

Problema 1

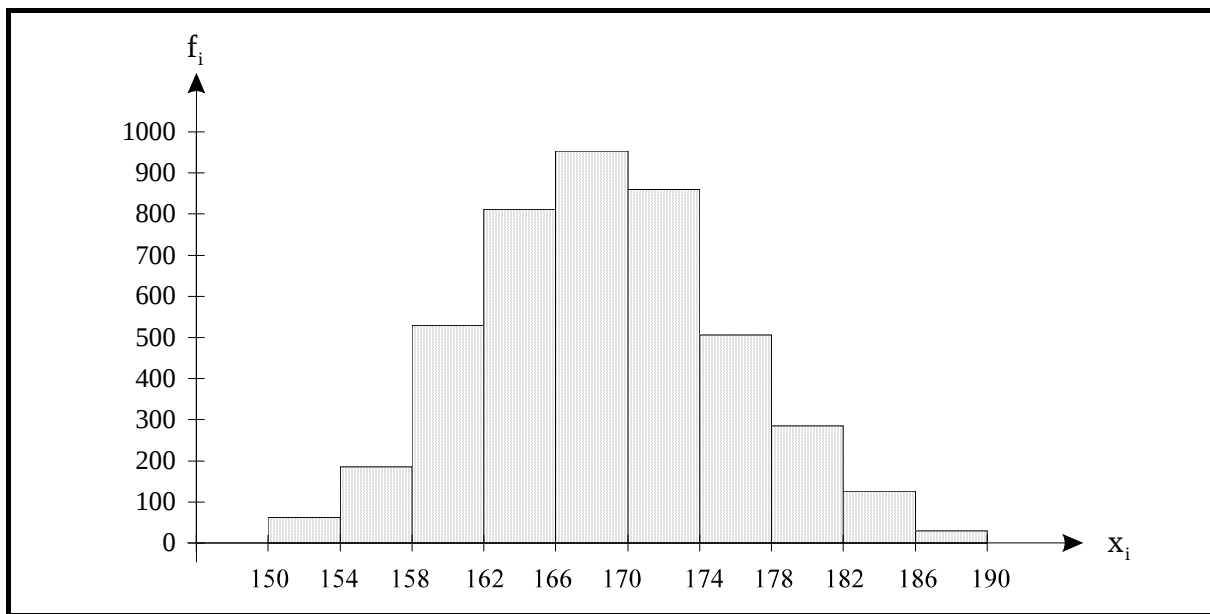
La siguiente tabla expresa la estatura en cm. de 4350 soldados:

Talla	152	156	160	164	168	172	176	180	184	188
Soldados	62	186	530	812	953	860	507	285	126	29

- a) Haz un histograma que represente la estatura en metros de los 4350 soldados. Para ello asocia a cada valor un intervalo de 4 cm. centrado en el punto que aparece en la tabla. Por ejemplo, a 152 le corresponde el intervalo $(150 - 154)$.
- b) Calcula la media, la moda y la desviación típica.
- c) ¿Cuántos soldados tienen su estatura en el intervalo $(\bar{x} - s, \bar{x} + s)$? ¿Cuántos en el intervalo $(\bar{x} - 2s, \bar{x} + 2s)$?
- d) Decimos que los soldados que tienen su estatura en $(\bar{x} + s, \bar{x} + 2s)$ son altos. ¿Cuántos soldados altos hay en total? ¿Qué porcentaje suponen?
- e) Decimos que los soldados que tienen su estatura en $(\bar{x} - 2s, \bar{x} - s)$ son bajos. ¿Qué porcentaje de soldados bajos hay en total?
- f) Calcula la mediana, los cuartiles y los centiles 10, 40, 70 y 90.

a)

Intervalos	x_i	f_i	F_i	$x_i f_i$	$x_i^2 f_i$	$F_i \cdot \frac{100}{4350}$
$[150 - 154)$	152	62	62	9.424	1.432.448	1'42
$[154 - 158)$	156	186	248	29.016	4.526.496	5'70
$[158 - 162)$	160	530	778	84.800	13.568.000	17'88
$[162 - 166)$	164	812	1.590	133.168	21.839.552	36'55
$[166 - 170)$	168	953	2.543	160.104	26.897.472	58'45
$[170 - 174)$	172	860	3.403	147.920	25.442.240	78'22
$[174 - 178)$	176	507	3.910	89.232	15.704.832	89'88
$[178 - 182)$	180	285	4.195	51.300	9.234.000	96'43
$[182 - 186)$	184	126	4.321	23.184	4.265.856	99'33
$[186 - 190)$	188	29	4.350	5.452	1.024.976	100
		4.350		733.600	123.935.872	



b)

Media: $\bar{x} = \frac{733.600}{4.350} = 168'64 \text{ cm.}$

Moda: La clase modal es $[166,170]$ porque corresponde a la que tiene mayor frecuencia, por lo tanto la moda es:

$$M_0 = 166 + 4 \cdot \frac{953 - 812}{953 - 812 + 953 - 860} = 168'41 \text{ cm}$$

Desviación típica: $s = \sqrt{\frac{123.935.872}{4350} - (168'64)^2} = 7'09$

c) En $(\bar{x} - s, \bar{x} + s)$ sabemos que hay el 68%, es decir, $0'68 \cdot 4350 = 2958$ soldados.

En $(\bar{x} - 2s, \bar{x} + 2s)$ sabemos que hay el 95%, es decir, $0'95 \cdot 4350 = 4132$ soldados.

d) Si en $(\bar{x} - s, \bar{x} + s)$ hay el 68% significa que en $(\bar{x}, \bar{x} + s)$ hay el 34%, por ser simétrica la distribución.

Si en $(\bar{x} - 2s, \bar{x} + 2s)$ hay el 95% significa que en $(\bar{x}, \bar{x} + 2s)$ hay el 47'5%, por ser simétrica la distribución.

Por tanto en $(x + s, x + 2s)$ hay el $47'5\% - 34\% = 13'5\%$, es decir, el número de soldados que hay en ese intervalo es $0'135 \cdot 4350 = 587$

e) Los mismos que en el apartado anterior.

f)

Mediana: $\frac{4350}{2} = 2175 > 1590$ luego la clase mediana es $[166, 170]$ y por tanto obtenemos para la mediana:

$$M_e = 166 + 4 \cdot \frac{2175 - 1590}{953} = 168'45 \text{ cm}$$

Cuartil inferior: Como $\frac{4350}{4} = 1087'5 > 778$ tenemos:

$$Q_1 = 162 + 4 \cdot \frac{1087'5 - 778}{812} = 163'52$$

Cuartil superior: Como $3 \cdot \frac{4350}{4} = 3262'5 > 2543$ tenemos:

$$Q_3 = 170 + 4 \cdot \frac{3262'5 - 2543}{860} = 173'34$$

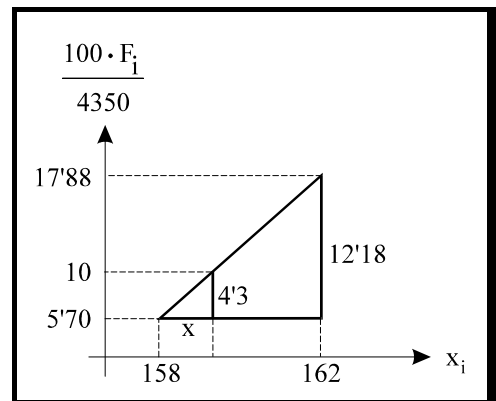
Centil 10: Como $10 \cdot \frac{4350}{100} = 435 > 248$ tenemos:

$$P_{10} = 158 + 4 \cdot \frac{435 - 248}{530} = 159'41$$

Si lo resolvemos gráficamente tenemos:

$$\frac{12'18}{4} = \frac{4'3}{x} \rightarrow x = 1'41$$

$$P_{10} = 158 + 1'41 = 159'41$$



Centil 40: Como $40 \cdot \frac{4350}{100} = 1740 > 1590$ tenemos:

$$P_{40} = 166 + 4 \cdot \frac{1740 - 1590}{953} = 166'62$$

Centil 70: Como $70 \cdot \frac{4350}{100} = 3045 > 2543$ tenemos:

$$P_{70} = 170 + 4 \cdot \frac{3045 - 2543}{860} = 172'33$$

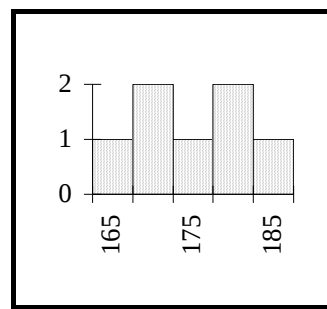
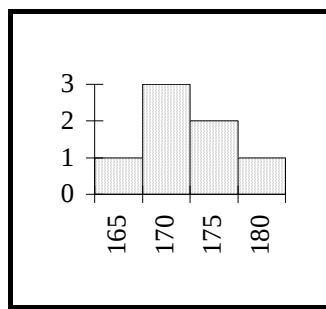
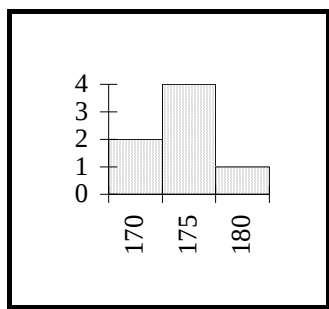
Centil 90: Como $90 \cdot \frac{4350}{100} = 3915 > 3910$ tenemos:

$$P_{90} = 178 + 4 \cdot \frac{3915 - 3910}{285} = 178'07$$

Problema 2

Las estaturas correspondientes de tres equipos de baloncesto A, B y C se distribuyen según las gráficas y con los parámetros que se dan a continuación:

	$\bar{x}$	s
A	175	6'5
B	174'3	3'2
C	172'1	4'5



¿Qué gráfica corresponde a cada equipo?

De la gráfica 1 deducimos la siguiente tabla

x_i	170	175	180
f_i	2	4	1

luego:

$$\bar{x} = \frac{170 \cdot 2 + 175 \cdot 4 + 180 \cdot 1}{2 + 4 + 1} = 174'28 \quad \text{y} \quad s = \sqrt{\frac{212700}{7} - (174'28)^2} = 3'19$$

De la gráfica 2 deducimos la siguiente tabla

x_i	165	170	175	180
f_i	1	3	2	1

luego:

$$\bar{x} = \frac{165 \cdot 1 + 170 \cdot 3 + 175 \cdot 2 + 180 \cdot 1}{7} = 172'14 \quad \text{y} \quad s = \sqrt{\frac{207575}{7} - (172'14)^2} = 4'51$$

De la gráfica 3 deducimos la siguiente tabla

x_i	165	170	175	180	185
f_i	1	2	1	2	1

$$\bar{x} = \frac{165 \cdot 1 + 170 \cdot 2 + 175 \cdot 1 + 180 \cdot 2 + 185 \cdot 1}{7} = 175 \quad \text{y} \quad s = \sqrt{\frac{214675}{7} - (175)^2} = 6'54$$

De todo lo anterior se deducen las siguientes correspondencias:

$$1 \longrightarrow B \qquad 2 \longrightarrow C \qquad 3 \longrightarrow A$$

Problema 3

Sabemos que en una clase, la calificación media de un examen ha sido 5 y la desviación típica 1'5. En esa misma clase, para otro examen, la calificación media ha sido, también, 5, y la desviación típica 1.

- a) Si un alumno ha obtenido un 8 en el primer examen y un 7'5 en el segundo, ¿qué nota te parece más meritoria? ¿Por qué?
- b) Si un alumno ha obtenido un 3'5 en el primer examen y un 4'5 en el segundo, ¿qué nota te parece más meritoria? ¿Por qué?
- a) Tenemos que comparar las notas del alumno con las de la clase, para lo cual utilizamos la puntuación típica. La puntuación típica mide la desviación respecto de la media del dato considerado, tomando como unidad la desviación típica, por lo que refleja cuán desviado se halla éste respecto de la media, independientemente de la unidad utilizada.

$$\text{Nota de 8} \longrightarrow \frac{8 - 5}{1'5} = 2$$

$$\text{Nota de 7'5} \longrightarrow \frac{7'5 - 5}{1} = 2'5$$

A mayor puntuación típica mejor puntuación en su distribución respecto de la otra distribución con quien se compara. Al ser la desviación respecto de la media mayor en el segundo caso que en el primero significa que en proporción es mejor nota el 7'5 del segundo examen que el 8 del primero aún cuando ésta última nota es más alta..

b)

$$\text{Nota de 3'5} \longrightarrow \frac{3'5 - 5}{1'5} = -1$$

$$\text{Nota de 4'5} \longrightarrow \frac{4'5 - 5}{1} = -0'5$$

Es más meritoria la nota de 4'5 puesto que es mayor $-0'5$ que -1 .

Problema 4

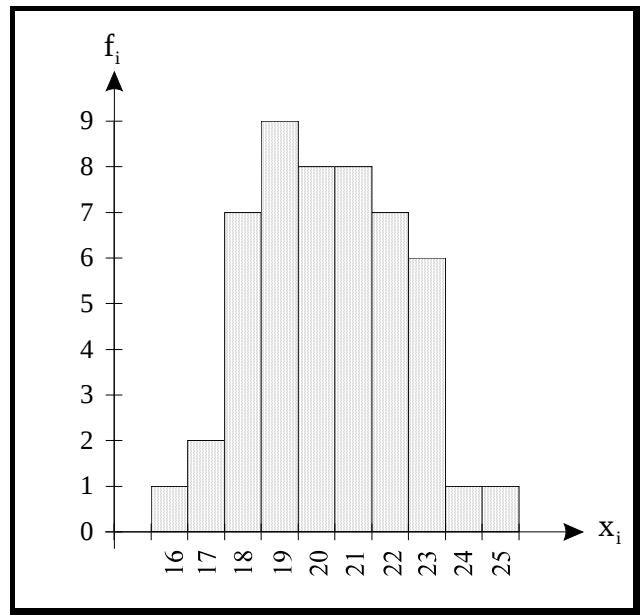
En una fábrica de tornillos se mide la longitud (en mm.) de algunos de ellos y se obtiene:

22, 20, 18, 21, 19, 17, 23, 23, 21 18, 22, 22, 19, 19, 20, 19, 23, 21, 23, 21, 19, 20, 18, 21, 19
22, 16, 19, 23, 18, 20, 22, 18, 25, 23, 21, 18, 24, 17, 20, 20, 19, 21, 20, 22, 18, 20, 22, 21, 19

- a) Haz una tabla de frecuencias, represéntala gráficamente y calcula la media y la desviación típica.
- b) Haz una nueva tabla de frecuencias agrupando los valores de dos en dos. Representala gráficamente y calcula la media y la desviación típica.
- c) Calcula las medianas, los cuartiles y el centil 90 (para ello haz previamente, el polígono de porcentajes acumulados).
- d) ¿Qué centil corresponde a la longitud 24 cm.?

a)

x_i	f_i	F_i	$x_i f_i$	$x_i^2 f_i$
16	1	1	16	256
17	2	3	34	578
18	7	10	126	2268
19	9	19	171	3249
20	8	27	160	3200
21	8	35	168	3528
22	7	42	154	3388
23	6	48	138	3174
24	1	49	24	576
25	1	50	25	625
	50		1016	20842

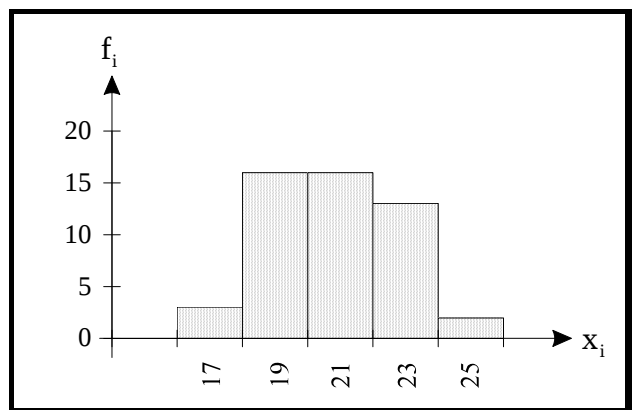


$$\bar{x} = \frac{1016}{50} = 20'32$$

$$s = \sqrt{\frac{20842}{50} - (20'32)^2} = 1'98$$

b)

Intervalos	x_i	f_i	F_i	$x_i f_i$	$x_i^2 f_i$
[16-18)	17	3	3	51	867
[18-20)	19	16	19	304	5776
[20-22)	21	16	35	336	7056
[22-24)	23	13	48	299	6877
[24-26)	25	2	50	50	1250
		50		1040	21826



$$\bar{x} = \frac{1040}{50} = 20'8$$

$$s = \sqrt{\frac{21826}{50} - (20'8)^2} = 1'96$$

c)

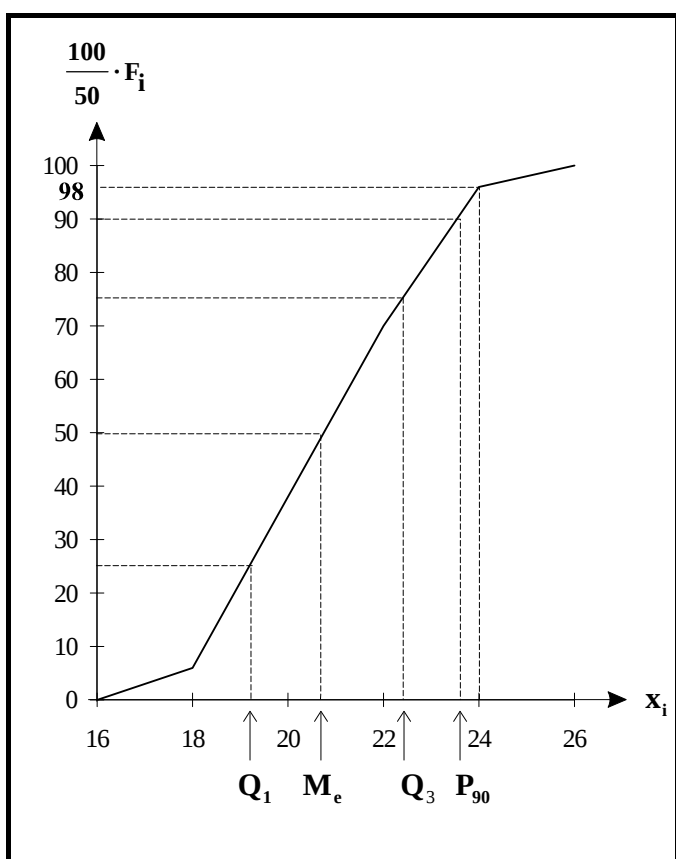
Intervalos	x_i	f_i	F_i	$x_i f_i$	$F_i \cdot \frac{100}{50}$
[16-18)	17	3	3	51	6
[18-20)	19	16	19	304	38
[20-22)	21	16	35	336	70
[22-24)	23	13	48	299	96
[24-26)	25	2	50	50	100
		50		1040	

Como $\frac{50}{2} = 25 > 19 \longrightarrow [20 - 22)$

$$M_e = 20 + 2 \cdot \frac{\frac{50}{2} - 19}{16} = 20'75$$

Como $\frac{50}{4} = 12'5 > 3 \longrightarrow [18 - 20)$

$$Q_1 = 18 + 2 \cdot \frac{\frac{50}{4} - 3}{16} = 19'18$$



Representamos el porcentaje de las frecuencias acumuladas.

Como $3 \cdot \frac{50}{4} = 37'5 > 35 \longrightarrow [22 - 24)$

$$Q_3 = 22 + 2 \cdot \frac{3 \cdot \frac{50}{4} - 35}{13} = 22'38$$

Como $90 \cdot \frac{50}{100} = 45 > 35 \longrightarrow [22, 24)$

$$P_{90} = 22 + 2 \cdot \frac{90 \cdot \frac{50}{100} - 35}{13} = 23'53$$

d) Veamos las distintas formas de resolver este apartado.

1. Si trabajamos con los datos del apartado a) tenemos:

Si aplicamos la fórmula tendremos que a 24 cm. le corresponde una frecuencia acumulada en porcentaje de $49 \cdot \frac{100}{50} = 98$. Es decir, corresponde el centil 98.

2. Si trabajamos con los datos agrupados en intervalos del apartado c) tenemos:

24 > 19 luego corresponde al intervalo [24 – 26 , por lo tanto:

$$P_x = 24 + 2 \cdot \frac{x \cdot \frac{50}{100} - 48}{2} = 24 \quad \Rightarrow \quad \frac{x}{2} - 48 = 0 \quad \Rightarrow \quad x = 96$$

Es decir, corresponde el centil 96.

3. Otra forma de resolver este apartado puede ser utilizando la semejanza de triángulos en el polígono de porcentajes acumulados, aunque en nuestro caso, por ser 24 límite de intervalos, se observa de la tabla que le corresponde el valor 96.

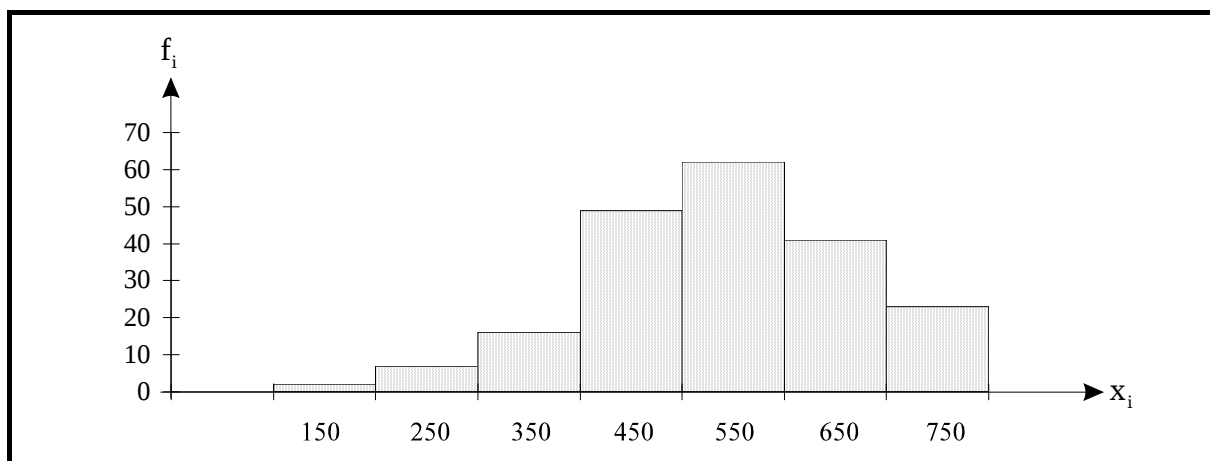
Problema 5

En una fábrica de bombillas se observaron 200 de ellas para estudiar su duración, y se obtuvieron los siguientes resultados:

Horas de duración	[100 – 200)	[200 – 300)	[300 – 400)	[400 – 500)	[500 – 600)	[600 – 700)	[700 – 800)
Nº de bombillas	2	7	16	49	62	41	23

- Representa la tabla mediante un histograma de frecuencias.
- Haz una tabla con las marcas de clase de cada intervalo y las frecuencias, y calcula con exactitud $\bar{x}$ y s .
- Obtén el porcentaje de bombillas cuya duración está en cada uno de los intervalos $(\bar{x} - s, \bar{x} + s)$ y $(\bar{x} - 2s, \bar{x} + 2s)$.
- Representa el polígono de porcentajes acumulados y obtén sobre él la mediana, los cuartiles y los centiles 10, 70 y 90. Obtenlos también analíticamente.
- ¿Qué centiles corresponden a la duración de 225 y 575 horas?
- ¿Qué porcentaje de bombillas duran más de 500 horas? ¿Y más de 675 horas? ¿Y menos de 275 horas?
- En otra fábrica se ha hecho una prueba con otro tipo de bombillas resultando una media de $\bar{x} = 552'6$ y una desviación típica de $s = 136'2$. ¿Qué bombilla será peor, una de la primera fábrica que dura 450 horas o una de esta segunda fábrica que dura 505 horas? ¿Y si las duraciones respectivas son de 500 y 520 horas respectivamente, cuál será mejor?

a)



b)

Horas de duración	Marcas de clase x_i	f_i	F_i	$x_i f_i$	$x_i^2 f_i$	$\frac{100 \cdot F_i}{200}$
[100-200)	150	2	2	300	45000	1
[200-300)	250	7	9	1750	437500	4'5
[300-400)	350	16	25	5600	1960000	12'5
[400-500)	450	49	74	22050	9922500	37
[500-600)	550	62	136	34100	18755000	68
[600-700)	650	41	177	26650	17322500	88'5
[700-800)	750	23	200	17250	12937500	100
		200		107700	61380000	

$$\bar{x} = \frac{107.700}{200} = 538'5$$

$$s = \sqrt{\frac{61.380.000}{200} - (538'5)^2} = 130'06$$

c)

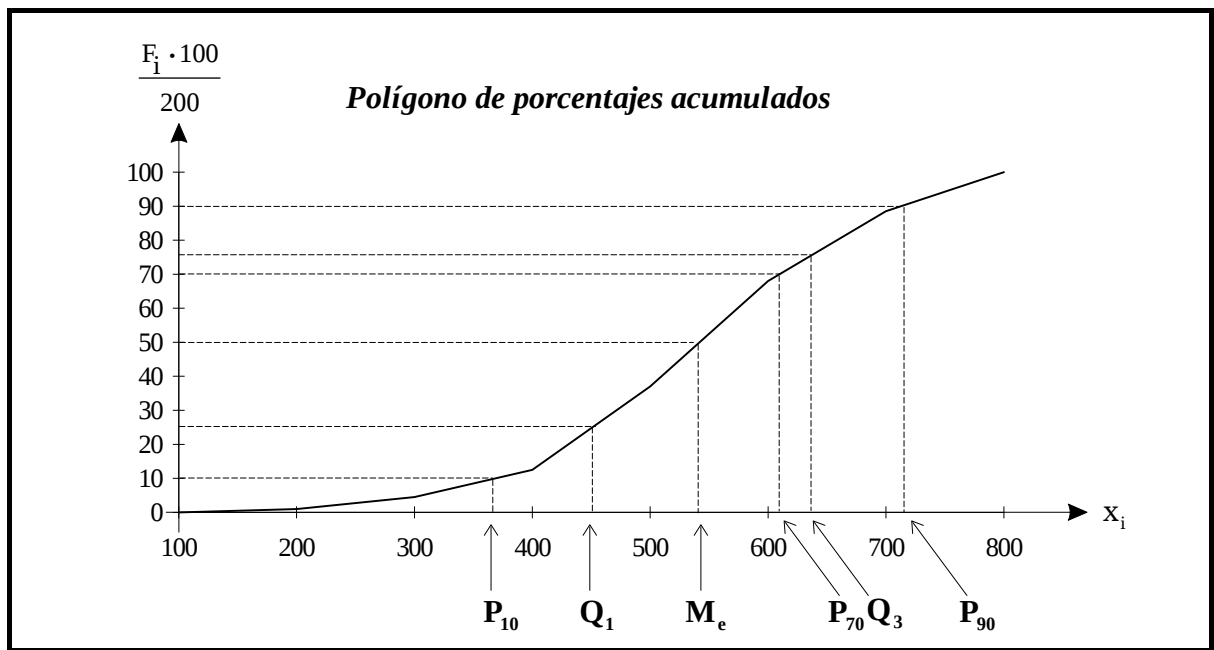
$$(\bar{x} - s, \bar{x} + s) \rightarrow (538'5 - 130'06, 538'5 + 130'06) \rightarrow (408'44, 668'56)$$

Suponiendo que se reparten por igual en los intervalos, si hacemos un recuento a ojo obtenemos aproximadamente $45 + 62 + 25 = 132$ que corresponde a $\frac{132}{200} = 0'66 = 66\%$

$$(\bar{x} - 2s, \bar{x} + 2s) \rightarrow (538'5 - 260'12, 538'5 + 260'12) \rightarrow (278'38, 798'62)$$

Haciendo el recuento en tabla obtenemos aproximadamente $1 + 16 + 49 + 62 + 41 + 22 = 191$ que corresponde a un porcentaje del $\frac{191}{200} = 0'955 = 95'5\%$

d)



Mediana $\frac{200}{2} = 100 > 74 \Rightarrow M_e = 500 + 100 \cdot \frac{\frac{200}{2} - 74}{62} = 541'93$

$Q_1 \longrightarrow \frac{200}{4} = 50 > 25 \Rightarrow Q_1 = 400 + 100 \cdot \frac{\frac{200}{4} - 25}{49} = 451'02$

$Q_3 \longrightarrow 3 \cdot \frac{200}{4} = 150 > 136 \Rightarrow Q_3 = 600 + 100 \cdot \frac{3 \cdot \frac{200}{4} - 136}{41} = 634'14$

$P_{10} \longrightarrow 10 \cdot \frac{200}{100} = 20 > 9 \Rightarrow P_{10} = 300 + 100 \cdot \frac{10 \cdot \frac{200}{100} - 9}{16} = 368'75$

$P_{70} \longrightarrow 70 \cdot \frac{200}{100} = 140 > 136 \Rightarrow P_{70} = 600 + 100 \cdot \frac{70 \cdot \frac{200}{100} - 136}{41} = 609'75$

$P_{90} \longrightarrow 90 \cdot \frac{200}{100} = 180 > 177 \Rightarrow P_{90} = 700 + 100 \cdot \frac{90 \cdot \frac{200}{100} - 177}{23} = 713'04$

e) Veamos las dos formas de calcularlos:

Para 225

Aplicando la fórmula:

$$225 = 200 + 100 \cdot \frac{x \cdot \frac{200}{100} - 2}{7} \longrightarrow 1575 = 1400 + 200x - 200 \longrightarrow x = 1'87 \approx 2$$

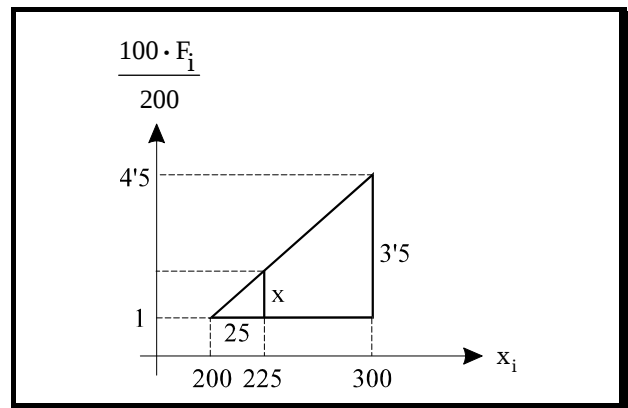
Luego corresponde el centil 2

Por semejanza de triángulos:

De la semejanza de triángulos obtenemos:

$$\frac{25}{x} = \frac{100}{3'5} \Rightarrow x = 0'875$$

por lo tanto $1 + 0'875 = 1'875 \approx 2$



Para 575

Aplicando la fórmula:

$$575 = 500 + 100 \cdot \frac{x \cdot \frac{200}{100} - 74}{62} \longrightarrow 35650 = 31000 + 200x - 7400 \longrightarrow x = 60'25 \approx 60$$

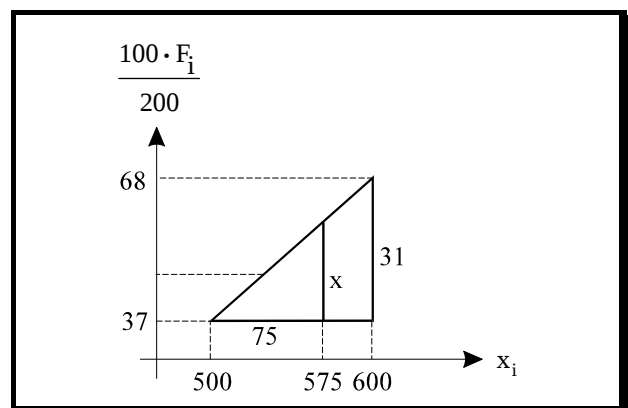
Luego corresponde al centil 60

Por semejanza de triángulos:

De la semejanza de triángulos obtenemos:

$$\frac{75}{x} = \frac{100}{31} \Rightarrow x = 23'25 \text{ por lo tanto}$$

$$37 + 23'25 = 60'25 \approx 60$$



f) **Para 500**

$$500 + 100 \cdot \frac{x \cdot \frac{200}{100} - 74}{62} = 500 \longrightarrow x = 37 \text{ luego duran más de 500 horas el } 63\%$$

Para 675

$$600 + 100 \cdot \frac{x \cdot \frac{200}{100} - 136}{41} = 675 \longrightarrow x = 83'37 \text{ luego duran más de 675 horas el } 16'62\%$$

Para 275

$$200 + 100 \cdot \frac{x \cdot \frac{200}{100} - 2}{7} = 275 \longrightarrow x = 3'62 \text{ luego duran menos de 275 horas el } 3'62\%$$

g) **Veamos cuál es peor, si la que dura 450 horas o la que dura 505 horas**

$$z_1 = \frac{450 - 538'5}{130'06} = -0'68 \quad z_2 = \frac{505 - 552'6}{136'2} = -0'34 \Rightarrow \text{es peor la primera que la}$$

segunda ya que la puntuación típica es menor.

Veamos cuál es mejor, si la que dura 500 horas o la que dura 505 horas

$$z_1 = \frac{500 - 538'5}{130'06} = -0'29 \quad z_2 = \frac{520 - 552'6}{136'2} = -0'23 \Rightarrow \text{es mejor la segunda que la}$$

primera, ya que la puntuación típica es mayor.

Problema 6

De un seguimiento realizado por un especialista en Pediatría, se obtuvo la siguiente tabla sobre los meses de edad de 50 niños de su consulta en el momento de andar por primera vez:

x_i	9	10	11	12	13	14	15
f_i	1	4	9	16	11	8	1

Obtener los siguientes parámetros:

a) Media, mediana, moda, cuartiles, centiles 33 y 67 y moda.

b) Rangos, desviación típica y coeficiente de variación de Pearson.

c) Coeficiente de asimetría y apuntamiento.

a) Construyamos la siguiente tabla que nos permite calcular varios de los parámetros pedidos:

x_i	f_i	$x_i f_i$	$x_i^2 f_i$	F_i	$(x_i - \bar{x})^3 \cdot f_i$	$(x_i - \bar{x})^4 \cdot f_i$
9	1	9	81	1	-32'768	104'8576
10	4	40	400	5	-42'592	93'7024
11	9	99	1.089	14	-15'552	18'6624
12	16	192	2.304	30	-0'128	0'0256
13	11	143	1859	41	5'632	4'5056
14	8	112	1568	49	46'656	83'9808
15	1	15	225	50	21'952	61'4656
	50	610	7526	190	-16'800	367'2000

Media $\bar{x} = \frac{610}{50} = 12'2$ meses

Mediana Corresponde a 12 meses, que es el que está en la mitad

Moda Corresponde a 12 meses, ya que es el que tiene la máxima frecuencia.

Q_1 El primer cuartil corresponde a la posición $\frac{\sum f_i}{4} = \frac{50}{4} = 12'5$. Por tanto se verifica que el cuartil primero es $Q_1 = 11$ meses, por ser éste el primer valor de la variable cuya frecuencia absoluta acumulada excede a la cuarta parte del número de datos.

Q_2 El segundo cuartil corresponde a la posición $2 \cdot \frac{\sum f_i}{4} = \frac{50}{4} = 25$. Por tanto se verifica que el cuartil segundo es $Q_2 = 12$ meses, por ser éste el primer valor de la variable cuya frecuencia absoluta acumulada excede a la mitad del número de datos.

Q_3 El tercer cuartil corresponde a la posición $3 \cdot \frac{\sum f_i}{4} = 3 \cdot \frac{50}{4} = 37'5$. Por tanto, se verifica que el cuartil tercero es $Q_3 = 13$ meses, por ser éste el primer valor de la variable cuya frecuencia absoluta acumulada excede a las tres cuartas partes del número de datos.

P₃₃ El centil 33 corresponde a la posición $33 \cdot \frac{\sum f_i}{100} = 33 \cdot \frac{50}{100} = 16'5$. Por tanto, se verifica que el centil de orden 33 es $P_{33} = 12$ meses, por ser éste el primer valor de la variable cuya frecuencia absoluta acumulada excede al 33% del total del número de datos.

P₆₇ El centil 67 corresponde a la posición $67 \cdot \frac{\sum f_i}{100} = 67 \cdot \frac{50}{100} = 33'5$. Por tanto, se verifica que el centil de orden 67 es $P_{67} = 13$ meses, por ser éste el primer valor de la variable cuya frecuencia absoluta acumulada excede al 67% del total del número de datos.

b) Rangos, desviación típica y coeficiente de variación de Pearson.

Rango $15 - 9 = 6$ meses

Rango intercuartílico $Q_3 - Q_1 = 13 - 11 = 2$

Desviación típica $s = \sqrt{\frac{7526}{50} - (12'2)^2} = 1'296$

Coeficiente de variación de Pearson $CV = \frac{s}{\bar{x}} = \frac{1'296}{12'2} = 0'11$

c) Coeficiente de asimetría y apuntamiento

Coeficiente de asimetría
$$g_1 = \frac{\frac{\sum_{i=1}^n (x_i - \bar{x})^3 \cdot f_i}{\sum_{i=1}^n f_i}}{s^3} = \frac{\frac{-16'8}{50}}{1'296^3} = -0'154$$

Coeficiente de curtosis
$$g_2 = \frac{\frac{\sum_{i=1}^n (x_i - \bar{x})^4 \cdot f_i}{\sum_{i=1}^n f_i}}{s^4} - 3 = \frac{\frac{365'2}{50}}{1'296^4} - 3 = 2'588$$