Feuille de TD n^{o} 2

Quelques corrections

Exercice 1
On considère la série statistique à deux variables suivante :

Valeurs de X \ Valeurs de Y	1	2	3	4
1	5	8	10	7
2	12	14	16	3
3	3	12	6	4

1) RAS
2)

Valeurs de X	1	2	3
Effectif	30	45	25

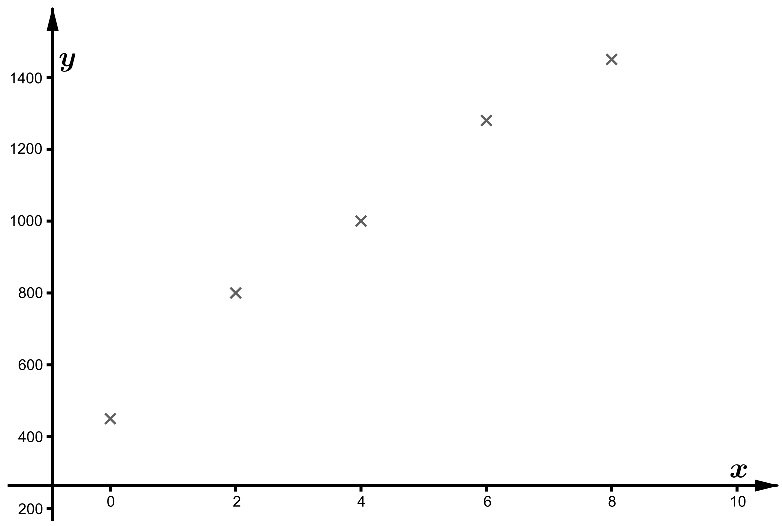
Valeurs de Y	1	2	3	4
Effectif	20	34	32	14

3) $M = (1,95; 2,40)$.

Exercice 2
Une collectivité territoriale s'intéresse à la quantité annuelle de déchets d'aluminium recyclés en tonnes. En 2014, cette collectivité dispose des données suivantes :

Année	2007	2009	2011	2013	2015
Rang de l'année X	0	2	4	6	8
Aluminium recyclé en tonnes Y	450	800	1 000	1 280	1 450

1)



Commentaire : Les points du nuage semblent quasiment alignés.

2) Une équation de la droite d'ajustement affine de y en x par la méthode des moindres carrés est donnée par

$$y = ax + b$$

avec

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

Explications des notations :

► $\overline{x}$ (resp. $\overline{y}$) est la moyenne de la série statistique X (resp. Y) ;
► $cov(X,Y)$ est la covariance de la série statistique double.
► $V(X)$ est la variance de la série statistique X.

On a :

$$cov(X,Y) = \frac{1}{N}\left(\sum_{i=1}^{N} x_i y_i\right) - \overline{x}\ \overline{y}$$

et

$$V(X) = \frac{1}{N}\left(\sum_{i=1}^{N} x_i^2\right) - \overline{x}^2$$

où

► N désigne l'effectif total. Ici $N = 5$.
► x_i (respectivement y_i) la valeur n^o i de la série X (respectivement Y).

Remarque : $V(X) = cov(X,X)$.

X	0	2	4	6	8
Y	450	800	1 000	1 280	1 450
X^2	0	4	16	36	64
Y^2	202 500	640 000	1 000 000	1 638 400	2 102 500
XY	0	1 600	4 000	7 680	11 600

De ce tableau on en déduit les calculs suivants :

$$\overline{x} = \frac{1}{N}\left(\sum_{i=1}^{N} x_i\right) = 4 \text{ et } \overline{y} = \frac{1}{N}\left(\sum_{i=1}^{N} y_i\right) = 996.$$

Donc

$$\overline{x}\ \overline{y} = 3\ 984;\ \overline{x}^2 = 16 \text{ et } \overline{y}^2 = 992\ 016.$$

Par ailleurs

$$\frac{1}{N}\left(\sum_{i=1}^{N} x_i y_i\right) = 4\ 976.$$

Donc

$$cov(X,Y) = \frac{1}{N}\left(\sum_{i=1}^{N} x_i y_i\right) - \overline{x}\ \overline{y} = 992 \text{ et } V(X) = \frac{1}{N}\left(\sum_{i=1}^{N} x_i^2\right) - \overline{x}^2 = 8.$$

Donc finalement

$$a = \frac{cov(X,Y)}{V(X)} = 124$$

et

$$b = \overline{y} - a\overline{x} = 500$$

Donc l'équation cherchée est

$$y = 124x + 500.$$

3) Le coefficient de corrélation linéaire $r_{X,Y}$ se calcule avec la formule

$$r_{X,Y} = \frac{cov(X,Y)}{\sqrt{V(X)V(Y)}}$$

où $V(Y)$ correspond à la variance de la série statistique Y. Après calcul on trouve :

$$V(Y) = \frac{1}{N}\left(\sum_{i=1}^{N} y_i^2\right) - \overline{y}^2 = 124\ 664.$$

Donc

$$r_{X,Y} \simeq 0,99.$$

Comme $|r_{X,Y}| \geq \frac{\sqrt{3}}{2}$ on en déduit que l'ajustement du nuage de points par la droite d'équation $y = 124x + 500$ est acceptable.

4) La quantité d'aluminium qui sera recyclée en 2017 sera de $124 \times 10 + 500 = 1\ 740$ tonnes.

5) Il s'agit de déterminer x entier tel que $124x + 500 \geq 1\ 900$. On trouve $x \geq 11,29$. On en déduit qu'à l'année de rang 12, c'est-à-dire en 2019, plus de 1 900 tonnes de déchets seront recyclés.

Exercice 3

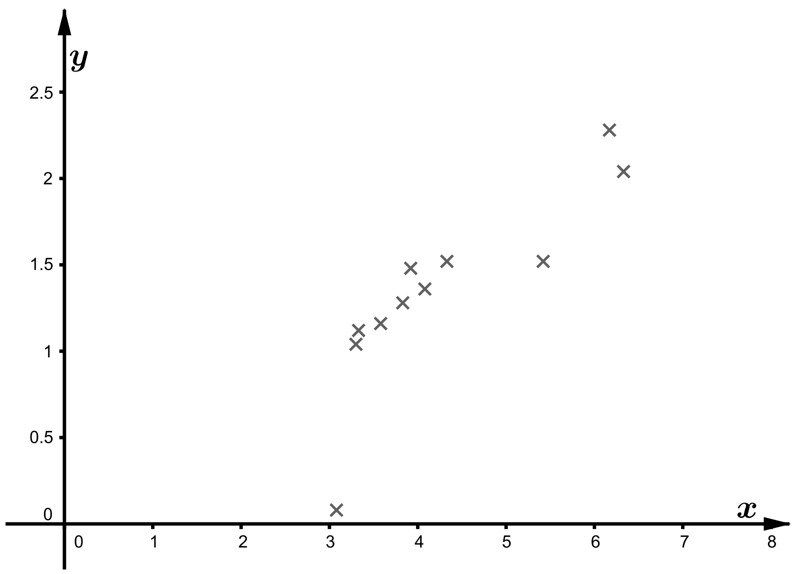
Pour la construction de maisons en Côte-d'Ivoire, on a utilisé des parpaings en géobeton, matériau obtenu en compressant de la terre.

Un laboratoire d'Abidjan a réalisé des essais de résistance à la compression.

Le tableau suivant donne les mesures obtenues pour 13 murs (notés $M_1, \ldots, M_{13}$) constitués de parpaings provenant de la même fabrication.

	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}	M_{11}	M_{12}	M_{13}
Résistance x des parpaings en Mégapascal	6,17	6,5	$3,08$	$3,58$	$3,33$	$4,08$	$3,3$	$3,92$	$4,33$	$4,5$	$6,33$	$3,83$	$5,42$
Résistance y du mur en Mégapascal	$2,28$	$2,25$	$0,8$	$1,16$	$1,12$	$1,36$	$1,04$	$1,48$	$1,52$	$1,36$	$2,04$	$1,28$	$1,52$

1)



2) Une équation de la droite d'ajustement affine de y en x par la méthode des moindres carrés est donnée par

$$y = ax + b$$

avec

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

Explications des notations :

- $\overline{x}$ (resp. $\overline{y}$) est la moyenne de la série statistique X (resp. Y) ;
- $cov(X,Y)$ est la covariance de la série statistique double.
- $V(X)$ est la variance de la série statistique X.

On a :

$$cov(X,Y) = \frac{1}{N}\left(\sum_{i=1}^{N} x_i y_i\right) - \overline{x}\ \overline{y}$$

et

$$V(X) = \frac{1}{N}\left(\sum_{i=1}^{N} x_i^2\right) - \overline{x}^2$$

où

- N désigne l'effectif total. Ici $N = 13$.
- x_i (respectivement y_i) la valeur n^o i de la série X (respectivement Y).

Remarque : $V(X) = cov(X,X)$.

	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}	M_{11}	M_{12}	M_{13}
X	6,17	6,5	3,08	3,58	3,33	4,08	3,3	3,92	4,33	4,5	6,33	3,83	5,42
Y	2,28	2,25	0,8	1,16	1,12	1,36	1,04	1,48	1,52	1,36	2,04	1,28	1,52
X^2	38,0689	42,25	9,4864	12,8164	11,0889	16,6464	10,89	15,3664	18,7489	20,25	40,0689	14,6689	29,3764
Y^2	5,1984	5,0625	0,64	1,3456	1,2544	1,8496	1,0816	2,1904	2,3104	1,8496	4,1616	1,6384	2,3104
XY	14,0676	14,625	2,464	4,1528	3,7296	5,5488	3,432	5,8016	6,5816	6,12	12,9132	4,9024	8,2384

De ce tableau on en déduit les calculs suivants :

$$\overline{x} = \frac{1}{N}\left(\sum_{i=1}^{N} x_i\right) = 4,49 \text{ Mg et } \overline{y} = \frac{1}{N}\left(\sum_{i=1}^{N} y_i\right) \simeq 1,4777 \text{ Mg.}$$

Donc

$$\overline{x}\ \overline{y} \simeq 6,6349 \text{ Mg}^2;\ \overline{x}^2 = 20,1601 \text{ Mg}^2 \text{ et } \overline{y}^2 \simeq 2,1836 \text{ Mg}^2.$$

Par ailleurs

$$\frac{1}{N}\left(\sum_{i=1}^{N} x_i y_i\right) \simeq 7,1213 \text{ Mg}^2.$$

Donc

$$cov(X,Y) = \frac{1}{N}\left(\sum_{i=1}^{N} x_i y_i\right) - \overline{x}\ \overline{y} \simeq 0,4864 \text{ Mg}^2 \text{ et } V(X) = \frac{1}{N}\left(\sum_{i=1}^{N} x_i^2\right) - \overline{x}^2 \simeq 1,3573 \text{ Mg}^2.$$

Donc finalement

$$a = \frac{cov(X,Y)}{V(X)} \simeq 0,3584 \text{ Mg}$$

et

$$b = \overline{y} - a\overline{x} \simeq -0,1315 \text{ Mg.}$$

Donc l'équation cherchée est

$$y = 0,3584x - 0,1315.$$

3) Le coefficient de corrélation linéaire $r_{X,Y}$ se calcule avec la formule

$$r_{X,Y} = \frac{cov(X,Y)}{\sqrt{V(X)V(Y)}}$$

où $V(Y)$ correspond à la variance de la série statistique Y. Après calcul on trouve :

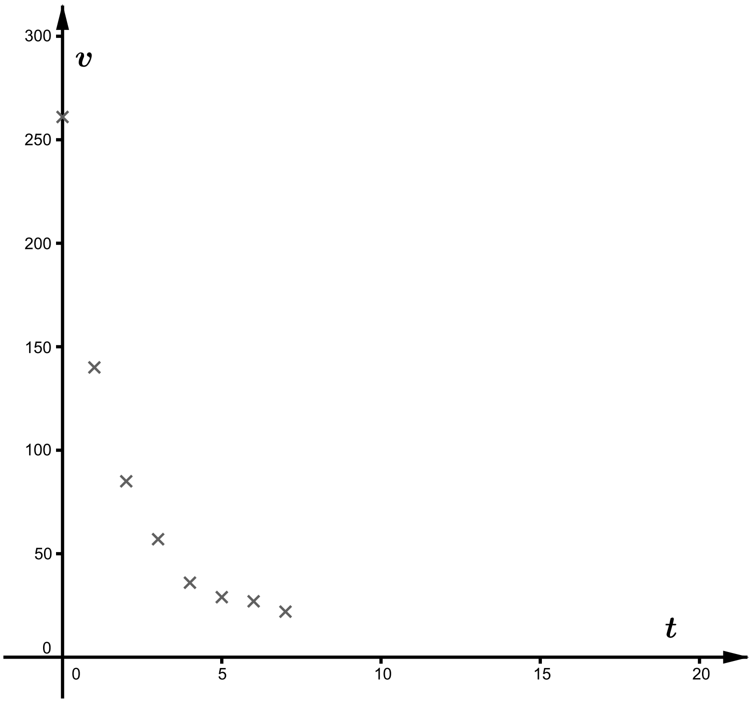
$$V(Y) = \frac{1}{N}\left(\sum_{i=1}^{N} y_i^2\right) - \overline{y}^2 \simeq 0,1928.$$

Donc
$$r_{X,Y} \simeq 0,9510.$$
Comme $|r_{X,Y}| \geq \dfrac{\sqrt{3}}{2}$ on en déduit que l'ajustement du nuage de points par la droite d'équation $y = 0,3584x - 0,1315$ est acceptable.
4) Si $x = 5$ alors $y = 0,3584 \times 5 - 0,1315 \simeq 1,6605$. Donc On peut s'attendre à avoir une résistance de compression de $1,6605$ Mg environ pour un mur constitué de parpaings ayant une résistance de compression de 5 Mg.

Exercice 4
Un mobile est propulsé à très grande vitesse sur un axe, puis il est ralenti. On s'intéresse à la vitesse de ce mobile durant le freinage. Dans tout l'exercice, les distances sont exprimées en mètre, les temps en seconde et les vitesses en mètre par seconde.
On a relevé les vitesses instantanées v_i de ce mobile aux instants t_i, pour i variant de 1 à 8.

t_i en s	0	1	2	3	4	5	6	7
v_i en m.s^{-1}	261	140	85	57	36	29	27	22

1)



Commentaire : Les points du nuage ne semblent pas alignés.
2) A l'aide de la calculatrice on trouve une équation de la droite d'ajustement affine de v en t par la méthode des moindres carrés :
$$v = -28,893t + 183,25.$$
3) A l'aide de la calculatrice on trouve :
$$r_{T,V} \simeq -0,86.$$
Comme $|r_{T,V}| < \dfrac{\sqrt{3}}{2}$ on en déduit que l'ajustement du nuage de points par la droite d'équation $v = -28,893t + 183,25$ n'est pas acceptable.
4)

T	0	1	2	3	4	5	6	7
V	261	140	85	57	36	29	27	22
$Y = \ln(V - 15)$	5,5	4,8	4,2	3,7	3,0	2,6	2,5	1,9

5) A l'aide de la calculatrice on trouve une équation de la droite d'ajustement affine de y en t par la méthode des moindres carrés :

$$y = -0,5t + 5,3.$$

6) Le coefficient de corrélation linéaire obtenu à l'aide de la calculatrice est :

$$r_{T,Y} \simeq -0,99.$$

Comme $|r_{T,V}| > \dfrac{\sqrt{3}}{2}$ on en déduit que l'ajustement du nuage de points de la série statistique (T, Y) par la droite d'équation $y = -0,5t + 5,3$ est acceptable.
On en déduit une expression approchée de la vitesse v en fonction du temps t :

$$y = -0,5t + 5,3 \Leftrightarrow \ln(v - 15) = -0,5t + 5,3 \Leftrightarrow v = e^{-0,5t+5,3} + 15.$$

7) La vitesse au bout de 9 secondes serait de $e^{-0,5\times9+5,3} + 15 \simeq 17,23$ m.s^{-1}.
8) Déterminer le temps au bout duquel on peut s'attendre à voir la vitesse du mobile passer sous $15,2$ m.s^{-1}, revient à résoudre l'inéquation

$$e^{-0,5t+5,3} + 15 \leqslant 15,2.$$

Or

$$\begin{aligned}
e^{-0,5t+5,3} + 15 \leqslant 15,2 \quad &\Rightarrow \quad e^{-0,5t+5,3} \leqslant 15,2 - 15 \\
&\Rightarrow \quad -0,5t + 5,3 \leqslant \ln(0,2) \quad \text{car la fonction ln est strictement croissante} \\
&\Rightarrow \quad -0,5t \leqslant \ln(0,2) - 5,3 \\
&\Rightarrow \quad t \geqslant \frac{\ln(0,2) - 5,3}{-0,5} \quad \text{car on divise l'inéquation par un réel strictement négatif} \\
&\Rightarrow \quad t > 13,81
\end{aligned}$$

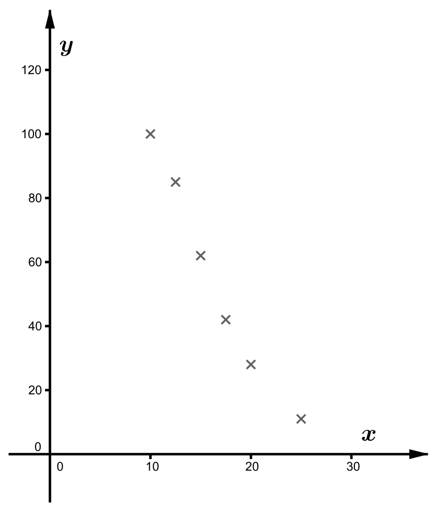
Donc au bout de $13,82$ secondes environ la vitesse du mobile passera sous $15,2$ m.s^{-1}.

Exercice 5

Une société veut vendre des machines destinées à des entreprises de terrassement. Le prix de vente minimal est fixé à 10 000 euros. Le nombre prévisible y de machines vendues est fonction du prix x proposé, en milliers d'euros. Une enquête auprès de clients potentiels a donné les résultats suivants :

x_i : prix proposé pour une machine en milliers d'euros	10	$12,5$	15	$17,5$	20	25
y_i : nombre prévisible de machines vendues au prix proposé	100	85	62	42	28	11

1)



Commentaire : Les points du nuage ne semblent pas alignés.
2)

X	10	$12,5$	15	$17,5$	20	25
Z	3,219	2,571	1,930	1,295	0,693	-0,547

3) La calculatrice permet de donner l'équation suivante :

$$z = -0,251x + 5,705.$$

4) La calculatrice permet de donner le coefficient de corrélation linéaire de la série (X, Z). On a :

$$r_{X,Z} \simeq -0,99.$$

Comme $|r_{X,Z}| > \dfrac{\sqrt{3}}{2}$ on en déduit que l'ajustement du nuage de points de la série statistique (X, Z) par la droite d'équation $z = -0,251x + 5,705$ est acceptable.
5) $z = -0,251x + 5,705 \Leftrightarrow \ln\left(\dfrac{y}{x-6}\right) = -0,251x + 5,705 \Leftrightarrow \dfrac{y}{x-6} = e^{-0,251x+5,705} \Leftrightarrow y = (x-6)e^{-0,251x+5,705}$.
6) On a

$$(23-6)e^{-0,251\times 23+5,705} \simeq 15,88.$$

Donc pour un prix proposé égal à 23 000 euros on peut estimer vendre 15 machines.