



El plan de muestreo

Oscar Federico Nave Herrera
Coordinador del Programa de
Asesoría Estadística para
Investigación -Digi-

¿Y ahora qué?

¿Cuántas muestras debo tomar?

¿Cuántas veces debo realizar los ensayos?

¿Tengo suficientes recursos y materiales?

¿De cuánto tiempo dispongo?

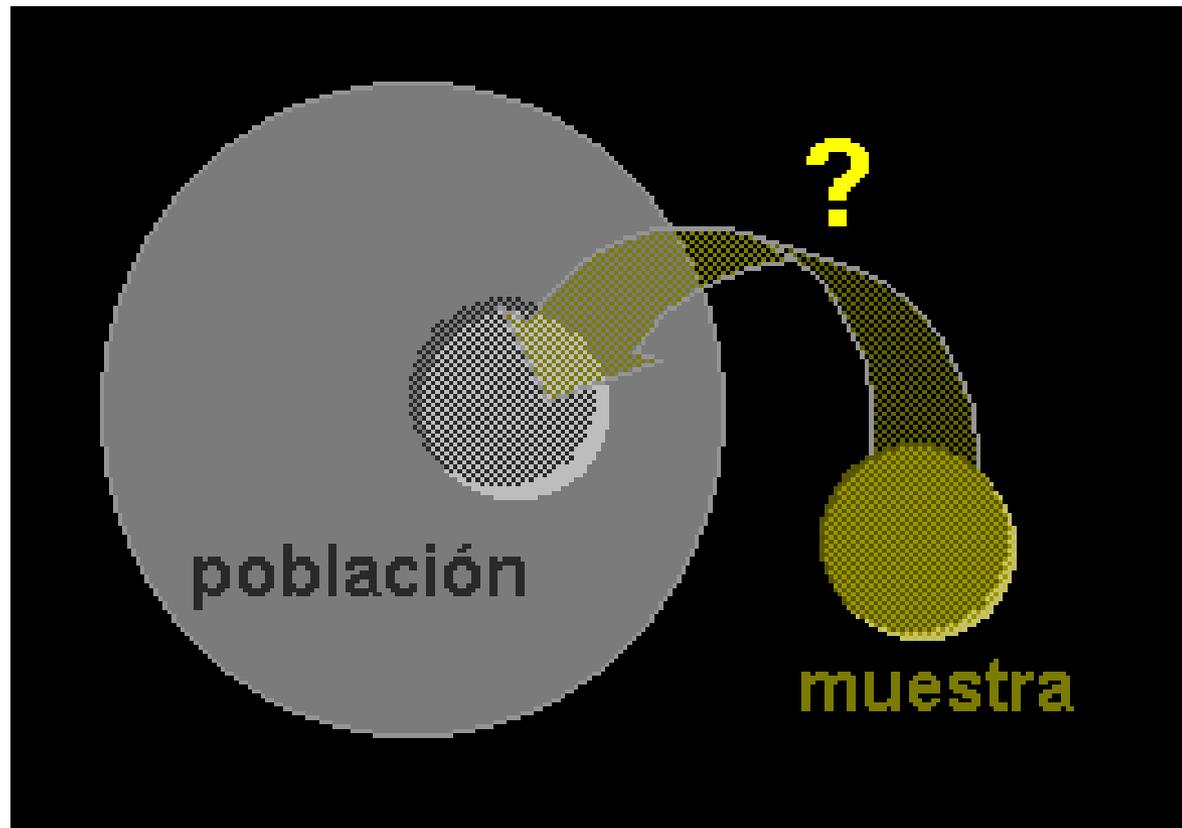
¿Podría tener otras limitaciones?

El plan de muestreo

Lo constituyen los elementos a considerar para extraer una muestra de una población.

Se debe definir el número de elementos a tomar y la forma en la cual se obtendrán las unidades muestrales.

La muestra



Ventajas de muestrear

- Mayor rapidez para investigar una variable
- Reducción de costos y recursos
- El estudio de una población puede ser muchas veces imposible

Muestra

- Representativa
- Estadísticamente representativa (incorrecto)

*Muestras por conveniencia

**Réplicas

Réplicas

Repetición del experimento. Es el número de ocasiones en que se efectúa una misma condición experimental.

Pseudorreplicación:

- Medidas repetidas de la misma unidad muestral o experimental.
- Sub-grupos de una misma unidad muestral o experimental.

¿Muestras o réplicas?

- Si la inferencia se pretende hacer sobre las poblaciones (estimación o pruebas de hipótesis), se requieren **muestras** de estas.
- Si se trata de un estudio experimental, en el que las unidades no representan una población, sino que se integrarán en grupos sobre los cuales se harán las inferencias (pruebas de hipótesis), se requieren **réplicas** (repeticiones).

Aspectos a tomar en cuenta

- Nivel de la investigación
- Distribución de los individuos en la población
- Fuentes de información
- Diseño muestral a aplicar
- Proporción de respuestas esperadas y de no respuesta
- **Tiempo y recursos disponibles**

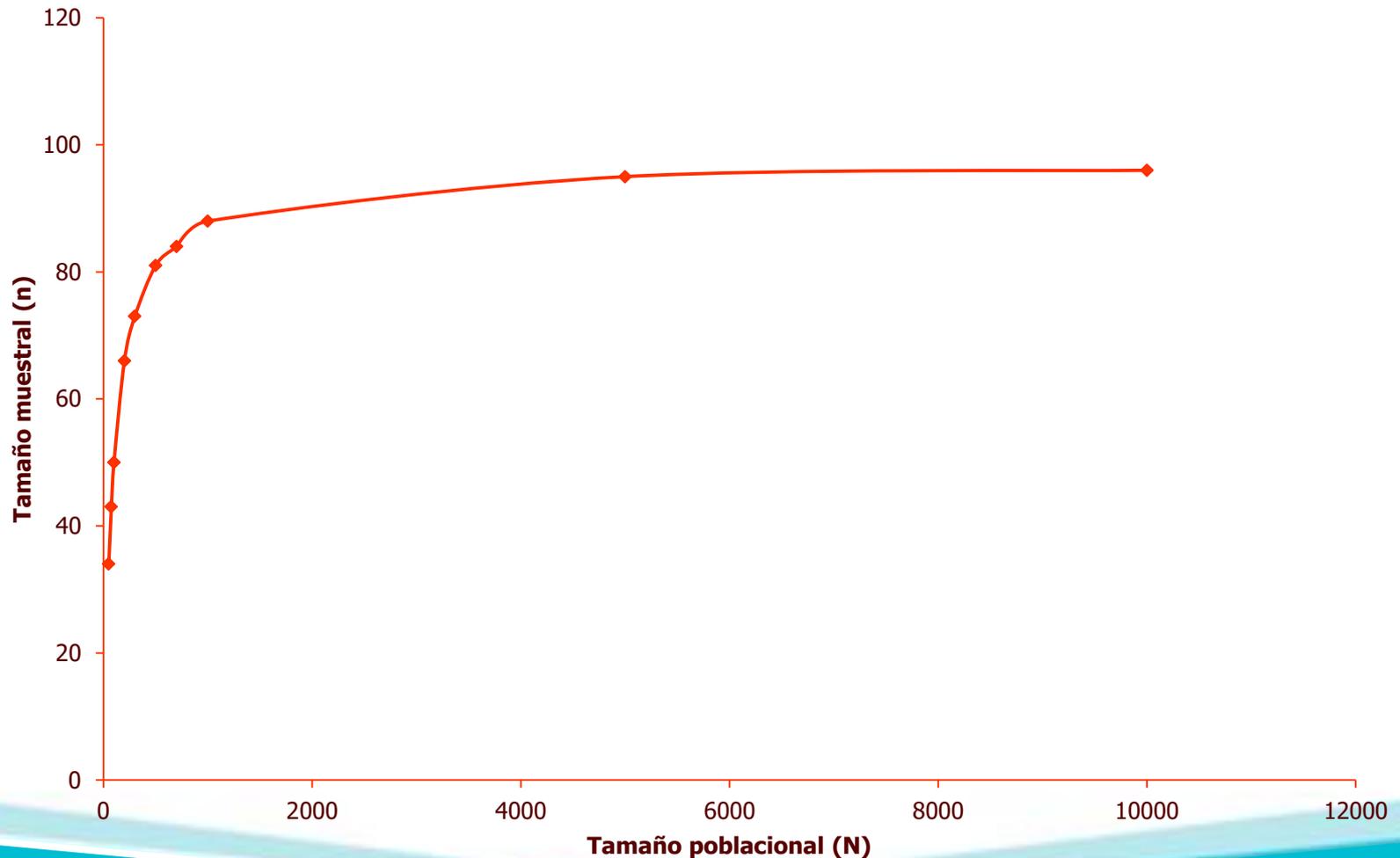
Factores a considerar

Tamaño de la población (N):

Se debe conocer si la población es finita (delimitada en número y ubicación) o infinita (indeterminada en número y ubicación)

La corrección de la muestra sobre el tamaño poblacional es útil hasta ciertos límites

Curva característica del tamaño muestral sobre poblaciones finitas para una variable binomial, con un nivel de confianza del 95%, suponiendo $p=q=0.5$ y un error máximo de 10%



Factores a considerar

Tipo de población:

Conocer cómo están distribuidos los individuos en la población, hay o no agrupaciones “naturales”, interesa controlar esas agrupaciones.

La población es homogénea o heterogénea.

Factores a considerar

Variable (s) a investigar:

¿Cuántas variables se van a medir?

¿Cuál de ellas será seleccionada?

La variable de diseño puede ser seleccionada considerando:

- ✓ Su importancia
- ✓ Su variabilidad

Factores a considerar

Variabilidad esperada (varianza):

Debe conocerse, considerarse, estimarse o suponerse cuánto va a variar la variable de diseño: Varianza, desviación estándar, coeficiente de variación.

Pueden ser útiles datos históricos, investigaciones y conocimientos previos.

Lo mejor: Realizar un estudio piloto

Factores a considerar

Niveles de confianza, significancia y poder:

- En caso se quiera realizar una estimación: Nivel de confianza (generalmente 95%).
- Cuando se desean hacer comparaciones o pruebas de hipótesis se debe establecer:
 - Nivel de significancia α (obligatorio): 0.05 (5%)
 - Poder de la prueba (70 a 80% regularmente)

En ambos casos se usa la distribución normal estándar (Z) que da los valores para las probabilidades indicadas.

Factores a considerar

Margen de error o precisión, tamaño del efecto:

Se establece sobre la medición que se desea realizar, con las mismas dimensionales.

Para estimación: Se interpreta como la máxima diferencia esperada sobre el valor real.

Para pruebas de hipótesis: Se interpreta como la mínima distancia entre los valores a partir de la cual se considerarán diferentes (tamaño del efecto).

Factores a considerar

- Tamaño de la población = N
- Tipo de población (estratos, grupos)
- Variable (s) a investigar
- Variabilidad estimada (precisión) = σ^2
- Nivel de confianza, nivel de significancia y/o poder = Z
- ¿Cuánto margen de error o precisión tendremos para aproximarnos al valor real? = Δ (D, d o e)

Fórmulas para estimación

Población infinita:

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

Población finita:

$$n = \frac{N \sigma^2}{\frac{(N - 1) \Delta^2}{Z^2} + \sigma^2}$$

Fórmula para pruebas de hipótesis

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2 \sigma^2}{\Delta^2}$$

El valor Z_{α} se incrementa según el número de grupos a comparar

Para comparaciones de dos o más grupos se multiplica por dos

Diseños de muestreo

En síntesis se pueden dividir en probabilísticos y no probabilísticos.

Probabilísticos

- Muestreo aleatorio simple
- Muestreo aleatorio sistemático
- Muestreo aleatorio estratificado
- Muestreo por conglomerados

No probabilísticos

- Muestreo por conveniencia o deliberado
- Muestreo según criterio o selección intencional
- Muestreo por cuotas

Cálculo práctico de la muestra

Calculadora muestral *on line* o descargable:

<https://www.imim.cat/ofertadeserveis/software-public/granmo/>

Programas estadísticos con módulos de muestreo (Epidat, Statgraphics)