### CAPÍTULO 3

### 3. TÉCNICAS DE MUESTREO

En el tercer capítulo, se presenta las definiciones estadísticas que serán utilizadas en el desarrollo del trabajo, además se expone la población objeto de estudio, así como: el marco muestral, la muestra piloto y el proceso que se efectuó para establecer el tamaño de la muestra a investigar.

**3.1. Definiciones básicas**

* **Población**

Es una colección finita o infinita de medidas, o el recuento de todas las unidades, que presentan una característica común.

* **Muestra**

Es un subconjunto del universo o de la población, dependiendo de que se haya seleccionado a un grupo de elementos o a un grupo de mediciones.

##### Unidades Estadísticas

##### Unidad de Investigación

Es la unidad mínima que mantiene la integridad de los datos que interesan estudiar y analizar. Es decir, el ente que contiene las partes que se van a analizar.

##### Unidad de análisis

Está definida como el elemento que se examina y del que se busca la información dentro de la unidad de investigación.

##### Unidades de muestreo

Son aquellas que contienen las unidades de análisis de la población y que se utilizarán para confeccionar o seleccionar la muestra. En general, es la selección de los conjuntos que serán tomados en cuenta para la conformar la muestra final en la investigación.

##### Muestreo

#####  Es la técnica empleada para la selección de elementos (unidades de investigación) representativos de la calidad y condiciones medias de un todo que conformarán una muestra.

Esta muestra puede ser: No probabilístico y probabilístico.

##### Muestreo no probabilístico

Es aquel utilizado en forma empírica, es decir, no se efectúa bajo normas probabilísticas de selección, por lo que sus procesos intervienen opiniones y criterios personales del investigador o muestrista o no existe norma bien definida o validada. Normalmente se acude a este tipo de muestreo cuando es difícil enumerar, listar o precisar el universo objeto de estudio o cuando no existen registros de los datos.

##### Muestreo probabilístico

Es cuando se puede determinar de antemano la probabilidad de selección de cada uno de los elementos de la población siendo esta distinta de cero. Este muestreo está basado en la teoría de la aleatoriedad o del azar, en la cual se fundamenta la estadística matemática.

Entre los tipos de muestreo de probabilístico tenemos:

Muestreo Aleatorio Simple

Muestreo Aleatorio Estratificado

Muestreo Sistemático

Muestreo por Conglomerados

##### Marco muestral

Es el proceso de definir y enumerar los elementos sobre los cuales se realizan las inferencias estadísticas en el muestreo probabilístico. Es importante la construcción de un marco muestral lo más perfecto posible a fin de que exista una correspondencia biunívoca entre las unidades muestrales poblacionales y las listas físicas que lo conforman. Entre los factores que contribuyen a distorsionar la calidad de un buen marco muestral están: a) Elementos faltantes, b) Unidades ocultas por estar pareadas con otras, c) Unidades muestrales repetidas y d) Elementos extraños.

* **Tipos de Muestreo**

Existen diferentes tipos de muestreo, a continuación se exponen los utilizados en este trabajo.

Es necesario indicar que **N** es el número de unidades de la población, y **n** es el número de unidades de la muestra.

**Muestreo Aleatorio Simple**

Es un procedimiento en donde al tomar un elemento para la muestra, cada uno de ellos tiene la misma probabilidad de ser elegido. Una muestra obtenida por este procedimiento se llama muestra aleatoria simple.

Uno de los métodos más utilizados para lograr que la muestra, de una población finita, sea aleatoria consiste en enumerar todos los N elementos y una vez fijado el tamaño n de la muestra se toman al azar n números.

**Muestreo Aleatorio Estratificado**

En este tipo de muestreo una población heterogénea con N unidades  se subdivide en L subpoblaciones lo más homogéneas posibles no solapadas denominadas *estratos*  de tamaños N1,N2....NL. La muestra estratificada de tamaño n se obtiene seleccionando nh elementos (h=1,2...L) de cada uno de los L estratos en que se subdivide la población de forma independiente. Si la muestra estratificada se obtiene seleccionando una muestra aleatoria simple en cada estrato de forma independiente, el muestreo se denomina muestreo aleatorio estratificado.

Los estratos pueden ser de igual o diferente tamaño, en este último caso, una manera posible de determinar el tamaño de la muestra al interior de cada estrato, es proporcional al tamaño de la misma. Este tipo de asignación se conoce como afijación proporcional, la que no siempre resulta ser la afijación más conveniente por cuanto los costos de muestreo en cada uno de los estratos pueden ser distintos; otro tipo es dependiendo de la variabilidad que presente la población.

**Muestreo por Conglomerados**

Consideramos una población finita con M unidades elementales ó últimas agrupadas en N unidades mayores llamadas conglomerados ó unidades primarias, de tal forma que no existan solapamientos entre los conglomerados y que éstos contengan en todo caso a la población en estudio. Consideramos como unidad de muestreo el conglomerado y extraemos de la población una muestra de n conglomerados a partir de la cual estimaremos los parámetros poblacionales.

El número de unidades elementales de un conglomerado se denomina tamaño del conglomerado. Los conglomerados pueden ser de igual ó de distinto tamaño y han de ser lo más heterogéneos posible dentro de ellos y lo más homogéneos posibles entre ellos, de tal forma que la situación ideal sería que un solo conglomerado pueda representar fielmente a la población.

Los casos frecuentes en los que se utiliza muestro por conglomerados son: la selección aleatoria de una población para efectuar un estudio de individuos dentro de ellas, la selección de granjas de una comarca para una investigación en que las unidades últimas fuesen cabezas de ganado, la selección de árboles ó matas de una plantación donde las unidades últimas fuesen los frutos.

* Se presenta a continuación algunos de los componentes que son comunes a las diferentes fórmulas, para calcular el tamaño de la muestra, dependiendo del método que se aplique.

**Grado de confianza**

Es fijado por el investigador de acuerdo con su experiencia y conocimiento que tenga de la población que va a investigar. Sin embargo, por lo general se trabaja con el 95% o 95.5%, correspondiendo un valor de Z=1.96 y Z=2.00, respectivamente.

**Grado de variabilidad**

Está dado por la varianza. Entre más variabilidad presente la característica, mayor será el tamaño de la muestra, necesario para que represente a la población.

**Error de muestreo**

Simbolizado por e o E, es determinado por el investigador teniendo en cuenta que a mayor error menor será el tamaño y, al contrario, a menor error mayor será el tamaño de la muestra.

**3.2. Población objetivo**

La población es un conjunto de datos que representa el centro del interés a investigar, el cual para este estudio son todas las personas que estén cursando el último año de los colegios fiscales y particulares ubicados en el sector urbano de la ciudad antes mencionada.

Es importante conocer el número de instituciones de educación media, ya que contienen las unidades de exploración. En la Tabla XVI, se puede apreciar el número de establecimientos que poseen el tercer año de diversificado, el cual es objeto de estudio.

#### TABLA XVI

COLEGIOS FISCALES Y PARTICULARES CON TERCER AÑO DE DIVERSIFICADO

|  |  |
| --- | --- |
| **CICLO** | **COLEGIOS** |
|  | **FISCALES** | **PARTICULARES** |
| Básico y Diversificado | 79 | 187 |
| Básico, Diversificado y Post Bachillerato | 6 | 0 |
| Básico, Diversificado y Post Básico | 2 | 1 |
| Diversificado | 3 | 0 |
| Diversificado y Post Bachillerato | 1 | 0 |
| TOTAL | 91 | 188 |

***Fuente****: Dirección Provincial del Guayas (Departamento de Estadística).*

Las instituciones ofrecen diferentes jornadas en el día, por lo que en realidad los elementos se encuentran repartidos en la siguiente cantidad de establecimientos, que podrán ser observados en la Tabla XVII, con el respectivo número de alumnos.

#### TABLA XVII

**NÚMERO DE INSTITUCIONES Y ESTUDIANTES DISTRIBUIDOS POR JORNADAS**

|  |  |  |
| --- | --- | --- |
|  | **FISCALES** | **PARTICULARES** |
| **Jornada de estudio** | **Total de Colegios** | **Total de Alumnos** | **Total de Colegios** | **Total de Alumnos** |
| Matutina | 47 | 7510 | 160 | 7565 |
| Vespertina | 33 | 2544 | 25 | 1194 |
| Nocturna | 36 | 1163 | 53 | 869 |
| Total | 116 | 11217 | 238 | 9628 |

***Fuente****: Dirección Provincial del Guayas (Departamento de Estadística).*

**3.3. Marco muestral**

El Marco Muestral, es una representación simbólica de la población, la cual para este estudio, es el listado de establecimientos educativos que posean el último año de diversificado y que se encuentren localizados en la zona urbana de la ciudad de Guayaquil.

En el ANEXO II, podrá visualizar el nombre, ciclo y jornada de los colegios fiscales y particulares (año lectivo 2001-2002), cuya información fue proporcionada por la Dirección Provincial del Guayas (Departamento de Estadística).

**3.4. Muestra Piloto**

La muestra piloto, se usa para obtener las estimaciones de determinadas características poblaciones que pueden utilizarse posteriormente en cálculos sobre el tamaño de la muestra.

Para obtener esta muestra se tuvo que escoger la variable que sea de mayor interés y calcular la varianza de sus observaciones. El tipo de muestreo a utilizarse en esta parte es *aleatorio simple*.

La muestra piloto se realizó a 150 estudiantes pertenecientes a colegios particulares y fiscales, distribuidos en las diversas jornadas, la variable que se consideró es la siguiente: Posee información necesaria sobre las carreras universitarias que se ofertan en el medio; cuyas respuestas pueden ser: 1. Total Desacuerdo, 2. Parcial Desacuerdo, 3. Indiferente, 4.Parcial Acuerdo, 5. Total Acuerdo. Las dos últimas opciones han sido tomadas como (proporción de estudiantes que poseen conocimiento sobre las carreras universitarias que se ofrecen) y las restantes  (proporción de alumnos que no poseen conocimiento); obteniendo como resultado para las proporciones estimadas los siguientes valores: =0.67 y =0.33.

**3.5. Determinación del tamaño de la muestra**

Para obtener el tamaño de la muestra mediante proporciones, es necesario conocer:

* Error máximo admisible (error de diseño)=0.04.
* Nivel de Confianza: (1-α)100% con α=0.05, se tiene entonces 95% de confianza.
* Zα/2=1.96, cuyo valor es obtenido de la tabla de Distribución Normal.
* Tamaño de la población, N = 20845
* Los valores de las estimaciones de =0.67 y =0.33 (obtenidos en la muestra piloto); así como la estimación del valor de la varianza (), la cual es: 0.2211.

Después de establecer estos valores, se utiliza el muestreo antes mencionado*, aleatorio simple,* para conocer el tamaño de la muestra n, para indicar este valor se debe calcular:





Al reemplazar los valores en las fórmulas antes descritas se obtiene lo siguiente:







Por lo tanto se tiene, que el tamaño de la muestra resultante para esta investigación es n = 518 estudiantes.

El total de la muestra obtenida se lo distribuye en los sostenimientos de los colegios particulares y fiscales, para esto se utiliza el *muestreo aleatorio estratificado* en el cual, las unidades de investigación que se encuentran dentro de cada estrato presentan características homogéneas, mientras que los que se sitúan entre ellos son heterogéneos.

Se tendrá dos estratos, (los cuales se llamaran PARTICULARES Y FISCALES); el primero representará a las instituciones particulares y el otro a los establecimientos fiscales que poseen el tercer año de diversificado, para determinar el número de unidades de cada estrato que serán seleccionadas aleatoriamente, y se utiliza la afijación proporcional*.*

En la Tabla XVIII, se muestra la cantidad de alumnos distribuidos proporcionadamente para los diferentes estratos, donde: n=518 (cantidad obtenida en el muestreo aleatorio simple) y N=20845 total de estudiantes.

#### TABLA XVIII

**TAMAÑO DE MUESTRA DISTRIBUIDO POR ESTRATO**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **h** | **ESTRATO** | **Total de Alumnos (Nh)** | **Peso del estrato****Wh=(Nh/N)**  | **Tamaño de la muestra** **nh=(n\* Wh)** |
| 1 | Fiscales | 11217 | 0.54 | 279 |
| 2 | Particulares  | 9628 | 0.46 | 239 |

La Tabla XIX presenta el total de alumnos que deben ser extraídos de cada tipo de colegio (particulares y fiscales), para lo cual se utiliza nuevamente la afijación proporcional; donde NF=11217, NP=9628 y el valor de n para colegios fiscales es 279, mientras que para los particulares es 239.

#### TABLA XIX

**TAMAÑO DE MUESTRA**

**DISTRIBUIDO POR JORNADAS**

**a.- Colegios Fiscales**

|  |  |  |  |
| --- | --- | --- | --- |
| **Jornada de estudio** | **Total de Alumnos** | **Peso del estrato****Wh = Nh/NF**  | **Tamaño de la muestra** **nh = n\* Wh** |
| Matutina | 7510 | 0.67 | 187 |
| Vespertina | 2544 | 0.23 | 63 |
| Nocturna | 1163 | 0.10 | 29 |
| Total | 11217 |  | 279 |

**b.- Colegios Particulares**

|  |  |  |  |
| --- | --- | --- | --- |
| **Jornada de estudio** | **Total de Alumnos** | **Peso del estrato****Wh = Nh/NP**  | **Tamaño de la muestra** **nh = n\* Wh** |
| Matutina | 7565 | 0.79 | 188 |
| Vespertina | 1194 | 0.12 | 30 |
| Nocturna | 869 | 0.09 | 21 |
| Total | 9628 |  | 239 |

A continuación se determina el número de instituciones de los cuales se extraerán los entes de investigación, para eso se utiliza el muestreo por conglomerados. Los establecimientos educativos muestran las unidades de estudio con características heterogéneas dentro de estos y homogéneas fuera de los mismos, estos grupos reciben el nombre de conglomerados.

Para determinar la cantidad de conglomerados (colegios), se utiliza la siguiente expresión:



donde:

* = Tamaño de la muestra en muestreo por conglomerados.
* = Tamaño de muestra necesario en muestreo aleatorio simple para obtener una precisión dada.
* = Coeficiente de correlación intraconglomerados.
* = Número de unidades elementales por conglomerado.
* La cantidad  por la que hay que multiplicar el tamaño de muestra por conglomerado  para que coincida con el tamaño de muestra necesario en muestreo aleatorio simple  para igual precisión en ambos tipos de muestreo, se denomina *Efecto del diseño*.

Al tratarse de dos estratos (particulares y fiscales), se obtendrá para cada uno de ellos el tamaño de conglomerado, para lo cual es necesario conocer los siguientes valores:

* El coeficiente de correlación intraconglomerados para cada estrato es: Particulares: =-0.0209 y para Fiscales: =-0.0120. Como es posible observar los dos valores de coeficiente son negativos lo cual nos indica que la precisión del muestreo por conglomerados es superior a la del muestreo aleatorio simple.
* Para calcular , se dividió el número de estudiantes para el número de colegios, esto se realizó por que cada institución no contaba con la misma cantidad de alumnos en el último año de diversificado. Los resultados obtenidos fueron: Particulares: =47 y Fiscales =83.
* El valor de  para particulares es 239, mientras que para el estrato de fiscales es 279.

Con los valores mencionados es posible calcular el tamaño de conglomerado, obteniendo lo siguiente:

*Particulares:* = 9.225

*Fiscales:* =4.5

La cantidad de conglomerados requeridos en el estrato Particulares es 9 colegios, mientras que en los Fiscales es 5 instituciones. En la Tabla XX, se observa el número de instituciones escogidos por cada jornada de estudio.

#### TABLA XX

**NÚMERO DE COLEGIOS**

 **DISTRIBUIDOS POR JORNADA**

|  |  |  |
| --- | --- | --- |
| **Jornada de estudio** | **FISCALES** | **PARTICULARES** |
| **Total de Colegios** | **Total de Colegios** |
| Matutina | 2 | 6 |
| Vespertina | 1 | 1 |
| Nocturna | 2 | 2 |
| Total | 5 | 9 |

Las personas que cursan el último año de colegio, son las unidades de investigación a quienes se les realizará el cuestionario, el cual será efectuado en las aulas de clase de las instituciones donde estas estudien.

Para seleccionar los colegios que forman parte de la muestra, se tuvo que enumerar en serie a cada institución que forma parte del marco muestral, tanto para los establecimientos fiscales como particulares; para luego generar números aleatorios, por medio de la función random, y así obtener los establecimientos educativos en los cuales, se debe investigar.

En la Tabla XXI, se podrá ver el nombre de las instituciones con el respectivo número de estudiantes, que fueron seleccionados aleatoriamente.

#### TABLA XXI

**NOMBRE DE INSTITUCIONES Y NÚMERO DE ALUMNOS SELECCIONADOS PARA LA MUESTRA**

|  |  |
| --- | --- |
| **FISCALES** | **PARTICULARES** |
| **Nombre de la Institución**  | **Número de estudiantes** | **Nombre de la Institución**  | **Número de estudiantes** |
| Vicente Rocafuerte | 99 | María Auxiliadora | 27 |
| Guayaquil (I.T.S.) | 88 | República del Ecuador | 27 |
| Ati II Pillahuaso | 63 | San Agustín | 64 |
| Juan de Dios Martínez Mera | 19 | Academia Naval Almirante Illingworth | 36 |
| Ana Villamil Icaza | 10 | Vicente León | 19 |
|  |  | Gloria Gorelik | 15 |
|  |  | Mercantil | 30 |
|  |  | Seis de Marzo | 12 |
|  |  | Trece de Abril | 9 |