

LA STATISTIQUE DESCRIPTIVE SIMPLIFIEE

CHAPITRE 2 LES PRANCIPALES CARACTERISTIQUES D'UNE SERIE

**AUTEUR : MATLAYA MOHAMED
PROFESSEUR AU COMPLEXE DE
FORMATION MAAMORA DE KENITRA**

CHAPITRE 2
LES PRINCIPALES CARACTERISTIQUES D'UNE SERIE

INTRODUCTION

Avec la représentation graphique nous avons vu comment synthétiser une série avec image.

Dans ce chapitre nous allons voir comment synthétiser une série par quelques chiffres. Ces nombres sont appelés caractéristiques d'une série.

Exemple : Soit les séries suivantes :

- Série1 : 78 - 79 - 80 - 83
- Série2 : 60 - 70 - 80 - 90 - 100
- Série3 : 1 - 1 - 1 - 1 - 396

Les séries ont toutes la moyenne 80 même si elles sont très différentes les unes que les autres.

Les valeurs de la 1^{ère} série sont proches de la moyenne alors que celles de la 3^{ème} sont éloignées de la moyenne.

Il y a donc nécessité, pour résumer une série de données de la présenter en 2 types de caractéristiques :

- Les caractéristiques de valeurs centrales.
- Les caractéristiques de dispersion.

SECTION 1 : LES CARACTERISTIQUES DE VALEUR CENTRALE :

I) LES MOYENNES :

A- LA MOYENNE ARITHMETIQUE :

1) Moyenne arithmétique simple :

Etant donnée n observations qu'on va appeler X₁, X₂, X₃,.....X_i,...X_n on appelle une moyenne arithmétique simple le nombre : \bar{x}

$$\bar{x} = \frac{\text{Somme de toutes les observations}}{\text{Le nombre d'observations}}$$

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_i + \dots + x_n}{n_1 + n_2 + \dots + n_i + \dots + n_n}$$

C'est la moyenne arithmétique simple :

$$\bar{x} = \frac{\sum x_i}{N}$$

Exemple : soit la série de notes suivante : 2 - 6 - 12 - 10 - 12 - 10 - 10 - 6

$$\bar{x} = \frac{2 + 6 + 12 + 10 + 12 + 10 + 10 + 6}{1 + 1 + 1 + 1 + 1 + 1 + 1 + 1}$$

$$\bar{x} = \frac{\sum x_i}{N} = \frac{68}{8} = 8,5$$

2) La moyenne arithmétique pondérée :

Lorsque les observations sont groupées c'est une moyenne arithmétique pondérée.

La moyenne arithmétique pondérée s'écrit :

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_i + \dots + x_n}{n_1 + n_2 + \dots + n_i + \dots + n_n}$$

Avec :

$$\bar{x} = \frac{\sum ni x_i}{\sum ni}$$

Exemple : soit la série des notes des élèves la classe qui peut être présentée de la manière suivante :

Notes des élèves	Elèves
8	1
10	8
14	9
15	6
16	4
18	2
Total	30

Le calcul de la moyenne se fait à l'aide du tableau suivant :

Notes des élèves x_i	Elèves n_i	$N_i \cdot x_i$
8	1	8
10	8	80
14	9	126
15	6	90
16	4	64
18	2	36
Total	30	404

$$\bar{x} = \frac{\sum ni x_i}{\sum ni} = \frac{404}{30} = 13,46$$

Donc la note moyenne de la classe est 13,46 points

3) Moyenne arithmétique de classes :

Lorsque les observations sont groupées dans une série de classes, il faut calculer d'abord les centres de classes : ces centres de classe prennent la variable x_i .

La moyenne arithmétique pondérée s'écrit :

$$\bar{x} = \frac{\sum ni x_i}{\sum ni}$$

Exemple :

Le tableau ci-dessous recense les montants des transactions effectuées par des clients pour la période du 15/09/n au 03/11/n dans un point de vente de l'entreprise :

Transactions	Nombre de clients
[0 - 200[34
[200 - 400[60
[400 - 600[23
[600 - 800[30
[800 - 1000[24
[1000 - 1200[13
1200 - 1400[11
Total	195

On aura pour calculer cette moyenne :

Transactions	Nombre de clients n_i	x_i	$n_i \cdot x_i$
[0 - 200[34	$(0 + 200) / 2 = 100$	3400
[200 - 400[60	$(200 + 400)/2 = 300$	18000
[400 - 600[23	$(400 + 600)/2 = 500$	11500
[600 - 800[30	$(600 + 800)/2 = 700$	21000
[800 - 1000[24	$(800 + 1000)/2 = 900$	21600
[1000 - 1200[13	$(1000 + 1200) / 2 = 1100$	14300
1200 - 1400[11	$(1100 + 1400)/2 = 1300$	14300
Total	195		104100

$$\bar{x} = \frac{\sum n_i x_i}{\sum n_i} = \frac{104100}{195} = 533.84$$

Les montants moyens des transactions effectuées par des clients est : 533,84dh

4) Méthode des simplifications des calculs

a) Moyenne arithmétique par changement d'origine :

Lorsque les calculs sont compliqués, on peut les simplifier en précédant à un changement de variable

Le changement d'origine consiste à poser : $x_i = x'_0 + x'_i$; Généralement x_0 correspond à la valeur la plus fréquente dans la série statistique c'ad celle qui a le plus grand effectif.

x'_0 : nouvelle origine

x'_i : nouvelle variable = $x_i - x_0$

Donc \bar{x} s'écrit alors :

$$\bar{x} = \frac{\sum ni x_i}{\sum ni} = \frac{\sum ni (x'_0 + x'_i)}{\sum ni}$$

$$\bar{x} = x_0 + \bar{x}'$$

Par changement d'échelle : Toute variable x_i peut s'écrire : $x_i = a x'_i$

Avec : a = nouvelle échelle et x'_i = nouvelle variable

Exemple :

Reprenant l'exemple précédent, on aura pour $x_0 = 300$ car elle a été répétée 60 fois

Transactions	Nombre de clients n_i	x_i	$x'_i = x_i - x_0$	$n_i x'_i$
[0 - 200[34	100	100 - 300 = (-200)	-6800
[200 - 400[60	300	300 - 300 = 0	0
[400 - 600[23	500	500 - 300 = 200	4600
[600 - 800[30	700	700 - 300 = 400	12000
[800 - 1000[24	900	900 - 300 = 600	14400
[1000 - 1200[13	1100	1100 - 300 = 800	10400
1200 - 1400[11	1300	1300 - 300 = 1000	11000
Total	195			45600

La moyenne arithmétique :

$$\bar{x}' = \frac{\sum ni x'_i}{\sum ni} = \frac{45\ 600}{195} = 233,84$$

$$\bar{x} = x_0 + \bar{x}' = 300 + 233,84 = 533,84$$

b) Moyenne arithmétique par changement d'origine et d'échelle :

$$\bar{x} = X_0 + a \bar{x}'$$

Avec :

X_0 = nouvelle origine,

a = est une constante (généralement c'est le diviseur commun),

x_i = nouvelle variable

Donc :

$$\bar{x} = \frac{\sum ni x_i}{\sum ni} = \frac{\sum ni (x_0 + ax'_i)}{\sum ni}$$

$$\bar{x} = x_0 + a \bar{x}'$$

Transactions	Nombre de clients n_i	x_i	$x'_i = \frac{x_i - x_0}{a}$	$n_i x'_i$
[0 - 200[34	100	$(100 - 300)/200 = (-1)$	-34
[200 - 400[60	300	$(300 - 300)/200 = 0$	0
[400 - 600[23	500	$(500 - 300)/200 = 1$	23
[600 - 800[30	700	$(700 - 300)/200 = 2$	60
[800 - 1000[24	900	$(900 - 300)/200 = 3$	72
[1000 - 1200[13	1100	$(1100 - 300)/200 = 4$	52
1200 - 1400[11	1300	$(1300 - 300)/200 = 5$	55
Total	195			228

X_0 = nouvelle origine = 300

a = est une constante (généralement c'est le diviseur commun) = 200

x'_i = nouvelle variable

$$\bar{x}' = \frac{\sum ni x'_i}{\sum ni} = 1,169$$

$$\bar{x} = x_0 + a \bar{x}'$$

$$\bar{x} = 300 + (200 \times 1,169)$$

$$\bar{x} = 300 + 233,84 = 533,8$$

5) Moyenne arithmétique des fréquences :

$$\bar{x} = \frac{\sum fi x_i}{\sum fi}$$

Exemple :

Transactions	Nombre de clients n_i	x_i	f_i	$f_i \cdot x_i$
[0 - 200[34	100	0,17	17,44
[200 - 400[60	300	0,31	92,31
[400 - 600[23	500	0,12	58,97
[600 - 800[30	700	0,15	107,69
[800 - 1000[24	900	0,12	110,76
[1000 - 1200[13	1100	0,07	73,33
1200 - 1400[11	1300	0,06	73,33
Total	195		1	533,84

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{533,84}{1} = 533,84$$

B) LA MOYENNE GEOMETRIQUE :

Étant donnée n observations connues individuellement ($x_1, x_2, x_3, \dots, x_n$) on appelle moyenne géométrique simple de N observations la grandeur G telle que :

$$G = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_i \cdot \dots \cdot x_n}$$

Exemple :

Nous avons la série suivante : 2 ; 6 ; 6 ; 10 ; 10 ; 10 ; 12 ; 12

$$G = \sqrt[8]{2 \cdot 6 \cdot 6 \cdot 10 \cdot 10 \cdot 10 \cdot 12 \cdot 12}$$

$$\text{ET } \log \text{ de } G = \frac{(\log 2 \cdot \log 6 \cdot \log 6 \cdot \log 10 \cdot \log 10 \cdot \log 10 \cdot \log 12 \cdot \log 12)}{8}$$

log de G = 0,87696 d'où G = 7,53

Lorsque les observations sont groupées ; chaque valeur **X_i** sera pondérée par l'effectif correspondant, la moyenne géométrique s'écrit :

$$G^{\sum n_i} = x_1^{n_1} \cdot x_2^{n_2} \cdot \dots \cdot x_k^{n_k}$$

$$\text{ou } G = \sqrt[\sum n_i]{x_1^{n_1} \cdot x_2^{n_2} \cdot \dots \cdot x_k^{n_k}}$$

$$\log G = \frac{\sum n_i \log x_i}{\sum n_i} =$$

Exemple :

Transactions	Nombre de clients ni	x_i	log xi	ni log xi
[0 - 200[34	100	2	68,00
[200 - 400[60	300	2,477121	148,63
[400 - 600[23	500	2,69897	62,08
[600 - 800[30	700	2,845098	85,35
[800 - 1000[24	900	2,954242	70,90
[1000 - 1200[13	1100	3,041392	39,54
1200 - 1400[11	1300	3,113943	34,25
Total	195		19,1308	508,75

$$\log G = \frac{\sum ni \log x_i}{\sum ni} = \frac{508,75}{195} = 2,608973$$

Et G = 406,41

C) LA MOYENNE HARMONIQUE :

Étant donnée n observations connues individuellement $x_1, x_2, x_3 \dots x_n$ on appelle moyenne harmonique le nombre H tel que :

$$\frac{1}{H} = \frac{1}{N} \left(\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \dots + \frac{1}{x_n} \right)$$

Si les observations sont groupées la moyenne harmonique pondérée s'écrit :

$$\frac{1}{H} = \frac{\sum ni (1/x_i)}{\sum ni}$$

$$\text{Et } H = \frac{\sum ni}{\sum ni (1/x_i)}$$

Exemple :

Transactions	Nombre de clients ni	x_i	$1/x_i$	$ni (1/x_i)$
[0 - 200[34	100	0,01000	0,340
[200 - 400[60	300	0,00333	0,200
[400 - 600[23	500	0,00200	0,046
[600 - 800[30	700	0,00143	0,043
[800 - 1000[24	900	0,00111	0,027
[1000 - 1200[13	1100	0,00091	0,012
1200 - 1400[11	1300	0,00077	0,008
Total	195		0,01955	0,676

$$H = \frac{\sum ni}{\sum ni (1/x_i)} = \frac{195}{0,676} = 288,54$$

D) LA MOYENNE QUADRATIQUE :

Etant donné n observations connues individuellement par : $x_1 ; x_2 ; x_3 \dots\dots x_n$

La moyenne quadratique simple c'est la quantité Q telle :

$$Q^2 = \frac{1}{N} (X^2_1 + X^2_2 + X^2_3 + \dots\dots\dots X^2_n)$$

Si les observations sont groupées, la moyenne quadratique s'écrit :

$$Q^2 = \frac{n_1x^2_1 + n_2x^2_2 + n_3x^2_3 + \dots\dots\dots n_nx^2_n}{N}$$

Donc: $Q^2 = \frac{\sum n_i x^2_i}{\sum ni}$

Avec Q = —

Exemple :

Transactions	Nombre de clients ni	x_i	x_i^2	$ni x_i^2$
[0 - 200[34	100	10000	340000
[200 - 400[60	300	90000	5400000
[400 - 600[23	500	250000	5750000
[600 - 800[30	700	490000	14700000
[800 - 1000[24	900	810000	19440000
[1000 - 1200[13	1100	1210000	15730000
1200 - 1400[11	1300	1690000	18590000
Total	195			79950000

$$Q^2 = \frac{\sum n_i x_i^2}{\sum n_i} = \frac{79950000}{195} = 410000$$

$$Q = \sqrt{\quad} = 640,31$$

CONCLUSIONS :

En théorie, aucune moyenne n'est meilleure que l'autre. L'utilisation de telle moyenne dépend du problème posé.

Une même distribution statistique peut permettre la détermination d'une moyenne arithmétique, d'une moyenne géométrique, d'une moyenne harmonique et d'une moyenne quadratique de mesures respectives différentes les unes des autres. Les calculs ne doivent être entrepris que s'ils conduisent à des résultats ayant une signification concrète.

Mais d'une manière générale, on retient la moyenne arithmétique.

On a toujours : $H < G < \bar{X} < Q$

II) LA MEDIANE (Me)

A) DEFINITION :

On appelle médiane d'une série classée par ordre croissant ou décroissant, la valeur du caractère qui partage les effectifs en deux parties égales.

En d'autres termes, c'est la valeur du caractère telle que la moitié des effectifs lui est supérieure et l'autre moitié lui est inférieure.

B) CALCUL DE LA MEDIANE :

1) Cas d'une variable discrète :

- Lorsque le nombre des observations d'une série statistique classées par ordre croissant ou décroissant est un nombre impair tels que par exemple : 18 - 22 - 33 - **45** - 55 - 60 - 65

Avec $N = 7$, la médiane correspond à la valeur située dans le rang de :

$(N+1)/2$ c'ad $(7+1)/2 =$ la valeur du 4^{ème} rang coïncide avec la valeur **45**.

Ou bien, la médiane correspond à la valeur qui se situe au centre du nombre des valeurs des observations de la série statistique : donc **Me = 45**

- Lorsque le nombre des observations d'une série statistique classées par ordre croissant ou décroissant est un nombre pair il n'y a pas de médiane. Il y'a un intervalle médian tels que par exemple :

12 - 16 - 21 - **26 - 30** - 36 - 40 - 50

Avec $N = 8$, l'intervalle médian est [26 - 30]

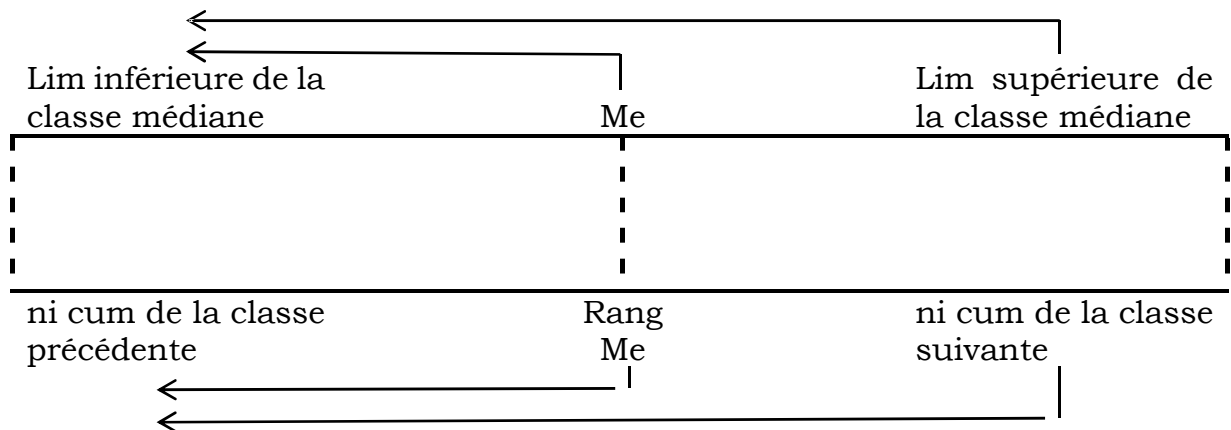
Il convient cependant de donner comme valeur à la médiane, le centre de l'intervalle médian.

Ainsi **Me = $\frac{26 + 30}{2} = 28$**

2) Cas d'une série de classes :

Pour calculer la médiane, il faut suivre les étapes suivantes :

- Calculer les effectifs cumulés croissants,
- Calculer le rang de la médiane : $\text{rang Me} = \sum ni/2$,
- Repérer le rang de la Me dans la colonne des ni cumulés croissants ou décroissants,
- Déterminer la classe médiane,
- Interpoler linéairement pour situer exactement la valeur de la médiane



Exemple :

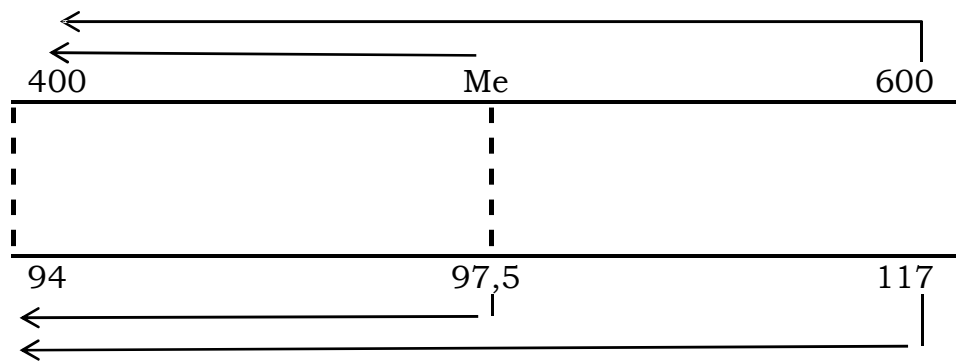
Partons de l'exemple suivant :

Soient les primes mensuelles de 195 employés d'une entreprise industrielle résumées dans le tableau ci-dessous.

Primes mensuelles		Employés ni	ni cumulés croissants
[0 - 200]		34	34
[200 - 400]	Me	60	94
[400 - 600]		23	117
[600 - 800]		30	147
[800 - 1000]		24	171
[1000 - 1200]		13	184
[1200 - 1400]		11	195
Totaux			195

97,5

- 1) Calculer les effectifs cumulés croissants : voir tableau ci-dessus,
- 2) Calculer le rang de la médiane : rang Me = $\sum ni/2 = 195/2 = 97,5$
- 3) Repérer le rang de la Me dans la colonne des ni cumulés croissants ou décroissants,
- 4) Déterminer la classe médiane : le rang de la médiane se situe entre la valeur 94 et 117 càd $94 < 97,5 < 117$ donc certainement la valeur de la médiane se trouve entre 400 et 600 donc la classe médiane est [400 - 600],
- 5) Il faut interpoler linéairement pour situer exactement la valeur de la médiane.



$$Me = \frac{Me - 400}{600 - 400} = \frac{97,5 - 94}{117 - 94}$$

$$Me = \frac{Me - 400}{200} = \frac{3,5}{23}$$

$$Me = 430,40dh$$

Signification :

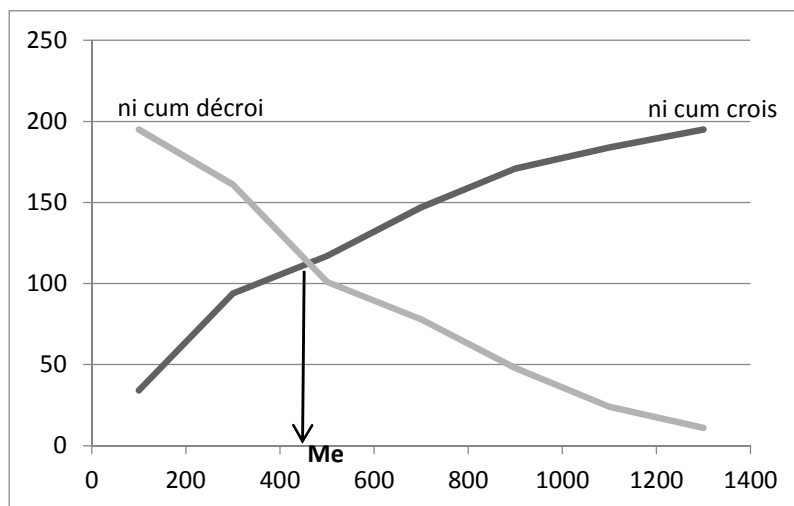
La moitié des employés touchent une prime mensuelle supérieure à 430,40dh et l'autre moitié des employés touchent une prime inférieure à 430,40dh.

C) DETERMINATION GRAPHIQUE DE LA MEDIANE :

On peut déterminer la médiane l'aide du graphique des séries cumulées croissantes et décroissantes,

Exemple :

Primes mensuelles	X_i	Employés n_i	ni cumulés croissants	ni cumulés décroissants
[0 - 200]	100	34	34	195
[200 - 400]	300	60	94	161
[400 - 600]	500	23	117	101
[600 - 800]	700	30	147	78
[800 - 1000]	900	24	171	48
[1000 - 1200]	1100	13	184	24
[1200 - 1400]	1300	11	195	11
Totaux		195		



III) LA MEDIALE :

A) DEFINITION :

On appelle médiale (MI), la valeur du caractère qui partage en deux parties égales le produit cumulé des valeurs de la variable par les fréquences correspondantes.

B) CALCUL DE LA MEDIALE :

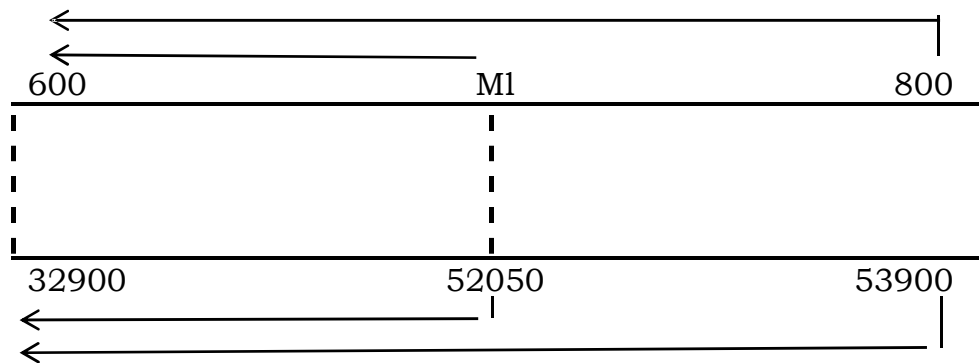
Pour calculer la médiale, il faut suivre les étapes suivantes :

- 1) Calculer les n_i , x_i cumulés croissants,
- 2) Calculer le rang de la médiale : $\text{rang Me} = \sum n_i x_i / 2$,
- 3) Repérer le rang de la MI dans la colonne des $n_i x_i$ cumulés croissants ou décroissants,
- 4) Déterminer la classe médiale,
- 5) Interpoler linéairement pour situer exactement la valeur de la médiale

Exemple :

Primes mensuelles		Xi	Employés ni	ni.xi	ni.xi cum croissants	
[0 - 200]	}	100	34	3400	3400	Rang MI
[200 - 400]		300	60	18000	21400	
[400 - 600]		500	23	11500	32900	
[600 - 800]		700	30	21000	53900	
[800 - 1000]		900	24	21600	75500	
[1000 - 1200]		1100	13	14300	89800	
[1200 - 1400]		1300	11	14300	104100	
Totaux				195	104100	

- 1) Calculer les ni.xi cumulés croissants : voir tableau ci-dessus,
- 2) Calculer le rang de la médiane :
rang MI = $\sum ni.xi / 2 = 104100/2 = 52050$
- 3) Repérer le rang de la MI dans la colonne des ni cumulés croissants ou décroissants,
- 4) Déterminer la classe médiale : le rang de la médiale se situe entre la valeur 32900 et 53900 càd $32900 < 52050 < 53900$ donc certainement donc la classe médiale est [600 - 800],
- 5) Il faut interpoler linéairement pour situer exactement la valeur de la médiale.



$$MI = \frac{MI - 600}{800 - 600} = \frac{52050 - 32900}{53900 - 32900}$$

$$MI = \frac{MI - 600}{200} = \frac{19150}{21000}$$

ML = 782,4dh

Signification :

782,4dh c'est la prime telle que la moitié de la masse salariale a permis de payer des employés qui touchent moins de 782,4dh et la moitié de la masse salariale a permis de payer des employés qui touchent plus de 782,4dh.

Remarque : Pour que la médiale ait un sens, il faut que les sommes des $n_i.x_i$ aient un sens (fortune ; revenus ; surface exploitée...).

IV) LE MODE : M_o

A) DEFINITION :

Le mode est la valeur de la variable qui présente l'effectif (ou la fréquence) le plus élevé. C'est la valeur du caractère la plus répétée dans une série statistique.

B) CALCUL DU MODE :

1) Cas du Caractère quantitatif discret : série simple

Exemple n°1 :

Dans la série suivante on a :

Nombre d'enfants à charge	Effectifs
0	4
1	15
2	29
3	18
4	10
5	3
6	1
Total	80

Le mode est égal à 2 enfants, car l'effectif correspondant est 29 qui le plus élevé de tous les effectifs observés.

Dans cette série on a un seul mode : c'est une **série uni modale**

Exemple n°2 :

Dans la série suivante on a :

Nombre d'enfants à charge	Effectifs
0	4
1	15
2	32
3	18
4	32
5	13
6	6
Total	120

Le mode est égal à 2 enfants et 4 enfants, car l'effectif correspondant à ces deux modes est 32 qui sont le plus élevé de tous les effectifs observés.

Dans cette série on a deux modes : c'est une **série bi modale**

Remarque :

On peut avoir autant de modes dans une seule série mais ils deviennent insignifiants.

2) Cas du caractère quantitatif continu : série de classes

Dans la série suivante on a :

Primes mensuelles	Employés ni
[0 - 200[34
[200 - 400[60
[400 - 600[23
[600 - 800[30
[800 - 1000[24
[1000 - 1200[13
[1200 - 1400[11
Totaux	195

Dans cette série on a le mode compris entre 200 et 400 : c'est une **classe modale** qui correspond au plus grand effectif de la série : soit **60**.

On peut chercher à connaître le mode d'une manière plus précise en utilisant l'une des trois méthodes ci-dessous :

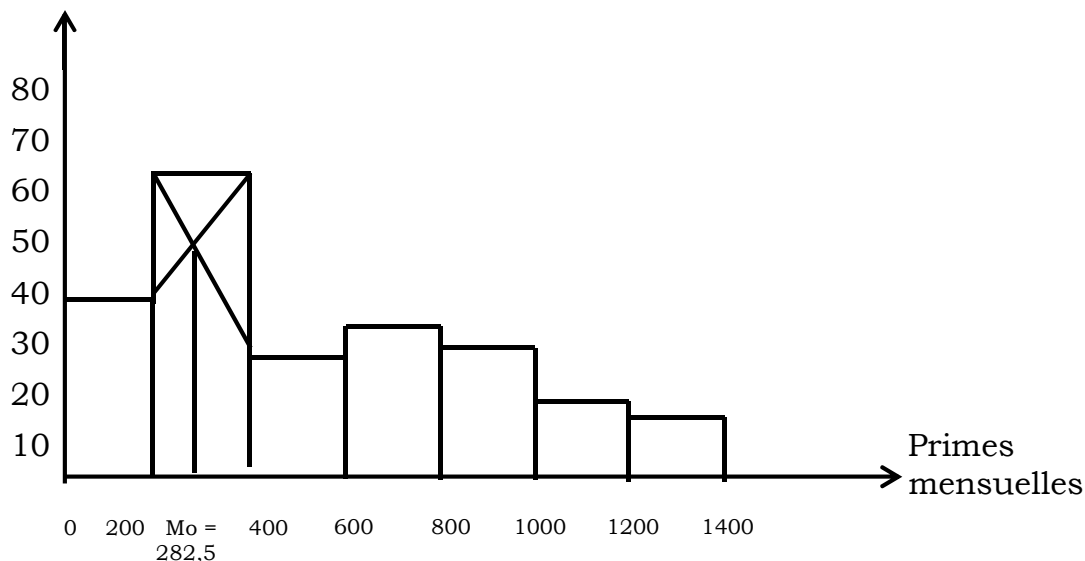
a) Détermination du mode en calculant le centre de la classe modale :

$$Mo = \frac{\text{limite inférieure de la classe modale} + \text{limite supérieure de la classe}}{2}$$

$$Mo = \frac{200 + 400}{2} = 300$$

b) Détermination graphique du mode :

Employés ni



Remarque :

Ne pas oublier en construisant l'histogramme de corriger les effectifs quand les classes ne sont pas égales.

c) Détermination du mode algébrique :

Cette méthode consiste à appliquer la formule suivante :

$$Mo = L_1 + \left[\frac{d_1}{d_1 + d_2} \times i \right]$$

Avec :

- **L₁** : La limite Inférieure de classe modale,
- **d₁** : La différence entre les effectifs de la classe modale et les effectifs de classe précédente,
- **d₂** : La différence entre les effectifs de classe modale et les effectifs de classe suivante
- **i** : L'intervalle de la classe modale ou l'amplitude de la classe modale.

Exemple :

Primes mensuelles	Employés ni	
[0 - 200[34	} d ₁ = 60 - 34 = 26
[200 - 400[60	
[400 - 600[23	} d ₂ = 60 - 23 = 37
[600 - 800[30	
[800 - 1000[24	
[1000 - 1200[13	
[1200 - 1400[11	
Totaux	195	

L₁ : 200

d₁ : 60 - 34 = 26

d₂ : 60 - 23 = 37

i : 400 - 200 = 200

Donc :

$$Mo = 200 + \frac{60 - 34}{(60 - 34) + (60 - 23)} \times 200$$

$$Mo = 200 + \frac{26}{26 + 37} \times 200$$

$$Mo = 282,5dh$$

Signification :

282,5dh c'est la prime que perçoit la majorité des employés.

C) AVANTAGES DU MODE

- Sa détermination est aisée par le graphique.
- Son intérêt est évident puisqu'il désigne la valeur de la variable qui revient le plus souvent.

D) INCONVENIENTS DU MODE :

- Il n'a de signification véritable que si l'effectif correspondant est nettement supérieur aux effectifs des autres valeurs de la variable.
- Il ne doit être retenu que s'il est unique (série uni modale)
- Une série peut être multi modale si plusieurs variables de la série présentent le même effectif. Le mode perd alors toute signification.