



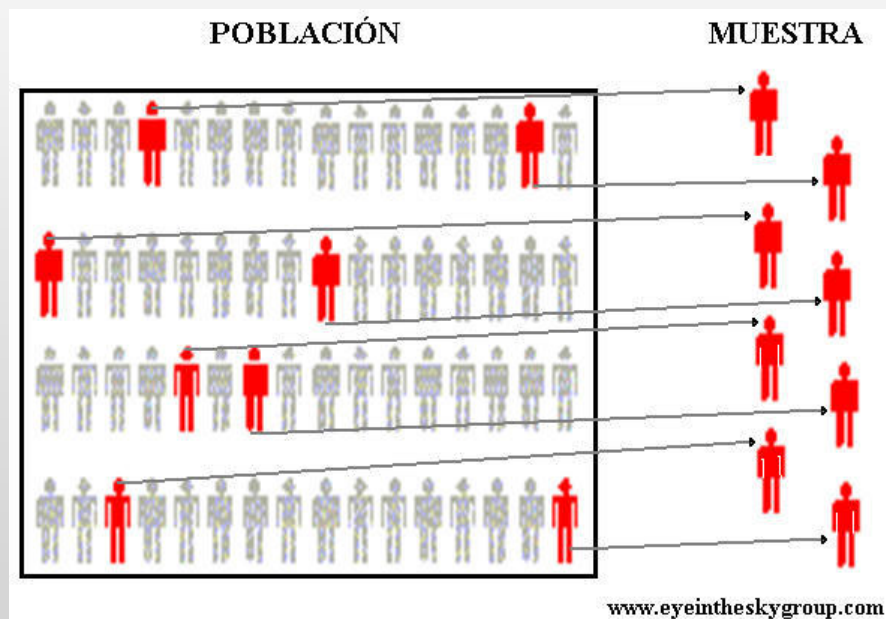
Muestreo Estadístico

Formación por Competencias

Por:
Rodolfo Romero

Muestreo

Muestreo: rama de la estadística que estudia los procesos estadísticos para seleccionar un subconjunto de observaciones o muestra de una población de interés, con el propósito de obtener conclusiones sobre toda la población.



Aplicación del muestreo

Selecciona parte de una población para medir características de interés, y hacer **inferencias** sobre las características de la población.

Cuando??

- ✓ Poblaciones grandes.
- ✓ Costo elevado para censar a cada unidad que forma parte de la población.
- ✓ Los errores no muestrales son menores que en un censo.
- ✓ Población Destructible: al uso se destruye.

Conceptos

- **Universo:** Conjunto de todos los elementos o individuos bajo estudio o sobre los cuales se va a extender las conclusiones a partir de la muestra. **“Población objetivo”**.
- **Población estadística:** Conjunto de mediciones hechas sobre los elementos de un universo.
- **Unidad elemental:** Unidad, elemento o individuo sobre el cual se desea realizar una observación, medición o conteo

Conceptos

- **Unidad de muestreo:** Unidades sobre las cuales es factible muestrear, para posteriormente, obtener la información correspondiente a las unidades elementales.

Puede abarcar más de una unidad elemental.

- **Nivel de confianza:** Probabilidad de que la estimación efectuada se ajuste a la realidad.

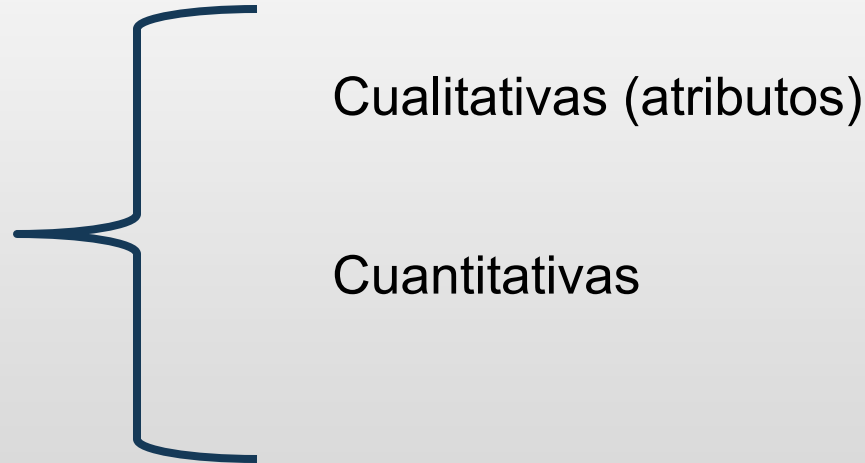
Conceptos

- **Marco de muestreo:** Lista completa de todas las unidades de muestreo, o cualquier material o mecanismo que permita delimitar o identificar en forma apropiada los elementos de una población (Ospina, 2001).
- **Varianza:** diferencia promedio que hay entre cada uno de los valores respecto a su punto central (*Media*). (S^2 para datos muestrales, y σ^2 para datos poblacionales)
- **Desviación estándar:** representa el promedio de diferencia que hay entre los datos y la media. (σ)

Conceptos

- **Variables de estudio:** Características a ser medidas, observadas o contadas en las unidades o elementos de la población de interés.

- Variables



Conceptos

- **Parámetros:** Valores que identifican la distribución de una variable en una población estadística. Los parámetros son las características poblacionales a estimar.
- **Tamaño de la población de muestreo:** El total de unidades elementales de la población accesible “N”
- **Tamaño de muestra:** El total de unidades elementales que serán seleccionadas de la población de acuerdo al diseño muestral “n”

Sesgos y errores en el muestreo

- Una buena muestra reproduce las características de interés que existen en la población de la manera más cercana posible (Lohr, 2000). **“Muestra representativa”**
- Cuando la muestra no es representativa, corresponde a una muestra sesgada

Tipos de sesgos

▪ Sesgo de selección:

- cuando alguna parte de la población objetivo no está representada en la muestra
- cuando no se especifica apropiadamente a la población objetivo, o no se incluye a todos los elementos de muestreo en el marco muestral
- cuando se sustituyen o cambian miembros de la muestra
- cuando se permite que la muestra conste sólo de voluntarios

Tipos de sesgos

▪ Sesgo de medición:

- cuando el instrumento con el que se mide tiende a desviarse del valor verdadero en alguna dirección
- la forma en que se llevan a cabo las mediciones o se aplican las encuestas

Errores

- **Errores de muestreo:** Es el resultado de considerar una muestra y no a toda la población. Estos errores se reportan en términos probabilísticos.

- **Errores que no son de muestreo:** Imprecisiones que no se pueden atribuir a la variabilidad entre las muestras.

Métodos de muestreo

- La elección depende de los objetivos de la investigación y del tipo de análisis que se quiera implementar.

Muestreo
probabilístico

- Selección aleatoria
- Cada unidad tiene una probabilidad conocida, casi siempre igual, de ser conocida

Muestreo no
probabilístico

- Selección sistemática o desordenada.
- No al azar

Métodos de muestreo

Probabilístico	No probabilístico
Aleatorio Simple	Intencional
Sistemático	Por Cuota
Estratificado	Casual o incidental
Conglomerado	Bola de Nieve

Decisión

¿Muestra probabilística o no probabilística?

Depende de:

- Objetivos de estudio
- Recursos económicos y de tiempo
- Diseño de la investigación

Muestreo aleatorio

- Se requiere conocer el total de la población
- Todos los sujetos de N tienen la misma posibilidad de ser seleccionados

Muestreo con reemplazo: un elemento puede ser seleccionado más de una vez en la muestra para ello se extrae un elemento de la población se observa y se devuelve a la población.

Muestreo sin reemplazo: no se devuelve los elementos extraídos a la población hasta que no se hallan extraídos todos los elementos

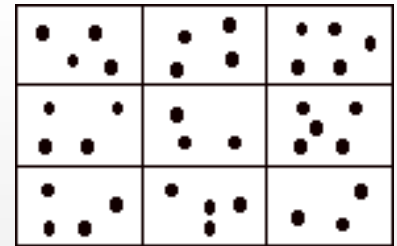
Muestreo sistemático

- Se divide $N/n = k$
- Se ordenan previamente los individuos de N ; después se elige uno de ellos al azar,
 - a intervalos constantes, se eligen todos los demás hasta completar la muestra.

- Ejemplo:
 $N=100$ elementos $n=25$ elementos intervalo de selección $100/25 = 4$.
Se toma aleatoriamente un número entre el 1 y el 4 (salió el 2),
Entonces 2,4,6,8...

Muestreo estratificado

- Organiza a la población en subconjuntos homogéneos (internamente)
- Diferentes entre ellos respecto a la variable a estudiar
- Cada estrato funciona independientemente
- Exige conocimiento detallado de la población
- Reduce el error muestral:
 - ✓ Asegurarse que cada estrato estará representado en la muestra



Ejemplo: Edades, sexo, profesiones

n=20 60% mujeres 40% hombres, entonces, muestra sería 12 mujeres y 8 hombres

Muestreo por conglomerado

- Cuando la población se encuentra dividida, de manera natural, en grupos que se supone que contienen toda la variabilidad de la población (conglomerado)
- Dentro de los grupos seleccionados se ubicarán las unidades elementales
- Simplifica el levantamiento de información muestral.

Ejemplo: unidades hospitalarias, nivel de audiencia de un canal

$n \div N^{\circ}$ de conglomerados = # de elementos por conglomerado



Métodos no probabilísticos

No existe la misma probabilidad de elegir los sujetos de la población



no certeza de representatividad



no se pueden hacer generalizaciones

Intencional o a criterio

Los elementos muestrales son seleccionados a criterio del investigador.

Muestreo representativo

- Muestreo subjetivo

Investigador debe conocer los elementos muestrales

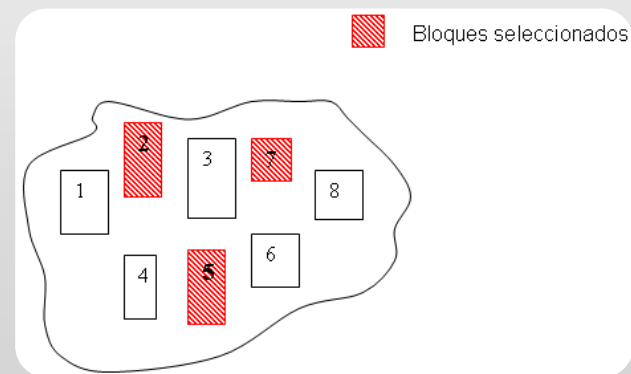
Ejemplo: escogencia de expertos en una determinada área

Por cuota

- Se fijan unas “cuotas” que consisten en un número de individuos con determinadas condiciones.

Ejemplo: 50 individuos de 20 a 45 años, de sexo femenino y residentes del Área Metropolitana.

Después se eligen los primeros que se encuentren que cumplan dichas características



Por casual o incidental

- El investigador selecciona directa e intencionadamente los individuos de la población.
- Muestra conformada por sujetos fácilmente accesibles y presentes en un lugar determinado, y en un momento preciso.
- Los sujetos se incluyen en el estudio a medida que se presentan, y hasta que la muestra alcance el tamaño deseado.

Ejemplo: encuestas a los estudiantes que asisten a una clase específica, un día determinado.

Tamaño de muestra para poblaciones infinitas (muy grandes)

Ejemplo: ¿A cuántas personas tendríamos que estudiar para conocer la prevalencia de diabetes?

Seguridad = 95%; Precisión = 3%; Proporción esperada = asumamos que puede ser próxima al 5%; si no tuviésemos ninguna idea de dicha proporción utilizaríamos el valor $p = 0,5$ (50%) que maximiza el tamaño muestral:

$$n = \frac{Z_{\alpha}^2 * p * q}{d^2}$$

donde:

- $Z_{\alpha}^2 = 1.96^2$ (ya que la seguridad es del 95%)
- p = proporción esperada (en este caso 5% = 0.05)
- $q = 1 - p$ (en este caso $1 - 0.05 = 0.95$)
- d = precisión (en este caso deseamos un 3%)

$$n = \frac{1.96^2 * 0.05 * 0.95}{0.03^2} = 203$$

Tamaño de muestra

Según diferentes seguridades el coeficiente de Z_{α} varía, así:

- Si el nivel de confianza fuese del 90% el coeficiente Z_{α} sería 1.645
- Si el nivel de confianza fuese del 95% el coeficiente Z_{α} sería 1.96
- Si el nivel de confianza fuese del 97.5% el coeficiente Z_{α} sería 2.24
- Si el nivel de confianza fuese del 99% el coeficiente Z_{α} sería 2.576

Tamaño de muestra con población finita (conocida)

Si la población es finita, es decir conocemos el total de la población y deseásemos saber cuántos del total tendremos que estudiar la respuesta sería:

$$n = \frac{N * Z_{\alpha}^2 * p * q}{d^2 * (N - 1) + Z_{\alpha}^2 * p * q}$$

donde:

- N = Total de la población
- $Z_{\alpha}^2 = 1.96^2$ (si la seguridad es del 95%)
- p = proporción esperada (en este caso 5% = 0.05)
- q = 1 - p (en este caso 1-0.05 = 0.95)
- d = precisión (en este caso deseamos un 3%).

¿A cuántas personas tendría que estudiar de una población de 15.000 habitantes para conocer la prevalencia de diabetes?

Seguridad = 95%; Precisión = 3%; proporción esperada = asumamos que puede ser próxima al 5% ; si no tuviese ninguna idea de dicha proporción utilizaríamos el valor p = 0.5 (50%) que maximiza el tamaño muestral.

$$n = \frac{15.000 * 1.96^2 * 0.05 * 0.95}{0.03^2 (15.000 - 1) + 1.96^2 * 0.05 * 0.95} = 200$$

Formación por Competencias

MÓDULO V

Producción de Bienes y Servicios Públicos

www.eap.ucr.ac.cr

Esta presentación se ha publicado en la página Web de la Escuela de Administración Pública de la Universidad de Costa Rica / Campus Bimodal.

Derechos Reservados 2012

Uso autorizado sólo para fines académicos en la Universidad de Costa Rica.

Bibliografía

- Población y muestra. Técnicas de muestreos. Management Mathematics for European Schools
- Elementos de muestreo. Universidad de Venezuela
- Muestreo. Universidad Nacional de Entre Ríos
- ***Estadística aplicada a las Ciencias Sociales. Tamaño necesario de la muestra: ¿Cuántos sujetos necesitamos?***. Universidad Pontificia Comillas • Madrid

E-Mail: [rodolfo.romero @ucr.ac.cr](mailto:rodolfo.romero@ucr.ac.cr)

