CAPÍTULO 19

Métodos no paramétricos

CONTENIDO

LA ESTADÍSTICA *en* LA PRÁCTICA

WEST SHELL REALTORS*
CINCINNATI, OHIO

La empresa West Shell Realtors fue fundada en 1958 y en ese entonces contaba con una oficina y un equipo de ventas formado por tres personas. En 1964 inició un programa de expansión a largo plazo durante el cual, casi anualmente, abrió nuevas oficinas. Con el tiempo, West Shell creció hasta convertirse en la mayor empresa inmobiliaria de Greater Cincinnati y ahora cuenta con oficinas en el suroeste de Ohio, en el sureste de Indiana y en el norte de Kentucky.

A las empresas de bienes raíces, como West Shell, el análisis estadístico les sirve para monitorear el curso de sus ventas. Cada mes se presenta un informe de cada una de las oficinas de West Shell, así como del total de la empresa. Resumen estadístico sobre la cantidad total de dólares en ventas, número de unidades vendidas y precio mediano de venta por unidad son esenciales para mantener informados, tanto a los gerentes de las distintas oficinas, como a los gerentes generales sobre el progreso y los problemas de la organización.

Además de los resúmenes mensuales sobre el curso de las operaciones, la empresa emplea diversas consideraciones estadísticas como guía para sus planes y estrategias. West Shell ha implementado una estrategia de expansión planeada. Cada vez que, debido a este plan de expansión, se quiere abrir una nueva oficina de ventas, la empresa tiene que decidir dónde abrir la nueva oficina. El tipo de datos usados para evaluar y comparar las distintas alternativas para la ubicación de una nueva oficina son los precios de venta de las casas, las tasas de facturación y los volúmenes de ventas pronosticados. Con el fin de identificar si existía alguna diferencia entre los patrones de ventas de estas dos áreas, West Shell se valió de métodos estadísticos no paramétricos.

En una ocasión West Shell tenía dos zonas posibles para abrir una nueva oficina; Clifton y Roselawn. Al comparar las dos zonas se tomaron en consideración diversos factores, entre ellos, los precios de venta de las casas.

Para mantener su competitividad, West Shell realiza análisis estadísticos de los precios de las casas. © Cortesía de Coldwell Banker West Shell.

A partir de muestras de 25 ventas en Clifton y 18 ventas en Roselawn se eligió la prueba de la suma de los rangos de Mann-Whitney Wilcoxon como la prueba estadística adecuada para las diferencias entre los patrones de ventas. Con un nivel de significancia de 0.05, la prueba de Mann-Whitney-Wilcoxon no permitió rechazar la hipótesis nula de que las dos poblaciones de precios de venta fueran idénticas. Por tanto, West Shell tuvo que buscar otros criterios, distintos a los precios de venta de las casas, para su proceso de selección de la ubicación de su nueva oficina.

En este capítulo se mostrará cómo aplicar pruebas estadísticas no paramétricas como la prueba de Mann-Whitney-Wilcoxon. También se discutirá la interpretación adecuada de dichas pruebas.

* Los autores agradecen a Rodney Fightmaster de West Shell Realtors por proporcionar este artículo para *La estadística en la práctica*.

Los métodos estadísticos hasta ahora presentados en este libro se conocen como *métodos paramétricos*. En este capítulo se presentan varios **métodos no paramétricos**. Estos métodos suelen ser aplicables en las situaciones en que los métodos paramétricos no lo son. Los métodos no paramétricos suelen requerir suposiciones menos restrictivas acerca del nivel de medición de los datos y menos suposiciones acerca de la forma de las distribuciones de probabilidad generadas por los datos muestrales.

Una de las consideraciones para determinar si lo apropiado es un método paramétrico o un método no paramétrico es la escala de medición empleada para generar los datos. Todos los datos son generados por una de las cuatro escalas de medición: nominal, ordinal, de intervalo o de razón. Por tanto, todos los análisis estadísticos se realizan con datos ya sea nominales, ordinales, de intervalo o de razón.

A continuación se definen y se proporcionan ejemplos de cada una de las cuatro escalas de medición.

1. *Escala nominal.* Una escala de medición es nominal si los datos son etiquetas o categorías que se usan para definir un atributo de un elemento. Los datos nominales pueden ser numéricos o no numéricos.

   Ejemplos. El mercado en el que cotiza una acción (NYSE, NASDAQ o AMEX) es un dato nominal no numérico. El número de seguro social de una persona es un dato nominal numérico.
2. *Escala ordinal*. Una escala de medición es ordinal si los datos pueden usarse para jerarquizar u ordenar las observaciones. Los datos ordinales pueden ser numéricos o no numéricos.

   Ejemplos. Las medidas pequeño, mediano y grande para dar el tamaño de un objeto son datos ordinales no numéricos. El lugar de los individuos en una clase 1, 2, 3, … son datos ordinales numéricos.
3. *Escala de intervalo.* Una escala de medición es de intervalo si los datos tienen las propiedades de los datos ordinales y los intervalos entre observaciones se expresan en términos de una unidad de medición fija. Los datos de intervalo tienen que ser numéricos.

   Ejemplos. Las mediciones de temperatura son datos de intervalo. Suponga que la temperatura en un lugar es de 21°C y en otro es de 4°C. Estos lugares se pueden jerarquizar de acuerdo con lo calurosos que son: el primero es más caliente que el segundo. La unidad fija de medición, 1°C, permite decir cuán más caliente es el primer lugar: 17°C.
4. *Escala de razón*. Una escala de medición es de razón si los datos tienen las propiedades de los datos de intervalo y el cociente (o razón) entre dos medidas tiene sentido. Los datos de razón tienen que ser numéricos.

   Ejemplos. Variables como la distancia, la altura, el peso y el tiempo se miden con una escala de razón. Las mediciones de temperatura no son datos de razón debido a que no existe un punto cero definido intrínsecamente. Por ejemplo, el punto de congelación del agua en la escala Fahrenheit es 32 grados y en la escala Celsius es 0 grados. Los cocientes entre datos de temperatura no tienen sentido. Por ejemplo, no tiene sentido decir que cuando la temperatura ambiente es de 20 grados hace el doble de calor que cuando es de 10 grados.

En el capítulo 1 se dijo que con las escalas nominal y ordinal se obtienen datos cualitativos. Con las escalas de intervalo y de razón se obtienen datos cuantitativos.

La mayor parte de los métodos estadísticos conocidos como métodos paramétricos requieren el uso de datos de las escalas de intervalo o de razón. Con estos niveles de medición, tienen sentido las operaciones aritméticas y medias, varianzas, desviaciones estándar, etc., pueden calcularse, interpretarse y usarse en el análisis. Con datos nominales y ordinales no es apropiado calcular medias, varianzas ni desviaciones estándar; por tanto, no pueden emplearse los métodos paramétricos. La única manera de analizar esos datos para obtener conclusiones estadísticas es emplear los métodos no paramétricos.

Si el nivel de medición de los datos es nominal u ordinal, calcular la media, la varianza y la desviación estándar no tiene sentido. Por tanto, con este tipo de datos, muchos de los procedimientos estadísticos discutidos previamente no pueden emplearse.

En general, para que un método estadístico se clasifique como método no paramétrico, debe satisfacer, por lo menos, una de las condiciones siguientes.*

1. Ser un método que pueda ser usado con datos nominales.
2. Ser un método que pueda ser usado con datos ordinales.
3. Ser un método que pueda ser usado con datos de intervalo o de razón cuando no sea posible hacer suposiciones acerca de la forma de la distribución de la población.

Si el nivel de medición de los datos es de intervalo o de razón y si las suposiciones necesarias sobre la distribución de probabilidad de la población son apropiadas, con los métodos paramétricos se obtienen procedimientos estadísticos más potentes y más refinados. En muchos de los casos en que se puede aplicar tanto un método paramétrico como un método no paramétrico, el método no paramétrico es casi tan bueno o casi tan potente como el método paramétrico. En los casos en que los datos son nominales y ordinales o en los casos en que las suposiciones requeridas por los métodos paramétricos son inapropiadas, sólo se cuenta con los métodos no pa-

*Véase W. J. Conover, *Practical Nonparametric Statistics*, 3. ed. (John Wiley & Sons, 1998).

ramétricos. Debido a que en los métodos no paramétricos se requieren mediciones de los datos menos restrictivas y menos suposiciones acerca de la distribución de la población, se considera que tienen una aplicación más general que los métodos paramétricos. Los métodos no paramétricos que se presentan en este capítulo son la prueba de los signos, la prueba de los rangos con signo de Wilcoxon, la prueba de Mann-Whitney-Wilcoxon, la prueba de Kruskal-Wallis y la correlación de los rangos de Spearman.

19.1 Prueba de los signos

En una aplicación de investigación de mercado de la **prueba de los signos** se usa una muestra de n clientes potenciales para que indiquen su preferencia por una de dos marcas de un producto, por ejemplo, de un café, de un detergente o de un refresco. Las n expresiones de preferencia son datos nominales, ya que el consumidor simplemente nombra una preferencia. Dados estos datos, el objetivo es determinar si existe diferencia en las preferencias entre los dos artículos que se comparan. Como se verá, la prueba de los signos es un procedimiento estadístico no paramétrico para responder esta pregunta.

Caso de muestras pequeñas

El caso de la muestra pequeña es siempre que $n \leq 20$. A continuación, mediante un estudio realizado para Sun Coast Farms, se ilustra el uso de la prueba de los signos para el caso de una muestra pequeña; Sun Coast produce un jugo de naranja comercializado bajo el nombre Citrus Valley. Un competidor de Sun Coast Farms produce también un jugo de naranja que comercializa bajo el nombre de Tropical Orange. En un estudio acerca de las preferencias de los consumidores respecto a estas dos marcas, a 12 individuos se les dieron muestras, sin marca, de cada uno de los productos. La marca que cada individuo probó primero fue seleccionada aleatoriamente. Después de probar los dos productos, se pidió a estas personas que indicaran su preferencia por una de las dos marcas. En este estudio, el objetivo es ver si hay una preferencia de los consumidores por uno de los dos productos. Sea p la proporción de la población de consumidores que prefiere Citrus Valley; las hipótesis que se quiere probar son las siguientes.

TABLA 19.1

PROBABILIDADES BINOMIALES CON $n = 12, p = 0.50$

Número de signos más	Probabilidad
0	0.0002
1	0.0029
2	0.0161
3	0.0537
4	0.1208
5	0.1934
6	0.2256
7	0.1934
8	0.1208
9	0.0537
10	0.0161
11	0.0029
12	0.0002

$$H_0: p = 0.50$$
$$H_a: p \neq 0.50$$

Si no se rechaza H_0, no se tendrán evidencias que indiquen la existencia de alguna diferencia en las preferencias de los consumidores por estas dos marcas de jugos de naranja. Sin embargo, si se rechaza H_0, se podrá concluir que las preferencias de los consumidores hacia estas marcas son diferentes. En ese caso, la marca seleccionada por el mayor número de consumidores se considerará que es la marca preferida.

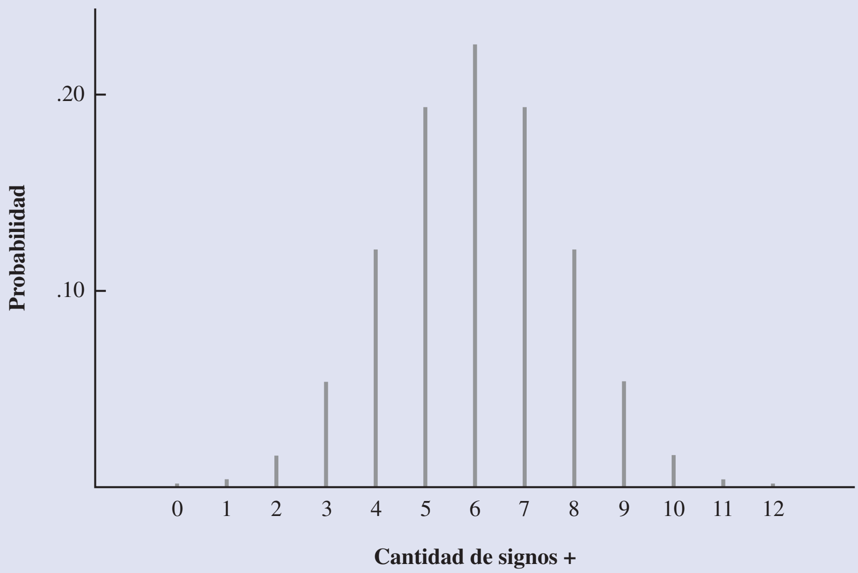
A continuación se muestra el uso de la versión para muestras pequeñas de la prueba de los signos al probar estas hipótesis para obtener una conclusión acerca de la preferencia de los consumidores. Para registrar los datos de la preferencia de los 12 individuos que participan en el estudio, se emplea un signo más si el individuo prefiere Citrus Valley y un signo menos si el individuo prefiere Tropical Orange. Debido a que los datos se registran en términos de signos más y menos, a esta prueba paramétrica se le conoce como prueba de los signos.

Las probabilidades binomiales exactas para tamaños de muestra menores o iguales a 20 se encuentran en la tabla 5 del apéndice B.

El número de signos más es el estadístico de prueba. Bajo la suposición de que H_0 es verdadera ($p = 0.50$), la distribución muestral del estadístico de prueba es una distribución binomial con $p = 0.50$. En la tabla 5 del apéndice B se encuentran las probabilidades de la distribución binomial para $n = 12$ y $p = 0.50$, las cuales se reproducen en la tabla 19.1. La figura 19.1 es la gráfica de esta distribución muestral binomial. En esta tabla se presenta la probabilidad para cada número de signos más, bajo la suposición de que H_0 es verdadera. Ahora se realiza la prueba para determinar si hay diferencia en las preferencias del público por estas marcas de jugos de naranja. Como nivel de significancia se usará 0.05.

En la tabla 19.2 se presentan los datos obtenidos sobre la preferencia. Los dos signos más, indican que dos consumidores prefirieron Citrus Valley. Ahora se pueden usar las distribuciones binomiales para determinar el valor-p de la prueba. Como es una prueba de dos colas, el valor-p se encuentra al duplicar la probabilidad en una cola de la distribución muestral binomial. El nú-

FIGURA 19.1 DISTRIBUCIÓN MUESTRAL BINOMIAL PARA EL NÚMERO DE SIGNOS MÁS CON $n = 12$ Y $p = 0.50$



mero de signos más para Sun Coast Farms (2) se encuentra en la cola inferior de la distribución. De manera que la probabilidad en esa cola es la probabilidad de 2, 1 y 0 signos más. Al sumar estas probabilidades se obtiene $0.0161 + 0.0029 + 0.0002 = 0.0192$. Al duplicar este valor se obtiene el valor-$p = 2(0.0192) = 0.0384$. Como el valor-$p \leq \alpha = 0.05$, se rechaza H_0. Esta prueba de sabores proporciona evidencias de que las preferencias de los consumidores difieren significativamente entre estas dos marcas de jugo de naranja. Se le informará a Sun Coast Farms que los consumidores prefieren Tropical Orange.

La prueba de hipótesis para Sun Coast Farms fue una prueba de dos colas. Como resultado el valor-p se halló al duplicar la probabilidad en una de las colas de la distribución binomial. También se puede hacer una prueba de signo de una cola. Si la prueba es de la cola inferior, el valor-p es la probabilidad de que el número de signos más sea menor o igual al número observado. Si la prueba es de la cola superior, el valor-p es la probabilidad de que el número de signos más sea mayor o igual al número observado.

TABLA 19.2 DATOS DE PREFERENCIAS EN LA PRUEBA DE SUN COAST FARMS

Individuo	Marca preferida	Dato registrado
1	Tropical Orange	−
2	Tropical Orange	−
3	Citrus Valley	+
4	Tropical Orange	−
5	Tropical Orange	−
6	Tropical Orange	−
7	Tropical Orange	−
8	Tropical Orange	−
9	Citrus Valley	+
10	Tropical Orange	−
11	Tropical Orange	−
12	Tropical Orange	−

En la prueba de sabores de Sun Coast Farms, los 12 individuos establecieron su preferencia por una de las dos marcas de jugo de naranja. En otras aplicaciones de la prueba de los signos, puede ocurrir que uno o más de los individuos de la muestra no puedan establecer su preferencia. Si una preferencia no puede ser establecida, se descarta de la muestra esa respuesta y la prueba de los signos se basará en la muestra de menor tamaño. Por último, las probabilidades binomiales que se presentan en la tabla 5 del apéndice B pueden usarse para pruebas de los signos con tamaños de muestras hasta $n = 20$. Para tamaños de muestra mayores, se usa la aproximación normal de las probabilidades binomiales.

Caso de muestras grandes

La prueba de los signos con muestras grandes es equivalente a una proporción poblacional con p = 0.50, como las presentadas en el capítulo 9.

Si la hipótesis nula es H_0: $p = 0.50$ y el tamaño de la muestra es $n > 20$, la distribución muestral del número de signos más se aproxima mediante una distribución normal.

APROXIMACIÓN NORMAL DE LA DISTRIBUCIÓN MUESTRAL DEL NÚMERO DE SIGNOS MÁS CUANDO H_0: $p = 0.50$

$$\text{Media: } \mu = 0.50n \tag{19.1}$$

$$\text{Desviación estándar: } \sigma = \sqrt{0.25n} \tag{19.2}$$

Forma de la distribución: aproximadamente normal siempre que $n > 20$.

A continuación se considera una aplicación de la prueba de los signos en la que se hace un sondeo político. En un sondeo realizado durante una campaña para elecciones presidenciales se pidió a 200 votantes registrados que evaluaran a los candidatos demócrata y republicano con relación a su política exterior. El resultado obtenido fue: 72 de los encuestados evaluaron mejor al candidato demócrata, 103 evaluaron mejor al republicano y 25 no encontraron diferencia entre los candidatos. ¿Con este sondeo puede observarse que exista una diferencia significativa, entre los candidatos, en términos de la opinión pública acerca de su política exterior?

Los casos de empate se eliminan del análisis

Se tiene que $n = 200 - 25 = 175$ fueron las personas que pudieron indicar qué candidato consideraban que tenía una mejor política exterior. Mediante la prueba de los signos y las ecuaciones (19.1) y (19.2) se puede hallar que la distribución muestral del número de signos más tiene las propiedades siguientes.

$$\mu = 0.50n = 0.50(175) = 87.5$$
$$\sigma = \sqrt{0.25n} = \sqrt{0.25(175)} = 6.6$$

Además, como $n = 175$, se puede asumir que la distribución muestral es aproximadamente normal. En la figura 19.2 se muestra esta distribución.

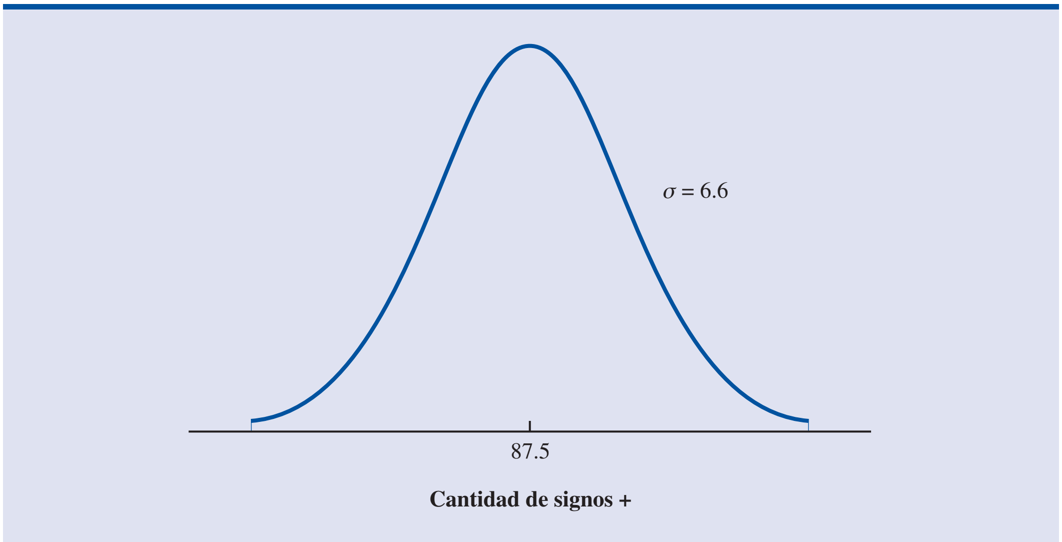
Ahora se procede a realizar la prueba de los signos con un nivel de significancia de 0.05, para obtener las conclusiones. Con base en el número de signos más ($x = 72$) que corresponden al número de personas que evaluaron como mejor la política exterior del candidato demócrata, se obtiene el valor siguiente para el estadístico de prueba

$$z = \frac{x - \mu}{\sigma} = \frac{72 - 87.5}{6.6} = -2.35$$

Si se usa el número de personas que evaluaron mejor al candidato republicano, z = 2.35, se llegará al mismo resultado.

En las tablas de probabilidad normal estándar se encuentra que el área en la cola, a la izquierda de $z = -2.35$ es 0.0094. Como se trata de una prueba de dos colas, el valor-p = 2(0.0094) = 0.0188. Como se obtiene que valor-$p \leq \alpha = 0.05$, se rechaza H_0. Como resultado de este estudio se encuentra que los candidatos difieren en términos de la opinión pública acerca de su política exterior.

FIGURA 19.2 DISTRIBUCIÓN DE PROBABILIDAD PARA EL NÚMERO DE SIGNOS MÁS EN UNA PRUEBA DE LOS SIGNOS EN LA QUE $n = 175$



Prueba de hipótesis acerca de la mediana

En el capítulo 9 se describió el uso de las pruebas de hipótesis para inferencias acerca de la media poblacional. Ahora se muestra cómo realizar una prueba de hipótesis acerca de la mediana poblacional. Recuerde que la mediana divide a la población de manera que 50% de los valores son mayores o iguales que la mediana y 50% de los valores son menores o iguales a la mediana. Cuando se utiliza la prueba de los signos se anota un signo más por cada dato muestral que sea mayor al valor de la mediana hipotética y un signo menos por cada dato muestral que sea menor al valor de la mediana hipotética. Los datos iguales al valor de la mediana hipotética, se descartan. Los cálculos en esta prueba de los signos se hacen igual.

Como ejemplo se realiza la siguiente prueba de hipótesis acerca del precio mediano de las casas nuevas en una determinada región.

$$H_0\text{: Mediana} = \$230\,000$$
$$H_a\text{: Mediana} \neq \$230\,000$$

En una muestra de 62 casas, 34 tuvieron un precio mayor al de la mediana, 26 tuvieron un precio menor al de la mediana y el precio de 2 de ellas fue exactamente $230 000.

Mediante las ecuaciones (19.1) y (19.2) con $n = 60$ casas, cuyos precios son diferentes a $230 000, se obtiene

$$\mu = 0.50\,n = 0.50(60) = 30$$
$$\sigma = \sqrt{0.25\,n} = \sqrt{0.25(60)} = 3.87$$

Como el número de signos más es $x = 34$, el estadístico de prueba es

$$z = \frac{x - \mu}{\sigma} = \frac{34 - 30}{3.87} = 1.03$$

Al aplicar las tablas de la probabilidad normal estándar con $z = 1.03$, se encuentra que el valor-p para dos colas es $2(1 - 0.8485) = 0.303$. Como el valor-$p > 0.05$, no se puede rechazar H_0. De acuerdo con los datos muestrales, no es posible rechazar la hipótesis nula que establece que el precio mediano de una casa nueva es $230 000.

Ejercicios

Métodos

Autoexamen

1. En la tabla siguiente se presentan las preferencias de 10 personas respecto a dos marcas de un producto.

Persona	Marca A frente a marca B	Persona	Marca A frente a marca B
1	+	6	+
2	+	7	−
3	+	8	+
4	−	9	−
5	+	10	+

Emplee $\alpha = 0.05$ y pruebe si existe alguna diferencia significativa en las preferencias por estas dos marcas. Un signo más indica preferencia por la marca A sobre la marca B.

Autoexamen

2. Realice la prueba de hipótesis siguiente.

$$H_0\text{: Mediana } \leq 150$$
$$H_a\text{: Mediana } > 150$$

En una muestra de tamaño 30 se obtuvieron 22 casos cuyo valor fue mayor que 150, tres cuyo valor fue exactamente 150 y cinco cuyo valor fue menor que 150. Con $\alpha = 0.01$ realice una prueba de hipótesis.

Aplicaciones

3. ¿Las divisiones de acciones son benéficas para los accionistas? La empresa SNL Securities estudió, a lo largo de 18 meses, las divisiones de acciones de la industria de la banca y encontró que las divisiones de las acciones tienden a incrementar el valor de las acciones de un individuo. Admita que en una muestra de 20 recientes divisiones de acciones, 14 hayan llevado a un aumento de su valor, cuatro hayan llevado a una disminución de su valor y dos no hayan ocasionado ningún cambio. Suponga que realiza un estudio para determinar si las divisiones de acciones aún benefician a los poseedores de acciones bancarias.
   a. ¿Cuáles son las hipótesis nula y alternativa?
   b. ¿A qué conclusión se llega con $\alpha = 0.05$?

Autoexamen

4. En un sondeo a 1 253 personas se les hizo una serie de preguntas acerca de la economía y del futuro de sus hijos. Una de las preguntas era, "¿Espera que sus hijos tengan una vida mejor a la que usted ha tenido, una vida peor o una vida más o menos igual de buena a la que usted ha tenido?" Las respuestas fueron, 34% mejor, 29% peor, 33% más o menos igual y 4% no supo contestar. Mediante la prueba de los signos y 0.05 como nivel de significancia, determine si el número de adultos que prevén un mejor futuro para sus hijos es mayor al número de adultos que prevén un futuro peor para sus hijos. ¿A qué conclusión llega?
5. La empresa Nielson Media Research identificó a *American Idol* y a *Dancing with the Stars* como los dos programas de televisión de mayor rating en febrero de 2006 (www.nielsenmedia.com, 10 de marzo de 2006). En un estudio local acerca del programa de televisión preferido, de 750 encuestados 330 votaron por *American Idol*, 270 por *Dancing with the Stars* y 150 por otro programa de televisión. Con 0.05 como nivel de significancia pruebe la hipótesis de que *American Idol* y *Dancing with the Stars* tienen el mismo nivel de preferencia. ¿A qué conclusión llega?

6. En el mercado de las computadoras personales la competencia es intensa. En una muestra de 500 compras, se encontró que 202 eran compras de la marca A, 158 de la marca B y 140 de otras marcas. Con un nivel de significancia de 0.05 pruebe la hipótesis de que las marcas A y B tienen la misma participación en el mercado de las computadoras personales. ¿Cuál es la conclusión?
7. El ingreso mediano anual de los suscriptores de la revista *Barron* es $131 000 (barrons.mag.com, 28 de julio de 2000). Suponga que en una muestra de 300 suscriptores a *The Wall Street Journal,* 165 suscriptores posean un ingreso mayor que $131 000 y 135 posean un ingreso menor que $131 000. ¿Puede concluir que hay diferencia entre los ingresos medianos de los dos grupos de suscriptores? Emplee $\alpha = 0.05$ como nivel de significancia, ¿a qué conclusión llega?
8. En una muestra de 150 partidos de básquetbol universitario, el equipo de casa ganó 98 partidos. Realice una prueba para determinar si los datos sustentan la hipótesis de que en el básquetbol universitario el equipo de casa tiene ventaja. ¿A qué conclusión llega con $\alpha = 0.05$?
9. El año pasado, en una determinada ciudad, la mediana del número de empleados de tiempo parcial en un restaurante de comida rápida era 15. Es posible que esta cantidad esté aumentando. En una muestra de nueve restaurantes de comida rápida se encontró que en siete de ellos trabajaban más de 15 empleados de tiempo parcial, en uno había exactamente 15 empleados que trabajaban de tiempo parcial y en otro más había menos de 15 empleados que trabajaban de tiempo parcial. Realice una prueba con $\alpha = 0.05$ para determinar si el número mediano de empleados que trabaja de tiempo parcial ha aumentado.
10. De acuerdo con un estudio nacional, el ingreso anual mediano que los adultos dicen haría realidad sus sueños es $152 000. Suponga que en Ohio, de 225 personas tomadas en una muestra, 122 indican que el ingreso necesario para hacer realidad sus sueños sea menor que $152 000, y 103 informen que esta cantidad sea mayor que $152 000. Pruebe la hipótesis nula de que en Ohio, el ingreso medio anual para que una persona haga realidad sus sueños es $152 000. Use $\alpha = 0.05$. ¿A qué conclusión llega?
11. El ingreso medio anual de los estudiantes con una licenciatura (en Estados Unidos) es $37 700 (*The New York Times Almanac*, 2006). A continuación se presentan los datos muestrales (en miles de dólares) de estudiantes universitarios en la zona de Chicago. Con los datos muestrales pruebe H_0: mediana ≤ 37.7 y H_a: mediana > 37.7 para la población de estudiantes con grado de licenciatura que trabajan en la zona de Chicago. Use $\alpha = 0.05$ como nivel de significancia. ¿Cuál es su conclusión?

archivo en **CD**
Annual

47.8	41.7	31.4	56.9	55.2
47.2	42.6	105.3	38.8	30.0
55.5	127.8	73.7	25.2	68.4
41.2	45.7	37.7	30.4	91.1
21.3	42.4	61.2	23.8	34.1
42.4	25.0	43.2	36.2	76.7
51.9	25.3	39.3	65.0	38.0
32.8	24.4	69.0	25.1	48.7
30.2	60.6	43.4	34.9	37.7
38.5	31.1	91.0	23.6	56.1

19.2 Prueba de los rangos con signo de Wilcoxon

La **prueba de los rangos con signo de Wilcoxon** es la alternativa no paramétrica al método paramétrico de las muestras por pares (o apareadas) presentado en el capítulo 10. En la situación de las muestras por pares, cada unidad experimental genera dos observaciones, una correspondiente a la población 1 y otra correspondiente a la población 2. Las diferencias entre los pares de observaciones permiten apreciar la diferencia entre las dos poblaciones.

En una fábrica se desea determinar cuál de dos métodos de producción difiere en el tiempo que se requiere para realizar una tarea. Se selecciona una muestra de 11 trabajadores y cada trabajador realiza la tarea con uno de estos dos métodos de producción. El método de producción

TABLA 19.3 TIEMPO EN MINUTOS PARA LA REALIZACIÓN DE UNA TAREA DE PRODUCCIÓN

	Método		
Trabajador	**1**	**2**	**Diferencia**
1	10.2	9.5	0.7
2	9.6	9.8	−0.2
3	9.2	8.8	0.4
4	10.6	10.1	0.5
5	9.9	10.3	−0.4
6	10.2	9.3	0.9
7	10.6	10.5	0.1
8	10.0	10.0	0.0
9	11.2	10.6	0.6
10	10.7	10.2	0.5
11	10.6	9.8	0.8

que usa primero cada trabajador es seleccionado de manera aleatoria. De manera que cada uno de los trabajadores de la muestra proporciona un par de observaciones como aparece en la tabla 19.3. Una diferencia positiva entre los tiempos de realización de la tarea indica que el método 1 requiere más tiempo, y una diferencia negativa entre los tiempos indica que el método 2 requiere más tiempo. ¿Los datos obtenidos indican que estos métodos son significativamente diferentes en términos del tiempo que se requiere para realizar la tarea?

En efecto, se tienen dos poblaciones de tiempos requeridos para realizar una tarea, cada población corresponde a cada uno de los métodos; las hipótesis a probar son las siguientes.

H_0: Las poblaciones son idénticas

H_a: Las poblaciones no son idénticas

Si no se puede rechazar H_0, no se contará con evidencia para concluir que los dos métodos difieren en los tiempos requeridos para realizar la tarea. Pero, si H_0 puede ser rechazada, se concluirá que los dos métodos difieren en los tiempos para realizar la tarea.

El primer paso en la prueba de los rangos con signo de Wilcoxon es ordenar los *valores absolutos* de las diferencias entre los dos métodos y asignarles un rango. Toda diferencia que sea igual a cero se descarta y las diferencias restantes se ordenan y se les asigna un rango. A las diferencias que tengan un mismo valor, el rango que se les asigna es el promedio de los números de sus posiciones en el conjunto de datos ordenados. En la última columna de la tabla 19.4 se muestran los rangos asignados a los valores absolutos de las diferencias. Observe que la diferencia cero obtenida por el trabajador 8 se descarta; después, a la diferencia absoluta más pequeña, que es 0.1 se le asigna el rango 1. Se continúa ordenando las diferencias absolutas hasta asignarle a la mayor diferencia absoluta, que es 0.9, el rango 10. El rango que se le asigna a cada una de las diferencias absolutas de los trabajadores 3 y 5, que son iguales, es el promedio de sus posiciones en el conjunto ordenado de las diferencias absolutas 3.5 y el rango para cada una de las diferencias absolutas iguales de los trabajadores 4 y 10 es el promedio de las posiciones que les corresponden en el conjunto ordenado de los datos, 5.5.

Una vez determinados los rangos de las diferencias absolutas, se les antepone el signo de la diferencia original entre los datos. Por ejemplo, a la diferencia 0.1 del trabajador 7, a la que se le ha asignado el rango 1 se le da el valor $+1$, ya que la diferencia observada entre los dos métodos es positiva. A la diferencia 0.2, que se le asignó el rango 2, se le da el valor -2 ya que la diferencia observada entre los dos métodos es negativa. En la última columna de la tabla 19.4 se encuentra la lista completa de todos los rangos así como la suma de todos ellos.

Ahora se vuelve a la hipótesis original de que las poblaciones de los tiempos necesarios para realizar la tarea, con cada uno de estos dos métodos, son iguales. Si las poblaciones que re-

TABLA 19.4 RANGOS DE LAS DIFERENCIAS ABSOLUTAS ACERCA DEL TIEMPO NECESARIO PARA REALIZAR UNA TAREA DE PRODUCCIÓN

Trabajador	Diferencia	Valor absoluto de la diferencia	Rango	Rango con signo
1	0.7	0.7	8	+ 8
2	−0.2	0.2	2	+ 2
3	0.4	0.4	3.5	+ 3.5
4	0.5	0.5	5.5	+ 5.5
5	−0.4	0.4	3.5	− 3.5
6	0.9	0.9	10	+10
7	0.1	0.1	1	+ 1
8	0.0	0.0	—	—
9	0.6	0.6	7	+ 7
10	0.5	0.5	5.5	+ 5.5
11	0.8	0.8	9	+ 9
			Suma de los rangos con signo	+44.0

presentan los tiempos requeridos para realizar la tarea con cada uno de los métodos fueran idénticas, se esperaría que los rangos positivos y los rangos negativos se compensaran unos con otros y se anularan, de manera que la suma de los valores de los rangos con signo sería aproximadamente cero. Por tanto, en la prueba de los rangos con signo de Wilcoxon, la prueba de significancia consiste en determinar si la suma de los rangos con signo (+44 en este caso) es significativamente distinta de cero.

Sea T la suma de los valores de los rangos con signo en una prueba de los rangos con signo de Wilcoxon. Si las dos poblaciones son idénticas y si el número de pares de datos es 10 o mayor, es posible demostrar que la distribución muestral de T puede ser aproximada mediante una distribución normal.

DISTRIBUCIÓN MUESTRAL DE T PARA POBLACIONES IDÉNTICAS

$$\text{Media: } \mu_T = 0 \tag{19.3}$$

$$\text{Desviación estándar: } \sigma_T = \sqrt{\frac{n(n+1)(2n+1)}{6}} \tag{19.4}$$

Forma de la distribución: aproximadamente normal siempre que $n \geq 10$.

En el ejemplo, después de descartar la observación en que la diferencia es cero (la del trabajador 8), se tiene $n = 10$. Por tanto, si emplea la ecuación (19.4), tiene

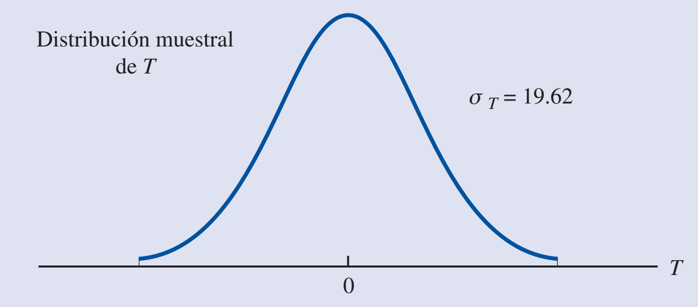
$$\sigma_T = \sqrt{\frac{10(11)(21)}{6}} = 19.62$$

En la figura 19.3 se presenta la distribución muestral de T bajo la suposición de que las dos poblaciones son idénticas. Ahora se procede a realizar la prueba de los rangos con signo de Wilcoxon con 0.05 como nivel de significancia, para llegar a una conclusión. Con la suma de los valores de los signos con rango $T = 44$, se obtiene el valor siguiente para el estadístico de prueba.

$$z = \frac{T - \mu_T}{\sigma_T} = \frac{44 - 0}{19.62} = 2.24$$

A partir de las tablas de probabilidad normal estándar y $z = 2.24$, se halla que para dos colas el valor-$p = 2(1 - 0.9875) = 0.025$. Como el valor-$p \leq \alpha = 0.05$, se rechaza H_0 y se concluye que

FIGURA 19.3 DISTRIBUCIÓN MUESTRAL DE LA *T* DE WILCOXON DEL TIEMPO PARA LA REALIZACIÓN DE UNA TAREA DE PRODUCCIÓN



las dos poblaciones no son idénticas y que los métodos difieren en el tiempo requerido para realizar la tarea. Como 8 trabajadores obtuvieron tiempos más cortos con el método 2, se concluye que el método 2 es el método de producción que se preferirá.

Ejercicios

Aplicaciones

Auto examen

12. Con objeto de determinar su efecto en el rendimiento de la gasolina en millas por galón en los automóviles de pasajeros, se prueban dos aditivos para gasolina. A continuación aparecen los resultados de esta prueba en 12 automóviles; en cada automóvil se probaron los dos aditivos. Use $\alpha = 0.05$ y la prueba de los rangos con signo de Wilcoxon para determinar si existe una diferencia significativa entre estos dos aditivos.

	Aditivo			Aditivo	
Automóvil	**1**	**2**	**Automóvil**	**1**	**2**
1	20.12	18.05	7	16.16	17.20
2	23.56	21.77	8	18.55	14.98
3	22.03	22.57	9	21.87	20.03
4	19.15	17.06	10	24.23	21.15
5	21.23	21.22	11	23.21	22.78
6	24.77	23.80	12	25.02	23.70

Auto examen

13. Para medir el tiempo que necesitaban para quedarse dormidos, un estudio probó el efecto de un relajante para hombres. Los datos siguientes corresponden a los minutos que requirieron cada uno de los 10 hombres de la muestra para quedarse dormidos. Use como nivel de significancia $\alpha = 0.05$ y determine si el relajante reduce el tiempo que se requiere para quedarse dormido. ¿Cuál es su conclusión?

Sujeto	Sin relajante	Con relajante	Sujeto	Sin relajante	Con relajante
1	15	10	6	7	5
2	12	10	7	8	10
3	22	12	8	10	7
4	8	11	9	14	11
5	10	9	10	9	6

14. En 10 de los principales aeropuertos se muestrearon los precios de la gasolina para automóviles rentados. A continuación se presentan los datos correspondientes a las empresas Avis y Budget (*USA Today,* 4 de abril de 2000).

Aeropuerto	Avis	Budget
Boston Logan	1.58	1.39
Chicago O'Hare	1.60	1.55
Chicago Midway	1.53	1.55
Denver	1.55	1.51
Fort Lauderdale	1.57	1.58
Los Ángeles	1.80	1.74
Miami	1.62	1.60
Nueva York (JFK)	1.69	1.60
Orange County, CA	1.75	1.59
Washington (Dulles)	1.55	1.54

Use $\alpha = 0.05$ para probar la hipótesis de que no hay diferencia entre las dos poblaciones. ¿Cuál es su conclusión?

15. Dos servicios nocturnos de paquetería fueron probados; se formaron dos muestras idénticas, de manera que a los dos servicios de paquetería se les notificara al mismo tiempo que se requerían sus servicios. A continuación se presentan los tiempos requeridos en cada entrega. ¿Estos datos sugieren que existe diferencia entre los tiempos que requiere cada uno de estos servicios?

	Servicio	
Entrega	1	2
1	24.5	28.0
2	26.0	25.5
3	28.0	32.0
4	21.0	20.0
5	18.0	19.5
6	36.0	28.0
7	25.0	29.0
8	21.0	22.0
9	24.0	23.5
10	26.0	29.5
11	31.0	30.0

16. El campeonato de los jugadores de la PGA tuvo lugar, del 23 al 26 de marzo de 2006, en el campo de golf TPC Sawgrass en Ponte Vedra Beach, Florida. A continuación se presentan las puntuaciones obtenidas, en la primera y segunda rondas, por 11 golfistas de una muestra. Use $\alpha = 0.05$ y determine si existe una diferencia significativa entre las puntuaciones obtenidas por los golfistas en la primera y en la segunda rondas. ¿Cuál es su conclusión?

Golfista	Primera ronda	Segunda ronda
Fred Couples	69	73
John Daly	70	73
Ernie Els	72	70
Jim Furyk	65	71
Phil Mickelson	70	73
Rocco Mediate	69	74
Nick Price	72	71
Vijay Singh	68	70
Sergio Garcia	70	68
Mike Weir	71	71
Tiger Woods	72	69

17. Como parte de una investigación de mercado que tenía por objeto evaluar la efectividad de una campaña de publicidad, se seleccionaron 10 ciudades para una prueba de mercado. Las ventas en dólares en cada una de estas ciudades, en la semana anterior a la campaña, se registraron. Después, se realizó la campaña durante dos semanas y se registraron las ventas que hubo en la primera semana, inmediatamente después de la campaña.

Ciudad	Ventas antes de la campaña	Ventas después de la campaña
Kansas City	130	160
Dayton	100	105
Cincinnati	120	140
Columbus	95	90
Cleveland	140	130
Indianapolis	80	82
Louisville	65	55
St. Louis	90	105
Pittsburgh	140	152
Peoria	125	140

Use $\alpha = 0.05$. ¿A qué conclusión llega acerca del valor de la campaña?

19.3 Prueba de Mann-Whitney-Wilcoxon

En esta sección se presenta otro método no paramétrico que se usa para determinar si hay diferencia entre dos poblaciones. Esta prueba, a diferencia de la prueba de los rangos con signo, no se basa en una muestra por pares. Aquí se usan dos muestras independientes, una de cada población. Esta prueba fue creada conjuntamente por Mann, Whitney y Wilcoxon. Algunas veces se le llama *prueba de Mann-Whitney* y otras veces *prueba de la suma de rangos de Wilcoxon.* Las dos versiones de esta prueba, la de Mann-Whitney y la de Wilcoxon son equivalentes. Aquí se le llamará **prueba de Mann-Whitney-Wilcoxon (MWW)**.

La prueba no paramétrica de MWW no requiere que los datos sean de intervalo ni tampoco que las poblaciones estén distribuidas normalmente. El único requisito es que la escala de medición de los datos sea por lo menos ordinal. Después, en lugar de probar las diferencias entre las medias de las dos poblaciones, la prueba de MWW determina si las dos poblaciones son idénticas. Las hipótesis en la prueba de MWW son las siguientes.

H_0: Las dos poblaciones son idénticas
H_a: Las dos poblaciones no son idénticas

Caso de muestras pequeñas

La prueba de MWW para el caso de muestras pequeñas se usa siempre que los tamaños de las muestras de ambas poblaciones sean menores o iguales a 10. El uso de la prueba de MWW para muestras pequeñas se ilustrará mediante un ejemplo sobre la preparación académica de los alumnos de la escuela Johnston. La mayoría de los alumnos de la escuela Johnston provienen de la escuela Garfield o de la escuela Mulberry. La cuestión que desean resolver los directivos de la escuela Johnston es si la población de los estudiantes que provenían de la escuela Garfield es idéntica, en términos de preparación académica, a la población de los estudiantes que provenían de la escuela Mulberry. Las hipótesis son las siguientes.

H_0: Las dos poblaciones son idénticas en términos de preparación académica

H_0: Las dos poblaciones no son idénticas en términos de preparación académica

TABLA 19.5 DATOS DE NIVEL ACADÉMICO

Escuela Garfield		Escuela Mulberry	
Estudiante	**Nivel académico**	**Estudiante**	**Nivel académico**
Fields	8	Hart	70
Clark	52	Phipps	202
Jones	112	Kirkwood	144
Tibbs	21	Abbott	175
		Guest	146

Los directivos de la escuela Johnston toman una muestra aleatoria de cuatro estudiantes provenientes de la escuela Garfield y otra muestra aleatoria de cinco estudiantes provenientes de la escuela Mulberry. De cada uno de los nueve estudiantes tomados para el estudio se registra su actual nivel académico. En la tabla 19.5 se presentan los niveles académicos de estos nueve estudiantes.

El primer paso en la prueba de MWW es *reunir en un solo conjunto* todos los datos y ordenarlos de menor a mayor. Al valor menor (nivel académico 8) se le da el rango 1 y al valor mayor (nivel académico 202) se le da el rango 9. En la tabla 19.6 se presentan los nueve estudiantes con sus rangos y ordenados de acuerdo con ellos.

El paso siguiente es sumar los rangos de cada muestra, por separado. Esto se muestra en la tabla 19.7. En la prueba de MWW se puede usar la suma de cualquiera de las muestras. Aquí se usará la suma de la muestra de los cuatro estudiantes de la escuela Garfield. Esta suma se denota con el símbolo T. De manera que en este ejemplo, $T = 11$.

¿Cuáles son las propiedades de la suma de los rangos en la muestra de Garfield? Puede ocurrir que los cuatro estudiantes en la muestra de Garfield sean los cuatro estudiantes que tengan los primeros rangos en este estudio, si este fuera el caso, $T = 1 + 2 + 3 + 4 = 10$ sería el menor valor que podría tener T, la suma de los rangos. Pero también puede ocurrir que los estudiantes de Garfield fueran los cuatro estudiantes que obtuvieran los últimos rangos, en cuyo caso $T = 6 + 7 + 8 + 9 = 30$ sería el mayor valor que podría tomar T. Por tanto, en la muestra de la escuela Garfield, el valor T estará entre 10 y 30.

Observe que valores de T cercanos a 10 significan que la escuela Garfield tiene los estudiantes significativamente mejores, o con rangos más altos, mientras que valores de T cercanos a 30 significan que la escuela Garfield tiene los estudiantes significativamente peores, o con rangos más bajos. Por tanto, si las dos poblaciones de estudiantes fueran idénticas, en términos de preparación académica, se esperaría que los valores de T fueran aproximadamente iguales al promedio de estos dos valores, o sea $(10 + 30)/2 = 20$.

En la tabla 8 del apéndice B se presentan los valores críticos para el estadístico T en la prueba de MWW para el caso en que los tamaños de ambas muestras son menores o iguales a 10. En esta tabla n_1 corresponde al tamaño de la muestra cuya suma de los rangos se está empleando en la prueba. El valor de T_L se lee directamente en la tabla y el valor de T_U se calcula con la ecuación (19.5).

TABLA 19.6 ESTUDIANTES ORDENADOS POR RANGOS

Estudiante	**Nivel académico**	**Rangos dados a las dos muestras juntas**	**Estudiante**	**Nivel académico**	**Rangos dados a las dos muestras juntas**
Fields	8	1	Kirkwood	144	6
Tibbs	21	2	Guest	146	7
Clark	52	3	Abbott	175	8
Hart	70	4	Phipps	202	9
Jones	112	5			

TABLA 19.7 SUMAS DE LOS RANGOS DE LOS ESTUDIANTES PROVENIENTES DE CADA UNA DE LAS ESCUELAS

Estudiantes de la escuela Garfield			Estudiantes de la esucela Mulberry		
Estudiante	**Nivel académico**	**Rango en la muestra**	**Estudiante**	**Nivel académico**	**Rango en la muestra**
Fields	8	1	Hart	70	4
Clark	52	3	Phipps	202	9
Jones	112	5	Kirkwood	144	6
Tibbs	21	2	Abbott	175	8
			Guest	146	7
Suma de los rangos		11			34

Ni los valores de T_L ni los de T_U se encuentran en la zona de rechazo. La hipótesis nula de que las poblaciones son idénticas debe rechazarse sólo si T es estrictamente menor que T_L o estrictamente mayor que T_U.

Por ejemplo, para el nivel de significancia 0.05, en la tabla 8 del apéndice B se encuentra que el valor crítico en la cola inferior para el estadístico de prueba en la prueba de MWW con $n_1 = 4$ (Garfield) y $n_2 = 5$ (Mulberry) es $T_L = 12$. El valor crítico en la cola superior para el estadístico de prueba en la prueba de MWW obtenido con la ecuación (19.5) es

$$T_U = 4(4 + 5 + 1) - 12 = 28$$

En consecuencia, la regla de decisión de esta prueba de MWW indica que la hipótesis nula de que las poblaciones son idénticas puede rechazarse si la suma de los rangos de la primera muestra (Garfield) es menor que 12 o mayor que 28. La regla de rechazo puede expresarse como

$$\text{Rechazar } H_0 \text{ si } T < 12 \text{ o } T > 28$$

Al realizar la prueba con la suma de los rangos de los estudiantes de la escuela Mulberry, se tiene $n_1 = 5$, $n_2 = 4$, $T_L = 17$, $T_U = 33$ y $T = 34$. Como $T > T_U$, se llega también a la misma conclusión, rechazar H_0.

En la tabla 19.7 se ve que $T = 11$. Por tanto, se rechaza la hipótesis H_0 y se concluye que la población de los estudiantes de la escuela Garfield es diferente de la población de los estudiantes de Mulberry en términos de preparación académica. Como los estudiantes de la escuela Garfield obtuvieron las mejores puntuaciones académicas, eso indica que los estudiantes de la escuela Garfield están mejor preparados que los estudiantes de la escuela Mulberry.

Caso de muestras grandes

Cuando los tamaños de las dos muestras son mayores o iguales a 10, para realizar la prueba de MWW se puede usar la aproximación normal de la distribución de T. Para ilustrar este caso de las muestras grandes se empleará un ejemplo del Third National Bank.

El Third National Bank tiene dos sucursales. En la tabla 19.8 se presentan los datos obtenidos de dos muestras aleatorias independientes, una de cada sucursal. ¿Estos datos muestran que son idénticas las poblaciones de saldos en las cuentas de cheques de las dos sucursales?

El primer paso en la prueba de MWW es *reunir en un solo conjunto,* todos los datos y ordenarlos de menor a mayor. En la tabla 19.8, de los 22 datos, se observa que el menor es $750 (sexto elemento de la muestra 2), a este dato se le asigna el rango 1. Al terminar con la asignación de rangos se llega a la lista siguiente.

Saldo ($)	Elemento	Rango asignado
750	6o. de la muestra 2	1
800	5o. de la muestra 2	2
805	7o. de la muestra 1	3

(*continúa*)

Saldo ($)	Elemento	Rango asignado
850	2o. de la muestra 2	4
.	.	.
.	.	.
.	.	.
1195	4o. de la muestra 1	21
1200	3o. de la muestra 1	22

Al ordenar los datos, una vez *reunidos en un solo conjunto*, puede ocurrir que los valores de dos o más datos sean iguales. En este caso, el rango que se le asigna a cada uno es el *promedio* de sus posiciones en el conjunto de los datos ordenados. Por ejemplo, al saldo $945 (octavo elemento de la muestra 1) se le asignará el rango 11. Pero, los siguientes dos datos del conjunto son iguales, su valor es $950 (sexto elemento de la muestra 1 y cuarto elemento de la muestra 2); a cada uno de estos dos datos, que les corresponden las posiciones 12 y 13, el rango que se les asigna es 12.5. Al siguiente dato cuyo valor es $955, continuando con el proceso de asignación de rangos, se le asigna el rango 14. En la tabla 19.9 se presenta el conjunto de datos con los rangos asignados a cada observación.

El paso siguiente en la prueba de MWW es sumar los rangos de cada muestra. Estas sumas se presentan en la tabla 19.9. La prueba se puede basar en la suma de los rangos de cualquiera de las muestras. Aquí se usará la suma de los rangos de la sucursal 1. Así, en este ejemplo, $T = 169.5$.

Dado que los tamaños de las muestras son $n_1 = 12$ y $n_2 = 10$ se puede usar la aproximación normal de la distribución muestral de la suma de los rangos T. La distribución muestral está determinada por las expresiones siguientes.

DISTRIBUCIÓN MUESTRAL DE T PARA POBLACIONES IDÉNTICAS

$$\text{Media: } \mu_T = \tfrac{1}{2} n_1(n_1 + n_2 + 1) \tag{19.6}$$

$$\text{Desviación estándar: } \sigma_T = \sqrt{\tfrac{1}{12} n_1 n_2(n_1 + n_2 + 1)} \tag{19.7}$$

Forma de la distribución: aproximadamente normal siempre que $n_1 \geq 10$ y $n_2 \geq 10$.

TABLA 19.8 SALDOS EN LAS CUENTAS DE DOS SUCURSALES DEL BANCO THIRD NATIONAL BANK

Sucursal 1		Sucursal 2	
Cuenta	Saldo ($)	Cuenta	Saldo ($)
1	1095	1	885
2	955	2	850
3	1200	3	915
4	1195	4	950
5	925	5	800
6	950	6	750
7	805	7	865
8	945	8	1000
9	875	9	1050
10	1055	10	935
11	1025		
12	975		

TABLA 19.9 RANGOS CORRESPONDIENTES A LOS DATOS (REUNIDOS EN UN SOLO CONJUNTO) DE LAS DOS MUESTRAS DEL THIRD NATIONAL BANK

	Sucursal 1			Sucursal 2	
Cuenta	**Saldo ($)**	**Rango**	**Cuenta**	**Saldo ($)**	**Rango**
1	1095	20	1	885	7
2	955	14	2	850	4
3	1200	22	3	915	8
4	1195	21	4	950	12.5
5	925	9	5	800	2
6	950	12.5	6	750	1
7	805	3	7	865	5
8	945	11	8	1000	16
9	875	6	9	1050	18
10	1055	19	10	935	10
11	1025	17		Suma de rangos	83.5
12	975	15			
	Suma de rangos	169.5			

Para la sucursal 1 se tiene

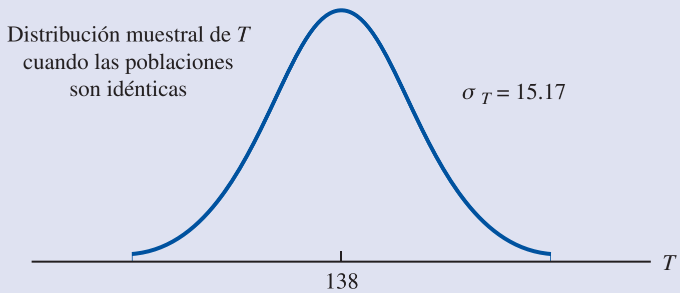
$$\mu_T = \tfrac{1}{2}\, 12(12 + 10 + 1) = 138$$
$$\sigma_T = \sqrt{\tfrac{1}{12}\, 12(10)(12 + 10 + 1)} = 15.17$$

La figura 19.4 es la distribución muestral de T. Ahora se procede a realizar la prueba de MWW con un nivel de significancia, para llegar a una conclusión, 0.05. Como para la sucursal 1, la suma de los rangos es $T = 169.5$, el valor del estadístico de prueba es el siguiente.

$$z = \frac{T - \mu_T}{\sigma_T} = \frac{169.5 - 138}{15.17} = 2.08$$

En la tabla de la distribución normal estándar, dado que $z = 2.08$, se encuentra que el valor-p para las dos colas es $2(1 - 0.9812) = 0.376$. Como el valor-$p \leq \alpha = 0.05$, se rechaza H_0 y se

FIGURA 19.4 DISTRIBUCIÓN MUESTRAL DE T PARA EL EJEMPLO DE THIRD NATIONAL BANK



concluye que estas dos poblaciones no son idénticas; las poblaciones de los saldos en las cuentas de las dos sucursales no son una misma población.

En resumen, la prueba de la suma de los rangos de Man-Whitney-Wilcoxon para determinar si dos muestras aleatorias independientes pertenecen a poblaciones idénticas consiste en los pasos siguientes.

1. Reunir en un solo conjunto las observaciones muestrales y ordenarlas de menor a mayor al asignarles un rango; a las observaciones muestrales que tengan un mismo valor se les asigna, a cada una, el promedio de los lugares que les corresponden en la lista ordenada de menor a mayor.
2. Calcular T, la suma de los rangos de la primera muestra.
3. En el caso de muestras grandes, para probar si existen diferencias significativas entre las dos poblaciones, el valor obtenido para T se compara con la distribución muestral de T para poblaciones idénticas con las ecuaciones (19.6) y (19.7). Para decidir si se rechaza H_0 se emplea el valor del estadístico de prueba estandarizado z y el valor-p. En el caso de muestras pequeñas, se usa la tabla 9 del apéndice B para hallar los valores críticos para la prueba.

NOTAS Y COMENTARIOS

La prueba no paramétrica vista en esta sección se utiliza para determinar si dos poblaciones son idénticas. Con las pruebas estadísticas paramétricas vistas en el capítulo 10 se prueba la igualdad de dos medias poblacionales. Cuando se rechaza la hipótesis de que las medias sean iguales, se concluye que las poblaciones difieren en sus medias. En la prueba de MWW, cuando se rechaza la hipótesis de que las poblaciones sean idénticas, no se puede decir en qué difieren. Las poblaciones pueden tener diferentes medias, diferentes medianas, diferentes varianzas o diferentes formas. No obstante, si se considera que las poblaciones son iguales en todos los aspectos con excepción de las medias, rechazar H_0 mediante este método no paramétrico implica que las medias son diferentes.

Ejercicios

Aplicaciones

Auto examen

18. Para probar el efecto de dos aditivos sobre el rendimiento de la gasolina, siete automóviles usan el aditivo 1 y nueve automóviles el aditivo 2. En los datos siguientes se presenta el rendimiento en millas por galón obtenido con cada uno de los dos aditivos. Use $\alpha = 0.05$ y la prueba de MWW para determinar si existe una diferencia significativa en el efecto que tienen los dos aditivos sobre el rendimiento.

Aditivo 1	Aditivo 2
17.3	18.7
18.4	17.8
19.1	21.3
16.7	21.0
18.2	22.1
18.6	18.7
17.5	19.8
	20.7
	20.2

Auto examen

19. A continuación se presentan los datos muestrales de los salarios iniciales de contadores públicos y planificadores financieros. Los salarios anuales están dados en miles de dólares.

Contador público	Planificador financiero	Contador público	Planificador financiero
45.2	44.0	50.0	48.6
53.8	44.2	45.9	44.7
51.3	48.1	54.5	48.9
53.2	50.9	52.0	46.8
49.2	46.9	46.9	43.9

a. Use 0.05 como nivel de significancia y pruebe la hipótesis de que no hay diferencia entre los salarios anuales iniciales de los contadores públicos y de los planificadores financieros.
b. Proporcione las medias muestrales de los salarios iniciales en estas dos profesiones.

20. La brecha entre los salarios de hombres y mujeres con la misma preparación disminuye cada vez más, pero aún no se ha cerrado totalmente. A continuación se presentan datos muestrales de siete hombres y siete mujeres con licenciatura. Los datos se dan en miles de dólares.

Hombre	30.6	75.5	45.2	62.2	38.2	49.9	55.3
Mujer	44.5	35.4	27.9	40.5	25.8	47.5	24.8

a. ¿Cuál es el salario mediano de los hombres? ¿Cuál el de las mujeres?
b. Use $\alpha = 0.05$ y realice una prueba de hipótesis para determinar si las dos poblaciones son iguales. Dé su conclusión.

21. Cada año, en diciembre, NRF/BIG Research realiza un estudio sobres el gasto que hacen las personas en las vacaciones de invierno. A continuación se presentan los datos muestrales sobre el gasto en las vacaciones de invierno en 2004 y 2005 (*USA Today*, 20 de diciembre de 2005).

2004	2005
623	752
687	582
748	781
638	805
713	723
645	728
726	674
700	766
794	908
662	737
814	796
674	724

a. Use $\alpha = 0.05$ y realice una prueba para determinar si en 2005 hubo un incremento en comparación con 2004. ¿Cuál es su conclusión?
b. Para cada uno de estos años calcule la media muestral del gasto en vacaciones. Dé el porcentaje en que aumentó o disminuyó el gasto en 2005.

22. *Business Week* publica estadísticas anuales sobre las 1 000 empresas más grandes. El cociente P/E (cociente de rendimiento por acción) de una empresa es el precio actual de las acciones de la empresa dividido entre la ganancia por acción en los últimos 12 meses. En la tabla 19.10 se presenta el cociente P/E de 10 empresas japonesas y 12 empresas estadounidenses de una muestra. ¿Es significativa la diferencia entre los dos países? Use la prueba de MWW y $\alpha = 0.01$ para dar sus conclusiones.

TABLA 19.10 COCIENTE P/E DE ALGUNAS EMPRESAS JAPONESAS Y ESTADOUNIDENSES

Japón		Estados Unidos	
Empresa	**Cociente P/E**	**Empresa**	**Cociente P/E**
Sumitomo Corp.	153	Gannet	19
Kinden	21	Motorola	24
Heiwa	18	Schlumberger	24
NCR Japan	125	Oracle Systems	43
Suzuki Motor	31	Gap	22
Fuji Bank	213	Winn-Dixie	14
Sumitomo Chemical	64	Ingersoll-Rand	21
Seibu Railway	666	American Electric Power	14
Shiseido	33	Hercules	21
Toho Gas	68	Times Mirror	38
		WellPoint Health	15
		Northern States Power	14

23. Los números de delitos por día reportados a la policía durante el verano y el invierno son los siguientes. Use 0.05, como nivel de significancia, para determinar si existe una diferencia significativa entre verano e invierno, en términos del número de crímenes reportados.

Invierno	Verano
18	28
20	18
15	24
16	32
21	18
20	29
12	23
16	38
19	28
20	18

24. Los hornos de microondas de una determinada marca se venden en Dallas y en San Antonio. Los precios se presentan a continuación. Use $\alpha = 0.05$ y pruebe si los precios en Dallas y en San Antonio son los mismos.

Dallas	San Antonio
445	460
489	451
405	435
485	479
439	475
449	445
436	429
420	434
430	410
405	422
	425
	459
	430

25. La National Association of Home Buiders proporciona datos sobre los más frecuentes proyectos de remodelación. Use la prueba de MWW para determinar si se puede concluir que el costo de remodelación de una cocina difiera del costo de remodelación de una recámara. Use 0.05 como nivel de significancia.

Cocina	Recámara
25 200	18 000
17 400	22 900
22 800	26 400
21 900	24 800
19 700	26 900
23 000	17 800
19 700	24 600
16 900	21 000
21 800	
23 600	

19.4 Prueba de Kruskal-Wallis

La prueba de MWW, vista en la sección 19.3 se puede usar para probar si dos poblaciones son idénticas. Kruskal y Wallis extendieron esta prueba a tres o más poblaciones. La hipótesis en la **prueba de Kruskal-Wallis** para $k \geq 3$ poblaciones se expresa como sigue.

H_0: Todas las poblaciones son idénticas

H_a: No todas las poblaciones son idénticas

La prueba de Kruskal-Wallis se basa en el análisis de muestras aleatorias independientes de cada una de las k poblaciones.

Esta prueba es una alternativa a la ANOVA presentada en el capítulo 13, en la que se prueba la igualdad de la media de k poblaciones.

En el capítulo 13 se mostró que el análisis de varianza (ANOVA) suele usarse para probar la igualdad de las medias de tres o más poblaciones. En el ANOVA se requieren datos de intervalo o de razón y se requiere suponer que las k poblaciones tienen una distribución normal.

La prueba no paramétrica de Kruskal-Wallis se puede usar tanto con datos ordinales como con datos de intervalo o de razón. Además, en la prueba de Kruskal-Wallis no es necesario suponer que las poblaciones tienen una distribución normal. De manera que siempre que los datos de $k \geq 3$ poblaciones sean ordinales o siempre que la suposición de que las poblaciones tengan una distribución normal sea cuestionable, la prueba de Kruskal-Wallis proporciona un método estadístico alternativo para probar si las poblaciones son idénticas. Esta prueba de Kruskal-Wallis se demostrará con un ejemplo de selección de empleados.

Los empleados que contrata la empresa Williams Manufacturings para su departamento administrativo provienen de tres universidades. Recién el departamento de personal de la empresa ha empezado a revisar el desempeño anual para determinar si hay diferencia en el desempeño de los empleados provenientes de estas tres universidades. Se cuenta con los datos de muestras independientes de clasificación de acuerdo con su desempeño de siete empleados provenientes de la universidad A, seis de la universidad B y siete de la universidad C. En la tabla 19.11 se presentan estos datos; la calificación de acuerdo con su desempeño se da en una escala de 0 a 100.

TABLA 19.11 CALIFICACIONES DE DESEMPEÑO DADAS A 20 EMPLEADOS DE WILLIAMS

Universidad A	Universidad B	Universidad C
25	60	50
70	20	70
60	30	60
85	15	80
95	40	90
90	35	70
80		75

Suponga que se desea probar si las tres poblaciones son idénticas respecto a las calificaciones por su desempeño. Como nivel de significancia se usará 0.05. El estadístico de prueba de Kruskal-Wallis se basa en la suma de los rangos de cada muestra y se calcula como se indica a continuación.

ESTADÍSTICO DE PRUEBA DE KRUSKAL-WALLIS

$$W = \left[\frac{12}{n_T(n_T + 1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i}\right] - 3(n_T + 1) \tag{19.8}$$

donde

k = número de poblaciones
n_i = número de elementos en la muestra i
$n_T = \Sigma n_i$ = número total de los elementos en todas las muestras
R_i = suma de los rangos de la muestra i

Kruskal y Wallis mostraron que bajo la suposición de la hipótesis nula de que las poblaciones son idénticas, la distribución muestral de W puede ser aproximada por una distribución chi-cuadrada con $k - 1$ grados de libertad. Esta aproximación suele ser aceptable siempre que el tamaño de cada una de las muestras sea mayor o igual a cinco. Si el valor del estadístico de prueba es grande, se rechaza la hipótesis nula de que las poblaciones son idénticas. De manera que se usa una prueba de la cola superior.

En la prueba de Kruskal-Wallis únicamente se usa el rango ordinal de los datos.

Para calcular el valor del estadístico W en esta prueba, primero es necesario ordenar los 20 elementos de menor a mayor y asignarles un rango. Al valor menor de estos datos, 15 y que se encuentra en la muestra de la universidad B, se le asigna el rango 1, mientras que al valor mayor, 95 y que se encuentra en la muestra de la universidad A, se le asigna el rango 20. En la tabla 19.12 se presentan los valores de estos datos, sus rangos correspondientes y la suma de los rangos de cada una de las tres muestras. Observe que a los elementos que tienen un mismo valor se les ha asignado un rango promedio;* por ejemplo, los valores 60, 70, 80 y 90 se encuentran repetidos.

Los tamaños de las muestras son

$$n_1 = 7 \qquad n_2 = 6 \qquad n_3 = 7$$

y

$$n_T = \Sigma n_i = 7 + 6 + 7 = 20$$

Mediante la ecuación (19.8) se calcula el estadístico W.

$$W = \frac{12}{20(21)}\left[\frac{(95)^2}{7} + \frac{(27)^2}{6} + \frac{(88)^2}{7}\right] - 3(20 + 1) = 8.92$$

En los procesos con computadora que se presentan en el apéndice F, al final del libro, se muestra el uso de Minitab y Excel para calcular el valor-p.

Ahora se emplea la tabla de la distribución chi-cuadrada (tabla 3 del apéndice B) para determinar el valor-p en esta prueba. Con $k - 1 = 3 - 1 = 2$ grados de libertad, se encuentra que para $\chi^2 = 7.378$, el área en la cola superior de la distribución chi-cuadrada es 0.025 y para $\chi^2 = 9.21$, el área en la cola superior de la distribución chi-cuadrada es 0.01. Como $W = 8.92$ se encuentra entre 7.378 y 9.21, se concluye que el área en la cola superior de la distribución está entre 0.025 y 0.01. Dado que se trata de una prueba de la cola superior, se concluye que el valor-p se encuentra entre 0.025 y 0.01. Con Minitab o con Excel se encuentra que el valor-p = 0.0116. Como el valor-$p \leq \alpha = 0.05$, se rechaza H_0 y se concluye que las poblaciones no son idénticas.

*Si hay muchos valores repetidos, es necesario modificar la ecuación (19.8); la fórmula modificada se encuentra en *Practical Nonparametric Statistics* de W. J. Conover.

TABLA 19.12 RANGOS PARA LOS 20 EMPLEADOS DE WILLIAMS

Universidad A	Rango	Universidad B	Rango	Universidad C	Rango
25	3	60	9	50	7
70	12	20	2	70	12
60	9	30	4	60	9
85	17	15	1	80	15.5
95	20	40	6	90	18.5
90	18.5	35	5	70	12
80	15.5			75	14
Suma de los rangos	95		27		88

El desempeño de los administradores difiere significativamente dependiendo de la universidad de que provienen. Además, dado que las calificaciones al desempeño de los empleados que provienen de la universidad B son las más bajas, sería prudente que la empresa dejara de reclutar empleados de la universidad B o por lo menos los evaluara con más cuidado.

NOTAS Y COMENTARIOS

En el ejemplo usado para ilustrar el procedimiento que se sigue en la prueba de Kruskal-Wallis, lo primero que se hizo fue obtener los datos de nivel de intervalo de las calificaciones del desempeño de los empleados. El procedimiento se hubiera podido realizar también si los datos fueran la clasificación ordinal de los 20 empleados. En ese caso, la prueba de Kruskal-Wallis se hubiera aplicado directamente a los datos originales; el paso de obtención de los rangos a partir de la calificación al desempeño se hubiera omitido.

Ejercicios

Métodos

Autoexamen

26. Las calificaciones dadas a tres productos por un panel de 15 consumidores son las siguientes.

Producto		
A	B	C
50	80	60
62	95	45
75	98	30
48	87	58
65	90	57

Use la prueba de Kruskal-Wallis y $\alpha = 0.05$ para determinar si existe una diferencia significativa entre las calificaciones dadas a los tres productos.

27. Para un examen de admisión se evalúan tres programas de preparación. Las calificaciones obtenidas por las 20 personas de una muestra empleada para probar los programas de preparación son las siguientes. Use la prueba de Kruskal-Wallis para determinar si hay una diferencia significativa entre los tres programas de preparación. Use $\alpha = 0.01$.

	Programa	
A	**B**	**C**
540	450	600
400	540	630
490	400	580
530	410	490
490	480	590
610	370	620
	550	570

Aplicaciones

Auto examen

28. Para bajar de peso basta con practicar una de las siguientes actividades tres veces por semana durante cuarenta minutos. En la tabla siguiente se muestra la cantidad de calorías que se quema con 40 minutos de cada una de estas actividades. ¿Estos datos indican que exista diferencia en la cantidad de calorías quemadas con cada una de estas actividades? Dé su conclusión.

Natación	Tenis	Andar en bicicleta
408	415	385
380	485	250
425	450	295
400	420	402
427	530	268

29. La revista *Condé Nast Traveler* realiza cada año un estudio para evaluar los 80 principales barcos cruceros del mundo (*Condé Nast Traveler,* febrero de 2006). A continuación se dan las evaluaciones dadas a los cruceros de una muestra de las líneas Holland America, Princess y Royal Caribbean; la evaluación máxima es 100. Use la prueba de Kruskal-Wallis con $\alpha = 0.05$ para determinar si hay diferencia significativa en las evaluaciones de los barcos de las tres líneas.

Holland America		Princess		Royal Caribbean	
Embarcación	**Evaluación**	**Embarcación**	**Evaluación**	**Embarcación**	**Evaluación**
Amsterdam	84.5	Coral	85.1	Adventure	84.8
Maasdam	81.4	Dawn	79.0	Jewel	81.8
Ooterdam	84.0	Island	83.9	Mariner	84.0
Volendam	78.5	Princess	81.1	Navigator	85.9
Westerdam	80.9	Star	83.7	Serenade	87.4

30. Una empresa grande envía a muchos de sus administrativos de primer nivel a un curso sobre habilidades de supervisión. Este curso se ofrece en cuatro centros educativos y la empresa desea determinar si éstos difieren en la calidad de la capacitación que ofrecen. Para lo cual toma una muestra de 20 de los empleados que han asistido a estos cursos y la muestra se ordena de acuerdo con sus habilidades para la supervisión, dando un rango a cada uno de los componentes de la muestra. Los resultados obtenidos se presentan a continuación.

Curso	Rango de acuerdo con sus habilidades como supervisor				
1	3	14	10	12	13
2	2	7	1	5	11
3	19	16	9	18	17
4	20	4	15	6	8

Observe que el supervisor que obtuvo el mejor rango asistió al curso 2 y el supervisor que obtuvo el peor rango asistió al curso 4. Use $\alpha = 0.05$ y realice una prueba para determinar si hay una diferencia significativa entre la capacitación ofrecida por estos cuatro cursos.

31. Los dulces más vendidos tienen muchas calorías. Los datos siguientes muestran el contenido de calorías en muestras de M&M, Kit Kat y Milky Way II. Pruebe si hay una diferencia significativa en el contenido de calorías de estos tres dulces. Emplee como nivel de significancia 0.05, ¿cuál es su conclusión?

M&Ms	Kit Kat	Milky Way II
230	225	200
210	205	208
240	245	202
250	235	190
230	220	180

19.5 Correlación de rangos

El coeficiente de correlación de rangos de Spearman es igual al coeficiente de correlación de Pearson, pero se emplea para datos ordinales.

El coeficiente de correlación es una medida de la relación lineal entre dos variables para las cuales se cuenta con datos de intervalo o de razón. En esta sección se estudia una medida de la relación entre dos variables en el caso de datos ordinales. El **coeficiente de correlación por rangos de Spearman** r_s se usa en estos casos.

COEFICIENTE DE CORRELACIÓN DE SPEARMAN

$$r_s = 1 - \frac{6\Sigma d_i^2}{n(n^2 - 1)} \tag{19.9}$$

donde

n = número de elementos o individuos a los que se les va a asignar un rango
x_i = rango del elemento i respecto de una variable
y_i = rango del elemento i respecto de la otra variable
$d_i = x_i - y_i$

A continuación se ilustra el uso del coeficiente de correlación por rangos de Spearman mediante un ejemplo. Una empresa desea determinar si las personas que, en el momento de ser contratadas, generaron expectativas de muy buenos vendedores, en realidad han tenido los mejores registros de ventas. Para esto, el gerente de personal revisa cuidadosamente las entrevistas de trabajo, los antecedentes académicos y las cartas de recomendación de 10 de los vendedores de la empresa. Después de esta revisión, ordena a estas 10 personas de acuerdo con su potencial de éxito, y les da un rango con base en la información disponible al momento de contratarlos. A continuación obtiene una lista del número de unidades vendidas por cada una de estas personas en el transcurso de los primeros dos años y los reordena con un rango de acuerdo con su desempeño real en ventas. En la tabla 19.13 se dan los datos relevantes y los dos rangos. La cuestión es-

TABLA 19.13 POTENCIAL DE VENTAS Y VENTAS REALIZADAS EN DOS AÑOS POR 10 VENDEDORES

Vendedor	Rango de acuerdo con su potencial	Ventas en dos años (unidades)	Rango de acuerdo con sus ventas en dos años
A	2	400	1
B	4	360	3
C	7	300	5
D	1	295	6
E	6	280	7
F	3	350	4
G	10	200	10
H	9	260	8
I	8	220	9
J	5	385	2

tadística es si los rangos, de acuerdo con su potencial de ventas al momento de la contratación, coinciden con los rangos de acuerdo con las ventas realizadas durante los dos primeros años.

Con los datos de la tabla 19.13 se calcula el coeficiente de correlación de rangos de Spearman. En la tabla 19.14 se resumen estos cálculos. Es claro que el coeficiente de correlación por rangos 0.73 es positivo. El coeficiente de correlación por rangos de Spearman varía de -1.0 a $+1.0$ y se interpreta igual que un coeficiente de correlación muestral, en que el valor positivo cercano a 1.0 indica una fuerte relación entre los rangos: si un rango crece el otro crece. Las correlaciones por rangos cercanas a -1.0 indican una fuerte relación pero negativa entre los rangos: cuando un rango crece el otro disminuye. El valor $r_s = 0.73$ indica una correlación positiva entre el desempeño potencial y real. Los individuos con un rango alto de potencial tienden a un alto desempeño.

TABLA 19.14 CÁLCULO DEL COEFICIENTE DE CORRELACIÓN POR RANGOS DE SPEARMAN ENTRE EL POTENCIAL DE VENTAS Y EL DESEMPEÑO EN VENTAS

Vendedor	x_i = Rango de acuerdo con el potencial	y_i = Rango de acuerdo con el desempeño en ventas	$d_i = x_i - y_i$	d_i^2
A	2	1	1	1
B	4	3	1	1
C	7	5	2	4
D	1	6	−5	25
E	6	7	−1	1
F	3	4	−1	1
G	10	10	0	0
H	9	8	1	1
I	8	9	−1	1
J	5	2	3	9
				$\Sigma d_i^2 = 44$

$$r_s = 1 - \frac{6\Sigma d_i^2}{n(n^2 - 1)} = 1 - \frac{6(44)}{10(100 - 1)} = 0.73$$

Prueba de significancia de la correlación de rangos

Hasta aquí se ha visto cómo usar los resultados muestrales para calcular el coeficiente de correlación por rangos. Como ocurre con muchos otros procedimientos estadísticos, se desea emplear los resultados muestrales para hacer inferencias acerca de la correlación por rangos poblacional ρ_S. Para hacer una inferencia acerca de la correlación por rangos poblacionales, se debe probar la hipótesis siguiente.

$$H_0: p_s = 0$$
$$H_\text{a}: p_s \neq 0$$

Bajo la hipótesis nula de que no existe correlación entre los rangos ($\rho_S = 0$), los rangos son independientes y la distribución muestral de r_S es la siguiente.

DISTRIBUCIÓN MUESTRAL DE r_S

$$\text{Media: } \mu_{r_s} = 0 \tag{19.10}$$

$$\text{Desviación estándar: } \sigma_{r_s} = \sqrt{\frac{1}{n-1}} \tag{19.11}$$

Forma de la distribución: aproximadamente normal, siempre que $n \geq 10$.

El coeficiente de correlación por rangos muestrales entre el potencial de ventas y el desempeño en ventas es $r_S = 0.73$. Con este valor se puede probar si hay una correlación por rangos significativa. De acuerdo con la ecuación (19.10) se tiene que $\mu_{r_s} = 0$ y de acuerdo con la ecuación (19.11) se tiene que $\sigma_{r_s} = \sqrt{1/(10-1)} = 0.33$. Si usa como estadístico de prueba la variable aleatoria normal estándar z, tiene

$$z = \frac{r_s - \mu_{r_s}}{\sigma_{r_s}} = \frac{0.73-0}{0.33} = 2.20$$

En las tablas de probabilidad normal estándar, se encuentra que para $z = 2.20$, el valor-p = $2(1 - 0.9861) = 0.0278$. Dado que el valor-$p \leq \alpha = 0.05$ se rechaza la hipótesis nula de que la correlación de los rangos sea cero. Por tanto, se puede concluir que hay una correlación de rangos significativa entre el potencial de ventas y el desempeño en ventas.

Ejercicios

Métodos

Autoexamen

32. Considere los siguientes conjuntos de rangos dados a los 10 elementos de una muestra.

Elemento	x_i	y_i	Elemento	x_i	y_i
1	10	8	6	2	7
2	6	4	7	8	6
3	7	10	8	5	3
4	3	2	9	1	1
5	4	5	10	9	9

a. Calcule el coeficiente de correlación por rangos de Spearman.
b. Use $\alpha = 0.05$ y pruebe la significancia de la correlación por rangos. Dé su conclusión.

33. Considere los siguientes seis conjuntos de rangos dados a seis objetos.

Caso uno			Caso dos		
Objeto	**Primer rango**	**Segundo rango**	**Objeto**	**Primer rango**	**Segundo rango**
A	1	1	A	1	6
B	2	2	B	2	5
C	3	3	C	3	4
D	4	4	D	4	3
E	5	5	E	5	2
F	6	6	F	6	1

Observe que en el primer caso los rangos son idénticos, mientras que en el segundo los rangos son exactamente opuestos. ¿Cuál es el valor que esperaría para el coeficiente de correlación por rangos de Spearman en cada caso? Explique. Para cada caso calcule el coeficiente de correlación por rangos.

Aplicaciones

Autoexamen

34. En la tabla siguiente se presentan los rangos dados para una muestra de 11 estados de acuerdo con el cociente alumnos-profesor (1 = más bajo, 11 más alto) y con los desembolsos por alumno (1 = más alto, 11 más bajo).

	Rango			Rango	
Estado	**Cociente alumnos-profesor**	**Desembolso por alumno**	**Estado**	**Cociente alumnos-profesor**	**Desembolso por alumno**
Arizona	10	9	Massachusetts	1	1
Colorado	8	5	Nebraska	2	7
Florida	6	4	North Dakota	7	8
Idaho	11	2	South Dakota	5	10
Iowa	4	6	Washington	9	3
Louisiana	3	11			

Emplee como nivel de significancia $\alpha = 0.05$, ¿parece haber relación entre el desembolso por alumno y el cociente alumnos-profesor?

35. En un estudio realizado por Harris Interactive, Inc. se evaluaron las principales empresas de Internet y se evaluó también su reputación. En la lista siguiente se muestra el ranking de 10 empresas de Internet en relación, por un lado, con su reputación y, por otro, con el porcentaje de entrevistados que dijeron estar dispuestos a comprar acciones de esa empresa.

	Reputación	**Probable compra**
Microsoft	1	3
Intel	2	4
Dell	3	1
Lucent	4	2
Texas Instruments	5	9
Cisco Systems	6	5
Hewlett-Packard	7	10
IBM	8	6
Motorola	9	7
Yahoo	10	8

a. Calcule la correlación por rangos entre reputación y probable compra.
b. Haga una prueba para determinar si existe una correlación por rangos positiva y significativa. ¿Cuál es el valor-*p*?
c. Emplee como nivel de significancia 0.05, ¿cuál es su conclusión?

36. A continuación se presenta el ranking de una muestra de golfistas profesionales respecto a "driving distance" y "putting" ¿Cuál es la correlación por rangos entre "driving distance" y "putting"? Como nivel de significancia emplee $\alpha = 0.10$.

Golfista profesional	Driving Distance	Putting
Fred Couples	1	5
David Duval	5	6
Ernie Els	4	10
Nick Faldo	9	2
Tom Lehman	6	7
Justin Leonard	10	3
Davis Love III	2	8
Phil Mickelson	3	9
Greg Norman	7	4
Mark O'Meara	8	1

37. En una determinada universidad, una organización de estudiantes entrevista tanto a estudiantes como a recién egresados para obtener información acerca de la calidad de la enseñanza. Al analizar las respuestas se llega a la siguiente clasificación de los profesores de acuerdo con su habilidad para la enseñanza. ¿Coincide la clasificación dada por los estudiantes con la clasificación dada por los recién egresados? Use $\alpha = 0.10$ y pruebe la significancia de la correlación por rangos.

	Clasificación de acuerdo con	
Profesor	**Estudiantes**	**Recién egresados**
1	4	6
2	6	8
3	8	5
4	3	1
5	1	2
6	2	3
7	5	7
8	10	9
9	7	4
10	9	10

Resumen

En este capítulo se presentaron varios métodos estadísticos que se clasifican como métodos no paramétricos. Debido a que los métodos no paramétricos pueden aplicarse a datos nominales y ordinales, así como a datos de intervalo o de razón y a que no se requieren suposiciones acerca de la distribución de la población, estos métodos amplían la clase de problemas que pueden ser sometidos al análisis estadístico.

La prueba de los signos es un método no paramétrico para identificar diferencias entre dos poblaciones, cuando los datos de que se dispone son datos nominales. En el caso de las muestras pequeñas, para determinar los valores críticos de los signos se emplea la distribución de probabilidad binomial; en el caso de las muestras grandes, se emplea una aproximación normal. La prueba de los rangos con signo de Wilcoxon es un método que se emplea para analizar pares de datos, siempre y cuando los datos de cada par sean datos de intervalo o de razón; no es necesario hacer suposiciones acerca de la distribución de la población. Con el método de Wilcoxon se prueba la hipótesis de que las dos poblaciones que se comparan son idénticas.

La prueba de Mann-Whitney-Wilcoxon es un método no paramétrico que se usa para probar la diferencia entre dos poblaciones con base en dos muestras aleatorias independientes; se presentaron tablas para el caso de muestras pequeñas y para el caso de muestras grandes se empleó la aproximación normal. La prueba de Kruskal-Wallis es el análogo no paramétrico a la prueba paramétrica ANOVA para las diferencias entre medias poblacionales.

En la última sección de este capítulo se presentó el coeficiente de correlación por rangos de Spearman, que es una medida de la relación entre dos conjuntos de elementos ordinales o datos ordenados por rangos.

Glosario

Métodos no paramétricos Métodos estadísticos que requieren pocas o ninguna suposición acerca de la distribución de probabilidad de la población y acerca del nivel de medición. Estos métodos suelen usarse cuando se cuenta con datos nominales y ordinales.

Prueba de los signos Prueba estadística no paramétrica para identificar diferencias entre dos poblaciones con base en el análisis de datos nominales.

Prueba de los rangos con signo de Wilcoxon Prueba estadística no paramétrica para identificar diferencias entre dos poblaciones con base en el análisis de dos muestras pareadas.

Prueba de Mann-Whitney-Wilcoxon (MWW) Prueba estadística no paramétrica para identificar diferencias entre dos poblaciones con base en el análisis de dos muestras independientes.

Prueba de Kruskal-Wallis Prueba no paramétrica para identificar diferencias entre tres o más poblaciones.

Coeficiente de correlación por rangos de Spearman Medida de la correlación que se basa en los datos ordenados por rangos de dos variables.

Fórmulas clave

Prueba de los signos (muestras grandes)

$$\text{Media: } \mu = 0.50n \tag{19.1}$$

$$\text{Desviación estándar: } \sigma = \sqrt{0.25n} \tag{19.2}$$

Prueba de los rangos con signo de Wilcoxon

$$\text{Media: } \mu_T = 0 \tag{19.3}$$

$$\text{Desviación estándar: } \sigma_T = \sqrt{\frac{n(n+1)(2n+1)}{6}} \tag{19.4}$$

Prueba de Mann-Whitney-Wilcoxon (muestras grandes)

$$\text{Media: } \mu_T = \tfrac{1}{2} n_1(n_1 + n_2 + 1) \tag{19.6}$$

$$\text{Desviación estándar: } \sigma_T = \sqrt{\tfrac{1}{12} n_1 n_2 (n_1 + n_2 + 1)} \tag{19.7}$$

Estadístico de prueba Kruskal-Wallis

$$W = \left[\frac{12}{n_T(n_T + 1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i}\right] - 3(n_T + 1) \quad \textbf{(19.8)}$$

Coeficiente de correlación por rangos de Spearman

$$r_s = 1 - \frac{6\Sigma d_i^2}{n(n^2 - 1)} \quad \textbf{(19.9)}$$

Ejercicios complementarios

38. En una encuesta se hizo la siguiente pregunta: ¿usted está a favor o en contra de proporcionar vales libres de impuestos o deducciones de impuestos a los padres que envían a sus hijos a escuelas privadas? De los 2 010 encuestados, 905 estuvieron a favor, 1 045 estuvieron en contra y 60 no tuvieron ninguna opinión al respecto. ¿Estos datos indican que existe una diferencia significativa entre las opiniones respecto al apoyo a los padres que envían a sus hijos a escuelas privadas? Como nivel de significancia use 0.05.

39. El precio mediano en Estados Unidos de una vivienda nueva unifamiliar es $230 000 (The Associated Press, 25 de marzo de 2006). Suponga que los siguientes datos son de las ventas de casas unifamiliares ya existentes en Houston y Boston.

	Más que $230 000	Igual que $230 000	Menos que $230 000
Houston	11	2	32
Boston	27	1	13

a. ¿Es el precio mediano de venta en Houston inferior a la mediana estadounidense? Use una prueba estadística con $\alpha = 0.05$ para respaldar su conclusión.

b. ¿Es el precio mediano de venta en Boston superior a la mediana estadounidense? Use una prueba estadística con $\alpha = 0.05$ para respaldar su conclusión.

40. A 12 amas de casa se les pidió que estimaran el precio de venta de dos modelos de refrigeradores. En la tabla siguiente se muestran las estimaciones que dieron. Use estos datos y 0.05 como nivel de significancia y haga una prueba para determinar si existe alguna diferencia entre los dos modelos en términos de la percepción que tienen las amas de casa sobre sus precios.

Ama de casa	Modelo 1	Modelo 2	Ama de casa	Modelo 1	Modelo 2
1	$650	$900	7	$700	$890
2	760	720	8	690	920
3	740	690	9	900	1000
4	700	850	10	500	690
5	590	920	11	610	700
6	620	800	12	720	700

41. Un estudio busca evaluar la ganancia potencial de peso con cierto alimento aviar. Una muestra de 12 aves se usó durante un periodo de seis semanas. Cada una de las aves se pesó antes y después de un periodo de prueba. Las diferencias entre los pesos antes y después observadas en las 12 aves son: 1.5, 1.2, −0.2, 0, 0.5, 0.7, 0.8, 1.0, 0, 0.6, 0.2, −0.01. Valores negativos indican pérdida de peso durante el periodo de prueba y los ceros indican que no hubo ninguna variación durante el

periodo de prueba. Use 0.05 como nivel de significancia y determine si este nuevo alimento parece ocasionar un aumento de peso en las aves.

42. Los datos siguientes son pesos de un producto en dos líneas de producción. Use $\alpha = 0.05$ y haga una prueba para determinar si existe diferencia entre los pesos de las dos líneas de producción.

Línea de producción 1	Línea de producción 2
13.6	13.7
13.8	14.1
14.0	14.2
13.9	14.0
13.4	14.6
13.2	13.5
13.3	14.4
13.6	14.8
12.9	14.5
14.4	14.3
	15.0
	14.9

43. Un cliente desea saber si hay una diferencia significativa entre los tiempos que se requieren para realizar un programa de evaluación por tres métodos diferentes. A continuación se presentan los tiempos (en horas) requeridos por cada uno de los 18 evaluadores para llevar a cabo el programa de evaluación.

Método 1	Método 2	Método 3
68	62	58
74	73	67
65	75	69
76	68	57
77	72	59
72	70	62

Use $\alpha = 0.05$ y realice una prueba para ver si existe una diferencia significativa entre los tiempos requeridos por los tres métodos.

44. Una muestra de 20 ingenieros que han estado empleados en una empresa durante tres años ha sido ordenada por rangos de acuerdo al potencial administrativo. Algunos de estos ingenieros han asistido a cursos de desarrollo dados por la empresa, otros han asistido a cursos de desarrollo dados por la universidad y los restantes no han asistido a ningún tipo de curso. Emplee $\alpha = 0.025$ y realice una prueba para ver si existe una diferencia significativa entre el potencial administrativo de los tres grupos.

Ningún curso	Curso de la empresa	Curso de la universidad
16	12	7
9	20	1
10	17	4
15	19	2
11	6	3
13	18	8
	14	5

45. A continuación se presentan las calificaciones dadas en la evaluación a cuatro profesores. Use $\alpha = 0.05$ y el método de Kruskal-Wallis para probar si existe diferencia significativa en sus evaluaciones.

Profesor	**Calificación en la evaluación del curso**								
Black	88	80	79	68	96	69			
Jennings	87	78	82	85	99	99	85	94	
Swanson	88	76	68	82	85	82	84	83	81
Wilson	80	85	56	71	89	87			

46. Los 15 alumnos de una muestra obtuvieron los rangos siguientes en el examen de mitad de semestre y en el examen final de un curso de estadística.

Rango		**Rango**		**Rango**	
Mitad	**Final**	**Mitad**	**Final**	**Mitad**	**Final**
1	4	6	2	11	14
2	7	7	5	12	15
3	1	8	12	13	11
4	3	9	6	14	10
5	8	10	9	15	13

Calcule el coeficiente de correlación por rangos de Spearman y emplee $\alpha = 0.10$, pruebe si hay una correlación significativa.