R1-hecha.pdf

ferluque

ESTADÍSTICA DESCRIPTIVA E INTROD A LA PROBABILIDAD

1º Doble Grado en Ingeniería Informática y Matemáticas

Facultad de Ciencias
Universidad de Granada

Relación 1 Definitiva

lunes, 2 de marzo de 2020 12:31

1. El número de hijos de las familias de una determinada barriada de una ciudad es una variable estadística de la que se conocen los siguientes datos:

x_i	n_i	N_i	f_i
0	80	80	0.16
1	110	190	0'22
2	130	320	0'26
3	90	410	0.18
4	40	450	0'08
5	30	480	0'06
6	20	500	0'04

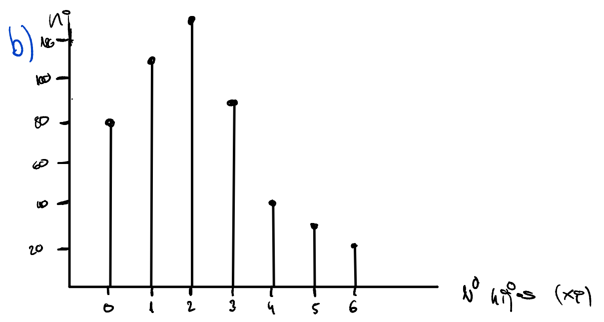
n_i : frecuencias absolutas
N_i : frecuencias absolutas acumuladas
f_i : frecuencias relativas

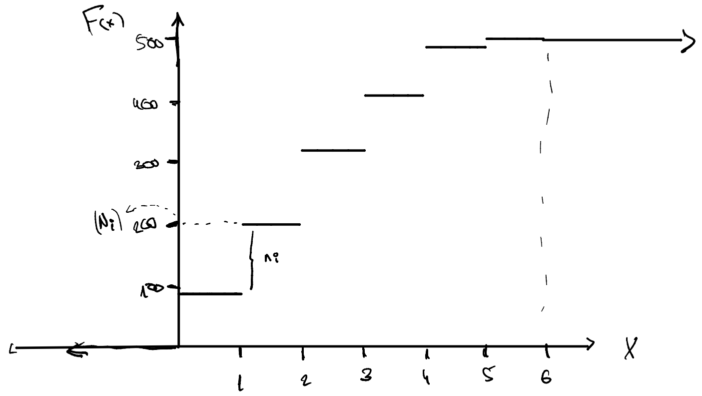
a) Completar la tabla de frecuencias.

b) Representar la distribución mediante un diagrama de barras y la curva de distribución.

c) Promediar los valores de la variable mediante diferentes medidas. Interpretarlas.

b)





c) $\bar{x} = \dfrac{\sum_{i=1}^{k} n_i x_i}{n} = 2'14$ hijos

Mo = 2 hijos

Me = 2 hijos

2. La puntuación obtenida por 50 personas que se presentaron a una prueba de selección, sumadas las puntuaciones de los distintos tests, fueron:

174, 185, 166, 176, 145, 166, 191, 175, 158, 156, 156, 187, 162, 172, 197, 181, 151,
161, 183, 172, 162, 147, 178, 176, 141, 170, 171, 158, 184, 173, 169, 162, 172, 181,
187, 177, 164, 171, 193, 183, 173, 179, 188, 179, 167, 178, 180, 168, 148, 173.

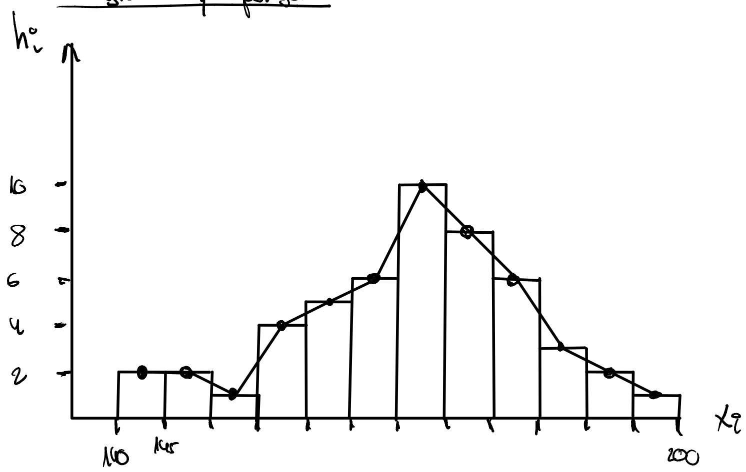
a) Agrupar los datos en intervalos de amplitud 5 desde 140 a 200 y dar la tabla de frecuencias.

b) Representar la distribución mediante un histograma, poligonal de frecuencias y curva de distribución.

a)

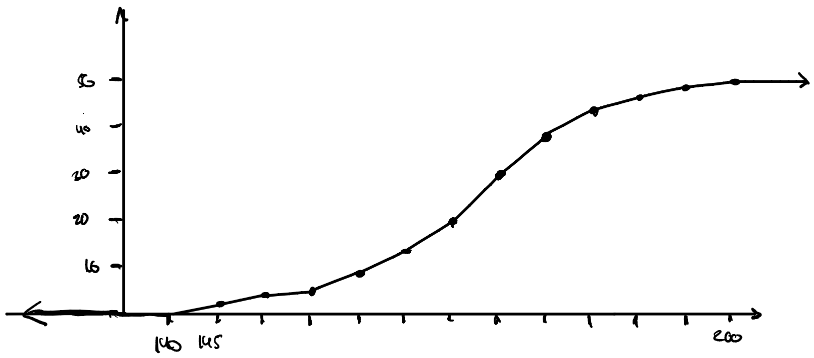
x_i	n_i	N_i	h_i
(140, 145]	2	2	2/5
(145, 150]	2	4	2/5
(150, 155]	1	5	1/5
(155, 160]	4	9	⋮
(160, 165]	5	14	
(165, 170]	6	20	
(170, 175]	10	30	
(175, 180]	8	38	
(180, 185]	6	44	
(185, 190]	3	47	
(190, 195]	2	49	
(195, 200]	1	50	

b) Histograma y poligonal




12
MARVEL STUDIOS
MS MARVEL
Serie Original
8 de junio solo en
Disney+
© 2022 MARVEL. Requiere suscripción. Se aplican términos y condiciones.

Curva de distribución



3. La distribución de la renta familiar en el año 2003 por comunidades autónomas se recoge en la siguiente tabla:

I_i	n_i	N_i	f_i	F_i	c_i	a_i	h_i
$(8300, 9300]$	2	2	2/18	2/18	8800	1000	0'002/18
$(9300, 10200]$	3	5	3/18	5/18	9750	900	0'003/18
$(10200, 11300]$	5	10	5/18	10/18	10750	1100	0'0045/18
$(11300, 12700]$	2	12	2/18	12/18	12000	1400	0'00008/18
$(12700, 13500]$	4	16	4/18	16/18	13100	800	0.005/18
$(13500, 14500]$	2	18	2/18	1	14000	1000	0.002/18

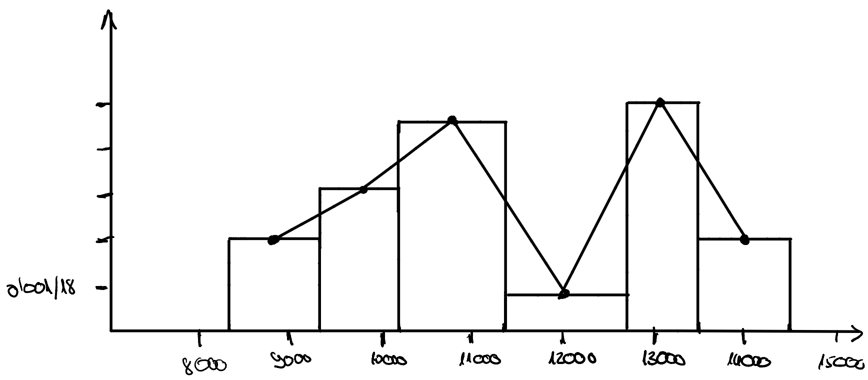
n_i : frecuencias absolutas
N_i : frec. absolutas acumuladas
f_i : frecuencias relativas
F_i : frec. relativas acumuladas
c_i : marcas de clase
a_i : amplitudes
h_i : densidades de frecuencia

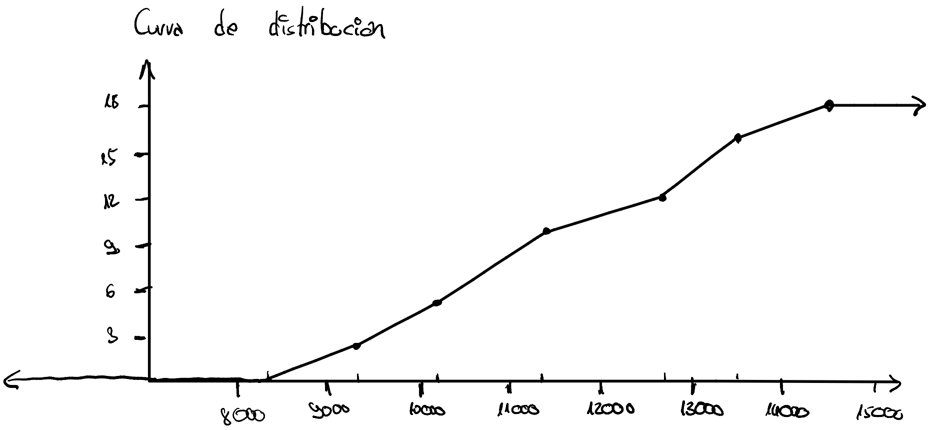
a) Completar la tabla.

b) Representar la distribución mediante un histograma, poligonal de frecuencias y curva de distribución.

c) ¿Cuántas comunidades presentan una renta menor o igual que 12700 euros? ¿Y cuántas superior a 11300 euros?

b) Histograma y poligonal





c) • $N_i \mid e_i \le 12700 \Rightarrow N_i = 12$]

• $N_i \mid e_i = 11300 \Rightarrow N_i = 10$

$n = 18 - 10 = 8$]

4. En una determinada empresa se realiza un estudio sobre la calidad de su producción. La distribución siguiente informa sobre el número de piezas defectuosas encontradas en 100 cajas examinadas con 50 unidades cada una de ellas:

Nº piezas defectuosas	0	1	2	3	4	5	6	7	8	9	10
Nº de cajas	6	9	10	11	14	16	16	9	4	3	2

a) Calcular el número medio de piezas defectuosas por caja.

b) ¿Cuantas piezas defectuosas se encuentran más frecuentemente en las cajas examinadas?

c) ¿Cuál es el número mediano de piezas defectuosas por caja?

d) Calcular los cuartiles de la distribución. Interpretarlos.

e) Calcular los deciles de orden 3 y 7. Interpretarlos.

f) Cuantificar la dispersión de la distribución utilizando diferentes medidas, interpretando los resultados y señalando las ventajas e inconvenientes de cada una.

a)

	x_i	n_i	N_i
	0	6	6
	1	9	15
	2	10	25
	3	11	36
Me	4	14	50
Mo	5	16	66
	6	16	82
	7	9	91
	8	4	95
	9	2	97
	10	3	100

$\bar{x} = \dfrac{\sum x_i \cdot n_i}{n} = 4'36$ piezas defectuosas / caja

$n = 100$ cajas

b) $Mo_1 = 5$ $\quad Mo_2 = 6$

c) $\dfrac{n}{2} = 50$ $\quad N_5 = 50 = n/2$

$Me = 4'5$

d) Los cuartiles son los valores para los que el 25% (1^{er} cuartil), 50% (2^{o} cuartil) o 75% (3^{er} cuartil) del total son inferiores o iguales a ellos.

1^{er} cuartil: $x_i \mid N_i \geq \frac{25}{100} \cdot n \Rightarrow x_3 = 2 \qquad N_3 = 25$

Como $N_3 = \frac{25}{100} \cdot n \Rightarrow P_{25} = \frac{x_3 + x_4}{2} = 3'5$

2^{o} cuartil: $x_i \mid N_i \geq \frac{50}{100} \cdot n \Rightarrow x_5 = 4 \qquad N_5 = 50$

Como $N_5 = \frac{50}{100} \cdot n \Rightarrow P_{50} = \frac{x_5 + x_6}{2} = 4'5$

3^{o} cuartil: $x_i \mid N_i \geq \frac{75}{100} \cdot n \Rightarrow x_7 = 6 \qquad N_7 = 82$

Como $N_7 > \frac{75}{100} \cdot n \Rightarrow P_{75} = x_7 = 6$

e) Decil de orden 3: $r = 30$

$x_i \mid N_i \geq \frac{30}{100} \cdot n \Rightarrow x_4 = 3 \; ; \; N_4 = 36$

Como $N_4 > \frac{36}{100} \cdot n \Rightarrow P_{30} = 3$

Decil de orden 7: $r = 70$

$x_i \mid N_i \geq \frac{70}{100} \cdot n \Rightarrow x_7 = 6 \; ; \; N_7 = 82$

Como $N_7 > \frac{70}{100} \cdot n \Rightarrow P_{70} = 6$

f) Medidas de dispersión absolutas

No pueden ser usadas para compararse entre variables pues, como vamos a observar, todas ellas llevan unidades

Recorrido:

$R = x_k - x_1 = 10$ cajas Nos indica la "amplitud" de las muestras

Recorrido intercuartílico

$R_I = P_{75} - P_{25} = Q_3 - Q_1 = 6 - 3'5 = 2'5$ cajas

Nos indica cómo de disperso está el 50% central de las muestras

Desviación absoluta media respecto a $\bar{x}$

$$D_{\bar{x}} = \frac{\sum_{i=1}^{k} |x_i - \bar{x}| n_i}{n} = 2'61$$

Nos indica cómo de representativa es la media
A mayor valor, menos representativa

Desviación absoluta media respecto a Mediana

$$D_{Me} = \frac{\sum_{i=1}^{k} |x_i - Me| n_i}{n} = 2$$

Nos indica cómo de representativa es la mediana
A mayor valor, menos representativa

$Me = 4'5$ piezas defectuosas

Varianza

$$\sigma^2 = \frac{\sum_{i=1}^{k} (x_i - \bar{x})^2 n_i}{n} = 5'87$$

Desviación típica

$$\sigma = +\sqrt{\sigma^2} = +\sqrt{5'87} = 2'4229 \text{ piezas defectuosas}$$

Medidas de dispersión relativas: Estas sí que pueden ser utilizadas para compararlas entre distintas variables, pues no tienen unidades.

Coeficiente de apertura

$$C_A = \frac{x_k}{x_1} =$$

Recorrido relativo

$$R_R = \frac{R}{\bar{x}} = \frac{10 \text{ ~~piezas defectuosas~~}}{4'36 \text{ ~~piezas defectuosas~~}} = 2'29$$

Nos da cuenta de la amplitud del intervalo que abarcan las distintas modalidades, pero sin unidades, lo que nos permite comparar con otras variables

Recorrido semi-intercuartílico

$$R_{SI} = \frac{R_I}{Q_3 + Q_1} = \frac{6 - 3'5}{6 + 3'5} = 0'26$$

Coeficiente de variación de Pearson

$$CV(x) = \frac{\sigma_x}{|\bar{x}|} = \frac{2'4229}{4'36} = 0'5557 \Big]$$

Indice de dispersion respecto a la Mediana

$$V_{Me} = \frac{D_{Me}}{Me} = \frac{2}{4'5} = 0'\hat{4} \Big]$$

5. Dadas las siguientes distribuciones:

$I_i^{(1)}$	$(0,1]$	$(1,2]$	$(2,3]$	$(3,4]$	$(4,5]$
$n_i^{(1)}$	12	13	11	8	6

$I_i^{(2)}$	$(0,1]$	$(1,3]$	$(3,6]$	$(6,10]$	$(10,12]$
$n_i^{(2)}$	1	6	7	12	2

Calcular para cada una de ellas:

a) Medias aritmética, armónica y geométrica.

b) El valor más frecuente.

c) El valor superado por el 50 % de las observaciones.

d) Recorrido, recorrido intercuartílico y desviación típica. Interpretarlos. ¿Qué distribución es más homogénea?

a) arit

$$\bar{x}^{(1)} = \frac{1}{n^{(1)}} \sum_{i=1}^{k} c_i n_i = 2'16$$

$$n^{(1)} = 56$$

$$\bar{x}^{(2)} = \frac{1}{n^{(2)}} \sum_{i=1}^{k} c_i n_i = 5'78$$

$$n^{(2)} = 28$$

geom

$$G^{(1)} = \sqrt[56]{\prod_{i=1}^{5} x_i^{n_i}} = 1'68$$

$$G^{(2)} = \sqrt[28]{\prod_{i=1}^{5} x_i^{n_i}} = 4'77$$

arm $\left\{ \begin{array}{l} H^{(1)} = \dfrac{n^{(1)}}{\sum_{k=1}^{5} \frac{n_i}{c_i}} = 1'23 \\ H^{(2)} = \dfrac{n^{(2)}}{\sum_{i=1}^{5} \frac{n_i}{c_i}} = 3'4 \end{array} \right.$

b) $M_o^{(1)} = e_{i-1} + \dfrac{h_i - h_{i-1}}{2h_i - h_{i-1} - h_{i+1}} (e_i - e_{i-1}) = 1 + \dfrac{13-12}{2 \cdot 13 - 12 - 11} (2-1) = \boxed{= 1'33}$

$I_i^{(1)}$	$h_i^{(1)}$
$(0,1]$	12
$(1,2]$	13 ←
$(2,3]$	11
$(3,4]$	8
$(4,5]$	6

$i = 2$

$M_{o_1}^{(2)} = e_{i-1} + \dfrac{h_i - h_{i-1}}{2h_i - h_{i-1} - h_{i-2}} (e_i - e_{i-1}) = 1 + \dfrac{3 - 1}{2 \cdot 3 - 1 - 7/3} (3-1) = \boxed{= \dfrac{5}{2} = 2'5}$

$I_i^{(2)}$	$h_i^{(2)}$
$(0,1]$	1
$(1,3]$	3
$(3,6]$	7/3
$(6,10]$	3
$(10,12]$	1

$i = 2$

$i = 4$

$M_{o_2}^{(2)} = e_3 \dfrac{h_4 - h_3}{2h_4 - h_3 - h_5} (e_4 - e_3) = \boxed{7}$

d) Recorrido

$R^{(1)} = e_k - e_0 = 5 - 0 = 5$

$R^{(2)} = e_k - e_0 = 12 - 0 = 12$

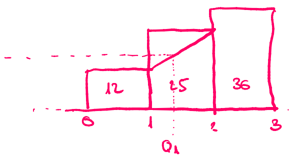
Si ambas son la misma variable estadística, la población 2 pres con mayor amplitud de modalidades

Recorrido intercuartílico

Q_1: Buscamos $I_i^{(1)} = (e_{i-1}, e_i] \mid N_{i-1} < \underset{12'5}{0'25\,n} \leq N_i$

$I_i^{(1)}$	$N_i^{(1)}$
(0, 1]	12
(1, 2]	25
(2, 3]	36
(3, 4]	44
(4, 5]	50

$I_2 = (1, 2] \Rightarrow N_2 > 12'5$

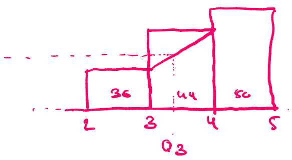


$$\frac{Q_1^{(1)} - 1}{12'5 - 12} = \frac{2-1}{25-12} \; ; \quad Q_1^{(1)} = 1'0385$$

Q_1: Buscamos $I_i^{(1)} = (e_{i-1}, e_i] \mid N_{i-1} < \underset{37'5}{0'75\,n} \leq N_i$

$I_i^{(1)}$	$N_i^{(1)}$
(0, 1]	12
(1, 2]	25
(2, 3]	36
(3, 4]	44
(4, 5]	50

$I_4 = (3, 4] \Rightarrow N_4 > 37'5$



$$\frac{Q_3^{(1)} - 3}{37'5 - 36} = \frac{4-3}{44-36} \quad Q_3^{(1)} = 3'1875$$

$$R_I^{(1)} = Q_3 - Q_1 = 3'1875 - 1'0385 = \boxed{2'149}$$

Q_1: Buscamos $I_i^{(2)} = (e_{i-1}, e_i] \mid N_{i-1} < \underset{7}{0'25\,n} \leq N_i$

$I_i^{(2)}$	$N_i^{(2)}$
(0, 1]	1
(1, 3]	7
(3, 6]	14
(6, 10]	26
(10, 12]	28

$I_2 = (1, 3] \quad N_2 = 7 = 0'25n$

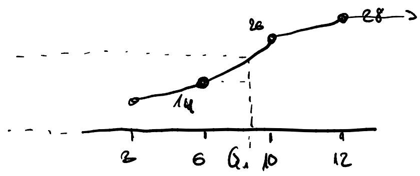
$\Downarrow$

$Q_1^{(2)} = 3$

Q_3: Buscamos $I_i^{(2)} = (e_{i-1}, e_i] \mid N_{i-1} < \underset{\Downarrow}{\overset{21}{\overset{\parallel}{0'75\, n}}} \leq N_i$

$I_i^{(2)}$	$N_i^{(2)}$
(0, 1]	1
(1, 3]	7
(3, 6]	14
(6, 10]	26
(10, 12]	28

$I_4 = (6, 10] \qquad N_4 > 21 = 0'25n$



$$\frac{Q_1^{(2)} - 6}{21 - 14} = \frac{10 - 6}{26 - 14} \qquad Q_3^{(2)} = 8'\hat{3}$$

$$R_I^{(2)} = Q_3^{(2)} - Q_1^{(2)} = 8'\hat{3} - 3 = 5'\hat{3}$$

Si ambas variables estadísticas son las mismas, podemos afirmar que en la segunda distribución, el 50% central de la población se encuentra más disperso que en la primera

Desviación típica

$$\sigma^{(1)} = \sqrt{\frac{\sum_{i=1}^{k} n_i (x_i - \bar{x})^2}{n}} = \sqrt{\frac{12(0'5-2'16)^2 + 13(1'5-2'16)^2 + 11(2'5-2'16) + 8(3'5-2'16)^2 + 6(4'5-2'16)^2}{50}} = 1'32$$

$$\sigma^{(2)} = \sqrt{\frac{\sum_{i=1}^{k} n_i (x_i - \bar{x})^2}{n}} = \sqrt{\frac{1(0'5-5'79)^2 + 6(2-5'79)^2 + 7(4'5-5'79)^2 + 12(8-5'79)^2 + 2(11-5'79)^2}{28}} = 2'9199$$

6. Un móvil efectúa un recorrido de 100 km en dos sentidos. En uno va a una velocidad constante de V_1=60 km/h y en el otro va a una velocidad constante de V_2=70 km/h. Calcular la velocidad media del recorrido.

Como la velocidad es un cociente entre dos magnitudes: $v = \frac{\Delta e}{\Delta t}$, la media apropiada será la media armónica:

$$H = \frac{n}{\sum_{i=1}^{2} \frac{n_i}{x_i}} = \frac{2}{\frac{1}{60} + \frac{1}{70}} = \boxed{64'62 \ \frac{km}{h}}$$

7. Las acciones de una empresa han producido los siguientes rendimientos netos anuales:

Año	Rentabilidad
1994	12 %
1995	10 %
1996	7 %
1997	6 %
1998	5 %

Obtener el rendimiento neto medio en esos cinco años.

Como el rendimiento tiene efectos acumulativos, la media apropiada será la geométrica:

$G = \sqrt[n]{\prod_{i=1}^{k} x_i^{n_i}}$ $\left\{ \begin{array}{l} n = 5 \\ k = 5 \\ n_i = 1 \end{array} \right.$ $G = \sqrt[5]{12 \cdot 10 \cdot 7 \cdot 6 \cdot 5} = 7'59\%$

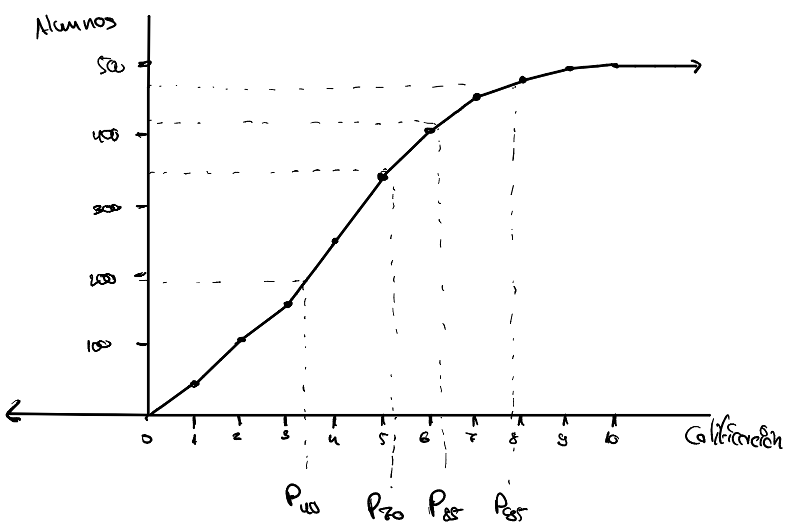
8. Un profesor califica a sus alumnos según el criterio siguiente: 40 % de suspensos, 30 % de aprobados, 15 % notables, 10 % sobresalientes y 5 % de matrículas. Las notas obtenidas son las siguientes:

(0, 1]	(1, 2]	(2, 3]	(3, 4]	(4, 5]	(5, 6]	(6, 7]	(7, 8]	(8, 9]	(9, 10]
34	74	56	81	94	70	41	28	16	4

Calcular las notas máximas para obtener cada una de las calificaciones.

Las notas máximas vendrán dadas por los percentiles de orden indicado para cada calificación:

I_i	n_i	N_i
(0, 1]	34	34
(1, 2]	74	108
(2, 3]	56	164
(3, 4]	81	245
(4, 5]	94	339
(5, 6]	70	409
(6, 7]	41	450
(7, 8]	28	478
(8, 9]	16	494
(9, 10]	4	498



$n = 498$ alumnos.

Suspensos: P_{40} : $199'2$

$I_i = (e_{i-1}, e_i] \mid N_{i-1} < \frac{40}{100} n \leq N_i$

$I_4 = (3, 4] \Rightarrow$ Como $N_i > \frac{40}{100} n$

$P_r = 3 + \frac{199'2 - 164}{81} (4-3) = 3'43$

Aquel que saque menos de $3'43 \Rightarrow$ Suspenso

Aprobados: P_{70} : $348'6$

$I_i = (e_{i-1}, e_i] \; ; \; N_{i-1} < \frac{70}{100} n \leq N_i$

$I_6 : (5, 6] \Rightarrow N_i > \frac{70}{100} n$

$P_r = 5 + \frac{348'6 - 339}{70} (6-5) = 5'13$ De $3'43$ a $5'13 \Rightarrow$ aprobado

Notable: P_{85} : $423'3$

$I_i = (e_{i-1}, e_i] : N_{i-1} < \frac{85}{100} \cdot n \leq N_i$

$\Downarrow$

$I_7 = (6, 7] \; ; \; N_i = 450 > 423'3$

$P_r = 6 + \frac{423'3 - 409}{41} (7-6) = 6'35$ De $5'13$ a $6'35 \Rightarrow$ Notable

Sobresaliente: P_{95} : $473'1$

$I_i = (e_{i-1}, e_i] \; ; \; N_{i-1} < \frac{95}{100} n \leq N_i$

$\Downarrow$

$I_8 = (7, 8] \Rightarrow$ Como $N_i = 478 > 473'1$

$P_{95} = 7 + \frac{473'1 - 450}{28} (8-7) = 7'827$ De $6'35$ a $7'827 \Rightarrow$ Sobresaliente

Matrícula: De $7'827$ a $10 \Rightarrow$ Matrícula

9. Se ha medido la altura de 110 jóvenes, obteniendo:

Altura	$(1.55, 1.60]$	$(1.60, 1.70]$	$(1.70, 1.80]$	$(1.80, 1.90]$	$(1.90, 2.00]$
Nº jóvenes	18	31	24	20	17

a) Si se consideran bajos el 3 % de los individuos de menor altura, ¿cuál es la altura máxima que pueden alcanzar?

b) Si se consideran altos el 18 % de los individuos de mayor altura, ¿cuál es su altura mínima?

c) ¿Qué altura es superada sólo por 1/4 de los jóvenes?

d) Calcular el número de jóvenes cuya altura es superior a 1.75.

e) Calcular la altura máxima de los 11 jóvenes más bajos.

f) Calcular la altura mínima de los 11 jóvenes más altos.

a) P_3 :

$I_i = (e_{i-1}, e_i] \mid N_{i-1} < \frac{3}{100} \cdot n \le N_i$

$n =$ 110 jóvenes

$\frac{3}{100} \cdot n = 3'3 \Downarrow$

$I_1 : (1'55, 1'60]$

$P_3 = 1'55 + \frac{3'3 - 1'55}{18}(1'60 - 1'55) = \boxed{1'55 \text{ m}}$

b)

I_i	N_i
$(1'55, 1'60]$	18
$(1'60, 1'70]$	49
$(1'70, 1'80]$	73
$(1'80, 1'90]$	93
$(1'90, 2'00]$	110

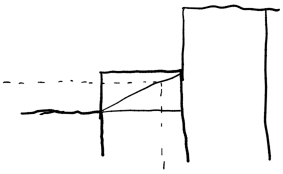
P_{82} :

$I_i = (e_{i-1}, e_i] \mid N_{i-1} < \frac{82}{100} n \le N_i$

$\frac{82}{100} n = 90'2$

$I_4 = (1'80, 1'90]$

$N_4 > 90'2$



P_{82}

$\frac{P_{82} - 1'80}{0'10} = \frac{90'2 - 73}{20}$

$P_{82} = \boxed{1'88 \text{ m}}$

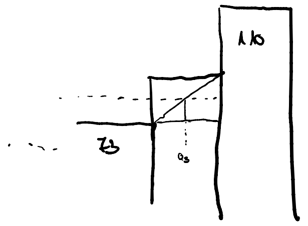
c) Q_3 :

$I_i = (e_{i-1}, e_i] \mid N_{i-1} < \frac{3}{4}n \le N_i$

$\frac{3}{4}n = 82'5$

I_i	N_i
(1'55, 1'60]	18
(1'60, 1'70]	49
(1'70, 1'80]	73
(1'80, 1'90]	93
(1'90, 2'00]	110

$I_4 = (1'80, 1'90]$

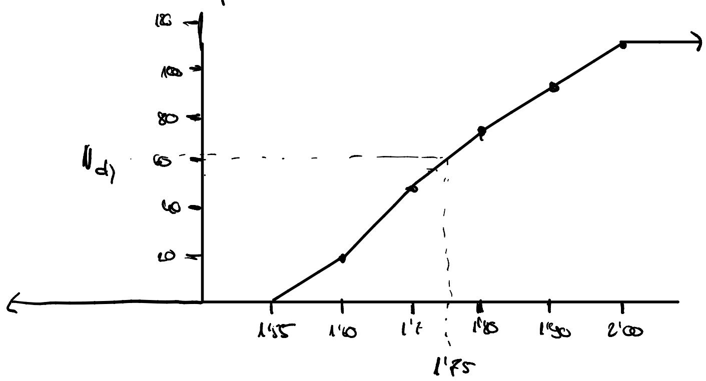


$$\frac{Q_3 - 1'80}{0'10} = \frac{82'5 - 73}{20}$$

$Q_3 = 1'85$ m

d) Nº de personas cuya altura es superior a 1'75 m

I_i	N_i
(1'55, 1'60]	18
(1'60, 1'70]	49
(1'70, 1'80]	73
(1'80, 1'90]	93
(1'90, 2'00]	110



La cifra que buscamos es $n - N_{d)}$. Calculamos $N_{d)}$ por interpolación

$$\frac{1'75 - 1'70}{N_{d)} - 49} = \frac{1'80 - 1'70}{73 - 49}$$

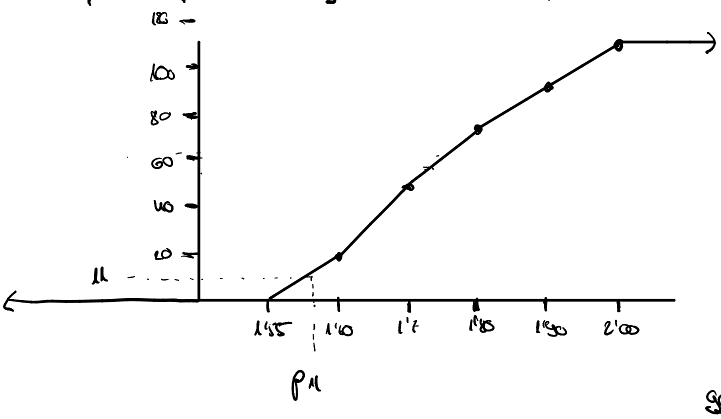
; $N_{d)} = 61$ personas miden menos de 1'75m

$0'1 N_{d)}$

Por tanto 49 personas miden más de 1'75 m

e) Encontrar $I_i = (e_{i-1}, e_i] \mid N_{i-1} < 11 \leq N_i$ Llamemos p_{11} a la cantidad buscada (Es el percentil 11)

$I_1 = (1'55, 1'60] \Rightarrow N_i = 18 > 11$

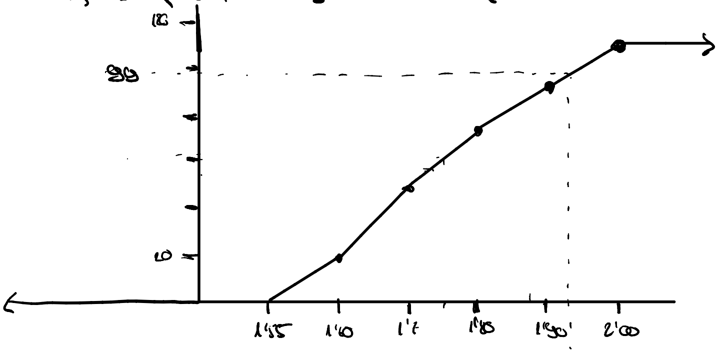


$$\frac{p_{11} - 1'55}{11 - 0} = \frac{1'60 - 1'55}{20 - 0}$$

$$\boxed{p_{11} = 1'58 \text{ m}}$$

f) Encontrar $I_i = (e_{i-1}, e_i] \mid N_{i-1} < \underbrace{110 - 11}_{99 \text{ personas}} \leq N_i$ Llamemos P_{99} a la cantidad buscada (Es el percentil 99)

$I_5 = (1'90, 2'00] \Rightarrow N_5 = 110 > 99$



$$\frac{P_{99} - 1'90}{99 - 93} = \frac{0'1}{110 - 93}$$

$$\boxed{P_{99} = 1'8353 \text{ m}}$$

10. Realizando una prueba para el estudio del cáncer a 150 personas se obtuvo la siguiente tabla según la edad de los enfermos:

Edad	(10, 30]	(30, 40]	(40, 50]	(50, 60]	(60, 90]
Nº enfermos	15	22	48	40	25

a) Calcular la edad más común de los individuos estudiados.

b) Calcular la edad mínima y máxima del 30 % central de los individuos.

c) Calcular el recorrido intercuartílico y la desviación típica.

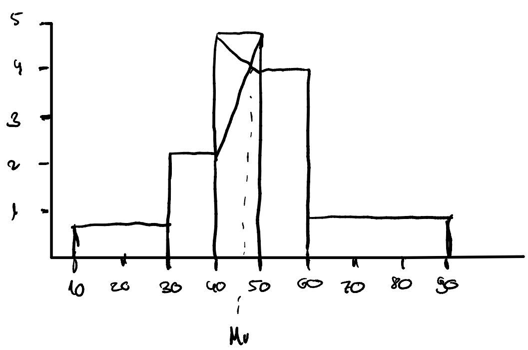
d) Calcular e interpretar los valores de los coeficientes de asimetría y curtosis.

a)

Edad	n_i	h_i
(10, 30]	15	0'75
(30, 40]	22	2'2
(40, 50]	48	4'8
(50, 60]	40	4
(60, 90]	25	$0'8\hat{3}$

M_o : $I_i = (e_{i-1}, e_i]$ tq

$h_i \geq h_j \quad \forall j \in \{1, \dots, k\}$



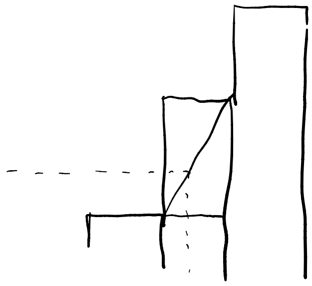
$$\frac{4'8-2'2}{4'8-4} = \frac{Mo-40}{50-Mo}$$

$\boxed{Mo = 47'647 \text{ años}}$

b) P_{35} : $\frac{35}{100} \cdot n = 52'5$ años

I_i	N_i
(10, 30]	15
(30, 40]	37
(40, 50]	85
(50, 60]	125
(60, 90]	150

$I_3 = (40, 50]$



$$\frac{P_{35}-40}{52'5-37} = \frac{10}{48}$$

$\boxed{P_{35} = 43'23 \text{ años}}$

P_{65} : $\frac{65}{100} n = 97'5$

$I_4 = (50, 60]$ $\qquad \frac{P_{65}-50}{97'5-85} = \frac{10}{40}$ $\qquad \boxed{P_{65} = 53'125 \text{ años}}$

30% de los profesores: (43'23 , 53'125)

c) $\underline{Q_1}$ $\quad 0'25\,n = 37'5$

$I_3 = (40, 50]$ $\qquad \frac{Q_1-40}{37'5-37} = \frac{10}{48}$ $\qquad Q_1 = 40'1$ años

$\underline{Q_3}$ $\quad 0'75\,n = 112'5$

$I_4 = (50, 60]$ $\qquad \frac{Q_3-50}{112'5-85} = \frac{10}{40}$ $\qquad Q_3 = 56'875$ años

$\boxed{R_I = 16'775 \text{ años}}$

σ

$$\bar{x} = \frac{\sum_{i=1}^{k} c_i n_i}{n} = 48'7 \text{ años}$$

$$\sigma^2 = \frac{\sum_{i=1}^{k} n_i (c_i - \bar{x})^2}{n} = \frac{15(20-48'7)^2 + 22(35-48'7)^2 + 48(45-48'7)^2 + 40(55-48'7)^2 + 25(75-48'7)^2}{150} = \boxed{240'14}$$

$$\sigma = +\sqrt{\sigma^2} = \boxed{15'5 \text{ años}}$$

d) Coeficiente de asimetría de Fisher

$$\bar{x} = \frac{\sum_{i=1}^{k} n_i c_i}{n} = 48'7 \text{ años}$$

$$\mu_3 = \frac{\sum_{i=1}^{k} n_i (c_i - \bar{x})^3}{n} = \frac{15\cdot(20-48'7)^3 + 22(35-48'7)^3 + 48(45-48'7)^3 + 40(55-48'7)^3 + 25(75-48'7)^3}{150} = 341'256$$

$$\sigma_x^3 = \left(\sqrt{\frac{\sum_{i=1}^{k} n_i (c_i - \bar{x})^2}{n}}\right)^3 = \left(\sqrt{\frac{15(20-48'7)^2 + 22(35-48'7)^2 + 48(45-48'7)^2 + 40(55-48'7)^2 + 25(75-48'7)^2}{150}}\right)^3 = 3721'3953$$

$$\frac{\mu_3}{\sigma_x^3} = \frac{341'256}{3721'3953} = \boxed{0'0917} \Rightarrow \text{asimétrica por la derecha}$$

Coeficiente de asimetría de Pearson

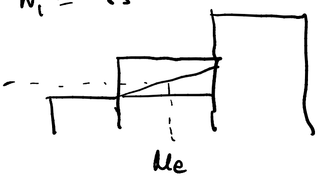
$$A_p = \frac{\bar{x} - M_o}{\sigma_x} = \frac{48'7 - 45'3}{15'5} = \boxed{0'22} \text{ Asimétrica por la derecha}$$

$$A_p^* = \frac{3(\bar{x} - M_e)}{\sigma_x} = \frac{3(48'7 - 47'8167)}{15'4966} = \boxed{0'1516} \Rightarrow \text{asimétrica por la derecha.}$$

Me: $I_i = (e_{i-1}, e_i]$ tq $N_{i-1} < 0'5n \leq N_i$

$0'5n = 75$ personas

$I_3 = [40, 50]$ $N_i = 85$



$$\frac{Me - 40}{75 - 37} = \frac{10}{48}$$

$Me = 47'9167$ años

Coeficiente de curtosis de Fisher

$$\gamma_2(X) = \frac{\mu_4}{\sigma^4} - 3 = \left(\frac{1}{n \cdot \sigma^4} \sum_{i=1}^{k} n_i (c_i - \bar{x})^4 \right) - 3 =$$

$$= \frac{1}{150 \cdot 57668'821} \cdot \left(15(20-48'7)^4 + 22(35-48'7)^4 + 48(45-48'7)^4 + \right.$$

$$\left. + 40(55-48'7)^4 + 25(75-48'7)^4 \right) - 3 = -0'3429 \Rightarrow \text{platicúrtica}$$

Coeficiente de curtosis de Kelley

$$K = \frac{1}{2} \frac{Q_3 - Q_1}{D_9 - D_1} - 0'263 = \frac{1}{2} \frac{56'875 - 40'1}{72 - 30} - 0'263 = -0'0633$$

$\Downarrow$ platicúrtica

$Q_1 = 40'1$ años
$Q_3 = 56'875$ años

D_1: $N_{i-1} < 0'1 \cdot 150 \leq N_i$

$0'1 \cdot 150 = 15$ enfermos

I_i	N_i
(10, 30]	15
(30, 40]	37
(40, 50]	85
(50, 60]	125
(60, 90]	150

$I_1 = (10, 30] \Rightarrow N_i = 15$ enfermos

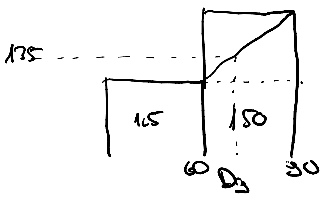
$D_1 = 30$ años

D_9: $N_{i-1} < 0'9 \cdot 150 \leq N_i$

$0'9 \cdot 150 = 135$ enfermos

I_i	N_i
(10, 30]	15
(30, 40]	37
(40, 50]	85
(50, 60]	125
(60, 90]	150

$I_5 = (60, 90]$ $N_5 = 150 > 135$



$$\frac{D_9 - 60}{135 - 125} = \frac{30}{25}$$

$D_9 = 72$ años