



# Estadística Descriptiva para Datos Agrupados

Organizando el Caos: De Grandes  
Listas a Información Valiosa

# La Necesidad de Agrupar Datos

## ¿Por qué agrupar?

Cuando manejamos muestras grandes (más de 30 o 40 datos) con valores muy diversos, una lista simple no nos dice nada.

Agrupamos para encontrar **patrones**.

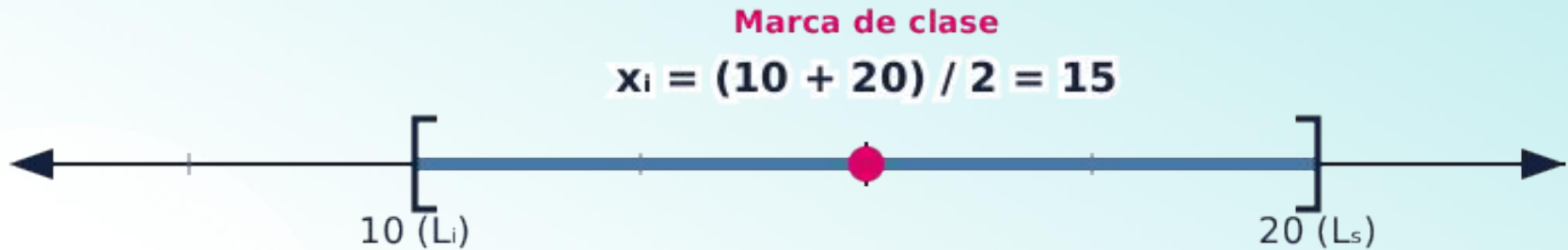
## Elementos Clave

- **Intervalo de Clase:** El 'rango' o cajón donde caen los datos.
- **Límites:** Cada intervalo tiene un **Límite Inferior (Li)** y un **Límite Superior (Ls)**.
- **Amplitud (A):** Es la diferencia entre los límites ( $A = Ls - Li$ ).



# La Marca de Clase (xi)

La marca de clase es el punto medio del intervalo. Es el valor que 'representa' a todos los datos que caen en ese grupo para poder hacer cálculos.



Se calcula sumando los límites y dividiendo para dos:  $x_i = (L_i + L_s) \div 2$



# Construyendo la Tabla de Frecuencias

1

## Frecuencia Absoluta ( $f_i$ )

Es el conteo de cuántos datos reales caen dentro de cada intervalo específico.

2

## Frecuencia Acumulada ( $F_i$ )

Suma sucesiva de frecuencias absolutas. Indica cuántos datos hay hasta ese punto.

3

## Frecuencia Relativa ( $h_i$ )

Proporción del total que representa cada intervalo ( $h_i = f_i \div N$ ). Útil para porcentajes.

# Desafío de Conceptos

1.

**Marca de Clase**

a) Representante o punto medio de un intervalo.

2.

**Amplitud**

b) El ancho o tamaño de cada intervalo de clase.

3.

**Límite Superior**

c) Número de veces que se repite un dato en un rango.

4.

**Frecuencia Absoluta**

d) El valor máximo permitido en un intervalo determinado.

# Desafío de Conceptos



1.

**Marca de Clase**

a) Representante o punto medio de un intervalo.

2.

**Amplitud**

b) El ancho o tamaño de cada intervalo de clase.

3.

**Límite Superior**

d) El valor máximo permitido en un intervalo determinado.

4.

**Frecuencia Absoluta**

c) Número de veces que se repite un dato en un rango.

# Cálculo de la Media Aritmética ( $\bar{x}$ )

Media aritmética

Sumatoria

Frecuencia absoluta

$\bar{x}$

=

$$\sum (f_i \cdot x_i)$$

Marca de clase

N

Número total de datos



# Mediana y Moda: Fórmulas de Precisión

Para datos agrupados, usamos fórmulas que consideran la posición dentro del intervalo:

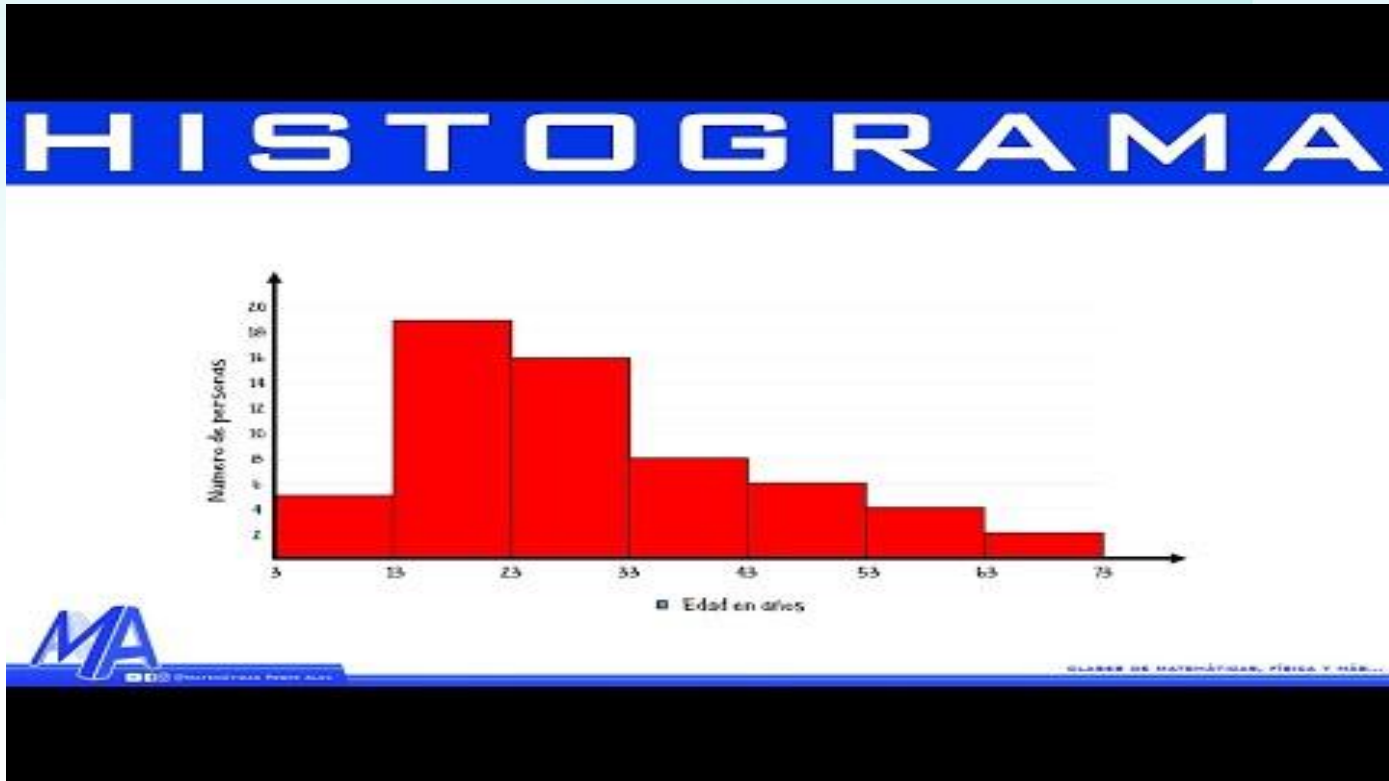
## La Mediana (Me)

Buscamos el intervalo donde la frecuencia acumulada ( $F_i$ ) supera el 50% de los datos ( $N/2$ ). Requiere  $L_i$ , amplitud y frecuencias.

## La Moda (Mo)

Es el intervalo con la frecuencia absoluta ( $f_i$ ) más alta. La fórmula interpola el valor exacto según diferencias con intervalos vecinos.

# Visualización de Datos Agrupados



# Histogramas vs. Polígonos



El Histograma usa barras pegadas. El ancho es la amplitud y el alto es la frecuencia.



El Polígono une los puntos medios (marcas de clase) para mostrar la tendencia.

# Resumen de la Unidad

## Puntos Clave para tu Examen

1. **Agrupar ahorra tiempo:** Convierte 1000 números en 7 u 8 filas manejables.
2. **La precisión importa:** Usa siempre la **Marca de Clase** para los cálculos de promedios.
3. **Interpretación:** Un histograma sesgado a la derecha nos dice que la mayoría de los datos son valores bajos.
4. **Verificación:** La suma de las frecuencias relativas siempre debe ser 1 (o 100%).

