

Estadística descriptiva de una variable

6 de mayo de 2009

Definiciones básicas

La estadística descriptiva trata de resumir la información existente en un conjunto de datos. Cuando se recoge información de interés conviene tener presentes las siguientes definiciones básicas:

- ▶ **Población:** Conjunto sobre el que se va a realizar el estudio. A cada elemento de la población se le denomina individuo.
- ▶ **Muestra:** Subconjunto de individuos de la población seleccionados con la finalidad de generalizar las conclusiones obtenidas en su análisis a toda la población. El número de individuos de una muestra se denomina tamaño muestral.

Cuando se realice un análisis descriptivo de un conjunto de datos no se distinguirá entre población y muestra. Se supondrá implícitamente que se dispone la información de todos los individuos de la población (censo).

Tipos de variables

Definición

Una variable es Variable es una característica que puede ser medida en cada uno de los individuos de la población.

Según su naturaleza, distinguiremos los siguientes tipos de variables:

- ▶ **Cuantitativas:** Son aquellas características de los individuos que tienen naturaleza numérica. Dentro de las variables cuantitativas distinguiremos dos tipos:
 - ▶ **Variables discretas:** Cuando el conjunto de posibles valores de la característica observada es finito o numerable como, por ejemplo, el número de hermanos de una persona o el número de clientes de un supermercado durante una hora.
 - ▶ **Variables continuas:** Cuando, al menos en teoría, el conjunto de posibles valores de la característica observada es infinito no numerable como, por ejemplo, la altura o el peso.
- ▶ **Cualitativas:** Serán aquellas variables que no tienen naturaleza numérica como, por ejemplo, el género de una persona, su religión, equipo, etc.

Distribución de frecuencias: frecuencias absolutas

Supongamos que hemos medido en un grupo de individuos una característica **numérica**. A esta variable la denotaremos con una letra mayúscula, X , y a los valores registrados con letras minúsculas

$$x_1 < x_2 < \cdots < x_k.$$

Los siguientes conceptos permiten estructurar la información disponible:

- ▶ **Frecuencia absoluta:** La frecuencia absoluta del dato x_i , $i = 1, \dots, k$, es el número de veces que se repite el valor x_i , y la denotaremos por n_i . Así, el tamaño de la muestra, que denotaremos por n , verificará

$$n = n_1 + n_2 + \cdots + n_k.$$

Si una variable es continua entonces, si no hay redondeos, no deberíamos tener valores repetidos. En este caso $k = n$ y $n_i = 1$, para todo $i = 1, \dots, n$.