

DISEÑO MUESTRAL DE LA ENCUESTA DE OPINIÓN SOBRE INTENCIÓN DE VOTO EN LAS ELECCIONES MUNICIPALES Y PRESIDENCIALES: UNMSM

*Olga Solano Dávila**, *Doris Gómez Ticerán***, *Ana Cárdenas Rojas****
*Ysabel Adriaola Cruz*****, *Felix Bartolo Gotarate******, *Blanca Martínez Portuguese******
& *Orlando Giraldo Laguna*

Resumen: En el presente trabajo presentamos la metodología del diseño muestral en la encuesta de opinión de intención de voto de los alumnos de la UNMSM en las elecciones municipales del 2010 y en las presidenciales del 2011. Utilizamos el muestreo aleatorio estratificado con afijación proporcional al tamaño (Scheffer y Mendenhall, 2003) y se consideraron los siguientes estratos: Facultad de Ciencias Administrativas, Facultad de Ciencias Biológicas entre otras, en total veinte facultades de la UNMSM. Las encuestas se llevaron a cabo en tres etapas, durante los meses de julio, agosto y setiembre. En cada etapa se consideró a 500 participantes los cuales fueron seleccionados de cada una de las veinte facultades, siguiendo el esquema de muestreo estratificado con afijación proporcional. Para el cálculo del tamaño de muestra se consideró un límite para el error de estimación del 4.41 %, con un nivel de confianza del 95 % y la información proporcionada por el Sistema Único de Matricula (SUM), de los alumnos matriculados el primer semestre del año académico 2010.

Palabras clave: Intención de voto, Elecciones Municipales y Presidenciales, Muestreo Aleatorio Estratificado, Tamaño de muestra.

SAMPLING DESIGN IN THE OPINION SURVEY ON INTENTION TO VOTE IN PRESIDENTIAL ELECTIONS TO STUDENTS OF THE UNMSM

Abstract: In the present work, we present the methodology of sample design, on the opinion poll of voting intentions of the students of San Marcos University in municipal elections of 2010 and presidential elections of 2011. We used stratified random sampling with allocation proportional to size (Scheffer y Mendenhall, 2003) and considered the following layers: Faculty of Administrative Sciences, Faculty of Biological Sciences, among others, in total twenty faculties of San Marcos University. The surveys were carried out in three stages, during the months of July, August and September. At each stage was considered to 500 participants who were selected from each of the twenty faculties, following the stratified sampling scheme with proportional allocation. To calculate the sample size was considered a limit to the estimation error of 4.41 % with a confidence level of 95 % and the information provided by the Unified Registration System (SUM) of students matriculated the first semester of 2010 academic year.

Key words: Voting intention, Municipal and Presidential Elections, Stratified Random Sampling, Size of sample.

1. INTRODUCCIÓN

La investigación de mercados a menudo comprende la estimación de una característica de alguna población de interés, por ejemplo, podría necesitarse información sobre las actitudes de los estudiantes

* UNMSMS, Facultad de Ciencias Matemáticas, e-mail: osolanod@unmsm.edu.pe

** UNMSMS, Facultad de Ciencias Matemáticas, e-mail: dorisgomezt@gmail.com

*** UNMSMS, Facultad de Ciencias Matemáticas, e-mail: acardenasr@unmsm.edu.pe

**** UNMSMS, Facultad de Ciencias Matemáticas, e-mail: radriazolac@unmsm.edu.pe

***** UNMSMS, Facultad de Ciencias Matemáticas, e-mail: fbartolog@unmsm.edu.pe

***** UNMSMS, Facultad de Ciencias Matemáticas, e-mail: bmartinezp@unmsm.edu.pe

***** UNMSMS, Facultad de Ciencias Matemáticas, e-mail: ogiraldol@unmsm.edu.pe

acerca de la propuesta de una instalación dentro de la escuela, o desear saber el porcentaje de electores que votaría por un determinado partido político. En cualquier caso, sería poco probable que se pudiera entrevistar a todos los miembros de la población. El contacto de toda la población, no valdría la pena desde el punto de vista de costo-beneficio. Además de ser costoso, en casi todos los casos sería innecesario, ya que una muestra por lo general es suficientemente confiable. El muestreo es útil si el tamaño de la población es grande y si tanto el costo como el tiempo asociados con la obtención de la información de la población son elevados. Además se podría perder la oportunidad de tomar una decisión rápida si se debe encuestar a una gran población.

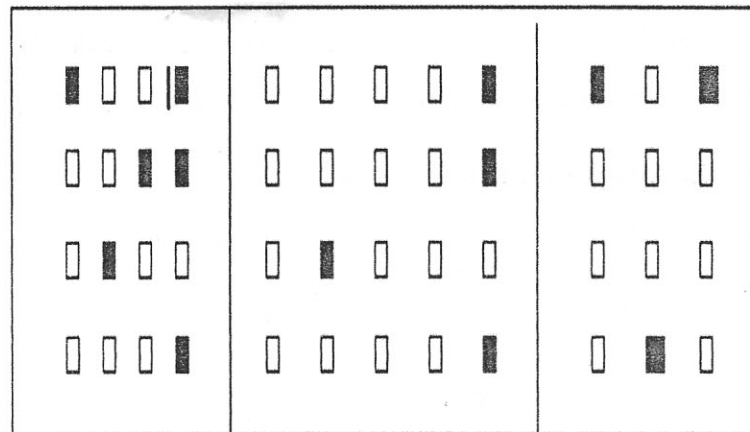
Asimismo, con el muestreo, en un espacio de tiempo determinado, se puede dedicar más tiempo a cada entrevista (personal), incrementando de esta forma la calidad de la respuesta. (Aaker et. al , 2007).

Existen varios esquemas de muestreo en la teoría, uno de ellos es por ejemplo el Muestreo Estratificado el cual se utiliza cuando el interés es por ejemplo estimar la proporción de votantes que tienen preferencia por un candidato en particular. En este caso se divide a la población en estratos (de acuerdo a alguna característica de interés, género, grupos de edad, etc.) se toma una muestra aleatoria de cada estrato. Luego se realizan las entrevistas para determinar la proporción de votantes que tienen preferencia por un candidato en particular, integrantes en el estrato i .

El objetivo del presente trabajo es estudiar la intención de voto en las elecciones municipales del 2010 y las presidenciales del 2011, de los alumnos de la Universidad Nacional Mayor de San Marcos, utilizando la teoría del muestreo estadístico, específicamente el Muestreo Aleatorio Estratificado. (Scheaffer et al 2003; Cochran, G., 1985; Lohr, S. 2000; Perez, C., 2006).

2. DISEÑO MUESTRAL: Muestreo Estratificado

Una muestra estratificada es obtenida mediante la separación de los elementos en grupos llamados estratos, y la selección posterior de una muestra aleatoria de cada estrato, por ejemplo, la Figura 1 muestra un esquema gráfico de muestreo estratificado de una población con tres estratos.



■ ELEMENTO DE LA MUESTRA

Figura 1: Esquema gráfico de un muestreo estratificado. Población con tres estratos

El procedimiento de selección de una muestra estratificada consiste en:

- i. Considerando L estratos independientes ($L > 1$) tal como se observa en la figura 2.

N_i : Tamaño del estrato i , ($i = 1, 2, \dots, L$)

$N = \sum_{i=1}^L N_i$: Tamaño de la población.

- ii. Seleccionar una muestra aleatoria independientemente de cada estrato, de tal manera que el tamaño total de la muestra es:

$$n = \sum_{i=1}^L n_i,$$

donde : n_i es el Tamaño de la muestra en el estrato i .

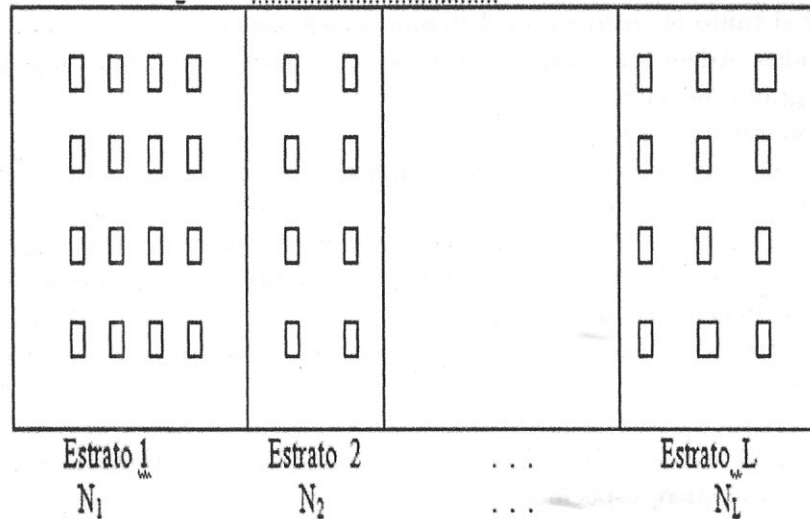


Figura 2: Estratificación de una población

Existen tres criterios que se deben de tener presente cuando se está planeando utilizar el muestreo estratificado.

1. Los estratos deben de formarse de tal manera que se garantice la independencia entre estratos. Es decir, los estratos deben ser completamente independientes en el proceso de selección y estimación.
2. Las mediciones dentro de los estratos deben ser homogéneas (baja variabilidad).
3. Las mediciones entre estratos deben ser heterogéneas (alta variabilidad).

2.1. Estimación de una proporción poblacional

Si lo que se desea estimar es la proporción poblacional P , de una característica determinada, se usa:

$$\begin{aligned} \hat{P}_{est} &= \frac{1}{N} (N_1 p_1 + N_2 p_2 + \dots + N_L p_L) \\ &= \frac{1}{N} \sum_{h=1}^L N_h p_h = \sum_{h=1}^L W_h p_h \end{aligned} \quad (2.1)$$

donde :

$$W_h = \frac{N_h}{N}$$

p_h : proporción de individuos con la característica de interés en el estrato h .

La varianza estimada del estimador de la proporción poblacional es :

$$\hat{V}(\hat{P}_{est}) = \frac{1}{N^2} \sum_{h=1}^L N_h^2 (1 - f_h) \left(\frac{p_h q_h}{n_h - 1} \right) \quad (2.2)$$

El límite para el error de estimación, de la proporción poblacional, es:

$$2\sqrt{\hat{V}(\hat{P}_{est})} = 2\sqrt{\frac{1}{N^2} \sum_{h=1}^L N_h^2 (1 - f_h) \left(\frac{p_h q_h}{n_h - 1} \right)} \quad (2.3)$$

El intervalo de confianza para la proporción de la población es:

$$\hat{P}_{est} \pm Z \sqrt{\hat{V}(\hat{P}_{est})}$$

donde :

$$\hat{P}_{est} = \sum_{h=1}^L W_h p_h$$

Z : es la abscisa de la distribución normal.

2.2. Selección del tamaño de muestra para estimar proporciones

Para estimar una proporción poblacional, primero indicamos qué tanta información deseamos mediante la especificación del tamaño del límite para el error de estimación. La fórmula para el tamaño de muestra n (para un límite para el error de estimación) está dada por

$$n = \frac{\sum_{h=1}^L N_h^2 p_h q_h / W_h}{N^2 D + \sum_{h=1}^L N_h p_h q_h} \quad (2.4)$$

donde :

W_h es fracción de observaciones asignadas al estrato h .

$D = \frac{B^2}{4}$ para estimar la proporción poblacional P_{est}

$2\sqrt{V(\hat{P}_{est})} = B$ (la estimación de P_{est} debe estar dentro de B unidades de la media poblacional, con una probabilidad aproximadamente igual a 0.95).

2.3. Afijación del tamaño de muestra en cada estrato

El objetivo del diseño de encuesta por muestreo es proporcionar estimadores con varianzas pequeñas al menor costo posible. Después de elegir el tamaño de muestra n , existen muchas maneras para dividir n entre los tamaños de muestra de los estratos individuales n_1, n_2, \dots, n_L . Cada división puede originar una varianza diferente para la media muestral. Por lo que nuestro objetivo es usar una asignación que presente una cantidad especificada de información a un costo mínimo.

De acuerdo a este objetivo, el mejor esquema de asignación está influenciado por tres factores. Ellos son los siguientes:

1. El número total de elementos en cada estrato.
2. La variabilidad de las observaciones dentro de cada estrato.
3. El costo por obtener una observación de cada estrato.

A continuación se presenta un resumen de las metodologías más conocidas para determinar en la muestra el tamaño de cada estrato.

2.3.1. Afijación aproximada que minimiza el costo para el valor fijo de $V(\bar{y}_{est})$ o que minimiza $V(\bar{y}_{est})$ para un costo fijo:

$$n_h = n \left(\frac{N_h \sigma_h / \sqrt{c_h}}{N_1 \sigma_1 / \sqrt{c_1} + N_2 \sigma_2 / \sqrt{c_2} + \dots + N_L \sigma_L / \sqrt{c_L}} \right)$$

$$n_h = n \left(\frac{N_h \sigma_h / \sqrt{c_h}}{\sum_{h=1}^L N_h \sigma_h^2} \right) \quad (2.5)$$

donde:

N_h : denota el tamaño del h -ésimo estrato,

σ_h^2 : denota la varianza poblacional para el h -ésimo estrato,

c_h : denota el costo para obtener una observación individual del h -ésimo estrato.

2.3.2 Afijación de Neyman

En algunos problemas de muestreo estratificado, el costo por obtener información es el mismo para todos los estratos. Si los costos son desconocidos, podríamos suponer que los costos por observación son iguales. Si $c_1 = c_2 = \dots = c_L$, entonces los términos de costo se cancelan en la ecuación (2.5) y

$$n_h = n \left(\frac{N_h \sigma_h}{\sum_{h=1}^L N_h \sigma_h} \right) \quad (2.6)$$

este método de seleccionar n_1, n_2, \dots, n_L se denomina afijación de Neyman.

2.3.3 Afijación Proporcional

Además de encontrar costos iguales, en algunas ocasiones resultan las mismas varianzas, $\sigma_1^2, \sigma_2^2, \dots, \sigma_L^2$. En tal caso se cancelan las σ_h en la Ecuación (2.6) y

$$n_h = n \left(\frac{N_h}{\sum_{h=1}^L N_h} \right) = n \left(\frac{N_h}{N} \right) \quad (2.7)$$

Este método de asignación de la muestra a los estratos es llamado afijación proporcional porque los tamaños de muestra n_1, n_2, \dots, n_L son proporcionales a los tamaños de los estratos N_1, N_2, \dots, N_L .

En la práctica la afijación proporcional puede y suele utilizarse cuando los costos y las varianzas de los estratos no son iguales.

3. Resultados

En esta sección se presentan los resultados de la investigación usando el método de muestreo estratificado con afijación proporcional.

Muestra piloto

La encuesta piloto se realizó en el mes de junio del año 2010 con la finalidad de detectar los posibles errores en el instrumento de medición diseñado para el estudio.

Después de la encuesta piloto el instrumento final quedó diseñado de la siguiente manera:

1. Parte I : Perfil del estudiante Sanmarquino: características de los estudiantes de pregrado como género, edad, EAP, facultad, ciclo de estudios y nivel socioeconómico.
2. Parte II: Intención de voto: recojo de opiniones del estudiante sobre intención de voto de las elecciones municipales y sobre algunas características de los postulantes de los diferentes partidos políticos

Muestra definitiva

Las facultades de:

1. CIENCIAS ADMINISTRATIVAS
2. CIENCIAS BIOLÓGICAS

3. CIENCIAS CONTABLES
4. CIENCIAS ECONÓMICAS
5. CIENCIAS FÍSICAS
6. CIENCIAS MATEMÁTICAS
7. CIENCIAS SOCIALES
8. DERECHO Y CIENCIA POLÍTICA
9. EDUCACIÓN
10. FARMACIA Y BIOQUÍMICA
11. INGENIERÍA DE SISTEMAS E INFORMÁTICA
12. INGENIERÍA ELECTRÓNICA
13. INGENIERÍA GEOLÓGICA, MINERA, METALÚRGICA Y GEOGRÁFICA
14. INGENIERÍA INDUSTRIAL
15. LETRAS Y CIENCIAS HUMANAS
16. MEDICINA
17. MEDICINA VETERINARIA
18. ODONTOLOGÍA
19. PSICOLOGÍA
20. QUÍMICA E INGENIERÍA QUÍMICA

de la UNMSM fueron tratados como estratos.

1. La unidad de análisis en todos los casos es el alumno.
2. La información de los alumnos matriculados en el Semestre 2010-I en cada facultad fue la siguiente:

Como deseamos conocer la proporción de votantes por algún partido político en cada facultad, optamos por utilizar el muestreo estratificado con afijación proporcional al tamaño de los estratos a fin de captar la intención de voto de los alumnos en cada facultad.

Así, para:

$P_i = 0,5$ ($i = 1, 2, 3, \dots, 20$): proporción de alumnos con intención de voto por algún partido político en la i -ésima facultad de la UNMSM;

$$W_i = \frac{N_i}{N}, \quad i = 1, 2, \dots, 20$$

N_i : Número de alumnos matriculados el Semestre 2010-I en la i -ésima Facultad de la UNMSM.
 $N = 28056$, total de alumnos matriculados el Semestre 2010-I en la UNMSM (Ver Cuadro 1).

$$W_1 = \frac{2556}{28056}, W_2 = \frac{711}{28056}, W_3 = \frac{1943}{28056}, W_4 = \frac{1680}{28056}, W_5 = \frac{949}{28056},$$

$$W_6 = \frac{1585}{28056}, W_7 = \frac{2023}{28056}, W_8 = \frac{2076}{28056}, W_9 = \frac{1277}{28056}, W_{10} = \frac{607}{28056},$$

$$W_{11} = \frac{1388}{28056}, W_{12} = \frac{1220}{28056}, W_{13} = \frac{1273}{28056}, W_{14} = \frac{1356}{28056}, W_{15} = \frac{1883}{28056},$$

$$W_{16} = \frac{2817}{28056}, W_{17} = \frac{313}{28056}, W_{18} = \frac{391}{28056}, W_{19} = \frac{768}{28056}, W_{20} = \frac{1240}{28056}.$$

Se elige $E = 0.0441$, límite para el error de estimación con un nivel de confianza del 95 %. Utilizando (2.4) el tamaño de muestra fue:

$$n = \frac{2556^2(0,5)(0,5)/(2556/28056) + \dots + 1240(0,5)(0,5)/1240(0,5)(0,5)/1240/28056}{28056^2(0,00194481/4) + 2556(0,5)(0,5) + 711(0,5)(0,5) + \dots + 1240(0,5)(0,5)}$$

Cuadro 1 : Distribución de los Estudiantes de Pre-Grado según Facultad- Semestre 2010-I - UNMSM

Facultades	Frecuencia	Porcentaje
Ciencias Administrativas	2556	9.11
Ciencias Biológicas	711	2.53
Ciencias Contables	1943	6.93
Ciencias Económicas	1680	5.99
Ciencias Físicas	949	3.38
Ciencias Matemáticas	1585	5.65
Ciencias Sociales	2023	7.21
Derecho y Ciencias Políticas	2076	7.40
Educación	1277	4.55
Farmacia y Bioquímica	607	2.16
Ingeniería de Sistemas e Informática	1388	4.95
Ingeniería Electrónica	1220	4.35
Ingeniería, Geológica,Minera, Metalúrgica y G.	1273	4.54
Ingeniería Industrial	1356	4.83
Letras y Ciencias Humanas	1883	6.71
Medicina	2817	10.04
Medicina Veterinaria	313	1.12
Odontología	391	1.39
Psicología	768	2.74
Química e Ingeniería Química	1240	4.42
Total	28056	100

$n = 500$. Es decir, el tamaño de muestra total fue de 500 unidades estadísticas. Estas 500 unidades, según la metodología de muestreo estratificado con afijación proporcional, usando (2.7) se repartirán proporcional al tamaño de cada estrato de la población, como se muestra a continuación.

a) Para la Facultad de Ciencias Administrativas:

$$n_1 = n \left(\frac{N_1}{N} \right) = 500 \left(\frac{2556}{28056} \right) = 46 \text{ alumnos de pregrado.}$$

b) Para la Facultad de Ciencias Biológicas:

$$n_2 = n \left(\frac{N_2}{N} \right) = 500 \left(\frac{711}{28056} \right) = 13 \text{ alumnos de pregrado.}$$

Asimismo calculamos para el resto de facultades. Finalmente, la distribución de la muestra según las diferentes facultades de la UNMSM se muestra en el Cuadro 2.

Cuadro 2: Distribución de la muestra requerido para cada facultad

Facultades	$n_i = n(N_i/N)$
Ciencias Administrativas	46
Ciencias Biológicas	13
Ciencias Contables	35
Ciencias Económicas	30
Ciencias Físicas	17
Ciencias Matemáticas	28
Ciencias Sociales	36
Derecho y Ciencias Políticas	37
Educación	22
Farmacia y Bioquímica	11
Ingeniería de Sistemas e Informática	25
Ingeniería Electrónica	22
Ingeniería, Geológica, Minera, Metalúrgica y G.	23
Ingeniería Industrial	24
Letras y Ciencias Humanas	34
Medicina	50
Medicina Veterinaria	6
Odontología	7
Psicología	14
Química e Ingeniería Química	22
Total	500

El muestreo aleatorio estratificado determinó seleccionar a 500 alumnos de pregrado, distribuidas de la siguiente manera: 46 alumnos de pregrado de la facultad de Ciencias Administrativas, 13 alumnos de pregrado de la facultad de Ciencias Biológicas, 35 alumnos de Ciencias Contables, y finalmente 22 alumnos de la facultad de Química e Ingeniería Química, como se muestra en el Cuadro 2.

4. Discusión y Conclusiones

El muestreo estratificado nos permitió poner en práctica la técnica estadística del muestreo, que resultaron eficientes en la selección y aplicación de la encuesta de intención de voto. Para el diseño del muestreo se utilizó el Muestreo Estratificado (Scheffer y Mendenhall, 2007; Cochran, 1977), con un nivel de confianza del 95% y un margen de error del 4.41%. el tamaño de muestra fue de 500 alumnos durante cada una de las tres etapas, que fueron durante los meses de julio, agosto y setiembre del año 2010.

Agradecimientos

Los autores expresamos nuestro agradecimiento al Consejo Superior de Investigaciones de la Universidad Nacional Mayor de San Marcos por el apoyo financiero para la ejecución del estudio motivo de la presente publicación.

Bibliografía

- [1] Aaker, A.; Kumar, V.; Day, G. (2007). Investigación de mercados. 4ta. Ed., Edit. Limusa, México.
- [2] Cochran, G.W. (1977). Técnicas de muestreo. 5ta. Ed. , Edit. Continental,. México.
- [3] Des, R. (1980). Teoría del muestreo. 1ra. Ed., Edit. Fondo de cultura económica. México.
- [4] Lohr, S. (2000). Muestreo: diseño y análisis. 2da. Ed., Edit. Thomson Editores. México.
- [5] Malhotra, N. (1997). Investigación de mercados. 2da. Ed., Edit. Prentice Hall. México.
- [6] Perez, C. (2006). Muestreo estadístico. 3ra. Ed., Edit. Pearson Educación. España.
- [7] Scheaffer, R.; Mendenhall, W. Y; Ott, L. (2007). Elementos de muestreo. 6ta. Ed., Edit. Iberoamérica. México.