

Distribución de Frecuencia

La distribución de frecuencia representa el conjunto de los posibles valores que una variable puede tener en una muestra. Es normalmente un listaje, ordenada por cantidades o por clases, que presenta el número de veces que una variable aparece en la muestra. Para trabajar con una muestra grande o cuando las variables pueden tener valores continuos es recomendable organizar los datos en distribuciones de frecuencia como presentadas abajo.

Suponga que la tabla abajo presente la distribución de frecuencia de las edades de un grupo de 100 personas:

Clase	f_i
0 — 10	20
10 — 20	30
20 — 30	40
30 — 40	10
	n=100

Vamos entender los datos fornecidos:

Clase – es el intervalo del grupo. En nuestro ejemplo quiere decir que 20 personas ten entre 0 y 10 años y que 40 personas ten entre 20 y 30 años. El primer grupo seria la primera clase (de cero hasta 10) la segunda línea seria la segunda clase (10 a 20) e así por delante.

0 – límite inferior (número de la derecha) de la primera clase.

10 – límite superior (número de la izquierda) de la primera clase.

“|—” significa que el límite inferior está incluido en la clase.

“—|” significa que el límite superior está incluido en la clase.

“—” significa que el límite (sea inferior, sea superior) no está incluido en la clase.

En nuestro ejemplo el limite inferior está incluido en la clase y el limite superior no (el 10 hace parte de la segunda clase y no de la primera, o sea, si alguien tuviera 10 años estará en la segunda clase y no en la primera).

f_i – frecuencia absoluta (número de individuos que pertenecen aquella clase).

n – número de observaciones (tamaño de la muestra)

PM – Punto medio de la clase. Es calculado de la siguiente forma:

$$PM_{Clase1} = \frac{(\text{Lim}_{(\text{superior})} + \text{Lim}_{(\text{inferior})})}{2} = \frac{10 + 0}{2} = 5$$

Clase	f _i	PM
0 — 10	20	5
10 — 20	30	15
20 — 30	40	25
30 — 40	10	35
	n=100	