

## TRATAMIENTO PARA DATOS AGRUPADOS.

Cuando la muestra consta de 30 o más datos, lo aconsejable es agrupar los datos en clases y a partir de estas determinar las características de la muestra y por consiguiente las de la población de donde fue tomada.

Antes de pasar a definir cuál es la manera de determinar las características de interés (media, mediana, moda, etc.) cuando se han agrupado en clases los datos de la muestra, es necesario que sepamos como se agrupan los datos.

Pasos para agrupar datos.

**a. Determinar el rango o recorrido de los datos.**

Rango = Valor mayor – Valor menor

**b. Establecer el número de clases (k) en que se van a agrupar los datos tomando como base para esto la siguiente tabla.**

Tamaño de muestra o No. De datos	Número de clases
Menos de 50	5 a 7
50 a 99	6 a 10
100 a 250	7 a 12
250 en adelante	10 a 20

El uso de esta tabla es uno de los criterios que se puede tomar en cuenta para establecer el número de clases en las que se van a agrupar los datos, existen otros para hacerlo.

Método Sturges:  $k = 1 + 3,332 \log n$

**Determinar la amplitud de clase para agrupar (C).**

$$C = \frac{\text{Rango}}{k}$$

**d. Formar clases y agrupar datos.**

Para formar la primera clase, se pone como límite inferior de la primera clase un valor un poco menor que el dato menor encontrado en la muestra y posteriormente se suma a este valor C, obteniendo de esta manera el límite superior de la primera clase, luego se procede a obtener los límites de la clase siguiente y así sucesivamente.

Ejemplo:

Los siguientes datos se refieren al diámetro en pulgadas de un engrane.

6.75	7.00	7.00	6.75	6.50	6.50	7.15	7.00
6.50	6.50	6.50	6.25	6.25	6.50	6.65	7.00
7.25	6.70	6.00	6.75	6.00	6.75	6.75	7.10
7.00	6.70	6.50	6.75	6.25	6.65	6.75	7.10
7.25	6.75	6.25	6.25	7.00	6.75	7.00	7.15

a) Agrupe datos en intervalos

b) Obtenga: media, mediana y moda..

Solución:

a) Agrupando datos;

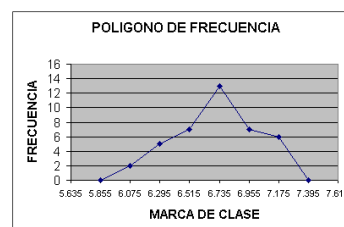
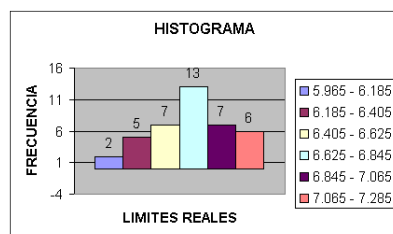
1.  $R = 7.25 - 6.00 = 1.25$

2.  $k = 1 + 3,332 \log 40 = 6,334 = 7$

3.  $A = \frac{R}{k} = \frac{1,25}{7} = 0,1786 = 0,18$

4. Formando clases.

LI	LS	Marca de clase (xi)	Frecuencia Absoluta (ni)	Frecuencia Absoluta Acumulada	Frecuencia relativa	Frecuencia Relativa acumulada
6,00 – 6,18		6,09	2	2	0,050	0,050
6,18 – 6,36		6,27	5	7	0,125	0,175
6,36 - 6,54		6,45	7	14	0,175	0,350
6,54 – 6,72		6,63	4	18	0,100	0,450
6,72 – 6,90		6,81	9	27	0,225	0,675
6,90 – 7,08		6,99	7	34	0,175	0,850
7,08 - 7,26		7,17	6	40	0,150	1,000



b) Gráficas:

c) **Media ( $\bar{x}$ ).**  $\bar{x} = \frac{\sum_{i=1}^k x_i \cdot n_i}{n}$

Donde:

k = número de clases

$x_i$  = marca de clase i

$n_i$  = frecuencia de la clase i

n = número de datos en la muestra

d) **Mediana ( $X_{med}$ ).**

$$X_{med} = Li + \left[ \frac{n/2 - N_{me-1}}{n_{me}} \right] A = 6.625 + \left[ \frac{40/2 - 14}{13} \right] (0.22) = 6.7265$$

Donde:

Li = límite real inferior de la clase que contiene a la mediana

$N_{me-1}$  = frecuencia absoluta acumulada de la clase en donde se encuentra la mediana

$n_{me}$  = frecuencia absoluta de la clase en donde se encuentra la mediana

A = amplitud real de la clase en donde se encuentra la mediana

n = número de datos en la muestra

e) **Moda ( $X_{mod}$ ).**

$$X_{mod} = Li + \left[ \frac{d_1}{d_1 + d_2} \right] A = 6.625 + \left[ \frac{6}{6 + 6} \right] (0.22) = 6.735 \text{ pulgadas}$$

Donde:

Li = límite real inferior de la clase que contiene a la moda

$$d_1 = |n_{mo} - n_{mo-1}| = |13 - 7| = 6$$

$$d_2 = |n_{mo} - n_{mo+1}| = |13 - 7| = 6$$

$n_{mo}$  = frecuencia de la clase que contiene a la moda

$n_{mo-1}$  = frecuencia de la clase anterior a la que contiene a la moda

$n_{mo+1}$  = frecuencia de la clase posterior a la que contiene a la moda

A = amplitud real de la clase que contiene a la moda

Ejercicios:

- Los datos que se dan a continuación corresponden a los pesos en Kg. de ochenta personas: (a) Obténgase una distribución de datos en intervalos (b) Calcúlese el porcentaje de personas de peso menor que 65 Kg. (c) ¿Cuántas personas tienen peso mayor o igual que 70 Kg. pero menor que 85? D) halle las m.t.c.

60; 66; 77; 70; 66; 68; 57; 70; 66; 52; 75; 65; 69; 71; 58; 66; 67; 74; 61; 63;  
 69; 80; 59; 66; 70; 67; 78; 75; 64; 71; 81; 62; 64; 69; 68; 72; 83; 56; 65; 74;  
 67; 54; 65; 65; 69; 61; 67; 73; 57; 62; 67; 68; 63; 67; 71; 68; 76; 61; 62; 63;  
 76; 61; 67; 67; 64; 72; 64; 73; 79; 58; 67; 71; 68; 59; 69; 70; 66; 62; 63; 66

- En una compañía aérea se sabe que, por término medio, el 65% de los vuelos tiene retraso. La distribución de los vuelos retrasados es la siguiente:

Duración de retrasos ( minutos )	# de vuelos
0 - 10	2000
10 - 20	3000
20 - 30	2500
30 - 40	2000
40 - 50	500

Hallar las m.t.c.