Université Abdelmalek Essaadi
Faculté des Sciences Juridiques
Economiques et Sociales
-Annexe Tétouan-

Année Universitaire: 2003/2004
Module: Méthodes quantitatives 1
Elément de module: Statistiques 1
Enseignant: M. El Merouani

Examen final (2ème semestre)

Durée 2h

Exercice n°1 : (7 points)

La répartition des 100 exploitations agricoles selon leurs superficies en hectare (ha) se présente comme suit :

Superficies (en ha) comprises entre	Nombre d'exploitations
0 – 5	5
5 – 10	24
10 – 20	38
20 – 50	26
50 – 100	7

1) Calculer les fréquences relatives, les centres et les amplitudes des classes.
2) Représenter l'histogramme et le polygône des fréquences.
3) Calculer les fréquences relatives cumulées croissantes et décroissantes. Représenter leurs courbes dans le même repère.
4) Quelle est la superficie la plus fréquente ?
5) Combien y-a-t-il des exploitations qui ont une superficie inférieure à 20 ha ?
6) Quel est le pourcentage des exploitations qui ont une superficie supérieure à 10 ha ?

Exercice n°2 : (13 points)

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) compris entre	Pourcentages des employés
0 – 30	20
30 – 60	28
60 – 90	36
90 – 120	16

1) Calculer les fréquences relatives et déduire les différents 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 annuel moyen.
5) Calculer la variance et l'écart-type.
6) Calculer la médiale (Ml) et donner son interprétation.
7) Que peut-on dire de la différence $\Delta M = Ml - Mé$? Comparer-la à l'étendue. Interpréter le résultat.
8) Tracer la courbe de concentration de Lorenz.
9) Calculer l'indice de Gini. Conclure.

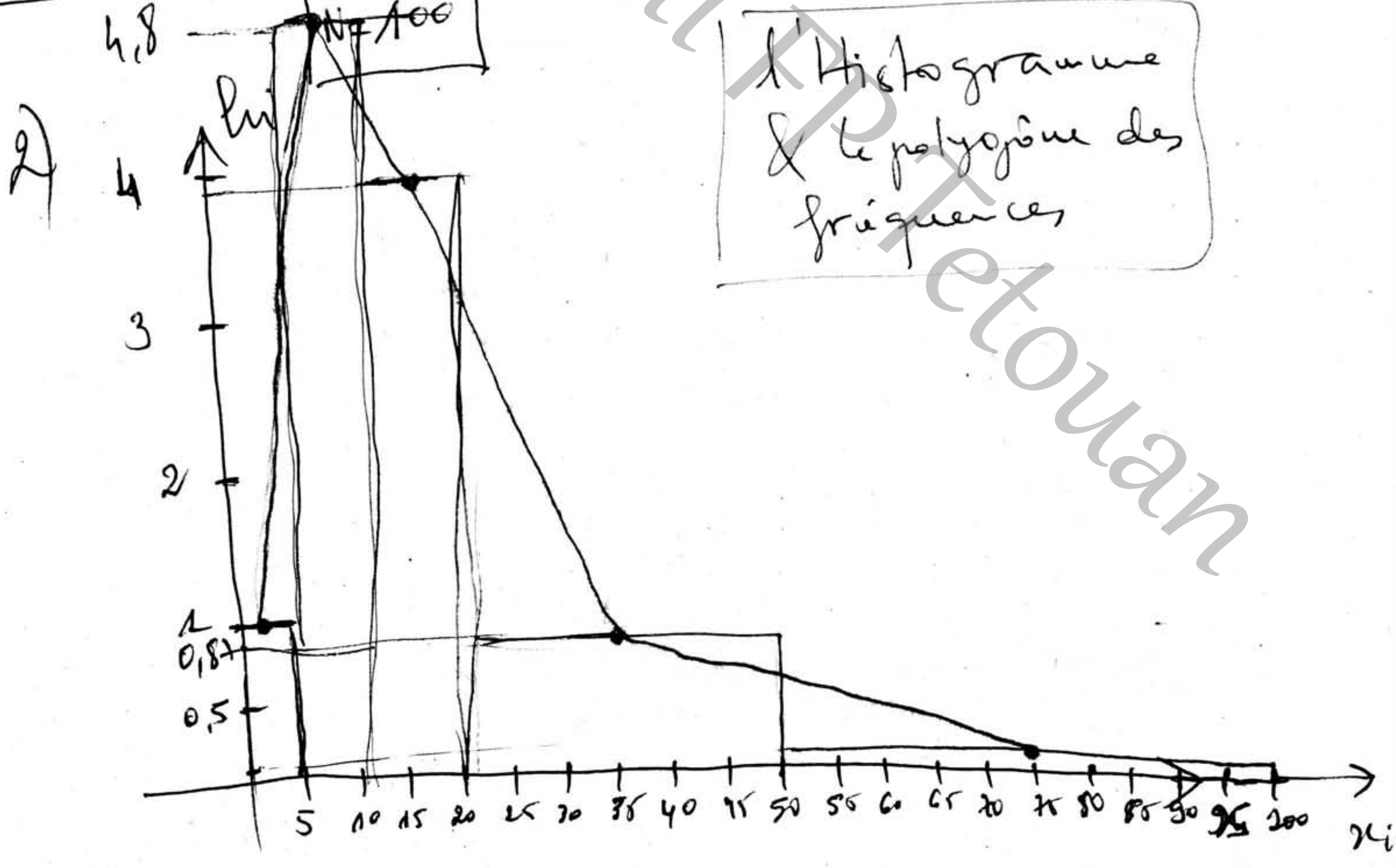
Examen Statistique

Exercice 1 :

1)

$[e_{i-1}, e_i[$	n_i	f_i	c_i	a_i	$h_i = \frac{n_i}{a_i}$
[0,5[	5	0,05	2,5	5	1
[5,10[	24	0,24	7,5	5	4,8
[10,20[	38	0,38	15	10	3,8
[20,50[	26	0,26	35	30	0,87
[50,100[	7	0,07	75	50	0,14
	N = 100				

2)

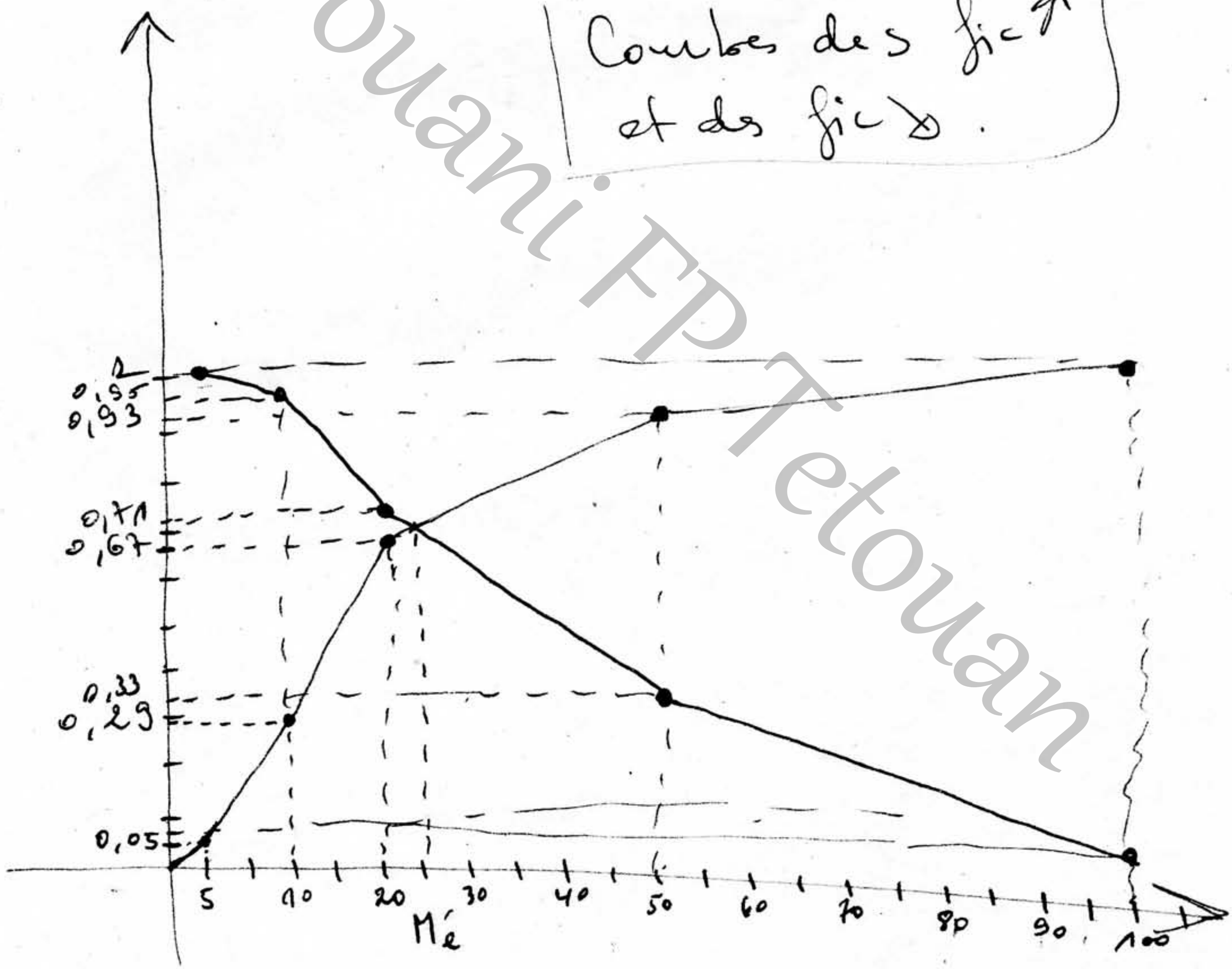


l'Histogramme & le polygône des fréquences

3) Les fréquences relatives cumulées ↗ et ↘.

$[e_{i-1}, e_i[$	f_i	f_{ic} ↗	f_{ic} ↘
[0, 5[	0,05	0,05	1
[5, 10[	0,24	0,29	0,95
[10, 20[	0,38	0,67	0,71
[20, 50[	0,26	0,93	0,33
[50, 100[	0,07	1	0,07

Courbes des f_{ic} ↗ et des f_{ic} ↘.



4) La surface la plus fréquente correspond au mode. La classe modale est $[5,10[$ celle qui correspond à li le plus grand. On applique la formule

$$\boxed{Mo = e_{i-1} + \frac{l_{i+1}}{l_{i-1} + l_{i+1}} a_i}$$

$$Mo = 5 + \frac{3,8}{1 + 3,8} 5$$

$$Mo = 8,958 \simeq 8,96 \in [5,10[$$

La superficie la plus fréquente est $\boxed{8,96 \text{ ha}}$

5) Inférieure à $\longrightarrow fic\uparrow$

$\downarrow \boxed{20 \text{ ha}} \longrightarrow 0,67 = fi = \frac{ni}{100}$

$$\Longrightarrow \boxed{ni = 67}$$

il y a 67 exploitation qui ont une superficie inférieure à 20 ha.

6) Supérieure à $\longrightarrow fic\downarrow$

$\downarrow \boxed{10 \text{ ha}} \longrightarrow 0,71 \longrightarrow 71\%$

71 % des exploitations ont une superficie supérieure à 10 ha

Exercice n° 2 :

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

$[e_{i-1}, e_i[$	n_i	f_i	$n_{ic}\nearrow$	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	840
	N=50				2970

2) Médiane ?

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

$\Rightarrow$ La classe médiane est [60,90[

$\Rightarrow$ On applique la formule

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

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

le salaire annuel médian est

$\boxed{Mé = 61667 \text{ DH}}$

$N/2 = 50/2 = 25$

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 $ni c\uparrow$, mais la 1ère valeur qui la dépasse est 24.

On applique la formule :

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

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

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

Q_2? $\frac{N}{4} . 2 = \frac{N}{2} = 25 \Rightarrow Q_2 = Mé = 61,67$

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

On applique la formule :

$$Q_3 = l_{i-1} + \frac{\left(\frac{N3}{4}\right) - n_{i-1}c\uparrow}{n_i} a_i$$

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

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

4) Moyenne arithmétique (Méthode directe)

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

Le salaire annuel moyen est 59400^{DH}.

5) Variance :

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

$[l_{i-1}, e_i[$	$(c_i - \bar{x})^2$	$n_i(c_i - \bar{x})^2$
[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

$$Var(X) = \frac{43632}{50} = \boxed{872,64}$$

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

6) Médiale (Ml) :

$[e_{i-1}, e_i[$	$(n_i c_i)c\nearrow$
[0,30[	150
[30,60[	780
[60,90[	2130 ×
[90,120[	2970

$$\frac{2970}{2} = 1485$$

cette valeur ne vient pas exactement dans la colonne des $(n_i c_i)c\nearrow$ la 1ère valeur qui la dépasse est : 2130. Donc la classe médiale est [60,90[.

On applique la formule :

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

$$Ml = 60 + \frac{1485 - 780}{1350} . 30$$

$$\boxed{Ml = 75,667} \in [60,90[$$

Interprétation : La masse salariale de l'entreprise, versée annuellement aux employés montre que le partage en 2 blocs égaux s'effectue par la valeur 75667 DH.

Autrement dit, la moitié de la masse salariale est reçue par le personnel qui a un salaire inférieur à 75667 DH et l'autre moitié est reçue par ceux qui ont un salaire supérieur à 75667 DH.

7) ΔM mesure la différence entre les valeurs centrales (correspondant à 50 %) correspondant d'un côté à la masse salariales totale et de l'autre côté au nombre d'employés. Nous avons $Ml = 75,667$
$Mé = 61,67$

$Ml > Mé \Rightarrow \Delta M = Ml - Mé > 0$

$\Delta M = 75,667 - 61,67 = 13,997$

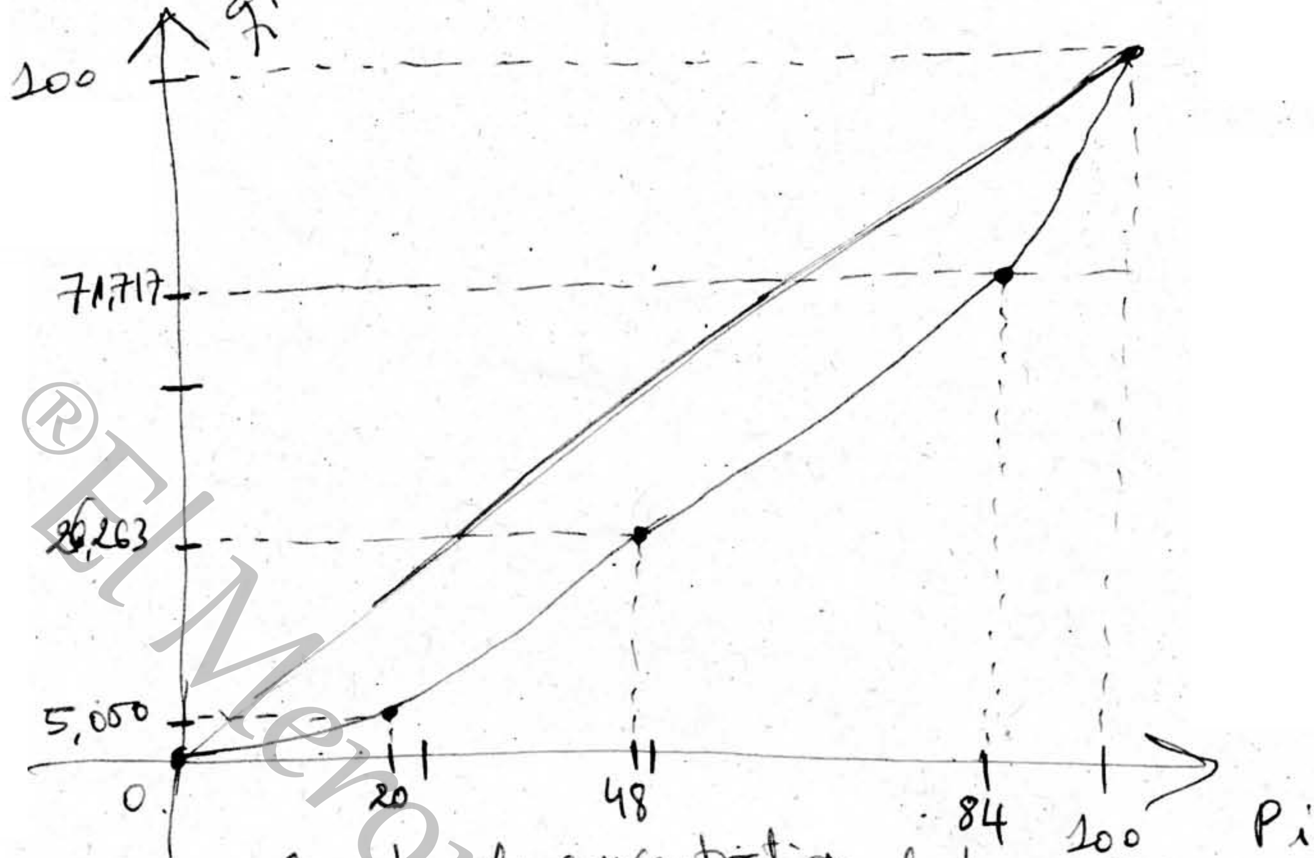
L'utilité de cette différence réside dans le fait que l'on peut la comparer à l'étendue de la série qui est l'intervalle de variation de la variable salaire : Etendue $= 120 - 0 = 120$

$I = \frac{\Delta M}{120} = \frac{13,997}{120} = 0,117$

cet indice veut dire que l'on a une faible concentration des salaires ou les salaires sont réparties d'une façon égalitaire.

8) Courbe de Lorenz :

$P_i = \frac{n_i c^{\uparrow}}{N} \cdot 100$	$q_i = \frac{(n_i c_i) c^{\uparrow}}{\sum_i n_i c_i} \cdot 100$
20	5,050
48	26,263
84	71,717
100	100
252	203,03



Courbe de concentration de Lorenz

9) $$I_G = 1 - \frac{\sum_{i=1}^{k-1} q_i}{\sum_{i=1}^{k-1} p_i} = 1 - \frac{103,03}{152} = 1 - 0,678$$

$$= 0,322$$

Conclusion :

faible concentration des salaires (équidistribution)