TEMA 58: POBLACIÓN Y MUESTRA. CONDICIONES DE REPRESENTATIVIDAD DE UNA MUESTRA. TIPOS DE MUESTREO. TAMAÑO DE UNA MUESTRA.

TEMA DE MUESTRA

I. POBLACIÓN Y MUESTRA

II. REPRESENTACIÓN DE UNA MUESTRA

III. TIPOS DE MUESTREO. TAMAÑO DE LA MUESTRA

III.1. MUESTREO PROBABILÍSTICO

III.1.A. MUESTREO CON PROBABILIDADES IGUALES

III.1.B. MUESTREO CON PROBABILIDADES DESIGUALES

III.1.C. MUESTREO ESTRATIFICABLE

III.1.D. MUESTREO POR CONGLOMERADOS

III.1.E. MUESTREO SISTEMÁTICO

III.1.F. OTROS MUESTREOS

III.2. MUESTREO INTENCIONAL U OPINATIVO

III.3. MUESTREO SIN NORMA

IV. BIBLIOGRAFÍA

TEMA 58: POBLACIÓN Y MUESTRA. CONDICIONES DE REPRESENTATIVIDAD DE UNA MUESTRA. TIPOS DE MUESTREO. TAMAÑO DE UNA MUESTRA.

Uno de los objetivos de la Estadística es hacer inferencias, es decir, extraer conclusiones acerca de conjuntos definidos por una propiedad a partir de un subconjunto determinado de los mismos. Uno de los problemas que se plantean, y que estudiaremos en este tema, es cómo elegir el subconjunto de forma que los datos que a partir de él se obtengan pueden ser extrapolados a toda la población.

Es evidente que cuantas más observaciones realicemos, mejor información tendremos de la distribución del conjunto del cual se extraen los datos, pero generalmente o es imposible o es costoso hacer extracciones experimentales de todos los individuos del grupo. Por ello, se debe buscar un método para determinar cuántas observaciones precisamos para poder conocer lo que deseamos con un error máximo fijado previamente. Este es el objetivo central de la Teoría del Muestreo.

I. POBLACIÓN Y MUESTRA

Los primeros conceptos fundamentales a considerar en la teoría de muestras son los de población y muestra.

Se denomina *población* o universo a cualquier colección finita o infinita de individuos o elementos sobre los que se quiere realizar un estudio.

Los elementos de una población no tienen por qué ser personas; pueden ser árboles, ranas, libros…

En una población se puede medir una o varias características y clasificar sus unidades de acuerdo con ellas.

A partir de estos resultados podemos hallar valores como la media, varianza, etc., a los que denominaremos parámetros o características poblacionales.

Se da el nombre de censo a la enumeración y anotación de ciertas características de todos los elementos de la población.

TEMA DE MUESTRA

A la población que se intenta investigar se le denomina población objetivo y estará compuesta por todos los individuos de la población. Sin embargo, en el mundo real existen omisiones, duplicaciones, unidades extrañas. Por otra parte quizás no pueda obtenerse información de algunas unidades por diferentes motivos, como que no están accesibles en ese momento o que se niegan a colaborar. Todo ello hace que el conjunto que realmente es objeto de la investigación sea diferente de la población objetivo y será denominada población investigada.

En muchos casos puede que no sea posible o conveniente obtener información de todos los individuos de la población, por lo que el estudio se reduce a una parte de la misma.

Se denomina *muestra* a una parte de la población que se obtiene con el fin de investigar las propiedades de la población.

Es deseable que la muestra represente lo mejor posible a los elementos de la población, en el sentido de que proporcione buena información sobre ella.

Para elegir las unidades que van a constituir la muestra es necesario disponer de un conjunto real de unidades que se ajuste lo mejor posible al conjunto ideal que constituye la población objetivo. Al listado de unidades que componen la población a partir del cual se selecciona la muestra lo llamaremos marco.

Llamamos *muestreo* al conjunto de reglas o al procedimiento mediante el cual se obtiene una o más muestras de la población.

II. REPRESENTACIÓN DE UNA MUESTRA

Cuando se extrae una muestra, los datos obtenidos a partir de ella nos permiten inferir unos valores aproximados de la población en su totalidad. A estos valores aproximados se les denomina *estimadores*, los cuales vendrán afectados por unos errores que llamaremos *errores de muestreo*. Cuanto menor sea el error de muestreo mayor será la precisión del estimador correspondiente.

Este tipo de errores siempre van asociados a las muestras, es decir, siempre cometeremos error de muestreo sea cual sea el método de elección de una muestra. Una muestra será tanto más representativa cuanto más pequeño sea el error de muestreo, entendiendo por muestra representativa aquella que nos sirve para juzgar con bastante certeza la característica que nos interesa de la población total.

TEMA DE MUESTRA

Vamos a tratar, pues, de saber más cosas del error de muestreo, viendo qué propiedades debe tener un buen estimador y cuál es el error de muestreo debido a dicho estimador.

Si $\hat{\theta}$ es un estimador de un parámetro desconocido θ de la población, una propiedad deseable de $\hat{\theta}$ es que sea insesgado, es decir, que $E(\hat{\theta}) = \theta$.

En caso de que $\hat{\theta}$ sea insesgado mediremos su precisión mediante la varianza (cuanto más pequeña, más preciso será $\hat{\theta}$). Si $\hat{\theta}$ no es insesgado, su precisión se mide con el error cuadrático medio:

$$ECM(\hat{\theta}) = E\left[(\hat{\theta} - \theta)^2\right]$$

(a menor error mayor precisión), que coincide con la varianza en el caso de ser insesgado.

En general, y partiendo de la expresión anterior, se deduce que

$$ECM(\hat{\theta}) = V(\hat{\theta}) + B^2$$

donde $V(\hat{\theta})$ es la varianza de $\hat{\theta}$ y $B = E(\hat{\theta}) - \theta$ es el llamado sesgo.

Entonces, como hemos visto, es la varianza de un estimador la que mide el error de muestreo. También se puede medir mediante la desviación típica, $\sigma_{\hat{\theta}}$, ya que es la raíz cuadrada de la varianza.

A veces también se considera el error de muestreo relativo o coeficiente de variación:

$$e.m.r = \frac{\sigma_{\hat{\theta}}}{E(\hat{\theta})}$$

Puede parecer que los errores de muestreo imponen una seria limitación a la utilización de muestras, pero esta impresión es injustificada, pues existen métodos que minimizan estos errores y procedimientos que permiten establecer el error de muestreo que estamos dispuestos a tolerar, y seleccionar la muestra de tal forma que las conclusiones que hagamos sobre los valores de la población no vengan afectados por errores de muestreo superiores a los establecidos.

TEMA DE MUESTRA

Además de los errores de muestreo existen otros errores ajenos al muestreo, como pueden ser:

- Errores de medida, que son la diferencia entre el valor observado y el valor real
- Errores de procesamiento, que son debidos a errores en la entrada, edición y análisis de datos.
- Errores de "falta de respuesta", que se producen cuando para ciertos elementos de la muestra no se obtiene la información deseada.

III. TIPOS DE MUESTREO. TAMAÑO DE LA MUESTRA

III.1. MUESTREO PROBABILÍSTICO

Es aquel en el que se puede calcularse de antemano la probabilidad de seleccionar una de las muestras posibles de la población.

La selección de la muestra se considera como un experimento aleatorio y por eso habla de muestras aleatorias en el sentido de muestras probabilísticas.

Se dice que el muestreo es probabilístico superior cuando conocemos la probabilidad de extracción de cierta parte de la población pero no de un elemento dentro de ella.

Por otra parte está el muestreo probabilístico inferior que es aquel en que se conoce la probabilidad de selección dentro de una parte de la población, pero no la probabilidad de selección de dicha parte.

Para que el muestreo sea probabilístico debe ser inferior y superior.

Vamos a estudiar varios tipos de muestreo probabilístico.

III.1.A. MUESTREO CON PROBABILIDADES IGUALES

Es el muestreo base para los muestreos probabilísticas; los conceptos fundamentales y las fórmulas importantes derivan de él. Por otra parte, sirve como referencia cuando se intenta evaluar la calidad de un método cualquiera.

La única hipótesis de la que se parte consiste en suponer la existencia de un marco perfecto, es decir sin omisiones ni duplicaciones.

Podemos distinguir dos tipos de muestreo con probabilidades iguales:

- Sin reemplazo o **muestreo aleatorio simple** (m.a.s)
- **Con reemplazo**

TEMA DE MUESTRA

El **m.a.s** es el más sencillo y el más utilizado. No necesita de ningún conocimiento a priori de la población. La muestra se obtiene unidad a unidad, sin reposición de éstas a la población después de cada selección, y con la misma probabilidad.

Supongamos que tengo una población P de tamaño N, cuyos elementos son $P=\{n_1,n_2,\ldots,n_N\}$. En ella estudiamos una variable X que toma el valor x_i en n_i.

Fijado el tamaño muestral n, considero todas las posibles muestras de tamaño n, y llamaré a ese conjunto S (es el espacio muestral).

Representaré con letras mayúsculas las características o parámetros de la población y con minúsculas las de la muestra. Así tengo:

N= tamaño poblacional n= tamaño muestral

$$\overline{X}=\frac{\sum_{i=1}^{N}x_i}{N}=\text{ media población} \qquad \overline{x}=\frac{\sum_{i=1}^{n}x_i}{n}=\text{ media muestral}$$

$$S^2=\frac{\sum_{i=1}^{N}\left(x_i-\overline{X}\right)^2}{N-1}=\text{ cuasivarianza pobl.} \qquad s^2=\frac{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2}{n-1}=\text{ cuasivarianza muestral}$$

$$\sigma^2=\frac{\sum_{i=1}^{N}\left(x_i-\overline{X}\right)^2}{N}=\text{ varianza pobl.} \qquad s'^2=\frac{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2}{n}=\text{ varianza muestral}$$

Lo que vamos a hacer será hallar estimadores del total poblacional (X), de la media poblacional ($\overline{X}$) y de la proporción de la población que está en un grupo determinado (P). Veremos si estos estimadores son buenos calculando sus varianzas y hallaremos el tamaño muestral óptimo en cada caso.

Antes de nada, veamos qué forma tomará un estimador insesgado de un parámetro cualquiera θ, que será de la forma $\theta=\sum_{i=1}^{N}a_ix_i$.

Un estimador de θ vendrá dado por $\hat{\theta}=\sum_{i=1}^{n}b_ix_i=\sum_{i=1}^{n}b_ix_iI_i$ donde

$$I_i(s)=\begin{cases}1 \text{ Si } n_i \text{ pertenece a la muestra } s\\ 0 \text{ Si } n_i \text{ no está en la muestra } s\end{cases} \qquad \text{con } s\in S$$

TEMA DE MUESTRA

Si imponemos que sea insesgado, entonces $E(\hat{\theta}) = \theta$:

$$E(\hat{\theta}) = E\left(\sum_{i=1}^{N} b_i x_i I_i\right) = \sum_{i=1}^{N} b_i x_i E(I_i) = (*)$$

Si represento por Π_i a la probabilidad de que el elemento n_i sea seleccionado en la muestra, entonces

$$E(I_i) = 1 \cdot P(n_i \in s) + 0 \cdot P(n_i \notin s) = \Pi_i$$

Entonces

$$(*) = \sum_{i=1}^{N} b_i x_i \Pi_i = \theta = \sum_{i=1}^{N} a_i x_i \text{ de donde } b_i \Pi_i = a_i \Rightarrow b_i = \frac{a_i}{\Pi_i}.$$

Por tanto $\hat{\theta} = \sum_{i=1}^{n} \frac{a_i}{\Pi_i} x_i$ es estimador insesgado de θ.

Es fácil comprobar que tanto si dos muestras con las mismas unidades en distinto orden coinciden, como si dos muestras con las mismas unidades en distinto orden son diferentes, se cumplen que $\Pi_i = \frac{n}{N}$.

También en ambos casos se comprueba que si Π_{ij} denota la probabilidad de que n_i y n_j sean seleccionadas conjuntamente en la muestra, entonces $\Pi_{ij} = \frac{n(n-1)}{N(N-1)}$. (Este valor se usará más tarde).

Ya estamos en condiciones de hallar los estimadores insesgados de X, $\bar{X}$ y P:

- $\bar{X} = \sum_{i=1}^{N} \frac{1}{N} x_i$, es decir, $a_i = \frac{1}{N} \quad \forall_i$. Por tanto, $\hat{\bar{X}} = \sum_{i=1}^{n} \frac{1}{N} \cdot \frac{1}{n/N} x_i = \sum_{i=1}^{n} \frac{1}{n} \cdot x_i = \bar{X}$
- $X = \sum_{i=1}^{N} x_i$, es decir, $a_i = 1 \quad \forall_i$. Por tanto, $\hat{X} = \sum_{i=1}^{n} \frac{1}{n/N} x_i = N \sum_{i=1}^{n} \frac{1}{n} \cdot x_i = N\bar{x}$
- $P = \frac{1}{N} \sum_{i=1}^{N} A_i$ donde $A_i = \begin{cases} 1 \text{ Si } n_i \in \text{ a la clase} \\ 0 \text{ Si } n_i \notin \text{ a la clase} \end{cases}$..

  Entonces, $\hat{P} = \sum_{i=1}^{n} \frac{1}{N} \cdot \frac{1}{n/N} A_i = \frac{1}{n} \sum_{i=1}^{n} A_i = p$

TEMA DE MUESTRA

(p= proporción de la muestra)

Estos estimadores son lineales e insesgados; si además tienen mínima varianza, entonces los valores de la variable están muy próximos al parámetro desconocido, y serán por tanto buenos estimadores.

Estudiemos pues, la varianza de los estimadores.

■ Para la media

Sabemos que $\hat{\bar{X}} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}\sum_{i=1}^{N} x_i I_i$

$$\operatorname{var}\left(\hat{\bar{X}}\right) = \operatorname{Var}\left(\frac{1}{n}\sum_{i=1}^{N} x_i I_i\right) = \frac{1}{n^2}\operatorname{var}\left(\sum_{i=1}^{N} x_i I_i\right) =$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{N}\operatorname{var}(x_i I_i) + \sum_{i\neq j}\operatorname{cov}(x_i I_i, x_j I_j)\right) =$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{N} x_i^2 \operatorname{var}(I_i) + \sum_{i\neq j} x_i x_j \operatorname{cov}(I_i, I_j)\right) = (*)$$

$$\left[\begin{array}{l}
E(I_i) = \Pi_i, \text{ como ya vimos} \\
E(I_i^2) = 1^2 \cdot P(n_i \in \text{a la muestra}) + 0^2 \cdot P(n_i \notin \text{a la muestra}) = \Pi_i \\
\operatorname{Var}(I_i) = E(I_i^2) - \left(E(I_i^2)\right)^2 = \Pi_i - \Pi_i^2 = \Pi_i(1-\Pi_i) \\
\operatorname{Cov}(I_i, I_j) = E(I_i, I_j) - E(I_i)E(I_j) \underset{\downarrow}{=} \Pi_{ij} - \Pi_i\Pi_j \\
I_i I_j = \left\{\begin{array}{ll} 1 & \text{si } n_i, n_j \in \text{ a la muestra} \\ 0 & \text{en otro caso} \end{array}\right\} \Rightarrow E(I_i I_j) = \Pi_{ij}
\end{array}\right.$$

$$(*) = \frac{1}{n}\left(\sum_{i=1}^{N} x_i^2 \Pi_i(1-\Pi_i) + \sum_{i\neq j} x_i x_j (\Pi_{ij} - \Pi_i\Pi_j)\right) = (*)$$

$$\Pi_i(1-\Pi_i) = \frac{n}{N}\left(1 - \frac{n}{N}\right) = \frac{n(N-n)}{N^2}$$

$$\Pi_{ij} - \Pi_i\Pi_j = \frac{n(n-1)}{N(N-1)} - \frac{n}{N}\cdot\frac{n}{N} = \frac{-n(N-n)}{N^2(N-1)}$$

TEMA DE MUESTRA

Sacando, en ambos miembros, factor común $\frac{(N-n)n}{N^2(N-1)}$

$$(*)=\frac{N-n}{nN^2(N-1)}\left((N-1)\sum_{i=1}^{N}x_i^2-\sum_{i\neq j}x_ix_j\right)=(*)$$

Desarrollamos σ^2:

$$\sigma^2=\frac{\sum_{i=1}^{N}(x_i-\bar{X})^2}{N}=\frac{\sum_{i=1}^{N}x_i^2}{N}-\bar{X}^2=\frac{\sum_{i=1}^{N}x_i^2}{N}-\frac{\left(\sum_{i=1}^{N}x_i\right)^2}{N^2}=$$

$$=\frac{\sum_{i=1}^{N}x_i^2}{N}-\frac{\sum_{i=1}^{N}x_i^2+\sum_{i\neq j}^{N}x_ix_j}{N^2}=\left(\frac{1}{N}-\frac{1}{N^2}\right)\sum_{i=1}^{N}x_i^2-\frac{1}{N^2}\sum_{i\neq j}^{N}x_ix_j=$$

$$=\frac{1}{N^2}\left((N-1)\sum_{i=1}^{N}x_i^2-\sum_{i\neq j}x_ix_j\right)$$

Entonces, nos queda que

$$(*)=\frac{N-n}{n(N-1)}\sigma^2=\frac{N-n}{n(N-1)}\cdot\frac{N-1}{N}\cdot S^2=\frac{N-n}{N}\cdot\frac{S^2}{n}=(1-f)\frac{S^2}{n}$$

con $f=\frac{n}{N}$. Al factor 1-f se le llama corrección para poblaciones finitas.

Conclusión:

$$Var\left(\hat{\bar{X}}\right)=(1-f)\frac{S^2}{n}$$

$$e.m.\left(\hat{\bar{X}}\right)=\sigma_{\bar{x}}=\sqrt{1-f}\,\frac{S}{\sqrt{n}}$$

$$c.v.\left(\hat{\bar{X}}\right)=\frac{\sigma_{\bar{x}}}{\bar{X}}=\sqrt{1-f}\,\frac{S/\bar{X}}{\sqrt{n}}$$

■ Para el total:

$$\hat{X}=N\bar{x}\Rightarrow Var\left(\hat{X}\right)=Var\left(N\bar{x}\right)=N^2Var\left(\hat{\bar{X}}\right)=N^2(1-f)\frac{S^2}{n}$$

El error de muestreo es $e.m.\left(\hat{X}\right)=N\sqrt{1-f}\,\frac{S}{\sqrt{n}}$

■ Para la proporción

TEMA DE MUESTRA

$$\hat{P} = p \Rightarrow Var\left(\hat{P}\right) = \frac{N-n}{N-1} \cdot \frac{\sigma^2}{n} = (*)$$

$$P = \frac{\Sigma A_i}{N} \text{ con } A_i \sim B(1,P)$$

$$\text{Luego } \sigma^2 = P(1-P) = PQ$$

$$(*) = \frac{N-n}{N-1} \cdot \frac{P \cdot Q}{n}$$

$$e.m.\left(\hat{P}\right) = \sqrt{\frac{N-n}{N-1} \cdot \frac{PQ}{n}}$$

El problema es que hemos obtenido que las varianzas de los estimadores dependen de la cuasivarianza poblacional, en general desconocida. Por tanto, tendremos que estimarla.

Se puede demostrar (de manera no muy fácil) que la cuasivarianza muestral es un estimador insesgado de la cuasivarianza poblacional.

Considerando lo anterior como válido, en todas las fórmulas anteriores se sustituye S por s, y obtenemos:

$$\widehat{Var}\left(\hat{\bar{X}}\right) = (1-f)\frac{s^2}{n}$$

$$\widehat{Var}\left(\hat{\bar{X}}\right) = N^2(1-f)\frac{s^2}{n}$$

$$\widehat{Var}\left(\hat{P}\right) = \frac{N-n}{N} \cdot \frac{s^2}{n} = \frac{N-n}{N} \cdot \frac{n}{n-1} \cdot \frac{\hat{P}\hat{Q}}{n} = \frac{N-n}{N} \cdot \frac{\hat{P}\hat{Q}}{n}$$

Por lo general no se puede obtener el tamaño exacto de una muerta pero sí una aproximación. La decisión sobre el tamaño de una muestra es importante ya que si tomamos una muestra muy grande puede implicar un desperdicio de recursos, y si la cogemos pequeña disminuye la utilizad de los resultados. La teoría de muestreo proporciona un planteamiento general que soluciona este problema.

Para ello fijamos un error máximo admisible $\left(e = K\sigma_{\hat{\theta}}\right)$, y fijo también un nivel de confianza p_k

TEMA DE MUESTRA

■ Si lo que quiero es estimar la media, el tamaño muestra óptimo será:

$$e = K\sigma_{\hat{\bar{x}}} = K\sqrt{(1-f)\frac{S^2}{n}} \Rightarrow e^2 = K^2(1-f)\frac{S^2}{n} \Rightarrow$$

$$e^2 = K^2\frac{S^2}{n} - K^2\frac{S^2}{N} \Rightarrow n = \frac{K^2S^2}{e^2 + \frac{K^2S^2}{N}} \Rightarrow$$

$$\Rightarrow n_{optimo} = \frac{K^2S^2}{1 + \frac{1}{N}\cdot\frac{K^2S^2}{e^2}}$$

■ Si quiero estimar totales, el tamaño muestral óptimo será:

$$e = K\sigma_{\hat{\bar{x}}} = K\sqrt{N^2(1-f)\frac{S^2}{n}} \Rightarrow e^2 = K^2N^2(1-f)\frac{S^2}{n} \Rightarrow$$

$$\Rightarrow n_{opt} = \frac{\frac{K^2N^2S^2}{e^2}}{1 + \frac{1}{N}\cdot\frac{K^2N^2S^2}{e^2}}$$

■ Si quiero estimar proporciones

$$e = K\sigma_{\hat{p}} = K\sqrt{\frac{(N-n)}{N-1}\cdot\frac{PQ}{n}} \Rightarrow e^2 = K^2\frac{N-n}{N-1}\cdot\frac{PQ}{n} \Rightarrow$$

$$\Rightarrow e^2 = K^2\frac{N}{N-1}\cdot\frac{PQ}{n} - K^2\frac{1}{N-1}\cdot PQ \Rightarrow$$

$$\Rightarrow n_{opt} = \frac{\frac{K^2N^2PQ}{N-1}}{e^2 + \frac{K^2PQ}{N-1}} = \frac{\frac{N}{N-1}\cdot\frac{K^2PQ}{e^2}}{1 + \frac{1}{N-1}\cdot\frac{K^2PQ}{e^2}}$$

Como vemos n está en función de S^2 que es desconocida. Lo que tendremos que hacer será estimar S^2 por datos de otros experimentos anteriores o mediante una encuesta piloto.

El **muestreo con reemplazo** se da cuando elegimos la muestra unidad a unidad, con reposición de éstas a la población después de cada selección.

El desarrollo que hemos hecho para el m.a.s se puede hacer aquí de forma análoga. Indiquemos, por ejemplo, que los estimadores son los mismos en ambos

TEMA DE MUESTRA

tipos de muestreo y que las varianzas son incluso más sencillas. Resaltamos que los estimadores de las varianzas en el m.a.s eran insesgados, sin embargo en el muestreo con reemplazo no tienen por qué serlo.

III.1.B. MUESTREO CON PROBABILIDADES DESIGUALES

Las extracciones con probabilidades iguales están basadas en la idea de que si no se conoce nada a priori de las variables estudiadas, o si no se puede utilizar la información que de ellas se posee, lo mejor es atribuir un peso igual a las unidades de la población.

Con frecuencia no se dan en la práctica estas características. Es normal que el organizador de la encuesta utilice datos o hipótesis anteriores a la realización de ésta. Estos datos le llevan a dar más importancia a ciertas unidades frente a otras, por lo que la probabilidad de selección de cada unidad sea distinta y variará según la importancia que ésta tenga.

III.1.C. MUESTREO ESTRATIFICABLE

Tiene como objetivo mejorar las estimaciones agrupando a los elementos más parecidos entre sí. Se divide la población en varias subpoblaciones o estratos, generalmente con elementos que tienen características comunes.

La selección de la muestra se hace tomando elementos de cada uno de los estratos:

E_i = Estrato i-ésimo

E_1	E_2	E_3
N_1	N_2	N_3

$N_1 + N_2 + N_3 = N$ (Tamaño población)

$n_1 + n_2 + n_3 = n$ (Tamaño muestra)

El reparto del tamaño de la muestra entre los diferentes estratos se llama afijación y puede ser de tres tipos:

- Uniforme: que consiste en repartir por igual la muestra en los diferentes estratos.
- Proporcional: que consiste en distribuir la muestra proporcionalmente al tamaño de cada estrato

TEMA DE MUESTRA

• Óptima: que hace que el reparto sea proporcional al número de elementos de cada estrato y a la desviación típica del mismo. Así se tiene en cuenta la falta de homogeneidad de la subpoblación.

Este tipo de muestreo tiene la ventaja de que pueden sacarse conclusiones en cada estrato.

III.1.D. MUESTREO POR CONGLOMERADOS

Consiste en dividir la población en partes (conglomerados) buscando que los elementos de cada conglomerado sean lo más heterogéneos posible, para que cada uno represente de alguna manera a la población. Se hace un listado de todos los conglomerados y se selecciona al azar un número determinado de esta lista.

III.1.E. MUESTREO SISTEMÁTICO

En este muestreo se enumeran los elementos de la población. Se toma $K = \frac{N}{n} = \frac{\text{tamaño población}}{\text{tamaño muestra}}$, y se elige aleatoriamente un elemento de entre los K primeros.

A continuación se toman los siguientes de K en K.

Cuando el origen no es aleatorio sino que se toman los elementos centrales de cada grupo de K consecutivos, el muestreo se denomina rígido o estrictamente sistemático y deja de ser probabilístico.

F. OTROS MUESTREOS

■ Muestreo doble o bifásico: Consta de dos fases.

En la primera se toma una muestra grande de forma rápida, sencilla y poco costosa, en la que se estudiará algún aspecto fundamental para la segunda fase. En la segunda fase se selecciona una muestra de entre los seleccionados en la muestra anterior por algún otro método.

Si se utilizan más de dos fases se llama muestreo múltiple.

■ Muestreo repetido: Consiste en la obtención de muestras de una misma población dinámica a intervalos regulares de tiempo, para estudiar la evolución de la población.

■ Muestreo secuencial: suele utilizarse en el campo industrial en modelos de inspección para aceptar o rechazar lotes de productos, entre otras cosas. No se fija

TEMA DE MUESTRA

de antemano el número de unidades que forman la muestra, sino que se va examinando cada una y a continuación se decide si seguir el muestreo o conformarse con la información ya obtenida para decidir si el lote se acepta o rechaza.

III.2. MUESTREO INTENCIONAL U OPINATIVO

En este tipo de muestreo es la persona que selecciona las muestras la que procura que ésta sea representativa. La representatividad depende, por tanto, de su intención u opinión, y la evaluación de ésta es totalmente subjetiva. La composición de la muestra puede estar bastante influenciada por las preferencias del individuo que las obtiene.

Este muestreo carece, por tanto, de una base teórica consistente. Sin embargo, su uso es bastante frecuente, por ejemplo en el denominado "muestreo por cuotas". Veamos en qué consiste:

Supongamos que el diseño de la encuesta ha seguido los principios del muestreo probabilístico hasta llegar al momento de seleccionar las personas que han de ser entrevistadas. Esta etapa final consiste, en el muestreo por cuotas, en asignar a cada entrevistador un número de entrevistas a personas en un determinado grupo de edad, sexo, nivel económico, zona de residencia, etc. Sujeto a estas restricciones, se deja en libertad al entrevistador para que elija a las personas que cumplen dichos requisitos.

El margen de libertad dado al entrevistador puede introducir sesgos en el proceso de selección que no podrán ser detectados.

El desconocimiento de las probabilidades de selección hace que no puedan estimarse los errores de muestreo.

III.3. MUESTREO SIN NORMA

Aquí se toma la muestra de cualquier manera, por comodidad o a capricho. Si la población es homogénea, la muestra obtenida puede ser representativa. En ocasiones la uniformidad puede ser sustituida por una buena mezcla de elementos antes de tomar muestras (por ejemplo, barajar las cartas, o girar el bombo de las bolas).

TEMA DE MUESTRA

IV. BIBLIOGRAFÍA

■ Cesar Pérez López. **Muestreo estadístico: conceptos y problemas resueltos.** Pearson Educación, 2005.

■ Virtudes Alba Fernández; Nuria Ruiz Fuentes. **Muestreo Estadístico**, Septem Ediciones, 2004.

■ Rodríguez Osuna, J. (1993). **Métodos de muestreo**. Madrid, CIS

TEMA DE MUESTRA