# Problemas del Tema 2: Estadística Descriptiva Unidimensional

Profesora: Carmen Elvira Ramos Domínguez

# Índice

# Problema 1

Problema 1. Una compañía eléctrica ha realizado una investigación estadística en relación con las causas que han provocado la interrupción del suministro de energía eléctrica en la red que abastece y se han obtenido los siguientes datos.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| AT | AP | FM | FF | OC | NA | AE | FF | OC | AT |
| FM | AE | FF | AT | FF | OC | AT | OC | AE | OC |
| FF | AT | NA | OC | AE | OC | FM | AT | AE | FF |
| AT | FM | OC | AE | AT | FF | OC | AP | OC | FM |
| OC | OC | AT | FF | OC | AT | FF | FM | AT | OC |
| FF | AT | AE | OC | FM | AE | FF | OC | OC | AE |
| NA | OC | FF | OC | AE | AT | OC | AT | FF | FM |
| AE | OC | OC | FF | AT | OC | OC | NA | FM | AT |
| OC | AT | OC | OC | OC | AT | AE | FF | OC | FF |
| OC | FF | OC | AE | FF | FM | AT | AP | NA | AT |

# Problema 1

donde las abreviaturas utilizadas significan:

- FF: Fallo Elemento
- AT: Atmosféricas
- AE: Agentes Externos
- FM: Falsas Maniobras
- AP: Mala Actuación de Protecciones
- OC: Otras Causas
- NA: No Aclaradas

Se pide:

# Problema 1

a).- Construir la tabla de frecuencias de la variable estadística en estudio. ¿De qué tipo se trata?.

Sea X = "Causas de la interrupción del suministro eléctrico". Se trata de un carácter cualitativo o atributo y su tabla de frecuencias sería la siguiente.

| $X_i$ | $n_i$ | $f_i$ | $\%_i$ |
|---|---|---|---|
| FF | 18 | 0,18 | 18% |
| AT | 20 | 0,20 | 20% |
| AE | 13 | 0,13 | 13% |
| FM | 10 | 0,10 | 10% |
| AP | 3 | 0,03 | 3% |
| OC | 31 | 0,31 | 31% |
| NA | 5 | 0,05 | 5% |

# Problema 1

b).- ¿Más de la mitad de las interrupciones se debieron a causas atmosféricas?

No, en la tabla de frecuencias se puede apreciar que tan sólo el 20% menos del 50% se debieron a causas atmosféricas.

c).- ¿El porcentaje de interrupciones no aclaradas es superior al debido a causas atmosféricas?

No es cierto, el 5% fue debido a no aclaradas que es menor que el 20% debido a causas atmosféricas.

d).- Representar un diagrama de rectángulos y uno circular.

# Problema 1

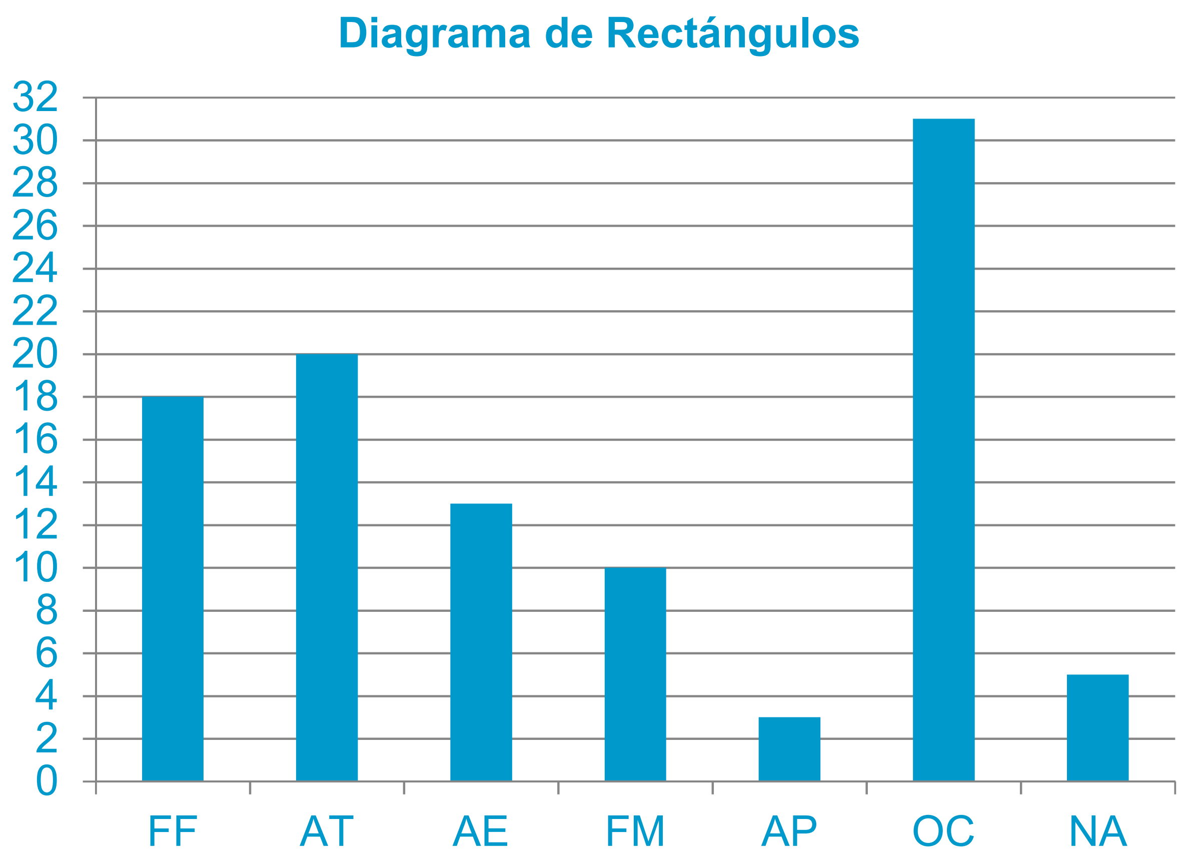
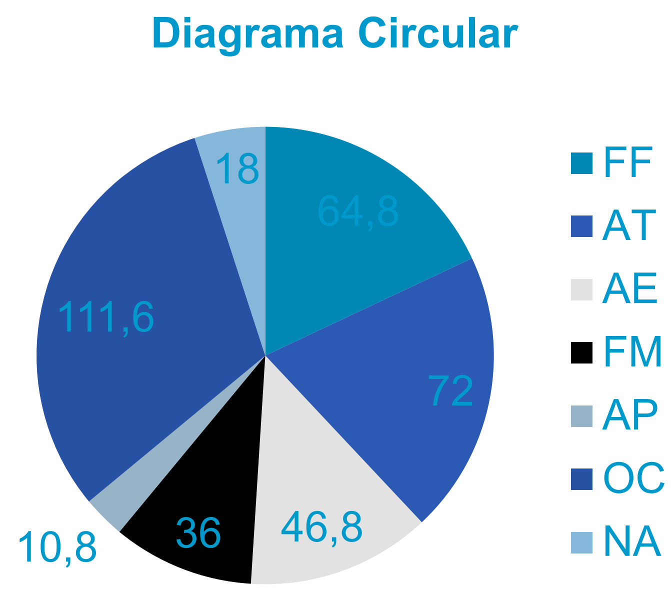


Diagrama de Rectángulos
32
30
28
26
24
22
20
18
16
14
12
10
8
6
4
2
0
FF
AT
AE
FM
AP
OC
NA

# Problema 1

| $X_i$ | $f_i$ | $360xf_i$ |
|---|---|---|
| FF | 0,18 | 64,8º |
| AT | 0,20 | 72º |
| AE | 0,13 | 46,8º |
| FM | 0,10 | 36º |
| AP | 0,03 | 10,8º |
| OC | 0,31 | 111,6º |
| NA | 0,05 | 18º |



# Problema 2

Problema 2: Las puntuaciones obtenidas por 100 opositores en el último ejercicio se presentan en el siguiente cuadro:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 7 | 3 | 2 | 4 | 5 | 1 | 8 | 6 | 1 | 5 |
| 3 | 2 | 4 | 9 | 8 | 1 | 0 | 2 | 4 | 1 |
| 2 | 5 | 6 | 5 | 4 | 7 | 1 | 3 | 0 | 5 |
| 8 | 6 | 3 | 4 | 0 | 10 | 2 | 5 | 7 | 4 |
| 0 | 2 | 1 | 5 | 6 | 4 | 3 | 5 | 2 | 3 |
| 9 | 7 | 3 | 4 | 3 | 5 | 7 | 4 | 6 | 5 |
| 6 | 1 | 0 | 5 | 7 | 8 | 5 | 2 | 3 | 10 |
| 4 | 6 | 2 | 1 | 1 | 2 | 6 | 7 | 4 | 5 |
| 4 | 7 | 6 | 3 | 5 | 0 | 2 | 8 | 2 | 7 |
| 8 | 5 | 2 | 7 | 1 | 4 | 6 | 3 | 5 | 6 |

Se pide:

# Problema 2

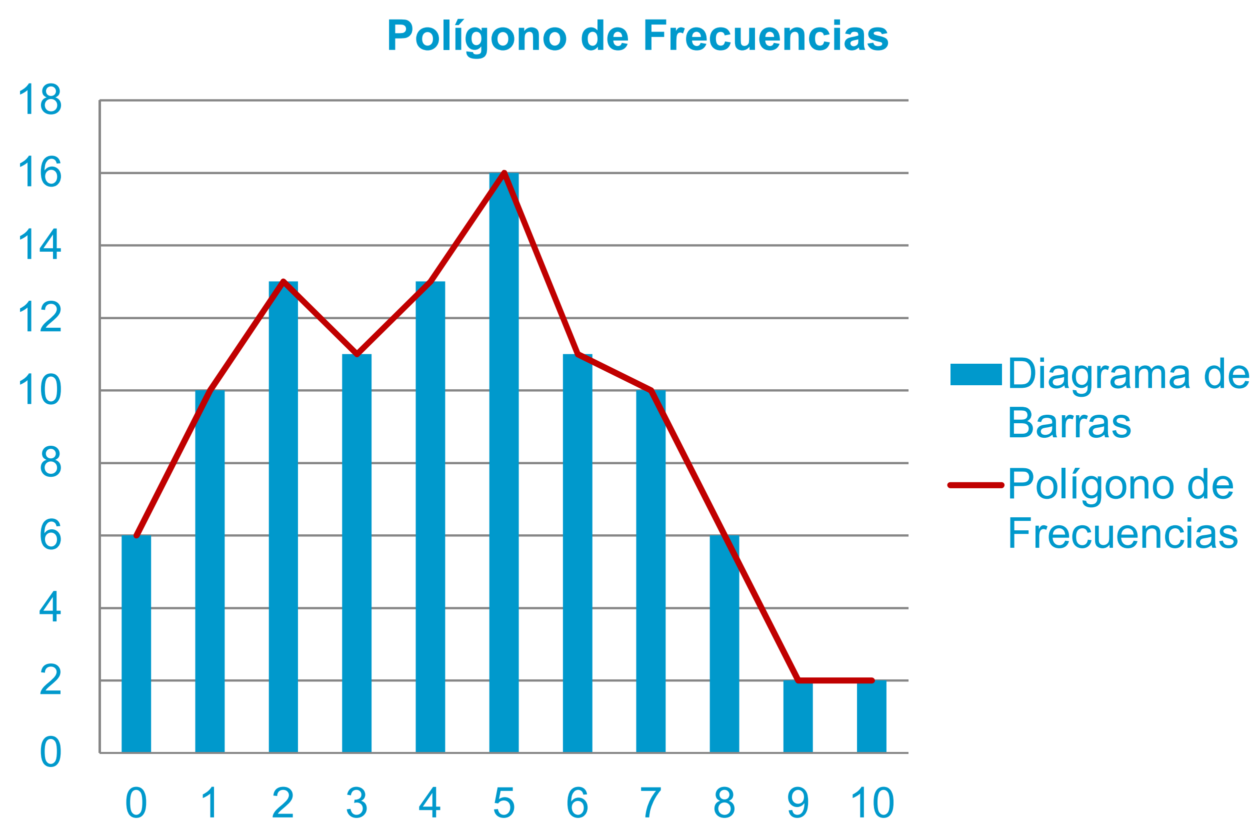
a).- Construir la tabla de fecuencias adecuada para las puntuaciones de los opositores.

Sea la variable X = “Puntuación obtenida por el opositor”. Construimos la tabla de frecuencias.

| $x_i$ | $n_i$ | $N_i$ | $f_i$ | $F_i$ | $\%_i$ |
|---|---|---|---|---|---|
| 0 | 6 | 6 | 0,06 | 0,06 | 6% |
| 1 | 10 | 16 | 0,10 | 0,16 | 10% |
| 2 | 13 | 29 | 0,13 | 0,29 | 13% |
| 3 | 11 | 40 | 0,11 | 0,40 | 11% |
| 4 | 13 | 53 | 0,13 | 0,53 | 13% |
| 5 | 16 | 69 | 0,16 | 0,69 | 16% |
| 6 | 11 | 80 | 0,11 | 0,80 | 11% |
| 7 | 10 | 90 | 0,10 | 0,90 | 10% |
| 8 | 6 | 96 | 0,06 | 0,96 | 6% |
| 9 | 2 | 98 | 0,02 | 0,98 | 2% |
| 10 | 2 | 100 | 0,02 | 1 | 2% |

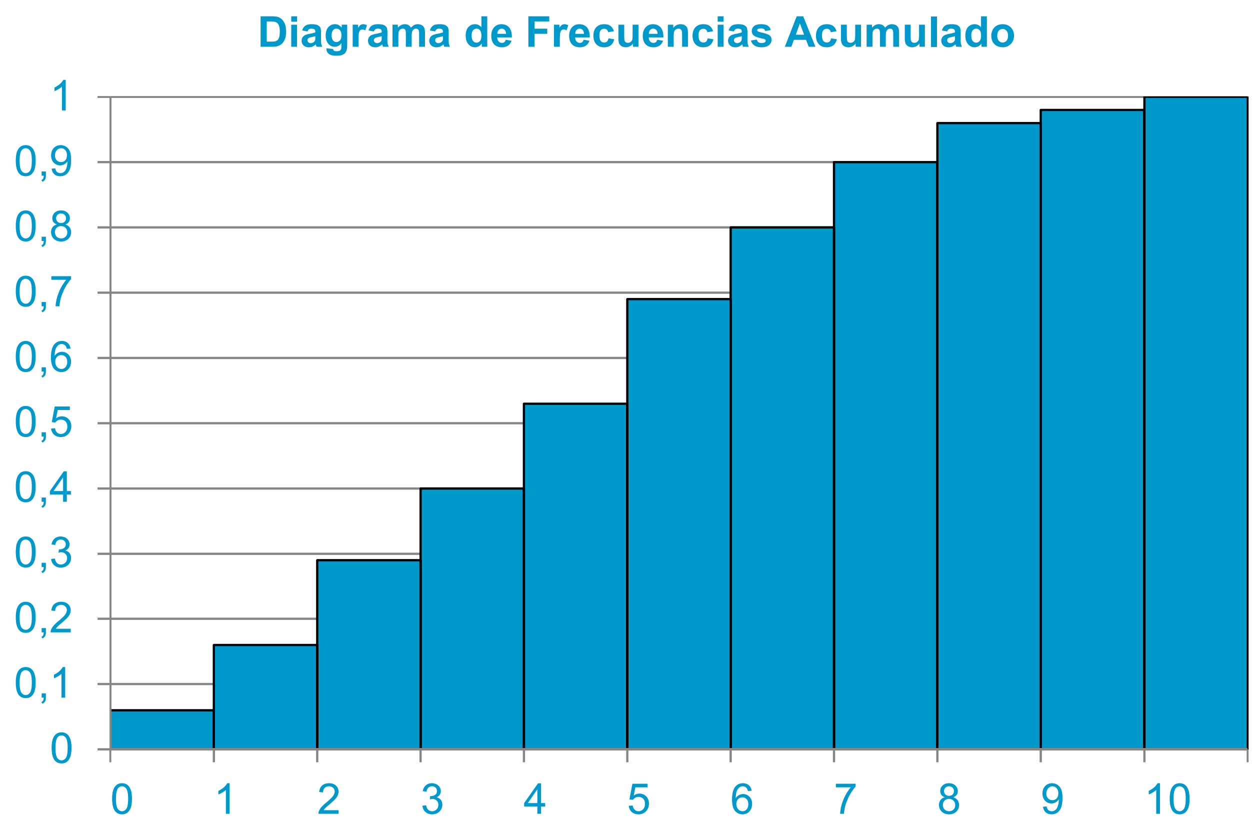
# Problema 2

b).- Dibujar el polígono de frecuencias y el diagrama de frecuencias relativa acumulado.



# Problema 2

Veamos ahora el diagrama de frecuencias acumulado o diagramas de barra acumulativo.



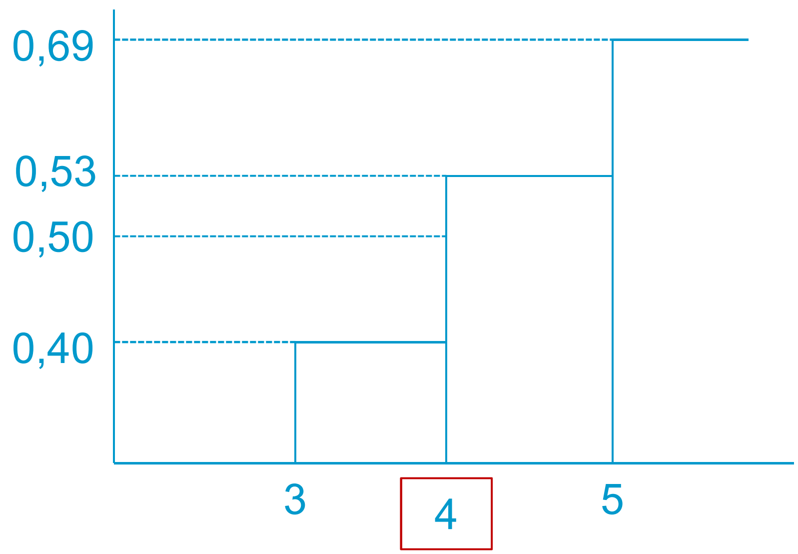
# Problema 2

c).- Calcular la media, la mediana y la moda de las puntuaciones obtenidas.

Veamos la puntuación media.

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i n_i}{n} = \frac{0x6+1x10+\cdots+9x2+10x2}{100} = \frac{423}{100} = 4{,}23$$

Veamos la mediana.




Me=4

Por último la moda sería el valor que más se repite, que en este caso es $M_d = 5$.

# Problema 2

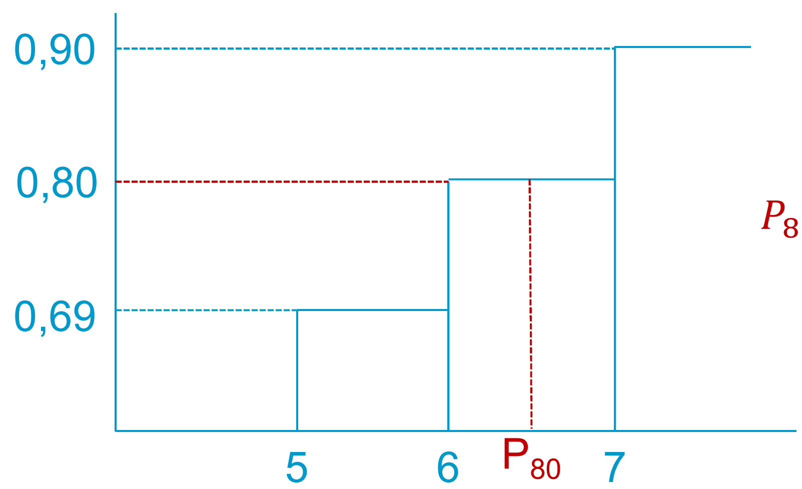
d).- ¿Cuál es el porcentaje de opositores que sacaron un 5 en la oposición? ¿y menos de un 5? Como se puede apreciar en la tabla el porcentaje de opositores que sacaron un 5 es el 16%. Y los que sacaron menos de un 5, esto es un 4 o menos, es el 53% ($F_i$ x100).

e).- Porcentaje de opositores que sacaron notas superiores a 4.

El porcentaje de opositores que sacaron más de un 4 serían, 100 menos los que sacaron menos o igual a un 4. Esto es, 100-53=47.

f).- Si sólo hay 20 plazas ¿En qué nota hay que situar el aprobado?

Buscamos la nota x que deja 20 opositores con una nota superior o igual, esto es, en este caso que hay 100, que deja el 20% por encima y el 80% por debajo. Ese valor es el percentil 80, $P_{80}$.




$$P_{80} = \frac{6+7}{2} = 6,5$$

# Problema 3

Problema 3. Considere las siguientes observaciones efectuadas en el flujo de información registrada en una red de ordenadores.

| | | | | | |
|---|---|---|---|---|---|
| 177,96 | 285,37 | 200,19 | 125,86 | 117,64 | 204,91 |
| 230,07 | 100,85 | 66,24 | 114,79 | 262,74 | 311,13 |
| 203,24 | 89,59 | 247,11 | 109,11 | 280,55 | 150,58 |
| 108,91 | 185,36 | 299,87 | 330,33 | 145,11 | 262,09 |
| 178,21 | 126,94 | 109,64 | 85,54 | 95,36 | 336,78 |
| 94,33 | 172,08 | 125,32 | 142,07 | 99,24 | 137,8 |

Se pide:

# Problema 3

a).- Construir la tabla de frecuencias agrupando en 5 intervalos cuyo límite inferior de la primera clase sea el mayor entero menor que el valor mínimo y el límite superior de la última clase sea el menor entero mayor que el valor máximo.

Sea la variable X = "Flujo de información registrada en una red de ordenadores". Veamos como construir la tabla de frecuencias.

$$\text{mín} = 66{,}24 \quad \lfloor 66{,}24 \rfloor = 66 \ \text{ y } \ \text{máx} = 336{,}78 \quad \lceil 336{,}78 \rceil = 337$$

$$\text{Rango} = 337 - 66 = 271 \quad \rightarrow \quad \text{Amplitud} = \frac{271}{5} = 54{,}2$$

| Clases | $m_i$ | $n_i$ | $N_i$ | $f_i$ | $F_i$ | $\%_i$ |
|---|---|---|---|---|---|---|
| (66, 120.2] | 93.1 | 12 | 12 | $0,\hat{3}$ | $0,\hat{3}$ | 33.3% |
| (120.2, 174.4] | 147.3 | 8 | 20 | $0,\hat{2}$ | $0,\hat{5}$ | 22.2% |
| (174.4, 228.6] | 201.5 | 6 | 26 | $0{,}1\hat{6}$ | $0{,}7\hat{2}$ | 16.6% |
| (228.6, 282.8] | 255.7 | 5 | 31 | $0{,}13\hat{8}$ | $0{,}86\hat{1}$ | 13.8% |
| (282.8, 337] | 309.9 | 5 | 36 | $0{,}13\hat{8}$ | 1 | 13.8% |

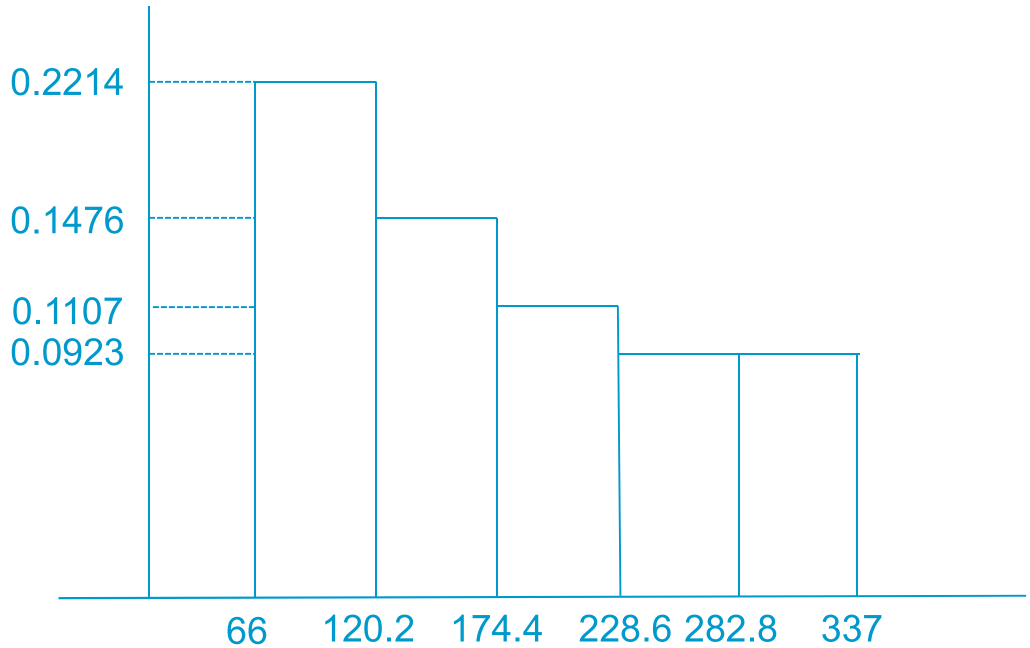
# Problema 3

b).- Representar el histograma de frecuencias absolutas y el polígono de frecuencias relativas acumulado.

Para representar el histograma de frecuencias absolutas, hemos de calcular primero las alturas, esto es, $h_i = \frac{n_i}{c_i}$.

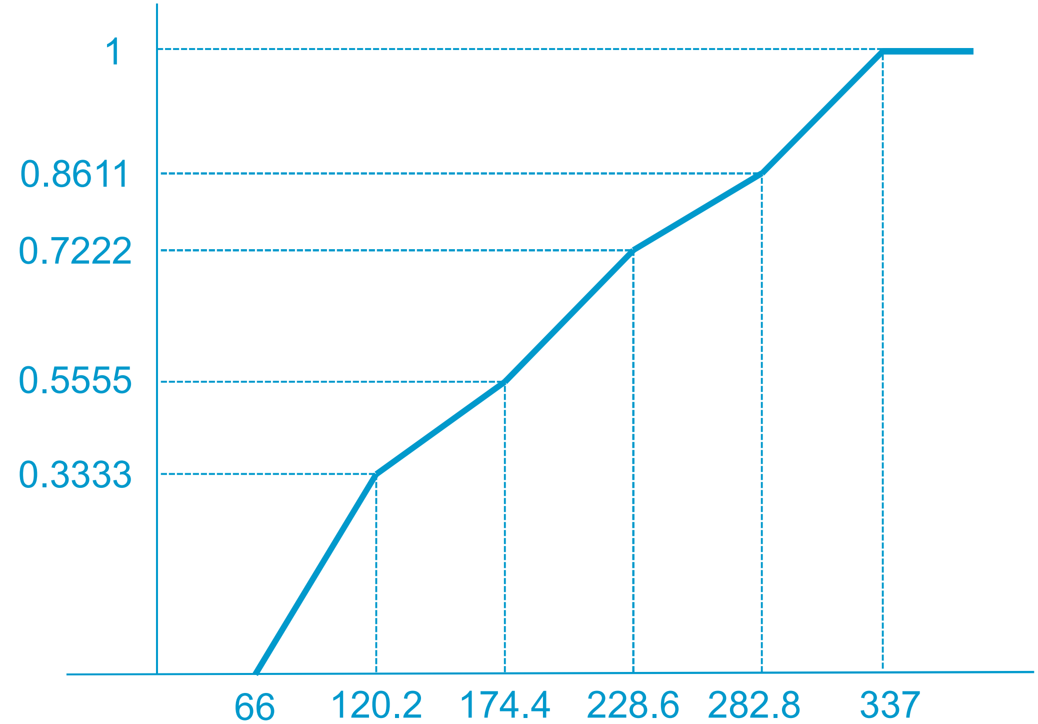
$$h_1 = \frac{12}{54{,}2} = 0.2214 \quad h_2 = \frac{8}{54{,}2} = 0.1476 \quad h_3 = \frac{6}{54{,}2} = 0.1107 \quad h_4 = h_5 = \frac{5}{54{,}2} = 0.0923$$

Entonces el histograma tendrá la siguiente forma:



# Problema 3

Veamos a continuación el polígono de frecuencias relativas acumulado.



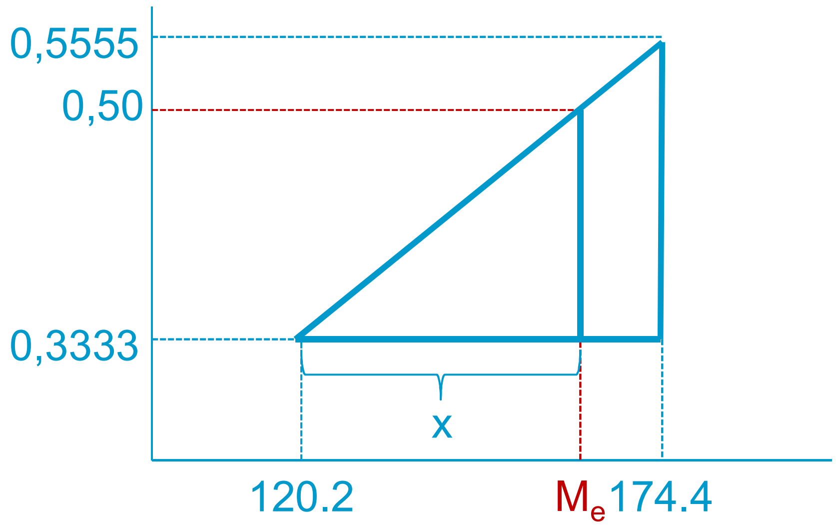

c).- Calcular la media, la mediana y la moda de la variable en estudio.

Veamos la media del flujo de información

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i n_i}{n} = \frac{93.1x12+147.3x8+\cdots+309.9x5}{36} = \frac{6332.6}{36} = 175.90\hat{5}$$

# Problema 3

Veamos ahora la mediana.




Aplicamos la relación de semejanza de triángulos

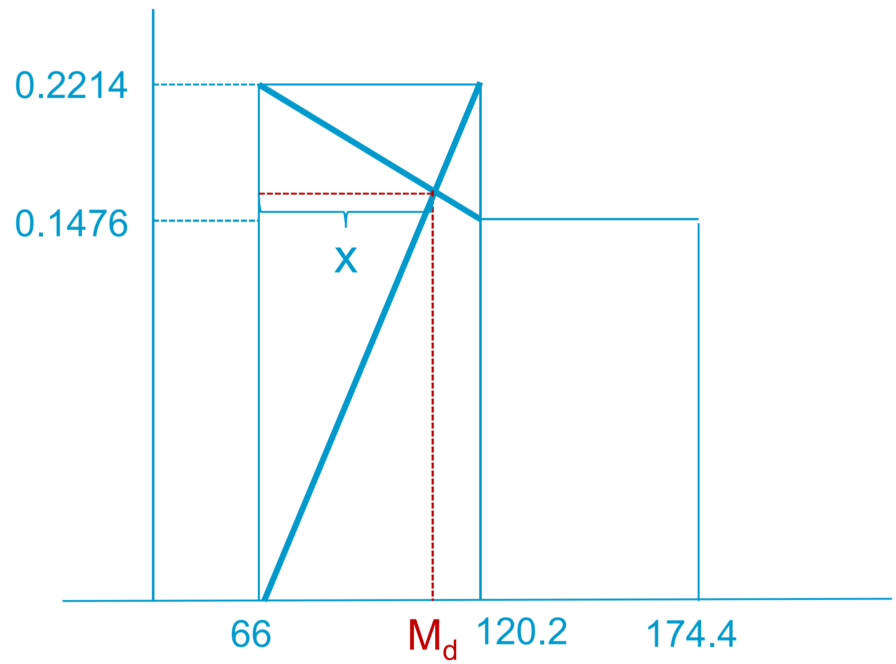
$$\frac{x}{0.5-0.\hat{3}}=\frac{174.4-120.2}{0.\hat{5}-0.\hat{3}}$$

$$\frac{x}{0.1\hat{6}}=\frac{54.2}{0.\hat{2}}$$

$$x=\frac{0.1\hat{6}\cdot 54.2}{0.\hat{2}}=40.64\hat{9}\ \rightarrow\ M_e=120.2+40.64\hat{9}=160{,}84\hat{9}$$

# Problema 3

Por último veamos la moda del flujo de información en la red. Para ello usamos la parte del histograma, donde se encuentra el intervalo modal, esto es, el rectángulo con mayor área por unidad de base.




Se aplica la relación de semejanza entre los triángulos de la pajarita.

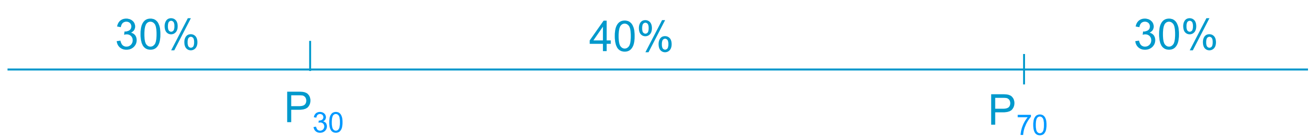
$$\frac{x}{0.2214} = \frac{54.2 - x}{0.2214 - 0.1476} \rightarrow \frac{x}{0.2214} = \frac{54.2 - x}{0.0738}$$

$$x = \frac{11.99988}{0.2952} = 40.65 \rightarrow \ M_d = 66 + 40.65 = 106.65$$

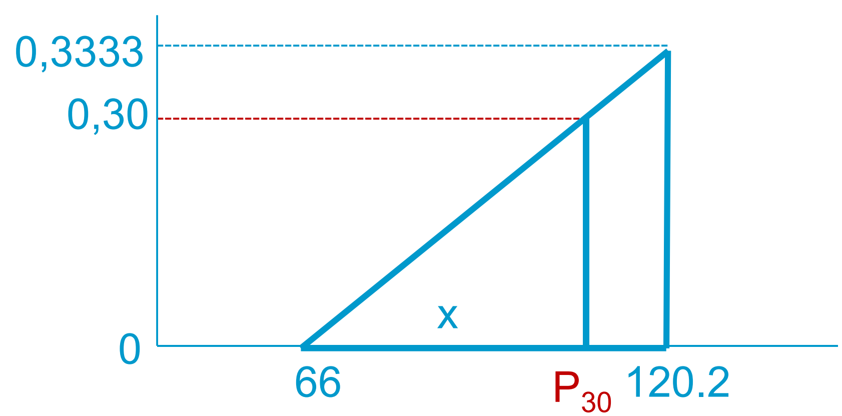
# Problema 3

d).- ¿Entre qué valores del flujo se encuentra el 40% de las observaciones centrales?

Para calcular los valores del flujo que dejan el 40% de las observaciones centrales procedemos como sigue. Si el 40% está interior, el 60% debe quedar por fuera y para que esté centrado el 40%, entonces el 60% debe ser 30% por ambos lados, como muestra el gráfico.




Como se puede apreciar en el gráfico buscamos el $P_{30}$ y el $P_{70}$.



# Problema 3

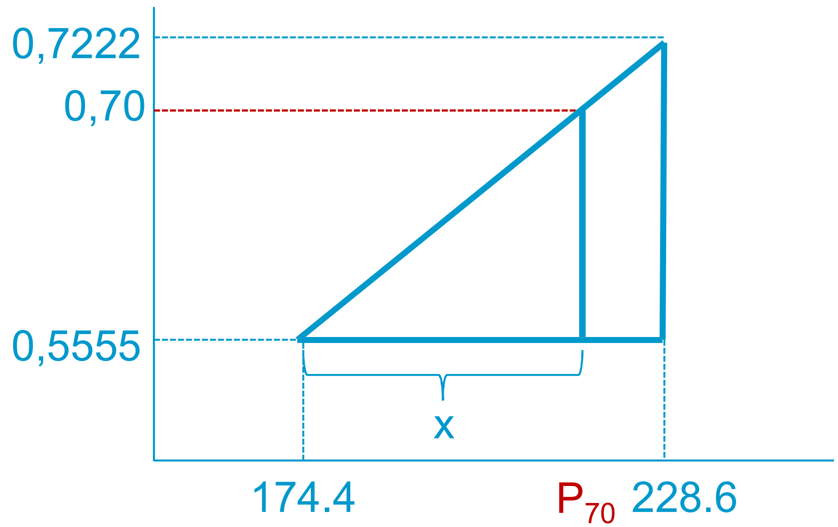
Aplicamos entonces la relación de semejanza de triángulos

$$\frac{x}{0.3-0}=\frac{120.2-66}{0.\hat{3}-0}$$

$$\frac{x}{0.3}=\frac{54.2}{0.\hat{3}}$$

$$x=\frac{0.3\cdot 54.2}{0.\hat{3}}=48.78\rightarrow \quad P_{30}=66+48.78=114.78$$

Veamos ahora el percentil 70.



# Problema 3

Aplicamos de nuevo la relación de semejanza de triángulos

$$\frac{x}{0.7-0.\hat{5}}=\frac{228.6-174.4}{0.7\hat{2}-0.\hat{5}}$$

$$\frac{x}{0.1\hat{4}}=\frac{54.2}{0.1\hat{6}}$$

$$x=\frac{0.1\hat{4}\cdot 54.2}{0.1\hat{6}}=46.97\hat{3}\rightarrow \quad P_{70}=174.4+46.97\hat{3}=221.37\hat{3}$$

e).- Calcular las medidas de dispersión: varianza y coeficiente de variación de Pearson. ¿Es la media representativa de los datos?

Veamos la varianza:

$$\sum_{\mathrm{i=1}}^{\mathrm{k}}\frac{(\mathrm{x_i})^2\mathrm{n_i}}{\mathrm{n}}-\bar{\mathrm{x}}^2=$$

$$=\frac{(93.1)^2\cdot 12+\ (147.3)^2\cdot 8+\ \ldots+(309.9)^2\cdot 5}{36}-\left(175.90\hat{5}\right)^2$$

$$=\frac{1328305.64}{36}-30942.7645=36897\ -\ 30942.7645=5954.6144$$

# Problema 3

Calculamos la desviación típica, tomando raíz cuadrada positiva.

$$S = \sqrt{5954.6144} = 77.166148$$

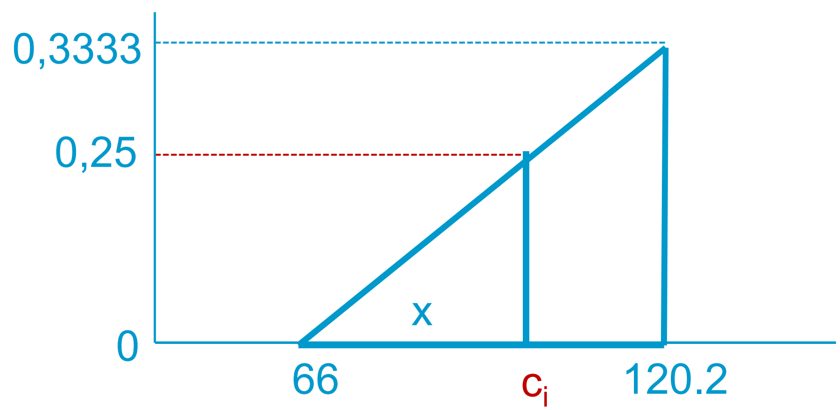
Entonces el coeficiente de variación de Pearson sería:

$$C.V = \frac{S}{|\bar{x}|} = \frac{77.166148}{175.90\hat{5}} = 0.438679$$

Como el coeficiente de variación de Pearson es menor que uno la media es representativa.

f).- Determinar el recorrido intercuartílico

Para determinar el recorrido intercuartílico tenemos que calcular el cuartil inferior y el cuartil superior.



# Problema 3

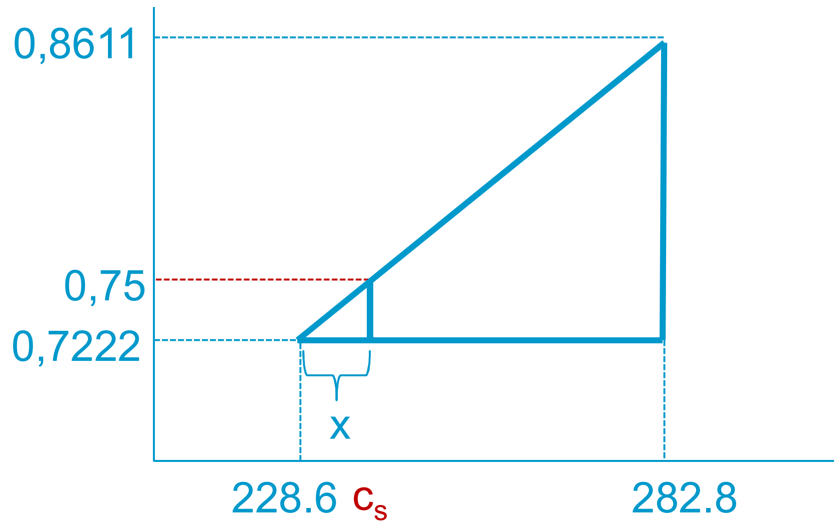
Aplicamos entonces la relación de semejanza de triángulos

$$\frac{x}{0.25-0} = \frac{120.2-66}{0.\hat{3}-0}$$

$$\frac{x}{0.25} = \frac{54.2}{0.\hat{3}}$$

$$x = \frac{0.25 \cdot 54.2}{0.\hat{3}} = 40.65 \rightarrow \; c_i = 66 + 40.65 = 106.65$$

Veamos ahora el cuartil superior



# Problema 3

Aplicamos de nuevo la relación de semejanza de triángulos

$$\frac{x}{0.75 - 0.7\hat{2}} = \frac{282.8 - 228.6}{0.861\hat{1} - 0.7\hat{2}}$$

$$\frac{x}{0.02\hat{7}} = \frac{54.2}{0.13\hat{8}}$$

$$x = \frac{0.02\hat{7} \cdot 54.2}{0.13\hat{8}} = 10.83\hat{9} \rightarrow \quad c_s = 228.6 + 10.83\hat{9} = 239.43\hat{9}$$

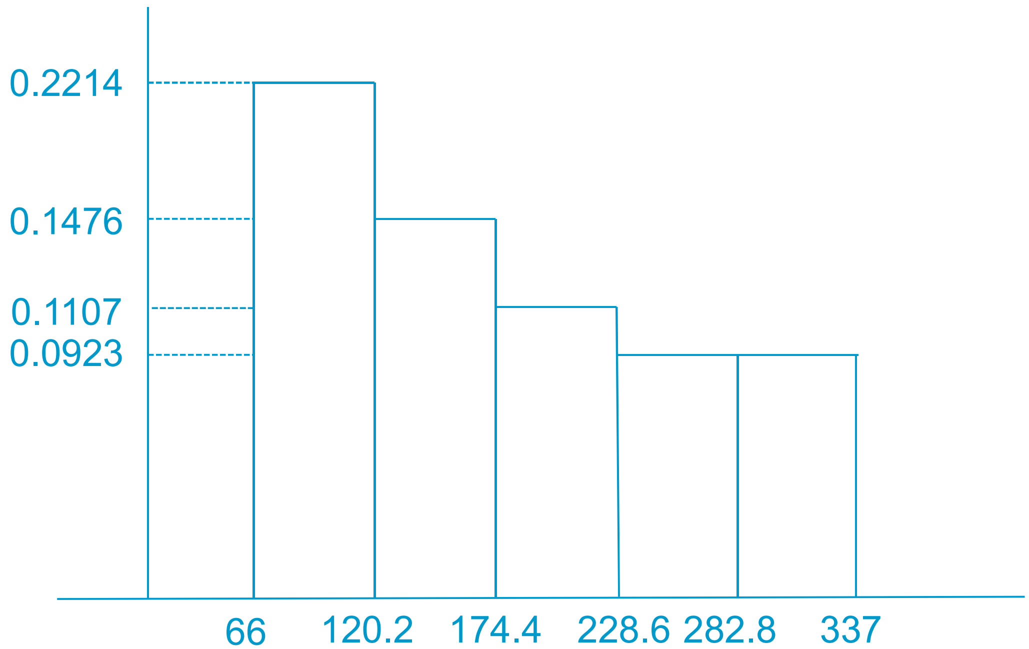
Entonces el recorrido intercuartílico sería:

$$c_s \; - \; c_i = 239.43\hat{9} \; - 106.65 = 132.789$$

g).- ¿Qué podemos decir de la simetría de la variable?

Si nos fijamos en el histograma, podemos decir que la variable presenta una asimetría a la derecha o positiva, porque dicha gráfica tiene una cola a la derecha.

# Problema 3




Para confirmarlo calculamos el coeficiente de simetría de Pearson.

$$A_p = \frac{\bar{x} - M_d}{S} = \frac{175.90\hat{5} - 106.65}{77.16615} = 0.8975 > 0$$

Al ser positivo entonces se confirma que el flujo de información presenta una asimetría a la derecha.

# Problema 4

Problema 4. La siguiente tabla presenta las velocidades de impresión de 40 ficheros de tipo PostScript, los tiempos vienen dados en

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 34.2 | 33.6 | 33.8 | 34.7 | 37.5 | 32.6 | 35.8 | 34.6 |
| 33.1 | 34.7 | 34.2 | 33.6 | 36.6 | 33.1 | 37.1 | 33.6 |
| 34.5 | 36.9 | 33.4 | 32.5 | 35.4 | 34.6 | 37.3 | 34.1 |
| 35.6 | 35.4 | 34.7 | 34.1 | 34.6 | 35.9 | 34.6 | 34.7 |
| 36.3 | 36.2 | 34.6 | 35.1 | 33.8 | 34.7 | 35.5 | 35.7 |

Se pide:

a).-Construir la tabla de frecuencias agrupando en 5 intervalos de igual amplitud cuyo límite inferior de la primera clase sea el valor mínimo y el límite superior de la última clase sea el valor máximo.

# Problema 4

Sea la variable X = “Velocidad de impresión de los ficheros PostScript”. Veamos la tabla de frecuencias de los datos de la muestra. $\text{mín} = 32.5$ y $\text{máx} = 37.5$

$$\text{Rango} = 37.5 - 32.5 = 5 \quad \rightarrow \quad \text{Amplitud} = \frac{5}{5} = 1$$

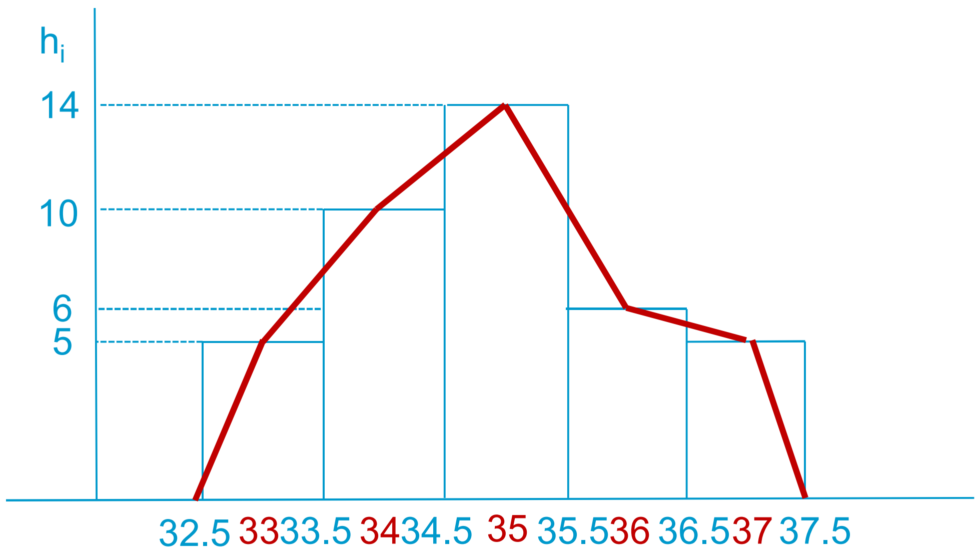
| Clases | $m_i$ | $n_i$ | $N_i$ | $f_i$ | $F_i$ | $\%_i$ |
|---|---|---|---|---|---|---|
| [32.5, 33.5] | 33 | 5 | 5 | 0.125 | 0.125 | 12.5% |
| (33.5, 34.5] | 34 | 10 | 15 | 0.25 | 0.375 | 25% |
| (34.5, 35.5] | 35 | 14 | 29 | 0.35 | 0.725 | 35% |
| (35.5, 36.5] | 36 | 6 | 35 | 0.15 | 0.875 | 15% |
| (36.5, 37.5] | 37 | 5 | 40 | 0.125 | 1 | 12.5% |
| Total | | 40 | | 1 | | 100% |

# Problema 4

b).- Representar el histograma y el polígono de frecuencias.

Para representar el histograma de frecuencias absolutas, hemos de calcular primero las alturas, esto es, $h_i = \frac{n_i}{c_i}$, siendo $n_i$ las frecuencias absolutas y $c_i$ las amplitudes de los intervalos. Como en este caso $c_i = 1$, entonces $h_i = n_i$.

Entonces el histograma y el polígono tendrán la siguiente forma:



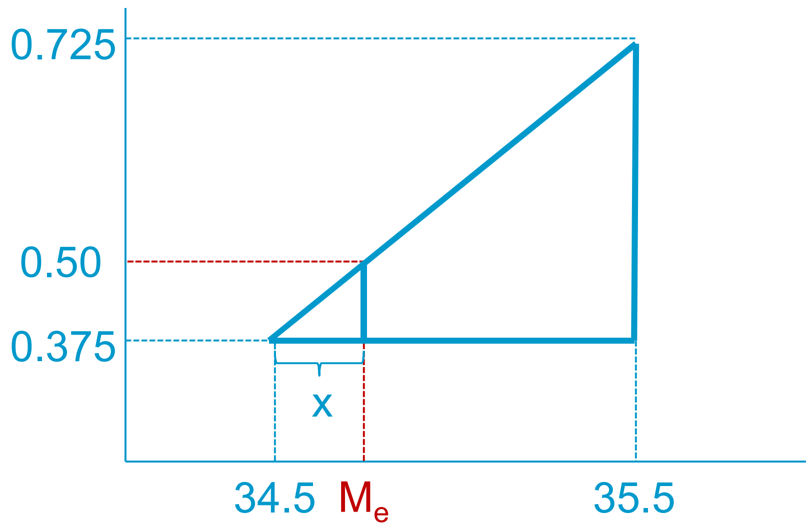
# Problema 4

c).- Calcular la media, la mediana y la moda de la velocidad de impresión.

Veamos la media de la velocidad de impresión.

$$\bar{x} = \frac{\sum_{i=1}^{k} m_i n_i}{n} = \frac{33x5+34x10+35x14+36x6+37x5}{40} = \frac{1396}{40} = 34.9$$

Veamos ahora la mediana




Aplicando entonces las relación de semejanza entre los triángulos rectángulos se tiene

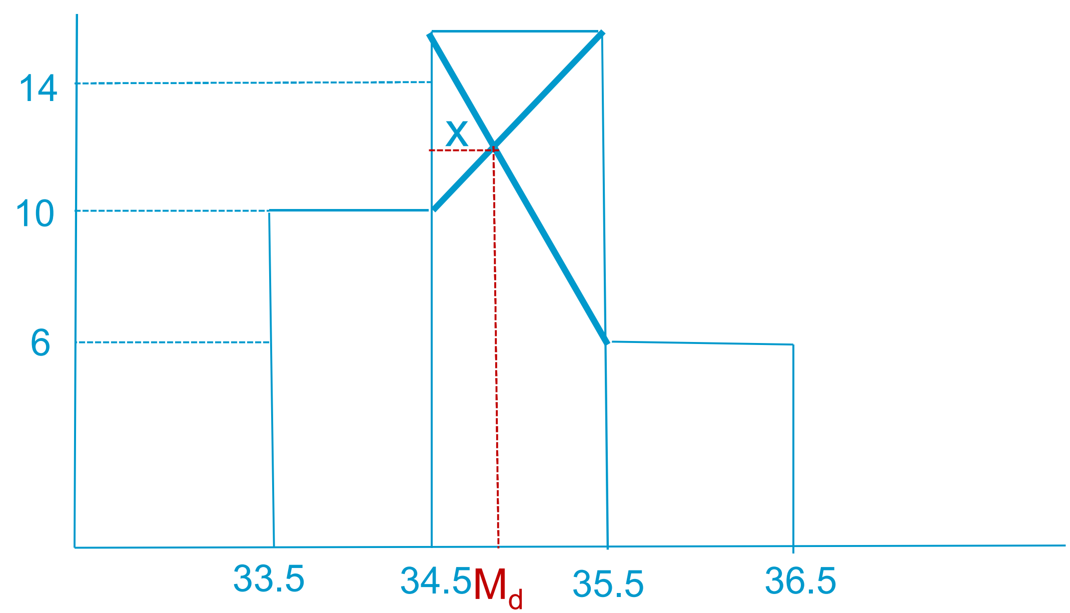
# Problema 4

$$\frac{x}{0.5 - 0.375} = \frac{35.5 - 34.5}{0.725 \ - 0.375}$$

$$\frac{x}{0.125} = \frac{1}{0.35}$$

$$x = \frac{0.125}{0.35} = 0.35714 \rightarrow \ \ M_e = 34.5 + 0.35714 = 34.85714$$

Seguimos con el cálculo de la moda. Para ello usamos la parte del histograma, donde se encuentra el intervalo modal.



# Problema 4

Aplicamos entonces la relación de semejanza entre los triángulos de la pajarita.

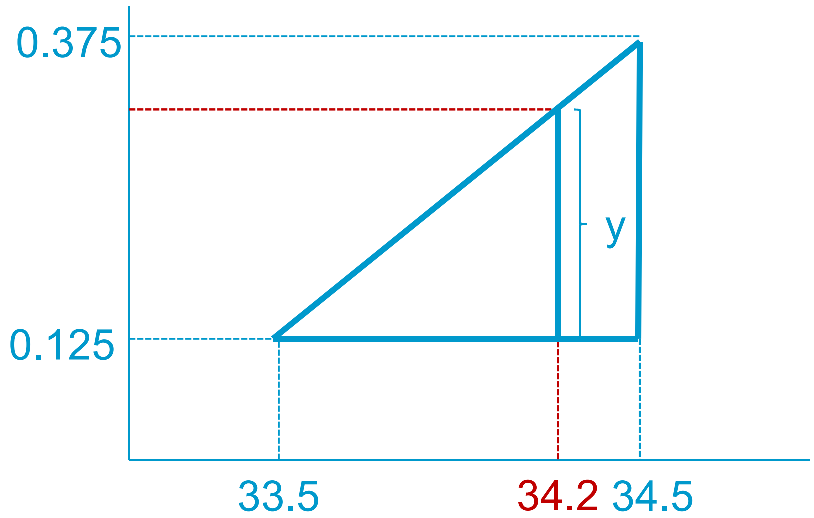
$$\frac{x}{14-10} = \frac{(35.5-34.5)-x}{14-6} \rightarrow \frac{x}{4} = \frac{1-x}{8} \rightarrow 4-4x = 8x \rightarrow 4 = 12x$$

$$x = \frac{4}{12} = 0.\hat{3} \rightarrow \quad M_d = 34.5 + 0.\hat{3} = 34.8\hat{3}$$

d).- ¿Qué porcentaje de ficheros postscript se han impreso con una velocidad inferior a 34.2 segundos? ¿Qué velocidad de impresión deja el 40% de las impresiones superiores a ella?

Para calcular el porcentaje de ficheros con velocidad inferior a 34.2, usamos la parte del polígono de frecuencias relativas acumulado en el que se encuentra 34.2.

# Problema 4




Aplicamos entonces la relación de semejanza entre los triángulos del gráfico.

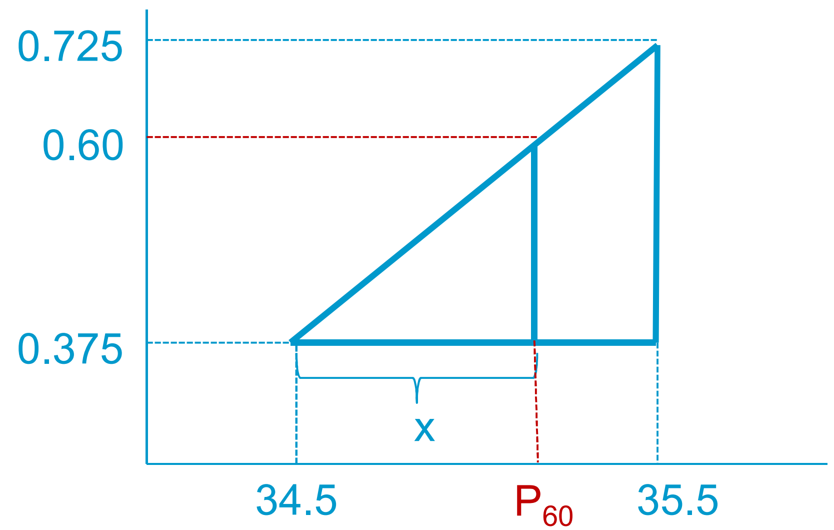
$$\frac{y}{34.2-33.5} = \frac{0.375-0.125}{34.5-33.5} \rightarrow \frac{y}{0.7} = \frac{0.25}{1} \rightarrow y = 0.25 \cdot 7 = 0.175$$

$$0.125 + 0.175 = 0.3$$

Por tanto, el 30% de los ficheros imprimen con una velocidad inferior a 34.2 segundos.

# Problema 4

El valor de la velocidad que deja un 40% de las impresiones superiores a éste, es el mismo que el que deja un 60% de las impresiones inferiores a él, esto es, el percentil 60, $P_{60}$.




De forma análoga, aplicamos entonces la relación de semejanza entre los triángulos del gráfico.

$$\frac{x}{0.6-0.375}=\frac{35.5-34.5}{0.725-0.375}\rightarrow\frac{x}{0.225}=\frac{1}{0.35}\rightarrow 0.35x=0.225$$

$$x=\frac{0.225}{0.35}=0.642857\ \rightarrow P_{60}=34.5\ +0.642857=35.142857$$

# Problema 5

Problema 5: Una empresa está realizando un estudio sobre la capacidad de ahorro de los individuos de una determinada región. Entre otros datos obtuvo la distribución del ahorro anual, habiéndose construido una tabla completa que posteriormente se extravió, quedando únicamente una fotocopia en mal estado con muchos datos borrosos. Reconstruya la tabla original a partir de la siguiente incompleta. Los datos aparecen en miles de pesetas.

# Problema 5

| Clases | $m_i$ | $n_i$ | $N_i$ | $f_i$ | $F_i$ | Amplitud |
|---|---|---|---|---|---|---|
| (20, 50] | 35 | 2 | 2 | | | |
| ( , 60] | | | 8 | | 0.04 | |
| (60, ] | | | | | 0.12 | 10 |
| ( , ] | 75 | | | 0.13 | | |
| ( , ] | | 44 | | | | |
| (100, ] | | | 119 | | | |
| ( , 150] | | | | 0.11 | | 25 |
| (150, ] | | | 171 | | | 50 |
| ( , 250] | | | 188 | | | |
| ( , 850] | 550 | 12 | 200 | | | |

# Problema 5

| Clases | $m_i$ | $n_i$ | $N_i$ | $f_i$ | $F_i$ | $\%_i$ | Amplitud |
|---|---|---|---|---|---|---|---|
| (20, 50] | 35 | 2 | 2 | 0.01 | 0.01 | 1% | 30 |
| (50, 60] | 55 | 6 | 8 | 0.03 | 0.04 | 3% | 10 |
| (60, 70] | 65 | 16 | 24 | 0.08 | 0.12 | 8% | 10 |
| (70, 80] | 75 | 26 | 50 | 0.13 | 0.25 | 13% | 10 |
| (80, 100] | 90 | 44 | 94 | 0.22 | 0.47 | 22% | 20 |
| (100, 125] | 112.5 | 25 | 119 | 0.125 | 0.595 | 12.5% | 25 |
| (125, 150] | 137.5 | 22 | 141 | 0.11 | 0.705 | 11% | 25 |
| (150, 200] | 175 | 30 | 171 | 0.15 | 0.855 | 15% | 50 |
| (200, 250] | 225 | 17 | 188 | 0.085 | 0.94 | 8% | 50 |
| (250, 850] | 550 | 12 | 200 | 0.06 | 1 | 6% | 600 |

# Problema 5

Se pide:

a).- Construir el histograma de frecuencias absolutas.

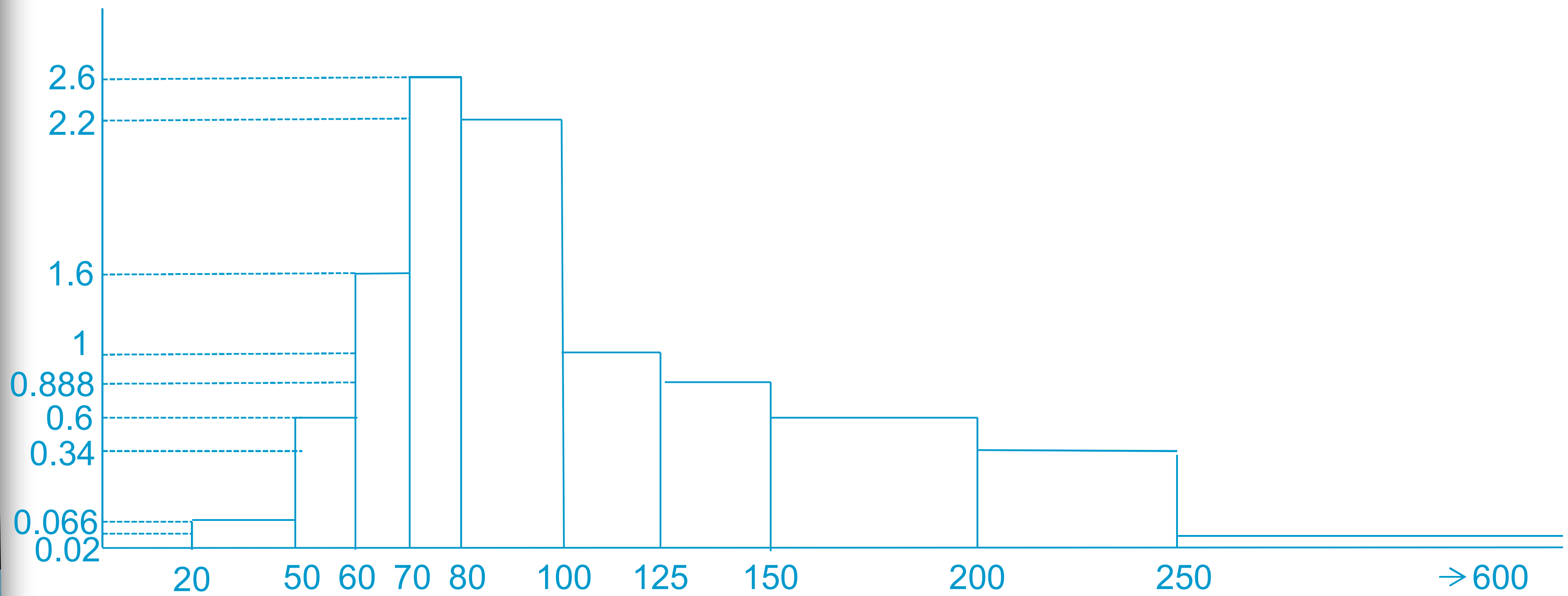
Sea X = "El ahorro anual de los individuos de una determinada región". Para construir el histograma, primero calculamos las alturas, esto es, $h_i = \frac{n_i}{c_i}$.

$$h_1 = \frac{2}{30} = 0.0\hat{6} \quad h_2 = \frac{6}{10} = 0.6 \quad h_3 = \frac{16}{10} = 1.6 \quad h_4 = \frac{26}{10} = 2.6 \quad h_5 = \frac{44}{20} = 2.2$$

$$h_6 = \frac{25}{25} = 1 \quad h_7 = \frac{22}{25} = 0.\hat{8} \quad h_8 = \frac{30}{50} = 0.6 \quad h_9 = \frac{17}{50} = 0.34 \quad h_{10} = \frac{12}{600} = 0.02$$

Entonces el histograma tendrá la siguiente forma:

# Problema 5



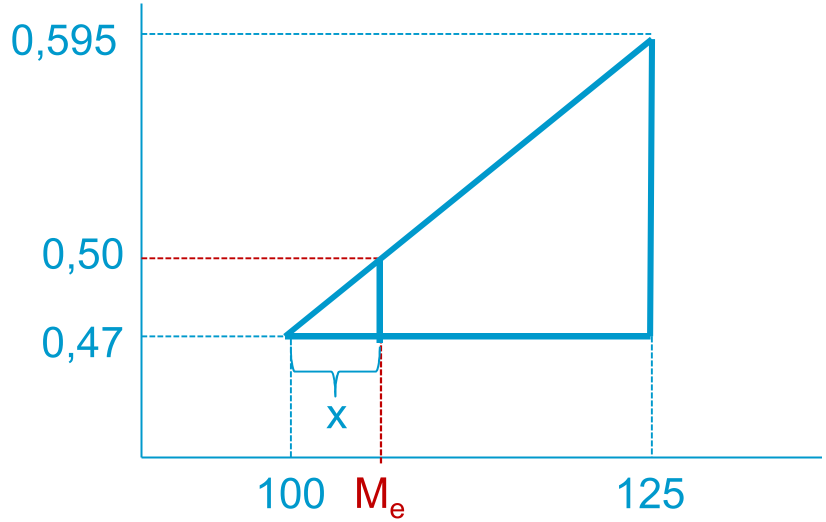

b).- Calcular la media, mediana y moda.

La media del ahorro anual sería:

$$\bar{x} = \frac{\sum_{i=1}^{k} m_i n_i}{n} = \frac{35x2+55x6 \;\; \cdots+550x12}{200} = \frac{28862{,}5}{200} = 144.3125$$

# Problema 5

Veamos ahora la mediana.




Aplicamos la relación de semejanza de triángulos

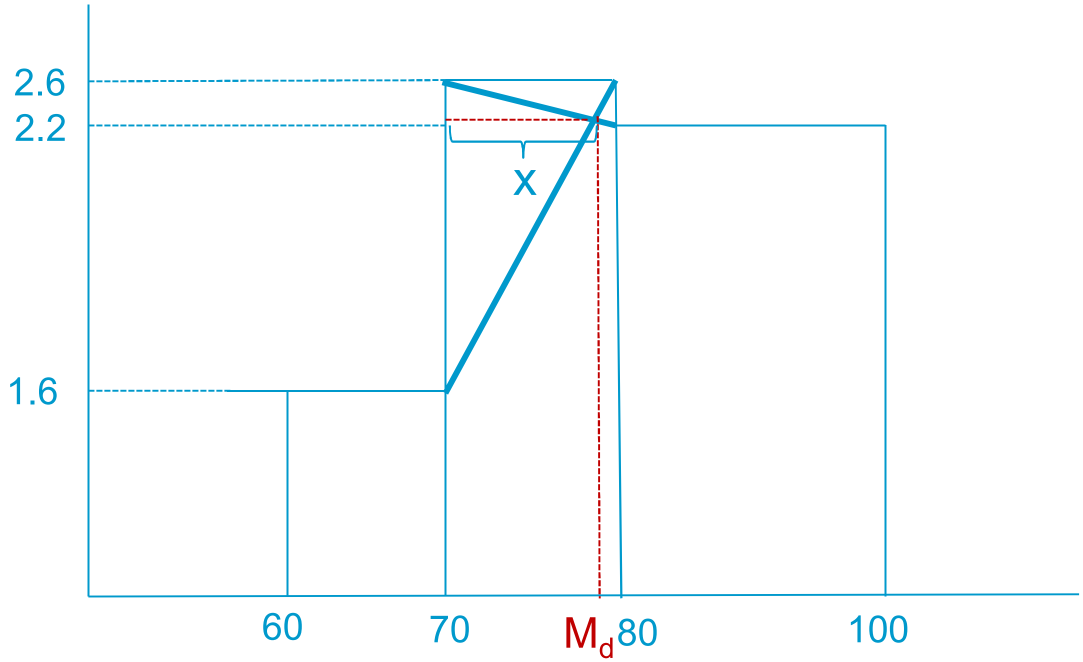
$$\frac{x}{0.5-0.47}=\frac{125-100}{0.595-0.47}$$

$$\frac{x}{0.03}=\frac{25}{0.125}$$

$$x=\frac{0.75}{0.125}=6 \;\rightarrow\; M_e=100+6=106$$

# Problema 5

Por último veamos la moda del ahorro anual de los individuos. Para ello usamos la parte del histograma, donde se encuentra el intervalo modal, esto es, el rectángulo con mayor área por unidad de base.




Se aplica la relación de semejanza entre los triángulos de la pajarita.

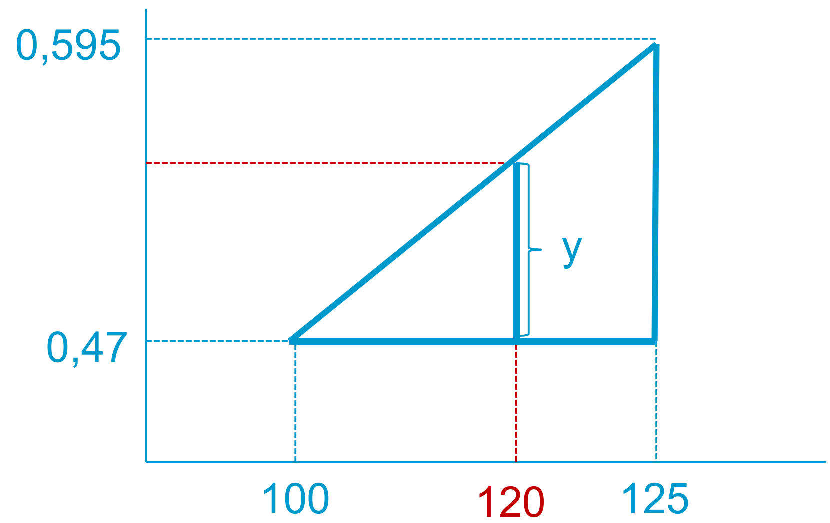
$$\frac{x}{2.6-1.6}=\frac{10-x}{2.6-2.2}\rightarrow\frac{x}{1}=\frac{10\ -x}{0.4}\rightarrow 10-x=0.4x$$

$$x=\frac{10}{1.4}=7.1428\rightarrow\ \ M_d=70+7.1428=77.1428$$

# Problema 5

c).- ¿Qué porcentaje de personas presentan un ahorro superior a 120000 ptas?

Para calcular el porcentaje de personas con un ahorro superior a 120000 pesetas, primero calculamos el porcentaje que tienen un ahorro inferior o igual a 120000 ptas.




Aplicamos entonces la relación de semejanza de triángulos

$$\frac{y}{120-100} = \frac{0.595-0.47}{125-100} \rightarrow \frac{y}{20} = \frac{0.125}{25}$$

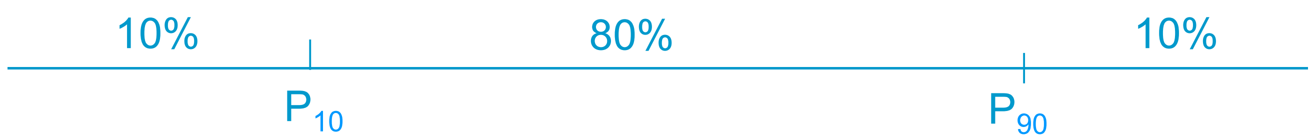
$$25y = 2.5 \rightarrow y = \frac{2.5}{25} = 0.1 \rightarrow \quad 0.47 + 0.1 = 0.57$$

# Problema 5

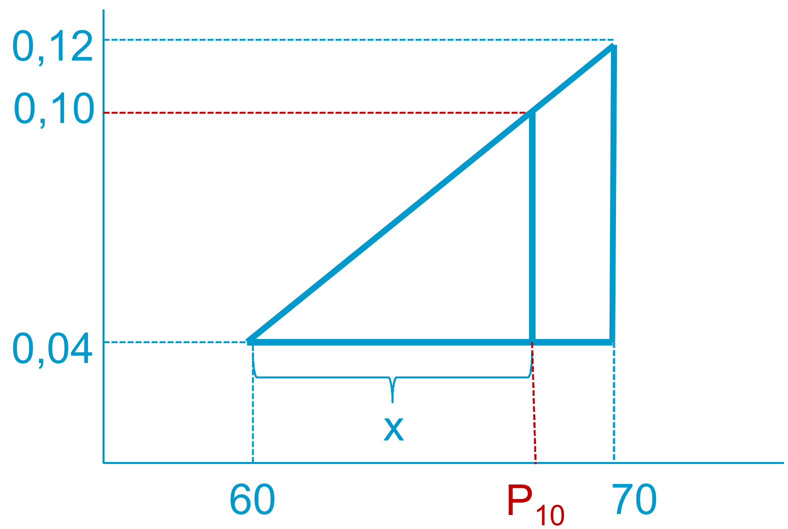
El 57% tienen un ahorro inferior a 120000 ptas, entonces 100-57=43% tienen un ahorro superior a 120000 ptas.

d).- ¿Entre qué valores de ahorro está situado el 80% central de las personas estudiadas?

Para determinar los valores de ahorro entre los cuales está el 80% central de las personas hemos de calcular




como vemos en el gráfico, los percentiles $P_{10}$ y $P_{90}$.



# Problema 5

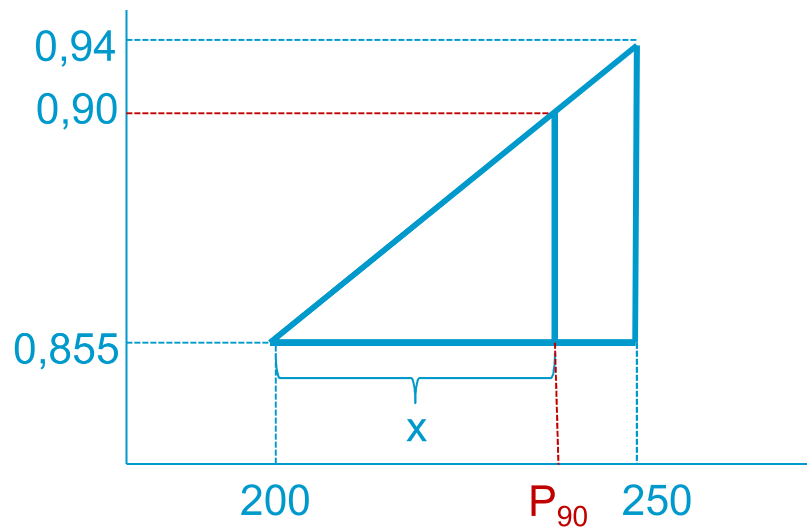
Aplicamos la relación de semejanza de triángulos

$$\frac{x}{0.10 - 0.04} = \frac{70 - 60}{0.12 - 0.04}$$

$$\frac{x}{0.06} = \frac{10}{0.08}$$

$$x = \frac{0.6}{0.08} = 7.5 \rightarrow \quad P_{10} = 60 + 7.5 = 67.5$$

Veamos ahora el percentil 90.



# Problema 5

Aplicamos de nuevo la relación de semejanza de triángulos

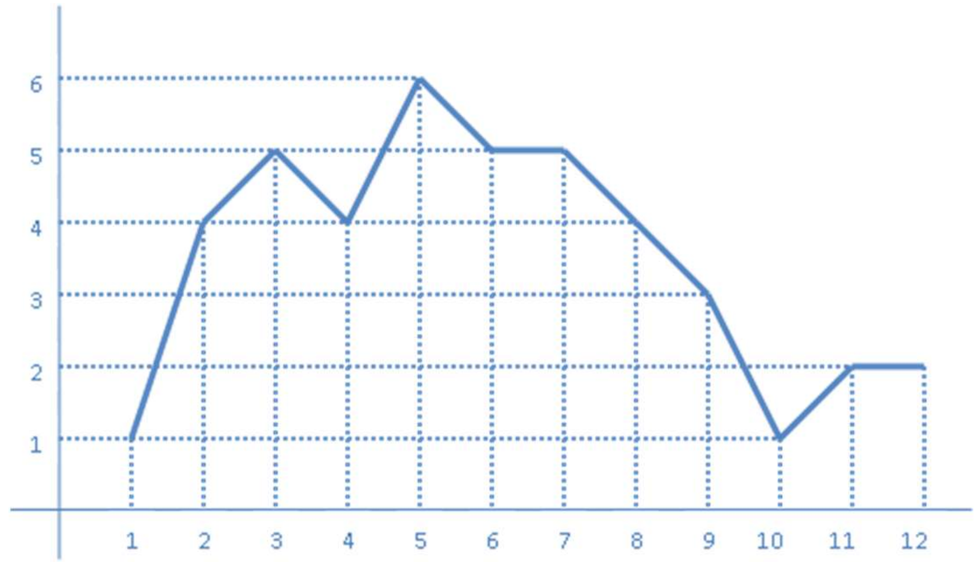
$$\frac{x}{0.90 - 0.855} = \frac{250 - 200}{0.94 - 0.855}$$

$$\frac{x}{0.045} = \frac{50}{0.085}$$

$$x = \frac{2{,}25}{0.085} = 26.47 \rightarrow \ P_{90} = 200 + 26.47 = 226.47$$

# Problema 6

Problema 6. El número de bits que circulan por un canal de comunicación durante diversos períodos de tiempo se muestra en el siguiente gráfico.




Se pide:

a).- Indicar qué tipo de gráfico es el representado anteriormente y qué tipo de variable se analiza en el mismo.

Sea X="El número de bits que circulan por un canal de comunicación". El gráfico es un polígono de frecuencias absolutas y la variable es discreta, ya que el polígono está despegado del eje X.

# Problema 6

b).- Construir la tabla de fecuencias.

La tabla de frecuencias sería como sigue:

| $x_i$ | $n_i$ | $N_i$ | $f_i$ | $F_i$ | $\%_i$ |
|---|---|---|---|---|---|
| 1 | 1 | 1 | 0,0238 | 0,0238 | 2,38% |
| 2 | 4 | 5 | 0,0952 | 0,1190 | 9,52% |
| 3 | 5 | 10 | 0,1190 | 0,2380 | 11,9% |
| 4 | 4 | 14 | 0,0952 | 0,3333 | 9,52% |
| 5 | 6 | 20 | 0,1428 | 0,4761 | 14,28% |
| 6 | 5 | 25 | 0,1190 | 0,5952 | 11,9% |
| 7 | 5 | 30 | 0,1190 | 0,7142 | 11,9% |
| 8 | 4 | 34 | 0,0952 | 0,8095 | 9,52% |
| 9 | 3 | 37 | 0,0714 | 0,8809 | 7,14% |
| 10 | 1 | 38 | 0,0238 | 0,9047 | 2,38% |
| 11 | 2 | 40 | 0,0476 | 0,9524 | 4,76% |
| 12 | 2 | 42 | 0,0476 | 1 | 4,76% |

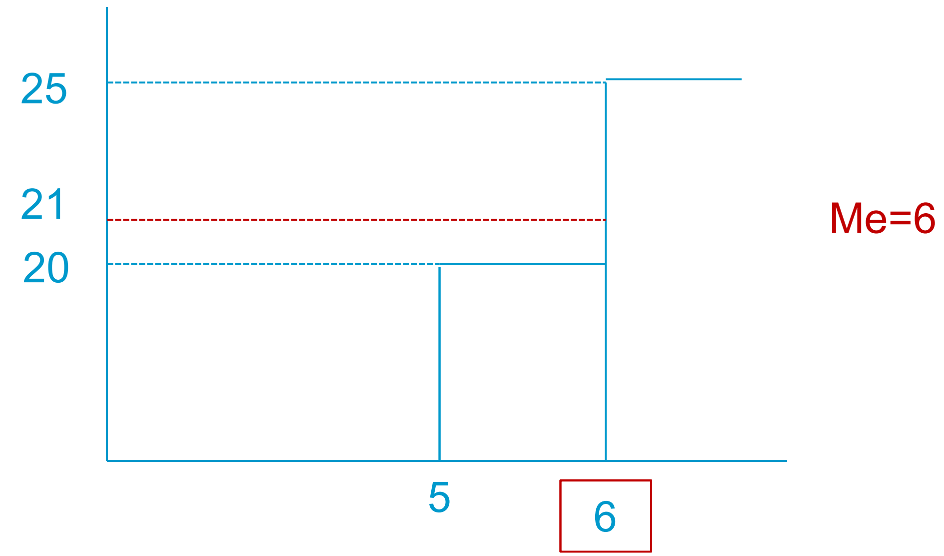
# Problema 6

c).- Calcular la media, la mediana y la moda.

Veamos la media aritmética.

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i n_i}{n} = \frac{1x1+2x4+\cdots+11x2+12x2}{42} = \frac{250}{42} = 5{,}9524$$

Veamos la mediana. Si calculamos $\frac{N}{2} = \frac{42}{2} = 21$. Como 21 está entre 20 y 25 en la columna de frecuencias absolutas acumuladas, entonces se puede ver en el gráfico que la mediana vale 6.

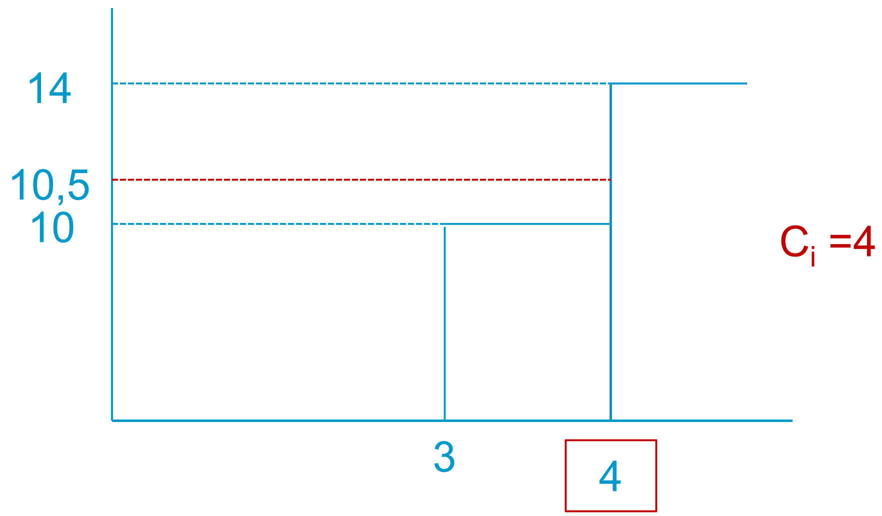



Por último la moda sería el valor que más se repite, que en este caso es $M_d = 5$.

# Problema 6

d).- Calcular los cuartiles

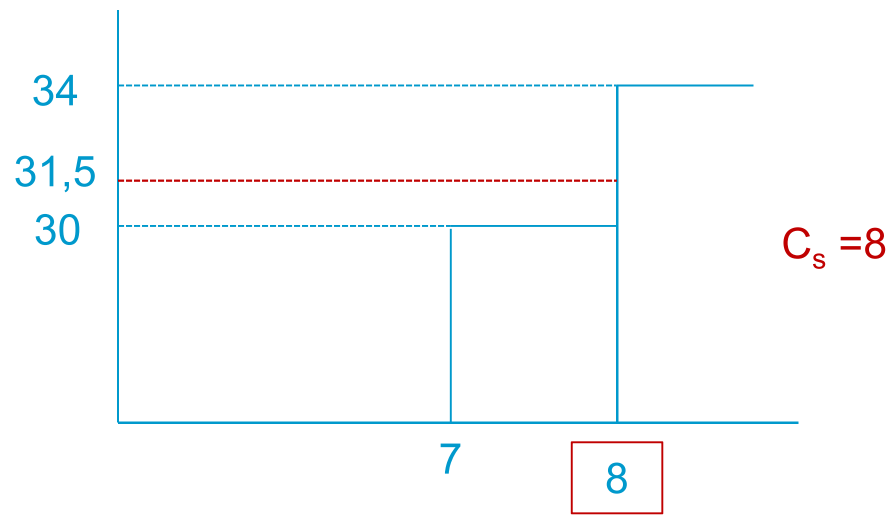
Veamos primero el cuartil inferior, $c_i$. Calculamos $\frac{N}{4} = \frac{42}{4} = 10{,}5$ y lo buscamos en la columna de frecuencias absolutas acumuladas. Se puede ver que está entre 10 y 14, entonces según el gráfico el cuartil inferior vale 4.




El segundo cuartil es la mediana, que ya vimos que vale 6.

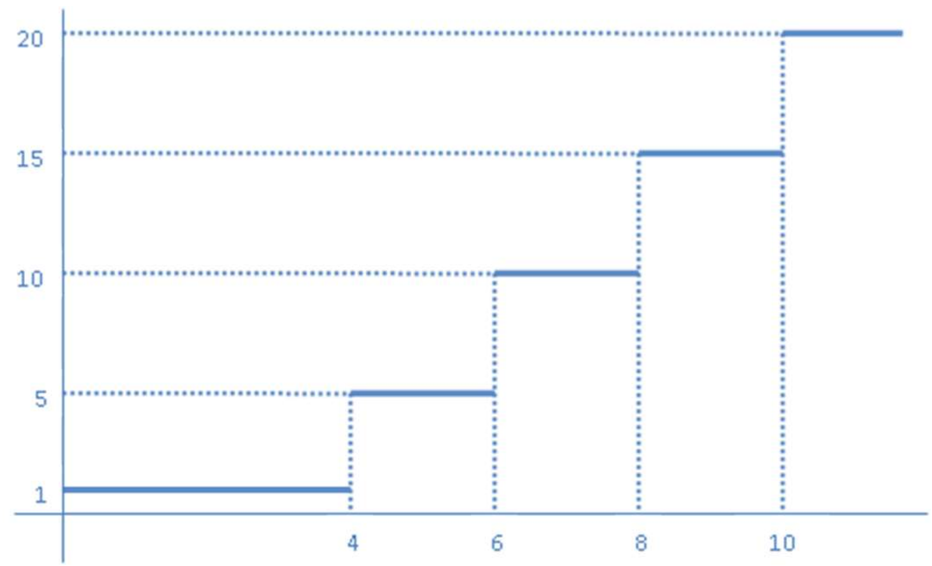
# Problema 6

El tercer cuartil o cuartil superior, $c_s$, se obtendría buscando $\frac{3N}{4} = \frac{3x42}{4} = 31{,}5$. Dicho valor se encuentra entre 30 y 34 en la columna de frecuencias absolutas acumuladas.



# Problema 7

Problema 7. A un grupo de alumnos se les realiza una prueba de rendimiento, siendo las puntuaciones obtenidas las que muestra el siguiente gráfico:




Se pide:

a).- ¿Qué tipo de gráfico recoge las puntuaciones de los alumnos?. Observando el mismo decir qué tipo de variable es la que se ha medido, y cuántos individuos consta la muestra.

Sea X = "Las puntuaciones de los alumnos". El gráfico que se analiza es un diagrama de frecuencias acumulado y pertenece a una variable es discreta.

# Problema 7

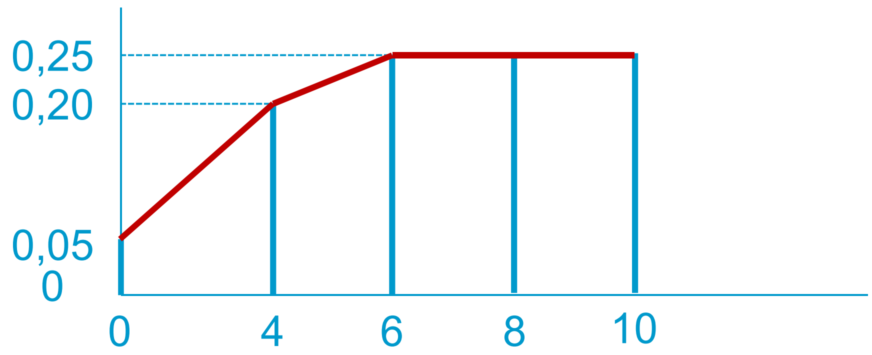
b).- Construir la tabla de fecuencias.

Veamos la tabla de frecuencias asociada al gráfico.

| $x_i$ | $n_i$ | $N_i$ | $f_i$ | $F_i$ | $\%_i$ |
|---|---|---|---|---|---|
| 0 | 1 | 1 | 0,05 | 0,05 | 5% |
| 4 | 4 | 5 | 0,20 | 0,25 | 20% |
| 6 | 5 | 10 | 0,25 | 0,50 | 25% |
| 8 | 5 | 15 | 0,25 | 0,75 | 25% |
| 10 | 5 | 20 | 0,25 | 1 | 25% |

c).- Representar gráficamente el polígono de frecuencias relativas.

El polígono de frecuencias relativas sería como sigue (rojo en el gráfico):



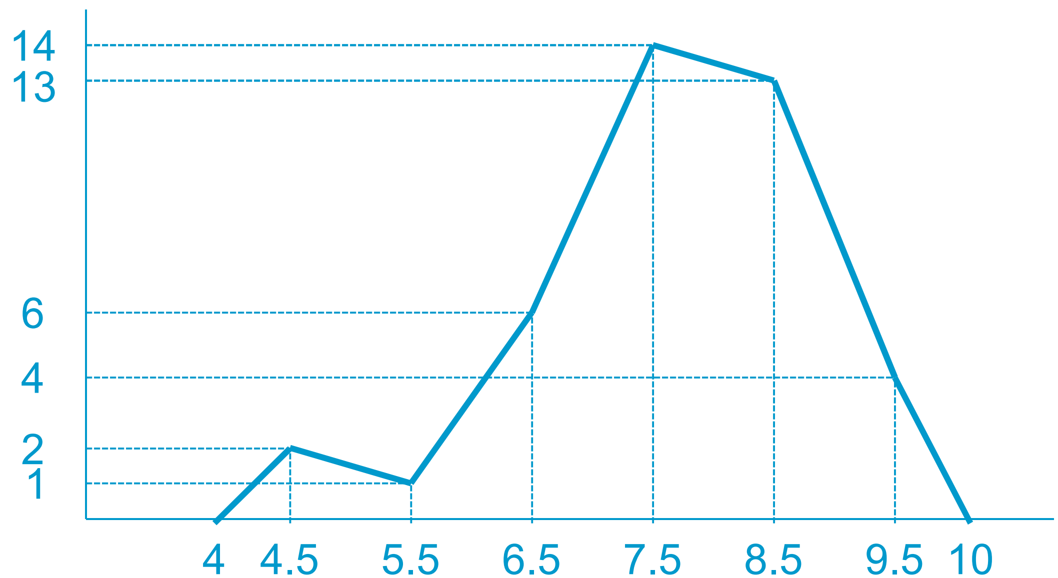
# Problema 7

d).- ¿Cuál es la frecuencia relativa acumulada del valor 7 de la variable?

La frecuencia absoluta acumulada del valor 7 de la variable es como vemos en el diagrama de frecuencias acumulada la misma que la del valor 6, esto es 10. Para calcular la relativa acumulada sólo debemos dividir por el número total de observaciones, 20, y nos da 0,5.

# Problema 8

Problema 8. Se ha analizado el tiempo que permanece conectado a Internet a lo largo de un día, un determinado equipo informático, obteniéndose el siguiente gráfico como resultado del estudio realizado.




Se pide:

a).- ¿Qué tipo de variable es la que se analiza?. Construir la tabla de frecuencias, que se corresponde con el gráfico anterior.

Sea X = “El tiempo de conexión a Internet”. Se trata de una variable de tipo continuo, ya que el polígono de frecuencias está pegado al eje de abcisas.

# Problema 8

Veamos la tabla de frecuencias asociada con el gráfico.

| Clases | $m_i$ | $h_i$ | $c_i$ | $n_i$ | $N_i$ | $f_i$ | $F_i$ | $\%_i$ |
|---|---|---|---|---|---|---|---|---|
| [4, 5] | 4,5 | 2 | 1 | 2 | 2 | 0,05 | 0,05 | 5% |
| [5, 6] | 5,5 | 1 | 1 | 1 | 3 | 0,025 | 0,075 | 2,5% |
| [6, 7] | 6,5 | 6 | 1 | 6 | 9 | 0,15 | 0,225 | 15% |
| [7, 8] | 7,5 | 14 | 1 | 14 | 23 | 0,35 | 0,575 | 35% |
| [8, 9] | 8,5 | 13 | 1 | 13 | 36 | 0,325 | 0,9 | 32,5% |
| [9, 10] | 9,5 | 4 | 1 | 4 | 40 | 0,10 | 1 | 10% |

b).- Calcular la media, mediana y moda del tiempo de conexión.

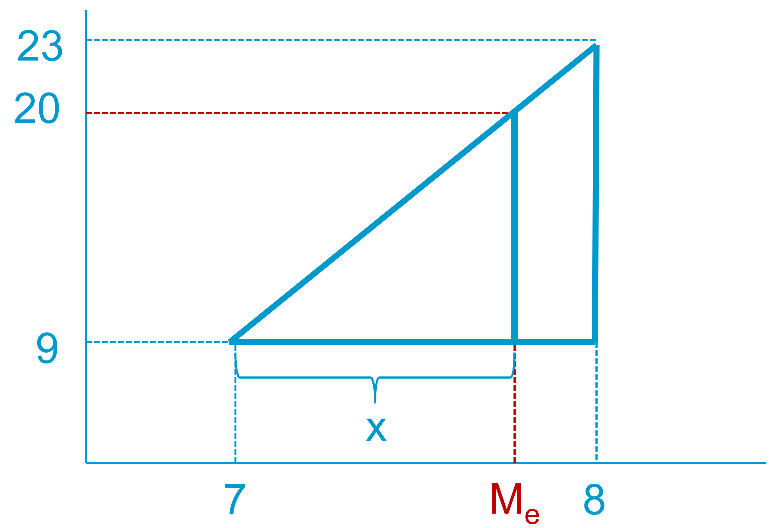
Comenzamos con la media aritmética.

$$\bar{x} = \frac{\sum_{i=1}^{k} m_i n_i}{n} = \frac{4.5x2+\ \ .5x1+\cdots+9.5x4}{40} = \frac{307}{40} = 7.675$$

Seguimos con la mediana. Para la cual buscamos $\frac{N}{2} = \frac{40}{2} = 20$ en la columna de frecuencias absolutas acumuladas.

# Problema 8

Entonces dibujamos el trozo de polígono acumulado en el que está 20.




Aplicamos la relación de semejanza de triángulos

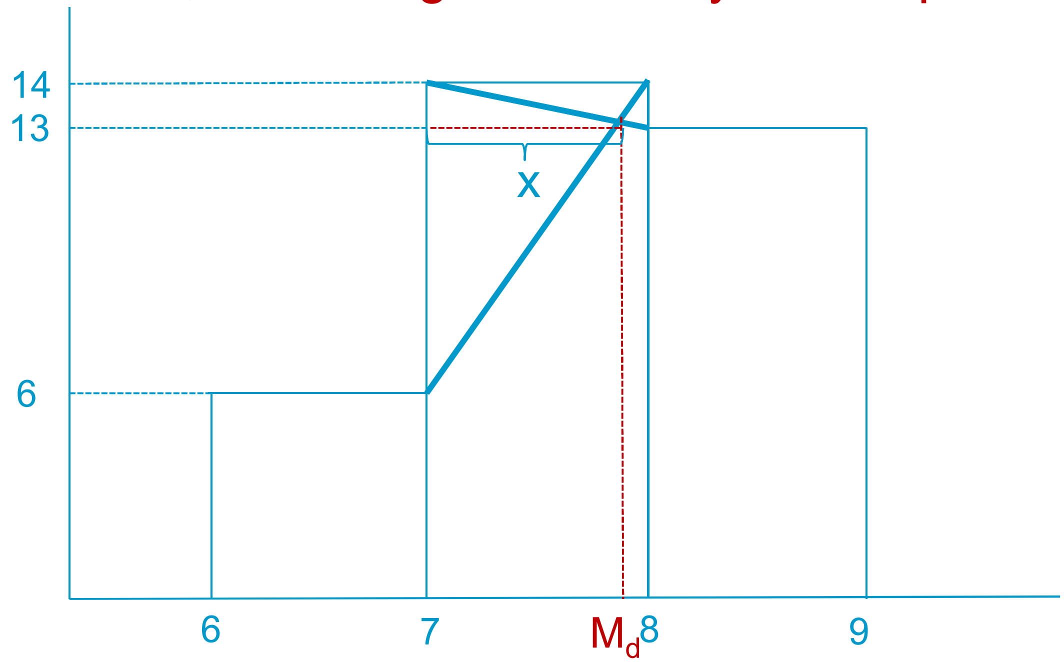
$$\frac{x}{20-9} = \frac{8-7}{23\ -9}$$

$$\frac{x}{11} = \frac{1}{14}$$

$$x = \frac{11}{14} = 0.7857 \rightarrow \ M_e = 7 + 7.7857 = 7.7857$$

# Problema 8

Por último, veamos la moda del tiempo de conexión a Internet. Para ello usamos la parte del histograma, donde se encuentra el intervalo modal, esto es, el rectángulo con mayor área por unidad de base.




Se aplica la relación de semejanza entre los triángulos de la pajarita.

$$\frac{x}{14-6} = \frac{1-x}{14-13} \rightarrow x = 8 \cdot (1-x) \rightarrow x = 8 - 8x \rightarrow 9x = 8$$

$$x = \frac{8}{9} = 0.\hat{8} \rightarrow \quad M_d = 7 + 0.\hat{8} = 7.\hat{8}$$

# Problema 8

c).- Hallar el coeficiente de variación de Pearson y el rango intercuartílico. ¿Se puede considerar que la media de la muestra es representativa? ¿Porqué?

Veamos la varianza:

| $m_i$ | 4,5 | 5,5 | 6,5 | 7,5 | 8,5 | 9,5 | |
|---|---|---|---|---|---|---|---|
| $n_i$ | 2 | 1 | 6 | 14 | 13 | 4 | 40 |
| $m_i x n_i$ | 9 | 5,5 | 39 | 105 | 110,5 | 38 | 307 |
| $m^2_i x n$ | 40,5 | 30,25 | 253,5 | 787,5 | 939,25 | 361 | 2412 |

$$S^2 = \sum_{i=1}^{k} \frac{(x_i)^2 n_i}{n} - \bar{x}^2 =$$

$$= \frac{(4,5)^2 \cdot 2 + (5,5)^2 \cdot 1 + \ldots + (9.5)^2 \cdot 4}{40} - (7.675)^2$$

$$= \frac{2412}{40} - 58,905625 = 60,3 - 58,905625 = 1,394375$$

$$S = 1,180836$$

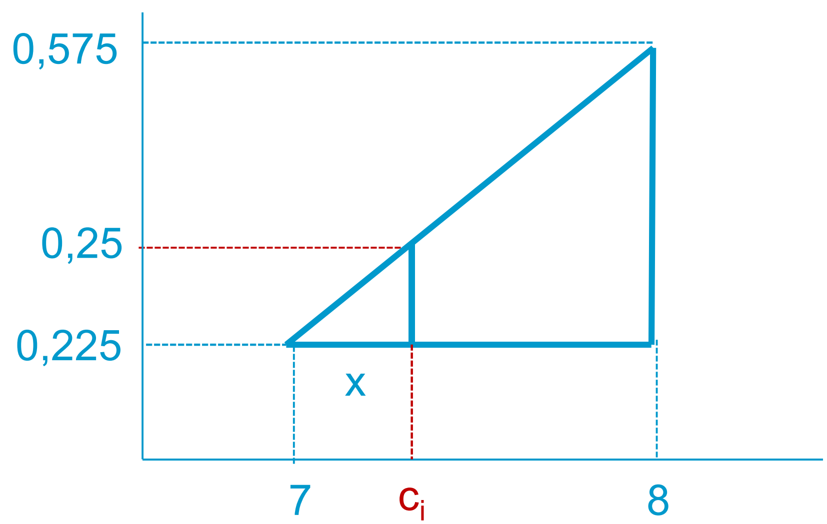
# Problema 8

Entonces el coeficiente de variación de Pearson sería:

$$C.V = \frac{S}{|\bar{x}|} = \frac{1,180836}{7,675} = 0.153855$$

Como el coeficiente de variación de Pearson es menor que uno la media es representativa.

Para determinar el recorrido intercuartílico tenemos que calcular el cuartil inferior y el cuartil superior.



# Problema 8

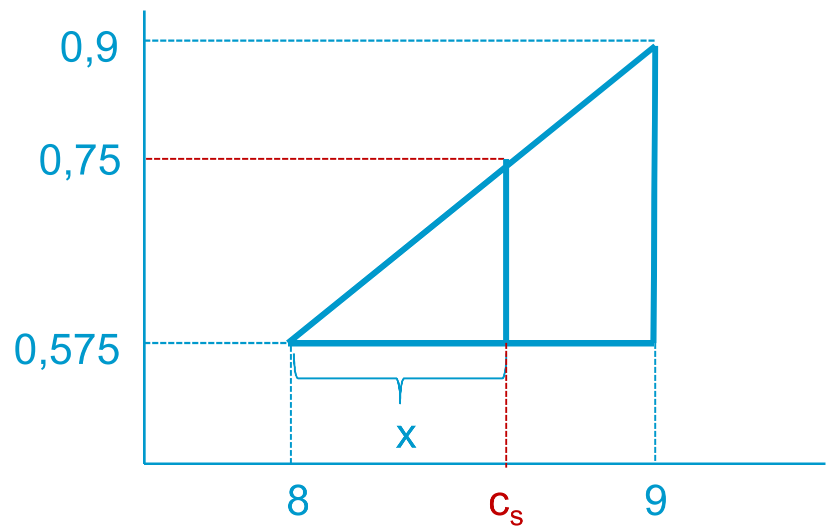
Aplicamos entonces la relación de semejanza de triángulos

$$\frac{x}{0.25 - 0.225} = \frac{8 - 7}{0.575 - 0.225}$$

$$\frac{x}{0.025} = \frac{1}{0.35}$$

$$x = \frac{0.025}{0.35} = 0.071428 \rightarrow \quad c_i = 7 + 0.071428 = 7.071428$$

Veamos ahora el cuartil superior



# Problema 8

Aplicamos de nuevo la relación de semejanza de triángulos

$$\frac{x}{0.75-0.575}=\frac{9-8}{0.9-0.575}$$

$$\frac{x}{0.175}=\frac{1}{0.325}$$

$$x=\frac{0.175}{0.325}=0.53846\rightarrow \quad c_s=8+0.53846=8.53846$$

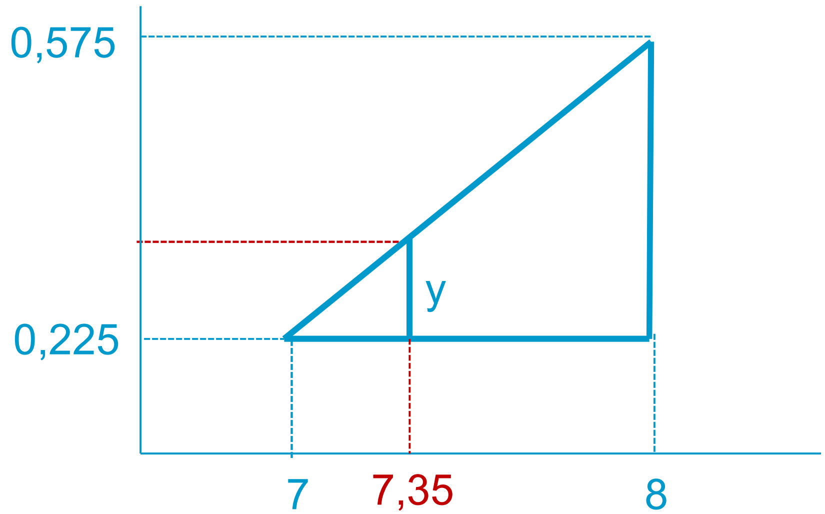
Entonces el recorrido intercuartílico sería:

$$c_s \ - \ c_i=8{,}53846 \ -7.071428=1{,}467032$$

d).- Calcular el porcentaje de días que el equipo permanece conectado menos de 7.35 horas . ¿Y más de 8.22?

Veamos primero el porcentaje de días que el equipo permanece conectado menos de 7.35 horas.

# Problema 8




Aplicamos entonces la relación de semejanza de triángulos

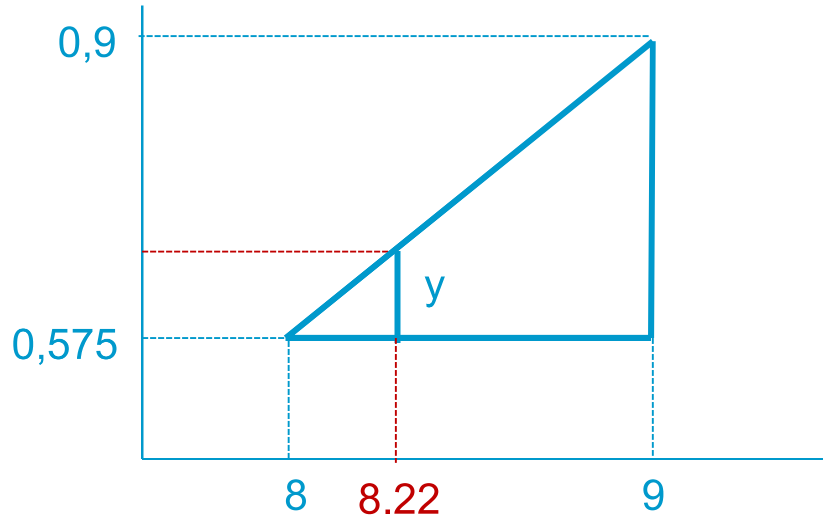
$$\frac{y}{7.35-7}=\frac{0.575-0.225}{8-7}$$

$$y=\ (0.35)^2=0.1225\rightarrow\ \ 0.225+0.1225=0.3475$$

Entonces el 34.75% de los días la conexión es menor a 7.35 horas.

Veamos ahora el porcentaje de días cuya conexión es mayor que 8.22 horas. Para ello primero calculamos el porcentaje que es menor que 8.22 horas y se le resta a 100.

# Problema 8




Aplicamos entonces la relación de semejanza de triángulos

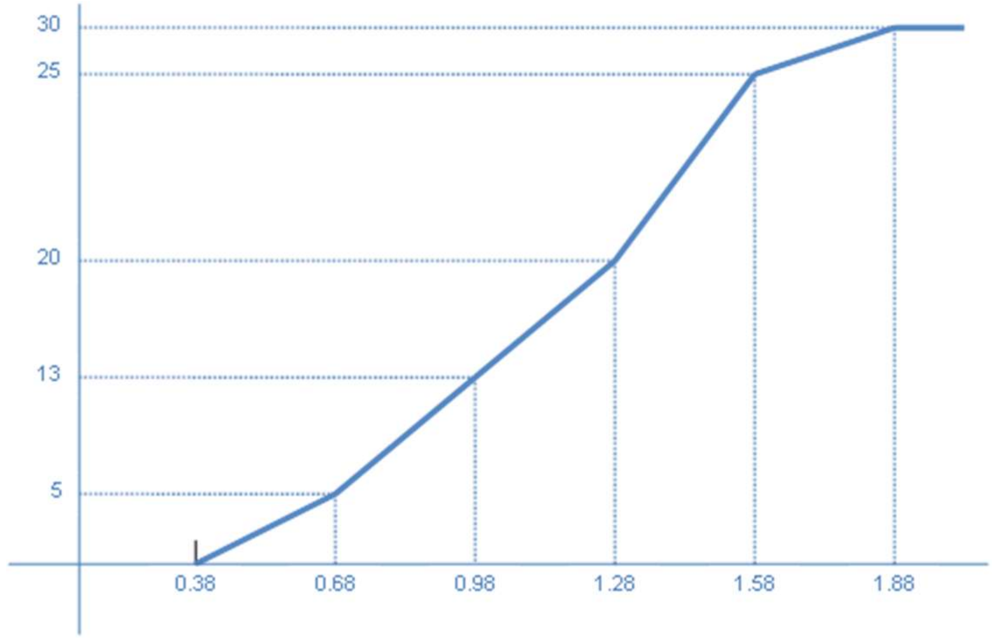
$$\frac{y}{8.22-8}=\frac{0.9-0.575}{9-8}$$

$$y=0.22\cdot 0.325=0.0715 \rightarrow \quad 0.575+0.0715=0.6465$$

El 64.65% quedan inferiores o iguales a 8.22 entonces superiores a 8.22 sería 100 - 64.65 = 35.35% de los días.

# Problema 9

Problema 9. Los costos de ejecución de programas de computadoras con el proceso de tiempo compartido varían de una sesión a otra. El siguiente gráfico recoge los costos por sesión para los diferentes usuarios (dado en Euros).




Se pide:

a).- ¿Qué gráfico es el anterior, y qué tipo de variable se analiza?. Construir la tabla de frecuencias asociada con dicho gráfico.

Sea X = "Costos de ejecución de los programas". El gráfico que se analiza es un polígono de frecuencias absolutas acumulado, ya que es creciente y por supuesto, es una variable continua.

# Problema 9

Veamos la tabla de frecuencias asociada con el gráfico.

| Clases | $m_i$ | $n_i$ | $N_i$ | $f_i$ | $F_i$ | $\%_i$ | $h_i$ |
|---|---|---|---|---|---|---|---|
| [0.38, 0.68] | 0.53 | 5 | 5 | $0.1\hat{6}$ | $0.1\hat{6}$ | 16.6% | $16.\hat{6}$ |
| [0.68, 0.98] | 0.83 | 8 | 13 | $0.2\hat{6}$ | $0.4\hat{3}$ | 26.6% | $26.\hat{6}$ |
| [0.98, 1.28] | 1.13 | 7 | 20 | $0.2\hat{3}$ | $0.\hat{6}$ | 23.3% | $23.\hat{3}$ |
| [1.28, 1.58] | 1.43 | 5 | 25 | $0.1\hat{6}$ | $0.8\hat{3}$ | 16.6% | $16.\hat{6}$ |
| [1.58, 1.88] | 1.73 | 5 | 30 | $0.1\hat{6}$ | 1 | 16.6% | $16.\hat{6}$ |

b).- Calcular la media, mediana y moda de la variable en estudio.

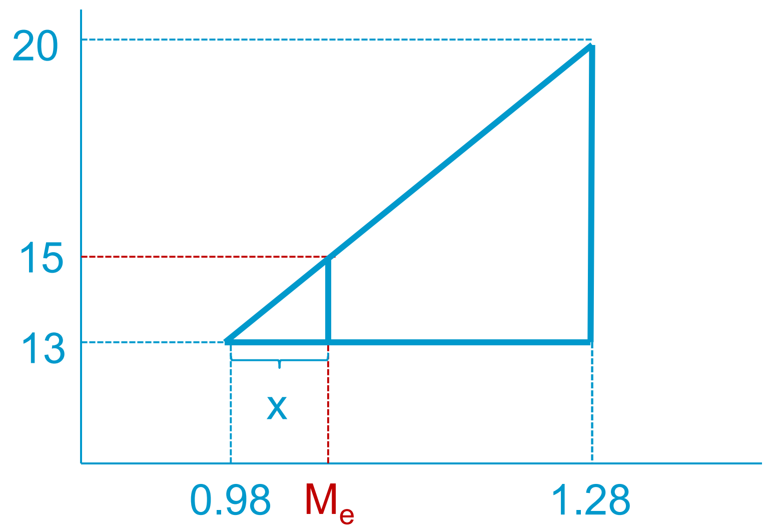
Comenzamos con la media aritmética.

$$\bar{x} = \frac{\sum_{i=1}^{k} m_i n_i}{n} = \frac{0.53x5+0.83x8+\cdots+1.73x5}{30} = \frac{33}{30} = 1.1$$

Seguimos con la mediana. Para la cual buscamos $\frac{N}{2} = \frac{30}{2} = 15$ en la columna de frecuencias absolutas acumuladas.

# Problema 9

Entonces dibujamos el trozo de polígono acumulado en el que está 15.




Aplicamos la relación de semejanza de triángulos

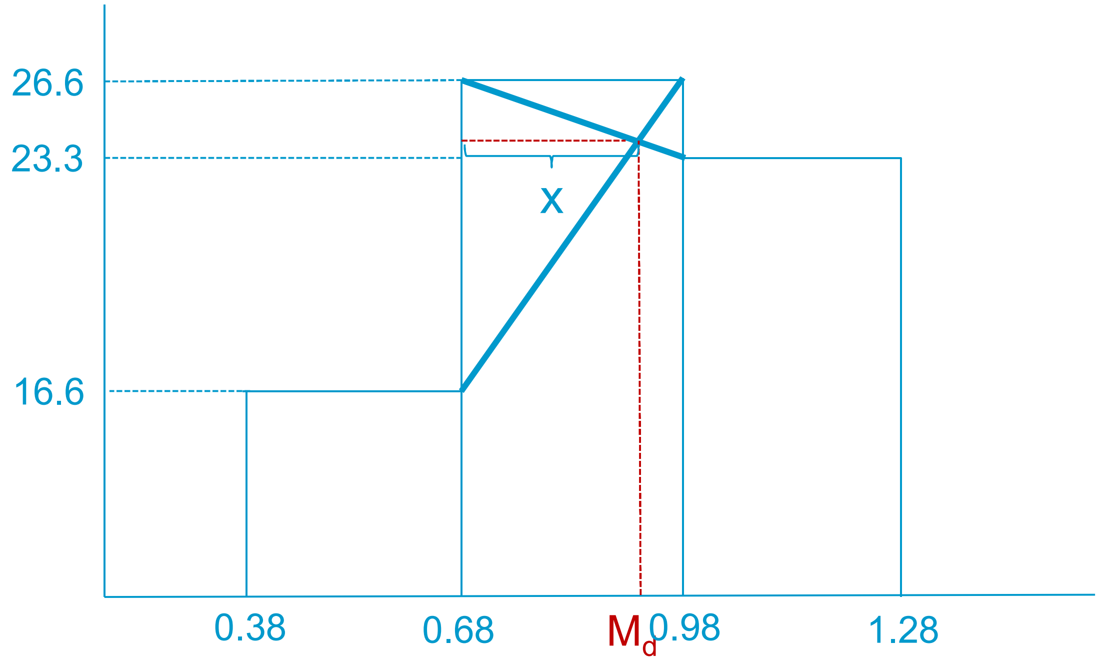
$$\frac{x}{15-13} = \frac{1.28-0.98}{20\ -13}$$

$$\frac{x}{2} = \frac{0.3}{7}$$

$$x = \frac{0,6}{7} = 0.0857 \rightarrow \ \ M_e = 0.98 + 0.0857 = 1.0657$$

# Problema 9

Por último, veamos la moda del costo de ejecución de los programas. Para ello usamos la parte del histograma, donde se encuentra el intervalo modal, esto es, el rectángulo con mayor área por unidad de base.




Se aplica la relación de semejanza entre los triángulos de la pajarita.

$$\frac{x}{26.\hat{6}-16.\hat{6}} = \frac{0.3-x}{26.\hat{6}-23.\hat{3}} \rightarrow \frac{x}{10} = \frac{0.3-x}{3.\hat{3}} \rightarrow 3.\hat{3}x = 3 - 10x \rightarrow 13.\hat{3}x = 3$$

$$x = \frac{3}{13.\hat{3}} = 0.225 \rightarrow \quad M_d = 0.68 + 0.225 = 0.905$$

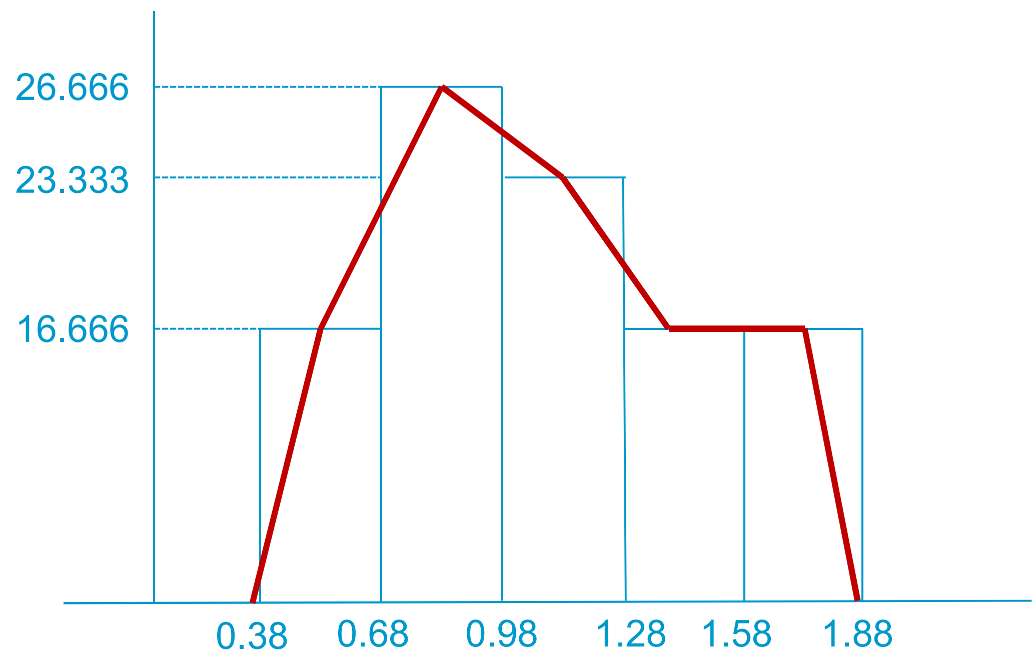
# Problema 9

b).- Representar el polígono de frecuencias de la variable.

Para representar el polígono de frecuencias, primero debemos dibujar el histograma, y para ello hemos de calcular las alturas, $h_i = \frac{n_i}{c_i}$.

$$h_1 = h_4 = h_5 = \frac{5}{0.3} = 16.\hat{6} \quad h_2 = \frac{8}{0{,}3} = 26.\hat{6} \quad h_3 = \frac{7}{0.3} = 23.\hat{3}$$

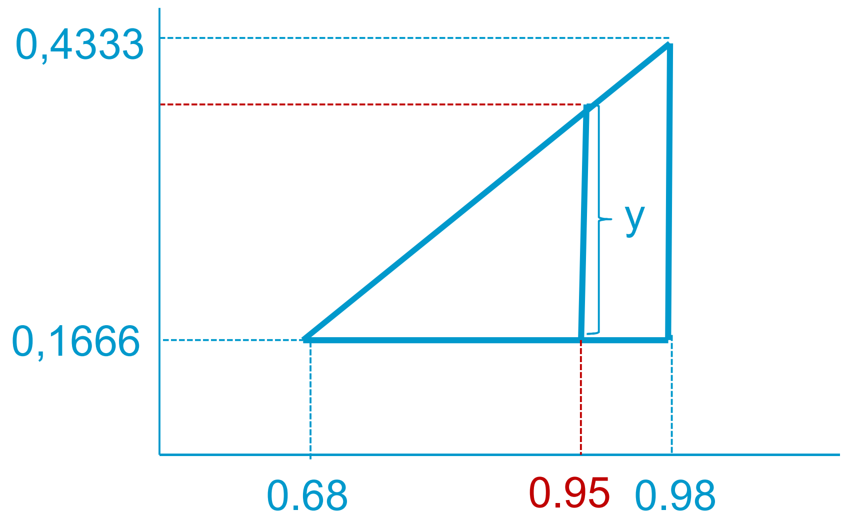
Entonces el histograma y el polígono tendrán la siguiente forma:



# Problema 9

d).- ¿Qué porcentaje de programas tienen un costo superior a 0.95 Euros? ¿Entre qué costos se encuentra el 70% de los programas con menor coste?

Veamos primero el porcentaje de programas con un costo inferior a 0.95 Euros, y luego se lo restamos a 100 para saber cuantos tienen un costo superior.




Aplicamos entonces la relación de semejanza de triángulos

$$\frac{y}{0.95-\ .68} = \frac{0.4\hat{3}-0.1\hat{6}}{0.98-0.68} \rightarrow \frac{y}{0.27} = \frac{0.2\hat{6}}{0.3}$$

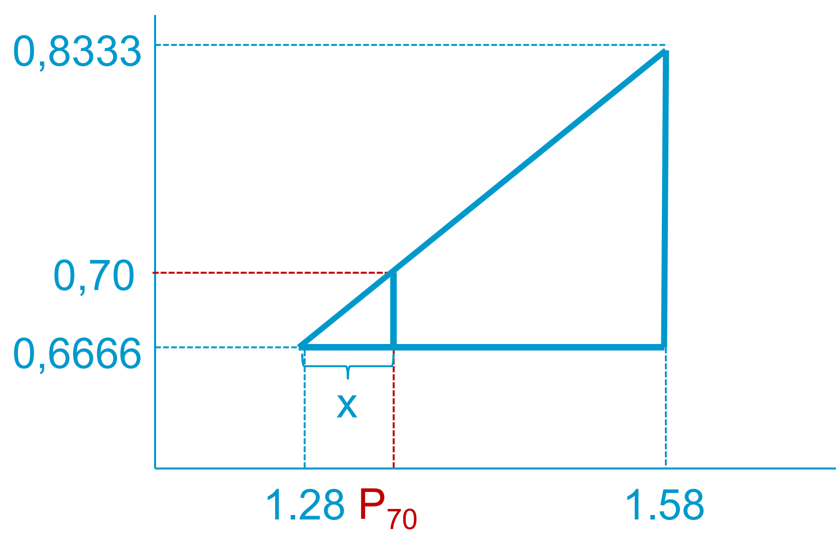
$$y = \frac{0.2\hat{6}\ \cdot 0{,}27}{0.3} = 0.24 \rightarrow\ \ 0.1\hat{6} + 0.24 = 0.40\hat{6}$$

# Problema 9

Entonces el $40.\hat{6}\%$ de los programas tienen un costo inferior a 0.95 y el 100 - $40.\hat{6}$ = $59.\hat{3}\%$ un costo superior.

Veamos ahora los costos entre los cuales se encuentra el 70% de las observaciones menores.


El mínimo valor es el 0.38, ahora calculamos el percentil 70.



# Problema 9

Aplicamos entonces la relación de semejanza de triángulos

$$\frac{x}{0.70 - 0.\hat{6}} = \frac{1.58 - 1.28}{0.8\hat{3} - 0.\hat{6}}$$

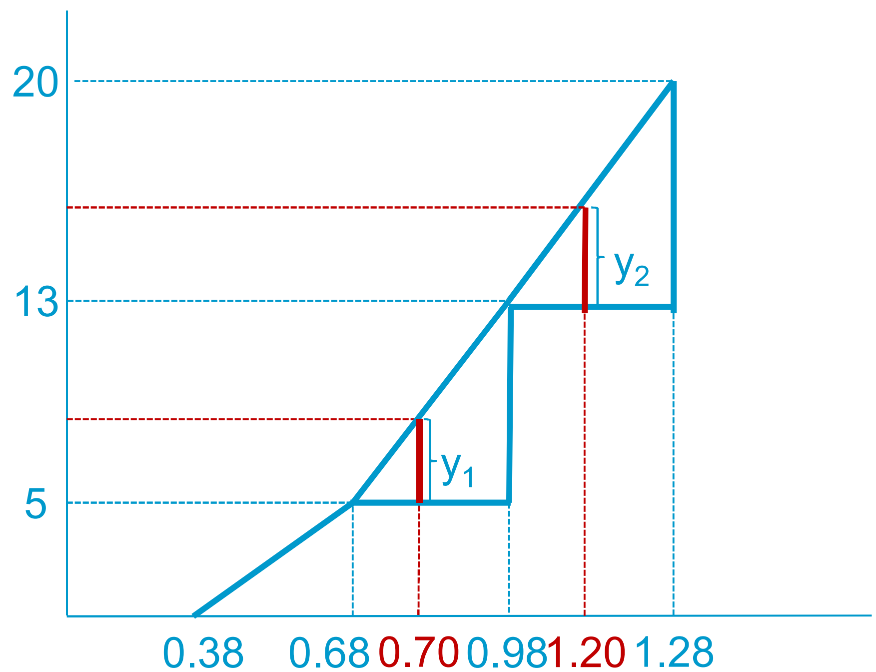
$$\frac{x}{0.0\hat{3}} = \frac{0.3}{0.1\hat{6}}$$

$$x = \frac{0.00\hat{9}}{0.1\hat{6}} = 0.05\hat{9} \;\rightarrow\; P_{70} = 1.28 + 0.05\hat{9} = 1.33\hat{9}$$

e).- ¿En cuántos programas el costo se mantuvo dentro el intervalo 0.70 y 1.25 euros?.

Como nos pregunta por el número de programas con costos entre 0.70 y 1.25, vamos a utilizar el polígono de frecuencias absolutas acumulado. Calculamos el número de programas con un coste inferior o igual a 1.25 y con un coste inferior o igual a 0.70 y los restamos para obtener los programas con un coste en medio de ambos costes.

# Problema 9

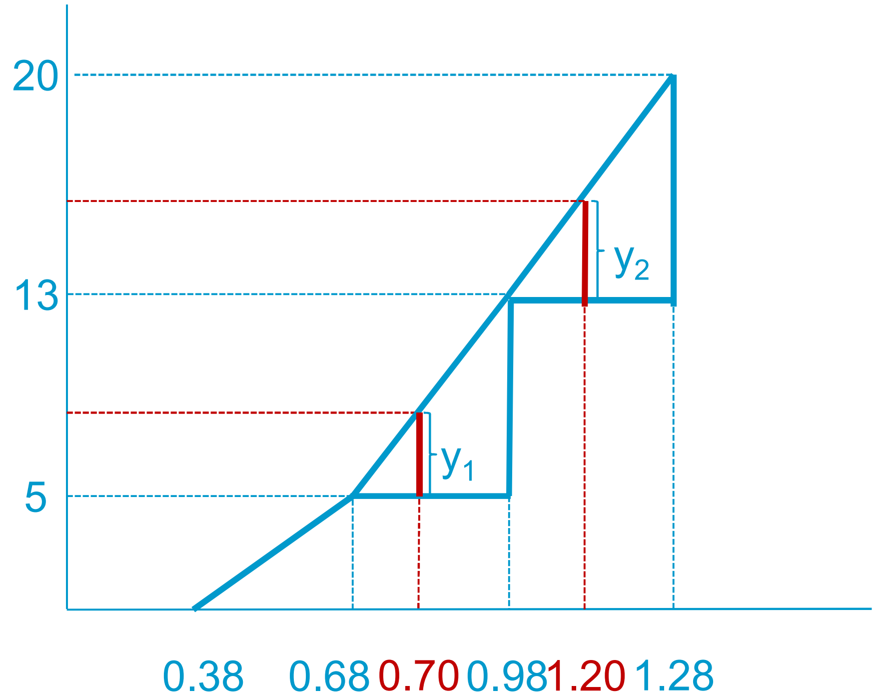



Aplicamos en ambos casos la relación de semejanza entre los triángulos rectángulos

$$\frac{y_1}{0.70-0.68} = \frac{13-5}{0.98-0.68} \rightarrow \frac{y_1}{0.02} = \frac{8}{0.3}$$

$$y_1 = \frac{0.02 \cdot 8}{0.3} = 0.5\hat{3} \rightarrow \quad 5 + 0.5\hat{3} = 5.5\hat{3}$$

# Problema 9




Para el caso de $y_2$ sería

$$\frac{y_2}{1.25-0.98} = \frac{20-13}{1.28-0.98} \rightarrow \frac{y_2}{0.27} = \frac{7}{0.3}$$

$$y_1 = \frac{0.27 \cdot 7}{0.3} = 6.3 \rightarrow 13 + 6.3 = 19.3$$

En medio estarían $19.3 - 5.5\hat{3} = 13.7\hat{6} \sim 14$ programas.

# Problema 10

Problema 10. Consideremos las siguientes muestras:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Muestra 1: | 10 | 9 | 8 | 7 | 8 | 6 | 10 | 6 |
| Muestra 2: | 10 | 6 | 10 | 6 | 8 | 10 | 8 | 6 |
| Muestra 3: | 9 | 7 | 9 | 7 | 10 | 8 | 8 | 7 |

Se pide:

a).- Calcular el recorrido para las distintas muestras. ¿Es posible concluir que todas las muestras exhiben la misma variabilidad?.

Veamos entonces los recorridos y rangos de las tres muestras.

Muestra 1: r = 10-6 =4; Muestra 2: 10-6=4; Muestra 3: 10-7= 3

No es posible de hablar de la misma variabilidad sólo con los rangos, aunque algunos coincidan.

# Problema 10

b).- Calcular la desviación típica de cada una de las muestras. ¿Estas cantidades sirven para indicar quién tiene mayor variabilidad?.

Veamos las desviaciones típicas.

Muestra 1:

$$\bar{x} = \frac{64}{8} = 8$$

$$S^2 = \frac{530}{8} - 8^2 = 66.25 - 64 = 2.25$$

$$S = 1.5$$

| $x_i$ | $n_i$ | $x_i\, n_i$ | $x_i^2\, n_i$ |
|---|---|---|---|
| 6 | 2 | 12 | 72 |
| 7 | 1 | 7 | 49 |
| 8 | 2 | 16 | 128 |
| 9 | 1 | 9 | 81 |
| 10 | 2 | 20 | 200 |
| | 8 | 64 | 530 |

# Problema 10

Muestra 2:

$$\bar{x} = \frac{64}{8} = 8$$

$$S^2 = \frac{536}{8} - 8^2 = 67 - 64 = 3$$

$$S = 1.7321$$

| $x_i$ | $n_i$ | $x_i\ n_i$ | $x_i^2\ n_i$ |
|---|---|---|---|
| 6 | 3 | 18 | 108 |
| 8 | 2 | 16 | 128 |
| 10 | 3 | 30 | 300 |
| | 8 | 64 | 536 |

Muestra 3:

$$\bar{x} = \frac{65}{8} = 8.125$$

$$S^2 = \frac{537}{8} - 8.125^2 =$$

$$= 67.125 - 66.015625 = 1.109375$$

$$S = 1.05327$$

| $x_i$ | $n_i$ | $x_i\ n_i$ | $x_i^2\ n_i$ |
|---|---|---|---|
| 7 | 3 | 21 | 147 |
| 8 | 2 | 16 | 128 |
| 9 | 2 | 18 | 162 |
| 10 | 1 | 10 | 100 |
| | 8 | 65 | 537 |

Aunque las desviaciones típicas sean distintas no podemos decir quién tiene mayor variabilidad.

# Problema 10

c).- ¿Qué coeficiente deberías calcular para hacer una comparación de la variabilidad de las distintas muestras?. Según dicho coeficiente indica de menor a mayor la variabilidad de las muestras.

Veamos los coeficientes de variación de Pearson, para poder comparar la dispersión o variabilidad de las muestras. $C.V. = \frac{S}{|\bar{x}|}$

Muestra 1: $C.V. = \frac{1.5}{8} = 0.1875$

Muestra 2: $C.V. = \frac{1.7321}{8} = 0.2165125$

Muestra 3: $C.V. = \frac{1.05327}{8.125} = 0.13145$

La relación de variabilidad de las muestras de menor a mayor es:

Muestra 3 < Muestra 1 < Muestra 2

0.13145 < 0.1875 < 0.2165125