



# UNIVERSIDAD DE MANAGUA

*Al más alto nivel*

## Estadística Inferencial

### Encuentro #7

#### Tema: Muestreo y Técnicas de muestreo



Prof.: MSc. Julio Rito Vargas A.

Grupos: CCEE y ADMVA /2016

## Técnicas de muestreo y determinación del tamaño de la muestra en Investigación

### Introducción

La lectura crítica de estudios de investigación incluye algunos interrogantes referidos a la delimitación de la población, la determinación de la muestra, si el tamaño elegido de la muestra es el adecuado y si el error de muestreo es apropiado. El muestreo es un elemento clave en la metodología de la investigación ya que implica seleccionar a un grupo de elementos que se utilizarán para dirigir un estudio. Por lo tanto es importante diseñar un plan de muestreo que defina el proceso de selección del grupo de elementos seleccionados. Es fundamental expresar claramente en todo trabajo de investigación los siguientes aspectos del diseño metodológico: la población y las características que deben poseer los elementos para formar parte del estudio, el número de elementos que conforman la población, si este número se conoce con certeza o se puede estimar, el tipo de muestreo y la técnica utilizada, el tamaño de la muestra y el error de muestreo establecido.

La teoría del muestreo permite determinar de manera efectiva la muestra que refleje con exactitud las características de la población sometida al estudio, ya que no siempre es posible tomar a todos los elementos que conforman la población.

Aquí se presenta una definición de población amplia que es aplicable a todos los ámbitos de la investigación científica.

El tamaño de la muestra a estudiar y la forma de selección de las unidades muestrales se debe justificar convenientemente de acuerdo al planteo del problema, la población, los objetivos y el propósito de la investigación.

Para acercar las herramientas que brinda la estadística a la investigación es necesario definir algunos conceptos previamente y a partir de ellos desarrollar las diferentes técnicas del muestreo acompañándolas de ejemplos y aplicaciones.

Se presentan las diferentes técnicas de muestreo y la determinación del tamaño de la muestra requerida en un trabajo de investigación.

La información presentada puede ser de utilidad a toda persona que desee realizar un trabajo de investigación científica, una tesina, una tesis, como a todo aquel que oriente el proceso, realice una lectura crítica o evalúe estudios científicos.

## **Definiciones**

### ***Universo***

Serie real o hipotética de elementos que comparten características definidas relacionadas con el problema de la investigación.

### ***Población***

Metodológicamente, en un trabajo de investigación, se denota a la población como un conjunto definido, limitado y accesible del universo que forma el referente para la elección de la muestra. Es el grupo al que se intenta generalizar los resultados del estudio. Comprende todos los elementos (personas, familias, grupos, objetos, organizaciones, etc.) que presentan características comunes que se definen a través de criterios establecidos para el estudio. Se debe definir la población con precisión, de modo que sea manifiesto cuándo cierto elemento pertenece o no a esa población. Para el enfoque cuantitativo, la población debe situarse claramente en torno de características de contenido, lugar y tiempo.

Otros conceptos teóricos que se deben aclarar en este contexto son el de población finita y población infinita. Se refiere a una población finita cuando la población tiene un tamaño establecido y limitado, esto es, existe un número denotado por  $N$  que indica cuántos elementos conforman la población. Por otra parte se refiere a población infinita cuando es teóricamente imposible, ya que se necesita tiempo y recursos, observar todos los elementos. Aunque la población sea excesivamente grande no existe una población infinita de objetos físicos, entonces se considera infinita cuando no se puede enumerar en un tiempo razonable.

### ***Censo***

Algunas veces resulta útil examinar a todos los elementos de la población, en este caso se denomina censo, como la realización de un censo demanda tiempo y recursos humanos, técnicos y económicos, sólo pueden hacerlo las grandes organizaciones como gobiernos, universidades u otros organismos nacionales e internacionales. Ejemplos: Censo de población y vivienda de los habitantes de la República de Nicaragua. Censo nacional Agropecuario. Censo de alumnos de la Universidad de Managua.

### ***Marco muestral***

Es el conjunto de las unidades de análisis delimitadas por los parámetros muestrales, aquellos que permiten realizar una selección asegurando la probabilidad de que todas sean elegidas. Existen los marcos de listas, donde los elementos o unidades muestrales son enumerados uno a uno.

Ejemplos: listado de poseedores de tarjetas de crédito, abonados de una guía telefónica o empleados de una organización.

### ***Marcos de áreas***

Aquí se representan las áreas geográficas en las que se distribuyen los habitantes de una población.

### ***Muestra***

Es una colección de individuos extraídos de la población a partir de algún procedimiento específico para su estudio o medición directa. Una muestra es una fracción o segmento de la población. La muestra es en cierta manera una réplica en miniatura de la población. Se estudian las muestras para describir a las poblaciones, ya que el estudio de muestras es más sencillo que el de la población completa, porque implica menor costo y demanda menos de tiempo.

### ***Muestra representativa***

Para que una muestra sea representativa, tiene que contener las características relevantes de la población en las mismas proporciones en que están incluidas en tal población. Para evaluar la representatividad de la muestra se compara la media muestral con la media poblacional, si este parámetro se desconoce se puede estimar tratando de encontrar las medias obtenidas en trabajos anteriores que han analizado las mismas variables. Siempre hay un grado de error en las estimaciones lo que constituye el **error muestral**, una de las maneras de minimizarlo es diseñando un plan de muestreo adecuado.

### ***Unidad de muestreo***

Un elemento se convierte en unidad de muestreo en el momento en que se encuentra disponible para su selección, en el instante en que es viable su selección práctica como fuente de datos o de información.

## **Tipos de muestreo**

### ***Muestras aleatorias o probabilísticas***

Las unidades de análisis tienen que estar dispuestas de modo tal que el proceso de selección dé una equiprobabilidad de selección a todas y cada una de las unidades que figuran en la población. Cada uno de los elementos de la población tiene la misma probabilidad de integrar parte de la muestra.

### ***Muestras no probabilísticas***

La elección de los elementos no depende de la probabilidad sino de las causas relacionadas con las características de la investigación o de quien establece la muestra.

Aquí el procedimiento no es mecánico, ni con base en fórmulas de probabilidad, sino que depende del proceso de toma de decisiones de una persona o de un grupo de personas y, desde luego, las muestras seleccionadas obedecen a otros criterios de investigación. Desde el enfoque cuantitativo y para determinado diseño, la utilidad de una muestra no probabilística reside no tanto en una "representatividad" de elementos, sino en una cuidadosa y controlada elección de sujetos con ciertas características definidas previamente en el planteamiento del problema.

### ***¿Cuál elegir?***

Elegir una muestra probabilística o no probabilística depende de los objetivos del estudio, del esquema de investigación y del alcance de sus contribuciones.

### **Técnicas de muestreo probabilístico**

Aleatorio simple, estratificado, por conglomerados, polietápico, aleatorio sistemático

#### ***Muestreo aleatorio simple***

Se utiliza cuando se conocen todos los elementos que conforman la población. Cada elemento de la población tiene la misma posibilidad de ser elegido para formar parte de la muestra. Se elabora una lista enumerativa de los elementos. Desde aquí se obtiene una serie de elementos que constituyen la muestra. El proceso consiste en trabajar con una tabla de dígitos al azar o una serie de números aleatorios generados en una planilla de cálculo y seleccionar a los que coinciden con el número al azar y el orden de la lista, o utilizando un programa estadístico que seleccione una cantidad de elementos aleatoriamente de una lista general que es la población.

Ejemplos:

Selección de la muestra (editorial, columna de opinión, etc.) para análisis de contenido.

Selección de la muestra del pasaje que se embarca en aeropuerto de Managua para determinar condiciones de seguridad aérea o motivo del viaje.

Selección de la muestra de asistentes a congreso sobre comunicación interna.

Selección de la muestra de usuarios de un servicio.

Selección de la muestra de pacientes que concurren a una unidad de salud..

#### ***Muestreo estratificado***

Es una variante del muestreo aleatorio simple. Se divide a la población en grupos homogéneos denominados estratos. Los elementos dentro de cada estrato se seleccionan aleatoriamente de acuerdo con una de las siguientes reglas:

- Un número específico de elementos se extrae de cada estrato, y corresponde a la proporción de ese estrato en la población
- Igual número de elementos se extraen de cada estrato, y los resultados son valorados de acuerdo con la porción del estrato de la población total.

Esta técnica de muestreo resulta apropiada cuando la población ya está dividida en grupos, porque refleja de forma más precisa las características de la población y permite efectuar comparaciones entre los estratos conformados.

Ejemplos de estratificación:

Grupos socioeconómicos, categorías laborales, grupos de edad, grupos étnicos, según profesión, según estado civil, según lugar de residencia, según sexo, según turno de trabajo o turno de estudio.

#### ***Muestreo por conglomerados o racimos***

Se utiliza cuando no es posible obtener una lista de todos los elementos de la población. Su empleo es adecuado si la población es muy grande y dispersa. Los conglomerados se caracterizan por ser homogéneos entre sí, pero internamente presentan un alto grado de heterogeneidad en sus componentes. La técnica consiste en lo siguiente: se divide a la población en grupos o racimos, luego se selecciona aleatoriamente algunos de esos grupos, por

considerar que cada uno de ellos es representativo de la población y posteriormente se toma una muestra aleatoria de cada uno de los grupos que se han seleccionado. Este procedimiento produce una muestra más precisa a un menor costo ya que se utiliza cuando hay variación considerable dentro de cada grupo, siendo los grupos similares entre sí. El conglomerado es común en los diseños polietápicos y en las muestras de zona geográfica. Cuando se muestrean conglomerados que contienen números de unidades desiguales, pueden utilizar el muestreo probabilístico proporcional al tamaño para que la probabilidad de selección del conglomerado sea igual a la proporción de unidades que contiene.

Ejemplos donde se puede utilizar el muestreo por conglomerados:

Si la población son escuelas, universidades, hospitales, distritos escolares o zonas geográficas, entonces las unidades muestrales dentro de cada conglomerado pueden ser alumnos, docentes, personal, pacientes o ciudadanos.

### ***Muestreo polietápico***

Muestreo en el que se procede por etapas: se obtiene una muestra de unidades primarias, más amplias que las siguientes; de cada unidad primaria se toman, para una submuestra, unidades secundarias, y así sucesivamente hasta llegar a las unidades últimas o más elementales.

Puede considerarse como una modificación del muestreo por conglomerados. Entonces, no forman parte de la muestra elementos o unidades de todos los conglomerados, sino que, una vez seleccionados los conglomerados aleatoriamente, se efectúan submuestras dentro de cada uno de ellos.

Ejemplo muestreo polietápico:

Múltiples etapas. Polietápico

1º Etapa: muestra de ciudades.

2º Etapa: muestra de familias.

3º Etapa: muestra de individuos.

En una investigación de mercado: el objetivo es determinar por muestreo el número promedio de televisores por hogar, en una ciudad grande.

Aplicación: "La Encuesta Permanente de Hogares (EPH) es una encuesta por muestreo.

Los hogares que serán encuestados son seleccionados de forma aleatoria en dos etapas de selección:

En la primer etapa, dentro de cada conglomerado, se selecciona una cantidad de áreas.

En la segunda etapa se listan todas las viviendas particulares de las áreas seleccionadas, para efectuar a partir de ese listado una selección aleatoria de viviendas. Los hogares que habitan esas viviendas son los hogares a encuestar.

### ***Muestreo sistemático***

Los elementos se seleccionan de la población en un intervalo uniforme que se mide respecto de tiempo, orden o espacio. Se emplea si existe una lista ordenada de los elementos de la población o cuando se sabe cuántos elementos componen esa población.

La técnica consiste en tomar cada  $k$  elementos de una lista que contiene todos los elementos de una población, eligiéndose al azar el primer elemento de la muestra. Para determinar el valor  $k$  se realiza el cociente entre el tamaño de la población  $N$  y el tamaño de la muestra, por ejemplo si se tiene una población de 150 personas y se desea tomar una muestra de 30 entonces

$$k = \frac{N}{n} = \frac{150}{30} = 5,$$

Esto es cada 5 personas se toma una y se puede elegir por cual comenzar.

Aplicación:

El gerente de Relaciones Públicas de una empresa de energía eléctrica ha instrumentado una campaña publicitaria institucional y desea conocer la opinión de sus clientes. Toma sistemáticamente cierta cantidad de personas de su listado de clientes de acuerdo al tamaño del listado total y las llama por teléfono para realizar la encuesta.

### **Técnicas de muestreo no probabilístico**

Por cuotas, opinático o intencional, casual o incidental, por redes o bola de nieve.

#### ***Muestreo por cuotas***

Consiste en formar estratos de la población sobre la base de ciertas características y en procurar que estén representadas en proporciones semejantes a las que existen en la población.

Principales características utilizadas: sexo, edad, ocupación, etc.

Una vez determinada la cuota se eligen los primeros que se encuentran y que cumplen esas características. Este tipo de muestreo tiene como beneficio que se pueden realizar estudios exploratorios rápidos y económicos.

Aplicación:

Una empresa quiere estimar la aceptación del sabor de un nuevo producto de la línea, para lo cual invita a la degustación del producto en un puesto comercial utilizando un muestreo por cuotas.

#### ***Muestreo opinático o intencional***

Es una técnica que se basa en la opinión del investigador para constituir una muestra de sujetos en función de su carácter típico, como en el estudio de casos extremos o marginales, o de los casos típicos. Permite estudiar fenómenos raros o inusitados.

Ejemplo:

Encuesta sobre autocuidado a personas que han recibido trasplante hepático en un hospital determinado.

#### ***Muestreo casual o incidental***

La muestra está 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 en vía pública que se realizan en un día y horario determinado.

#### ***Muestreo por redes (bola de nieve)***

Consiste en localizar a algunos individuos según determinadas características. Se utiliza en poblaciones marginales o de difícil acceso. Se basa en redes sociales, en las amistades. Cuando se encontró el primer representante, éste puede conducir a otro, y éste a un tercero, y así sucesivamente hasta conseguir una muestra suficiente.

Ejemplos:

Consumidores de productos *premium*. Personas que padecen determinada enfermedad.

Miembros de una secta. Adictos que rechazan la instancia de rehabilitación. Mujeres golpeadas. Aficionado a prácticas culturales inusuales.

## **Determinación del tamaño de la muestra en investigación cuantitativa**

Los factores que determinan el tamaño de la muestra son la distribución de la población, el nivel de confianza y el margen de error permitido.

### **Distribución de la población**

La población puede ser homogénea o heterogénea, se puede realizar un estudio piloto para tener alguna medida estadística descriptiva como la media y la desviación estándar (S). Si no se poseen datos previos, se puede adoptar los supuestos de posición conservadora, esto es, considera un universo infinito por lo tanto los valores de la proporción en máxima incertidumbre son  $P=Q=0.50$ , entonces  $S=0.5$ .

### **Nivel de confianza**

Se refiere a la probabilidad de que la estimación efectuada se ajuste a la realidad.

Niveles de confianza: 0.90; 0.95; 0.99. Los valores de la tabla Normal para Z de acuerdo al nivel de confianza son: 1.645; 1.96 y 2.575 respectivamente.

En Ciencias Sociales puede aceptarse el trabajo con un nivel de confianza de 0.95.

Cuanto más grande es el nivel de confianza, mayor será la garantía de que la estimación realizada a través de la muestra se aproxime a la realidad.

### **Error de muestreo permitido**

Para fijar el tamaño de muestra adecuado a cada investigación es preciso determinar el porcentaje de error que estamos dispuestos a tolerar. El error es el porcentaje de incertidumbre, es el riesgo estimado de que la muestra elegida no sea representativa. A medida que se incrementa el tamaño de la muestra, el error muestral tiende a reducirse.

### **Tamaño de la Muestra**

Si se establece una muestra probabilística y se conoce el tamaño de la población, de ahora en adelante denotado por N se procede a determinar por fórmula el tamaño de la muestra adecuado. No siempre se tiene el dato del tamaño de la población y entonces existe otra fórmula para obtenerlo. ¿Cuál es el menor número de unidades muestrales (personas, familias, grupos, organizaciones, etc.) que se necesitan para conformar una muestra (n) que asegure un error de muestreo menor de 0.01, 0.03 o 0.05?

Tamaño de la muestra denotado por n.

### **Muestreo aleatorio simple**

$$n = \frac{S^2}{\frac{\varepsilon^2}{Z^2} + \frac{S^2}{N}}$$

Si se conoce el tamaño de la población:

N = tamaño de la población

n = tamaño necesario de la muestra

Z = margen de confiabilidad o número de unidades de desviación estándar en la distribución normal que producirá un nivel deseado de confianza

S = desviación estándar de la población conocida o estimada a partir de anteriores estudios o de una prueba piloto.

E = error o diferencia máxima entre la media muestral y la media de la población que se está dispuesto a aceptar con un nivel de confianza que se ha definido.

### **Muestreo aleatorio simple**

Si no se conoce el tamaño de la población:

$$n = \frac{Z^2_{\alpha/2} S^2}{\varepsilon^2}$$

n = tamaño necesario de la muestra

Z = margen de confiabilidad o número de unidades de desviación estándar en la distribución normal que producirá un nivel deseado de confianza

S = desviación estándar de la población conocida o estimada a partir de anteriores estudios o de una prueba piloto.

E = error o diferencia máxima entre la media muestral y la media de la población que se está dispuesto a aceptar con un nivel de confianza que se ha definido.

### **Muestreo proporcional**

$$n = \frac{Z^2_{\alpha/2} P Q N}{\varepsilon^2 (N - 1) + Z^2 P Q}$$



n = tamaño necesario de la muestra  
Z = margen de confiabilidad  
P = probabilidad de que el evento ocurra  
Q = probabilidad de que el evento no ocurra  
E = error de estimación  
N = tamaño de la población

### Muestreo estratificado.

Suponiendo que se trabaja con tres estratos A, B y C el tamaño de cada estrato proporcional al tamaño de la población en cada estrato se determina con la siguiente fórmula:

$$n_A = \frac{n(N_A)(S_A)}{(N_A)(S_A) + (N_B)(S_B) + (N_C)(S_C)}$$

n = tamaño total de la muestra  
n<sub>A</sub> = tamaño óptimo de la muestra que se extrae del estrato A  
N<sub>A</sub> = número de elementos en el estrato A  
S<sub>A</sub> = desviación estándar de los elementos del estrato A.  
N<sub>B</sub> = número de elementos en el estrato B  
S<sub>B</sub> = desviación estándar de los elementos del estrato B.  
N<sub>C</sub> = número de elementos en el estrato C  
S<sub>C</sub> = desviación estándar de los elementos del estrato C.

#### Aplicación:

Se realiza un estudio sobre prevalencia de consumo de tabaco en adolescentes con el objetivo de revelar el perfil de consumo de tabaco en adolescentes de la Ciudad de Buenos Aires. La población está conformada por adolescentes escolarizados en el nivel medio en escuelas públicas de la Ciudad de Buenos Aires.

La muestra se determina partiendo de la matrícula de 90240 alumnos registrada en el periodo 2008 en las 141 escuelas de educación media de la Ciudad de Buenos Aires, se tomará una muestra de 1055 alumnos estableciendo un nivel confianza del 95% y un margen de error deseado del 3%. Se seleccionarán aleatoriamente tres escuelas que tengan más de 500 alumnos y se realizará un muestreo estratificado proporcional de acuerdo a la matrícula de cada una y el tamaño de la muestra seleccionado.

Fórmulas utilizadas en diferentes casos: a) si se conoce el tamaño de la población; b) si no se conoce N; c) muestreo proporcional al tamaño muestral:

$$a)n = \frac{S^2}{\frac{\varepsilon^2}{Z^2} + \frac{S^2}{N}} = \frac{0.5^2}{\frac{0.03^2}{1.96^2} + \frac{0.5^2}{90240}} = 1054,64$$

$$b)n = \frac{Z^2_{\alpha/2} S^2}{\varepsilon^2} = \frac{1.96^2 (0.5)^2}{0.03^2} = 1067,11$$

$$c)n = \frac{Z^2_{\alpha/2} PQN}{\varepsilon^2(N-1) + Z^2 PQ} = \frac{1.96^2 (0.5)(0.5)(90240)}{0.03^2(90239) + 1.96^2 (0.5)(0.5)} = 1054,64$$

En ambos casos se trabajó con una dispersión  $S=0.5$ . El tamaño de la muestra se redondea al entero siguiente, para el caso a) y c)  $n=1055$  y para el caso b)  $n=1068$ .

### Conclusión

La investigación constituye una actividad fundamental para los profesionales. Para evitar los sesgos de selección que son los errores referidos al proceso de identificación de la población a estudiar y selección de un espacio muestral donde se realizará el estudio, es necesario recurrir a las técnicas de muestreo para fundamentar y ofrecer verificabilidad al estudio realizado.

### Referencias

Nélida Mónica Cantoni Rabolini – autora del artículo.

### Bibliografía

- Burns N., Grove S. K. *Investigación en enfermería*. Madrid: Elsevier, 2004.  
 Levin R. I., Rubin D. S. *Estadística para administración y economía*. México: Pearson Prentice Hall, 2004.  
 Pérez López, C. *Muestreo estadístico: conceptos y problemas resueltos*. Madrid: Pearson Prentice Hall, 2005.  
 Vieytes, R., *Metodología de la investigación en organizaciones, mercado y sociedad: epistemología y técnicas*. Buenos Aires: Editorial de las Ciencias, 2004.