

Module : Méthodes Quantitatives I

Élément : Statistique I

Enseignant : M. El Merouani.

Contrôle continu final (Durée 2 heures)

Problème 1 : (6 points)

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 1000 DH 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.

Problème 2 : (14 points)

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(Q_3) et le cinquième décile (D_5).
- 5) Donner l'interprétation et la valeur de la médiale (Ml).
- 6) Calculer la différence $\Delta M = Ml - Mé$. Calculer l'indice de concentration. Interpréter le résultat.
- 7) Construire la courbe de Lorenz.
- 8) Déterminer l'indice de concentration de Gini. Conclure.



$[e_{i-1}, e_i]$	c_i	n_i	$n_i c_i$	c_i^2	$n_i c_i^2$	Problème n°1
$[10, 20]$	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	

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

$$\text{Var}(x) = \left(\frac{1}{N} \sum c_i^2 n_i \right) - \bar{x}^2 = \frac{981,25}{32} - 900 \\ = 81,25$$

$$\sigma(x) = \sqrt{\text{Var}(x)} = \sqrt{81,25} = 9,014$$

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

$[e_{i-1}, e_i]$	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	10	300	625	7500
$[30, 40]$	35	14	490	1225	17150
$[40, 50]$	45	6	270	2025	12150
		36	1190		37700

$$\bar{y} = 31,11 \Rightarrow \bar{y}^2 = 967,832$$

$$\text{Var}(y) = 79,39 \Rightarrow \sigma(y) = 8,91$$

$$CV = \frac{\sigma(y)}{y} = \frac{8,91}{31,11} = 0,286$$

$[e_{i-1}, e_i]$	f_i	n_i	$f_i c_d$	a_i	$h_i = \frac{n_i}{a_i}$	nocc	c_i	$f_i c_i$	\bar{c}_i^2	$c_i^2 f_i$	
$[0, 10]$	0,22	45	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	112	
$[30, 50]$	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	550	1265	
$N = 200$		$\bar{x} = 34,75$		1826							

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

Problème n° 2

1/ On cherche par les $f_i c_d$ la valeur qui correspond à la classe $[30, 50]$, c'est 0,5, on la multiplie par 200 et on obtient le pourcentage des exploitations agricoles qui ont une SAV supérieure à 30 ha, c'est 50 %.

2/ La classe modale est $[0, 10]$ car elle correspond à $h_i = 4,4$ le plus grand.

On applique la formule : $M_o = e_{i-1} + \frac{h_{i+1}}{h_{i+1} + h_{i+2}} a_i$

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

$$\text{donc } M_o = 10$$

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

3/ La médiane :

$\frac{N}{2} = \frac{200}{2} = 100$, cette valeur existe exactement parmi les nocc

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

$$\text{On prend } M_e = e_i = 30$$

4/

 $Q_3 ?$

$$\frac{N \cdot 3}{4} = \frac{200 \times 3}{4} = 150$$

On cherche 150 parmi les ncc

⇒ il n'existe pas exactement, mais 154 est la 1^{ère} valeur qui le dépasse, on applique alors la formule :

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

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

$$Q_3 = 48,52 \text{ ha.}$$

 $D_5 ?$

$$\frac{N \cdot 7}{10} = \frac{200 \times 7}{10} = 140.$$

On cherche 140 parmi les ncc

cette valeur apparaît dans le tableau, alors on prend $D_5 = e_2 = 30 \text{ ha.}$

5% of échale (Ml) :

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

$$\frac{6950}{2} = 3475$$

On cherche cette valeur parmi les $(n_i c_i) c_i \uparrow$
la 1^{ère} valeur qui le dépasse est 3500
Donc la classe médiale est [30, 50]

on applique la formule :

(3)

$$M\ell = e_{i-1} + \frac{\sum_{i=1}^n n_i c_i - (n_{i-1} q_{i-1}) c_i}{n_i c_i} o_i$$

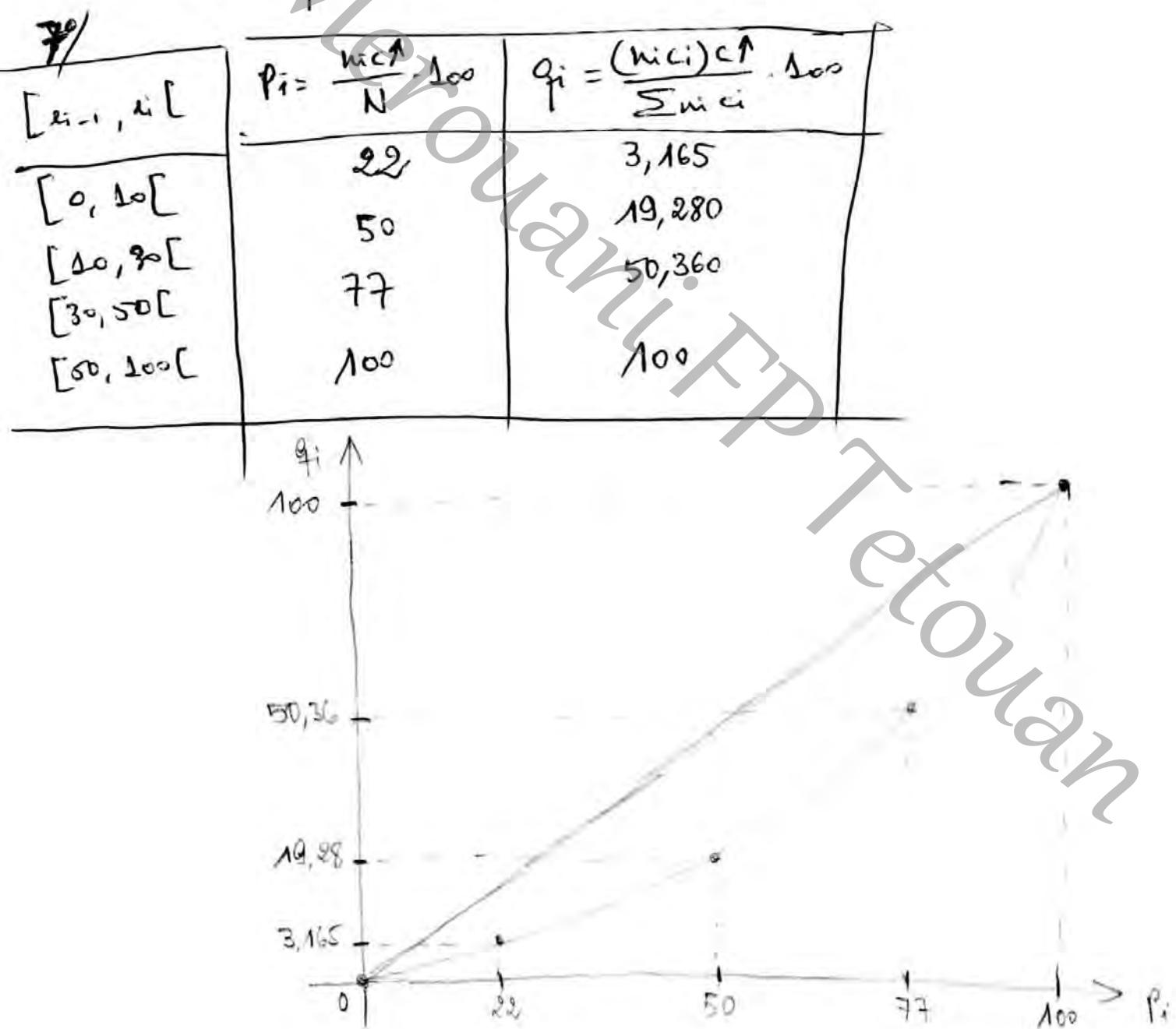
$$= 30 + \frac{3475 - 1340}{2160} 20 = 49,768$$

8) $\Delta M = M\ell - M_e = 49,768 - 30 = 49,768$

$$E = 100$$

$$I = \frac{\Delta M}{E} = \frac{49,768}{100} = 0,49768 \approx 0,498$$

\Rightarrow Equi-distribution.



9%

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

\Rightarrow équidistribution