Université Mohamed Boudiaf - M'sila
Faculté des Sciences et Technologies
Département de Genie Civil
Année Universitaire 2021-2022

Matière: Probabilités-Statistiques

Corrigé de Travaux Dirigés N° 2

Séries statistiques à deux variables

Exercice n°1 :

1. /La population est 28 personnes.
   Les caractère étudiés sont: 1) le sexe. Ce caractère est qualitatif nominal.
   2) la couleur des yeux : Ce caractère est qualitatif nominal.

2. Compléter le tableau de contingence:

Sexe / Couleur des yeux	Noir	Vert	Marron	total
Homme	1	4	7	12
Femme	5	6	5	16
total	6	10	12	28

3. **Rappel**:.
   f_{ij} : est donnée par $f_{ij} = \dfrac{n_{ij}}{N}$

$f_{i\bullet}$: la fréquence marginale de x_i est donné par $f_{i\bullet} = \dfrac{n_{i\bullet}}{N}$.
$f_{\bullet j}$:la fréquence marginale de y_j est donné par $f_{\bullet j} = \dfrac{n_{\bullet j}}{N}$.
Donc les fréquences du coulpe (x_i, y_j) sont

$$\begin{cases} f_{11} = \dfrac{1}{28} \\ f_{21} = \dfrac{5}{28} \end{cases}, \begin{cases} f_{12} = \dfrac{4}{28} \\ f_{22} = \dfrac{6}{28} \end{cases} \; et \; \begin{cases} f_{13} = \dfrac{7}{28} \\ f_{23} = \dfrac{5}{28} \end{cases}.$$

les fréquences marginaux de x_i sont

$$\begin{cases} f_{1\bullet} = \dfrac{1+4+7}{28} = \dfrac{12}{28} \\ f_{2\bullet} = \dfrac{5+6+5}{28} = \dfrac{16}{28} \end{cases}$$

les fréquences marginaux de y_i sont

$$\begin{cases} f_{\bullet 1} = \dfrac{1+5}{28} = \dfrac{6}{28} \\ f_{\bullet 2} = \dfrac{4+6}{28} = \dfrac{10}{28} \\ f_{\bullet 3} = \dfrac{12}{28} \end{cases}$$

4. **Rappel:** On dit que les variables X et Y sont indépendantes si pour tout le couple (i,j)
$$f_{ij} = f_{i\bullet} \times f_{\bullet j}$$
Comme $\frac{12}{28} \times \frac{6}{28} \neq \frac{1}{28}$ c-à-d $f_{11} \neq f_{\bullet 1} \times f_{1\bullet}$, donc Les deux variables X et Y ne sont pas indépendantes .

<u>**Exercice n°2 :**</u>
1) La population =100 candidats, sa taille=100, le type des variables étudiées=quantitatives continues
2) Détermine les distributions marginales de X et de Y .
On a le tableau de contingence suivant:

$X \backslash Y$	$[0,4]$	$]4,8]$	$]8,12]$	$]12,16]$	$]16,20]$	total
$[0,4]$	3	4	2	0	0	**9**
$]4,8]$	6	9	7	4	0	**26**
$]8,12]$	1	8	15	12	8	**44**
$]12,16]$	0	1	7	7	3	**18**
$]16,20]$	0	0	1	0	2	**3**
total	**10**	**22**	**32**	**23**	**13**	100

Donc, la distributions marginales de X est

X	$[0,4]$	$]4,8]$	$]8,12]$	$]12,16]$	$]16,20]$	total
$n_{i\bullet}$	9	26	44	18	3	100
x_i le centre	2	6	10	14	18	/

et la distributions marginales de Y est

Y	$[0,4]$	$]4,8]$	$]8,12]$	$]12,16]$	$]16,20]$	total
$n_{\bullet j}$	**10**	**22**	**32**	**23**	**13**	100
y_i le centre	2	6	10	14	18	/

3) La moyenne marginale de X :
$$\bar{x} = \frac{1}{N}\sum_{i=1}^{5} n_{i\bullet}x_i = \frac{1}{100}\left(9\times 2 + 26\times 6 + 44\times 10 + 18\times 14 + 3\times 18\right) = 9,2$$
La moyenne marginale de Y::
$$\bar{y} = \frac{1}{N}\sum_{i=1}^{5} n_{\bullet j}y_j = \frac{1}{100}\left(10\times 2 + 22\times 6 + 32\times 10 + 23\times 14 + 13\times 18\right) = 10,28$$
La variances marginales;de X
$$\begin{aligned} V(X) &= \left(\frac{1}{N}\sum_{i=1}^{k} n_{i\bullet}x_i^2\right) - \bar{x}^2 = \frac{1}{100}\left(9\times 2^2 + 26\times 6^2 + 44\times 10^2 + 18\times 14^2 + 3\times 18^2\right) - 9.2^2 \\ &= 14.08. \end{aligned}$$

L'écart-type marginale de X :

$$\sigma(X) = \sqrt{V(X)} = \sqrt{14.08} \approx 3.75.$$

La Variance marginale de Y :

$$\begin{aligned} V(Y) &= \left(\frac{1}{N}\sum_{i=1}^{l} n_{\bullet j} y_j^2\right) - \bar{y}^2 \\ &= \frac{1}{100}\left(10\times 2^2 + 22\times 6^2 + 32\times 10^2 + 23\times 14^2 + 13\times 18^2\right) - 10.28^2 \\ &\approx 21.84 \end{aligned}$$

L'écart-type marginale de Y

$$\begin{aligned} \sigma(Y) &= \sqrt{V(Y)} = \sqrt{21.84} \\ &\approx 4.67. \end{aligned}$$

4)] La distribution conditionnelle de Y sachant que X est dans l'intervalle $]4;8]$.

Y	$[0,4]$	$]4,8]$	$]8,12]$	$]12,16]$	$]16,20]$	total
n_{2j}	6	9	7	4	0	**26**
y_i le centre	2	6	10	14	18	/

5) la moyenne de la distribution conditionnelle de Y sachant que X est dans l'intervalle $]4;8]$.

$$\bar{y}_2 = \frac{1}{n_{2\bullet}}\sum_{j=1}^{5} n_{ij} y_j = \frac{1}{26}\left(6\times 2 + 9\times 6 + 7\times 10 + 4\times 14 + 0\times 18\right) \approx 7.38$$

Exercice n°3 :

1. Le nuage de points: On pose x le poids en kg et y la taille en cm

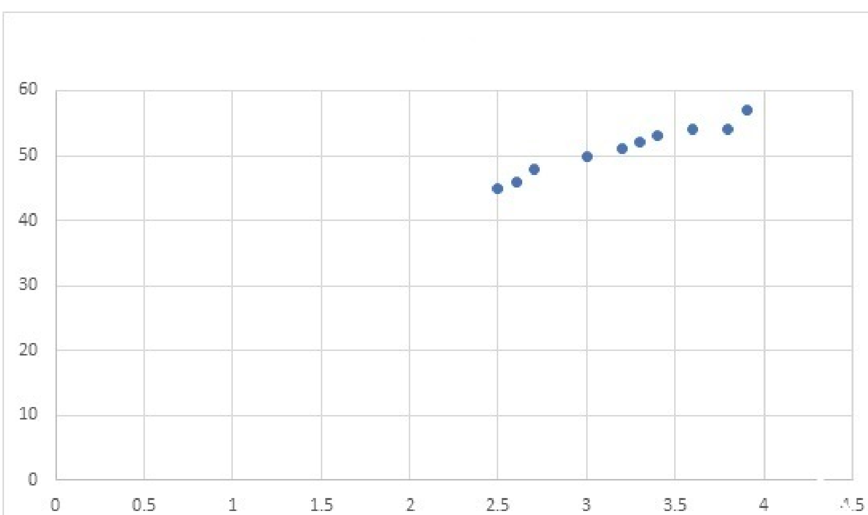

Figure 1: Nuage de points

Nuage très allongé, donc l'ajustement affine est justifié

Le tableau est dans l'ordre croissant des x on fait donc deux sous nuages, le premier avec les nouveaux nés n° 1 à 5 et le second avec les nouveaux nés n° 6 à 10.

Enfant	1	2	3	4	5
Masse en kg	2.5	2.6	2.7	3	3.2
Taille en cm	45	46	48	50	51

,

Enfant	6	7	8	9	10
Masse en kg	3.3	3.4	3.6	3.8	3.9
Taille en cm	52	54	54	54	57

Pour le point G_1 $\begin{cases} \bar{x}_1 = \dfrac{2.5+2.6+2.7+3+3.2}{5} = 2.8 \\ \bar{y}_1 = \dfrac{45+46+48+50+51}{5} = 48 \end{cases}$, on a donc le point $G_1(2,8;48)$

pour le point G_2 $\begin{cases} \bar{x}_2 = \dfrac{3.3+3.4+3.6+3.8+3.9}{5} = 3.6 \\ \bar{y}_2 = \dfrac{52+54+54+54+57}{5} = 54.2 \end{cases}$, on a donc le point $G_2(3,6;54,2)$

La droite (D) passe par le point G_1 et G_2 elle a donc pour équation $y = ax + b$ avec

$$a = \frac{54.2 - 48}{3.6 - 2.8} = 7.75$$

Le point $G_2(3,6;54,2)$ est un point de cette droite , alors
$b = \bar{y}_2 - a\bar{x}_2 = 54.2 - 7.75 \times 3.6 = 26.3$.
donc (D) : $y = 7.75x + 26.3$.

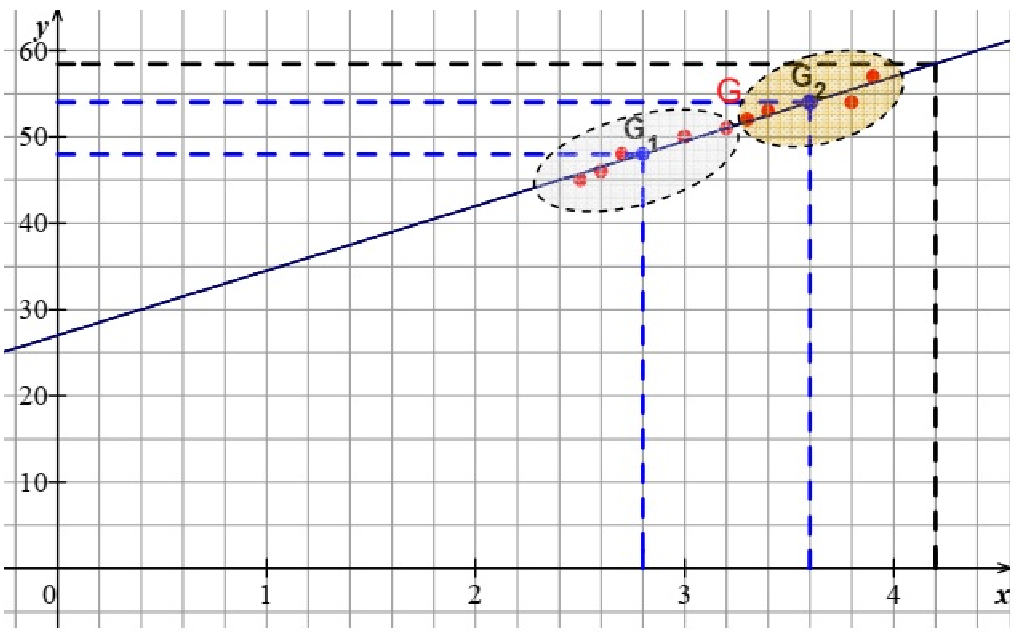

Figure 2: La droite de Mayer

3. Les coordonnés du point moyen $G\,(\bar{x}, \bar{y})$:
$\begin{cases} \bar{x} = \dfrac{\bar{x}_1 + \bar{x}_2}{2} = \dfrac{2.8+3.6}{2} = 3.2 \\ \bar{y} = \dfrac{\bar{y}_1 + \bar{y}_2}{2} = \dfrac{48+54.2}{2} = 51.1 \end{cases}$

$7.75 \times 3.2 + 26.3 = 51.1$, donc $G \in (D)$.

4. pour $x = 4,2$, $y = 7.75 \times 4.2 + 26.3 = 58.85$.

5. La droite par la méthode des moindres carrés est donnée par

$$(\Delta) : Y = aX + b,$$

avec

$$a = \frac{cov\,(X,Y)}{V\,(X)} \quad \text{et} \quad b = \bar{y} - a\bar{x}$$

On a

$$\begin{aligned} \bar{x} &= \frac{1}{10}\sum_{i=1}^{10} x_i = \frac{1}{10}\left(2.5 + 2.6 + 2.7 + 3 + 3.2 + 3.3 + 3.4 + 3.6 + 3.8 + 3.9\right) \\ &= 3.2 \end{aligned}$$

$$\begin{aligned} \bar{y} &= \frac{1}{10}\sum_{i=1}^{10} y_i \\ &= \frac{1}{10}\left(45 + 46 + 48 + 50 + 51 + 52 + 54 + 54 + 54 + 57\right) \\ &= 51.1 \end{aligned}$$

$$\begin{aligned} V\,(X) &= \left(\frac{1}{N}\sum_{i=1}^{10} x_i^2\right) - \bar{x}^2 \\ &= \frac{1}{10}\left(2.5^2 + 2.6^2 + 2.7^2 + 3^2 + 3.2^2 + 3.3^2 + 3.4^2 + 3.6^2 + 3.8^2 + 3.9^2\right) - 3.2^2 \\ &= 10.46 - 10.24 \\ &= 0.22. \end{aligned}$$

$$cov\,(X,Y) = \sigma_{xy} = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N} x_i y_j - \bar{y}\bar{x} = 1.68.$$

Alors,

$$a = \frac{cov\,(X,Y)}{V\,(X)} = \frac{1.68}{0.22} = 7.6364 \approx 7.64$$

et

$$b = \bar{y} - a\bar{x} = 51.1 - 7.64 \times 3.2 \approx 26.65.$$

Donc, donc

$$(\Delta) : y = 7.64x + 26.65.$$

Pour x = 4,2 on a

$$y = 7.64 \times 4.2 + 26.65 = 58.738$$

On retrouve un résultat proche de celui obtenu par la méthode de Mayer.

Le responsable de la matière: Merini Abdelaziz