

Tema 2: Muestreo.

Estadística. 4º Curso.

Licenciatura en Ciencias Ambientales

- 1 Introducción
- 2 Muestreo Aleatorio Simple
- 3 Muestreo Aleatorio Estratificado
- 4 Muestreo Sistemático
- 5 Muestreo por conglomerados
- 6 Otros diseños muestrales

1. Introducción

Necesidad de realizar un muestreo

- En poblaciones de gran tamaño es imposible estudiar a todos y cada uno de los individuos.
- Se debe seleccionar un subconjunto de dicha población, al que denominaremos *muestra*, cuyos individuos serán los objetos de la investigación.
- El proceso de selección de los individuos no debe estar condicionado por las ideas previas del experimentador. Por el contrario, los individuos deben ser seleccionados de forma aleatoria, formando lo que se denomina una *muestra aleatoria*. A este proceso de selección aleatoria de la muestra se le denomina *muestreo*.
- La muestra debe ser representativa de la población, de modo que las conclusiones del estudio puedan ser extrapoladas a dicha población. Con esta finalidad, se han desarrollado una serie de técnicas que tienen como objetivo hacer más eficiente el proceso de muestreo.

2. Muestreo Aleatorio Simple

Se caracteriza porque cada individuo de la población tiene la misma probabilidad de ser elegido en la muestra.

En la práctica se desarrolla numerando a los individuos de la población y extrayendo una serie de números aleatorios que determinarán los individuos seleccionados.

Ventajas del Muestreo Aleatorio Simple

- Es el tipo de muestreo más sencillo de llevar a cabo y el que más se utiliza.
- El azar es el único elemento que interviene en la elección de la muestra, lo que evita cualquier tipo de subjetividad.

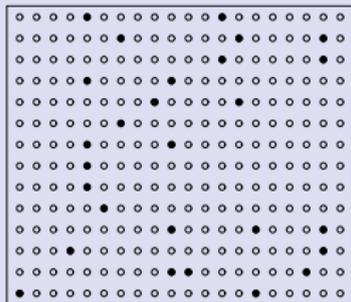
Inconvenientes del Muestreo Aleatorio Simple

- Para llevarlo a cabo es necesario disponer de una copia del censo (listado de individuos de la población).
- Al dejarlo todo al azar la muestra puede no ser suficientemente representativa de la población que estamos estudiando.

2. Muestreo Aleatorio Simple

Ejemplo 1

Se desea determinar la concentración media, en mg/cm^2 , de cierto producto químico en el suelo de un determinado cementerio de coches para saber si se ha de proceder a su limpieza antes de que pase a las aguas subterráneas. En la siguiente figura aparece un plano de dicho cementerio.



Se ha superpuesto un enrejado con 280 áreas. Tenemos la posibilidad de medir la concentración en 27 de dichas áreas. Diseñamos una M.A.S. generando en el ordenador 27 números aleatorios entre 1 y 280.

3. Muestreo Aleatorio Estratificado

Se divide la población en subconjuntos llamados *estratos*. Los individuos de un estrato son muy similares entre ellos en relación al carácter que nos interesa. Por contra los individuos de un estrato son muy diferentes de los individuos de los otros estratos. Se aplica en poblaciones muy heterogéneas frente al carácter que se está estudiando.

Para extraer la muestra se lleva a cabo un muestreo aleatorio simple en cada estrato. El número de individuos de cada estrato que van a formar parte de la muestra debe ser proporcional al tamaño del estrato.

Si en la población hay N individuos divididos en estratos de N_1, N_2, \dots individuos y en la muestra deseamos seleccionar n individuos, el número de individuos seleccionados de cada estrato será

$$n_1 = \frac{N_1}{N} \cdot n \quad \text{del estrato 1} \quad , \quad n_2 = \frac{N_2}{N} \cdot n \quad \text{del estrato 2} \quad , \dots$$

3. Muestreo Aleatorio Estratificado

Ventajas del Muestreo Aleatorio Estratificado

- Contempla una fuente de variabilidad, lo cual disminuirá el error al extrapolar las conclusiones del estudio al conjunto de toda la población.
- Es posible estudiar cada estrato como una subpoblación en sí mismo, llegando a conclusiones sobre el comportamiento del carácter bajo estudio en dicho estrato.

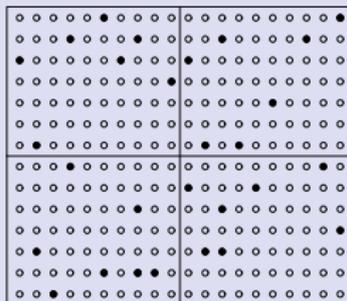
Inconvenientes del Muestreo Aleatorio Estratificado

- Si a la hora de formar los estratos nos basamos en supuestos incorrectos, esto puede llevar a conclusiones erróneas en el estudio. Por ello es imprescindible que los estratos se formen en base a consideraciones obvias como localización espacial, altitud, edad, etc.
- Al igual que en un M.A.S. hay que disponer de un censo de cada uno de los estratos en que hemos dividido la población.

3. Muestreo Aleatorio Estratificado

Ejemplo 1 (continuación)

Supongamos que el suelo del cementerio tiene 4 zonas diferenciadas. Un Muestreo Aleatorio Estratificado podría dar el siguiente resultado:



Se han dividido los 280 posibles puntos de muestreo en 4 estratos de 70 que se corresponden con las zonas antes aludidas. En cada zona se ha seleccionado, mediante un M.A.S., una muestra de 7 puntos.

4. Muestreo Sistemático

Se aplica cuando la población es bastante irregular respecto al carácter que estamos estudiando y deseamos que en la muestra se refleje toda esta variabilidad.

Para llevarlo a la práctica se numera la población y se calcula el paso del muestreo mediante la fórmula $k = N/n$, siendo N el tamaño de la población y n el de la muestra. Se elige un individuo al azar entre los k primeros. A partir de dicho individuo se elige un nuevo elemento cada k unidades hasta completar la muestra.

Ventajas del Muestreo Sistemático

- El proceso de elección de la muestra es más sencillo.
- Si elegimos k pequeño, la muestra resultante cubre toda la variabilidad de la población respecto al carácter bajo estudio.

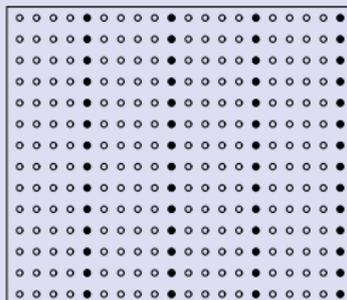
Inconvenientes del Muestreo Sistemático

- Depende de la ordenación que le demos a los individuos de la población.
- Puede dar problemas si la población presenta un comportamiento cíclicos.
- Al igual que los anteriores se requiere un censo de la población para poder llevarlo a cabo.

4. Muestreo Sistemático

Ejemplo 1 (continuación)

Se ha ordenado la población (de izquierda a derecha y de arriba a abajo). Se ha calculado $k = 5$ y se ha elegido un número al azar entre 1 y 5 que ha sido el 5. A partir de aquí se toma un nuevo punto de muestreo cada 5 unidades.



5. Muestreo por Conglomerados

Se divide la población en *conglomerados*. Un conglomerado es un subconjunto de la población que, en términos generales, posee una variabilidad parecida a la del global de dicha población respecto al carácter que estamos estudiando (podría decirse que un conglomerado es una copia en pequeño de la población).

La metodología a emplear es la siguiente:

- Dividir la población en conglomerados no excesivamente grandes. Esta división se realiza habitualmente atendiendo a divisiones geográficas, administrativas, o de otro tipo que resulten evidentes.
- Consideramos el conjunto de los conglomerados como una nueva población, procediéndose a la selección de una serie de conglomerados mediante un M.A.S.
- Realizamos las mediciones de las variables bajo estudio en todos los individuos de los conglomerados seleccionados.

5. Muestreo por Conglomerados

Ventajas del Muestreo por Conglomerados

Muy aconsejable cuando no disponemos de un censo de la población o los recursos económicos o temporales son limitados.

Inconvenientes del Muestreo por Conglomerados

Si los conglomerados no reproducen exactamente la variabilidad de la población, los errores al extrapolar los resultados de la investigación serán mayores.

6. Otros diseños muestrales

Muestreo Multietápico

Se trata de combinar dos o más diseños muestrales de los anteriormente descritos. Cada vez que aplicamos un diseño muestral diremos que hemos llevado a cabo una *etapa* del muestreo.

Habitualmente la primera etapa en un muestreo multietápico consiste en la aplicación de un muestreo por conglomerados. Una vez seleccionados los conglomerados, en cada uno de ellos se puede aplicar un muestreo aleatorio simple, estratificado, sistemático o un segundo muestreo por conglomerados.

Como puede deducirse, en dos o tres etapas (no suele haber más) hay bastantes posibilidades de combinar los diseños muestrales estudiados en este tema.

Muestreo aleatorio con probabilidades desiguales

En esencia se desarrolla como un Muestreo Aleatorio Simple, pero a diferencia de este, no todos los individuos tienen la misma probabilidad de ser seleccionados para la muestra. Conlleva el riesgo de que se produzca un sesgo porque la representación en la muestra de determinados grupos puede ser mayor de la que correspondería a dichos grupos por su peso en la población.