de 0.01, ¿puede concluir que por la autopista de cuota de Nueva Jersey manejaba una proporción mayor de hombres que lo indicado por las estadísticas nacionales?

24. Un artículo reciente de *USA Today* informó que sólo hay un trabajo disponible por cada tres nuevos graduados de universidad. Las principales razones fueron una sobrepoblación de graduados universitarios y una economía débil. Una encuesta de 200 recién graduados reveló que 80 estudiantes tenían trabajo. Con un nivel de significancia de 0.02, ¿puede concluir que una proporción mayor de estudiantes de su escuela tienen empleo?

25. Chicken Delight afirma que 90% de sus pedidos se entrega en 10 minutos desde que se hace el pedido. Una muestra de 100 pedidos mostró que 82 se entregaron en el tiempo prometido. Con un nivel de significancia de 0.10, ¿puede concluir que menos de 90% de los pedidos se entregó en menos de 10 minutos?

26. Una investigación de la Universidad de Toledo indica que 50% de los estudiantes cambia de área de estudios después del primer año. Una muestra aleatoria de 100 estudiantes de la Facultad de Administración reveló que 48 habían cambiado de área de estudio después del primer año del programa de estudios. ¿Hubo una reducción significativa en la proporción de estudiantes que cambian de área el primer año en este programa? Realice una prueba con un nivel de significancia de 0.05.

10.10 Error tipo II

OA9 Calcular la probabilidad de un error tipo II.

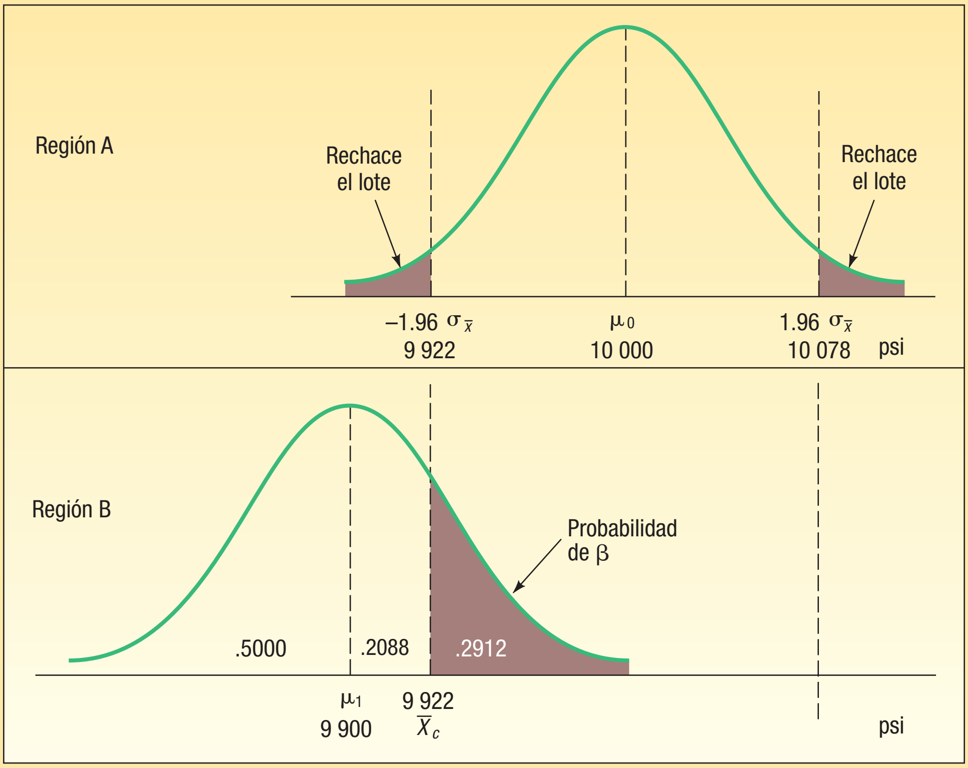
Recuerde que el nivel de significancia, identificado con el símbolo α, es la probabilidad de que la hipótesis nula se rechace cuando es verdadera. Esto recibe el nombre de *error tipo I.* Los niveles de significancia más comunes son 0.05 y 0.01, y los establece el investigador desde el inicio de la prueba.

En un caso de prueba de hipótesis también existe la posibilidad de que no se rechace una hipótesis nula cuando en realidad es falsa. Es decir, se acepta una hipótesis nula falsa. Esto recibe el nombre de *error tipo II.* La probabilidad de un error tipo II se identifica con la letra griega beta (β). Los siguientes ejemplos ilustran los detalles de la determinación del valor de β.

Ejemplo

Western Wire Products compra barras de acero para hacer clavijas. La experiencia indica que la fuerza media de tensión de las cargas que llegan es de 10 000 psi, y que la desviación estándar, σ, es de 400 psi.

Con el fin de tomar una decisión sobre las cargas de barras de acero que llegan, el fabricante establece la siguiente regla para que el inspector de control de calidad se apegue a ella: "Tome una muestra de 100 barras de acero. Si la fuerza media ($\overline{X}$) se encuentra entre 9 922 y 10 078 psi con un nivel de significancia de 0.05, acepte el lote. De lo contrario, debe rechazarlo." La gráfica 10-9, región A, muestra la región en que se rechaza cada lote y en la que no se rechaza. La media de esta distribución se representa mediante μ_0. Las colas de la curva representan la probabilidad de cometer un error tipo I, es decir, de rechazar el lote de barras de acero que ingresa cuando, en realidad, se trata de un buen lote, con una media de 10 000 psi.



GRÁFICA 10-9 Gráficas que muestran los errores tipo I y tipo II

Suponga que la media poblacional desconocida de un lote que llega, designada μ, es en realidad de 9 900 psi. ¿Cuál es la probabilidad de que el inspector de control de calidad no rechace la carga (error tipo II)?

Solución

La probabilidad de cometer un error tipo II, según representa el área sombreada en la gráfica 10-9, región B, se calcula al determinar el área bajo la curva normal que se localiza sobre 9 922 libras. El cálculo de las áreas bajo la curva normal se analizó en el capítulo 7. Un breve repaso: es necesario determinar primero la probabilidad de que la media muestral caiga entre 9 900 y

9 922. Después, se resta esta probabilidad de 0.5000 (que representa toda el área más allá de la media de 9 900) para llegar a la probabilidad de cometer un error tipo II en este caso.

El número de unidades estándares (valor de z) entre la media del lote que llega (9 900), designada μ_1, y $\overline{X}_c$, que representa el valor crítico de 9 922, se calcula de la siguiente manera:

ERROR TIPO II

$$z = \frac{\overline{X}_c - \mu_1}{\sigma/\sqrt{n}} \tag{10-4}$$

Si $n = 100$ y $\sigma = 400$, el valor de z es 0.55:

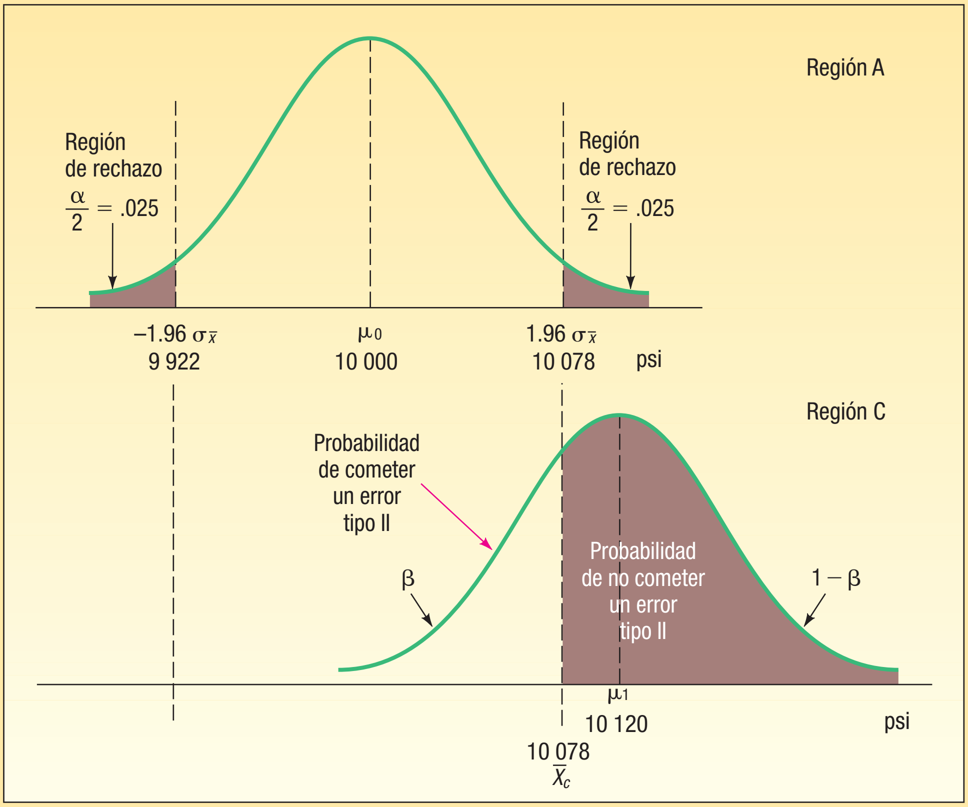
$$z = \frac{\overline{X}_c - \mu_1}{\sigma/\sqrt{n}} = \frac{9\,922 - 9\,900}{400/\sqrt{100}} = \frac{22}{40} = 0.55$$

El área bajo la curva entre 9 900 y 9 922 (un valor z de 0.55) es 0.2088. El área bajo la curva más allá de 9 922 libras es 0.5000 − 0.2088 o 0.2912; tal es la probabilidad de cometer un error tipo II, es decir, de aceptar el ingreso de un lote de barras de acero cuando la media poblacional es de 9 900 psi.

Otra ilustración, en la gráfica 10-10, región C, describe la probabilidad de aceptar un lote cuando la media poblacional es de 10 120. Para determinar la probabilidad:

$$z = \frac{\overline{X}_c - \mu_1}{\sigma/\sqrt{n}} = \frac{10\,078 - 10\,120}{400/\sqrt{100}} = -1.05$$

La probabilidad de que z sea menor que −1.05 es 0.1469, que se determina al calcular 0.5000 − 0.3531. Por lo tanto, β, o la probabilidad de cometer un error tipo II, es 0.1469.



GRÁFICA 10-10 Errores tipo I y tipo II (otro ejemplo)

De acuerdo con las técnicas que se ilustran en las gráficas 10-9, región B, y 10-10, región C, puede determinarse la probabilidad de aceptar una hipótesis como verdadera cuando en realidad es falsa para cualquier valor de μ_1.

Las probabilidades de cometer un error tipo II aparecen en la columna central de la tabla 10-4 para valores selectos de μ, dados en la columna de la izquierda. La columna derecha proporciona la probabilidad de no cometer un error tipo II, que también se conoce como la fuerza de una prueba.

TABLA 10-4 Probabilidades de cometer un error tipo II con μ_0 = 10 000 libras y medias alternativas seleccionadas, nivel de significancia 0.05

Media alternativa seleccionada (libras)	Probabilidad de cometer un error tipo II (β)	Probabilidad de no cometer un error tipo II ($1 - \beta$)
9 820	.0054	.9946
9 880	.1469	.8531
9 900	.2912	.7088
9 940	.6736	.3264
9 980	.9265	.0735
10 000	—*	—
10 020	.9265	.0735
10 060	.6736	.3264
10 100	.2912	.7088
10 120	.1469	.8531
10 180	.0054	.9946

* No es posible cometer un error tipo II cuando $\mu = \mu_0$.

Autoevaluación 10-6

Repase el ejemplo anterior. Suponga que la media real de un lote de barras de acero que llega es de 10 180 psi. ¿Cuál es la probabilidad de que el inspector de control de calidad acepte las barras como si tuvieran una media de 10 000 psi? (Parece poco probable que las barras de acero se rechacen si la fuerza de tensión es mayor que la especificada. No obstante, puede ser que la clavija tenga una doble función en un motor fuera de borda. Tal vez esté diseñada para que no se desprenda si el motor golpea un objeto pequeño, aunque sí lo haga si golpea una roca. Por consiguiente, el acero no debe ser demasiado fuerte.)

El área no sombreada de la gráfica 10-10, región C, representa la probabilidad de aceptar por error la hipótesis que indica que la fuerza de tensión media de las barras de acero es de 10 000 psi. ¿Cuál es la probabilidad de cometer un error tipo II?

Ejercicios

connect

27. Consulte la tabla 10-4 y el ejemplo anterior. Si $n = 100$, $\sigma = 400$, $\bar{X}_c = 9\,922$ y $\mu_1 = 9\,880$, verifique que la probabilidad de cometer un error tipo II sea de 0.1469.

28. Consulte la tabla 10-4 y el ejemplo anterior. Si $n = 100$, $\sigma = 400$, $\bar{X}_c = 9\,922$ y $\mu_1 = 9\,940$, verifique que la probabilidad de cometer un error tipo II sea de 0.6736.

Resumen del capítulo

I. El objetivo de la prueba de hipótesis consiste en verificar la validez de una afirmación relacionada con un parámetro de la población.

II. Los pasos para llevar a cabo una prueba de hipótesis son los siguientes:

- **A.** Se formula la hipótesis nula (H_0) y la hipótesis alternativa (H_1).
- **B.** Se selecciona el nivel de significancia.
  1. El nivel de significancia es la probabilidad de rechazar una hipótesis nula verdadera.
  2. Los niveles de significancia más frecuentes son 0.01, 0.05 y 0.10, pero es posible cualquier valor entre 0 y 1.00.

C. Se selecciona el estadístico de prueba.
   1. Un estadístico de prueba es un valor que se calcula a partir de la información de una muestra para determinar si se rechaza la hipótesis nula.
   2. En este capítulo se consideraron dos estadísticos de prueba.
      ***a*)** La distribución normal estándar se utiliza cuando la población sigue la distribución normal y se conoce la desviación estándar de la población.
      ***b*)** La distribución t de Student se emplea cuando la población sigue la distribución normal y se desconoce la desviación estándar de la población.

D. Se establece la regla de decisión.
   1. La regla de decisión indica la condición o condiciones en que se rechaza la hipótesis nula.
   2. En una prueba de dos colas, la región de rechazo se divide uniformemente entre las colas izquierda y derecha de la distribución.
   3. En una prueba de una cola, toda la región de rechazo se encuentra en la cola izquierda o en la cola derecha.

E. Se selecciona una muestra, se calcula el valor del estadístico de la prueba, se toma una decisión respecto de la hipótesis nula y se interpretan los resultados.

III. Un valor p es la probabilidad de que el valor del estadístico de prueba sea tan extremo como el valor calculado cuando la hipótesis nula es verdadera.

IV. Al probar una hipótesis sobre la media de la población:
   A. Si se conoce la desviación estándar de la población, σ, el estadístico de prueba es la distribución normal estándar, y se determina a partir de:

$$z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \quad \textbf{(10-1)}$$

   B. Si no se conoce la desviación estándar de la población, pero hay por lo menos 30 observaciones en la muestra, s se sustituye por σ. El estadístico de prueba es la distribución t, y su valor se determina de acuerdo con:

$$t = \frac{\overline{X} - \mu}{s/\sqrt{n}} \quad \textbf{(10-2)}$$

   Las principales características de la distribución t de Student son:
   1. Es una distribución continua.
   2. Tiene forma de campana y es simétrica.
   3. Es plana o más amplia que la distribución normal estándar.
   4. Existe una familia de distribuciones t, según el número de grados de libertad.

V. Cuando se prueba la proporción de una población:
   A. Deben cumplirse las condiciones binomiales.
   B. Tanto $n\pi$ como $n(1 - \pi)$ deben ser al menos 5.
   C. El estadístico de prueba es

$$z = \frac{p - \pi}{\sqrt{\dfrac{\pi(1 - \pi)}{n}}} \quad \textbf{(10-3)}$$

VI. Existen dos tipos de errores que se pueden presentar en una prueba de hipótesis.
   A. Un error tipo I, cuando se rechaza una hipótesis nula.
      1. La probabilidad de cometer un error tipo I es igual al nivel de significancia.
      2. Esta probabilidad se designa con la letra griega α.
   B. Un error tipo II, cuando no se rechaza una hipótesis nula falsa.
      1. La probabilidad de cometer un error tipo II se designa con la letra griega β.
      2. La probabilidad de cometer un error tipo II se determina por medio de

$$z = \frac{\overline{X}_c - \mu_1}{\sigma/\sqrt{n}} \quad \textbf{(10-4)}$$

Clave de pronunciación

SÍMBOLO	SIGNIFICADO	PRONUNCIACIÓN
H_0	Hipótesis nula	*H, subíndice cero*
H_1	Hipótesis alternativa	*H, subíndice uno*
$\alpha/2$	Nivel de significancia de dos colas	*Alfa sobre 2*
$\overline{X}_c$	Límite de la media muestral	*X barra, subíndice c*
μ_0	Media supuesta de la población	*Mu, subíndice cero*

11

Pruebas de hipótesis de dos muestras

La familia Damon posee un viñedo grande en el oeste de Nueva York a orillas de lago Erie. Las vides deben fumigarse al inicio de la temporada de cultivo para protegerlas contra diversos insectos y enfermedades. Dos nuevos insecticidas acaban de salir al mercado: Pernod 5 y Action. Para probar su eficacia, se seleccionaron tres hileras y se fumigaron con Pernod 5, y otras tres se fumigaron con Action. Cuando las uvas maduraron, se revisaron 400 vides tratadas con Pernod 5 para saber si no estaban infectadas. De igual forma, se revisó una muestra de 400 plantas fumigadas con Action. Con un nivel de significancia de 0.05, ¿se puede concluir que existe una diferencia entre la proporción de vides infectadas empleando Pernod 5 en comparación con las fumigadas con Action? (Vea el ejercicio 9, objetivo 2.)

Objetivos de aprendizaje

Al concluir el capítulo, será capaz de:

OA1 Realizar la prueba de hipótesis de que dos medias poblacionales independientes con desviaciones estándar conocidas son iguales.

OA2 Efectuar la prueba de la hipótesis de que dos proporciones de poblaciones son iguales.

OA3 Ejecutar la prueba de hipótesis de que dos medias poblacionales independientes son iguales, bajo el supuesto de desviaciones estándares poblacionales iguales pero desconocidas.

OA4 Ejecutar una prueba de hipótesis de que dos medias poblacionales independientes son iguales, bajo el supuesto de desviaciones estándares poblacionales desiguales pero desconocidas.

OA5 Comprender la diferencia entre muestras dependientes e independientes.

OA6 Efectuar una prueba de hipótesis acerca de la diferencia media entre observaciones apareadas y dependientes.

11.1 Introducción

En el capítulo 10 se inició el estudio de las pruebas de hipótesis. Se describió su naturaleza y se realizaron algunas pruebas de hipótesis en las cuales se compararon los resultados de una sola muestra con un valor poblacional. Es decir, se seleccionó una sola muestra aleatoria de una población y se realizó una prueba para ver si era razonable el valor propuesto de la población. Recuerde que en el capítulo 10 se seleccionó una muestra del número de escritorios ensamblados por semana en la Jamestown Steel Company para determinar si había un cambio en la tasa de producción. De modo similar, se muestrearon votantes en un área de un estado para determinar si la proporción de la población que apoyaría al gobernador para su reelección era menor que 0.80. En ambos casos, se compararon los resultados estadísticos de una *sola* muestra con un parámetro de la población.

En este capítulo se amplía la idea de pruebas de hipótesis para dos muestras. Se seleccionan muestras aleatorias de dos poblaciones distintas para determinar si son iguales las medias o las proporciones de la población. Algunas interrogantes por probar son:

1. ¿Hay alguna diferencia entre el valor medio de los bienes raíces residenciales que vendieron los agentes hombres y los que negociaron las mujeres en el sur de Florida?
2. ¿Hay alguna diferencia entre los números medios de defectos producidos en los turnos matutino y vespertino en Kimble Products?
3. ¿Hay alguna diferencia entre el número de días de ausentismo de los trabajadores jóvenes (menores de 21 años de edad) y los trabajadores mayores (mayores de 60 años) en la industria de comida rápida?
4. ¿Hay alguna diferencia entre la proporción de estudiantes de maestría de la Ohio State University y la University of Cincinnati que aprobaron el examen de certificación de contador público en el primer intento?
5. ¿Hay un aumento de la tasa de producción si se toca música en el área de producción?

Este capítulo se inicia con el caso en el que se seleccionan muestras aleatorias de dos poblaciones independientes y se desea investigar si tienen la misma media.

11.2 Pruebas de hipótesis de dos muestras: muestras independientes

Un especialista en planeación urbana de Florida desea saber si hay alguna diferencia entre los salarios medios por hora de los plomeros y los electricistas en el centro de ese estado. Un contador financiero quiere saber si la tasa de recuperación media de los fondos mutualistas de alto rendimiento es distinta que la tasa de recuperación media de los fondos mutualistas globales. En cada uno de estos casos hay dos poblaciones independientes. En el primero, los plomeros representan una población, y los electricistas, otra. En el segundo caso, los fondos mutualistas de alto rendimiento son una población, y los fondos mutualistas globales, otra.

En cada uno de los casos, para despejar la duda, se debería seleccionar una muestra aleatoria de cada población y calcular la media de las dos muestras. Si las dos medias poblacionales son iguales, es decir, si el salario medio por hora de los plomeros y los electricistas es igual, se esperaría que la *diferencia* entre las dos medias poblacionales fuese de cero. Pero, ¿qué pasaría si los resultados produjeran una diferencia distinta de cero? ¿La diferencia se debe a la casualidad o a que existe una diferencia real entre los salarios por hora? Una prueba de las medias de dos muestras ayudará a responder la pregunta.

Estadística en acción

La elección presidencial de Estados Unidos en 2000 fue una de las más cerradas de la historia. Los medios de información fueron incapaces de hacer una proyección del ganador. La decisión final, con recuentos y decisiones judiciales, tardó más de cinco semanas. Ésta no fue la única elección en la cual hubo controversia. Poco antes de la elección presidencial de 1936, el *New York Times* publicó el encabezado: "La encuesta de *Digest* da a Landon 32 estados: Landon va ganando 4-3." Sin embargo, Alfred Landon, de Kansas, no resultó electo presidente. En realidad, Roosevelt ganó por más de 11 millones de votos y recibió 523 votos en el Colegio Electoral. ¿Por qué el encabezado estuvo tan errado?

El *Literary Digest* recopiló una muestra de votantes entre las listas de números telefónicos, registros automovilísticos y sus lectores. En 1936 no muchas personas tenían teléfono o automóvil. Además, quienes leían el *Digest* solían ser más ricos y votaban por los republicanos. Por todo ello, la población de la muestra no representaba a la población de votantes. Un segundo problema fue la falta de respuestas. Se enviaron encuestas a más de 10

(*continúa*)

millones de personas y cerca de 2.3 millones las respondieron. Sin embargo, no se tomó en cuenta si las personas que respondieron formaban una muestra representativa de los votantes.

Con las computadoras y los métodos modernos de encuestas, las muestras se seleccionan y verifican con cuidado para tener la seguridad de que sean representativas. ¿Qué sucedió con *Literary Digest*? Cerró el negocio poco después de la elección de 1936.

Es necesario regresar a los resultados del capítulo 8. Recuerde que se demostró que una distribución de las medias suele aproximarse a la distribución normal. Es necesario, una vez más, suponer que una distribución de las medias de muestras seguirá una distribución normal. Es posible demostrar en forma matemática que la distribución de las diferencias entre medias muestrales de dos distribuciones normales también es normal.

Esta teoría se ejemplifica en términos del especialista en planeación urbana de Tampa, Florida. Para iniciar, dé por cierta información que por lo general no está disponible. Suponga que la población de plomeros tiene un salario medio de \$30.00 por hora y una desviación estándar de \$5.00 por hora. La población de electricistas tiene un salario medio de \$29.00 y una desviación estándar de \$4.50. Ahora, a partir de esta información, es claro que las dos medias poblacionales no son iguales. Los plomeros ganan \$1.00 por hora más que los electricistas. Pero no se puede esperar que se descubra esta diferencia cada vez que tomen muestras de las dos poblaciones.

Suponga que selecciona una muestra aleatoria de 40 plomeros y otra de 35 electricistas, y que calcula la media de cada muestra. Después determina la diferencia entre las medias muestrales. Esta diferencia entre las medias muestrales es la que llama la atención. Si las poblaciones tienen la misma media, es de esperar que la diferencia entre las dos medias muestrales sea cero. Si hay alguna diferencia entre las medias poblacionales, debería existir una diferencia entre las medias muestrales.

Para comprender la teoría, necesita tomar varios pares de muestras, calcular la media de cada una, determinar la diferencia entre las medias muestrales y estudiar la distribución de las diferencias entre las medias muestrales. Del estudio de la distribución de las diferencias entre las medias muestrales del capítulo 8, sabe que la distribución de ellas sigue la distribución normal. Si las dos distribuciones de las medias muestrales siguen la distribución normal, la distribución de sus diferencias también debe seguir la distribución normal. Éste es el primer obstáculo.

El segundo se refiere a la media de esta distribución de las diferencias. Si determina que la media de esta distribución es cero, esto implica que no hay diferencia entre las dos poblaciones. Por otro lado, si la media de la distribución de las diferencias es igual a algún valor distinto de cero, ya sea positivo o negativo, concluirá que las dos poblaciones no tienen la misma media.

OA1 Realizar la prueba de hipótesis de que dos medias poblacionales independientes con desviaciones estándar conocidas son iguales.

Para reportar algunos resultados concretos, recuerde al especialista en planeación urbana de Tampa, Florida. En la tabla 11-1 aparece el resultado de la selección de 20 muestras diferentes de 40 plomeros y 35 electricistas, luego de calcular la media de cada muestra y determinar la diferencia entre dos medias muestrales. En el primer caso, la muestra de 40 plomeros tiene una media de \$29.80, y la de los electricistas es de \$28.76. La diferencia entre las medias muestrales es de \$1.04. Este proceso se repitió 19 veces más. Observe que en 17 de los 20 casos la media de los plomeros es mayor que la de los electricistas.

El obstáculo final es que se necesita saber algo acerca de la *variabilidad* de la distribución de las diferencias. En otras palabras, ¿cuál es la desviación estándar de esta distribución de las diferencias? En la teoría estadística se demuestra que cuando se tienen poblaciones independientes, como en este caso, la distribución de las diferencias tiene una varianza (desviación estándar elevada al cuadrado) igual a la suma de dos varianzas individuales. Esto significa que se pueden sumar las varianzas de dos distribuciones muestrales. En otras palabras, la varianza de la diferencia entre medias muestrales ($\bar{X}_1 - \bar{X}_2$) es igual a la suma de la varianza de los plomeros y de la varianza de los electricistas.

VARIANZA DE LA DISTRIBUCIÓN DE LAS DIFERENCIAS ENTRE MEDIAS

$$\sigma^2_{\bar{X}_1-\bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \qquad \textbf{(11-1)}$$

TABLA 11-1 Medias de muestras aleatorias de plomeros y electricistas

Muestra	Plomeros	Electricistas	Diferencia
1	$29.80	$28.76	$1.04
2	30.32	29.40	0.92
3	30.57	29.94	0.63
4	30.04	28.93	1.11
5	30.09	29.78	0.31
6	30.02	28.66	1.36
7	29.60	29.13	0.47
8	29.63	29.42	0.21
9	30.17	29.29	0.88
10	30.81	29.75	1.06
11	30.09	28.05	2.04
12	29.35	29.07	0.28
13	29.42	28.79	0.63
14	29.78	29.54	0.24
15	29.60	29.60	0.00
16	30.60	30.19	0.41
17	30.79	28.65	2.14
18	29.14	29.95	−0.81
19	29.91	28.75	1.16
20	28.74	29.21	−0.47

El término $\sigma^2_{\bar{X}_1-\bar{X}_2}$ parece complejo, pero no es difícil interpretarlo. La parte σ^2 indica que es una varianza, y el subíndice, $\bar{X}_1 - \bar{X}_2$, que es una distribución de las diferencias de las medias muestrales.

Es posible representar esta ecuación en forma más práctica con la raíz cuadrada, de modo que se obtenga la desviación estándar de la distribución o "error estándar" de las diferencias. Por último, se estandariza la distribución de las diferencias. El resultado es la ecuación siguiente.

PRUEBA DE DOS MEDIAS DE MUESTRAS σ CONOCIDA

$$z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \qquad \textbf{(11-2)}$$

Antes de presentar un ejemplo, repase las suposiciones necesarias para emplear la fórmula (11-2).

- Las dos poblaciones siguen distribuciones normales.
- Las dos muestras no deben estar relacionadas, es decir, deben ser independientes.
- Debe conocerse la desviación estándar de las dos poblaciones.

En el ejemplo siguiente se muestran los detalles de la prueba de hipótesis de dos medias poblacionales.

Ejemplo

Los clientes de los supermercados FoodTown tienen una opción al pagar por sus compras. Pueden pagar en una caja registradora normal operada por un cajero, o emplear el nuevo procedimiento: Fast Lane. Cuando eligen la primera alternativa, un empleado registra cada artículo, lo pone en una banda transportadora pequeña de donde otro empleado lo toma y lo pone en una bolsa, y después en el carrito de víveres. En el procedimiento Fast Lane, el cliente registra cada artículo, lo pone en una bolsa y coloca las bolsas en el carrito. Este procedimiento

está diseñado para reducir el tiempo que los clientes pierden en la fila de la caja.

El aparato de Fast Lane se acaba de instalar en la sucursal de la calle Byrne de FoodTown. La gerente de la tienda desea saber si el tiempo medio de pago con el método tradicional es mayor que con Fast Lane, para lo cual reunió la información siguiente sobre la muestra. El tiempo se mide desde el momento en que el cliente ingresa a la fila hasta que sus bolsas están en el carrito. De aquí que el tiempo incluye tanto la espera en la fila como el registro. ¿Cuál es el valor *p*?

Tipo de cliente	Media muestral	Desviación estándar de la población	Tamaño de la muestra
Tradicional	5.50 minutos	0.40 minutos	50
Fast Lane	5.30 minutos	0.30 minutos	100

Solución

Para responder la pregunta anterior emplee el procedimiento de prueba de hipótesis de cinco pasos.

Paso 1: Formule las hipótesis nula y alternativa. La hipótesis nula es que no hay diferencia entre los tiempos medios de pago de los dos grupos. En otras palabras, la diferencia de 0.20 minutos entre el tiempo medio de pago con el método tradicional y el tiempo medio de pago con Fast Lane se debe a la casualidad. La hipótesis alternativa es que el tiempo medio de quienes utilizan el método tradicional es mayor. Si μ_s se refiere al tiempo medio de pago de la población de clientes tradicionales y μ_f al tiempo medio de pago de los clientes que emplean Fast Lane, las hipótesis nula y alternativa son:

$$H_0: \mu_s \leq \mu_f$$
$$H_1: \mu_s > \mu_f$$

Paso 2: Seleccione el nivel de significancia. Éste es la probabilidad de que rechace la hipótesis nula cuando en realidad es verdadera. Esta posibilidad se determina antes de seleccionar la muestra o de realizar algún cálculo. Los niveles de significancia 0.05 y 0.01 son los más comunes, pero también se emplean otros valores, como 0.02 y 0.10. En teoría, se puede seleccionar cualquier valor entre 0 y 1 para el nivel de significancia. En este caso se seleccionó el nivel de significancia 0.01.

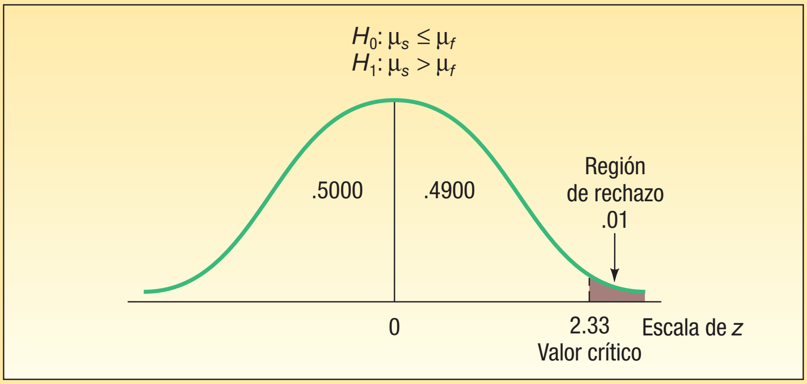
Paso 3: Determine el estadístico de prueba. En el capítulo 10 empleó la distribución normal estándar (es decir, z) y t como estadísticos de prueba. En este caso se usa la distribución z como el estadístico de prueba debido a que las desviaciones estándares de las dos poblaciones se conocen.

Paso 4: Formule una regla de decisión. Esta regla se basa en las hipótesis nula y alternativa (es decir, prueba de una o dos colas), en el nivel de significancia y en el estadístico de prueba empleado. Seleccionó el nivel de significancia 0.01 y la distribución z como el estadístico de prueba, y desea determinar si el tiempo medio de pago es mayor con el método tradicional. Se formula la hipótesis alternativa que indica que el tiempo medio de pago de quienes emplean el método tradicional es mayor. De aquí, la región de rechazo se encuentra en la cola superior de la distribución normal (una prueba de una cola). Para determinar el valor crítico, coloque 0.01 del área total en la cola superior. Esto significa que 0.4900 (0.5000 − 0.0100) del área se ubica entre el valor z de 0 y el valor crítico. Después, busque en el cuerpo del apéndice B.1 un valor ubicado cerca de 0.4900, que es 2.33.

Estadística en acción

¿Vive para trabajar o trabaja para vivir? Una encuesta reciente entre 802 trabajadores estadounidenses reveló que, entre quienes consideran su trabajo como una profesión, el número medio de horas que trabajan por día es de 8.7. Entre los que consideraban su trabajo como un empleo, el número medio de horas trabajadas por día era de 7.6.

Por lo tanto, su regla de decisión es rechazar H_0 si el valor calculado a partir del estadístico de prueba es mayor que 2.33. En la gráfica 11-1 aparece la regla de decisión.



GRÁFICA 11-1 Regla de decisión de una prueba de una cola con un nivel de significancia 0.01

Paso 5: Tome la decisión respecto de H_0 e interprete el resultado. Emplee la fórmula (11-2) para calcular el valor del estadístico de prueba.

$$z = \frac{\bar{X}_s - \bar{X}_f}{\sqrt{\dfrac{\sigma_s^2}{n_s} + \dfrac{\sigma_f^2}{n_f}}} = \frac{5.5 - 5.3}{\sqrt{\dfrac{0.40^2}{50} + \dfrac{0.30^2}{100}}} = \frac{0.2}{0.064} = 3.13$$

El valor calculado, 3.13, es mayor que el valor crítico 2.33; en consecuencia, debe rechazar la hipótesis nula y aceptar la alternativa. La diferencia de 0.20 minutos entre el tiempo medio de pago con el método tradicional es demasiado grande para deberse a la casualidad. En otras palabras, la conclusión es que el método Fast Lane es más rápido.

¿Cuál es el valor p del estadístico de prueba? Recuerde que el valor p es la probabilidad de determinar un valor del estadístico de prueba así de excepcional cuando la hipótesis nula es verdadera. Para calcular el valor p es necesaria la probabilidad de un valor z mayor que 3.13. En el apéndice B.1 no aparece la probabilidad asociada con 3.13. El mayor valor disponible es 3.09. El área que corresponde a 3.09 es 0.4990. En este caso, el valor p es menor que 0.0010, calculado mediante 0.5000 – 0.4900. La conclusión es que hay muy pocas probabilidades de que la hipótesis nula sea verdadera.

En resumen, los criterios para emplear la fórmula (11-2) son:

1. *Las muestras son de poblaciones independientes*. Esto significa, por ejemplo, que el tiempo de pago de los clientes que emplean Fast Lane no está relacionado con el tiempo de pago de los demás clientes. Por ejemplo, el tiempo del señor Smith no afecta ningún otro tiempo de pago de otros clientes.
2. *Ambas poblaciones siguen la distribución normal.* En el ejemplo FoodTown, esto significa que la población de tiempos tanto en la fila estándar como en la de Fast Lane siguen la distribución normal.
3. *Las dos desviaciones estándares de las poblaciones se conocen*. En el ejemplo de FoodTown, la desviación estándar de la población de los tiempos de pago con Fast Lane fue 0.30 minutos. La desviación estándar de los tiempos de pago tradicionales fue 0.40 minutos. Emplee la fórmula (11-2) para determinar el valor del estadístico de prueba.

Autoevaluación 11-1

Tom Sevits, propietario de Appliance Patch, observó una diferencia en el total en dólares de las ventas entre los hombres y las mujeres que emplea como agentes de ventas. Una muestra de 40 días reveló que los hombres venden una media de $1 400 por concepto de venta de aparatos por día. En una muestra de 50 días, las mujeres vendieron una media de $1 500 por concepto de venta de aparatos por día. Suponga que la desviación estándar de los hombres es de $200 y la de las mujeres de $250. Con un nivel de significancia de 0.05, ¿puede el señor Sevits concluir que la cantidad media que venden por día las mujeres es mayor?

a) Formule las hipótesis nula y alternativa.
b) ¿Cuál es la regla de decisión?
c) ¿Cuál es el valor del estadístico de prueba?
d) ¿Cuál es su decisión respecto de la hipótesis nula?
e) ¿Cuál es el valor *p*?
f) Interprete el resultado.

Ejercicios

connect

1. Considere una muestra de 40 observaciones de una población con una desviación estándar de la población de 5. La media muestral es 102. Otra muestra de 50 observaciones de una segunda población tiene una desviación estándar de la población de 6. La media muestral es 99. Realice la prueba de hipótesis siguiente con el nivel de significancia de 0.04.

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

***a*)** ¿Se trata de una prueba de una o de dos colas?
***b*)** Formule la regla de decisión.
***c*)** Calcule el valor del estadístico de prueba.
***d*)** ¿Cuál es su decisión respecto de H_0?
***e*)** ¿Cuál es el valor *p*?

2. Considere una muestra de 65 observaciones de una población con una desviación estándar de la población de 0.75. La media muestral es 2.67. Otra muestra de 50 observaciones de una segunda población tiene una desviación estándar de la población de 0.66. La media muestral es 2.59. Realice la prueba de hipótesis siguiente con el nivel de significancia de 0.08.

$$H_0: \mu_1 \leq \mu_2$$
$$H_1: \mu_1 > \mu_2$$

***a*)** ¿Se trata de una prueba de una o de dos colas?
***b*)** Formule la regla de decisión.
***c*)** Calcule el valor del estadístico de prueba.
***d*)** ¿Cuál es su decisión respecto de H_0?
***e*)** ¿Cuál es el valor *p*?

Nota: Para resolver los ejercicios siguientes utilice el procedimiento de prueba de hipótesis de cinco pasos.

3. La compañía Gibbs Baby desea comparar el aumento de peso de bebés que consumen su producto en comparación con el producto de su competidor. Una muestra de 40 bebés que consumen los productos Gibbs reveló un aumento de peso medio de 7.6 libras en sus primeros tres meses de vida, con una desviación estándar de la población de la muestra de 2.3 libras. Una muestra de 55 bebés que consumen la marca del competidor reveló un aumento medio de 8.1 libras, con una desviación estándar de la población de 2.9 libras. Con un nivel de significancia de 0.05, ¿es posible concluir que los bebés que consumieron la marca Gibbs ganaron menos peso? Calcule el valor *p* e interprételo.

4. Como parte de un estudio de empleados corporativos, el director de recursos humanos de PNC, Inc., desea comparar la distancia que deben cubrir para ir al trabajo los empleados de su oficina del centro de Cincinnati con la distancia que recorren quienes trabajan en el centro de Pittsburgh. Una muestra de 35 empleados de Cincinnati muestra que viajan una media de 370 millas al mes. Por su parte, una muestra de 40 empleados de Pittsburgh indica que viajan una media de 380 millas al mes. La desviación estándar de la población de los empleados de Cincinnati y Pittsburgh es de 30 y 26 millas, respectivamente. Con un nivel de significancia de 0.05, ¿existe alguna diferencia entre el número medio de millas recorrido al mes entre los empleados de Cincinnati y los de Pittsburgh?

5. Se sospecha que la altura de las mujeres es un factor para tener partos difíciles; esto es, una mujer más bajita tiene más probabilidades de necesitar una cesárea. Un investigador médico encontró, en una muestra de 45 mujeres que habían tenido un parto normal, que su estatura media era de 61.4 pulgadas. Una segunda muestra de 39 mujeres que fueron sometidas a cesárea tuvo una estatura media de 60.6 pulgadas. Suponga que la población de estaturas relacionadas con los partos normales tiene una desviación estándar de 1.2 pulgadas. También, que las estaturas de la población de mujeres que tuvieron partos por cesárea tiene una desviación estándar de 1.1 pulgadas. ¿Eran más bajas las que tuvieron parto por cesárea? Utilice un nivel de significancia de 0.05. Encuentre el valor p y explique lo que significa.

6. Mary Jo Fitzpatrick es la vicepresidenta de servicios de enfermería del hospital Luke's Memorial. Hace poco observó que en las ofertas de trabajo para enfermeras sindicalizadas, los sueldos son más altos que para las no sindicalizadas. Decidió investigar y reunió la información siguiente.

Grupo	Salario medio	Desviación estándar de la población	Tamaño de la muestra
Sindicalizadas	\$20.75	\$2.25	40
No sindicalizadas	\$19.80	\$1.90	45

¿Es razonable concluir que las enfermeras sindicalizadas ganan más? Utilice un nivel de significancia de 0.02. ¿Cuál es el valor p?

11.3 Prueba de proporciones de dos muestras

En la sección anterior se consideró una prueba de medias poblacionales. Sin embargo, con frecuencia también se tiene interés en saber si dos proporciones de muestras provienen de poblaciones iguales. A continuación se presentan algunos ejemplos.

- El vicepresidente de recursos humanos desea saber si hay alguna diferencia entre la proporción de empleados asalariados por hora que faltan más de 5 días de trabajo por año en las plantas de Atlanta y Houston.
- General Motors considera un diseño nuevo para el Chevy Malibú. El diseño se muestra a un grupo de compradores potenciales menores de 30 años de edad y a otro grupo de mayores de 60 años. La compañía quiere saber si hay alguna diferencia entre la proporción de los dos grupos a quienes les gusta el diseño nuevo.
- Un asesor de la industria de aerolíneas está investigando el miedo a volar entre los adultos. En específico, la compañía desea saber si hay alguna diferencia entre la proporción de hombres con respecto a mujeres que temen viajar en avión.

OA2 Efectuar la prueba de la hipótesis de que dos proporciones de poblaciones son iguales.

En los casos anteriores, cada elemento o individuo muestreado se clasifica como "éxito" o "fracaso". Es decir, en el ejemplo del Chevy Malibú, cada comprador potencial se clasifica como "le gusta el diseño nuevo" o "no le gusta el diseño nuevo". Después, se compara la proporción del grupo de menores de 30 años de edad con la proporción del grupo de mayores de 60 años que indique el gusto por el diseño nuevo. ¿Las diferencias se deben a la casualidad? En este estudio no se obtiene ninguna medida, sólo se clasifican los individuos u objetos.

Para realizar la prueba, suponga que la muestra es lo bastante grande para que la distribución normal sirva como una buena aproximación a la distribución binomial. El estadístico de prueba sigue la distribución normal estándar. El valor de z se calcula a partir de la fórmula siguiente:

PRUEBA DE PROPORCIONES DE DOS MUESTRAS

$$z = \frac{p_1 - p_2}{\sqrt{\dfrac{p_c(1 - p_c)}{n_1} + \dfrac{p_c(1 - p_c)}{n_2}}} \qquad \textbf{(11-3)}$$

La fórmula (11-3) es la misma que la (11-2) con las proporciones muestrales respectivas en lugar de las medias muestrales, y con $p_c(1 - p_c)$ en lugar de las dos varianzas. Además:

n_1 es el número de observaciones en la primera muestra.
n_2 es el número de observaciones en la segunda muestra.
p_1 es la proporción en la primera muestra que posee la característica.
p_2 es la proporción en la segunda muestra que posee la característica.
p_c es la proporción conjunta que posee la característica en las muestras combinadas. Se denomina estimación conjunta de la proporción poblacional y se calcula a partir de la fórmula siguiente.

PROPORCIÓN CONJUNTA

$$p_c = \frac{X_1 + X_2}{n_1 + n_2} \qquad \textbf{(11-4)}$$

Donde:
X_1 es el número que posee la característica en la primera muestra.
X_2 es el número que posee la característica en la segunda muestra.

En el ejemplo siguiente se ilustra la prueba de proporciones de dos muestras.

Ejemplo

La compañía de perfumes Manelli desarrolló una fragancia nueva que planea comercializar con el nombre de Heavenly. Varios estudios de mercado indican que Heavenly tiene buen potencial de mercado. El departamento de ventas de Manelli tiene interés en saber si hay alguna diferencia entre las proporciones de mujeres jóvenes y mayores que comprarían el perfume si saliera al mercado. Hay dos poblaciones independientes, una de mujeres jóvenes y la otra de mujeres mayores. A cada una de las mujeres muestreadas se le pedirá que huela el perfume e indique si le gusta lo suficiente para comprar un frasco.

Solución

Utilizará el procedimiento usual de prueba de hipótesis de cinco pasos.

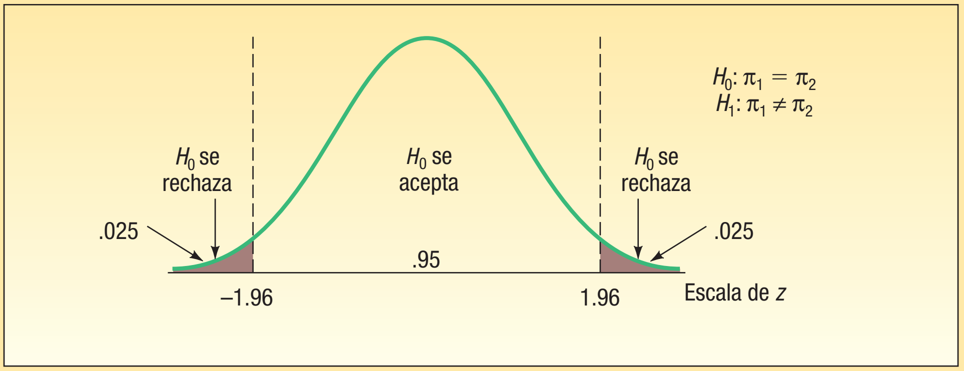
Paso 1: Formule H_0 y H_1. En este caso, la hipótesis nula es: “No hay diferencia en la proporción de mujeres jóvenes y mayores que prefieren Heavenly.” Designe a π_1 como la proporción de mujeres jóvenes que comprarían Heavenly y π_2 como la proporción de mujeres mayores que lo comprarían. La hipótesis alternativa es que las dos proporciones no son iguales.

$$H_0: \pi_1 = \pi_2$$
$$H_1: \pi_1 \neq \pi_2$$

Paso 2: Seleccione el nivel de significancia. En este ejemplo se elige un nivel de significancia de 0.05.

Paso 3: Determine el estadístico de prueba. El estadístico de prueba sigue la distribución normal estándar. El valor del estadístico de prueba se calcula a partir de la fórmula (11-3).

Paso 4: Formule la regla de decisión. Recuerde que la hipótesis alternativa del paso 1 no indica una dirección, de modo que ésta es una prueba de dos colas. Para determinar el valor crítico, divida el nivel de significancia a la mitad y coloque esta cantidad en cada cola de la distribución z. Después, reste esta cantidad al área total a la derecha de cero, es decir, $0.5000 - 0.0250 = 0.4750$. Por último, busque en el cuerpo de la tabla z (apéndice B.1) el valor más cercano, que es 1.96. Los valores críticos son -1.96 y $+1.96$. Como antes, si el valor calculado de z se encuentra en la región entre $+1.96$ y -1.96, no se rechaza la hipótesis nula. En tal caso, se supone que cualquier diferencia entre las proporciones de las dos muestras se debe a la variación casual. Esta información aparece en la gráfica 11-2.



GRÁFICA 11-2 Reglas de decisión de la prueba de la fragancia Heavenly, nivel de significancia 0.05

Paso 5: Seleccione una muestra y tome una decisión. Una muestra aleatoria de 100 mujeres jóvenes reveló que a 19 les gustó la fragancia Heavenly lo suficiente para comprarla. De manera similar, una muestra de 200 mujeres mayores reveló que a 62 les gustó la fragancia lo suficiente para comprarla. Se designa p_1 como el número de mujeres jóvenes y p_2 como el de las mujeres mayores.

$$p_1 = \frac{X_1}{n_1} = \frac{19}{100} = .19 \qquad p_2 = \frac{X_2}{n_2} = \frac{62}{200} = .31$$

La pregunta de investigación es si la diferencia de 0.12 en las dos proporciones de las dos muestras se debe a la casualidad o si hay alguna diferencia entre las proporciones de mujeres jóvenes y mayores a quienes les gusta la fragancia Heavenly.

Después, se combinan o se conjuntan las proporciones de las muestras. Se emplea la fórmula (11-4).

$$p_c = \frac{X_1 + X_2}{n_1 + n_2} = \frac{19 + 62}{100 + 200} = \frac{81}{300} = 0.27$$

Observe que la proporción conjunta se aproxima más a 0.31 que a 0.19 debido a que se muestrearon más mujeres mayores que jóvenes.

Con la fórmula (11-3) se determina el valor del estadístico de prueba.

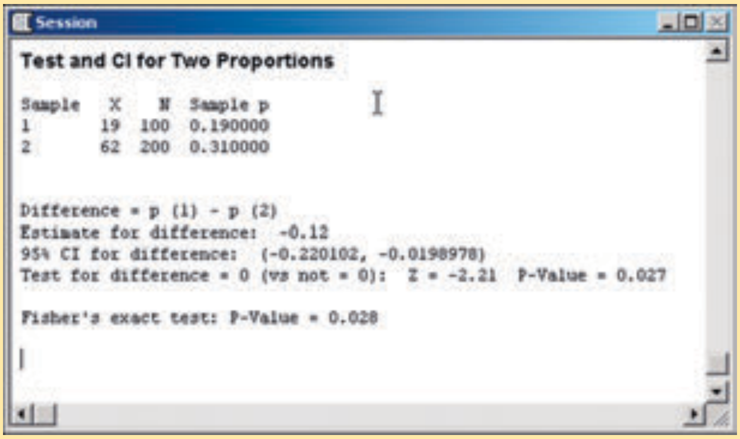
$$z = \frac{p_1 - p_2}{\sqrt{\dfrac{p_c(1 - p_c)}{n_1} + \dfrac{p_c(1 - p_c)}{n_2}}} = \frac{.19 - .31}{\sqrt{\dfrac{.27(1 - .27)}{100} + \dfrac{.27(1 - .27)}{200}}} = -2.21$$

El valor calculado de -2.21 se encuentra en el área de rechazo; es decir, está a la izquierda de -1.96. Por lo tanto, rechace la hipótesis nula en el nivel de significancia 0.05. En otras palabras, se rechaza la hipótesis nula de que la proporción de mujeres jóvenes que comprarían la fragancia es igual a la proporción de mujeres mayores que también la comprarían. Es improbable que la diferencia entre las dos proporciones de las muestras se deba a la casualidad. Para determinar el valor p, consulte el apéndice B.1 y encuentre la probabilidad de un valor z menor que -2.21 o mayor que 2.21. El valor z que corresponde a 2.21 es 0.4864. Por ello, la probabilidad de determinar que el valor del estadístico de prueba sea menor que −2.21 o mayor que 2.21 es:

$$\text{Valor } p = 2(.5000 - .4864) = 2(.0136) = .0272$$

El valor p de 0.0272 es menor que el nivel de significancia 0.05, por lo cual debe rechazar la hipótesis nula. Una vez más, la conclusión es que hay una diferencia entre las proporciones de mujeres jóvenes y mayores que comprarían la fragancia Heavenly.

El sistema Minitab tiene un procedimiento para determinar de forma rápida el valor del estadístico de prueba y calcular el valor p. Los resultados son los siguientes.



```
Session

Test and CI for Two Proportions

Sample   X    N  Sample p
1       19  100  0.190000
2       62  200  0.310000


Difference = p (1) - p (2)
Estimate for difference:  -0.12
95% CI for difference:  (-0.220102, -0.0198978)
Test for difference = 0 (vs not = 0):  Z = -2.21  P-Value = 0.027

Fisher's exact test: P-Value = 0.028
```

Observe que en el resultado de Minitab aparecen dos proporciones de las muestras, el valor de z y el valor p.

Autoevaluación 11-2

De 150 adultos que probaron un nuevo pastel sabor durazno, 87 lo calificaron como excelente. De 200 niños muestreados, 123 lo calificaron como excelente. Con un nivel de significancia de 0.10, ¿puede concluir que existe una diferencia significativa entre la proporción de adultos y la de niños que calificaron al nuevo sabor como excelente?

a) Formule las hipótesis nula y alternativa.
b) ¿Cuál es la probabilidad de un error tipo I?
c) ¿Se trata de una prueba de una o dos colas?
d) ¿Cuál es la regla de decisión?
e) ¿Cuál es el valor del estadístico de prueba?
f) ¿Cuál es su decisión respecto de la hipótesis nula?
g) ¿Cuál es el valor p? Explique qué significa en términos de este problema.

Ejercicios

connect™

7. Las hipótesis nula y alternativa son:

$$H_0\colon \pi_1 \leq \pi_2$$
$$H_1\colon \pi_1 > \pi_2$$

Una muestra de 100 observaciones de la primera población indicó que X_1 es 70. Una muestra de 150 observaciones de la segunda población reveló que X_2 es 90. Utilice un nivel de significancia de 0.05 para probar la hipótesis.

a) Formule la regla de decisión.
b) Calcule la proporción conjunta.
c) Calcule el valor del estadístico de prueba.
d) ¿Cuál es su decisión respecto de la hipótesis nula?

8. Las hipótesis nula y alternativa son:

$$H_0\colon \pi_1 = \pi_2$$
$$H_1\colon \pi_1 \neq \pi_2$$

Una muestra de 200 observaciones de la primera población indicó que X_1 es 170; otra, de 150 observaciones de la segunda población, reveló que X_2 es 110. Utilice el nivel de significancia 0.05 para probar la hipótesis.

a) Formule la regla de decisión.
b) Calcule la proporción conjunta.
c) Estime el valor del estadístico de prueba.
d) ¿Cuál es su decisión respecto de la hipótesis nula?

Nota: Para resolver los ejercicios siguientes utilice el procedimiento de prueba de hipótesis de cinco pasos.

9. La familia Damon posee un viñedo grande en el oeste de Nueva York a orillas de lago Erie. Los viñedos deben fumigarse al inicio de la temporada de cultivo para protegerlos contra diversos insectos y enfermedades. Dos nuevos insecticidas acaban de salir al mercado: Pernod 5 y Action. Para probar su eficacia, se seleccionaron tres hileras y se fumigaron con Pernod 5, y otras tres se fumigaron con Action. Cuando las uvas maduraron, se revisaron 400 vides tratadas con Pernod 5 para saber si no estaban infectadas. De igual forma, se revisó una muestra de 400 vides fumigadas con Action. Los resultados son:

Insecticida	Número de vides revisadas (tamaño de la muestra)	Número de vides infectadas
Pernod 5	400	24
Action	400	40

Con un nivel de significancia de 0.05, ¿se puede concluir que existe una diferencia entre la proporción de vides infectadas empleando Pernod 5 en comparación con las fumigadas con Action?

10. GfK Custom Research North America realizó encuestas idénticas en un intervalo de cinco años. Una pregunta para las mujeres fue: "¿La mayoría de los hombres son amables, gentiles y considerados?" La primera encuesta reveló que, de las 3 000 mujeres encuestadas, 2 010 dijeron que sí. La última encuesta reveló que 1 530 de las 3 000 mujeres a las cuales se les formuló la pregunta pensaban que los hombres eran amables, gentiles y considerados. Con un nivel de significancia de 0.05, ¿se puede concluir que las mujeres consideran que los hombres son menos amables, gentiles y considerados en la última encuesta en comparación con la primera?

11. A una muestra nacional de republicanos y demócratas influyentes se les preguntó, como parte de una encuesta muy amplia, si estaban en favor de relajar las normas ambientales para que se pudiera quemar carbón con alto contenido de azufre en las plantas eléctricas. Los resultados fueron:

	Republicanos	Demócratas
Número en la muestra	1 000	800
Número en favor	200	168

Con un nivel de significancia 0.02, ¿puede concluir que hay una proporción mayor de demócratas en favor de relajar las normas? Determine el valor *p*.

12. El departamento de investigación de la oficina matriz de la New Hampshire Insurance realiza investigaciones continuas sobre las causas de accidentes automovilísticos, las características de los conductores, etc. Una muestra aleatoria de 400 pólizas de personas solteras reveló que 120 habían protagonizado al menos un accidente en el periodo anterior de tres años. De forma similar, una muestra de 600 pólizas de personas casadas reveló que 150 habían estado involucradas en al menos un accidente. Con un nivel de significancia de 0.05, ¿hay una diferencia significativa entre las proporciones de personas solteras y casadas involucradas en un accidente durante un periodo de tres años? Determine el valor *p*.

11.4 Comparación de medias poblacionales con desviaciones estándares desconocidas

En las dos secciones anteriores se describieron las condiciones en que la distribución normal estándar, es decir, *z*, se empleó como el estadístico de prueba. En un caso se trabajó con una variable (cálculo de la media) y en el segundo con un atributo (cálculo de una proporción). En el primer caso se deseaba comparar dos medias muestrales de poblaciones independientes para determinar si provenían de las mismas poblaciones o de poblaciones iguales. En ese caso se supuso que la población seguía la distribución de probabilidad normal y que se conocía la desviación estándar de la población. En muchos casos, de hecho en la mayoría, no se conoce la desviación estándar de la población. Este problema se soluciona, igual que en el caso de una muestra en el capítulo anterior, al sustituir la desviación estándar de la muestra (s) por la desviación estándar de la población (σ). Vea la fórmula (10-2) en la página 348.

Desviaciones estándares poblacionales iguales

En esta sección se describe otro método para comparar las medias muestrales de dos poblaciones independientes y determinar si las poblaciones muestreadas pueden tener, de forma razonable, la misma media. Dicho método *no* requiere que se conozcan las desviaciones estándares de las poblaciones. Esto proporciona más flexibilidad cuando se investiga la diferencia entre las medias de las muestras. Hay dos diferencias importantes entre esta prueba y la descrita antes en este capítulo.

1. Las poblaciones muestreadas tienen desviaciones estándares iguales pero desconocidas. Debido a esta suposición, las desviaciones estándares de las muestras se combinan, o "agrupan".
2. Se utiliza la distribución t como el estadístico de prueba.

OA3 Ejecutar una prueba de hipótesis de que dos medias poblacionales independientes son iguales, bajo el supuesto de desviaciones estándares poblacionales iguales pero desconocidas.

La fórmula para determinar el valor del estadístico de prueba t es similar a la fórmula (11-2), pero es necesario un cálculo adicional. Las dos desviaciones estándares de las muestras se agrupan para formar una sola estimación de la desviación estándar desconocida de la población. En esencia, se calcula una media ponderada de las dos desviaciones estándares de las dos muestras y se emplea este valor como una estimación de la desviación estándar desconocida de la población. Las ponderaciones son los grados de libertad que proporciona cada muestra. ¿Por qué es necesario agrupar las desviaciones estándares de las muestras? Como supuso que las dos poblaciones tienen desviaciones estándares iguales, la mejor estimación posible de ese valor es combinar o agrupar toda la información de las muestras que se tenga acerca del valor de la desviación estándar de la población.

La fórmula siguiente se emplea para agrupar las desviaciones estándares de las muestras. Observe que participan dos factores: el número de observaciones en cada muestra y las propias desviaciones estándares de las muestras.

VARIANZA CONJUNTA

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \tag{11-5}$$

donde:

s_1^2 es la varianza (desviación estándar elevada al cuadrado) de la primera muestra.

s_2^2 es la varianza de la segunda muestra.

El valor de t se calcula a partir de la ecuación siguiente.

PRUEBAS DE MEDIAS DE DOS MUESTRAS σ DESCONOCIDAS

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \tag{11-6}$$

donde :

$\bar{X}_1$ es la media de la primera muestra.

$\bar{X}_2$ es la media de la segunda muestra.

n_1 es el número de observaciones en la primera muestra.

n_2 es el número de observaciones en la segunda muestra.

s_p^2 es la estimación conjunta de la varianza de la población.

El número de grados de libertad de la prueba es el número total de elementos muestreados menos el número total de muestras. Como hay dos muestras, hay $n_1 + n_2 - 2$ grados de libertad.

En resumen, la prueba respeta tres requisitos o suposiciones.

1. Las poblaciones muestreadas siguen la distribución normal.
2. Las poblaciones muestreadas son independientes.
3. Las desviaciones estándares de las dos poblaciones son iguales.

En el ejemplo/solución siguiente se explican los detalles de la prueba.

Ejemplo

Owens Lawn Care, Inc., fabrica y ensambla podadoras de césped que envía a distribuidores instalados en Estados Unidos y Canadá. Se han propuesto dos procedimientos distintos para el montaje del motor al chasis de la podadora. La pregunta es: ¿existe una diferencia entre ellos con respecto al tiempo medio para montar los motores al chasis de las podadoras? El primer procedimiento lo desarrolló Herb Welles, un antiguo empleado de Owens (designado como procedimiento 1), y el otro lo desarrolló William Atkins, vicepresidente de ingeniería de Owens (designado como procedimiento 2). Para evaluar los dos métodos, se decidió realizar un estudio de tiempos y movimientos. Se midió el tiempo de montaje en una muestra de cinco empleados según el método de Welles y seis con el método de Atkins. Los resultados, en minutos, aparecen a continuación. ¿Hay alguna diferencia entre los tiempos medios de montaje? Utilice un nivel de significancia de 0.10.

Welles (minutos)	Atkins (minutos)
2	3
4	7
9	5
3	8
2	4
	3

Solución

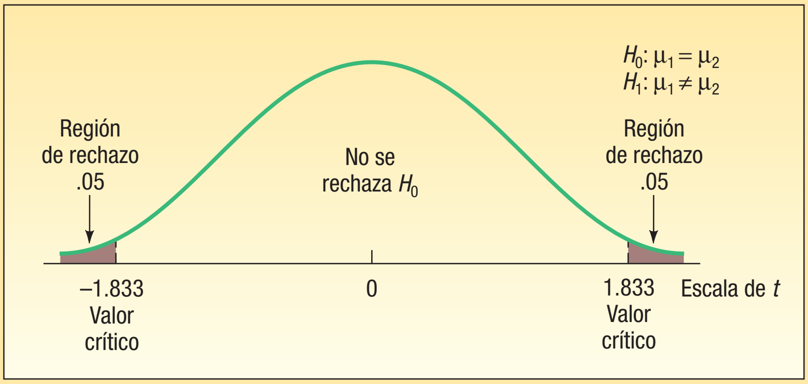
Al seguir el procedimiento de los cinco pasos, la hipótesis nula establece que no hay diferencia entre los tiempos medios de montaje de ambos procedimientos. La hipótesis alternativa indica que sí existe una diferencia.

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

Las suposiciones son:

- Las observaciones incluidas en la muestra de Welles son *independientes* de las observaciones de la muestra de Atkins.
- Las dos poblaciones siguen la distribución normal.
- Las dos poblaciones tienen desviaciones estándares iguales.

¿Hay alguna diferencia entre los tiempos medios de ensamblado con los métodos de Welles y Atkins? Los grados de libertad son iguales al número total de elementos muestreados menos el número de muestras, en este caso, $n_1 + n_2 - 2$. Cinco trabajadores utilizaron el método de Welles y seis el de Atkins. Por lo tanto, hay 9 grados de libertad, calculados así: $5 + 6 - 2$. Los valores críticos de t, del apéndice B.2 de $gl = 9$, una prueba de dos colas y el nivel de significancia de 0.10, son -1.833 y 1.833. La regla de decisión se ilustra en la gráfica 11-3. No se rechaza la hipótesis nula si el valor calculado de t se encuentra entre -1.833 y 1.833.



GRÁFICA 11-3 Regiones de rechazo, prueba de dos colas, $gl = 9$ y nivel de significancia 0.10

Se emplean tres pasos para calcular el valor de t.

Paso 1: Calcule las desviaciones estándar de las muestras. Para calcular la desviación estándar de la muestra usaremos la fórmula (3-11) de la página 84. Vea los detalles a continuación.

Método de Welles		Método de Atkins	
X_1	$(X_1 - \bar{X}_1)^2$	X_2	$(X_2 - \bar{X}_2)^2$
2	$(2-4)^2 = 4$	3	$(3-5)^2 = 4$
4	$(4-4)^2 = 0$	7	$(7-5)^2 = 4$
9	$(9-4)^2 = 25$	5	$(5-5)^2 = 0$
3	$(3-4)^2 = 1$	8	$(8-5)^2 = 9$
2	$(2-4)^2 = 4$	4	$(4-5)^2 = 1$
20	34	3	$(3-5)^2 = 4$
		30	22

$$\bar{X}_1 = \frac{\Sigma X_1}{n_1} = \frac{20}{5} = 4 \qquad \bar{X}_2 = \frac{\Sigma X_2}{n_2} = \frac{30}{6} = 5$$

$$s_1 = \sqrt{\frac{\Sigma(X_1 - \bar{X}_1)^2}{n_1 - 1}} = \sqrt{\frac{34}{5-1}} = 2.9155 \qquad s_2 = \sqrt{\frac{\Sigma(X_2 - \bar{X}_2)^2}{n_2 - 1}} = \sqrt{\frac{22}{6-1}} = 2.0976$$

Paso 2: Agrupe las varianzas de las muestras. Emplee la fórmula (11-5) para agrupar las varianzas de las muestras (desviaciones estándares al cuadrado).

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(5-1)(2.9155)^2 + (6-1)(2.0976)^2}{5 + 6 - 2} = 6.2222$$

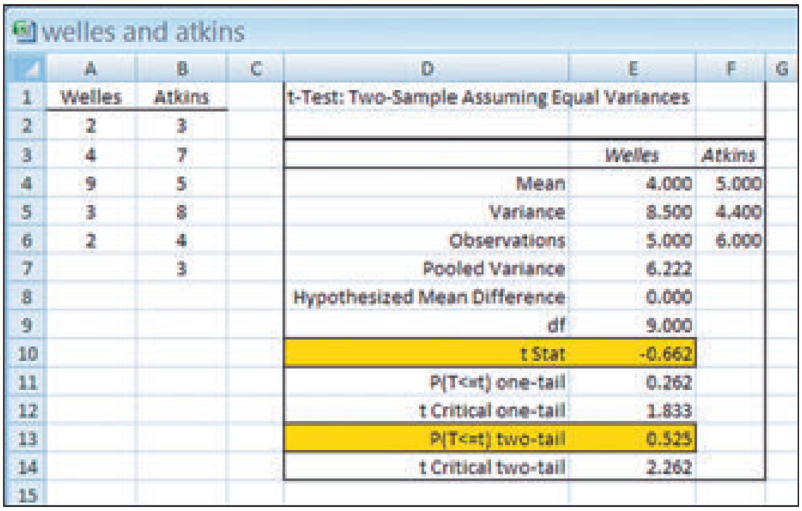
Paso 3: Determine el valor de *t*. El tiempo medio de montaje del método de Welles es de 4.00 minutos, determinado mediante $\bar{X}_1 = 20/5$. El tiempo medio de montaje del método de Atkins es de 5.00 minutos, que se determinó mediante $\bar{X}_2 = 30/6$. Se utiliza la fórmula (11-6) para calcular el valor de t.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{4.00 - 5.00}{\sqrt{6.2222\left(\frac{1}{5} + \frac{1}{6}\right)}} = -0.662$$

La decisión es no rechazar la hipótesis nula, porque -0.662 se encuentra en la región entre -1.833 y 1.833. Se concluye que no existe diferencia entre los tiempos medios necesarios para montar el motor en el chasis con ambos métodos.

Estime también el valor p con el apéndice B.2. Localice la fila con 9 grados de libertad y utilice la columna de prueba de dos colas. Encuentre el valor t, sin considerar el signo, el cual está más cercano al valor calculado de 0.662. Es 1.383, que corresponde a un nivel de significancia de 0.20. Así, aunque se hubiera utilizado el nivel de significancia de 20%, no habría rechazado la hipótesis nula de medias iguales. El valor p es mayor que 0.20.

Excel tiene un procedimiento denominado “Prueba t: dos muestras si las varianzas son iguales” para realizar los cálculos de las fórmulas (11-5) y (11-6), así como para determinar las medias y varianzas de las muestras. Los datos se ingresan en las dos primeras columnas de la hoja de cálculo de Excel y se identifican como “Welles” y “Atkins”. A continuación se presenta la captura de pantalla. El valor de t, denominado “t Stat”, es −0.662, y el valor p de dos colas es 0.525. Como sería de esperar, el valor p es mayor que el nivel de significancia de 0.10. La conclusión es no rechazar la hipótesis nula.


welles and atkins

	A	B	C	D	E	F	G
1	Welles	Atkins		t-Test: Two-Sample Assuming Equal Variances			
2	2	3					
3	4	7			Welles	Atkins	
4	9	5		Mean	4.000	5.000	
5	3	8		Variance	8.500	4.400	
6	2	4		Observations	5.000	6.000	
7		3		Pooled Variance	6.222		
8				Hypothesized Mean Difference	0.000		
9				df	9.000		
10				t Stat	-0.662		
11				P(T<=t) one-tail	0.262		
12				t Critical one-tail	1.833		
13				P(T<=t) two-tail	0.525		
14				t Critical two-tail	2.262		
15							

Autoevaluación 11-3

El gerente de producción de Bellevue Steel, fabricante de sillas de ruedas, desea comparar el número de sillas de ruedas defectuosas producidas en el turno matutino con el del turno vespertino. Una muestra de la producción de 6 turnos matutinos y 8 vespertinos reveló el número de defectos siguiente.

Matutino	5	8	7	6	9	7		
Vespertino	8	10	7	11	9	12	14	9

Con un nivel de significancia de 0.05, ¿hay alguna diferencia entre el número medio de defectos por turno?

a) Formule las hipótesis nula y alternativa.
b) ¿Cuál es la regla de decisión?
c) ¿Cuál es el valor del estadístico de prueba?
d) ¿Cuál es su decisión respecto de la hipótesis nula?
e) ¿Cuál es el valor p?
f) Interprete el resultado.
g) ¿Cuáles son las suposiciones necesarias de esta prueba?

Ejercicios

connect™

En los ejercicios 13 y 14: *a*) formule la regla de decisión, *b*) calcule la estimación conjunta de la varianza de la población, *c*) calcule el estadístico de prueba, *d*) tome una decisión respecto de la hipótesis nula y *e*) calcule el valor p.

13. Las hipótesis nula y alternativa son:

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

Una muestra aleatoria de 10 observaciones de una población reveló una media muestral de 23 y una desviación estándar de 4. Una muestra aleatoria de 8 observaciones de otra población reveló una media muestral de 26 y una desviación estándar de la muestra de 5. Con un nivel de significancia de 0.05, ¿hay alguna diferencia entre las medias poblacionales?

14. Las hipótesis nula y alternativa son:

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

Una muestra aleatoria de 15 observaciones de la primera población reveló una media muestral de 350 y una desviación estándar de la muestra de 12. Una muestra aleatoria de 17 observaciones de la segunda población reveló una media de 342 y una desviación estándar de la muestra de 15. Con un nivel de significancia de 0.10, ¿hay alguna diferencia entre las medias poblacionales?

Nota: En los ejercicios siguientes utilice el procedimiento de prueba de cinco pasos.

15. A continuación se enlistan los salarios en miles de dólares de los 25 jugadores de la jornada inicial del equipo de los Yanquis de Nueva York, de las Ligas Mayores de Béisbol. Estos datos aparecen también en el capítulo 4, ejercicio 22.

Jugador	Salario ($000)	Posición
Aceves, Alfredo	435.7	Pitcher
Burnett, A.J.	16 500.0	Pitcher
Cano, Robinson	9 000.0	Segunda base
Cervelli, Francisco	410.8	Catcher
Chamberlain, Joba	488.0	Pitcher
Gardner, Brett	452.5	Jardinero
Granderson, Curtis	5 500.0	Jardinero
Hughes, Phil	447.0	Pitcher
Jeter, Derek	22 600.0	Receptor de pase corto
Johnson, Nick	5 500.0	Primera base
Marte, Damaso	4 000.0	Pitcher
Mitre, Sergio	850.0	Pitcher
Park, Chan Ho	1 200.0	Pitcher
Pena, Ramiro	412.1	Defensa
Pettitte, Andy	11 750.0	Pitcher
Posada, Jorge	13 100.0	Catcher
Rivera, Mariano	15 000.0	Pitcher
Robertson, David	426.7	Pitcher
Rodriguez, Alex	33 000.0	Tercera base
Sabathia, CC	24 285.7	Pitcher
Swisher, Nick	6 850.0	Jardinero
Teixeira, Mark	20 625.0	Primera base
Thames, Marcus	900.0	Jardinero
Vazquez, Javier	11 500.0	Pitcher
Winn, Randy	1 100.0	Jardinero

Divida a los jugadores en dos grupos: pitchers y no pitchers (jugadores de posición). Asuma que existen varianzas poblacionales iguales para ambos. Pruebe la hipótesis de que los salarios medios de los pitchers y los jugadores de posición son los mismos comparados con la hipótesis alternativa de que no lo son. Utilice un nivel de significancia de 0.01.

16. En un estudio reciente se comparó el tiempo que pasan juntas las parejas en que sólo trabaja uno de los cónyuges con las parejas en que ambos trabajan. De acuerdo con los registros que llevaron las esposas durante el estudio, la cantidad media de tiempo que pasan juntos viendo televisión las parejas en que sólo trabaja uno de los cónyuges fue 61 minutos por día, con una desviación estándar de 15.5 minutos. Las parejas en que los dos trabajan, el número medio de minutos que ven televisión fue de 48.4 minutos, con una desviación estándar de 18.1 minutos. Con un nivel de significancia de 0.01, ¿se puede concluir que, en promedio, las parejas en que sólo trabaja uno de los cónyuges pasan más tiempo juntos viendo televisión? En el estudio había 15 parejas en que sólo uno trabaja y 12 en que trabajan los dos.

17. Lisa Monnin es la directora de presupuestos de Nexos Media, Inc. Ella quiere comparar los gastos diarios en viáticos del personal de ventas con los gastos del personal de auditoría, para lo cual recopiló la información siguiente sobre las muestras.

Ventas (dólares)	131	135	146	165	136	142	
Auditoría (dólares)	130	102	129	143	149	120	139

Con un nivel de significancia de 0.10, ¿puede Monnin concluir que los gastos diarios medios del personal de ventas son mayores que los del personal de auditoría? ¿Cuál es el valor de *p*?

18. La Area Chamber of Commerce de Tampa Bay (Florida) quería saber si el salario semanal medio de las enfermeras era mayor que el de los maestros de escuela. Para esta investigación recopiló la información siguiente sobre las cantidades que ganó la semana pasada una muestra de maestros y enfermeras.

Maestros de escuela (dólares)	845	826	827	875	784	809	802	820	829	830	842	832
Enfermeras (dólares)	841	890	821	771	850	859	825	829				

¿Es razonable concluir que es mayor el salario semanal medio de las enfermeras? Utilice un nivel de significancia de 0.01. ¿Cuál es el valor *p*?

Medias poblacionales con desviaciones estándares desiguales

En las secciones anteriores fue necesario suponer que las poblaciones tenían desviaciones estándares iguales. En otras palabras, no se conocían las desviaciones estándares de las poblaciones, sino que se suponían iguales. En muchos casos, ésta es una suposición razonable, pero ¿qué sucede si no son iguales? En el capítulo siguiente se presenta un método formal para probar esta suposición de varianzas iguales.

OA4 Ejecutar la prueba de hipótesis de que dos medias poblacionales independientes son iguales, bajo el supuesto de desviaciones estándares poblacionales desiguales pero desconocidas.

Si no es razonable suponer que las desviaciones estándares poblacionales son iguales, se emplea un estadístico muy similar a la fórmula (11-2). Las desviaciones estándares de las muestras, s_1 y s_2, se emplean en lugar de las desviaciones estándares de las poblaciones respectivas. Además, los grados de libertad se ajustan hacia abajo mediante una fórmula de aproximación compleja. El efecto es reducir el número de grados de libertad de la prueba, lo cual requerirá un valor mayor del estadístico de prueba para rechazar la hipótesis nula.

La fórmula del estadístico *t* es:

ESTADÍSTICO DE PRUEBA DE MEDIAS SIN DIFERENCIA, VARIANZAS DESIGUALES

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} \qquad \textbf{(11-7)}$$

Los grados de libertad estadística se determinan mediante:

GRADOS DE LIBERTAD PARA PRUEBA CON VARIANZA DESIGUAL

$$gl = \frac{[(s_1^2/n_1) + (s_2^2/n_2)]^2}{\dfrac{(s_1^2/n_1)^2}{n_1 - 1} + \dfrac{(s_2^2/n_2)^2}{n_2 - 1}} \qquad \textbf{(11-8)}$$

donde n_1 y n_2 son los tamaños muestrales respectivos, y s_1 y s_2, las desviaciones estándares de las muestras respectivas. Si es necesario, esta fracción se redondea hacia abajo a un valor entero. En el ejemplo siguiente se ilustran los detalles.

Ejemplo

El personal en un laboratorio de pruebas del consumidor evalúa la absorción de toallas de papel. Se desea comparar un conjunto de toallas de una marca con un grupo similar de toallas de otra marca. De cada una de ellas se sumerge una pieza del papel en un tubo con un fluido, se deja que el papel escurra en una charola durante dos minutos y después se evalúa la cantidad de líquido que el papel absorbió de la charola. Una muestra aleatoria de 9 toallas de papel de la primera marca absorbió las cantidades siguientes de líquido en milímetros.

8	8	3	1	9	7	5	5	12

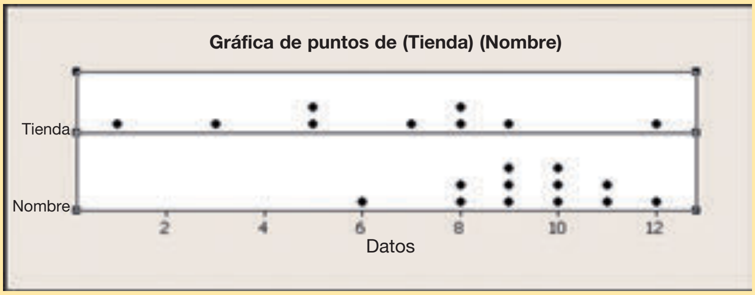
Una muestra aleatoria independiente de 12 toallas de la otra marca absorbió las cantidades siguientes de líquido en milímetros.

12	11	10	6	8	9	9	10	11	9	8	10

Utilice el nivel de significancia de 0.10 y pruebe si existe una diferencia entre las cantidades medias de líquido que absorbieron los dos tipos de toallas.

Solución

Para iniciar se supone que las cantidades de líquido absorbido siguen la distribución de probabilidad normal de las toallas de la segunda marca como de las de la primera. No se conocen las desviaciones estándares de las poblaciones, por lo que se empleará la distribución t como estadístico de prueba. No parece razonable la suposición de desviaciones estándares de las poblaciones iguales. La cantidad de absorción en la primera marca varía de 1 ml a 12 ml. En el caso de la segunda, la cantidad de absorción varía de 6 ml a 12 ml. Es decir, existe más variación en la cantidad de absorción de la primera marca que de la segunda. Se observa la diferencia en la variación en la gráfica de puntos siguiente que se obtuvo con Minitab. Los comandos del software para crear una gráfica de puntos en Minitab se dan en la página 135.



Por lo tanto, se decide emplear la distribución t y suponer que las desviaciones estándares de las poblaciones no son iguales.

En el procedimiento de prueba de hipótesis de cinco pasos, el primero es formular las hipótesis nula y alternativa. La hipótesis nula es que no hay diferencia en la cantidad media de líquido que absorben ambos tipos de toallas. La hipótesis alternativa es que sí hay una diferencia.

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

El nivel de significancia es 0.10, y el estadístico de prueba sigue la distribución *t*. Como no se desea suponer desviaciones estándares de las poblaciones iguales, se ajustan los grados de libertad con la fórmula (11-8). Para hacer ello se necesita determinar las desviaciones estándares de las muestras. El sistema Minitab es útil para determinar rápidamente estos resultados. También se encontrará la tasa de absorción media, la cual se empleará en breve. Los tamaños muestrales respectivos son $n_1 = 9$ y $n_2 = 12$, y las desviaciones estándares respectivas, 3.32 ml y 1.621 ml.

```
Estadísticos  descriptivos:  Tienda,  Nombre

Variable          N    Media       Desv.est.
Tienda            9     6.44            3.32
Nombre           12    9.417           1.621
```

Al sustituir esta información en la fórmula (11-8):

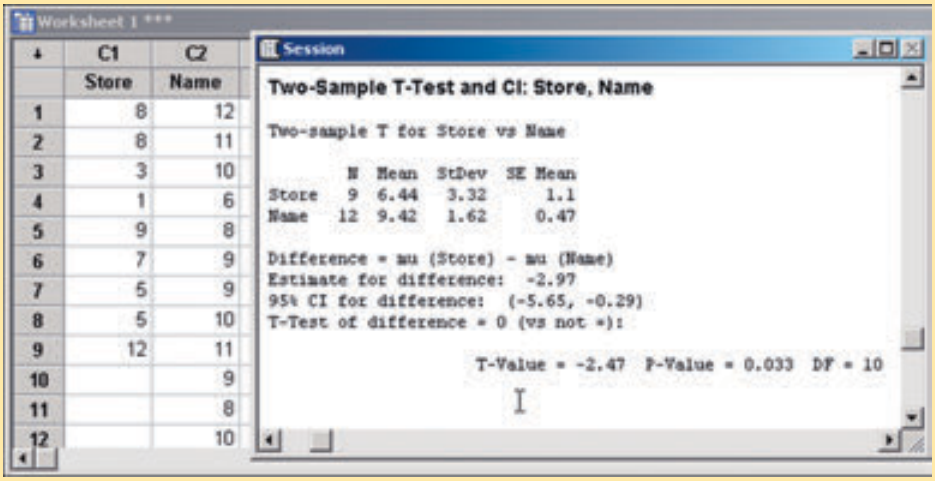
$$gl = \frac{[(s_1^2/n_1) + (s_2^2/n_2)]^2}{\dfrac{(s_1^2/n_1)^2}{n_1 - 1} + \dfrac{(s_2^2/n_2)^2}{n_2 - 1}} = \frac{[(3.32^2/9) + (1.621^2/12)]^2}{\dfrac{(3.32^2/9)^2}{9 - 1} + \dfrac{(1.621^2/12)^2}{12 - 1}} = \frac{1.4436^2}{.1875 + .0043} = 10.88$$

La práctica común es redondear hacia abajo a un entero, por lo que se emplean 10 grados de libertad. Del apéndice B.2 con 10 grados de libertad, una prueba de dos colas y un nivel de significancia de 0.10, los valores *t* críticos son -1.812 y 1.812. La regla de decisión es rechazar la hipótesis nula si el valor calculado de *t* es menor que -1.812 o mayor que 1.812.

Para determinar el valor del estadístico de prueba se emplea la fórmula (11-7). Recuerde, de la salida Minitab anterior, que la cantidad de absorción de las toallas de papel de primera marca es 6.44 ml, y 9.417 ml de la otra.

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} = \frac{6.44 - 9.417}{\sqrt{\dfrac{3.32^2}{9} + \dfrac{1.621^2}{12}}} = -2.478$$

El valor calculado de *t* es menor que el valor crítico menor, por lo que la decisión es rechazar la hipótesis nula. Se concluye que la tasa de absorción media de las dos toallas no es la misma. La salida de Minitab para este ejemplo es la siguiente.



Autoevaluación 11-4

Con frecuencia para las compañías es útil saber quiénes son sus clientes y cómo se convirtieron en lo que son. Una compañía de tarjetas de crédito tiene interés en saber si el tarjetahabiente la solicitó por interés propio o si fue contactado por teléfono por un agente. La compañía obtuvo la información muestral siguiente respecto de los saldos al final del mes de los dos grupos.

Fuente	Media	Desviación estándar	Tamaño de la muestra
Solicitantes	$1 568	$356	10
Contactados	1 967	857	8

¿Es razonable concluir que el saldo medio de los tarjetahabientes que fueron contactados por teléfono es mayor que el de quienes solicitaron la tarjeta por cuenta propia? Suponga que las desviaciones estándares de las poblaciones no son iguales. Utilice el nivel de significancia 0.05.

a) Formule las hipótesis nula y alternativa.
b) ¿Cuántos grados de libertad hay?
c) ¿Cuál es la regla de decisión?
d) ¿Cuál es el valor del estadístico de prueba?
e) ¿Cuál es su decisión respecto de la hipótesis nula?
f) Interprete el resultado.

Ejercicios

En los ejercicios 19 y 20 suponga que las poblaciones muestrales no tienen desviaciones estándares iguales y utilice el nivel de significancia 0.05: *a*) determine el número de grados de libertad, *b*) formule la regla de decisión, *c*) calcule el valor del estadístico de prueba y *d*) tome su decisión acerca de la hipótesis nula.

19. Las hipótesis nula y alternativa son:

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

Una muestra aleatoria de 15 elementos de la primera población reveló una media de 50 y una desviación estándar de 5. Una muestra de 12 elementos para la segunda población reveló una media de 46 y una desviación estándar de 15.

20. Las hipótesis nula y alternativa son:

$$H_0: \mu_1 \leq \mu_2$$
$$H_1: \mu_1 > \mu_2$$

Una muestra aleatoria de 20 elementos de la primera población reveló una media de 100 y una desviación estándar de 15. Una muestra de 16 elementos de la segunda población reveló una media de 94 y una desviación estándar de 8. Utilice un nivel de significancia de 0.05.

21. En un artículo reciente en *The Wall Street Journal* se comparó el costo de adopción de niños de China con el de Rusia. En una muestra de 16 adopciones de China, el costo medio fue $11 045, con una desviación estándar de $835. En una muestra de 18 adopciones de niños de Rusia, el costo medio fue $12 840, con una desviación estándar de $1 545. ¿Puede concluir que el costo medio de adoptar niños es mayor en Rusia? Suponga que las dos desviaciones estándares poblacionales no son iguales. Utilice el nivel de significancia de 0.05.

22. Suponga que usted es un experto en la industria de la moda y desea reunir información para comparar la cantidad mensual que ganan las modelos que vistieron ropa de Liz Claiborne con respecto a las que modelaron ropa de Calvin Klein. La siguiente es la cantidad (en miles de dólares) que gana al mes por una muestra de modelos de Liz Claiborne:

$5.0	$4.5	$3.4	$3.4	$6.0	$3.3	$4.5	$4.6	$3.5	$5.2
4.8	4.4	4.6	3.6	5.0					

La siguiente es la cantidad (en miles de dólares) que gana una muestra de modelos de Calvin Klein:

\$3.1	\$3.7	\$3.6	\$4.0	\$3.8	\$3.8	\$5.9	\$4.9	\$3.6	\$3.6
2.3	4.0								

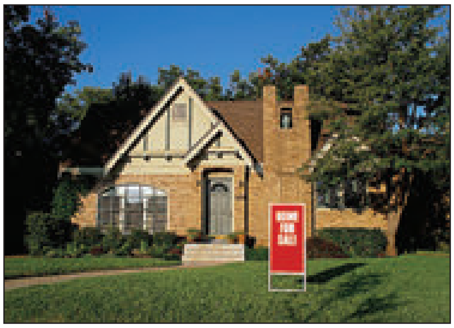
¿Es razonable concluir que las modelos de Claiborne ganan más? Utilice un nivel de significancia de 0.05 y suponga que las desviaciones estándares de las poblaciones no son iguales.

11.5 Pruebas de hipótesis de dos muestras: muestras dependientes

OA5 Comprender la diferencia entre muestras dependientes e independientes.

En la página 383 se probó la diferencia entre las medias de dos muestras independientes. Se comparó el tiempo medio que se requiere para montar un motor según el método de Welles con el de Atkins. Las muestras eran *independientes*, lo que significa que la muestra de los tiempos de ensamblado del método de Welles no estaba de ninguna manera relacionada con la muestra de los tiempos que insumía el de Atkins.

Sin embargo, hay situaciones en que las muestras no son independientes. En otras palabras, las muestras son *dependientes* o están *relacionadas*. Como ejemplo, la compañía Nickel Savings and Loan recurre a dos empresas, Schadek Appraisals y Bowyer Real State, para valuar los bienes raíces sobre los cuales se hacen los préstamos. Es importante que los avalúos de estas dos empresas contemplen valores similares. Para revisar la consistencia de las dos empresas, Nickel Savings selecciona en forma aleatoria 10 casas y pide a Schadek Appraisals y a Bowyer Real State que las valúen. De cada una se harán dos avalúos; cada casa tendrá un avalúo de Schadek Appraisals y otro de Bowyer Real State. Los avalúos dependen o están relacionados con la casa seleccionada. A esto también se le conoce como **muestra apareada**.



Para la prueba de hipótesis el interés recae en la distribución de las *diferencias* entre los valores de avalúo de cada casa. De aquí, sólo hay una muestra. En palabras más formales, se investiga si la media de la distribución de las diferencias entre los avalúos es 0. La muestra se compone de las *diferencias* entre los avalúos determinados por Schadek Appraisals y los de Bowyer Real State. Si las dos empresas reportan estimaciones similares, entonces algunas veces los avalúos de Schadek serán los de valor mayor y otras veces lo serán los de Bowyer Real State. Sin embargo, la media de la distribución de las diferencias será 0. Por otro lado, si una de las empresas reporta de manera consistente los avalúos más altos, la media de la distribución de las diferencias no será 0.

OA6 Efectuar una prueba de hipótesis acerca de la diferencia media entre observaciones apareadas y dependientes.

Se empleará el símbolo μ_d para indicar la media poblacional de la distribución de las diferencias. Se supone que la distribución de las diferencias de la población sigue la distribución normal. El estadístico de prueba sigue la distribución *t,* y su valor se calcula a partir de la fórmula siguiente:

PRUEBA *t* APAREADA

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} \qquad \textbf{(11-9)}$$

Hay $n - 1$ grados de libertad y

$\bar{d}$ es la media de la diferencia entre las observaciones apareadas o relacionadas.
s_d es la desviación estándar de las diferencias entre las observaciones apareadas o relacionadas.
n es el número de observaciones apareadas.

La desviación estándar de las diferencias se calcula mediante la conocida fórmula de la desviación estándar, excepto que X se sustituye por d. La fórmula es:

$$s_d = \sqrt{\frac{\Sigma(d - \bar{d})^2}{n - 1}}$$

En el ejemplo siguiente se ilustra esta prueba.

Ejemplo

Recuerde que Nickel Savings and Loan desea comparar las dos compañías que contrata para valuar las casas. Nickel Savings seleccionó una muestra de 10 propiedades y programa los avalúos de las dos empresas. Los resultados, en miles de dólares, son:

Casa	Schadek	Bowyer
1	235	228
2	210	205
3	231	219
4	242	240
5	205	198
6	230	223
7	231	227
8	210	215
9	225	222
10	249	245

Con un nivel de significancia de 0.05, ¿se puede concluir que hay una diferencia entre los avalúos medios de las casas?

Solución

El primer paso es formular las hipótesis nula y alternativa. En este caso es adecuada una alternativa de dos colas porque se tiene interés en determinar si hay una *diferencia* entre los avalúos. No existe interés en demostrar si una empresa en particular valúa las propiedades con un valor mayor que la otra. La pregunta es si las diferencias en la muestra entre los avalúos pueden provenir de una población con una media de 0. Si la media de las diferencias de la población es 0, se concluye que no hay diferencia entre los avalúos. Las hipótesis nula y alternativa son:

$$H_0: \mu_d = 0$$
$$H_1: \mu_d \neq 0$$

Hay 10 casas valuadas por las dos empresas, por lo que $n = 10$, y $gl = n - 1 = 10 - 1 = 9$. Se tiene una prueba de dos colas, y el nivel de significancia es 0.05. Para determinar el valor crítico consulte el apéndice B.2, y vea la fila con 9 grados de libertad hasta la columna de una prueba de dos colas y el nivel de significancia 0.05. El valor en la intersección es 2.262. Este valor aparece en el cuadro de la tabla 11-2. La regla de decisión es rechazar la hipótesis nula si el valor calculado de t es menor que -2.262 o mayor que 2.262. Éstos son los detalles del cálculo.

Casa	Schadek	Bowyer	Diferencia, d	$(d - \bar{d})$	$(d - \bar{d})^2$
1	235	228	7	2.4	5.76
2	210	205	5	0.4	0.16
3	231	219	12	7.4	54.76
4	242	240	2	−2.6	6.76

(*continúa*)

Casa	Schadek	Bowyer	Diferencia, d	$(d - \bar{d})$	$(d - \bar{d})^2$
5	205	198	7	2.4	5.76
6	230	223	7	2.4	5.76
7	231	227	4	−0.6	0.36
8	210	215	−5	−9.6	92.16
9	225	222	3	−1.6	2.56
10	249	245	4	−0.6	0.36
			46	0	174.40

$$\bar{d} = \frac{\Sigma d}{n} = \frac{46}{10} = 4.60$$

$$s_d = \sqrt{\frac{\Sigma(d - \bar{d})^2}{n - 1}} = \sqrt{\frac{174.4}{10 - 1}} = 4.402$$

Con la fórmula (11-9), el valor del estadístico de prueba es 3.305, determinado por

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} = \frac{4.6}{4.402/\sqrt{10}} = \frac{4.6}{1.3920} = 3.305$$

Como el valor calculado de t se encuentra en la región de rechazo, se rechaza la hipótesis nula. La distribución de las diferencias de la población no tiene una media de 0. Se concluye que hay una diferencia entre los avalúos medios de las casas. La diferencia mayor de $12 000 es en la casa 3. Quizás éste sería un buen lugar para iniciar una revisión más detallada.

Para determinar el valor p, consulte el apéndice B.2 y la sección de una prueba de dos colas. Busque en la fila con 9 grados de libertad y encuentre los valores de t que se aproximen al valor calculado. Para un nivel de significancia de 0.01, el valor de t es 3.250. El valor calculado es mayor, pero menor que el valor de 4.781 que corresponde al nivel de significancia de 0.001. De aquí, el valor p es menor que 0.01. Esta información se resalta en la tabla 11-2.

TABLA 11-2 Parte de la distribución t del apéndice B.2

	Intervalos de confianza					
	80%	90%	95%	98%	99%	99.9%
	Nivel de significancia de una prueba de una cola					
gl	0.10	0.05	0.025	0.01	0.005	0.0005
	Nivel de significancia de una prueba de dos colas					
	0.20	0.10	0.05	0.02	0.01	0.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.599
3	1.638	2.353	3.182	4.541	5.841	12.924
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.869
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.408
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	$\boxed{2.262}$	2.821	$\boxed{3.250}$	$\boxed{4.781}$
10	1.372	1.812	2.228	2.764	3.169	4.587

Excel tiene un procedimiento denominado “Prueba *t*: Dos muestras apareadas para medias” que realiza los cálculos de la fórmula (11-9). La captura de pantalla de este procedimiento aparece a continuación.

El valor calculado de *t* es 3.305, y el valor *p* de dos colas, 0.009. Como el valor *p* es menor que 0.05, se rechaza la hipótesis de que la media de la distribución de las diferencias entre los avalúos es cero. De hecho, este valor *p* se encuentra entre 0.01 y 0.001. Hay una pequeña posibilidad de que la hipótesis nula sea verdadera.

paired t test

	A	B	C	D	E	F	G
1	Home	Schadek	Bowyer		t-Test: Paired Two Sample for Means		
2	1	235	228				
3	2	210	205			*Schadek*	*Bowyer*
4	3	231	219		Mean	226.800	222.200
5	4	242	240		Variance	208.844	204.178
6	5	205	198		Observations	10.000	10.000
7	6	230	223		Pearson Correlation	0.953	
8	7	231	227		Hypothesized Mean Difference	0.000	
9	8	210	215		df	9.000	
10	9	225	222		t Stat	3.305	
11	10	249	245		P(T<=t) one-tail	0.005	
12					t Critical one-tail	1.833	
13					P(T<=t) two-tail	0.009	
14					t Critical two-tail	2.262	
15							

11.6 Comparación de muestras dependientes e independientes

Con frecuencia, los estudiantes principiantes confunden la diferencia entre las pruebas de muestras independientes [fórmula (11-6)] con las pruebas de muestras dependientes [fórmula (11-9)]. ¿Cómo distinguir la diferencia entre muestras dependientes e independientes? Hay dos tipos de muestras dependientes: 1) las que se caracterizan por una medición, una intervención de algún tipo y después otra medición, y 2) una relación o agrupación de las observaciones. Para explicarlo con más detalle:

1. El primer tipo de muestra dependiente se caracteriza por una medición seguida de una intervención de alguna clase y después otra medición. Esto se puede denominar un estudio de “antes” y “después”. Dos ejemplos ayudarán a explicarlo mejor. Suponga que desea demostrar que, al colocar bocinas en el área de producción y tocar música relajante, aumenta la producción. Comienza con la selección de una muestra de trabajadores y una medición de sus resultados en las condiciones actuales. Después instala las bocinas en el área de producción y vuelve a medir la producción de los mismos trabajadores. Hay dos mediciones: antes de colocar las bocinas en el área de producción y después. La intervención es la colocación de las bocinas en el área de producción.

   Un segundo ejemplo comprende una empresa educativa que ofrece cursos diseñados para incrementar las calificaciones en los exámenes y la capacidad para leer (SAT). Suponga que la empresa quiere ofrecer un curso que ayudará a los alumnos de primer año de preparatoria a aumentar sus puntajes en el SAT. Para iniciar, cada estudiante presenta el SAT en el primer año de preparatoria. Durante el verano, entre los años primero y último, participan en el curso que les proporciona consejos para presentar exámenes. Para finalizar, durante el otoño del último año de preparatoria, vuelven a presentar el SAT. Una vez más, el procedimiento se caracteriza por una medición (presentar el SAT como estudiante de primer año), una intervención (los talleres de verano) y otra medición (presentar el SAT durante su último año).

2. El segundo tipo de muestra dependiente se caracteriza por relacionar o aparear observaciones. En el ejemplo anterior, Nickel Savings es una muestra dependiente de este tipo. Se seleccionó una propiedad para su valuación y después obtuvo dos valuaciones sobre ella. Como segundo ejemplo, suponga que una psicóloga industrial desea estudiar las similitudes intelectuales de parejas recién casadas, para lo cual selecciona una muestra de recién casados. Después, administra una prueba de inteligencia estándar tanto al hombre como a la mujer para determinar la diferencia entre las calificaciones. Observe la relación que ocurrió: se comparan las calificaciones apareadas o relacionadas por un matrimonio.

¿Por qué se prefieren las muestras dependientes a las independientes? Cuando se emplean muestras dependientes, se reduce la variación en la distribución del muestreo. Para ilustrar este ejemplo se utilizará el caso de Nickel Savings and Loan. Suponga que se tienen dos muestras independientes de propiedades de bienes raíces para su avalúo y se realiza la prueba de hipótesis siguiente, con la fórmula (11-6). Las hipótesis nula y alternativa son:

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

Ahora hay dos muestras independientes de 10 cada una. Así, el número de grados de libertad es $10 + 10 - 2 = 18$. Del apéndice B.2, en el nivel de significancia de 0.05, H_0 se rechaza si t es menor que -2.101 o mayor que 2.101.

Se emplean los mismos comandos de Excel que en la página 100 del capítulo 3 para determinar la media y la desviación estándar de las dos muestras independientes, y los comandos de Excel de la página 408 de este capítulo para encontrar la varianza agrupada y el valor de “t Stat”. Estos valores están resaltados con color amarillo.

Independent t test

	A	B	C	D	E	F	G	H
1	Home	Schadek	Bowyer		t-Test: Two-Sample Assuming Equal Variances			
2	1	235	228					
3	2	210	205			Schadek	Bowyer	
4	3	231	219		Mean	226.800	222.200	
5	4	242	240		Variance	208.844	204.178	
6	5	205	198		Observations	10.000	10.000	
7	6	230	223		Pooled Variance	206.511		
8	7	231	227		Hypothesized Mean Difference	0.000		
9	8	210	215		df	18.000		
10	9	225	222		t Stat	0.716		
11	10	249	245		P(T<=t) one-tail	0.242		
12					t Critical one-tail	1.734		
13	Mean =	226.80	222.20		P(T<=t) two-tail	0.483		
14	S =	14.45	14.29		t Critical two-tail	2.101		
15								

La media del avalúo de las 10 propiedades de Schadek es de \$226 800, y la desviación estándar, \$14 500. La media de los avalúos de Bowyer Real State es de \$222 200, y la desviación estándar, \$14 290. Para facilitar los cálculos, se emplean miles de dólares en lugar de dólares. El valor de la estimación agrupada de la varianza a partir de la fórmula (11-5) es

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(10 - 1)(14.45^2) + (10 - 1)(14.29)^2}{10 + 10 - 2} = 206.50$$

De la fórmula (11-6), t es 0.716.

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{226.8 - 222.2}{\sqrt{206.50\left(\frac{1}{10} + \frac{1}{10}\right)}} = \frac{4.6}{6.4265} = 0.716$$

El valor calculado de t (0.716) es menor que 2.101, de manera que la hipótesis nula no se rechaza. No es posible demostrar que hay una diferencia entre los avalúos medios. ¡Ésta no es la misma conclusión a la que se llegó antes! ¿Por qué? El numerador es el mismo que en la prueba de observaciones apareadas (4.6). Sin embargo, el denominador es menor. En la prueba por pares el denominador es 1.3920 (vea los cálculos en la página 394). En el caso de las muestras independientes, el denominador es 6.4265. La variación o incertidumbre son mayores. Esto explica la diferencia entre los valores t y la diferencia entre las decisiones estadísticas. El denominador mide el error estándar de la estadística. Cuando las muestras *no* se aparean, se presentan dos clases de variación: diferencias entre las dos empresas valuadoras y la diferencia en el valor del bien raíz. Las propiedades 4 y 10 tienen valores comparativamente altos, en tanto que el del número 5 es relativamente bajo. Estos datos muestran lo diferentes que son los avalúos de las propiedades, pero lo que interesa en realidad es la diferencia entre las dos empresas valuadoras.

La estrategia es aparear los valores para reducir la variación entre las propiedades. En la prueba apareada sólo se emplea la diferencia entre las dos empresas valuadoras para la misma propiedad. Así, la estadística apareada o dependiente se enfoca sobre la variación entre Schadek Appraisals y Bowyer Real State. Por lo tanto, su error estándar siempre es menor. Esto, a su vez, conduce a una estadística de prueba mayor y a una probabilidad mayor de rechazar la hipótesis nula. Por lo tanto, siempre que sea posible se deben aparear los datos.

Aquí hay una mala noticia. En la prueba de observaciones apareadas, los grados de libertad son la mitad de lo que serían si no se apareasen las muestras. En el ejemplo de bienes raíces, los grados de libertad disminuyen de 18 a 9 cuando las observaciones están apareadas. Sin embargo, en la mayoría de los casos, éste es un precio pequeño que se debe pagar por una prueba mejor.

Autoevaluación 11-5

La publicidad que realiza Sylph Fitness Center afirma que, al terminar su entrenamiento, las personas bajarán de peso. Una muestra aleatoria de ocho participantes recientes reveló los pesos siguientes antes y después de terminar el entrenamiento. Con un nivel de significancia de 0.01, ¿se puede concluir que los participantes bajan de peso?

Nombre	Antes	Después
Hunter	155	154
Cashman	228	207
Mervine	141	147
Massa	162	157
Creola	211	196
Peterson	164	150
Redding	184	170
Poust	172	165

a) Formule las hipótesis nula y alternativa.
b) ¿Cuál es el valor crítico de t?
c) ¿Cuál es el valor calculado de t?
d) Interprete el resultado. ¿Cuál es el valor p?
e) ¿Qué suposición necesita acerca de la distribución de las diferencias?

Ejercicios

23. Las hipótesis nula y alternativa son:

$$H_0: \mu_d \leq 0$$
$$H_1: \mu_d > 0$$

En la información muestral siguiente aparece el número de unidades defectuosas que producen los turnos matutino y vespertino en una muestra de cuatro días durante el mes pasado.

	Día			
	1	**2**	**3**	**4**
Turno matutino	10	12	15	19
Turno vespertino	8	9	12	15

Con un nivel de significancia de 0.05, ¿se puede concluir que se producen más defectos en el turno vespertino?

24. Las hipótesis nula y alternativa son:

$$H_0: \mu_d = 0$$
$$H_1: \mu_d \neq 0$$

Las observaciones apareadas siguientes muestran el número de multas de tránsito por conducir a exceso de velocidad de los oficiales Dhondt y Meredith, de la South Carolina Highway Patrol, durante los últimos cinco meses.

	Día				
	Mayo	**Junio**	**Julio**	**Agosto**	**Septiembre**
Oficial Dhondt	30	22	25	19	26
Oficial Meredith	26	19	20	15	19

Con un nivel de significancia de 0.05, ¿hay alguna diferencia entre los números medios de multas que dieron los dos oficiales?

Nota: Para resolver los ejercicios siguientes utilice el procedimiento de prueba de hipótesis de cinco pasos.

25. La gerencia de Discount Furniture, cadena de mueblerías de descuento del noreste de Estados Unidos, diseñó un plan de incentivos para sus agentes de ventas. Para evaluar este plan innovador, se seleccionaron a 12 vendedores al azar, y se registraron sus ingresos anteriores y posteriores al plan.

Vendedor	Antes	Después
Sid Mahone	$320	$340
Carol Quick	290	285
Tom Jackson	421	475
Andy Jones	510	510
Jean Sloan	210	210
Jack Walker	402	500
Peg Mancuso	625	631
Anita Loma	560	560
John Cuso	360	365
Carl Utz	431	431
A. S. Kushner	506	525
Fern Lawton	505	619

¿Hubo algún aumento significativo en el ingreso semanal de un vendedor debido al innovador plan de incentivos? Utilice el nivel de significancia 0.05. Calcule el valor p e interprételo.

26. Hace poco, el gobierno federal estadounidense otorgó fondos para un programa especial diseñado para reducir los delitos en áreas de alto riesgo. Un estudio de los resultados del programa en ocho áreas de alto riesgo de Miami, Florida, produjo los resultados siguientes.

	Número de delitos por área							
	A	B	C	D	E	F	G	H
Antes	14	7	4	5	17	12	8	9
Después	2	7	3	6	8	13	3	5

¿Hubo alguna disminución en el número de delitos desde la inauguración del programa? Utilice el nivel de significancia 0.01. Calcule el valor p.

Resumen del capítulo

I. Al comparar dos medias poblacionales se desea saber si pueden ser iguales.
   A. Se investiga si la distribución de la diferencia entre las medias puede tener una media de 0.
   B. El estadístico de prueba sigue la distribución normal estándar si se conocen las desviaciones estándares de las poblaciones.
      1. No se requiere de ninguna suposición acerca de la forma de las poblaciones.
      2. Las muestras son de poblaciones independientes.
      3. La fórmula para calcular el valor z es

$$z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \qquad \textbf{(11-2)}$$

II. También se puede comprobar si dos muestras provienen de poblaciones con la misma proporción de éxitos.
   A. Las dos proporciones muestrales se agrupan con la fórmula siguiente:

$$p_c = \frac{X_1 + X_2}{n_1 + n_2} \qquad \textbf{(11-4)}$$

   B. Se calcula el valor del estadístico de prueba a partir de la fórmula siguiente:

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_c(1 - p_c)}{n_1} + \frac{p_c(1 - p_c)}{n_2}}} \qquad \textbf{(11-3)}$$

III. El estadístico de prueba para comparar dos medias es la distribución t, si no se conocen las desviaciones estándares poblacionales.
   A. Las dos poblaciones deben seguir la distribución normal.
   B. Las poblaciones deben tener desviaciones estándares iguales.
   C. Las muestras son independientes.
   D. La determinación del valor de t requiere dos pasos.
      1. El primer paso es agrupar las desviaciones estándares de acuerdo con la fórmula siguiente:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \qquad \textbf{(11-5)}$$

      2. El valor de t se calcula a partir de la fórmula siguiente:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \qquad \textbf{(11-6)}$$

      3. Los grados de libertad de la prueba son $n_1 + n_2 - 2$.

IV. Si no es posible suponer que las desviaciones estándares de la población son iguales,

A. Utilice la distribución t como el estadístico de prueba, pero ajuste los grados de libertad mediante la fórmula siguiente:

$$gl = \frac{[(s_1^2/n_1) + (s_2^2/n_2)]^2}{\dfrac{(s_1^2/n_1)^2}{n_1 - 1} + \dfrac{(s_2^2/n_2)^2}{n_2 - 1}} \quad \textbf{(11-8)}$$

B. El valor del estadístico de prueba se calcula a partir de la fórmula siguiente:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} \quad \textbf{(11-7)}$$

V. Para muestras dependientes, se supone que la distribución de las diferencias apareadas entre las poblaciones tiene una media de 0.

A. Primero se calcula la media y la desviación estándar de las diferencias muestrales.

B. El valor del estadístico de prueba se calcula a partir de la fórmula siguiente:

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} \quad \textbf{(11-9)}$$

Clave de pronunciación

SÍMBOLO	SIGNIFICADO	PRONUNCIACIÓN
p_c	Proporción conjunta	*p subíndice c*
s_p^2	Varianza conjunta de la muestra	*s subíndice p al cuadrado*
$\bar{X}_1$	Media de la primera muestra	*X barra subíndice* 1
$\bar{X}_2$	Media de la segunda muestra	*X barra subíndice* 2
$\bar{d}$	Media de la diferencia entre observaciones dependientes	*d barra*
s_d	Desviación estándar de la diferencia entre observaciones dependientes	*s subíndice d*

Ejercicios del capítulo

connect™

27. Un estudio reciente se enfocó en el número de veces que los hombres y las mujeres que viven solos compran comida para llevar en un mes. La información se resume a continuación.

Estadístico	Hombres	Mujeres
Media de la muestra	24.51	22.69
Desviación estándar de la **población**	4.48	3.86
Tamaño de la muestra	35	40

Con un nivel de significancia de 0.01, ¿hay alguna diferencia entre el número medio de veces que los hombres y las mujeres piden comida para llevar en un mes? ¿Cuál es el valor p?

28. Clark Heter es un ingeniero industrial en Lyons Products, y le gustaría determinar si se producen más unidades en el turno nocturno que en el matutino. Suponga que la desviación estándar de la población del número de unidades producidas en el turno matutino es 21 y 28 en el nocturno. Una muestra de 54 trabajadores del turno matutino reveló que el número medio de unidades producidas fue 345. Una muestra de 60 trabajadores del turno nocturno reveló que el número medio de unidades producidas fue 351. Con un nivel de significación de 0.05, ¿es mayor el número de unidades producidas en el turno nocturno?

29. Fry Brothers Heating and Air Conditioning, Inc., emplea a Larry Clark y George Murnen para ofrecer por teléfono servicios de reparación de chimeneas y unidades de aire acondicionado en casas. Al propietario, Tom Fry, le gustaría saber si hay alguna diferencia entre los números medios de lla-