Université A.MIRA–Béjaia
Faculté de Technologie
Département de Technologie-2ème Année

Examen Final de Probabilités et Statistiques

Exercice 1 (09.50 points) :
Soit la répartition des salaires (10^2 DA) journaliers des employés d'une usine :

Salaires	[5,6[	[6,7[	[7,8.5[	[8.5, 9[	[9,10[
Fréquences cumulées croissantes	0.18	0.43	0.79	0.94	1

1. Déterminer :
   i- La population étudiée (en précisant l'unité statistique).
   ii- Le caractère étudié et sa nature.
2. Construire le tableau représentatif de cette distribution statistique.
3. Représenter graphiquement cette distribution.
4. Donner la fonction de répartition et tracer son graphe.
5. Calculer le mode, la médiane, la moyenne et l'intervalle interquartile.
6. Quelle est la proportion des employés qui perçoivent un salaire compris entre 6 et 8 (10^2 DA) par jour ?

Barème Détaillé de l'Exercice 1 : 01.00+01.00+00.50+02.00+03.50+01.50

Exercice 2 (10.50 points) :
On a étudié pendant les années 1995 à 2004 la production de blé X (en tonnes/ha) en fonction du nombre de jours de pluie Y, et obtenu :

X_i	13	13.4	14.5	15.2	16.5	16.5	17.2	17.4	18	18.3
Y_i	40	40	50	50	50	60	60	70	70	60

1. Préciser la nature des variables statistiques étudiées.
2. Calculer la moyenne de X et de Y.
3. Calculer le coefficient de corrélation linéaire. Conclure.
4. Déterminer l'équation de la droite de Mayer.

Maintenant, on décide de répartir la variable X en trois classes.

1. Compléter le tableau suivant :

X/Y	40	50	60	70
[13, 15[	2			
[15, 17[		2		
[17, 19[			2	

2. Déterminer les deux distributions marginales de X et Y.
3. Les deux variables sont-elles indépendantes ? Justifier votre réponse.
4. Déterminer la distribution conditionnelle de $Y/X < 15$ et calculer sa moyenne.
5. Déterminer l'équation de la droite de régression de X en Y.
6. Peut-on prévoir la production de blé pour 75 jours de pluies ?

Barème Détaillé de l'Exercice 2 :
00.50+00.50+01.50+02.25+00.75+01.00+00.50+01.00+01.50+01.00

Bonne Chance
© Mr Boualem

Corrigé de l'examen Maths 04

2013 - 2014

Exo1 : 1) Déterminer :

i) Population : Les employés d'une usine
Unité statistique : 1 employé.

ii) Caractère : Le salaire
Nature du caractère : quantitatif continue.

(1)

2) Construire le tableau représentatif de la série :

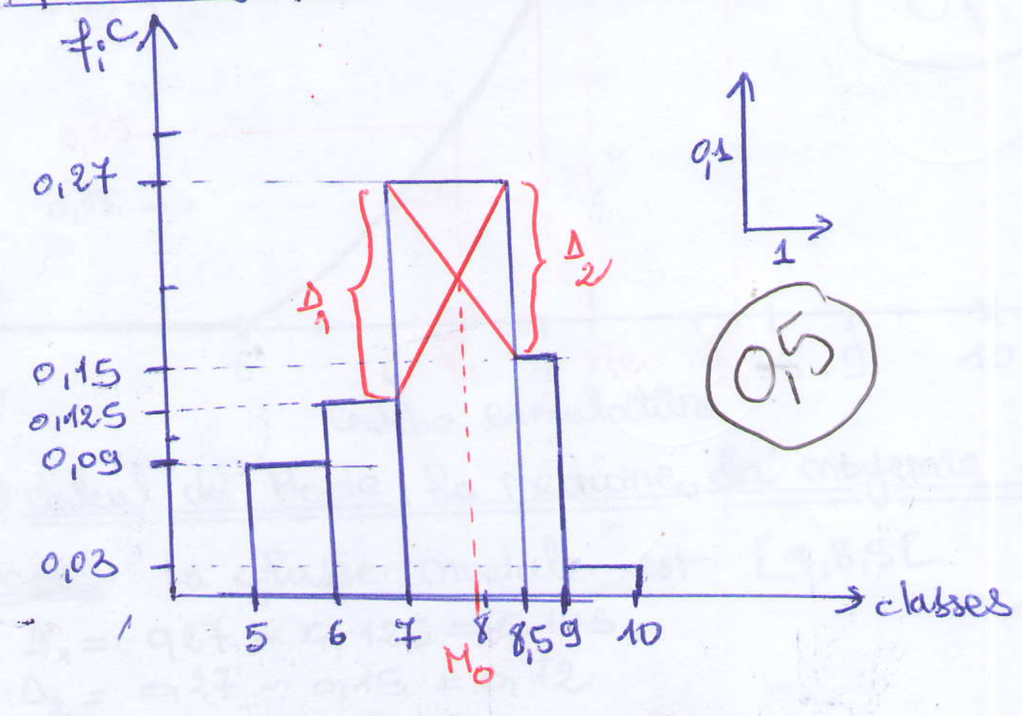
Classes	$F_i \nearrow$	f_i	a_i	f_i^c	x_i	$f_i x_i$
[5,6[	0,18	0,18	1	0,09	5,5	0,99
[6,7[	0,43	0,25	1	0,125	6,5	1,625
[7, 8.5[	0,79	0,36	1,5	0,27	7,75	2,79
[8.5, 9[	0,94	0,15	0,5	0,15	8,75	1,3125
[9, 10[	1	0,06	1	0,03	9,5	0,57
Total	/	1	/	/	/	7,2875

(1)

$f_i^c = \frac{f_i}{a_i} \times a$

$a = PGCD(a_i)$
$= PGCD(0,5, 1, 1,5)$
$= 0,5$

3) Représenter graphiquement la série :



(0,5)

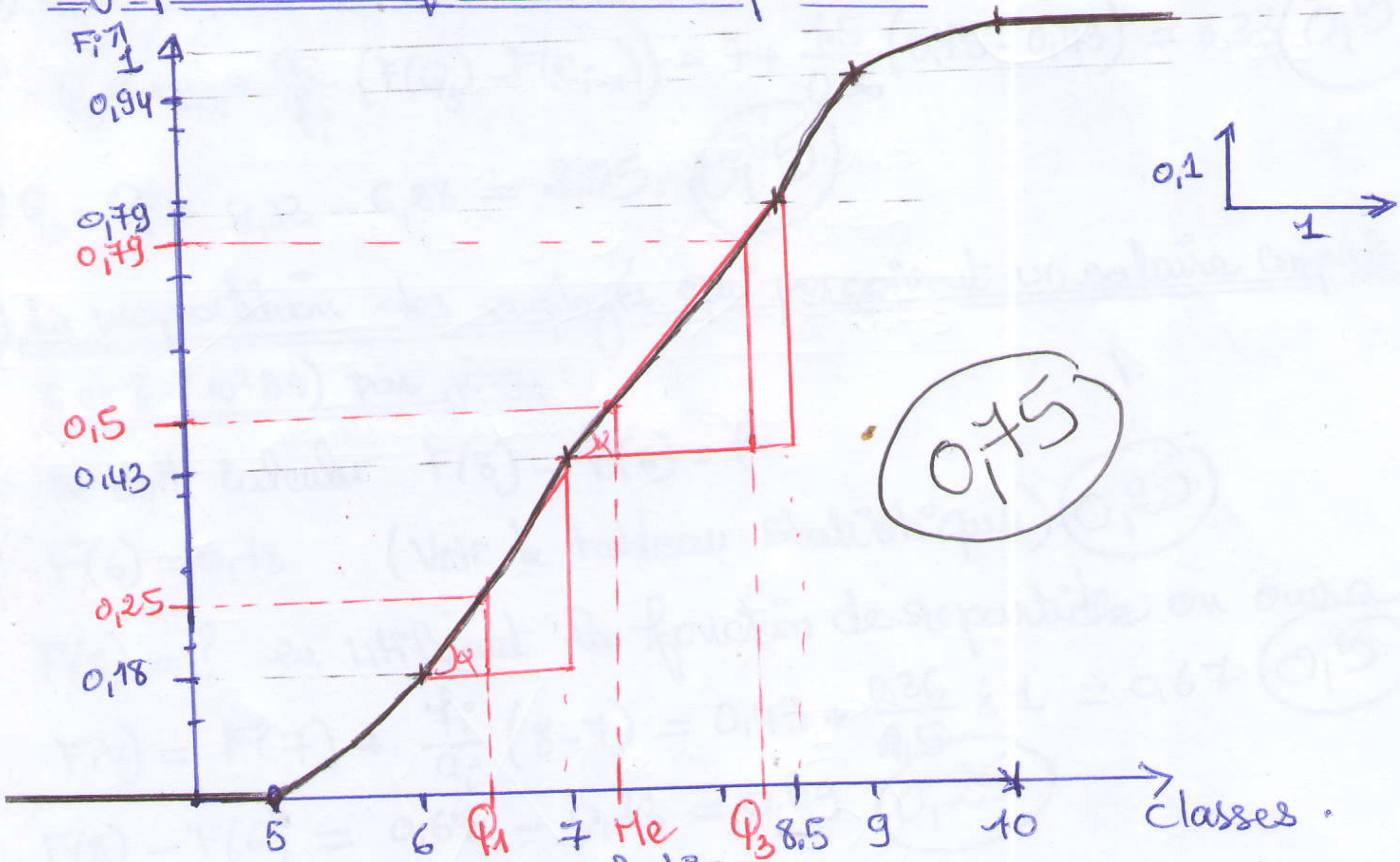
la fonction de répartition :

$$F(x) = \begin{cases} 0 & x < 5 \\ F(e_{i-1}) + \dfrac{f_i}{a_i}(x - e_{i-1}) & x \in [e_{i-1}, e_i[\\ 1 & x \geq 10 \end{cases}$$

$$= \begin{cases} 0 & x < 5 \\ 0{,}18\,(x-5), & 5 \leq x < 6 \\ 0{,}18 + 0{,}25(x-6), & 6 \leq x < 7 \\ 0{,}43 + 0{,}24(x-7), & 7 \leq x < 8.5 \\ 0{,}79 + 0{,}3\,(x-8.5), & 8.5 \leq x < 9 \\ 0{,}94 + 0{,}06\,(x-9), & 9 \leq x < 10 \\ 1 & x \geq 10 \end{cases}$$

1,25

Le graphe de la fonction de répartition :



Courbe cumulative

0,75

5) calcul du Mode, la mediane, la moyenne et les quartiles :

Mode : La classe modale est $[7, 8{,}5[$

$\Delta_1 = 0{,}27 - 0{,}125 = 0{,}145$

$\Delta_2 = 0{,}27 - 0{,}15 = 0{,}12$

P2

$$M_0 = e_{i-1} + a_i \frac{\Delta_1}{\Delta_1 + \Delta_2} = 7 + 1.5\ \frac{0,145}{0,145+0,12} = 7,82$$ 0,75

$$\boxed{M_0 = 7,82}$$

<u>La Mediane</u>: la classe médiane est $[7, 8.5[$

$$M_e = e_{i-1} + \frac{a_i}{f_i}\left(F(M_e) - F(e_{i-1})\right) = 7 + \frac{1.5}{0,36}(0,5 - 0,43) = 7,2916$$ 0,75

<u>Moyenne</u>: $\bar{x} = \sum_{i=1}^{5} f_i x_i = 7,2875$. 0,5

<u>L'intervalle interquartiles</u>:

- $Q_1 \in [6, 7[$

$$Q_1 = e_{i-1} + \frac{a_i}{f_i}\left(F(Q_1) - F(e_{i-1})\right) = 6 + \frac{1}{0,25}(0,25 - 0,18) = 6,28$$ 0,5

- $Q_3 \in [7, 8.5[$

$$Q_3 = e_{i-1} + \frac{a_i}{f_i}\left(F(Q_3) - F(e_{i-1})\right) = 7 + \frac{1.5}{0,36}(0,75 - 0,43) = 8,33$$ 0,5

- $Q_3 - Q_1 = 8,33 - 6,28 = 2,05$ 0,5

6) <u>La proportion des employés qui perçoivent un salaire compris entre 6 et 8 (10^2 DA) par jour</u>:

On doit calculer $F(8) - F(6) = ?$

$F(6) = 0,18$ (Voir le tableau statistique) 0,25

$F(8) = ?$ en utilisant la fonction de repartition on aura

$$F(8) = F(7) + \frac{f_i}{a_i}(8-7) = 0,43 + \frac{0,36}{1,5} \times 1 = 0,67$$ 0,5

$$F(8) - F(6) = 0,67 - 0,18 = 0,49$$ 0,25

Donc, il y'a 49 % des employés de l'usine ayant un salaire compris entre 6 et 8 (10^2 DA). 0,5

Exo2 : I).

x_i	13	13.4	14.5	15.2	16.5	16.5	17.2	17.4	18	18.3
y_i	40	40	50	50	50	60	60	70	70	60

1) Préciser la nature des variables statistiques étudiées :

X : variable statistique Quantitative Continue.
Y : " " " Discrète. 0,5

2) Calculer la moyenne de X et Y :

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{10}(13+13.4+\dots+18+18.3) = \frac{160}{10} = 16$$

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{10}(40+40+\dots+70+60) = \frac{550}{10} = 55$$

0,5

3) Calculer le coefficient de corrélation :

$$\rho = \frac{cov(X,Y)}{\sigma_X \times \sigma_Y}$$

0,25

$$V(X) = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{X}^2 = \frac{2591,84}{10} - (16)^2 = 259,184 - 256 = 3,184$$

$$V(Y) = \frac{1}{n}\sum_{i=1}^{n} y_i^2 - \bar{Y}^2 = \frac{31300}{10} - (55)^2 = 3130 - 3025 = 105$$

0,25

$$Cov(X,Y) = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{X}\bar{Y} = \frac{8964}{10} - (16)(55) = 896,4 - 880 = 16,4$$

$$Cov(X,Y) = 16,4$$ 0,5

0,25

$$\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{16,4}{\sqrt{3,184}\sqrt{105}} = \frac{16,4}{1,784 \times 10,247} = 0,897$$

$\rho = 0,897 \Rightarrow$ Donc il y a une forte corrélation linéaire entre X et Y. 0,25

4). Déterminer l'équation de la droite de Mayer :

○ $E_1 = \{(13,40), (13.4,40), (14.5,50), (15.2,50), (16.5,50)\}$

$G_1 = (\bar{X}_1, \bar{Y}_1) = ?$

$\bar{X}_1 = \frac{1}{5}(13 + 13,4 + 14,5 + 15,2 + 16.5) = \frac{72,5}{5} = 14,52$

$\bar{Y}_1 = \frac{1}{5}(40 + 40 + 50 + 50 + 50) = \frac{230}{5} = 46$

Donc $G_1(\bar{X}_1, \bar{Y}_1) = (14,52 , 46)$ 0,5

$E_2 = \{(16.5, 60), (17.2, 60), (17.4, 70), (18, 70), (18.3, 60)\}$

$G_2(\bar{X}_2, \bar{Y}_2) = ?$

$\bar{X}_2 = \frac{1}{5}(16.5 + 17.2 + 17.4 + 18 + 18.3) = \frac{87.4}{5} = 17.48$

$\bar{Y}_2 = \frac{1}{5}(60 + 60 + 70 + 70 + 60) = \frac{320}{5} = 64$

$G_2(\bar{X}_2, \bar{Y}_2) = (17.48 , 64)$ 0,5

La droite de Mayer est de la forme suivante

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

$G_1 \in (\Delta) \Rightarrow \bar{Y}_1 = a\bar{X}_1 + b$ --- ①

$G_2 \in (\Delta) \Rightarrow \bar{Y}_2 = a\bar{X}_2 + b$ --- ②

① - ② $\Rightarrow \bar{Y}_1 - \bar{Y}_2 = a(\bar{X}_1 - \bar{X}_2) \Rightarrow a = \frac{\bar{Y}_1 - \bar{Y}_2}{\bar{X}_1 - \bar{X}_2}$

$\Rightarrow a = \frac{46 - 64}{14,52 - 17,48} = \frac{-18}{-2,96} = 6,08$ 0,5

En remplaçant dans ① on obtient

① $\Rightarrow \bar{Y}_1 = a\bar{X}_1 + b \Rightarrow b = \bar{Y}_1 - a\bar{X}_1 = 46 - (6,08)(14,52)$

$b = 46 - 88,2816 = -42,2816.$ 0,5

La droite de Mayer est

$(\Delta) : Y = 6,08X - 42,2814$ 0,25

II. 1) Compléter le tableau :

X \ Y	40	50	60	70	$n_{i\cdot}$
[13,15[	2	1	0	0	3
[15,17[	0	2	1	0	3
[17,19[	0	0	2	2	4
$n_{\cdot j}$	2	3	3	2	10

0,75

2) Les deux distributions marginales de X et Y :

Selon X :

0,5

X	$n_{i\cdot}$	$f_{i\cdot}$
[13,15[	3	0,3
[15,17[	3	0,3
[17,19[	4	0,4
Total	10	1

Selon Y : 0,5

Y	$n_{\cdot j}$	$f_{\cdot j}$
40	2	0,2
50	3	0,3
60	3	0,3
70	2	0,2
Total	10	1

3) Les variables X et Y sont indépendantes $\Rightarrow \forall i,j \quad n_{ij} = \frac{n_{i\cdot} \times n_{\cdot j}}{n}$

$i=2$, $j=1$, $n_{21}=0$, $\frac{n_{2\cdot} \times n_{\cdot 1}}{n} = \frac{3 \times 2}{10} = 0,6$ 0,5

$\exists (i,j) = (2,1)$ tq $n_{21} \neq \frac{n_{2\cdot} \times n_{\cdot 1}}{n}$ $\Rightarrow$ X et Y ne sont pas indépendantes

4) Distributions Conditionnelles :

$Y/X<15 : \equiv Y / X \in [13,15[$

Y	n_{1j}	f_{1j}
40	2	0,67
50	1	0,33
60	0	0
70	0	0
Total	3	1

0,5

$$\bar{X}_1 = \frac{1}{n_{1\cdot}} \sum_{i=1}^{4} n_{1j} y_j = \frac{1}{3}(80+50+0+0)$$

$$= \frac{1}{3}(130) = 43,33. \quad 0,5$$

P6

5) <u>Déterminer l'équation de la droite de regression de X en Y</u>

La droite est de la forme $X = \alpha Y + \beta$

$\alpha = \frac{Cov(X,Y)}{V(Y)}$. $\beta = \bar{X} - \alpha \bar{Y}$

$\bar{X} = \frac{1}{n} \sum_{i=1}^{3} n_{i.} x_i = \frac{1}{10}(3\times14 + 3\times16 + 4\times18) = \frac{162}{10} = 16,2$ 0,25

$\bar{Y} = \frac{1}{n} \sum_{i=1}^{4} n_{.j} y_j = \frac{1}{10}(2\times40 + 3\times50 + 3\times60 + 2\times70) = \frac{550}{10} = 55$

$Cov(X,Y) = \frac{1}{n} \sum_{i=1}^{3}\sum_{j=1}^{4} n_{ij} x_i y_j - \bar{X}\bar{Y}$

$= \frac{1}{10}(1120 + 700 + 0 + 0 + 0 + 1600 + 960 + 0 + 0 + 0 + 2160 + 2520)$
$- (16,2)(55) = 15$ 0,5

$V(Y) = 105$

$\alpha = \frac{Cov(X,Y)}{V(Y)} = \frac{15}{105} = 0,143$ 0,25

$\beta = \bar{X} - \alpha\bar{Y} = 16,2 - (0,143)(55) = 8,335$ 0,25

La forme de la droite de régression est

$(\Delta') : X = 0,143 Y + 8,335$. 0,25

6) Peut-on prévoir la production de blé pour 75 jours de pluies ?

<u>Réponse</u>: Oui on peut prévoir puisque il y'a une forte Corrélation entre X et Y. 0,5

$Y = 75$ jours.

$X = 0,143 Y + 8,335 = 0,143(75) + 8,335$
$= 19,06$. 0,5