Terminale A

Mathématiques

CÔTE D'IVOIRE – ÉCOLE NUMÉRIQUE

Thème : ORGANISATION DES DONNEES

Leçon 6 : STATISTIQUE À DEUX VARIABLES

Durée : 18 heures **Code :**

A. SITUATION D'APPRENTISSAGE

Dans le cadre des recherches pour un exposé, des élèves d'une classe de Terminale ont été accrochés par les informations suivantes :
La prévision météorologique est une science en pleine évolution. Elle a pour objectif de prédire un ensemble de paramètres comme la pluviométrie, la pression atmosphérique, la température, etc.
Le tableau suivant donne les pluviométries et températures moyennes de septembre 2018 à août 2019 dans une ville.

	Sept 2018	Oct. 2018	Nov. 2018	Déc 2018	Jan 2019	Fév. 2019	Mar 2019	Avr. 2019	Mai 2019	Juin 2019	Juillet 2019	Août 2019
Pluviométrie (en mm)	13	23	49	49	50	64	79	48	40	10	5	6
Température (en °C)	23	17	14	10	10	11	13	15	17	23	27	28

Dans l'affiche la température moyenne d'octobre 2019 était de 32 °C. Les élèves veulent connaitre la pluviométrie du mois d'octobre 2019. Un des élèves affirme que la pluviométrie n'est pas liée à la température et qu'on ne peut savoir la pluviométrie d'octobre. Ce que réfute certains. Toute la classe ayant été saisi, décide de chercher à savoir si la pluviométrie est liée à la température et si c'est le cas, de prévoir la pluviométrie d'octobre 2019.

B. CONTENU DE LA LECON

I. Présentation de la série statistique double

1. Définition

On considère deux caractères quantitatifs X et Y sur une même population de *n* individus.
On note : x_1 , x_2 , x_3 ,…, x_p les valeurs(ou les modalités) du caractère X ; y_1, y_2 , y_3 ,…, y_q les valeurs du caractère Y et n_{ij} l'effectif du couple (x_i ,y_j).
On appelle série statistique double de caractère (X, Y), l'ensemble des triplets (x_i , y_j ,n_{ij}).

Exemple

Une étude statistique porte sur une population de 100 ménages. Deux caractères X et Y sont étudiés :
- le caractère X est le nombre d'enfants,
- le caractère Y est le nombre de pièces de l'appartement occupé.

On obtient le tableau suivant qui représente la série statistique de caractère (X, Y).

Y / X	1	2	3	4
0	6	3	1	0
1	4	11	3	1
2	1	10	16	3
3	0	5	13	5
4	0	1	4	8
5	0	0	1	4

Sur la 1ère ligne, sont inscrites les valeurs du caractère Y ; soit $y_1 = 1$; $y_2 = 2$; $y_3 = 3$; $y_4 = 4$.
La 1ère colonne affiche les valeurs du caractère X qui sont : $x_1 = 0$; $x_2 = 1$; $x_3 = 2$; $x_4 = 3$; $x_5 = 4; x_6 = 5$.
Les nombres du tableau qui ne sont pas sur la 1ère ligne et la 1ère colonne, représentent les différents effectifs n_{ij} des couples (x_i, y_j).
Ainsi considérons le nombre 6 dans ce tableau. On constate qu'il est dans la colonne de la valeur 1 du caractère Y et dans la ligne de la valeur 0 du caractère X. On dit alors qu'il y a 6 ménages qui n'ont aucun enfant et qui occupent un appartement d'une pièce.
On dit que l'effectif du couple (0 ; 1) est 6.

Combien de ménages ont deux enfants et occupent un appartement de trois pièces ?
On va donc considérer la ligne ayant la valeur 2 du caractère X et la colonne ayant la valeur 3 du caractère Y. L'intersection de cette ligne et de cette colonne donne 16.
16 ménages ont donc deux enfants et occupent un appartement de trois pièces.
On dit que l'effectif du couple (2 ; 3) est 16.
Exercez-vous à la maison avec le reste des n_{ij}.
Le tableau ci-dessus est un tableau à double entrée appelé tableau de contingence.

2. Tableau de séries marginales (*Série A1 seulement*)

Reprenons l'exemple précédent.
Il est question de trouver l'effectif de chaque modalité du caractère X et l'effectif de chaque modalité du caractère Y.

Y / X	1	2	3	4	Total
0	6	3	1	0	10
1	4	11	3	1	19
2	1	10	16	3	30
3	0	5	13	5	23
4	0	1	4	8	13
5	0	0	1	4	5
Total	11	30	38	21	100

Considérons le caractère X.
Pour trouver l'effectif de la valeur 0, on additionne tous les n_{ij} qui se trouvent sur la ligne de la valeur 0 du caractère X c'est-à-dire : 6 + 3 + 1+ 0 = 10.

Quel est l'effectif de la valeur 3 du caractère X ?

Pour trouver l'effectif de la valeur 3 du caractère X, on additionne tous les n_{ij} qui se trouvent sur la ligne de la valeur 3 du caractère X c'est-à-dire : 0 + 5 + 13 + 5 = 23.

On procède de la même manière pour trouver l'effectif des autres modalités du caractère X. Ainsi à chaque valeur du caractère X, on a son effectif dans la dernière colonne.
D'où le tableau linéaire associé à X :

x_i	0	1	2	3	4	5
n_i	10	19	30	23	13	5

La série ainsi obtenue est appelée série marginale du caractère X.

En faisant de même avec les colonnes, on obtient :
L'effectif de la modalité 1 du caractère Y en additionnant les n_{ij} de la colonne où se trouve cette modalité. Soit 6 + 4 + 1 + 0 + 0 + 0 = 11.
On obtient ainsi l'effectif de chaque modalité du caractère Y dans la dernière ligne du tableau.

D'où le tableau linéaire associé à Y :

y_i	1	2	3	4
n_i	11	30	38	21

La série ainsi obtenue est appelée série marginale du caractère Y.

• Dressons le tableau des fréquences marginales du caractère X.
On rappelle que la fréquence est l'effectif de la modalité sur l'effectif total.
On obtient le tableau suivant :

x_i	0	1	2	3	4	5
f_i	$\frac{10}{100}$	$\frac{19}{100}$	$\frac{30}{100}$	$\frac{23}{100}$	$\frac{13}{100}$	$\frac{5}{100}$

De la même manière, définis le tableau des fréquences marginales du caractère Y.

y_i	1	2	3	4
f_i	$\frac{11}{100}$	$\frac{30}{100}$	$\frac{38}{100}$	$\frac{21}{100}$

3. Nuage de points

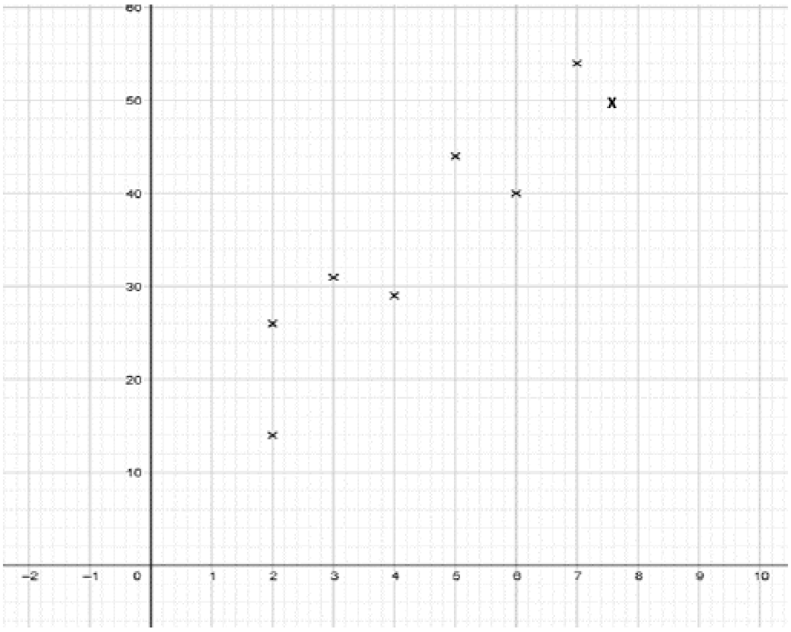
<u>Définition</u>
On considère deux caractères quantitatifs X et Y sur une même population de *n* individus.
On note x_1 , x_2 , x_3 ,…, x_p les valeurs du caractère X,
y_1 , y_2 , y_3 ,…, y_q les valeurs du caractère Y,
On appelle nuage de points associé à la série statistique double de caractère (X, Y), la représentation dans un repère orthogonal des points de couple de coordonnées $(x_i; y_j)$ d'effectifs non nuls.

Exercice de fixation
Le tableau suivant donne le nombre d'exploitations agricoles d'une région selon leur superficie en hectares.

Superficie X	2	2	3	4	5	6	7	7,6
Nombre d'exploitations Y	14	26	31	29	44	40	54	50

Représente le nuage de points associé à cette série.

Solution



Remarque

Dans la suite, les séries doubles considérées seront comme la série de l'exemple précédent ; c'est-à-dire l'effectif n_i de chaque couple $(x_i\ ,y_i)$ vaut 1.

4. Point moyen

Définition

On appelle point moyen d'un nuage de n points M_i de coordonnées $(x_i; y_i)$ le point G de coordonnées $(x_G; y_G)$ telles que :

$x_G = \overline{X} = \frac{x_1+x_2+\cdots+x_n}{n}$; $y_G = \overline{Y} = \frac{y_1+y_2+\cdots+y_n}{n}$.

Exercice de fixation

Détermine les coordonnées du point moyen du nuage de points de la série statistique suivante :

Superficie X	2	2	3	4	5	6	7	7,6
Nombre d'exploitations Y	14	26	31	29	44	40	54	50

Solution

C'est le point de coordonnées $(\overline{X}\ ;\overline{Y})$.

On a : $\overline{X} = \frac{2+2+3+4+5+6+7+7{,}6}{8} = \frac{36{,}6}{8} = 4{,}575$

et $\overline{Y} = \frac{14+26+31+29+44+40+54+50}{8} = \frac{288}{8} = 36.$

Donc G(4,575 ; 36).

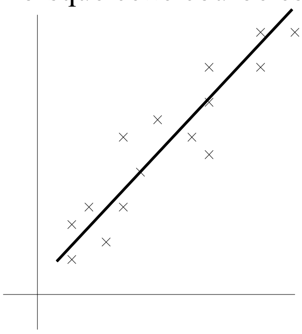
Exercice de maison
On considère la série statistique double suivante :

x_i	0	1	2	3	4	5	6	7	8
y_i	160	110	100	72	36	29	20	10	3

1) Représente le nuage de points de cette série statistique.
2) Détermine les coordonnées du point moyen de cette série statistique.

II. Ajustement
1. Présentation

Soit un nuage de points associé à une série statistique double représenté dans un repère orthogonal.
Faire un ajustement de ce nuage de points, c'est trouver une courbe qui passe le plus près « possible » du maximum de points de ce nuage.
Lorsque cette courbe est une droite, on dit que l'ajustement est affine ou linéaire.



Exemple d'ajustement par une droite.

2. Ajustement linéaire par la méthode de Mayer

a. Droite d'ajustement
Pour déterminer la droite d'ajustement linéaire d'un nuage de points, on peut procéder comme suit :
- On range la série statistique double (X; Y) suivant les valeurs croissantes des x_i.
- Si le nombre n de points du nuage de points est pair, on partage la série statistique en deux séries statistiques de même effectif :

$(x_1; y_1), (x_2; y_2), \cdots, (x_p; y_p)$ et $(x_{p+1}; y_{p+1}), (x_{P+2}; y_{P+2}), \cdots, (x_n; y_n)$, tel que p= $\frac{n}{2}$.
- Si le nombre n de points du nuage de points est impair, alors on partage le nuage de points en deux sous-nuages d'effectif $\frac{n+1}{2}$ et $\frac{n+1}{2} - 1$.
- On détermine le point moyen G_1 du premier sous-nuage et le point moyen G_2 du deuxième sous-nuage.
- La droite (G_1G_2) est la droite d'ajustement par la méthode de Mayer.

Remarque :
- La droite (G_1G_2) passe par le point moyen G du nuage de points.

Exercice de fixation
Partage la série statistique ci-dessous en deux séries et détermine le point moyen de chacune d'elles.

Superficie X	2	2	3	4	5	6	7	7,6
Nombre d'exploitations Y	14	26	31	29	44	40	54	50

Solution

Les valeurs du caractère X sont rangées dans l'ordre croissant.
L'effectif total de la série est 8.
On va partager la série en deux séries d'effectif 4 chacune.

Série 1

X	2	2	3	4
Y	14	26	31	29

Point moyen G_1
G_1 $(\overline{X_1};\overline{Y_1})$ avec $\overline{X_1} = \frac{2+2+3+4}{4} = 2{,}75$ et $\overline{Y_1} = \frac{14+26+31+29}{4} = 25$
Donc : $G_1(2{,}75 ; 25)$

Série 2

x_i	5	6	7	7,6
y_i	44	40	54	50

Point moyen G_2
G_2 $(\overline{X_2};\overline{Y_2})$, avec $\overline{X_2} = \frac{5+6+7+7{,}6}{4} = 6{,}4$ et $\overline{Y_2} = \frac{44+40+54+50}{4} = 47$.
Donc : $G_2(6{,}4 ; 47)$

b. Equation

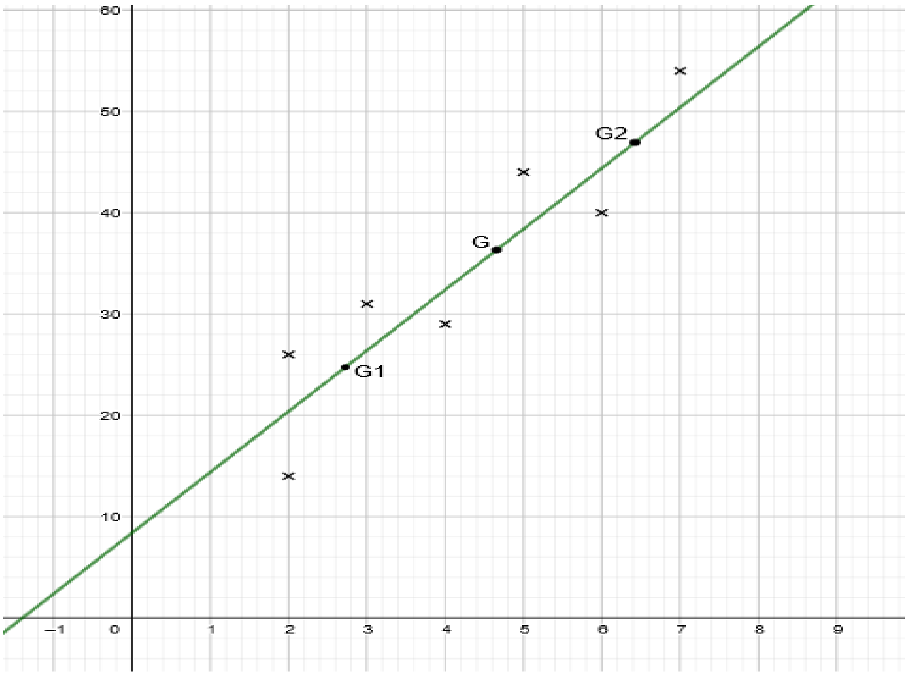
Soit $G_1(\overline{X}_1;\overline{Y}_1)$ et $G_2(\overline{X}_2;\overline{Y}_2)$ les points moyens des sous-nuages.
On détermine une équation de la droite (G_1G_2) à l'aide d'un vecteur directeur ou du coefficient directeur.

Exercice de fixation

On considère la série statistique précédente.
Détermine une équation de la droite d'ajustement linéaire du nuage de points par la méthode de Mayer.
Trace cette droite.

Solution

C'est la droite (G_1G_2) avec $G_1(2{,}75 ; 25)$ et $G_2(6{,}4 ; 47)$.
Elle a pour équation $y = ax + b$;
avec $a = \frac{\overline{Y_2} - \overline{Y}_1}{\overline{X_2} - \overline{X}_1}$ *ou* $a = \frac{\overline{Y_1} - \overline{Y}_2}{\overline{X_1} - \overline{X}_2}$ *et* $b = \overline{Y}_1 - a \times \overline{X}_1$ *ou* $b = \overline{Y}_2 - a \times \overline{X}_2$.
D'où $a = \frac{47-25}{6{,}4-2{,}75} = \frac{440}{73}$ et $b = 25 - \frac{440}{73} \times 2{,}75 = \frac{615}{73}$
Donc (G_1G_2) : $y = \frac{440}{73}x + \frac{615}{73}$



3. Ajustement linéaire par la méthode des moindres carrés (*Série A1 seulement*)

a. Covariance

Définition

On appelle covariance de la série statistique double de caractère (X ; Y), le nombre réel noté Cov(X ; Y) tel que :

$\text{Cov}(X, Y) = \frac{1}{n}\sum(x_i - \overline{\text{X}})(y_i - \overline{\text{Y}})$ ou $\text{Cov}(X, Y) = \frac{\sum x_i y_i}{n} - \overline{\text{X}}\,\overline{\text{Y}}$.

Exercice de fixation

Calcule la covariance de la série statistique suivante sachant que G(4,575 ; 36).

Superficie X	2	2	3	4	5	6	7	7,6
Nombre d'exploitations Y	14	26	31	29	44	40	54	50

Solution

La covariance de cette série statistique est telle que: $\text{Cov}(\text{X}, \text{Y}) = \frac{\sum x_i y_i}{n} - \overline{\text{X}}\,\overline{\text{Y}}$.

On a:

$\text{Cov}(\text{X}, \text{Y}) = \frac{2\times14+2\times26+3\times31+4\times29+5\times44+6\times40+7\times54+7{,}6\times50}{8} - 4{,}575 \times 36$

$\text{Cov}(\text{X}, \text{Y}) = \frac{1503}{8} - 164{,}7$.

Donc : $\text{Cov}(\text{X}, \text{Y}) = 23{,}675$

b. Coefficient de corrélation linéaire

<u>Définition</u>

Soit V(X) la variance de la série statistique de caractère X, V(Y) la variance de la série statistique de caractère Y et Cov(X ; Y) la covariance de la série statistique (X ; Y).

On appelle coefficient de corrélation linéaire de la série statistique double (X ; Y), le nombre réel noté r tel que : $r = \frac{\text{Cov}(\text{X},\text{Y})}{\sqrt{\text{V}(\text{X})}\sqrt{\text{V}(\text{Y})}}$.

Rappel : $\mathrm{V(X)} = \frac{\sum x_i^2}{n} - \left(\overline{\mathrm{X}}\right)^2$ et $\mathrm{V(Y)} = \frac{\sum y_i^2}{n} - \left(\overline{\mathrm{Y}}\right)^2$.

Exercice de fixation
Calcule le coefficient de corrélation linéaire de la série statistique suivante.

Superficie X	2	2	3	4	5	6	7	7,6
Nombre d'exploitations Y	14	26	31	29	44	40	54	50

Solution
Le coefficient de corrélation linéaire r de cette série statistique est tel que: $r = \frac{\mathrm{Cov(X,Y)}}{\sqrt{\mathrm{V(X)}}\sqrt{\mathrm{V(Y)}}}$

On a:

- $\mathrm{V(X)} = \frac{\sum x_i^2}{n} - \left(\overline{\mathrm{X}}\right)^2 = \frac{2^2+2^2+3^2+4^2+5^2+6^2+7^2+(7{,}6)^2}{8} - 4{,}575^2$
  $\mathrm{V(X)} = \frac{200{,}76}{8} - 4{,}575^2 \approx 4{,}16$

- $\mathrm{V(Y)} = \frac{\sum y_i^2}{n} - \left(\overline{\mathrm{Y}}\right)^2 = \frac{14^2+26^2+31^2+29^2+44^2+40^2+54^2+50^2}{8} - 36^2$
  $\mathrm{V(Y)} = \frac{11626}{8} - 36^2 = 157{,}25$

Donc : $r = \frac{23{,}675}{\sqrt{4{,}16 \times 157{,}25}} \approx 0{,}92$

Remarques

- Le coefficient de corrélation linéaire permet de voir la dépendance linéaire des deux caractères X et Y.
- Le coefficient de corrélation linéaire r est un nombre réel de même signe que COV(X, Y) et on a : $-1 \leq r \leq 1$.
- Si $|r|$ est proche de 1, c'est-à-dire en pratique : $0{,}87 \leq r \leq 1$ ou $-1 \leq r \leq -0{,}87$, alors on dit qu'il y a une bonne corrélation linéaire ou une forte corrélation linéaire entre les deux caractères X et Y.

Exercice de fixation
Interprète le coefficient de corrélation linéaire de l'exercice de fixation précèdent.

Solution
On sait que : $r = 0{,}92$.
Comme $0{,}87 \leq r \leq 1$, il y a une forte corrélation entre la superficie et le nombre d'exploitations agricoles de cette région.

c. Droites de régressions

Propriété
Soit V(X) la variance de la série statistique de caractère X, V(Y) la variance de la série statistique de caractère Y et Cov(X, Y) la covariance de X et Y.
On suppose qu'il y a une forte corrélation entre les caractères X et Y .

i. Droite de régression de Y en X.
La droite (D) d'équation : $y = ax + b$ où $a = \frac{\mathrm{Cov(X,Y)}}{\mathrm{V(X)}}$ et $b = \overline{Y} - a\overline{X}$ est appelée la droite de régression de Y en X par la méthode des moindres carrés.

ii. Droite de régression de X en Y.

La droite (D') d'équation : $x = a'y + b'$ avec : $a' = \frac{\text{Cov(X,Y)}}{\text{V(Y)}}$ et $b' = \overline{X} - a'\overline{Y}$ est appelée la droite de régression de X en Y par la méthode des moindres carrés.

Exercice de fixation

On considère la série statistique suivante :

Superficie X	2	2	3	4	5	6	7	7,6
Nombre d'exploitations Y	14	26	31	29	44	40	54	50

On sait que: $\text{Cov}(X, Y) = 23{,}675$,V(X)=4,6 ; V(Y)= 157,25 et 0,87≤ $r \leq 1$.

1. Détermine une équation de la droite d'ajustement linéaire de Y en X par la méthode des moindres carrés ; (On donnera les arrondis d'ordre 2 de a et b.).

2.Détermine une équation de la droite d'ajustement linéaire de X en Y par la méthode des moindres carrés. (On donnera les arrondis d'ordre 2 de a'et b')

Solution

Comme 0,87≤ $r \leq 1$, il y'a une bonne relation entre X et Y.

1.Déterminons la droite d'ajustement linéaire de Y en X par la méthode des moindres carrés.

C'est la droite (D) d'équation : $y = ax + b$ avec $a = \frac{\text{Cov(X,Y)}}{\text{V(X)}}$ et $b = \overline{\text{Y}} - a\overline{\text{X}}$

$a = \frac{\text{Cov(X,Y)}}{\text{V(X)}} = \frac{23{,}675}{4{,}16} = 5{,}69$ et $b = \overline{\text{Y}} - a\overline{\text{X}}$ = 36 – 5,69 ×4,575 = 9,97

Donc (D) : $y = 5{,}69x + 9{,}97$

2. Déterminons la droite d'ajustement linéaire de X en Y par la méthode des moindres carrés.

C'est la droite (D') d'équation : $x = a'y + b'$ avec $a' = \frac{\text{Cov(X,Y)}}{\text{V(Y)}}$ et $b' = \overline{\text{X}} - a\overline{\text{Y}}$

$a' = \frac{\text{Cov(X,Y)}}{\text{V(Y)}} = \frac{23{,}675}{157{,}25} = 0{,}15$ et $b' = \overline{\text{X}} - a\overline{\text{Y}}$ = 4,575 – 0,15×36 = –0,825

Donc : (D') : $x = 0{,}15y$ –0,83

Remarques

- Les droites (D) et (D') passent par le point moyen G du nuage de points.
- Si r est le coefficient de corrélation linéaire on a :
- $aa' = r^2$ et $|r| = \sqrt{aa'}$
- Si a > 0 et a' > 0, alors $r = \sqrt{aa'}$.
- Si a < 0 et a' < 0, alors $r = -\sqrt{aa'}$.
- Si $r^2 = 1$, alors $a = \frac{1}{a'}$ et les deux droites sont confondues.

III. Estimation

• La droite d'ajustement tracée du nuage de points permet graphiquement une estimation de y connaissant x (resp. x connaissant y).

• L'équation de la droite d'ajustement permet de calculer une estimation de y connaissant x (resp. x connaissant y).

Exercice de fixation

On considère la série statistique suivante :

Superficie X	2	2	3	4	5	6	7	7,6
Nombre d'exploitations Y	14	26	31	29	44	40	54	50

Soit (D) : $y = 5{,}69x + 9{,}97$, la droite de régression de y en x.

En considérant que la tendance se poursuit ainsi, détermine le nombre d'exploitations agricoles pour une superficie de 9 h

Solution

Par la méthode de Mayer

Une superficie de 9 ha correspond à $x = 9$.

En utilisant l'équation de la droite de Mayer, on a : $y = 6 \times 9 + 8{,}4 = 62{,}4$.

Donc pour une superficie de 9 ha, le nombre d'exploitations est estimé à 63.

Par la méthode des moindres carrés

Série A1 seulement: Avec l'équation de la droite de régression de y en x obtenue par la méthode des moindres carrés, on a : $y = 5{,}69x + 9{,}97$,

$y = 5{,}69 \times 9 + 9{,}97 = 61{,}8$

Donc pour une superficie de 9 ha, le nombre d'exploitations agricoles est estimé à 62.

C. SITUATION COMPLEXE

Le tableau ci-dessous donne le nombre total d'adhérents au club littéraire d'un lycée au cours de l'année civile 2020.

Mois	Janv	Fév	Mars	Avr	Mai	Juin	Juil	Août	Sept	Oct	Nov	Déc
Rang x_i	1	2	3	4	5	6	7	8	9	10	11	12
Nombre d'adhérents y_i	1100	1160	1220	1370	1620	1550	1600	1500	1790	1940	2060	1980

Une Organisation Non Gouvernementale promet d'octroyer une aide financière considérable au club si le nombre d'adhérents dépasse les 3000 élèves. L'élève de la Terminale A qui dirige le club désire connaître la date à laquelle ce don pourra se faire. Il te sollicite pour l'aider.

Détermine la date (mois et année) probable de la réception de ce don.

Solution.

- ➢ **Pour trouver la date, nous allons utiliser les statistiques à deux variables,**
- ➢ **Je détermine la droite de régression linéaire,**
- ➢ **J'estime la date.**

Résolution pour les classes de TA2

- Partageons la série en deux sous séries

S_1

Mois	Janv	Fév	Mars	Avr	Mai	Juin
Rang x_i	1	2	3	4	5	6
Nombre d'adhérents y_i	1100	1160	1220	1370	1620	1550

S_2

Mois	Juil	Août	Sept	Oct	Nov	Déc
Rang x_i	7	8	9	10	11	12
Nombre d'adhérents y_i	1600	1500	1790	1940	2060	1980

- **Je calcule la moyenne de chaque sous série**

Pour la série S_1
$\overline{x}_1 = \frac{1+2+3+4+5+6}{6} = 3,5$; $\overline{y}_1 = \frac{1100+1160+1220+1370+1620+1550}{6} = 1338,33$
Donc $G_1(3,5; 1338,33)$
Pour la série S_2
$\overline{x}_2 = \frac{7+8+9+10+11+12}{6} = 9,5$; $\overline{y}_2 = \frac{1600+1500+1790+1740+2060+1980}{6} = 1778,33$
Donc $G_2(9,5; 1778,33)$

- **Je détermine une équation de la droite de régression par la méthode de Mayer**
  Elle a pour équation $y = ax + b$;

$$a = \frac{\overline{Y_2} - \overline{Y}_1}{\overline{X_2} - \overline{X}_1} = \frac{1778,33 - 1338,33}{9,5 - 3,5} = \frac{440}{6} = 73,33$$
$$b = \overline{Y}_2 - a \times \overline{X}_2 = 1778,33 - 73,33 \times 9,5 = 1081,66$$

Donc $y = 73,33x + 1081,66$
- Je déduis le rang du mois pour $y = 3000$

$y = 3000$ équivaut à $x = \frac{3000-1081,66}{73,33} = 26,16$
Le rang cherché est sensiblement égal à 27.
- Je donne la date (mois et année) probable de la réception de ce don.

La date probable de la réception de ce don est mars 2022.

Résolution pour les classes de TA1

- **Je détermine une équation de la droite de régression de Y en X.**

Soit (D) cette droite.
Une équation de (D) est sous la forme: $y = ax + b$ *avec* $a = \frac{\mathrm{Cov(X,Y)}}{\mathrm{V(X)}}$ et $b = \overline{Y} - a\overline{X}$.

- Les coordonnées du point moyen G
On a:
$$\overline{X} = \frac{\sum x_i}{n} = \frac{78}{12} = 6,5$$
$$\overline{Y} = \frac{\sum y_i}{n} = \frac{18890}{12} = 1574,167$$

- La variance de X
$$V(X) = \frac{\sum {x_i}^2}{n} - \overline{\mathrm{X}}^2$$
$$V(X) = \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2 + 7^2 + 8^2 + 9^2 + 10^2 + 11^2 + 12^2}{12} - 6,5^2$$
$$\mathrm{V(X)} = \frac{650}{12} - (6.5)^2 = 11,917$$

- La variance de Y

$$V(Y) = \frac{\sum {y_i}^2}{n} - \overline{\mathrm{Y}}^2$$

$$V(Y) = \frac{1100^2 + 1160^2 + 1220^2 + 1370^2 + 1620 + 1550^2 + 1600^2 + 1500^2 + 1790^2 + 1940^2 + 2060^2 + 1980^2}{12} - (1574{,}167)^2$$

$$V(Y) = \frac{30889500}{12} - (1574{,}167)^2 = 96123{,}256$$

- La covariance de X et Y

$$\mathrm{Cov}(X, Y) = \frac{\sum x_i y_i}{n} - \overline{\mathrm{X}}\,\overline{\mathrm{Y}}$$

$$\mathrm{Cov}(X, Y) = \frac{1100 + 2320 + 3660 + 5480 + 8100 + 9300 + 11200 + 12000 + 16110 + 19400 + 22660 + 23760}{12} - 6{,}5 \times 1574{,}167$$

$$\mathrm{COV(X, Y)} = \frac{135090}{12} - 6{,}5 \times 1574{,}167 = 1025{,}4145$$

- Une équation de la droite (D): $y = ax + b$

$$a = \frac{1025{,}4145}{11{,}917} = 86{,}046$$

$$b = 1574{,}167 - 86{,}046 \times 6{,}5 = 1014{,}868$$

D'où (D): $\boldsymbol{y} = 86{,}046x + 1014{,}868$

- Je déduis le rang du mois pour $y = 3000$

$y = 3000$ équivaut à $x = \frac{3000 - 1014{,}868}{86{,}046} = 23{,}071$

Le rang cherché est sensiblement égal à 24.

- Donne la date (mois et année) probable de la réception de ce don.

La date probable de la réception de ce don est Décembre 2021.

D. EXERCICES

1. Exercices de renforcement

Exercice 1

Dans le tableau ci-dessous, on donne la taille moyenne (en cm) des nouveau-nés en fonction du nombre de l'âge gestationnel (en semaines).

Âge gestationnel (semaines)	30	31	32	33	34	35	36	37
Taille (cm)	47,5	48,5	49	49,7	50	50,5	50,8	51,2

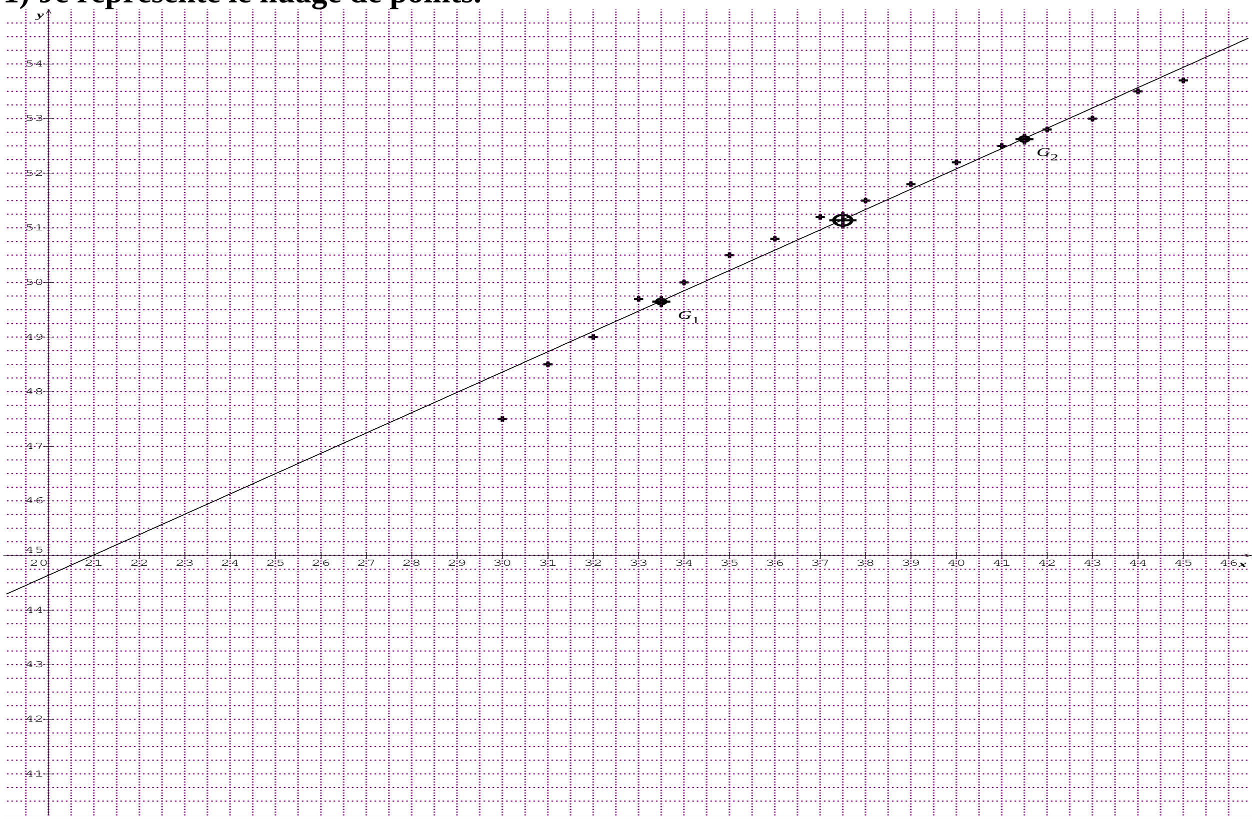
Âge gestationnel (semaines)	38	39	40	41	42	43	44	45
Taille (cm)	51,5	51,8	52,2	52,5	52,8	53	53,5	53,7

1) Représente le nuage de points dans un repère orthogonal en prenant comme unités :
En abscisses : 1 cm pour 1 semaine (commencer la graduation à 20 semaines).

En ordonnées : 2 cm par unité (commencer la graduation à 45 cm).
2) On se propose de tracer la droite d'ajustement de ce nuage de points.
a) Calcule les coordonnées des points moyens G_1 et G_2 ;
b) Trace la droite d'ajustement passant par les points G_1 et G_2.
3) Détermine une équation de cette droite d'ajustement.

SOLUTION

1) Je représente le nuage de points.



2.a) Je calcule les coordonnés des points moyens G_1 et G_2

On a:

$\overline{X_1} = \frac{\sum x_i}{n} = \frac{30+31+32+33+34+35+36+37}{8} = 33{,}5$

$\overline{Y_1} = \frac{\sum y_i}{n} = \frac{47{,}5+48{,}5+49+49{,}7+50+50{,}5+50{,}8+51{,}2}{8} = 49{,}65$

$\overline{X_2} = \frac{\sum x_i}{n} = \frac{38+39+40+41+42+43+44+45}{8} = 41{,}5$

$\overline{Y_2} = \frac{\sum y_i}{n} = \frac{51{,}5+51{,}8+52{,}2+52{,}5+52{,}8+53+53{,}5+53{,}7}{8} = 52{,}625$

Donc: $\boldsymbol{G_1}(\mathbf{33{,}5};\ \mathbf{49{,}65})$ et $\boldsymbol{G_2}(\mathbf{41{,}5};\ \mathbf{52{,}625})$

b) Tracé de la droite (G_1G_2): *(Voire nuage de points).*

***3)* Je détermine une équation de la droite (G_1G_2).**

Une équation la droite (G_1G_2) est sous la forme: $y = ax + b$ avec $a = \frac{\overline{Y_2}-\overline{Y_1}}{\overline{X_2}-\overline{X_1}}$ et $b = \overline{Y_1} - a\overline{X_1}$

$a = \frac{52{,}625 - 49{,}65}{41{,}5 - 33{,}5} = 0.372 \quad et \ \ b = 49{,}65 - 0.372 \times 33{,}5 = 37{,}188.$

D'où **(G_1G_2):** $\boldsymbol{y = 0.372x + 37{,}188}$

Exercice 2

Le tableau suivant donne le chiffre d'affaires X (en millions de francs) réalisé au cours des 6 derniers mois par un site de vente en ligne en fonction du nombre de commandes Y reçues.

Nombre de commandes (x_i)	6 400	8 350	9 125	9 600	10 050	12 000
Chiffre d'affaires mensuel (y_i)	250	320	335	350	370	400

<u>Partie A</u>

1) Représente le nuage de points associé à cette série statistique (X, Y).
2) Détermine les coordonnées du point moyen de ce nuage.
3) a) Détermine une équation de la droite d'ajustement linéaire du nuage de points par la méthode de Mayer.
   b) Trace cette droite.
4) En considérant que la tendance se poursuit ainsi, détermine le chiffre d'affaires pour 15 000 commandes reçues.

<u>Partie B</u> (Série A1 seulement)

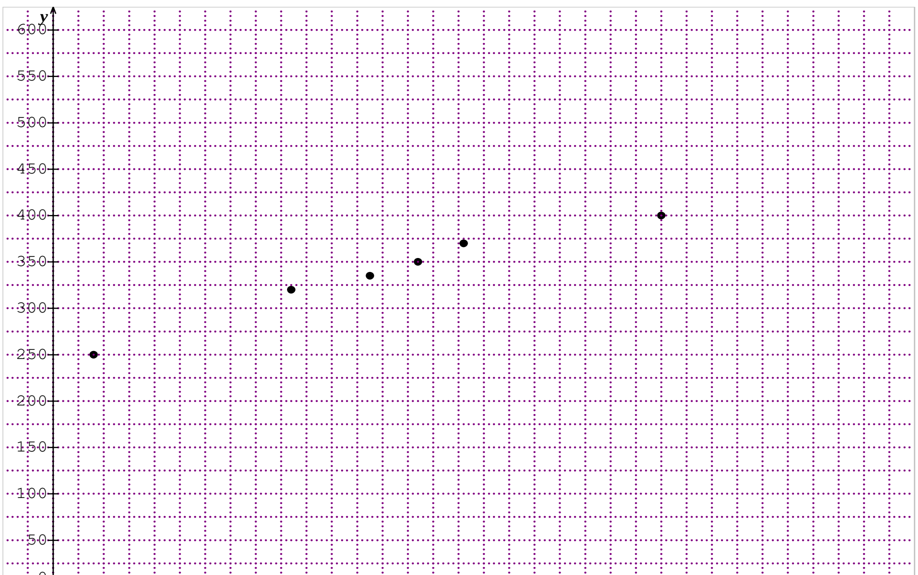
1) Calcule la covariance de la série statistique (X, Y).
2) Calcule le coefficient de corrélation linéaire de la série statistique (X, Y).
3) Interprète le coefficient de corrélation linéaire.
4) Détermine une équation de la droite d'ajustement linéaire de Y en X du nuage de points par la méthode des moindres carrés.
5) Détermine une équation de la droite d'ajustement linéaire de X en Y du nuage de points par la méthode des moindres carrés.
6) En considérant que la tendance se poursuit ainsi, détermine le chiffre d'affaires pour 15 000 commandes reçues.

SOLUTION

<u>Partie A</u>

1) Je représente le nuage de points associé à la série

Nombre de commandes (x_i)	6 400	8 350	9 125	9 600	10 050	12 000
Chiffre d'affaires mensuel (y_i)	250	320	335	350	370	400



2) **Je détermine les coordonnées du point moyen**.

Soit G($\bar{X}$; $\bar{Y}$) le point moyen . On a :

$$\bar{X} = \frac{\sum x_i}{n} = \frac{6400 + 8350 + 9125 + 9600 + 10050 + 12000}{6} = 9254{,}17$$

$$\bar{Y} = \frac{\sum y_i}{n} = \frac{250 + 320 + 335 + 350 + 370 + 400}{6} = 337{,}5$$

Donc : G(9254,17 ; 337,5)

3.a) Je détermine une équation de la droite d'ajustement linéaire du nuage de points par la méthode de Mayer

Série 1

x_i	6400	8350	9125
y_i	250	320	335

S

Soit $G_1(\bar{X}_1; \bar{Y}_1)$ le point moyen de la série 1. On a :

$$\bar{X}_1 = \frac{6400 + 8350 + 9125}{3} = 7958{,}33$$

$$\bar{Y}_1 = \frac{250 + 320 + 335}{3} = 301{,}66$$

Ainsi : G_1 (7958,33 ; 301,66).

Série 2

x_i	9600	10050	12000
y_i	350	370	400

Soit $G_2(\bar{X}_2; \bar{Y}_2)$ le point moyen de la série 2. On a :

$$\bar{X}_2 = \frac{9600 + 10050 + 12000}{3} = 10550$$

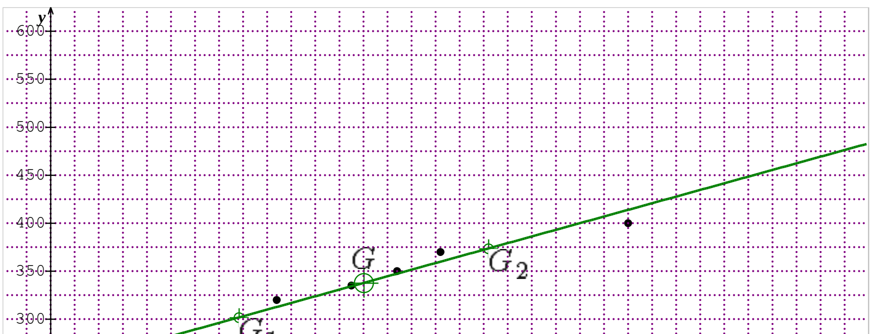
$$\bar{Y}_2 = \frac{350 + 370 + 400}{3} = 373{,}33$$

Ainsi : $G_2(10550\,;373{,}33)$

(G_1G_2): $y = ax+b$ avec a = $\frac{\bar{Y}_1-\bar{Y}_2}{\bar{X}_1-\bar{X}_2} = \frac{301{,}66-373{,}33}{7958{,}33-10550}$ = 0 ,028 et b = $\bar{Y}_2 - a\,\bar{X}_2 = 78$

C'est-à-dire (G_1G_2): y = 0,028x + 78

b) Tracé de la droite d'ajustement linéaire



4) Je détermine le chiffre d'affaires pour 15 000 commandes reçues.

Une commande de 15 000 correspond à x = 15 000.
L'équation de la droite de Mayer est : $y = 0{,}028x + 78$.
On remplace x par 15 000 dans cette équation, on obtient $y = 498$.
Le chiffre d'affaires est d'environ 498 000 000 francs.

<u>Partie B</u> (Série A1 seulement)

1) Je calcule la covariance de la série (X ; Y)
Pour calculer les moyennes, les variances et la covariance, on peut se servir du tableau suivant :
On a:

$$\mathrm{COV}(X,Y) = \frac{\sum x_i y_j}{n} - \overline{\mathrm{X}}\,\overline{\mathrm{Y}}$$

$$COV(X,Y) = \frac{6400 \times 250 + 8350 \times 320 + 9125 \times 335 + 9600 \times 350 + 10050 \times 370 + 12000 \times 400}{6}$$

$$= \frac{19207375}{6} - 9254{,}17 \text{ x } 337{,}5 = 77\,947{,}9$$

2) Je calcule le coefficient de corrélation r

$$V(X) = \frac{\sum {x_i}^2}{n} - \overline{\mathrm{X}}^2$$

$$V(X) = \frac{6400^2 + 8350^2 + 9125^2 + 9600^2 + 10050^2 + 1200^2}{6} - (9254{,}167)^2$$

$$= \frac{531\,110\,625}{6} - (9254{,}167)^2 = 2\,878\,824{,}824$$

Donc $V(X) = 2\,878\,824{,}824$

$$V(Y) = \frac{\sum {y_i}^2}{n} - \overline{\mathrm{Y}}^2$$

$$V(Y) = \frac{250^2 + 320^2 + 335^2 + 350^2 + 370^2 + 400^2}{6} - (337{,}5)^2$$

$$V(Y) = \frac{696\,525}{6} - (337{,}5)^2 = 2\,181{,}254$$

Donc
$V(Y) = 2\,181{,}254$

$$r = \frac{\mathrm{COV(X,\ Y)}}{\sqrt{\mathrm{V(X)}}\sqrt{\mathrm{V(Y)}}} = \frac{77\,947{,}9}{\sqrt{2\ \ 878\,824{,}824} \times \sqrt{2181{,}254}} = 0{,}98$$

3) J'interprète le coefficient de corrélation linéaire r

On remarque que : 0,87$\leq r < 1$, ainsi, on peut conclure qu'il y a une forte corrélation entre le nombre de commandes reçues et le chiffre d'affaires réalisé au cours des 6 derniers mois.

4) Je détermine une équation de la droite d'ajustement linéaire de Yen X par la méthode des moindres carrés

Puisqu'il y a une forte corrélation entre les commandes reçues et le chiffre d'affaires alors,

(D) a pour équation : $y = ax + b$ avec $a = \frac{\text{Cov(X,Y)}}{\text{V(X)}}$ et $b = \overline{Y} - a\overline{X}$.

$a = \frac{77947{,}9}{2878824{,}824}$ = 0,027 et b = 337,5- 0,027 x 9254, 17 = 87,63

Soit (D) : y = 0,027x + 87,63

5) **Je détermine une équation de la droite d'ajustement linéaire de X en Y par la méthode des moindres carrés**

La droite (D') a pour équation : $x = a'y + b'$ avec : $a' = \frac{\text{Cov(X,Y)}}{\text{V(Y)}}$ et $b' = \overline{X} - a'\overline{Y}$.

$a' = \frac{77947{,}9}{2181{,}254} = 35{,}735$ et b' = 9254,17-35,735 x 337, 5 = -2806,39

Soit (D'): x = 35,735y - 2806,39

6) Je détermine une estimation du chiffre d'affaires pour 15 000 commandes reçues

Ici, x = *15 000* et en utilisant l'équation de la droite de régression de Y en X, on a :

$y = 0{,}027 \times 15000 + 87{,}63$

$y = 492{,}63$

Pour 15 000 commandes reçues, le site de vente en ligne fera un chiffre d'affaires d'environ 492 630 000 francs.

2. Exercices d'approfondissement

EXERCICE 1

Un chef d'entreprise reçoit de la part de ses collaborateurs la demande d'obtenir des véhicules de fonction plus confortables et plus puissants. Il sollicite alors son comptable afin que celui-ci examine la demande et sa faisabilité.

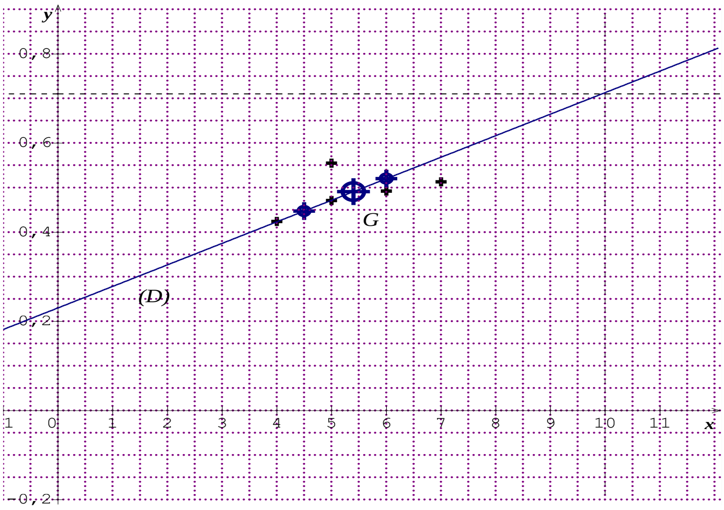
Le comptable utilise le tableau ci-dessous, donnant le prix de revient kilométrique (PRK) des véhicules d'une puissance fiscale de 4 à 8 CV et en fait une projection sur les véhicules plus puissants.

Puissance fiscale des véhicules (CV)	4	5	6	7	8
Prix de revient kilométrique (PRK)	0,424	0,471	0,492	0,513	0,555

1) Représente cette série statistique par un nuage.
2) Calcule les coordonnées du point moyen G.
3) On admet que la droite d'ajustement de cette série a pour équation : $y = 0{,}03x + 0{,}311$
   a. Justifie que le point G appartient à cette droite.
   b. Trace cette droite dans le repère précédent.
4) En utilisant la droite d'ajustement, détermine le prix de revient d'une voiture de 10 CV. (Laisser apparents les traits nécessaires à la lecture).
5). Le comptable fixe le prix de revient kilométrique maximum à 0,650 €.
Calcule la puissance maximale du véhicule qui correspond à cette exigence.

SOLUTION

1) Je représente le nuage de cette série statistique.



2) Je calcule les coordonnées du points moyen G

On a: $\overline{\mathrm{X}} = \frac{\sum x_i}{n} = \frac{4+5+6+7+8}{5} = 6$

$\overline{\mathrm{Y}} = \frac{\sum y_i}{n} = \frac{0{,}424 + 0{,}471 + 0{,}492 + 0{,}513 + 0{,}555}{5} = 0{,}491$

Donc G(6; 0,491)

3.a) Je justifie que le point G appartient à la droite d'ajustement (D) de cette série.

On a: (D): $y = 0.03x + 0.311$

Pour x = 6 on a: $y = 0.03 \times 6 + 0.311 = 0.491$.

Donc le point G appartient à la droite (D).

3.b) Tracé de (D): *(Voir nuage de point)*

4) Graphiquement pour $x = 10$, $on\ trouve\ y = 0{,}71$

Donc le prix de revient kilométrique d'une voiture de 10 CV est 0.71 €.

5) Je calcule la puissance maximale du véhicule qui a un prix de revient kilométrique de 10 CV.

On a: (D): $y = 0.03x + 0.311$.

Pour $y = 0.650$ $on\ a$: $x = \frac{0.650 - 0.311}{0.03} = 11.3$

Donc la puissance maximale du véhicule qui a un prix de revient kilométrique de 0.650 € est de 11.3 CV.

EXERCICE 2 (*Série A1 seulement*)

La consommation d'une voiture, *z*, est donnée en fonction de sa vitesse, *x*, par le tableau suivant :

x (en km/h)	80	90	100	110	120
z (en litres/ 100 km)	4	5	6,5	8	10

1) Examine la proportionnalité des deux grandeurs variables : consommation et vitesse. Justifie ta réponse.

2) Complète le tableau ci-dessus, après l'avoir reproduit, par une ligne : $y = \ln z$ dont on donnera les valeurs approchées à 6 décimales (les meilleurs possibles).

3) Dans un repère d'origine 0 ($x_0 = 70$; $y_0 = 1{,}30$ en prenant comme unités 1 cm pour 10 km/h en abscisse et 1 cm pour 0, 10 en ordonnée, représente le nuage de points (x, y).

4) Détermine une équation de la droite d'ajustement des points de coordonnées *(x, y)* du nuage, par la méthode des moindres carrés. Donne cette équation sous la forme y = *Ax* + *B*, avec les valeurs approchées de A et B (les meilleures possibles) à 3 décimales.
5) Estime *y* pour une vitesse de 140 km/h.
Estime la consommation aux 100 km pour cette vitesse de 140 km/h, à 0,5 L près comme dans le tableau initialement donné.

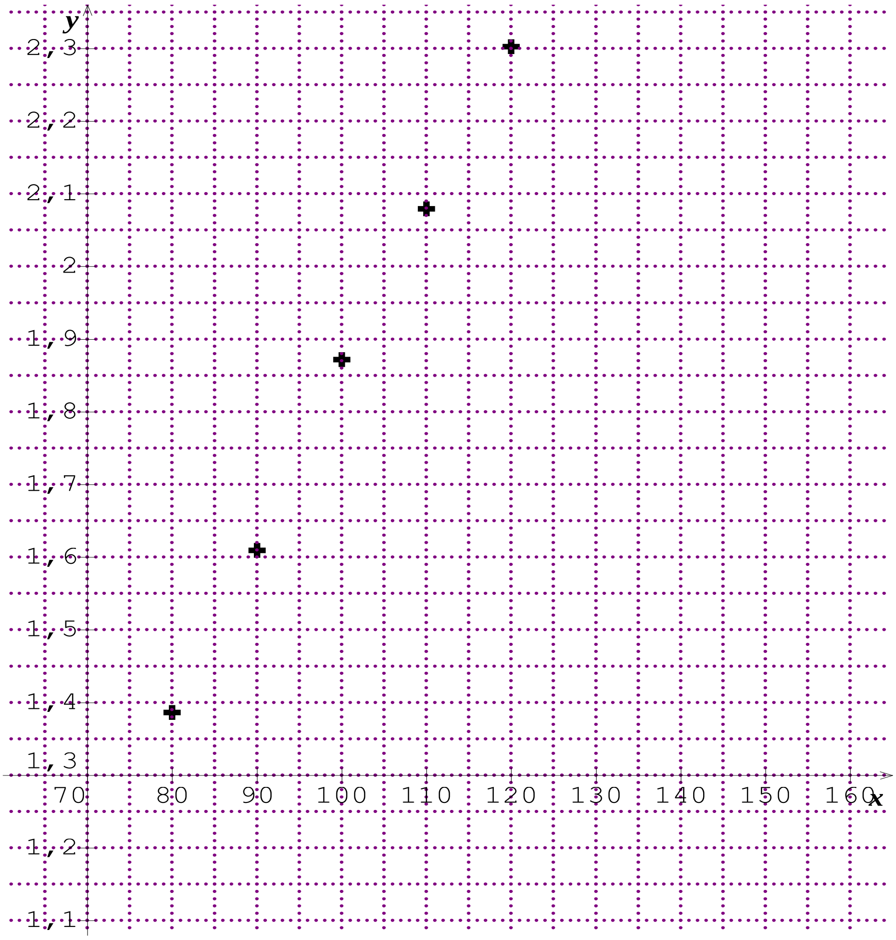
SOLUTION

1) On a: $80 \times 5 = 400 \quad alors\ que\ 90 \times 4 = 360$
Don ce n'est pas un tableau de proportionnalité.

2) Je complète le tableau

x (en km/h)	80	90	100	110	120
z (en litres/ 100 km)	4	5	6,5	8	10
y = ln(z)	1,386294	1,609438	1,871802	2,079442	2,302585

3) Je représente le nuage de points



4) Je détermine une équation de la droite d'ajustement des points de coordonnées (*x*, *y*) du nuage de points par la méthode des moindres carrés.

- Les coordonnées du points moyen G

On a: On a: $\overline{\mathrm{X}} = \frac{\sum x_i}{n} = \frac{500}{5} = 100$

$\overline{\mathrm{Y}} = \frac{\sum y_j}{n} = \frac{9{,}249561}{5} = 1{,}8499122$

Donc G(100; $\mathbf{1{,}8499122}$)

- La variance de X

On a: $V(X) = \frac{\sum {x_i}^2}{n} - \overline{\mathrm{X}}^2$

$$V(X) = \frac{80^2 + 90^2 + 100^2 + 110^2 + 120^2}{5} - 100^2$$

$$\mathrm{V(X)} = \frac{51000}{5} - (100)^2 = 200$$

- La variance de Y

On a: $V(Y) = \frac{\sum {y_i}^2}{n} - \overline{\mathrm{Y}}^2$

$$V(Y) = \frac{80^2 + 90^2 + 100^2 + 110^2 + 120^2}{5} - (1{,}8499122)^2$$

$$= \frac{17{,}641743}{5} - (1{,}8499122)^2 = 0{,}106173$$

- La covariance de X et Y

On a: $\mathrm{COV}(X, Y) = \frac{\sum x_i y_j}{n} - \overline{\mathrm{X}}\,\overline{\mathrm{Y}}$

$$COV(X,Y) = \frac{80\times1{,}386294+90\times1{,}609438+100\times1{,}871802+110\times2{,}079442+120\times2{,}302585}{5} - 100 \text{ x } 1{,}8499122$$

$$COV(X,Y) = \frac{947{,}98196}{5} - 100 \times 1{,}8499122 = 4{,}605172$$

- Le coefficient de corrélation linéaire

On a: $r = \frac{\mathrm{COV(X,Y)}}{\sqrt{\mathrm{V(X)}}\sqrt{\mathrm{V(Y)}}} = \frac{4{,}605172}{\sqrt{200}\times\sqrt{0{,}106173}} = 0{,}999364$

Il y a une bonne corrélation entre X et Y.

Une équation de la droite de régression (D) de Y en X est sous la forme $y = ax + b$

avec $a = \frac{Cov(X,Y)}{V(X)}$ $\;et\;$ $b = \bar{Y} - a\bar{X}$

$$a = \frac{4{,}605172}{200} = 0{,}023 \;\; et \;\; b = 1{,}8499122 - 0.023 \times 100 = -0{,}45$$

D'où $(\boldsymbol{D})$: $\boldsymbol{y = 0{,}023x - 0{,}45}$

5) Je donne une estimation de y pour une vitesse de *140 km/h*

Pour $\boldsymbol{x = 140}$ ***on a***: $\boldsymbol{y = 0.023 \times 140 - 045 = 2{,}77}$.

* J'estime la consommation aux 100 km pour cette vitesse de 140 km /h à 0,5 L.

On a: $y = ln(z)$ donc $z = e^y$.

Donc pour $\boldsymbol{y = 2{,}77}$ **on a :** $\boldsymbol{z = e^{2{,}77} = 15{,}959}$

Donc pour une vitesse de **140 km/h** la consommation aux 100 km sera d'environ **16 L**.

4. Situation complexe

Dans le cadre des recherches pour un exposé, des élèves d'une classe de Terminale ont été accrochés par les informations suivantes :
La prévision météorologique est une science en pleine évolution. Elle a pour objectif de prédire un ensemble de paramètres comme la pluviométrie, la pression, la température, etc.
Le tableau suivant donne les pluviométries et températures moyennes de septembre 2018 à août 2019 d'une ville.

	Sept 18	Oct. 18	Nov. 18	Déc 18	Jan 19	Fév. 19	Mars 19	Avril 19	Mai 19	Juin 19	Juillet 19	Août 19
Pluviométrie (en mm)	13	23	49	49	50	64	79	48	40	10	5	6
Température (en °C)	23	17	14	10	10	11	13	15	17	23	27	28

La température moyenne d'octobre 2019 était de 32 °C.
Ils décident alors de chercher à savoir si la pluviométrie est liée à la température et dans ce cas, prévoir la pluviométrie d'octobre 2019.
À l'aide des outils mathématiques au programme, justifie que la pluviométrie est liée à la température et détermine une estimation de la pluviométrie d'octobre 2019.

SOLUTION :
Pour répondre à ces questions, je vais utiliser la méthode des moindres carrés appliquée à la statistique à deux variables.

Soit:
- X la variable représentant la pluviométrie et prenant les valeurs x_i *(avec* i $\in \{1;\ 2;\ 3;\ \dots;\ 12\}$.
- Y la variable représentant la Température et prenant les valeurs y_i *(avec* i $\in \{1;\ 2;\ 3;\ \dots;\ 12\}$.

1) Je dresse le tableau des calculs.

	Sept. 18	Oct. 18	Nov. 18	Déc 18	Jan 19	Fév. 19	Mars 19	Avril 19	Mai 19	Juin 19	Juillet 19	Août 19	Totaux
x_i	13	23	49	49	50	64	79	48	40	10	5	6	436
y_i	23	17	14	10	10	11	13	15	17	23	27	28	208
x_iy_i	299	391	686	490	500	704	1027	720	680	230	135	168	6030
x_i^2	169	529	2401	2401	2500	4096	6241	2304	1600	100	25	36	22402
y_i^2	529	289	196	100	100	121	169	225	289	529	729	784	4060

$$\mathrm{Cov}(X, Y)$$

2) **Je calcule les coordonnées du point moyen G**.
On a: $\bar{X} = \frac{\sum x_i}{n} = \frac{436}{12} \approx 36{,}333$ et $\bar{Y} = \frac{\sum y_i}{n} = \frac{208}{12} \approx 17{,}333$.
Donc G(36,333; 17,333).

3) Je calcule :

- La variance de X

On a: $V(X) = \frac{\sum x_i^2}{n} - \bar{X}^2$

$$V(X) = \frac{13^2 + 23^2 + 49^2 + 49^2 + 50^2 + 64^2 + 79^2 + 48^2 + 40^2 + 10^2 + 5^2 + 6^2}{12} - 36.333^2$$

$V(X) = \frac{22402}{12} - 36.333^2 \approx 546{,}746\,.\quad Donc\ \ V(X) = 546{,}746$

- La variance de Y

  On a: $V(Y) = \frac{\sum y_i^2}{n} - \bar{Y}^2$

$$V(Y) = \frac{23^2 + 17^2 + 14^2 + 10^2 + 10^2 + 11^2 + 13^2 + 15^2 + 17^2 + 23^2 + 27^2 + 28^2}{12} - 17.333^2$$

$V(Y) = \frac{4060}{12} - 17.333^2 \approx 37{,}9$. Donc $V(Y) = 37{,}9$

- La covariance de X et Y

On a: $\text{Cov}(X, Y) = \frac{\sum x_i y_i}{n} - \bar{X}\,\bar{Y} = \frac{6030}{12} - 36{,}333 \times 17{,}333 \approx -127{,}26.$

Donc $\text{Cov}(X, Y) = -127{,}26$

4) Je calcule le coefficient de corrélation linéaire r

On a: $r = \frac{Cov(X,Y)}{\sqrt{V(X) \times V(Y)}} = \frac{-127{,}26}{\sqrt{546{,}746 \times 37{,}9}} \approx -0{,}884$. Donc $r = -0{,}884$

Je remarque que: $0{,}87 \leq |r| < 1$.

Donc il y a une bonne corrélation linéaire entre les variable X et Y.

Conclusion : La pluviométrie est liée à la température.

5) Je détermine une estimation de la pluviométrie d'octobre 2019 sachant que la température moyenne d'Octobre 2019 était de 32 °C.

Ici on a $y = 32$. Je doit chercher à déterminer la valeur de x en utilisant l'équation de la droite de régression de Y en X.

Soit (D) cette droite.

L'équation réduite de (D) est sous la forme $y = ax + b$ avec $a = \frac{Cov(X,Y)}{V(X)}$ et $b = \bar{Y} - a\bar{X}$.

On a: $a = \frac{-127{,}26}{546{,}746} = -0{,}233$ et $b = 17{,}333 + 0.233 \times 36{,}333 = 25{,}798$

D'où $(D): y = -0{,}233x + 25{,}798$

Donc pour $y = 32$, on a: $x = \frac{y - 25{,}798}{-0.233} = \frac{32 - 25{,}798}{-0.233} \approx -26{,}62.$

Conclusion : Pour une température moyenne de 32 °C en Octobre 2019, la pluviométrie d'Octobre 2019 est quasi-nulle.