

## TABLA DE FRECUENCIA

Todo deseo de reunir, organizar y describir datos es motivado por el deseo de conocer algo (estado de naturaleza) de una población. Los datos normalmente son un conjunto de valores observados de de  $n$  observaciones sobre  $p$  variables, a partir de los cuales se desea analizar para posterior realizar inferencias de estos resultados a toda la población de referencia.

Una tabla de frecuencia es un resumen tal de la información, que conserva las características esenciales, tales como:

- Tendencia Central
- Posición
- Variabilidad
- Forma

Por lo tanto a partir de una tabla de frecuencia podemos determinar medidas descriptivas y construir graficos.

OBSERVACION : Una tabla de frecuencia solo tiene sentido construirla a una variable de naturaleza cuantitativa.

### Forma General de Una Tabla de Frecuencia

Linf - Lsup	$m_i$	$f_i$	$F_i$	$h_i=fr$	$H_i=Fr$
$Linf_1 - Lsup_1$	$m_1$	$f_1$	$F_1$	$h_1$	$H_1$
$Linf_2 - Lsup_2$	$m_2$	$f_2$	$F_2$	$h_2$	$H_2$
$Linf_i - Lsup_i$	$m_i$	$f_i$	$F_i$	$h_i$	$H_i$
$Linf_k - Lsup_k$	$m_k$	$f_k$	$n$	$h_k$	$1$
$\Sigma$		$n$		$1$	

### NOMENCLATURA

**Linf<sub>i</sub>** : Límite inferior del i-ésimo intervalo o clase

**Lsup<sub>i</sub>**: Límite superior del i-ésimo intervalo o clase

**m<sub>i</sub>** : Marca de clase.  $m_i = \frac{Lsup_i + Linf_i}{2}$

**f<sub>i</sub>** : Frecuencia Absoluta, se determina contando el número de observaciones que pertenecen o estan contenidas en el i-ésimo intervalo o clase.

$F_i$  : Frecuencia Absoluta Acumulada, se determina :  $F_i = \sum_{j=1}^i f_j$  Observe que la suma va desde  $j = 1$  hasta  $i$ , donde  $i$  representa el intervalo en el cual se encuentra.

$h_i$  : Frecuencia relativa  $h_i = \frac{f_i}{n}$

$H_i$  : Frecuencia Relativa Acumulada, se determina:

$$H_i = \sum_{j=1}^i h_j = \sum_{j=1}^i \frac{f_j}{n} = \frac{1}{n} \sum_{j=1}^i f_j = \frac{1}{n} F_i$$

## CONSTRUCCION DE UNA TABLA DE FRECUENCIA

### PASO 1

$$\text{RANGO} = \text{MAXIMO} - \text{MINIMO}$$

### PASO 2

$$\text{NUMERO DE INTERVALOS} = \text{APROX E. SUP}(K - \text{STURGE})$$

$$K\text{-STURGE} = 1 + 3.32 \text{ LOG}_{10} N \quad (\text{aproximar al entero superior})$$

### PASO 3

DETERMINAR LA AMPLITUD DE CADA INTERVALO O CLASE.

$$\text{AMPLITUD DE INTERVALO} = \frac{\text{MAXIMO} - \text{MINIMO}}{\text{NUMERO DE INTERVALOS}} = \text{APROX AL LA UNIDAD DE PRECISIÓN SUPERIOR.}$$

El concepto de "Unidad de Precisión" esta relacionado con la precisión con que son medidos los datos, es decir, valores enteros, o con un decimal o con 2 decimales, etc. A modo de ejemplo, los datos del laboratorio N: 1 estan medidos con un decimal, en consecuencia la unidad de precisión será la décima.

### PASO 4

DETERMINAR EL LIMITE INFERIOR DEL PRIMER INTERVALO

$$\text{Lin}_1 = \text{minimo} - \frac{1}{2} \text{ Unidad de Precisión}$$

### PASO 5

$$\text{Lsup}_1 = \text{Lin}_1 + \text{amplitud}(A)$$

$$\text{Lin}_i = \text{Lsup}_{(i-1)}$$

$$\text{Lsup}_i = \text{Lin}_i + \text{amplitud}(A)$$

**EJEMPLO** : Construir la tabla de frecuencia para los 99 valores de mejora del rendimiento dados en el laboratorio-1

**PASO 1**

$$\begin{aligned}\text{RANGO} &= \text{MAXIMO} - \text{MINIMO} \\ \text{RANGO} &= 22,4 - (-11,4) \\ \text{RANGO} &= 33,8\end{aligned}$$

**PASO 2**

$$K\text{-STURGE} = 1 + 3.32 \text{ LOG}_{10}99 = 7,626 \text{ Aproximado a } 8$$

**PASO 3**

$$A = \frac{\text{MAXIMO} - \text{MINIMO}}{\text{NUMERO DE INTERVALOS}} = \frac{33,8}{8} = 4,225 \text{ aproximar a } 4,3$$

Recuerde que la unidad de precisión en este caso es la décima, por tanto la décima superior es 4,3

**PASO 4**

$$\boxed{\text{Linf}_1 = \text{minimo} - \frac{1}{2} \text{ Unidad de Precisión} = -11,4 - \frac{1}{2}(0,1) = -11,45}$$

**PASO 5**

$$\text{Lsup}_1 = -11,45 + 4,3 = -7,15$$

En consecuencia, los restantes límites inferiores y superiores para los restantes 7 intervalos, se determinan de forma analoga.

La primera marca de clase es:

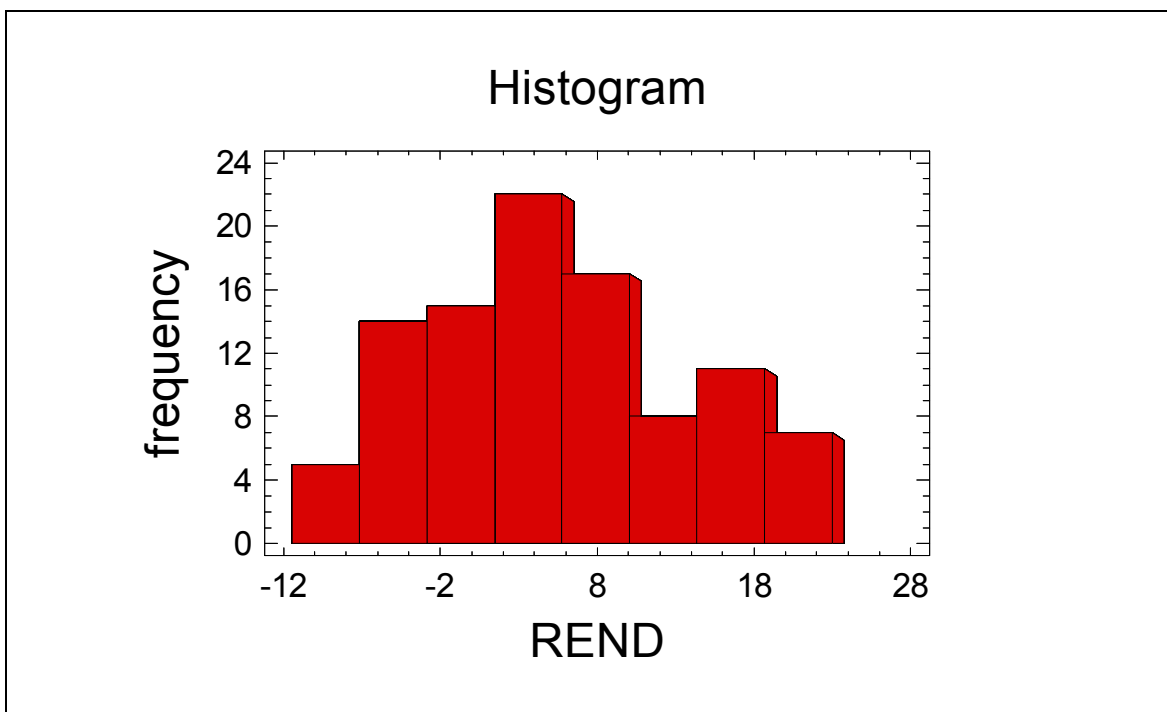
$$m_1 = \frac{\text{Lsup}_1 + \text{Linf}_1}{2} = \frac{-11,45 - 7,15}{2} = -9,3$$

Para calcular restantes 7 marcas de clase, no es necesario realizar este mismo procedimiento, sino simplemente sumar a la marca de clase anterior la amplitud A. Esto es posible porque todos los intervalos tienen la misma amplitud.

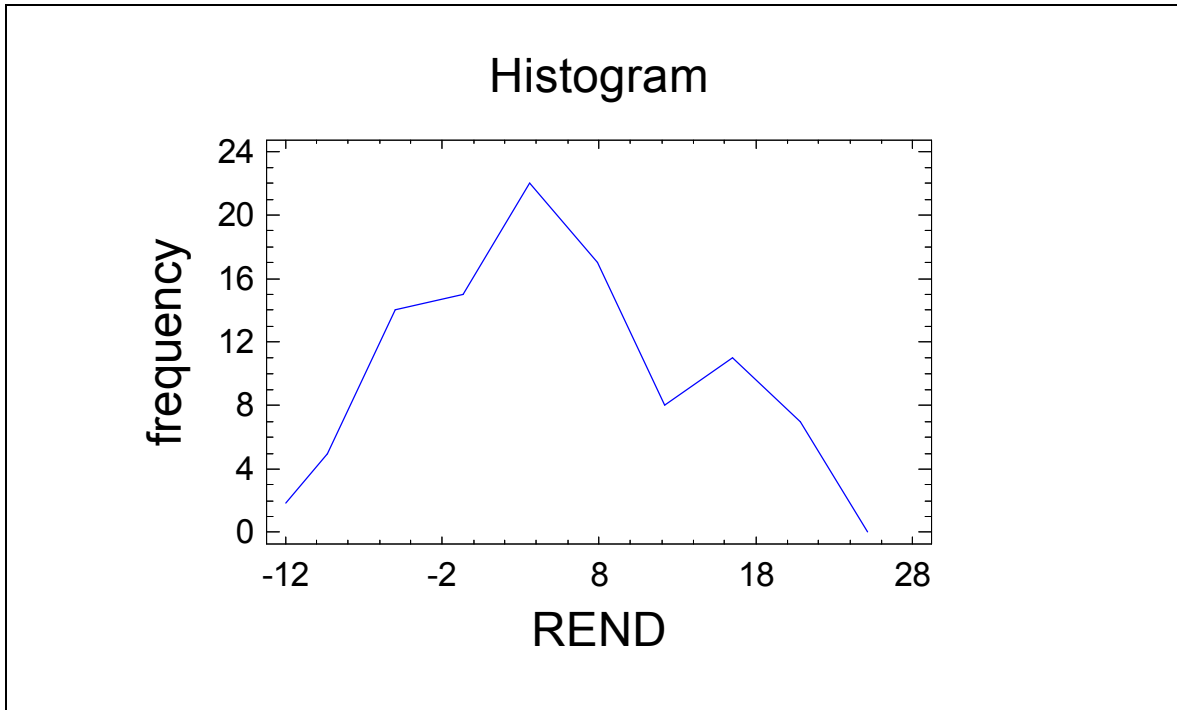
Tabla de frecuencia

Linf	Lsup	Mi	fi	Fi	hi	Hi
-11,45	-7,15	-9,3	5	5	0,051	0,05050505
-7,15	-2,85	-5	14	19	0,141	0,19191919
-2,85	1,45	-0,7	15	34	0,152	0,34343434
1,45	5,75	3,6	22	56	0,222	0,56565657
5,75	10,05	7,9	17	73	0,172	0,73737374
10,05	14,35	12,2	8	81	0,081	0,81818182
14,35	18,65	16,5	11	92	0,111	0,92929293
18,65	22,95	20,8	7	99	0,071	1
			99		1	

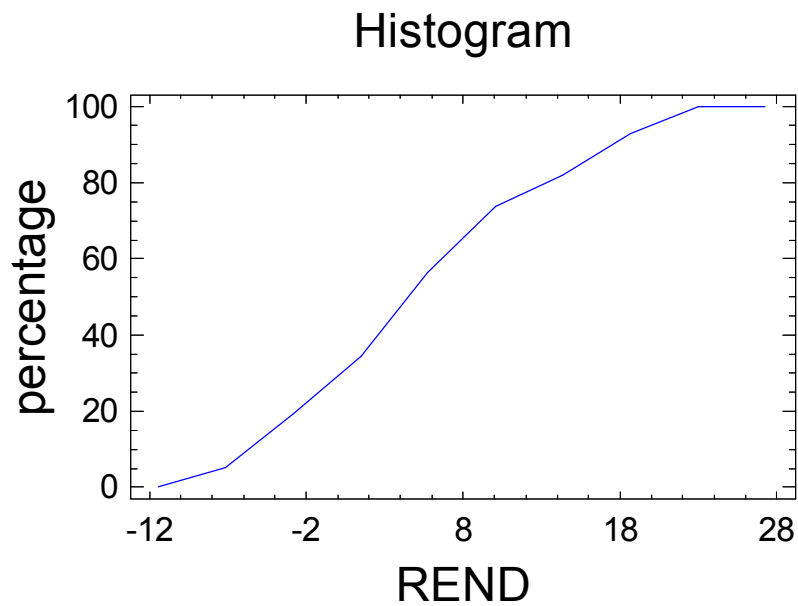
El Histograma asociado a la tabla de frecuencia es



Poligono de frecuencia

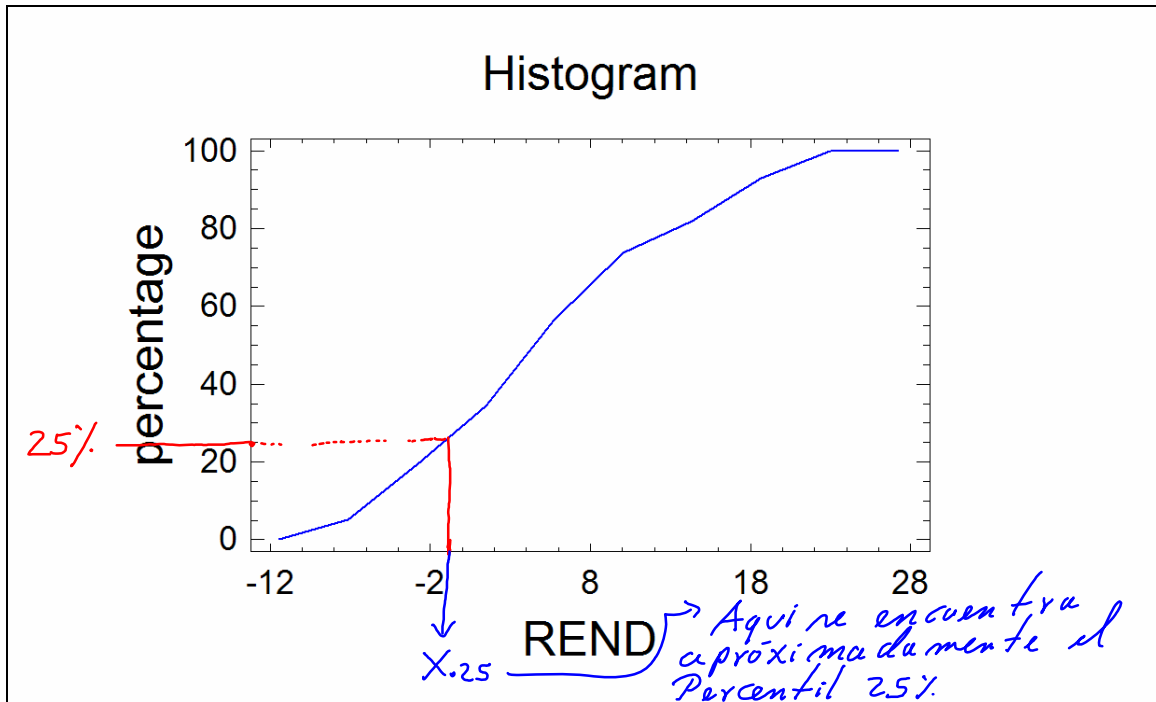


El poligono de frecuencia relativa acumulado se denomina "OJIVA", y tiene la importante propiedad de permitir determinar percentiles graficamente. Para esta tabla de frecuencia, la "OJIVA" es la siguiente:

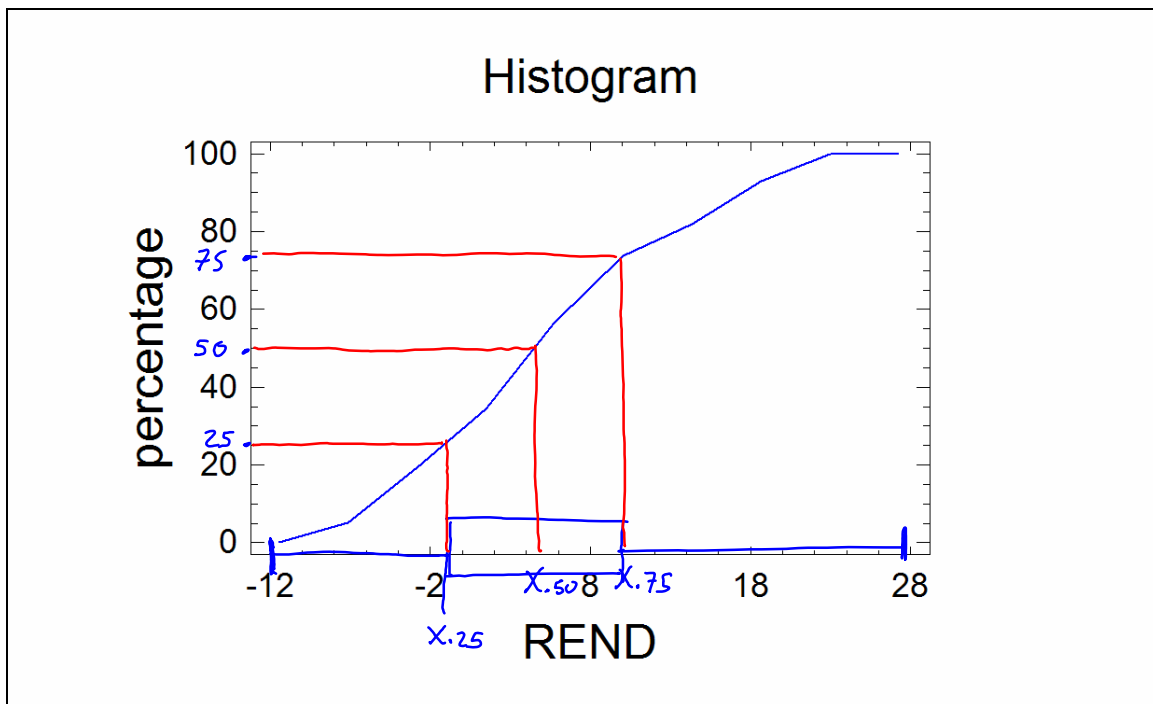


## MAS SOBRE TABLAS DE FRECUENCIA

Como se indico anteriormente, la “OJIVA” es un poligono de frecuencia relativa acumulada, por este motivo es posible estimar por interpolación sobre el PERCENTILES. A modo de ejemplo, estimaremos el percentil 25 sobre la “OJIVA” para la tabla de frecuencia asociada a los datos del Laboratorio-1.



*OJO: si se pueden estimar por interpolación percentiles, entonces es posible dibujar sobre la OJIVA el gráfico de caja, que es como sigue:*



Seguro que concuerdan que esto de la estadística es muy divertido.

## MEDIDAS DESCRIPTIVAS PARA DATOS TABULADOS

De la misma forma con que determinamos medidas descriptivas para datos no tabulados, también es posible determinar las mismas medidas para datos tabulados (TABLAS DE FRECUENCIA). Estas medidas son:

### Percentiles

$$X_{.p} = \text{Limf}_{(.p)} + \left[ \frac{\frac{n \cdot p + p}{100} - F_{(.p-1)}}{f_{(.p)}} \right] \cdot a$$

### Nomenclatura

$\text{Limf}_{(.p)}$ : Se denomina límite inferior del intervalo en que se encuentra el p-ésimo percentil. Para determinar dicho intervalo, debemos buscar en la Columna de Frecuencia absoluta acumulada ( $F_i$ ) de la tabla de frecuencia, el primer valor que contenga la cantidad

$$\frac{n \cdot p + p}{100}$$

$F_{(.p-1)}$ : Corresponde a la frecuencia absoluta Acumulada anterior al intervalo en que se encuentra ubicado el p-ésimo percentil.

$f_{(.p)}$ : Corresponde a la frecuencia absoluta del intervalo en que se encuentra ubicado el p-ésimo percentil.

$a$ : Amplitud de Intervalo.



A modo de Ejemplo, construiremos a mano el gráfico de caja, dado que esto implica calcular los percentiles  $X_{.25}$ ,  $X_{.50}$  y  $X_{.75}$

**Solución:**

i) Percentil 25%  $\Rightarrow$  Calcular  $1^{ro} \frac{n \cdot p + p}{100} = \frac{99 \cdot 25 + 25}{100} = 25$

$\rightarrow$  Esto implica que debemos buscar en la columna  $F_i$ , el primer valor que contenga **25**

Linf	Lsup	Mi	fi	Fi	hi	Hi
-11,45	-7,15	-9,3	5	5	0,051	0,05050505
-7,15	-2,85	-5	14	19	0,141	0,19191919
-2,85	1,45	-0,7	15	34	0,152	0,34343434
1,45	5,75	3,6	22	56	0,222	0,56565657
5,75	10,05	7,9	17	73	0,172	0,73737374
10,05	14,35	12,2	8	81	0,081	0,81818182
14,35	18,65	16,5	11	92	0,111	0,92929293
18,65	22,95	20,8	7	99	0,071	1
			99		1	

Intervalo que contiene el Percentil 25  $\rightarrow$  Intervalo Modal

$$X_{.25} = -2,85 + \left[ \frac{25 - 19}{15} \right] \cdot 4,3 = -1,13$$

ii) Percentil 50

$1^{ro}$  Calcular  $\frac{n \cdot p + p}{100} = \frac{99 \cdot 50 + 50}{100} = 50$

$$X_{.50} = 1,45 + \left[ \frac{50 - 34}{22} \right] \cdot 4,3 = 4,58$$

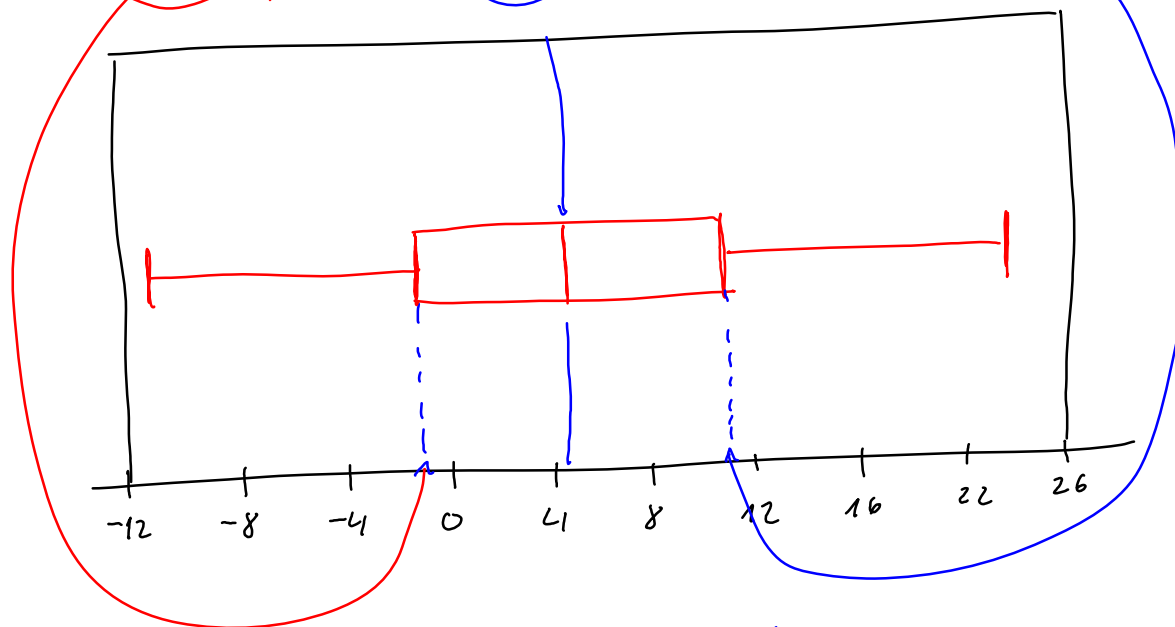
iii) Percentil 75%

1ºº Calculamos  $\frac{n \cdot p + p}{100} = \frac{99 \cdot 75 + 75}{100} = 75$

$$X_{.75} = 10,05 + \left[ \frac{75 - 73}{8} \right] \cdot 4,3 = 11,125$$

Ahora el Plato de Fondo  $\Rightarrow$  Gráfico de Caja.

$$X_{.25} = -1,13 ; X_{.50} = 4,58 ; X_{.75} = 11,1$$



Resultado ✓✓

Moda ( $X_m$ )

$$X_m = \text{Limf}(m) + \left[ \frac{|f(m) - f(m-1)|}{|f(m) - f(m-1)| + |f(m) - f(m+1)|} \right] \cdot a$$

Recuerde que  $|| \Rightarrow$  Valor absoluto

Nomenclatura

$\text{Limf}(m)$ : Límite inferior del intervalo modal, es decir, donde se encuentra la mayor frecuencia absoluta ( $f_i$ )

$f(m)$ : frecuencia absoluta del intervalo modal

$f(m-1)$ : frecuencia absoluta del intervalo anterior al modal.

$f(m+1)$ : frecuencia absoluta del intervalo posterior al intervalo modal.

Ejemplo: Determine la moda para la tabla de frecuencia

$$X_m = 1,45 + \left[ \frac{|22 - 15|}{|22 - 15| + |22 - 17|} \right] \cdot 4,3 = 3,95 \text{ OK}$$

OTO La Mediana =  $X_{.50}$  ✓✓

## Media ( $\bar{X}$ )

$$\bar{X} = \frac{1}{n} \sum_{i=1}^k m_i \cdot f_i ; \text{ donde } n = \sum_{i=1}^k f_i$$

Corresponde al número de intervalos

**Ejemplo:** Determine la media muestral para la tabla de frecuencia de Vel. Sube...

$$\bar{X} = \sum_{i=1}^8 m_i \cdot f_i = (m_1 \cdot f_1 + m_2 \cdot f_2 + m_3 \cdot f_3 + m_4 \cdot f_4 + m_5 \cdot f_5 +$$

$$m_6 \cdot f_6 + m_7 \cdot f_7 + m_8 \cdot f_8) \cdot \frac{1}{n}$$

$$= \frac{(-9,3) \cdot 5 + (-5) \cdot 14 + (-0,7) \cdot 15 + (3,6) \cdot 22 + (7,9) \cdot 17 + (12,2) \cdot 8 + (16,5) \cdot 11 + (20,8) \cdot 7}{99}$$

$$= \frac{511,2}{99} = 5,16 \quad \checkmark$$

## Varianza ( $S^2$ )

$$S^2 = \frac{n \cdot \sum_{i=1}^k (m_i)^2 \cdot f_i - \left( \sum_{i=1}^k m_i \cdot f_i \right)^2}{n \cdot (n-1)}$$

Observe

1) que la cantidad  $\left( \sum_{i=1}^k m_i \cdot f_i \right)$  la utiliza para calcular la media

Ejemplo: Determine la varianza para la tabla de frecuencia de los datos que Ud. sabe.....

OTO: Para este calculo, necesitamos sólo 3 datos:

$n$ , que lo tenemos

$\sum_{i=1}^8 m_i \cdot f_i$ , que también lo tenemos

$$\begin{aligned}\sum_{i=1}^8 (m_i)^2 \cdot f_i &= m_1^2 \cdot f_1 + m_2^2 \cdot f_2 + m_3^2 \cdot f_3 + \dots + m_7^2 \cdot f_7 + m_8^2 \cdot f_8 \\ &= (9,3)^2 \cdot 5 + (5)^2 \cdot 14 + (0,7)^2 \cdot 15 + (3,6)^2 \cdot 22 + (7,9) \cdot 17 + (12,2)^2 \cdot 8 + \\ &\quad (16,5)^2 \cdot 11 + (20,8)^2 \cdot 7 = 9349,84\end{aligned}$$

En consecuencia:

$$S^2 = \frac{99 \cdot (9349,84) - (511,2)^2}{99 \cdot 98} = 68,47 \checkmark \checkmark$$

Si lo hiciste bien, entonces puedes estar en los primeros posibles



o



## Preguntas Posibles

- 1: ¿Se puede calcular el coef de Variación? R: Si
- 2: ¿Se puede calcular el Coef de Asimetría? R: Si
- 3: ¿Se puede calcular el coef de Curtosis? R: Si

## Fundamental

Este profesor, por ningún motivo pediría en una prueba que construyan una tabla de frecuencia, sino que me intirase que la comprenden y conozcan lo que contiene, esto por si las moscas.