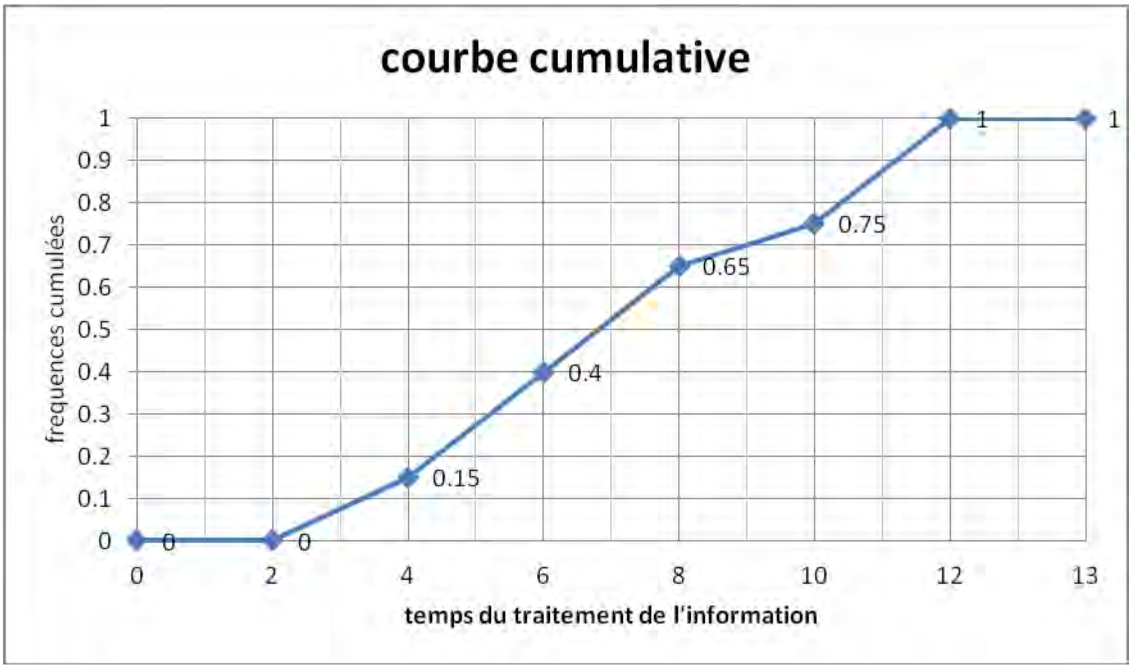
Université AboubekrBelkaid Tlemcen | **2016/2017**
Faculté des sciences | **Cours : statistiques**
Département de mathématiques | **MI**
Durée 1h30

Contrôle en statistiques
09/03/2017

Exercice N°1 (7pts)

On veut comparer l'efficacité du traitement de l'information de plusieurs logiciels pour cela on compare le temps (en centièmes de seconde) que met chaque logiciel à traiter une information commune donnée. On obtient la courbe cumulative suivante :



1. Déterminer la population , la variable X et son type
2. Donner le tableau statsitique associé
3. Calculer le 1ére quartile.
4. Déterminer la proportion des logiciels dont le temps du traitement de l'information est inférieur ou égal a 5.5centièmesde seconde.
5. Calculer la moyenne.

Exercice N°2 (5pts)

On teste un nouveau traitement (médicament) sur 100 malades, on demande à ces malades d'énumérer (donner le nombre) des effets indésirables(nausée, fièvre, douleur, sommeil) qu'ils pensent avoir eu pendant la période de traitement. On pose X la variable statistique qui représente le nombre de ces effets indésirables ainsi on obtient le tableau suivant :

xi	0	1	2	3	4
ni	33	22	23	12	10

1. Calculer le mode et la médiane que représente ces valeurs par rapport a notre étude
2. Déterminer la fonction cumulative , tracer son graphe

Exercice N°3 (Questions de cours) (8pts)

I] Dites si les assertions suivantes sont vraies ou fausses tout en justifiant vos réponses :

- Soit la variables X qui a pour valeurs 100,95,90,85,…,10,5. Alors son écart type σ =-5

- La fonction F(x)= e^x $\forall x \in \mathbb{R}$ définit bien une fonction cumulative (fonction de répartition)
- Dans une population de taille N=10, le nombre de classes possible k en utilisant la formule de Sturge est de k=5

II] Démontrez ce qui suit:

- $\frac{1}{N}\sum_{i=1}^{k} n_i(x_i - \overline{X}) = 0$

- Cov(aX, aY)= a²cov(X,Y)

III] Donnez les formules de a et b afin de déterminer la droite de régression y=ax+b

Bon Courage

2016/2017

MI

Solution Proposée
Pour le contrôle de Statistiques

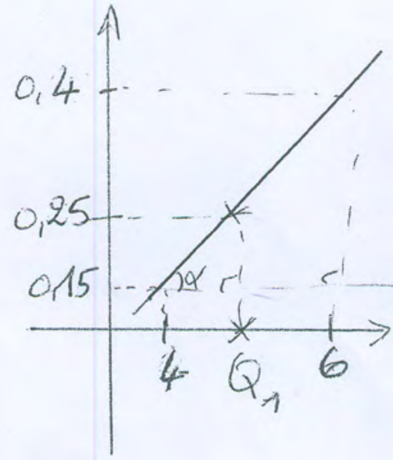
Exo1 :

1) la Population $\longrightarrow$ les logiciels
la variable X $\longrightarrow$ Temps de Traitement de l'informati
type $\longrightarrow$ quantitatif continu.

2) Tableau

C_i	[2,4[	[4,6[	[6,8[	[8,10[	[10,12[	Σ
f_{ic}	0,15	0,4	0,65	0,75	1	/
f_i	0.15	0,25	0,25	0,1	0,25	1
x_i	3	5	7	9	11	/
$f_i \times x_i$	0.45	1,25	1,75	0,9	2,75	

3) le 1er quartile Q_1



$$tg\alpha = \frac{0,4-0,15}{6-4} = \frac{0,25}{2} = 0,125$$

$$tg\alpha = \frac{0,25-0,15}{Q_1-4} = \frac{0,1}{Q_1-4}$$

$$tg\alpha = tg\alpha \Leftrightarrow \frac{0,1}{Q_1-4} = 0,125$$

$$\Rightarrow Q_1 = 4 + \frac{0.1}{0.125} = 4,8$$

4) Fonction cumulative (ou Fonction de répartition)

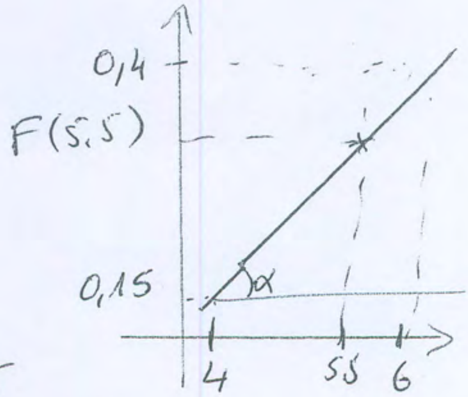
$$F : \mathbb{R} \longrightarrow [0,1]$$
$$x \longmapsto F(x) = P(X \leq x)$$ qui représente

la Proportion des individus dont la valeur de la variable X est $\leq$ à x

donc $F(5,5) = P(X \leq 5,5)$

$$tg\alpha = \frac{0{,}4-0{,}15}{6-4} = \frac{0{,}25}{2} = 0{,}125$$

$$tg\alpha = \frac{F(5.5)-0{,}15}{5.5-4} = \frac{F(5.5)-0{,}15}{1{,}5}$$



$$tg\alpha = tg\alpha \Rightarrow \frac{F(5.5)-0{,}15}{1{,}5} = 0{,}125$$

$$\Rightarrow F(5.5) = 0{,}15 + 0{,}125 \times 1{,}5$$

$$\Rightarrow F(5.5) = 0{,}3375 \simeq 33{,}75\%$$

5) La moyenne $\overline{x}$

$$\overline{x} = \sum_i f_i \times x_i = 0{,}45 + . + 2{,}75 = 7{,}1.$$

Exo2 :

1)

x_i	0	1	2	3	4
n_i	33	22	23	12	10
n_{ic}	33	55	78	90	100
f_{ic}	0,33	0,55	0,78	0,9	1

Le mode M_o

le + grand effectif $n_i = 33 \Rightarrow \boxed{M_o = 0}$

Le mode, nous renseigne que pour cette population la majorité des malades ont eu "0" effets indésirables à ce médicament (ou traitement).

La médiane M_e

le rang $\longrightarrow R_g(M_e) = \frac{N+1}{2} = \frac{100+1}{2} = 50{,}5$

$\hookrightarrow x_{50} < M_e < x_{51} \Rightarrow M_e = \frac{x_{50}+x_{51}}{2}$

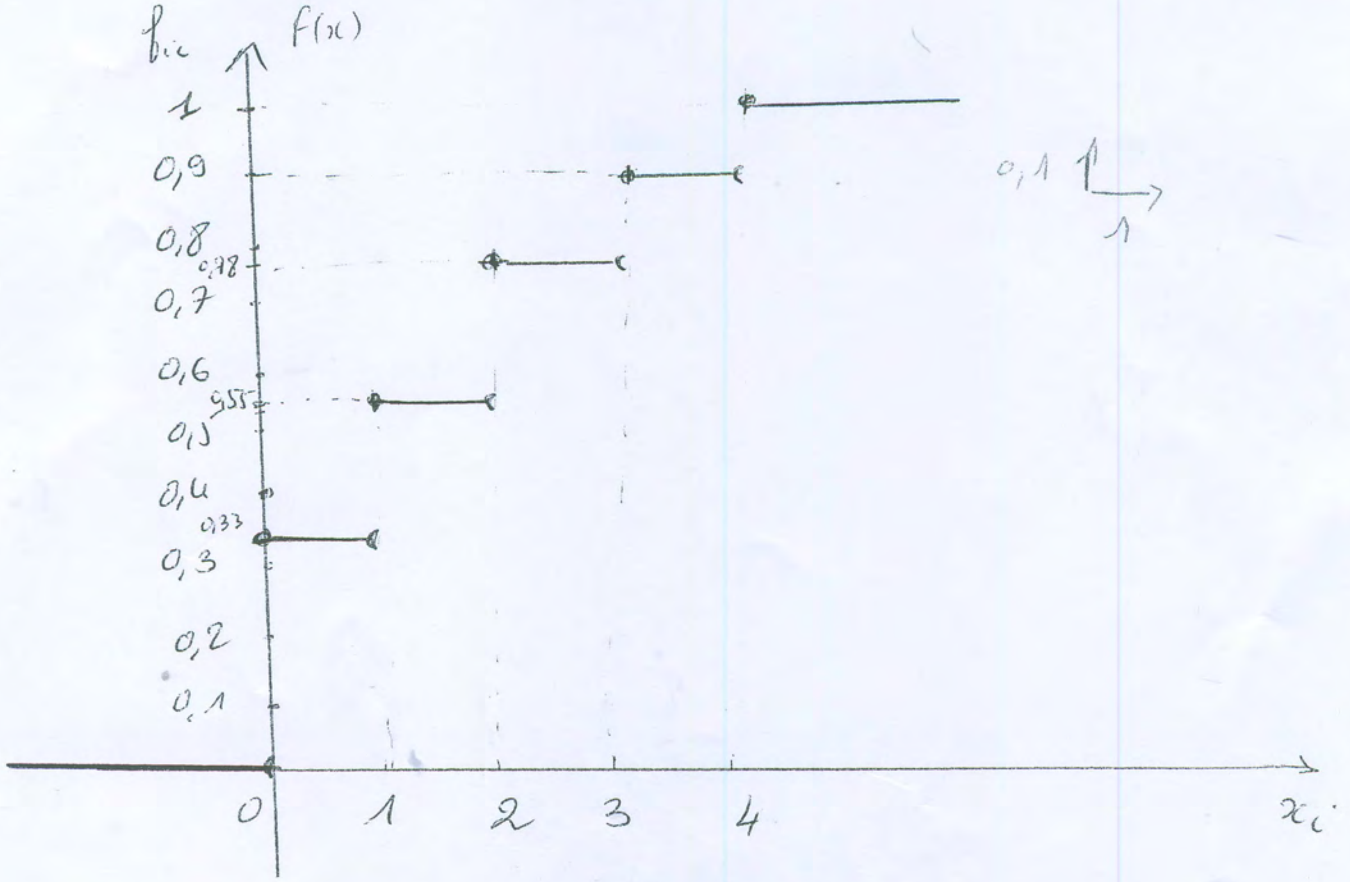
Or $x_{50} = x_{51} = 1 \quad \Rightarrow \quad M_e = \frac{1+1}{2} = 1$

La mediane Me , nous informe qu'arrivé à 50% de notre population (la moitie) on a 1 seul effet indésirable constaté

2) Fonction cumulative

$$F : \mathbb{R} \longrightarrow [0,1]$$
$$x \longmapsto F(x) = P(X \leq x) = \sum_{x_i \leq x} f_i$$

$$F(x) = \begin{cases} 0 & \text{Si } x < 0 \\ 0,33 & \text{Si } 0 \leq x < 1 \\ 0,55 & \text{Si } 1 \leq x < 2 \\ 0,78 & \text{Si } 2 \leq x < 3 \\ 0,9 & \text{Si } 3 \leq x < 4 \\ 1 & \text{Si } x \geq 4 \end{cases}$$



Courbe cumulative

Exo3

I]

1) Faux : car $\sigma_X = \sqrt{V(X)} \geq 0$ toujours positif

2) Faux : car $F(x) \in [0,1]$ donc $\lim_{x \to +\infty} F(x) = 1$

or $\lim_{x \to +\infty} F(x) = \lim_{x \to +\infty} e^x = +\infty \neq 0$

donc $F(x)$ n'est pas une fonction de répartition

3) Vrai Sturge : $k = 1 + 3{,}3 \log_{10} N = 4{,}3 \simeq 5$ car $k \in \mathbb{N}^*$

II]

1) $\frac{1}{N} \sum_i n_i (x_i - \overline{x}) = \frac{1}{N} \sum_i (n_i x_i - n_i \overline{x})$

$= \frac{1}{N} \left(\sum_i n_i x_i - \sum n_i \overline{x} \right)$

$= \frac{1}{N} \sum_i n_i x_i - \frac{1}{N} \overline{x} \sum n_i$ or $\sum n_i = N$

$= \frac{1}{N} \sum_i n_i x_i - \frac{1}{N} \overline{x} \cdot N$

$= \overline{x} - \overline{x} = 0$ CQFD

2) $\mathrm{cov}(aX, aY) = \overline{(aX)(aY)} - \overline{aX} \cdot \overline{aY}$

$= \overline{a^2 XY} - a \overline{X} \cdot a \overline{Y}$ linéarité

$= a^2 \overline{XY} - a^2 \overline{X} \cdot \overline{Y}$

$= a^2 (\overline{XY} - \overline{X} \cdot \overline{Y})$

III] Formule de la droite de régression

$(\Delta) : y = ax + b$

$a = \frac{\mathrm{cov}(X,Y)}{V(X)}$ $b = \overline{Y} - a \overline{X}$.