**MUESTREO**

En muchas ocasiones cuando debemos analizar una población el análisis de todos los individuos resulta poco costo-efectivo. En estos casos resulta indicado realizar un muestreo. Denominamos muestreo a la selección de un grupo de observaciones (muestra) de una población, con el fin de obtener información. Existen dos tipos de muestreo: **el** **probabilístico y el no probabilístico**.

Las **técnicas de muestreo probabilísticas** se caracterizan por tener una probabilidad conocida y mayor a cero de que cada individuo sea incluido en la muestra. Esta probabilidad se calcula dependiendo del diseño muestral seleccionado: *muestreo aleatorio simple, estratificado, sistemático o por conglomerados, polietápico*, entre otros.

En cambio, en las **técnicas de muestreo de tipo no probabilístico** la selección de los sujetos depende de ciertas características, criterios, u otros factores, que el cliente o investigador considere adecuado. Por lo cual, en términos estadísticos, carecen de validez, confiabilidad y reproducibilidad; debido a que este tipo de muestras no dan certeza que cada sujeto a estudio represente a la población objetivo.

En este curso nos enfocamos en datos secundarios (ya recolectados), que se encuentran generalmente en una base de datos, y por consiguiente el problema del muestreo es mucho menos complejo. Sin embargo, resulta interesante revisar un poco este tema y ver cómo se realiza cuando tenemos datos primarios (por recolectar).

**Muestreo probabilístico**

El muestreo probabilístico tiene por objeto estudiar los métodos de selección de muestras representativas de la población con el fin de hacer inferencias sobre la población. La representatividad de una muestra se garantiza por medio de la construcción matemática de la metodología de selección de las unidades experimentales.

**Muestreo aleatorio simple**

En el *muestreo aleatorio simple (M.A.S)* cada elemento de la población tiene la misma probabilidad de ser seleccionado para la muestra. Sin embargo, en el muestreo aleatorio simple *la selección de observaciones puede ser con repetición o sin repetición*. El término **repetición** se refiere a que *una observación puede ser seleccionada más de una vez*; es decir, no es descartada después de ser seleccionada. Por lo contrario, en el muestreo **sin repetición** no se tienen individuos repetidos, puesto que *una observación no puede ser seleccionada más de una vez*.

**Aleatorio sistemático**

Cuando la población está ordenada, ya sea de manera teórica (en una lista) o de manera práctica (en una fila) es posible utilizar el muestreo sistemático. En este tipo de muestreo se recorre la población de manera uniforme y se selecciona cada *k-ésimo* individuo a partir de un individuo aleatorio escogido entre *1* y *k*. Para comprenderlo mejor, veamos el método en detalle.

El método de muestreo sistemático se utiliza en muestras ordenadas del *1* al *N*. Consiste en lo siguiente:

* Se asume que tenemos una población de *N* individuos ordenados del *1* al *N*. Se quiere seleccionar una muestra de tamaño *n*.
* Sea *k* el entero más próximo a *N/n.*
* Se escoge al azar un número *i* entre *1* y *k*.
* La muestra será el elemento *i* y los elementos *i + k*, *i + 2k*, … hasta terminar. Es decir, el elemento *k* y los elementos a intervalos fijos *k* hasta conseguir los n sujetos:

*M = (i, i+k, i+2k, …, i+(n-1) k)*

**Muestreo por conglomerados**

El muestreo por conglomerados consiste en que la población está dividida en grupos de manera natural (barrios, regiones, ciudades, etc.) que se caracterizan por ser muy similares entre sí (intercambiables) y muy diversos en su interior (pueblos, ciudades, escuelas). Estos grupos se denominan conglomerados. Entonces en primera instancia seleccionamos aleatoriamente una muestra de conglomerados y posteriormente seleccionamos unidades experimentales de cada conglomerado.

**Muestreo estratificado**

Al contrario de los conglomerados, *los estratos son grupos de individuos uniformes en su interior, pero muy distintos entre sí (grupos etarios, segmentos de clientes, profesiones).* Por esta razón es necesario utilizar todos los estratos en el proceso de muestreo. De no hacerlo se ignoraría una parte importante de la población.

Como parte de este muestreo se toma una submuestra a partir de cada estrato mediante un muestreo aleatorio simple, de modo que para obtener la muestra general se combinan las submuestras de todos los estratos.

Suponemos que hay *k* estratos de tamaños N1, N2, …, Nk, de forma que:

N = N1 + N2 + … + Nk

En cada estrato se toman n1, n2, …, nk elementos para la muestra, de manera que se toman en total *n* individuos, es decir:

n = n1 + n2 + … + nk

A menudo se toma una razón de muestreo igual para todos los estratos. Es decir que las unidades en la muestra se asignan entre los estratos y en proporción con el número relativo de unidades en cada estrato de la población. A una muestra seleccionada de esta forma se le denomina muestra estratificada proporcional.

**Polietápico**

*El muestreo polietápico o en etapas, se recomienda cuando los tamaños de la muestra de los conglomerados o estratos son muy variables*, *cuando la población resulta muy compleja o cuando el tamaño poblacional es muy grande.* Consiste en realizar varias etapas, cada una con un diseño muestral propio, que puede ser aleatorio simple, por conglomerados o estratificado. En cada etapa se escogen grupos más y más pequeños, hasta llegar a las unidades experimentales.

**Este es el diseño muestral más frecuente para encuestas a nivel nacional, por la complejidad de las poblaciones.**

**Muestreo no probabilístico**

El muestreo no probabilístico se caracteriza por carecer de sustento probabilístico, por lo cual no es posible garantizar muestras representativas. No obstante, constituye una poderosa herramienta de recolección de información. Este tipo de muestreo es utilizado en ciencias sociales, en entornos donde las condiciones no permiten el uso de muestreos probabilísticos, o donde la representatividad no resulta relevante.

**Intencional**

La selección intencional se utiliza en casos donde la población objetivo está constituida por individuos muy poco frecuentes. En este caso se incluye la mayor cantidad de sujetos sin importar el sistema de selección. Al crear un panel de expertos o examinar una enfermedad extraña es necesario conseguir individuos de análisis cuya aleatorización resulta muy costosa.

**Por conveniencia**

La selección de estos casos se debe a la accesibilidad de la inclusión del elemento y proximidad de este sujeto con el investigador. Es útil cuando es necesario disponer de información de manera rápida y poco costosa, sin importar si es generalizable a la población. Cuando una persona nos aborda en la calle para llenar una encuesta es muestreo por conveniencia.

**Muestreo de bola de nieve**

El muestreo de bola de nieve o también conocido como muestreo por referidos se realiza sobre poblaciones en las que no se conoce a sus individuos o es muy difícil acceder a ellos, por lo que cada sujeto estudiado propone a otros, de tal manera que genera un efecto acumulativo de observaciones. Es ampliamente utilizado en políticas públicas y estudios sociales. Este es el sistema más frecuente para encontrar información sobre crimen organizado.

**Muestreo por cuotas**

El muestreo de cuotas, se basa en seleccionar la muestra después de dividir la población en estratos. Los sujetos dentro de cada grupo son elegidos por métodos no probabilísticos. Además, la cantidad de elementos seleccionados se deben a un número arbitrario (cuotas) a partir de los cuales se construye una muestra relativamente proporcional a la población.

**Conclusión**

El muestreo es muy importante debido a que estudiar la población en su totalidad no siempre es óptimo. Es necesario que tengamos muy claro que para cada diseño muestral probabilístico existen distintas fórmulas y ecuaciones que determinan los tamaños muestrales, tanto a nivel de muestra total como de muestras parciales (en cada etapa).En ese sentido es apropiado investigar los tamaños muestrales.