CAPÍTULO 3

1. TEORÍA DE MUESTREO
	1. **INTRODUCCIÓN**

En este capítulo se describirá el diseño muestral utilizado, acompañado de algunos conceptos básicos sobre muestreo estadístico que faciliten al lector la mejor comprensión de este capítulo.

* 1. **DISEÑO MUESTRAL**

Para los propósitos de seleccionar una muestra debe ser posible dividir la población en lo que se denomina unidades de muestreo de forma no ambigua. Esto implica que todo elemento de la población debe pertenecer a una sola unidad de muestreo[[1]](#footnote-2)(2). Esta no es una tarea fácil, debido a que siempre hay casos dudosos y deben establecerse algunas reglas arbitrarias para manejarlos.

Una vez que las unidades de muestreo han sido clasificadas sin ambigüedades, los problemas técnicos que recibirán la más cuidadosa atención serán la forma en que se seleccionará la muestra requerida y la estimación de las características de la población y el posterior análisis estadístico.

Para escoger la forma en que se seleccionará la muestra existen varias técnicas de muestreo que se pudieran utilizar, los cuales dependerán de las características de la población de la que provienen los datos a investigar. En base a esto, es que para el presente estudio se ha diseñado las etapas de muestreo utilizando: Muestreo aleatorio simple y muestreo Estratificado, que es el método que más se ajusta a las necesidades e intereses de la empresa proveedora de servicios de telefonía celular, cada uno de estos tipos de muestreo se definirán a continuación.

* + 1. **MUESTREO ALEATORIO SIMPLE**

Es un procedimiento de selección con probabilidades iguales que consiste en obtener la muestra unidad a unidad de forma aleatoria sin reposición a la población de las unidades previamente seleccionadas, teniendo presente además que el orden de clasificación de los elementos en las muestras no interviene, es decir, muestras con los mismos elementos colocados en orden distinto se consideran iguales. De esta forma las muestras con elementos repetidos son imposibles. Como el procedimiento de selección es con probabilidades iguales todas las muestras son equiprobables, asimismo se cumple que todos las unidades de la población van a tener la misma probabilidad de pertenecer a la muestra. Si se supone en todo momento que el tamaño de la población es N y el tamaño de la muestra es n. Como la muestra se selecciona sin reposición, se realiza la selección sucesiva de las unidades para la muestra con probabilidades para valores de t = 0,1,...,n. En el caso particular del presente estudio las unidades de muestreo son seleccionadas de una población objetivo que está constituida por los clientes de telefonía celular de PORTA en la ciudad de Guayaquil.

* + - 1. **Determinación del tamaño de la muestra para un error relativo dado**

El error relativo de muestreo prefijado es igual a:



y dada las características del cuestionario que se va a aplicar a los clientes, el mismo que está compuesto de variables cualitativas, el error relativo se expresará en función de la proporción poblacional *P*, quedando de la siguiente forma:



donde  es la varianza de *P* en términos de la cuasivarianza de la población *S2*. Por tanto, para determinar el tamaño de la muestra *n* es necesario conocer el tamaño de la población *N*, la cuasivarianza *S2*, la cual es estimada a través de *s2* utilizando la muestra piloto, debe considerarse además el error relativo *er*, que dependerá de la precisión que requiera establecer.

Desarrollando la expresión correspondiente al error relativo *er*, se tiene:



SI, entonces



Si se supone que, donde:

 , y



Entonces se obtiene la expresión final que determina el tamaño de la muestra:



* + - 1. **Ejemplo ilustrativo**

Con el propósito de ilustrar un poco más claramente lo antes explicado se presenta el siguiente ejercicio cuyo enunciado es como sigue:

**Ejercicio:** Con el objeto del análisis de la divisibilidad de un conjunto de números consideramos la población virtual *X*i *= {4, 13, 17, 23, 6, 1}, i= 1...6,* mediante muestreo irrestricto aleatorio se extraen muestras de tamaño 2: a) Se trata de estimar el parámetro poblacional PROPORCIÓN DE NÚMEROS PRIMOS, mediante su estimador insesgado basado en las muestras del espacio muestral. Hallar la distribución en el muestreo de dicho estimador y de la estimación insesgada de su varianza. Comprobar que se cumple la relación:



Además se deberá hallar el tamaño de la muestra necesario para que el error relativo de muestreo sea del 2%.

**Desarrollo:**

Al tratarse de muestreo aleatorio irrestricto, estamos ante muestreo aleatorio simple sin reposición, donde el orden de colocación de los elementos en las muestras de tamaño 2 no interviene, con lo que se obtienen

Por otra parte en este problema se considera la clase A de los números primos, con lo que se asociará a los *Xi* los *Ai* que valen cero cuando *Xi* no es primo y valen uno cuando *Xi* es primo. Las probabilidades iniciales de selección de los elementos de la población para la muestra valdrán 1/6 y las probabilidades de las muestras serán todas iguales a 1/15. Las muestras resultantes posibles, los estadísticos de orden y la proporción se muestran en la Tabla 2. Para este ejemplo el tamaño de la muestra es *n=2* y el de la población *N=6*.

La varianza del estimador de la proporción también puede calcularse mediante:

valor que coincide con el que se obtiene a partir de la distribución en el muestreo que se muestra en la Tabla 2.

|  |
| --- |
| ***TABLA 2***Valores de estadísticos de orden, proporción muestral de cada una de las muestras de tamaño n = 2 |
|

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Muestras****n = 2** | **X(1)** | **X(2)** | **A1** | **A2** | **Pij** | **P** | **Cuasi****varianza** |
| 2 | 13 | 4 | 13 | 0 | 1 | 0.16667 | 0.5 | 0.5 |
| 2 | 17 | 4 | 17 | 0 | 1 | 0.16667 | 0.5 | 0.5 |
| 2 | 23 | 4 | 23 | 0 | 1 | 0.16667 | 0.5 | 0.5 |
| 2 | 6 | 4 | 6 | 0 | 0 | 0.16667 | 0.0 | 0.0 |
| 2 | 1 | 4 | 1 | 0 | 1 | 0.16667 | 0.5 | 0.5 |
| 13 | 17 | 13 | 17 | 1 | 1 | 0.16667 | 1.0 | 0.0 |
| 13 | 23 | 13 | 23 | 1 | 1 | 0.16667 | 1.0 | 0.0 |
| 13 | 6 | 13 | 6 | 1 | 0 | 0.16667 | 0.5 | 0.5 |
| 13 | 1 | 13 | 1 | 1 | 1 | 0.16667 | 1.0 | 0.0 |
| 17 | 23 | 17 | 23 | 1 | 1 | 0.16667 | 1.0 | 0.0 |
| 17 | 6 | 17 | 6 | 1 | 0 | 0.16667 | 0.5 | 0.5 |
| 17 | 1 | 17 | 1 | 1 | 1 | 0.16667 | 1.0 | 0.0 |
| 23 | 6 | 23 | 6 | 1 | 0 | 0.16667 | 0.5 | 0.5 |
| 23 | 1 | 23 | 1 | 1 | 1 | 0.16667 | 1.0 | 0.0 |
| 6 | 1 | 6 | 1 | 0 | 1 | 0.16667 | 0.5 | 0.5 |

 |
| ***Fuente:*** *Pérez César, Técnicas de Muestreo Estadístico****Elaboración:*** *José L. Mora M.* |

Para hallar el tamaño de muestra necesario para estimar la proporción con un error relativo de muestreo igual al 2% (*er*=0,02), se procede de la siguiente manera:



lo cual indica que se tomará como tamaño de muestra necesario *n=6*, es decir, se necesitará tomar como muestra toda la población, esto se da porque el error fijado es demasiado pequeño.

* + - 1. **Tamaño de la muestra mediante Muestreo Aleatorio Simple**

El tamaño *n* de la muestra se lo obtuvo mediante el software SAMPLE el cual proporciona tamaños de muestra que pueden tomarse para un error relativo deseado a partir de poblaciones de un tamaño dado. Para nuestro caso en particular el tamaño de la población objetivo que se obtuvo de parte de PORTA fue *N = 558000* clientes (ver Tabla 2), y el resultado que se obtuvo al someterlo al software SAMPLE es el que se muestra a continuación en la Tabla 3a.

|  |
| --- |
| TABLA 3aIntervalos de tamaño de muestra para la población de N=558.000 clientes |
|

|  |  |  |
| --- | --- | --- |
| **Niveles** | **Tamaño de muestra** | **Precisión Relativa** |
| Mínimo | 96 | 10% |
| Sugerido | 300 | 10% |
| Máximo | 384 | 5% |
| **TOTAL(N)** | **558.000** |  |

 |
|  *Fuente y Elaboración: José L. Mora, mediante software SAMPLE* |

Debido a que el objetivo principal del presente estudio no es la aplicación de técnicas de muestreo en sí, no se realizará el cálculo del tamaño de muestra sino más bien, se realizará una comparación de la precisión relativa escogida frente a la precisión relativa obtenida de los datos recolectados por medio de las encuestas, lo que servirá para comprobar que el tamaño de muestra con el error relativo de muestreo seleccionados está dentro de los parámetros correctos.

**Prueba del Tamaño de muestra mediante el error o precisión relativa.-** Esta comprobación se la realiza a partir de los datos recolectados, para lo que se escogen las variables que se someterán a esta prueba. Como el parámetro poblacional estimado es la proporción P, las variables que deben escogerse para la prueba son aquellas que cumplan con esta característica y que sean comunes entre los tres estratos; siendo escogidas las siguientes variables:

**Tarifarios:**

*X8:*¿Ha visitado alguna vez un Centro de Atención al Cliente?

*X18:*¿Conoce usted, el Sistema \*611, para Servicio al Cliente?

*X47:*¿Estaría usted dispuesto a Cambiarse a la Competencia?

**Autocontrol:**

*Y8:* ¿Ha visitado alguna vez un Centro de Atención al Cliente?

*Y18:* ¿Conoce usted, el Sistema \*611, para Servicio al Cliente?

*Y47:* ¿Estaría usted dispuesto a Cambiarse a la Competencia?

**Prepago:**

*Z8:* ¿Ha visitado alguna vez un Centro de Atención al Cliente?

*Z18:* ¿Conoce usted, el Sistema \*611, para Servicio al Cliente?

*Z39:* ¿Estaría usted dispuesto a Cambiarse a la Competencia?

Como estas variables son comunes entre si, se agrupan en una sola quedando de la siguiente forma:

P8: (*X8, Y8,Z8)*

P18: *X18, Y18, Z18)*

P47: *X47, Y47, , Z39)*

Los valores obtenidos con estas variables son los que se muestran a continuación:

|  |
| --- |
| TABLA 3bProporciones obtenidas para las variables seleccionadas |
|

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variables** | n | **Proporción****(P)** | **Std. Deviation****(S)** | **S2** | **P2** |
| P8 | 300 | 0,73 | 0,44 | 0,194 | 0,533 |
| P18 | 300 | 0,78 | 0,42 | 0,176 | 0,608 |
| P47 | 300 | 0,38 | 0,49 | 0,240 | 0,144 |

 |
|  *Fuente y Elaboración: José L. Mora* |

Luego, se someten estos valores a las fórmulas empleadas para calcular *er*, que tiene la siguiente forma:





Así tenemos los resultados que se muestran a continuación en la Tabla 3c

|  |
| --- |
| TABLA 3cError relativo obtenido de las variables seleccionadas |
|

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variables** | n | **Proporción****(P)** | ***er2*** | ***er*** |
| P8 | 300 | 0,73 | 0,0012 | 0,0348 |
| P18 | 300 | 0,78 | 0,0010 | 0,0311 |
| P47 | 300 | 0,38 | 0,0055 | 0,0744 |

 |
|  *Fuente y Elaboración: José L. Mora* |

Como se pudo apreciar (ver Tabla 3b) la variable que se debía seleccionar de las tres previamente escogidas debía ser *P47* por ser la de mayor varianza, de la que se tiene que el error relativo resultante es :

****

* + 1. **MUESTREO ESTRATIFICADO**

El muestreo estratificado es un método mediante el cual se divide la población en subpoblaciones o estratos, dentro de cada uno de los cuales se hace una selección aleatoria con el objeto de mejorar las estimaciones mediante el previo agrupamiento de los elementos más parecidos entre sí.

En el muestreo estratificado se supone una población heterogénea con N unidades  la misma que se subdivide en *L* subpoblaciones lo más homogéneas posibles no solapadas que se denominan 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, lo que puede expresarse gráficamente de la siguiente forma:

Población



Muestra



* + - 1. **Afijación de la Muestra**

Se llama afijación de la muestra al reparto, asignación adjudicación, adscripción o distribución del tamaño muestral *n* entre los diferentes estratos. Esto es, a la determinación de los valores *nh* que verifiquen que . Pueden establecerse muchas afijaciones, pero para los efectos del estudio en mención se utilizarán: afijación proporcional y afijación uniforme, con el propósito de distribuir los 300 clientes que constituyen la muestra seleccionada entre los 3 estratos que se han obtenido (ver Tabla 4). Luego, de las afijaciones resultantes se seleccionará la más adecuada y que mejor se ajuste a las necesidades de la empresa proveedora del servicio de telefonía celular, PORTA.

* + - * 1. **Afijación Proporcional**

Método de distribución del tamaño n de la muestra, que consiste es asignar a cada estrato un número de unidades muestrales proporcional a su tamaño. Las n unidades de la muestra se distribuyen proporcionalmente a los tamaños de los estratos expresados en número de unidades, de la siguiente forma:



donde *nh* es el tamaño de muestra para cada estrato, *Nh* es el tamaño de cada estrato, *N* es el tamaño de la población y *Wh* es la proporción para cada estrato el mismo que se multiplica por *n* que es el tamaño de la muestra, con lo que se obtiene la afijación que corresponderá a cada estrato.

Para este tipo de afijación, la varianza del estimador de la proporción muestral será:



* + - * 1. **Afijación Uniforme**

Este método de distribución del tamaño *n* de la muestra, consiste en asignar a cada estrato el mismo número de unidades muestrales, con lo que se tomará todos los n*h* iguales a *n/L*, aumentando o disminuyendo este tamaño en una unidad si *n* no fuese múltiplo de *L*, como se muestra a continuación:



donde E denota la parte entera.

La afijación uniforme aplicada a cada estrato se realiza de la siguiente manera:



La varianza del estimador de la proporción muestral para este tipo de afijación, será:



Con este tipo de afijación se da la misma importancia a todos los estratos, en cuanto a tamaño de la muestra, con lo cual se favorece a los estratos de menor tamaño y se perjudica a los grandes en cuanto a precisión.

* + - 1. **Tamaño de la muestra mediante Muestreo Estratificado**

Como complemento a lo anteriormente expuesto, se procede a determinar el tamaño n de la muestra estratificada para el tipo de afijación proporcional, para lo cual se parte de un error de muestreo dado, de la siguiente forma:



* 1. **DESARROLLO Y APLICACIÓN DE LAS TÉCNICAS DE MUESTREO**

Revisados los métodos de cálculo del tamaño de muestra mediante muestro estratificado, el siguiente paso será seleccionar la afijación del tamaño n de la muestra que mejor se ajuste a las necesidades de la empresa proveedora de los servicios de telefonía celular, PORTA.

Para esto se partirá identificando cuales son los estratos que se pueden establecer dentro de la población de estudio, la misma que se describe en la Tabla 4. Luego se fijará como tamaño de la muestra el sugerido del rango obtenido mediante software (ver Tabla 3a), el cual es *n= 300* clientes. Estos 300 clientes serán sometidos a las afijaciones antes explicadas entre los estratos que se muestran en la Tabla 5, donde también se observan los tamaños de muestra *nh* para cada estrato de la población los cuales se obtuvieron mediante el proceso de afijación proporcional.

|  |
| --- |
| *TABLA 4*Segmentación de Clientes a Nivel Nacional y Regional |
|

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Segmentación | **Nacional** | **GYE** | **Proporción****(Wh\*100%)** | **UIO** | **Proporción****(Wh\*100%)** |
| Tarifario | 45.563 | 28.249 | 5,06% | 17.314 | 5,06% |
| Autocontrol | 55.689 | 34.527 | 6,19% | 21.162 | 6,19% |
| Prepago | 798.748 | 495.224 | 88,75% | 303.524 | 88,75% |
| **TOTAL** | 900.000 | **558.000** | 100,0% | **342.000** | 100,0% |

 |
| ***Fuente:*** *CONECEL S.A.****Elaboración:*** *José L. Mora M.* |

|  |
| --- |
| *TABLA 5*Determinación de Estratos y Afijación del Tamaño de Muestra para n =300 Clientes mediante Afijación Proporcional |
|

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Est | **Segmentación** | **Nacional** | **GYE****(Nh)** | **Proporción****(Wh\*100%)** | **T. Muestra****(nh)** |
| 1 | Tarifario | 45.563 | 28.249 | 5,06% | 16 |
| 2 | Autocontrol | 55.689 | 34.527 | 6,19% | 18 |
| 3 | Prepago | 798.748 | 495.224 | 88,75% | 266 |
|  | **TOTAL** | 900.000 | **558.000** | 100,0% | **300** |

 |
| ***Elaboración:*** *José L. Mora M.* |

Lo siguiente sería aplicar la afijación uniforme de la muestra para los 3 estratos y un tamaño de muestra n =300 clientes, de lo que se obtiene lo siguiente (ver Tabla 6):

|  |
| --- |
| *TABLA 6*Afijación del Tamaño de Muestra para *n =300* Clientes mediante Afijación Uniforme |
|

|  |  |  |  |
| --- | --- | --- | --- |
| **Segmentación** | **Nacional** | **GYE** | **T. Muestra** |
| Tarifario | 45.563 | **28.249** | 100 |
| Autocontrol | 55.689 | **34.527** | 100 |
| Prepago | 798.748 | **495.224** | 100 |
| **TOTAL** | 900.000 | **558.000** | **300** |

 |
| ***Elaboración:*** *José L. Mora M.* |

Entre las dos afijaciones aplicadas al mismo tamaño de muestra n se tiene que la más precisa es la proporcional dado que asigna a cada estrato un número proporcional a su contribución sobre toda la población, pero al observar la cantidad de unidades con las que contribuye sobre la muestra el estrato 1 y 2 (ver Tabla 5) que entre los dos suman 34 clientes, se considera que este valor no brindará la suficiente información para los fines de la empresa, dado que aunque estos dos estratos juntos representen el 11,25% de la población, no es menos cierto que entre los dos representan aproximadamente el 65% o 70% de los ingresos que la compañía percibe por los servicios de telefonía celular. Por tanto, se decide seleccionar el método de afijación uniforme el cual asigna igual cantidad de unidades de muestreo a cada uno de los estratos, que si bien es cierto, este método al dar la misma importancia a todos los estratos está favoreciendo a los estratos de menor tamaño y perjudicando a los estratos más grandes en cuanto a su precisión, también es el método que más se ajusta a las necesidades del proveedor de servicios de telefonía celular cuyos objetivos están orientados hacia la satisfacción de todos sus clientes, pero mantiene un especial interés sobre los clientes que les producen sus mayores ingresos, como es el caso de los clientes Tarifario y Autocontrol.

* + 1. **POBLACIÓN OBJETIVO**

Una tarea importante para el investigador es definir cuidadosa y completamente la población antes de recolectar la muestra. Inicialmente una población es una colección de elementos acerca de los cuales se desea hacer alguna inferencia.

Esta población inicial que se desea investigar se denomina población objetivo. La misma que para los fines del estudio que aquí se desarrolla está compuesta por los clientes de un proveedor de telefonía celular de la ciudad de Guayaquil[[2]](#footnote-3)(3), para lo cual se escogió a PORTA Celular como empresa proveedora y a sus clientes de la ciudad de Guayaquil como la población objetivo (ver Tabla 7).

|  |
| --- |
| TABLA 7Cantidad de Clientes a Nivel Nacional y Regional |
|

|  |  |  |  |
| --- | --- | --- | --- |
| **Segmentación** | **Nacional** | **GYE** | UIO |
| Tarifario | 45.563 | **28.249** | 17.314 |
| Autocontrol | 55.689 | **34.527** | 21.162 |
| Prepago | 798.748 | **495.224** | 303.524 |
| **TOTAL** | 900.000 | **558.000** | 342.000 |

 |
| ***Fuente:*** *CONECEL S.A., enero 2003****Elaboración:*** *José L. Mora M.* |

De esto se determinó que PORTA tiene a nivel nacional aproximadamente 900000 clientes, de los cuales 558000 aproximadamente pertenecen a la región ciudad de Guayaquil, la misma que está compuesta de 28249 clientes Tarifarios, 34527 clientes Autocontrol y 495224 clientes Prepago aproximadamente.

|  |
| --- |
| GRÁFICO 3.1Porcentaje de Clientes de Telefonía Celular de PORTA por regiones a Nivel Nacional |
|  |
| ***Fuente:*** *CONECEL S.A.****Elaboración:*** *José L. Mora M.* |

|  |
| --- |
| GRÁFICO 3.2Porcentaje de Clientes de Telefonía Celular de PORTA de la ciudad de Guayaquil por Estratos |
|  |
| ***Fuente:*** *CONECEL S.A.****Elaboración:*** *José L. Mora M.* |

* + 1. **MARCO MUESTRAL**

El marco muestral en el estudio que aquí se presenta está compuesto por los estratos Tarifario, Autocontrol y Prepago, los cuales han sido seleccionados mediante las técnicas de muestreo descritas en este capítulo. El número de clientes seleccionados de cada estrato es el que se presenta en la Tabla 6 de acuerdo con las consideraciones descritas y realizadas previamente, donde siempre debe considerarse que no debe existir solapamientos, es decir, que no puede darse el caso de que un mismo cliente pertenezca a dos o más estratos.

1. (2) Por ejemplo, si la unidad es la familia, debe definírsela de tal forma que una persona no pertenezca a dos familias diferentes ni dejarse fuera a cualquier persona que pertenezca a la población. [↑](#footnote-ref-2)
2. (3) CONECEL S.A. establece dos regiones donde ofrece sus servicios de telefonía celular: Guayaquil y Quito. Las cuales agrupan a otras ciudades como en el caso de la región Guayaquil la cual agrupa a las ciudades Guayaquil, Machala, Manta, Quevedo y Cuenca. [↑](#footnote-ref-3)