



Sèrie 1:

1)

a) (1 punt).

Grams CO emesos	Nombre de vehicles
30-50	3
50-70	12
70-90	12
90-100	15
100-150	16
150-200	12

b) (1 punt).

El percentatge de vehicles diesel Euro 5 és $(3/26) \cdot 100 = 11,54\%$

c) (1,5 punts). *Mitjana aritmètica dels vehicles diesel i benzina: 0,75 punts cadascuna.*

$$\text{Mitjana aritmètica Diesel: } \bar{x} = \frac{\sum_{i=1}^n c_i n_i}{n} = \frac{1880}{26} = 72,30$$

$$\text{Mitjana aritmètica Benzina: } \bar{x} = \frac{\sum_{i=1}^n c_i n_i}{n} = \frac{5445}{44} = 123,75$$

d) (4,5 punts). *Variància dels dels vehicles diesel i benzina: 1 punt cadascuna; desviació típica: 0,25 cadascuna; coeficient de variació: 0,25 cadascun; interpretació: 1,5 punts.*

Variància dels vehicles diesel:

$$S_x^2 = \frac{\sum_i c_i^2 n_i}{n} - \bar{x}^2 = \frac{146350}{26} - (72,30)^2 = 5628,85 - 5227,29 = 401,56$$

$$\text{Desviació típica dels vehicles diesel: } S_x = \sqrt{401,56} = 20,04$$

Coeficient de variació de Pearson dels vehicles diesel:

$$V_x = 100 \cdot 20,04/72,30 = 27,72\%$$



Variància dels vehicles benzina:

$$S_x^2 = \frac{\sum_i c_i^2 n_i}{n} - \bar{x}^2 = \frac{731325}{44} - (123,75)^2 = 16621,02 - 15314,06 = 1306,96$$

Desviació típica dels vehicle benzina: $S_x = \sqrt{1306,96} = 36,15$

Coefficient de variació de Pearson dels vehicle benzina:

$$V_x = 100 \cdot 36,15/123,75 = 29,21\%$$

La mitjana dels vehicles diesel és més representativa ja que té el coeficient de variació de Pearson més petit.

- e) (2 punts). *Interpretació de l'enunciat: 0,5 punts; identificació de l'interval que conté el tercer quartil: 0,5 punts; càlcul del primer quartil: 1 punt.*

S'ha de calcular el primer quartil ja que per sota d'aquest valor hi ha un 25% de vehicles que emeten menys CO.

Identificació de l'interval:

Calculem la taula de freqüències absolutes acumulades:

Grams CO emesos	N_i
30-50	3
50-70	15
70-90	27
90-100	42
100-150	58
150-200	70

El primer quartil és a l'interval 70-90 ja que té la primera freqüència absoluta acumulada que supera el valor $n/4=17,5$.

Una aproximació del primer 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),

$$Q_1 = L_{i-1} + \frac{\frac{n}{4} - N_{i-1}}{n_i} \cdot a_i = 70 + \frac{17,5 - 15}{12} \cdot 20 = 74,17.$$



2)

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

Mitjana aritmètica:

$$\bar{x} = \frac{\sum_{i=1}^n c_i n_i}{n} = \frac{6128}{2470} = 2,48$$

Variància:

$$S_x^2 = \frac{\sum_i c_i^2 n_i}{n} - \bar{x}^2 = \frac{18718}{2470} - (2,48)^2 = 7,58 - 6,15 = 1,43$$

Desviació estàndard: $S_x = \sqrt{1,43} = 1,20$

- b) (3 punts). *Interpretació de l'enunciat: 1 punt; interval modal: 1 punt ; càlcul de la moda: 1 punt.*

La nota més freqüent és la moda, ja que és el valor que més vegades es repeteix.

L'interval modal és el que té la major freqüència, és a dir, interval 2-4

Una aproximació de la moda s'obté amb la següent fórmula (n_{i-1} i n_{i+1} són la densitat de freqüència dels intervals anterior i posterior a l'interval modal, respectivament),

$$Mo = L_{i-1} + \frac{n_{i+1}}{n_{i-1} + n_{i+1}} a_i = 2 + \frac{172}{832 + 172} 2 = 2,34$$

- c) (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 usuaris que donen la nota més alta.

Per identificar l'interval del tercer quartil calculem la freqüència absoluta acumulada

Número de recetas	N_i
0-2	832
2-4	2289
4-6	2461
6-8	2469
8-10	2470

El tercer quartil és a l'interval 2-4 ja que té la primera freqüència absoluta acumulada que supera el valor $3n/4=1852,5$.



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),

$$Q_3 = L_{i-1} + \frac{\frac{3n}{4} - N_{i-1}}{n_i} \cdot a_i = 2 + \frac{1852,5 - 832}{1457} \cdot 2 = 3,40.$$

- d) (2 punts). *identificació de l'interval que conté la mediana: 1 punt; càlcul de la mediana: 1 punt*

La mediana és a l'interval 2-4 ja que té la primera freqüència absoluta acumulada que supera el valor $n/2=1235$.

Una aproximació de la mediana 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),

$$Me = L_{i-1} + \frac{\frac{n}{2} - N_{i-1}}{n_i} \cdot a_i = 2 + \frac{1235 - 832}{1457} \cdot 2 = 2,55.$$

3)

- a) (7 punts). *Mitjana aritmètica de les dues variables: 1,5 punts; variàncies: 2 punts; desviacions estàndards: 0,5 punts; covariància: 1,25 punts coeficients de correlació: 0,75 punts; interpretació: 1 punt.*

Les mitjanes de nombre vehicles i velocitat:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{495}{10} = 49,5 \quad \bar{y} = \frac{\sum_{i=1}^n y_i}{n} = \frac{640}{10} = 64.$$

La variàncies:

$$S_x^2 = \frac{\sum_{i=1}^n x_i^2}{n} - \bar{x}^2 = \frac{38475}{10} - (49,5)^2 = 3847,5 - 2450,25 = 1397,25$$

i

$$S_y^2 = \frac{\sum_{i=1}^n y_i^2}{n} - \bar{y}^2 = \frac{58050}{10} - (64)^2 = 5805 - 4096 = 1709.$$

Les desviacions estàndard: $S_x = \sqrt{1397,25} = 37,38$ i $S_y = \sqrt{1709} = 41,34$.



La covariància entre les dues variables:

$$S_{xy} = \frac{\sum_{i=1}^n x_i y_i}{n} - \bar{x}\bar{y} = \frac{17125}{10} - 49,5 \cdot 64 = 1712,5 - 3168 = -1455,5.$$

El coeficient de correlació: $r_{xy} = \frac{S_{xy}}{S_x S_y} = \frac{-1455,5}{37,38 \cdot 41,34} = -0,94.$

Interpretació: entre les dues variables hi ha una relació lineal negativa que és molt intensa.

- b) (3 punts). Càlcul del pendent de la recta: 1,25 punts; càlcul del terme independent de la recta: 1,25 punts; especificació de la recta: 0,5 punts

Pendent de la recta: $b = \frac{S_{xy}}{S_x^2} = \frac{-1455,5}{1397,25} = -1,04$

Terme independent de la recta:

$$a = \bar{y} - b\bar{x} = 64 + 1,04 \cdot 49,5 = 64 + 51,48 = 115,48$$

La recta de regressió:

$$\hat{y}_i = 115,48 - 1,04 x_i$$

4)

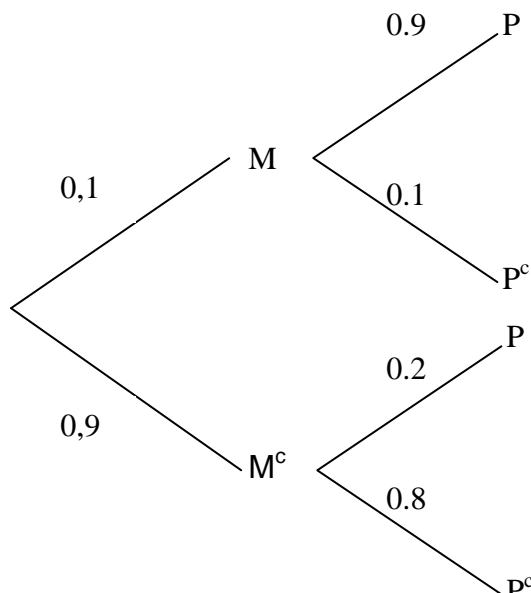
- a) (3 punts) *Diagrama d'arbre: 1,5 punts; probabilitats: 1,5 punts.*

$M =$ "desenvolupa malaltia", $P =$ "Diagnostic positiu";

$M^c =$ "no desenvolupa malaltia", $P^c =$ "Diagnostic negatiu";

$P(M) = 0,1$, $P(M^c) = 0,9$

$P(P|M) = 0,9$, $P(P^c|M^c) = 0,8$





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

$$P(P \cap M) = P(P|M) P(M) = 0.9 * 0.1 = 0.09$$

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

$$P(P \cap M) + P(P^c \cap M^c) = 0.09 + 0.72 = 0.81$$

on:

$$P(P \cap M) = P(P|M) P(M) = 0.9 * 0.1 = 0.09$$

$$P(P^c \cap M^c) = P(P^c | M^c) P(M^c) = 0.8 * 0.9 = 0.72$$

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

$$P(M^c | P^c) = P(M^c \cap P^c) / P(P^c) = 0.72 / 0.73 = 0.986$$

on:

$$P(P^c \cap M^c) = P(P^c | M^c) P(M^c) = 0.8 * 0.9 = 0.72$$

$$P(P^c) = P(P^c | M) P(M) + P(P^c | M^c) P(M^c) = 0.1 * 0.1 + 0.8 * 0.9 = 0.73$$

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

$$P(M | P) = P(M \cap P) / P(P) = 0.09 / 0.27 = 0.333$$

on:

$$P(M \cap P) = P(P|M) P(M) = 0.9 * 0.1 = 0.09$$

$$P(P) = P(P | M) P(M) + P(P | M^c) P(M^c) = 0.9 * 0.1 + 0.2 * 0.9 = 0.27$$

5)

X="Nombre de seguidors de twitter"

X té una distribució normal: $X \sim N(200, 50^2)$

Suposant $Z \sim N(0, 1)$

- a) (1,5 punts). *Si el procediment correcte: 1 punt; si la cerca del valor a les taules i el càlcul és correcte: 0,5 punts.*

$$P(X > 240) = P(Z > (240 - 200) / 50) = P(Z > 0,8) =$$

$$= 1 - P(Z < 0,8) = 1 - 0,7881 = 0,2119$$

El percentatge que té més de 240 seguidors és un 21,19%

- b) (1,5 punts). *Si el procediment correcte: 1 punt; si la cerca del valor a les taules i el càlcul és correcte: 0,5 punts.*

$$P(X < 100) = P(Z < (100 - 200) / 50) = P(Z < -2) = P(Z > 2) =$$

$$= 1 - P(Z < 2) = 1 - 0,9772 = 0,0228$$

El percentatge que té menys de 100 seguidors és un 2,28 %



- c) (2 punts). *Si el procediment correcte: 1,5 punt; si la cerca del valor a les taules i el càlcul és correcte: 0,5 punts.*

$$P(180 < X < 240) = P(X < 240) - P(X < 180) = 0.7881 - 0.3446 = 0.4435$$

- i. $P(X < 180) = P(Z < (180 - 200)/50) = P(Z < -0.4) = P(Z > 0.4) = 1 - P(Z < 0.4) = 1 - 0.6554 = 0.3446$
- ii. $P(X < 240) = 1 - P(X > 240) = 1 - 0.2119 = 0.7881$ o bé $P(X < 240) = P(Z < (240 - 200)/50) = P(Z < 0.8) = 0.7881$

El percentatge que té entre 180 i 240 seguidors és un 44,35%

- d) (2,5 punts). *Si el procediment correcte: 2 punts; si la cerca del valor a les taules i el càlcul és correcte: 0,5 punts.*

Cal trobar el valor de x de manera que
 $P(X > x) = P(Z > (x - 200)/50) = 0.1$

Dit altrament

$$P(Z < (x - 200)/50) = 1 - 0.1 = 0.9$$

Així, si fem servir les taules $(x - 200)/50 = 1,28$; aïllant la solució és igual a

$$x = 1.28 * 50 + 200 = 264$$

(fent servir les taules també es podria utilitzar el valor 1,29 en comptes de 1,28)

- e) (2,5 punts). *Si el procediment correcte: 2 punts; si la cerca del valor a les taules i el càlcul és correcte: 0,5 punts.*

Cal trobar el valor de x de manera que
 $P(X < x) = P(Z < (x - 200)/50) = 0.3$

Ara bé,

$$P(Z < (x - 200)/50) = P(Z > -(x - 200)/50) = 1 - P(Z < -(x - 200)/50)$$

Per tant, el valor de x ha de complir

$$P(Z < -(x - 200)/50) = 1 - 0.3 = 0.7$$

Així, si fem servir les taules $-(x - 200)/50 = 0,52$; aïllant la solució és igual a $x =$

$$0.52 * 50 + 200 = 174$$

(fent servir les taules també es podria utilitzar el valor 0,53 en comptes de 0,52)