TD 2

Université Abdelmalek Essaâdi, **Faculté .S.J.E.S à Tétouan**,
Licence fondamentale en Sciences Economiques et Gestion.
Année universitaire 2020 - 2021

Travaux Dirigés 2 : Statistique Descriptive

Exercice 1

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 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

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 $(en 10^3 DH)$	[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 .

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

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

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

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

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

Exercice 4

Une étude sur le chiffre d'affaires d'une population de PME a permis d'obtenir les résultats suivants (en milliers de dirhamrs).

minimum	3500
moyenne	4900
Ecart-type	650
mode	4550
Intervalle intercartile	1100
médiane	6400
premier quartile	4100
étendue	5000

1- Parmi ces paramétres, indiquer celles qui sont caractéristique de position ou de tendance centrale et celle qui sont des caractéristique de dispersion.

2- Queles le chiffre d'affaires le plus fréquent . et le plus grand dans cette population de PME ?

3- Quel est le coefficient de variation de cette population PME ? la population étudiée est-elle homogène ?

4- Donner le deuxième et le troisième quartile . Représenter le diagrame de Box-Wiskers ou la Boite de Tukey.

Exercice 5

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.

5- Donner l'interprétation et la valeur de la médiale (Ml).

6- Que peut-on dire de la différence ΔM =Ml-Mé ? . Comparer-la à l'étendue.Interpréter le résultat.

7- Construire la Courbe de LorenZ.

8- déterminer l'indice de concentration de Gini .Conclure.

Exercice 6

Le tableau ci-dessous donne la répartition de 200 exploitations agricoles . dans un certain pays, selon la SAU (surface agricole utilisée) exprimée en hectares.

SAU(en ha)	Fréquences(en %)
de 0 à 10	22
de 10 à 30	28
de 30 à 50	27
de 50 à 100	23

1- Déterminer le pourcentage des exploitations agricoles qui ont une SAU supérieure à 30 ha.

2- Déterminer la SAU la plus fréquente.

3- Donner la valeur et l'interprétation de la médiane (Mé).

4- Calculer le troisième quartile (Q3) et le cinquiéme décile(D5).

5- Donner l'interprétation et la valeur de la médiale (Ml).

6- Calculer la différence ΔM =Ml-Mé ? .Calculer l'indice de concentration .Interpréter le résultat.

7- Construire la Courbe de Lorenz et Déterminer l'indice de concentration de Gini .Conclure.

Exercice 7

Une association de la défense des droits de la femme a mené une enquete auprès du personnel ouvrier d'un secteur industriel. Les résultats concernant les salaires annuels nets en milliers de dirhams sont résumés dans le tableaux suivant.

Salaire annul (en millier de DH)	Nombre d'ouvrières
[30,36[	10
[36,42[	8
[42,54[	4
[54,60[	n_4
Total	N

1- Calculer l'étendue de cette distribution statistique.
2- Sachant que le salaire annuel moyenne des femmes enquetées est égal à 39500DH. déterminer l'éffectif (n_4) de la dernière classe de la distribution du salaire de ces femmes, ansi que l'effectif total N.
3- Quel est le pourcentage des ouvrières qui ont un salaire annuel inférieur à 42000DH .
4- Déterminer les quartiles Q1;Q2 et Q3.Calculer l'écart-interquartille relatif.
5- Représenter le diagramme de Box-Wiskers correspondant. l'étendue est-il est un bon outil de mesure la dispersion. Pourquoi?
6- Calculer la variance et l'écart-type de cette distribution. Déterminer le coefficient de variation ? Conclure ?.

Exercice 8
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 .
5- Calculer la médiale (Ml) et donner son interprétation .
6- Que peut-on dire de la différence ΔM =Ml-Mé ? . Comparer-la à l'étendue.Interpréter le résultat.
7- Construire la Courbe de Lorenz et Déterminer l'indice de concentration de Gini .Conclure.

Exercice 9
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 exprimée en 1° .

Ex : 1 — T.D : N° 2 : Solution

$\boxed{S_1}$

$[e_{i-1}, e_i[$	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		31400

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

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

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

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

$[e_{i-1}, e_i[$	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

$\bar{Y} = 31,11 \Rightarrow \boxed{\bar{Y}^2 = 967,83}$

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

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

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

Ex: 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 : $\boxed{\bar{X} = 12,7935}$

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

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

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

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

$\Rightarrow$ tong est plus dispersé que Tét.

EX : 3

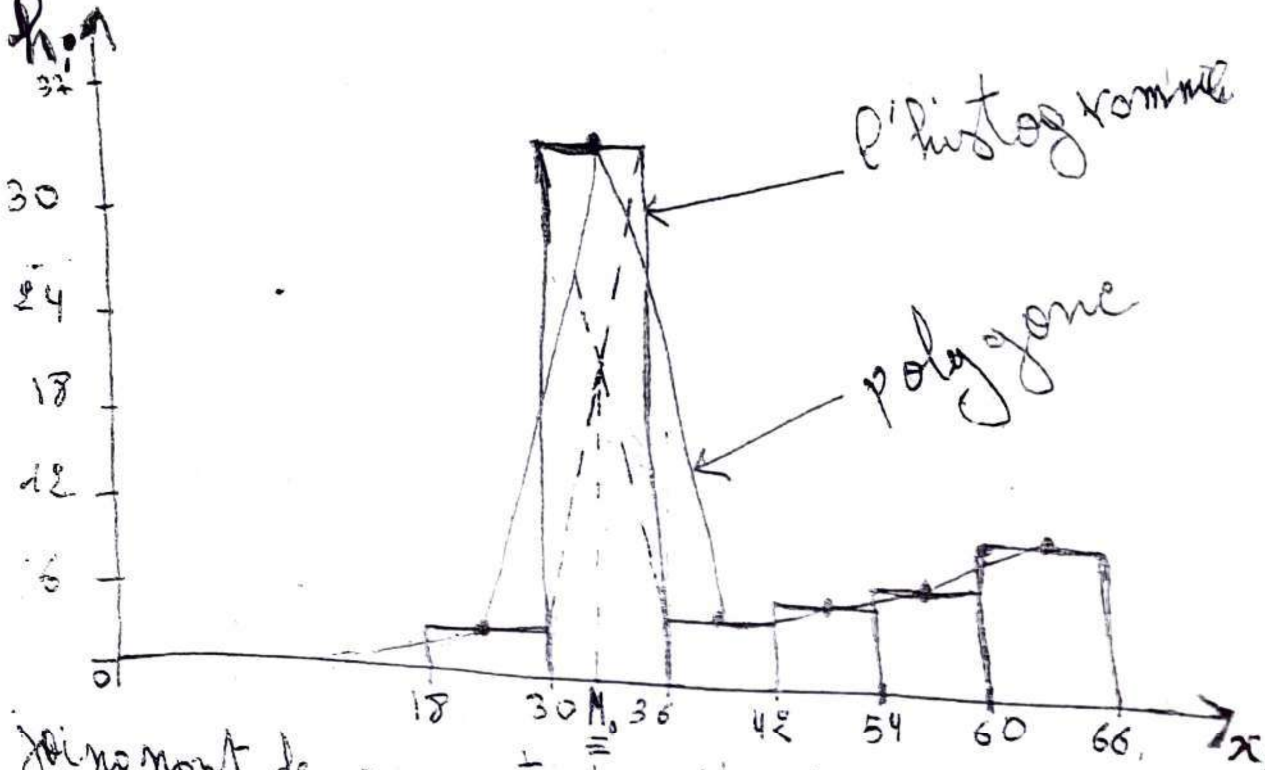
1)

$[e_{i-1}, e_i[$	n_i	c_i	$n_i c_i$	a_i	$h_i = \frac{n_i}{a_i}$	$n_i cc$ ↗ ou $n_i cc$
[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			

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

3) Détermination du Mode M_0 par la Méthode graphique :

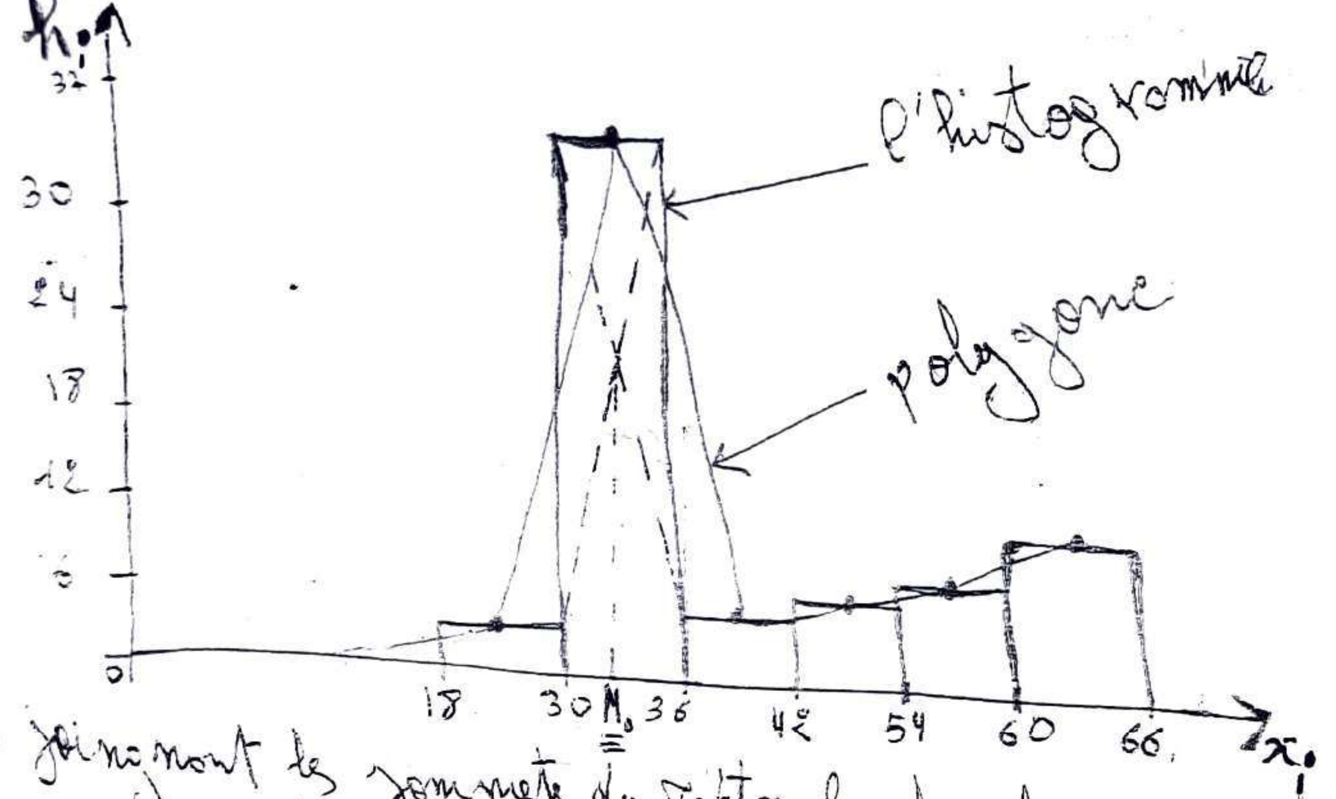
Cela se fait sur l'histogramme. Donc d'abord on trace l'histogramme. Mais comme c'est une série statistique quantitative continue avec amplitudes de classe différentes, donc on trace l'histogramme pour les $h_i = \frac{n_i}{a_i}$



-) 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 point de rencontre des diagonales obtenues (voir graphe en haut) donne la position de M_o parmi les x_i : on a $M_o \in [30,36[$

4) Détermination de M_o par le calcul :

Comme les amplitudes de classes sont différentes on définit la classe modale comme étant celle de 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.



En joignant les sommets du rectangle le plus élevé et le sommet 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_o parmi les x_i : on à $\underline{M_o \in [30,36[}$

4) Détermination de M_o par le calcul :

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

⇓

suite 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 : $M_é = e_{i-1} + \frac{\frac{N}{2} - n_{(i-1)}cc}{n_i} a_i$

$$M_é = 30 + \frac{21{,}5 - 13}{219} \times 6 = 30 + 5{,}224$$
$$\boxed{= 35{,}534}$$

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

$\bar{x} = 43{,}123$
$M_o = 34{,}528$
$M_é = 35{,}534$

$\Rightarrow$ 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: 4 Données : $x_{min} = 3500$

$\bar{x} = 4900$

$\sigma = 650$

$M_o = 4550$

$Q_3 - Q_1 = 1100$

$M_é = 4600$

$Q_1 = 4100$

$E = x_{max} - x_{min} = 5000.$

1)

paramètres de position	paramètres de dispersion
$\bar{x}$ M_o $M_é$	σ $Q_3 - Q_1$ Q_1 $E = x_{max} - x_{min}$

2) on a : $x_{max} = 5000 + x_{min}$

$= 5000 + 3500 = 8500$

→ Donc le chiffre d'affaires le plus grand est :

$\boxed{8500.000 \text{ DH}}$

→ le chiffre d'affaire le plus fréquent est donné par le mode : $\boxed{M_o = 4550}$

donc c'est 4550.000 DH.

3) $C_V = \frac{\sigma(x)}{\overline{X}} = \frac{650}{4900} = 0{,}133$

En pourcentage $C_V = \frac{\sigma}{\overline{X}} \times 100 = 13{,}3\%$.

Comme $C_V < 30\%$, donc la population étudiée est homogène.

4) le deuxième quartile $Q_2 = Mé$.

d'où $\boxed{Q_2 = 4600}$

on a : $Q_3 - Q_1 = 1100$

$\Rightarrow Q_3 = 1100 + Q_1 = 1100 + 4100 = 5200$

$\Rightarrow \boxed{Q_3 = 5200}$

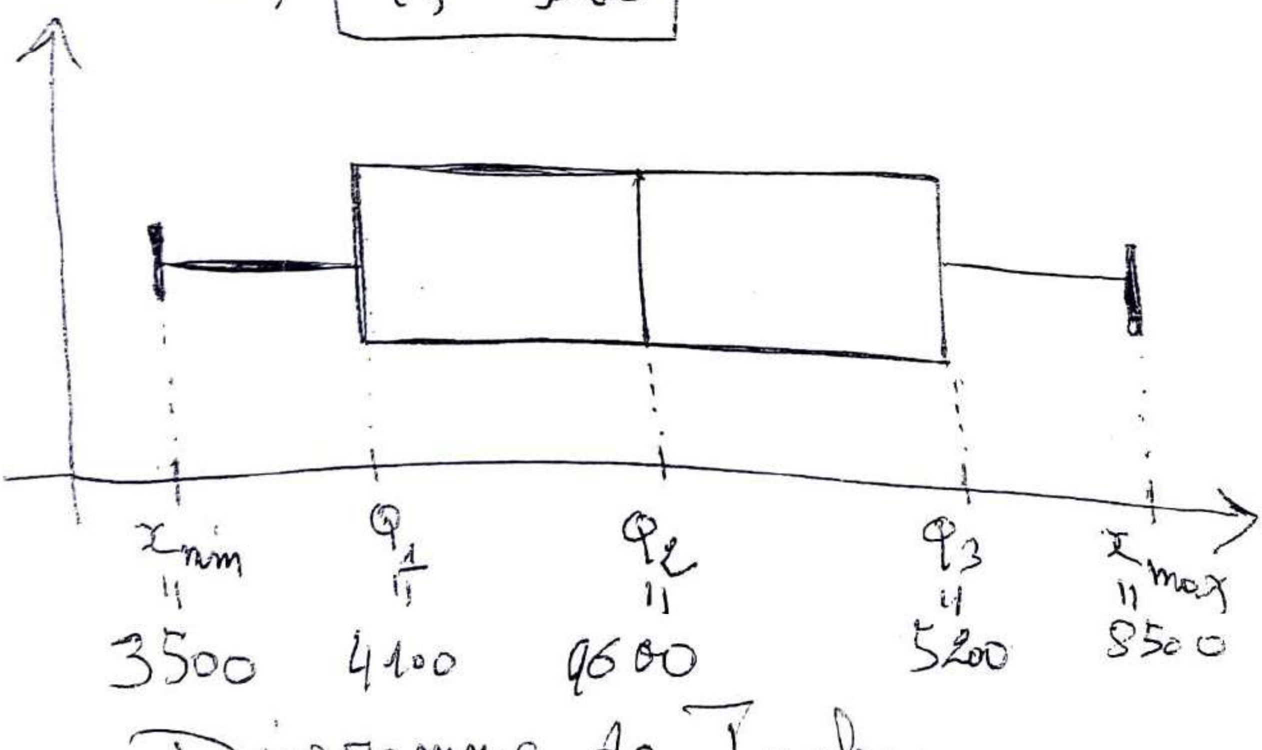


Diagramme de Tuckey

Ex : 5

1)

$[e_{i-1}; e_i[$	n_i	c_i	$n_i c_i$	$n_i c\uparrow$	$(c_i-\overline{x})^2$	$n_i(c_i-\overline{x})^2$
$[X, 50[$	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

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

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

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

2) la médiane est le salaire pour lequel on a 50 employés ont un salaire inférieur au salaire médiane et entre 50 employés ont un salaire supérieur au salaire médiane.

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

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

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

$\Rightarrow$ 90 est la 1ère valeur supérieure à 75 $\Rightarrow$

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

$$\boxed{Q_3 = 125}$$

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

* $\frac{N}{100} \times 35 = 35 \Rightarrow$ la 1ère valeur supérieure à 35 et 70

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

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

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

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

5) La Médiale (M_e) est la valeur de la variable qui divise vers la masse salariale totale (9400) en deux blocs égaux de 4700 chacun. $\Rightarrow \frac{9400}{2} = 4700$.

$$M_e = e_{i-1} + \frac{\frac{\sum n_i c_i}{2} - (n_{i-1} c_{i-1})c\uparrow}{n_i c_i} a_i$$

$[e_{i-1} ; e_i[$	$(n_i c_i)c\uparrow$
[; [	900
[; [	3900
[100 ; 200[	6900*
[; [	9400

la classe médiale

$$\Rightarrow M_e = 100 + \frac{4700 - 3900}{3000} \cdot 100$$

$$\Rightarrow M_e = 100 + \frac{800 \times 100}{3000} = 126,67$$

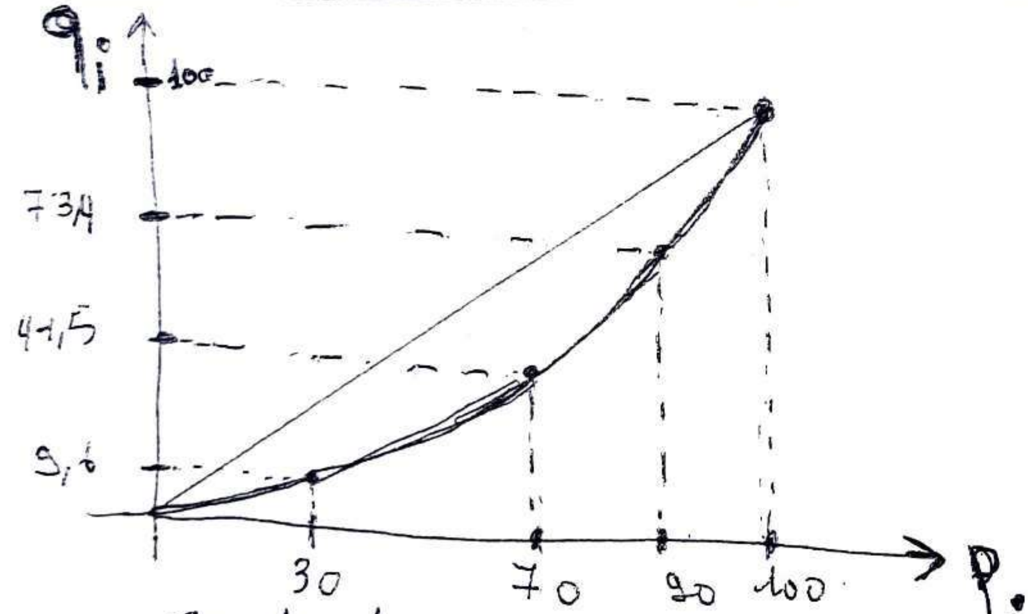
6) ΔM mesure la différence entre les valeurs centrales (correspondant à 50%) correspondant d'un côté à la masse salariale totale et de l'autre côté au nombre d'employés. Nous avons $M_e > M_{\acute{e}}$ d'où $\Rightarrow \Delta M > 0$; Dans notre cas $\Delta M = 126,67 - 75 = 51,67$.
L'utilité de cette différence réside dans le fait qu'on peut la comparer à l'étendue de la série qui est l'intervalle de variation de la variable salaire : Etendue $= 300 - 10 = 290$.

$$\frac{\Delta M}{290} = \frac{51,67}{290} \simeq 0,178.$$

cet indice signifie qu'on a une faible concentration des salaires.

7)

$(n_i c_i)_c \uparrow$	$p_i = \frac{(n_i)_c \uparrow}{N} \cdot 100$	$q_i = \frac{(n_i c_i)_c \uparrow}{\sum n_i c_i} \times 100$
900	30	9,574
3 900	70	41,489
6 900	90	73,404
9 400	100	100
Total	290	224,467



Courbe de concentration de Lorenz.

8) $$I_G = 1 - \frac{\sum_{i=1}^{K-1} q_i}{\sum_{i=1}^{K-1} p_i} = 1 - \frac{124,5}{190} = 0,345$$

$\Rightarrow$ faible concentration des salaires $\rightarrow$.
(équidistribution).

Ex : 6

$f_i = \frac{n_i}{N}$: ← d'après le cours :

$[e_{i-1}; e_i[$	f_i	n_i	f_i cd	a_i	$h_i = \frac{n_i}{a_i}$	$n_i c\uparrow$	c_i	$f_i c_i$	c_i^2	$c_i^2 f_i$
[0; 10[	0,22	44	1	10	4,4	44	5	1,1	25	5,5
[10; 30[	0,28	56	0,78	20	2,8	100	20	5,6	400	1,12
[30; 50[	0,27	54	0,5	20	2,7	154	40	10,8	1600	432
[50; 100[	0,23	46	0,23	50	0,92	200	75	17,25	5550	1276,5
		$N = 200$					$\bar{X} =$	34,75		1826

$f_i = \frac{n_i}{N}$

1) on cherche par les f_i cd la valeur qui correspond à la classe [30, 50[, c'est 0,5, on la multiplie par 100 et on obtient le pourcentage des exploitations agricoles qui ont un SAU supérieur à 30 ha, c'est $\boxed{50\%}$.

2) la classe modale est [0, 10[car elle correspond à $h_i = 4,4$ le plus grand. on applique la formule :

$$M_0 = e_{i-1} + \frac{h_{i+1}}{h_{i-1} + h_{i+1}} \times a_i \text{ , avec } [e_{i-1}, e_i[$$

$e_{i-1} = 0$; $a_i = 10$; $h_{i-1} = 0$; $h_{i+1} = 2,8$

donc $\boxed{M_0 = 10}$

d'où la SAU la plus fréquente est 10 ha.

3) La médiane ($M_é$) :

$\frac{N}{2} = \frac{200}{2} = 100$, cette valeur existe exactement parmi les $(n_i)c\uparrow$

$\Rightarrow$ la classe médiane est $[10, 30[= [e_{i-1} ; e_i[$

on prend $\Rightarrow$ $\boxed{M_e = e_i = 30}$

4°) Q_3 ?

$\frac{N}{4} \cdot 3 = \frac{200 \times 3}{4} = 150$ $\Rightarrow$ on cherche 150 parmi les $(n_i)c\uparrow$

$\Rightarrow$ il n'existe pas exactement, mais 154 est la 1ère valeur qui la dépasse, on applique alors la formule :

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

$$= 30 + \frac{150 - 100}{54} \times 20 = 30 + \frac{50}{54} \times 20$$

$\Rightarrow$ $\boxed{Q_3 = 48{,}52 \text{ ha}}$

$\boxed{D_5}$?

$\frac{N \cdot 5}{10} = \frac{200 \times 5}{10} = 100$; on cherche 100 parmi les $(n_i)c\uparrow$

cette valeur apparaît dans le Tableau, alors on prend :

$\boxed{D_5 = e_2 = 30 \text{ ha}}$

5) Médiale ($M\ell$) :

$[e_{i-1} ; e_i[$	n_i	c_i	$n_i c_i$	$(n_i c_i)c\uparrow$
[0, 10[	44	5	220	220
[10, 30[	56	20	1120	1340
[30, 50[	54	40	2160	3500 *
[50, 100[	46	75	3450	6950

5°) Médiale (Ml) : $\frac{6950}{2} = 3475$

on cherche cette valeur parmi les $(n_i c_i)c\nearrow$ la 1er valeur qui la dépasse est 3500 Donc la classe médiale est $[30, 50[$

on applique la formule : $M_l = e_{i-1} + \frac{\frac{\sum n_i c_i}{2} - (n_{i-1} \cdot c_{i-1})c\nearrow}{n_i c_i} \times a_i$

$$= 30 + \frac{3475 - 1340}{2160} \times 20 = 49{,}768$$

6°) $\Delta M = M_l - M_e = 49{,}768 - 30 = 19{,}768$

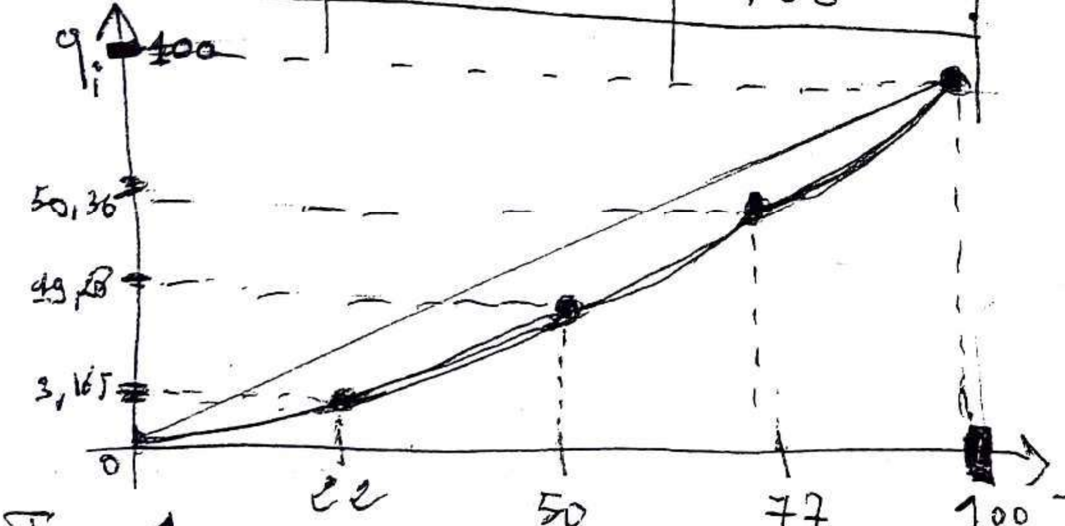
$E = 100$

$\Rightarrow I = \frac{\Delta M}{E} = \frac{19{,}768}{100} = 0{,}19768 \simeq 0{,}198$

$\Rightarrow$ Équi-distribution.

7°)

$[e_{i-1} ; e_i[$	$p_i = \frac{n_i c\nearrow}{N} \times 100$	$q_i = \frac{(n_i c_i)c\nearrow}{\sum n_i c_i} \times 100$
$[0, 10[$	22	3,165
$[10, 30[$	50	19,280
$[30, 50[$	77	50,360
$[50, 100[$	100	100



$I_G = 1 - \frac{\sum_{i=1}^{k-1} q_i}{\sum_{i=1}^{k-1} p_i} = 1 - \frac{72{,}805}{149} = 1 - 0{,}489 = 0{,}511$

$\Rightarrow$ Équi-distribution.

Ex : 7

1)

$[e_{i-1};e_i[$	n_i	c_i	n_{icc} ou $n_{ic}\uparrow$	a_i	$(c_i-\bar{x})^2$	$n_i(c_i-\bar{x})^2$	f_i	$f_{ic}\uparrow$
[30;36[	10	33	10	6	42,25	422,5	0,4	0,4
[36;42[	8	39	18	6	0,25	2	0,3	0,7
[42;54[	4	48	22	12	72,25	289	0,2	0,9
[54;60[	$n_4 = 2$	57	24	6	306,25	612,5	0,1	1
	N=24					1326	1	

on a : $\boxed{\bar{x} = 39,5}$

1°) l'étendue ou l'intervalle de variation est la façon la plus simpliste de mesurer la dispersion.

on a : $E = x_{max} - x_{min}$

= borne sup de la dernière classe _ borne inf de la $1^{ère}$.

$= 60 - 30 = 30$

2°) $\bar{x} = 39,5$ en milliers de dirhams.

on a : $10 \times 33 + 8 \times 39 + 4 \times 48 + 57 \times \underline{\underline{n_4}} = N \times 39,5$

mais $N = 10 + 8 + 4 + n_4$

alors : $834 + 57 n_4 = 39,5 \times (22 + n_4)$

$\Rightarrow 834 + 57 \times n_4 = 869 + 39,5\, n_4$

$\Rightarrow 17,5 \times n_4 = 35 \Rightarrow \boxed{n_4 = 2}$ et $\boxed{N = 24}$

3) on lit directement du tableau des calculs dans la colonne des f_{icc} en face de la classe [36;42[on a la valeur 0,7.

4) $\underline{\underline{Q_1?}}$: $\frac{N}{4} = \frac{24}{4} = 6$ on cherche ce 6 parmi les $(n_i)c\uparrow$ $\Rightarrow$ il n'existe pas exactement, mais $\underline{\underline{10}}$ est la $1^{ère}$ valeur qui dépasse ce 6 Alors on applique la formule :

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

$$= 30 + \frac{6-0}{10} \times 6 = 30 + 3{,}6 = \underline{33{,}6}$$

$\underline{\underline{Q_2?}}$: $\frac{N}{4} \times 2$ (càd que $\underline{\underline{i=2}}$) $\Rightarrow \frac{N}{4} \times 2 = 12$ cette valeur n'existe pas parmi les $n_i cc$, mais 18 est la $1^{ère}$ valeur qui le dépasse, on applique la formule :

$$\underline{\underline{Q_2}} = e_{i-1} + \frac{\frac{N}{4} \times 2 - n_{i-1} cc}{n_i} \times a_i.$$

$$= 36 + \frac{12-10}{8} \times 6 = 36 + 1{,}5 = \underline{37{,}5}$$

$\underline{\underline{Q_3?}}$ $\frac{N}{4} \times 3$ $(i=3) \rightarrow \frac{N}{4} \times 3 = 18$ cette valeur existe parmi les $n_i cc$ alors on prend :

$Q_3 = e_i = \underline{\underline{42}}$ (car $18 \in [36 ; 42[= [e_{i-1}, e_i[$)

$\Rightarrow$ la $\underline{\underline{\text{dérivation}}}$ $\underline{\text{quartile}}$ ou $\underline{\underline{\text{semi-interquartile}}}$ est :

$$\Rightarrow : \frac{Q_3 - Q_1}{2} = \frac{42 - 33{,}6}{2} = \underline{\underline{4{,}2}}$$