

INTRODUCCION

Estadística: Es una disciplina tecnológica que engloba varias herramientas (mayormente matemáticas) para la toma de decisiones a partir de información generada a través de datos.



A esta disciplina podemos dividirla en dos ramas: la Estadística descriptiva por un lado, y la Estadística inferencial por el otro.

Ejemplo: Un profesor de historia calcula la calificación promedio de uno de sus grupos. Como la estadística describe el desempeño del grupo pero no hace ninguna generalización acerca de los diferentes grupos podemos decir que el profesor está utilizando estadística descriptiva. Los gráficos, tablas y diagramas que muestran los datos de manera más clara y elocuente son ejemplos de estadística descriptiva. Ahora bien, si el profesor decide utilizar el promedio de calificación obtenido por uno de sus grupos en una unidad para estimar la calificación promedio del grupo en las 10 unidades del curso, se trataría de estimación concerniente a la estadística inferencial. Cualquier conclusión a la que llegue sobre el grupo de las 10 unidades estará basada en una generalización que va más allá de los datos en la unidad evaluada, y esta puede ser no completamente válida, de modo que el profesor debe establecer que probabilidad hay de que sea cierta.

Dato: Son los valores asociados con un rasgo o con una propiedad que sirven para distinguir la ocurrencia de algo. Un valor se convierte en un dato siempre y cuando yo lo relacione con la variable que estoy estudiando, es decir, siempre y cuando lo contextualice en el experimento.

Ejemplo: “Levine” y “Krehbiel” son valores que ayudan a distinguir un autor de un libro de otro autor. Si la variable de estudio de mi experimento está definida como “Nombres de autores de libros” estos valores son datos, pero, si en cambio mi experimento es sobre la edad de los alumnos del curso de estadística de la UBA estos valores, Levine y Krehbiel, no son datos.

Variable: Es una característica de un objeto o individuo.

Ejemplo: “Levine” y “Krehbiel” son valores para una variable que podría denominarse nombre o nombre del autor.

Una variable puede ser *cualitativa* o *cuantitativa*:

Las variables cualitativas o categóricas son aquellas que expresan características y que no pueden ser expresadas con números. Por ejemplo: series de Netflix más vistas en Argentina, El color de ojos de los actores de una película, etc. Sin embargo, puede ocurrir que una respuesta categórica se disfraze de una respuesta numérica: a veces aparecen “números”. Por ejemplo: al momento de realizar una encuesta se debe responder con 0 para muy malo, 1 para bueno y 2 para muy bueno. Esos aparentan ser números pero no se los puede operar numéricamente, por eso, en el caso del ejemplo, la variable sería cualitativa ordinal.

Las variables cuantitativas o numéricas son aquellas variables que se expresan mediante un número, por lo tanto se pueden realizar operaciones matemáticas con ellas. Por ejemplo: El peso de las vacas de un campo, La estatura de los ciudadanos, El número de personas que visitan una página web en un día.

A su vez, dependiendo de qué tipo de escala de medición puede utilizarse para la variable de estudio, podemos clasificar a las variables en:

- Variables cualitativas medibles en escala nominal: Cuando los datos cualitativos se presentan por como esta dada la respuesta se mide en escala nominal. Son aquellas variables que presentan valores no numéricos y en las que no existe un orden. Ejemplo: El estado civil de una persona; los valores serian: soltero, casado, divorciado, viudo. Otro ejemplo podría ser el lugar de nacimiento de tus amigos.
- Variables cualitativas medibles en escala ordinal: Cuando los datos cualitativos pueden ser ordenados, a dichos datos les corresponde la escala ordinal de medición. Ejemplo: La encuesta de un local de comida rápida: ¿Qué te pareció la comida? Muy mala, mala, regular, buena, muy buena.
- Variables cuantitativas medibles en escala de intervalo: Cuando los datos provienen de procesos de medición, donde los elementos forman un conjunto infinito no numerable y el cero significa ausencia de valor, a la variable cuantitativa se la denomina continua y se la puede medir en escala de intervalo. Ejemplo: temperatura medida en grados centígrados, altura en centímetros de una persona.
- Variables cuantitativas medibles en escala de razón: Cuando los datos provienen de procesos de conteo, a la variable cuantitativa se la denomina discreta y se la puede medir en escala de razón siempre y cuando el cero signifique ausencia de valor. No necesariamente los valores de una variable discreta son números enteros (o naturales), sino que lo que importa es que todos los resultados constituyan un conjunto numerable (finito o infinito).

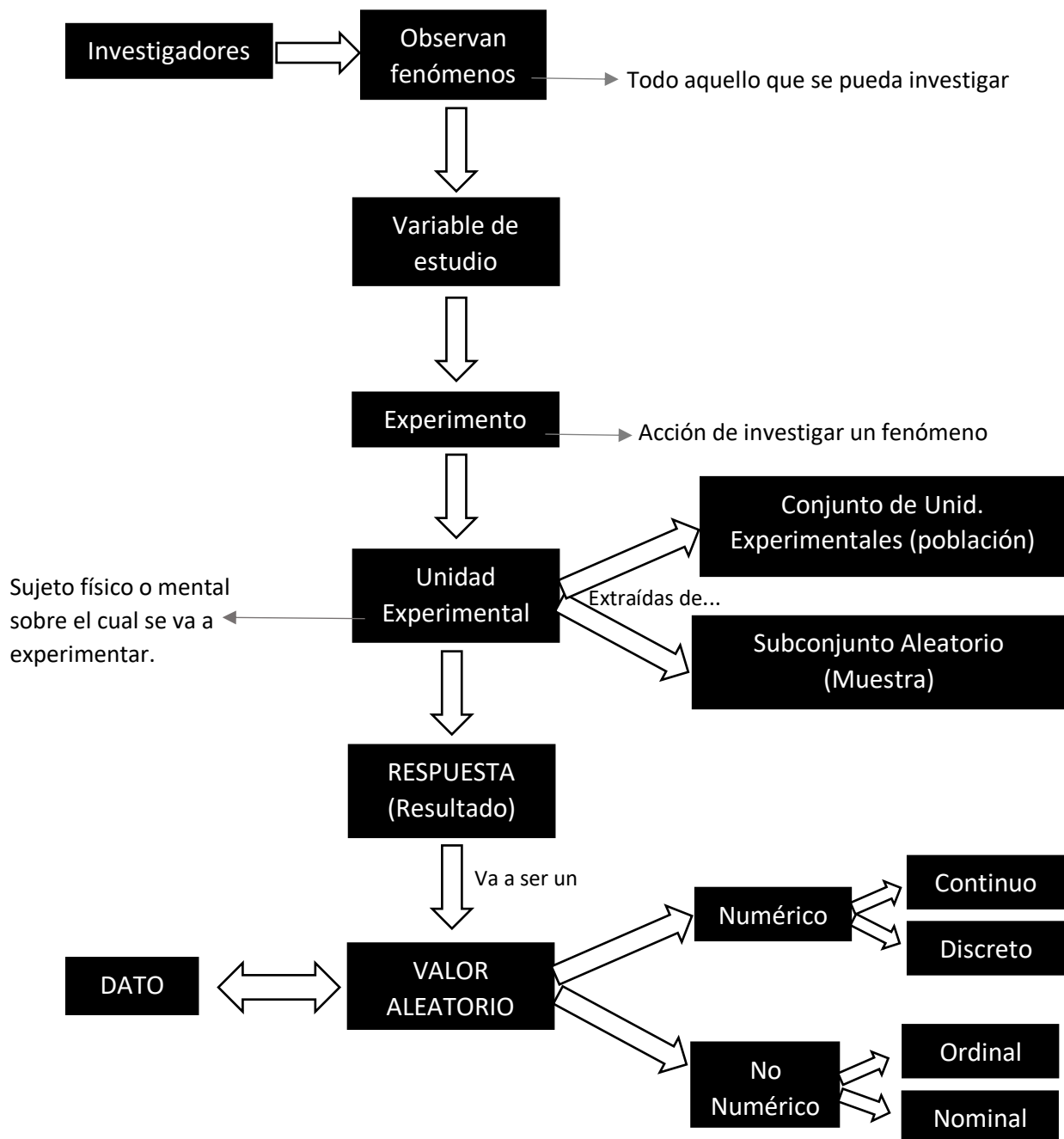
Las propiedades de las escalas de medición mencionadas son acumulativas; es decir, las propiedades de la escala nominal las tiene la ordinal, las tiene la de intervalo y todas las de intervalo las tiene la de razón.

Escala nominal —> Etiqueta datos.

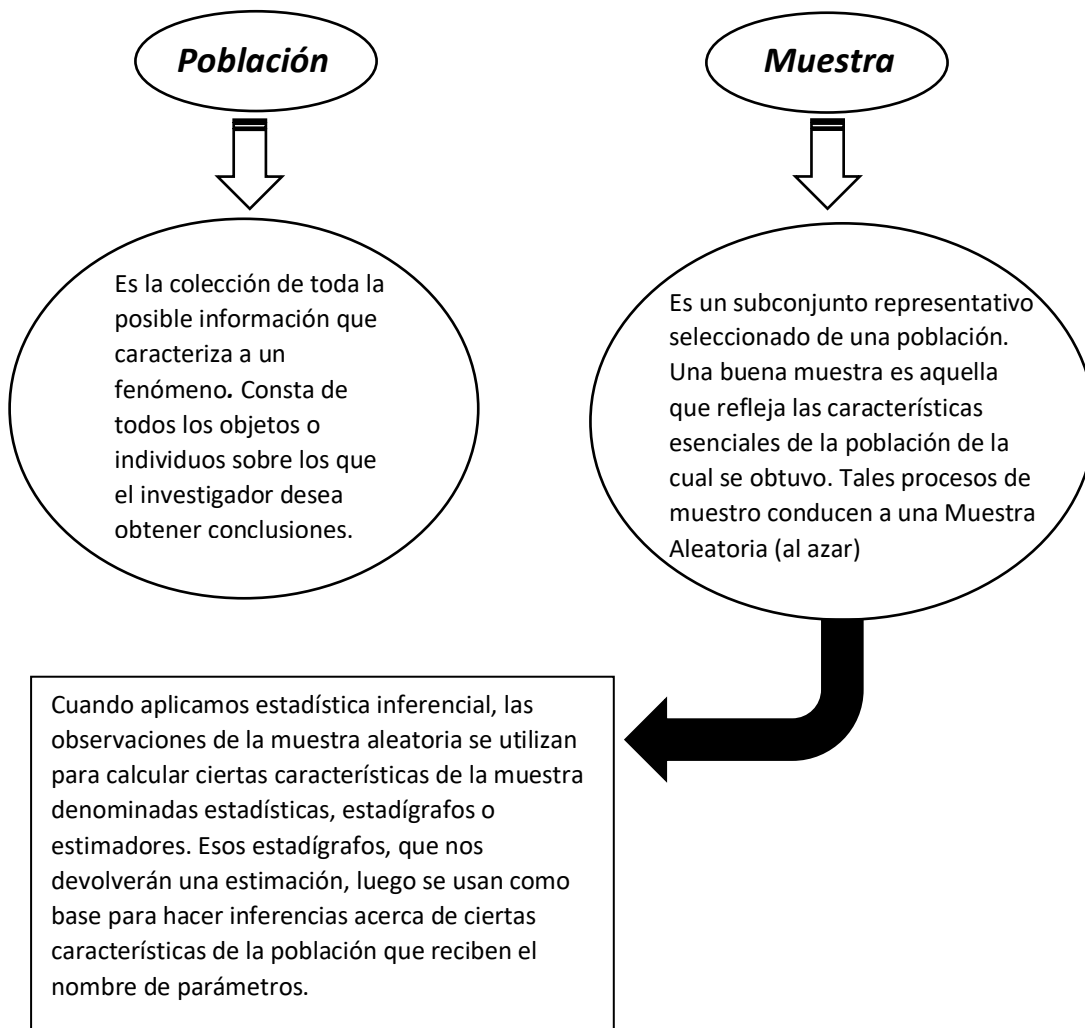
Escala ordinal —> Etiqueta y ordena datos.

E. de intervalo —> Etiqueta, ordena y opera matemáticamente datos, pero no puede establecer proporción entre ellos.

Escala de razón —> Etiqueta, ordena, opera matemáticamente y establece proporción entre los datos.



La estadística descriptiva se ubica en la última parte del diseño del experimento y nos brinda herramientas para recolectar, presentar y resumir datos. Se aplica tanto a la población (censo) como a la muestra (encuesta muestral).



Espacio de Variabilidad: *Es el conjunto de todos los distintos valores que puede tomar la variable, tanto sea numérica como no numérica. Este espacio de variabilidad lo puedo definir ex post (a posteriori), o ex ante (a priori), dependiendo de las características físicas del experimento.*

Ex post → Una vez que se tienen los valores se modifica el experimento. Ejemplo: estoy haciendo un experimento en el que veo el color de los autos que pasan. Luego de ver los autos que pasaron, y de anotar sus respectivos colores, como investigador decido agrupar los autos que eran azul oscuro y azul claro en una sola categoría; azul. En ese caso modifique el espacio de variabilidad ex post, realizando una reconfiguración del experimento.

Ex ante → No modifico el experimento o lo modifico antes de tener los datos.
Ejemplo: Las caras de un dado son 6, tengo un espacio de variabilidad dado donde no puedo sacar ni poner nada.

SIMBOLOGIA

X → Variable de estudio (se representa con cualquier letra mayúscula).

x_i → Observación i-esima (la misma letra con la que definí la variable de estudio pero en minúscula y un subíndice i).

N → Tamaño poblacional.

n → Tamaño muestral.

Ω → Espacio de variabilidad.

$\# \Omega$ → Cardinal del espacio de variabilidad. (Cantidad de elementos que tiene el espacio de variabilidad).