La aproximación a la normal se realiza con la media y la desviación típica, calculadas según las siguientes expresiones:

$$\mu = np; \qquad \mu = 0{,}5n \tag{22.13}$$

$$\sigma = \sqrt{npq}; \qquad \sigma = 0{,}5\sqrt{n} \tag{22.14}$$

En la expresiones anteriores, n es el número de casos, p la probabilidad de obtener signo $+$ y q la probabilidad de obtener signo $-$, bajo la hipótesis nula $p = q = 0{,}5$; en las expresiones anteriores p y q se sustituyen por estos valores. A partir de los parámetros anteriores, se calcula el valor de Z:

$$Z = \frac{X - 0{,}5n}{0{,}5\sqrt{n}} \tag{22.15}$$

En la expresión anterior, X representa el número de signos positivos observado. La mayoría de los autores indican que debe efectuarse una corrección de continuidad, sumando o restando, según los casos; en la expresión anterior, 0,5 a X, número de signos $+$ observados. Aplicando esta corrección:

$$Z = \frac{(X \pm 0{,}5) - 0{,}5n}{0{,}5\sqrt{n}} \tag{22.16}$$

En la expresión anterior, se suma 0,5 cuando X es menor que $n/2$ y se resta 0,5 cuando X es mayor que $n/2$.

EJEMPLO 22.4

Las puntuaciones correspondientes a 15 alumnos antes y después de realizar un curso de estadística son las siguientes:

Antes:	5	6	6	8	7	5	4	3	7	5	6	6	3	5	5
Después:	6	6	7	9	6	4	6	3	8	8	4	7	2	7	8
Signos:	+	0	+	+	−	−	+	0	+	+	−	+	−	+	+

Se observa que 9 mejoran, 4 empeoran y 2 repiten puntuación.

Aplicando la distribución binomial, puesto que $n < 25$. Se calcula la probabilidad de que bajo los supuestos de la hipótesis nula, $p = q = 0{,}5$, haya 4 o menos signos negativos, y 9 o más signos positivos, ambas probabilidades se suman:

La probabilidad de 4 o menos signos negativos se denota por $P(4)_{(-)}$, y se calcula de la manera siguiente: $P(4)_{(-)} = P(4) + P(3) + P(2) + P(1) + P(0)$; $P(4)_{(-)} = 0{,}134$; $P(9)_{(+)} = P(9) + P(10) + P(11) + P(12) + P(13) + P(14) + P(15)$; $P(9)_{(+)} = 0{,}133$.

Las probabilidades anteriores se han calculado mediante la distribución binomial. La probabilidad total es 0,267; como es mayor que 0,05, no se rechaza la hipótesis nula, y se concluye que no hay evidencia de que el curso haya servido para mejorar la nota; debe tenerse en cuenta que la muestra es pequeña.

La prueba de los signos con SPSS

Seleccione en el menú análisis Pruebas no paramétricas y en el listado de pruebas disponible: 2 muestras relacionadas. Aparece la pantalla siguiente:

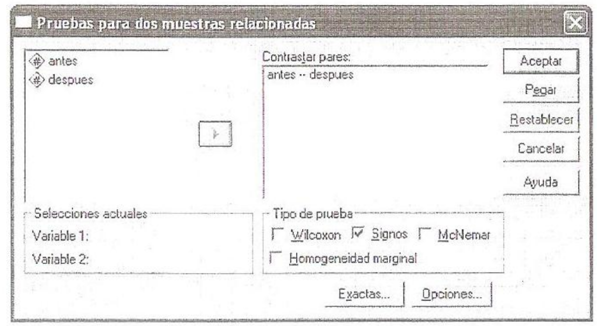

FIGURA 22.8

Se han introducido los datos correspondientes al ejemplo anterior.

Se seleccionan las dos variables a la vez y se pasan a la pantalla Contrastar pares. En tipo de prueba marcar solamente Signos, pulsar Aceptar, se obtienen los siguientes resultados:

Pruebas no paramétricas

Prueba de los signos

Frecuencias

		N
Después − antes	Diferencias negativas [a]	4
	Diferencias positivas [b]	9
	Empates [c]	2
	Total	15

[a] Después < antes. [b] Después > antes. [c] Después = antes.

Estadísticos de contraste [b]

	Después – antes
Sig. exacta (bilateral)	0,267 [a]

[a] Se ha usado la distribución binomial.
[b] Prueba de los signos.

En la tabla encabezada por frecuencias se indican las diferencias negativas, las positivas y los empates observados. A continuación se muestra la probabilidad asociada a los datos observados y cómo se ha calculado, en este caso ha sido mediante la distribución binomial.

Las teclas virtuales: Exactas y Opciones, ofrecen las mismas posibilidades que se han comentado en la prueba binomial en el apartado 22.2.1.

22.3.2. La prueba de Wilcoxon

La prueba de Wilcoxon es aplicable en los mismos supuestos que en el caso anterior, es decir, variables cuantitativas relacionadas. Esta prueba es más potente que la de los signos, pues tiene en cuenta el aumento o disminución de la variable y la magnitud del cambio.

La prueba consiste en calcular las diferencias entre los valores y ordenarlas de menor a mayor por valor absoluto. Una vez ordenadas las diferencias, se numeran de 1 a n, siendo n el número de elementos de la muestra; al número asignado se le denomina *rango*. El rango 1 se asigna a la mínima diferencia observada en valor absoluto, y así sucesivamente hasta n, cuyo rango corresponde a la máxima diferencia. Si hay empate, se asigna a cada diferencia empatada la media de los rangos implicados en el empate; por ejemplo, si hay tres elementos empatados a los que les corresponderían los rangos 4, 5 y 6, se asigna a los tres el rango medio, es decir, 5.

Una vez ordenados los datos, se suman los rangos de las diferencias positivas, $W+$, y negativas, $W-$, y se elige el menor de los dos. Los casos en los que la diferencia es cero se ignoran.

En la mayoría de las tablas y estadísticos se usa la suma de rangos menor.

La prueba se basa en que, si no hay diferencias entre las dos variables relacionadas, los rangos deben estar repartidos de forma homogénea, y tan probable es encontrar un rango grande positivo como negativo. Por lo tanto, si se suman los rangos correspondientes a diferencias positivas, $W+$, y los rangos correspondientes a diferencias negativas, $W-$, deben ser similares y se encontrarán pequeñas diferencias debidas al azar. Si las diferencias entre las suma de rangos son grandes, indica que entre las variables hay diferencias debidas a causas distintas al azar.

Las hipótesis en la prueba de Wilcoxon se pueden enunciar de la manera siguiente:

$$H_0 \quad W(+) = W(-)$$
$$H_1 \quad W(+) \neq W(-) \qquad \alpha$$

El contraste se resuelve para muestras pequeñas, consultando las tablas de Wilcoxon (Tabla XIII de Anexos), en las que se representan las máximas o mínimas sumas de rangos consideradas aceptables. Se rechaza la hipótesis nula, en caso de que la suma de rangos observada sea superior o inferior a estos valores. Para muestras mayores de 25 se puede hacer una aproximación a la normal, con la media y desviación típica definidas por las siguientes expresiones:

$$\mu_W = \frac{n(n+1)}{4} \qquad [22.17]$$

$$\sigma_W = \sqrt{\frac{n(n+1)(2n+1)}{24}} \qquad [22.18]$$

A partir de las expresiones anteriores se deduce la expresión para el cálculo de Z:

$$Z = \frac{W - \left[\frac{n(n+1)}{4}\right]}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} \qquad [22.19]$$

En la expresión anterior, W es la mínima suma de rangos entre las sumas de rangos de las diferencias positivas y negativas.

EJEMPLO 22.5

A continuación se resuelve, mediante la prueba de Wilcoxon, el ejercicio anterior, el 22.5.

Antes	Después	Diferencias	Rangos	Rangos con signo
5	6	1	4,5	−4,5
6	6	0		
6	7	1	4,5	−4,5
8	9	1	4,5	−4,5
7	6	1	4,5	+4,5
5	4	1	4,5	+4,5
4	6	2	10	−10
3	3	0		
7	8	1	4,5	−4,5
5	8	3	12,5	−12,5
6	4	2	10	+10
6	7	1	4,5	−4,5
3	2	1	4,5	+4,5
5	7	2	10	−10
5	8	3	12,5	−12,5

En la tabla anterior, las diferencias entre las variables antes y después están calculadas en valor absoluto. En la columna RANGOS, se han ordenado las diferencias según los siguientes criterios: hay dos diferencias iguales a cero, las cuales se han ignorado; la diferencia mínima observada en valor absoluto es 1, pero hay 8 unos, les corresponden los rangos del 1 al 8, se asigna el rango medio, 4,5, a cada uno de ellos; a continuación hay tres diferencias con valor 2, a las que corresponden los rangos 9, 10 y 11, se asigna el rango medio, 10, a las tres; por último, hay dos diferencias iguales a tres, a las que les corresponden los rangos 12 y 13, se les asigna el rango medio, 12,5.

En la columna Rangos con signo, se asigna el signo menos a las diferencias negativas y el signo más a las diferencias positivas. El signo (−), en este caso, significa que la puntuación ha aumentado, puesto que al restar ANTES-DESPUÉS las puntuaciones que han aumentado tienen diferencia negativa. El signo en esta prueba es un símbolo diferenciador y debe tenerse cuidado con su interpretación.

Sumamos los rangos con signo positivo $W+ = 23,5$ y los negativos $W- = 67,5$. Como valor W se considera el menor, es decir, 23,5.

En la tabla se observa que para $n = 15$, el punto crítico para una significación de 0,05, es 25, como el valor W obtenido es 23,5 que es menor se rechaza la hipótesis nula y se concluye que hay diferencias estadísticamente significativas entre las dos variables y, consecuentemente, el cursillo ha tenido influencia en las puntuaciones. Observe que con los mismos datos la prueba de los signos no detectó diferencias estadísticamente significativas, pero esto no es una contradicción, la prueba de Wilcoxon es estadísticamente más potente porque usa más información, tiene en cuenta el signo y la magnitud de las diferencias.

La prueba de Wilcoxon con SPSS

Seleccione en el menú Análisis, Pruebas no paramétricas, y en el listado de Pruebas disponible: 2 muestras relacionadas. Aparece la pantalla siguiente:

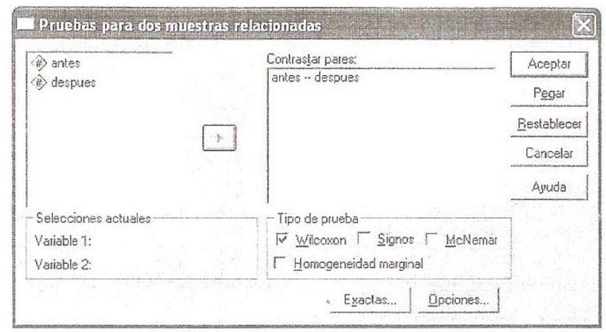

FIGURA 22.9

En la pantalla anterior se ha marcado Wilcoxon. Se realiza la prueba con los datos del ejemplo 22.6. Pulsando Aceptar se obtienen los resultados siguientes:

Pruebas no paramétricas

Prueba de los rangos con signo de Wilcoxon

Rangos

		N	Rango promedio	Suma de rangos
Después − antes	Rangos negativos	4[a]	5,88	23,50
	Rangos positivos	9[b]	7,50	67,50
	Empates	2[c]		
	Total	15		

[a] Después < antes. [b] Después > antes. [c] Después = antes.

Estadísticos de contraste[b]

	Después − antes
Z	−1,581[a]
Sig. asintót. (bilateral)	0,114

[a] Basado en los rangos negativos.
[b] Prueba de los rangos con signo de Wilcoxon.

En la primera tabla encabezada por rangos, se muestran los rangos positivos, los rangos negativos y los empates.

En la tabla encabezada por estadísticos de contraste se indica el valor Z obtenido. SPSS ha realizado la aproximación a la normal, aunque el tamaño de la muestra es menor que 25, en este caso la aproximación no es adecuada, por eso el valor no es estadísticamente significativo. Para muestras menores que 25 hay que consultar las tablas.

Las teclas virtuales: Exactas y Opciones, ofrecen las mismas posibilidades que se han comentado en la prueba binomial en el apartado 22.2.1.

22.4. PRUEBAS NO PARAMÉTRICAS PARA DOS MUESTRAS INDEPENDIENTES

En este apartado se estudian pruebas para analizar casos en los que las dos variables son independientes. Se considera que las variables son independientes cuando los valores de cada variable proceden de elementos muestrales distintos, con características diferentes. Las pruebas analizan si estas características influyen en el valor de la variable.

Los campos de aplicación son múltiples. Por ejemplo, si las puntuaciones en anatomía de dos muestras de alumnos procedentes de universidades distintas

son diferentes; si dos tratamientos antihipertensivos tienen efectos diferentes en dos muestras de pacientes con características distintas; si la proporción de votantes a un determinado partido político difieren en dos regiones de un determinado país, etc.

Las pruebas disponibles que se estudian en este apartado son las siguientes:

- Prueba de la mediana.
- Prueba de Mann-Whitney.
- Prueba de Kolmogorov-Smirnov.
- Prueba de Wald-Wolfowitz.
- Prueba de Moses.

22.4.1. Prueba de la mediana para dos muestras independientes

Esta prueba es adecuada para comparar dos variables cuantitativas. Los tamaños de las dos muestras pueden ser distintos; se denotan mediante n_1 y n_2.

La prueba se fundamenta en analizar si las medianas de las dos poblaciones son distintas. En primer lugar se ordenan de menor a mayor todos los datos como si pertenecieran a un solo grupo, aunque identificando el grupo de pertenencia, y se calcula la mediana. A continuación, en una tabla dos por dos, se disponen el número de casos de cada muestra, que son mayores o menores que la mediana global. Si las medianas poblacionales son iguales, la proporción de casos de cada muestra que son menores o mayores que la mediana global deben ser similares, salvo diferencias debidas al azar; por el contrario, si las medianas son diferentes, las proporciones de casos mayores y menores de la mediana global serán significativamente distintas en ambas muestras.

		Grupo	
		A	*B*
Variable	> Mediana	a	b
	≦ Mediana	c	d

Si no hay diferencias entre las medianas poblacionales las proporciones $\frac{a}{a+c}$ y $\frac{b}{b+d}$ deben ser similares.

Las hipótesis se pueden enunciar de la manera siguiente:

$$H_0\ M1 = M2$$
$$H_1\ M1 \neq M2 \qquad \alpha$$

La hipótesis nula indica que las medianas poblacionales son iguales y la alternativa que son distintas. También pueden plantearse contrastes unilaterales.

Si el número total de casos, suma de las dos muestras, es mayor que 20 y ninguna frecuencia esperada es menor que 5, se aplica la prueba de la Chi-cuadrado de Pearson. Cuando la muestra tiene menos de 20 casos hay que hacer la corrección de Yates[1]. En el caso de que el número total de casos sea menor que 20, o alguna frecuencia teórica menor que 5, se aplica la prueba de Fisher.

EJEMPLO 22.6 Se carga en 1 colum. y tipo (1/2)

Se prueban dos tratamientos *A* y *B*, en pacientes que tienen el colesterol basal elevado. Se asignan aleatoriamente a los grupos 70 pacientes, 35 a cada uno de ellos. La mediana del colesterol basal de los 70 valores es 230 mg por 100 mililitros. Una vez clasificados los resultados se obtiene la tabla siguiente:

		Tratamiento	
		A	*B*
Colesterolemia basal	> Mediana	20	14
	≦ Mediana	15	21

Plantear las hipótesis y resolver el contraste con $\alpha = 0{,}05$.

$$H_0 \quad M_A = M_B$$
$$H_1 \quad M_A \neq M_B \qquad \alpha = 0{,}05$$

El valor de Chi-cuadrado con la corrección de Yates es 1,4, el punto crítico para las pruebas basadas en la distribución Chi-cuadrado con un grado de libertad es 3,84, como el valor es menor no hay pruebas estadísticas para rechazar la hipótesis nula. Se concluye que no hay diferencias estadísticamente significativas, según la prueba de la mediana, de que haya diferencias entre los tratamientos.

La prueba de la mediana con SPSS

En el menú Análisis seleccione Estadística no paramétrica, y en la lista de éstas pruebas seleccione *K* muestras independientes: la prueba de la mediana no se puede realizar a partir de 2 muestras independientes, pero sí en *K* muestras independientes, en este caso $K = 2$. Una vez realizadas las selecciones aparece la pantalla siguiente:

[1] En el Capítulo 14 se analizan las pruebas basadas en la distribución Chi-cuadrado incluyendo la corrección de Yates.

FIGURA 22.10

Se ha marcado la prueba de la mediana. Para que junto a la variable que define los grupos figuren los valores 1 y 2, hay que pulsar la tecla virtual definir rango y en mínimo y máximo especificar 1 y 2 respectivamente. Pulsando aceptar se obtienen los siguientes resultados que corresponden al ejemplo 22.7:

Pruebas no paramétricas

Prueba de la mediana

Frecuencias

		Tratamiento	
		A	*B*
Colesterolemia basal	> Mediana	20	14
	≦ Mediana	15	21

Estadísticos de contraste [a]

		Colesterolemia basal
N		70
Mediana		230,00
Chi-cuadrado		2,059
Gl		1
Sig. asintót.		0,151
Corrección por continuidad de Yates	Chi-cuadrado	1,430
	Gl	1
	Sig. asintót.	0,232

[a] Variable de agrupación: tratamiento.

En la primera tabla se clasifican los resultados según sean mayores o menores que la mediana.

En la tabla siguiente se muestran el número de datos, la mediana del colesterol basal de los 70 casos, el valor de Chi-cuadrado sin la corrección de Yates, los grados de libertad y la significación estadística; las tres últimas filas muestran el valor de Chi-cuadrado con la corrección de Yates, sus grados de libertad y la significación estadística.

En este caso la prueba a considerar es la corrección de Yates puesto que el tamaño de la muestra es menor que 200. Las conclusiones son las mismas que las del ejemplo 22.7.

Las teclas virtuales: Exactas y Opciones, ofrecen las mismas posibilidades que se han comentado en la prueba binomial en el apartado 22.2.1.

22.4.2. La prueba de Mann-Whitney

Esta prueba es aplicable para comparar los valores de dos variables cuantitativas independientes, también se puede aplicar a variables ordinales. Las dos muestras pueden tener tamaños distintos. Es la prueba no paramétrica considerada más potente para comparar los valores de dos variables cuantitativas independientes.

El procedimiento es el siguiente: se agrupan los datos de las dos muestras en un sólo grupo, se ordenan los datos de menor a mayor, asignándole a cada dato el rango correspondiente a su orden. Si no hay diferencias entre las dos variables, se espera que los rangos estén uniformemente repartidos entre los dos grupos; por el contrario, si hay diferencias entre las dos variables, se espera que los rangos menores se asocien con una de las muestras y los mayores con la otra.

Las hipótesis pueden enunciarse de la manera siguiente:

H_0 No hay diferencias entre las variables.
H_1 Hay diferencias entre las variábles. α

Si existen diferencias mayores de las esperadas por efecto del azar entre los valores de las variables, los detectaría la prueba propuesta por Mann-Whitney, basada en la suma de los rangos correspondientes a cada muestra.

Se dispone de datos cuantitativos correspondientes a dos muestras aleatorias, con tamaños n_1 y n_2; la suma de los rangos correspondientes a cada grupo se denotan mediante R_1 y R_2. Los estadísticos U_1 y U_2 se obtienen mediante las expresiones siguientes:

$$U_1 = n_1 n_2 + \left[\frac{n_1(n_1+1)}{2}\right] - R_1 \qquad [22.20]$$

$$U_2 = n_1 n_2 + \left[\frac{n_2(n_2+1)}{2}\right] - R_2 \qquad [22.21]$$

Una vez calculados los parámetros anteriores, se elige el menor; a este valor se le denomina U y, mediante la tabla XIV de Anexos, se comprueba si las diferencias entre los valores de las variables son estadísticamente significativas.

Cuando las muestras tienen más de 20 casos, se consigue una buena aproximación a una distribución normal con media y desviación típica definidas según las siguientes expresiones:

$$\mu_U = \frac{n_1 n_2}{2} \qquad [22.22]$$

$$\sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}} \qquad [22.23]$$

A partir de la media y desviación típica definidas en las expresiones anteriores, se calcula Z, valor normal tipificado, mediante la siguiente expresión:

$$Z = \frac{U - \left(\frac{n_1 n_2}{2}\right)}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}} \qquad [22.24]$$

En la expresión anterior U, es el menor entre U_1 y U_2.

EJEMPLO 22.7

Se trata durante tres meses a dos grupos de pacientes hipercolesterolémicos seleccionados aleatoriamente con dos tratamientos distintos 1 y 2; al finalizar el tratamiento se mide el colesterol basal en ambos grupos y se comparan los resultados. Los resultados obtenidos son los siguientes:

1: 210 232 225 198 205 207
2: 230 275 280 220 301 332 340

Los tamaños muestrales son distintos, $n_2 = 7$ y $n_1 = 6$.

Se decide comparar los valores mediante la prueba de Mann-Whitney, con $\alpha = 0{,}05$, para lo cual se agrupan todos los valores en un sólo conjunto y se ordenan de menor a mayor:

Valor:	198	205	207	210	220	225	230	232	275	280	301	332	340
Rango:	1	2	3	4	5	6	7	8	9	10	11	12	13
Grupo:	1	1	1	1	2	1	2	1	2	2	2	2	2

La suma de rangos para el grupo 1, $R_1 = 24$.
La suma de rangos para el grupo 2, $R_2 = 67$.

A partir de los datos anteriores, se calcula U_1 y U_2 mediante las expresiones 22.20 y 22.21.

$$U_1 = 3 \qquad U_2 = 33$$

El menor de los dos es $U_1 = 3$, y este valor es el que corresponde al parámetro $U = 3$.

Consultando en la tabla XIV de Anexos cuando $n_1 = 6$, $n_2 = 7$ y $U = 3$, encontramos que la significación para los valores anteriores es 0,004; por lo tanto, se rechaza la hipótesis nula y se concluye que los tratamientos tienen efectos distintos. El tratamiento 2 es el mejor de los dos, puesto que los valores de colesterol son significativamente menores en los pacientes que lo han seguido.

La prueba de Mann-Whitney con SPSS

En el menú Análisis seleccione Estadística no paramétrica, y en la lista de estas pruebas seleccione 2 muestras independientes, aparece la pantalla siguiente:

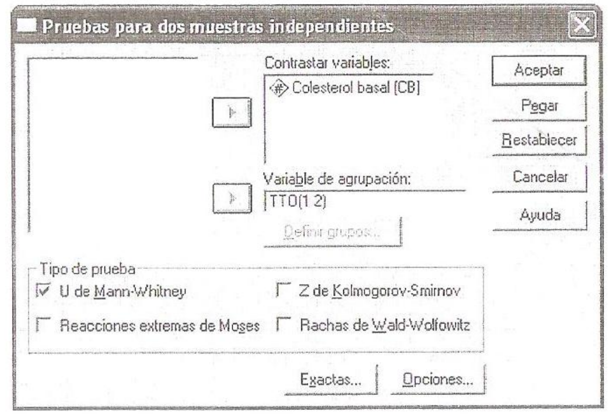

FIGURA 22.11

En la pantalla anterior la variable que se quiere contrastar, Colesterol basal, se pasa a la ventana encabezada por «Contrastar variables», y la variable que define los grupos, Tto, a la encabezada por «Variable de agrupación», en esta variable hay que definir los valores que definen los tratamientos, 1 y 2 en este caso pulsando «Definir grupos»; después se marca en «Tipo de prueba» U de Mann-Whitney. Pulsando en Aceptar se obtienen los siguientes resultados:

Pruebas no paramétricas

Prueba de Mann-Whitney

Rangos

	Tratamiento	*N*	Rango promedio	Suma de rangos
Colesterol basal	1	7	9,57	67,00
	2	6	4,00	24,00
	Total	13		

Estadísticos de contraste[b]

	Colesterol basal
U de Mann-Whitney	3,000
W de Wilcoxon	24,000
Z	−2,571
Sig. asintót. (bilateral)	0,010
Sig. exacta [2.ª (Sig. unilateral)]	0,008[a]

[a] No corregidos para los empates.
[b] Variable de agrupación: tratamiento.

En la primera tabla se muestran, para cada grupo, el número de casos, el rango promedio y la suma de rangos.

En la segunda tabla se muestra el valor de U; la W de Wilcoxon que corresponde a una prueba para comparar dos muestras independientes, aunque no ha tenido tanta aceptación como là homónima correspondiente a dos muestras relacionadas, en este caso se ignora; la Z calculada mediante la expresión 22.24 y la significación estadística que le corresponde, como las muestras son menores que 20 la aproximación a la normal es inadecuada; en la última fila se muestra la significación estadística, según las tablas, 0,008, que es la correcta, se muestra la significación correspondiente a un contraste unilateral, como el contraste, en éste caso es bilateral la significación estadística es 0,004.

Las teclas virtuales: Exactas y Opciones, ofrecen las mismas posibilidades que se han comentado en la prueba binomial en el apartado 22.2.1.

22.4.3. La prueba de Kolmogorov-Smirnov para dos variables independientes

Además de la prueba no paramétrica para una muestra estudiada en este mismo capítulo, Kolmogorov y Smirnov idearon otra prueba, válida para comparar dos variables independientes. Las variables que comparar deben ser cuantitativas. Los fundamentos de esta prueba son similares a los de la prueba aplicable para una sola muestra.

En el caso de dos variables independientes, la prueba pretende comprobar si las distribuciones poblacionales de las dos variables son iguales o distintas. La prueba de dos colas es sensible a diferencias en tendencia central, dispersión y colocación.

Las hipótesis se pueden enunciar de la manera siguiente:

H_0 Las distribuciones son iguales.
H_1 Las distribuciones son distintas. α

El estadístico de contraste es D, que es la máxima diferencia entre las frecuencias relativas acumuladas calculadas para cada valor. El parámetro D se puede calcular mediante la siguiente expresión:

$$D = \text{Máx}\ [F_1(x) - F_2(x)] \qquad [22.25]$$

En la expresión anterior, F_1 es la frecuencia relativa acumulada de valores de la primera muestra, que son iguales o menores que x, F_2 es la proporción de valores de la segunda muestra que son iguales o menores que x. La diferencia anterior se calcula para todos los valores y el valor de la diferencia máxima es el parámetro D.

El parámetro D está tabulado y, consultando las correspondientes tablas, se puede comprobar si las diferencias son o no estadísticamente significativas. Cuando las dos muestras son mayores que 40 casos, se puede utilizar el siguiente estadístico para resolver el contraste de hipótesis:

$$\chi^2 = 4D^2\left(\frac{n_1 n_2}{n_1 + n_2}\right) \qquad [22.26]$$

El estadístico anterior sigue una distribución Chi-cuadrado con 2 grados de libertad.

También se utiliza la denominada Z de Kolmogorov-Smirnov, que se calcula mediante la siguiente expresión:

$$Z_{K-S} = \text{Máx}\ |D| \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}} \qquad [22.27]$$

La Z_{K-S} se distribuye según una normal tipificada.

EJEMPLO 22.8

Se seleccionan aleatoriamente dos muestras de 10 individuos cada una de ellas, M_1 y M_2, en dos poblaciones distintas, la variable de interés es la tensión arterial sistólica. Los datos obtenidos son los siguientes:

M_1:	120	125	133	152	126	134	152	140	136	164
M_2:	162	150	163	153	127	144	161	123	138	152

H_0 Las distribuciones son iguales.
H_1 Las distribuciones son distintas. $\alpha = 0{,}05$

La D máxima es 0,4, aplicando la expresión 22.27, se obtiene la Z de Kolmogorov Smirnof:

$$Z_{K-S} = 0{,}4 \cdot \sqrt{\frac{10 \cdot 10}{10 + 10}}; \qquad Z_{K-S} = 0{,}894$$

Los puntos críticos son $-1{,}96$ y $1{,}96$, como $0{,}894$ está entre los puntos críticos, no hay pruebas estadísticas para rechazar la hipótesis nula y, por lo tanto, de que las distribuciones sean distintas.

La prueba de Kolmogorov-Smirnov para dos muestras independientes con SPSS

En el menú Análisis seleccione Estadística no paramétrica, y en la lista de Pruebas seleccione 2 muestras independientes, aparece la pantalla siguiente:

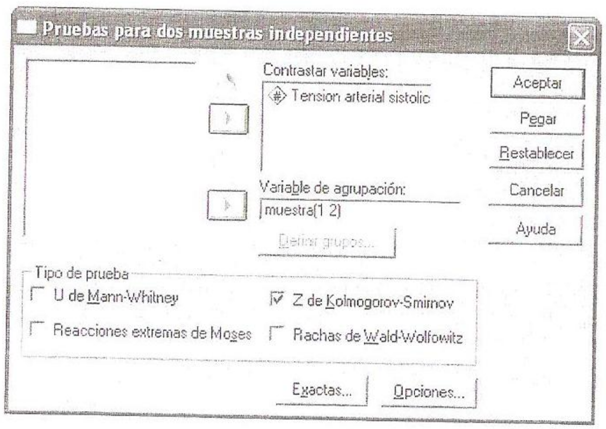

FIGURA 22.12

La variable cuya distribución se quiere contrastar se pasa a la ventana «Contrastar variables»; la que define los grupos denominada muestras en «variable de agrupación» pulsando en «Definir grupos» se indican los valores que definen los grupos, 1 y 2 en este caso. Se resuelve mediante SPSS el ejemplo 22.9.

En «Tipo de prueba» se marca Z de Kolmogorov-Smirnov. Pulsando Aceptar se obtienen los siguientes resultados:

Pruebas no paramétricas

Prueba de Kolmogorov-Smirnov para dos muestras

Frecuencias

	Tratamiento	N
Tensión arterial sistólica	1	10
	2	10
	Total	20

Estadísticos de contraste[a]

		Tensión arterial sistólica
Diferencias más extremas	Absoluta	0,400
	Positiva	0,400
	Negativa	−0,400
Z de Kolmogorov-Smirnov		0,894
Sig. asintót. (bilateral)		0,400

[a] Variable de agrupación: tratamiento.

En la primera tabla se indican el número de casos de cada muestra, en la tabla siguiente se muestran las diferencias absoluta, positiva y negativa; la Z de Kolmogorov-Smirnov, que es 0,894 y su significación estadística bilateral, que como es mayor que 0,05, no es estadísticamente significativa.

Las teclas virtuales: Exactas y Opciones, ofrecen las mismas posibilidades que se han comentado en la prueba binomial en el apartado 22.2.1

22.4.4. La prueba de las rachas de Wald-Wolfowitz para dos variables independientes

Esta prueba analiza las distribuciones de dos variables independientes y puede detectar diferencias en la tendencia central, dispersión u oblicuidad.

Los datos de las dos variables se agrupan en un solo conjunto de datos, ordenándolos de menor a mayor y contando a continuación las rachas referidas a los grupos. Si el primer dato es del grupo 1, se cuenta una racha; si los datos segundo y tercero pertenecen al grupo 2, se cuenta otra racha, etc. Cada sucesión de datos correspondiente a un grupo se cuenta como una racha. La prueba se basa en que si las distribuciones de los datos son iguales, las rachas estarán uniformemente repartidas; por el contrario, si hay diferencias entre ellas, las rachas de cada grupo se encontrarán repartidas de forma desequilibrada.

Las hipótesis son.

H_0 Distribuciones iguales.
H_1 Distribuciones distintas. $\alpha = 0{,}05$

El contraste anterior se resuelve para muestras pequeñas, menores o iguales que 20, consultando las rachas obtenidas en la correspondiente tabla. Para muestras mayores que 20, se puede realizar una aproximación a la normal con media y desviación típica definidas mediante las siguientes expresiones:

$$\mu_r = \frac{2n_1n_2}{n_1 + n_2} + 1 \qquad [22.28]$$

$$\sigma_r = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}} \qquad [22.29]$$

A partir de las expresiones para la media y la desviación típica anteriores, se deduce la expresión para Z:

$$Z = \frac{r - \left(\frac{2n_1n_2}{n_1 + n_2} + 1\right)}{\sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}} \qquad [22.30]$$

En la expresión anterior, r es el número de rachas observado. Algunos autores indican que en la expresión anterior debe realizarse una corrección de continuidad para mejorar la aproximación. La corrección por continuidad consiste en restar 0'5 al valor absoluto del numerador en la expresión anterior, que queda de la siguiente manera:

$$Z = \frac{\left|r - \left(\frac{2n_1n_2}{n_1 + n_2} + 1\right)\right| - 0{,}5}{\sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}} \qquad [22.31]$$

La prueba de las rachas de Wald-Wolfowitz para dos variables independientes con SPSS

En el menú Análisis seleccione Estadística no paramétrica, y en la lista de Pruebas seleccione 2 muestras independientes, aparece la pantalla siguiente:

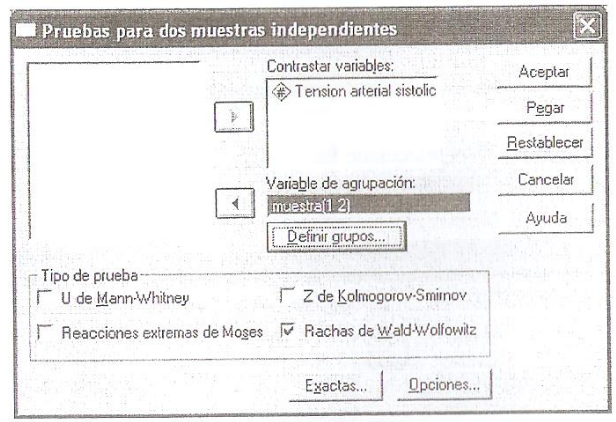

FIGURA 22.13

La variable cuya distribución se quiere contrastar se pasa a la ventana «Contrastar variables»; la que define los grupos, denominada muestra en «variable de agrupación» pulsando en «Definir grupos» se indican los valores que definen los grupos, 1 y 2 en este caso. Se resuelve mediante la prueba de Wald-Wolfowitz con SPSS el ejemplo 22.8.

En «Tipo de prueba» se marca rachas de Wald-Wolfowitz. Pulsando Aceptar se obtienen los siguientes resultados:

Pruebas no paramétricas

Prueba de Wald-Wolfowitz

Frecuencias

	Tratamiento	*N*
Tensión arterial sistólica	1	10
	2	10
	Total	20

Estadísticos de contraste [b, c]

		Número de rachas	*Z*	Sig. exacta (unilateral)
Tensión arterial sistólica	Mínimo posible	11 [a]	0,000	0,586
	Máximo posible	13 [c]	1,149	0,872

[a] Hay 1 empates inter-grupos que implican tres casos.
[b] Prueba de Wald-Wolfowitz.
[c] Variable de agrupación: Tratamiento.

En la primera tabla se indica el número de casos de cada grupo. En la segunda tabla se muestra el número de rachas, el valor de *Z* y la significación estadística, que en este caso es mayor que 0,05; consecuentemente no hay pruebas estadísticas de que las distribuciones sean diferentes en los dos grupos.

El número de rachas está entre 11 y 13 porque hay 3 empates, y dependiendo del orden en que se pongan puede haber 11, 12 o 13 rachas; en cualquier caso la significación estadística es mayor que 0,05.

Las teclas virtuales: Exactas y Opciones, ofrecen las mismas posibilidades que se han comentado en la prueba binomial en el apartado 22.2.1.

22.4.5. La prueba de los valores extremos de Moses

Las pruebas que comparan medias no son significativas si los valores medios son similares, pero puede ocurrir que los valores de dos variables tengan medias parecidas por causas distintas. Si los valores de una variable son todos parecidos; los de la otra variable la mitad son pequeños y la otra mitad grandes, las medias pueden ser iguales, sin embargo los valores extremos pueden ser diferentes. Por ejemplo, un tratamiento hipotensor disminuye la tensión arterial sistólica de manera uniforme en una muestra, y disminuye mucho la tensión arterial sistólica en

la mitad de los integrantes de la otra muestra y poco en la otra mitad, las tensiones arteriales medias pueden ser iguales, pero el comportamiento del tratamiento en las dos muestras es diferente.

La prueba de Moses trata de determinar si el comportamiento en los valores extremos, mayores y menores, de dos variables es igual o distinto. A uno de los grupos se le denomina *grupo experimental*, y al otro, *grupo control*.

La prueba se basa en agrupar en un solo conjunto los datos y ordenarlos de menor a mayor, asignándoles los rangos correspondientes y evaluando el parámetro *S*, al que se denomina *amplitud*, que es el número más pequeño de posiciones que incluye todos los valores del grupo control; se calcula mediante la siguiente expresión:

$$S = RM - Rm + 1 \qquad [22.32]$$

En la expresión anterior *RM* es el máximo rango obtenido por los valores del grupo control y *Rm* es el mínimo rango obtenido por los valores del grupo control.

Si no hay diferencias significativas entre los valores extremos se espera que los rangos mayores y menores estén repartidos equitativamente entre los dos grupos; en caso contrario, habrá diferencias que podrán ser detectadas por la prueba de Moses. Se recortan los datos extremos el mayor y el menor del grupo control, en este caso el parámetro $h = 1$; puede decidirse recortar más datos, si se recortan los dos mayores y los dos menores $h = 2$, etc. Moses propuso evaluar la amplitud recortada, *S'*, en lugar de *S*; la amplitud recortada se calcula mediante la expresión 22.32, pero ignorando los valores extremos, el mayor y el menor, del grupo control si $h = 1$.

$$S' = RM' - Rm' + 1 \qquad [22.33]$$

En la expresión anterior *RM'* es el rango máximo recortado y *Rm'* el rango mínimo recortado.

EJEMPLO 22.9

Se dispone de dos muestras aleatorias de glucemias basales correspondientes a dos grupos de diabéticos tratados con dos tipos de insulinas distintas, *A* y *B*, a los que se denominan *experimental* (E) y *control* (C), respectivamente.

Experimental:	126	99	100	150	120	125
Control:	80	110	104	101	132	135

Los valores anteriores los disponemos en un solo grupo y les asignamos el correspondiente rango.

Valores:	80	99	100	101	104	110	120	125	126	132	135	150
Grupos:	C	E	E	C	C	C	E	E	E	C	C	E

El máximo rango obtenido por los valores del grupo control es 11, que es el correspondiente al valor 135, y el mínimo 1, que es el correspondiente a 80. Por tanto, $S = 11 - 1 + 1 = 11$; esto indica que el número más pequeño de posiciones que incluye a todos los valores es 11. Ignorando los valores extremos del grupo control, que son 132 y 80, el rango máximo recortado es 10 y el mínimo 4; $S' = 10\text{-}4 + 1 = 7$.

Las hipótesis en la prueba de Moses son:

H_0 No hay diferencias en los valores extremos.
H_1 Hay diferencias en los valores extremos. α

El estadístico de contraste mediante el que se calcula la significación estadística es el siguiente:

$$P(S' \leq n_c - 2h + g) = \frac{\sum_{i=0}^{g} \binom{i + n_c - 2h - 2}{i} \cdot \binom{n_E + 2h + 1 - i}{i}}{\binom{n_c + n_E}{n_c}} \qquad [22.34]$$

En la expresión anterior n_c es el tamaño de la muestra del grupo control, n_E es el tamaño de la muestra del grupo experimental, h es el número de elementos que se recorta de cada extremo, lo más frecuente es que $h = 1$, g es el valor de S' que excede a $(n_c - 2h)$, i es un contador que oscila entre 0 y g.

La prueba de Moses con SPSS

En el menú Análisis seleccione Estadística no paramétrica, y en la lista de Pruebas seleccione 2 muestras independientes, aparece la pantalla siguiente:

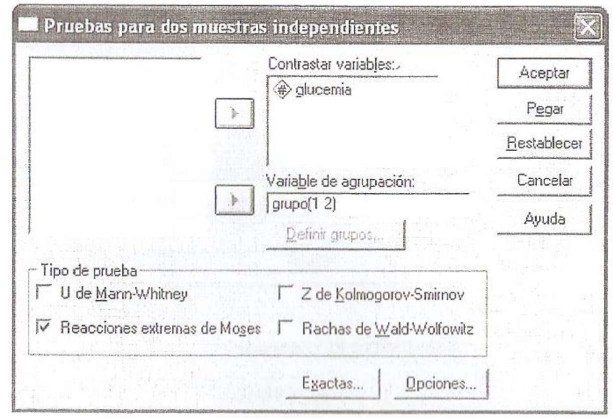

FIGURA 22.14

La variable cuya distribución se quiere contrastar, glucemia, se pasa a la ventana «Contrastar variables»; la que define los grupos denominada grupo en «variable de agrupación» pulsando en «Definir grupos» se indican los valores que definen los grupos, 1 y 2 en este caso. Se resuelve mediante la prueba de reacciones extremas de Moses, el ejemplo 22.9.

Se han obtenido los siguientes resultados:

Pruebas no paramétricas

Prueba de Moses

Frecuencias

	Tratamiento	N
Glucemia	CONTROL (Control)	6
	EXPERIMENTAL (Experimental)	6
	Total	12

Estadísticos de contraste [a, b]

		Glucemia
Amplitud observada del grupo control		11
	Sig. (unilateral)	0,773
Amplitud recortada del grupo control		7
	Sig. (unilateral)	0,716
Valores atípicos recortados de cada extremo		1

[a] Prueba de Moses.
[b] Variable de agrupación: grupo.

En la primera tabla se indica el número de casos que hay en cada grupo, en la segunda tabla se muestra el valor de S, al que se denomina *amplitud observada del grupo control* y el valor de P calculada mediante 22.34 sin recortar los valores extremos, y el valor de S' al que se denomina *amplitud recortada del grupo control*, calculada mediante 22.34, recortando el valor máximo y el valor mínimo del grupo control, también indica el valor de h, al que se denomina *valores atípicos recortados de cada extremo*. Como la significación estadística es mayor que 0,05, no se rechaza la hipótesis nula y se concluye que no hay diferencias en los valores extremos.

Las teclas virtuales: Exactas y Opciones, ofrecen las mismas posibilidades que se han comentado en la prueba binomial en el apartado 22.2.1.

22.5. PRUEBAS PARA k VARIABLES RELACIONADAS

En este apartado se estudian las pruebas no paramétricas más utilizadas para comparar más de dos variables relacionadas. En muchos trabajos de investigación se comparan datos de una variable en varias circunstancias distintas. Se considera que K grupos de datos son independientes cuando los elementos de la muestra de cada grupo son distintos; y dependientes o relacionados, cuando los elementos son los mismos en circunstancias distintas. La circunstancia diferenciadora más frecuente suele ser el tiempo, aunque hay otras; por ejemplo, en una muestra de pacientes comparar la concentración de una sustancia en varios organos.

Los ejemplos de este tipo de variables son múltiples. Supongamos que se quiere estudiar la concentración de un tóxico en distintos órganos de peces: cerebro, corazón y sangre; con objeto de evaluar el alcance de la contaminación de un río, y los órganos más afectados. Para llevar a cabo este experimento se podrían extraer tres muestras de peces y, en cada una de ellas, estudiar la concentración de tóxico en un órgano determinado; pero debido a la posible dispersión del hábitat, podría ocurrir que algunas de las diferencias fueran debidas a que algunos peces

vivieran en regiones más o menos contaminadas, y no que hubiera diferentes concentraciones en los órganos, es decir, habría una variabilidad entre peces y otra dentro de cada pez. Un experimento en el que podrían obtenerse medidas con menos errores de medida sería extraer una sola muestra de peces y estudiar en cada pez la concentración de tóxico en cada órgano; en este caso se eliminaría la variabilidad entre peces. De esta manera, las tres variables, cada una de ellas correspondiente a la concentración de tóxico en un órgano, están relacionadas.

Las pruebas más utilizadas para comparar K variables relacionadas son:

- La prueba de Friedman.
- La prueba de Kendall.
- La prueba de Cochran.

22.5.1. Prueba de Friedman

Esta prueba permite comparar K variables cuantitativas u ordinales relacionadas. Se recomienda hacer una tabla en la que las K variables, es decir, las K medidas, estén en las columnas y los n elementos en las filas, de esta manera la tabla tendrá K columnas y n filas.

A la prueba de Wilcoxon también se le denomina *análisis de la varianza de bloques no paramétricos*, porque puede aplicarse en los supuestos de ANOVA bloques cuando no se cumplen las condiciones paramétricas, pero esto es contradictorio porque el análisis de la varianza se denomina así porque contrasta las hipótesis mediante un cociente de varianzas: los cuadrados medios, para lo cual se tienen que cumplir condiciones paramétricas.

Las filas, en ciertas ocasiones, pueden representar grupos de elementos, bloques, aunque lo más frecuente es que sean elementos.

A los valores de cada fila se les asigna un número del 1 a K, según el orden de magnitud de menor a mayor; a este número se le denomina *rango*. Si no hay diferencia entre las variables los rangos deben estén repartidos en cada columna de manera uniforme, sólo habría entre ellas pequeñas diferencias debidas al azar; en caso de que las diferencias sean demasiado grandes, sería debido a diferencias entre los valores de las variables independientes del azar.

Las hipótesis son:

$$\begin{array}{lll} H_0 & R_1 = R_2 = \ldots = R_k & \\ H_1 & R_i \neq R_j \text{ para algún } ij & \alpha \end{array}$$

R_i representa la suma de rangos correspondientes a la i-ésima variable; según la disposición de los datos indicada también es la suma de rangos correspondientes a la i-ésima columna.

La hipótesis nula indica que las sumas de rangos, de todas las variables son iguales, salvo diferencias explicables por el azar. La hipótesis alternativa indica