Université de Tlemcen
Faculté des Sciences
Département de Mathématiques
1ère Année LMD —Math et M.I—
Année Universitaire 2022 / 2023

Examen Final de Statistique

Exercice 1 : (4 Pts)

Soit un caractère discret X, définie sur une population Ω, qui prend les modalités $x_1, \dots, x_n$ avec leurs fréquences partielles respectives $f_1, \dots, f_n$. On pose $\bar{x}$ et σ_X respectivement la moyenne et l'écart type du caractère X. Calculer la moyenne et la variance du caractère Y définie par : $Y(\omega) = \frac{X(\omega)-\bar{x}}{\sigma_X}$ (c'est-à-dire pour tout $i = 1, \dots n$: $y_i = \frac{x_i-\bar{x}}{\sigma_X}$).

Exercice 2 : (10 Pts)

Soit S, le caractère continu, égal à la note (sur 20) obtenue par chacun des 450 étudiants lors d'un examen final de statistique. Les valeurs prises par S avec leurs fréquences partielles respectives, sont données par le tableau suivant :

Classe	[0 ; 5[	[5 ; 10[	[10 ; 15[	[15 ; 20[
f_i	0,15	0,30	0,45	0,10

1) Compléter le tableau :

Classe	c_i Centre de Classe	f_i fréquence partielle	F_i fréquence cumulée

2) Calculer le nombre d'étudiants dont les notes se trouvent dans l'intervalle $[Q_1 ; Q_3]$

3) Déterminer la valeur de a pour que 35% des étudiants aient une note dans l'intervalle $[a ; Q_3]$

4) Déterminer la classe de la médiane et calculer la valeur de la médiane M_e

5) Déterminer la classe modale C_{M_o} et calculer le mode M_o

Exercice 3 : (Questions de cours) (6 Pts) [Les parties : I) et II) sont indépendantes]

I) Soit Y, un caractère continu, définie sur une population Ω de 50 individus.

$y_{min} = 140$ et $y_{max} = 171$. En utilisant la méthode de STURGE, répartir les valeurs du caractère Y en classes.

II) Soit le couple de caractère (X, T) qui prend les valeurs (x_i, t_j) ; pour $i = 1, \dots, k$ et $j = 1, \dots, l$; avec leurs fréquences partielles f_{ij}

a) Pour tout $j = 1, \dots, l$ on pose $\bar{x}_j = \sum_{i=1}^{i=k} \frac{f_{ij}}{f_{\cdot j}} x_i$. Démontrer que $\bar{x} = \sum_{j=1}^{j=l} f_{\cdot j}\, \bar{x}_j$

b) On suppose que pour tout $i = 1, \dots, k$ et $j = 1, \dots, l$: $f_{ij} = f_{i\cdot} f_{\cdot j}$

Calculer la valeur du coefficient de corrélation linaire $\rho_{X,T}$

NB :(Important pour la notation) **Toutes vos réponses doivent etre argumentées et justifiées**

- Corrigé succint de l'épreuve finale de Statistique -

Exercice 1 : 4 pts

Le caractere discret X prend les valeurs $x_1, \dots, x_n$ avec leurs frequences partielles respectives $f_1, \dots, f_n$

or $y_i = \frac{x_i - \bar{x}}{\sigma_x}$; d'où la frequence partielle de y_i est f_i

$$\left(\mathrm{Card}\{\omega \in \Omega : X(\omega) = x_i\} = \left\{\omega \in \Omega : \underbrace{\frac{X(\omega) - \bar{x}}{\sigma_x}}_{Y(\omega)} = \underbrace{\frac{x_i - \bar{x}}{\sigma_x}}_{y_i}\right\}\right.$$

d'où le tableau

y_i	y_1	- - - -	y_i	- - -	y_n
f_i	f_1		f_i		f_n

Or $\bar{x} = \sum f_i x_i$ 0,5pt et $\sigma_x = \sqrt{\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}$ 0,5pt

Or $\bar{y} = \sum_{i=1}^{n} f_i y_i = \sum_{i=1}^{n} f_i \frac{x_i - \bar{x}}{\sigma_x} = \frac{1}{\sigma_x}\left[\sum f_i x_i - \sum f_i \bar{x}\right]$

$\Rightarrow \bar{y} = \frac{1}{\sigma_x}\left[\bar{x} - \bar{x}\right] = 0$ 1pt

Par ailleurs

$\mathrm{Var}(Y) = \sum f_i (y_i - \bar{y})^2$ 0,5pt

$\Rightarrow \mathrm{Var}(Y) = \sum f_i \left(\frac{x_i - \bar{x}}{\sigma_x}\right)^2 = \frac{1}{\sigma_x^2} \sum f_i (x_i - \bar{x})^2$

$\Rightarrow \mathrm{Var}(Y) = \frac{1}{\sigma_x^2} \mathrm{Var}(X) = \frac{\mathrm{Var}(X)}{\mathrm{Var}(X)} = 1$

1,5pt

I/4

ercice 2 (10pts)

1°) Dressons le tableau

classes	[0;5[	[5;10[	[10;15[	[15;20[
c_i	2,5	7,5	12,5	17,5
f_i	0,15	0,30	0,45	0,10
F_i	0,15	0,45	0,90	1,00

1pt

2°) On sait (d'après le cours) que $F_X(b) - F_X(a)$ = pourcentage des individus (ω) tel que $X(\omega) \in [a;b]$ 0,5pt

or $F_X(Q_3) - F_X(Q_1) = 0,75 - 0,25 = 0,5$ 0,5pt

d'où 50% des etudiants ont leurs notes dans l'intervalle $[Q_1; Q_3]$ 0,5pt

$\Longrightarrow \frac{450}{2} = 225$ etudiants ont leurs notes dans l'intervalle $[Q_1; Q_3]$ 0,5pt

3°) Pour les mêmes raisons que 2°)

$F_X(Q_3) - F_X(a) = 0,35 \Longrightarrow F_X(a) = 0,75 - 0,35 = 0,40$ 0,5pt

D'après le tableau $a \in [5;10[$ 0,5pt

$$\Longrightarrow a = 5 + 5\,\frac{0,40 - 0,15}{0,45 - 0,15} = 5 + 5\,\frac{0,25}{0,30} = 9,167$$ 1pt

4°) Comme (d'après le tableau) :

$F_X(10) = 0,45$ et $F_X(15) = 0,90$; Alors $Me \in [10;15[$ 1pt

or Me vérifie $F_X(Me) = 0,5$ 0,5pt

d'où $Me = 10 + 5\,\frac{0,50 - 0,40}{0,90 - 0,45} = 10,55$ 1pt

II / 4

D'après le tableau ; $C_{Mo} = [10 ; 15[$ (C'est la classe qui a la plus grande fi) 0,5 pt

or $M_o = a_i + h \frac{\Delta_1}{\Delta_1 + \Delta_2}$ avec $a_i = 10$; $h = 5$; 1 pt

$\Delta_1 = 0,45 - 0,30$ et $\Delta_2 = 0,45 - 0,10$

$\Rightarrow M_o = 10 + 5 \frac{0,15}{(0,15) + (0,35)} = 11,5$ 1 pt

Exercice 3 (Questions de cours) 6 pts

I) Soit e = etendue = $y_{max} - y_{min} = 171 - 140 = 31$ 0,5 pt

D'après STURGE n = le nombre de classe

$\simeq 1 + 3,3 \log_{10} n$ (où $n = 50$)

$\simeq 6,64$ 0,5 pt

on prend $n = 7$ (par exemple)

or h (longueur d'une classe) $> \frac{e}{n} = \frac{31}{7} = 4,42$

on prend $h = 4,5$ (par exemple) 0,5 pt

d'où les classes sont : 0,5 pt

$[140 ; 144,5[$; $[144 ; 149[$; $[149 ; 153,5[$; $[153,5 ; 158[$

$[158 ; 162,5[$; $[162,5 ; 167[$; $[167 ; 171,5[$.

III/a

a) On sait que $\bar{x} = \sum_{i=1}^{k} f_{i\cdot} x_i$ avec $f_{i\cdot} = \sum_{j=1}^{l} f_{ij}$ (0,5pt)

$\Rightarrow \bar{x} = \sum_{i=1}^{k} \sum_{j=1}^{l} f_{ij} x_i = \sum_{j=1}^{l} \sum_{i=1}^{k} \frac{f_{ij}}{f_{\cdot j}} f_{\cdot j} x_i$ (0,5pt)

$\Rightarrow \bar{x} = \sum_{j=1}^{l} f_{\cdot j} \underbrace{\sum_{i=1}^{k} \frac{f_{ij}}{f_{\cdot j}} x_i}_{\bar{x}_j}$ (0,5pt)

$\Rightarrow \bar{x} = \sum_{j=1}^{l} f_{\cdot j} \bar{x}_j$ (0,5pt)

b) On a (d'après le cours) :

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y}$$ (0,5pt)

or $cov(X,Y) = \sum_{i=1}^{k} \sum_{j=1}^{l} f_{ij} (x_i - \bar{x})(y_j - \bar{y})$

avec $\bar{x} = \sum_{i=1}^{k} f_{i\cdot} x_i$; $\bar{y} = \sum_{j=1}^{l} f_{\cdot j} y_j$;

$f_{i\cdot} = \sum_{j=1}^{l} f_{ij}$ et $f_{\cdot j} = \sum_{i=1}^{k} f_{i\cdot}$ (0,5pt)

On a par hypothèse que $f_{ij} = f_{i\cdot} f_{\cdot j}$

$\Rightarrow cov(X,Y) = \sum_{i=1}^{k} \sum_{j=1}^{l} f_{i\cdot} f_{\cdot j} (x_i - \bar{x})(y_j - \bar{y})$

$= \underbrace{\sum_{i=1}^{k} f_{i\cdot} (x_i - \bar{x})}_{0} \cdot \underbrace{\sum_{j=1}^{l} f_{\cdot j} (y_j - \bar{y})}_{0}$

$\Rightarrow cov(X,Y) = 0$ (0,5pt)

$\Rightarrow \rho_{X,Y} = 0$ (0,5pt)

IV / 4