

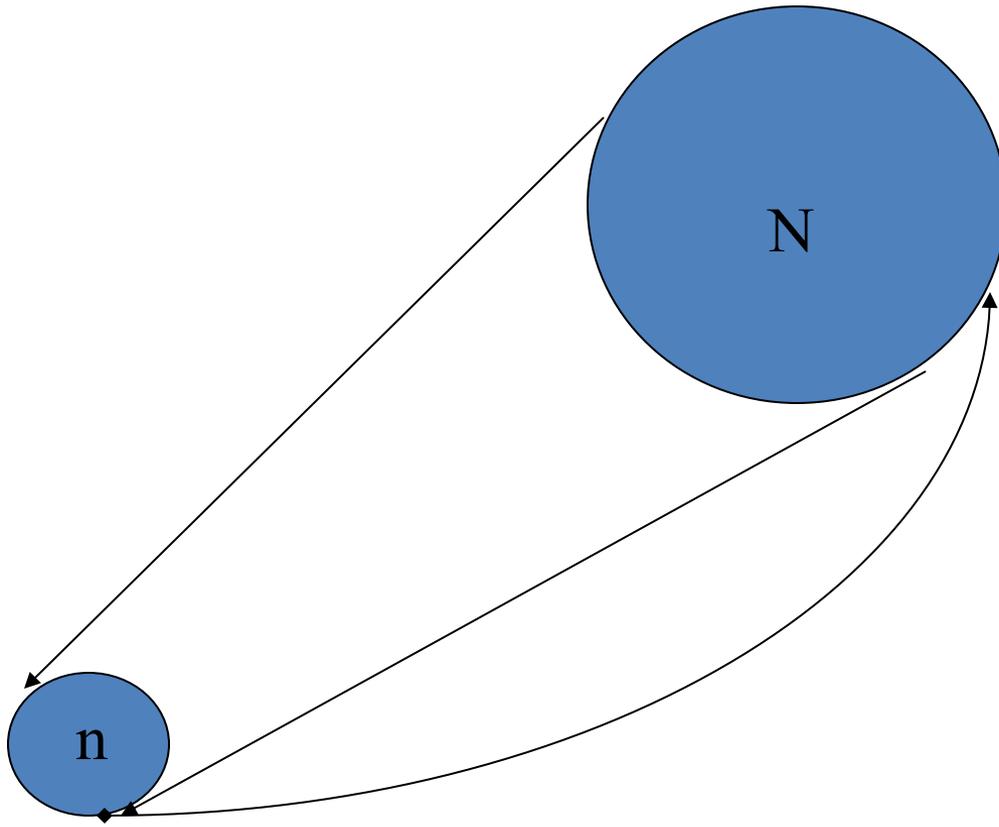
# Intervalos de confianza para medias y proporciones

Preparado por  
Juan José García García

# Introducción

- En el campo del análisis estadístico, uno de los aspectos fundamentales que se desea reportar de los datos estudiados, o que se encuentra en los resultados que ofrecen los autores de una investigación, es generalizar los hallazgos observados en una muestra a la población de la cual fue seleccionada, proceso al que llamaremos inferencia o estimación.

# Inferencia



# Requisito

- El elemento central para poder llevar a cabo esta estimación es el hecho de que la muestra estudiada haya sido seleccionada por métodos probabilísticos:
  - muestreo aleatorio simple,
  - muestreo sistemático
  - muestreo estratificado,
  - muestreo por conglomerados.

- Con lo anterior, se espera que las diferentes unidades de muestreo hayan tenido la misma oportunidad de ser incluidas, y que ha sido el azar el encargado de la selección.
- Queda con ello establecido que cuando se utilizan métodos no probabilísticos no puede llevarse a cabo la generalización de resultados.

- No obstante, el proceso de selección de las muestras posibles a partir de una población se enfrenta a la variabilidad entre las mismas, al que denominaremos error de muestreo, o error estándar.

# Intervalo de confianza: Concepto

- Es un rango que se encuentra delimitado por dos valores, inferior y superior, dentro del cual, con un cierto grado de probabilidad establecido (arbitraria, pero generalmente de 95 %), se espera que se encuentre el parámetro o valor poblacional de la variable o indicador estudiado en la muestra (estimador puntual).

- Lo anterior implica que, si se repitiera el estudio 100 veces, en 95 de ellas el parámetro de interés quedaría incluido entre los límites establecidos por el intervalo calculado.

- A mayor tamaño muestral, se tendrá mayor precisión, por lo que la amplitud del intervalo de confianza será más pequeña.
- Entre más estrecho sea este intervalo, en términos generales, puede decirse que resulta más útil para apoyar la toma de decisiones.

- Siempre existe la probabilidad de que el parámetro analizado se encuentre en realidad fuera del rango calculado, pero este hecho no puede conocerse.

# Cálculo del intervalo

- Podemos identificar que la forma en que se obtiene el intervalo de confianza tiene una estructura similar independientemente del estimador estudiado:
- Valor del estimador puntual  $\pm$  coeficiente de confiabilidad  $\times$  error estándar

- El estimador puntual es el valor de interés obtenido directamente en la muestra: una proporción, un promedio, una medida de asociación.
- El coeficiente de confiabilidad, suponiendo una distribución normal, está basado en un valor  $z$ , que, para una confiabilidad de 95 %, se convierte, prácticamente en una constante, 1.96. (2.58, para una confiabilidad de 99 %).
- Y el error estándar es obtenido según el caso.

- A mayor tamaño muestral, el error estándar es más pequeño, y viceversa, si la muestra es pequeña, el error estándar es más grande.
- Este valor, al multiplicarse por el coeficiente de confiabilidad genera la amplitud que tendrá la estimación del parámetro, es decir, su precisión.

- Su cálculo difiere según el estimador buscado, es decir, se obtiene de manera distinta para el estudio de un promedio que el de una proporción, aunque su construcción comparte elementos comunes:

$$\sqrt{\frac{\text{Variabilidad}}{\text{Tamaño de la muestra}}}$$

- En el caso de la estimación de un promedio, la medida de variabilidad que se emplea es la desviación estándar ( $s$ ) o la varianza,  $s^2$
- Así entonces, el error estándar =

$$\frac{s}{\sqrt{n}} \dots\dots \text{ó} \dots\dots \sqrt{\frac{s^2}{n}}$$

$$\sqrt{\frac{pq}{n}}$$

- En la estimación de una proporción, la variabilidad está dada por el producto  $pq$ , donde  $p$  es la proporción de la muestra con la característica de interés (enfermedad o factor de riesgo, por ejemplo), y  $q$  es el complemento, es decir,  $1 - p$ .
- La suma de  $p$  y  $q$  es igual a la unidad.
- Y el error estándar es =  $\sqrt{\frac{pq}{n}}$

- Debe notarse que la magnitud del error estándar se desprende de los valores de sus componentes, de tal forma que la variabilidad es un aspecto inherente al fenómeno estudiado y no es manipulable por el investigador, pero el tamaño muestral sí es sujeto de modificaciones.

# Ejercicio 1

- Calcule el intervalo de confianza (95 %) para un promedio, suponiendo:
- $n = 500$
- Promedio de índice de masa corporal = 23.5
- Desviación estándar = 1.5
  
- Interprete.
- Qué ocurre con el intervalo, si el tamaño muestral se reduce a 50?
- Y si es de 5000?

- Las respuestas se muestran con los resultados obtenidos a través del programa Epidat 3.1

# Inferencia sobre una media

- Media: 23.5
- Desviación estándar: 1.5
- Tamaño de muestra: 500
- Nivel de confianza: 95.0%

- Media                   IC (95.0%)
- -----
- 23.500           23.368           23.632

- Con un 95 % de confianza puede afirmarse que el promedio del IMC en la población total se encuentra entre 23.4 y 23.6.

# IC con $n = 50$

- Media: 23.500
- Desviación estándar: 1.500
- Tamaño de muestra: 50
- Nivel de confianza: 95.0%

•	Media	IC (95.0%)	
•	-----	-----	
•	23.500	23.074	23.926

# IC con $n = 5000$

- Media: 23.500
- Desviación estándar: 1.500
- Tamaño de muestra: 5000
- Nivel de confianza: 95.0%

	Media	IC (95.0%)	
•	-----	-----	-----
•	23.500	23.458	23.542

# Ejercicio 2

- Calcule un intervalo de confianza para una proporción, suponiendo:
- $n = 300$
- $p = 0.35$

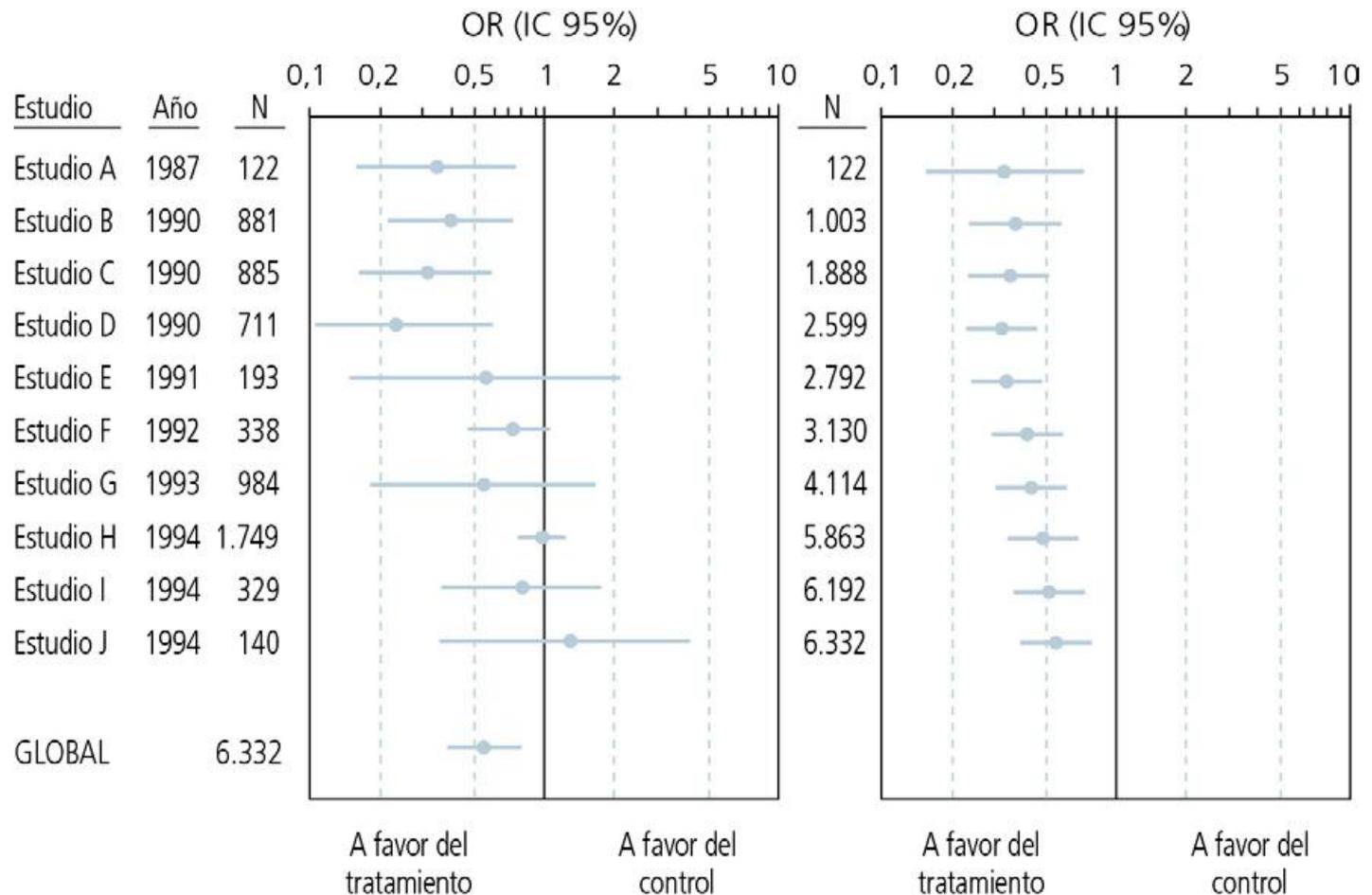
# Inferencia sobre una proporción

- Número de casos: 105
- Tamaño de muestra: 300
- Nivel de confianza: 95.0%

Proporción (%)	IC (95.0%)
-----	-----
35.000	29.436 - 40.564

- Con un 95 % de confianza puede afirmarse que la prevalencia de la enfermedad en la población se encuentra entre 29.4 y 40.6 %.

# IC en un metaanálisis



A la derecha se observan resultados en forma acumulativa