

Teoría – Tema 10

Estadística unidimensional

¿Qué es una variable estadística?

Supongamos que medimos, en una clase de Educación Secundaria, la altura de los alumnos en centímetros. Es lógico pensar que habrá alumnos más altos, otros más bajos, alumnos que medirán igual, etc. El valor de la altura x es una variable, que en cada alumno tomará un valor concreto x_i .

Los factores que afectan a esta variable x son muchos: la genética, la alimentación, el clima, el estado de salud... Sí es claro que el valor de x varía (de ahí el nombre de variable) dentro de la población que hemos tomado (una clase de Educación Secundaria). Y el conjunto de valores x_i de cada alumno nos ofrece información estadística (en término medio) del valor de x .

Si trabajamos con una única variable estadística hablamos de estadística unidimensional, donde se definen los siguientes parámetros:

- **Tamaño de la muestra:** número de medidas realizadas. Se representa por la letra N .
- **Recorrido:** Diferencia entre el mayor y el menor valor obtenido en los valores x_i .
- **Media aritmética:** Suma de todos los valores x_i dividido por el número de muestras tomadas.

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{N}$$

- **Frecuencia absoluta del valor x_i :** Número de veces que aparece el valor x_i . Se representa como f_i . La suma de todas las frecuencias absolutas coincide con el tamaño N de la muestra.

$$\sum_{i=1}^n f_i = N$$

- **Frecuencia relativa del valor x_i :** Es el cociente de cada frecuencia absoluta entre el número de muestras tomadas.

$$n_i = \frac{f_i}{N}$$

El valor de la frecuencia relativa n_i viene dado en tanto por 1. Si lo multiplicamos por 100 lo tendremos en tanto por ciento (%).

- **Frecuencia absoluta acumulada del valor** x_i : Suma de las frecuencias absolutas de los valores de la variable menores o iguales a x_i .

$$F_i = f_1 + f_2 + f_3 + \dots + f_i = \sum_{j=1}^i f_j$$

- **Frecuencia relativa acumulada del valor** x_i : Es la suma de las frecuencias relativas de los valores de la variable menores o iguales a x_i .

$$N_i = n_1 + n_2 + n_3 + \dots + n_i = \sum_{j=1}^i n_j$$

El valor de la frecuencia relativa acumulada N_i viene dado en tanto por 1. Si lo multiplicamos por 100 lo tendremos en tanto por ciento (%).

- **Moda**: Valor de la variable x_i con mayor frecuencia absoluta.
- **Mediana**: El menor valor de la variable x_i que acumula el 50% de la frecuencia relativa acumulada en %.
- **Varianza**: La media aritmética de los cuadrados de las desviaciones respecto a la media. La varianza siempre es un valor positivo o nulo. Una varianza muy pequeña indica que el valor medio \bar{x} es un buen representante del conjunto de los valores particulares x_i , y viceversa. Se representa por s^2 o por σ^2 .

$$s^2 = \frac{f_1(x_1 - \bar{x})^2 + f_2(x_2 - \bar{x})^2 + \dots + f_n(x_n - \bar{x})^2}{f_1 + f_2 + \dots + f_n} = \frac{\sum_{i=1}^n f_i(x_i - \bar{x})^2}{N}$$

- **Desviación típica**: Es la raíz cuadrada positiva de la varianza. Se representa por s o por σ .