

ESTIMACION DEL

TAMAÑO DE LA

MUESTRA Y DE LA

POTENCIA

HIPOTESIS Y PRINCIPIOS

Sabemos a quién y qué vamos a estudiar.
Ahora hay que decidir cuántos individuos contendrá la muestra.

- **Hipótesis nula** (H_0) → En la población no existe asociación entre las variables predictoras y de desenlace.
- **Hipótesis alternativa** (H_1) → Proposición de que existe una asociación. No puede examinarse directamente: se acepta por rechazo de H_0 .

HIPOTESIS Y PRINCIPIOS

- **H1 de una cola** → (Unilateral) especifica la dirección de la asociación entre variables predictoras y de desenlace → predicción de que los pacientes con cáncer de páncreas consumen más tabaco que los controles.
- **H1 de dos colas** → (Bilateral) afirma que hay asociación; no especifica la dirección → los pacientes con cáncer de páncreas tendrán una tasa diferente de consumo de café (más alta o más baja).

Principios estadísticos subyacentes

- **Error tipo I** (Falso positivo) → Rechazo H_0 que es verdadera en la población.
- **Error tipo II** (Falso negativo) → Aceptación de H_0 cuando es falsa en la población.
- Los errores tipo I y II no son totalmente evitables pero mejoran (Disminuyen) aumentando la muestra.

Principios estadísticos subyacentes

- **Tamaño del efecto:** Si la magnitud de asociación entre variables predictora y de desenlace en la población es grande, será fácil detectarla en la muestra y viceversa.
- No conocemos la magnitud real de la asociación y debemos elegir el tamaño de la asociación que nos *gustaría* ser capaces de detectar en la muestra. Esta cantidad se denomina tamaño del efecto. Seleccionar el efecto de tamaño adecuado es el aspecto más difícil al planificar el tamaño de la muestra.

Principios estadísticos subyacentes

- Lo basamos en otras experiencias y en el sentido común, eligiendo el % que creamos clínicamente relevante.
- La probabilidad máxima de cometer el error tipo I se llama α .
- La probabilidad máxima de cometer el error tipo II se llama β .
- Normalmente los fija el investigador antes de comenzar el estudio.

Principios estadísticos subyacentes

- α = Nivel de significación estadística.
- $1-\beta$ = Potencia o poder.
- En general elegiremos un valor pequeño de α cuando queramos evitar los FP (falsos positivos) y pequeño de β cuando queramos evitar los FN (falsos negativos).
- Valor $p \rightarrow$ Se rechaza la H_0 para $p < \alpha$ y hablamos de "significación".

Variabilidad: A mayor variabilidad o diseminación de la variable de desenlace, mayor posibilidad de superposición y más difícil demostrar una diferencia global entre ambos grupos.

Importante en variables continuas (Figura)

- Los abandonos o pérdidas de seguimiento no cuentan en el tamaño de la muestra y debe preverse al planificar.
- Atención a las variables de confusión

Técnicas de cálculo de la muestra en estudios analíticos y experimentales

Diversas variantes, pero tienen *pasos en común*:

- 1.- Plantear H_0 y H_1 (de una o dos colas)
- 2.- Seleccionar la prueba estadística
- 3.- Elegir un tamaño razonable del efecto y variabilidad
- 4.- Fijar α y β (para α unilateral o bilateral)
- 5.- Usar la tabla o fórmula.

Prueba χ^2 o estadígrafo Z

Para comparar la *proporción* de individuos en cada grupo que tienen cierto desenlace dicotómico.

En un experimento o en un estudio de cohortes:

- P1 → proporción esperada de individuos que presentarán el desenlace en un grupo
- P2 → proporción esperada en el otro grupo.

En casos y controles:

- P1 → proporción de casos esperados con un riesgo específico
- P2 → proporción esperada de controles con el factor de riesgo

Deberemos:

1.- Plantear H_0 y H_1 (con una o dos colas)

2.- Estimar el tamaño del efecto y la variabilidad en:

$P_1 \rightarrow$ proporción con el desenlace en un grupo.

$P_2 \rightarrow$ proporción con el desenlace en otro grupo.

3.- Fijar α y β

4.- Aplicar la fórmula (en general):

$$N = [Z_{\alpha} \sqrt{P(1-P)(1/q_1 + 1/q_2)} + Z_{\beta} \sqrt{P_1(1-P_1)(1/q_1) + [P_2(1-P_2)(1/q_2)]}]^2 / (P_1 - P_2)^2$$

q_1 = proporción de individuos del grupo 1

q_2 = proporción de individuos del grupo 2

N = número total de individuos

$$P = q_1 P_1 + q_2 P_2$$

Prueba de la t

Si el valor medio de la variable continua en un grupo difiere del valor medio del otro. Se supone que la distribución de la variable continua en los dos grupos se aproxima a la Normal.

Pasos:

- 1.- Plantear H_0 y H_1 (decidir si una o dos colas).
- 2.- Tamaño del efecto \rightarrow diferencia entre los valores medios de variables de desenlace de los grupos de estudio.
- 3.- Variabilidad (como desviación estándar)
- 4.- Fijar α y β
- 5.- Aplicar la fórmula:

$$N = [(1/q_1 + 1/q_2)S^2(Z_\alpha + Z_\beta)^2] \div E^2$$

$Z_\alpha = 1,96$ ($\alpha = 0,05$); $1,645$ ($\alpha = 0,10$)

$Z_\beta = 0,84$ ($\beta = 0,20$); $1,282$ ($\beta = 0,10$)

E = Tamaño estandarizado del efecto (tablas)

q_1 = proporción de individuos del grupo 1

q_2 = proporción de individuos del grupo 2

N = número total requerido de individuos

Coeficiente de correlación

No se utiliza habitualmente para determinar el tamaño de la muestra, pero puede ser útil para ello cuando las dos variables son continuas.

Pasos:

- 1.- H_0 y H_1 (una o dos colas).
- 2.- Tamaño del efecto \rightarrow menor valor absoluto de r que se desea detectar. (La variabilidad está en función de r y está incluida en la tabla).
- 3.- Fijar α y β .
- 4.- Mirar tabla

Situaciones especiales

A veces no hay técnicas disponibles o son difíciles de usar como en regresión logística, por ejemplo. Calcular el tamaño como si fuese un análisis más simple.

En variables ordinales:

- * si tienen bastantes categorías → como continua.
- * si tienen 3 ó 4 categorías → dicotomizar.

Estudios descriptivos

Se basa en principios algo diferentes.

Se utilizan intervalos de confianza (IC) y nivel de confianza.

VARIABLES CONTINUAS → IC alrededor de la media.

- 1.- Estimar la DS (desviación estándar)
- 2.- Especificar la precisión deseada (amplitud total) del IC.
- 3.- Seleccionar el nivel de confianza.
- 4.- Mirar tabla

Variables dicotómicas

IC alrededor de la proporción.

- 1.- Estimar la proporción (p) esperada (si $p > 0,5$, se utiliza q)
- 2.- Especificar la precisión (amplitud) del IC
- 3.- Seleccionar el nivel de confianza
- 4.- Mirar tabla

Estrategias para reducir al mínimo el tamaño muestral

Cuando el tamaño de la muestra es mayor que el número de individuos que se pueden estudiar de manera realista:

- repetir el cálculo ya que es fácil equivocarse
- revisar tamaño del efecto y variabilidad
- $\uparrow \alpha$ ó β
- H1 de bilateral a unilateral
- \downarrow nivel de confianza y estrechar o ampliar el IC

Estrategias para ↓ al mínimo el tamaño muestral

- Varias estrategias para ↓ el tamaño (o ↑ potencia)
- ¿Se continúa respondiendo a la pregunta pral.
- Uso de variables continuas
- Uso de variables más precisas → ↓ variabilidad
- Uso de mediciones apareadas: Una basal y otra al concluir el estudio → variable de desenlace es el cambio
- Usar grupos de tamaño desigual → Gralmente ↑ los controles
- Desenlace más frecuente → para desenlaces dicotómicos

Estrategias para reducir al mínimo el tamaño muestral

Ante información insuficiente → Cuando el investigador utiliza un instrumento diseñado por él mismo o usa una muestra no seleccionada antes.

- Buscar datos anteriores → revisión bibliográfica
- Contacto con otros investigadores.
- Realizar un pequeño estudio (con n° reducido de individuos)
- Dicotomizar variables continuas con media y DE dudosas