

2. TIPOS DE MUESTREO PROBABILÍSTICO

2.1. Muestreo aleatorio simple

Es el método más simple para elegir la muestra. Básicamente es un sorteo realizado de tal manera que se garantice que cada una de las unidades elementales tenga la misma probabilidad conocida.

La elección, por lo tanto, es aleatoria y se puede llevar a cabo mediante un ánfora con balotas con números o identificadores de cada uno de los elementos de la población. También, puede realizarse usando una tabla denominada Tabla de números aleatorios y la lista completa de los elementos de la población.

Se puede mencionar dos cualidades:

- Independencia: La elección de cada unidad de muestreo es independiente de la elección de otra, dicho de otra manera, la elección de un elemento no influye en la probabilidad de elección del resto de unidades. Esto es posible si la elección se realiza con reemplazo.
- Representatividad: El conjunto de elementos elegidos es un fiel reflejo de su población.

Es muy eficiente con poblaciones pequeñas, desarrolla una muestra representativa sobre todo si la población es bastante homogénea. A cambio se puede asegurar que no es tan eficiente con poblaciones heterogéneas, dado que la elección aleatoria podría producir sesgo y no se puede realizar con poblaciones infinitas.

Ejemplo 2:

Los ingresos anuales, grado de instrucción, la categoría laboral de los 120 empleados de Mariana S. A., empresa dedicada a la producción de ropa para bebé, se muestran en el Apéndice A Tabla de datos de empleados de Mariana S. A.

Con esta base de datos, obtendremos una muestra aleatoria de tamaño 10 utilizando la Tabla de Números aleatorios, del Apéndice B.

En la tabla de números aleatorios, elegimos de manera aleatoria las columnas C7, C2, C11.

Datos:

$$N = 120 \text{ (tamaño de población)}$$

$$n = 10 \text{ (tamaño de muestra)}$$

De esta manera, seguimos buscando y los elegidos son los siguientes: 012; 014; 092; 007

1.º Tomamos la columna C7 y le añadimos 2 columnas consecutivas, C8 y C9.

Esto porque $N = 120$ es un número que tiene 3 dígitos.

2.º Buscamos en las tres columnas números menores o iguales a 120. Por ejemplo, 012 = 12.

El primer elemento elegido es el empleado n.º 12:



Figura 7. Muestreo por criterio. Tomada de Graphic Resources LLC, 2015.

Tabla n.^a 1. Números aleatorios.

C7	C8	C9
4	0	9
0	1	2
5	6	7
5	0	6
2	4	4
4	5	9
3	2	0
3	3	7
3	9	3
7	3	1
8	1	0

Fuente: Tomada de Manual autoformativo de Estadística II Cerrón, 2013.

De esta manera, seguimos buscando y los elegidos son los siguientes:

012; 014; 092; 007 Tabla 1.

3.^º Hacemos lo mismo con la columna C2 a la que añadimos C3 y C4, y se obtiene lo siguiente:

036; 058

Con C11, añadimos C12 y C13:

086; 011; 088; 067

Por tanto, los elementos elegidos son 12, 14, 92, 7, 36, 58, 36, 58, 86, 11, 88, 67.

2.2. Muestreo sistemático

Se realiza mediante un proceso que se desarrolla mediante los siguientes pasos:

1.^º Se calcula el tamaño de salto o intervalo a dar en la elección de los elementos que se muestran a continuación:

$$K = \frac{N}{n} , \text{ donde } N \text{ es el tamaño de la población y } n \text{ es el tamaño de la muestra a elegir.}$$

2.^º De manera aleatoria, se elige al primer elemento entre los "k primeros". Se puede usar la tabla de números aleatorios.

3.^º Se elige a los siguientes dando "saltos" de k en k elementos.



Figura 8. Muestreo sistemático.



Este tipo de muestreo tiene la ventaja de una mayor posibilidad de elegir una muestra representativa. La elección de cada elemento es independiente de la probabilidad de elección de otros elementos.

En su contra, se puede mencionar que la población debe estar ordenada como en una lista. Si existe alguna forma de periodicidad en la lista, entonces el sesgo es mucho mayor que en el muestreo aleatorio simple. Esto puede producirse, por ejemplo, en datos de ventas que son muy susceptibles a las fechas de pago en las empresas.

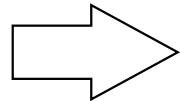
Es muy sensible a la influencia de circunstancias sistemáticas, es decir, a sucesos que inciden en el desarrollo de los datos de manera indirecta como, por ejemplo, un engranaje defectuoso y que este defecto produzca cambios en la producción.

Ejemplo 3:

Del caso de la empresa Mariana S. A., desarrolle un muestreo para elegir una muestra de 10 personas de la lista de 120 empleados.

Datos:

$$\begin{aligned} N &= 120 \\ n &= 10 \end{aligned}$$



1.^º Calculamos el intervalo de selección así:

$$K = \frac{N}{n} = \frac{120}{10} = 12$$

2.^º Elegimos el punto de partida.

Para ello, se debe elegir un elemento de los primeros $k = 12$. Luego, tomamos en la tabla de números aleatorios la columna C5 y C6 (aleatoriamente), y buscamos el primer número menor o igual a 12. El resultado fue el siguiente:

$$1.^{\text{er}} \text{ elegido} = 04$$

La primera persona elegida es la que ocupa el cuarto lugar en la lista tal como se aprecia en la siguiente tabla:

Tabla n.^º 2. Lista de datos de empleados (tabla de datos Mariana)

id	Sexo	Grand Inst	N. ^º de Hijos	Cat. Lab.	Ingreso (miles)	Tiemp. Empl.
1	Hombre	Universidad inconcluso	6	Admin.	\$40.20	98
2	Mujer	Primaria	7	Admin.	\$21.90	98
3	Hombre	Técnico	5	Admin.	\$32.10	98
4	Mujer	Secundaria	2	Admin.	\$21.90	98
5	Mujer	Secundaria	6	Admin.	\$24.00	98
6	Mujer	Universidad inconcluso	4	Admin.	\$30.30	98
7	Hombre	Técnico	3	Admin.	\$27.75	98

3.^º A partir de esta persona, se eligen los siguientes casos dando saltos con $k = 12$.

Tabla n.^º 3. Tabla de datos Mariana.

id	Sexo	Grand Inst	Nº de Hijos	Cat. Lab.	Ingreso (miles)	Tiemp. Empl.
1	Hombre	Universidad inconcluso	6	Admin.	\$40.20	98
2	Mujer	Primaria	7	Admin.	\$21.90	98
3	Hombre	Técnico	5	Admin.	\$32.10	98
4	Mujer	Secundaria	2	Admin.	\$21.90	98
5	Mujer	Secundaria	6	Admin.	\$24.00	98
6	Mujer	Universidad inconcluso	4	Admin.	\$30.30	98
7	Hombre	Técnico	3	Admin.	\$27.75	98
8	Mujer	Primaria	5	Admin.	\$31.35	96
9	Hombre	Técnico	2	Admin.	\$31.35	96
10	Hombre	Secundaria	3	Admin.	\$23.25	95
11	Mujer	Técnico	2	Admin.	\$22.35	95
12	Mujer	Secundaria	8	Admin.	\$30.00	95
13	Hombre	Secundaria	4	Admin.	\$35.55	94
14	Hombre	Secundaria	8	Admin.	\$25.05	94
15	Hombre	Primaria	7	Admin.	\$22.50	94
16	Hombre	Primaria	6	Admin.	\$21.90	93
17	Hombre	Técnico	4	Admin.	\$41.10	93
18	Mujer	Secundaria	3	Admin.	\$26.40	93
19	Mujer	Primaria	6	Admin.	\$25.05	93
20	Mujer	Secundaria	4	Admin.	\$28.50	92
21	Hombre	Técnico	8	Admin.	\$33.45	90
22	Mujer	Técnico	2	Admin.	\$32.55	90
23	Mujer	Secundaria	8	Admin.	\$33.30	90
24	Hombre	Secundaria	4	Admin.	\$27.30	90

id	Sexo	Grand Inst	Nº de Hijos	Cat. Lab.	Ingreso (miles)	Tiemp. Empl.
25	Mujer	Técnico	2	Admin.	\$26.65	88
26	Hombre	Universidad inconcluso	2	Admin.	\$52.65	86
27	Hombre	Secundaria	7	Admin.	\$26.70	86
28	Hombre	Técnico	1	Admin.	\$37.50	84
29	Mujer	Secundaria	4	Admin.	\$16.50	84
30	Mujer	Secundaria	8	Admin.	\$24.75	84
31	Mujer	Secundaria	5	Admin.	\$24.00	83
32	Mujer	Secundaria	8	Admin.	\$20.40	83
33	Hombre	Técnico	2	Admin.	\$30.15	82
34	Mujer	Secundaria	0	Admin.	\$33.90	82
35	Hombre	Técnico	8	Admin.	\$22.50	82
36	Mujer	Primaria	5	Admin.	\$27.45	81
37	Mujer	Secundaria	3	Admin.	\$27.30	81
38	Mujer	Secundaria	6	Admin.	\$23.10	81
39	Mujer	Secundaria	2	Admin.	\$23.10	81
40	Hombre	Secundaria	3	Admin.	\$25.50	81
41	Hombre	Primaria	5	Admin.	\$21.30	80
42	Mujer	Secundaria	3	Admin.	\$23.40	80
43	Hombre	Técnico	2	Admin.	\$28.65	79
44	Hombre	Universidad inconcluso	5	Admin.	\$40.35	78
45	Hombre	Técnico	2	Admin.	\$25.95	78
46	Hombre	Universidad inconcluso	3	Admin.	\$26.55	78
47	Hombre	Técnico	6	Admin.	\$30.75	77
48	Hombre	Técnico	6	Admin.	\$34.60	77

Así resultan elegidos los siguientes casos:

4, 16, 28, 40, 52, 64, 76, 88, 100, 112.

2.3 Muestreo por estratos

Cuando las poblaciones son extensas y/o se quiere asegurar la representatividad de los grupos en la muestra, entonces la población se separa en estratos:

Los estratos se pueden formar de acuerdo con características (variables) muy importantes para el desarrollo de la investigación. Por ejemplo, en el caso de una encuesta de opinión, se estila separar a las personas en estratos como sus ingresos, nivel educativo, número de hijos.

Una vez que se han definido los estratos, se procede a un muestreo aleatorio simple o sistemático en cada estrato.

La elección puede darse tras obtener una muestra de igual tamaño de cada estrato o al tomar una cantidad proporcional de cada uno de ellos. Luego, se puede o no otorgar pesos relativos de acuerdo con la significancia de cada estrato.

Estas son algunas ventajas de este tipo de muestras:

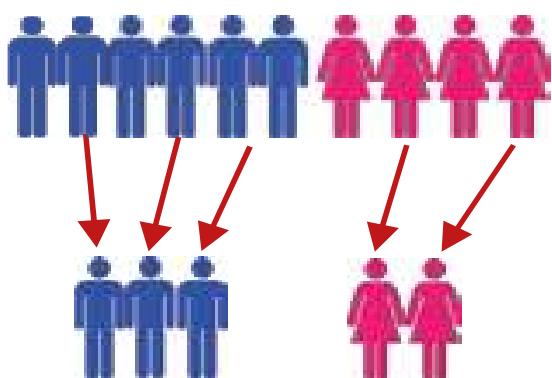


Figura 9. Muestreo por estratos.



- Una mayor representatividad.
- Una mejor precisión en las estimaciones.
- Menor costo en el recojo de datos.
- Ayuda a obtener estimaciones por estratos.

El procedimiento básico por excelencia es el siguiente:

- a. Desarrolle una lista de variables que sean de importancia para definir los estratos.
- b. Construya los estratos cuidando que sean mutuamente excluyentes; es decir, ningún elemento de la población puede pertenecer a dos estratos.
- c. Determine el número de elementos a elegir en cada estrato. Si se trata de una elección proporcional, entonces se puede usar la siguiente fórmula:

$$n_1 = n \left(\frac{N_1}{N} \right)$$

donde

N = Tamaño de la población

$$n_2 = n \left(\frac{N_2}{N} \right)$$

N1, N2, N3... son los tamaños de los estratos en la población.

$$n_3 = n \left(\frac{N_3}{N} \right)$$

n = Tamaño de muestra

n1, n2, n3... son los tamaños de muestra a obtener de cada estrato.

- d. Seleccione de preferencia muestras aleatorias de cada estrato.

Es recomendable tener en cuenta que muchos estratos generarían trabajo innecesario y redundante, y un número pequeño no ayudaría a reducir el sesgo. Se recomienda usar mínimo tres y máximo ocho estratos.

Ejemplo 4:

Con los datos del ejemplo n.^o 1, se requiere realizar la elección de una muestra de 30 individuos.

Datos:

N = 120 (población)

n = 30 (muestra)

Resumiendo, los datos:

Tabla n.^o 4. Muestra de individuos y categoría laboral.

Categoría Laboral	N _i
Directivo	22
Administrativo	90
Seguridad	8
N =	120

$$\begin{aligned} &= N_1 \\ &= N_2 \\ &= N_3 \end{aligned}$$



Tabla n.^o 5. Categoría laboral y resultados de n.

Categoría Laboral	n ($\frac{N_i}{N}$)	
Directivo	$30 \left(\frac{22}{120} \right) = 5,499 \rightarrow n_1 = 5$	
Administrativo	$30 \left(\frac{90}{120} \right) = 22,5 \rightarrow n_2 = 23$	
Seguridad	$30 \left(\frac{8}{120} \right) = 2,000 \rightarrow n_3 = 2$	

Lo que nos indica que se tiene que realizar un muestreo aleatorio en cada estrato para obtener una muestra de 5 Directivos, 23 Administrativos y 2 de Seguridad, que sumarían en total 20 personas.

- Directivos (C3 y C4) → de los 22 casos se eligen los valores que se muestran a continuación: 13°, 8°, 5°, 18° y 14°.
- Administrativos (C7, C8) → de los 90 casos se eligen los siguientes: 35, 41, 16, 66, 27, 84, 40, 1, 56, 50, 24, 45, 32, 33, 39, 73, 81, 91, 59, 91, 39, 55, 18.
- Seguridad (C12) → de los 8 casos se eligen dos: 1° y 8°.

2.4. Muestreo por conglomerados

Un muestreo por conglomerados es el que se realiza en poblaciones muy extensas. La extensión se encuentra dividida y solo basta con reconocer los límites de cada división que se denomina conglomerado.

Además, se realiza un sorteo entre conglomerados para elegir uno o dos de ellos, los cuales constituirán la muestra.

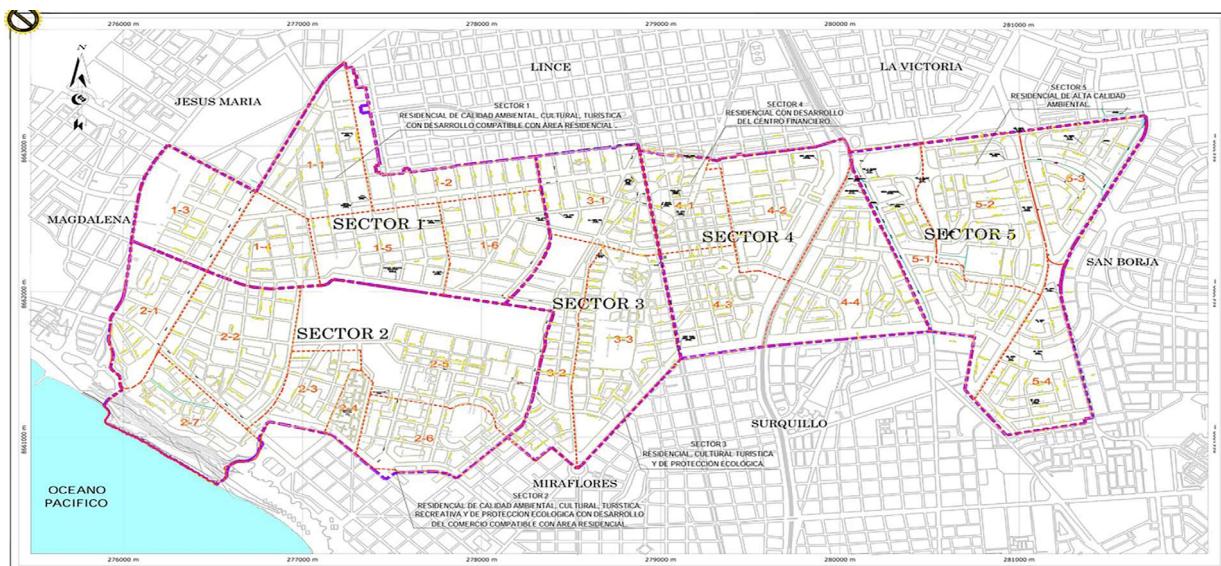


Figura 10. Muestreo por conglomerados (Municipalidad de San Isidro, 2015).

¿Conglomerados o estratos?

Alguien puede considerar similitudes entre conglomerados y estratos, pero son muy distantes; tan solo en el hecho de la formación de los grupos en estratos se deben usar variables para definir los estratos, mientras que en los conglomerados no se requieren de variables, sino que simplemente se toma la ubicación de los elementos y se identifican los límites.

Lectura seleccionada n.^o 1

Diario Perú 21. (2015, 23 abril). INEI: Pobreza en el Perú disminuyó solo 1,2 puntos porcentuales en 2014. Perú 21. Lima. Disponible en: <https://goo.gl/hXXKiq>