

## **OBTENCIÓN DE DATOS – MUESTREO ALEATORIO**

Cuando se necesita determinar el valor de alguna característica poblacional no siempre es posible ni deseable obtenerlo midiendo todos los individuos de la población. Por ej. si queremos saber la altura promedio de una plantación de pinos, o el peso promedio de los individuos de una población de roedores, no resultaría fácil ni eficiente en términos de tiempo, trabajo y/o dinero, el intentar medir la altura de todos los árboles de la plantación o el peso de todos los roedores de la población. Es por eso que se recurre a las técnicas de muestreo.

Al muestrear estudiamos una parte de la población (conjunto de “n” observaciones) y extrapolamos sus características a la totalidad, debiendo tomar un **número de muestras representativo** a fin de que reflejen adecuadamente las características de la población.

Los muestreos probabilísticos, son aquellos en los que se utiliza algún sistema de selección aleatoria para garantizar que cada unidad de la población tenga una probabilidad específica de ser seleccionada, cumpliéndose así dos requisitos:

1. Toda unidad tiene una probabilidad de ser elegida
2. Esa probabilidad es conocida de antemano.

Las principales técnicas probabilísticas son:

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

### **Muestreo Aleatorio Simple.**

Es el método conceptualmente más sencillo de muestreo; sin embargo es el menos utilizado ya que requiere contar a priori con un listado de todas las unidades de la población accesible (lo que se denomina marco muestral) a partir del cual se seleccionará la muestra. El procedimiento para realizar un muestreo aleatorio simple consiste en:

1. Dispongamos de la lista de todos los elementos de la población y a continuación los numeramos consecutivamente.
2. Mediante la utilización de una tabla de números aleatorios o mediante programas informáticos diseñados al efecto obtendremos aleatoriamente los “n” números que componen nuestra muestra en función del tamaño deseado.
3. Iremos al listado donde tenemos numerados los elementos de la población y tomaremos cada uno de los “n” elementos de la población, que correspondan a los números obtenidos anteriormente de forma aleatoria y esos elementos serán los que constituirán nuestra muestra.

El muestreo aleatorio se trata de un método equiprobabilístico pues todos los elementos tienen una probabilidad de ser elegidos, esa probabilidad es conocida de antemano y además esa probabilidad es la misma para todos los elementos.

Por ejemplo: si deseáramos conocer la incidencia de caries en alumnos de 1º de Educación Polimodal de una Zona Básica de Salud, a la hora de decidir tomar una muestra de esa población podemos optar por un muestreo aleatorio simple. Para ello necesitaremos un listado de los alumnos matriculados en ese curso en los colegios de la zona, posteriormente numeramos consecutivamente a todos los alumnos obteniendo así el marco muestral, es decir el listado de todos los niños de 1º de Educación Polimodal (que en nuestro ejemplo tiene un tamaño de  $N = 200$ ), mediante la utilización de una tabla de números aleatorios, o de programas de informática que nos generan aleatoriamente números, obtendremos “n” números aleatorios entre el 1 y

el 200 (en nuestro ejemplo  $n = 50$ ), iremos posteriormente al listado donde tenemos numerados a los niños matriculados en 1º y seleccionamos los 50 niños que correspondan con los números obtenidos aleatoriamente y ellos constituirán nuestra muestra.

|---o-----o-----o-----o-----o-----o-----o-----|

Los “o” simbolizan las unidades seleccionadas con un muestreo aleatorio simple, y cada “-” una posible unidad a seleccionar.

### Muestreo sistemático.

Consiste en seleccionar a los individuos según una regla o proceso periódico. Para ello en primer lugar se debe calcular la **constante de muestreo** (K) o razón de muestreo, dividiendo el total de la población elegible por el tamaño de la muestra deseado ( $N/n = K$ )

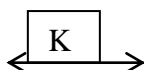
La primera unidad (r) se extrae tomando un número al azar entre 1 y la constante de muestreo (K) y a partir de ahí se va sumando la constante de muestreo consecutivamente hasta completar el tamaño de la muestra, (siendo el primer individuo r, el segundo  $r + K$ , el tercero  $r + 2K$  y así sucesivamente hasta completar los “n” individuos).

Este tipo de muestreo tiene la ventaja de no necesitar tener la lista de la población cerrada de antemano para seleccionar la muestra, por ejemplo si deseáramos conocer como se recogen los hábitos tóxicos en las historias clínicas que se abrirán durante el próximo año, teniendo un cálculo aproximado del total de historias clínicas que se van a abrir podemos calcular la constante de muestreo y el punto de arranque e ir seleccionando las historias clínicas que correspondan para nuestra muestra según se vayan abriendo.

Sin embargo este tipo de muestreo no es aconsejable en las situaciones en que las unidades de muestreo están ordenadas por algún criterio periódico y la constante de muestreo puede coincidir con ello.

Por ejemplo : si queremos conocer la demanda en un centro de salud y elegimos como unidad de muestreo el día de la semana, puede ocurrir que obtengamos como constante de muestreo (K) el número 7, en ese caso estaríamos describiendo siempre la demanda que se produce un día determinado, por ejemplo el sábado, teniendo una elevada probabilidad que la demanda que se produce el sábado sea diferente de la demanda que se produce a lo largo de toda la semana, no siendo por tanto la muestra obtenida representativa de la población que queremos estudiar.

Si en el ejemplo que desarrollamos anteriormente sobre la incidencia de caries en alumnos de 1º de Educación Polimodal de una Zona Básica de Salud optásemos por un muestreo sistemático, necesitaríamos calcular primero la constante de muestreo (si  $N = 200$  y  $n = 50$ ) la constante de muestreo K será 4 ( $200/50 = 4$ ) elegiremos aleatoriamente un número entre 1 y el 4 (supongamos que es el 3,  $r = 3$ ) y ese será nuestro punto de arranque. Tomaremos el listado de los alumnos matriculados en ese curso en los colegios de la zona numerados consecutivamente y el primer niño que tomaremos será el que ocupa el puesto 3 ( $r = 3$ ), el segundo 7 ( $r + K$ , es decir  $3 + 4$ ), el tercero 11 ( $r + 2K$ , es decir  $3 + 2 \times 4$ ), el cuarto 15 ( $r + 3K$  es decir  $3 + 3 \times 4$ ) y así sucesivamente hasta tomar los 50 niños que constituyen nuestra muestra.



|---x-----|---x-----|---x-----|---x-----|---x-----|  
 0            1K            2K            3K            4K

Las “x” simbolizan las unidades seleccionadas con un muestreo sistemático.

**Muestreo Estratificado:** Se divide a la población en estratos, niveles o grupos según criterios prefijados; las unidades incluidas en cada estrato son relativamente Homogéneas con respecto a las características que se vayan a estudiar y los estratos son diferentes unos de otros.

Se toma una sub-muestra a partir de cada estrato mediante un procedimiento aleatorio simple. Para obtener la muestra general se combinan las sub-muestras correspondientes a todos los estratos.

El muestreo estratificado se emplea con frecuencia en el manejo de poblaciones heterogéneas. Un ejemplo son las encuestas de opinión, donde los elementos (personas) son heterogéneos en razón a su sexo, edad, profesión, etc. Interesa en estos casos que la muestra tenga composición análoga a la población y esto se consigue con un muestreo estratificado.

Cuando la muestra se toma asignando una proporción de miembros a cada estrato y escogiendo los elementos dentro de cada estrato por muestreo aleatorio simple se denomina “Muestra estratificada proporcional”.

Algunas veces la razón de muestreo en cada estrato está inversamente relacionada con la homogeneidad de las unidades en el estrato; mientras más homogéneo sea el estrato, menor será la proporción incluida en la muestra. A una muestra obtenida de esta forma se la denomina “muestra estatificada desproporcionada”.

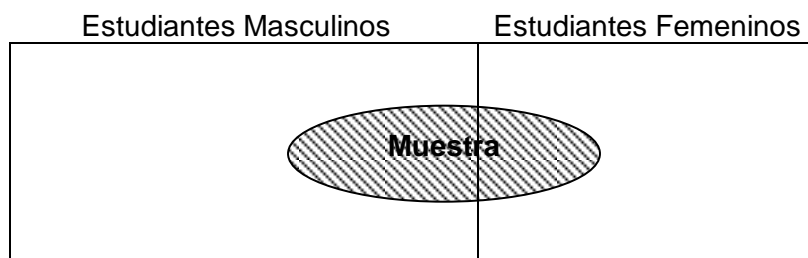
Ejemplo: Supongamos que realizamos un estudio sobre la población de estudiantes de una Universidad, en el que a través de una muestra de 10 de ellos queremos obtener información sobre el uso de lápices de labios.

En primera aproximación lo que procede es hacer un muestreo aleatorio simple, pero en su lugar podemos reflexionar sobre el hecho de que el comportamiento de la población con respecto a este carácter no es homogéneo, y atendiendo a él, podemos dividir a la población en dos estratos:

- Estudiantes masculinos (60% del total);
- Estudiantes femeninos (40% restante).

de modo que se repartan proporcionalmente ambos grupos el número total de muestras, en función de sus respectivos tamaños (6 varones y 4 mujeres). Esto es lo que se denomina asignación proporcional.

Si observamos con más atención, nos encontramos (salvo sorpresas de probabilidad reducida) que el comportamiento de los varones con respecto al carácter que se estudia es muy homogéneo y diferenciado del grupo de las mujeres.



**Muestreo por conglomerados:** requiere elegir una muestra aleatoria simple de unidades heterogéneas entre sí de la población llamadas conglomerados. Cada elemento de la población pertenece exactamente a un conglomerado, y los elementos dentro de cada conglomerado son usualmente heterogéneos o disímiles.

En el muestreo por conglomerados, éstos se forman para representar, tan fielmente como sea posible, a toda la población; entonces se usa una muestra aleatoria simple de conglomerados para estudiarla. Los estudios de instituciones sociales como iglesias, hospitales, escuelas y prisiones se realizan, generalmente, con base en el muestreo por conglomerados.

Las unidades de muestreo suelen ser un grupo de elementos que comúnmente es llamado conglomerado de elementos. El muestreo de estas unidades (grupo de elementos) es llamado “muestreo por conglomerados”.

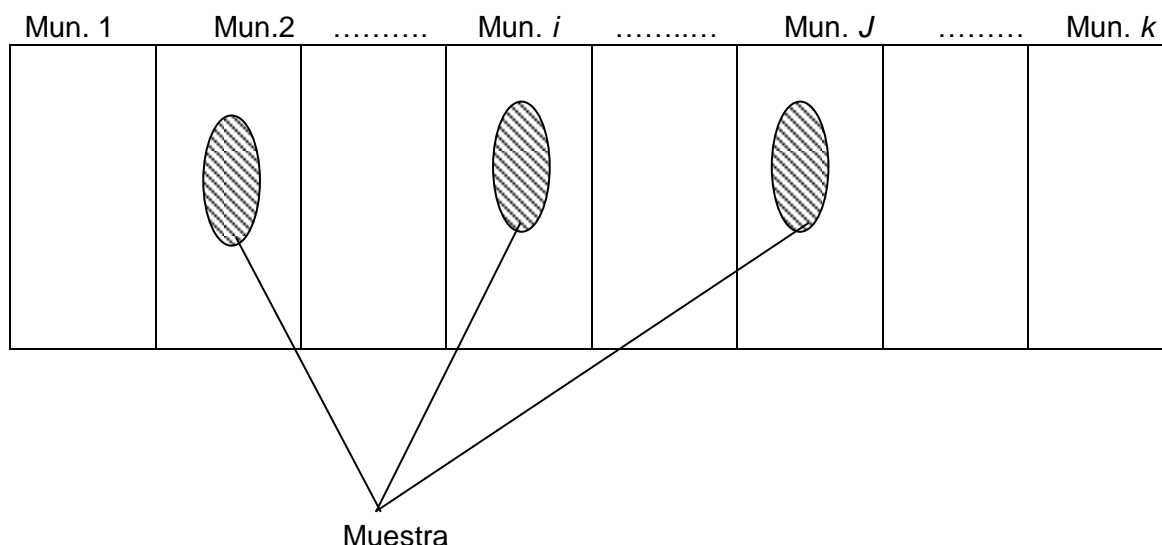
Teóricamente, los conglomerados deben formarse conservando una fuerte heterogeneidad entre sus elementos y homogeneidad entre conglomerados. Des Raj, afirma que un conglomerado es una especie de miniatura de la población.

**Ejemplo:** Se desea estudiar la característica aleatoria: cantidad de errores cometidos en la planilla del último censo. Realizando solamente el estudio en los municipio de las ciudades de la provincia del Chubut.

Se considera que la característica es similar en cada uno de los municipios y que éste comportamiento es análogo a como se comporta en la provincia del Chubut en su conjunto.

En este caso la población de planillas de censo se encuentra dividida en grupos homogéneos entre sí respecto a la característica en estudio.

Primeramente se determinan los municipios a muestrear ( se puede emplear una selección aleatoria simple) y después en cada municipio se toma una muestra aleatoria de planillas del último censo.



### Muestreo por etapas

Una muestra “bietápica” se obtiene seleccionando primero una muestra de UNIDADES y luego una muestra aleatoria de elementos de cada UNIDAD seleccionada.

La unidad de muestreo que forma parte de la selección en la primera etapa recibe el nombre de unidad primaria de muestreo (UPM); la subsiguiente es la unidad secundaria de muestreo (USM) y así sucesivamente si es que se dan más etapas.

La generalización de este muestreo en más etapas se llama “muestreo multietápico o polietápico”.