Bioestadística y uso de software científico

TEMA 6

CÁLCULO DEL TAMAÑO MUESTRAL

Índice

- Requisitos generales para el cálculo del tamaño muestral
- Estimación de una proporción
- Estimación de una media
- Comparación de dos proporciones
- Comparación de dos medias
- Cálculo de la potencia del estudio

Requisitos para estimar el tamaño muestral

- Para hacer los cálculos, el investigador debe fijar antes 4 parámetros:
  - Error α
  - Error β
  - Desviación estándar prevista
  - Diferencia mínima que se desea detectar en la comparación

Requisitos para estimar el tamaño muestral

- Error α
  - Habitualmente 0,05 (5%)
  - Cuanto menor sea α, mayor es el tamaño muestral necesario

Requisitos para estimar el tamaño muestral

- Error β
  - Habitualmente 0,20 (20%) ó 0,10 (10%)
  - Cuanto menor sea β, mayor es el tamaño muestral necesario

Requisitos para estimar el tamaño muestral

- Desviación estándar prevista
  - Se puede obtener de estudios de otros investigadores
  - O de un estudio piloto en la misma población
  - Cuanto mayor sea la desviación estándar, mayor es el tamaño muestral necesario
  - En muchos estudios falta tamaño muestral porque la desviación estándar real fue mayor que la prevista

Requisitos para estimar el tamaño muestral

- Diferencia mínima que se desea detectar en la comparación
  - “Magnitud del efecto”
  - ¿En cuánto bajará la tensión arterial con el medicamento?
  - Lo fija el investigador; debe tener sentido biológico
  - Cuanto menor sea la diferencia que se quiere detectar, mayor será el tamaño muestral necesario

Requisitos para estimar el tamaño muestral

Parámetro	Si el parámetro disminuye…
Error α	$\uparrow$ n
Error β	$\uparrow$ n
Desviación estándar prevista (s)	$\downarrow$ n
Diferencia mínima (d)	$\uparrow$ n

Índice

- Requisitos generales para el cálculo del tamaño muestral
- Estimación de una proporción
- Estimación de una media
- Comparación de dos proporciones
- Comparación de dos medias
- Cálculo de la potencia del estudio

Estimación de una proporción

- ¿Cuántos partos debemos estudiar para conocer la frecuencia de síndrome de bajo peso al nacer con error α = 5%, precisión = 4%, si suponemos que la frecuencia que obtendremos es 0,2?
  - Error α: 5% $\longrightarrow z_{\alpha/2}=1{,}96$
  - Desviación estándar esperada:

$$s=\sqrt{p(1-p)}=\sqrt{0,2\times 0,8}=0,4$$

  - Precisión: d = 0,04

Estimación de una proporción

- ¿Cuántos partos debemos estudiar para conocer la frecuencia de síndrome de bajo peso al nacer con error α = 5%, precisión = 4%, si suponemos que la frecuencia que obtendremos es 0,2?
  - Calcular:

$$n = \frac{z_{\alpha/2}^2 \, p(1-p)}{d^2}$$

$$n = \frac{1,96^2 \times 0,2 \times 0,8}{0,04^2} = 384,2 \rightarrow 385 \; \textit{partos necesarios}$$

Estimación de una proporción

- ¿Y si la precisión deseada hubiera sido del 2%?
  - Error α: 5% $\longrightarrow z_{\alpha/2}=1{,}96$
  - Desviación estándar esperada:

$$s=\sqrt{p(1-p)}=\sqrt{0,2\times 0,8}=0,4$$

  - Precisión: d = 0,02

Estimación de una proporción

- ¿Y si la precisión deseada hubiera sido del 2%?
  - Calcular:

$$n = \frac{z_{\alpha/2}^2 p(1-p)}{d^2}$$

$$n = \frac{1,96^2 \times 0,2 \times 0,8}{0,02^2} = 1536,6 \rightarrow 1537 \ \textit{partos necesarios}$$

  - Reducir la precisión a la mitad ⟶Multiplicar por 4 el tamaño muestral

Estimación de una proporción

- ¿Y si la frecuencia esperada es 50%?
  - Error α: 5%$\longrightarrow z_{\alpha/2}$=1,96
  - Desviación estándar esperada:

$$s = \sqrt{p(1-p)} = \sqrt{0,5 \times 0,5} = 0,5$$

  - Precisión: d = 0,04

Estimación de una proporción

- ¿Y si la frecuencia esperada es 50%?
  - Calcular:

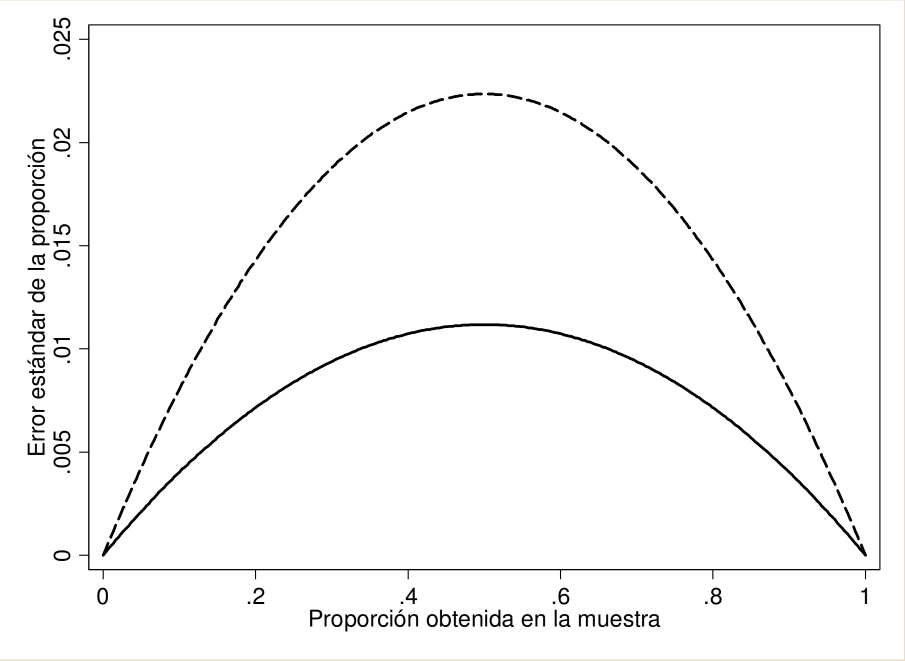
$$n = \frac{z_{\alpha/2}^2 p(1-p)}{d^2}$$

$$n = \frac{1{,}96^2 \times 0{,}5 \times 0{,}5}{0{,}04^2} = 600{,}3 \rightarrow 601 \textit{ partos necesarios}$$

  - El mayor tamaño muestral se produce cuando la frecuencia esperada es 50% ("máxima indeterminación")

Estimación de una proporción

- ¿Y si la frecuencia esperada es 50%?
  - El mayor tamaño muestral se produce cuando la frecuencia esperada es 50% (“máxima indeterminación”)



Estimación de una proporción

- ¿Y si el error α es 1%?
  - Error α: 1%⟶$z_{\alpha/2}$=2,6
  - Desviación estándar esperada:

$$s = \sqrt{p(1-p)} = \sqrt{0,2 \times 0,8} = 0,4$$

  - Precisión: d = 0,04

Estimación de una proporción

- ¿Y si el error α es 1%?
  - Calcular:

$$n = \frac{z_{\alpha/2}^2 p(1-p)}{d^2}$$

$$n = \frac{2,6^2 \times 0,2 \times 0,8}{0,04^2} = 676 \ \textit{partos necesarios}$$

  - Disminuir α⟶aumenta n

Índice

- Requisitos generales para el cálculo del tamaño muestral
- Estimación de una proporción
- Estimación de una media
- Comparación de dos proporciones
- Comparación de dos medias
- Cálculo de la potencia del estudio

Estimación de una media

- Número de sujetos necesario para medir la concentración de PCBs en sangre en la población cántabra; queremos que el intervalo de confianza al 90% tenga una anchura de 0,3µg; un estudio previo realizado en Cataluña indica que la desviación estándar es 1µg.
  - IC 90% $\longrightarrow \alpha=10\% \longrightarrow z_{\alpha/2}=1{,}64$
  - $s=1$
  - Anchura=0,3 $\longrightarrow$ precisión: $d=0{,}15$

Estimación de una media

- IC 90%⟶α=10%⟶$z_{\alpha/2}$=1,64
- s=1
- Anchura=0,3⟶precisión: d=0,15
- Calcular:

$$n = \frac{z_{\alpha/2}^2 s^2}{d^2}$$

$$n = \frac{1,64^2 \times 1^2}{0,15^2} = 119,5 \rightarrow 120 \textit{ sujetos necesarios}$$

Estimación de una media

$$n = \frac{z_{\alpha/2}^2 s^2}{d^2}$$

- Menor $\alpha \longrightarrow$ mayor $z_{\alpha/2} \longrightarrow$ mayor n
- Mayor s$\longrightarrow$mayor n
- Menor anchura$\longrightarrow$menor d$\longrightarrow$mayor n

Estimación de una media

- Tamaño muestral para estimar una media

- Influencia del tamaño muestral en el error estándar de la media

Índice

- Requisitos generales para el cálculo del tamaño muestral
- Estimación de una proporción
- Estimación de una media
- Comparación de dos proporciones
- Comparación de dos medias
- Cálculo de la potencia del estudio

Comparación de dos proporciones

- Queremos saber si la mortalidad con el fármaco A es menor que con el fármaco B. Esperamos una mortalidad media del 10%, consideramos que un descenso del 1% es relevante. Queremos medirlo con error α del 5% y error β del 20%.
  - $\alpha=5\% \longrightarrow z_{\alpha/2}=1{,}96$
  - $\beta=20\% \longrightarrow z_{\beta}=0{,}84$
  - $p=0{,}1 \longrightarrow$ $s=\sqrt{p(1-p)}=\sqrt{0{,}1\times 0{,}9}=0{,}3$
  - $d=0{,}01$

Comparación de dos proporciones

- $\alpha=5\% \longrightarrow z_{\alpha/2}=1,96$
- $\beta=20\% \longrightarrow z_{\beta}=0,84$
- $p=0,1 \longrightarrow s=\sqrt{p(1-p)}=\sqrt{0,1\times 0,9}=0,3$
- $d=0,01$
- Calcular:

$$n=\frac{2\left(z_{\alpha/2}+z_{\beta}\right)^2 s^2}{d^2}$$

$$n=\frac{2\times(1,96+0,84)^2\times 0,3^2}{0,01^2}=14112 \textit{ sujetos con cada fármaco}$$

Comparación de dos proporciones

$$n = \frac{2\left(z_{\alpha/2} + z_{\beta}\right)^2 s^2}{d^2}$$

- Menor $\alpha \longrightarrow$ mayor $z_{\alpha/2} \longrightarrow$ mayor n
- Menor $\beta \longrightarrow$ mayor $z_{\beta} \longrightarrow$ mayor n
- Mayor s $\longrightarrow$ mayor n
- Menor d $\longrightarrow$ mayor n

Índice

- Requisitos generales para el cálculo del tamaño muestral
- Estimación de una proporción
- Estimación de una media
- Comparación de dos proporciones
- Comparación de dos medias
- Cálculo de la potencia del estudio

Comparación de dos medias

- Queremos comparar la capacidad de dos antiarrítmicos (A y B) para bajar la frecuencia cardíaca. Consideramos relevante detectar un descenso de 5 latidos por minuto, con error a = 10% y error b = 10%. Un estudio piloto indica que la desviación estándar puede ser 8 latidos/minuto
  - α=10%$\longrightarrow z_{\alpha/2}$=1,64
  - β=10%$\longrightarrow z_{\beta}$=1,28
  - s=8
  - d=5

Comparación de dos medias

- $\alpha=10\% \longrightarrow z_{\alpha/2}=1{,}64$
- $\beta=10\% \longrightarrow z_{\beta}=1{,}28$
- s=8
- d=5
- Calcular

$$n=\frac{2\left(z_{\alpha/2}+z_{\beta}\right)^2 s^2}{d^2}$$

$$n=\frac{2\times(1,64+1,28)^2\times 8^2}{5^2}=43,7\rightarrow 44\ en\ cada\ grupo$$

Índice

- Requisitos generales para el cálculo del tamaño muestral
- Estimación de una proporción
- Estimación de una media
- Comparación de dos proporciones
- Comparación de dos medias
- Cálculo de la potencia del estudio

Cálculo de la potencia estadística

- Cuando el estudio no obtiene diferencias significativas, el investigador (o el lector) debe preguntarse: ¿el estudio tenía capacidad para detectar estas diferencias?

- Eso es calcular la potencia estadística:
  - Potencia = $1-\beta$
  - Potencia = capacidad del estudio para rechazar la hipótesis nula cuando es falsa

Cálculo de la potencia estadística

- En el estudio anterior, la n necesaria era 44 por grupo. Al llevarlo a cabo, sólo se pudieron reclutar 30 por grupo. ¿Qué potencia tenía el estudio para detectar una diferencia de 5 latidos/minuto con error α = 5%?
  - $\alpha=5\% \longrightarrow z_{\alpha/2}=1{,}96$
  - $s=8$
  - $d=5$
  - $n=30$
  - ¿$1-\beta$?

Cálculo de la potencia estadística

- $\alpha=5\% \longrightarrow z_{\alpha/2}=1,96$
- s=8
- d=5
- n=30
- ¿1-β?

$$n=\frac{2\left(z_{\alpha/2}+z_{\beta}\right)^{2}s^{2}}{d^{2}} \Rightarrow$$

$$z_{\beta}=\sqrt{\frac{nd^{2}}{2s^{2}}}-z_{\alpha/2}$$

Cálculo de la potencia estadística

- $\alpha=5\% \longrightarrow z_{\alpha/2}=1,96$
- s=8
- d=5
- n=30
- ¿1-β?

$$z_\beta = \sqrt{\frac{nd^2}{2s^2}} - z_{\alpha/2}$$

$$z_\beta = \sqrt{\frac{30\times5^2}{2\times8^2}} - 1,96 = 0,46 \rightarrow \beta = 0,32 \rightarrow potencia = 0,68$$

Cálculo de la potencia estadística

- Para calcular la potencia en un estudio sobre diferencia de proporciones, la fórmula es la misma teniendo en cuenta que $s^2=p(1-p)$

$$z_{\beta} = \sqrt{\frac{nd^2}{2s^2}} - z_{\alpha/2} = \sqrt{\frac{nd^2}{2p(1-p)}} - z_{\alpha/2}$$