Université Abdelmalek Essaâdi ,
Faculté Polydisciplinaire de Tétouan ,
Licence fondamentale en Sciences Economiques et Gestion.
Année universitaire 2016-2017.

Travaux Dirigés 2 : Statistique Descriptive

Exercice 1

Soit la série statistique des salaires d'une entreprie:

Salaires	Nombre demployés
X -50	30
50-100	40
100 -200	20
200 -300	10

1-Retrouver la borne inférieure X de la première calsse sachant que le salaire moyenne est de 94.
(pour la suite des calcule, retenez la valeur trouvée à la première question.)
2-Donner l'interprétation et la valeur de la médiane(Mé)
3-Calculer le troisième quantile, le septième décile et le percentile 35.
4-Déterminer la variance et l'écart-type. 4 10 100

Exercice 2

Une enquete sur la mobilité a donné la réparation suivant (exprimée en %) pour une population d'individus domiciliés à la région Tanger-Tétouan, selon la distance entre le domicile et le lieu de travail (distance exprimée en kilomètres):

Distance (en km)	Tétouanais(%)	Tangérois(%)
[0, 2[	5.7	3.1
[2, 6[	12.0	7.7
[6, 10[	6.3	6.8
[10, 15[	8.1	4.6
[15, 25[	11.0	6.8
[25, 50[	11.2	5.4
[50, 60[	7.8	3.5

1- Calculez la distance moyenne parcourue par un travailleur tétouanais et par un tangérois.
2- Déterminez la variance et l'écart-type des distances parcourues par un travailleur tétounais et par un tangérois.
3- Calculer le cefficient de variation pour les distances parcurues par un travailleur tétouanais et pour celles parcourue par un tangérois. Conclure.

Exercice 3

Les salaires annuels (en 1000 DH) des employés d'une entreprise composée de deux filiales X et Y sont réparties selon le tableau suivant:

Salaires en 1000DH compris entre	Nombre d'employés de la filiale X	Nombre d'employés de la filiale Y
10 et 20	5	4
20 et 30	10	12
30 et 40	13	14
40 et 50	4	6

1-Déterminer le salaire moyen de la filiale X et de la filiale Y .
2- Déterminer la variance et l'écart-type des salaires de la filiale X et de la filiale Y .
3- Comparer la dispersion des salaires de la filiale X et de la filiale Y .

Exercice 4

On étudie les revenus (Annuels en miliers de dirhams) d'un ensemble de familles d'un quartier de Tétouan, les données sont regroupées dans le tableau suivant:

Revenus annuels $(en10^3DH)$	$[18,30[$	$[30,36[$	$[36,42[$	$[42,54[$	$[54,60[$	$[60,66[$
Effectifs	13	219 •	20	46	50	82

1- Préciser les caractéristique de cette série (populaion,taille ou l'effectif total, individu, caractère étudié, type de caractére et modalités .

1- Calculer la moyenne $\bar{x}$ de cette série statistique.

2- Dresser l'histogramme de cette série statistique puis représenter son polygone.

3- Déterminer le mode M_o de cette série, graphiquement et par calcul.

4- Calculer la médiane M_e de cette série statistique en explicitons vos calculs.

5- La série étudiée est-elle symétrique ou asymétrique ? Justifier votre réponse.Pouvait-on prévoir ce résultat ?.

Exercice 5

La distribution des salaires horaires, en dirhams, des N employés d'une grande entreprise et donnée par:

Classes	Effectifs
$[50,100[$	10
$[100,150[$	14
$[150,200[$	16
$[200,250[$	n

Ces données sont incomplet car, à la suit d'un incident, l'effectif de la dernière classe est illisible; alors, on a décidé da la noter provisoirement par n. Mais, on sait que la médiane de cette série statistique est 153,125 DH.

1- Exprimer la moyenne arithmétique de cette distribution en fonction de n.

2- Exprimer la médiane de cette série statistique en fonction de n, sachant que la valeur $\frac{N}{2}$ n'a pas été trouvée exactement parmi les effectifs cumulées croissants.

3- Retrouver la valeur numérique de la moyenne arithmétique en remplaçant la valeur de n trouvée dans l'expression de la moyenne ewprimée en 1° .

Exercice 6

La distribution, en pourcentage, des 50 employés d'une entreprise selon leurs salaires annuels (en 1000 dirhams) est donnée par le tableau suivant :

Salaires annuels (en 1000 DH) comprises entre	Pourcentages des employés
0 -30	20
30-60	28
60-90	36
90 -120	16

1- Calculer les fréquence relatives et déduire les différences effectifs.

2- Quel est le salaire médian (Mé)? Interpréter le résultat.

3- Calculer les trois quartiles Q_3, Q_2 et Q_1.

4- Déterminer le salaire annule moyen et Calculer la variance et l'écart-type .

Exercice 7

Le chifre d'affaire d'une entreprise a augmenté de 5% les deux premières années, de 7% les trois années suivantes et de 4% l'année d'apres. Quelle est, en percentage, son augmentation annuelle moyenne?

Exercice 8

Le chiffre d'affire d'une entreprise a subit les augmentations annuelles suivantes :

Année	Augmentation en %
2003	4%
2003	5%
2004	6%
2005	5%
2006	4%

Calculer son taux de croissance moyen.

||Ex 1:|| T.D. N°2 Solution

1)

$[e_{i-1}; e_i[$	n_i	c_i	$n_i c_i$	$n_i C\uparrow$	$(c_i-\bar{x})^2$	$n_i(c_i-\bar{x})^2$
$[X, 50[$	30	c_1	$30c_1$	30	4096	122880
$[50, 100[$	40	75	3000	70	361	14440
$[100, 200[$	20	150	3000	90	3136	62720
$[200, 300[$	10	250	2500	100	24336	243360
	$N=100$		9400			443400

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{4} n_i c_i = \frac{30c_1}{100} + \frac{8500}{100} = 94$$

$$\Rightarrow 30c_1 = 100(94-85) = 900 \Rightarrow c_1 = \frac{900}{30} = 30$$

$$\frac{X+50}{2} = 30 \Rightarrow X = 60-50 = 10$$

2) la médiane est le salaire pour lequel on a 50 employés ont un salaire inférieure au salaire médiane et autres 50 employés ont un salaire supérieure au salaire médiane.

$$\frac{N}{2} = \frac{100}{2} = 50 \Rightarrow M_é = e_{i-1} + \frac{\frac{N}{2} - n_{i-1}\uparrow}{n_i} \times a_i$$

$\Rightarrow$ 70 est la 1ère valeur supérieure à 50 $\Rightarrow [50, 100[$ la classe médiane $\Rightarrow M_é = 50 + \frac{50-30}{40} \times 50 = 75$

Alg $\boxed{M_é = 75}$

3) * $\frac{N}{4}\times 3 = \frac{100}{4}\times 3 = 75$

$$Q_3 = e_{i-1} + \frac{\frac{N}{4}\times 3 - n_{i-1}c}{n_i}\times c_i$$

⇒ 90 est la 1ère valeur supérieur à 75 ⇒

$$\Rightarrow Q_3 = 100 + \frac{75-70}{20}\cdot 100 = 100 + \frac{5}{20}\times 100 = 125$$

$$\boxed{Q_3 = 125}$$

* $\frac{N}{10}\cdot 7 = \frac{100}{10}\times 7 = 70$ cette valeur apparait dans le tableau alors on prend $D_7 = 100$

* $\frac{N}{100}\times 35 = 35$ ⇒ la 1ère valeur supérieur à 35 est 70

$$\Rightarrow P_{35} = 50 + \frac{35-30}{40}\cdot 50$$

$$\Rightarrow P_{35} = 50 + 6,25 = 56,25$$

4) $V(X) = \frac{1}{N}\sum_{i=1}^{k}(c_i - \bar{X})^2 n_i = 4434$

$$\sigma(X) = \sqrt{V(X)} = \sqrt{4434} \simeq 66,59.$$

Ex : 2 |

Distance (en km)	f_i Tétouan	f_i Tanger	c_i
[0, 2[	0,057	0,031	1
[2, 6[	0,12	0,077	4
[6, 10[	0,063	0,068	8
[10, 15[	0,081	0,046	12,5
[15, 25[	0,11	0,068	20
[25, 50[	0,112	0,054	37,5
[50, 60[	0,078	0,035	55

$f_i c_i$ Tet	$f_i c_i$ Tanger
0,057	0,031
0,48	0,308
0,504	0,544
1,0125	0,575
2,2	1,36
4,2	2,026
4,29	1,925

c_i^2	$c_i^2 f_i$ Tet	$c_i^2 f_i$ Tanger
1	0,057	0,031
16	1,92	1,232
64	4,032	4,352
156,25	12,656	7,1875
400	44	27,2
1406,25	157,5	75,9375
3025	235,95	105,875
$\Sigma =$	456,115	221,815

1) $\bar{X} = \frac{1}{N} \sum x_i n_i = \sum f_i c_i$

Tétouan : $\boxed{\bar{X} = 12,7435}$.

Tanger : $\boxed{\bar{X} = 6,768}$

(3)

2) $V(X) = \left(\sum c_i^2 f_i\right) - \bar{x}^2$

Tét: $\bar{X}^2 = 162{,}397$

$V(X) = 456{,}115 - 162{,}397 = 293{,}718$

$\sigma(X) = \sqrt{V(X)} = \sqrt{293{,}718} = 17{,}138$

Tong $\bar{X}^2 = 45{,}806$

$V(x) = 176{,}009 \Rightarrow \sigma(X) = 13{,}267$

3) $CV(\text{tét}) = \frac{\sigma}{\bar{X}} = \frac{17{,}138}{12{,}7435} = 1{,}345$

$CV(\text{tongl}) = \frac{\sigma}{\bar{X}} = \frac{13{,}267}{6{,}768} = 1{,}960$

$\Rightarrow$ tong et plus dispersée que Tét.

$[e_{i-1}, e_i[$	c_i	n_i	$n_i c_i$	c_i^2	$n_i c_i^2$
[10, 20[	15	5	75	225	1125
[20, 30[	25	10	250	625	6250
[30, 40[	35	13	455	1225	15925
[40, 50[	45	4	180	2025	8100
		32	960		3140

$$\bar{X} = \frac{1}{N} \sum n_i c_i = \frac{960}{32} = \boxed{30} \Rightarrow \bar{X}^2 = 900$$

$$V(X) = \left(\frac{1}{N} \sum c_i^2 n_i\right) - \bar{X}^2 = 981,25 - 900$$
$$= \boxed{81,25}$$

$$\sigma(X) = \sqrt{V(X)} = \sqrt{81,25} = 9,014$$

$$CV = \frac{\sigma(X)}{\bar{X}} = \frac{9,014}{30} = 0,300$$

$[e_{i-1}, e_i[$	c_i	n_i	$n_i c_i$	c_i^2	$n_i c_i^2$
[10, 20[	15	4	60	225	900
[20, 30[	25	12	300	625	7500
[30, 40[	35	14	490	1225	17150
[40, 50[	45	6	270	2025	12150
		36	1120		37700

$$\bar{Y} = 31,11 \Rightarrow \bar{Y}^2 = 967,8$$

$$V(Y) = \boxed{79,39}$$

$$\Rightarrow \sigma(y) = \boxed{8,91}$$

$$CV = \frac{\sigma(Y)}{\bar{Y}} = \frac{8,91}{31,11} = \boxed{0,281}$$

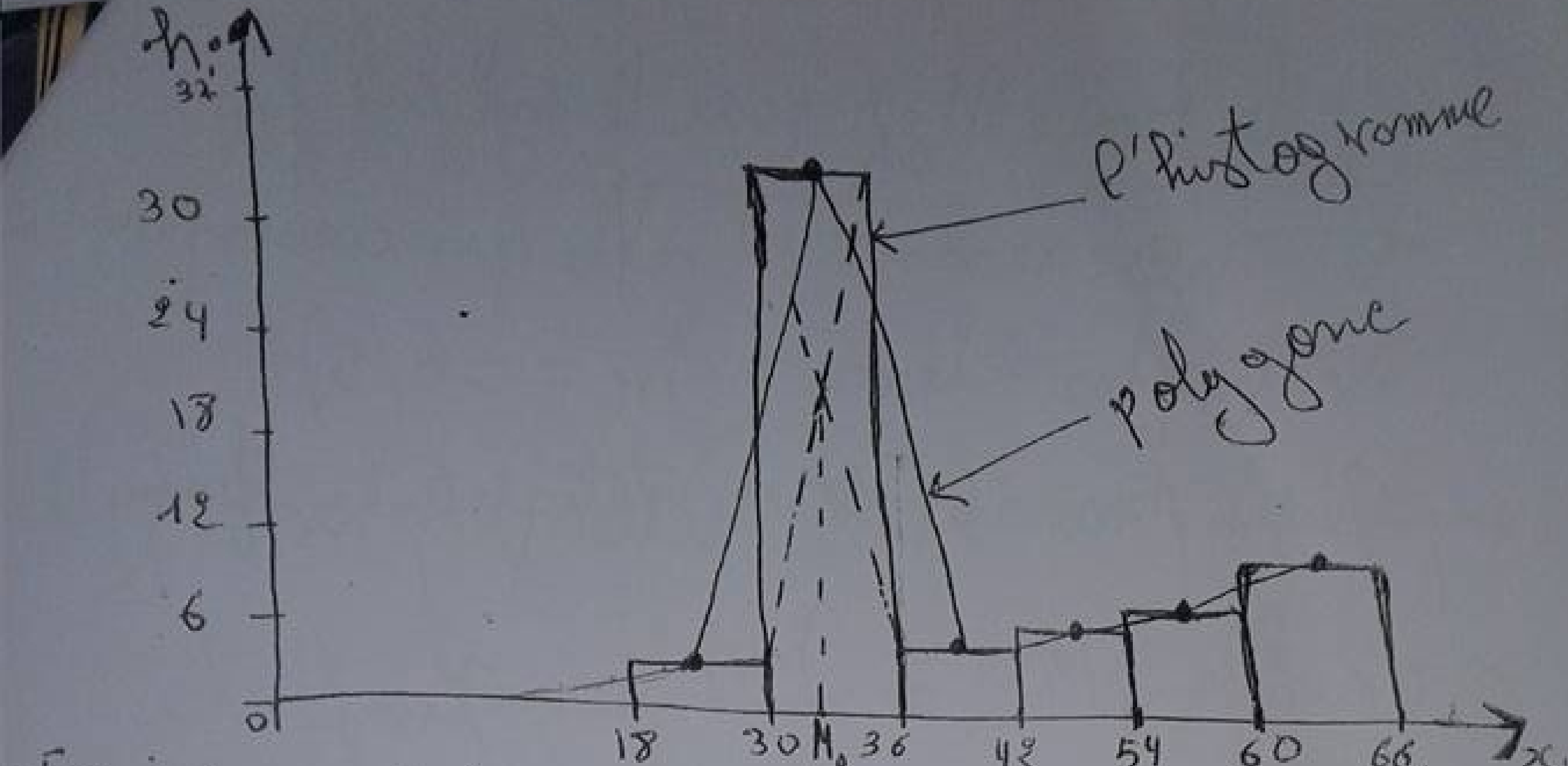
Ex : ...

$[e_{i-1}, e_i[$	n_i	c_i	$n_i c_i$	a_i	$h_i = \frac{n_i}{a_i}$	n_i cc
$[18, 30[$	13	24	312	12	1,083	13
$[30, 36[$	219	33	7227	6	36,5	232
$[36, 42[$	20	39	780	6	3,333	252
$[42, 54[$	46	48	2208	12	3,833	298
$[54, 60[$	50	57	2850	6	8,333	348
$[60; 66[$	82	63	5166	6	13,667	430
	$N = 430$		18543			

$$\bar{x} = \frac{1}{N} = \sum_{i=1}^{6} n_i c_i = 43,123$$

3) Détermination du Mode M_0 par la Méthode graphique :

Cela se fait sur l'histogramme. Donc d'abord on trace l'histogramme. Mais comme c'est une série statistique quantitative continue avec amplitudes de classe différentes, alors on trace l'histogramme pour les $h_i = \frac{n_i}{a_i}$



·) En joignant les sommets du rectangle le plus élevé, et les sommets du rectangle juste avant et le suivant, la projection sur l'axe des x du points de rencontre des diagonales obtenues (voir graphe en haut) donne la position de M_0 parmi les x_i : on a $M_0 \in [30,36[$

4) **Détermination de M_0 par le calcul :**

Comme les amplitudes de classes sont différentes on définit la classe modale comme étant celle du plus grande hauteur dans l'histogramme (qui correspond à h_i le plus grand). La classe modale est $[30,36[$, et on applique la formule.

↓↓

suit 4°) le mode M_o (: par le calcul)

la classe modale est celle qui correspond à h_i la plus élevée ($h_i = 36,5$) :

c'est [30, 36[et on applique la formule :

$$M_o = e_{i-1} + \frac{h_{i+1}}{h_{i-1} + h_{i+1}} a_i$$

$$M_o = 30 + \frac{3,333}{1,083 + 3,333} \times 6$$

$$= 30 + \frac{3,333}{4,416} \times 6 = 30 + 4,528 = 34,528$$

5) la médiane $M_é$:

$\frac{N}{2} = \frac{430}{2} = 215$ cette valeur ne se trouve pas exactement parmi les $n_i cc$, mais la 1ère valeur qui la dépasse est 232, elle correspond à [30, 36[(la classe médiane) et on applique la formule : $M_é = e_{i-1} + \frac{\frac{N}{2} - n_{(i-1)}^{cc}}{n_i} a_i$

$$M_é = 30 + \frac{215-13}{219} \times 6 = 30 + 5,534$$

$$\boxed{= 35,534}$$

6) Pour que cette série soit symétrique, il faut que $M_o = \bar{x} = M_é$, mais on a :

$$\left.\begin{array}{l} \bar{x} = 43,123 \\ M_o = 34,528 \\ M_é = 35,534 \end{array}\right\} \Rightarrow \text{elle n'est pas symétrique car } M_o \neq \bar{x} \neq M_é.$$

Bien sûr, on voit sur l'histogramme que la représentation graphique n'est pas symétrique.

EX :

1) $f_i = \frac{n_i}{N} \Rightarrow n_i = f_i \times N$

$[e_{i-1}, e_i[$	n_i	f_i	$n_i c\nearrow$	c_i	$n_i c_i$
$[0,30[$	10	0,2	10	15	150
$[30,60[$	14	0,28	24	45	630
$[60,90[$	18	0,36	42	75	1350
$[90,120$	8	0,16	50	105	540
	$N=50$				2970

2) Médian ?

$\frac{N}{2} = \frac{50}{2} = 25 \Rightarrow$ Cette valeur ne vient pas exactement dans la colonne des $n_i c\nearrow$, mais la 1ère valeur qui la dépasse est $\underline{\underline{42}}$

$\Rightarrow$ la classe médiane est $[60,90[$.

$\Rightarrow$ on applique la formule :

$$M_é = e_{i-1} + \frac{\frac{N}{2} - n_{(i-1)}c\nearrow}{n_i} a_i$$

$$M_é = 60 + \frac{25-24}{18} \times 30 = 61,667 \simeq 61,67 \in [60,90[$$

le salaire annuel médian est

$$\boxed{Mé = 61667 \text{ DH}}$$

$\frac{N}{2} = \frac{50}{2} = 25$

Interprétation :

=> Il y a 25 employés qui ont un salaire inférieur à 61667 DH et 25 autres qui ont un salaire supérieur à 61667 DH

3) Quartiles :

$Q_1 : \frac{N}{4} = \frac{50}{4} = 12,5 \Rightarrow$ cette valeur ne vient pas exactement dans la colonne des $n_{i}c\uparrow$, mais la $1^{ère}$ valeur qui la dépasse est 24.

on applique la formule :

$$Q_1 = e_{i-1} + \frac{\frac{N}{4} - n_{i-1}c\uparrow}{n_i} a_i$$

$$Q_1 = 30 + \frac{12,5 - 10}{14} \cdot 30$$

$$\Rightarrow \boxed{Q_1 = 35,357} \in [30,60)$$

Q_2 ? $\frac{N}{4} \cdot 2 = \frac{N}{2} = 25 \Rightarrow Q_2 = M_e$

$\Rightarrow Q_2 = M_e = \boxed{61,67}$

Q_3 ? $\frac{N}{4} \times 3 = 37,5 \Rightarrow$ cette valeur n'existe pas dans la colonne des $n_i c\uparrow$, mais la 1ère valeur qui la dépasse est : $\underline{\underline{42}}$

on applique la formule :

$$Q_3 = e_{i-1} + \frac{\frac{N \cdot 3}{4} - n_{i-1} c\uparrow}{n_i} a_i$$

$$Q_3 = 60 + \frac{37,5 - 24}{18} \cdot 30$$

$\boxed{Q_3 = 82,5} \in [60, 90[$.

4) Moyenne arithmétique (Méthode directe).

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{k} n_i c_i = \frac{2970}{50} = 59,4$$

$\Rightarrow$ le salaire annuel moyen est $\underline{\underline{59400 DH}}$

5) variance : $V(X) = \frac{1}{N}\sum_{i=1}^{k} n_i (c_i - \bar{x})^2$.

$[e_{i-1}, e_i[$	$(c_i - \bar{x})^2$	$n_i(c_i - \bar{x})^2$
$[0, 30[$	1971,36	19713,6
$[30, 60[$	207,36	2903,04
$[60, 90[$	243,36	4380,48
$[90, 120[$	2079,36	16634,88
		43632

$$V(X) = \frac{43632}{5} = \boxed{872,64}$$

$$\sigma(X) = \boxed{29,540}$$

"Ex : 6"

l'augmentation moyenne annuelle est une moyenne géométrique :

$$G = \sqrt[6]{(1{,}05)^2 (1{,}07)^3 (1{,}04)}$$

$$G = 1{,}0582$$

soit : un taux de croissance 5,82 % approximativement

"Ex : 7"

l'augmentation annuelle moyenne est donnée par

$$G = \sqrt[5]{(1{,}04)^2 (1{,}06)(1{,}05)^2} \quad \left(\text{moyenne géométrique}\right)$$

$$G \simeq 1{,}048$$

soit un taux de croissance de 4,8 % approximativement.