

Muestreo en estudios observacionales

César Gutiérrez Villafuerte

Departamento Académico de Medicina Preventiva y Salud Pública

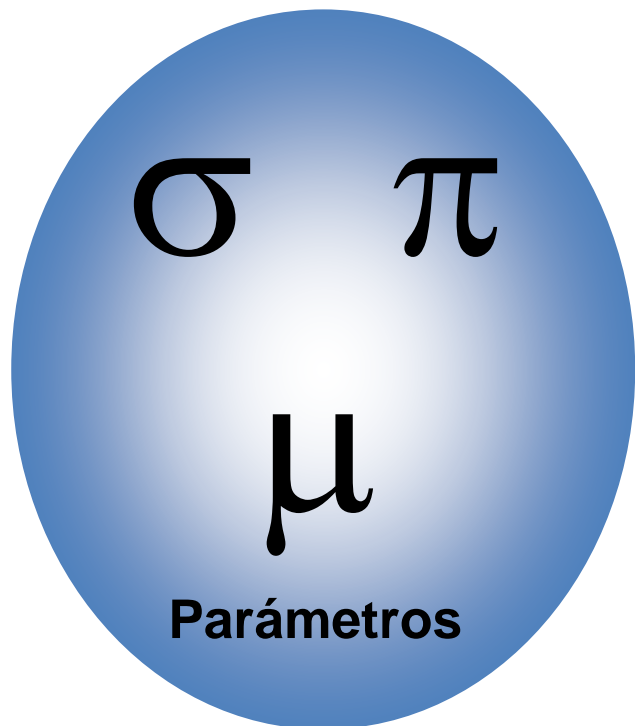
Facultad de Medicina – UNMSM

cgutierrezv@unmsm.edu.pe

Concepto de muestreo

Se denomina muestreo al procedimiento mediante el cual se selecciona un conjunto de elementos de la población, para, a partir de este conjunto de elementos, inferir el valor de una o varias características de la población.

Población (N)



Muestreo

Muestra (n)



Inferencia

Población

Conjunto de elementos (objetos, situaciones o sujetos) que comparten características similares definidas por el investigador.

Una población no siempre se refiere a seres humanos (p.e. historias clínicas en un determinado hospital, muestras de sangre almacenadas en un laboratorio, viviendas en una comunidad, etc.).

Unidad de análisis, muestreo y marco muestral

Unidad de análisis: elementos de la población objetivo de nuestra investigación.

Unidad de muestreo: elementos que nos permiten llegar a las unidades de análisis.

Marco muestral: listado de cada una de las unidades de muestreo.

No siempre la unidad de análisis se corresponde con la unidad de muestreo.

No siempre es posible contar con un marco muestral.

Marco Muestral

Dado que el diseño para la Encuesta ENDES Continua 2012 - 2014 es bietápica, se utiliza un marco muestral en cada una de las etapas de la selección de las unidades de muestreo.

Para la primera etapa: selección de UPM (conglomerados), se utiliza la información del Censo de Población y Vivienda del 2007. Para la segunda etapa: selección de USM (viviendas), se utiliza el marco muestral proveniente de la actualización cartográfica y registro de edificios y viviendas realizada previamente a las entrevistas, con el objetivo de identificar y registrar cambios en las áreas seleccionadas. Producto de este proceso se tiene un registro actualizado de viviendas que en ese momento existían en el área, el cual constituirá el marco de muestreo para la selección de viviendas.

Apéndice A. Diseño y cobertura de la muestra. ENDES 2013.

<http://proyectos.inei.gob.pe/web/biblioineipub/bancopub/Est/Lib1075/index.html>

Selección de muestras: consideraciones

- Forma de selección
- Tamaño de la muestra

Van a depender de los objetivos, diseño de la investigación y características de la población en estudio.

Tipos de diseños de selección de muestras

- Muestreo Probabilístico
 - Aleatorio Simple
 - Sistemático
 - Estratificado
 - Conglomerados
- Muestreo No Probabilístico
 - Por conveniencia
 - Por casos consecutivos
 - Por cuota
 - Bola de nieve

Cálculo del tamaño de la muestra

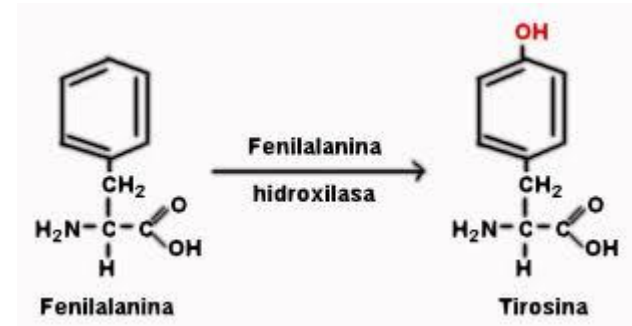
- Nivel de confianza
- Potencia (cuando se comparan grupos)
- Varianza de la variable
- Margen de Error [o Precisión]

$$n = \frac{(\text{Nivel de Confianza} + \text{Potencia}) \times \text{Varianza}}{\text{Margen de Error [o Precisión]}}$$



Si usted va de compras al mercado y quiere comprar uvas, ¿cuántas uvas toma para probar?

Si desea determinar la frecuencia de fenilcetonuria en recién nacidos, ¿necesitará una muestra muy grande o no muy grande?



Si desea determinar la frecuencia de bajo peso para la edad gestacional en recién nacidos, ¿necesitará una muestra muy grande o no muy grande?



Cálculo del tamaño de muestras para la estimación de parámetros

Estimación de una proporción

$$n = \frac{Z^2 pq}{d^2}$$

Z = Valor Z para el nivel de confianza establecido.

p = proporción del evento de interés.

q = 1 - p

d = error absoluto.

Estimación de una media

$$n = \frac{Z^2 \sigma^2}{d^2}$$

Z = Valor Z para el nivel de confianza establecido.

σ^2 = varianza de la variable de interés.

d = error absoluto.

Cálculo del tamaño de muestras para estudios de cohortes

$$n = \frac{r + 1}{r(\lambda - 1)^2 \pi^2} \left[z_\alpha \sqrt{(r + 1)p_c(1 - p_c)} + z_\beta \sqrt{\lambda\pi(1 - \lambda\pi) + r\pi(1 - \pi)} \right]^2$$

r = relación número de expuestos / número de no expuestos

λ = riesgo relativo a detectar

z_α = Nivel de confianza

z_β = Potencia

p_c = proporción (incidencia) común entre ambos grupos

π = proporción (incidencia) en el grupo de referencia (no expuestos)

Cálculo del tamaño de muestras para estudios casos y controles

$$n = \frac{(r + 1) (1 + (\lambda - 1)P)^2}{rP^2 (P - 1)^2 (\lambda - 1)^2} \left[z_{\alpha} \sqrt{(r + 1)p_c(1 - p_c)} \right]$$

r = relación número casos / número de controles

λ = odds ratio a detectar

z_{α} = Nivel de confianza

z_{β} = Potencia

p_c = proporción (de la exposición) común entre ambos grupos

P = proporción (de la exposición) en el grupo de referencia (control)

Fuentes de sesgo en el muestreo

- No respuesta.
- Estudio solamente de voluntarios.
- Muestreo sólo de los registrados.
- Pérdida de casos de corta duración.
- Sesgo estacional.
- Accesibilidad de áreas seleccionadas.

¿Preguntas?
Muchas gracias por su atención

César Gutiérrez Villafuerte

Departamento Académico de Medicina Preventiva y Salud Pública

Facultad de Medicina – UNMSM

cgutierrezv@unmsm.edu.pe