

Introducción a la estadística

Introducción

Eva Romero Ramos
evarom03@ucm.es

Universidad Complutense de Madrid

- 1 Concepto de estadística
- 2 Étaps de un estudio estadístico
- 3 Algunos conceptos básicos
- 4 Tipos de datos

Concepto de estadística

Podemos definir a la estadística como *la parte de las matemáticas que proporciona las herramientas necesarias para analizar, entender e interpretar un conjunto de datos.*

Para ello, la estadística abarca desde el planteamiento del problema y la recogida, organización y análisis de los datos hasta la obtención de las conclusiones finales.

La estadística es aplicable a la mayoría de las ciencias, ya que en todas ellas existen datos que nos interesará analizar de forma objetiva.

Podemos aplicarla a ciencias como la **medicina**, la **sociología**, el **marketing**, la **ingeniería** o las **finanzas** entre otras.

Concepto de estadística

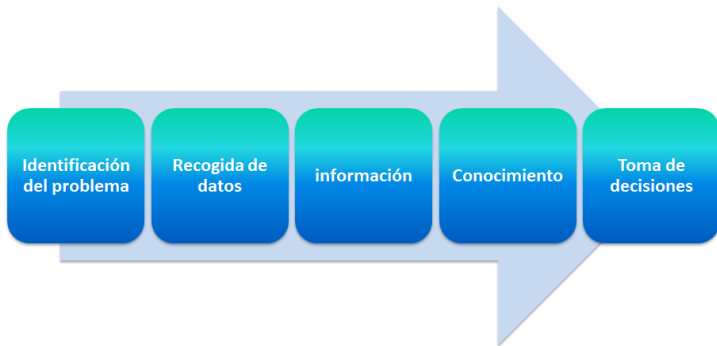
En el ámbito de la informática, la estadística está cobrando cada día más relevancia, ya que las herramientas informáticas facilitan el tratamiento de datos.

Muchos ingenieros en informática en la actualidad trabajan aspectos muy relacionados con la gestión de los datos, el almacenamiento o la organización de flujos de datos, ya que los datos se han convertido en un importante activo para las empresas y su gestión en un elemento clave para el crecimiento y posicionamiento empresarial.

La estadística nos proporciona herramientas para entender, analizar e interpretar diferentes tipos de datos, lo que nos proporcionará la información necesaria para tomar decisiones en ambientes de incertidumbre.

Étapas de un estudio estadístico

Podemos decir que las etapas que tiene un estudio estadístico son las siguientes:



Étapas de un estudio estadístico

Un estudio estadístico comienza con la *identificación de un problema a resolver* mediante el análisis de datos.

Es importante tener claro qué queremos analizar y definir adecuadamente los parámetros del problema.

Una vez definidos podemos pasar a la *recogida de datos*.

La recogida de datos es un proceso difícil de definir, ya que en la mayoría de los casos no tendremos acceso a todos los individuos objeto de estudio y tendremos que conformarnos con seleccionar una muestra.

Para seleccionar una muestra adecuadamente es necesario recurrir al *procedimiento de muestreo* que se adecue mejor a cada situación específica.

La mayoría de los procedimientos de muestreo que existen recurren a la aleatoriedad para formar la muestra de un modo u otro.

Étapas de un estudio estadístico

Una vez que tenemos la muestra podemos pasar al siguiente paso al que hemos denominado *información*.

Este paso consiste en calcular las medidas de la estadística descriptiva que sean de nuestro interés en cada caso.

Es recomendable calcular medidas de *posición*, de *dispersión* y de *forma* para conocer en profundidad el comportamiento de la variable considerada para los datos de la muestra.

Étapas de un estudio estadístico

Con la información de la muestra y aplicando el proceso de modelización adecuado podemos inferir información de toda la población en estudio.

Hemos denominado a este paso conocimiento, porque *nos permite pasar de la información de la muestra al conocimiento de la población que es de nuestro interés.*

Algunos conceptos básicos

Definición.- Población

En estadística se denomina población al conjunto de todos los individuos que son objeto de estudio y sobre los que se desea obtener información.

Llamaremos N al tamaño de la población y buscaremos tener información para los N individuos incluidos en la población.

Definición.- Muestra

Una muestra es el conjunto de individuos a cuya información realmente tendremos acceso dentro de la población.

Debe ser un buen reflejo de la población de la que proviene. Si es así diremos que la muestra es **representativa**.

Denominaremos n al tamaño muestral, y a partir de la información de los n individuos de la muestra, tendremos que inferir el comportamiento de la población de N individuos.

Algunos conceptos básicos

Definición.- Elemento

Un elemento es cada uno de los componentes de la población en estudio.

Los elementos a menudo se denominan individuos, aunque las poblaciones pueden tener elementos de distintas naturalezas (personas, empresas, países, productos ...).

Definición.- Variable

En estadística se denomina variable a cualquier característica del elemento o ítem en estudio.

Las variables estadísticas no son siempre numéricas y pueden ser también de carácter cualitativo. A las variables cualitativas las denominaremos **atributos** o simplemente **variables cualitativas**.

Algunos conceptos básicos

Las variables cuantitativas a su vez pueden ser **continuas** o **discretas**:

- Las variables **continuas** puede tomar infinitos valores dentro de un intervalo.
- Las variables **discretas** tomarán valores dentro de un número finito de valores. En este caso podremos hablar de valores consecutivos.

Definición.- Dato

Un dato es el valor concreto que toma una variable para un individuo de la muestra. Dependiendo de la naturaleza de la variable que consideremos los datos pueden ser numéricos o no.

Cuando el dato es número hablaremos de su **valor**, pero si el dato esta asociado a un atributo hablaremos de su **modalidad**.

Dependiendo de su naturaleza el conjunto de datos puede ser:

- Los **datos históricos** o series temporales nos ofrecen información para una variable y para un mismo elemento o ítem a lo largo del tiempo.
- Los **datos de corte transversal** nos ofrecen información sobre una variable para diferentes elementos.
- Los **datos de panel** nos ofrecen información sobre una variable para diferentes elementos a lo largo del tiempo.

- **Escala nominal:** la información se puede clasificar en categorías no numéricas, mutuamente excluyentes, entre las cuales no hay ninguna relación de orden. Son datos correspondientes a atributos (modalidades).
- **Escala ordinal:** la información se puede clasificar en categorías no numéricas, entre las cuales hay algún tipo de orden. Son datos correspondientes a atributos (modalidades).
- **Escala de intervalos:** es una escala cuantitativa, las observaciones tienen una unidad de medida y podemos cuantificar la distancia entre dos observaciones. El cero es arbitrario y la proporción entre dos datos no tiene sentido. Son datos correspondientes a variables (valores).

- **Escala de proporción**: es una escala cuantitativa, las observaciones tienen una unidad de medida y podemos cuantificar la distancia entre dos observaciones, además tiene sentido fijar un punto de origen que marque el cero absoluto. Son datos correspondientes a variables (valores).

Según el tipo de escala, elegimos los métodos de estadística más adecuados.