USB
Laboratorio 1 (Práctica nº 1)
Estadística para Ingenieros/Matemáticos
Prof. Javier Martínez y Prof. Liliana Córdova

Estadística Descriptiva Solución Práctica nº 1

Problema nº 1

Caso	Población	Objetivo Inferencial	Obtener Muestra
a)	Casas unifamiliares de la ciudad.	Estimación del consumo semanal promedio de agua en casas unifamiliares.	Dirigirse al ente pertinente para obtener los datos de consumo semanal de los hogares unifamiliares.
b)	Votantes del Estado.	Determinar si la mayoría de los votantes votaría por el candidato.	Aplicar encuestas bien distribuidas en todo el Estado.
c)	Lote de transistores.	Determinar si la duración promedio de cierto transistor supera las 500 horas.	Seleccionar una muestra al azar del lote de transistores lo suficientemente grande y hacerle pruebas de rendimiento.

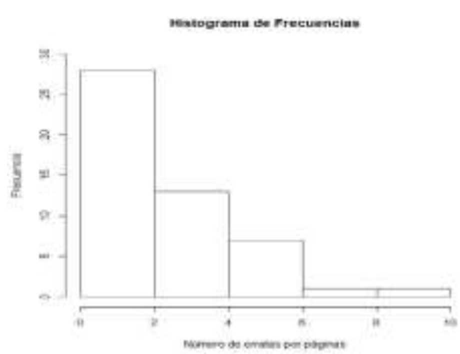
Problema nº 2

Los datos son de tipo variable cuantitativa discreta. Luego de ordenar los datos de manera creciente, se tiene la siguiente tabla de datos no agrupados:

Tabla de Datos No Agrupados

Variable y_i	Frecuencia **fi**	Frecuencia Relativa **hi** **fi / n**	Frecuencia Acumulada **Fi**	Frecuencia Relativa Acumulada **Hi**	(fi * y_i)/n	fi * $(y_i - \bar{y})^2/n-1$
0	6	0,12	6	0,12	0,000	0,879
1	10	0,2	16	0,32	0,200	0,576
2	12	0,24	28	0,56	0,480	0,113
3	8	0,16	36	0,72	0,480	0,017
4	5	0,1	41	0,82	0,400	0,178
5	4	0,08	45	0,9	0,400	0,439
6	3	0,06	48	0,96	0,360	0,675
8	1	0,02	49	0,98	0,160	0,578
10	1	0,02	50	1	0,200	1,094
Número de Datos **n** =	**50**				**2,680**	**4,549**

En el histograma se puede observar que la mayor frecuencia absoluta se encuentra en el intervalo [0,2], la moda y la mediana se encuentran en dicho intervalo. Así mismo, las frecuencias decrecen conforme al aumento del número de las erratas.



$$\boldsymbol{Media} = \sum \frac{(\mathrm{fi} * \mathrm{yi})}{\mathrm{n}} = \mathbf{2.68}$$

$$\boldsymbol{Varianza} = \sum \mathrm{fi} * \frac{(\mathrm{yi} - \bar{\mathrm{y}})^2}{\mathrm{n}} = \mathbf{4.55}$$

$$\boldsymbol{Desviaci\acute{o}n\ Estandar} = \sqrt{\mathrm{Varianza}} = \sqrt{4.4576} = \mathbf{2.11}$$

Formula de la Mediana	
Si n Par	**Si n Impar**
$\dfrac{\left(\boldsymbol{y}_{\left(\frac{n+1}{2}\right)} + \boldsymbol{y}_{\left(\frac{n+1}{2}+1\right)}\right)}{\mathbf{2}}$	$\boldsymbol{y}_{\left(\frac{n+1}{2}\right)}$

Moda = Termino de mayor frecuencia = **2**

$$\boldsymbol{Mediana} = \frac{\left(\mathrm{y}_{(25)} + \mathrm{y}_{(26)}\right)}{2} = \frac{(2+2)}{2} = \mathbf{2}$$

Como la media es igual a 2,68; se espera obtener en promedio de 2 a 3 erratas por hoja. Los datos están un poco dispersos por tener varianza 4.55. La mediana es 2 cuyo valor donde se obtiene el 50% de la proporción y la mayor cantidad de erratas por página fue 2 (Moda).

Fórmula para Cuartiles
$\boldsymbol{q_i} = \boldsymbol{y}_{\left[\!\left[\frac{i(n+1)}{4}\right]\!\right]} + \left(\frac{\boldsymbol{i(n+1)}}{\mathbf{4}} - \left[\!\left[\frac{\boldsymbol{i(n+1)}}{\mathbf{4}}\right]\!\right]\right)\left(\boldsymbol{y}_{\left[\!\left[\frac{i(n+1)}{4}+1\right]\!\right]} - \boldsymbol{y}_{\left[\!\left[\frac{i(n+1)}{4}\right]\!\right]}\right)$

$$\boldsymbol{q_1} = y_{\left[\!\left[\frac{50+1}{4}\right]\!\right]} + \left(\frac{(50+1)}{4} - \left[\!\left[\frac{(50+1)}{4}\right]\!\right]\right)\left(y_{\left[\!\left[\frac{(50+1)}{4}+1\right]\!\right]} - y_{\left[\!\left[\frac{(50+1)}{4}\right]\!\right]}\right)$$

$$= y_{12} + (12.75 - 12)(y_{13} - y_{12}) = 1 + 0.75(1-1) = \mathbf{1}$$

$$\boldsymbol{q_3} = y_{\left[\!\left[\frac{3\,(50+1)}{4}\right]\!\right]} + \left(\frac{3(50+1)}{4} - \left[\!\left[\frac{3(50+1)}{4}\right]\!\right]\right)\left(y_{\left[\!\left[\frac{3(50+1)}{4}+1\right]\!\right]} - y_{\left[\!\left[\frac{3(50+1)}{4}\right]\!\right]}\right)$$

$$= y_{38} + (38.25 - 38)(y_{39} - y_{38}) = 4 + 0.25(4-4) = \mathbf{4}$$

$$\boldsymbol{Rango\ Intercuartil} = q_3 - q_1 = 4 - 1 = \mathbf{3}$$

$$\boldsymbol{N\acute{u}mero\ de\ Clases} = \sqrt{\mathrm{n}} = \sqrt{50} \sim \mathbf{7}$$

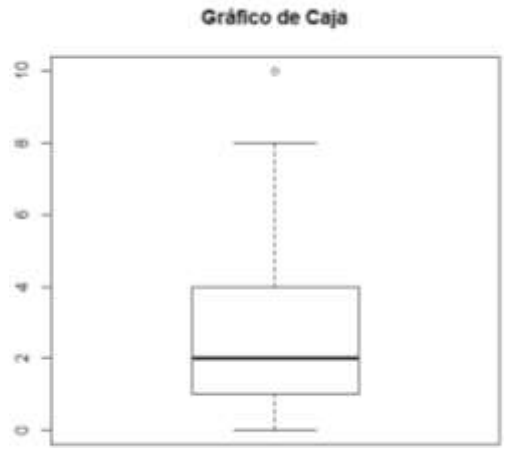
$$\boldsymbol{Rango} = \mathrm{y}_{\max} - \mathrm{y}_{\min} = 10 - 0 = \mathbf{10}$$

$$\boldsymbol{Amplitd} = \frac{Rango}{N\acute{u}mero\ de\ Clases} = \frac{10}{7} \sim \mathbf{2}$$

Tabla de Datos Agrupados					
Intervalo de Clase	**Marca de Clase (Punto Medio)**	Frecuencia **fi**	Frecuencia Relativa **hi** **fi / n**	Frecuencia Acumulada **Fi**	Frecuencia Relativa Acumulada Hi
[0,2)	1	16	0,32	16	0,32
[2,4)	3	20	0,4	36	0,72
[4,6)	5	9	0,18	45	0,9
[6,8)	7	3	0,06	48	0,96
[8,10]	9	2	0,04	50	1
	Número de Datos **n** =	50			

Limites Admisibles

$\boldsymbol{LI} = q_1 - \mathrm{f} * (q_3 - q_1) = 1 - (1.5) * 3 = \mathbf{-3.5}$ con **f = 1.5**
$\boldsymbol{LS} = q_3 + \mathrm{f} * (q_3 - q_1) = 4 + (1.5) * 3 = \mathbf{8.5}$



Problema 3

Tabla de Datos Agrupados							
Intervalo de Clase	**Marca de Clase m_i**	**Frecuencia fi**	**Frecuencia Relativa hi**	**Frecuencia Acumulada Fi**	**Frecuencia Relativa acumulada Hi**	**(fi * m_i)/n**	**fi * $(m_i - \bar{X})^2$/(n-1)**
[0,1)	0,5	6	0,3	6	0,3	0,15	0,32
[1,2)	1,5	9	0,45	15	0,75	0,68	0,00
[2,3)	2,5	4	0,2	19	0,95	0,50	0,21
[3,4)	3,5	1	0,05	20	1	0,18	0,21
	Número de Datos n =	**20**				**1,50**	**0,74**

a) Se Espera que los pacientes esperen en promedio 1.5 horas para ser atendidos. Con una variabilidad de 0.74 horas.

$$\boldsymbol{Media} = \sum \frac{(\mathrm{fi} * \mathrm{y_i})}{\mathrm{n}} = \mathbf{1.5}$$

$$\boldsymbol{Varianza} = \sum \mathrm{fi} * \frac{(\mathrm{m_i} - \bar{\mathrm{X}})^2}{\mathrm{n} - 1} = \mathbf{0.74}$$

$$\boldsymbol{Desviaci\acute{o}n\ Estandar} = \sqrt{\mathrm{Varianza}} = \sqrt{0.74} = \mathbf{0.86}$$

b) El paciente esperará ya que la mitad de los pacientes son atendidos en 1.44 horas.

$$\boldsymbol{Mediana} = \boldsymbol{LI} + \frac{\left(\frac{\boldsymbol{n}}{\mathbf{2}} - \boldsymbol{F_{i-1}}\right)}{\boldsymbol{fi}}(\boldsymbol{Amplitud}) = \mathbf{1} + \frac{\left(\frac{\mathbf{20}}{\mathbf{2}} - \mathbf{6}\right)}{\mathbf{9}}(\mathbf{1}) = \mathbf{1.44}$$

c) El tiempo de mayor frecuencia para el cual un paciente es atendido es 1.375.

$$\boldsymbol{Moda = LI + \frac{(f_i - f_{i-1})}{(f_i - f_{i-1}) + (f_i - f_{i+1})}(Amplitud)} = 1 + \frac{(9-6)}{(9-6)+(9-4)}(1) = 1 + \frac{3}{3+5} = \mathbf{1.375}$$

d) El tiempo de espera de una persona para ser pasada a otra consulta debe ser mayor a 2.25 horas.

$$\boldsymbol{p_{80} = LI + \frac{\left(\frac{80.n}{100} - F_{i-1}\right)}{fi}(Amplitud)} = 2 + \frac{\left(\frac{80(20)}{100} - 15\right)}{4}(1) = \mathbf{2.25}$$

Problema nº 4

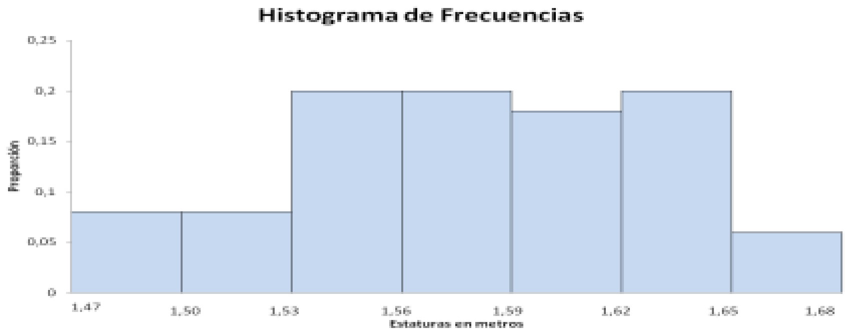
Los datos son variables cuantitativas continuas. Luego de ordenar de manera creciente, se procede a conformar la tabla de datos agrupados.

$$\boldsymbol{Número\ de\ Clases} = \sqrt{\mathrm{n}} = \sqrt{50} = \mathbf{7.07}$$

$$\boldsymbol{Rango} = \mathrm{y_{max}} - \mathrm{y_{min}} = 1.68 - 1.47 = \mathbf{0.21}$$

$$\boldsymbol{Amplitd} = \frac{Rango}{Número\ de\ Clases} = \frac{0.21}{7.07} = \mathbf{0.03}$$

Tabla de Datos Agrupados							
Intervalo de Clase	**Marca de Clase y_i**	Frecuencia **fi**	Frecuencia Relativa **hi fi / n**	Frecuencia Acumulada **Fi**	Frecuencia Relativa Acumulada Hi	(fi * y_i)/n	fi * $\frac{(yi-\bar{y})^2}{n}/n-1$
[1.47,1.50)	1,485	4	0,08	4	0,08	0,12	0,00073
[1.50,1.53)	1,515	4	0,08	8	0,16	0,12	0,00034
[1.53,1.56)	1,545	10	0,2	18	0,36	0,31	0,00025
[1.56,1.59)	1,575	10	0,2	28	0,56	0,32	0,00000
[1.59,1.62)	1,605	9	0,18	37	0,74	0,29	0,00012
[1.62,1.65)	1,635	10	0,2	47	0,94	0,33	0,00062
[1.65,1.68]	1,665	3	0,06	50	1	0,10	0,00044
	Número de Datos **n** =	**50**				**1,58**	**0,00251**



$$\boldsymbol{Mediana = LI + \frac{\left(\frac{n}{2} - F_{i-1}\right)}{fi}(Amplitud)} = 1.56 + \frac{\left(\frac{50}{2} - 18\right)}{10}(0.03) = 1.58$$

$$\boldsymbol{Media} = \sum \frac{(\mathrm{fi} * \mathrm{yi})}{\mathrm{n}} = \mathbf{1.58}$$

$$\boldsymbol{Varianza} = \sum \mathrm{fi} * \frac{(\mathrm{yi} - \bar{\mathrm{y}})^2}{\mathrm{n}} = \mathbf{0.00246}$$

$$\boldsymbol{Desviaci\acute{o}n\ Estandar} = \sqrt{\mathrm{Varianza}} = \sqrt{0.00246} = \mathbf{0.05}$$

Multimodal

$$\boldsymbol{Moda\ 1 = LI + \frac{(f_i - f_{i-1})}{(f_i - f_{i-1}) + (f_i - f_{i+1})}(Amplitud)} = 1.53 + \frac{(10-4)}{(10-4) + (10-10)}(0.03) = \mathbf{1.56}$$

$$\boldsymbol{Moda\ 2 = LI + \frac{(f_i - f_{i-1})}{(f_i - f_{i-1}) + (f_i - f_{i+1})}(Amplitud)} = 1.56 + \frac{(10-10)}{(10-10) + (10-9)}(0.03) = \mathbf{1.56}$$

$$\boldsymbol{Moda\ 3 = LI + \frac{(f_i - f_{i-1})}{(f_i - f_{i-1}) + (f_i - f_{i+1})}(Amplitud)} = 1.62 + \frac{(10-9)}{(10-9) + (10-3)}(0.03) = \mathbf{1.62}$$

Cuartiles

$$\boldsymbol{q_1 = LI + \frac{\left(\frac{i\,n}{4} - F_{i-1}\right)}{fi}(Amplitud)} = 1.53 + \frac{\left(\frac{50}{4} - 8\right)}{10}(0.03) = \mathbf{1.54}$$

$$\boldsymbol{q_3 = LI + \frac{\left(\frac{i\,n}{4} - F_{i-1}\right)}{fi}(Amplitud)} = 1.62 + \frac{\left(\frac{3(50)}{4} - 37\right)}{10}(0.03) = \mathbf{1.62}$$

$$\boldsymbol{Rango\ Intercuartil} = q_3 - q_1 = 1.62 - 1.54 = 0.08$$

Limites Admisibles

$\boldsymbol{LI} = \ q_1 - \ \mathrm{f}(q_3 - \ q_1) = \ 1.54 - (1.5)0.08 = \mathbf{1.42}$

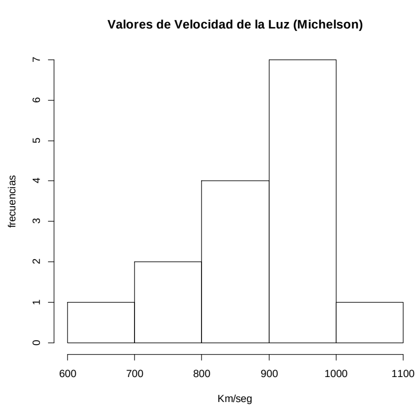
$\boldsymbol{LS} = \ q_3 + \ \mathrm{f}(q_3 - \ q_1) = 1.62 - (1.5)0.08 = \mathbf{1.74}$

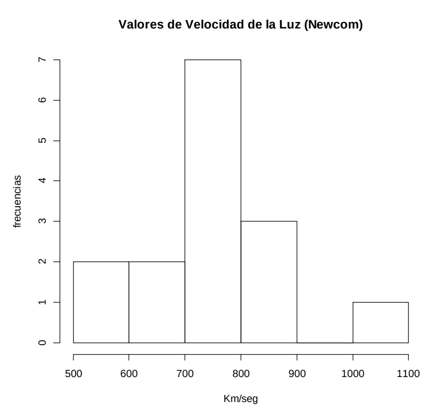
Problema 5

Tabla de Datos No Agrupados (Michelson)						
Variable yi	**Frecuencia fi**	**Frecuencia Relativa hi fi / n**	**Frecuencia Acumulada Fi**	**Frecuencia Relativa Acumulada Hi**	**(fi * yi)/n**	**fi *$(y_i - \bar{y})^2/n - 1$**
650	1	0,067	1	0,067	43,333	4346,032
740	1	0,067	2	0,133	49,333	1753,175
760	1	0,067	3	0,200	50,667	1334,127
850	2	0,133	5	0,333	113,333	311,111
880	1	0,067	6	0,400	58,667	19,841
900	1	0,067	7	0,467	60,000	0,794
930	2	0,133	9	0,600	124,000	158,730
950	1	0,067	10	0,667	63,333	203,175
980	3	0,200	13	0,867	196,000	1488,095
1000	1	0,067	14	0,933	66,667	762,698
1070	1	0,067	15	1,000	71,333	2146,032
Número de Datos n =	**15**				**896,667**	**12523,810**

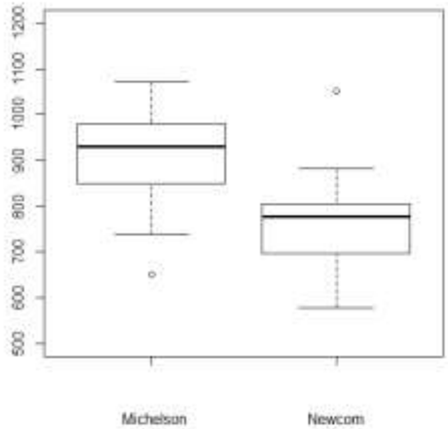
Tabla de Datos No Agrupados (Newcomb)						
Variable yi	**Frecuencia fi**	**Frecuencia Relativa hi**	**Frecuencia Acumulada Fi**	**Frecuencia Relativa Acumulada Hi**	**(fi * yi)/n**	**fi * $(y_i - \bar{y})^2/n-1$**
578	1	0,067	1	0,067	38,533	2449,931
599	1	0,067	2	0,133	39,933	1925,831
611	1	0,067	3	0,200	40,733	1654,631
682	1	0,067	4	0,267	45,467	470,960
711	1	0,067	5	0,333	47,400	194,631
772	1	0,067	6	0,400	51,467	5,531
774	1	0,067	7	0,467	51,600	8,331
778	1	0,067	8	0,533	51,867	15,646
781	1	0,067	9	0,600	52,067	22,631
796	1	0,067	9	0,667	53,067	76,846
796	1	0,067	11	0,733	53,067	76,846
816	1	0,067	12	0,800	54,400	199,131
820	1	0,067	13	0,867	54,667	230,446
883	1	0,067	14	0,933	58,867	1025,146
1051	1	0,067	15	1,000	70,067	5916,346
Número de Datos n =	**15**				**763,200**	**14272,886**

a)





b)	***Michelson***	**Newcomb**
Media	896,667	763,2
Desviación Estándar	111.909	119.469
Mediana	$y_{\left(\frac{15+1}{2}\right)} = y_{8=930}$	$y_{\left(\frac{15+1}{2}\right)} = y_{8=778}$



Problema 6

Tabla de Datos No Agrupados						
Julio-Agosto	**Frecuencia fi**	**Frecuencia Relativa hi fi / n**	**Frecuencia Acumulada Fi**	**Frecuencia Relativa Acumulada Hi**	**(fi * yi)/n**	**fi * $(y_i - \bar{y})^2/n-1$**
11	1	0,083	1	0,083	0,917	823,336
76	1	0,083	2	0,167	6,333	82,730
79	1	0,083	3	0,250	6,583	67,093
94	1	0,083	4	0,333	7,833	13,457
95	1	0,083	5	0,417	7,917	11,336
102	1	0,083	6	0,500	8,500	1,578
108	1	0,083	7	0,583	9,000	0,306
112	1	0,083	8	0,667	9,333	3,093
129	1	0,083	9	0,750	10,750	47,396
144	1	0,083	10	0,833	12,000	130,124
162	2	0,167	12	1,000	27,000	566,793
Número de Datos n =	**12**				**106,167**	**1747,242**

Diciembre	**Frecuencia fi**	**Frecuencia Relativa hi fi / n**	**Frecuencia Acumulada Fi**	**Frecuencia Relativa Acumulada Hi**	**(fi * yi)/n**	**fi *** $(\boldsymbol{y_i} - \overline{\boldsymbol{y}})^{\mathbf{2}} / \boldsymbol{n} - \mathbf{1}$
22	1	0,083	1	0,083	1,833	19,334
25	1	0,083	2	0,167	2,083	12,198
27	1	0,083	3	0,250	2,250	8,349
30	1	0,083	4	0,333	2,500	3,940
32	1	0,083	5	0,417	2,667	1,910
33	1	0,083	6	0,500	2,750	1,167
40	1	0,083	7	0,583	3,333	1,061
41	3	0,250	10	0,833	10,250	5,320
43	1	0,083	11	0,917	3,583	3,743
64	1	0,083	12	1,000	5,333	68,334
Número de Datos n =	**12**				**36,583**	**125,356**

Problema 8

$$\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} (\boldsymbol{x_i} - \overline{\boldsymbol{x}}) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \bar{x} = n \sum_{i=1}^{n} \frac{x_i}{n} - \sum_{i=1}^{n} \bar{x} = n\bar{x} - n\bar{x} = \mathbf{0}$$

Problema 9

$$n(\bar{x}^2 - 2\bar{x}^2) = -n\bar{x}^2 \quad con\ n \in \mathbb{N} - \{0\}$$

$$n\bar{x}^2 - 2\bar{x}\ n\bar{x} = -n\bar{x}^2$$

$$-2\bar{x}\, n \sum_{i=1}^{n} \frac{x_i}{n} + \sum_{i=1}^{n} \bar{x}^2 = -\frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2$$

$$\sum_{i=1}^{n} {x_i}^2 - \sum_{i=1}^{n} 2\bar{x}\, x_i + \sum_{i=1}^{n} \bar{x}^2 = \sum_{i=1}^{n} {x_i}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} {x_i}^2\right)^2$$

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} {x_i}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} {x_i}^2\right)^2$$

$$\boldsymbol{s^2} = \frac{\mathbf{1}}{\boldsymbol{n} - \mathbf{1}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} (\boldsymbol{x_i} - \overline{\boldsymbol{x}})^{\mathbf{2}} = \frac{\mathbf{1}}{\boldsymbol{n} - \mathbf{1}} \left(\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{x_i}^{\mathbf{2}} - \frac{\mathbf{1}}{\boldsymbol{n}} \left(\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{x_i}^{\mathbf{2}} \right)^{\mathbf{2}} \right)$$

Fórmula para Tabla de Datos Agrupados	
Media	$\sum \frac{(\mathrm{fi} * \ \mathrm{yi})}{\mathrm{n}}$
Varianza	$\sum \mathrm{fi} * \frac{(\mathrm{yi} - \ \tilde{\mathrm{y}})^2}{\mathrm{n}}$
Mediana	$\boldsymbol{LI} + \frac{\left(\frac{\boldsymbol{n}}{\boldsymbol{2}} - \boldsymbol{F_{i-1}}\right)}{\boldsymbol{fi}}(\boldsymbol{Amplitud})$
Moda	$\boldsymbol{LI} + \frac{(\boldsymbol{f_i} - \boldsymbol{f_{i-1}})}{(\boldsymbol{f_i} - \boldsymbol{f_{i-1}}) + (\boldsymbol{f_i} - \boldsymbol{f_{i+1}})}(\boldsymbol{Amplitud})$
$\boldsymbol{q_i}$	$\boldsymbol{LI} + \frac{\left(\frac{\boldsymbol{i\,n}}{\boldsymbol{4}} - \boldsymbol{F_{i-1}}\right)}{\boldsymbol{fi}}(\boldsymbol{Amplitud})$
$\boldsymbol{d_i}$	$\boldsymbol{LI} + \frac{\left(\frac{\boldsymbol{i\,n}}{\boldsymbol{10}} - \boldsymbol{F_{i-1}}\right)}{\boldsymbol{fi}}(\boldsymbol{Amplitud})$
$\boldsymbol{p_i}$	$\boldsymbol{LI} + \frac{\left(\frac{\boldsymbol{i\,n}}{\boldsymbol{100}} - \boldsymbol{F_{i-1}}\right)}{\boldsymbol{fi}}(\boldsymbol{Amplitud})$