22 Estadística no paramétrica

En este capítulo se estudian un conjunto de pruebas estadísticas, agrupadas bajo el epígrafe de Estadística no paramétrica o métodos de distribución libre. Algunas pruebas como la binomial o la bondad del ajuste ya se han estudiado en otros capítulos, no obstante, se hace una introducción a ellas y se remite al lector a los capítulos correspondientes donde se analizan con más extensión. La resolución de las pruebas no paramétricas con SPSS se hace al final del apartado correspondiente a cada una de ellas, en lugar de todas juntas al final del capítulo.

22.1. ESTADÍSTICA PARAMÉTRICA Y NO PARAMÉTRICA

La mayoría de las pruebas estadísticas analizadas en capítulos anteriores para que su aplicación sea correcta, tienen que cumplir una serie de requisitos, como que las variables estadísticas, o la media aritmética muestral, sigan una distribución normal, que las varianzas sean homogéneas, etc. Las pruebas estadísticas que exigen el cumplimiento de requisitos a las variables o a alguno de sus parámetros se denominan *paramétricas*.

El incumplimiento de las exigencias de las pruebas paramétricas, sobre todo la normalidad, es más fácil que ocurra en muestras pequeñas, menores de 30 casos. Por otra parte, las pruebas paramétricas aumentan su potencia estadística en relación con las no paramétricas según aumenta el tamaño de la muestra. En muestras menores de once [1] casos, la potencia estadística de ambos tipos de pruebas es equivalente. Teniendo en cuenta que las exigencias para las pruebas no paramétricas son menores y que detectar violaciones de las condiciones de aplicabilidad

[1] En estadística, sobre todo la analítica, pocas veces son útiles muestras tan pequeñas, pero en ocasiones se utilizan.

en las pruebas paramétricas es más difícil cuando las muestras son pequeñas, se hace aconsejable utilizar pruebas estadísticas no paramétricas cuando las muestras tienen menos que once casos, aunque la potencia estadística de estas muestras es muy pequeña, sea cual sea la prueba que se utilice. Las pruebas no paramétricas pueden usarse para muestras de cualquier tamaño cuando no se cumplan los criterios de aplicabilidad de las pruebas paramétricas. Éstas son preferibles cuando cumplen los requisitos de aplicabilidad porque su potencia es mayor y, también, su plasticidad.

Otra circunstancia que influye en la popularidad creciente de estas pruebas es que los conocimientos estadísticos para comprenderlas son mínimos, mientras que los conocimientos estadísticos necesarios para comprender y aplicar correctamente las pruebas paramétricas son mucho mayores.

Las pruebas no paramétricas, también se denominan *pruebas de distribución libre*, debido a que en la mayoría de ellas no es importante cómo se distribuyan los datos. Las pruebas no paramétricas pueden aplicarse a datos que sigan distribuciones normales o no, lo cual implica que estas pruebas pueden realizarse incluso en casos en los que se cumplen las condiciones de aplicabilidad de las pruebas paramétricas.

Hay un requisito imprescindible para realizar cualquier prueba estadística analítica, paramétrica y no paramétrica: que la muestra sea aleatoria.

Las pruebas no paramétricas se clasifican para su estudio, según el número de muestras, y la relación existente entre ellas:

- Una sola muestra.
- Dos muestras relacionadas (dependientes).
- Dos muestras independientes.
- Varias muestras relacionadas (dependientes).
- Varias muestras independientes.

En este capítulo las pruebas no paramétricas se analizan en el mismo orden del listado anterior.

22.2. PRUEBAS PARA UNA SOLA MUESTRA

En este apartado se estudian las pruebas no paramétricas aplicables a una sola muestra. Son muy útiles para contrastar hipótesis sobre la posibilidad de que en la población muestreada las variables tengan unas características determinadas.

Las pruebas no paramétricas más utilizadas basadas en una sola muestra son: binomial, Chi-cuadrado, Kolmogorov-Smirnov y la prueba de las rachas.

22.2.1. Prueba binomial

En ciencias de la salud y ciencias sociales es frecuente el uso de variables dicotómicas, las cuales tienen dos valores posibles; por ejemplo, hombre y mujer,

enfermo y no enfermo, sí y no, etc. También se pueden dicotomizar, para análisis concretos, variables cuantitativas; por ejemplo, si se dispone de datos de la tensión arterial de una muestra de personas, puede interesar conocer la proporción que tienen un valor menor o igual que ciento cuarenta milímetros de Hg, y la proporción que tienen un valor mayor, es decir, para este análisis concreto la variable cuantitativa es dicotómica; al valor que divide en dos los valores de la variable se le denomina *punto de corte*.

La prueba binomial [1] permite comprobar hipótesis sobre la proporción de casos que hay en una población con un determinada característica. Por ejemplo, si la proporción de mujeres en una población es distinta del 50%, o si la proporción de personas que han padecido de catarro común en los últimos seis meses es mayor que el 40%, si la proporción de individuos con colesterol basal mayor que 225 mg/100 ml es mayor que el 60% etc.

Si se selecciona una muestra aleatoria de tamaño n de una población, y se define una variable dicotómica, con valores posibles A o B, siendo p la proporción de A, y $q = (1 - p)$, es decir, la proporción de B. La probabilidad de obtener K casos con valor A, si se selecciona una muestra aleatoria de n elementos se puede calcular mediante la siguiente expresión:

$$P(K) = \binom{n}{K} p^K q^{(n-K)} \qquad [22.1]$$

En la expresión anterior, $P(K)$ es la probabilidad de que de n casos K tengan el valor A; evidentemente, $K \leq n$.

La expresión anterior es aplicable, si $p \geq 0{,}05$ y $q \leq 0{,}95$. Si la probabilidad es pequeña, menor que 0,05, la distribución de probabilidad aplicable es la de Poisson.

EJEMPLO 22.1

Si se tira al aire seis veces seguidas una moneda perfecta: ¿cuál es la probabilidad de sacar dos caras?

El resultado de tirar una moneda al aire puede ser cara o cruz (variable dicotómica); si la moneda es perfecta, la probabilidad de sacar cara (p) es 0,5 y la probabilidad de sacar cruz (q) también es 0,5. La probabilidad de sacar exactamente dos caras es:

$$P(2) = \binom{6}{2} 0{,}5^2 \cdot 0{,}5^4; \qquad P(2) = 15 \cdot 0{,}25 \cdot 0{,}0625; \qquad P(2) = 0{,}234$$

[1] En el Capítulo 7 se estudia con detalle la distribución binomial, se resuelven varios ejemplos.

Si el tamaño de la muestra n es mayor que 30 [1], se puede aproximar la distribución de probabilidad a una distribución normal, con $\mu = np$ y $\sigma^2 = npq$. Con estos parámetros, la abscisa normal tipificada z es:

$$Z = \frac{((X \pm 0{,}5) - np)}{\sqrt{npq}} \qquad [22.2]$$

Una vez obtenida z, se calcula su probabilidad consultando las correspondientes tablas de la distribución normal tipificada. Sumar o restar 0,5 a X es una corrección por continuidad; se suma 0,5 a x cuando es menor que np, y se resta 0,5 a x cuando es mayor que np.

Las hipótesis sobre la proporción poblacional de una determinada característica puede ser de dos colas o de una cola:

$$H_0 \quad p = a$$
$$H_1 \quad p \neq a \qquad \alpha$$

Si se quiere contrastar la hipótesis de que en una población la proporción de mujeres es distinta de 0,5, con $\alpha = 0{,}05$, el contraste se plantea de la manera siguiente:

$$H_0 \quad p = 0{,}5 \qquad \alpha = 0{,}05$$
$$H_1 \quad p \neq 0{,}5$$

El contraste de hipótesis anterior es de dos colas; se rechaza la hipótesis nula, tanto si p es significativamente mayor que 0,5 como si es significativamente menor.

Los contrastes de una cola que se pueden plantear son los siguientes:

$$H_0 \quad p \leq a$$
$$H_0 \quad p > a \qquad \alpha$$

$$H_0 \quad p \geq a$$
$$H_1 \quad p < a \qquad \alpha$$

Si se tiene la hipótesis de que en una población la proporción de hipertensos es mayor que 0,12, el contraste de hipótesis se plantea de la siguiente manera:

$$H_0 \quad p \leq 0{,}12$$
$$H_1 \quad p > 0{,}12 \qquad \alpha = 0{,}05$$

[1] SPSS hace automaticamente la aproximación a la normal si n es mayor que 25.

En el caso anterior sólo se considera la posibilidad de que p sea mayor o igual que 0,12.

EJEMPLO 22.2

Se tiene la hipótesis de que la proporción de hipertensos en una población es mayor que 0,18. Para comprobarlo se selecciona una muestra aleatoria de 100 personas, de las cuales 27 son hipertensas. Plantear las hipótesis y resolver el contraste con $\alpha = 0{,}05$.

Las hipótesis son las siguientes:

$$H_0 \quad p \leq 0{,}18 \qquad \alpha = 0{,}05$$
$$H_1 \quad p > 0{,}18$$

El tamaño de la muestra es mayor que 30, se puede hacer la aproximación a la normal. El valor de z que deja a la derecha un área de 0,05 es 1,645. Si el valor del estadístico de contraste es mayor que 1,645, se rechaza la hipótesis nula y se concluye que la proporción de hipertensos en la población es mayor que 0,18; si es menor se considera que no hay pruebas estadísticas de que la proporción sea mayor que 0,18.

Aplicando la expresión 22.2:

$$Z = \frac{(27 + 0{,}5) - (100 \cdot 0{,}18)}{\sqrt{100 \cdot 0{,}18 \cdot 0{,}82}}; \qquad Z = 2{,}47$$

Como el valor del estadístico de contraste es mayor que 1,65 se rechaza la hipótesis nula y se concluye que la proporción de hipertensos es mayor que 0,18, con $P < 0{,}0068$.

La prueba binomial con SPSS

Como ejemplo para estudiar la prueba binomial con SPSS, se utilizan los siguientes datos:

Fumar: 1, 2, 1, 2, 1, 2, 1, 2, 2, 2, 1, 2, 1, 2, 1, 2, 1, 2, 2, 2, 2, 2, 2

La variable anterior se ha codificado con un uno para SÍ y un dos para NO. Introduzca los datos en la base de datos de SPSS.

En el menú analizar seleccione Pruebas no paramétricas y en el listado de pruebas binomial. Aparece la siguiente pantalla:

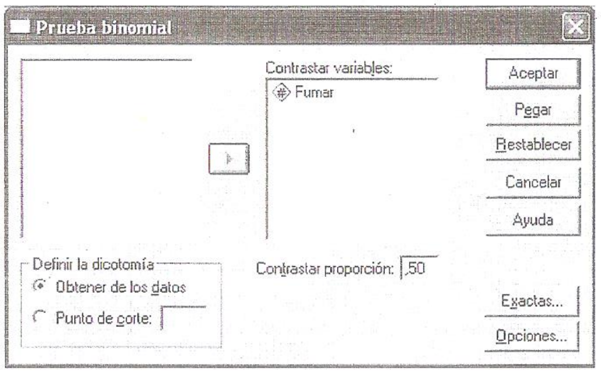


FIGURA 22.1

La variable que se quiere contrastar, en este caso Fumar, se pasa a la ventana encabezada por contrastar variables.

En la parte inferior a la izquierda «Definir la dicotomía» se puede optar por que se obtenga la dicotomía de los datos o mediante punto de corte. Obtener de los datos se utiliza cuando la variable que se quiere contrastar es dicotómica, es decir, sólo tiene dos valores definidos. La opción punto de corte se utiliza para variables numéricas con más de dos valores, se consideran pertenecientes a la primera categoría los valores menores o iguales que el valor especificado a la derecha de «Punto de corte», pertenecientes a la otra categoría se consideraran los valores mayores que el punto de corte.

En mitad de la pantalla en la parte inferior está «Contrastar proporción» por defecto el valor indicado es 0,50, este valor puede cambiarlo el usuario.

Dejando el valor 0,50 en contrastar proporción, las hipótesis que se contrastan son las siguientes:

$$H_0 \quad P(\text{SI}) = P(\text{NO}) = 0{,}5$$
$$H_1 \quad P(\text{SI}) \neq P(\text{NO}) \qquad \alpha = 0{,}05$$

En este caso las dos proporciones son iguales, cuando son diferentes, la proporción especificada en «Contrastar proporción» es la correspondiente a la primera categoría.

Pulsando en «Aceptar» se obtienen los siguientes resultados:

Pruebas no paramétricas

Prueba binomial

		Categoría	*N*	Proporción observada	Prop. de prueba	Sig. exacta (bilateral)
Fumar	Grupo 1	Sí	8	0,33	0,50	0,152
	Grupo 2	No	16	0,67		
	Total		24	1,00		

En la tabla anterior se muestran los valores observados en cada categoría: 8 para SÍ y 16 para NO, la proporción de cada una de ellas, 0,33 y 0,67 respectivamente, la proporción de prueba que es la asignada para realizar el contraste, y la significación estadística que en este caso es 0,152, como es mayor que 0,05, no se rechaza la hipótesis nula y se concluye que no hay pruebas estadísticas de que la proporción de la primera categoría sea diferente de 0,5. Recuerde que no rechazar la hipótesis nula no quiere decir que se haya demostrado, de hecho, tampoco se rechazaría para muchos otros valores; pruebe con 0,33, 0,40, 0,38..., verá que tampoco se rechaza la hipótesis nula para estos valores, sería absurdo considerar que se ha demostrado que la proporción es igual a todas ellas.

En la parte inferior derecha de la pantalla 22.1, hay dos posibilidades: «Exactas» y «Opciones».

La primera no se incluye en todas las versiones y permite realizar los cálculos mediante el método de Montecarlo, puede ser útil para hacer simulaciones.

TÉCNICA COMBINA CONCEPTOS (MUESTREO ALEATORIO)

Pulsando opciones aparece la siguiente pantalla:

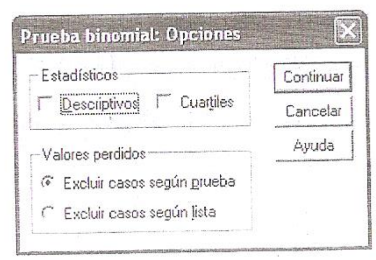


FIGURA 22.2

Si se marca «Descriptivos», en los resultados aparece una tabla con la media, la desviación típica, el número de casos, el máximo y el mínimo. Si se marca

«Cuartiles» se obtienen los valores correspondientes a los cuartiles. Tanto los descriptivos como los cuartiles tienen sentido si la variable es cuantitativa continua, o discreta con muchas categorías. Esta opción puede ser útil cuando se usan variables dicotomizadas mediante punto de corte.

En valores perdidos hay dos posibilidades: «Excluir casos según prueba» y «Excluir casos según lista». La opción por defecto es la primera; en este caso si se especifican varias variables para hacer contrastes, se excluyen los casos con valores ausentes al realizar cada uno de ellos; si se marca «Excluir casos según lista», se excluyen de la realización de todas las pruebas los casos que tengan valores ausentes en alguna de las variables.

22.2.2. Bondad del ajuste: prueba χ^2

Una de las pruebas más utilizadas para analizar el ajuste de datos experimentales a distribuciones teóricas es la basada en la distribución χ^2. En el Capítulo 14 se analiza esta prueba con detalle y se realizan varios ejemplos; también se estudia cómo realizar ajustes de datos a distribuciones teóricas con SPSS.

La prueba compara las frecuencias observadas con las esperadas, bajo la hipótesis de que en la población muestreada la variable se ajusta a una distribución teórica determinada. El estadístico de contraste es el siguiente:

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \qquad [22.3]$$

En la expresión anterior, O_i representa las frecuencias relativas observadas y E_i a las frecuencias relativas esperadas, K es el número de categorías que se comparan. Los grados de libertad del estadístico son $K - 1$.

La prueba Chi-cuadrado no es aplicable si más del 20% de las frecuencias esperadas son menores que 5.

Las hipótesis se plantean de la manera siguiente:

H_0 Los datos se ajustan a la distribución teórica.
H_1 Los datos no se ajustan a la distribución teórica.

El problema de estos contrastes de hipótesis es que si no se rechaza la hipótesis nula [1], no puede considerarse demostrado que los datos se ajusten a la distribución teórica; sin embargo, si se rechaza la hipótesis nula sí puede considerarse que los datos no se ajustan a la distribución propuesta. Es frecuente que en artículos publicados en revistas e incluso en libros considerar demostrada la hipótesis nula si no se ha rechazado, y se consideran los datos como si se ajustaran a la distribución teórica propuesta, esto es un error importante y muchas veces trascendente.

[1] En el Capítulo 11 se analizan las características de los contrastes de hipótesis.

22.2.3. Pruebas de Kolmogorov-Smirnov; Kolmogorov-Smirnov-Lilliefors y Shapiro Wilks

Esta prueba, desarrollada por Kolmogorov para una muestra y junto con Smirnov para dos muestras, se conoce en ambos casos como prueba de Kolmogorov-Smirnov, y originalmente se desarrolló para comprobar si la distribución empírica de una variable cuantitativa, es decir, la distribución observada de una variable, se ajusta a una distribución teórica conocida; por ejemplo, distribución normal, distribución exponencial, etc. Después se comenzó a aplicar a distribuciones discretas.

La prueba de Kolmogorov-Smirnov se basa en comparar los valores absolutos de las diferencias entre las frecuencias relativas acumulativas experimentales u observadas, $F_0(x)$, y las frecuencias relativas acumulativas teóricas o esperadas, $F_e(x)$. Las hipótesis son:

$$H_0 \quad F_0(x) = F_e(x)$$
$$H_1 \quad F_0(x) \neq F_e(x) \qquad \alpha$$

El estadístico de contraste es el siguiente:

$$\text{Máx } (D = |F_0(x) - F_e(x)|) \qquad [22.4]$$

Entre las frecuencias observadas y esperadas se permiten pequeñas diferencias explicables por el azar. El punto crítico, PC, del contraste se obtiene de una tabla si $n \leq 35$; si $n > 35$ el punto crítico para $\alpha = 0{,}05$, se obtiene mediante la siguiente expresión:

$$\text{PC} = \frac{1{,}36}{\sqrt{n}} \qquad [22.5]$$

Si n es mayor de 35 y $\alpha = 0{,}01$ el punto crítico, PC, se obtiene mediante la siguiente expresión:

$$\text{PC} = \frac{1{,}63}{\sqrt{n}} \qquad [22.6]$$

Si la máxima diferencia observada es mayor que el punto crítico se rechaza la hipótesis nula, en caso contrario no se rechaza.

El principal problema de esta prueba es que se considera que los datos se ajustan a la distribución teórica si no se rechaza la hipótesis nula. Ya se ha comentado en múltiples ocasiones en este libro que las hipótesis nulas no se pueden demostrar estadísticamente; pero si se rechaza la hipótesis nula se considera demostrada la alternativa, es decir, se considera demostrado que no se ajusta a la distribución teórica propuesta, aunque indicando la significación estadística. La prueba de Kolmogorov-Smirnov ha sido muy utilizada para evaluar el ajuste de variables continuas a una distribución normal, en este caso las hipótesis son las siguientes:

H_0 La variable se distribuye normalmente.
H_1 La variable no se distribuye normalmente.

Es relativamente frecuente encontrar artículos «científicos», incluso libros de estadística, que digan que se ha demostrado que los datos se distribuyen normalmente o que se ajustan a una normal en la población muestreada; esto es claramente erróneo, se podría decir que no se ha rechazado la hipótesis de normalidad, pero nunca que se ha demostrado la normalidad de los datos. Otra cosa es que si no se rechaza la hipótesis de normalidad y el tamaño de la muestra no es muy pequeño se apliquen pruebas que exigen la normalidad, teniendo en cuenta que no se ha rechazado dicha hipótesis.

La prueba de Kolmogorov-Smirnov se considera muy conservadora, es decir, es difícil rechazar la hipótesis nula, sobre todo si la variable es discreta y, aunque sea continua, si hay que estimar los parámetros de la distribución, como la media y la desviación típica en el caso de la distribución normal.

EJEMPLO 22.3

Un determinado tumor pulmonar se clasifica en cinco tipos distintos, en cuanto a la diferenciación celular, se cree que las cinco se presentan en la misma proporción, es decir, un veinte por ciento, un quinto. Se selecciona una muestra al azar de veinte tumores, obteniéndose las siguientes frecuencias absolutas.

Tipo celular:	1	2	3	4	5
Frecuencia:	4	8	2	2	4

Resolver el contraste con $\alpha = 0{,}05$.

H_0 Los cinco tipos celulares tienen la misma proporción.
H_1 Al menos una proporción es distinta de las demás.

La tabla de frecuencias relativas acumuladas es la siguiente:

Tipo celular	$F_0(x)$	$F_e(x)$	$D = \lvert F_0(x) - F_e(x) \rvert$
1	1/5 = 0,2	4/20 = 0,2	0
2	2/5 = 0,4	12/20 = 0,6	0,2
3	3/5 = 0,6	14/20 = 0,7	0,1
4	4/5 = 0,8	16/20 = 0,8	0
5	5/5 = 1	20/20 = 1	0

La máxima diferencia entre las frecuencias relativas acumuladas es 0,2, consultando en la tabla XI de los Anexos correspondiente a la prueba de Kolmogorov, el punto crítico para $n = 20$ y $\alpha = 0{,}05$, es 0,294, como el valor experimental es menor, no se puede rechazar la hipótesis de que las proporciones de los tipos celulares sean iguales.

La prueba de Kolmogorov-Smirnov con SPSS

Seleccione en el menú análisis pruebas no paramétricas y en el listado de pruebas disponible: *K-S* de 1 muestra. Aparece la pantalla siguiente:

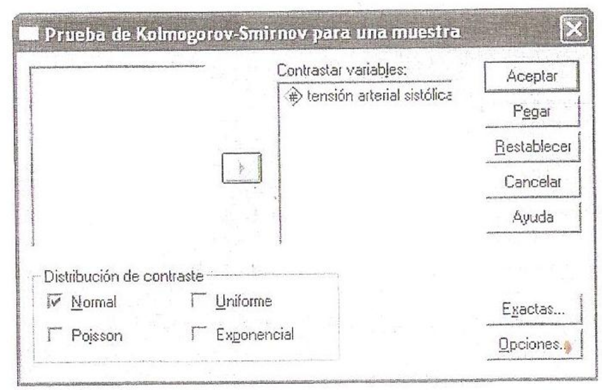


FIGURA 22.3

En la parte inferior a la izquierda en «Distribución de contraste» se puede realizar el ajuste mediante la prueba de Kolmogorov-Smirnov para las distribuciones normal, uniforme, Poisson y exponencial.

Se va a realizar un ajuste a la distribución normal con los 31 datos siguientes que son tensiones arteriales sistólicas, TAS, medidas en milímetros de Hg:

TAS: 115 120 125 125 130 132 132 136 139 139 141 142 142 142 142 144
145 146 146 151 152 152 154 155 160 161 162 162 164 165 171

Una vez introducidos los datos, pase la variable TAS a la ventana encabezada por «contrastar variables» y pulse Aceptar, se obtienen los siguientes resultados:

Pruebas no paramétricas

Prueba de Kolmogorov-Smirnov para una muestra

		Tensión arterial
N		31
	Media	144,90
Parámetros normales [a, b]	Desviación típica	14,032
Diferencias más extremas	Absoluta	0,085
	Positiva	0,082
	Negativa	−0,085
Z de Kolmogorov-Smirnov		0,472
Sig. asintót. (bilateral)		0,979

[a] La distribución de contraste es la Normal.
[b] Se han calculado a partir de los datos.

En la tabla de resultados en la primera fila se muestra el tamaño de la muestra, que es 31; en la segunda y tercera filas la media y la desviación típica de los datos, que son los parámetros que se utilizan para la distribución normal a la que se ajustan los datos; en las tres filas siguientes se muestran las máximas diferencias entre las frecuencias relativas acumuladas en valor absoluto, la mayor positiva y la mayor negativa; la que se usa para realizar el contraste es la máxima en valor absoluto, es decir 0,85; en la fila siguiente se muestra la Z de Kolmogorov-Smirnov y la significación, que como es mayor que 0,05 no permite rechazar la hipótesis nula que es la de normalidad.

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

Ajustes de normalidad: las pruebas de Kolmogorov-Smirnov Lilliefors y de Shapiro Wilks

En la actualidad las pruebas más aplicadas para el ajuste de los datos a una distribución normal son la de Kolmogorov-Smirnov, con la modificación de Lilliefors si el tamaño de la muestra es mayor que 30 y la de Shapiro Wilks si el tamaño de la muestra es menor o igual que 30.

Lilliefors realizó una modificación en la prueba de Kolmogorov-Smirnov para el ajuste de datos de una variable cuantitativa a una distribución normal, mejorando su potencia estadística. Observe que la prueba de Kolmogorov-Smirnov se desarrolló para ajustar los datos a cualquier distribución de variables continuas, también se utilizó para el ajuste a distribuciones de variables discretas, mientras que la modificación de Lilliefors es válida únicamente para el ajuste a distribuciones normales.

Las hipótesis que contrastar son:

H_0 La variable se distribuye normalmente.
H_1 La variable no se distribuye normalmente. α

El estadístico de contraste es el mismo que en la prueba de Kolmogorov-Smirnov: Máx ($D = |F_0(x) - F_e(x)|$). Cambia la tabla de puntos críticos. La estimación de los parámetros de la curva normal: media y desviación típica, se realiza a partir de los parámetros muestrales. Es la prueba más adecuada si $n > 30$.

La prueba de Shapiro-Wilks es la más utilizada si $n \leq 30$. Su fundamento es comparar cuantil [1] a cuantil, el valor esperado bajo la hipótesis de que los datos se distribuyen según una normal con la media y la desviación típica de los datos, con el cuantil observado:

$$E\left(\frac{x_i - \mu}{\sigma}\right) = c_{i,n} \quad [22.7]$$

Despejando en la ecuación anterior y teniendo en cuenta que el valor esperado de una constante es ella misma:

$$E(x_i) = \mu + c_{i,n}\sigma \quad [22.8]$$

La expresión anterior es la ecuación de una recta en la que la variable dependiente es el valor esperado bajo la hipótesis de normalidad y la variable independiente el cuantil observado.

El estadístico de contraste se basa en el cuadrado del coeficiente de correlación entre la variable dependiente y la independiente de la ecuación anterior.

[1] Cuantil a cuantil quiere decir dato a dato.

Con SPSS se va a estudiar el ajuste a la normal mediante las dos pruebas anteriores con los datos ajustados en la prueba de Kolmogorov-Smirnov con SPSS.

Las pruebas de Shapiro-Wilks y de Kolmogorov-Smirnov Lilliefors con SPSS se obtienen seleccionando en analizar estadísticos descriptivos y dentro de las opciones «Explorar» pulsando se obtiene la pantalla siguiente.

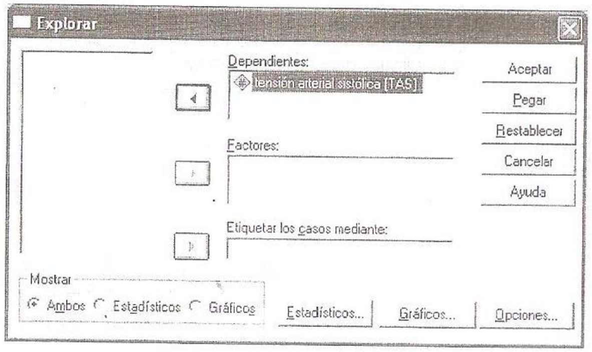


FIGURA 22.5

En la ventana «Dependientes» debe estar la variable que se quiere ajustar. Pulsando la tecla «Gráficos» se obtiene la pantalla siguiente:

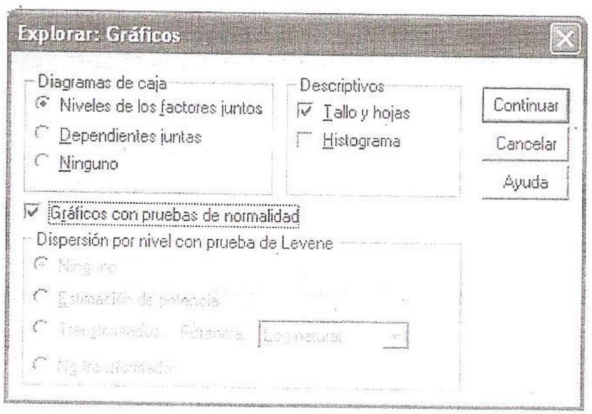


FIGURA 22.6

Si se quiere realizar el ajuste a la normal hay que marcar «Graficos con pruebas de normalidad».

Se obtienen tablas de estadísticos descriptivos y un histograma. Además, que es lo que se va a comentar en este apartado, una tabla con las pruebas que se quieren realizar y dos gráficos de normalidad:

Pruebas de normalidad

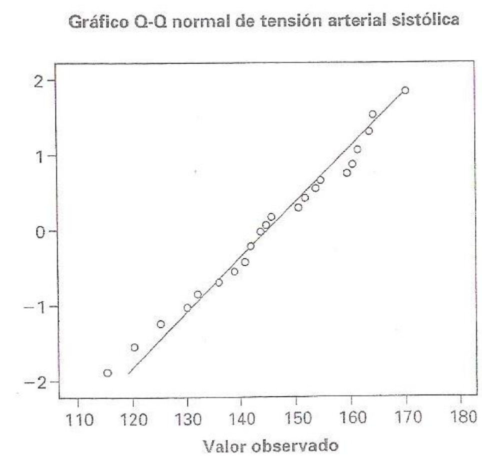
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Estadístico	**gl**	**Sig.**	**Estadístico**	**gl**	**Sig.**
Tensión arterial sistólica	0,085	31	0,200*	0,979	31	0,781

* Este es un límite inferior de la significación verdadera.
[a] Corrección de la significación de Lilliefors.

Como hay 31 casos, la prueba aplicable es la de Kolmogorov-Smirnov Lilliefors. La máxima diferencia observada entre las frecuencias relativas acumuladas observada y esperada bajo la hipótesis de normalidad es 0,085, que, según la tabla de Lilliefors, corresponde a una significación estadística de 0,2; como es mayor que 0,05 no se rechaza la hipótesis nula, que es la de normalidad.

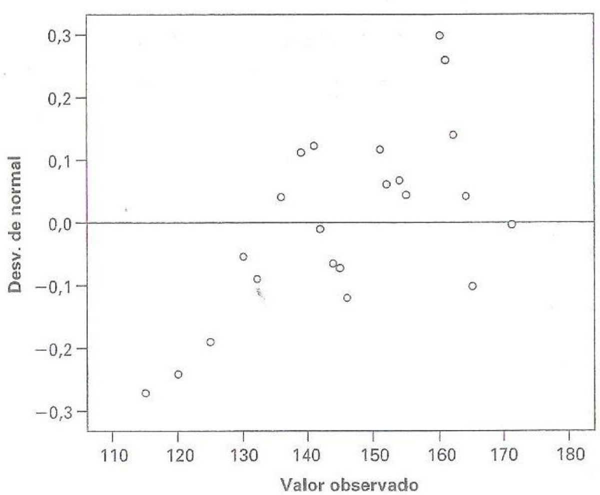
El estadístico de Shapiro-Wilks es 0,979, que se corresponde con una significación de 0,781, que es mayor que 0,05, consecuentemente, no se puede rechazar la hipótesis de normalidad.

No siempre coinciden las dos pruebas en el rechazo o no de la hipótesis nula, se debe utilizar la que corresponda al tamaño de la muestra.



En el caso de un ajuste perfecto a una distribución normal los puntos estarían en la recta.

Gráfico Q-Q normal sin tendencias de tensión arterial sistólica



En una normal perfecta los datos deben distribuirse por encima y debajo de la línea cero al azar.

22.2.4. Pruebas de aleatoriedad: prueba de las rachas

Las tres pruebas analizadas anteriormente analizan la bondad del ajuste de un conjunto de datos a distribuciones teóricas; tienen en cuenta si las frecuencias observadas difieren significativamente de las frecuencias esperadas en el supuesto de que la variable siga una determinada distribución estadística.

La prueba de las rachas analiza si el orden del muestreo es compatible con la aleatoriedad, tiene en cuenta las frecuencias y el orden de los valores en el muestreo. Podría ocurrir que de 20 observaciones 10 tuvieran un valor y 10 otro, pero su orden de observación no fuera el adecuado para considerar la aleatoriedad del muestreo. Para poder realizar esta prueba, es necesario conocer el orden de observación de los datos. Si se ha alterado este orden por haber clasificado los datos

según otros criterios, de menor a mayor o por cualquier otro, la prueba no tendrá valor.

La correcta aplicación de las pruebas de aleatoriedad exige que el muestreo sea aleatorio, que se apliquen sobre el orden de observación y que se planteen las hipótesis estadísticas antes de realizar el experimento.

Existen varias pruebas de aleatoriedad, la más utilizada es la de las rachas, *runs* en inglés. Se denomina racha a una sucesión de observaciones con el mismo valor. La prueba se aplica a dos valores, si la variable no es dicotómica hay que dicotomizarla para realizarla. En el caso de variables cuantitativas se considera a la mediana punto de corte, puesto que por definición, es el valor que divide los datos en dos partes iguales en cuanto al número de casos; la mitad de los valores son menores que la mediana y la otra mitad mayores, se pueden denotar con un signo − a los datos menores que la mediana, y con un signo + a los datos mayores que la mediana.

Como se ha dicho, una racha es una sucesión de valores idénticos. Por ejemplo, si se tira una moneda al aire 15 veces se anota el orden de aparición de los valores cara, *C*, y cruz, *S*; si se obtiene la siguiente secuencia, por orden de observación: *C C S C S C S S S S C C C C C*, las dos primeras observaciones, *C C*, constituyen una racha, puesto que son dos observaciones idénticas, la tercera observación es *S*, que es otra racha, la cuarta observación, *C*, constituye otra racha, la quinta observación, *S*, constituye otra racha, la sexta observación, *C*, constituye otra racha, la séptima, octava, novena y décima observaciones, *S S S S*, constituyen otra racha, y las cinco ultimas observaciones, *C C C C C*, constituyen otra racha; por lo tanto, la secuencia anterior consta de 7 rachas. Una racha comienza a contarse cuando cambia el valor de la observación y se considera hasta que vuelve a haber un cambio en el valor de la variable. Independientemente del número de valores repetidos que salgan, se considera racha tanto una sucesión de un sólo valor como una sucesión de 10 valores: en ambos casos se contabiliza una racha.

En la prueba de las rachas las hipótesis son:

H_0 La muestra es aleatoria.
H_1 La muestra no es aleatoria. α

Tanto si hay pocas rachas como si hay muchas, puede deberse a dependencia entre los resultados. En caso de que haya pocas rachas, podría deberse a que un determinado valor favorece que éste se repita. Si hay muchas rachas, podría ser porque si se obtiene un valor disminuye la probabilidad de que éste se repita.

El máximo y el mínimo número de rachas permitido para aceptar la aleatoriedad está tabulado (Tabla XII de Anexos), depende de n_1 y n_2, que representan el número de veces que se ha observado cada valor de la variable. Recuérdese que la prueba de las rachas se aplica sobre variables dicotómicas. En el ejemplo anterior, $n_1 = 9$ y $n_2 = 6$, puesto que se han observado 9 caras y seis cruces. En el apéndice de tablas hay dos para evaluar la prueba de las rachas, para un nivel de significación 0,05. La primera indica el número mínimo de rachas aceptable, y la segunda el número máximo. Si el valor observado del número de rachas está entre los dos valores, se considera que no hay pruebas estadísticas para rechazar la hipótesis nula y que el orden del muestreo es compatible con la aleatoriedad; si es mayor que el número mínimo o que el número máximo, se considera que la probabilidad de que el muestreo sea aleatorio es menor que 0,05, y se puede rechazar la hipótesis nula. En el ejemplo de las 15 tiradas de monedas en el que se observan 7 rachas, consultando la tabla I con $n_1 = 6$ y $n_2 = 9$, el número mínimo de rachas es 4; la tabla II indica que el número máximo de tiradas es 13; como el número de rachas observado es 7, que está entre los dos valores, no se rechaza la hipótesis de aleatoriedad.

Las tablas para la prueba de las rachas están hechas para muestras relativamente pequeñas. Es difícil, encontrar tablas con valores para n_1 o n_2 mayores que 20. Si alguno de los valores se repite más de 20 veces, se puede utilizar un estadístico de contraste que se distribuye normalmente. Los parámetros de este estadístico, media y desviación típica se calculan según las siguientes expresiones:

$$\mu_r = \frac{2n_1n_2}{n_1 + n_2} + 1 \qquad [22.9]$$

$$\sigma_r = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}} \qquad [22.10]$$

$$Z = \frac{r - \mu_r}{\sigma_r} \qquad [22.11]$$

En la expresión anterior r es el número de rachas.

$$Z = \frac{r - \left(\dfrac{2n_1n_2}{n_1 + n_2} + 1\right)}{\sqrt{\dfrac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}} \qquad [22.12]$$

En las expresiones anteriores, n_1 y n_2 indican el número de observaciones de cada valor de la variable, y r el número de rachas observadas. La media y la desviación típica anteriores, representan el número medio de rachas y la desviación típica esperadas, bajo el supuesto de que los valores de la variable son independientes, es decir, el orden es aleatorio. Se admiten diferencias respecto a estos valores que explique el azar y, por tanto, que la probabilidad de obtenerlos no sea demasiado pequeña; si la probabilidad es menor que 0,05, este es el valor de la

significación estadística más utilizado, se rechaza la hipótesis nula y se concluye que las observaciones no son aleatorias. Se evalúa el contraste en función de los valores de Z obtenidos; para un contraste bilateral con $\alpha = 0{,}05$, se rechaza la hipótesis nula si Z es mayor o menor de 1,96.

La prueba de las rachas con SPSS

Seleccione en el menú análisis pruebas no paramétricas y en el listado de pruebas disponible: Rachas. Aparece la pantalla siguiente:

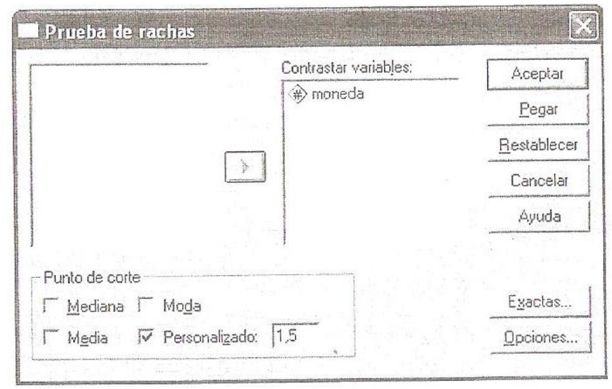


FIGURA 22.7

Se observa que en la parte inferior a la izquierda hay cuatro opciones para definir el punto de corte: la Mediana, la Moda, la Media y Personalizado. En el caso de variables cuantitativas se selecciona la Mediana, en el caso de variables dicotómicas seleccionar Personalizado y un valor entre los valores asignados a las dos categorías, se recomienda el punto medio.

Para realizar el ejemplo de las 15 tiradas, se ha denominado a la variable Moneda, y se ha codificado cara con un uno y cruz con un dos. Observe que si se indica que se considere a la mediana como punto de corte que en este caso es 1, no se puede hacer la prueba, no siempre ocurre esto en variables dicotómicas, pero a veces sí y éste es un ejemplo; por eso se aconseja desmarcar la mediana y dar un valor entre los dos valores codificados, observe, que en la pantalla de ejemplo se ha marcado la opción personalizar y se ha especificado el valor 1,5. Se obtienen los resultados siguientes:

Pruebas no paramétricas

Prueba de rachas

	Moneda
Valor de prueba[a]	1,50
Casos en total	15
Número de rachas	7
Z	−0,392
Sig. asintót. (bilateral)	0,695

[a] Especificado por el usuario.

En la tabla anterior se indica el punto de corte, denominado valor de prueba, el resultado es el mismo especificando cualquier valor mayor que uno y menor que dos; en la segunda fila se indica el número de casos; en la tercera el número de rachas; en la cuarta el valor de Z según la expresión 22.12, y la significación estadística, es decir, la P que es 0,695, como es mayor que 0,05 no se rechaza la hipótesis nula.

Observe que SPSS realiza la prueba de las rachas evaluando Z según la expresión 22.12, y no según las tablas. La evaluación de la Z permite ser más preciso en el cálculo de la significación estadística.

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. PRUEBAS NO PARAMÉTRICAS CON DOS VARIABLES RELACIONADAS

En muchos trabajos de investigación se comparan datos de una variable en dos circunstancias distintas. Se considera que dos 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.

En el caso de muestras relacionadas, la circunstancia diferenciadora suele ser el tiempo, aunque no siempre. Por ejemplo, comparar la tensión arterial sistólica en un grupo de pacientes antes y después de un determinado tratamiento; comparar la puntuación de un grupo de alumnos en matemáticas antes y después de realizar un determinado curso; comparar la calidad de un determinado producto antes y después de modificar las técnicas de producción, etc.

El planteamiento de estudios sobre muestras relacionadas tiene la ventaja de poder eliminar la influencia de variables extrañas, al ser cada individuo el control de sí mismo, aunque en función del estudio, hay que evaluar convenientemente las condiciones ambientales y asegurarnos de que algunos cambios no son producto del transcurrir del tiempo, sobre todo si entre el antes y el después el tiempo es largo; de hecho, excepto en casos muy concretos, la evaluación terapéutica no se suele hacer con muestras relacionadas debido a lo engañosos que pueden ser los resultados[1].

Las pruebas no paramétricas para comparar dos muestras relacionadas, son las siguientes:

- Prueba de los signos.
- Prueba de Wilcoxon.

La prueba de Mcnemar, que muchos autores incluyen entre las pruebas no paramétricas con dos muestras relacionadas, se estudia en el Capítulo 14.

22.3.1. Prueba de los signos para dos variables relacionadas

Esta prueba es aplicable a variables cuantitativas, después de una circunstancia que puede influir en sus valores; se evalua el resultado con un signo más, +, si un valor ha aumentado y con un signo menos, −, si ha disminuido. Por ejemplo, un grupo de estudiantes obtiene una determinada puntuación en matemáticas, después de un cursillo se les vuelve a evaluar, se indica con un signo, +, que han mejorado su puntuación anterior; con un signo, −, que ha disminuido y con cero los casos en los que no haya cambio; si un alumno pasa de un 3 a un 8, y otro de 4 a 5, a los dos se les califica con +, puesto que la nota ha subido, pero sin tener en cuenta cuánto ha subido.

En caso de que la circunstancia diferenciadora no tenga efecto global sobre la variable, se espera encontrar una proporción similar de signos positivos y negativos, es decir, 0,5. Son admisibles pequeñas diferencias explicables por el azar. Si la diferencia entre la proporción de signos es estadísticamente significativa, se concluye que los valores de la variable son estadísticamente diferentes en las dos muestras.

Las hipótesis de la prueba de los signos se pueden enunciar de la manera siguiente:

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

Si la muestra es pequeña: menos de 25 casos, se resuelve el contraste anterior mediante la distribución binomial y, si la muestra es mayor que 25, se realiza una aproximación a la normal.

[1] Véase Capítulo 25: características estadísticas de los ensayos clínicos.