

5. Vocabulario - jerga

Los conceptos requieren de palabras específicas para ser identificados.

La estadística tiene su propio vocabulario. Veremos algunos términos básicos, que volveremos a encontrar más adelante, además, seguiremos incorporando términos a lo largo de todo el libro.



¿Población? ¿Unidades?
¿Variables? ¿Muestra aleatoria?

Con la intención de fijar ideas, retomemos la investigación para saber quién ganará las próximas elecciones como presidente del Club Grande de Fútbol.

El primer paso es determinar el **grupo de personas a ser estudiadas**, o sea determinar la **población en estudio**. En este caso es la totalidad de los socios del Club Grande de Fútbol en condiciones de votar.

El segundo paso es **recolectar los datos**. Aquí aparecen varias cuestiones que nos permiten ilustrar más términos específicos. ¿Cuáles **individuos** serán encuestados?, esto es, ¿cuál será la **muestra**? ¿Se los elegirá en forma **aleatoria** de manera que todos los socios tengan la misma oportunidad de ser seleccionados? ¿Qué **variables** (edad, género) serán importantes en relación al tema central de la encuesta (candidato preferido)?

□ 5.1 Unidades muestrales

A los objetos de interés de un estudio se los denomina **unidades muestrales** o simplemente unidades. Muchas veces, las unidades muestrales son **individuos**: tornillos, personas, tubos de pasta dentífrica, lamparitas. Otras veces, las unidades están compuestas por **muchos individuos**: ciudades, escuelas, lotes (de tornillos) etc.

□ 5.2 Variables

Las **variables** son **características** que pueden **cambiar** de una **unidad** muestral a otra, como la **edad** de las personas, la población de cada ciudad, el **porcentaje** de alumnos reprobados de una escuela, la **preferencia** de una comida balanceada para un animal, la **intensidad** de emisión de rayos X de cada televisor, la **capacidad** de almacenamiento de un disco rígido, la **longitud** de un tornillo, la **duración** o el consumo de una lamparita.



No confundir una **unidad** muestral como **objeto** completo y diferenciado que se encuentra dentro de un conjunto (una docena tiene doce unidades) con las unidades que se utilizan para valorar una magnitud (el metro es una unidad de longitud).

□ 5.3 Población

Para cualquier pregunta que interese responder, primero es necesario dirigir la atención a un **grupo particular de unidades** muestrales: personas, ciudades, animales, televisores, discos rígidos, tornillos o lamparitas.

- ¿Qué piensan los porteños sobre el Sistema de Evaluación Permanente de Conductores?
- ¿Qué porcentaje de familias de la ciudad de Santa Fe tienen mascotas?
- ¿Cuál es la expectativa de vida de los diabéticos?
- ¿Qué porcentaje de todos los tubos de pasta dentífrica son llenados de acuerdo a sus especificaciones?
- ¿Cuál es la duración promedio de las lámparas de bajo consumo de una determinada marca?
- ¿Los jóvenes deportistas consumen menos alcohol que los sedentarios?

En cada uno de los ejemplos, se plantea una pregunta y se puede identificar uno o más grupos específicos de unidades que interesa estudiar: los porteños (habitantes de la ciudad de Buenos Aires), las familias de la ciudad de Santa Fe, los diabéticos, los tubos de pasta dentífrica, las lámparas de bajo consumo, los deportistas y los sedentarios.

Se llama **población a todo el grupo de unidades muestrales** (generalmente son individuos) que interesa estudiar con el fin de responder una pregunta de investigación. Las poblaciones, sin embargo, pueden ser difíciles de definir. En un buen estudio, los investigadores deben **definir la población con toda claridad**.

La pregunta respecto a si los jóvenes que practican deportes consumen menos alcohol, sirve de ejemplo para ver lo difícil que puede ser definir con precisión la población. ¿Cómo definir un joven? ¿Los menores de 18 años de edad? ¿Los menores de 30 años? ¿Cómo definiría un sedentario? ¿Interesa estudiar los jóvenes de la República Argentina

o los de todo el mundo? Los resultados pueden ser diferentes para los menores de 18 que para los mayores, para los latinoamericanos comparados con los europeos, y así otras clasificaciones.

Muchas veces, los investigadores quieren estudiar y sacar conclusiones sobre una **población amplia** pero, con el fin de ahorrar tiempo, dinero, o simplemente porque no se les ocurre nada mejor, sólo estudian una **población muy restringida**. Esto puede conducir a serios problemas al momento de sacar conclusiones.

Supongamos que un profesor universitario quiere estudiar si los jóvenes que practican deportes consumen menos alcohol. Basa su estudio en un grupo de sus alumnos, que participan porque al hacerlo se les da cinco puntos adicionales en su puntaje final. Este grupo de alumnos constituye una muestra; pero los resultados no pueden generalizarse a toda la población de jóvenes, ni siquiera a todos los estudiantes.

□ 5.4 Muestra

¿Qué hacemos para probar la sopa? Revolvemos la olla con una cuchara, sacamos una porción -una muestra- la saboreamos y sacamos una conclusión sobre toda la sopa de la olla sin haber en realidad probado toda. Si la muestra ha sido tomada adecuadamente -sin elegir tramposamente la parte buena- tendremos una buena idea del sabor de la totalidad de la sopa. Esto se hace en estadística, más específicamente en **inferencia estadística**.

Los investigadores quieren averiguar algo sobre una población, pero no tienen tiempo o dinero para estudiar a todos los individuos que la conforman. Por lo tanto, ¿qué hacen? Seleccionan una **cantidad pequeña de unidades muestrales de la población** (esto se llama una **muestra**), estudian esas unidades, generalmente individuos, y utilizan esa información para sacar conclusiones sobre toda de la población.

5.4.1 Muestra representativa

Nos interesa obtener “**buenas muestras**”.



Una buena **muestra** debe ser **representativa** de la población. Esto significa, que todas las características importantes de la población tienen que estar en la muestra en **la misma proporción que en la población**.

Una muestra tiene, en pequeño y lo más parecidas posibles, las características de la población.

Podremos sacar conclusiones respecto de la población total a partir de una muestra -esto es, realizar una inferencia-, para todas aquellas características en las cuales la muestra representa a la población.

Ejemplo:

Consideremos un ejemplo simple, una población constituida por personas que difieren entre sí en una única característica con dos categorías:

Característica: el peso

Categorías: gordo, flaco

La figura 5.1 muestra una población hipotética que tiene 18 individuos que son gordos, o flacos (no hay gente con peso normal en esta población) y una muestra representativa.

En la figura 5.1 podemos ver la población total y la muestra. Respecto de la población, podemos decir que: 5 de cada 9 personas son gordas. Esta relación se repite en la muestra representativa.

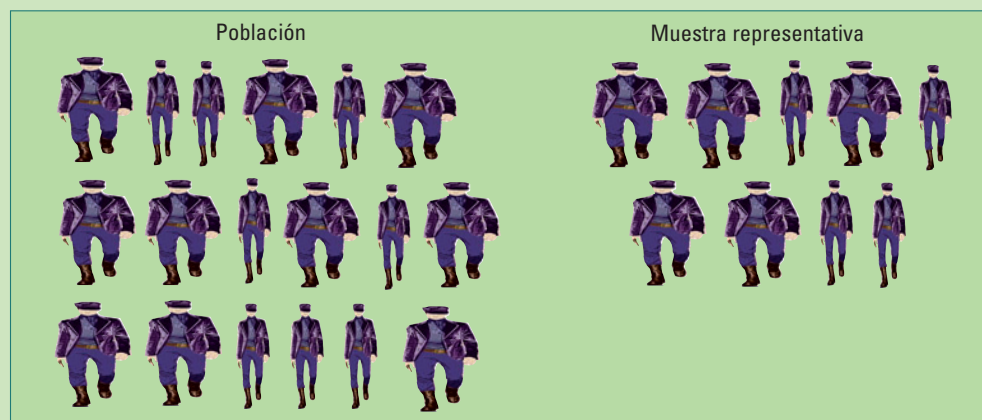


Figura 5.1 Muestra representativa de una población que sólo tiene una característica, el peso, con dos posibilidades: gordo y flaco.

En la vida real, es muy difícil que una muestra tenga proporciones idénticas a las poblacionales, pero **deberían ser muy parecidas** en todas las características que se puedan conocer.

Si se quiere realizar un estudio para averiguar a qué edad caminan los bebés de la Argentina, la muestra debería tener una distribución geográfica, por provincia o región, similar a la del último censo disponible de población. Si se considera que hay otros factores que pueden influir además de la región, por ejemplo el tipo de vivienda (casa o departamento), la muestra también tendrá que ser representativa de esos otros factores.