Examen de Statistique et Probabilités

Université : Laghouat
Département : Sciences et Techniques
Section : $2^{ème}$ Année ST, Semestre 3, 2017-2018
Examen préparé par : Dr. Mohand Bentobache
Date : 27/01/2018. **Durée** : 1h30

EXERCICE 1 (10pts) :

Les poids des employés d'une entreprise (en Kg) sont distribués de la manière suivante :

Poids en (Kg)	[50, 55[	[55, 57[	[57, 59[	[59, 61[	[61,63[
Nombre d'employés	15	34	26	58	36

1- Représenter cette série par un histogramme. **(2.5 pts)**
2- Calculer les effectifs cumulés décroissants (N_i^-) et représentez-les graphiquement. **(2.5 pts)**
3- Calculer le mode et la médiane analytiquement. **(3pts)**
4- Calculer le poids moyen des employés (moyenne arithmétique $\bar{X}$) et la variance $V(X)$. **(2pts)**

EXERCICE 2 (5pts) : On donne le tableau d'observations suivant :

x_i	5	13	11	17
y_i	9	4	7	3

1- Trouver la droite d'ajustement de y en x par la méthode des moindres carrées. **(3.5pts)**
2- Déterminer le coefficient de corrélation linéaire et déduire la qualité de l'ajustement. **(1.5pts)**

EXERCICE 3 (5pts) : Soient β et α des nombres réels strictement positifs et f la fonction définie sur $\mathbb{R}$ par

$$f(x) = \begin{cases} 0, & \text{si } x < \beta\,; \\ \alpha\frac{\beta^\alpha}{x^{\alpha+1}}, & \text{si } x \geq \beta. \end{cases}$$

1- Démontrer que f est bien une densité de probabilité. **(1pt)**
Soit X une variable aléatoire de densité f. On dit que X suit une loi de Pareto de paramètre (α, β).
2- Déterminer la fonction de répartition de X. **(2.25pts)**
3- On suppose que $\alpha > 2$. Calculer l'espérance de X. **(1.75pts)**

Bonne chance

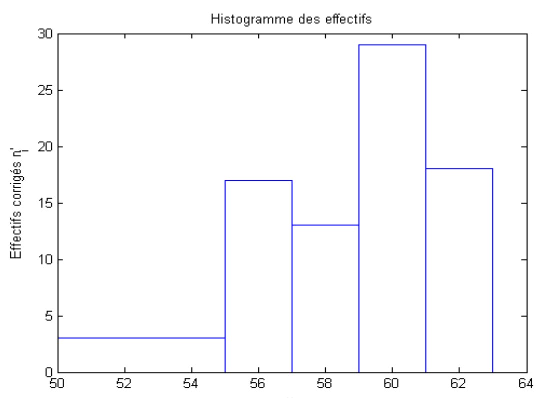
Corrigé-Type de l'Examen de Statistique et Probabilités

Université : Laghouat. **Département** : ST. **Section** : 2ème Année ST, Semestre 3, 2017-2018.
Préparé par : Dr. Mohand Bentobache. **Date** : 27/01/2018.

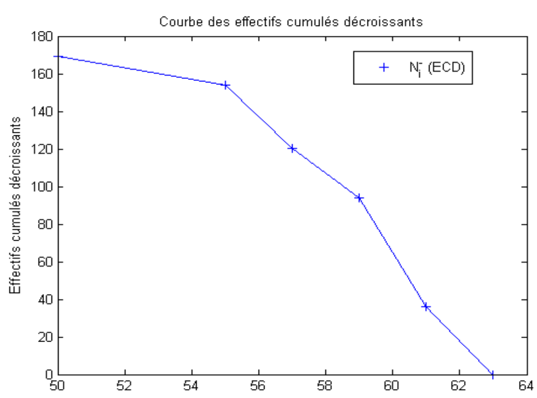
EXERCICE 1 (10pts) :
Le tableau statistique est :

Classes				**0.5 pts**		**0.5 pts**			
$[e_{i-1}$	$e_i[$	n_i	a_i	n'_i	N_i^+	N_i^-	x_i	$n_i x_i$	$n_i x_i^2$
[50	55[	15	5	3	15	169	52,5	787,50	41343,75
[55	57[	34	2	17	49	154	56	1904	106624
[57	59[	26	2	13	75	120	58	1508	87464
[59	61[	58	2	29	133	94	60	3480	208800
[61	63[	36	2	18	169	36	62	2232	138384
		169						9911,50	582615,75

1) Représentation graphique des effectifs par un histogramme (**2pts)** :



2) Représentation graphique des effectifs cumulés (**2pts)** :



3) Calcul du mode et de la médiane analytiquement :

▶ Comme les amplitudes sont différentes, l'effectif corrigé maximum est $n'_i = n'_4 = 29$.

La classe modale est $CMo = [e_{i-1}, e_i[= [59, 61[$. **(0.25pts)**

L'effectif corrigé et l'amplitude de la classe modale sont : $n'_i = 29$ et $a_i = 2$.

Les distances d_1 et d_2 sont : $d_1 = n'_i - n'_{i-1} = 16$ et $d_2 = n'_i - n'_{i+1} = 11$.

Le mode est alors :

$$Mo = e_{i-1} + a_i \times \frac{d_1}{d_1 + d_2} = 60.19. \text{ (\textbf{1.25pts})}$$

▶ On a

$$N^+_{i-1} = N^+_3 = 75 \leq \frac{N}{2} = 84.50 \leq N^+_i = N^+_4 = 133.$$

Donc la classe médiane est : $CMe = [e_{i-1}, e_i[= [59, 61[$. **(0.25pts)**

L'effectif et l'amplitude de la classe médiane sont : $n_i = 58$ et $a_i = 2$.

La médiane est alors :

$$Me = e_{i-1} + a_i \times \frac{\frac{N}{2} - N^+_{i-1}}{n_i} = 59.33. \text{ (\textbf{1.25pts})}$$

4) Calcul de la moyenne, la variance et l'écart-type :

▶ La moyenne de la variable statistique X est : $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} n_i x_i = 58.65$. **(1pts)**

▶ La variance de la variable statistique X est : $V(X) = \frac{1}{n}\sum_{i=1}^{n} n_i x_i^2 - \bar{x}^2 = 7.85$. **(1pts)**

EXERCICE 2 (5pts) :

1) Calcul des coefficients de la droite de regression (Δ) :

x_i	y_i	x_i^2	y_i^2	$x_i y_i$
5	9	25	81	45
13	4	169	16	52
11	7	121	49	77
17	3	289	9	51
46	23	604	155	225

▶ La moyenne de la variable statistique x est : $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 11.50$. **(0.5pts)**

▶ La moyenne de la variable statistique y est : $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = 5.75$. **(0.5pts)**

▶ La variance de la variable statistique x est : $V(x) = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2 = 18.75$. **(0.5pts)**

▶ La variance de la variable statistique y est $V(y) = \frac{1}{n}\sum_{i=1}^{n} y_i^2 - \bar{y}^2 = 5.69$. **(0.5pts)**

▶ La covariance des de x et y est : $\text{Cov}(x, y) = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y} = -9.88$. **(0.5pts)**

▶ La pente de la droite de regression de y en x est : $a = \frac{Cov(x,y)}{V(x)} = -0.53$. **(0.5pts)**

▶ Le point d'intersection de la droite (Δ) avec l'axe des y est : $b = \bar{y} - a\bar{x} = 11.81$. **(0.5pts)**

▶ L'équation de la droite de regression de y en x est : $y = -0.53x + 11.81$.

2) Calcul du coefficient de corrélation linéaire :

▶ L'écart-type la variable statistique x est : $\sigma_x = \sqrt{V(x)} = 4.33$.

▶ L'écart-type la variable statistique y est : $\sigma_y = \sqrt{V(y)} = 2.38$.

▶ Le coefficient de corrélation linéaire est : $\rho = \frac{Cov(x,y)}{\sigma_x \sigma_y} = -0.96$. **(1pts)**

▶ Le coefficient ρ est proche de -1, donc l'ajustement linéaire est bon. **(0.5pts)**

Exercice 03 (05 pts)
1) La fonction f est positive ou nulle sur $\mathbb{R}$, continue sauf en β. On a, pour tout $x \geq \beta$,

$$\int_{-\infty}^{+\infty} f(t)dt = \lim_{x\to+\infty} \int_{\beta}^{x} f(t)dt$$

$$= \lim_{x\to+\infty} \int_{\beta}^{x} \alpha \frac{\beta^{\alpha}}{t^{\alpha+1}} dt$$

$$= \lim_{x\to+\infty} [-\beta^{\alpha} t^{-\alpha}]_{\beta}^{x} = \lim_{x\to+\infty} -\beta^{\alpha} x^{-\alpha} + 1.$$

On a $\lim_{x\to+\infty} x^{-\alpha} = 0$, car $\alpha > 0$. Donc $\int_{-\infty}^{+\infty} f(t)dt = 1$. **(1pt)**
2) Fonction de répartition de X :
La fonction de répartition de X est donnée par :

$$F(x) = \int_{-\infty}^{x} f(t)dt. \text{ (0.25pt)}$$

Cas 1 : pour $x < \beta$, $F(x) = 0$. **(0.5pt)**
Cas 1 : pour $x \geq \beta$, on a

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{\beta}^{x} f(t)dt = \int_{\beta}^{x} \alpha \frac{\beta^{\alpha}}{t^{\alpha+1}} dt$$

$$= [-\beta^{\alpha} t^{-\alpha}]_{\beta}^{x} = -\beta^{\alpha} x^{-\alpha} + 1 = 1 - (\frac{\beta}{x})^{\alpha}. \text{ (1.5pt)}$$

D'où

$$F(x) = \begin{cases} 0, & x < \beta\,; \\ 1 - (\frac{\beta}{x})^{\alpha}, & x \geq \beta. \end{cases}$$

3) On suppose que $\alpha > 2$:
On a

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx, \text{ (0.25pt)}$$

$$E(X) = \lim_{x\to+\infty} \int_{\beta}^{x} t f(t)dt = \lim_{x\to+\infty} \int_{\beta}^{x} \alpha t \frac{\beta^{\alpha}}{t^{\alpha+1}} dt$$

$$= \alpha\beta^{\alpha} \lim_{x\to+\infty} \int_{\beta}^{x} \frac{1}{t^{\alpha}} dt = \alpha\beta^{\alpha} \lim_{x\to+\infty} [\frac{1}{-\alpha+1} t^{-\alpha+1}]_{\beta}^{x}.$$

$$= \frac{-\alpha}{\alpha-1} \lim_{x\to+\infty} [\frac{\beta^{\alpha}}{t^{\alpha-1}}]_{\beta}^{x} = \frac{-\alpha}{\alpha-1} \lim_{x\to+\infty} [\frac{\beta^{\alpha}}{x^{\alpha-1}} - \frac{\beta^{\alpha}}{\beta^{\alpha-1}}].$$

$$E(X) = \frac{\alpha\beta}{\alpha-1}. \text{ (1.5pt)}$$

Remarques :

- Les formules sont notées sur 0.25, si l'étudiant a bien compris comment les utiliser.
- On donne (0.5 pts) si l'on représente l'histogramme avec les n_i au lieu des n'_i.
- On donne (0.5 pts) pour la courbe des effectifs cumulés s'il l'on constate un décalage.