T. 4 − Contraste de hipótesis

1. Definición y conceptos básicos

1.1. Definición de contraste de hipótesis

• Proceso mediante el cual se intenta comprobar si una afirmación sobre alguna propiedad poblacional puede ser sostenida a la luz de la información muestral disponible (Pardo y San Martín, 1998).

• El contraste de hipótesis se enmarca en el proceder habitual del método científico:

(1) Laguna de conocimiento / incertidumbre.

Ejemplo: Parecen existir diferencias entre mujeres y hombres en la capacidad para orientarse en el espacio.

(2) Conjetura explicativa de esa incertidumbre que pueda ser verificada a partir de datos obtenidos de forma empírica → Hipótesis científica (los conceptos o atributos implicados en la misma deben aparecer expresados de forma precisa, así como el modo en que éstos van a ser medidos.)

Ejemplo: La orientación en el espacio, entendida como [...] y medida a través de [...], es distinta en mujeres y hombres.

(3) Expresión en términos estadísticos de la hipótesis científica → Hipótesis estadística (H_e)

Ejemplo: Existen diferencias estadísticamente significativas en la puntuación media de hombres y mujeres en la prueba *X* de orientación espacial → $H_e: \mu_{X_{Mujeres}} \neq \mu_{X_{Hombres}}$

(4) Contraste de hipótesis: proceso orientado a comprobar si la H_e planteada es compatible con la evidencia empírica obtenida a partir de una muestra de la población de interés.

1.2. De la hipótesis científica a la hipótesis estadística

Una hipótesis estadística es la expresión de una hipótesis científica en términos de parámetros.

A modo de **ejemplo** de expresión de una hipótesis científica en forma de hipótesis estadística:

Hipótesis científica → En el actual plan de estudios de la licenciatura de Psicología de la UVEG, el rendimiento académico (notas) en las asignaturas de 2º es mayor que en las de 1º.

Hipótesis estadística → H_e: $\mu_{Notas\,2^o} > \mu_{Notas\,1^o}$ $(o\ Md_{Notas\,2^o} > Md_{Notas\,1^o}, u\ otras\ posibles)$.

Otro **ejemplo**:

Hipótesis científica → En la población de trabajadores de España la prevalencia de "mobbing" es inferior al 10%.

Hipótesis estadística → H_e: $\pi_{mobbing} < 0{,}10$.

Nótese que, mientras que en la 1ª hipótesis hay implicados dos parámetros (o, lo que es lo mismo, dos poblaciones), la 2ª hipótesis se refiere al valor de un único parámetro (y hay implicada una única población).

• Apunte terminológico: cuando en la H_e aparecen los operadores relacionales < o > se habla de contraste unilateral, mientras que cuando aparece el ≠ se habla de contraste bilateral.

Ejercicio 1: Expresar las siguientes hipótesis científicas en forma de H_e. Determina, para cada una de ellas, si hay implicados uno o dos parámetros, y si son contrastes bilaterales o unilaterales.

(1) La proporción de sujetos con ideación suicida en la población de personas con depresión es superior a 0,40.

(2) El tiempo medio de reconocimiento visual de palabras monosilábicas reales (no pseudo-palabras) está por debajo de los 230 msg.

(3) La incidencia del fracaso escolar es diferente en zonas rurales o urbanas.

(4) La variabilidad de las puntuaciones en la escala de inteligencia de Weschler es distinta para los asiáticos y para los europeos.
(5) Los esquizofrénicos tienen mayor capacidad en el reconocimiento de patrones visuales repetitivos que los que no están diagnosticados de esta patología.
(6) Existe algún tipo de relación entre el nivel de inseguridad y la agresividad de las personas.

1.3. La lógica del contraste de hipótesis

• El contraste de hipótesis representa un desarrollo de la estadística inferencial que permite, a partir de la obtención de información en muestras, poner a prueba hipótesis a nivel poblacional. Obviamente, no tendrá sentido su aplicación cuando se tenga acceso a la información poblacional, caso en el que la verificación de la H_e es inmediata.

> **Ejemplo**: Sea el caso de la hipótesis "En el actual plan de estudios de la licenciatura de Psicología de la UVEG, el rendimiento académico (notas) en las asignaturas de 2º es mayor que en las de 1º, cuyo contraste se va a plantear según la H_e: $\mu_{Notas\,2º} > \mu_{Notas\,1º}$. ¿Cómo se llevaría a cabo en la práctica?
>
> La comprobación de H_e sería inmediata cuando se dispusiese de de las notas de 1º y 2º para todos los estudiantes en el actual plan de estudios: bastaría con obtener la nota media en ambas poblaciones y comparar ambas medias. Esto no es lo más habitual en la práctica, así que en el caso en que sólo dispusiésemos de las notas para una muestra de estudiantes -dos en realidad, una muestra de estudiantes de 1ª y otra de 2ª-, habría que plantear el contraste hipótesis derivado de la aplicación de la estadística inferencial que permita inferir si esas 2 medias son diferentes a nivel poblacional.

• El modo de proceder en el contraste de hipótesis se sustenta sobre la lógica falsacionista, esto es, se asume a priori que es cierta la hipótesis que plantea la relación contraria a la planteada en la H_e. A esta hipótesis se le suele denominar como *hipótesis nula* (H_o) dado que en la misma se recoge la igualdad o inexistencia de relación entre los atributos implicados en la H_e; por su parte, a esta última se le suele denominar también como *hipótesis alternativa* (H_1), dada su complementariedad con la H_o. Ambas hipótesis son exhaustivas, en tanto que recogen todos los posibles resultados, y mutuamente excluyentes.

Ejemplos para las hipótesis formuladas anteriormente,

$$H_e\text{: } \mu_{Notas\,2^o} > \mu_{Notas\,1^o} \qquad \rightarrow H_o\text{: } \mu_{Notas\,2^o} \leq \mu_{Notas\,1^o}$$

$$H_e\text{: } \pi_{mobbing} < 0{,}10 \qquad \rightarrow H_o\text{: } \pi_{mobbing} \geq 0{,}10$$

El caso es que si la evidencia empírica recogida (resultados obtenidos a partir de una muestra de la población) *discrepa* en gran medida de lo que plantea la H_o, se rechaza ésta y se acepta la H_1. Si, por el contrario, los resultados empíricos no discrepan significativamente de lo planteado en la H_o, entonces se mantiene ésta. En definitiva, que es la H_o la que realmente se somete a contraste con la evidencia empírica. En el siguiente apartado se verá cómo se hace operativa esta discrepancia.

• Consecuencias de la lógica que subyace al contraste de hipótesis:

(1) Cuando se rechaza la H_o podemos afirmar que hemos conseguido probar, en términos probabilísticos, que esa hipótesis es falsa y, en consecuencia, que los resultados obtenidos apoyan la H_e, Ello suele expresarse en los informes de investigación diciendo que se ha obtenido "una diferencia / relación / resultado estadísticamente significativo" -en cuanto que las H_e siempre plantean algún tipo de resultado consistente en una diferencia o en una relación.

(2) Cuando no se rechaza (o se mantiene) la H_o en un contraste de hipótesis, ello implica obtener evidencia empírica a favor de que la H_e es falsa. Ahora bien, ello no significa que se pruebe que la H_o es verdadera. Si se mantiene la H_o, lo único que se puede afirmar es que se ha obtenido evidencia empírica compatible con la misma.
En los informes de investigación se suele expresar este tipo de resultado diciendo que se ha obtenido "una diferencia / relación / resultado que no es estadísticamente significativo", o bien, diciendo que "los resultados no son concluyentes" –aludiendo a que no se ha obtenido evidencia empírica que apoye la H_e, que es la que refleja lo que se quiere poner a prueba.

Ejercicio 2: Formular las H_o correspondientes a las H_e planteadas en el ejercicio anterior. Tener en cuenta los siguientes requisitos en la formulación de una H_o: (1) que aparezca el signo de igualdad en la misma (=, ≥ o ≤); y (2) que las H_o y H_e sean complementarias.

2. Aplicación del contraste de hipótesis

• Se pueden distinguir dos estrategias generales a la hora de llevar a cabo un contraste de hipótesis y, más concretamente, en el análisis de la discrepancia entre lo que asume la H_o y los resultados obtenidos a partir de los datos recogidos en una muestra: (1) el basado en la realización de pruebas de significación estadística (muy divulgado y utilizado, sigue siendo la estrategia implementada por la mayoría de los paquetes estadísticos y la que tiene más presencia en las publicaciones científicas); (2) el basado en la utilización de intervalos de confianza (más intuitivo e informativo, su divulgación en la docencia de la estadística, así como su aparición en las publicaciones científicas, se va extendiendo de forma paulatina). Ambos permiten tomar una decisión sobre el mantenimiento o el rechazo de la H_o y, en consecuencia, sobre el apoyo o no a la H_e.

A continuación se va a describir el fundamento de ambas aproximaciones, así como su aplicación en la realización de algunos de los contrastes de hipótesis más utilizados. En temas sucesivos, sin embargo, se volcará la atención hacia la estrategia basada en pruebas de significación, no por una mayor predilección hacia la misma, sino por ser coherentes con la mayor presencia de éstas en programas informáticos y publicaciones.

2.1. Contraste de hipótesis basado en pruebas de significación

• El contraste de hipótesis basado en pruebas de significación consiste en calcular la probabilidad de que se dé un resultado como el obtenido en la muestra, partiendo del supuesto de que es cierto lo que aparece expresado en la H_o. Si esa probabilidad, conocida como nivel de significación (*P* o *Sig*), es muy pequeña se rechaza la H_o y se acepta la H_e. El valor por debajo del cual se considera lo suficientemente pequeña esa probabilidad es una convención en la práctica de la estadística, es denominado α o nivel de riesgo, y se establece generalmente en 0,05 o 0,01. En consecuencia, la prueba de significación conduce a un rechazo de la H_o siempre que $P < \alpha$.

• En la práctica de la ejecución de una prueba de significación, la obtención del nivel de significación (*P* o *Sig*) asociado al resultado obtenido en una muestra para un determinado estadístico (por ejemplo, el valor de la media aritmética de una variable), asumiendo que la H_o es cierta, supone:

> 1º) Transformar ese resultado muestral de acuerdo a una expresión conocida como estadístico de contraste, de los que existe un amplio repertorio adecuado a los distintos parámetros que pueden aparecer implicados en una H_e.
>
> 2º) Dicho estadístico es una variable aleatoria que tiene una distribución muestral que se ajusta a un patrón conocido, suponiendo que es cierta la H_o. Las distribuciones muestrales para los

estadísticos de contraste más utilizados se ajustan a las siguientes distribuciones de probabilidad: la curva normal, la distribución binomial, la distribución ji-cuadrado, la distribución t y la distribución F. Para estas distribuciones existen tablas que se pueden encontrar en los apéndices de la mayoría de los libros de Estadística.

3º) Obtener, en la tabla estadística que se corresponda con la distribución muestral del estadístico de contraste utilizado, el nivel de significación (*Sig*), es decir, la probabilidad asociada al valor del estadístico de contraste obtenido en la muestra.

2.1.1. Caso del contraste de hipótesis sobre una proporción

• El objetivo de este contraste de hipótesis es decidir si una hipótesis sobre un determinado valor (k) de proporción de una variable en una población es apoyada, o no, por la evidencia empírica obtenida a partir de una muestra de esa población.

• Procedimiento:

1. Se decide el nivel de riesgo (α) que se desea asumir en el contraste de hipótesis (habitualmente, 0,05 o 0,01) y se plantean las hipótesis estadística y nula. Tres posibles hipótesis en este caso:

$$H_e : \pi_X > k \rightarrow H_o : \pi_X \leq k \quad (contraste\ unilateral)$$
$$H_e : \pi_X \neq k \rightarrow H_o : \pi_X = k \quad (contraste\ bilateral)$$
$$H_e : \pi_X < k \rightarrow H_o : \pi_X \geq k \quad (contraste\ unilateral)$$

2. Explorar si la proporción empírica obtenida en la muestra parece apoyar, en principio, la hipótesis estadística planteada. Si no la apoya no tiene sentido continuar y se mantendría la H_o. El resultado puede resultarnos de interés a la hora de plantear hipótesis estadísticas más afinadas en el futuro.

3. Se obtiene el siguiente estadístico de contraste (este estadístico no es más que la tipificación del valor de la proporción obtenido en la muestra, tal y como se ilustrará en el siguiente ejemplo):

$$z = \frac{p - k}{\sqrt{k \cdot (1 - k)/n}}$$

4. Se obtiene en la distribución binomial o, en el caso de que la muestra sea grande, en la distribución normal estandarizada, la probabilidad de obtener un valor como el obtenido con el

estadístico de contraste o más extremo, esto es, el nivel de significación (*P* o *Sig*). Multiplicar por dos el valor obtenido si el contraste que se lleva a cabo es bilateral, esto es,

$$Sig\ (\text{bilateral}) = Sig\ (\text{unilateral}) \times 2$$

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

Ejemplo de contraste de hipótesis acerca de una proporción poblacional (adaptada de Losilla y cols., 2005). Supóngase que en la presente legislatura el gobierno se plantea como objetivo que más del 40% de los colegios públicos deberían contar con pizarras digitales. En caso de que no se llegue a este mínimo, el gobierno propondrá una partida presupuestaria para cubrir esta deficiencia. Para averiguar si este mínimo se cumple, se selecciona aleatoriamente una muestra de 20 colegios públicos del país, donde se observa que 13 de ellos disponen de pizarras digitales en las aulas.

1. Nivel de riesgo (α) = 0,05 -o lo que es lo mismo, nivel de confianza (1- α) = 0,95. La hipótesis estadística es unilateral ya que el problema establece que la proporción de colegios con pizarras digitales sea *superior* al 40%

$$H_e : \pi > 0{,}40 \ \rightarrow H_0 : \pi \leq 0{,}40 \quad (\textit{contraste unilateral})$$

2. La proporción obtenida en la muestra parece apoyar la hipótesis del investigador, ya que la proporción de colegios públicos con pizarras digitales (p = 13/20 = 0,65) es superior a 0,40.
3. Se obtiene el valor del estadístico de contraste:

$$z = \frac{0{,}65 - 0{,}40}{\sqrt{\dfrac{0{,}40 \cdot 0{,}60}{20}}} = 2{,}27$$

Veamos el origen de este estadístico de contraste: Si la H_o es cierta, esto es, en la población de colegios públicos de España hay, como máximo, un porcentaje del 40% de colegios con pizarras digitales (π=0,40), se sigue que si obtenemos al azar todas las posibles muestras de 20 colegios de esa población y obtenemos el valor de la proporción (*p*) de colegios con pizarra digital en cada una de las muestras, la tª de la probabilidad determina que la distribución muestral del estadístico proporción seguirá una distribución normal (se cumple la condición de muestra grande ya que $n \cdot \pi = 20 \cdot 0{,}4 = 8 > 5$ y $n \cdot (1-\pi) = 20 \cdot 0{,}6 = 12 > 5$) con parámetros:

$$\mu_{P_{Xi}} \; [E(p_{Xi})] = 0,40; \qquad \sigma_{P_{Xi}} \; [EE(p_{Xi})] = \sqrt{\frac{0,40 \cdot 0,60}{20}} = 0,11$$

La proporción obtenida en la muestra (p=0,65) es una de entre todas las que se podrían haber obtenido y pertenece a la distribución normal anterior. Por tanto, si queremos obtener cuál es la probabilidad de obtener, en una muestra determinada, una proporción como la que hemos obtenido sólo tenemos que estandarizar esta proporción (transformarla en una puntuación z) para poder utilizar las tablas de la distribución normal unitaria o estandarizada. Así pues, el estadístico de contraste anterior no es más que la tipificación de la proporción obtenida en la muestra utilizando como distribución de referencia la que teóricamente se obtendría si es cierta la H_o.

4. El nivel de significación (*P* o *Sig*) se obtiene de las tablas de la distribución normal estandarizada y es igual a:

$$Sig = P(z \geq 2,27) = 0,0116$$

La interpretación de *Sig* es la siguiente: Si la H_o es cierta, esto es, si en la población de colegios públicos de España hay, como máximo, una proporción de 0,40 que tienen pizarras digitales, la probabilidad de obtener en una muestra de 20 colegios de esa población una proporción de 0,65 o superior es de 0,0116. En consecuencia, es poco verosímil la H_o, los datos empíricos no la apoyan.

5. Decisión: dado que 0,0116 < 0,05 se rechaza la H_o y se mantiene, en consecuencia, la hipótesis alternativa de que la proporción de colegios con pizarra digital es superior a 0,4. Nótese que si hubiera fijado α en 0,01, la decisión hubiera sido mantener la H_o dado que 0,0116 > 0,01.

La siguiente figura ilustra el desarrollo de la prueba de significación estadística para el ejemplo anterior:

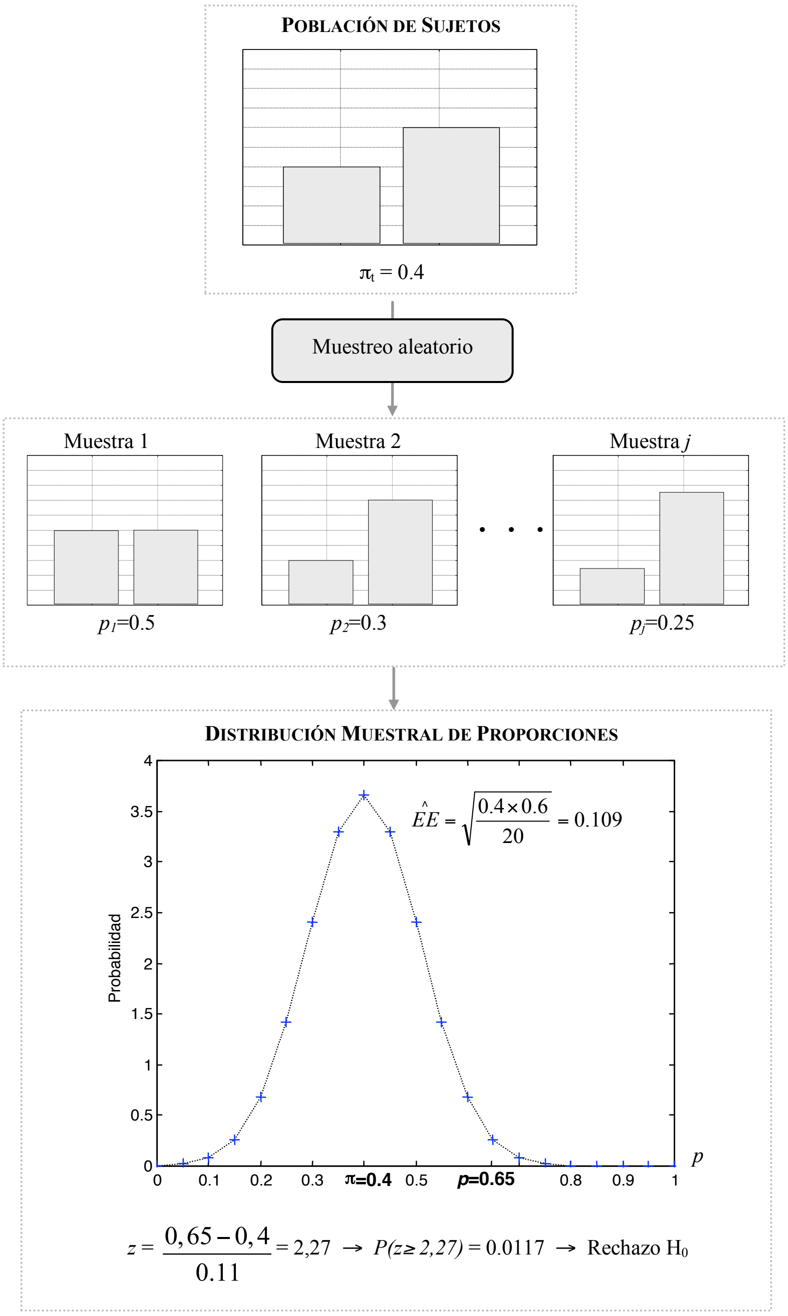


Figura 1. Prueba de conformidad para la comparación de una proporción observada p con una teórica $\pi_t = 0.4$ en muestras de tamaño $n = 20$ (Losilla y cols., 2005; inspirada en la figura presentada por Pardo y San Martín, 1991, p. 62).

Ejemplo con SPSS a partir de los datos obtenidos con el Cuestionario de Vida Académica:

La proporción de mujeres entre los estudiantes de Estadística en Psicología de la UVEG en una muestra de n = 174 sujetos de dicha materia y titulación es de 0,816. Sabiendo que la proporción de mujeres entre los estudiantes de Estadística en Psicología de la Universidad

Complutense de Madrid (UCM) es del 80% ¿Podemos afirmar que la proporción de mujeres en la población de estudiantes de Estadística de la UVEG es distinta con α=0.05?

$$H_e\text{: } \pi_{(\text{UVEG})} \neq 0{,}80 \;\rightarrow\; H_o\text{: } \pi_{(\text{UVEG})} = 0{,}80 \text{ (contraste bilateral);}$$

$$p = 0{,}816; \quad n=174; \qquad EE(p) = \sqrt{\frac{0.80 \cdot 0.20}{174}} = 0{,}03$$

$$\text{Prueba de significación: } z = \frac{0{,}816 - 0{,}80}{\sqrt{\frac{0{,}80 \cdot 0{,}20}{174}}} = \frac{0{,}016}{0{,}03} = 0{,}53$$

$$P\,(Sig) = P\,(Z \leq -0{,}53) + P\,(Z \geq 0{,}53) = 0{,}2981 \cdot 2 = 0{,}5962 \;\rightarrow\; \text{Se mantiene } H_o$$

SPSS: Pruebas no paramétricas| Binomial (contrastar proporción=0,80)

Prueba binomial

		Categoría	N	Proporción observada	Prop. de prueba	Sig. asintót. (unilateral)	Sig. exacta (unilateral)
sexo	Grupo 1	Mujer	142	,816	,8	,337[a]	,337
	Grupo 2	Hombre	32	,184			
	Total		174	1,0			

a. Basado en la aproximación Z.

Nota: El resultado obtenido con SPSS no coincide exactamente con el obtenido al realizar la prueba de significación. Para obtener con el programa la significación bilateral hay que multiplicar por dos la significación unilateral: (P = 0,337×2) = 0,674

Y si la hipótesis alternativa fuera que la proporción de mujeres en la población de estudiantes de UVEG es superior a la de la UCM ¿Cuál sería el nivel de significación correcto?

$$H_e\text{: } \pi_{(\text{UVEG})} > 0{,}80 \text{ (unilateral)} \;\rightarrow\; H_o\text{: } \pi_{(\text{UVEG})} \leq 0{,}80$$

$$Sig = \text{P}\,(Z \geq 0{,}53) = 0{,}2981 \text{ (SPSS: Prueba exacta: } Sig = 0{,}337)$$

2.1.2. Caso del contraste de hipótesis sobre una media

• El objetivo de este contraste de hipótesis es decidir si una determinada hipótesis sobre el valor (k) de la media de una variable en una población es apoyada, o no, por la evidencia empírica ($\bar{X}$) obtenida a partir de una muestra de esa población. Veamos como abordar este contraste de hipótesis a partir de la realización de una prueba de significación.

• Procedimiento:

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. Tres posibles hipótesis en este caso:

$$H_e: \mu_X > k \rightarrow H_o: \mu_X \leq k \ \ (contraste\, unilateral)$$
$$H_e: \mu_X \neq k \rightarrow H_o: \mu_X = k \ \ (contraste\, bilateral)$$
$$H_e: \mu_X < k \rightarrow H_o: \mu_X \geq k \ \ (contraste\, unilateral)$$

2. Explorar si la media empírica obtenida parece 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_o. El resultado puede resultarnos de interés a la hora de plantear hipótesis estadísticas más afinadas en el futuro.
3. Se calcula uno de los dos siguientes estadísticos de contraste:
   - En el caso de ser conocida la varianza en la población: $Z = \dfrac{\bar{X} - k}{\frac{\sigma_X}{\sqrt{n}}}$
   - En el caso más habitual de no conocerse la varianza en la población: $T = \dfrac{\bar{X} - k}{\frac{s'_X}{\sqrt{n}}}$
4. Se obtiene en la distribución normal estandarizada, para el primer caso, o en la distribución *t* con *n-1* grados de libertad, para el segundo, la probabilidad de obtener un valor como el obtenido con el estadístico de contraste o más extremo, esto es, el nivel de significación (*Sig*). Multiplicar por dos el valor obtenido si el contraste que se lleva a cabo es bilateral. (Nota: si la muestra es grande (> 30), se puede utilizar la distribución normal en vez de la distribución *t*)
5. Decisión: se mantiene la H_o si $Sig > \alpha$; por contra, se rechaza si $Sig < \alpha$.

Ejemplo: Diferentes trabajos sobre memoria icónica han mostrado que el promedio de letras recordadas por sujetos normales en presentación taquistoscópica de 1 segundo es de 4,5 letras (μ = 4,5), con una desviación típica de 1,4 (σ = 1,4). Sin embargo, un grupo de investigadores sospecha que tal afirmación puede no ser correcta y, para comprobarlo, selecciona a una muestra de 25 sujetos a los que se les aplica la citada tarea, obteniéndose un promedio de palabras recordadas de 5,1 ($\bar{X} = 5{,}1$). ¿Apoya este estudio la hipótesis de este grupo de investigadores de que la media de letras recordadas a nivel poblacional es distinta de 4,5?

Prueba de significación:

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

$$H_e: \mu \neq 4{,}5 \ \rightarrow \ H_o: \mu = 4{,}5$$

2. El resultado muestral ($\bar{X} = 5,1$) parece apoyar, en un principio, la hipótesis de que la media de letras recordadas no es de 4,5.
3. Cálculo del estadístico de contraste:

$$z_{\bar{X}} = \frac{\bar{X} - k}{\frac{\sigma_X}{n}} = \frac{5,1 - 4,5}{\frac{1,4}{\sqrt{25}}} = 2,14$$

Al igual que en el caso de la proporción, este estadístico de contraste es una tipificación de la media obtenida en la muestra ya que, si es cierta H_o ($\mu = 4,5$), la distribución muestral de las medias obtenidas en muestras de n=25 extraídas al azar de la población anterior seguirá una distribución normal con parámetros:

$$\mu = 4,5;\ EE = \frac{1,4}{\sqrt{25}} = 0,28$$

4. $Sig = P(z \geq 2,14) = 0,016$

Como se trata de un contraste bilateral → $Sig = 0,016 \cdot 2 = 0,032$

5. Decisión: 0,032 < 0,05, por tanto, se rechaza la H_o

Ejemplo con SPSS a partir de los datos obtenidos con el Cuestionario de Vida Académica:
La edad media de los estudiantes de Estadística en Psicología de la UVEG en una muestra de n = 174 sujetos de dicha materia y titulación se situó en 21,15 años y la cuasi-desviación típica en 5,06 años. Sabiendo que la media de edad entre los estudiantes de Estadística en Psicología de la Universidad complutense de Madrid es de 22 años ¿Podemos afirmar que la media en la UVEG es diferente con α=0,05?

H_e: $\mu_{\text{(UVEG)}} \neq 22$ (bilateral); H_o: $\mu_{\text{(UVEG)}} = 22$

$\bar{X} = 21,15;\ s'_X = 5,06;\ n = 174$

Prueba de significación:

$$EE(\bar{X}) = \frac{5,06}{\sqrt{174}} = 0,384$$

$$t = \frac{21,15 - 22}{0,384} = -2,21$$

$Sig = P(t \geq 2,21) = 0,0136 \times 2 = 0,027$ → Rechazo H_o

Obsérvese la equivalencia con los resultados obtenidos con SPSS:

SPSS: Analizar | Prueba T para una muestra (Valor de prueba=22):

Estadísticos para una muestra

	N	Media	Desviación típ.	Error típ. de la media
edad	174	21,15	5,060	,384

Prueba para una muestra

	Valor de prueba = 22					
					95% Intervalo de confianza para la diferencia	
	t	gl	Sig. (bilateral)	Diferencia de medias	Inferior	Superior
edad	-2,217	173	,028	-,851	-1,61	-,09

- ¿Cuál hubiera sido la decisión considerando α=0,01?

- Y si la hipótesis fuera que la media en la UVEG es inferior a la media en la UCM ¿Cuál sería el nivel de significación correcto?

$$H_e\text{: } \mu_{\text{(UVEG)}} < 22 \rightarrow H_o\text{: } \mu_{\text{(UVEG)}} \geq 22 \quad \text{(contraste unilateral)}$$

P ($t \leq -2{,}21$) = 0,014 (o, también, dividiendo por 2 la significación bilateral proporcionada en el output de SPSS: 0,028/2 = 0,014) → Rechazo H_o (para α=0,05)

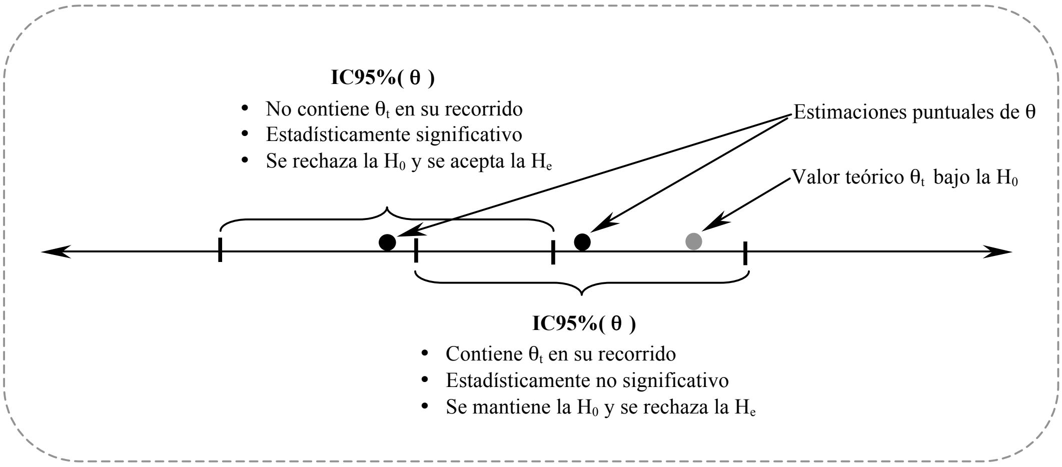
2.2. Contraste de hipótesis basado en intervalos de confianza

• Esta segunda aproximación al contraste de hipótesis supone obtener el *IC* del parámetro implicado en la H_o en base a los datos obtenidos empíricamente (muestra) y de acuerdo al nivel de riesgo asumido en el contraste de hipótesis. Una vez obtenido el *IC*, se comprueba si el resultado obtenido sustenta el mantenimiento o el rechazo de dicha hipótesis.

2.2.1. Caso del contraste de hipótesis basado en intervalos de confianza acerca de un parámetro poblacional (media, proporción, varianza...)

• Procedimiento:

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.
2. Se obtiene el *IC* a partir del valor del estadístico obtenido en la muestra.
3. Se decide el rechazo de la H_o en el caso en que el *IC* no contenga en su recorrido el valor teórico bajo la H_o, en caso contrario se mantiene la H_o.



Losilla y cols. (2005): Contraste de hipótesis en base al intervalo de confianza del parámetro poblacional.

Ejemplo de contraste de hipótesis acerca de una media poblacional: Diferentes trabajos sobre memoria icónica han mostrado que el promedio de letras recordadas por sujetos normales en presentación taquistoscópica durante 1 segundo es de 4,5 letras (μ = 4,5 letras). Sin embargo, un grupo de investigadores sospecha que tal afirmación puede no ser correcta y, a fin de comprobarlo, selecciona a una muestra de 35 sujetos a los que se les aplica la citada tarea, obteniéndose un promedio de palabras recordadas de 5,1 ($\bar{X} = 5,1$ letras), con una cuasi-desviación típica de 1,4 (s'_x = 1,4 letras).

1. Nivel de riesgo (α) = 0,05, o lo que es lo mismo, un nivel de confianza ($1-\alpha$) = 0,95

$$H_e : \mu \neq 4,5 \rightarrow H_o : \mu = 4,5$$

2. $IC(1-\alpha)(\mu_X) = \left[\bar{X} \pm z_{1-\alpha/2} \cdot \frac{s'_X}{\sqrt{n}} \right]$

$$IC(0,95)(\mu_X) = \left[5,1 + \pm 1,96 \cdot \frac{1,4}{\sqrt{35}} \right] = [4,64; 5,56]$$

3. Decisión: La H_o (μ = 4,5) no aparece sustentada por la evidencia empírica procedente de la muestra (el *IC*), por tanto, se rechaza la H_o y se acepta la H_e.

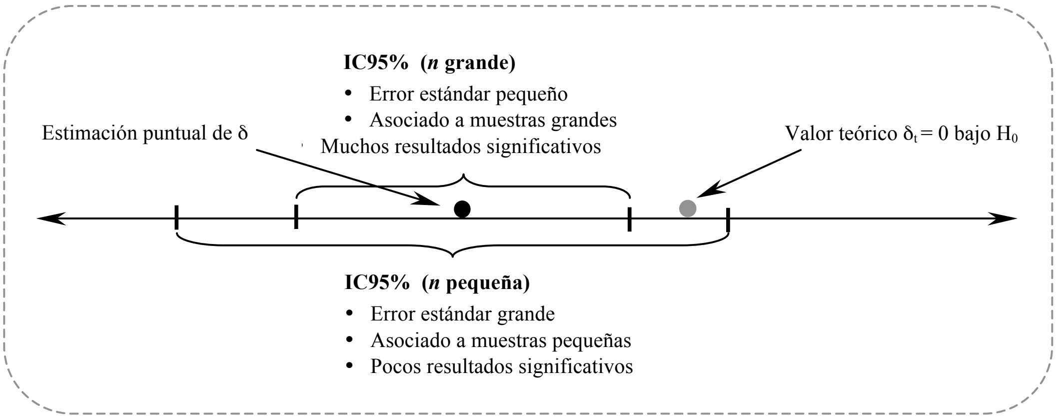
3. Factores que influyen en el rechazo de la hipótesis nula

• El rechazo de la H_o en un contraste de hipótesis dependerá fundamentalmente del grado de discrepancia entre el enunciado de la H_o y el resultado muestral. Ahora bien, una misma discrepancia puede ser estadísticamente significativa o no en función de factores como: el tamaño de la muestra, el riesgo de error α fijado a priori para el contraste y que este contraste sea unilateral o bilateral.

3.1. Significación estadística y tamaño muestral

• Un problema asociado al contraste de hipótesis es que la decisión de mantener o rechazar la H_o viene determinada, aparte de por el grado de discrepancia entre la H_o y el resultado obtenido a nivel empírico (muestra), por el tamaño de la muestra utilizada en el estudio. Así, cuanto mayor es el tamaño de la muestra, disminuye el *EE* de la distribución muestral del estadístico (proporción o media) y ello influye tanto en el *IC* obtenido como en la prueba de significación.

• Por lo que respecta al *IC*, al aumentar la muestra éste se hace más estrecho (preciso) y, por tanto, más fácil es que el valor del parámetro establecido en la H_o quede fuera del *IC* y que, en consecuencia, se rechace la H_o.



Losilla y cols. (2005). Relación entre la significación estadística, la magnitud de la diferencia observada y el tamaño muestral.

Ejemplo de la orientación vocacional y la madurez (adaptado de Pardo y San Martin, 1998): De acuerdo con los datos recogidos durante los últimos años por un psicólogo escolar, los estudiantes de COU obtienen una puntación media de 190 en una determinada prueba de madurez. El psicólogo cree que si los estudiantes recibiesen orientación vocacional obtendrían una puntuación superior en la mencionada prueba. Para obtener evidencia sobre ello, toma una muestra aleatoria de *n* estudiantes de Bachiller a los que proporciona orientación vocacional y, a posteriori, les aplica la citada prueba de madurez obteniendo una puntuación media de 195 y una cuasi-desviación típica de 24. ¿Se puede considerar que hay una mejora en la puntuación en la prueba para los estudiantes que reciben orientación vocacional?

$\mu_x = 190 \rightarrow$ Contraste de hipótesis: $\alpha = 0,05$ $\quad H_e : \mu_X > 190 \quad H_o : \mu_X \leq 190$

Supónganse los dos casos siguientes en que sólo varía el tamaño de muestra considerado:

Caso 1 $\rightarrow$ Muestra A: $n = 30 \quad \bar{X} = 195 \quad S'_X = 24$

$IC(0,95)(\mu_x)$ = [186,4 ; 203,6] => Se mantiene la H_o

Caso 2 $\rightarrow$ Muestra B: $n = 100 \quad \bar{X} = 195 \quad S'_X = 24$

$IC(0,95)(\mu_x)$ = [190,3 ;199,7] => Se rechaza la H_o

• De lo anterior se puede deducir que se podría llegar a forzar el rechazo de una H_o simplemente eligiendo un tamaño muestral lo suficientemente grande. Ello no es del todo así porque la reducción en la anchura de un *IC* no está linealmente relacionada con el incremento del tamaño de la muestra. Por otro lado, no hay que olvidar que cuanto mayor sea el tamaño muestral considerado, mayor será también el coste asociado a la recogida de los datos.

Ejemplo de la orientación vocacional y la madurez:

$\mu_x = 190 \rightarrow$ Contraste de hipótesis: $\alpha = 0,05$ $\quad H_e : \mu_X > 190 \quad H_o : \mu_X \leq 190$

Muestras A, B, C y D: $\bar{X} = 195 \quad S'_X = 24$

Muestra A (n = 30) => $IC(0,95)(\mu_x)$ = [186,4; 203,6]

Muestra B (n = 100) => $IC(0,95)(\mu_x)$ = [190,3; 199,7]

Muestra C (n = 170) => $IC(0,95)(\mu_x)$ = [191,4; 198,6]

Muestra D (n = 240) => $IC(0,95)(\mu_x)$ = [191,9; 198,1]

• Por otra parte, la prueba de significación también se ve influida por el tamaño de la muestra ya que, al disminuir el *EE* de la distribución muestral del estadístico (media o proporción), aumentará el valor del estadístico de contraste y, en consecuencia, disminuirá el valor de *Sig*.

• Se ha planteado alguna estrategia alternativa a la decisión dicotómica de considerar la relación planteada en la H_o como estadísticamente significativa o no (o sea, rechazo de la H_o vs. mantenimiento de la H_o). Una de ellas se basa en la utilización de indicadores de tamaño del efecto, esto es, un tipo de indicador continuo (no dicotómico) de la magnitud de la diferencia o de la relación planteada en la H_o y que tiene la particularidad de no estar influido por el tamaño de la muestra.

Un problema asociado a los mismos viene determinado a la hora de establecer una decisión relativa a la magnitud de esos índices (¿cuándo se puede decir que es pequeño, o grande, el valor de los mismos?).

3.2. Otros factores

• El rechazo de la H_o en un contraste de hipótesis dependerá también del riesgo de error α fijado a priori para el contraste y de que este contraste sea unilateral o bilateral. Así, por ejemplo, si al contrastar una determinada H_o se obtiene que P=0,03, ésta se rechazará si α= 0,05, pero no si fijamos el valor α= 0,01.

• En cuanto a que el contraste sea unilateral o bilateral si, por ejemplo, al realizar el contraste de una hipótesis unilateral obtenemos que P = 0,03 se rechazaría H_o (siendo α= 0,05), pero si el contraste anterior fuera bilateral el valor P sería igual a 0,06 (0,03×2) y la decisión sería mantener la H_o (siendo α= 0,05). Así pues, un contraste bilateral es siempre más conservador que un contraste unilateral, de manera que si una hipótesis se rechaza siendo el contraste bilateral, también se rechazaría si fuera unilateral, pero no a la inversa.

4. Significación estadística y relevancia práctica

• El rechazo de la H_o en un contraste de hipótesis dependerá fundamentalmente del grado de discrepancia entre el enunciado de la H_o y el resultado muestral. Ahora bien, que una discrepancia sea estadísticamente significativa (rechazo de la H_o) no implica que sea importante o relevante en la práctica.

• La valoración de la relevancia práctica de un resultado sólo tiene sentido cuando previamente se ha rechazado la H_o, pero no si ésta se mantiene. La forma de valorar esa relevancia, cuando proceda, es a partir del *IC* del parámetro al que se refiere la hipótesis.

Ejemplo de la adherencia al tratamiento con esquizofrénicos: La principal causa de recaída de los esquizofrénicos es la interrupción o disminución del tratamiento farmacológico al que, habitualmente de por vida, deben someterse. Con el fin de mejorar la adherencia al tratamiento, en 1995 se inició una campaña informativa sobre las consecuencias de la disminución del tratamiento, ya que en esa fecha se sabía que, en promedio, los enfermos tomaban únicamente el 71% de la dosis prescrita. En

1998 se decidió evaluar la eficacia de dicha campaña, para lo que se seleccionó al azar una muestra de 70 sujetos afectos de esquizofrenia, de quienes se obtuvo el porcentaje que representaba el consumo real de fármacos respecto al consumo dictado por el terapeuta, es decir, la adherencia al tratamiento. Se halló que, la media de la dosis prescrita que tomaban los sujetos era del 76%, con una cuasi-desviación típica de 15.

$\mu_x = 71$ → Contraste de hipótesis: $\alpha = 0{,}05$ $H_e : \mu_X > 71$ $H_o : \mu_X \leq 71$

Muestra: $n = 70$ $\bar{X} = 76$ $S'_X = 15$ → $IC(0{,}95)(\mu_x) = [72{,}5;\ 79{,}5]$

Decisión: Se rechaza la H_o => Diferencia estadísticamente significativa; pero, ¿diferencia relevante a nivel práctico? - Dependerá de la diferencia que se considere como relevante en la práctica.

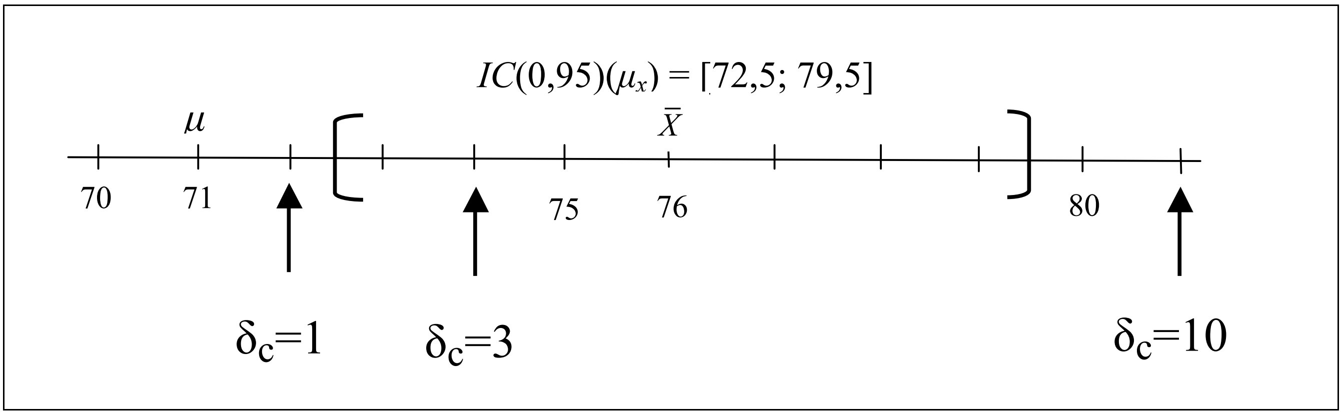
• Establecimiento de un criterio de relevancia o de significación práctica (δ_c o δ*) que permita determinar cuándo la diferencia es lo suficientemente grande como para que se puede considerar como relevante en la práctica. Ese valor se puede fijar de acuerdo a criterios prácticos, clínicos, socio-económicos, u otros que se considere importantes, si bien, se puede dar el caso de que se carezca de argumentos para asignar un valor concreto como criterio de relevancia o que no tenga mucho sentido hacerlo.

Ejemplo de la adherencia al tratamiento con esquizofrénicos:

$\mu_x = 71$ →Contraste de hipótesis: $\alpha = 0{,}05$ $H_e : \mu_X > 71$ $H_o : \mu_X \leq 71$

Muestra: $n = 70$ $\bar{X} = 76$ $S'_X = 15$ → $IC(0{,}95)(\mu_x) = [72{,}5;\ 79{,}5]$

- Si se considerase, por los motivos que sea, que $\delta_c = 1$ (una unidad por encima del parámetro de referencia), entonces, aparte de una diferencia estadísticamente significativa, podremos decir que la diferencia es relevante en la práctica (y, por tanto, relevante el efecto de la campaña informativa).



- Si, por ejemplo, se considera como criterio de relevancia práctica $\delta_c = 10$, entonces, tendríamos una diferencia que sigue siendo estadísticamente significativa pero que no es relevante en la práctica.
- Y, como último ejemplo, en el caso que se considerase $\delta_c = 3$, nada cambia respecto a la significación estadística, pero no se podría afirmar nada respecto a la significación práctica, tan sólo decir que el resultado es no concluyente al respecto y sugerir que se repita el estudio y, a poder ser, con un mayor tamaño muestral a fin de incrementar la precisión del intervalo de confianza.

• Como se deriva de la exposición anterior, aunque *IC* y prueba de significación son dos procedimientos alternativos que nos permiten tomar una decisión sobre el rechazo o no de la H_o, la aproximación basada en el *IC* es más informativa ya que, además, nos permite estimar el valor del parámetro poblacional de interés y, por tanto, valorar la relevancia práctica del resultado obtenido en el estudio.

Ejercicio 3 (basado en el enunciado del ejercicio de la orientación vocacional y la prueba de madurez con una muestra de tamaño $n = 100$, media muestral 195 y cuasi-desviación típica 8). De acuerdo a los resultados del contraste de hipótesis, señalar cuál es la significación estadística y práctica en cada uno de los cuatro casos siguientes:

(1) Los "*teóricos de la madurez*" consideran como un efecto relevante de la formación en orientación vocacional, que en la media de la prueba de madurez haya una mejora de dos puntos.

(2) …haya una mejora de cuatro puntos.

(3) …haya una mejora de seis puntos.

(4) …haya una mejora de ocho puntos.

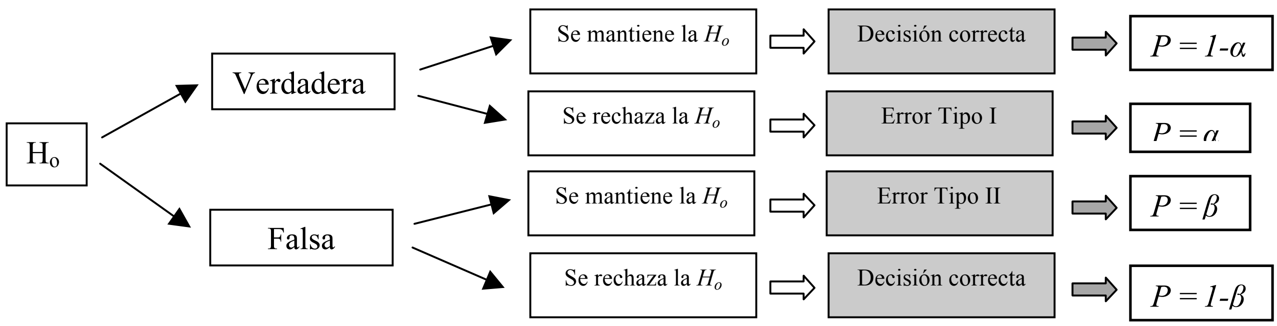
$\mu_x = 190$ → Contraste de hipótesis: $\alpha = 0,05$ $H_e : \mu_X > 190$ $H_o : \mu_X \leq 190$

Muestra: $n = 100$ $\bar{X} = 195$ $S'_X = 8$ → $IC(0,95)(\mu_x) = [193,4; 196,6]$

5. Errores asociados al contraste de hipótesis

• El rechazo o no de la hipótesis nula en un contraste de hipótesis es una decisión que lleva asociada una determinada probabilidad de error. Así, al rechazar una H_o nula cuando el nivel de significación *P* es inferior al valor α fijado a priori podemos cometer un error, dado que podría ser cierta la H_o. Así, podría ocurrir que, siendo cierta la hipótesis nula, obtuviéramos por azar en la muestra un valor del estadístico de contraste cuya probabilidad fuera inferior a 0,05, y esto podría ocurrir en un 5% de las posibles muestras obtenidas. A este riesgo de error de rechazar H_o siendo que es verdadera se le conoce como error tipo I y la probabilidad de ocurrencia del mismo es igual al nivel de riesgo α fijado a priori. Al complementario de este valor se le suele conocer como nivel de confianza ($1-\alpha$) y, consecuentemente, representa la probabilidad de tomar una decisión acertada, esto es, mantener la H_o siendo que ésta es en realidad la verdadera.

• Si un posible error en un contraste de hipótesis consiste en rechazar la H_o siendo ésta cierta (error tipo I), el otro error posible, complementario al anterior, consiste en mantener la H_o siendo que es falsa (o lo que es lo mismo, no aceptar la H_e siendo que es la hipótesis verdadera). A este error se le conoce como error tipo II y la probabilidad de ocurrencia del mismo se expresa con la letra β.



• La probabilidad asociada a la decisión complementaria al error tipo II (esto es, la decisión de aceptar la H_e cuando la H_o es falsa) se le conoce como potencia del contraste de hipótesis y se expresa simbólicamente como ($1-\beta$). Su magnitud depende de diversos factores, entre ellos, el tamaño de la muestra y el nivel de riego (α) asumido en el contraste de hipótesis. Cuanto mayor son éstos, mayor es también (*1−β*). El cálculo de la potencia es facilitado por el programa SPSS sólo para algunas pruebas de contraste de hipótesis.

Referencias

Pardo, A. y San Martín, R. (1998). *Análisis de datos en psicología II (2ª ed.).* Madrid: Pirámide.

Losilla, J. M., Navarro, B., Palmer, A., Rodrigo, M. F., y Ato, M. (2005). *Del contraste de hipótesis al modelado estadístico.* Documenta Universitaria. [www.edicionsapeticio.com]