Sèrie 3

1)

a) (4 punts). *Ordenació de les dades: 1,5 punt; identificació dels quartils: 1 punt (0,5 punts cadascun); interpretació: 1,5 punts (0,75 punts cadascun).*

La mostra ordenada és: 353, 366, 372, 375, 394, 408, 425, 438, 470, 471, 489, 524.

El primer quartil és la mitjana del valor que és a la posició n/4=12/4=3 i el valor posterior: Q_1=(372+375)/2=373,5. El 25% dels anys el nombre de trasplantaments és inferior a 373,5 i el 75% dels anys és superior.

El tercer quartil és la mitjana del valor que és a la posició 3n/4=36/4=9 i el valor posterior: Q_2=(470+471)/2=470,5. El 75% dels anys el nombre de trasplantaments és inferior a 470,5 i el 25% dels anys és superior.

b) (1 punt).

El percentatge d'anys que el nombre de trasplantaments ha estat superior a 400 és igual a (7/12)·100=58,3%.

c) (2 punts). *Mitjana aritmètica: 0,75 punts; variància: 1 punt; desviació estàndard: 0,25 punts.*

La mitjana del nombre de trasplantaments és igual a

$$\bar{x} = \frac{\sum_i x_i}{n} = \frac{5085}{12} = 423{,}75 \ .$$

La variància és igual a

$$S_x^2 = \frac{\sum_i x_i^2}{n} - \bar{x}^2 = \frac{2188181}{12} - (423{,}75)^2 = 182348{,}41\widehat{6} - 179564{,}0625$$
$$= 2784{,}3541\widehat{6}.$$

La desviació estàndard del nombre de trasplantaments és igual a

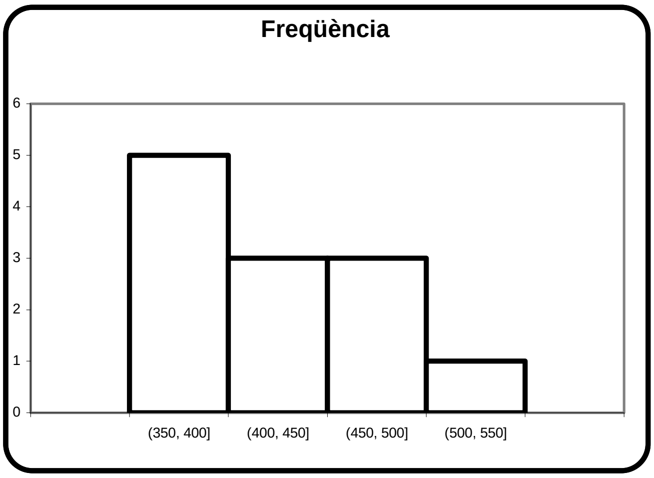
$$S_x = \sqrt{2784{,}3541\widehat{6}} = 52{,}7670 \ .$$

d) (3 punts). *Agrupació en intervals i recompte de freqüències: 2 punts; representació de l'histograma: 1 punt.*

La taula de freqüències amb els intervals és la següent:

350-400	5
400-450	3
450-500	3
500-550	1

Per tant, l'histograma és



2)

a) (1 punt).
La freqüència absoluta s'obté de la relació $n_i=n\cdot f_i$ = 4500000$\cdot f_i$. Per tant,

		Audiències de ràdio	
		Fins a 2,5 hores diàries	Més de 2,5 hores diàries
Edat	15-30	720000	270000
	30-45	945000	495000
	45-65	765000	495000
	65-100	450000	360000

b) (2 punts). *Mitjana aritmètica: 0,75 punts; variància: 1 punt; desviació estàndard: 0,25 punts.*

Edat	n_i
15-30	270000
30-45	495000
45-65	495000
65-100	360000

Mitjana aritmètica de l'edat: $\overline{x} = \frac{\sum_{i=1}^{n} c_i n_i}{n} = \frac{81562500}{1620000} = 50{,}35$

Variància de l'edat:

$$S_X^2 = \frac{\sum_i c_i^2 n_i}{n} - \overline{x}^2 = \frac{4780406250}{1620000} - (50{,}35)^2 = 2950{,}87 - 2534{,}84 = 416{,}03$$

Desviació estàndard de l'edat: $S_x = \sqrt{416{,}03} = 20{,}40$

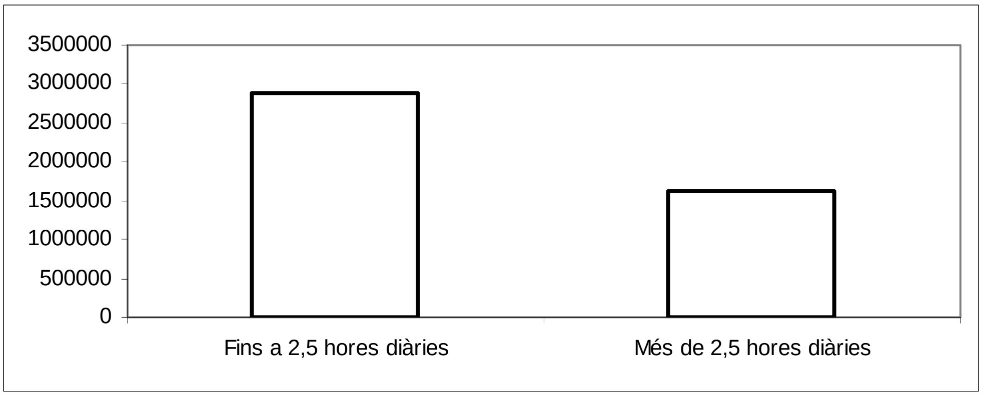
c) (1 punt).

El percentatge de menors de 45 anys entre els oïdors que escolten la ràdio menys de 2,5 hores diàries és igual a

100·(720000+945000)/2880000=57,8%.

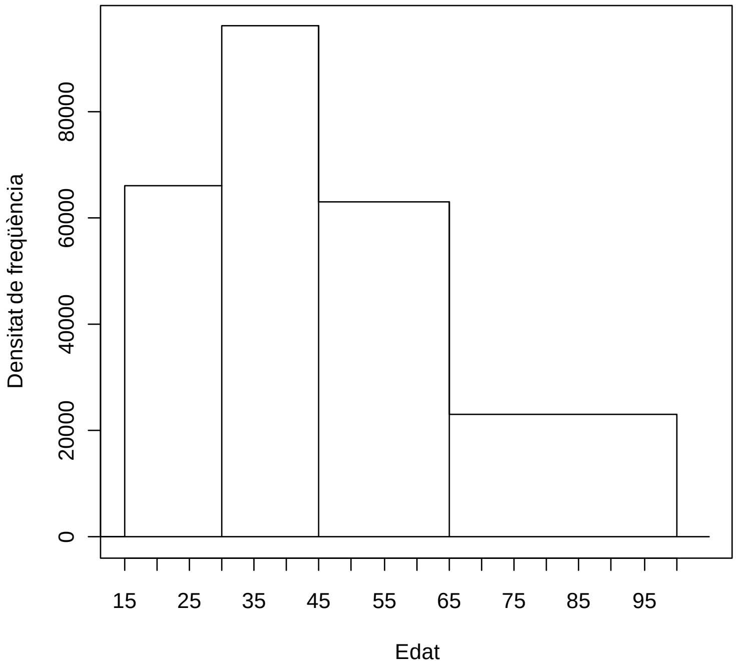
d) (1 punt) *Obtenció de la distribució marginal: 0,5 punts; gràfic de barres: 0,5 punts.*

Audiències de ràdio	n_i
Fins a 2,5 hores diàries	2880000
Més de 2,5 hores diàries	1620000



e) (2 punts). *Obtenció de la distribució marginal de l'edat: 0,5 punts; càlcul de la densitat de freqüència: 0,5 punts; representació de l'histograma: 1 punt.*

Edat	n_i	a_i	h_i
15-30	990000	15	66000
30-45	1440000	15	96000
45-65	1260000	20	63000
65-100	810000	35	23142,9



f) (3 punts). *Interpretació de l'enunciat: 1 punt; identificació de l'interval que conté el tercer quartil: 1 punt; càlcul del tercer quartil: 1 punt.*

S'ha de calcular el tercer quartil ja que per sobre d'aquest valor hi ha un 25% dels oïdors i són els que tenen més edat.

El tercer quartil és a l'interval 45-65 ja que té la primera freqüència absoluta acumulada que supera el valor 3n/4=3375000.

Una aproximació del tercer quartil s'obté amb la següent fórmula (L_{i-1} és el límit inferior de l'interval que conté el quartil, n_i és la seva freqüència absoluta, N_{i-1} és la freqüència absoluta acumulada de l'interval anterior, a_i és l'amplitud),

$$Q_3 = L_{i-1} + \frac{\frac{3n}{4} - N_{i-1}}{n_i} \cdot a_i = 45 + \frac{3375000 - 2430000}{1260000} \cdot 20 = 60 \,.$$

3)

a) (5 punts). *Mitjana aritmètica de les dues variables: 1,5 punts; variàncies: 2 punts; desviacions estàndards: 0,5 punts; coeficients de variació i interpretació: 1 punt.*

Les mitjanes de la renda i preu:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{488}{10} = 48{,}8 \qquad \bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{51}{10} = 5{,}1 .$$

La variàncies:

$$S_x^2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2 = \frac{26670}{10} - (48{,}8)^2 = 2667 - 2381{,}44 = 285{,}56$$

i

$$S_y^2 = \frac{\sum_{i=1}^{n} y_i^2}{n} - \bar{y}^2 = \frac{275{,}92}{10} - (5{,}1)^2 = 27{,}592 - 26{,}01 = 1{,}582 .$$

Les desviacions estàndard: $S_x = \sqrt{285{,}56} = 16{,}90$ i $S_y = \sqrt{1{,}582} = 1{,}26$.

Els coeficients de variació són igual a $V_x = 16{,}90/48{,}1 = 0{,}35$ i $V_y = 1.26/5{,}1 = 0{,}25$. Per tant, la mitjana del preu és una mica més representativa.

b) (3 punts). *Covariància: 1,25 punts; coeficient de correlació: 0,75 punts; interpretació: 1 punt.*

La covariància entre les dues variables:

$$S_{xy} = \frac{\sum_{i=1}^{n} x_i y_i}{n} - \bar{x}\bar{y} = \frac{2682}{10} - 48{,}8 \cdot 5{,}1 = 268{,}2 - 248{,}88 = 19{,}32 .$$

El coeficient de correlació: $r_{xy} = \frac{S_{xy}}{S_x S_y} = \frac{19{,}32}{16{,}90 \cdot 1{,}26} = 0{,}91$.

Interpretació: entre les dues variables hi ha una relació lineal positiva o directa que és molt intensa. La renda familiar està molt relacionada amb el preu de l'habitatge, rendes familiars elevades en un districte es relaciona amb preus elevats i, a la inversa, rendes familiars baixes es relaciona amb preus baixos.

c) (1 punt).

Entre els districtes amb una renda familiar inferior a 45000 euros, el percentatge que té un preu de l'habitatge nou inferior a 5000 euros el metre quadrat és igual a (5/6)·100=83,3%.

d) (1 punt).

Hi ha 3 districtes amb una renda familiar superior a 45000 euros i un preu de l'habitatge nou superior a 5000 euros el metre quadrat

4)

a) (1,5 punts). *Si el plantejament i el desenvolupament és correcte: 1 punt; si el càlcul és correcte: 0,5 punts.*

$P(30 < X < 45) = F(45) - F(30) = \frac{45}{90+45} - \frac{30}{90+30} = 0{,}08\widehat{3}$. Per tant, un 8,33% de les visites duren entre 30 i 45 segons.

b) (1 punt). *Si el plantejament i el desenvolupament és correcte: 0,75 punts; si el càlcul és correcte: 0,25 punts.*

$P(X \leq 60) = F(60) = \frac{60}{90+60} = 0{,}4$. Per tant, un 40% de les visites duren menys de 60 segons.

c) (1 punt). *Si el plantejament i el desenvolupament és correcte: 0,75 punts; si el càlcul és correcte: 0,25 punts.*

$P(X > 120) = 1 - F(120) = 1 - \frac{120}{90+120} = 0{,}4286$. Per tant, un 42,86% de les visites duren més de 120 segons.

d) (2,5 punts) *Si el plantejament i el desenvolupament és correcte: 1,5 punts; si la resolució del valor és correcta: 1 punt.*

$$0{,}05 = P(X > a) = 1 - F(a) = 1 - \frac{a}{90+a} \Rightarrow \frac{a}{90+a} = 0{,}95 \Rightarrow a = 1710$$

Per tant, només un 5% de les visites duren més de 1710 segons.

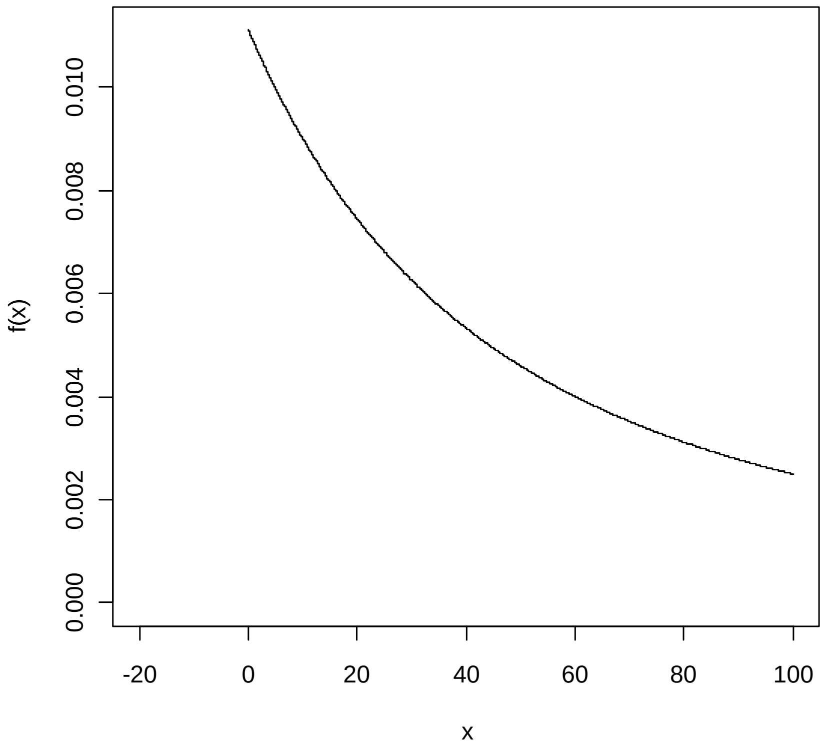
e) (2,5 punts) *Si el plantejament i el desenvolupament és correcte: 1,5 punts; si la resolució del valor és correcta: 1 punt.*

$$0{,}50 = P(X < a) = F(a) = \frac{a}{90+a} \Rightarrow a = 90$$

Per tant, un 50% de les visites no superen el 90 segons.

f) (1,5 punts). *Funció de densitat: 1 punt; representació gràfica aproximada: 0,5 punts.*

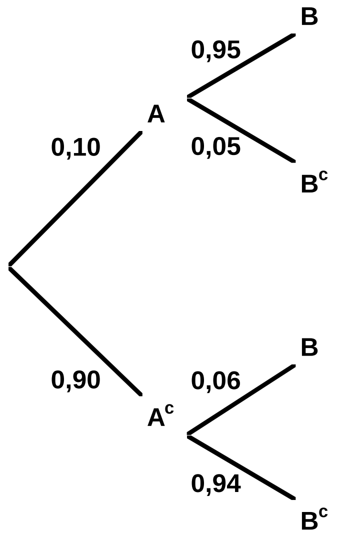
$$f(x) = F'(x) = \begin{cases} \dfrac{90}{(90+x)^2} & si \ x \geq 0 \\ 0 & si \ x < 0 \end{cases}$$



5)

a) (3 punts) *Diagrama d'arbre: 1,5 punts; probabilitats: 1,5 punts.*

A="la persona té la malaltia", B="la prova resulta positiva";
$P(A)=0{,}10$, $P(B|A^c)=0{,}05$, $P(B^c|A)=0{,}06$



b) (0,5 punts).

$P(B^c|A^c)=1- P(B|A^c)=1-0{,}06=0{,}94$.

c) (0,5 punts).

$P(B|A)=1- P(B^c|A)=1-0{,}05=0{,}95$.

d) (2 punts). *Si el plantejament i desenvolupament és correcte: 1,5 punts; si el càlcul és correcte: 0,5 punts.*

$P(A|B)=P(A\cap B)/P(B)=P(A)\cdot P(B|A)/ P(B)=0{,}10\cdot 0{,}95/0{,}149=0{,}638$.

e) (2 punts). *Si el plantejament i desenvolupament és correcte: 1,5 punts; si el càlcul és correcte: 0,5 punts.*

$P(B)=P(A)\ P(B|A)+ P(A^c)\ P(B|A^c)=0{,}10\cdot 0{,}95+0{,}90\cdot 0{,}06=0{,}149$. És a dir, el 14,9% de les vegades aquesta prova diagnòstica resulta positiva.

f) (2 punts). *Si el plantejament i desenvolupament és correcte: 1,5 punts; si el càlcul és correcte: 0,5 punts.*

B_i="la prova resulta positiva en la i-èsima persona"

$P(\{B_1 \cap B_2^c \cap B_3^c\} \cup \{B_1^c \cap B_2 \cap B_3^c\} \cup \{B_1^c \cap B_2^c \cap B_3\})=$
$=P(B_1 \cap B_2^c \cap B_3^c)+P(B_1^c \cap B_2 \cap B_3^c)+P(B_1^c \cap B_2^c \cap B_3)=$
$=0{,}149 \cdot 0{,}851 \cdot 0{,}851+0{,}851 \cdot 0{,}149 \cdot 0{,}851+0{,}851 \cdot 0{,}851 \cdot 0{,}149=0.324$

Resolució alternativa:

X="Nombre de proves amb resultat positiu". X té una distribució binomial, X~B(3;0,149). Per tant, $P(X=1)=\binom{3}{1} \cdot (0{,}149)^1 \cdot (0{,}851)^2=0{,}324$