Introducción al cálculo del tamaño muestral

Marta Roqué

Barcelona, 15/10/2010

Talleres virtuales para investigadores

Organización Panamericana de la Salud
Red Cochrane Iberoamericana

Por favor,

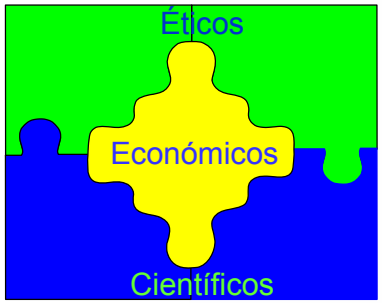
- No connecten su micrófono de Elluminate
- Desactiven la opción de video de Elluminate: enlentece la conexión
- Apaguen sus móviles: les causarán interferencias
- Todos los mensajes de chat son visibles para los moderadores
- Hay moderación para problemas técnicos
- La presentación se está grabando y se colgará en la web posteriormente

Tamaño muestral

Crédito: Iñaki Perez

Tamaño muestral: ¿Por qué?

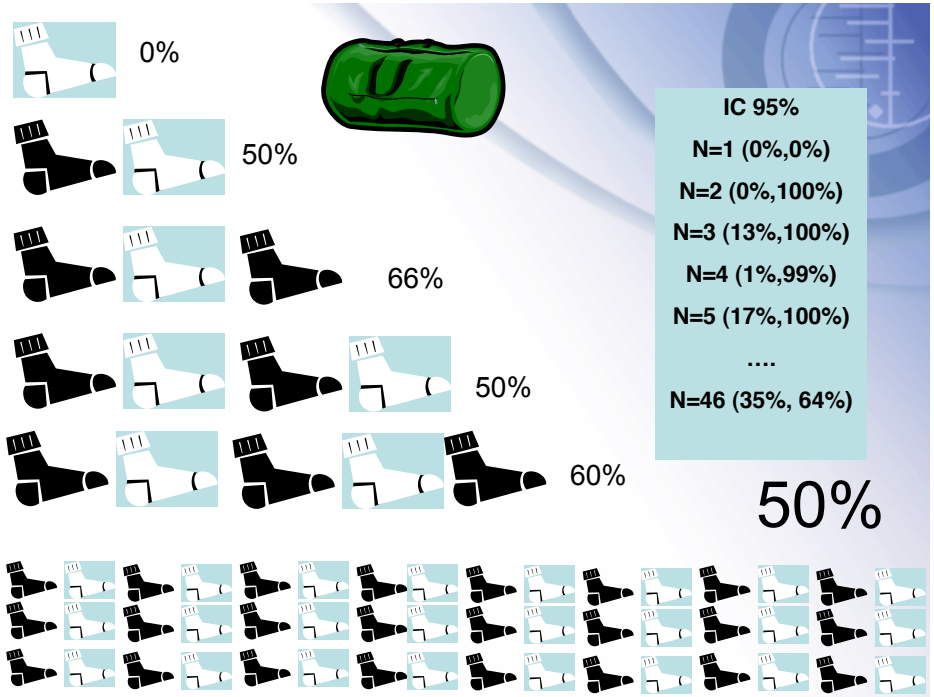
- El principio general que justifica trabajar con muestras es que resulta más barato, más rápido y más fácil que hacerlo con poblaciones completas



El número de pacientes necesario para contestar adecuadamente las preguntas

Suficiente para detectar las diferencias si existen realmente

Incrementar pacientes incrementa proporcionalmente el coste del estudio



El tamaño de la muestra depende de...

- Las características, objetivos y diseño del estudio
- La(s) variable(s) principal(es) y distribución de referencia
- La magnitud del efecto del tratamiento (δ) que queremos ser capaces de detectar
- La variabilidad de la medida
- Tipo de contraste de hipótesis (pruebas de hipótesis)
- Los errores de Tipo I y II y la potencia (α, β y 1-β), o nivel de confianza y potencia
- La tasa de retiradas del estudio y pérdidas de seguimiento

Características, objetivos y diseño

Estimación de parámetros poblacionales

- Prevalencia de diabetes tipo 2 en la población
- Promedio de las cifras de colesterol en la población infantil

Contraste de hipótesis

- Podría un programa de vacunación reducir la mortalidad infantil en más de un 15%
- Es diferente la eficacia analgésica de A en relación a B

Crédito: Iñaki Perez

- Describimos las características (parámetros) de la población a partir de las estimaciones generadas con la información de la muestra, con un nivel de incertidumbre pre-establecido.

Información de la muestra → Características de la población

¿Qué parámetros estimamos?

Variable categórica:
proporción, diferencia de proporciones

Variable continua:
media, diferencia de medias

Otros tipos de variables

Estimación de parámetros: proporción

$$n = \frac{k^2 p_1 (1 - p_1)}{w^2}$$

- K: Factor relacionado con la confianza
- p_1: previsión del resultado
- w: Imprecisión (error) admisible en la estimación

Crédito: Iñaki Perez

Estimación de parámetros: media

$$n = \frac{z_{\alpha/2}^2 \sigma^2}{w^2}$$

- K: $z_{\alpha/2}$: nivel de confianza
- σ : desviación típica
- w: imprecisión máxima aceptable
- ¡ la media esperada no cuenta !

- ¿Qué es la imprecisión?
- ¿Qué es el nivel de confianza?
- ¿Qué significa la variabilidad?

Valores k y niveles de confianza

K	Nivel de confianza
3.29	0.999
2.8	0.995
2.58	0.99
2.32	0.98
2.24	0.97
1.96	0.95
1.64	0.90
1.28	0.80
1.03	0.70
0.84	0.60
0.67	0.50
0.52	0.40
0.38	0.30
0.25	0.20

Ejemplo: proporción + imprecisión

¿Cuál es el número de niños en edad escolar que necesitamos para estimar la prevalencia de dislexia?

- Nivel de confianza 0.99 -> K=2.58
- p_1: Estimación inicial 10%
- w: Imprecisión admisible +-3%

N=664

La amplitud del intervalo de confianza será del 6%

p.e. p=14%

IC 99% (11%, 17%)

Ejemplo: proporción + confianza

¿Cuál es la incidencia de recidiva al primer año con tratamiento antiulceroso?

- Nivel de confianza 0.95 -> K=1.96
- Estimación inicial 15% de recidivas
- Imprecisión máxima admisible +-6%

N=136

P=15% IC 95% (9%, 21%)

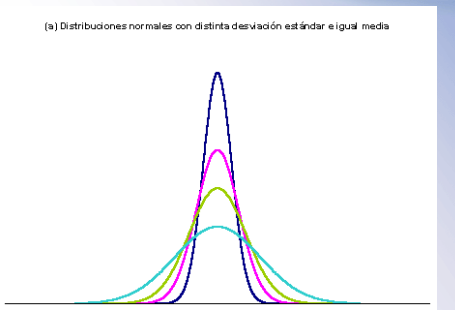
- Nivel de confianza 0.99

N=235; p=15% IC 99% (9%, 21%)

Ejemplo: media + variabilidad

¿Cuáles son los niveles de hemoglobina (g/l) en una determinada área de salud?

- Desviación típica es de 10 g/l (σ=10).
- Error de estimación +/- 1 g/l (precisión=1)
- Nivel de confianza del 95%
- Entonces, N=385
- Pero si σ=15, N=865
- Y si σ=5, N=96



Características, objetivos y diseño

Estimación de parámetros poblacionales

Prevalencia de diabetes tipo 2 en la población

Promedio de las cifras de colesterol en la población infantil

Contraste de hipótesis

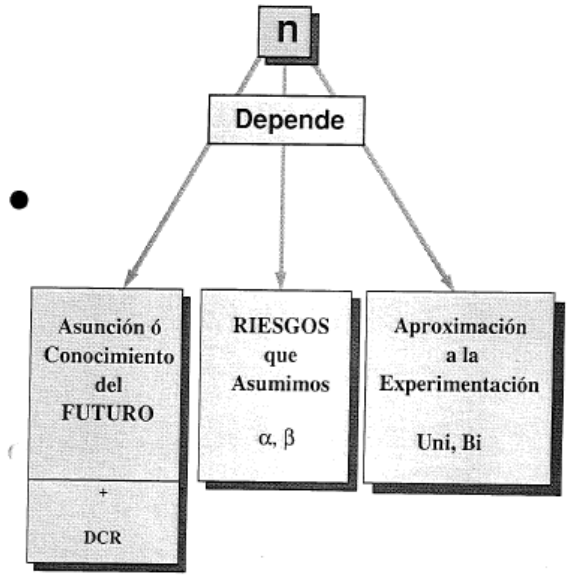
Podría un programa de vacunación reducir la mortalidad infantil en más de un 15%

Es diferente la eficacia analgésica de A en relación a B

Crédito: Iñaki Perez

El tamaño de la muestra depende de...

- Las características, objetivos y diseño del estudio
- La(s) variable(s) principal(es) y distribución de referencia
- La magnitud del efecto del tratamiento (δ) a detectar
- La variabilidad de la medida o imprecisión
- Contraste de hipótesis (pruebas de hipótesis)
- Los errores de Tipo I y II y el poder (α, β y 1-β)
- La tasa de retiradas del estudio y pérdidas de seguimiento



Crédito: Ignasi Gich

Contraste de hipótesis: comparación de dos proporciones

$$n_t = n_c = \frac{p_1(1-p_1)+p_2(1-p_2)}{\delta^2}(Z_{\alpha/2}+Z_\beta)^2$$

- $F(\alpha,\beta)$: $(Z_{\alpha/2} + Z_\beta)^2$: nivel de confianza
- P_1, p_2 : proporciones comparadas
- δ: p_1-p_2: magnitud del efecto a detectar

Contraste de hipótesis: comparación de dos medias

$$n_t = n_c = \frac{\left(Z_{\alpha/2}+Z_\beta\right)^2 2\sigma^2}{\delta^2}$$

- $F(\alpha,\beta)$: $(Z_{\alpha/2} + Z_\beta)^2$: nivel de confianza
  σ : desviación típica (variabilidad medida)
  δ : magnitud del efecto a detectar

Magnitud de las diferencias

- Diferencias clínicamente relevantes factibles de detectar
- Valor realista, basado en el conocimiento de la enfermedad y en nuestra disponibilidad de medios
- Diferencias moderadas 10%, 15%, 20%, 25% en términos relativos

Ejemplo: comparamos un antibiótico nuevo con otro standard, respecto a la tasa de curación a 2 semanas de infección urinaria

- El atb std tiene tasa curación de 80%. Se considerará superior el nuevo producto con una tasa de 85% o mejor.
- P_1: eficacia del atb std 80%
- p_2 : eficacia del atb nuevo 85%
- δ: p_1-p_2: 0.05
- Cambio relativo 6%

Ejemplo: comparamos fitoestrógenos con placebo en la reducción de sofocos menopáusicos

- Las mujeres que no responden al placebo tendrán una media de 30 sofocos semanales. Esperamos que los fitoestrógenos puedan reducir esta media en un 30% (cambio relativo).
- Cambio absoluto: de 30 a 21.

  δ : magnitud del efecto a detectar: 9 sofocos semanales

Ejemplo: Tiempo de disolución de 2 formulaciones de tabletas

- Se considera muy importante identificar si las formulaciones diferian en 10 minutos o más en el tiempo de alcanzar la disolución al 80%
- δ : magnitud del efecto a detectar: 10 minutos

Variabilidad

- Si los valores de base de la variable de interés difieren poco entre los pacientes (poca variabilidad), el tamaño necesario será menor
- La variabilidad real sólo se conoce con exactitud cuando se han obtenido los resultados

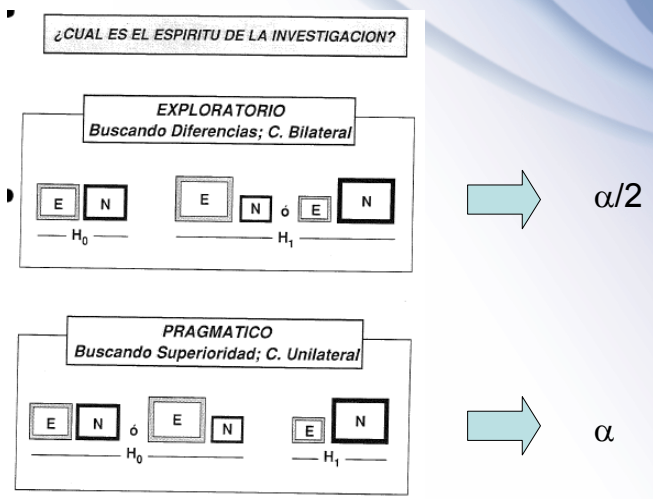
Variabilidad (II)

- En la fase de diseño, aproximamos:
  - a partir de datos obtenidos en ensayos previos
  - realizando una prueba piloto
- $\text{Variabilidad}_t = \text{Var}_1 + \text{Var}_2$
- Para una determinada magnitud de las diferencias, el número de pacientes puede reducirse disminuyendo la variabilidad.

Estrategias para reducir la variabilidad (y disminuir la n)

- En la selección del criterio de evaluación
  - Criterios cuantitativos
  - Variables precisas
  - Criterio de presentación frecuente
- Medición de las variables
  - Medidas repetidas
  - Técnicas de medición precisas
  - Estandarización de criterios, situacions y técnicas
- Población de estudio
  - Criterios de selección estrictos (pobl homogénea)
  - Población en la que sea + probable aparición respuesta
- Diseño del estudio
  - Asignación aleatoria
  - Medidas apareadas
  - Diseños especiales (cruzados, etc)
  - Estratificación por factores pronósticos

Prueba uni- o bilateral



Crédito: Ignasi Gich

Ejemplo:

Bilateral (dos colas)

H_o: $\theta_E - \theta_C = 0$
H_1: $\theta_E - \theta_C \neq 0$

Antibiótico A vs antibiótico B en infección urinaria

H_o: Tasa A = Tasa B
H_1: Tasa A ≠ Tasa B

Unilateral (una cola)

H_o: $\theta_E - \theta_C = 0$
H_1: $\theta_E - \theta_C > 0$ ó H_1: $\theta_E - \theta_C < 0$

Fármaco A + medidas preventivas reinfección vs Fármaco A en oxiuriasis

Ho: Tasa A = Tasa B
H1: Tasa A > Tasa B

Nivel de confianza y toma de decisiones

- Antibiótico A vs antibiótico B en infección urinaria

- Ho: Tasa A = Tasa B
- H1: Tasa A ≠ Tasa B

DECISION	REALIDAD: H_0 Verdadero	REALIDAD: H_0 Falso
H_0 No Rechazado	OK $(1-\alpha)$	Riesgo β F- Error tipo II
H_0 Rechazado H_1 Aceptado	Riesgo α F+ Error tipo I	OK $(1-\beta)$ Poder

$$n = f(\alpha, \beta)$$

Crédito: Ignasi Gich

Crédito: Iñaki Perez

Los errores de Tipo I, II y el poder

Bilateral (dos colas)

H_0: $\theta_E - \theta_C = 0$

H_1: $\theta_E - \theta_C \neq 0$

		Realidad: Ttos. Iguales	Realidad: Ttos. Diferentes
Conclusión	Ttos. Iguales	Acierto	Error tipo II (β)
	Ttos. Diferentes	Error tipo I (α)	Acierto (Potencia)

- Error tipo I (α): La probabilidad de rechazar la hipótesis nula siendo verdadera
- El valor del error tipo I ó α es igual o inferior a 0.05 (5%)
- Error tipo II (β): La probabilidad de aceptar la hipótesis nula siendo falsa
- El valor del error tipo II ó β es igual o inferior a 0.20 (20%)

Ejemplo:

- Antibiótico A vs antibiótico B en infección urinaria tienen tasa de respuesta 80% y 85%
- Error T1: Los 2 atb tienen la misma eficacia pero yo concluyo que B es mejor.
- Error T2: El atb B es mejor pero yo concluyo que es igual de eficaz que el atb A

Nivel de confianza: errores T1 y T2

- Para controlar el primer tipo de error se mantiene el nivel de significación estadística (α) a un nivel de probabilidad bajo
- El error de tipo 2 se controla definiendo la diferencia a detectar, y la probabilidad de detectarla que deseamos tener en el ensayo (potencia, 1-β)

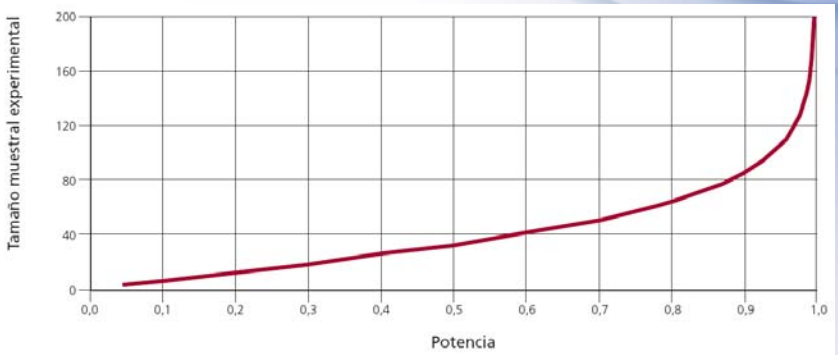
¿Qué α elegir?

α

0,01 <——————————————> 0,05

Unico	<—— Ensayo ——>	Múltiples
Elevado	<—— Riesgo terapeútico ——>	Bajo
Caro	<—— Método alternativo ——>	Barato
Conservadora	<—— Posición ——>	Arriesgada
Administración	<—— Promotor ——>	Interesado

¿Qué β elegir?



0.05 / 0.10 / 0.20

Nivel de confianza $\left(Z_{\alpha/2}+Z_{\beta}\right)^2$

TABLA 15.3. Valores de la función $f = (Z_\alpha + Z_\beta)^2$ para el cálculo del tamaño de la muestra

α (uni) α (bi)	0,05 0,10	0,025 0,05	0,01 0,02	0,005 0,01	0,001 0,005
β					
0,20	6,18	7,84	10,03	11,67	15,45
0,10	8,56	10,50	13,01	14,87	19,11
0,05	10,82	12,99	15,77	17,81	22,42
0,025	12,99	15,36	18,37	20,57	25,50
0,01	15,77	18,37	21,64	24,03	29,33
0,005	17,81	20,57	24,03	26,53	32,10
0,001	22,42	25,50	29,33	32,10	38,19

Para una n determinada, incrementar la alfa causa una disminución de la beta

Ejemplo: comparamos un antibiótico nuevo con otro standard, respecto a la tasa de curación a 2 semanas de infección urinaria

- El atb std tiene tasa curación de 80%. Se considerará superior el nuevo producto con una tasa de 85% o mejor. Es crucial poder demostrar esta superioridad. -> β=0.05
- α= 0.05
- Contraste bilateral
- P_1: eficacia del atb std 80%
- p_2 : eficacia del atb nuevo 85%
- δ: p_1-p_2: 0.05

N=1498 por grupo

Ejemplo: Tiempo de disolución de 2 formulaciones de tabletas

- Se considera muy importante identificar si las formulaciones diferian en 10 minutos o más en el tiempo de alcanzar la disolución al 80%
- $\alpha = 0.05$
- $\beta = 0.1$
- Contraste bilateral
- σ : desviación típica estimada e
- δ : magnitud del efecto a de nutos

¡ Depende del software !

N=10 por grupo ; N= grupo

Pérdidas durante el seguimiento

- Aumentamos el tamaño teórico según el porcentaje esperado de pérdidas.

$$n' = \frac{n}{1-d}$$

Ejemplo: antibiótico en infección urinaria

- $\beta = 0.05$
- $\alpha = 0.05$
- Contraste bilateral
- P_1: eficacia del atb std 80%
- p_2 : eficacia del atb nuevo 85%
- δ: p_1-p_2: 0.05
- Pérdidas: 5%

N=1498 por grupo -> N=1577 por grupo

Comparación de dos proporciones / Dos muestras independientes / No igualdad (3)

Tamaño de Muestra | Potencia | Gráfico

Parámetros

Parámetro	Valor
Nivel de Significación (α)	5,00
Unilateral / Bilateral (c)	2
Proporción Grupo de Referencia (p_1)	80,00
Proporción Grupo Experimental (p_2)	85,00
Proporción de la Muestra en el Grupo de Referencia (w_1)	50,00
Potencia (pow = 1-β)	95,00
Porcentaje de abandonos (ab)	5,00

Resultados

Resultado	Valor
Tamaño de Muestra Grupo de Referencia (n_1)	1498
Tamaño de Muestra Grupo Experimental (n_2)	1498
Tamaño de muestra a reclutar Grupo de Referencia (nr_1)	1577
Tamaño de muestra a reclutar Grupo Experimental (nr_2)	1577

Implicaciones de las pérdidas

- ¿Cuántas son respecto a los pacientes observados?
- ¿Se van a perder más en un grupo que en otro?
- ¿Por qué motivo se pierden?
- ¿El valor de la variable de desenlace es el mismo que en los observados?
- ¿Cómo los vamos a analizar?

Casos especiales

- Muestras finitas (muestra>5% población total)
- Más de 2 grupos
- Grupos de diferente tamaño
- Diseños especiales (cruzados)
- Equivalencia de tratamientos
- Pruebas no paramétricas
- Evaluación de supervivencia
- Corrección por comparaciones múltiples
- Estudios de cohortes -> incidencia de la enfermedad en pacientes no expuestos -> RR
- Estudios caso-control -> prevalencia de la exposición entre los controles -> OR

Software

- ENE 3.0
- http://www.e-biometria.com/
- GRANMO http://www.imim.es/ofertadeserveis/es_software p_blic.html
- EPIDAT
- http://www.sergas.es/MostrarContidos_N3_T02.aspx?IdPaxina=50114&uri=http://dxsp.sergas.es/ApliEdatos/Epidat/gal/default.asp

¡Gracias por vuestra atención!

Y además, gracias a

- Ignasi Gich (Hospital de la Santa Creu i Sant Pau, Barcelona) y Iñaki Perez (Hospital Clinic, Barcelona) por sus materiales
- Marrugat J, Vila J, Pavesi M, Sanz F. MedClin, 1998; 111(7):267-76
- Bolton S. Pharmaceutical Statistics. 1984