

## Epreuve finale de Statistiques

### Exercice 1: (4 pts)

Le caractère  $X$  représente le parcours (en km) par jour d'un taxi.

Les valeurs prises par  $X$ , réparties en classes avec leurs fréquences cumulées sont données par le tableau suivant:

$C_i$	[75, 80[	[80, 85[	[85, 90[	[90, 95[	[95, 100[	[100, 105[	[105, 110[
$F_i$	0,05	0,15	0,35	0,71	0,86	0,94	1

- 1) Déterminer la fonction de répartition  $F$  de  $X$
- 2) Déterminer le quantile  $Q_3$  et donner son interprétation
- 3) Calculer le pourcentage des taxis qui parcourent entre 86 et 105 km par jour.

### Exercice 2: (8 pts)

La v.s  $Z = (X, Y)$  est définie par le tableau de contingence des effectifs partiels  $n_{ij}$  suivant

$x_i$	2	8	12	18
$y_j$				
6	8	1	1	0
9	1	10	2	0
11	1	2	14	1
14	0	0	2	7

- 10) Calculer  $\bar{x}, \bar{y}, \sigma_x, \sigma_y$
- 20) Calculer  $\rho$  le coefficient de corrélation entre  $X$  et  $Y$
- 30) Déterminer la droite de régression de  $Y$  en  $X$  et estimer la valeur de  $Y$  si  $X = 4$

### Exercice 3 : (5pts)

On lance un dé à six faces numérotées de 1 à 6.

On note  $p_i$  la probabilité de la face  $i$ .

Ce dé est truqué de telle sorte que les probabilités de sortie des faces sont:

$$P_1 = 0,1 ; P_2 = 0,2 ; P_3 = 0,3 ; P_4 = 0,1 ; P_5 = 0,15$$

Quelle est la probabilité ?

a) d'obtenir 6

b) d'avoir un nombre pair

On lance 5 fois ce dé équilibré

Quelle est la probabilité ?

c) d'avoir au moins un "3"

d) exactement deux "4"

### (Q.C.) Exercice 4 : (3pts)

Soit  $Z = (x_i, y_i)$   $i = \{1, \dots, N\}$  une variable

statistique double.

$$\text{Montrer que } \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{N} \sum_{i=1}^N x_i y_i - \bar{x} \bar{y}$$

---

Bon courage.

Toute réponse doit être justifiée

Ex 1

10)

$C_i$	$[75, 80[$	$[80, 85[$	$[85, 90[$	$[90, 95[$	$[95, 100[$	$[100, 105[$	$[105, 110[$	
$f_i$	0,05	0,1	0,2	0,36	0,15	0,08	0,06	
$F_i$	0,05	0,15	0,35	0,71	0,86	0,94	1	

(0,5)

D'après le cours  $F(x) =$

$$\left\{ \begin{array}{l}
 0 \text{ si } x \leq 75 \\
 \frac{0,05}{5}(x-75) = 0,01x - 0,75 \quad \text{si } x \in C_1 \\
 \frac{0,1}{5}(x-80) + 0,05 = 0,02x - 1,55 \quad \text{si } x \in C_2 \\
 \frac{0,2}{5}(x-85) + 0,15 = 0,04x - 3,25 \quad \text{si } x \in C_3 \\
 \frac{0,36}{5}(x-90) + 0,35 = 0,072x - 6,13 \quad \text{si } x \in C_4 \\
 \frac{0,15}{5}(x-95) + 0,71 = 0,03x - 2,14 \quad \text{si } x \in C_5 \\
 \frac{0,08}{5}(x-100) + 0,86 = 0,016x - 0,74 \quad \text{si } x \in C_6 \\
 \frac{0,06}{5}(x-105) + 0,94 = 0,012x - 0,32 \quad \text{si } x \in C_7 \\
 1 \text{ si } x \geq 110
 \end{array} \right.$$

(1,5)

20)  $Q_3 \in C_5 = [95, 100[$   $F_4 \leq 0,75$  et  $F_5 > 0,75$

d'où  $0,03Q_3 - 2,14 = 0,75$  donc  $Q_3 = \frac{0,75 + 2,14}{0,03} = 96,33$

Interprétation: 75% des taxis parcourent un nombre moyen de km par jour  $\leq 96,33$  km (1)

30)  $F(86) = 0,04 \times 86 - 3,25 = 0,19$   
 $F(105) = 0,94$   
 $F(105) - F(86) = 0,94 - 0,19 = 0,75$

donc 75% des taxis parcourent entre 86 et 105 km en moyenne par jour. (1)

## Exercice 2

$y_j \backslash x_i$	2	8	12	18	$n_{.j}$
6	8	1	1	0	10
9	1	10	2	0	13
11	1	2	14	1	18
14	0	0	2	7	9
$n_{i.}$	10	13	19	8	$N=50$

(1)

$$N = \sum_{i=1}^4 n_{i.} = \sum_{j=1}^4 n_{.j} = 50$$

$$1^o) \bar{x} = \frac{1}{N} \sum_{i=1}^4 n_{i.} x_i = \frac{496}{50} = 9,92 \quad (0,5)$$

$$\bar{y} = \frac{1}{N} \sum_{j=1}^4 n_{.j} y_j = \frac{501}{50} = 10,02 \quad (0,5)$$

$$\text{Var}(X) = \frac{1}{N} \sum_{i=1}^4 n_{i.} x_i^2 - \bar{x}^2 = \frac{6200}{50} - (9,92)^2 = 25,39 \quad \text{d'où } \sigma_x = 5,04 \quad (1)$$

$$\text{Var}(Y) = \frac{1}{N} \sum_{j=1}^4 n_{.j} y_j^2 - \bar{y}^2 = \frac{5355}{50} - (10,02)^2 = 6,70 \quad \text{d'où } \sigma_y = 2,59 \quad (1)$$

$$2^o) \text{Cov}(X, Y) = \frac{1}{N} \sum_{i=1}^4 \sum_{j=1}^4 n_{ij} x_i y_j - \bar{x} \bar{y} = \frac{5540}{50} - (9,92)(10,02) = 10,88 \quad (1)$$

$$\rho = \frac{\text{Cov}(X, Y)}{\sigma_x \cdot \sigma_y} = \frac{10,88}{5,04 \times 2,59} = 0,83 \quad (1)$$

3^o)  $y - \bar{y} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} (x - \bar{x})$  est l'équation de la droite

de régression de Y en X

$$\text{d'où } y - 10,02 = \frac{10,88}{25,39} (x - 9,92) \quad \text{ce qui donne}$$

$$\boxed{y = 0,43x + 5,75} \quad (1)$$

donc pour  $x=4$  on a  $y=7,47$  est la valeur estimée de Y si  $x=4$  (1)

### Exercice 3:

a) soit A l'événement "obtenir un 6"

$$P(A) = P_6 = 1 - \sum_{i=1}^5 p_i = 1 - 0,85 = \underline{0,15} \quad (1)$$

b) soit B l'événement "obtenir un nombre pair"

$$P(B) = P(2 \text{ ou } 4 \text{ ou } 6) = P(2) + P(4) + P(6) = P_2 + P_4 + P_6 = 0,2 + 0,1 + 0,15$$

$$\underline{P(B) = 0,45} \quad (1)$$

c) un événement élémentaire est un arrangement avec répétition de 5 nombres parmi 6

$$\underline{\text{Card } \Omega = 6^5 = 7776} \quad (0,15)$$

Soit C = "avoir au moins un 3"

$$P(C) = 1 - P(\bar{C}) = 1 - \frac{\text{Card } \bar{C}}{\text{Card } \Omega}$$

$$= 1 - \frac{5^5}{7776} = 1 - \frac{3125}{7776} = 1 - 0,4 = \underline{0,6} \quad (1)$$

$$\text{d) Card } D = C_5^2 P(4,4, \bar{4}, \bar{4}, \bar{4}) = C_5^2 \times 5^3 = 1250$$

$$P(D) = \frac{\text{Card } D}{\text{Card } \Omega} = \frac{1250}{7776} = 0,16 \quad (1,15)$$

### Exercice 4

Z est donnée sous forme d'un nuage de points

$$\text{Cov}(X, Y) = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{N} \sum_{i=1}^N [x_i y_i - x_i \bar{y} - \bar{x} y_i + \bar{x} \bar{y}]$$

$$= \frac{1}{N} \sum_{i=1}^N x_i y_i - \frac{\bar{y}}{N} \sum_{i=1}^N x_i - \frac{\bar{x}}{N} \sum_{i=1}^N y_i + \frac{\bar{x} \bar{y}}{N} \sum_{i=1}^N 1$$

$$= \frac{1}{N} \sum_{i=1}^N x_i y_i - \bar{y} \bar{x} - \bar{x} \bar{y} + \bar{x} \bar{y} \quad (3)$$

$$= \frac{1}{N} \sum_{i=1}^N x_i y_i - \bar{x} \bar{y}$$