USTHB-Faculté de **M**athématiques
Module : Probabilités et Statistique.

Corrigé série d'exercices N°2

Exercice 1

Y \ X	$[5,7[$ 6	$[7,9[$ 8	$[9,11[$ 10	$n_{i.}$	$f_{i.}$	$n_{i.}X_i$
13 $[11,15[$	1	0	0	**1**	$\frac{1}{94}$	**13**
17 $[15,19[$	9	1	0	**10**	$\frac{10}{94}$	**170**
21 $[19,23[$	16	13	0	**29**	$\frac{29}{94}$	**609**
25 $[23,27[$	8	32	14	**54**	$\frac{54}{94}$	**1350**
$n_{.j}$	34	46	14	**94**		**2142**
$f_{.j}$	$\frac{34}{94}$	$\frac{46}{94}$	$\frac{14}{94}$		**1**	
$n_{.j}Y_j$	204	368	140	712	**712**	

1) **Combien d'enfants ayant moins de 9 ans ont un poids inférieur à 19 kg ?**

Il y a (1+9+1= 11) **11** enfants ayant moins de 9 ans ont un poids inférieur à 19 kg

2) **Quelle est la proportion d'enfants ayant moins de 9 ans ?**

$$\frac{34+46}{94} = \mathbf{0.851}$$

La proportion d'enfants ayant moins de 9 ans est$\mathbf{0.851}$

D'où le pourcentage $\mathbf{85,1\%}$

3) **Calculer le poids moyen et l'âge moyen des enfants.**

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} n_{i\bullet}x_i = \frac{1}{94}\sum_{i=1}^{4} n_{i\bullet}X_i = \frac{2142}{94} = \mathbf{22.79}$$

$$\bar{y} = \frac{1}{n}\sum_{j=1}^{l} n_{\bullet j}y_j = \frac{1}{94}\sum_{j=1}^{3} n_{\bullet j}y_j = \frac{1}{94}712 = \mathbf{7.57}$$

4) Donner la distribution conditionnelle de Y sachant que $X \in [23, 27[$ et calculer sa moyenne.

La distribution de $Y|_{X \in [23,27[}$:

Classes	[5, 7[	[7,9[	[9, 11[	Σ
$\boldsymbol{Y_j}$	6	8	10	
$\boldsymbol{n_{4\,j}}$	8	32	14	$n_{4\bullet} = 54$
$\boldsymbol{n_{4\,j}Y_j}$	48	256	140	**444**

$$\bar{y}_{/X\in[23,27[} = \frac{1}{n_{4\bullet}} \sum_{j=1}^{l} n_{4j}Y_j = \frac{1}{54} \sum_{j=1}^{3} n_{4j}Y_j = \frac{1}{54}\, 444 = \mathbf{8.22}$$

5) Donner la distribution conditionnelle de X sachant que $Y \in [5, 7[$

La distribution de $X|_{Y \in [5,\ 7[}$:

Classes	[11, 15[	[15,19[	[19, 23[	[23, 27[	Σ
$\boldsymbol{X_i}$	13	17	21	25	
$\boldsymbol{n_{i\,1}}$	1	9	16	8	$n_{\bullet 1} = 34$
$\boldsymbol{n_{i1}X_i}$	13	153	336	**200**	**702**

$$\bar{x}_{/Y\in[5,\ 7[} = \frac{1}{n_{\bullet 1}} \sum_{i=1}^{k} n_{i1}X_i = \frac{1}{34} \sum_{i=1}^{4} n_{i1}X_i = \frac{1}{34}\, 702 = \mathbf{20.64}$$

6) X et Y sont-elles indépendantes ?

On a $\bar{x}_{/Y\in[5,\ 7[} = 20.64 \neq 22.79 = \bar{x}$ donc X dépend de Y.

Exercice 02 : On a relevé la production de blé (**X** en quintaux) et le nombre de jours de pluie (**Y**). On a obtenu les résultats suivants :

Valeurs de X	200	184	225	250	240	195	210	225	250	220
Valeurs de Y	50	30	70	90	50	30	50	60	70	70

1- Compléter le tableau de contingence suivant :

X \ Y	30	50	60	70	90	$n_{i\cdot}$	$n_{i\cdot}X_i$	$n_{i\cdot}X_i^2$	$\sum_{j=1}^{5} n_{ij}X_iY_j$
$[184, 206[$ **195**	2	1	0	0	0	3	585	114075	21450
$[206, 228[$ **217**	0	1	1	2	0	4	868	188356	54250
$[228, 250]$ **239**	0	1	0	1	1	3	717	171363	50190
$n_{\cdot j}$	2	3	1	3	1	10	**2170**	**473794**	**125890**
$n_{\cdot j}Y_i$	60	150	60	210	90	**570**			
$n_{\cdot j}Y_i^2$	1800	7500	3600	14700	8100	**35700**			

2- X et Y sont-elles indépendantes ?

X et Y sont dites indépendantes si : $\boldsymbol{n \times n_{ij} = n_{i\bullet} \times n_{\bullet j}}$, $\forall \boldsymbol{i}$ **et** $\forall \boldsymbol{j}$.

Pour $\boldsymbol{i = 2}$ et $\boldsymbol{j = 1}$ on a

$n \times n_{21} = 10 \times 0 = 0$ et $n_{2\bullet} \times n_{\bullet 1} = 4 \times 2 = 8 \neq 0$ donc X et Y sont dépendants

3- Calculer la covariance entre les variables X et Y.

$$S_{XY} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{l} n_{ij}X_iY_j - \bar{x}\bar{y}$$

Calculons $\bar{x}$ et $\bar{y}$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{3} n_{i\bullet}X_i = \frac{1}{10}2170 = \mathbf{217}$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{5} n_{\bullet j}Y_j = \frac{1}{10}570 = \mathbf{57}$$

Donc $S_{XY} = \frac{1}{10}125890 - 217 \times 57 = \mathbf{220}$

4- Calculer le coefficient de corrélation et interpréter le résultat obtenu.

$$\boldsymbol{\rho_{XY}} = \frac{\boldsymbol{S_{XY}}}{\boldsymbol{\sigma_X \sigma_Y}}$$

Calculons $\boldsymbol{\sigma_X}$ et $\boldsymbol{\sigma_Y}$

$$\sigma_X^2 = \left(\frac{1}{n}\sum_{i=1}^{5} n_{i\bullet} X_i^2\right) - \bar{x}^2 = \mathbf{290.4}\text{donc } \sigma_X = \mathbf{17.04}$$
$$\sigma_Y^2 = \left(\frac{1}{n}\sum_{j=1}^{3} n_{\bullet j} Y_j^2\right) - \bar{y}^2 = \mathbf{321}\text{donc } \sigma_Y = \mathbf{17.91}$$

Ainsi

$$\rho_{XY} = \frac{S_{XY}}{\sigma_X \sigma_Y} = \frac{220}{17.04 \times 17.91} = \mathbf{0.72} > 0.7 \Rightarrow \text{L'ajustement affine est justifié}$$

5- Donner la droite de régression de Yen X

On a :$\boldsymbol{D_y(x)}: \hat{y} = \hat{a}\, x + \hat{b}$ avec

$$\hat{\boldsymbol{a}} = \frac{S_{XY}}{\sigma_X{}^2} = \frac{220}{290.4} = \mathbf{0.75}\text{et } \hat{b} = \bar{y} - \hat{a}\bar{x} = 57 - 0.75\,(217) = -\mathbf{107.39}$$

Donc :

$$\boxed{\boldsymbol{D_y(x)}: \hat{y} = 0.75\, x - 107.39}$$

6- Prévoir la production de blé pour un nombre de jours de pluie égal à 55 jours.

On a $y_0 = 55$ et on cherche $x_0 = ?$

On a besoin de trouver la droite de régression de **X en Y** :$D_x(y): \hat{x} = \hat{a}' y + \hat{b}'$ avec

$$\hat{\boldsymbol{a}}' = \frac{\boldsymbol{S_{XY}}}{\boldsymbol{\sigma_Y{}^2}} = \frac{220}{321} = \mathbf{0.68}$$

$$\hat{\boldsymbol{b}}' = \bar{\boldsymbol{x}} - \hat{\boldsymbol{a}}'\bar{y} = 217 - 0.68(57) = \mathbf{178.24}$$

Donc :

$$\boxed{D_x(y): \hat{x} = 0.68y + 178.24}$$

$$x_0 = 0.68\,(y_0) + 178.24 = 0.68\,(55) + 178.24 = 215.64$$

La production de blé pour un nombre de jours de pluie égal à 55 jours est :$\mathbf{215.64}$ **quintaux.**

Exercice 3 :

Xla quantité de blé et Y la quantité d'orge.

1) Tableau de contingence

Nombre de classes pour X , $\boldsymbol{k} = \mathbf{3}$

$$a = \frac{e}{k} = \frac{250 - 190}{3} = \frac{60}{3} = \mathbf{20}$$

Nombre de classes pour Y , $\boldsymbol{l} = \mathbf{2}$

$$a = \frac{e}{l} = \frac{115 - 75}{3} = \frac{40}{2} = \mathbf{20}$$

X \ Y	$[\mathbf{75}, \mathbf{95}[$ $\mathbf{85}$	$[\mathbf{95}, \mathbf{115}[$ $\mathbf{105}$
$[\mathbf{190}, \mathbf{210}[$ $\mathbf{200}$	2	2
$[\mathbf{210}, \mathbf{230}[$ $\mathbf{220}$	0	2
$[\mathbf{230}, \mathbf{250}]$ $\mathbf{240}$	2	2

2) Les distributions marginales :

X \ Y	$[\mathbf{75}, \mathbf{95}[$ $\mathbf{85}$	$[\mathbf{95}, \mathbf{115}[$ $\mathbf{105}$	$\boldsymbol{n_{i\cdot}}$
$[\mathbf{190}, \mathbf{210}[$ $\mathbf{200}$	2	2	4
$[\mathbf{210}, \mathbf{230}[$ $\mathbf{220}$	0	2	2
$[\mathbf{230}, \mathbf{250}]$ $\mathbf{240}$	2	2	4
$\boldsymbol{n_{\cdot j}}$	4	6	**10**

3) Calculer le coefficient de corrélation entre X et Y.

X \ Y	$[75,95[$ **85**	$[95,115[$ **105**	$n_{i\cdot}$	$n_{i\cdot}X_i$	$\sum_{j=1}^{2} n_{ij}X_iY_j$
$[190,210[$ **200**	2	2	4	800	76000
$[210,230[$ **220**	0	2	2	440	46200
$[230,250]$ **240**	2	2	4	960	91200
$n_{\cdot j}$	4	6	**10**	**2200**	$\sum_{i=1}^{3}\sum_{j=1}^{2} n_{ij}X_iY_j$= **213 400**
$n_{\cdot j}Y_j$	340	630	**970**	/	/ / /

$\rho_{XY} = \frac{S_{XY}}{\sigma_X\sigma_Y}$avec $S_{XY} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{l} n_{ij}X_iY_j - \bar{x}\bar{y}$

Calculons $\bar{x}$ et $\bar{y}$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{3} n_{i\bullet}X_i = \frac{1}{10}\,2200 = \mathbf{220}$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{2} \boldsymbol{n_{\cdot j}Y_j} = \frac{1}{10}\,970 = \mathbf{97}$$

Donc

$$S_{XY} = \frac{1}{10}\sum_{i=1}^{3}\sum_{j=1}^{2} n_{ij}X_iY_j - \bar{x}\bar{y} = \frac{1}{10}\mathbf{213\,400} - \mathbf{220} \times \mathbf{97} = \mathbf{0}$$

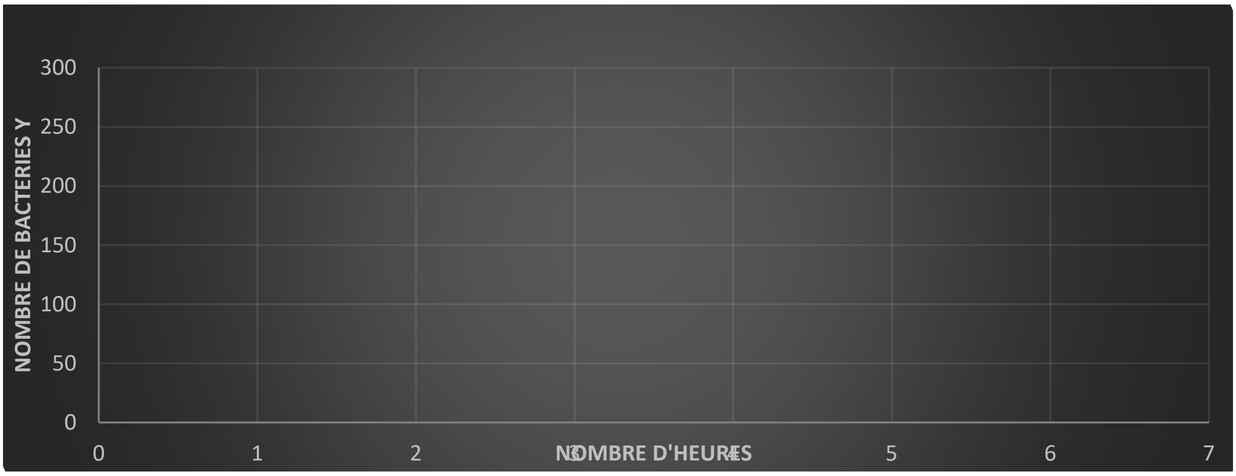
Ainsi $\boldsymbol{\rho_{XY}} = \mathbf{0}$ **.**

<u>Exercice 4 :</u>

Le nombre Y de bactéries présentes par unité de volume après X heures :

Nombre d'heures (X)	0	1	2	3	4	5	6
Nombre de bactéries par unité de volume (Y)	32	47	65	92	132	190	275

1/ Nuage de points.



2/ Ajuster ces points au sens des moindres carrés par la courbe d'équation $\boldsymbol{Y = a\,b^X}$.

On a : $\boldsymbol{Y = a\,b^X \Longrightarrow \ln Y = X \ln b + \ln a}$ **,** on obtient alors une équation d'une droite de la forme :

$\boldsymbol{V = BX + A}$ avec $V = \ln Y$, $B = \ln b$, $A = \ln a$

X	Y	$V = \ln Y$	XV	X^2
0	32	3.45	0	0
1	47	3.85	3.85	1
2	65	4.17	8.34	4
3	92	4.52	13.56	9
4	132	4.88	19.52	16
5	190	5.24	26.2	25
6	275	5.61	33.66	36
21		**31.73**	**105.13**	**91**

Calculons A et B

$$\boldsymbol{B} = \frac{S_{XV}}{\sigma_X^2} = \frac{\boldsymbol{cov}\,(\boldsymbol{X},\boldsymbol{V})}{\sigma_X^2} \quad \text{et} \quad \boldsymbol{A} = \overline{\boldsymbol{V}} - \boldsymbol{B}\overline{\boldsymbol{X}}$$

- $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{7}\sum_{i=1}^{7} X_i = \frac{1}{7}\ 21 = \mathbf{3}$
- $\bar{V} = \frac{1}{n}\sum_{i=1}^{n} V_i = \frac{1}{7}\sum_{i=1}^{7} V_i = \frac{1}{7}\ 31.73 = \mathbf{4.53}$
- $cov\,(X,V) = \frac{1}{n}\sum_{i=1}^{n} X_iV_i - \bar{X}\bar{V} = \frac{1}{7}\ 105.13 - \ 3 \times 4.53 = \mathbf{1.42}$
- $\sigma_X^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \ \bar{X}^2 = \frac{1}{7}\ 91 - 9 = \mathbf{4}$

Donc $B = \frac{cov\,(X,V)}{\sigma_X^2} = \frac{1.42}{4} = \mathbf{0.355}$

$A = \bar{V} - B\bar{X} = 4.53 - 0.355 \times 3 = \mathbf{3.465}$

Ainsi l'équation de la droite de régression est :

$$\boxed{\boldsymbol{V} = \mathbf{0.355}\ \boldsymbol{X} + \mathbf{3.465}}$$

Déterminons l'ajustement selon **la courbe d'équation $\boldsymbol{Y} = \boldsymbol{a}\ \boldsymbol{b}^{\boldsymbol{X}}$**

On a

- $A = lna \Longrightarrow a = e^A = \ e^{3.465} = \mathbf{9.41}$
- $B = lnb \ \Longrightarrow b = e^B = e^{0.355} = \mathbf{0.96}$

Donc

$$\boxed{\boldsymbol{Y} = \mathbf{9.41} \times (\mathbf{0.96})^{\boldsymbol{X}}}$$

Exercice 5 :

La distance de freinage d'un véhicule sur une route sèche, en fonction de sa vitesse est donnée par le tableau suivant :

V : vitesse(km/h)	30	40	50	60	70	80	90	100	110	120
D : distance (m)	12	14	28	36	48	64	80	96	114	148

On donne : $\sum_{i=1}^{10} V_i = 750 \quad \sum_{i=1}^{10} D_i = 640 \quad \sum_{i=1}^{10} V_i^2 = 64500 \quad \sum_{i=1}^{10} D_i^2 = 59336 \quad \sum_{i=1}^{10} V_i\, D_i = 60060$

1/ Calculer le coefficient de corrélation entre V et D

On a: $\boldsymbol{\rho_{VD}} = \frac{\boldsymbol{S_{VD}}}{\boldsymbol{\sigma_V \sigma_D}} = \frac{\boldsymbol{cov\ (V,D)}}{\boldsymbol{\sigma_V \sigma_D}}$

- $\bar{V} = \frac{1}{n}\sum_{i=1}^{n} V_i = \frac{1}{10}\sum_{i=1}^{10} V_i = \frac{1}{10}\ 750 = \mathbf{75}$
- $\bar{D} = \frac{1}{n}\sum_{i=1}^{n} D_i = \frac{1}{10}\sum_{i=1}^{10} D_i = \frac{1}{10}\ 640 = \mathbf{64}$
- $S_{VD} = \left(\frac{1}{10}\sum_{i=1}^{10} V_i D_i\right) - \bar{V}\bar{D} = \left(\frac{1}{10} 60\ 060\right) - 75 \times 64 = \mathbf{1206}$
- $\sigma_V^2 = \left(\frac{1}{10}\sum_{i=1}^{10} V_i^2\right) - \bar{V}^2 = \frac{1}{10} 64\ 500 - 75^2 = \mathbf{825}$donc $\sigma_V = \mathbf{28.72}$
- $\sigma_D^2 = \left(\frac{1}{10}\sum_{i=1}^{10} D_i^2\right) - \bar{D}^2 = \frac{1}{10} 59\ 336 - 64^2 = 1837.6$donc $\sigma_D = \mathbf{42.86}$

Ainsi

$\rho_{VD} = \frac{S_{VD}}{\sigma_V \sigma_D} = \frac{1206}{28.72 \times 42.86} = \mathbf{0.97} > 0.7$ ⇒L'ajustement affine est justifié et il y a une très forte corrélation linéaire entre V et D

2/ Donner les équations des deux droites de régression.

- **La droite de régression de D en V :** $\hat{D} = \hat{a}\ V + \hat{b}$

$$\hat{\boldsymbol{a}} = \frac{S_{VD}}{{\sigma_V}^2} = \frac{1206}{825} = \mathbf{1.46}$$

$\hat{\boldsymbol{b}} = \bar{D} - \hat{a}\bar{V} = 64 - 1.46(75) = \mathbf{-45.5}$

Donc $\boxed{\hat{\boldsymbol{D}} = \mathbf{1.46}\ \boldsymbol{V} - \mathbf{45.5}}$

- **La droite de régression de V en D :** $\hat{V} = \hat{a}'D + \hat{b}'$

$$\hat{\boldsymbol{a}}' = \frac{S_{VD}}{{\sigma_D}^2} = \frac{1206}{1837.6} = \mathbf{0.66}$$

$\hat{\boldsymbol{b}}' = \bar{V} - \hat{a}'\bar{D} = 75 - 0.66(64) = \mathbf{-32.76}$

Donc $\boxed{\hat{\boldsymbol{V}} = \mathbf{0.66}\ \boldsymbol{D} - \mathbf{32.76}}$

3/ Peut-on estimer la distance de freinage, si le véhicule roule à 150km/h ?

V= 150 D= ?

$\hat{D} = 1.46(150) - 45.5 = 173.5$

Donc, si le véhicule roule à 150 km/h, la distance de freinage est **173.5 mètres**.