الكلية متعددة التخصصات الناظور
ⵜⴰⵢⵙⵍⴰⵏⵜ ⵜⴰⴳⵜⵙⵏⵉⵏⵜ ⵏ ⵏⴰⴹⵓⵕ
Faculté Pluridisciplinaire de Nador

Université Mohamed Premier
Faculté Pluridisciplinaire -Nador
Département de Mathématiques

Correction de TD N° 2 de statistique descriptive

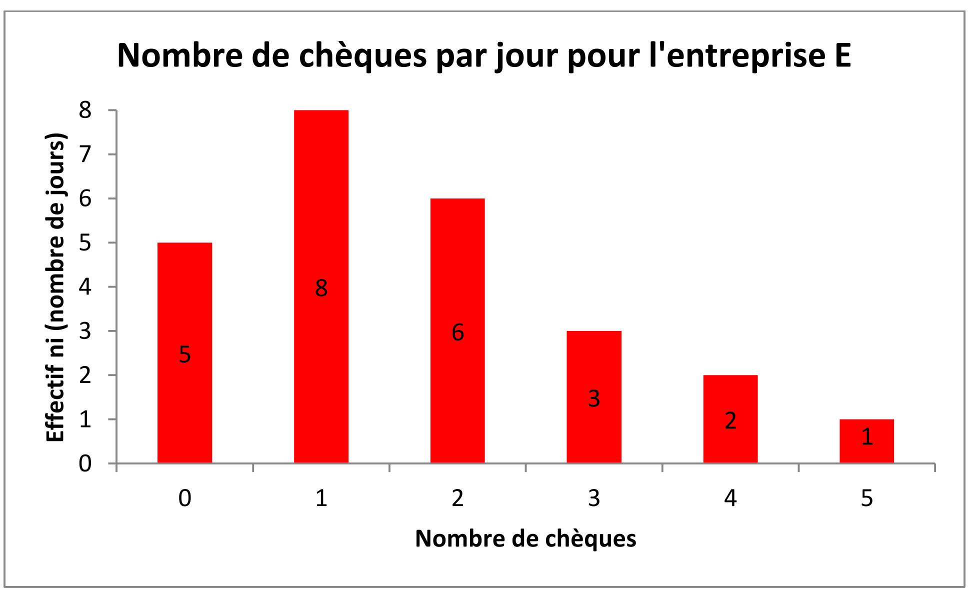
Exercice 1 :

1. **Population :** Les jours de fonctionnement de l'entreprise E.
Taille de l'échantillon: N=25 jours d'observation.
L'individu : 1 jour d'observation.
Le caractère : le nombre des chèques de ''montant élevé : supérieur à 50000 DH''
Nature de caractère : quantitatif discret

2. Distribution en termes d'effectifs

Nombre des chèques x_i	0	1	2	3	4	5	Total
Effectifs ni	5	8	6	3	2	1	25

3.



4. **Les paramètres de tendance centrale :**

<u>Le mode</u> : $M_0=1$

<u>La moyenne</u> : $m=\overline{X}=\frac{\sum_{i=1}^{6} n_i x_i}{N}=\frac{0\times5+1\times8+2\times6+3\times3+4\times2+5\times1}{25}=1.68$

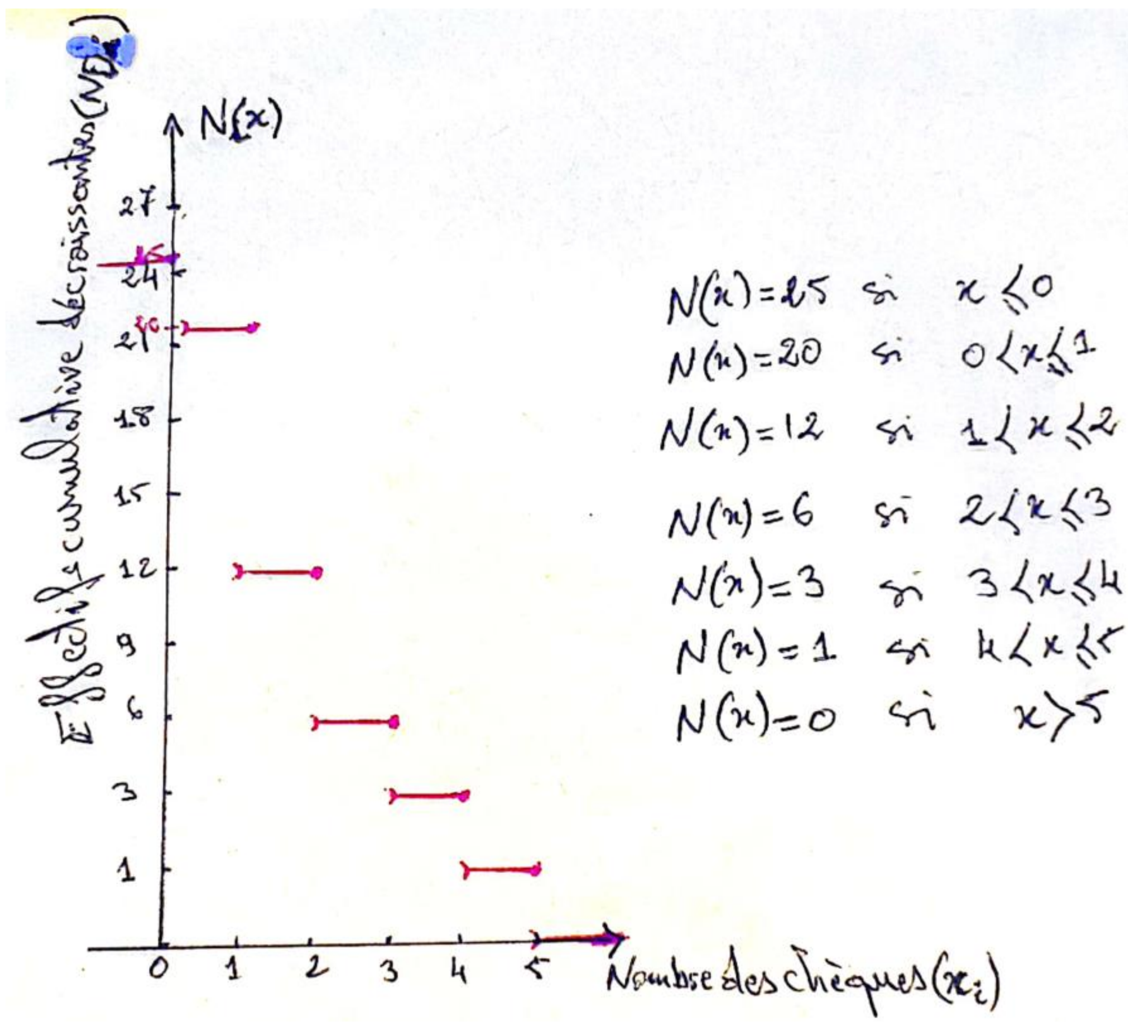
<u>La médiane</u> : (25+1)/2=13 ; La médiane se trouve au rang 13 après avoir rangé les caractères dans un ordre croissant

0 0 0 0 0 1 1 1 1 1 1 1 1 2 2 2 2 2 2 3 3 3 4 4 5

Alors, Me=1

5. Courbe cumulative décroissante :

Nombre des chèques	0	1	2	3	4	5	Total
Effectifs ni	5	8	6	3	2	1	25
ND_i décroissant	25	20	12	6	3	1	



6. Calcul des quantiles

Q_1 :

0 0 0 0 0 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 3 3 3 4 4 5

$N/4=25/4=6.25$

Le premier quartile Q_1 est le terme de rang immédiatement supérieur à 6.25 qui est 7. Alors $Q_1=1$

D_2 :

0 0 0 0 0 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 3 3 3 4 4 5

$2N/10=50/10=5$

Le deuxième décile D_2 est le terme de rang 5. Alors $D_2=3$.

D_7 :

0 0 0 0 0 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 3 3 3 4 4 5

$7N/10=175/10=17.5$

Le septième décile D_7 est le terme de rang immédiatement supérieur à 17.5 qui est 18. Alors $D_7=2$.

C_{35} :

0 0 0 0 0 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 3 3 3 4 4 5

$35N/100=875/100=8.75$

Le trente cinquième centile C_{35} est le terme de rang immédiatement supérieur à 8.75 qui est 9. Alors $C_{35}=1$.

7. L'écart interdécile

0 0 **0** 0 0 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 3 3 3 **4** 4 5

$N/10=25/10=2.5$

Le premier décile D_1 est le terme de rang immédiatement supérieur à 2.5 qui est 3. Alors $D_1=0$. Et on a,

$9N/10=225/10=22.5$

Le neuvième décile D_9 est le terme de rang immédiatement supérieur à 22.5 qui est 23. Alors $D_9=4$.

Alors l'écart interdécile est : $EI=D_9-D_1=4-0=4$.

8. L'étendue :

E=5-0=5

La variance : $S_x^2 = \frac{1}{N}\sum_{i=0}^{5} n_i x_i^2 - \overline{X}^2$

Nombre des chèques xi	0	1	2	3	4	5	Total
x_i^2	0	1	4	9	16	25	
Effectifs ni	5	8	6	3	2	1	25
$n_i*x_i^2$	0	8	24	27	32	25	116

$$S_x^2 = \frac{116}{25} - 1.68^2 \approx 1.82$$

L'écart type :

$$S_x = \sqrt{S_x^2} \approx \sqrt{1.82} \approx 1.35$$

Coefficient de variation :

Le coefficient de variation CV d'une variable statistique est le ratio entre l'écart-type et la moyenne exprimé sous la forme d'un pourcentage:

$$CV = \frac{\sigma}{\mu} \text{ estimé par } CV = \frac{S_x}{m} = \frac{1.35}{1.68} = 80\%$$

Le coefficient de variation est un indicateur de l'homogénéité de la population. On considère qu'un coefficient de variation inférieur à 15% indique que la population est homogène, tandis qu'un coefficient supérieur à 15% indique que les valeurs sont relativement dispersées. Le coefficient de variation est une mesure sans unité et indépendante de l'ordre de grandeur. On peut donc l'utiliser pour comparer la dispersion de variables statistiques avec des ordres de grandeur et des unités différentes.

9. Coefficient d'asymétrie de Yule

$$A_Y = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}$$

3N/4=3*25/4=18.75

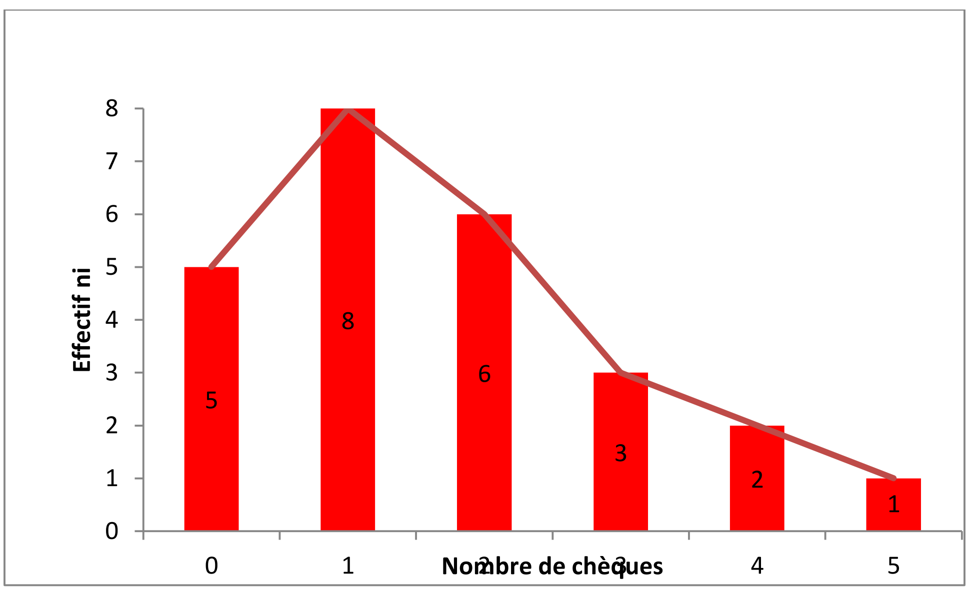
Le troisième quartile Q_3 est le terme de rang immédiatement supérieur à 18.75 qui est 19.

0 0 0 0 0 1 1 1 1 1 1 1 1 2 2 2 2 2 **2** 3 3 3 4 4 5

Alors Q_3=2

Par la suite, $$A_Y = \frac{2+1-2\times 1}{2-1} = 1$$

<u>Interprétation</u> : Puisque A_Y >0, la distribution est allongée à droite (asymétrie en i)

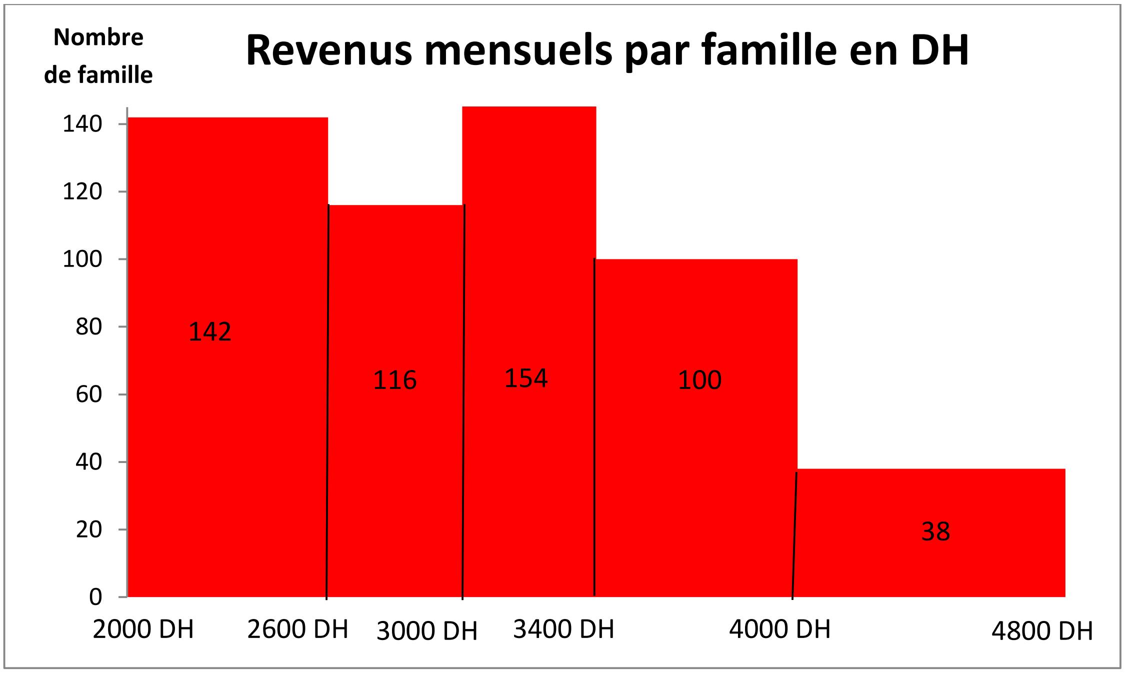


<u>Exercice 2 :</u>

1. **Population :** Les familles d'un quartier d'une ville.
   Taille de l'échantillon: N=550 familles.
   L'individu : une famille de quartier.
   Le caractère : le revenu mensuel en dirhams
   Nature de caractère : quantitatif continu

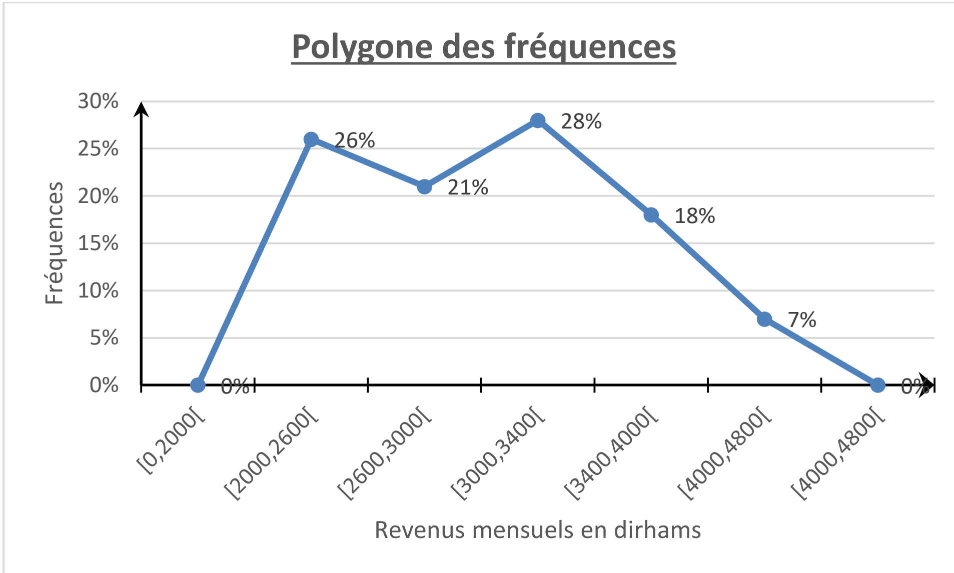
2.

Revenus	[2000,2600[	[2600,3000[	[3000,3400[	[3400,4000[	[4000,4800[
Effectif	142	116	154	100	38



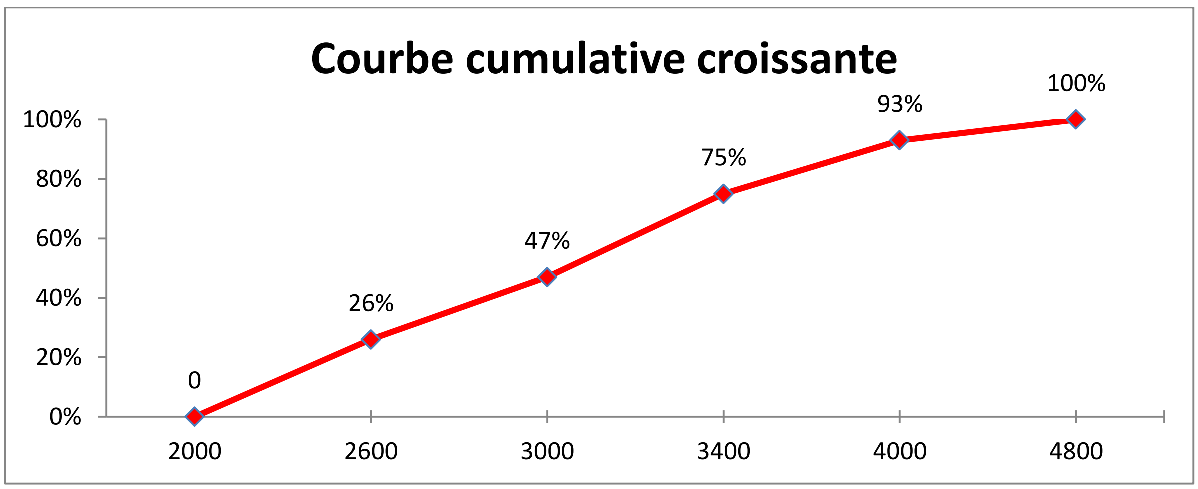
3. Polygone des fréquences

Revenus x_i	[2000,2600[	[2600,3000[	[3000,3400[	[3400,4000[	[4000,4800[	Total
Effectif n_i	142	116	154	100	38	550
Centre c_i	2300	2800	3200	3700	4400	
Fréquence f_i	26%	21%	28%	18%	7%	100%
Fréquence cumulée croissante FCC_i	26%	47%	75%	93%	100%	
Fréquence cumulée décroissante FCD_i	100%	74%	53%	25%	7%	



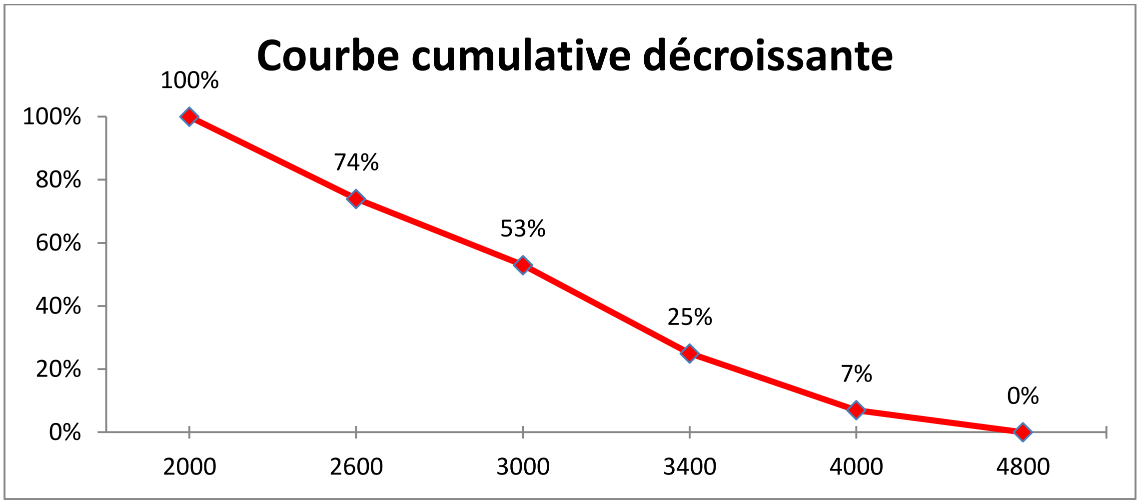
Polygone des fréquences
Fréquences
30%
25%
20%
15%
10%
5%
0%
26%
21%
28%
18%
7%
0%
0%
[0,2000[
[2000,2600[
[2600,3000[
[3000,3400[
[3400,4000[
[4000,4800[
[4000,4800[
Revenus mensuels en dirhams

4.



Courbe cumulative croissante
100%
80%
60%
40%
20%
0%
0
26%
47%
75%
93%
100%
2000
2600
3000
3400
4000
4800

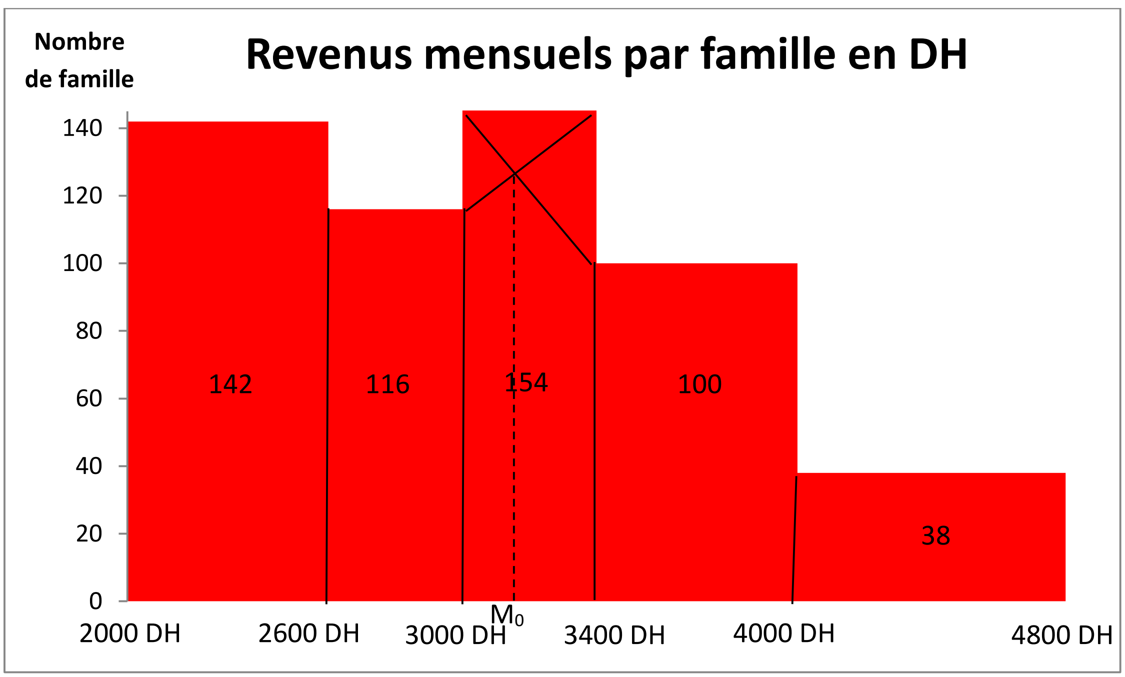
5.



6. <u>Les paramètres de position (paramètres de position)</u>

<u>Le mode</u>

La classe modale : [3000,3400[



$$M_0 = x_i + \frac{D1}{D1 + D2}(x_{i+1} - x_i)$$

$$M_0 = 3000 + \frac{154 - 116}{(154 - 116) + (154 - 100)}(3400 - 3000)$$

Alors,

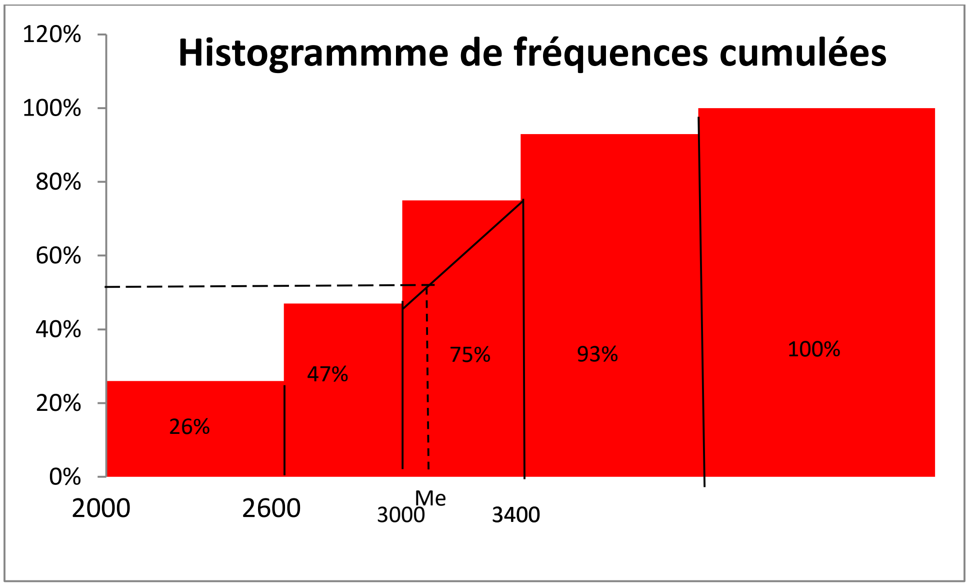
$$M_0 = 3165,22Dhs$$

La moyenne : $m = \overline{X} = \frac{\sum_{i=1}^{6} n_i c_i}{N}$

Revenus x_i	**[2000,2600[**	**[2600,3000[**	**[3000,3400[**	**[3400,4000[**	**[4000,4800[**	**Total**
Effectif n_i	142	116	154	100	38	550
Centre de classe c_i	2300	2800	3200	3700	4400	
n_i*c_i	326600	324800	492800	370000	167200	1681400

$$m = \overline{X} = \frac{1681400}{550} = 3057.09\text{Dhs}$$

La médiane :



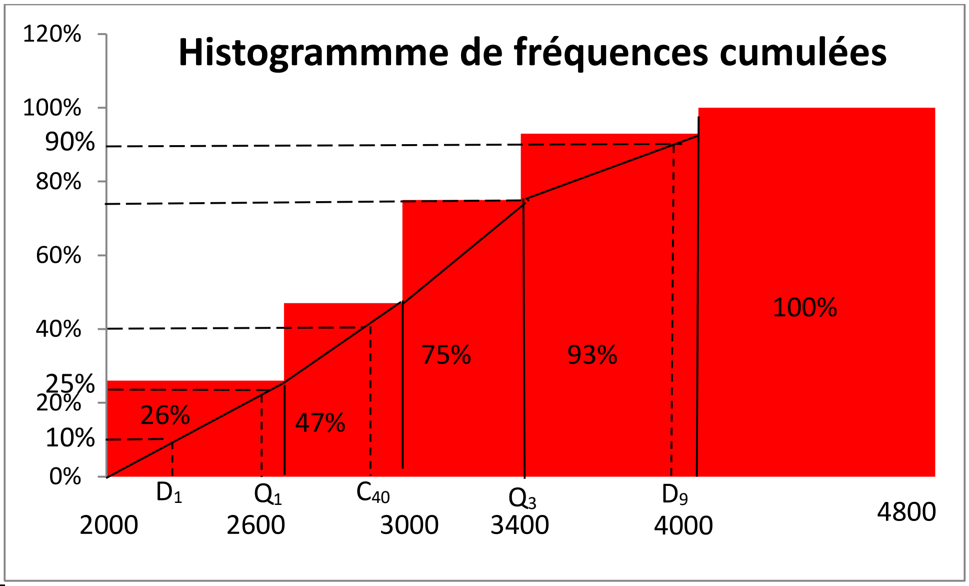
$$\frac{3400 - 3000}{75 - 47} = \frac{M_e - 3000}{50 - 47}$$

$$\frac{400}{28} = \frac{M_e - 3000}{3}$$

Alors

$$M_e = 3042,86Dhs$$

7. Les quantiles



-

$$\frac{2600-2000}{26-0}=\frac{Q1-2000}{25-0}$$

$$\frac{600}{26}=\frac{Q1-2000}{25}$$

Alors

$$Q_1=2576,92Dhs$$

Et on a,

$$Q_3=3400Dhs$$

Et

$$Q_2=M_e=3042,86Dhs$$

On a aussi

$$\frac{3000-2600}{47-26}=\frac{C40-2600}{40-26}$$

$$\frac{400}{21}=\frac{C40-2600}{14}$$

Alors

$$C_{40}=2866,67Dhs$$

Par la même manière,

$$\frac{2600-2000}{26-0}=\frac{D1-2000}{10-0}$$

$$\frac{600}{26} = \frac{D1 - 2000}{10}$$

Alors

$$D_1 = 2230,77Dhs$$

Et

$$\frac{4000 - 3400}{93 - 75} = \frac{D9 - 3400}{90 - 75}$$

$$\frac{600}{18} = \frac{D9 - 3400}{15}$$

Alors

$$D_9 = 3900Dhs$$

8. <u>Ecart interdécile</u> :

$$EID = D_9 - D_1 = 3900 - 2230,77 = 1669,23Dhs$$

9. <u>Etendue</u>

$$E = 4800 - 2000 = 2800Dhs$$

<u>La variance</u>

Revenus x_i	[2000,2600[	[2600,3000[	[3000,3400[	[3400,4000[	[4000,4800[	Total
Effectif n_i	142	116	154	100	38	550
Centre de classe c_i	2300	2800	3200	3700	4400	
$n_i*c_i^2$	751180000	909440000	1576960000	1369000000	735680000	5342260000

$$S_x^2 = \frac{1}{N}\sum_{i=1}^{q} n_i c_i^2 - \overline{X}^2$$

$$S_x^2 = \frac{5342260000}{550} - 3057.09^2 = 367400,73$$

<u>L'écart type :</u> $S_x = \sqrt{S_x^2} = 606,14$

<u>Le coefficient de variation :</u>

$$CV = \frac{\sigma}{\mu} \text{ estimé par } CV = \frac{s_x}{m} = \frac{606,14}{3057,09} = 19,83\%$$

10. Le coefficient d'asymétrie de Pearson

$$A_p = \frac{\overline{X} - M_o}{S_x}$$

Alors

$$A_p = \frac{3057.09 - 3165,22}{606,14} = -0.18$$

- Interprétation : Puisque le coefficient d'asymétrie est négatif alors la distribution est allongée à gauche. Cette allongement est légère, on peut dire aussi que la distribution est symétrique.