Université Abdelmalek Essaâdi
Faculté Polydisciplinaire de Tétouan
Département de Statistique et Informatique

Année universitaire : 2010/2011
Sciences économiques et Gestion
Semestre 1

Contrôle final de Statistique Descriptive I
Durée: 2heures

Exercice 1 :
On a relevé pendant 50 quinzaines successives les niveaux de ventes, exprimés en milliers d'unités d'un certain produit cosmétique. Les résultats sont les suivants :

Niveau de vente	[0 ; 5[	[5 ; 10[	[10 ; 12[	[12 ; 20[
Nombre de quinzaines	5	20	15	10

1. Calculer les effectifs cumulés croissants et les fréquences cumulées croissantes de cette série statistique.
2. Donner sa fonction de répartition $F(x)$ et représenter la courbe cumulative.
3. Calculer sa moyenne arithmétique et donner son interprétation économique.
4. Déterminer le mode Mo de cette série, graphiquement et par le calcul.
5. Calculer la médiane Me de cette série statistique en explicitant vos calculs.
6. La série étudiée est-elle symétrique ? Justifier votre réponse.

Exercice 2 :
Une association de la défense des droits de la femme a mené une enquête auprès du personnel ouvrier d'un secteur industriel. Les résultats concernant les salaires annuels nets en milliers de dirhams sont résumés dans le tableau suivant :

Salaire annuel (en milliers 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 moyen des femmes enquêtées est égal à 39500 DH, déterminer l'effectif n_4 de la dernière classe de la distribution du salaire de ces femmes, ainsi que l'effectif total N.
3. Quel est le pourcentage des ouvrières qui ont un salaire annuel inférieur à 42000 DH.
4. Déterminer les quartiles Q_1, Q_2 et Q_3. Calculer l'écart interquartile relatif.
5. Représenter le diagramme de Box & Wiskers correspondant. L'étendue est-il est un bon outil de mesuer la dispersion ? Pourquoi ?
6. Calculer la variance et l'écart-type de cette distribution.
7. Déterminer le coefficient de variation. Conclure.

Bon courage !

2010/2011

Correction du contrôle final de Statistique Descriptive I

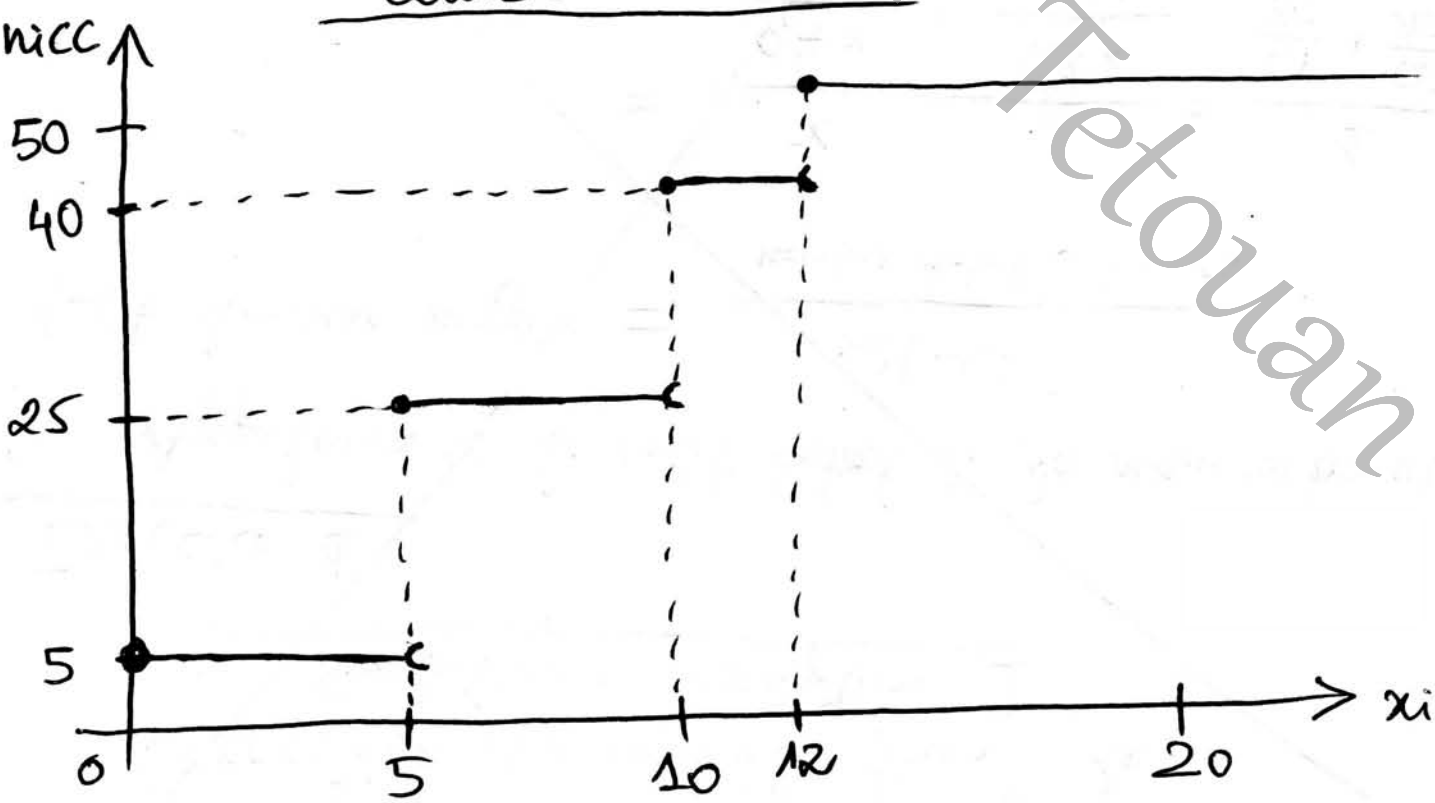
Exercice 1 :

1)

$[e_{i-1}, e_i[$	n_i	n_{icc}	f_i	f_{icc}	c_i	$n_i c_i$	a_i	$h_i = \frac{n_i}{a_i}$
[0,5[	5	5	0,1	0,1	2,5	12,5	5	1
[5,10[	20	25	0,4	0,5	7,5	150	5	4
[10,12[	15	40	0,3	0,8	11	165	2	7,5
[12,20[	10	50	0,2	1	16	160	8	1,25
	N=50		$\Sigma f_i=1$			487,5		

2)

Courbe cumulative



2) La fonction de répartition $F(x)$ est définie par :

$$F(x) = f_i cc \quad \text{pour chaque } e_{i-1} \leq x < e_i$$

ou encore $F(x) = \sum_{x<e_i} f_i$ avec $e_{i-1} \leq x < e_i$

Alors, dans le cas de notre série :

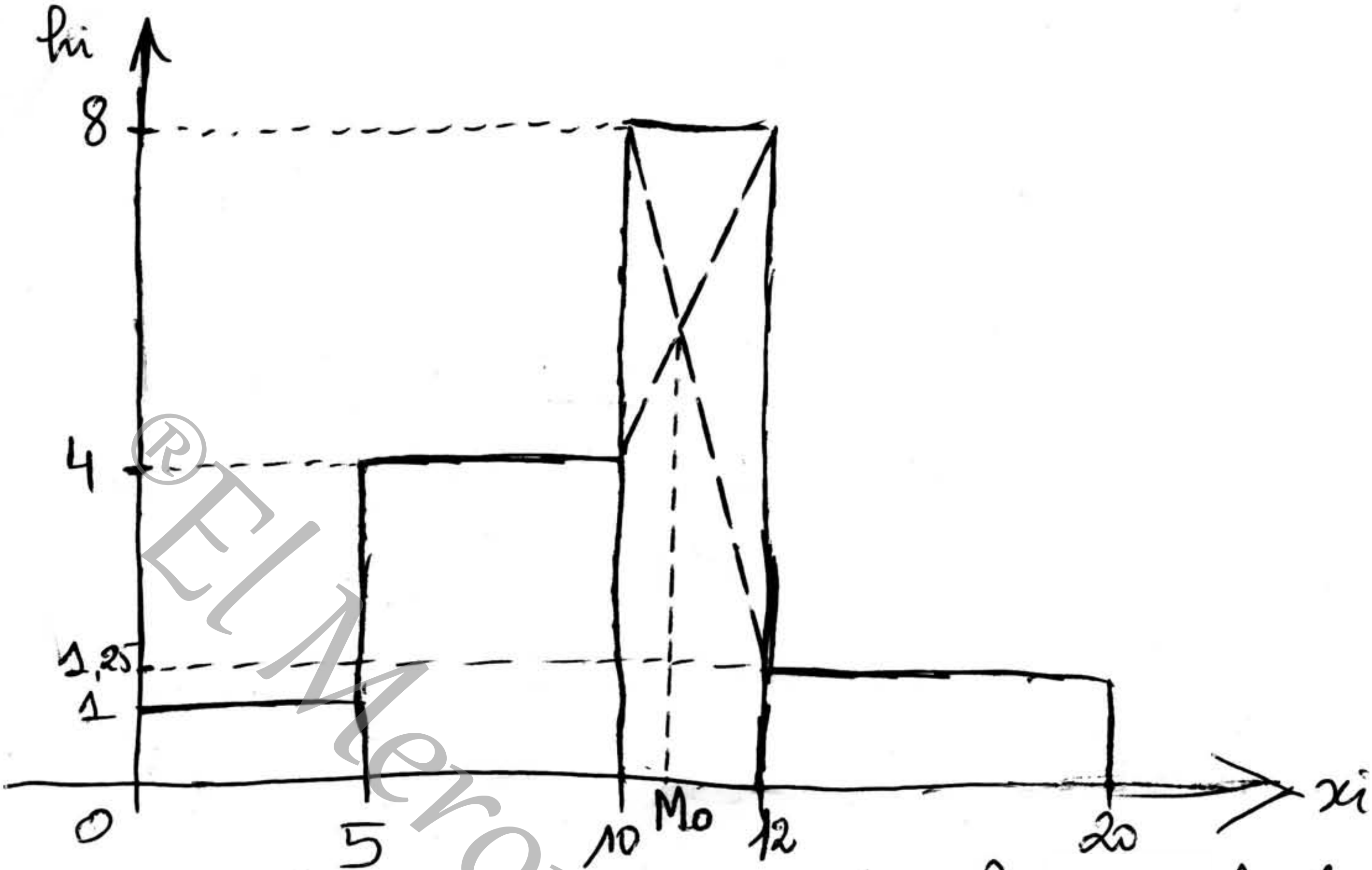
$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 0,1 & \text{si } x \in [0,5[\\ 0,5 & \text{si } x \in [5,10[\\ 0,8 & \text{si } x \in [10,12[\\ 1 & \text{si } x > 12 \end{cases}$$

3) <u>Moyenne arithmétique :</u>

$$\overline{X} = \frac{1}{N}\sum_{i=1}^{4} n_i c_i = \frac{487,5}{50} = 9,75$$ est le niveau de vente moyen pendant ces 50 quinzaines.

4) <u>Détermination du Mode Mo par la méthode graphique :</u>

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 classes 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 sommet du rectangle juste avant et le suivant, la projection sur l'axe des x du point de rencontre des diagonals obtenues (voir graphe en haut) donne la position de Mo parmi les x_i. On a $Mo \in [10,12[$.

Détermination de Mo 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 la plus grande). La classe modale est $[10,12[$, et on applique une des formules suivantes :

(I) $$Mo = e_{i-1} + \frac{h_{i+1}}{h_{i-1} + h_{i+1}} a_i$$

ou (II) $$Mo = e_{i-1} + \frac{h_i - h_{i-1}}{(h_i - h_{i+1}) + (h_i - h_{i-1})} a_i$$

Par application de (I) :

$$M_o = 10 + \frac{1,25}{4 + 1,25} \times 2 = \underline{\underline{10,48}}$$

Par utilisation de (II) :

$$M_o = 10 + \frac{7,5 - 4}{(7,5 - 1,25) + (7,5 - 4)} \times 2 = \underline{\underline{10,72}}$$

5) <u>Calcul de la médiane :</u>

$\frac{N}{2} = \frac{50}{2} = 25$, cette valeur existe exactement parmi les nicc $\Longrightarrow$ la classe médiale est $[5,10] = [e_{i-1}, e_i]$

Alors, on prend $\underline{\underline{M_é = e_i = 10}}$

6) On a : $\bar{x} = 9,75 \neq M_o = 10,72 \neq M_é = 10$

Donc cette série est asymétrique.

Exercice 2 :

$[e_{i-1}, e_i[$	n_i	c_i	n_{icc}	a_i	$(c_i - \bar{x})^2$	$n_i(c_i - \bar{x})^2$	f_i	f_{icc}
[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°) $\overline{x} = 39{,}5$ en milliers de dirhams

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

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

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

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

$\Rightarrow 17{,}5 n_4 = 35$

$\Rightarrow \boxed{n_4 = 2}$ et $\boxed{N = 24}$

4°) Q_1 ?

$\frac{N}{4} = \frac{24}{4} = 6$ on cherche ce 6 parmi les nicc

$\Rightarrow$ il n'existe pas exactement, mais 10 est la $1^{ère}$ valeur qui dépasse ce 6 alors on applique la formule :

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

$$Q_1 = 30 + \frac{6-0}{10} \times 6$$

$$Q_1 = 30 + 3,6 = \underline{33,6}$$

Q_2 ?

$\frac{N}{4} \times 2 = 12$ cette valeur n'existe pas parmi les nicc, mais 18 est la 1ère valeur qui le dépasse, on applique la formule

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

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

Q_3 ? $\frac{N}{4} \times 3 = 18$ cette valeur existe parmi les nicc alors on prend $Q_3 = e_i = \underline{42}$

<u>La dérivation quartile ou le semi_interquartile est</u>

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

<u>L'écart interquartile relatif :</u>

$$\frac{Q_3 - Q_1}{Q_2} = \frac{42 - 33,6}{37,5} = \underline{\underline{0,224}}$$

5)

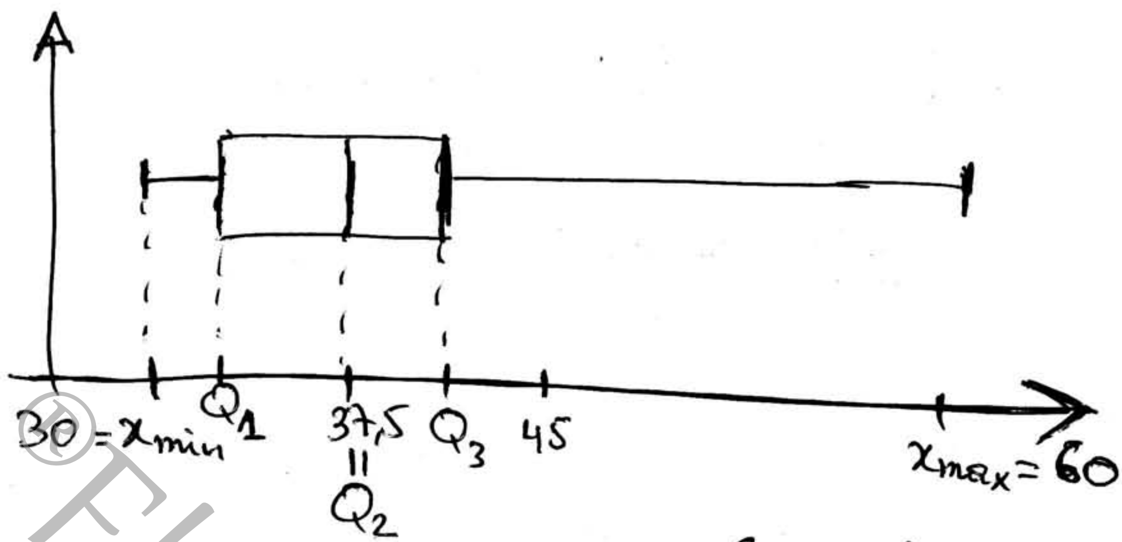


Diagramme de Box & Wiskers

6) Variance :

$$Var(X) = \frac{1}{N} \sum n_i (x_i - \bar{x})^2 = \frac{1326}{24} = \underline{\underline{55,25}}$$

$$\sigma(X) = \sqrt{Var(X)} = \underline{7,43}$$

7°)

$$C_v = \frac{\sigma(X)}{\bar{X}} = \frac{7,43}{39,5} = 0,188$$

$$C_v = 18,8 \%$$

Cette série est peu dispersée (homogène).

3°) On lit directement du tableau des calculs dans la colonne des $f_i cc$ en face de la classe [36,42[on a la valeur 0,7.

Donc le pourcentage des ouvrières qui ont un salaire inférieure à 42000 DH est **70%**