

T. 3 – Inferencia estadística acerca de la relación entre variables (I)

1. El caso de dos variables categóricas

2. El caso de una variable categórica y una variable cuantitativa

3. El caso de dos variables cuantitativas

- Tras haberse tratado en el capítulo anterior el contraste de hipótesis de parámetros univariados –en concreto, la media y la proporción–, se aborda ahora una serie de pruebas de contraste de hipótesis acerca de la relación entre dos variables. Todas ellas tienen en común un aspecto aplicado: su amplia utilización en la práctica del análisis estadístico.

1. El caso de dos variables categóricas

- Se trata en este caso de valorar la asociación entre dos variables categóricas, considerando si existe una relación entre ambas variables a nivel poblacional. La información a partir de la que valorar dicha asociación viene expresada habitualmente en forma de tabla de contingencia, donde las casillas de la misma reflejan la distribución conjunta de ambas variables categóricas, tal como se vio en el tema “Organización y representación gráfica de datos multivariados”.

• A continuación vamos a centrarnos en la prueba de significación estadística del índice *ji-cuadrado* de Pearson (χ^2), la más utilizada en la práctica para llevar a cabo el contraste de hipótesis acerca de la relación entre dos variables categóricas. El índice *ji-cuadrado* de Pearson ya fue presentado en un tema previo; ahora vamos a ver el procedimiento para aplicar la prueba de significación correspondiente a este índice estadístico, la conocida como prueba *ji-cuadrado* de Pearson, la cual nos va a permitir contrastar la hipótesis de relación nula entre las variables. Cabe señalar que esta prueba se aplicará de la misma manera independientemente de que la relación de las variables sea simétrica o asimétrica –recordamos que, en este segundo caso, se asume que una de las variables adopta el rol de variable explicativa y la otra de variable de respuesta.

• Pasos en la prueba de significación *ji-cuadrado* de Pearson:

1. Se decide el nivel de riesgo (α) que se desea asumir en el contraste de hipótesis y se plantean las hipótesis estadística y nula. Así, siendo *A* y *B* dos variables categóricas con *I* y *J* categorías, respectivamente, la hipótesis estadística plantea que existe relación a nivel poblacional entre ambas variables, mientras que la hipótesis nula conjetura lo contrario, esto es, que las dos variables son independientes:

$$H_e : \chi_{AB}^2 \neq 0$$

$$H_o : \chi_{AB}^2 = 0$$

2. Se calcula el estadístico de contraste *ji-cuadrado* de Pearson (χ^2), el cual se basa en la comparación, para cada casilla de la tabla, de las frecuencias observadas (n_{ij}) con las frecuencias esperadas suponiendo cierta la hipótesis nula (m_{ij}):

$$\chi^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{(n_{ij} - m_{ij})^2}{m_{ij}}$$

La obtención de las frecuencias esperadas es el resultado de la aplicación del teorema de la probabilidad según el cual dos sucesos son independientes si su probabilidad conjunta es igual al producto de sus probabilidades simples. Así, para cada casilla de la tabla:

$$m_{ij} = \frac{n_{i+} \cdot n_{+j}}{n}$$

donde n_{i+} y n_{+j} son las respectivas frecuencias marginales y n es el nº total de casos.

3. Se obtiene, en la distribución *ji-cuadrado* con grados de libertad igual a $(I-1) \times (J-1)$, el nivel de significación (*Sig*) asociado al estadístico χ^2 obtenido, es decir, la probabilidad de obtener un valor del estadístico de contraste como el obtenido o superior.

Podemos obtener este valor de probabilidad fácilmente con el programa *Excel* si introducimos, en una casilla cualquiera de la hoja de cálculo, la siguiente fórmula con los valores entre paréntesis que nos interese:

=DISTR.CHI(χ^2 ;grados_de_libertad)

4. Decisión: se mantiene la H_0 si $Sig > \alpha$; por contra, se rechaza si $Sig < \alpha$.

Ejemplo: Se recogieron datos en una muestra de 500 personas mayores de 70 años sobre las variables “Estado de ánimo” y “Vivir en una residencia” (ejemplo recogido en el capítulo “Organización y representación gráfica de datos multivariados”). La variable “Estado de ánimo” se midió utilizando una escala que reflejaba 3 categorías ordenadas de estado de ánimo: malo, regular y bueno. En este caso se podría considerar que las variables tienen un rol asimétrico, siendo la variable “Vivir en una residencia” la variable explicativa, pero ello no entraña ningún cambio en la realización de la prueba de significación que nos ocupa.

Supóngase que los resultados obtenidos fueron los presentados en esta tabla de contingencia:

	Sí	No	Total
–	48	70	118
±	42	105	147
+	60	175	235
Total	150	350	500

Prueba de significación estadística:

1. Nivel de riesgo (α) = 0,05

$$H_e : \chi_{AB}^2 \neq 0 \rightarrow H_o : \chi_{AB}^2 = 0$$

2. El estadístico de contraste χ^2 requiere obtener, en primer lugar, las frecuencias esperadas en cada casilla de la tabla de contingencia suponiendo cierta la hipótesis nula ($\chi_{AB}^2 = 0$), esto es, que no existiese relación entre ambas variables.

$$m_{11} = \frac{150 \cdot 118}{500} = 35,4$$

$$m_{21} = \frac{150 \cdot 147}{500} = 44,1$$

$$m_{31} = \frac{150 \cdot 235}{500} = 70,5$$

$$m_{12} = \frac{350 \cdot 118}{500} = 82,6$$

$$m_{22} = \frac{350 \cdot 147}{500} = 102,9$$

$$m_{32} = \frac{350 \cdot 235}{500} = 164,5$$

En la siguiente tabla obtenida con SPSS se pueden observar las frecuencias observadas y esperadas en cada casilla de la tabla:

Tabla de contingencia Estado ánimo * Vivir residencia

			Vivir residencia		Total
			Sí	No	
Estado ánimo	Negativo	Recuento	48	70	118
		Frecuencia esperada	35,4	82,6	118,0
	Neutro	Recuento	42	105	147
		Frecuencia esperada	44,1	102,9	147,0
	Positivo	Recuento	60	175	235
		Frecuencia esperada	70,5	164,5	235,0
Total		Recuento	150	350	500
		Frecuencia esperada	150,0	350,0	500,0

Así, el estadístico de contraste será igual a:

$$\chi_{AB}^2 = \frac{(48 - 35,4)^2}{35,4} + \frac{(42 - 44,1)^2}{44,1} + \dots + \frac{(175 - 164,5)^2}{164,5} = 8,784$$

3. La distribución muestral del estadístico χ^2 es la distribución ji-cuadrado con 2 grados de libertad (1×2). Obteniendo con Excel el correspondiente valor [=DISTR.CHI(8,784;2)]:

$$Sig = P(\chi^2 \geq 8,784) = 0,012$$

4. Decisión: $0,012 < 0,05$, por tanto, se rechaza la H_0 de independencia entre ambas variables.

- La aplicación de la prueba de significación anterior es problemática en el caso de frecuencias esperadas bajas, más concretamente, se suele utilizar como criterio que no más del 20% de las frecuencias esperadas de la tabla de contingencia sea menor de 5 ($m_{ij} < 5$). Cuando no se cumple este criterio es cuestionable que el estadístico χ^2 se distribuya según la distribución ji-cuadrado y, por tanto, ya no tiene sentido buscar el nivel de significación (*Sig*) en dicha tabla. En la práctica, cuando no se satisfaga el criterio citado, una solución habitual consiste en agrupar categorías de alguna de las variables, siempre y cuando tenga sentido esa agrupación, hasta conseguir que estas frecuencias esperadas bajas (< 5) no superen el 20% de las casillas.
- El estadístico χ^2 tiene el problema de estar influido, además de por el grado de asociación entre las variables, por el tamaño de la muestra (n), de manera que a mayor n , aún siendo la relación entre dos variables la misma, se obtienen valores superiores del estadístico χ^2 y, por tanto, el valor de *Sig* disminuye, haciéndose más probable el rechazo de la hipótesis nula. Para compensar este problema en la aplicación de la prueba *ji-cuadrado* de Pearson, es conveniente complementar esta prueba con otros índices orientados a medir la intensidad de la asociación entre dos variables categóricas, pero que no dependan de n . Estos índices aparecen también denominados en la literatura estadística como índices de tamaño del efecto y, para el caso de la asociación de dos variables categóricas, dos de los más utilizados son el coeficiente de contingencia *phi de Pearson* y el coeficiente *V de Cramer* (ver capítulo “Estadísticos de asociación entre variables”).

• En el caso de obtenerse una relación estadísticamente significativa, será normalmente también necesario interpretar en qué consiste esa relación a fin de dotar de significado al informe de nuestros resultados. Para tal fin pueden resultar de gran ayuda la obtención de los porcentajes (o proporciones) condicionados, ya sea por fila o por columna, tal como ya se describió en el capítulo de “Estadísticos de asociación entre variables”.

• Los resultados obtenidos con SPSS a través del comando *Tablas de contingencia* del menú *Análisis* permite obtener la tabla de contingencia con las frecuencias observadas, las frecuencias esperadas y los porcentajes condicionados para cada casilla, así como los resultados de la *prueba ji-cuadrado* y las medidas complementarias de intensidad de la asociación.

SPSS: Analizar > Estadísticos descriptivos > Tablas de contingencia

Tabla de contingencia Estado ánimo * Vivir residencia

			Vivir residencia		Total
			Sí	No	
Estado ánimo	Negativo	Recuento	48	70	118
		Frecuencia esperada	35,4	82,6	118,0
		% dentro de Vivir residencia	32,0%	20,0%	23,6%
	Neutro	Recuento	42	105	147
		Frecuencia esperada	44,1	102,9	147,0
		% dentro de Vivir residencia	28,0%	30,0%	29,4%
	Positivo	Recuento	60	175	235
		Frecuencia esperada	70,5	164,5	235,0
		% dentro de Vivir residencia	40,0%	50,0%	47,0%
Total	Recuento	150	350	500	
	Frecuencia esperada	150,0	350,0	500,0	
	% dentro de Vivir residencia	100,0%	100,0%	100,0%	

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	8,784 ^a	2	,012
Razón de verosimilitudes	8,507	2	,014
Asociación lineal por lineal	7,788	1	,005
N de casos válidos	500		

a. 0 casillas (.0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 35,40.

Medidas simétricas

		Valor	Sig. aproximada
Nominal por nominal	Phi	,133	,012
	V de Cramer	,133	,012
	Coefficiente de contingencia	,131	,012
N de casos válidos		500	

Ejercicio 1: A partir de una encuesta realizada a un grupo de 173 estudiantes universitarios se quiere evaluar si existe relación entre las variables ‘Sexo’ y ‘Compaginar estudios y trabajo’ (escala de respuesta: no trabajo; trabajo a tiempo completo; trabajo a tiempo parcial). La siguiente tabla de contingencia muestra las frecuencias observadas (recuentos) y las frecuencias esperadas para ambas variables, así como los porcentajes de “Compaginar...” condicionados, para hombres y mujeres.

Tabla de contingencia Compaginar estudios con trabajo * sexo

			sexo		Total
			Hombre	Mujer	
Compaginar estudios con trabajo	no	Recuento	20	96	116
		Frecuencia esperada	20,8	95,2	116,0
		% dentro de sexo	64,5%	67,6%	67,1%
	sí, tiempo completo	Recuento	5	10	15
		Frecuencia esperada	2,7	12,3	15,0
		% dentro de sexo	16,1%	7,0%	8,7%
	sí, tiempo parcial	Recuento	6	36	42
		Frecuencia esperada	7,5	34,5	42,0
		% dentro de sexo	19,4%	25,4%	24,3%
Total	Recuento	31	142	173	
	Frecuencia esperada	31,0	142,0	173,0	
	% dentro de sexo	100,0%	100,0%	100,0%	

¿Qué se puede decir respecto a la relación entre ambas variables? Para contestar, realiza la prueba de significación estadística correspondiente ($\alpha = 0,05$) y obtén también algún índice que informe del tamaño del efecto de esa asociación (*phi* o *V de Cramer*). Redacta un párrafo resumiendo los resultados obtenidos. Tras realizar los cálculos de forma manual, intenta obtener esos mismos índices haciendo uso del paquete estadístico *SPSS*.

Ejercicio 2: Ídem. para las variables ‘Con quien convives’ y ‘Compaginar estudios con trabajo’. Aprovecha los siguientes resultados obtenidos con *SPSS* para contestar.

Tabla de contingencia Compaginar estudios con trabajo * ¿Con quien convives?

			¿Con quien convives?				Total
			solo	padres, familiares	estudiantes, amigos	pareja y/o hijos	
Compaginar estudios con trabajo	no	Recuento	1	75	39	0	115
		Frecuencia esperada	3,3	72,9	34,8	4,0	115,0
	sí, tiempo completo	Recuento	3	5	2	5	15
		Frecuencia esperada	,4	9,5	4,5	,5	15,0
	sí, tiempo parcial	Recuento	1	29	11	1	42
		Frecuencia esperada	1,2	26,6	12,7	1,5	42,0
	Total	Recuento	5	109	52	6	172
		Frecuencia esperada	5,0	109,0	52,0	6,0	172,0

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	63,789 ^a	6	,000
Razón de verosimilitudes	35,950	6	,000
Asociación lineal por lineal	,021	1	,886
N de casos válidos	172		

a. 7 casillas (58,3%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es ,44.

2. El caso de una variable categórica y una variable cuantitativa

• Se aborda ahora el contraste de hipótesis acerca de la relación entre una variable categórica y una cuantitativa, si bien, nos vamos a centrar ahora en el caso en que la variable categórica sea dicotómica. Este tipo de contraste, siguiendo la estrategia planteada en el análisis de la relación entre una variable categórica dicotómica y una cuantitativa (ver capítulo “Estadísticos de asociación entre variables”), se basa en comprobar la existencia de diferencias entre las medias en la variable cuantitativa en los dos subgrupos definidos por la variable categórica, eso sí, se trata ahora de extraer conclusiones a nivel poblacional. O, dicho de otra manera, se trata ahora de hipótesis acerca de la igualdad entre dos parámetros (las medias en ambas poblaciones). Al igual que en el caso de las hipótesis en que aparece implicado un único parámetro, estas hipótesis o contrastes pueden ser unilaterales, si se especifica cuál de las medias será superior a la otra, o bilaterales, si únicamente se especifica que dichas medias serán diferentes.

Ejemplo: La hipótesis de que, en la población de personas de la 3ª edad, la media de satisfacción vital es superior entre los que viven con sus familiares que entre los que viven en una residencia se formularía como:

$$H_e: \mu (\text{familiares}) > \mu (\text{residencia}); \quad H_o: \mu (\text{familiares}) \leq \mu (\text{residencia})$$

Una formulación alternativa a partir de la diferencia entre ambas medias poblacionales (δ) sería:

$$H_e: \delta > 0; \quad H_o: \delta \leq 0 \quad (\text{donde } \delta = \mu (\text{familiares}) - \mu (\text{residencia}))$$

Ejercicio 3: Formula estadísticamente las siguientes hipótesis (H_e y H_o) de las dos formas anteriores:

- La satisfacción laboral es superior entre los trabajadores del sector de la industria que en el sector servicios.
- La relación con los compañeros es valorada diferencialmente por hombres y mujeres.
- Los accidentes laborales se producen más entre los trabajadores con contrato parcial que fijo.
- Tras el entrenamiento recibido los sujetos muestran una mayor autoestima.

2.1. El caso de dos muestras independientes

- Se trata del contraste de hipótesis acerca de la diferencia entre dos medias obtenidas para una misma variable cuantitativa en dos muestras que representan a dos poblaciones distintas, siendo precisamente la variable categórica la que va a determinar la pertenencia a una u a otra población. De ahí que a este contraste se haga referencia como contraste de hipótesis de dos muestras independientes, pues cada media ha sido obtenida a partir de una muestra de casos totalmente independiente de la otra muestra.
- El objetivo de este contraste de hipótesis es decidir si las medias empíricas obtenidas en las dos muestras definidas por la variable categórica proceden, o no, de poblaciones con idéntica media. En el caso de rechazar la hipótesis nula (hipótesis de igualdad de las medias), ello significaría que la evidencia empírica apoya que existe relación entre ambas variables a nivel poblacional. Veamos cómo abordar este contraste de hipótesis a partir de la realización de una prueba de significación estadística y, posteriormente, también a partir de la obtención del intervalo de confianza (*IC*) para la diferencia entre ambas medias.

2.1.1. Contraste de hipótesis basado en prueba de significación estadística: la prueba *T* de Student para muestras independientes.

- Pasos en la prueba *T* de Student para muestras independientes:

1. Se decide el nivel de riesgo (α) que se desea asumir en el contraste de hipótesis y se plantean las hipótesis estadística y nula.

Si *A* y *B* representan a las dos sub-poblaciones definidas por la variable categórica, una posible hipótesis estadística sería, por ejemplo:

$$H_e : \mu_B > \mu_A \quad \text{y, complementariamente,} \quad H_o : \mu_B \leq \mu_A$$

Otra forma de plantear esas mismas hipótesis que podemos encontrar en la literatura estadística es utilizando un único parámetro que representa la diferencia entre las dos medias. Para nuestro ejemplo anterior:

$$H_e : \delta > 0 \quad (\text{donde } \delta = \mu_B - \mu_A) \quad \text{y, complementariamente,} \quad H_o : \delta \leq 0 \quad (\text{donde } \delta = \mu_B - \mu_A)$$

- En el caso de tratarse de un contraste bilateral, la H_o y la H_e se expresarán como:

$$H_e : \mu_B \neq \mu_A \quad (\text{o } H_e : \delta \neq 0) \quad \text{y, complementariamente,} \quad H_o : \mu_B = \mu_A \quad (\text{o } H_o : \delta = 0)$$

2. Explorar si las medias empíricas obtenidas parecen apoyar, en principio, la hipótesis estadística planteada. En caso contrario, no tiene sentido continuar con los siguientes pasos del contraste de

hipótesis y se mantendría la H_0 . En cualquier caso, el resultado muestral puede ser aprovechado a la hora de plantear hipótesis estadísticas más afinadas en el futuro.

3. Se calcula el siguiente estadístico de contraste:

$$T = \frac{|d|}{EE(d)}$$

donde $|d|$ es el valor de la diferencia entre las dos medias muestrales en valor absoluto y $EE(d)$ es el error estándar de la de la distribución muestral del estadístico de la diferencia entre esas dos medias bajo el supuesto de la hipótesis nula. Este último se calcula como:

$$EE(d) = \sqrt{\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

4. Se obtiene en la distribución t de Student con n_1+n_2-2 grados de libertad, la probabilidad de obtener un valor como el obtenido con el estadístico de contraste o más extremo, esto es, el nivel de significación (Sig).

Se puede obtener este valor de probabilidad con el programa *Excel* si introducimos, en una casilla cualquiera de la hoja de cálculo, la siguiente fórmula con los valores entre paréntesis que nos interese:

$$=DISTR.T(T;grados_de_libertad;1)$$

En el caso de tratarse de un contraste bilateral, basta con multiplicar el valor de Sig obtenido por 2, o bien, cambiar en la anterior fórmula de Excel el último 1 entre paréntesis por un 2.

(Nota: si ambas muestras son de tamaño superior a 30, se puede utilizar la distribución normal en vez de la distribución t para buscar el nivel de significación del estadístico de contraste).

5. Decisión: se mantiene la H_0 si $Sig > \alpha$; por contra, se rechaza si $Sig < \alpha$.

Ejemplo: Loftus y Burns (1982) realizaron un experimento para comprobar en qué medida un choque emocional puede alterar el recuerdo. Para ello proyectaron a dos grupos de sujetos una misma película con dos versiones de la misma: en una de ellas aparecía una escena de gran violencia orientada a producir un choque emocional en los sujetos, pero esta escena no aparecía en la otra película. Posteriormente se hacía una prueba de memoria a fin de medir el recuerdo de los sujetos de ambos grupos. Supóngase que los resultados obtenidos con SPSS fueron:

Informe

Recuerdo			
Choque_Emocional	Media	Desv. típ.	N
No	15,00	2,121	5
Sí	11,00	1,581	5
Total	13,00	2,749	10

En este ejemplo, la variable cuantitativa es la puntuación obtenida en la prueba de memoria, mientras que la variable categórica es el haber visto una película u otra, esto es, haber recibido el choque emocional o no. A continuación se muestran los pasos en la prueba de significación orientada a contrastar si existen diferencias estadísticamente significativas entre las medias de los dos grupos a nivel poblacional:

1. Nivel de riesgo (α) = 0,05

$$H_e: \mu_{NO} \neq \mu_{SI} \quad (\text{o } H_e: \delta \neq 0) \quad \leftarrow \text{contraste de hipótesis bilateral}$$

$$H_o: \mu_{NO} = \mu_{SI} \quad (\text{o } H_o: \delta = 0)$$

2. Las medias muestrales (15 y 11) parecen apoyar, en un principio, la hipótesis de que el recuerdo difiere en ambos grupos.
3. Cálculo del estadístico de contraste T :

$$EE(d) = \sqrt{\frac{2,12^2(5-1) + 1,58^2(5-1)}{5+5-2} \left(\frac{1}{5} + \frac{1}{5} \right)} = 1,183$$

$$T = \frac{|d|}{EE(d)} = \frac{4}{1,183} = 3,38$$

(Señalar que, al igual que en el caso del contraste de hipótesis sobre una media, este estadístico de contraste es una estandarización de la diferencia de las medias obtenidas en la muestra ya que, si es cierta la H_o , la distribución muestral de la diferencia de medias obtenidas en muestras de $n = 5$ extraídas al azar de una población en que los dos subgrupos tienen la misma media seguirá una distribución t con parámetros $E(d) = 0$ y $EE(d) = 1,183$)

4. La distribución muestral del estadístico T para este ejemplo será la distribución t de Student con 8 grados de libertad ($5+5-2$). Obteniendo con *Excel* el correspondiente valor de *Sig* para un contraste bilateral [=DISTR.T(3,38;8;2)], se obtiene que $Sig = 0,0096$
5. Decisión: $0,0096 < 0,05$, por tanto, se rechaza la H_o y se puede concluir, con un nivel de riesgo del 5%, que ambas muestras proceden de poblaciones con diferente media y, en consecuencia, que hay una relación estadísticamente significativa entre ambas variables. Ello implica que el choque emocional sí que influye en el recuerdo, más concretamente, cuando éste se da, el recuerdo promedio se reduce de forma estadísticamente significativa respecto a cuando no existe choque emocional.

- La correcta aplicación de la prueba de significación estadística anterior requiere del cumplimiento de una serie de supuestos, entre ellos, uno de gran importancia es el conocido como supuesto de homogeneidad de las varianzas, esto es, las varianzas de la variable cuantitativa en ambas

poblacionales debe ser igual. El cumplimiento de este supuesto es evaluado habitualmente mediante la *prueba de Levene*, prueba de significación en la que contrasta la hipótesis estadística:

$$H_e: \sigma_A \neq \sigma_B \quad \text{y, complementariamente, } H_o: \sigma_A = \sigma_B$$

En caso de que no se satisfaga este supuesto (rechazo de la H_o), es necesario realizar un cálculo diferente del $EE(d)$, así como de los grados de libertad de su distribución muestral (véase, por ejemplo, Pardo y San Martín, 1998).

- La salida de SPSS al ejecutar la prueba T de Student para la comparación de medias en muestras independientes incluye el resultado de la prueba de Levene. Véase en el análisis realizado para el ejemplo de las variables ‘Choque emocional’ y ‘Recuerdo’:

SPSS: Analizar > Comparar medias > Prueba T para muestras independientes

Prueba de muestras independientes

		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						
		F	Sig.	t	gl	Sig. (bilateral)	Diferencia de medias	Error típ. de la diferencia	95% Intervalo de confianza para la diferencia	
									Inferior	Superior
Recuerdo	Se han asumido varianzas iguales	,400	,545	-3,381	8	0,0096	-4,000	1,183	-6,729	-1,271
	No se han asumido varianzas iguales			-3,381	7,396	0,0108	-4,000	1,183	-6,768	-1,232

En este **ejemplo**, por lo que respecta al supuesto de homogeneidad de las varianzas, dado el nivel de significación obtenido para la prueba de Levene ($Sig = 0,545$) y asumiendo un nivel de riesgo $\alpha = 0,05$, se mantiene la hipótesis nula de igualdad de varianzas en la población, por lo que podemos considerar como satisfecho el cumplimiento de este supuesto. En caso contrario (si el nivel de significación de la prueba de Levene fuese inferior a 0,05), nos deberemos fijar en los valores de la fila inferior de la tabla de resultados, donde se muestran los resultados para la prueba T corregida para el incumplimiento del supuesto de homogeneidad de las varianzas.

Comentar también, en relación al ejemplo anterior y a la salida de *SPSS* para el mismo, que el valor de T obtenido por *SPSS* tiene signo negativo, mientras que el que hemos obtenido nosotros es el mismo pero de signo positivo. Ello es debido a que en la fórmula que hemos planteado para calcular T , éste se obtiene como el valor absoluto de la diferencia de las medias, por lo que siempre dará lugar a un valor de T positivo. En cualquier caso, el valor que obtendremos para el nivel de significación será el mismo y, por lo tanto, se llegará a una misma decisión en el contraste de hipótesis.

• La prueba T nos informará acerca de la significación estadística de la diferencia entre dos medias y , en consecuencia, acerca de si la relación entre una variable dicotómica y una variable cuantitativa es estadísticamente significativa; sin embargo, no resulta obvio interpretar a partir del valor de T cuál es la magnitud de la relación. Como indicador de la intensidad de la relación (tamaño del efecto), resulta más apropiado el índice d de Cohen, el cual fue presentado en el capítulo de *Estadísticos de asociación entre variables*. Si disponemos del valor de T , podemos obtener fácilmente el valor de d a partir de la siguiente expresión:

$$d = T \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

(Comprobar a partir del ejercicio siguiente, donde la T da 2,27)

• Otro indicador del tamaño del efecto de la relación entre una variable categórica dicotómica y una variable cuantitativa es el conocido como coeficiente de correlación biserial-puntual (r_{bp}), el cual tiene la ventaja de oscilar entre 0 y 1. Puede ser calculado con el programa *SPSS* obteniendo el coeficiente de correlación de Pearson entre ambas variables (elimínese el signo si da negativo, pues carece de sentido en este caso), o bien, si se dispone del valor de T , a través de la siguiente fórmula:

$$r_{bp} = \frac{T}{\sqrt{T^2 + n - 2}}$$

2.1.2. Contraste de hipótesis basado en intervalo de confianza

• Pasos en la elaboración de un *IC* de la diferencia de dos medias independientes:

1. El *IC* para este tipo de contraste se plantea para el valor de la diferencia entre las dos medias poblacionales (δ) como:

$$IC(1-\alpha)(\delta) = \left[|d| + t_{(n_1+n_2-2)}(\alpha/2) \cdot EE(d) ; |d| + t_{(n_1+n_2-2)}(1-\alpha/2) \cdot EE(d) \right]$$

– El $EE(d)$ se obtiene tal y como se expuso más arriba al describir la prueba de significación.

– Si el tamaño de la muestra es superior a 30, se pueden utilizar para crear el *IC* los valores de $z(\alpha/2)$ y $z(1-\alpha/2)$, en vez de los de $t_{(n_1+n_2-2)}(\alpha/2)$ y $t_{(n_1+n_2-2)}(1-\alpha/2)$, respectivamente.

2. Se decide el rechazo de la H_0 cuando el *IC* de la diferencia entre las medias no se corresponde con la diferencia expresada en la H_0 ; en caso contrario, se mantiene la H_0 . Un aspecto positivo de la creación del *IC* es que, además de permitirnos llevar a cabo el contraste de hipótesis, resulta informativo acerca de la magnitud de la diferencia de las medias en la población.

Ejemplo para el estudio orientado a analizar la relación entre choque emocional y recuerdo: el *IC* de la diferencia de medias para un nivel de confianza del 95% se obtendrá, teniendo en cuenta que $t_{8; 0,025} = -2,31$ y $t_{8; 0,975} = 2,31$, como:

$$IC(95\%)(\delta) = [4 - 2,31 \cdot 1,18 ; 4 + 2,31 \cdot 1,18] = [1,27 ; 6,73]$$

- Véase también este *IC* en los resultados obtenidos con *SPSS* que fueron presentados en el apartado previo.
- Dado que el intervalo anterior no incluye el valor 0, se rechazaría la H_0 . Ahora bien, un *IC* nos ofrece información añadida al contraste de hipótesis propiamente dicho, pues plantea una horquilla de valores, entre 1,27 y 6,73 en este caso, que, con un nivel de confianza del 95%, contendrá la diferencia entre las dos medias de las poblaciones. Por otra parte, puede resultar igualmente interesante saber qué valores es muy poco probable que tome la diferencia entre esas dos medias (en nuestro ejemplo, es muy poco probable que la diferencia entre las medias sea inferior a 1 o superior a 7).
- Por último, no hay que olvidar que cuando se encuentre una relación estadísticamente significativa, como ha sido el caso de nuestro ejemplo, debemos interpretar el sentido de esa relación a fin de completar el informe de nuestro análisis. Así, podemos decir que la relación encontrada, aparte de estadísticamente significativa, consiste en la existencia de un recuerdo superior cuando no hay choque emocional que en la presencia del mismo.

Ejercicio 4: Se ha llevado a cabo un experimento para determinar si los dibujos actúan como facilitadores o entorpecedores del aprendizaje de palabras en niños de 3 años. Se han seleccionado aleatoriamente a 72 niños de una escuela infantil y a la mitad de ellos se ha planteado enseñarles nuevas palabras sin utilizar ilustraciones (grupo A) y a la otra mitad lo mismo pero con dibujos sencillos que representan esas nuevas palabras (grupo B). Tras el entrenamiento se ha evaluado el número de palabras aprendidas por cada niño y se han obtenido los siguientes resultados en cada grupo: $\bar{X}_A = 24$ $s_A^2 = 16$ $\bar{X}_B = 30$ $s_B^2 = 21$. ¿Qué se puede decir respecto a la cuestión de partida? Realiza la prueba de significación estadística correspondiente, así como el intervalo de confianza de la diferencia de medias ($\alpha = 0,05$).

- Una vez concluido que la diferencia entre medias es estadísticamente significativa nos hemos de plantear la importancia o relevancia práctica de esa diferencia. El *IC* de la diferencia entre las medias nos permite valorar esta relevancia. Para ello debemos contar con un criterio de relevancia práctica (δ^*) y compararlo con los límites del *IC* obtenido.

Ejemplo para el estudio orientado a analizar la relación entre choque emocional y recuerdo: consideremos 3 criterios diferentes para valorar la relevancia de la diferencia entre medias. Recordemos que el *IC* de la diferencia entre las medias obtenido para este ejemplo fue:

$$IC(95\%)(\delta) = [4 - 2,31 \cdot 1,18 ; 4 + 2,31 \cdot 1,18] = [1,27 ; 6,73]$$

Si consideramos que la diferencia entre las medias, para ser considerada relevante en la práctica, ha de ser de al menos:

- 1 unidad ($\delta^*=1$): la diferencia entre medias obtenida se consideraría importante o relevante
- 7 unidades ($\delta^*=7$): la diferencia entre medias obtenida se consideraría no importante o relevante
- 3 unidades ($\delta^*=3$): el resultado obtenido no sería concluyente al respecto

- En conclusión, el *IC* de la diferencia de medias es una alternativa a la prueba de significación a la hora de decidir sobre el rechazo o no de la *H₀*. Además, es más informativo que la prueba de significación, ya que permite conocer entre qué valores se encontrará en la población la diferencia entre las medias, a la vez que nos permite valorar la relevancia práctica de esa diferencia. Por ello, en los informes de investigación o informes técnicos debería incluirse siempre, además de la prueba de significación, el *IC* de la diferencia entre las medias.

- En el informe técnico o de investigación deberían redactarse los resultados de la prueba de significación de alguna de estas formas equivalentes:

- ...hay un efecto estadísticamente significativo del choque emocional sobre el recuerdo ($t_8 = -3,38$; $p=0,0096$).
- ...hay diferencias estadísticamente significativas en el recuerdo en las condiciones de presencia y ausencia de choque emocional ($t_8 = -3,38$; $p=0,0096$).
- ...la presencia de un choque emocional disminuye significativamente el recuerdo de los sujetos ($t_8 = -3,38$; $p=0,0096$).

La redacción anterior se debe completar añadiendo el *IC* de la diferencia de medias:

... con una confianza del 95% la diferencia entre las medias se sitúa entre 1,27 y 6,73 unidades.

... con una confianza del 95%, la presencia de choque emocional disminuye el recuerdo entre 1,27 y 6,73 unidades.

2.2. El caso de dos muestras relacionadas

- Se trata del contraste de hipótesis acerca de la relación entre una variable categórica dicotómica y una variable cuantitativa en que esta última es medida en alguna de las siguientes circunstancias: (1) en un mismo grupo de sujetos antes y después de la aplicación de una determinada acción (intervención, tratamiento...) que viene representada por la variable dicotómica (antes: no acción; después: sí acción); (2) en dos grupos de sujetos relacionados entre sí, esto es, cada sujeto en uno de los grupos tiene un par en el otro grupo con el que tiene algún tipo de equivalencia en terceras variables –un caso paradigmático de diseño de investigación en que se da esta circunstancia es aquél en que los dos grupos están constituidos por pares de gemelos.
- A este tipo de contraste se le suele denominar como contraste de hipótesis de dos muestras relacionadas y se hace operativo a través de la comparación de las dos medias obtenidas en la variable cuantitativa, bien para una misma muestra de sujetos medida antes y después de una determinada intervención, bien para dos grupos de sujetos relacionados entre sí. Veamos cómo llevar a cabo este contraste de hipótesis a partir de la realización de una prueba de significación estadística y, posteriormente, a partir del intervalo de confianza para la diferencia entre ambas medias.

2.2.1. Contraste de hipótesis basado en prueba de significación estadística: la prueba *T* de Student para muestras relacionadas.

- Pasos en la prueba *T* de Student para muestras relacionadas:

1. Se decide el nivel de riesgo (α) que se desea asumir en el contraste de hipótesis y se plantean las hipótesis estadística y nula.

- Si en dos muestras relacionadas son obtenidas sus medias en una misma variable cuantitativa o, lo que es más habitual en la práctica, en una misma muestra son obtenidas esas medias en dos momentos temporales *A* y *B* distintos, una posible hipótesis estadística sería:

$$H_e : \mu_B > \mu_A \text{ y, complementariamente, } H_o : \mu_B \leq \mu_A$$

- Análogamente a como se hizo para la prueba de significación para dos muestras independientes, las hipótesis anteriores se pueden expresar con un único parámetro (δ^*) que representa la diferencia, a nivel poblacional, entre esas dos medias relacionadas:

$$H_e : \delta^* > 0 \text{ (donde } \delta^* = \mu_B - \mu_A \text{) y, complementariamente, } H_o : \delta^* \leq 0 \text{ (donde } \delta^* = \mu_B - \mu_A \text{)}$$

- En el caso de plantearse una hipótesis estadística de desigualdad (contraste bilateral), sería:

$$H_e : \mu_B \neq \mu_A \text{ (o } H_e : \delta^* \neq 0 \text{) y, complementariamente, } H_o : \mu_B = \mu_A \text{ (o } H_o : \delta^* = 0 \text{)}$$

2. Explorar si la diferencia de las medias obtenidas para A y para B en la muestra (d^*) parece apoyar, en principio, la hipótesis estadística planteada. Si esa diferencia es 0 (o muy próxima a 0) o de signo contrario al esperado, no tendrá sentido continuar con el contraste de hipótesis y se mantendría la H_0 .
3. Se calcula el siguiente estadístico de contraste:

$$T = \frac{|d^*|}{EE(d^*)}$$

donde $|d^*|$ es el valor de la diferencia entre las dos medias muestrales en valor absoluto y $EE(d^*)$ es el error típico de la distribución muestral de la diferencia de esas dos medias relacionadas. Este último se calcula como:

$$EE(d^*) = \frac{s'_{DIF}}{\sqrt{n}} \quad \text{siendo} \quad s'_{DIF} = \sqrt{\frac{\sum (DIF_i - \overline{DIF})^2}{n-1}}$$

La última fórmula representa la cuasi-desviación estándar de DIF , la variable resultante de obtener, para cada sujeto, la diferencia entre su valor en una variable y su valor en la otra variable (ver ejemplo a continuación).

4. Se obtiene en la distribución t de Student con $n-1$ grados de libertad, la probabilidad de obtener un valor como el obtenido con el estadístico de contraste o más extremo, esto es, el nivel de significación (Sig). Esta probabilidad la podremos obtener con el programa *Excel* tal como ya fue descrito para la prueba de significación estadística para dos muestras independientes. (Nota: si el tamaño de la muestra es superior a 30, algo no infrecuente, se puede utilizar la distribución normal en vez de la distribución t a la hora de buscar el nivel de significación).
5. Decisión: se mantiene la H_0 si $Sig > \alpha$; por contra, se rechaza si $Sig < \alpha$.

Ejemplo: Supongamos un estudio en que 6 sujetos son sometidos a un tratamiento contra la ansiedad y que tenemos medidas de un determinado test de ansiedad (rango de puntuaciones posible: 0 a 10) que fue aplicado antes (pre-test) y después (post-test) de que el tratamiento se llevara a cabo. Se desea saber si hay una relación estadísticamente significativa entre el nivel de ansiedad y la aplicación del tratamiento, esto es, si hay diferencias entre las medias de ansiedad antes y después del tratamiento. Las puntuaciones obtenidas en ansiedad y la diferencia entre ellas (DIF) para los 6 participantes en este estudio fueron las siguientes:

<i>ID</i>	<i>Pre-test</i>	<i>Post-test</i>	<i>DIF</i>
S1	9	6	3
S2	10	8	2
S3	8	7	1
S4	7	6	1
S5	8	4	4
S6	6	5	1

siendo $Media(Pre-test) = 8$, $Media(Post-test) = 6$, $Media(DIF) = 2$ y $S'_{DIF} = 1,265$

Pasos en la prueba de significación orientada a contrastar si existen diferencias estadísticamente significativas entre las medias de ansiedad en el pre-test y en el post-test:

1. Nivel de riesgo (α) = 0,05. Contraste de hipótesis unilateral:

$$H_e : \mu_{POST} \neq \mu_{PRE} \quad (\text{o } H_e : \delta^* \neq 0) \quad \leftarrow \text{contraste de hipótesis bilateral}$$

$$H_o : \mu_{POST} = \mu_{PRE} \quad (\text{o } H_o : \delta^* = 0)$$

2. La diferencia de las medias muestrales ($d^* = 8 - 6 = 2$) sustenta, en un principio, la hipótesis de que este valor es diferente de 0 a nivel poblacional.
3. Cálculo del estadístico de contraste T para medias de muestras relacionadas:

$$EE_{DIF} = \frac{1,265}{\sqrt{6}} = 0,516 \quad \rightarrow \quad t = \frac{2}{0,516} = 3,87$$

(Al igual que en el caso del contraste sobre dos muestras independientes, este estadístico de contraste es una estandarización de la media de las diferencias obtenida en la muestra ya que, si es cierta H_o , la distribución muestral de la diferencia de dos medias relacionadas en muestras de $n = 6$ extraídas al azar de una población seguirá una distribución t con parámetros $E(d^*) = 0$ y $EE(d^*) = 0,516$)

4. La distribución muestral del estadístico T para este ejemplo será la distribución t de Student con 5 grados de libertad (*g.l.*: 6-1). Obteniendo con *Excel* el correspondiente valor de *Sig* para un contraste bilateral [=DISTR.T(3,87;5;2)], se obtiene que $Sig = 0,012$.
5. Decisión: $0,012 < 0,05$, por tanto, se rechaza la H_o .

Aparte de diferentes desde un punto de vista estadístico, el hecho de que la media post-test sea inferior a la media pre-test evidencia que el tratamiento contra la ansiedad aplicado es positivo, pues ha supuesto una reducción en la media del nivel de ansiedad.

- Los resultados obtenidos con SPSS al ejecutar la prueba T para muestras relacionadas de nuestro ejemplo se muestran en la siguiente tabla:

SPSS: Analizar > Comparar medias > Prueba T para muestras relacionadas

Prueba de muestras relacionadas

		Diferencias relacionadas				t	gl	Sig. (bilateral)	
		Media	Desviación típ.	Error típ. de la media	95% Intervalo de confianza para la diferencia				
					Inferior				Superior
Par 1	Pretest - Posttest	2,000	1,265	,516	,673	3,327	3,873	5	,012

2.2.2. Contraste de hipótesis basado en intervalo de confianza

- Pasos en la elaboración de un *IC* de la diferencia de dos medias relacionadas:

1. El *IC* para este tipo de contraste se plantea para el valor de la diferencia entre las dos medias relacionadas (δ^*) como:

$$IC(1-\alpha)(\delta^*) = \left[|d^*| + t_{(n-1);(\alpha/2)} \cdot EE(d^*) ; |d^*| + t_{(n-1);1-(\alpha/2)} \cdot EE(d^*) \right]$$

- El $EE(d^*)$ se calcula tal y como se expuso previamente al describir la prueba de significación.
- Si el tamaño de la muestra es superior a 30, se pueden utilizar para crear el *IC* los valores de $z_{1-(\alpha/2)}$ y $z_{1-(\alpha/2)}$, en vez de los de $t_{(n-1);1-(\alpha/2)}$ y $t_{(n-1);1-(\alpha/2)}$, respectivamente.

2. Se decide el rechazo de la H_0 cuando el *IC* de la media de las diferencias no se corresponde con la diferencia expresada en la H_0 ; en caso contrario, se mantiene la H_0 .

Como ya se comentó previamente, una ventaja de los *IC* es que, además de permitirnos llevar a cabo el contraste de hipótesis, proporcionan información sobre la magnitud de la diferencia de las medias, no limitándose a una decisión dicotómica de mantener o rechazar la hipótesis nula como ocurre con la prueba de significación estadística.

Ejemplo: En el caso del ejemplo anterior sobre el estudio de la efectividad del tratamiento de la ansiedad, el *IC* de la diferencia de las medias para un nivel de confianza del 95% se obtendrá, teniendo en cuenta que $t_{5;0,025} = -2,57$ y $t_{5;0,975} = 2,57$:

$$IC(95\%)(\delta^*) = [2 - 2,57 \cdot 0,516 ; 2 + 2,57 \cdot 0,516] = [0,67 ; 3,33]$$

Véase también este *IC* en los resultados de *SPSS* presentados en el apartado anterior.

Dado que el intervalo anterior no incluye el valor 0, se rechazaría la H_0 . Además, este *IC* permite afirmar que, a nivel poblacional (con un nivel de confianza del 95%), la media de ansiedad en el post-test se encuentra entre 0,67 y 3,33 unidades por debajo de la media en el pre-test.

Ejercicio 5: Según sugieren algunos trabajos, los niños con problemas psicomotores severos aumentan su rendimiento en los tests de inteligencia no verbal si se les proporciona un entrenamiento adecuado para realizar las tareas de estos tests que implican habilidades psicomotrices. Con el fin de obtener evidencia adicional sobre esta afirmación, en una investigación se seleccionó una muestra aleatoria de 20 niños con problemas psicomotores severos y se les pasó un test de inteligencia no verbal para obtener una medida inicial en el test. Después se les entrenó durante dos meses en tareas psicomotrices, pero no iguales, a las planteadas por el test y, terminado el entrenamiento, se les volvió a pasar el test de inteligencia para obtener una medida post-tratamiento. La media de las puntuaciones del test antes fue de 78,8, mientras que la media después fue de 81,8 ¿Hay suficiente evidencia empírica para afirmar, con un α del 0,01, que el tipo de entrenamiento efectuado consigue aumentar el rendimiento en el test de inteligencia de los niños con problemas psicomotores severos?

Realiza la prueba de significación correspondiente, así como el intervalo de confianza de la diferencia de medias. Datos de interés para la resolución del ejercicio: $t_{(19)(0,005)} = -2,86$; $t_{(19)(0,995)} = 2,86$; $EE(d^*) = 1,4$

Ejercicio 6: Realiza el mismo ejercicio considerando que la muestra hubiese sido de 200 niños, en cuyo caso $EE(d^*) = 0,34$.

3. El caso de dos variables cuantitativas.

- Se describe en esta sección el contraste de hipótesis relativo al parámetro del coeficiente de correlación de Pearson (ρ_{XY}) –el índice estadístico más utilizado a la hora de evaluar la asociación entre dos variables cuantitativas–, así como el de los parámetros de la ecuación de regresión lineal simple. Como veremos, el contraste del parámetro de la pendiente de la ecuación de regresión se encuentra directamente vinculado al contraste del coeficiente de correlación de Pearson.

3.1. Contraste de hipótesis acerca de la relación entre dos variables cuantitativas

- Este contraste está orientado a comprobar, para una determinada población, la hipótesis estadística $H_e: \rho_{XY} \neq 0$, frente a la hipótesis nula $H_o: \rho_{XY} = 0$ (hipótesis de independencia entre las dos variables), a partir del valor del coeficiente de correlación de Pearson (r_{XY}) obtenido en una muestra de esa población para dos variables cuantitativas X e Y . Veamos cómo abordar este contraste de hipótesis a partir de la realización de una prueba de significación estadística y, alternativamente, a través de la creación del intervalo de confianza correspondiente.

3.1.1. Contraste de hipótesis basado en prueba de significación estadística: la prueba T para el coeficiente de correlación de Pearson

- Pasos en la prueba T para el coeficiente de correlación de Pearson

1. Se decide el nivel de riesgo (α) que se desea asumir en el contraste de hipótesis y se plantean las hipótesis estadística y nula. En el caso de acometer un contraste bilateral, éstas serían:

$$H_e : \rho_{XY} \neq 0 \quad \text{y, complementariamente,} \quad H_o : \rho_{XY} = 0$$

2. Explorar si el coeficiente de correlación de Pearson obtenido en la muestra apoya, en principio, la hipótesis estadística planteada, esto es, que no sea nulo (o prácticamente nulo) o de signo contrario al esperado, pues en ese caso no tiene mucho sentido continuar con los siguientes pasos y se mantendría la H_o .
3. Se calcula el estadístico de contraste correspondiente a esta prueba de significación:

$$T = \frac{|r_{XY}|}{EE(r_{XY})}$$

donde $|r_{XY}|$ es el valor absoluto del coeficiente de correlación de Pearson obtenido a nivel muestral y $EE(r_{XY})$ es el error estándar de la distribución muestral del estadístico del coeficiente de correlación bajo el supuesto de que sea cierto lo expresado en la hipótesis nula ($H_o : \rho_{XY} = 0$), el cual se obtiene de acuerdo a la siguiente fórmula:

$$EE(r_{XY}) = \sqrt{\frac{1 - r_{XY}^2}{n - 2}}$$

4. Se obtiene en la distribución t con $n-2$ grados de libertad el nivel de significación (Sig) correspondiente al valor de T calculado, esto es, la probabilidad de obtener un valor como T o superior. En caso de contraste bilateral, se multiplica por dos el valor Sig obtenido. Tal vez resulte más cómodo obtener ese valor con *Excel* haciendo uso de la función que ofrece este programa para tal fin:

Contraste unilateral \rightarrow =DISTR.T(T ;grados_de_libertad;1)

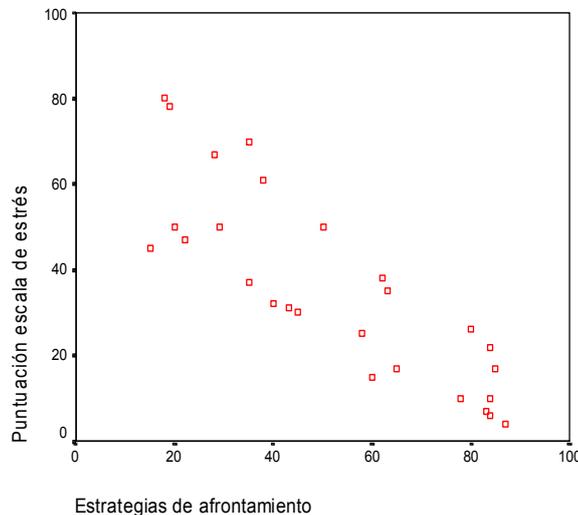
Contraste bilateral \rightarrow =DISTR.T(T ;grados_de_libertad;2)

Como en otras pruebas de significación estadística basadas en la distribución t , si el tamaño de la muestra es superior a 30, se puede utilizar la distribución normal para obtener el valor de Sig .

5. Decisión: se mantiene la H_o si $Sig > \alpha$; por contra, se rechaza si $Sig < \alpha$.

En el caso que se rechace la H_o , se suele expresar este resultado diciendo que la relación entre X e Y es estadísticamente significativa o, equivalentemente, que la relación entre ambas variables a nivel poblacional es, con un nivel de confianza del $(1 - \alpha)\%$, distinta de cero.

Ejemplo: A continuación se presenta el diagrama de dispersión obtenido en un estudio realizado con una muestra de 27 personas adultas, cuyo objetivo fue estudiar si existe relación entre el nivel de estrategias de afrontamiento (X) y el nivel de estrés (Y). El valor del coeficiente de correlación de Pearson obtenido entre ambas variables fue igual a $-0,847$



Vamos a realizar a continuación la prueba de significación estadística que permita comprobar si existe una relación estadísticamente significativa entre ambas variables:

1. Nivel de riesgo (α): 0,05 Hipótesis: $H_e : \rho_{XY} \neq 0 \rightarrow H_o : \rho_{XY} = 0$
2. El coeficiente de correlación de Pearson obtenido a nivel muestral ($r_{XY} = -0,847$) es muy distinto de 0, por tanto, la evidencia empírica parece apoyar la hipótesis de que existe relación entre ambas variables a nivel poblacional.
3. Obtención del estadístico de contraste T :

$$EE(r_{XY}) = \sqrt{\frac{1 - (-0,847)^2}{27 - 2}} = 0,1065 \quad \rightarrow \quad T = \frac{0,847}{0,1065} = 7,95$$

(Al igual que para otras pruebas de significación ya tratadas, este estadístico de contraste consiste en una estandarización del estadístico obtenido en la muestra. Así, si es cierta la H_o ($\rho_{XY} = 0$), la distribución muestral del coeficiente de correlación de Pearson obtenido en muestras de 27 casos extraídas al azar de la población, seguirá una distribución t con 25 grados de libertad y parámetros $E(r_{XY}) = 0$ y $EE(r_{XY}) = 0,1065$.)

4. La distribución muestral del estadístico T en este ejemplo se ajustará a la de la distribución t de Student con 25 grados de libertad ($g.l. = 27 - 2 = 25$). Al buscar en esta distribución cuál es la probabilidad asociado al valor de T , o bien, al obtenerlo con la función de Excel [=DISTR.T(7,95;25;2)] da igual a un valor muy bajo: $Sig = 0,000000026$ ($=2,6E-8$)

5. Decisión: $0,000000026 < 0,05$, por tanto, se rechaza la H_0 y se concluye que hay una relación estadísticamente significativa entre ambas variables o, dicho de otro modo, que la muestra procede de una población en la que $\rho_{XY} \neq 0$.

- Cuando se obtiene en *SPSS* el coeficiente de correlación de Pearson entre 2 variables, los resultados son presentados en forma de matriz –la conocida como matriz de correlaciones–, donde para todas las variables analizadas se muestra el coeficiente de correlación por pares. En cada casilla se muestra también el nivel de significación resultante de aplicar la prueba de significación y, en el caso de que la prueba de significación evidencie una relación estadísticamente significativa, se añade un asterisco ($Sig < 0,05$) o dos ($Sig < 0,01$) junto al valor del correspondiente coeficiente.

Para el **ejemplo** de las variables “Estrategias de afrontamiento” y “Estrés”, la matriz inferior muestra el resultado proporcionado por *SPSS* al obtener el coeficiente de correlación de Pearson entre ambas variables.

SPSS: Analizar > Correlaciones > Bivariadas (Pearson)

		Correlaciones	
		Puntuación escala de estrés	Estrategias de afrontamiento
Puntuación escala de estrés	Correlación de Pearson	1	-,847**
	Sig. (bilateral)		,000
	N	27	27
Estrategias de afrontamiento	Correlación de Pearson	-,847**	1
	Sig. (bilateral)	,000	
	N	27	27

** . La correlación es significativa al nivel 0,01 (bilateral).

3.1.2. Contraste de hipótesis basado en intervalo de confianza

- Pasos en la elaboración del *IC* del coeficiente de correlación de Pearson:

1. Se establece el nivel de riesgo (α) y se plantean las hipótesis estadística y nula:

$$H_e : \rho_{XY} \neq 0 \rightarrow H_o : \rho_{XY} = 0$$

2. Obtención del *IC*: Se aplica primero la *transformación z de Fisher* sobre el valor de correlación muestral. Este coeficiente de correlación transformado se expresa simbólicamente como $Z_{r_{XY}}$

$$z_{r_{XY}} = 0,5 \cdot \ln \left[\frac{1+r_{XY}}{1-r_{XY}} \right]$$

Se calcula el IC transformado de acuerdo a la siguiente expresión:

$$IC(1-\alpha)(z_{\rho_{XY}}) = \left[z_{r_{XY}} + z_{\alpha/2} \cdot \frac{1}{\sqrt{n-3}} ; z_{r_{XY}} + z_{(1-\alpha/2)} \cdot \frac{1}{\sqrt{n-3}} \right]$$

Finalmente, tras calcular el IC de $z_{\rho_{XY}}$, debemos realizar la transformación inversa a fin de obtener el IC de ρ_{XY} :

$$IC(1-\alpha)(\rho_{XY}) = \left[\frac{e^{2 \cdot I_{\text{inf}}} - 1}{e^{2 \cdot I_{\text{inf}}} + 1} ; \frac{e^{2 \cdot I_{\text{sup}}} - 1}{e^{2 \cdot I_{\text{sup}}} + 1} \right]$$

3. Se decide el rechazo de la H_0 cuando el IC de la correlación no contiene el valor 0 expresado en la H_0 . En caso contrario, se mantiene la H_0 .

Para el **ejemplo** de las variables “Estrategias de afrontamiento” y Estrés”, el IC de la correlación para un nivel de confianza del 0,95 se obtendría tal como se muestra a continuación:

$$z_{r_{XY}} = 0,5 \cdot \ln \left[\frac{1-0,847}{1+0,847} \right] = -1,24$$

$$IC(0,95)(z_{\rho_{XY}}) = \left[-1,24 \pm 1,96 \cdot \frac{1}{\sqrt{24}} \right] = [-1,64 ; -0,84]$$

$$IC(1-\alpha)(\rho_{XY}) = \left[\frac{e^{2 \cdot (-1,64)} - 1}{e^{2 \cdot (-1,64)} + 1} ; \frac{e^{2 \cdot (-0,84)} - 1}{e^{2 \cdot (-0,84)} + 1} \right] = \left[\frac{0,038-1}{0,038+1} ; \frac{0,186-1}{0,186+1} \right] = [-0,926 ; -0,657]$$

A partir del resultado obtenido, se decide el rechazo de la H_0 porque el IC de la correlación no contiene el valor expresado en la H_0 , esto es, cero. En consecuencia, se concluye que hay una relación estadísticamente significativa entre ambas variables o, dicho de otro modo, que la muestra procede de una población en la que $\rho_{XY} \neq 0$.

Ejercicio 7: Se sospecha que existe algún tipo de relación entre la dosis ingerida de un determinado ansiolítico ($mg.$) y el tiempo de reacción ante señales acústicas ($ms.$). Para comprobarlo, se realizó un estudio con una muestra de 98 sujetos, obteniéndose un coeficiente de correlación de Pearson entre ambas variables de 0,20. ¿Qué se puede concluir, a partir de la evidencia empírica obtenida, respecto a la existencia de relación entre ambas variables a nivel poblacional? Para contestar, realizad la prueba de significación.

3.2. Contraste de hipótesis acerca del parámetro de la pendiente de la ecuación de regresión

• Se pueden plantear contrastes de hipótesis para los dos parámetros del modelo de regresión lineal simple, sin embargo, nos vamos a centrar aquí únicamente en el parámetro asociado a la variable predictora (β_1) -la pendiente de la ecuación de regresión- y no en el parámetro del origen de la ecuación (β_0), menos relevante desde un punto de vista aplicado. Así, el contraste de hipótesis más habitual en la práctica es acerca de si β_1 es significativamente distinto de 0 ($H_e: \beta_1 \neq 0$), como medio de valorar si se puede considerar como significativamente distinta de cero la contribución de la variable explicativa (X) en la predicción de la variable de respuesta (Y). En cambio, no suele resultar ya tan relevante en la práctica el evaluar si el parámetro de la constante de la ecuación de regresión (β_0) es significativamente distinto de 0 ($H_e: \beta_0 \neq 0$).

3.2.1. Contraste de hipótesis basado en prueba de significación estadística

• Pasos en la prueba T para la pendiente de la ecuación de regresión:

1. Se decide el nivel de riesgo (α) que se desea asumir en el contraste de hipótesis y se plantean las hipótesis estadística y nula. Si realizamos un contraste bilateral:

$$H_e : \beta_1 \neq 0 \quad \text{y, complementariamente,} \quad H_o : \beta_1 = 0$$

2. Explorar si el estadístico de la pendiente obtenido en la muestra (b_1) apoya, en un principio, la hipótesis estadística planteada. Si este valor es 0 (o muy próxima a 0) o de signo contrario al esperado, no tendrá sentido continuar y se mantendría la H_o .

3. Se calcula el estadístico de contraste correspondiente a esta prueba de significación:

$$T = \frac{|b_1|}{EE(b_1)}$$

donde $|b_1|$ es el valor absoluto de la pendiente obtenido en la muestra y $EE(b_1)$ (el error estándar de la distribución muestral del estadístico de la pendiente bajo el supuesto de que sea cierto lo expresado en la hipótesis nula) se obtiene según la siguiente fórmula:

$$EE(b_1) = \sqrt{\frac{SCE}{(n-1) \cdot (n-2) \cdot s_X^2}}$$

(SCE : suma de cuadrados de los errores o, también, de los residuales. Este concepto ha sido introducido en el capítulo sobre el modelo de regresión lineal)

4. Dado que el estadístico de contraste se distribuye de acuerdo a la distribución t con $n-2$ grados de libertad, se obtiene en la citada distribución el nivel de significación (Sig) correspondiente al valor de T obtenido, esto es, la probabilidad de obtener un valor como T o superior. En caso de contraste bilateral, se multiplica por dos el valor de Sig obtenido. También podremos obtenerlo con *Excel* haciendo uso de la función que ofrece este programa para tal fin:

Contraste unilateral \rightarrow =DISTR.T(T ;grados_de_libertad;1)

Contraste bilateral \rightarrow =DISTR.T(T ;grados_de_libertad;2)

De nuevo aquí, si la muestra es de tamaño superior a 30, se puede utilizar la distribución normal en vez de la distribución t a la hora de obtener el valor de Sig .

5. Decisión: se mantiene la H_0 si $Sig > \alpha$; por contra, se rechaza si $Sig < \alpha$.

En el caso que se rechace la H_0 , se suele expresar este resultado diciendo que el valor de la pendiente es estadísticamente significativo, esto es, que el valor de la pendiente de la ecuación de regresión a nivel poblacional es, con un nivel de confianza del $(1 - \alpha)\%$, distinto de cero.

También puede aparecer expresado, en un lenguaje más directo, que la variable explicativa (X) es un predictor estadísticamente significativo de la variable de respuesta (Y).

Ejemplo: Siguiendo con el ejemplo de las variables “Estrategias de afrontamiento” y “Estrés” utilizado para el contraste de hipótesis del coeficiente de correlación de Pearson, supongamos ahora que se haya planteado un modelo predictivo en que la variable “Estrategias de afrontamiento” es considerada como variable explicativa (X) y la variable “Estrés” como variable de respuesta (Y), y que este modelo predictivo se haya plasmado en la obtención, a partir de datos empíricos ($n = 27$), en la siguiente ecuación de regresión lineal:

$$\text{Estrés} = 75,42 - 0,763 \cdot \text{Afrontamiento}$$

Los resultados obtenidos con *SPSS* asociados a este análisis de regresión son los siguientes:

Coefficientes^a

Modelo	Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.	Intervalo de confianza para B al 95%	
	B	Error típ.	Beta			Límite inferior	Límite superior
1 (Constante)	75.425	5.532		13.634	.000	64.031	86.819
Estrategias de afrontamiento	-.763	.096	-.847	-7.951	.000	-.961	-.566

a. Variable dependiente: Puntuación escala de estrés

ANOVA^b

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	9320.650	1	9320.650	63.216	.000 ^a
	Residual	3686.017	25	147.441		
	Total	13006.667	26			

a. Variables predictoras: (Constante), Estrategias de afrontamiento

b. Variable dependiente: Puntuación escala de estrés

A partir de la estimación del parámetro de la pendiente de la ecuación de regresión ($b_1 = -0,763$), vamos a realizar la prueba de significación estadística que permita contrastar la hipótesis de que ese parámetro es distinto de 0 a nivel poblacional:

1. Nivel de riesgo (α) = 0,05. Hipótesis asumiendo un contraste bilateral:

$$H_e : \beta_1 \neq 0 \rightarrow H_o : \beta_1 = 0$$

2. El valor de la pendiente obtenido a nivel muestral ($-0,763$), distinto de 0, apoya la hipótesis estadística planteada por lo que continuaremos con los siguientes pasos.
3. Obtención del estadístico de contraste T :

(Un dato necesario para poder obtener el $EE(b_1)$ que no aparece en el enunciado de este ejemplo: la cuasi-varianza de la variable “Estrategias de afrontamiento” es igual a 615,03)

$$EE(b_1) = \sqrt{\frac{3686,02}{26 \cdot 25 \cdot 615,03}} = 0,096$$

$$T = \frac{0,763}{0,096} = 7,95$$

(Al igual que para otras pruebas de significación ya tratadas, este estadístico de contraste consiste en una estandarización del estadístico obtenido en la muestra. Así, la distribución muestral del estadístico de la pendiente de la ecuación de regresión obtenido en muestras de 27 casos extraídas al azar de una población en que ese parámetro sea 0 ($\beta_1 = 0$), seguirá una distribución t con parámetros $E(\beta_1) = 0$ y $EE(\beta_1) = 0,096$.)

4. La distribución muestral del estadístico T en este ejemplo será la distribución t de Student con 25 grados de libertad. Al buscar en esta distribución cuál es la probabilidad de obtener un valor como el obtenido para T o más extremo se obtiene para el caso de un contraste bilateral (introduciendo en *Excel* la función: =DISTR.T(7,95;25;2)): $Sig = 0,000000026$.
5. Decisión: $0,000000026 < 0,05$, por lo tanto, se rechaza la H_o y se concluye que el valor del parámetro de la pendiente es distinto de 0 (con un nivel de confianza del 95%) o, en otros términos más aplicados, que la variable “Estrategias de afrontamiento” es un predictor estadísticamente significativo del “Estrés“.

Nota importante sobre esta prueba de significación: Si comparamos el valor del estadístico de contraste obtenido en este ejemplo con el calculado en el ejemplo anterior de la prueba de significación del coeficiente de correlación de Pearson, llegaremos a una conclusión inmediata: se trata de un contraste equivalente. Así, si para el coeficiente de correlación entre dos variables obtenemos un resultado estadísticamente significativo, también lo obtendremos para la pendiente de la ecuación de regresión de una variable sobre la otra.

3.2.2. Contraste de hipótesis basado en intervalo de confianza

• Pasos en la elaboración del *IC* de la pendiente de la ecuación de regresión:

1. Se establece el nivel de riesgo (α) y la hipótesis a contrastar. En el caso de un contraste de hipótesis bilateral:

$$H_e : \beta_1 \neq 0 \rightarrow H_o : \beta_1 = 0$$

2. El *IC* de la pendiente de la ecuación de regresión se obtiene de acuerdo a la siguiente expresión:

$$IC(1-\alpha)(\beta_1) = \left[b_1 + t_{(n-2)}(\alpha/2) \cdot EE(b_1) ; b_1 + t_{(n-2);(1-\alpha/2)} \cdot EE(b_1) \right]$$

donde:

$$EE(b_1) = \sqrt{\frac{SCE}{(n-1) \cdot (n-2) \cdot s_x^2}}$$

Recordar que a partir de $n \geq 30$, la distribución normal y la distribución t son muy próximas entre sí y, por tanto, se pueden utilizar los valores de la distribución normal asociados al nivel de confianza que se establezca.

3. Se decide el rechazo de la H_o cuando el *IC* de la pendiente de regresión [$IC(1-\alpha)(\beta_1)$] no contiene el valor 0 expresado en la H_o . En caso contrario, se mantiene la H_o .

Para el **ejemplo** de las variables “Estrategias de afrontamiento” y Estrés”, el *IC* del parámetro de la pendiente de regresión vendría dado por la siguiente expresión (siendo $\alpha = 0,05$):

(Algunos datos de interés para su cálculo: $EE(b_1) = 0,096$; $t_{25;0,975} = 2,06$)

$$IC(0,95)(\rho_{XY}) = [-0,763 \pm 2,06 \cdot 0,096] = [-0,961 ; -0,566]$$

Dado que el *IC* no contiene el valor 0, se rechaza la H_o y, por lo tanto, se considera el parámetro de la pendiente como significativamente distinto de cero, es más, el *IC* obtenido evidencia una relación alta entre ambas variables, pues el límite inferior de dicho *IC* es ya un valor elevado de por sí.

Ejercicio 8: En un estudio en enseñanza primaria en que se pretendía poner de manifiesto la posible influencia de las expectativas que de los estudiantes tienen los profesores (X) sobre el rendimiento académico de los mismos (Y), se obtuvo a partir de una muestra de 200 estudiantes la siguiente ecuación de regresión de Y sobre X : $Y' = 5,7 + 0,20 \cdot X$. A partir de este resultado, ¿se puede considerar estadísticamente significativa la contribución de la variable “Expectativas del profesor” (X) a la hora de predecir el “Rendimiento académico” (Y) de los estudiantes?

Realiza la prueba de significación estadística correspondiente, así como el intervalo de confianza de la pendiente de la ecuación de regresión ($\alpha = 0,05$). Datos de interés: $s_X^2 = 9$; $SCE = 796$

Referencias

- Losilla, J. M., Navarro, B., Palmer, A., Rodrigo, M. F., y Ato, M. (2005). *Del contraste de hipótesis al modelado estadístico*. Tarrasa: CBS (www.edicionsapeticio.com).
- Loftus, E. F., y Burns, T. E. (1982). Mental shock can produce retrograde amnesia. *Memory and Cognition*, 10, 318-323.
- Pardo, A., Ruiz, M.A. y San Martín, R. (2009). *Análisis de datos en ciencias sociales y de la salud I*. Madrid: Síntesis.