que, al menos, la suma de rangos de una variable es significativamente distinta de alguna de las otras.

Para resolver el contraste de hipótesis anterior, Friedman propuso un estadístico que se distribuye como una Chi-cuadrado con $K - 1$ grados de libertad, siendo K el número de variables relacionadas; se calcula mediante la siguiente expresión:

$$\chi^2_{FR} = \frac{12}{nK(K+1)} \sum_{i=1}^{k} R_i^2 - 3n \cdot (K+1) \qquad [22.35]$$

En la expresión anterior, n representa el número de elementos o de bloques, K el número de variables relacionadas y R_i representa la suma de rangos de la i-ésima variable.

Comparaciones múltiples

Si se rechaza la hipótesis nula se concluye que al menos el rango medio de una de las variables es distinta de las demás, pero no se sabe cuál o cuáles son. Se podrían realizar comparaciones dos a dos de todas las parejas posibles mediante la prueba de Wilcoxon, pero habría que tener en cuenta el error tipo I global. En el Capítulo 16, correspondiente al análisis de la varianza de una vía, se abordó con detalle este tema.

Como en la comparación de variables relacionadas los tamaños muestrales son iguales, es aconsejable aplicar las pruebas de Tukey o Dunnet, según los casos, con las correspondientes modificaciones para el caso no paramétrico. Las características generales de estas pruebas se estudian detalladamente en el Capítulo 16, y en el Capítulo 17 la aplicación para K variables relacionadas; no obstante, en este apartado se recuerdan los aspectos principales y las características específicas no paramétricas. En el caso no paramétrico se comparan los rangos medios y no los valores medios de las variables, de esta manera, aunque los valores de las variables estén medidos en la escala ordinal, los rangos son una variable cuantitativa discreta. Estas pruebas permiten calcular la diferencia mínima significativa, DMS, entre dos rangos medios. Si la diferencia entre dos rangos medios es mayor que el DMS correspondiente se considera que la diferencia entre esas variables es estadísticamente significativa. Muchos paquetes estadísticos no permiten realizar pruebas de comparación múltiple no paramétricas, entre ellos SPSS.

Si se quieren realizar todas las posibles combinaciones de dos medias, se recomienda aplicar la prueba de Tukey. Si se quieren comparar los rangos medios con uno al que se considera grupo control, se recomienda aplicar la prueba de Dunnett; por ejemplo, si el primer grupo se considera el basal y el resto de medidas interesa compararlas sólo respecto a los valores basales.

Prueba de Tukey

La diferencia mínima significativa, DMS, mediante la prueba de Tukey, se puede calcular mediante la siguiente expresión:

$$DMS_T = \frac{q_{\alpha;K,\infty}}{\sqrt{2}} \cdot \sqrt{\frac{K(K+1)}{6n}} \qquad [22.36]$$

El cálculo de intervalos de confianza para la diferencia de rangos es un tema controvertido porque la variable es discreta, y si el tamaño de la muestra es pequeño el intervalo es muy grande y poco útil desde el punto de vista técnico. Un intervalo de confianza $1 - \alpha$, para la diferencia de rangos medios entre dos variables, $D\overline{R}_{i,j}$, se puede calcular mediante la siguiente expresión:

$$D\overline{R}_{i,j} \in (\hat{D}\overline{R}_{i,1} - DMS_T; \quad \hat{D}\overline{R}_{i,1} + DMS_T) \qquad [22.37]$$

En la expresión anterior y en las siguientes $D\overline{R}_{i,j}$, es la diferencia entre los rangos medios poblacionales entre la i-ésima y la j-ésima variables; $\hat{D}\overline{R}_{i,1}$ es la diferencia de rangos estimada entre la i-ésima y la j-ésima variables, es decir, la diferencia observada entre los rangos de la i-ésima y la j-ésima variables.

En la expresión anterior $q_{\alpha;k,\infty}$ es el valor del rango estudentizado con K e ∞ grados de libertad, que deja a su derecha un área bajo la curva igual a α; n es el tamaño de la muestra, los valores están tabulados, tabla VIII de Anexos.

Prueba de Dunnett

La diferencia mínima significativa, DMS, mediante la prueba de Dunnett, se puede calcular mediante la siguiente expresión:

$$DMS_D = D_{\alpha/2;K,\infty} \cdot \sqrt{\frac{K(K+1)}{6n}} \qquad [22.38]$$

En la expresión anterior $D_{\alpha/2;K,\infty}$ es el valor del estadístico de Dunnet (tabla IX) con K e ∞ grados de libertad, que deja a su derecha un área bajo la curva igual a $\alpha/2$.

Un intervalo de confianza $1 - \alpha$, para la diferencia de rangos medios entre dos variables, $\overline{D}_{R;i,j}$, se puede calcular mediante la siguiente expresión:

$$D\overline{R}_{i,j} \in (\hat{D}\overline{R}_{i,1} - DMS_D; \quad \hat{D}\overline{R}_{i,1} + DMS_D) \qquad [22.39]$$

EJEMPLO 22.10

Con objeto de estudiar la diferencia de concentración de un tóxico en distintos órganos de peces, se extrae una muestra aleatoria de peces de un río y se estudia en cada uno de ellos la concentración del tóxico en cerebro corazón y sangre. El objetivo del estudio es conocer si la concentración del tóxico en los tres órganos es igual o distinta. Los resultados obtenidos son los siguientes:

Concentración de tóxico en mg/1.000 g

Cerebro	Corazón	Sangre
164 (3)	96 (2)	51 (1)
105 (2)	115 (3)	41 (1)
150 (3)	100 (2)	46 (1)
145 (3)	75 (1)	79 (2)
139 (3)	88 (2)	52 (1)
144 (3)	64 (1)	70 (2)
139 (3)	97 (2)	46 (1)
98 (2)	101 (3)	52 (1)
146 (3)	99 (2)	55 (1)
153 (3)	91 (2)	39 (1)
138 (3)	94 (2)	41 (1)
99 (2)	105 (3)	46 (1)

En la tabla anterior los rangos por filas figuran entre paréntesis; por ejemplo en la primera fila, que recoge los datos del primer pez, la menor concentración está en la sangre a la que se asigna el menor rango, es decir el uno, de la fila; el segundo rango corresponde al corazón y la mayor concentración y el tercer rango al cerebro.

Las hipótesis son:

$$H_0 \quad R_1 = R_2 = R_3$$
$$H_1 \quad R_i \neq R_j \text{ para algún } ij \qquad \alpha = 0{,}05$$

Las sumas de rangos correspondientes a cada órgano, a cada variable, a cada columna son:

$$R_1 = 33 \qquad R_2 = 25 \qquad R_3 = 14$$

Dividiendo las sumas de rangos anteriores por 12 se obtienen los rangos medios:

$$\overline{R}_1 = 2{,}75 \qquad \overline{R}_2 = 2{,}08 \qquad \overline{R}_3 = 1{,}17;$$

Las diferencias entre las concentraciones medias en los órganos son técnicamente significativas. Aplicando la expresión 22.35, se obtiene el valor del estadístico de contraste: $\chi^2_{FR} = 15{,}17$; el punto crítico para una distribución Chi-cuadrado con 2 grados de libertad es 5,99, como el valor obtenido es mucho mayor, hay pruebas estadísticas suficientes para rechazar la hipótesis nula y concluir que las concentraciones del tóxico en los diferentes órganos son estadísticamente significativas; la significación estadística la P es 0,00051.

$$D\overline{R}_{1,2} = 2{,}75 - 2{,}08; \qquad D\overline{R}_{1,2} = 0{,}67$$
$$D\overline{R}_{1,3} = 2{,}75 - 1{,}17; \qquad D\overline{R}_{1,3} = 1{,}58$$
$$D\overline{R}_{2,3} = 2{,}08 - 1{,}17; \qquad D\overline{R}_{2,3} = 0{,}91$$

Ahora hay que dilucidar entre que órganos están las diferencias.

Se calcula la diferencia mínima significativa según Tukey aplicando la expresión 22.36:

$$DMS_T = \frac{3{,}314}{\sqrt{2}} \cdot \sqrt{\frac{3 \cdot (3+1)}{6 \cdot 12}}; \qquad DMS_T = 0{,}96$$

Se considera que hay diferencias estadísticamente significativas entre las variables que su diferencia de rangos medios sea mayor que 0,96.

Solo hay diferencias estadísticamente significativas en la concentración del tóxico entre cerebro y sangre. Entre corazón y sangre hay diferencias aparentemente importantes que no son significativas, debe tenerse en cuenta que la muestra es pequeña.

La prueba de Friedman con SPSS

En el menú análisis seleccione estadística no paramétrica, y en la lista de estas pruebas seleccione *K* muestras relacionadas, aparece la pantalla siguiente:

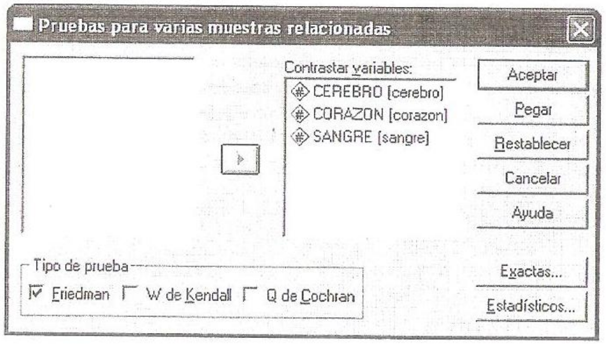

FIGURA 22.15

Se resuelve el ejemplo 22.11 con SPSS, una vez introducidos los datos, las variables que se quieren contrastar Cerebro, Corazón y Sangre en este caso se pasan a la ventana «Contrastar variables»; se marca en «Tipo de prueba» Friedman, pulsando Aceptar se obtienen los resultados siguientes:

Pruebas no paramétricas

Prueba de Friedman

Rangos

	Rango promedio
Cerebro	2,75
Corazón	2,08
Sangre	1,17

Estadísticos de contraste [a]

N	12
Chi-cuadrado	15,167
gl	2
Sig. asintót.	0,001

[a] Prueba de Friedman.

En la primera tabla se muestran los rangos medios correspondientes a cada variable. En la segunda tabla se muestran el número de casos, el valor del estadístico de contraste, los grados de libertad y la significación estadística, que es aproximada $P < 0{,}001$. Las conclusiones son las mismas que se expusieron anteriormente.

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

22.5.2. Coeficiente de concordancia de Kendall

El coeficiente de concordancia de Kendall, al que la mayoría de los autores simboliza por la letra W, es una técnica de análisis estadístico muy utilizada en ciencias de la salud y en sociología.

Mide el grado de concordancia entre un grupo de elementos y un grupo de características. Si la concordancia es la máxima posible, $W = 1$, el máximo valor que puede tener el coeficiente W es la unidad; por el contrario, si la concordancia es la mínima posible, es decir, no hay concordancia $W = 0$. Por lo tanto, el coeficiente puede oscilar entre 0 y 1.

El coeficiente W de Kendall, se puede calcular mediante la siguiente expresión:

$$W = \frac{12 \sum_{i=1}^{k} S_i^2 - 3n^2K(K+1)^2}{n^2K^2(K-1)} \qquad [22.40]$$

En la expresión anterior, K es el número de características que evaluar, n es el número de elementos que intervienen en el estudio, S es la suma de las puntuaciones de cada característica evaluada, que correspondería a la suma de las columnas; si se disponen las características en columnas y los elementos de la muestra en filas. El máximo valor posible para W es 1, máxima concordancia, y el mínimo cero, para una falta total de concordancia.

En caso de empates en las puntuaciones, la expresión para el cálculo de W es la siguiente:

$$W = \frac{12\sum_{j=1}^{K} S_i^2 - 3n^2K(K+1)^2}{n^2K(K^2-1) - n\sum_{i=1}^{K}(t_i^3 - t_i)} \qquad [22.41]$$

En la expresión anterior, t_i es el número de observaciones empatadas con el mismo rango; puede tomar valores entre 2 y K. El valor K lo tomaría si un evaluador diera a todas las características el mismo rango.

Las hipótesis en el caso del coeficiente de Kendall son las siguientes:

H_0 $W = 0$ No hay concordancia

H_1 $W > 0$ Hay concordancia α

El estadístico que utilizar para resolver el contraste de hipótesis anterior es el siguiente:

$$\chi^2_{KEN} = n(K-1)W \qquad [22.42]$$

El estadístico anterior sigue una distribución Chi-cuadrado, con $K - 1$ grados de libertad.

Los campos más frecuentes de aplicación son los siguientes:

a) Homogeneidad entre circunstancias y elementos muestrales.
b) Acuerdo entre expertos y características.

a) Conocer el grado de homogeneidad de un grupo de elementos respecto a un grupo de circunstancias. Por ejemplo, se quiere comprobar si los alumnos universitarios formados en un mismo centro tienen un grado de formación homogéneo respecto a varias disciplinas o, por el contrario, si hay disciplinas en las que los alumnos, en general, destacan y otras en las que la preparación de la mayoría es deficiente. La comprobación anterior puede realizarse seleccionando una muestra de n alumnos, evaluando la aptitud de cada uno de ellos en cada una de las K disciplinas seleccionadas; para cada alumno se clasifica el orden de aptitud en cada una de las K disciplinas de 1 a K; si un alumno tiene la mejor nota en Matemáticas, se le asigna el número 1 a esta disciplina en la fila correspondiente, si la nota siguiente corresponde a Anatomía, se asigna el número 2 a esta disciplina, y así hasta K, número que le corresponderá a la disciplina en que ese alumno tenga peor nota. Esta clasificación se realiza con todos los alumnos seleccionados.

EJEMPLO 22.11

Se quiere estudiar la concordancia entre disciplinas y alumnos formados en una determinada facultad de Medicina. Se obtuvieron los resultados siguientes:

Alumnos	Matemáticas	Anatomía	Física	Biología
1	1	2	4	3
2	4	3	1	2
3	3	1	4	2
4	3	2	4	1
5	4	1	3	2
6	3	1	4	2
7	3	2	4	1
8	4	2	3	1
Rango medio	3,13	1,75	3,38	1,75

El primer alumno ha obtenido la mejor calificación en Matemáticas, la segunda mejor nota la obtuvo en Anatomía, la tercera en Biología y la cuarta en Física. Las puntuaciones anteriores no son las notas; en Matemáticas, el primer alumno no ha obtenido un 1. Por otra parte, cuando se dice que el primer alumno ha obtenido la mejor nota en Matemáticas, no se refiere a todos los alumnos, sino a la clasificación de las disciplinas en cada alumno.

Si en el ejemplo anterior hubiera concordancia entre la preparación de los alumnos, como parece que la hay, puesto que la mayoría obtienen las calificaciones más altas en Anatomía y Biología, se espera que la suma de las columnas sea distinta.

Las hipótesis que contrastar son las siguientes:

H_0 $W = 0$ No hay diferencias entre las materias

H_1 $W > 0$ Hay diferencias entre las materias $\alpha = 0{,}05$

Aplicando la expresión 22.40 se obtiene el coeficiente $W = 0{,}46$. Ahora hay que dilucidar si es estadísticamente significativo; el valor de Chi-cuadrado se obtiene mediante la expresión 22.42 $\chi^2_{FR} = 10{,}95$; el punto crítico para una distribución Chi-cuadrado con 3 grados de libertad es 7,81; como el valor obtenido es mayor se rechaza la hipótesis nula y se concluye que hay diferencias en preparación entre las materias.

En el ejemplo anterior, la concordancia indica que la preparación de los alumnos está desnivelada. En general, los alumnos están mejor preparados en unas disciplinas que en otras, lo cual puede ser por un deficiente funcionamiento de las disciplinas en las que la aptitud es menor, o bien porque el grado de interés de los alumnos en general se centra en unas disciplinas determinadas.

b) Otro campo de aplicación muy importante es estudiar el grado de acuerdo de un grupo de expertos (médicos, psicólogos, jueces, etc.), sobre un conjunto de temas. Por ejemplo, se podría plantear a un grupo de de n médicos, que clasifique en orden de importancia un conjunto de K síntomas, colocando primero el que considere el más importante, al que se asigna el número 1, y con K al que considere menos importante. Igual que en el ejemplo anterior, la clasificación de la valoración de importancia de los síntomas se hace para cada médico de forma independiente de los demás. Si hay acuerdo entre la valoración de los síntomas, se espera encontrar un desequilibrio en la suma de las puntuaciones de cada síntoma. Esta suma será menor para el síntoma más importante y mayor para el menos importante.

La prueba W de Kendall con SPSS

En el menú análisis seleccione estadística no paramétrica, y en la lista de estas pruebas seleccione K muestras relacionadas, aparece la pantalla siguiente:

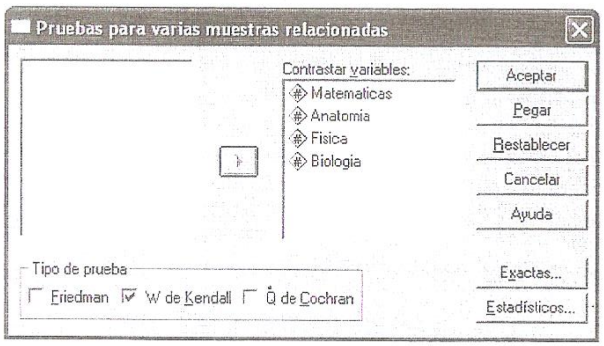

FIGURA 22.16

Se resuelve el ejemplo 22.12 con SPSS, una vez introducidos los datos, las variables que se quieren contrastar Matemáticas, Anatomía, Física y Biología se pasan a la ventana «Contrastar variables»; se marca en «Tipo de prueba» W de Kendall, pulsando Aceptar se obtienen los resultados siguientes:

Pruebas no paramétricas

Prueba W de Kendall

Rangos

	Rango promedio
Matemáticas	3,13
Anatomía	1,75
Física	3,38
Biología	1,75

Estadísticos de contraste

N	8
W de Kendall[a]	0,456
Chi-cuadrado	10,950
gl	3
Sig. asintót.	0,012

[a] Coeficiente de concordancia de Kendall.

En la primera tabla se muestran los rangos medios correspondientes a cada variable. En la segunda tabla se muestran el número de casos, el valor del estadístico de W de Kendall, el valor del estadístico de contraste los grados de libertad y la significación estadística, que es aproximada $P < 0{,}012$. Las conclusiones son las mismas que se expusieron anteriormente.

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

22.5.3. La prueba de la Q de Cochran

Esta prueba es válida para evaluar si la respuesta de un grupo de elementos ante un conjunto de características es homogénea, o por el contrario existen diferencias entre los elementos estudiados. La respuesta es dicotómica: sólo puede tener dos valores, éxito o fracaso, sí o no, etc., que se denotan mediante 0 y 1.

Los campos de aplicación de esta prueba son múltiples. Por ejemplo, n individuos son sometidos a k pruebas; cada una de ellas sólo puede evaluarse con éxito o fracaso. La prueba de la Q permite estudiar si las diferencias entre las características son estadísticamente significativas.

Las hipótesis se pueden enunciar de la manera siguiente:

H_0 No hay diferencias entre las características.
H_1 Hay diferencias entre las características. α

El estadístico de contraste es el siguiente:

$$Q = \frac{K(K-1)\sum_{i=1}^{k} T_i^2 - [(K-1)(\sum_{j=1}^{n} S_j)^2]}{K\sum_{j=1}^{n} S_j - \sum_{j=1}^{n} S_j^2} \qquad [22.43]$$

La expresión anterior, representa la Q de Cochran. K es el número de pruebas o características, n es el número de casos, S_j es la suma de las puntuaciones otorgadas para cada caso y T_i es la suma de las puntuaciones de cada prueba.

El estadístico Q sigue una distribución Chi-cuadrado, con $K - 1$ grados de libertad.

DICOTOMICA 1 BUENO / 0 MALO

EJEMPLO 22.12

A un grupo de 10 expertos se les pide que lean cuatro artículos científicos, A, B, C y D, cada uno de los cuales deben evaluar como bueno (1) o como malo (0). Los resultados obtenidos son los siguientes:

SUMA PUNTUACIONES OTORGADAS P/ CADA CASO

Expertos	Artículos					
	A	B	C	D	S_j	S_j^2
1	0	0	1	0	1	1
2	1	0	1	1	3	3^2 9
3	1	0	1	1	3	3^2 9
4	1	1	1	1	4	4^2 16
5	0	1	0	0	1	1
6	1	0	1	1	3	9
7	1	0	1	1	3	9
8	1	0	1	0	2	4
9	1	0	1	0	2	4
10	0	0	0	0	0	0
T_i	7	2	8	5		
T_i^2	49	4	64	25		

(2) S_j

SUMA DE PUNTUACIONES P/ CADA PRUEBA

7^2 2^2 8^2 5^2

$$K = 4 \qquad \sum_{i=1}^{4} T_i^2 = (49 + 4 + 64 + 25) = 142$$

$$\sum_{j=1}^{10} S_j = 1 + 3 + 3 + 4 + 1 + 3 + 3 + 2 + 2 + 0 = 22$$

$$\sum_{j=1}^{10} S_j^2 = 1 + 9 + 9 + 16 + 1 + 9 + 9 + 4 + 4 + 0 = 62$$

En la tabla anterior, T_i representa el total de las puntuaciones de cada artículo, y S_j representa el total de las puntuaciones otorgadas por cada experto.

Las hipótesis en este caso son:

H_0 Los artículos son similares.
H_1 Los artículos son distintos. α

$$Q = \frac{4 \cdot 3\,(49 + 4 + 64 + 25) - 3 \cdot 22^2}{4 \cdot 22 - 62}$$

?

En el ejemplo anterior, Q es igual a 9,69; el punto crítico para una distribución Chi-cuadrado con 3 grados de libertad y una significación de 0,05 es 7,81; como el valor de la Q es mayor que el punto crítico, se rechaza la hipótesis nula y se acepta que el valor de los artículos, según los expertos, es distinto.

La prueba Q de Cochran con SPSS

En el menú análisis seleccione estadística no paramétrica, y en la lista de estas pruebas seleccione K muestras relacionadas, aparece la pantalla siguiente:

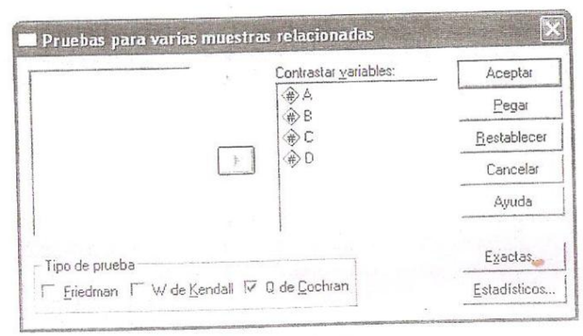

FIGURA 22.17

Se resuelve el ejemplo 22.13 con SPSS, una vez introducidos los datos, las variables que se quieren contrastar A, B, C y D se pasan a la ventana «Contrastar variables»; se marca en «Tipo de prueba» Q de Cochran, pulsando Aceptar se obtienen los resultados siguientes:

Pruebas no paramétricas

Prueba de Cochran

Frecuencias

	Valor	
	0	1
A	3	7
B	8	2
C	2	8
D	5	5

BUENO T_i

Estadísticos de contraste

N	10
Q de Cochran	9,692 [a]
gl	3
Sig. asintót.	0,021

[a] 0 se trata como un éxito.

P > 0,05 ACEPTO H_0
P < 0,05 LO RECHAZO
?

En la primera tabla se muestran los ceros y unos que han correspondido a cada artículo. En la segunda tabla se muestran el número de casos, el valor del estadístico de Q de Cochran, los grados de libertad y la significación estadística, que es aproximada $P < 0{,}021$. Las conclusiones son las mismas que se expusieron en el ejemplo.

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

22.6. PRUEBAS NO PARAMÉTRICAS PARA *k* VARIABLES INDEPENDIENTES

En este apartado se analizan las pruebas no paramétricas más utilizadas para comparar los valores de K variables independientes.

Los campos de aplicación son múltiples. Por ejemplo, si se quieren comparar simultáneamente el colesterol basal en tres grupos: no fumadores, fumadores de menos de 10 cigarrillos y fumadores de más de 10 cigarrillos; no sería correcto hacer la comparación aplicando de manera sucesiva las pruebas para dos muestras, pues la probabilidad de cometer error tipo I sería muy grande.

Las pruebas disponibles en SPSS para evaluar K muestras independientes son:

- La prueba de Kruskal-Wallis.
- La prueba de la mediana para K variables.

22.6.1. La prueba de Kruskal Wallis

Esta prueba es válida para comparar simultáneamente los valores de K variables cuantitativas u ordinales.

Las hipótesis son:

H_0 Los valores de las K variables son similares.
H_1 Los valores de las K variables son diferentes. α

La prueba se basa en agrupar los datos de las K variables en un solo grupo, ordenado de menor a mayor, asignando a cada dato el correspondiente rango. Si los valores son similares, los datos de las K variables se repartirán de manera homogénea en el grupo común ordenado, y la suma de los rangos asignados a cada grupo tendrá valores próximos. Por el contrario, si los valores son distintos son de esperar diferencias entre las sumas de rangos más grandes que las explicables por el azar.

El estadístico de contraste para esta prueba se puede calcular mediante la siguiente expresión:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{K} \frac{R_i^2}{n_i} - 3(N+1) \qquad [22.44]$$

En la expresión anterior, K es el número de grupos, n_i es el número de casos del i-ésimo grupo, R_i es la suma de rangos del i-ésimo grupo y N es el número total de sujetos que intervienen en la prueba.

Para muestras pequeñas, la significación de los valores de H está tabulada. Según aumenta el tamaño de la muestra, H se aproxima a una distribución Chi-cuadrado con $K - 1$ grados de libertad. La aproximación a la Chi-cuadrado puede hacerse para muestras con más de ocho elementos.

Cuando hay empates (dos o más datos tienen los mismos valores), se resuelven asignando a cada dato implicado en el empate el rango medio correspondiente a todos los rangos implicados en dicho empate. En caso de empates, el estadístico H debe ser corregido, se calcula mediante la siguiente expresión:

$$H = \frac{\frac{12}{N(N+1)} \sum_{i=1}^{K} \frac{R_i^2}{n_i} 3(N+1)}{1 - \left[\frac{\sum_{s=1}^{r} (t_s^3 - t_s)}{(N^3 - N)} \right]} \quad [22.45]$$

En la expresión anterior, s indica el s-ésimo empate (téngase en cuenta que puede haber varios empates), y r el número total de empates; t_s es el número de sujetos empatados en el s-ésimo empate.

El parámetro H se distribuye según una Chi-cuadrado con $(K - 1)$ grados de libertad, excepto para muestras pequeñas que no tienen interés práctico.

EJEMPLO 22.13

Para comparar tres tratamientos antihipertensivos A, B y C, se seleccionan aleatoriamente tres muestras de pacientes hipertensos, a cada una de las cuales se asigna un tratamiento distinto durante tres meses. Al cabc de ese tiempo se mide la TAS (tensión arterial sistólica) a todos los pacientes. Los resultados obtenidos son los siguientes:

A: 122 130 155 115 140 125
B: 145 138 152 137 148
C: 165 136 130 145 150 138 160

Los datos anteriores se ordenan de menor a mayor en un solo grupo, asignándoles los correspondientes rangos.

En la siguiente tabla se muestran los resultados:

Valor	Rango	Grupo	Valor	Rango	Grupo
115	1	1	140	10	1
122	2	1	145	11,5	2
125	3	1	145	11,5	3
130	4,5	1	148	13	2
130	4,5	3	150	14	3
136	6	3	152	15	2
137	7	2	155	16	1
138	8,5	2	160	17	3
138	8,5	3	165	18	3

En la tabla anterior se observan varios empates; el primer empate lo forman dos valores 130, uno del grupo 1 y otro del grupo 3; a estos datos les corresponden los rangos 4 y 5, asignándoles a cada uno el rango medio de los rangos empatados 4,5. El resto de los empates se resuelven de la misma manera.

La suma de rangos del primer grupo es 36,5, la suma de rangos del segundo grupo es 55, y la suma de rangos del tercer grupo es 79,6.

El valor de H sin corregir por empates es 3,69 y el valor de H corregido por empates es 3,714.

El punto crítico para Chi-cuadrado con dos grados de libertad y $\alpha = 0{,}05$ es 5,99; como $H = 3{,}71$ es menor, no se puede rechazar la hipótesis nula y se concluye que no hay pruebas estadísticas de que el efecto hipotensor de los tratamientos sea diferente.

Comparaciones múltiples

Si se rechaza la hipótesis nula se concluye que al menos la mediana de una de las variables es distinta de las demás, pero no se sabe cuál o cuáles son. Se podrían realizar comparaciones dos a dos de todas las parejas posibles mediante la prueba de Mann-Whitney, pero habría que tener en cuenta el error tipo I global. En el Capítulo 16, correspondiente al análisis de la varianza de una vía se abordó con detalle este tema.

Es aconsejable aplicar las pruebas de Scheffé, Tukey o Dunnet, según los casos, con las correspondientes modificaciones para el caso no paramétrico. Las características de estas pruebas se estudian detalladamente en el Capítulo 16; no obstante, en este apartado se recuerdan los aspectos principales y las características específicas no paramétricas. En el caso no paramétrico se comparan los rangos medios y no los valores medios de las variables, de esta manera aunque los valores de las variables estén medidos en la escala ordinal, los rangos son una variable cuantitativa discreta. Estas pruebas permiten calcular intervalos de confianza y la diferencia mínima significativa, DMS, entre dos rangos medios. Si la diferencia entre dos rangos medios es mayor que el DMS correspondiente se considera que la diferencia entre los rangos medios es estadísticamente significativa. Muchos paquetes estadísticos no permiten realizar pruebas de comparación múltiple no paramétricas, entre ellos SPSS.

Si se quieren realizar todas las posibles combinaciones de dos medias, se recomienda aplicar la prueba de Scheffé si los grupos tienen distinto número de elementos. En caso de que todos los grupos tengan el mismo número de elementos, se recomienda aplicar la prueba de Tukey. Si se quieren comparar los rangos medios con uno al que se considera grupo control, se recomienda aplicar la prueba de Dunnett.

El cálculo de intervalos de confianza para la diferencia entre rangos es un tema controvertido, no obstante se dan las fórmulas para su cálculo en las tres pruebas que se estudian a continuación.

Prueba de Scheffé

La diferencia mínima significativa, DMS, mediante la prueba de Scheffé, se puede calcular mediante la siguiente expresión:

$$DMS_S = \sqrt{\chi^2_{\alpha,\, K-1}} \cdot \sqrt{\frac{N(N+1)}{6} \cdot \left(\frac{1}{n_i} + \frac{1}{n_j}\right)} \qquad [22.46]$$

En las expresión anterior $\chi^2_{\alpha,\, K-1}$ es el valor de Chi-cuadrado con $K - 1$ grados de libertad, siendo K el número de grupos, que deja a su derecha un área bajo la curva igual a α; n_i y n_j son el número de elementos de los grupos que se comparan; N es el número total de elementos, es la suma de los tamaños muestrales de todos los grupos:

$$N = \sum_{i=1}^{K} n_i \qquad [22.47]$$

Un intervalo de confianza $1 - \alpha$, para la diferencia de rangos medios entre dos variables, $D\overline{R}_{i,j}$, se puede calcular mediante la siguiente expresión:

$$D\overline{R}_{i,j} \in (\overline{\hat{D}R}_{i,j} - DMS; \quad \overline{\hat{D}R}_{i,j} + DMS) \qquad [22.48]$$

En la expresión anterior y en las siguientes $D\overline{R}_{i,j}$, es la diferencia entre los rangos medios poblacionales entre la i-ésima y la j-ésima variables; $\overline{\hat{D}R}_{i,j}$ es la diferencia de rangos estimada entre la i-ésima y la j-ésima variables, es decir, la diferencia observada entre los rangos de la i-ésima y la j-ésima variables.

Prueba de Tukey

La diferencia mínima significativa, DMS, mediante la prueba de Tukey, se puede calcular mediante la siguiente expresión:

$$DMS_T = \frac{q_{\alpha;\, K,\, \infty}}{\sqrt{2}} \cdot \sqrt{\frac{K(N+1)}{6}} \qquad [22.49]$$

En la expresión anterior $q_{\alpha;\, K,\, \infty}$ es el valor del rango estudentizado con K e ∞ grados de libertad, que deja a su derecha un área bajo la curva igual a α (tabla VIII de Anexos). Un intervalo de confianza $1 - \alpha$, para la diferencia de rangos medios entre dos variables, $\overline{D}_{R:\, i,\, j}$, se puede calcular mediante la siguiente expresión:

$$D\overline{R}_{i,j} \in (\overline{\hat{D}R}_{i,j} - DMS_T; \quad \overline{\hat{D}R}_{i,j} + DMS_T) \qquad [22.50]$$

Prueba de Dunnett

La diferencia mínima significativa, DMS, mediante la prueba de Dunnett, se puede calcular mediante la siguiente expresión:

$$DMS_D = D_{\alpha/2;\, K,\, \infty} \cdot \frac{K(N+1)}{6} \qquad [22.51]$$

En la expresión anterior $D_{\alpha/2;\, K,\, \infty}$ es el valor del estadístico de Dunnet con K e ∞ grados de libertad, que deja a su derecha un área bajo la curva igual a $\alpha/2$ (tabla IX).

Un intervalo de confianza $1 - \alpha$, para la diferencia de rangos medios entre dos variables, $\overline{D}_{R:\, i,\, j}$, se puede calcular mediante la siguiente expresión:

$$D\overline{R}_{i,j} \in (\overline{\hat{D}R}_{i,j} - DMS_D; \quad \overline{\hat{D}R}_{i,j} + DMS_D) \qquad [22.52]$$

La prueba de Kruskal Wallis con SPSS

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

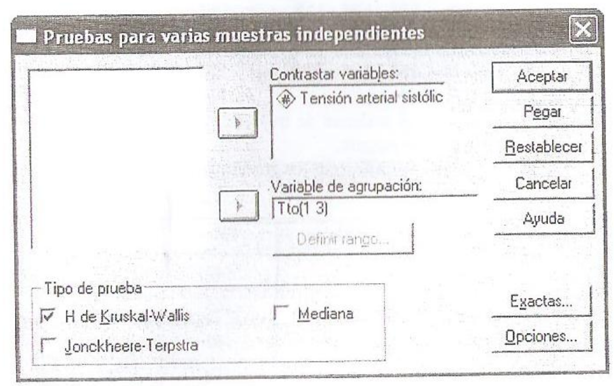

FIGURA 22.18

Se resuelve el ejemplo 22.14 con SPSS, una vez introducidos los datos, las variables que se quieren contrastar Tensión arterial sistólica se pasan a la ventana «Contrastar variables»; la variable Tto se pasa a la ventana «Variable de agrupación», hay que pulsar «Definir rango» para asignar los valores 1 y 3 que definen el intervalo de definición de los tratamientos; pulsando Aceptar se obtienen los resultados siguientes:

Pruebas no paramétricas

Prueba de Kruskal-Wallis

Rangos

	Tratamiento	N	Rango promedio
Tensión arterial sistólica	A	6	6,08
	B	5	11,00
	C	7	11,36
	Total	18	

Estadísticos de contraste [a, b]

	Tensión arterial sistólica
Chi-cuadrado	3,711
gl	2
Sig. asintót.	0,156

[a] Prueba de Kruskal-Wallis.

[b] Variable de agrupación: tratamiento.

En la primera tabla se muestran el número de casos y el rango promedio que corresponde a cada grupo. En la segunda tabla se muestran el número de casos, el valor del estadístico H, los grados de libertad y la significación estadística, que es aproximada $P < 0{,}156$, como es mayor que 0,05 no se rechaza la hipótesis nula. Las conclusiones son las mismas que se expusieron en el ejemplo.

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

22.6.2. La prueba de la mediana para k variables

La prueba de la mediana para K variables es una ampliación de la prueba de la mediana para dos variables; los conceptos fundamentales de ambas coinciden.

Esta prueba es adecuada para comparar si K variables cuantitativas tienen la misma mediana. Los tamaños de las K poblaciones pueden ser distintos. Se denotan $n_1 \ldots n_k$.

En primer lugar se calcula la mediana para todos los datos. Se disponen todos los datos en un mismo grupo y se calcula la mediana global. A continuación, en una tabla 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, la proporción de casos por encima o debajo de la mediana global será significativamente distinta para cada variable.

Las hipótesis son:

$$H_0 \quad M1 = M2 = \ldots MK$$
$$H_1 \quad Mi \neq Mj \text{ para algún } ij \qquad \alpha$$

La hipótesis nula enuncia que las medianas poblacionales son iguales, y la alternativa que al menos la mediana de una de las variables es distinta de las demás. También pueden plantearse contrastes unilaterales.

El contraste de hipótesis anterior se resuelve aplicando la clásica prueba de la Chi-cuadrado, teniendo en cuenta que para la correcta aplicación de la prueba, no debe haber más del 20% de las casillas con valores esperados menores que 5. En caso de que haya más del 20% de las casillas con valores menores de 5, la prueba no es aplicable.

La prueba de la mediana se utiliza poco porque es más potente la de Kruskall-Wallis.

La prueba de la mediana para K muestras independientes con SPSS

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

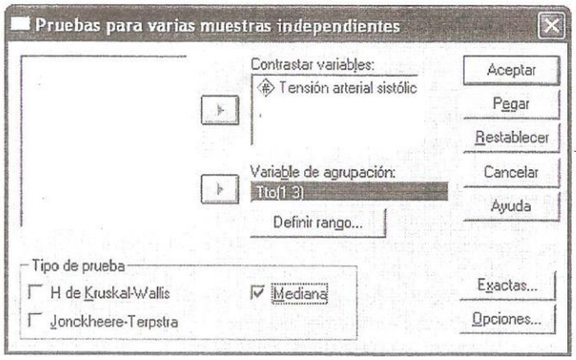

FIGURA 22.19

Se resuelve el ejemplo 22.13 con SPSS, una vez introducidos los datos, las variables que se quieren contrastar; Tensión arterial sistólica se pasa a la ventana «contrastar variables»; la variable Tto se pasa a «Ventana de agrupación», hay que pulsar «Definir rango» para asignar los valores 1 y 3 que definen el intervalo de definición de los tratamientos; pulsando Aceptar se obtienen los resultados siguientes:

Pruebas no paramétricas

Prueba de la mediana

Frecuencias

		A	*B*	*C*
Tensión arterial sistólica	> Mediana	2	3	4
	≦ Mediana	4	2	3

Estadísticos de contraste [b]

	Tensión arterial sistólica
N	18
Mediana	139,00
Chi-cuadrado	1,010 [a]
gl	2
Sig. asintót.	0,604

[a] 6 casillas (100,0%) tienen frecuencias esperadas menores que 5. La frecuencia de casilla esperada mínima es 2,5.

[b] Variable de agrupación: tratamiento.

En la primera tabla se muestra el número de casos de cada grupo que están por encima y por debajo de la mediana global de los datos. Esta tabla es muy útil aunque sólo sea como estadística descriptiva. Si no se puede realizar la prueba analítica porque la muestra es pequeña, permite observar si las diferencias entre los grupos son grandes o pequeñas. En este caso las diferencias no son importantes.

En la segunda tabla se muestran el número de casos; el valor de la mediana global, el valor del estadístico Chi-cuadrado, los grados de libertad y la significación estadística. Debajo de la tabla se indica en porcentaje de casillas con frecuencias esperadas menores que 5, que en este caso es el 100%. No es correcto aplicar la prueba analítica de la mediana porque más del 20% de las casillas tienen frecuencias esperadas menores que 5, pero sirve la tabla como descriptiva. En este caso se recomienda aplicar la prueba de Kruskall-Wallis.

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