



Guía 10

Deduciendo la media de la población a partir de las medias muestrales

Objetivos

Emplear elementos del muestreo aleatorio simple para inferir sobre la media de una población.

Actitud

Muestra compromiso al trabajar con información cuantitativa de diversos contextos.

En esta guía estudiaremos los conceptos de Muestra y Población. Además, aprenderemos a producir Muestras Aleatorias, con el propósito de analizar la relación existente entre las Medias Muestrales y la Media Poblacional.

Actividad de exploración

- Junto a tu profesor y compañeros, analicen la situación de la “Encuesta de Tenencia Responsable de las Mascotas”¹.
- Dar ejemplos de muestras representativas y otras que no sean representativas.

Población: totalidad de los individuos, objetos u observaciones que poseen al menos una característica en común.

Muestra: subconjunto o subgrupo de la población desde la cual recogimos información.

Muestreo: proceso de extraer o seleccionar una muestra de una población.

Diseño muestral o plan de muestreo: es una estrategia o método sistemático para seleccionar la muestra.

Advertencias

- Nunca se debe olvidar que el muestreo permite obtener conclusiones de la población basada en una información bastante limitada, es decir, a partir de muestras.
- Usar planes de muestreo peligrosos.
 - El muestreo por conveniencia:** usar elementos de la población fáciles de conseguir. Es muy fácil que la opinión de aquellos más cercanos a nosotros difiera mucho de la opinión general.
 - Respuestas voluntarias:** aquellos que respondan pueden estar muy interesados en dar su opinión o hacerlo buscando un premio, por lo que sus opiniones suelen ser no representativas.

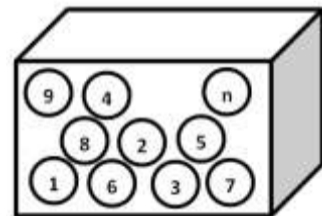


Producción de muestras aleatorias

En la extracción secuencial de los elementos para la muestra, lo más habitual es que no se pueda seleccionar un elemento más de una vez. Cuando este es el caso decimos que el muestreo o que la muestra es **sin reposición**. Cuando no hay restricciones en el número de veces que puede aparecer un elemento de la población en la muestra decimos que el muestreo o la muestra es **con reposición** o **con reemplazo**.

Observación

Los calificativos aplicados a las muestras (con o sin reposición) provienen del denominado **modelo de urna**, el cual es útil en diversos contextos. En este modelo, se reemplaza la población real por un conjunto de “N” fichas en una urna, de las cuales se desea extraer “n” de ellas. Si cada ficha lleva impresa una etiqueta (por ejemplo, un número) que identifica a cada elemento de la población, esto es equivalente a seleccionar una **muestra de tamaño “n”** de la población. Después de seleccionar una ficha, hay dos posibilidades, que definen dos tipos de muestreo:



- Se deja la ficha fuera de la caja. (**Muestreo sin reposición**)
- Se repone la ficha en la caja o se la reemplaza por una idéntica. (**Muestreo con reposición**)

Se supone que todo elemento de la población tiene la **misma probabilidad de ser seleccionada**.

¹ Actividad del texto de 4º Medio, Editorial Santillana, Año 2012, p. 200.



Ejemplo de muestras en un juego de lotería con ayuda de Excel

Supongamos un juego de lotería que consta de 90 fichas, numeradas del 1 al 90. Vamos a obtener diferentes muestras con ayuda del archivo Excel [9 Aleatorio](#):

- Experimentar y conocer cómo funciona el archivo.
- Obtener una muestra de 5 elementos sin reposición y ordenarla de menor a mayor.
- Obtener una muestra de 5 elementos con reposición y ordenarla de menor a mayor.
- Modificar el archivo para obtener una muestra de 10 elementos sin reposición, a partir de un total de 30 fichas.
- Experimentar con otras muestras y total de fichas, con reposición y sin reposición.

Un muestreo probabilístico es aquel donde **cada muestra** tiene una probabilidad conocida de ser seleccionada. Por ser el caso más frecuente, restringimos el estudio al caso de muestreo sin reposición.

El método de muestreo sin reposición que asigna la misma probabilidad a todas las muestras se llama **muestreo aleatorio simple**. El resultado de este plan de muestreo es una **muestra aleatoria simple**.

Muestras aleatorias con SIMPUC² en Excel³

El objetivo de esta actividad es seleccionar 5 muestras aleatorias simples, a partir de un listado de 25 jóvenes en el cual se indican nombre, apellido, edad y área de interés para futuros estudios.

Para identificar a cada persona usaremos una numeración correlativa, que aparece en la primera columna. Lo más sencillo es, entonces, seleccionar 5 elementos de esta columna, los que identifican las filas de la tabla. Esto equivale a elegir, al azar, 5 números des del 1 al 25, lo que se hará utilizando el archivo Excel llamado SIMPUC.

Primero abrir el archivo Word llamado [10 Actividad Simpuc](#) para ver los datos de los 25 estudiantes. Luego abrir el archivo SIMPUC, para activar las funcionalidades de este archivo; en “**advertencia de seguridad**”, “**opciones**”, seleccionar “**habilitar este contenido**”. El resto de las instrucciones consisten en la selección de ciertas opciones en el cuadro que aparece al presionar el botón “**inicio**”:

- Muestreo → Sin reposición y equiprobables.
- Elementos → Números enteros → De: 1 hasta 25.
- Tamaño de muestra → 11.
- Total de muestra → 5
- Y presionar **Generar**.

Ver vídeo explicativo en YouTube de uso de SIMPUC: [3 Ejemplo uso SIMPUC](#)

Relación entre medias muestrales y media de la población

1° Con apoyo de SIMPUC

Vamos a investigar, con ayuda de SIMPUC, la relación que existe entre la media aritmética de las medias de varias muestras de igual tamaño de una población y la media poblacional. Además, veremos qué sucede con esta relación a medida que aumenta el tamaño de las muestras.

En SIMPUC realizar las siguientes actividades:

- En la hoja “**Frecuencias**” registrar las notas obtenidas por 30 alumnos de un curso de 1° Año Medio en una prueba de Lenguaje:

4,8	5,3	6,1	5,0	5,1	7,0	3,4	4,9	5,1	3,6
3,2	2,2	4,7	4,8	2,8	4,0	5,1	5,4	3,4	2,0
3,5	7,0	3,0	4,1	5,8	4,7	6,4	5,9	4,8	5,1

- Estas notas constituirán la Población. Usando las funcionalidades de Excel, determine la media de esta población.
- Generar 20 muestras de tamaño 5 sin reposición y equiprobables. Luego calcular la media aritmética de cada muestra y, por último, el promedio de las 20 medias muestrales ([Ver vídeo de YouTube Promedio de las medias muestrales versus la media poblacional](#)). Registre los resultados en la tabla de la página 3.
- Repetir la actividad anterior, de modo de completar la tabla de la página 3.

Nota importante: cada vez que generes un conjunto de muestras, cierra SIMPUC sin guardar y vuelve abrir el archivo.

² Extraído desde <http://www.mat.puc.cl/articulo81.html>

³ Actividad extraída desde el texto Análisis Estadístico, Mineduc, 2003.



Tipos de muestras	Promedio de las medias muestrales	Media aritmética de la población
20 muestras de tamaño 5, sin reposición y equiprobables.		
20 muestras de tamaño 10, sin reposición y equiprobables.		
20 muestras de tamaño 15, sin reposición y equiprobables.		
20 muestras de tamaño 25, sin reposición y equiprobables.		
30 muestras de tamaño 5, sin reposición y equiprobables.		
40 muestras de tamaño 15, sin reposición y equiprobables.		
50 muestras de tamaño 20, sin reposición y equiprobables.		
100 muestras de tamaño 25, sin reposición y equiprobables.		

Registra en tu cuaderno las respuestas a las siguientes preguntas:

- ¿Qué relación observas entre el promedio de las medias de las muestras y la media aritmética de la población?
- A medida que aumentas el tamaño de las muestras, ¿qué sucede con el promedio de las medias de las muestras y la media aritmética de la población?
- A medida que aumentas el tamaño de las muestras y el número de muestras, ¿qué sucede con el promedio de las medias de las muestras y la media aritmética de la población?

2° Generando muestras a mano y sacando conclusiones

Considera como población los cinco primeros números naturales: $Población = \{1,2,3,4,5\}$

Desarrolla en tu cuaderno la siguiente actividad:

- Calcula el promedio de la población.
- Escribe todas las muestras de tamaño 2 y sin reposición.
- Calcula el promedio de cada una de las muestras.
- Determina la media aritmética de los promedios de todas las muestras.
- ¿Qué relación observas entre el promedio de las medias de las muestras y la media aritmética de la población?

3° La gran Media

Supongamos que queremos encontrar la media global de una población, donde ella se ha dividido en "k" muestras de tamaños $n_1, n_2, n_3, \dots, n_k$ y $\bar{x}_1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_k$ son las medias correspondientes. La gran media de la población se define como:

$$\bar{\bar{x}} = \frac{n_1 \cdot \bar{x}_1 + n_2 \cdot \bar{x}_2 + n_3 \cdot \bar{x}_3 + \dots + n_k \cdot \bar{x}_k}{n_1 + n_2 + n_3 + \dots + n_k}$$

Desarrolla en tu cuaderno:

Considerando la población $\{1,2,3,4,5\}$ de la actividad anterior y sus respectivas muestras. Calcula la gran media y compárala con la media poblacional.

Para Finalizar una idea clave:

Si se selecciona al **azar** una muestra **de gran tamaño**, los **porcentajes muestrales** se parecerán a los **poblacionales** y lo mismo sucederá con los **promedios**.

