

Chapitre STATISTIQUES

Cour et exercices M me GUESMIA AZIZA

I / Introduction

La statistique est une science d'observation qui, en classant des données, décrit des phénomènes (des événements) au moyen d'un certain nombre de valeurs numériques.

• **Population** : c'est l'ensemble étudié, ses éléments sont appelés des individus.

Exemple : si l'étude statistique porte sur les élèves d'un lycée, la population est alors l'ensemble des élèves du lycée et les individus sont les élèves.

• **Individu** : c'est un élément de la population.

• **Effectif total** : c'est le nombre total d'individus.

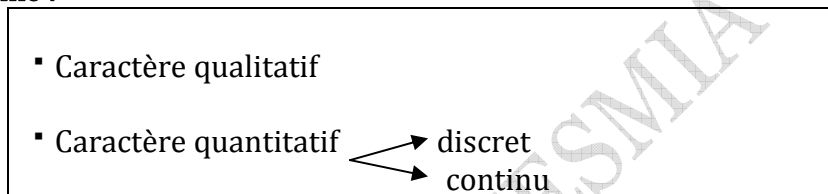
• **Caractère** : c'est la propriété étudiée.

On distingue les caractères **qualitatifs** ; non mesurables (couleur, marque d'un appareil....) et les caractères **quantitatifs** qui sont mesurables (taille, poids, notes d'un devoirsalaires....) Les caractères quantitatifs sont :

Discrets qui ne peuvent prendre qu'un nombre fini de valeurs (notes à un devoir...).

Continus dont on regroupe les valeurs par intervalles (taille, durée d'écoute...).

En résumé :



II / Le vocabulaire utilisé en statistiques

1) Caractère d'une population

Les outils et les méthodes des études statistiques s'appliquent à des ensembles d'éléments nommés populations (exemple : l'ensemble des élèves d'un lycée, l'ensemble des pièces fabriquées en une heure par une machine, l'ensemble des trajets journaliers des élèves d'un Lycée).

Chaque élément de la population étudiée est : une unité statistique ou un individu (élève, pièce fabriquée, trajet journalier)

Le caractère ou variable statistique d'une population est la propriété sur laquelle porte l'étude statistique.

Le caractère statistique peut être :

Qualitatif (couleur d'une voiture, marque d'un appareil électroménager)

Quantitatif (Dans ce cas le caractère statistique est mesurable, il peut être noté par une variable statistique.)

Une variable statistique peut être :

Discrète (Elle prend un nombre fini de valeurs comme par exemple un nombre de voitures par puissance fiscale ou un nombre de personnes par foyer.)

Continue (Elle prend toutes les valeurs à l'intérieur d'un intervalle donné.)

L'étude statistique d'une population par rapport à une variable continue impose de regrouper le grand nombre de valeurs en tranches ou classes. (Classes d'âge pour une population de personnes ; classes du montant des achats pour une population de clients d'une grande surface.)

2) Classes et effectifs

Une classe, c'est la portion de l'intervalle auquel appartiennent les valeurs de caractère. Une série statistique associe à chaque valeur x_i du caractère le nombre d'individus correspondant, appelé effectif partiel et noté n_i .

Exemple : les notes obtenues par tous les candidats aux épreuves de baccalauréat.

3) L'effectif de la population

L'effectif n d'une classe est le nombre d'individus qu'elle contient.

Exemple : dire que la valeur 200 a pour effectif 13 signifie qu'il ya 13 observations pour la valeur du caractère 200.

Remarque : Si le tableau des effectifs est donné, la définition permet toujours de calculer le tableau des fréquences. Inversement, si l'on donne le tableau des fréquences, on peut reconstituer le tableau des effectifs à condition de connaître **l'effectif total** de la série.

▪ On appelle **effectif cumulé croissant** de la valeur x le nombre d'individus ayant des valeurs du caractère **inférieurs ou égaux** à x .

▪ On appelle **effectif cumulé décroissant** de la valeur x le nombre d'individus ayant des valeurs du caractère **supérieures ou égales** à x .

Exemple : On a relevé les vitesses de passage en Km /h de 50 voitures.

Classes	[70 ; 80[[80 ; 90[[90 ; 100[[100 ; 110[[110;120[[120 ; 130[
Effectifs	2	5	8	20	10	5
Effectifs cumulés croissant	2	2+5=7	7+8=15	15+20=35	35+10=45	45+5=50

35 voitures sont passées à une vitesse moins de 110 Km/h

15 voitures sont passées à une vitesse moins de 100 Km/h

Classes	[70 ; 80[[80 ; 90[[90 ; 100[[100 ; 110[[110;120[[120 ; 130[
Effectifs	2	5	8	20	10	5
Effectifs cumulés décroissant	50	50-2=48	48-5=43	43-8=35	35-20=15	15-10=5

43 voitures sont passées à 90 Km/h au moins.

35 voitures sont passées à 100 Km/h au moins.

4) La fréquence d'une valeur x_i du caractère est le quotient de l'effectif n_i de ce caractère par l'effectif total N : $f_i = \frac{n_i}{N}$

Remarque : La somme des fréquences est égale à 1.

5) Le pourcentage

Le pourcentage p d'une classe s'obtient en multipliant sa fréquence f par 100.

$$p = \frac{n}{N} \times 100.$$

Remarque : La somme des pourcentages est égale à 100.

6) Fréquences cumulées : La fréquence cumulée correspondant à une modalité est la somme de la fréquence correspondant à cette modalité et des fréquences correspondants à toutes les valeurs précédentes.

Exemple : Répartition des élèves d'un lycée.

Modalités	1 ^{ère}	2 ^{ème}	3 ^{ème}	4 ^{ème}	total
Effectifs	140	155	104	101	500
Fréquences	$\frac{140}{500} = 0,28$	$\frac{155}{500} = 0,31$	$\frac{104}{500} = 0,208$	$\frac{101}{500} = 0,202$	1
Pourcentage	28%	31%	20,8%	20,2%	100%
Fréquences cumulées croissantes	0,28	$0,28+0,31=0,59$	$0,59+0,208=0,798$	$0,798+0,202=1$	

20,2 % des élèves de ce lycée sont en terminale (4^{ème})

Les élèves de 1^{ère} représentent 28 % du nombre total des élèves de ce lycée.

III / Représentations graphiques utilisées en statistiques

Diagramme à barres : Pour des modalités quantitatives, on peut utiliser un système d'axe.

- Dans le graphique cartésien, on porte :
- En abscisse, les valeurs du caractère.
 - En ordonné, les effectifs correspondants. Ou les fréquences correspondantes

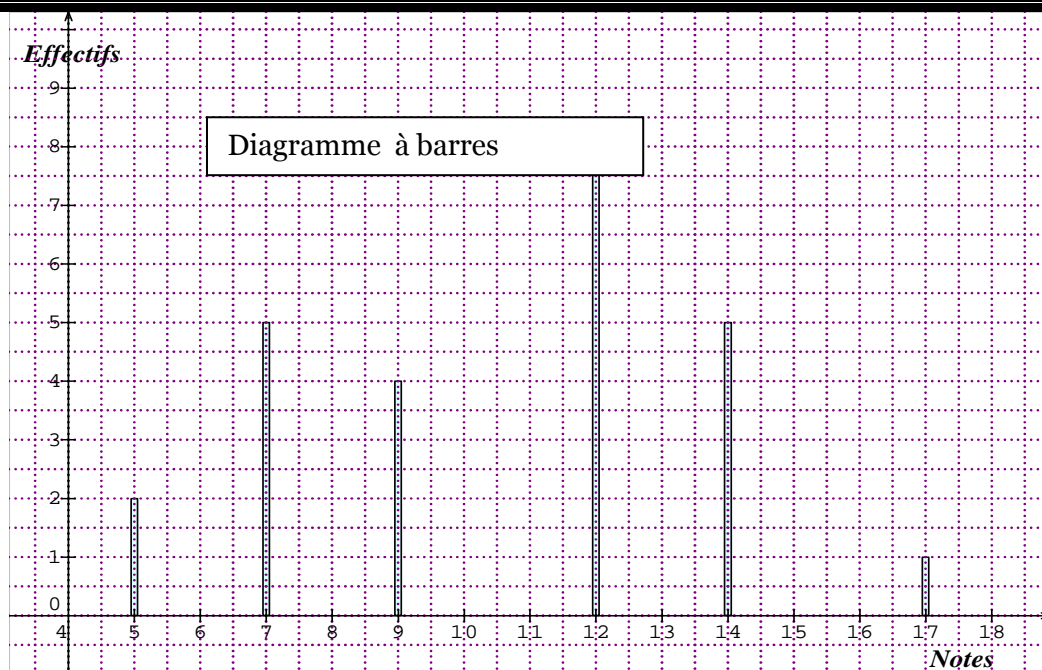
La ligne polygonale obtenue en joignant les points représentatifs par des segments de droites s'appelle **le polygone des effectifs** ou **le polygone des fréquences**.

Un diagramme à barres est une représentation de données statistiques à l'aide de rectangles de largeur constante.

Exemple n° 1 : Voici les notes d'un devoir de sciences physiques d'une classe.

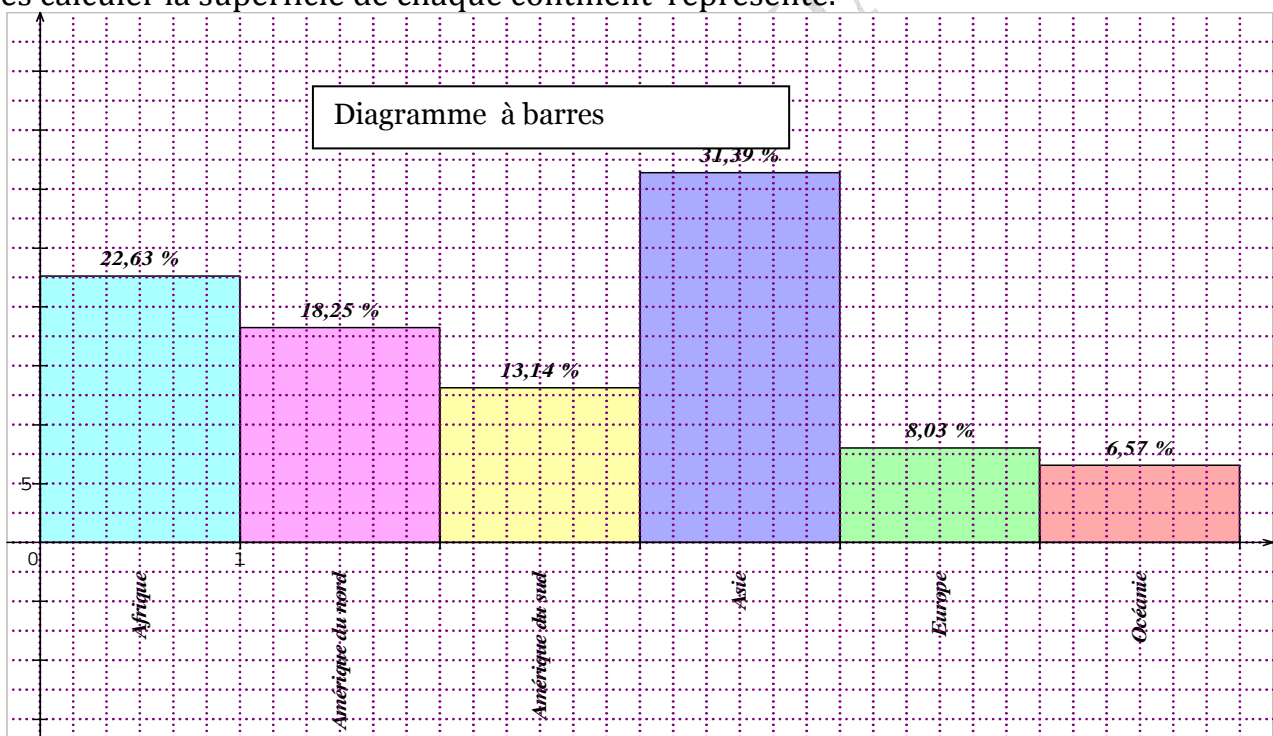
Notes	5	7	9	12	14	17
Effectifs	2	5	4	8	5	1

Le diagramme ci-dessous est un diagramme à barres, la longueur d'une barre dépend de l'effectif qui lui est associé. La plus petite barre est celle de la note 17 car l'effectif de 17 est 1. La barre la plus longue correspond à la note 12 car l'effectif de la note 12 est 8 : c'est le plus grand effectif de cette série.



Exemple n° 2 : Caractère qualitatif

Le schéma ci-dessous est un diagramme à barres représentant les pourcentages des superficies des blocs continentaux. Si la superficie de l'Afrique est 30 millions de kilomètre carrés calculer la superficie de chaque continent représenté.



Réponse

La superficie de l'Afrique est 30 millions de kilomètre carrés et elle représente 22,63 %, donc l'Amérique du nord a une superficie : $S_{Amn} = \frac{30 \times 18,25}{22} = 24,886$ millions de kilomètre carrés.

L'Amérique du sud a une superficie : $S_{Ams} = \frac{30 \times 13,14}{22} = 18$ millions de kilomètre carrés.

L'Asie a une superficie : $S_{Asie} = \frac{30 \times 31,39}{22} = 42,8$ millions de kilomètre carrés.

L'Europe a une superficie : $S_{Eu} = \frac{30 \times 8,03}{22} = 10,95$ millions de kilomètre carrés.

L'Océanie a une superficie : $S_{Océ} = \frac{30 \times 6,57}{22} = 8,96$ millions de kilomètre carrés.

1) Diagramme en bâtons:

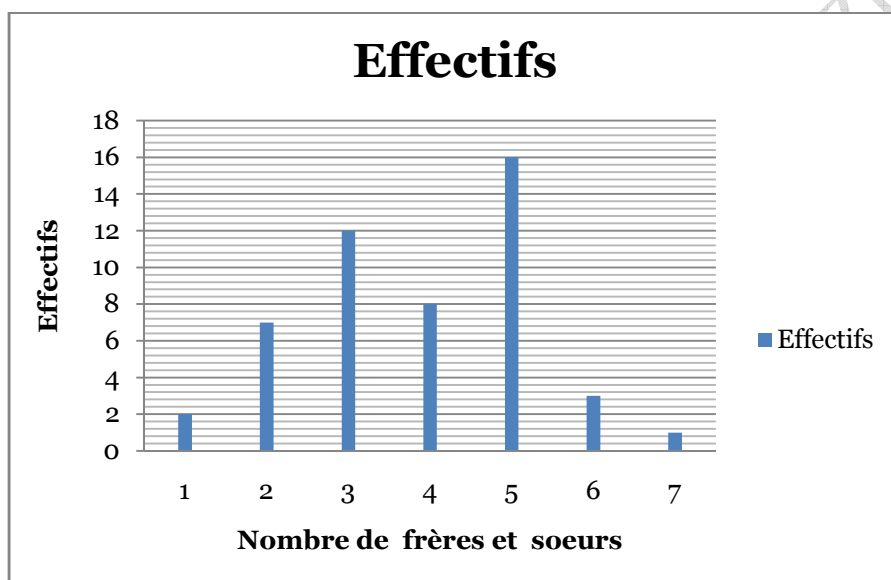
Pour les caractères quantitatifs discrets, on utilise le diagramme en bâton :
 Dans un repère orthogonal, pour chaque valeur de la série statistique on trace un trait vertical dont la hauteur est proportionnelle à l'effectif (dans l'unité choisie).

Lorsque le caractère est discret une représentation peut être un diagramme à barres ou encore mieux un diagramme en bâtons qui représente les données statistiques à l'aide de segments.

Exemple n° 3 : On a relevé le nombre de frères et sœurs d'un groupe de 49 élèves :

Nombre de frères et sœurs	1	2	3	4	5	6	7
Nombre d'élèves	2	7	12	8	16	3	1

Cette série statistique peut être représentée par le diagramme en bâtons ci-dessous :



Remarque : Le diagramme à barres de l'exemple n° 1 peut être remplacé par un diagramme à bâtons.

2) Polygone des fréquences :

Un polygone de fréquence est obtenu en joignant par des segments de droites les extrémités des bâtons.

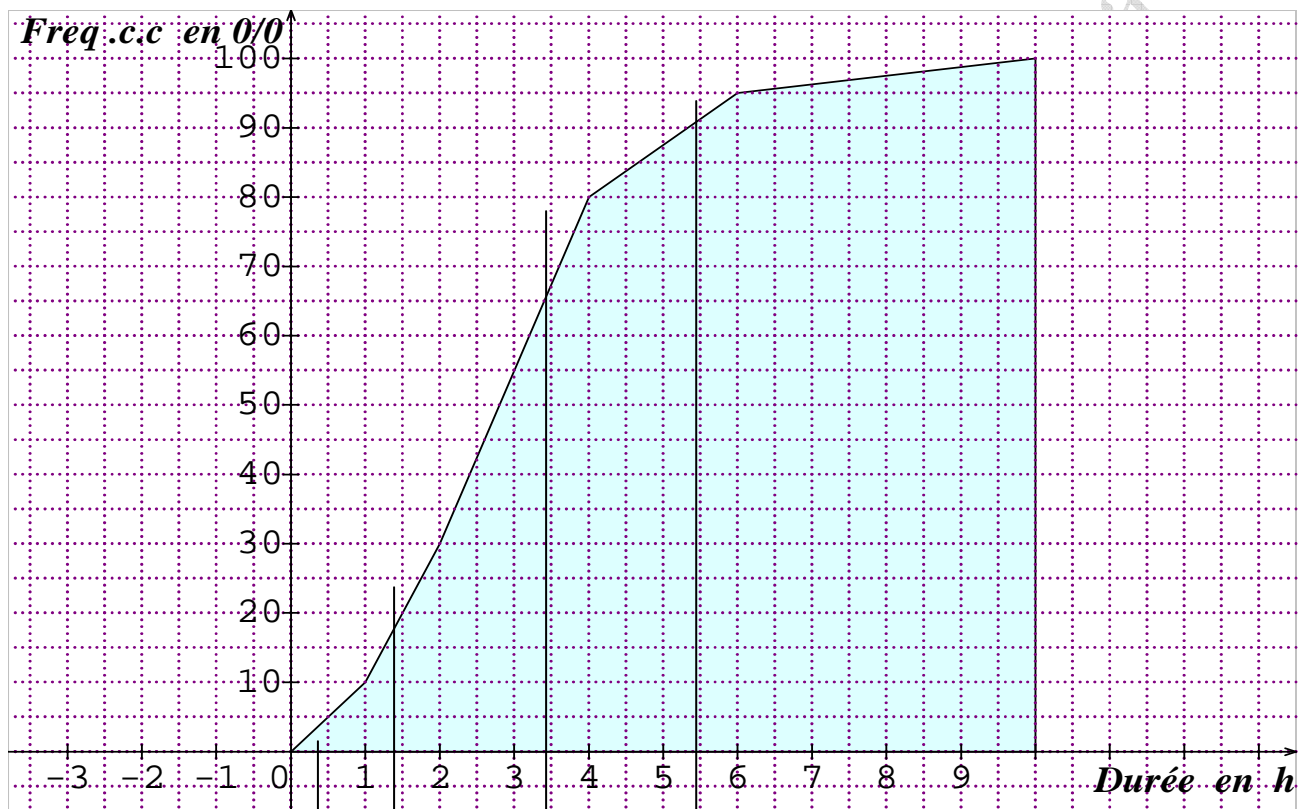
Exemple : Une enquête sur 1000 personnes porte sur la durée en heure(s) passée chaque jour devant la télévision, les résultats sont donnés dans le tableau suivant.

Durée	[0 ; 1[[1 ; 2[[2 ; 4[[4; 6[[6 ; 10[Total
Effectifs	100	200	500	150	50	1000

Question : construire le polygone des fréquences cumulées croissantes.

Réponse : pour tracer le polygone des fréquences cumulées croissantes on utilise le tableau ci-dessous :

Durée	[0 ; 1[[1 ; 2[[2 ; 4[[4 ; 6[[6 ; 10[Total
Effectifs	100	200	500	150	50	1000
Fréquences	0,1	0,2	0,5	0,15	0,05	1
Fréquences cumulées croissantes	0,1	0,3	0,8	0,95	1	
Fréquences .C. C en %	10	30	80	95	100	



Remarque : Dans le polygone des fréquences cumulées c'est plus facile de représenter sur l'axe des ordonnées les fréquences cumulées croissantes en %.

3) Histogramme :

On utilise ce procédé lorsque le caractère est quantitatif continu :

- Chaque classe est représentée par un rectangle dont l'aire est proportionnelle à l'effectif.
- Un des cotés des rectangles appartient à un axe, et ces cotés sont proportionnels aux amplitudes des classes.

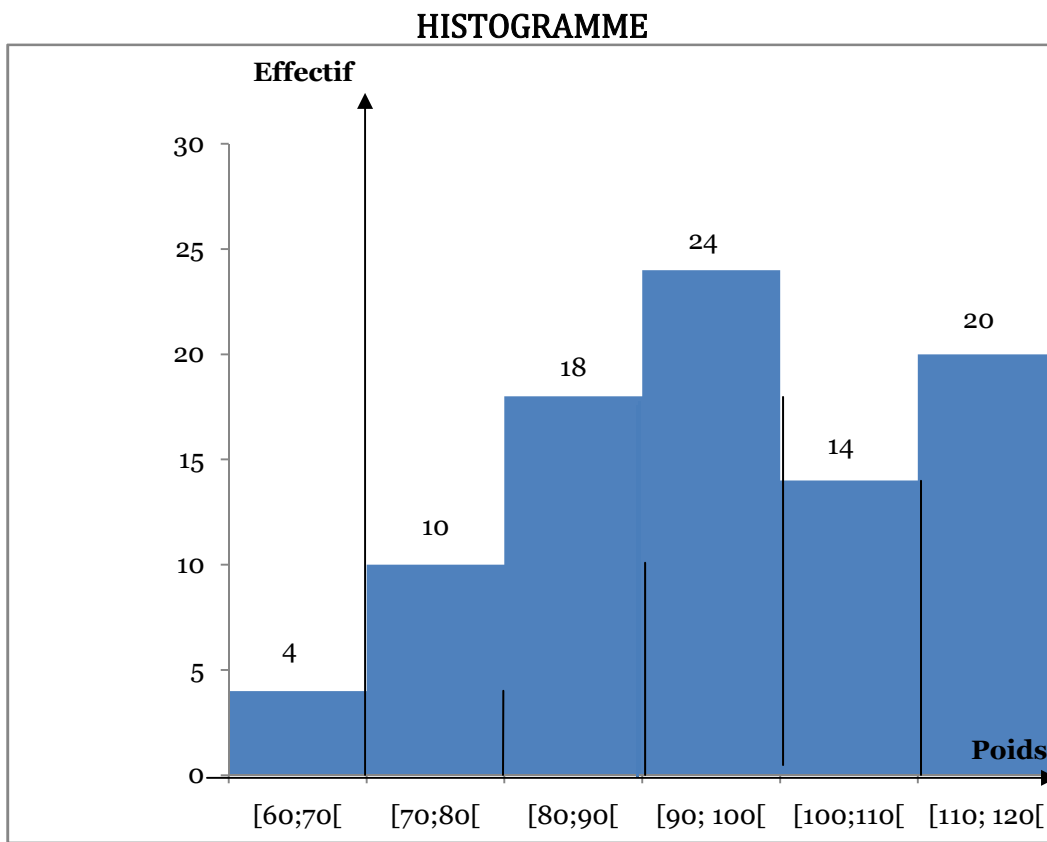
• Si les classes sont d'égale amplitude, la hauteur de chaque rectangle est donc proportionnelle à l'effectif.

Exemple :

On a pesé 80 joueurs, les résultats exprimés en kilogramme sont :

Poids	[60; 70[[70; 80[[80; 90[[90; 100[[100; 110[[110; 120[
Nombre de joueurs	4	10	18	24	14	10

Cette série est représentée par l'histogramme suivant :



Si les classes sont d'amplitudes différentes une légère modification s'impose pour conserver la proportionnalité de l'aire.

4) Diagramme circulaire :

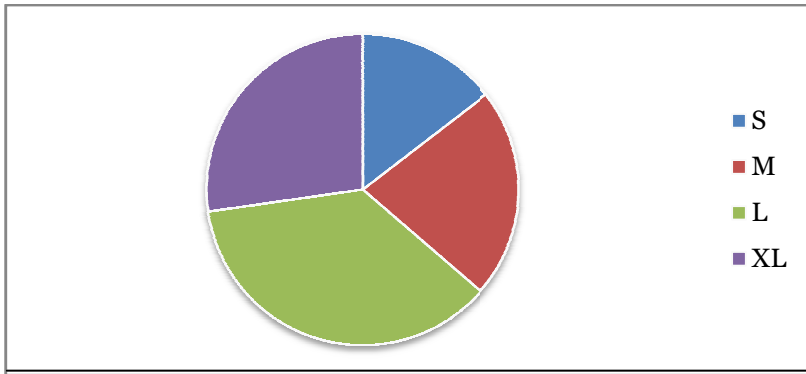
Chaque classe est représentée par un secteur circulaire dont la surface est proportionnelle à l'effectif. Il en est donc de même pour l'angle au centre.

Exemple :

Dans une boutique les tee-shirts sont vendus dans quatre tailles différentes : S, M, L, et XL. Les ventes mensuelles sont données par le tableau suivant :

Taille	S	M	L	XL
Nombre de ventes	16	24	40	30

Cette série statistique est représentée par le diagramme circulaire ci-dessous :



Interprétation :

Les angles des secteurs s'obtiennent par proportionnalité.
 L'effectif de la population étant 110, la taille L occupe presque 36,4 % du disque.
 Par conséquent, l'angle du secteur angulaire représentant la taille L occupera 36,4 % de 360° c'est-à-dire presque 131°.

5) Polygone des effectifs cumulés croissants :

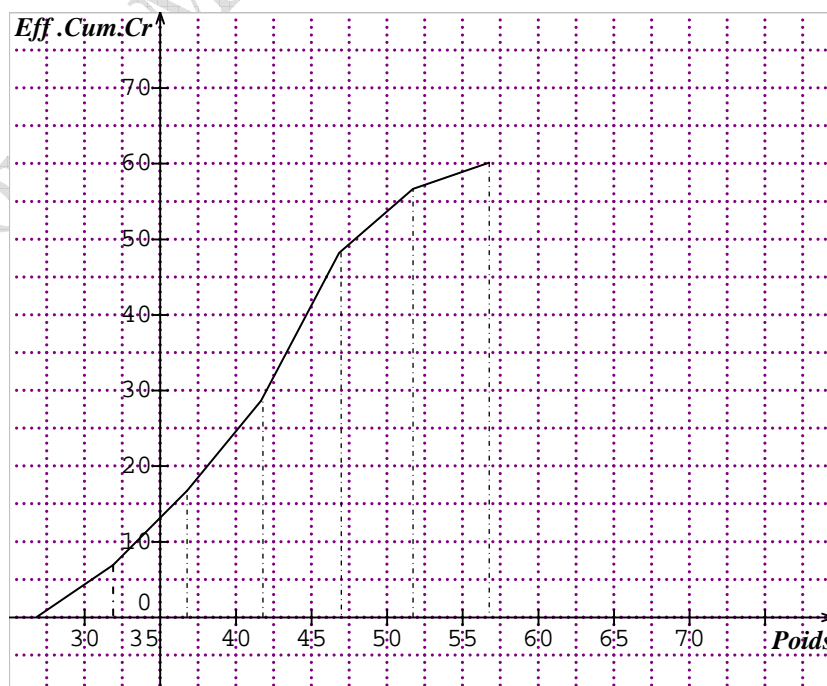
Graphiquement, on obtient un ensemble un ensemble de points que l'on relie par des segments de droites. On trace ainsi le polygone des effectifs cumulés croissants(ou décroissants).

Exemple :

On a pesé 60 garçons voici les résultats :

Poids	[40 ; 45[[45 ; 50[[50 ; 55[[55 ; 60[[60 ; 65[[65 ; 70[
Effectifs	6	10	12	19	9	4
Effectifs cumulés croissants	6	16	28	47	56	60

Polygone des effectifs cumulés croissants :



D'après ce polygone des effectifs cumulés croissants on peut déduire que moins de 50 % des garçons pesés ont un poids inférieur à 55 kilogrammes.

IV/ Paramètres d'un caractère statistique :

On distingue deux types de paramètres sur un caractère statistique.

♠ Les paramètres de position

♠ Les paramètres de dispersion

Les paramètres de position

{Le mode La médiane La moyenne Quartile

Les paramètres de dispersion

{L'étendue L'écart type Variance Intervalles, Interquartiles

A / Paramètres de position

1°) **Mode** : ou valeur dominante.

a) Pour un caractère discret :

Le mode est une valeur du caractère qui correspond à l'effectif le plus grand.

Exemple n° 1

Pour la série statistique ci dessous, le mode est 5 (l'effectif de 5 est 16).

Nombre de frères et sœurs	1	2	3	4	5	6	7
Nombre d'élèves	2	7	12	8	16	3	1

b) Pour un caractère qualitatif :

Exemple n °2

Le tableau suivant donne la répartition des mariages dans une salle de fête durant 2004.

Mois	Jan	Fév	Mars	Avr	Mai	Jui	Juil	Août	Sep	Oct	Nov	Déc
Nombre de mariages	0	3	3	10	7	20	36	18	15	11	5	1

Le mode de cette série est Juillet.

c) Pour un caractère quantitatif continue: (série classée) : On définit la classe modale, la classe dont l'effectif est relativement le plus élevé et on attribut au mode la valeur centrale de cette classe.

Exemple: On a pesé 80 joueurs, les résultats exprimés en kilogramme sont :

Poids	[60; 70[[70; 80[[80; 90[[90; 100[[100; 110[[110; 120[
Nombre de joueurs	4	10	18	24	14	10

Pour cette série la classe modale est [90 ; 100[car son effectif est le plus élevé : 24.

Le mode de cette série est : 95 ; $(\frac{90+100}{2} = 95)$.

▪ Si le polygone des effectifs ne présente qu'une seule pointe, la série est dite série **uni-modale**.

▪ Si le polygone des effectifs présente plusieurs pointes la série a plusieurs classes modales (ou plusieurs modes). La série est donc **multimodale**.

2°) Moyenne :

On considère un caractère quantitatif : (discret ou continu)

Valeur du caractère ou centre de l'intervalle	x_1	x_2	x_p
Effectif	n_1	n_2	n_p

La moyenne arithmétique est alors le réel \bar{x} défini par :

$$\bar{x} = \frac{n_1 x_1 + n_2 x_2 + \dots + n_p x_p}{N} \quad \text{avec } N = n_1 + n_2 + n_3 + \dots + n_p \text{ (N: effectif total)}$$

Exemple n°1 :

Deux classes A et B ont fait le même devoir de contrôle de maths :

Classe A	10 19 5 18 10 18 10 3 9 3
	19 8 17 5 11 10 11 18 8 14
	16 7 10 7 5

Classe B	11 10 4 7 12 3 6 17 9
	19 13 11 10 6 7 19 18
	11 10 7 3 4 14 9 8

Calculer les moyennes des séries des deux classes.

Soit \bar{x}_A : la moyenne de la classe A et \bar{x}_B : la moyenne de la classe B.

$$\bar{x}_A = \frac{3 \times 2 + 5 \times 3 + 7 \times 2 + 8 \times 2 + 9 + 10 \times 5 + 11 \times 2 + 14 + 16 + 17 + 18 \times 3 + 19 \times 2}{25} = 10,84$$

$$\bar{x}_B = \frac{3 \times 2 + 4 \times 2 + 6 \times 2 + 7 \times 3 + 8 \times 1 + 9 \times 2 + 10 \times 3 + 11 \times 3 + 12 + 13 + 14 + 17 + 18 + 19 \times 2}{25} = 9,92$$

PROPRIÉTÉ

- Si on ajoute à toutes les valeurs d'une série statistique le même nombre b, on augmente la moyenne de cette série par b.
- Si les valeurs d'une série statistique sont multipliées ou divisées par un même nombre a, la moyenne de cette série est aussi multipliée ou divisée par a.

PROPRIÉTÉ

Si une population d'effectif N est composée d'une partie d'effectif N_1 et de moyenne x_1 et d'une autre partie d'effectif N_2 et de moyenne x_2 la moyenne \bar{x} de la population totale est telle que :

$$\bar{x} = \frac{N_1 x_1 + N_2 x_2}{N}$$

Exemple n°2 :

Si dans une classe, les 15 garçons d'une classe mesurent en moyenne 182 cm et si les 20 filles mesurent en moyenne 168 cm, alors la taille moyenne d'un élève de cette classe est égale à

$$\bar{x} = \frac{15 \times 182 + 20 \times 168}{15 + 20} = 174 \text{ cm.}$$

Exemple n°3 : Dans une entreprise, la répartition des salaires est la suivante :

Salaire en dinars	[300 ; 500[[500 ; 700[[700 ; 1000[[1000 ; 1500[
Nombre de salariés	50	25	12	3

Pour calculer le salaire moyen, on dresse d'abord un tableau donnant les centres de classe :

Classe salaire	Centre x_i	Effectifs n_i	$x_i \times n_i$
[300 ; 500[400	50	20 000
[500 ; 700[600	25	15 000
[700 ; 1000[850	12	10 200
[1000 ; 1500[1250	3	3750
Total	////////////////	90	48 950

La moyenne de cette série est $\bar{x} = \frac{48\,950}{90} = 543,889$ dinars

Exemple n°4

En utilisant les fréquences (tableau ci-dessus)

Salaire en dinars	[300 ; 500[[500 ; 700[[700 ; 1000[[1000 ; 1500[
Nombre de salariés	50	25	12	3

Classe Salaire	Centre x_i	Effectifs n_i	Fréquences f_i
[300 ; 500[400	50	$\frac{50}{90} = 0,55$
[500 ; 700[600	25	$\frac{50}{90} = 0,27$
[700 ; 1000[850	12	$\frac{50}{90} = 0,13$
[1000 ; 1500[1250	3	$\frac{50}{90} = 0,033$
Total	////////////////	90	////////////////

$$\bar{x} = x_1 f_1 + x_2 f_2 + x_3 f_3 + x_4 f_4 = 543,889 \text{ dinars}$$

Remarque ; moyenne mode

- La moyenne d'un caractère quantitatif n'est pas nécessairement une valeur de ce caractère, alors que le mode est une valeur : c'est la valeur qui a l'effectif maximum.
- La moyenne est unique alors qu'il peut exister plusieurs modes.

3°) Médiane :

a) Définition :

L'idée générale est que la médiane est une valeur du caractère qui partage la population en deux parties de même effectif. De façon plus précise, on appelle médiane d'une série

statistique **discrète** toute valeur M du caractère telle qu'au moins 50% des individus aient une valeur du caractère inférieure ou égale à M et au moins 50% des individus aient une valeur du caractère supérieure ou égale à M.

b) Recherche pratique de la médiane : caractère quantitatif : discret

On range les valeurs du caractère une par une dans l'ordre croissant (chaque valeur du caractère doit apparaître un nombre de fois égal à l'effectif correspondant).

Si l'effectif total est impair, la médiane M est la valeur du caractère située au milieu.

Si l'effectif total est pair, la médiane M est la demi-somme des 2 valeurs situées au milieu.

Exemple:

On considère la série statistique suivante :

Valeur du caractère x_i	7	8	9	10	11	14	16
Caractère x_i	2	1	1	1	2	1	2

• Liste des valeurs du caractère : 7 ; 7 ; 8 ; 9 ; **10 ; 11** ; 11 ; 14 ; 16 ; 16

• L'effectif total est pair : la médiane M est la demi-somme des 2 valeurs situées au milieu.

D'où, $M = \frac{10+11}{2} = 10,5$

Exemple 2 :

On considère la série statistique suivante :

Valeur du caractère x_i	12	13	15	17	19	22
Caractère x_i	2	1	1	2	3	2

Liste des valeurs du caractère : 12 ; 12 ; 13 ; 15 ; 17 ; **17** ; 19 ; 19 ; 19 ; 22 ; 22

• L'effectif total est impair : la médiane M est la valeur située au milieu. D'où, $M = 17$.

c) Recherche pratique de la médiane : caractère quantitatif : continu

Dans le cas d'une série d'un caractère continu, l'intervalle de variation du caractère x est partagé en classes, la médiane est la valeur du caractère correspondant à un **effectif cumulé** égale à la moitié de l'effectif total.

La médiane partage la série classée en deux parties de **même effectif**.

1ère Méthode graphique pour déterminer la médiane:

Dans ce cas, la médiane est déterminée par le polygone des effectifs cumulés croissants ou décroissants, on trace la droite **horizontale** correspondant à la **moitié** de l'effectif total.

La médiane est l'abscisse du point du polygone dont l'ordonnée est égale à $\frac{N}{2}$ (où N est l'effectif total de la série).

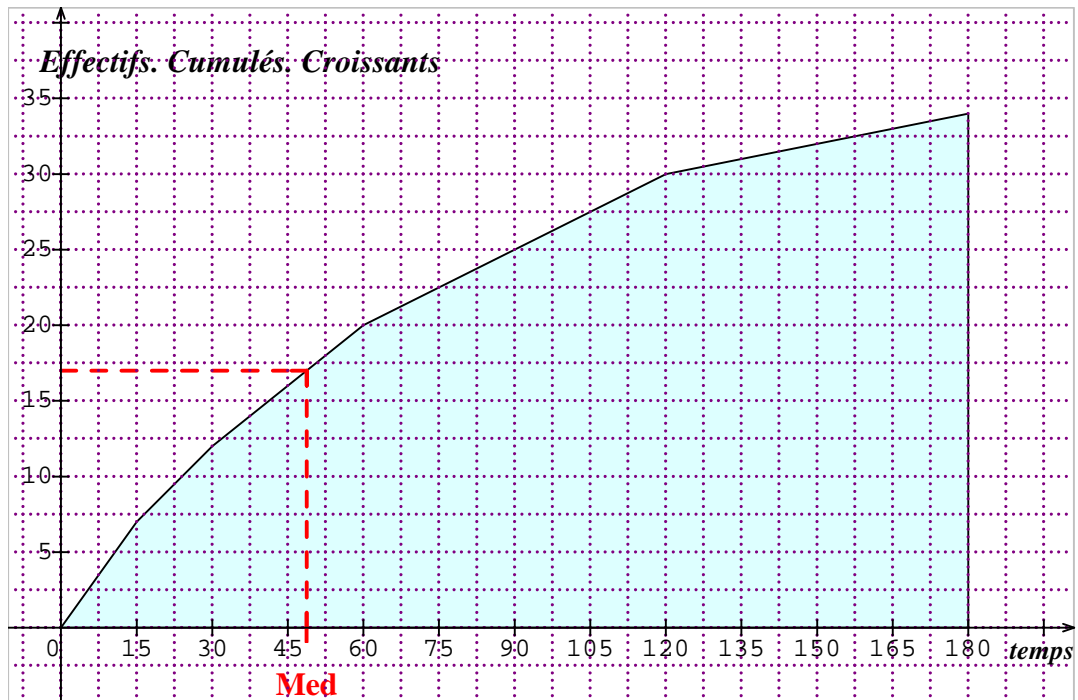
Exemple n°2 :

Temps en minutes passé devant la télévision par 34 élèves pendant une certaine journée.

Temps en minutes	[0 ; 15[[15 ; 30[[30 ; 60[[60 ; 120[[120 ; 180[
Nombre d'élèves	7	5	8	10	4

Effectifs cumulés croissants	7	12	20	30	34
------------------------------	---	----	----	----	----

Le polygone des effectifs cumulés croissants correspondant :



L'effectif total est 34, la moitié de l'effectif total est 17 donc d'après le graphique ci-dessus, la médiane : Med est l'abscisse du point du polygone dont l'ordonnée est égale $\frac{34}{2} = 17$.

Graphiquement on lit : **Med = 48,75** ou **Med = 49 mn et 15 secondes**.

Interprétation :

La moitié de ce groupe d'élèves passe un temps inférieur à 49 mn et 15 secondes devant la télévision.

2^{ème} Méthode graphique pour déterminer la médiane :

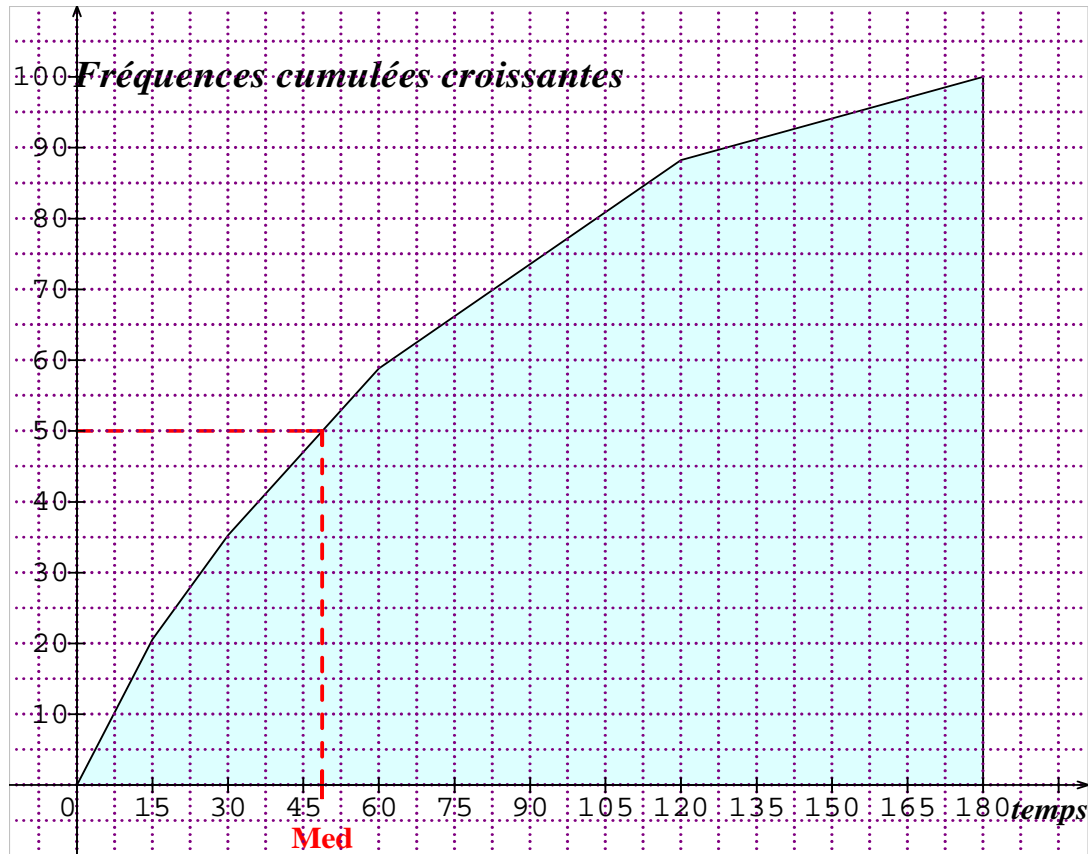
Dans ce cas, la médiane est déterminée par le polygone des fréquences cumulées croissantes ou décroissantes, on trace la droite **horizontale** correspondant à la **fréquence cumulées 0,5**. La médiane est l'abscisse du point du polygone dont l'ordonnée est égale **0,5**.

Exemple n°3:

Temps en minutes passé devant la télévision par 34 élèves pendant une certaine journée.

Temps en minutes	[0 ; 15[[15 ; 30[[30 ; 60[[60 ; 120[[120 ; 180[
Nombre d'élèves	7	5	8	10	4
Fréquences	0,21	0,15	0,23	0,29	0,12
Fréquences cumulées croissantes	0,21	0,36	0,59	0,88	1

Fréquences. Cumul. crois. en %	21	36	59	88	100
--------------------------------	----	----	----	----	-----

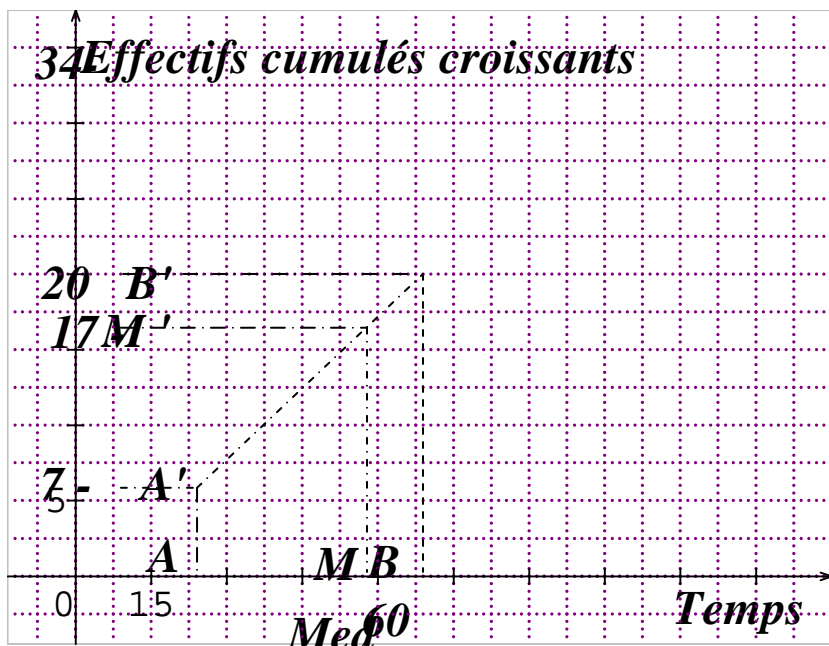


Graphiquement on lit : $Med = 48,75$ ou $Med = 49 \text{ mn et } 15 \text{ secondes}$.

3^{ème} Méthode graphique pour déterminer la médiane: Méthode de calcul.

La classe médiane est $[30 ; 60]$ car d'après le tableau des effectifs cumulés croissants, il ya 20 élèves donc dans cette classe figure la moitié des élèves 17 qui correspond à 50% de l'effectif total.

Calculons la médiane par interpolation linéaire en s'aidant du schéma suivant :



$$\frac{AM}{AB} = \frac{A'M'}{A'B'} \quad \text{donc} \quad \frac{Med-15}{60-15} = \frac{7-17}{7-20} \quad \text{d'où} \quad Med - 15 = 45 \times \frac{10}{13}; \quad Med = 15 + 36,6$$

Med = 49,6

D'une manière pratique, on utilise le tableau ci-dessous : (avec x : abscisse de y)

x	y
15	7
Med	17
60	20

$$\frac{Med-15}{17-7} = \frac{60-15}{20-7} \quad Med - 15 = \frac{10 \times 45}{13} \quad \mathbf{Med = 49,6}$$

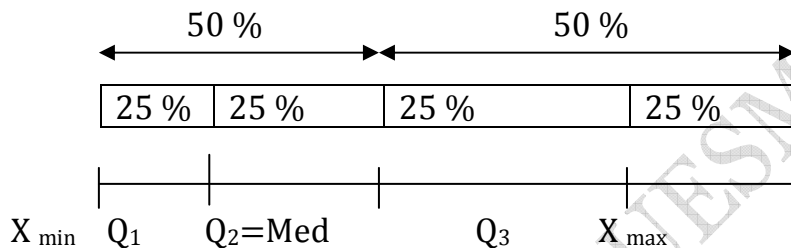
4°) Quartiles :

a) **Vocabulaires :**

La médiane partage la série en deux groupes de même effectif. Les quartiles partagent la série en quatre groupes de même effectif.

Ils sont donc au nombre de trois Q_1 ; Q_2 ; Q_3 où Q_2 est la médiane. Les intervalles qu'ils définissent contiennent chacun 25 % des observations, soit un quart de l'effectif total, $\frac{N}{4}$

comme le montre le schéma :



Q_1 laisse 25 % des observations « avant » et 75 % « après ».

Q_3 laisse 75 % des observations « avant » et 25 % « après ».

♠ Le 1^{er} **quartile** est le nombre noté Q_1 et qui est égal à la plus petite valeur du caractère tel qu'au moins 25 % de l'effectif de la série prennent des valeurs inférieures ou égales à Q_1 .

♠ Le 3^{ème} **quartile** est le nombre noté Q_3 et qui est égal à la plus petite valeur du caractère tel qu'au moins 75 % de l'effectif de la série prennent des valeurs inférieures ou égales à Q_3 .

♠ L'intervalle $[Q_1 Q_3]$ s'appelle **intervalle interquartile**.

♠ le réel $Q_3 - Q_1$ s'appelle **l'écart interquartile**.

b) **Recherche pratique des quartiles : série à caractère discret :**

Exemple n°1

On considère la série ordonnée suivante : 4 5 6 7 11 13 14 15 16.

$\begin{array}{ccccccc} & & & \downarrow & & \downarrow & & \downarrow \\ & & & Q_1 & & \text{Med} & & Q_3 \end{array}$

La série comporte 9 observations (N : l'effectif total est impair) donc la médiane correspond à la 5^{ème} valeur de la série ordonnée. La médiane est 11.

La série comportant 9 termes, $\frac{9}{4} = 2,25$; on dépasse 2 termes de la série donc on prend Q_1 le 3^{ème} terme de la série c'est à dire $Q_1 = 6$.

Q_3 est symétrique de Q_1 par rapport à la médiane d'où $Q_1 = 6$; $Q_2 = \text{Med} = 11$; $Q_3 = 14$.

Exemple n°2 caractère quantitatif : discret

On considère la série ordonnée suivante :

4 ; 5 ; 5	6 ; 7 ; 11 ; 13 ; 13 ; 14	15 ; 16 ; 18
-----------	---------------------------	--------------

$$Q_1 = 5 \quad \text{Med} = 12 \quad Q_2 = 14,5$$

La série comporte 12 observations (N : l'effectif total est pair) donc la médiane correspond à la moyenne de la 6^{ème} et la 7^{ème} valeur de la série ordonnée.

La médiane est : $\text{Med} = \frac{11+13}{2} = 12.$

La série comportant 12 termes, $\frac{12}{4} = 3$; donc Q_1 est la valeur telle que au moins 25 % de l'effectif total prenne des valeurs inférieures ou égales à Q_1 . On prend $Q_1 = 5$.

$12 \times \frac{3}{4} = 9$. On prend $Q_3 = 14$.

Conclusion $Q_1 = 5$; $Q_2 = \text{Med} = 12$; $Q_3 = 14$.

c) Recherche pratique des quartiles : série à caractère continu :

1^{ère} Méthode :

Au moyen d'une interpolation linéaire en utilisant le polygone des effectifs cumulés (croissant ou décroissant).

2^{ème} Méthode :

Dans le cas d'une série classée, le procédé de détermination des quartiles est identique à celui de détermination de la médiane, **la méthode graphique est la plus employée.**

♠ **Le quartile inférieur Q_1** est l'abscisse du point d'ordonné $\frac{1}{4}$ sur le polygone des fréquences cumulées et $\frac{N}{4}$ sur le polygone des effectifs cumulés.

♠ **Le quartile supérieur Q_3** est l'abscisse du point d'ordonné $\frac{3}{4}$ sur le polygone des fréquences cumulées et $\frac{3N}{4}$ sur le polygone des effectifs cumulés.

Exemple : caractère quantitatif : continu

On donne la série suivante des salariés d'une entreprise selon leurs classes des salaires mensuels nets exprimés en centaines de dinars.

Classes	[2 ; 4[[4 ; 6[[6 ; 8[[8 ; 10[[10 ; 12[[12 ; 14[Total
Effectifs	5	8	10	7	6	4	40

On veut déterminer les quartiles de cette série. Pour cela on utilise le tableau des effectifs cumulés croissants ci-dessous et par deux méthodes :

1^{ère} Méthode :

Classes	[2 ; 4[[4 ; 6[[6 ; 8[[8 ; 10[[10 ; 12[[12 ; 14[Total
Effectifs	5	8	10	7	6	4	40
Effectifs cumulés croissants	5	13	23	30	36	40	/////

Calcul du 1^{er} quartile : le quart de l'effectif total est $\frac{40}{4} = 10$.

Au moyen d'une interpolation linéaire en utilisant le polygone des effectifs cumulés croissant ci-dessous, on peut : trouver Q_1

x	y
4	5
Q_1	10
6	13

$$\frac{Q_1 - 4}{10 - 5} = \frac{6 - 4}{13 - 5} \quad \text{d'où} \quad Q_1 = \frac{10}{8} + 4 \quad ; \quad Q_1 = 5,25 \text{ en centaines de dinars} \quad Q_1 = 525 \text{ D.}$$

Calcul du 2^{ème} quartile : Médiane : la moitié de l'effectif total est $\frac{40}{2} = 20$.

De même en utilisant le tableau.

x	y
6	13
Me = Q_2	20
8	23

$$\frac{Q_2 - 6}{8 - 6} = \frac{20 - 13}{23 - 13} \quad \text{d'où} \quad Q_2 = \frac{7}{10} + 6 \quad ; \quad Q_2 = 7,4 \text{ en centaines de dinars} ; Q_2 = 740 \text{ D.}$$

Me = 7,4 en centaines de dinars ; Me = 740 D.

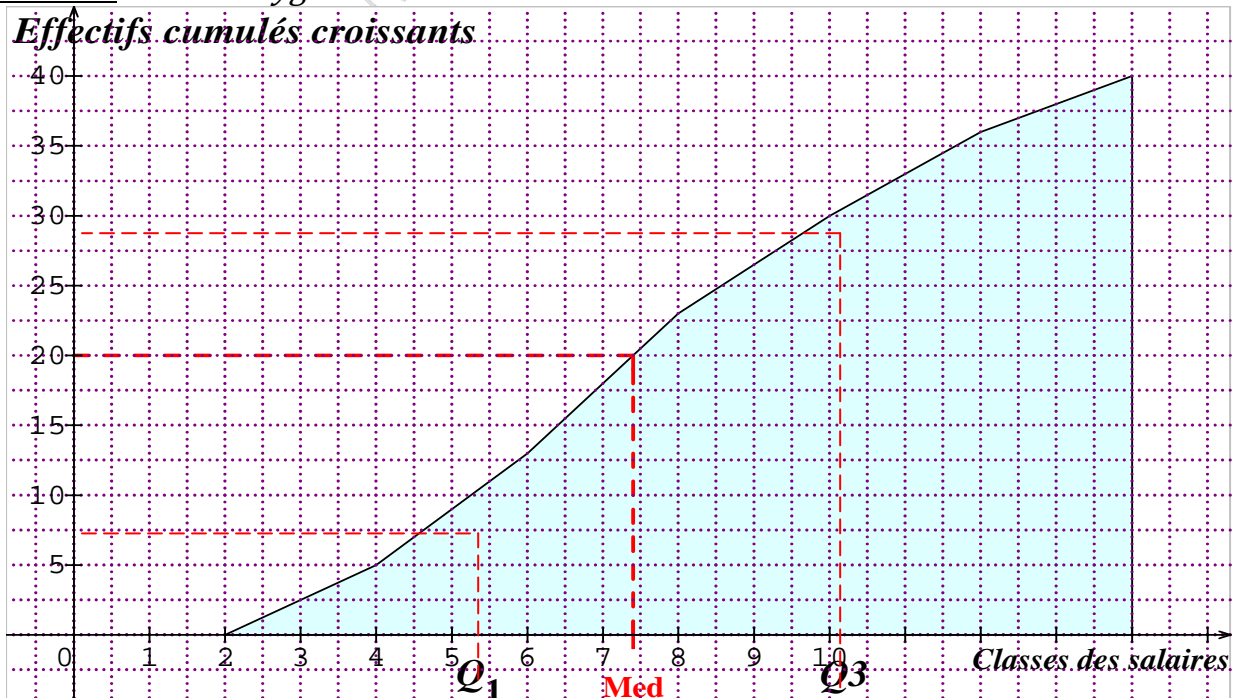
Calcul du 3^{ème} quartile : $\frac{3N}{4} = 30$

Sur le tableau on lit 30 personnes chacun a un salaire inférieur à 10 centaines de dinars.

Classes	[8 ; 10[
Effectifs	7
Effectifs cumulés croissants	30

Donc $Q_3 = 10$ en centaines de dinars ; $Q_3 = 1\ 000 \text{ D.}$

2^{ème} Méthode : Polygone des effectifs cumulés croissants



Le quartile inférieur Q_1 est l'abscisse du point d'ordonné $\frac{N}{4}$ sur le polygone des effectifs cumulés croissants ou décroissants. Donc Q_1 est l'abscisse du point d'ordonné 10.

Graphiquement $Q_1 = 5,25$ centaines de dinars ; $Q_1 = 525$ D.

Le quartile Q_2 est la médiane Me donc Q_2 est l'abscisse du point d'ordonné $\frac{N}{2}$ sur le polygone des effectifs cumulés croissants ou décroissants. Donc Q_2 est l'abscisse du point d'ordonné 20.

Graphiquement $Me = Q_2 = 7,4$ centaines de dinars ; $Q_2 = 740$ D.

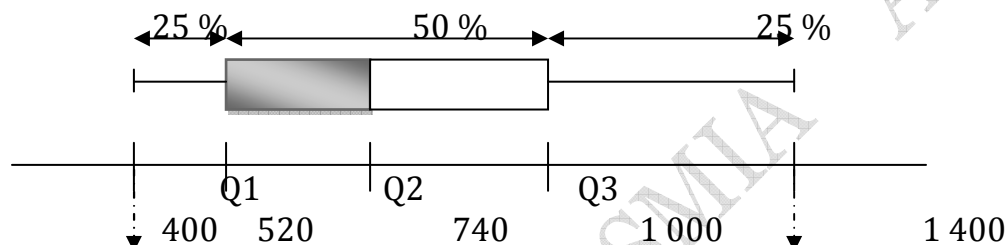
Le quartile supérieur Q_3 est l'abscisse du point d'ordonné $\frac{3N}{4}$ sur le polygone des effectifs cumulés croissants ou décroissants. Donc Q_3 est l'abscisse du point d'ordonné 30.

Graphiquement $Q_3 = 10$ centaines de dinars ; $Q_3 = 1\ 000$ D.

L'intervalle interquartile est $[520 ; 1\ 000]$.

L'écart interquartile est égal à $1\ 000 - 520 = 480$ D.

d) Diagramme en boîte



Les trois quarts de l'effectif sont 30 ce sont tous les salariés dont le salaire est inférieur à 1 000D.

50 % des salariés ont un salaire entre 520 D et 1 000 D.

B / Paramètres de dispersion :

1°) Etendue:

Définition : L'étendue d'une série statistique est la différence entre ses deux valeurs extrêmes (la plus grande et la plus petite valeur du caractère) ;

Exemple n°1: caractère quantitatif : discret

Modalités	1	2	3	4	5	6	7	8
Effectifs	90	60	30	60	30	0	30	10

L'étendue est : $8 - 1 = 7$.

Exemple : caractère quantitatif : continu

Classes	[10 ; 50[[50 ; 100[[100 ; 200[[200 ; 300[[300 ; 500[
Effectifs	18	104	47	23	6

L'étendue est : $500 - 10 = 490$.

Remarque : Moyenne et étendue

La moyenne ne résume pas toujours toute l'information qu'on veut extraire d'une série statistique.

Les deux séries ci-dessous, de même moyenne ($\bar{x} = 0$), semblent inégalement dispersées.

-2,4	-0,4	1,6	-3,4	-2,4	1,6	3,6	-1,4	0,6	2,6
6	-3	-4	8	6	-12	7	5	4	-17

La première série a pour étendue $3,6 - (-3,4) = 7$

La deuxième série a pour étendue $8 - (-17) = 25$.

La seconde série est plus dispersée que la première.

2°) Variance:

Définition : la variance notée V est la différence entre la moyenne des carrés et le carré de la moyenne.

a) **caractère quantitatif : discret**

$$V = \frac{n_1(x_1 - \bar{x})^2 + n_2(x_2 - \bar{x})^2 + n_3(x_3 - \bar{x})^2 + \dots + n_p(x_p - \bar{x})^2}{n_1 + n_2 + n_3 + \dots + n_p} - \bar{x}^2$$

Pour le calcul pratique : $V = \frac{n_1 x_1^2 + n_2 x_2^2 + n_3 x_3^2 + \dots + n_p x_p^2}{n_1 + n_2 + n_3 + \dots + n_p} - \bar{x}^2$

b) **caractère quantitatif : continu**

La variance V est : $V = \frac{n_1(c_1 - \bar{x})^2 + n_2(c_2 - \bar{x})^2 + n_3(c_3 - \bar{x})^2 + \dots + n_p(c_p - \bar{x})^2}{n_1 + n_2 + n_3 + \dots + n_p} - \bar{x}^2$

Avec c_1 : le centre de 1ère classe , c_2 : le centre de 2ème classe

Pour le calcul pratique : $V = \frac{n_1 c_1^2 + n_2 c_2^2 + n_3 c_3^2 + \dots + n_p c_p^2}{n_1 + n_2 + n_3 + \dots + n_p} - \bar{x}^2$

3°) Ecart-type:

Définition : l'écart-type est la racine carrée de la variance, on le note : σ . $\sigma = \sqrt{V}$.

L'écart-type d'une série statistique est une mesure de la plus ou moins grande dispersion des valeurs de la série. Un écart-type important signifie que les valeurs de la série s'éloignent souvent et de façon importante de la moyenne.

Exemple caractère quantitatif : discret

On considère la série statistique suivante :

Modalités	15	18	19	20	21	22
Effectifs	1	1	2	2	4	6

x_i	n_i	$n_i x_i$	$n_i x_i^2$
15	1	15	225
18	1	18	324
19	2	38	722
20	2	40	800
21	4	84	1764
22	6	132	2904
Total	16	327	6739

$\bar{x} = 20,44$;

$$V = \frac{n_1(x_1 - \bar{x})^2 + n_2(x_2 - \bar{x})^2 + n_3(x_3 - \bar{x})^2 + \dots + n_p(x_p - \bar{x})^2}{n_1 + n_2 + n_3 + \dots + n_p} - \bar{x}^2$$

Pour le calcul pratique :

$$V = \frac{n_1 x_1^2 + n_2 x_2^2 + n_3 x_3^2 + \dots + n_p x_p^2}{n_1 + n_2 + n_3 + \dots + n_p} - \bar{x}^2$$

$$N = n_1 + n_2 + n_3 + \dots + n_p$$

$$V = \frac{15^2 + 18^2 + 2 \times 19^2 + 2 \times 20^2 + 4 \times 21^2 + 6 \times 22^2}{16} - (20,4)^2$$

$$V = \frac{225 + 324 + 722 + 800 + 1764 + 2904}{16} - (20,4)^2$$

$$V = 421,1875 - 417,7936 = 3,3939; \quad \text{l'écart-type : } \sigma = \sqrt{V}; \quad \sigma = 1,84.$$

Remarque : On peut obtenir N , \bar{x} , V , et σ directement à partir de la calculatrice ; pour

Exemple caractère quantitatif : continu

On considère la série statistique suivante :

Classes	[0 ; 1[[1 ; 2[[2 ; 4[[4 ; 6[[6 ; 10[
Effectifs	100	200	500	150	50

Pour calculer la variance V et l'écart-type σ , on dresse le tableau ci-dessous :

Classes : x_i	Centres : c_i	Effectifs : n_i	Carrés des c_i : c_i^2	Produits : $n_i \times c_i^2$
[0 ; 1[0,5	100	0,25	25
[1 ; 2[1,5	200	2,25	450
[2 ; 4[3	500	9	4 500
[4 ; 6[5	150	25	3 750
[6 ; 10[8	50	64	3 200
Total		1 000		11 925

La variance V est :
$$V = \frac{n_1 c_1^2 + n_2 c_2^2 + n_3 c_3^2 + \dots + n_p c_p^2}{n_1 + n_2 + n_3 + \dots + n_p} - \bar{x}^2$$

$$\bar{x} = \frac{0,5 \times 100 + 1,5 \times 200 + 3 \times 500 + 5 \times 150 + 8 \times 50}{1\,000} = 3 \quad \bar{x}^2 = 9$$

$$V = \frac{11\,925}{1\,000} - 9 = 2,925; \quad \text{l'écart-type : } \sigma = \sqrt{V}; \quad \sigma = 1,71$$

Notons que l'écart-type sera d'autant plus grand que la série est dispersée autour de la moyenne. Il est nul dans le cas où toutes les données sont identiques.

V/ Série Chronologiques:

▪ Le coefficient multiplicateur qui permet de passer de l'année b à l'année n est égal au quotient de la valeur de l'année n par la valeur de l'année b .

$$C = \frac{\text{Valeur de l'année } n}{\text{Valeur de l'année } b}$$

▪ L'indice I de l'année n , base 100 en l'année b est égal à $I = C \times 100$.