



Educaguía
.com

ESTADÍSTICA

MEDIDAS DE CENTRALIZACIÓN

MEDIDAS DE CENTRALIZACIÓN

Las medidas de centralización son estadísticos que reflejan algún valor global de la serie estadística.

Las principales medidas de centralización son:

- Media aritmética simple.
- Media aritmética ponderada.
- Mediana.
- Moda.
- Percentiles.
- Media geométrica.
- Media armónica.
- Media cuadrática.

Las medidas de centralización deben de cumplir los siguientes requisitos:

- Su valor debe estar comprendido entre los valores mínimo y máximo de la serie.
- En su cálculo deben intervenir todos y cada uno de los valores de la serie.

Media aritmética simple

La media aritmética simple es igual al cociente que resulta de dividir la suma de todos los datos de la serie entre el número de ellos.

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum_1^n x_i}{n}$$

Cuando alguno de los valores de la serie aparecen repetidos, la media aritmética es igual a la suma de los productos de cada valor de la variable por el número de veces que el mismo se repite, dividida entre el número

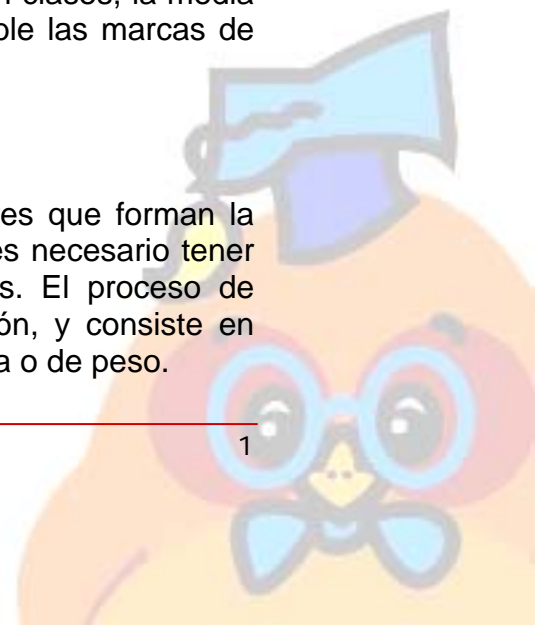
$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + x_3 f_3 + \dots + x_n f_n}{N} = \frac{\sum_1^n x_i f_i}{N}$$

total de observaciones.

Cuando el recorrido de la variable estadística se divide en clases, la media aritmética se obtiene tomando como valores de la variable las marcas de clase.

Media aritmética ponderada

Cuando queremos hallar la media aritmética y los valores que forman la serie estadística no tienen todos la misma importancia, es necesario tener en cuenta la importancia relativa de los distintos datos. El proceso de valorar la importancia de los datos se llama ponderación, y consiste en asignar a cada valor un coeficiente llamado de importancia o de peso.



$$\bar{x} = \frac{x_1 f_1 k_1 + x_2 f_2 k_2 + \dots + x_n f_n k_n}{f_1 k_1 + f_2 k_2 + \dots + f_n k_n} = \frac{\sum_1^n x_i f_i k_i}{\sum_1^n f_i k_i}$$

Mediana

Se considera mediana al valor que ocupa el lugar central, cuando los valores de la serie están ordenados, es decir, cuando el número de datos superiores e inferiores coinciden.

Si existiesen dos elementos centrales, por tratarse de un N par, la mediana se calcularía con la media aritmética de los valores centrales correspondientes.

El método más práctico para encontrar la mediana es calcular la columna de las frecuencias acumuladas, y buscamos un valor N_i mayor que $N/2$, el valor obtenido será la media.

Cuando las distribuciones vienen dadas por intervalos, puede hallarse la mediana por cada uno de los métodos siguientes:

- Transformar la distribución mediante la marca de clase y hallar la mediana como se explicó anteriormente.
- Una vez hallada la columna de las frecuencias acumuladas, calculamos la fila superior a $N/2$ y después aplicamos la siguiente fórmula en el intervalo correspondiente a la misma.

$$M_e = L_{i-1} + c \frac{\frac{N}{2} - F_{i-1}}{f_i}$$

Donde L_{i-1} es el límite inferior del intervalo en cuestión, F_{i-1} la frecuencia acumulada anterior a la clase mediana, c amplitud de la clase mediana y f_i frecuencia absoluta de la clase mediana.

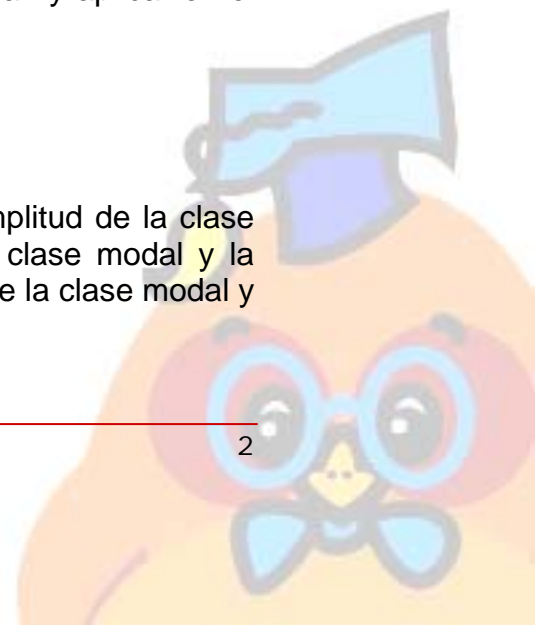
Moda

La moda es el valor de la variable que más veces se repite, es decir el que tiene la máxima frecuencia. Cuando la serie viene dada por intervalos pueden seguirse también dos caminos:

- Transformarla mediante la marca de clase y comprobar quien tiene la máxima frecuencia.
- Observar el intervalo de mayor frecuencia "clase modal" y aplicar en él la siguiente fórmula:

$$M_o = L_{i-1} + c \frac{D_1}{D_1 + D_2}$$

Donde L_{i-1} es el límite inferior de la clase modal, c la amplitud de la clase modal, D_1 diferencia entre la frecuencia absoluta de la clase modal y la clase anterior, D_2 diferencia entre la frecuencia absoluta de la clase modal y la clase siguiente.



Percentiles

Percentil de una serie estadística es el valor que verifica la propiedad de que el % del total de valores de la serie son inferiores a él. Se representa por p_m .

Relacionando este concepto con el de mediana, es evidente que el percentil 50 de la serie es la mediana, ya que por ocupar el valor central el 50% de los valores son inferiores a ella.

$$p_m = L_{i-1} + c \frac{\frac{mN}{100} - F_{i-1}}{f_i}$$

Al igual que existen los percentiles existen también los cuartiles, quintiles, deciles,... La forma de resolverlos es la misma que la de los centiles, con la salvedad de que en lugar de estar dividido por cien está dividido por cuatro, cinco, diez,... según el cuartil que se calcule.

Media Geométrica

Es la raíz de orden N-ésimo del producto de los N valores pertenecientes a la serie estadística.

$$G = \sqrt[N]{x_1 x_2 x_3 \dots x_n}$$

Cuando alguno o algunos de los valores de la serie aparecen repetidos, la media geométrica es:

$$G = \sqrt[N]{x_1^{f_1} x_2^{f_2} x_3^{f_3} \dots x_n^{f_n}}$$

Cuando la variable viene dada por intervalos, la media se calcula tomando como valores de la variable, la marca de clase.

Media armónica

Es el recíproco de la media aritmética de los valores recíprocos de la

$$H = \frac{N}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_N}}$$

variable.

Cuando los valores aparecen repetidos:

$$H = \frac{N}{\frac{1}{x_1} f_1 + \frac{1}{x_2} f_2 + \dots + \frac{1}{x_n} f_n}$$



Media cuadrática

Es la raíz cuadrada de la media aritmética de los valores de la variable elevados al cuadrado.

$$C = \sqrt{\frac{\sum_1^N x_i^2}{N}} \quad C = \sqrt{\frac{\sum_1^n x_i^2 f_i}{N}}$$

Cuando los valores están repetidos, se aplicará la segunda fórmula.

