

## Formulario de Estadística

Dado el cuadro de datos:

Clases	$x_i$	$f_i$	$F_i$	$h_i$ %	$H_i$ %
[3.3, 3.8)	3.55	3	3	7.5 %	7.5 %
[3.8, 4.3)	4.05	8	11	20 %	27.5 %
[4.3, 4.8)	4.55	14	25	35 %	62.5 %
[4.8, 5.3)	5.05	6	31	15 %	77.5 %
[5.3, 5.8)	5.55	4	35	10 %	87.5 %
[5.8, 6.3)	6.05	5	40	12.5 %	100 %

Donde:

- $x_i$ : Marca de clase (valor del dato).
- $f_i$ : Frecuencia absoluta.
- $F_i$ : Frecuencia acumulada.
- $h_i$ : Frecuencia relativa.
- $H_i$ : Frecuencia acumulada relativa.

Considerese:  $n = 40$  (Total de datos)

Marca de Clase:  $x_i$  \_\_\_\_\_

Donde: 
$$x_i = \frac{L_i + L_s}{2}$$

- $L_i$ : Límite inferior de la clase.
- $L_s$ : Límite superior de la clase.

**Nota:** La marca de clase es el valor del dato, en el caso de datos agrupados y continuos se calcula de dicha manera, sin embargo, cuando los valores son discretos o cualitativos, la marca de clase es el valor del dato en sí mismo.

Frecuencia Absoluta:  $f_i$  \_\_\_\_\_

Donde: 
$$f_i = \sum_{i=1}^n x_i$$

- $n$ : Total de datos.
- $x_i$ : Marca de clase.

**Nota:** La frecuencia absoluta es la cantidad de veces que se repite un mismo dato de la muestra.

Frecuencia Relativa:  $h_i$  \_\_\_\_\_

Donde: 
$$h_i = \frac{f_i}{n} \times 100$$

- $f_i$ : Frecuencia absoluta.
- $n$ : Total de datos.

**Nota:** Se muestra la frecuencia relativa porcentual, generalmente la más usada. Puede que se solicite en algún caso particular la frecuencia relativa en forma decimal.

Construcción de Clases \_\_\_\_\_

$$R = x_{max} - x_{min}$$

$$k = 1 + 3,322 \times \log_{10}(n)$$

$$A = \frac{R}{k} = C$$

Donde:

- $R$ : Rango de la muestra.
- $x_{max}$ : Máximo valor de la muestra.
- $x_{min}$ : Mínimo valor de la muestra.
- $n$ : Total de datos.
- $k$ : Número de clases.
- $A, C$ : Amplitud o ancho de clase.

**Nota:** Esta parte del formulario indica el proceso que se debe tomar para determinar, dado un conjunto de datos continuos, la cantidad de clases y el ancho de las mismas, esto con el objetivo de calcular las medidas pertinentes.

Media Aritmética:  $\bar{x}$  \_\_\_\_\_

$$\bar{x} = \frac{\sum_{i=1}^n x_i \cdot f_i}{n}$$

Donde:

- $x_i$ : Marca de clase.
- $f_i$ : Frecuencia absoluta.
- $n$ : Total de datos.

**Nota:** La media aritmética es el promedio de los datos, se calcula multiplicando cada dato por su frecuencia y dividiendo la suma de estos entre el total de datos.

Mediana de Datos Agrupados \_\_\_\_\_

$$Me = L_i + \left( \frac{\frac{n}{2} - F_{i-1}}{f_i} \right) \cdot C$$

Donde:

- $L_i$ : Límite inferior de la clase.
- $F_{i-1}$ : Frecuencia acumulada anterior.
- $f_i$ : Frecuencia absoluta.
- $C$ : Ancho de clase.
- $n$ : Total de datos.

**Nota:** La mediana es el valor que divide a la muestra en dos partes iguales, se calcula con la fórmula anterior, donde se busca la clase que contiene la mediana y se despeja el valor de la misma. En esta fórmula se escoge la clase que contiene a la mediana escogiendo al  $F_i$  que contenga a  $\frac{n}{2}$

## Moda de Datos Agrupados

Donde: 
$$Mo = L_i + \left( \frac{f_i - f_{i-1}}{2f_i - f_{i-1} - f_{i+1}} \right) \cdot C$$

- $L_i$ : Límite inferior de la clase.
- $f_i$ : Frecuencia absoluta.
- $f_{i-1}$ : Frecuencia absoluta anterior.
- $f_{i+1}$ : Frecuencia absoluta siguiente.
- $C$ : Ancho de clase.

**Nota:** La moda es el valor que más se repite en la muestra, se calcula con la fórmula anterior, donde se busca la clase que contiene a la moda y se despeja el valor de la misma. En esta fórmula se escoge la clase que contiene a la moda aligiendo al  $f_i$  más grande.

## Variancia y Desviación Estándar

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot f_i}{n - 1}$$

Donde: 
$$S = \sqrt{S^2}$$

- $x_i$ : Marca de clase.
- $\bar{x}$ : Media aritmética.
- $f_i$ : Frecuencia absoluta.
- $n$ : Total de datos.

**Nota:** La variancia es una medida de dispersión que indica cuánto se alejan los datos de la media, se calcula con la fórmula anterior. La desviación estándar es la raíz cuadrada de la variancia.

## Coeficiente de Variación

Donde: 
$$CV = \frac{S}{\bar{x}} \times 100$$

- $S$ : Desviación estándar.
- $\bar{x}$ : Media aritmética.

Además:

- $CV < 10\%$ : Datos homogéneos.
- $10\% \leq CV < 30\%$ : Variabilidad aceptable.
- $CV \geq 30\%$ : Datos heterogéneos.

**Nota:** El coeficiente de variación es una medida de dispersión relativa, se calcula con la fórmula anterior, se expresa en porcentaje y se interpreta como el porcentaje de variabilidad que tiene la muestra en relación a la media.

## Cuantiles

$$Q_p = L_i + \left( \frac{\frac{p \cdot n}{\%} - F_{i-1}}{f_i} \right) \cdot C$$

Donde:

- $L_i$ : Límite inferior de la clase.
- $F_{i-1}$ : Frecuencia acumulada anterior.
- $p$ : Porcentaje del cuantil.
- $C$ : Ancho de clase.
- $n$ : Total de datos.
- $\%$ : Valor del cuantil (100, 10, 4)

**Nota:** Los cuantiles son valores que dividen a la muestra en partes iguales, se calculan con la fórmula anterior, donde se busca la clase que contiene al cuantil y se despeja el valor del mismo. En esta fórmula se escoge la clase que contiene al cuantil escogiendo la clase de  $F_i$  que contenga la posición dada por:  $\frac{p \cdot n}{100}$ .