COMO DETERMINAR EL TAMAÑO DE LA MUESTRA EN ESTUDIOS CLÍNICOS

Martín Alonso Rondón Sepúlveda
MSc Bioestadística

Pontificia Universidad JAVERIANA Bogotá

| VIGILADA MINEDUCACIÓN |

Departamento de Epidemiología Clínica y Bioestadística

INTRODUCCIÓN

- Es quizá la pregunta más frecuente que se hacen los investigadores al momento de realizar un estudio

- Desafortunadamente, es una pregunta que no tiene una única respuesta, ya que su resultado depende de múltiples factores.

- Siempre se espera contar un tamaño de muestra óptimo, de tal manera que se puedan hacer inferencias a la población con la mayor certidumbre posible, ojalá tratando de minimizar el error en la estimación y asegurando un alto nivel de confianza.

INTRODUCCIÓN

- La pregunta entonces es, ¿cuál debe ser el tamaño mínimo necesario de unidades de observación que se deben seleccionar, de tal manera que permita a los investigadores encontrar resultados válidos, confiables y ojalá precisos?

INTRODUCCIÓN

- Grandes tamaños de la muestra, mejoran la precisión de la estimación, aunque no son sinónimos de resultados más confiables, así como tamaños de muestra pequeños no significan que el estudio no presenta validez.

- Una muestra se recolecta más rápido y es más económica

- Lo ideal es determinar para cada estudio específico, cuál debería ser el número mínimo de unidades de observación que se deben considerar.

OBJETIVO DE LA PRESENTACIÓN

- Hacer una introducción al tema del cálculo del tamaño de la muestra en investigación.

- Presentar conceptos básicos tales como: pregunta de investigación, variabilidad, tamaño del efecto, validez, precisión, confiabilidad, etc.

- Fórmulas y ejemplos para los casos más frecuentes.

Aspectos importantes

- Parámetro: Características de interés del total de la población (ejemplo: promedios poblacionales, proporciones poblacionales, etc.)

- Estadístico: Son datos o medidas que se obtienen sobre una muestra (ejemplo: promedios, medianas, proporciones)

- Estimador: Estadístico usado para estimar un parámetro poblacional

¿Cómo estudiar la característica de interés?

- Hay dos alternativas:
  - Se examina la característica de interés en toda la población: *censo.*
  - Se examina la característica de interés en un subconjunto de los elementos de la población: *muestra*.

Estimación de parámetros

¿Se puede creer en el resultado de este estimador?

- Representatividad
  - ¿La muestra es una representación de la población?
- Validez
- Confiabilidad

Validez

- La distribución de la característica en la muestra debe representar insesgadamente la distribución de la característica en la población blanco

  - Sólo mujeres

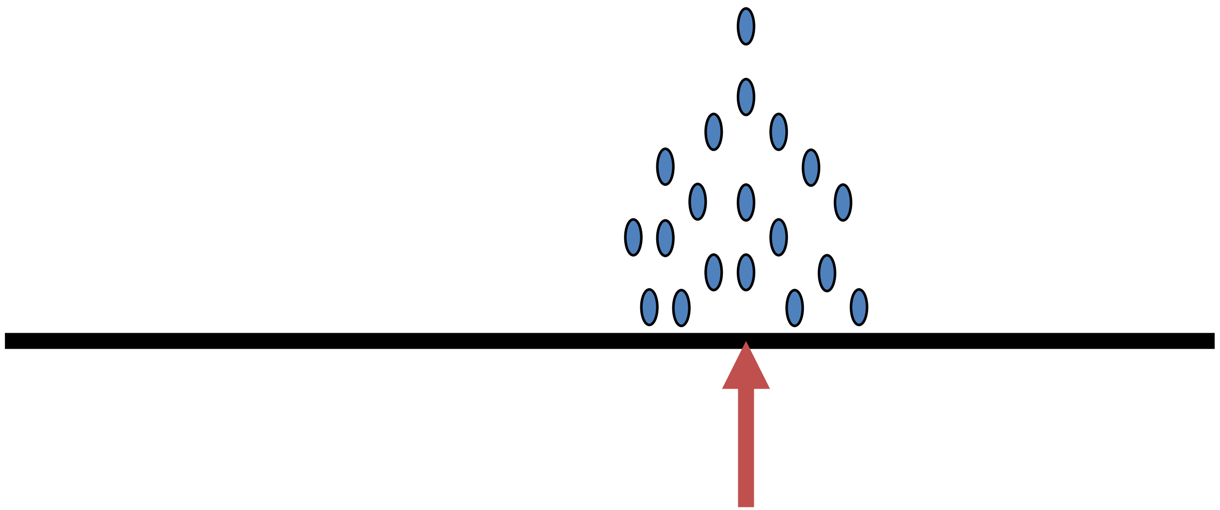
  - Sólo estudiantes de medicina

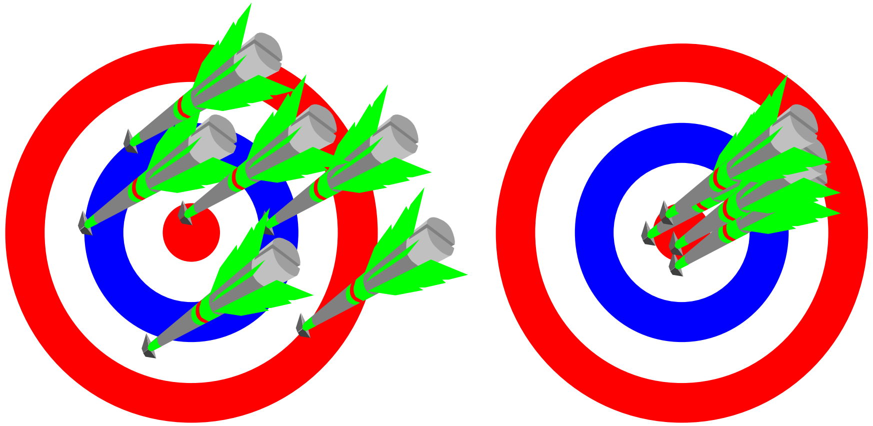
Precisión

- La precisión de la estimación hecha con una muestra aumenta a medida que se incrementa el número de elementos incluidos.

- Si se incluyen todos los elementos (censo) la muestra se hace igual a la población blanco.

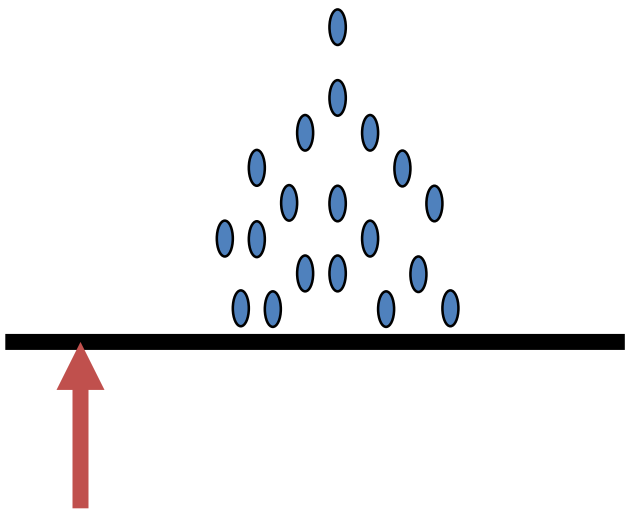
Error Aleatorio

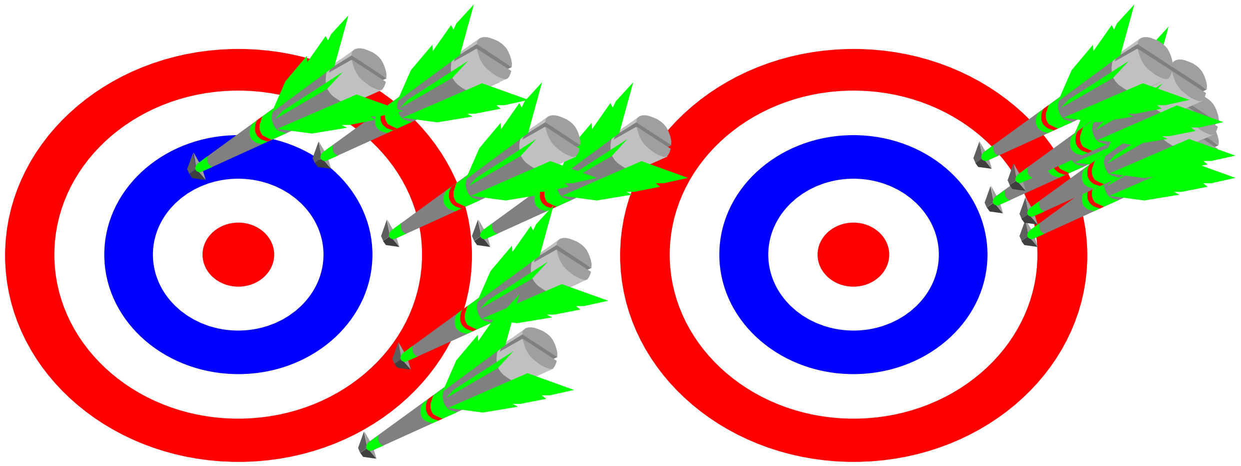




Las observaciones del estudio difieren de la verdad por efecto del azar

Error Sistemático (Sesgo)





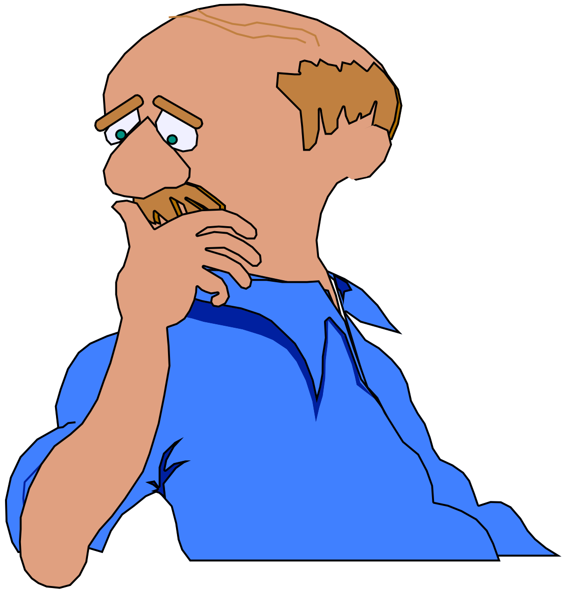
Las observaciones del estudio difieren de la verdad de manera sistemática

Tamaño de la muestra: Definición

¿Cuántos elementos necesito en mi muestra?

Es el número (*n*) de elementos o unidades de la población, que se seleccionarán para hacer parte de la muestra.

Tamaño de la muestra: *Métodos para determinarlo*



- Criterio propio
- Todos los que pueda recolectar
- Similar a otros estudios
- Cálculo basado en niveles de precisión, confianza y poder

Tamaño de la muestra:

Justificación

“Si al realizar un estudio no logramos demostrar diferencias, no significa que en la población no existan”

Importancia del tamaño de la muestra

Muestras demasiado pequeñas pueden no encontrar diferencias estadísticamente significativas, aunque estas diferencias existan (mayor error tipo II)

Muestras demasiado grandes pueden encontrar diferencias estadísticamente significativas, aún con diferencias numéricas mínimas (sin relevancia clínica)

Requisitos para calcular el tamaño de la muestra

- Pregunta de investigación, hipótesis y tipo de estudio
- Relación y número de grupos a medir y/o comparar
- Tipo de variables
- Diferencia que se considera importante o significativa
- Variabilidad de la(s) población(es) de estudio
- Confiabilidad que se espera del estudio

Pregunta de investigación

- La pregunta de investigación es el elemento esencial del cual parte el estudio de interés.

  - Para preguntas similares entre investigaciones
    - No necesariamente la metodología es la misma
    - El tamaño de la muestra depende de las características de la población

Planteamiento de las hipótesis estadísticas

- Generalmente las estimaciones del tamaño de la muestra se basan en las hipótesis estadísticas de interés

  - Se deben especificar los parámetros de interés a estimar.
  - Cuál es la direccionalidad de la hipótesis alterna
  - Se deben especificar el grupo o número de grupos que se seleccionarán.

Tipo de Estudio (Diseño)

- El diseño planeado esta directamente ligado a la pregunta de investigación, por lo tanto, el tamaño de la muestra puede diferir según la estrategia de diseño empleada.

- Existe una gran variedad de diseños de investigación: estudios transversales, estudios de casos y controles, estudios de cohorte prospectiva o retrospectiva, pruebas diagnosticas, de antes y después, de correlación, datos longitudinales, por mencionar solo algunos. Cada uno requiere una estimación diferente del tamaño de la muestra.

Relación y número de grupos a medir y/o comparar

- Para el calculo del tamaño de la muestra, es importante considerar si sólo se desea hacer la estimación para una población o se desean comparar dos o más grupos.

- A medida que el numero de grupos a comparar aumente, también lo hará el tamaño de la muestra, ya que cada grupo debe tener el mínimo de información, para poder hacer correctamente la comparación.

Requisitos para calcular el tamaño de la muestra

Error Tipo I

- Probabilidad de rechazar la hipótesis nula siendo que esta es verdadera
  - Generalmente valores de: 0.01, 0.05, 0.10
- *Error Tipo II*
  - Probabilidad de No rechazar la hipótesis nula siendo que esta es falsa
    - Generalmente valores entre: 0.10 – 0.20

Requisitos para calcular el tamaño de la muestra

Confianza

- Seguridad de obtener resultados similares si se repite el estudio
  - Generalmente 90%, 95% y 99%
- *Poder*
  - Seguridad de obtener resultados en el estudio que representen la realidad, es decir, capacidad del estudio para detectar diferencias significativas entre los grupos si estas existen
    - Generalmente 80% y 90%

Tipos de Diseños

- ESTUDIOS DESCRIPTIVOS
  - ESTUDIOS DE CORTE TRANSVERSAL
  - ESTUDIOS DE CORRELACIÓN

- ESTUDIOS ANALÍTICOS
  - ESTUDIOS OBSERVACIONALES
    - ESTUDIOS DE CASOS Y CONTROLES
    - ESTUDIO DE COHORTES
  - ESTUDIOS EXPERIMENTALES

Estudios descriptivos

- ESTUDIOS DE CORTE TRANSVERSAL

  - Promedios
    - Un grupo
    - Dos Grupos (Dependientes e Independientes)
    - Mas de Dos Grupos (Dependientes e Independientes)

  - Proporciones
    - Un grupo
    - Dos Grupos (Dependientes e Independientes)
    - Mas de Dos Grupos (Dependientes e Independientes)

Tamaño de la muestra para promedios un grupo

$$n = \frac{Z_{1-\alpha/2}^{2} S^{2}}{d^{2}}$$

Factores a considerar:

- d = Precisión esperada (distancia a la media poblacional)
- α = Nivel de confianza

Ejemplo

Se quiere probar la hipótesis de que el peso promedio en pacientes adultos hombres de una determinada población es de 75 ± 10 kg. Calcule el tamaño de la muestra necesario con un poder del 80% y confiabilidad del 95%, si se desea una precisión en la estimación del 3%.

Factores a considerar:

- d = Precisión esperada (distancia a la media poblacional)
- d = 0.03(75) = 2.15
- α =0.05
- $Z_{(1-\alpha/2)}$ = 1.96
- S = 10

$$n = \frac{(1.96)^2 (10)^2}{(2.15)^2}$$

$$n = \frac{3.84 * 100}{4.6225} = \frac{384}{4.6225} \approx 84$$

Comparación de las medias de dos grupos (muestras independientes y una diferencia específica).

La diferencia de las medias μ_1 y μ_2

- Se asume que las varianzas poblacionales son iguales (homocedásticas) pero desconocidas.

- Entonces, si se desea estimar el tamaño de la muestra para la diferencia se toma

$$d = \mu_1 - \mu_2$$

Comparación de las medias de dos grupos (muestras independientes y una diferencia específica).

$$n = \frac{\left(Z_{1-\alpha} + Z_{1-\beta}\right) 4(\sigma^2)}{d^2}$$

σ: Desviación estándar estimada. A priori.

$Z_{1-\alpha}$:valor crítico de la distribución normal estándar (una cola).

En caso de dos colas buscar $Z_{1-\frac{\alpha}{2}}$

d = μ_1 - μ_2, diferencia entre las medias de los dos grupos.

Ejemplo

- Se desea llevar a cabo un estudio para evaluar un complemento dietético para mujeres embarazadas bajo la hipótesis de que este suplemento aumentará el peso al nacer de los bebés en al menos 100 gramos, siendo que en promedio el peso al nacer de los bebes en esa población es de 3000 gramos. Un estudio piloto reportó una desviación estándar del peso al nacer de 500 gramos, la cual se asume igual en ambos grupos. La hipótesis será probada al 5% de nivel de significación y se desea tener un poder del 80% (Lemeshow S., Hosmer D., Klar J. y Lwanga S., 1990).

α =0.05.

β = 0,2

σ =500.

d =3100-3000=100

$$n = \frac{(1.6450+0.8416)^2(4*(500^2))}{100^2} \approx 619$$

Se necesitan 620 mujeres embarazadas en total, es decir, se requieren 310 mujeres embarazadas por grupo

Tamaño de la Muestra
Proporciones

- Proporción Puntual/prevalencia
- Diferencia de proporciones
- Más de dos proporciones

Estimación Puntual

- Precisión en términos de porcentaje alrededor de la proporción
- Precisión deseada en unidades absolutas
- Estimación puntual de prevalencia
- Hipótesis nula de una población específica

Estimación Puntual

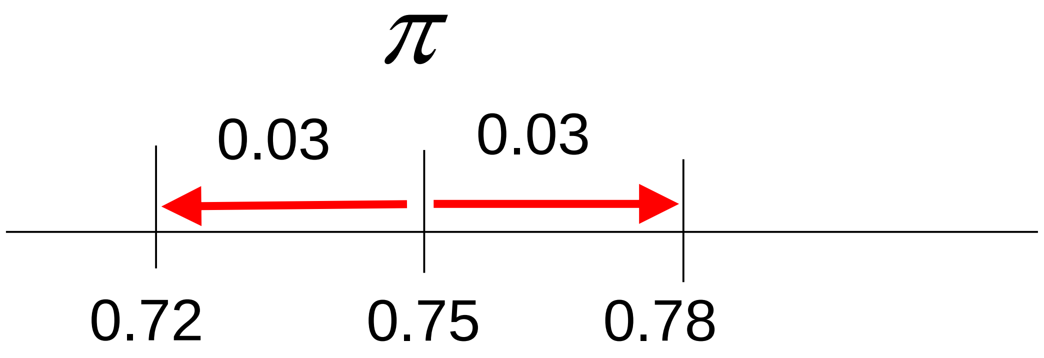
$$n = \frac{\left(Z_{1-\frac{\alpha}{2}}\right)^2 P(1-P)}{d^2}$$

Ejemplo:

- 20000 individuos atendidos en la EPS Q.
- Se desea determinar la proporción de individuos quienes recibieron atención adecuada.
- Se tomarán medidas remediales si el porcentaje es < 80%.
- Se cree que la verdadera proporción es del 75%

Estimación Puntual - ejemplo

3% de precisión de la verdadera proporción

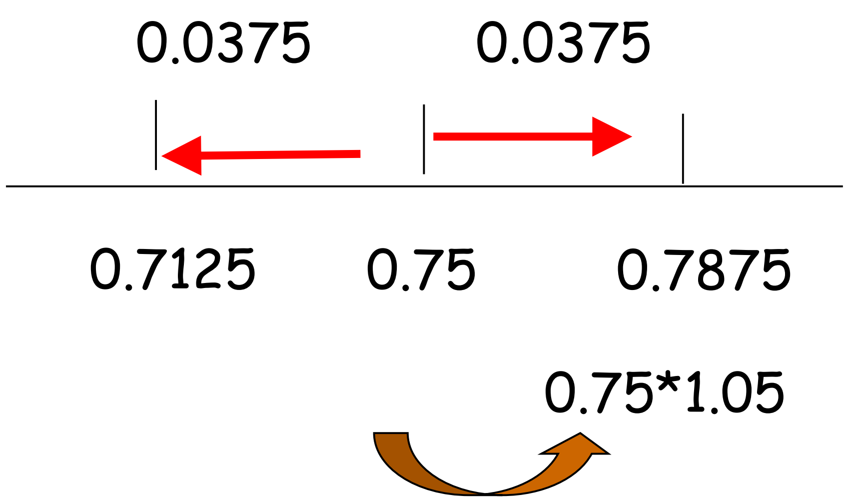


$$n = \frac{3.84(0.75)(0.25)}{0.03^2} \approx 800$$

Error tipo I=0.05
Dos colas: Z=1.96
p=0.75
1-p = 0.25
d=0.03

Precisión en términos de porcentaje alrededor de la proporción

π



5% alrededor de la proporción verdadera

$$n = \frac{Z^2_{1-\alpha/2}(1-P)}{\varepsilon^2 P}$$

Error tipo I=0.05

π=0.75

Porcentaje=5%

n=513

Estimación Puntual de la Prevalencia con poblaciones Finitas

- Se ajusta por el tamaño de poblacional

$$n = \frac{n_1}{1 + \frac{n_1}{N}}$$

- En el caso del ejemplo, asumiendo una población de 10000 personas.

$$n = \frac{n_1}{1 + \frac{n_1}{N}} = \frac{513}{1 + \frac{513}{10000}} \approx 488$$

Estimación Puntual – Fórmula Exacta
Muestreo Aleatorio Simple

$$n = \frac{\left(Z_{1-\frac{\alpha}{2}}\right)^2 (1-P)}{d^2(P)}$$

En el ejercicio anterior

$$n = \frac{1.96^2(0.25)}{(0.03^2)(0.75)} \approx 1.423$$

Comparación de 2 grupos Independientes

- Involucra error tipo I y II
- Bastantes formulas disponibles
- Todas son aproximaciones (datos discretos a continuos) vs Exacta

Comparación de 2 grupos Independientes Una formula

- Diferencias entre proporciones (%)

$$n = \frac{p_1(100 - p_1) + p_2(100 - p_2)}{(p_1 - p_2)^2} \quad f(\alpha, \beta)$$

p_1 y p_2 estimaciones del porcentaje de respuesta del grupo control y tratamiento respectivamente

f(a,b):Constante de cómputo de los niveles de a y b a usar. Ver tabla.

Comparación de 2 grupos Independientes - Tabla f(α,β)

	Nivel α			
Poder	**Una Cola**		**Dos Colas**	
(1- β)	**0.05**	**0.01**	**0.05**	**0.01**
0.5	2.71	5.41	3.84	6.63
0.7	4.71	8.13	6.17	9.61
0.8	6.18	10.04	7.85	11.68
0.9	8.56	13.02	10.51	14.88

Tabla modificada de Dennis, R. (1989) Cómo estimar el tamaño de la muestra en investigaciones con humanos. Acta Médica Colombiana. Vol 14(2), 92-99

Comparación de 2 grupos Independientes - Ejemplo

Se desea calcular un tamaño de muestra para estimar si la proporción de hombres que fuma es mayor a la proporción de mujeres que fuma (es decir a una cola), en una población de pacientes que asisten a una central de urgencias. Según los investigadores se espera que la proporción de hombres que fuma sea cercana al 25%, mientras la proporción de mujeres que fuma se espera que se encuentre alrededor del 18%. Use un nivel de confianza del 95% y un poder del 80%.

Comparación de 2 grupos Independientes Ejemplo - Una formula

- Diferencias entre proporciones (%)

$$n = \frac{0.25(0.75) + 0.18(0.82)}{(0.07)^2} \quad 6.18 \approx 423 \text{ personas por grupo}$$

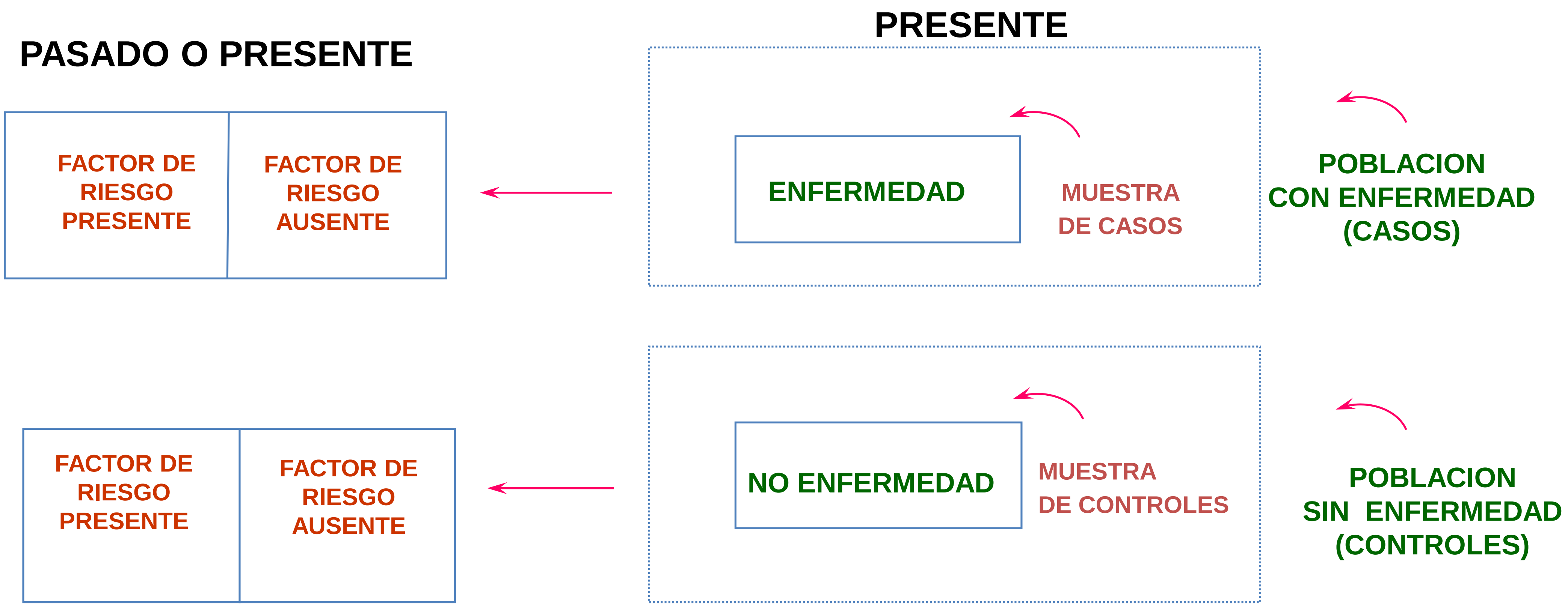
Casos y Controles

- Investigación epidemiológica
- Analítico-observacional
- Sujetos seleccionados con base en si/no tienen una enfermedad particular.
- Comparación : presencia del factor de riesgo

Casos y Controles: Requisitos

- Cuidadosa definición y selección de:
  - Casos (problema a estudiar)
  - Controles (Aproximación para estudiar y analizar el problema)
- Definición aceptada de la enfermedad

Diseño de Casos y Controles



Casos y Controles.
Número de Grupos de Control

- Un solo grupo control
- Varios grupos de control
    - Necesario medir resultados por cada grupo control
    - Excepcionales
    - Mayor trabajo
    - Análisis más sofisticado
    - Búsqueda exhaustiva de posibles diferencias entre los controles

Casos y Controles: Odds Ratio Crudo

Factor de Riesgo	Enfermedad	
	Presente (Casos)	Ausente (Controles)
Presente	a	b
Ausente	c	d
Total	a+c	b+d

OR=ad/bc
Estimador en variable respuesta “rara”

$$H_0 : OR = 1$$

$$H_a : OR \neq 1$$

OR=1 significa no asociación
OR>1 factor de riesgo
OR<1 factor protector

Casos y Controles: 2 grupos Independientes

$$n = \frac{1}{(\lambda_0 - \Omega)^2}\left\{Z_{1-\frac{\alpha}{2}}\sqrt{\left(1+\frac{1}{C}\right)\Pi(1-\Pi)} + Z_{1-\beta}\sqrt{\frac{\lambda_0(1-\lambda_0)}{C} + \Omega(1-\Omega)}\right\}^2$$

$$\Omega = \left(\frac{\lambda_0 + \delta}{1+\delta}\right) \; y \; \Pi = \frac{\lambda_0}{1+C}\left[C + \frac{\Omega}{\lambda_0}\right]$$

$$\mathrm{m} = cn \quad y \quad N = m + n = n(C+1)$$

$$\lambda_1 = \frac{(OR)\,\lambda_0}{1-\lambda_0(1-OR)}$$

Factores a considerar:

- λ_0 = proporción de la exposición entre los controles
- λ_1 = proporción de la exposición entre los casos
- Ω = diferencia entre las proporciones de expuestos y no expuestos
- Π = Proporción de la población general
- OR = Riesgo de presentar el evento de interés entre expuestos y no expuestos
- C = Razón entre número de expuestos y numero de no expuestos
- α = Nivel de confianza
- β = Poder del estudio.

Casos y controles: Ejemplo

- Si la prevalencia de cáncer en la población de no fumadores es del 25% y del 43% en el grupo de fumadores, calcule el tamaño de la muestra necesario con un poder del 80%, confiabilidad del 95% y una asignación igual entre los grupos (a una cola).

$\alpha = 0.05$

$\beta = 0.20$

$C = 1$

$\lambda_0 = 0.25$

$\lambda_1 = 0.43$

$Z_{1-\frac{\alpha}{2}} = 1.645$

$Z_{1-\beta} = 0.84$

$\delta = 0.43\text{-}0.25 = 0.18$

$$\Omega = \left(\frac{0.43}{1.18}\right) = 0.3644$$

$$\Pi = \frac{0.25}{2}\left[1 + \frac{0.3644}{0.25}\right] = 0.3072$$

$$n = \frac{1}{(0.25-0.3644)^2}\left\{1.645\sqrt{\left(1+\frac{1}{1}\right)(0.31)(0.69)} + 0.8416\sqrt{\frac{0.25(0.75)}{1} + 0.364(0.636)}\right\}^2 \Rightarrow$$

201 sujetos

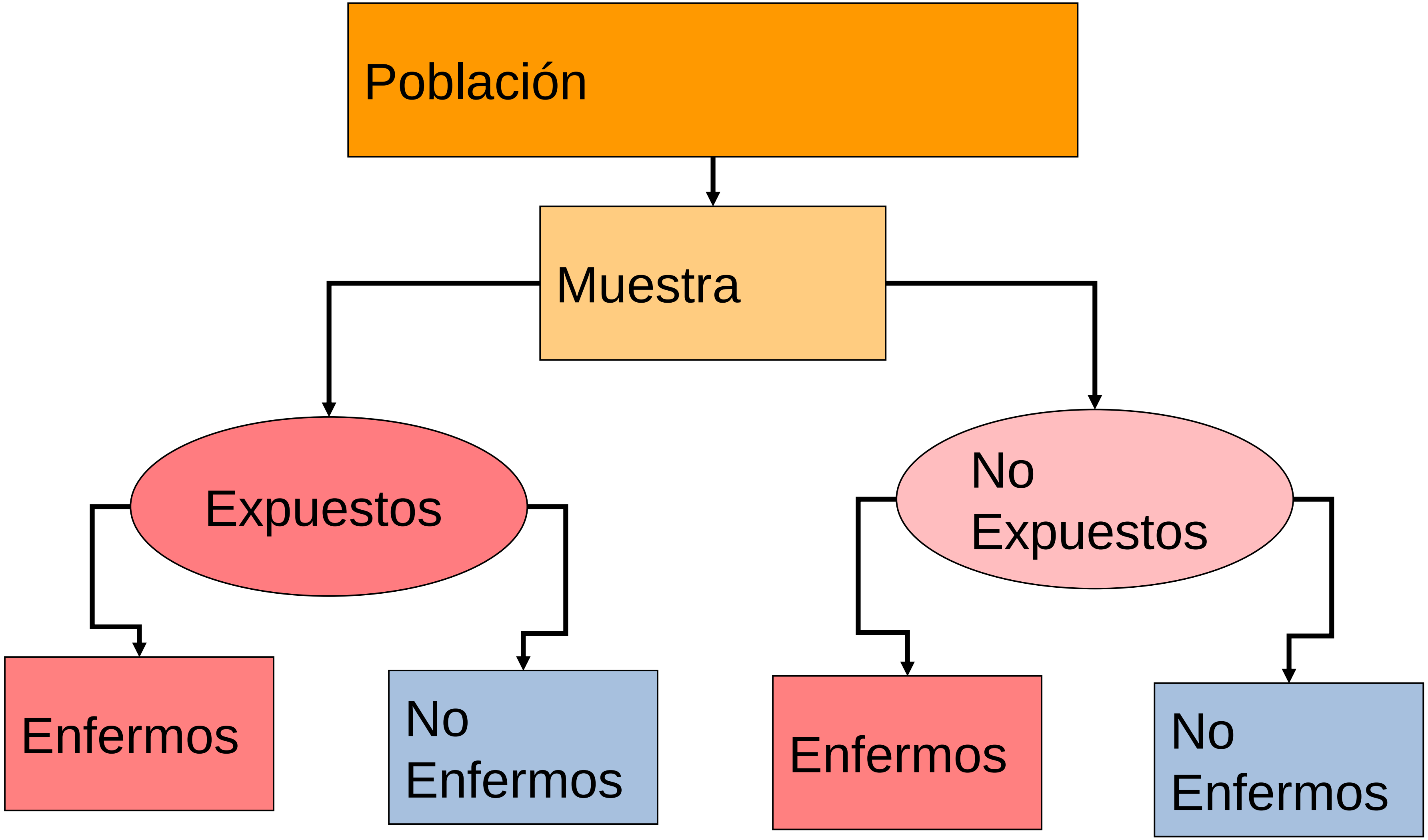
El tamaño de la muestra total es de 402 pacientes, es decir, se necesitan 201 casos y 201 controles

Tamaño de la Muestra en estudios de Cohortes

Cohorte: Definición

- Grupo bien definido de personas que han tenido una experiencia común (exposición), que son seguidos en el tiempo, para determinar la incidencia de una enfermedad

Estudio de Cohorte



Estudios de Cohorte: Ventajas

- Único estudio que permite estimar tasas de riesgo (RR).
- Permite obtener información sobre incidencia y prevalencia.
- Clara relación temporal entre exposición y enfermedad.
- Particularmente eficiente para estudios de exposición “rara”.
- Produce información sobre múltiples exposiciones.
- Produce información sobre múltiples respuestas de una misma exposición.
- Minimiza los sesgos.

Estudios de Cohorte: Desventajas

- Gran inversión de tiempo
- Frecuentemente requiere grandes tamaños de muestra.
- Costoso
- No eficiente para el estudio de enfermedades poco frecuentes.
- Perdidas en el seguimiento disminuye la validez del estudio.
- Cambios en los métodos de diagnóstico pueden sesgar los resultados.

Estudios de Cohorte: Tamaño de la Muestra

- Estimación de una proporción
- Estimación de dos proporciones
- Transformación logarítmica para cohortes
- IC para la estimación del riesgo relativo.
- Tiempo al evento

Estudios de Cohorte: Tamaño de la Muestra

- Estimación de una proporción
  - Prevalencias o porcentajes
  - En unidades absolutas.
  - Prueba de hipótesis
  - Intervalos de Confianza

- Estimación de dos proporciones
  - Diferencia de proporciones poblacionales bajo una hipótesis
  - Intervalos de confianza para la estimación del riesgo relativo

Estudios de Cohorte: Tamaño de la Muestra - Fórmula

$$n = \frac{\left\{Z_{1-\frac{\alpha}{2}}\sqrt{(C+1)\bar{\lambda}(1-\bar{\lambda})} + Z_{1-\beta}\sqrt{\lambda_0(1-\lambda_0) + C(\lambda_0+\delta)(1-\lambda_0-\delta)}\right\}^2}{C(\delta^2)}$$

$$\bar{\lambda} = \frac{[C\lambda_0 + (\lambda_0+\delta)]}{1+C}$$

$$m = cn \quad y \quad N = m + n = n(C+1)$$

$$\lambda_1 = RR^*\lambda_0$$

Factores a considerar:

- $\lambda_0 =$ proporción del evento de interés en los no expuestos
- $\lambda_1 = (\lambda_0+\delta)$ = proporción del evento de interés en los expuestos
- δ = diferencia entre las proporciones de expuestos y no expuestos
- $\bar{\lambda}$ = proporción del evento de interés en el total de la población
- RR = Riesgo de presentar el evento de interés entre expuestos y no expuestos
- C = Razón entre número de expuestos y numero de no expuestos
- α = Nivel de confianza
- β = Poder del estudio.

Estudios de Cohorte: Tamaño de la Muestra - Ejemplo

- Se cree que el riesgo relativo de desarrollar cáncer de pulmón en los fumadores (Expuestos) es de 1.8 comparado con los no fumadores (No Expuestos) para una población determinada. Si la proporción de personas que se espera desarrollen cáncer en la población de no fumadores es del 15%, calcule el tamaño de la muestra necesario con un poder del 80%, confiabilidad del 95% y una asignación igual entre los grupos (a una cola).

$\alpha = 0.05$
$\beta = 0.20$
$RR = 1.8$
$C = 1$
$\lambda_0 = 0.15$
$\lambda_1 = 1.8*0.15 = 0.27$
$\bar{\lambda} = 0.21$
$Z_{1-\alpha} = 1.645$
$Z_{1-\beta} = 0.84$
$\delta = 0.27\text{-}0.15=0.12$

$$n = \frac{\left\{1.645\sqrt{(1+1)(0.21)(0.79)} + 0.84\sqrt{0.15(0.85) + 1(0.27)(0.73)}\right\}^2}{1(0.12^2)}$$

$$n = \frac{\{0.9476 + 0.479\}^2}{0.0144} = \frac{1.43^2}{0.0144} = \frac{2.05}{0.0144} \Rightarrow \text{142 sujetos}$$

El tamaño de la muestra total es de 284 pacientes, es decir, se necesitan 142 expuestos y 142 no expuestos

Estudios de Cohorte: Tamaño de la Muestra - Ejemplo

- Como cambia el tamaño de la muestra si el RR fuera de 1.5?

$\lambda_0 = 0.15$

$\lambda_1 = 1.5*0.15 = 0.225$

$\bar{\lambda} = 0.1875$

$\delta = 0.225 - 0.15 = 0.075$

$$n = \frac{\left\{1.645\sqrt{(1+1)(0.1875)(0.8125)} + 0.84\sqrt{0.15(0.85) + 1(0.225)(0.775)}\right\}^2}{1(0.075^2)}$$

$$n = \frac{\{0.908 + 0.4615\}^2}{0.005625} = \frac{1.369^2}{0.005625} = \frac{1.876}{0.005625}$$

334 sujetos

El tamaño de la muestra total es de 668 pacientes, es decir, se necesitan 334 expuestos y 334 no expuestos

MEDICIÓN DEL RIESGO SEGÚN LA EXPOSICIÓN

		Desenlace	
		Si	No
Exposición	+	a	b
	-	c	d

Riesgo expuestos (Re) = a/(a+b)
Riesgo no expuestos (Rne) = c/(c+d)

RIESGO RELATIVO: RR = Re / Rne

MEDICIÓN DEL RIESGO SEGÚN LA EXPOSICIÓN

RR = 1 : no hay diferencias

RR > 1 : exposición aumenta el riesgo

RR < 1 : exposición reduce el riesgo

Comparación de 2 grupos Pareados

Prueba de McNemar

- Involucra error tipo I y II
- Estimador del Riesgo Relativo Indirecto
- Proporción de discordantes

$$n = \frac{\left\{ Z_{1-\alpha/2}(OR+1) + Z_{1-\beta}\sqrt{(OR+1)^2 - (OR-1)^2 \pi_{Disc}} \right\}^2}{(OR-1)^2 \pi_{Disc}}$$

Comparación de 2 grupos pareados

Ejemplo: Comportamiento sexual seguro en mujeres brasileras. Se desea medir la concordancia entre el comportamiento sexual seguro antes y después de conocer el resultado de SIDA en sus compañeros sexuales.

Guimaraes M, et al.(2001) International Journal of Std & Aids

Ejemplo

	Después		
Antes	Seguro	No seguro	Total
Seguro	197	26	223
No Seguro	97	8	105
Total	294	34	328

$$\hat{\pi}_{dis} = (26+97)/328 = 0.375$$

$$\hat{OR} = 26/97 = 0.27$$

Ejemplo

- Repetir el estudio en Colombia, nivel de error tipo I= 0.05, tipo II=0.20.
- Proporción discordantes antes 6%
- Proporción discordantes después 24%
- Discordantes
- Riesgo relativo indirecto estimado

$$\hat{\pi}_{dis} = 0.3$$

$$OR = 4$$

$$n = \frac{\left\{1.96(4+1) + 0.8416\sqrt{(4+1)^2 - (4-1)^2(0.3)}\right\}^2}{(4-1)^2(0.3)}$$

$$n = \frac{\{9.78 + 3.97\}^2}{2.7} = \frac{189.73}{2.7} \approx 71 pares$$

Tamaño de la Muestra para el coeficiente de Correlación

Coeficiente de Correlación de:

- **Intraclase**
  - medir la confiabilidad a través de la medición de acuerdo y desacuerdo en diferentes mediciones
  - Se basa en un análisis de varianza sobre medidas repetidas

- **Pearson**
  - Mide la magnitud y la dirección de la **relación lineal** entre dos variables, para evaluar el grado de relación entre ellas

Coeficiente de Correlación Intraclase: Interpretación

- Proporción de la variabilidad total que es debida a la variabilidad de las observaciones.

$CCI < 0.4$ Confiabilidad baja
$0.4 \leq CCI < 0.75$ Confiabilidad regular
$CCI \geq 0.75$ Confiabilidad Excelente

Coeficiente de Correlación Intraclase: Tamaño de la muestra

- Fórmula basada en la amplitud del intervalo de confianza .
- Fórmula para medir la reproducibilidad del CCI (estimación puntual) .

Amplitud del Intervalo de Confianza: Tamaño de la Muestra

- Objetivo: estimación del CCI cuando se realiza un diseño test-retest, dada una amplitud deseada del intervalo de confianza.

$$n = \left(\frac{z_{1-\alpha/2}}{\left(0.5 \ln \frac{1+r}{1-r}\right) - \left(0.5 \ln \frac{1+r-ICC_h}{1-r+ICC_h}\right)} \right)^2 + 3$$

Amplitud del Intervalo de Confianza: Ejemplo

Se desea calcular el tamaño de la muestra requerido para determinar la magnitud del CCI con un nivel de significación de 0.05, a partir de una estimación muestral de 0.70 y una amplitud de ± 5%.

$\alpha = 0.05$ ICh= 0.05 r= 0.7

$$n = \left(\frac{1.96}{\left(0.5 \ln \frac{1+0.7}{1-0.7}\right) - \left(0.5 \ln \frac{1+0.7-0.05}{1-0.7+0.05}\right)} \right)^2 + 3$$

$$n = \left(\frac{1.96}{0.8673 - 0.7753} \right)^2 = (21.30)^2 + 3$$

$n \geq 457$

Coeficiente de Correlación de Pearson: Tamaño de la Muestra

- Comparación de valores del coeficiente de correlación (dos poblaciones).
- Prueba de hipótesis para el coeficiente de correlación (una población).
- Intervalo de confianza para estimar el coeficiente de correlación.

Prueba de hipótesis para ρ
(una población)

Se desea probar la hipótesis $H_0 : \rho = r_c$

$$n = 2 \times \left(\frac{Z_{1-\alpha/2} + Z_{1-\beta}}{q_c} \right)^2 + 3$$

$$q_c = \frac{1}{2}\log\left[\frac{1+\rho}{1-\rho}\right] - \frac{1}{2}\log\left[\frac{1+r}{1-r}\right]$$

Prueba de hipótesis para ρ
Ejemplo

Encontrar el tamaño de la muestra necesario para probar las hipótesis $H_0 : \rho = 0.60$

Los valores de error tipo I y tipo II aceptables son 0.05 y 0.20, respectivamente.
Correlación muestral r = 0.80

$$q_c = \frac{1}{2}\ln\left[\frac{1+0.6}{1-0.6}\right] - \frac{1}{2}\ln\left[\frac{1+0.8}{1-0.8}\right] = 0.6931 - 1.099 = -0.4055$$

$$n = 2 \times \left(\frac{1.96+0.8416}{-0.4055}\right)^2 + 3$$

99 sujetos por grupo

Programas para calculo de muestra

- Stata
- Epidat
- Tamaño de la muestra
- MedCalc
- PASS: Sample Size and Power
- PS: Power and Sample Size Calculation
- R
- SAS
- Y Muchos más

- Gracias!