

DISTRIBUCIÓN DE FRECUENCIAS

FILA DE DATOS. - Son los datos recolectados que no han sido organizados numéricamente, así por ejemplo los pesos de 50 trabajadores en orden del abecedario.

ORDENACIONES. - Es un conjunto de datos numéricos en orden ascendente o descendente, la deferencia entre el mayor y el menor se llama RANGO (AMPLITUD) de ese conjunto de datos. Así por ejemplo al hablar de los 50 trabajadores el peso mayor era 588 N. y el menor 220 N. Entonces el rango es $(588 - 220) N = 368 N$.

FRECUENCIA DE DATOS AGRUPADOS:

Una tabla de frecuencias para datos puede constar de: clases, límites reales, frecuencias absolutas, porcentuales, acumuladas, puntos medios (Marca de clase) $= \frac{ls+li}{2}$.

Para determinar estos elementos, cuando el número de datos es relativamente grande, se aconseja seguir el siguiente ALGORITMO:

- Se calcula la amplitud o rango: Amplitud = Máximo – mínimo (en datos discretos Máximo – mínimo +1)
- El número de clases C debe estar en lo posible entre 5 y 20; una fórmula que se utiliza es la siguiente.

$$C = 1 + 3.33 \cdot \log_{10} n, \text{ siendo } n \text{ el número total de datos. Sturges o la } \sqrt{n}$$

- El ancho del intervalo c se calcula así (cuando todas las clases van a tener el mismo ancho)

$$ai = \frac{\text{Amplitud}}{C}; \text{ donde } C \text{ es el número de clases}$$

- En cada clase se colocan los límites inferiores y superiores (simples o absolutos)
- En cada clase se colocan los límites reales inferiores y superiores (el límite real inferior es 0.5 menos del límite inferior simple, y el límite real superior es 0,5 más del límite superior simple, siempre que estos sean enteros, si estos tienen un decimal el límite real inferior es 0,05 más del límite superior simple y el límite real superior es 0,05 más del límite superior simple, y así sucesivamente).
- Se pone los puntos medios de cada clase que son la semisuma de los límites reales correspondientes.

- Se colocan las frecuencias (absolutas o simples) de cada clase, que son el número de observaciones que corresponden a cada clase.
- Se puede colocar las frecuencias acumuladas. La frecuencia acumulada de cada clase es la suma de todas las frecuencias inferiores incluida la de la clase.
- Se puede colocar las frecuencias relativas que son las frecuencias absolutas o simples de cada clase divididas para n. Se puede colocar también las frecuencias relativas acumuladas.
- En fin, se colocan las desviaciones o cualquier otro dato según los estadígrafos que vayamos a calcular.

Ejemplo:

22 23 14 26 11 18 21 20 22 19 14 19 21 26 27 24 23 21
 12 25 20 15 11 17 20 19 24 21 17 21 23 24 26 25 25 22
 13 15 10 18 16 15 19 28 23 18 27 25 22 20 17 16 13 20
 16 17 19 22 29 24 18 13 16 30 26 : n =65

- Amplitud (rango) = Max - Mín+1
- Numero de clases:
- Ancho del intervalo:

Límites reales	Límites	f _i	f%	Puntos Medios (X _i)	f% a
	10-12				
	13-15				
	16-18				
	19-21				
	22-24				
	25-27				
	28-30				
TOTAL					

MEDIDAS DE CENTRALIZACIÓN PARA DATOS AGRUPADOS

a) **MEDIA ARITMÉTICA.** - para calcular la media aritmética se debe multiplicar el punto medio de cada clase por la frecuencia absoluta de la misma, luego se suman todos los productos y se divide para el número total de datos:

$$\bar{X} = \frac{\sum f_i \cdot X_i}{n} = \frac{1306}{65} = 20.09$$

Calcule la media aritmética de los datos agrupados de la tabla anterior:

Clases	f_i	Puntos medios(X_i)	$f_i X_i$
10 -12			
13-15			
16-18			
19-21			
22-24			
25-27			
28-30			
Total: Σ			

b) **MEDIANA.** - Para calcular la mediana de datos agrupados, necesitamos primero localizar la clase mediana o llamado también intervalo mediano.

El intervalo mediano contiene a la primera frecuencia acumulada mayor o igual a $\frac{n}{2} = 32.5$

; luego se aplica la siguiente formula:

$$Md = L_i + \left[\frac{\frac{n}{2} - f_{a-1}}{fmd} \right] ai$$

De donde:

L_i = Límite real inferior de la clase mediana.

n = número total de datos

f_{a-1} = frecuencia acumulada anterior a la del intervalo mediano

fmd = frecuencia de la clase mediana

ai = ancho del intervalo = limite real superior – limite real inferior.

Ejemplo calcule la mediana de los datos agrupadas de la tabla de nuestro estudio:

Límites reales	Clases	f_i	$f \%$	Puntos medios	f_i acumulada

El intervalo mediano es el cuarto, porque la cuarta frecuencia acumulada es la mínima

mayor o igual a $\frac{n}{2} =$; además:

$L_i = f_{a-1} =$ $fmd =$ $c =$ luego:

$$Md = L_i + \left[\frac{\frac{n}{2} - f_{a-1}}{fmd} \right] ai \Rightarrow Md =$$

$Md =$

c) MODA. - para hallar la moda de datos agrupados, debemos ubicar la clase modal, que es la que tiene mayor frecuencia absoluta, luego aplicamos la fórmula.

$$Mo = L_i + \left[\frac{d_1}{d_1 + d_2} \right] ai$$

Donde:

d_1 = frecuencia absoluta de la clase modal – frecuencia absoluta de clase inferior.

d_2 = frecuencia absoluta de la clase modal – frecuencia absoluta de la clase superior.

ai = ancho del intervalo de clase.

si hay dos frecuencias más altas la distribución es bimodal.

Ejemplo: calcule la moda de los datos agrupados del ejercicio en clase

Limites reales	clases	f _i

La clase de mayor frecuencia es la cuarta, por lo tanto, es la clase modal.

$L_i =$ $d_1 =$ $d_2 =$ $ai =$

Luego el valor calculado de la moda es:

$$Mo = L_i + \left[\frac{d_1}{d_1 + d_2} \right] ai = \quad Mo =$$