# CHAPITRE II
# LES PARAMETRES DE TENDANCE CENTRALE

Le statisticien ou le chercheur dispose encore d'une multitude de techniques de synthèse qui permettent de renseigner sur le phénomène étudié, à travers la caractérisation de la série. Cette caractérisation, consiste à représenter la série statistique par des chiffres illustratifs ou représentatifs, appelés paramètres, que le chercheur calcule lui-même, selon la nature du phénomène qu'il veut mettre en évidence.

Il existe en statistique descriptive trois grands groupes de paramètres :

- Les paramètres de tendance centrale ou de position,
- Les paramètres de dispersion,
- Les paramètres de concentration.

Un paramètre statistique pour être pertinent, objectif et correctement utilisé doit remplir les (6) conditions suivantes émises par le statisticien britannique Yule (Goldfarb. B et Pardoux. C, 2013).

1. Etre défini de façon objective ou précise,
2. Dépendre de toutes les valeurs (modalités) de la série,
3. Avoir une signification concrète facile à concevoir,
4. Etre simple à calculer,
5. Etre peu sensible aux fluctuations d'échantillonnage,
6. Se prêter aisément au calcul algébrique.

Ainsi, plus un paramètre satisfait un grand nombre de ces conditions, plus il est considéré comme pertinent, voire parfait.

L'objet de ce troisième chapitre est l'étude des paramètres du premier groupe, à savoir ; les paramètres de tendance centrale. Nous y étudions successivement : le mode (section I), la médiane (section II), la moyenne arithmétique (section III) .

## Section I : Le mode

### I.1-Définition

Le mode d'une série ou distribution statistique, noté « Mo », est la modalité dominante ou la plus fréquente, c'est-à-dire la valeur de la variable qui correspond au plus grand effectif ou plus grande fréquence relative. Par exemple, dans la série : {5, 6, 7, 7, 8, 8, 8, 9, 10} ; Mo=8, 8 est la valeur qui se répète le plus (3 fois).

### I.2-Determination pratique

Le mode se détermine de manière algébrique et graphique, et la détermination diffère selon qu'il s'agisse de variable statistique discrète (VSD) ou variable statistique continue (VSC).

#### 2.1-Cas de la variable statistique discrète (VSD)

Lorsque la variable est discrète, le mode peut être repéré :

- **Dans le tableau :** c'est la modalité « xi » pour laquelle l'effectif ou la fréquence relative est le plus élevé ;
- **Sur le diagramme en bâtons :** c'est la modalité xi correspondant au bâton le plus haut.

#### 2.1.1-Par la méthode analytique ou algébrique

Il s'agit de repérer dans la colonne (ni) ou (fi) du tableau statistique l'effectif le plus élevé (ou le chiffre le plus élevé), la modalité (xi) (dans la colonne des xi) correspondant à celui-ci est le mode.

**Exemple 01 :** Soit la distribution suivante :

| Xi | ni |
|---|---|
| 1 | 4 |
| 2 | 3 |
| 3 | 8 |
| 4 | 3 |
| 5 | 2 |
| Total | 20 |

**-Calcul du mode de cette distribution.**

Il suffit de regarder la colonne (ni), le plus grand nombre c'est ni=8, la modalité correspondante dans la colonne (xi) c'est 3. Donc le Mode = 3.

**Exemple 02 :** Soit la distribution suivante :

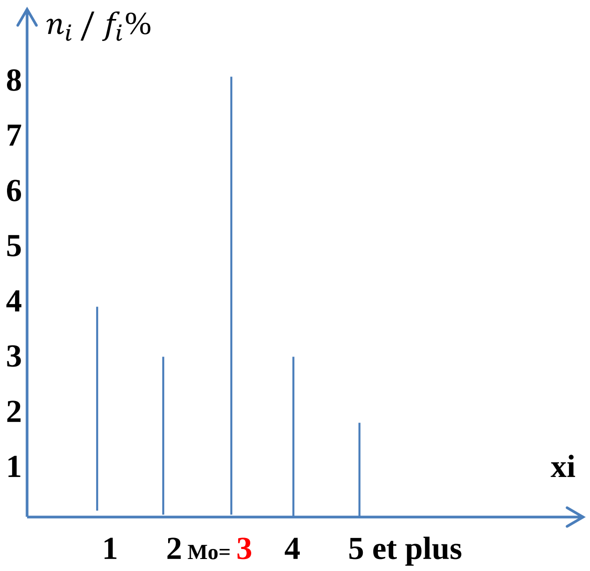
| Nombre d'enfants | Nombre de ménages | $f_i\%$ |
|---|---|---|
| 0 | 5 | 5.62 |
| 1 | 17 | 19.10 |
| 2 ← | 31← | 34.83 |
| 3 | 20 | 22.47 |
| 4 | 11 | 12.36 |
| 5 et plus | 5 | 5.62 |
| ∑ | 89 | 100% |

$M_o$=2 enfants puisque "$n_i$" le plus élevé est 31 correspondant à la fréquence relative $f_i$ la plus élevée aussi34.83%.

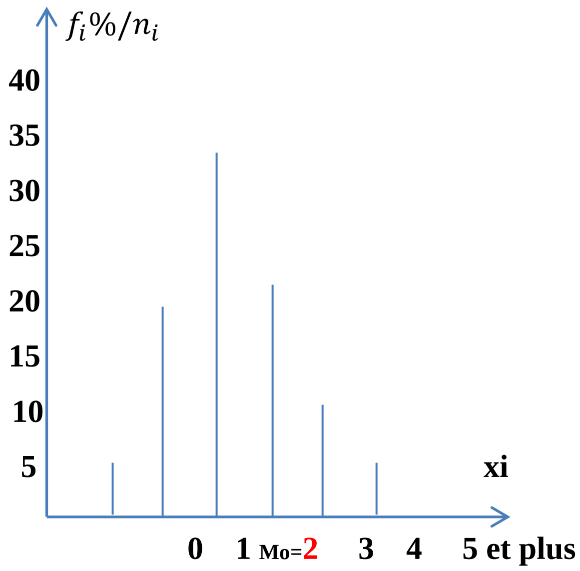
### 2.1.2-Par la méthode graphique :

Il s'agit d'une variable statistique discrète (valeurs individuelles ou isolées), donc le graphique correspondant est le diagramme en bâtons. Le principe est de repérer le bâton le plus long (dont la hauteur indique l'effectif), la valeur (xi) correspondante sur l'axe horizontal (l'axe xx`) est le mode. On aura alors le graphique suivant :

Le mode graphiquement pour l'exemple n°01




Le mode graphiquement pour l'exemple n°02




### 2.2-Cas de la variable statistique continue (VSC)

Ici, les données sont regroupées sous forme de classes. La valeur (xi) correspondant au mode appartient forcément à une classe. Cette classe s'appelle classe modale. Elle est repérée comme la classe « la plus dense », c'est-à-dire celle correspondant à la plus grande densité (di). Deux situations peuvent se présenter :

### 2.2.1-Cas où les amplitudes de classes constantes ou égales

On détermine d'abord la classe modale. Elle correspond à l'effectif le plus élevé et à la fréquence relative la plus élevée, puis on repère à l'intérieur de cette classe la valeur du mode. Dans ce cas, la classe la plus dense est celle aussi correspondant au plus grand effectif, car dans ce cas $d_i = \frac{n_i}{a_i}$ revient à diviser tous les effectifs par la même amplitude, donc la classe qui a le plus grand effectif sera aussi la plus dense.

De manière algébrique, la détermination du mode suit les étapes suivantes :

- Repérer dans la colonne (ni) ou (fi) la plus grande valeur (le plus grand effectif) ;
- Repérer la classe modale correspondant au plus grand effectif,
- Appliquer la formule du mode réservée au cas de la variable continue :

$$\mathrm{M_0} = x_0 + a_i \frac{(\mathrm{n_{mo}} - n_{mo-1})}{(\mathrm{n_{mo}} - n_{mo-1}) + (\mathrm{n_{mo}} - n_{mo+1})}$$

Ou bien :

$$\mathrm{M_0} = x_0 + a_i \frac{(f_{i_{\mathrm{mo}}} - f_{i_{mo-1}})}{(f_{i_{\mathrm{mo}}} - f_{i_{mo-1}}) + (f_{i_{\mathrm{mo}}} - f_{i_{mo+1}})}$$

Avec :

- $x_0$ = borne inférieure de la classe modale.
- $a_i$ = amplitude de la classe modale.
- $\mathrm{n_{mo}}$= effectif de la classe modale.
- $n_{mo-1}$ = effectif de la classe avant ou précédent la classe modale.
- $n_{mo+1}$= effectif ou fréquence de la classe après ou suivant la classe modale.
- $f_{i_{\mathrm{mo}}}$ = fréquence de la classe modale.
- $f_{i_{mo-1}}$ = fréquence de la classe avant ou précédent la classe modale.
- $f_{i_{mo+1}}$ = fréquence de la classe après ou suivant la classe modale.

**Exemple :** voici les dépenses mensuelles d'un échantillon de 100 ménages.

| **Dépenses $(10^3)$ DA** | **Nombre de ménages** | $a_i$ | $f_i$ |
|---|---|---|---|
| [10-30[ | 10 | 20 | 0.1 |
| [30-50[ | 20 | 20 | 0.2 |
| [50-70[ | ← 45 | 20 | 0.45 |
| [70-90[ | 17 | 20 | 0.17 |
| [90-110[ | 8 | 20 | 0.08 |
| $\sum$ | 100 | / | 1 |

$n_i$ le plus élevé est 45 donc la classe modale est [50-70[

$$M_0 = x_0 + a_i \frac{(n_{mo} - n_{mo-1})}{(n_{mo} - n_{mo-1}) + (n_{mo} - n_{mo+1})}$$

$$M_0 = 50 + 20 \frac{(45 - 20)}{(45 - 20) + (45 - 17)}$$

$\rightarrow \mathbf{M_o = 59.43 \times 10^3}$ **DA**

Ou bien :

$$M_o = x_0 + a_i \frac{(f_{i_{mo}} - f_{i_{mo-1}})}{(f_{i_{mo}} - f_{i_{mo-1}}) + (f_{i_{mo}} - f_{i_{mo+1}})}$$

$$M_o = 50 + 20 \frac{(0.45 - 0.2)}{(0.45 - 0.2) + (0.45 - 0.17)}$$

$\rightarrow \mathbf{M_o = 59.43 \times 10^3}$ **DA**

### 2.2.2-Cas où les amplitudes des classes sont inégales ou non constantes (ai # Cte)

Lorsque les amplitudes sont inégales, il faut corriger les effectifs ou les fréquences relatives. Le mode sera calculé en utilisant ces effectifs ou fréquences corrigés.

En ce sens, la classe la plus dense n'est pas forcément celle qui correspond au plus grand effectif. En effet, $d_i = \frac{n_i}{a_i}$, (ai # Cte), signifie que l'on divise les effectifs ou les fréquences par des valeurs de $a_i$ différentes, ce qui implique que l'on peut tomber sur des situations où la classe correspondant au plus grand effectif n'est pas la classe la plus dense.

Aussi, avant de calculer le mode, il faut comme précédemment, calculer les densités (ou calculer les effectifs corrigés nic).

La détermination du mode suit les étapes suivantes :

- Calculer les densités ($d_i$) ou les effectifs corrigés ($n_{ic}$),
- Repérer dans la colonne ($d_i$) ou ($n_{ic}$) la plus grande valeur (plus grande densité),
- Repérer la classe modale correspondant à cette plus grande valeur,
- Appliquer la formule du mode réservée au cas continu avec amplitude non constante :

$$\mathbf{M_0 = x_0 + a_i \frac{(n_{icmo} - n_{icmo-1})}{(n_{icmo} - n_{icmo-1}) + (n_{icmo} - n_{icmo+1})}}$$

Avec :

- $x_0$ = borne inférieure de la classe modale.
- $a_i$ = amplitude de la classe modale.
- $n_{icmo}$ = effectif ou fréquence de la classe modale.
- $n_{icmo-1}$ = effectif ou fréquence de la classe avant ou précédent la classe modale.
- $n_{icmo+1}$ = effectif ou fréquence de la classe après ou suivant la classe modale

Ou bien :

$$\mathbf{M_0} = x_0 + a_i \frac{(\mathbf{f}_{icmo} - \mathbf{f}_{icmo-1})}{(\mathbf{f}_{icmo} - \mathbf{f}_{icmo-1}) + (\mathbf{f}_{icmo} - \mathbf{f}_{icmo+1})}$$

Ou bien encore avec les densités :

$$\mathbf{M_0} = x_0 + a_i \frac{(\mathbf{d}_{imo} - d_{imo-1})}{(\mathbf{d}_{imo} - d_{imo-1}) + (\mathbf{d}_{imo} - d_{imo+1})}$$

Avec :

- $x_0$ = borne inférieure de la classe modale.
- $a_i$ = amplitude de la classe modale.
- $\mathbf{d}_{imo}$= densité de la classe modale.
- $d_{imo-1}$ = densité de la classe avant ou précédent la classe modale.
- $d_{imo+1}$= densité de la classe après ou suivant la classe modale.

**Remarque :**

L'étudiant a le choix dans les applications (exercices) de choisir, parmi les formules proposées.

**<u>Exercice n°01 (« ai » constante)</u> :** Calculer le mode de la distribution suivante et représenter le graphiquement :

| Classes | Effectifs |
|---|---|
| [10-20[ | 10 |
| [20-30[ | 10 |
| [30-40[ | 15 |
| [40-50[ | 5 |

Il faut d'abord, au préalable, calculer les amplitudes et voir si elles sont constantes ou pas. On ajoute alors une colonne (ai) au tableau, on obtient le tableau ci-après :

| **Classes** | $n_i$ | $f_i$ | $a_i$ |
|---|---|---|---|
| [10-20[ | 10 | 0,250 | 10 |
| [20-30[ | 10 | 0,250 | 10 |
| [30-40[ | 15 | 0,375 | 10 |
| [40-50[ | 5 | 0,125 | 10 |
| **Total** | 40 | 1 | - |

On constate donc que l'amplitude (ai) est constante (ai=10). Par conséquent, la classe modale est celle qui correspond au plus grand effectif. Si on regarde dans la colonne (ni), la plus grande valeur c'est 15, (ni=15). La classe correspondant (c'est-à-dire sur la même ligne) c'est [30-40], c'est la classe modale, donc $x_0$ = 30.

On applique la formule du mode développée précédemment ((1) ou (2)) :

$$M_0 = x_0 + a_i \frac{(n_{mo} - n_{mo-1})}{(n_{mo} - n_{mo-1}) + (n_{mo} - n_{mo+1})}$$

$$M_0 = 30 + 10 \frac{(15 - 10)}{(15 - 10) + (15 - 5)}$$

$M_0 = 33.33$ DA

Ou bien :

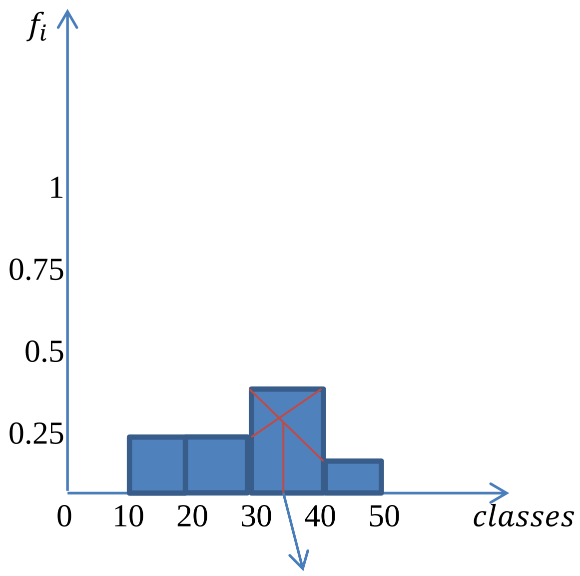
$$M_0 = x_0 + a_i \frac{(f_{i_{mo}} - f_{i_{mo-1}})}{(f_{i_{mo}} - f_{i_{mo-1}}) + (f_{i_{mo}} - f_{i_{mo+1}})}$$

$$M_0 = 30 + 10 \frac{(0.375 - 0.250)}{(0.375 - 0.250) + (0.375 - 0.125)}$$

$M_0 = 33.33$ DA

**De manière graphique**, il s'agit d'une VSC, donc la fonction de distribution est représentée graphiquement par l'histogramme. Compte tenu de l'amplitude de classe (constante ou pas), le mode se détermine de la même manière à partir du rectangle de la classe modale (la plus dense) dans l'histogramme).

Il suffit pour cela de relier par une ligne (ou un segment) la borne supérieure du rectangle de la classe modale à la borne supérieure du rectangle de la classe précédente. Ensuite, relier par une autre ligne ou segment la borne inférieure du rectangle de la classe modale à la borne inférieure de la classe suivante. Le point de croisement des deux lignes ou deux segments est ensuite projeté sur l'axe des abscisses (l'axe des classes ou horizontal). La valeur ($x_i$) correspondante sur cet axe est le mode. Voyons cela dans la représentation graphique suivante :




$\mathbf{M_o = 33.33 \times DA}$

**<u>Exercice 02 (« ai » inconstante)</u> :** Calculer le mode de la distribution suivante :

| Classes | Effectifs |
|---|---|
| [10-20[ | 10 |
| [20-30[ | 20 |
| [30-50[ | 30 |
| [50-80[ | 25 |

Pour calculer le mode, il faut d'abord vérifier les amplitudes (ai). On construit comme précédemment une colonne (ai) de laquelle dépendra le nombre de colonnes à ajouter au tableau, selon que « ai » soit Cte ou pas.

| Classes | Effectifs | ai | $\mathbf{d_i}$ | $\boldsymbol{n_{ic}}$ |
|---|---|---|---|---|
| [10-20[ | 10 | 10 | 1 | 10 |
| [20-30[ | 20 | 10 | 2 | 20 |
| [30-50[ | 30 | 20 | 1.5 | 15 |
| [50-80[ | 25 | 30 | 0.83 | 8.33 |
| **Total** | 85 | - | - | / |

Si on regarde la colonne (ai), on constate que l'amplitude de classe n'est pas constante. La classe modale est donc la classe la plus dense, c'est-à-dire celle qui a la plus grande densité ou le plus grand effectif corrigé, et qui n'est pas forcément celle qui a le plus grand effectif. L'amplitude de base (la plus petite amplitude), d'après le tableau est $a_0 = 10$. On en déduit les $(n_{ic})$. Ou bien, on calcule directement les densités $(d_i)$.

Dans la colonne $(d_i)$, la plus grande valeur c'est 2. Sur la même ligne, la plus grande valeur correspondante dans la colonne $(n_{ic})$ doit automatiquement être la plus grande valeur de la colonne$(n_{ic})$. Si ce n'est pas le cas, cela voudrait dire qu'il y a erreur de calcul quelque part que l'étudiant doit vérifier et corriger.

Dans notre tableau, la densité la plus élevée $d_2 = 2$, correspond également sur la même ligne, dans la colonne$(n_{ic})$, à la plus grande valeur $n_{ic}$, $n_{ic} = 20$. La classe correspondante, à savoir ; [20-30[, est la classe modale ou la plus dense.

$$\mathbf{M_0} = x_0 + a_i \frac{(\mathbf{f}_{icmo} - \mathbf{f}_{icmo-1})}{(\mathbf{f}_{icmo} - \mathbf{f}_{icmo-1}) + (\mathbf{f}_{icmo} - \mathbf{f}_{icmo+1})}$$

$$\mathbf{M_0} = 20 + 10 \frac{(20 - 10)}{(20 - 10) + (20 - 15)}$$

$$\mathbf{M_0} = 26.67\ \mathbf{DA}$$

Ou bien :

$$\mathbf{M_0} = x_0 + a_i \frac{(\mathbf{d}_{imo} - d_{imo-1})}{(\mathbf{d}_{imo} - d_{imo-1}) + (\mathbf{d}_{imo} - d_{imo+1})}$$

$$\mathbf{M_0} = 20 + 10 \frac{(2-1)}{(2-1)+(2-1.5)}$$

$$\mathbf{M_0} = 26.67\ \mathbf{DA}$$

**Exercice 03 :** Voici les dépenses hebdomadaires de 100 ménages.

| **Dépenses** $(10^2\ DA)$ | $\mathbf{n_i}$ | $\mathbf{a_i}$ | $\mathbf{n_{ic}}$ | $\mathbf{d_i}$ |
|---|---|---|---|---|
| [30-50[ | 10 | 20 | 10 | 0.5 |
| [50-80[ | 30 | 30 | 20 | 1 |
| [80-100[ | 25 | 20 | 25 | 1.25 |
| [100-130[ | 27 | 30 | 18 | 0.9 |
| [130-150[ | 8 | 20 | 8 | 0.4 |
| $\sum$ | 100 | / | / | / |

La classe modale est [80-100[car l'effectif corrigé le plus élevé est $n_{3c} = 25$, elle correspond aussi à la densité la plus élevée $d_3 =$ 1.25. Une fois la classe modale déterminée, on applique les formules du carré ci-après.

$$\mathbf{M_0} = x_0 + a_i \frac{(\mathbf{n}_{icmo} - \mathbf{n}_{icmo-1})}{(\mathbf{n}_{icmo} - \mathbf{n}_{icmo-1}) + (\mathbf{n}_{icmo} - \mathbf{n}_{icmo+1})}$$

$$\mathbf{M_0} = 80 + 20 \frac{(25 - 20)}{(25 - 20) + (25 - 18)}$$

$$\mathbf{M_0} = 88.33\ \mathbf{DA}$$

Ou bien :

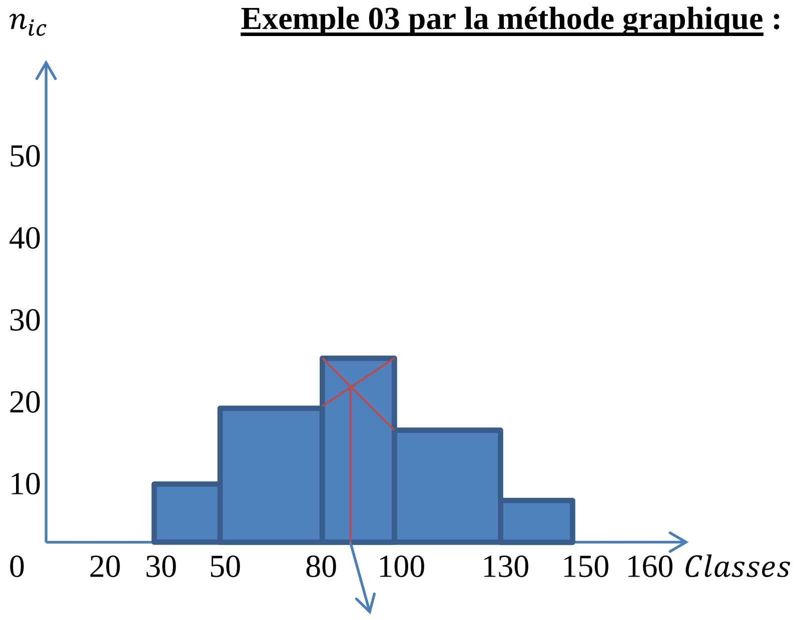
$$\mathbf{M_0} = x_0 + a_i \frac{(\mathbf{d}_{imo} - d_{imo-1})}{(\mathbf{d}_{imo} - d_{imo-1}) + (\mathbf{d}_{imo} - d_{imo+1})}$$

$$\mathbf{M_0} = 80 + 20 \frac{(1.25-1)}{(1.25-1)+(1.25-0.9)}$$

$$\mathbf{M_0} = 88.33\ \mathbf{DA}$$

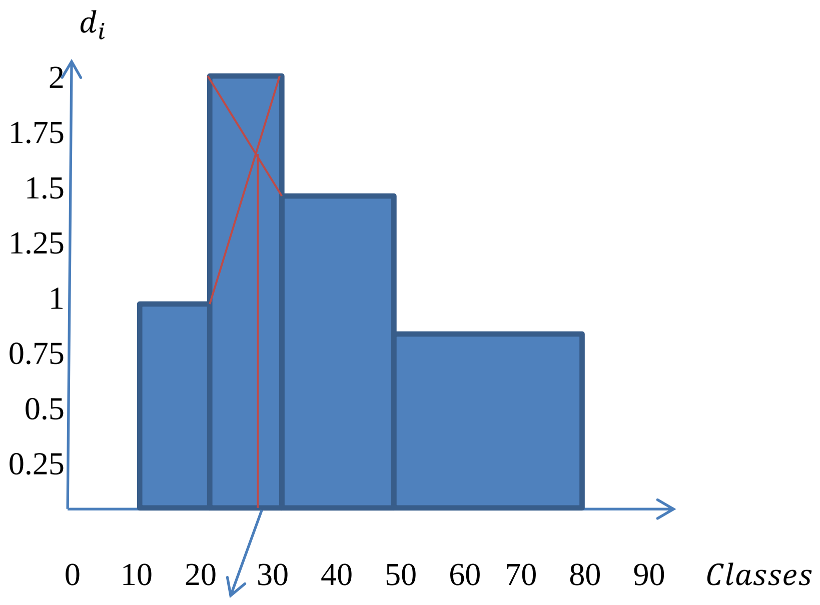
**<u>La manière graphique</u> :** Reprenant l'exemple 3 précédent et traçant le graphique correspondant. Il s'agit d'une VSC, donc le graphique correspondant est un histogramme, comme suit :

**<u>Exemple 03 par la méthode graphique</u> :**




$\mathbf{M_o}$ **= DA 88.33 DA**

**<u>Exemple 02 par la méthode graphique</u> :**




$\mathbf{M_o = 26.67 \times DA}$

**Remarque :**

- Le mode satisfait aux 1ere, 3éme, 4éme conditions de Yule mais pas aux autres conditions.
- Son principal avantage est d'être déterminé très rapidement et sa signification est simple. On l'utilise surtout dans le cas des variables statistiques discrètes et pour faire apparaitre facilement une première estimation de valeur centrale de la distribution.

- Une série statistique peut présenter un, deux ou plusieurs modes à la fois, comme elle peut ne pas présenter du tout (ni= constante), on l'appelle alors une distribution ou série « amodale ».

## Section II : La médiane « $M_e$ »

### II.1-Définition

Etymologiquement[1], médiane signifie « Milieu ». C'est le milieu de la distribution ordonnée. Notée « Me, elle est la modalité $(x_i)$ qui est située dans la série ordonnée de telle sorte que le nombre de modalités situées avant la médiane (ou inférieure à la médiane) est le même que celui des modalités situées après la médiane (ou supérieure à la médiane). Autrement dit, la médiane est la modalité qui partage l'effectif total des modalités en deux groupes de même taille (50% chacun).

C'est donc la valeur $(x_i)$ pour laquelle 50% des individus présentent des modalités inférieures à Me, et 50% des individus restant présentent des modalités supérieures à Me. Les modalités étant, bien entendu, obligatoirement ordonnées par ordre croissant. La médiane est donc un paramètre de position.

### II.2-Détermination pratique de la médiane :

La médiane se détermine de deux manières ; algébrique et graphique, et la détermination diffère selon qu'il s'agisse de VSD ou VSC.

### 2.1-Cas de la variable statistique discrète

### 2.1.1-Determination de la médiane par la méthode algébrique

Dans ce cas la détermination de la médiane dépend du nombre d'individus (N). Avant de déterminer la médiane, les individus de la population doivent être rangés dans un ordre croissant ou décroissant.

#### 2.1.1.1-Si N est un nombre impair

Donc N = 2K+1, la médiane correspond à la valeur de K+1$^{éme}$ observation (le rang de $M_e$ est K+1).

Dans ce cas il existe une modalité, parmi toutes les modalités de la série, qui divise la série en deux groupes de même effectif (N/2). Cette modalité est la médiane. On la détermine comme suit :

N impair, ceci implique que mathématiquement N s'écrit : N = 2k+1.

[1] Du point de vue de l'origine ou des racines du mot.

**NB/-** k est un effectif cumulé croissant. Il donne la position (ou le numéro) de la médiane dans la série ordonnée.

Dans ce cas la médiane est la modalité correspondant à la position numéro (k+1)$^{éme}$. Il suffit alors de calculer k.

**Exemple 01 :**

Soit la série ordonnée suivante : 3 – 6 – 12 – 18 – 20 – 23 – 28. Déterminer la médiane. Il y a sept (7) modalités ou individus dans la série, soit N=7.

7 étant un nombre impair, il s'écrit donc N = 2k+1 = 7 $\longrightarrow$ k = N-1/2 =3

Donc la médiane est la modalité numéro (k+1)$^{éme}$, soit 3+1=4 ou la 4$^{éme.}$

Dans la série ordonnée, on constate que la 4$^{éme}$ modalité c'est 18. Donc Me = 18.

**Exemple 02 :** Soit la série des notes suivantes :{14, 7, 10, 5, 9, 8, 11}

Ordonnons la série :{ 5, 7, 8, 9, 10, 11, 14}

N=7→N est impair → N = 2K+1→K = $\frac{N-1}{2} = \frac{7-1}{2} \rightarrow K = 3\ et\ K + 1 = 4$; $M_e$correspond à la valeur de la 4éme observation, donc $M_e = 9$.

**2.1.1.2-Si N est un nombre pair**

Dans ce cas il existe non pas une modalité médiane mais un intervalle médian délimité par [k$^{éme}$ et (k+1)$^{éme}$ ] modalités .

**<u>Exemple 01</u> :**

Soit la série ordonnée suivante : 3 – 6 – 12 – 18 – 20 – 23 – 28 - 30. Déterminer la médiane. Il y a sept (8) modalités ou individus dans la série, soit N=8.

7 étant un nombre impair, il s'écrit donc N = 2k = 8 → k = N/2 =4 → l'intervalle médian est délimité par la 4éme et (4+1= 5$^{éme}$) modalités.

La 4$^{éme}$ correspond à la modalité 18, et la 5éme modalité à 20. Donc l'intervalle médian est : [18-20[, soit Me ≈ (18+20/2)=19

**<u>Exemple 02</u> :** Soit la série des notes suivantes : :{ 5, 7, 8, 9, 10, 11, 12, 14}

N=8→N est pair → N = 2K→K = $\frac{N}{2} = \frac{8}{2} \rightarrow K = 4\ et\ K + 1 = 5$; $M_e \in [4^{éme}, 5^{éme}]$ →$M_e \in$ [9, 10[→$M_e = \frac{9+10}{2} = 9.5$

### 2.1.2-Determination de la médiane dans le tableau statistique

Dans ce cas où N est trop élevé, on a recours alors au tableau statistique, où la colonne (Ni) nous permet de repérer les positions des modalités qu'on cherche, en suivant la même logique. La médiane est déterminée à partir des $f_i^{\nearrow}$ ou des $n_i^{\nearrow}$. On choisit la valeur $f_i^{\nearrow}$ égale ou immédiatement supérieure à 0.5, ou la valeur des $n_i^{\nearrow}$ égale ou immédiatement supérieure à ( $^{N}/_{2}$). Après, on suit le sens des flèches comme indiqué dans les tableaux suivants :

**Exemple 01 :** Déterminer la médiane de la distribution suivante :

| $x_i$ | $n_i$ | $N_i$ |
|---|---|---|
| 0 | 10 | 10 |
| 1 | 32 | 42 |
| 2 | 36 | 78 |
| 3 | 15 | 93 |
| 4 | 5 | 98 |
| 5 | 2 | 100 |
| **Total** | 100 | - |

Dans cet exemple on déterminera la médiane en utilisant directement la colonne des $N_i$. N=100, donc N est un nombre pair. Il existe donc un intervalle médian.

N=2k → k = 50 et (k+1) = 51. Me ∈ $[50éme; 51éme[$ modalités. C'est-à-dire Me se situe à la position entre la 50$^{éme}$ et la 51$^{éme}$ modalités du tableau. K étant un effectif cumulé (Ni), on peut lire dans le tableau que la 50$^{éme}$ et la 51$^{éme}$ se situent entre les effectifs cumulés $N_2$= 42 et $N_3$= 78, c'est-à-dire entre la 42éme et la 78éme modalités. Or, d'après le tableau on peut lire que de la 43éme jusqu'à 78éme ce sont toutes des modalités égale à 2.

Autrement dit, les deux modalités que nous cherchons (la 50éme et51éme) font partie de ces modalités puisqu'elles sont situées entre la 43éme et 78éme modalités. Donc notre intervalle médian est $[2 - 2]$, ce qui donne Me = 2. Ou bien de manière plus simple encore, il suffit de calculer N/2, celui-ci nous donne l'effectif cumulé ou la position de la modalité médiane. Si cette position se situe entre deux effectifs cumulés ou entre deux lignes de la colonne Ni, on prend directement la ligne du bas, c'est-à-dire la valeur immédiatement supérieure ou égale à (N/2).

C'est ce même principe qu'on utilise pour la détermination graphique.

**Exemple 02 :** Soit le nombre d'erreurs d'assemblage sur un ensemble d'appareils

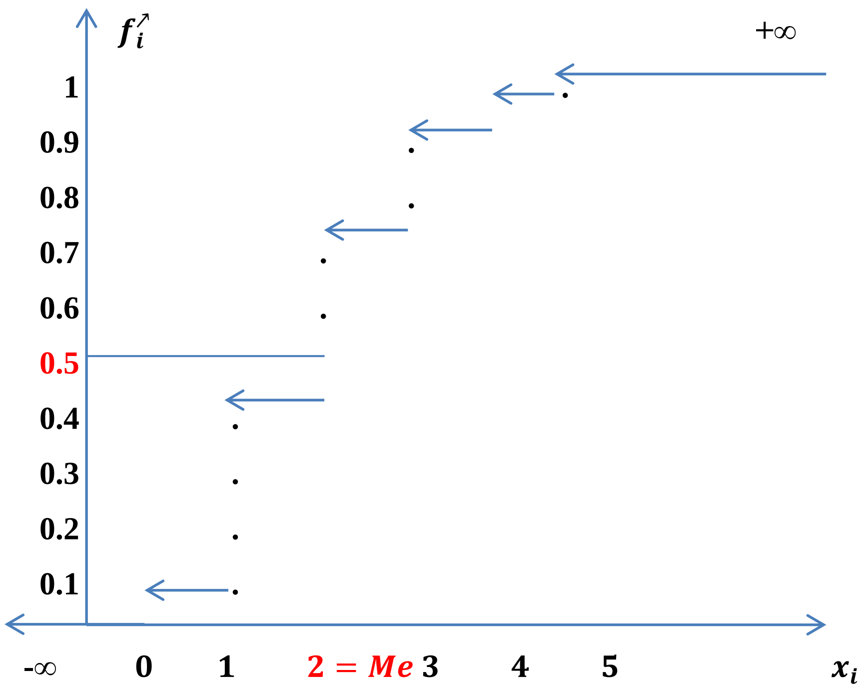
| Nombre d'erreurs | Nombre d'appareils | $f_i$ | $f_i^{\nearrow}$ | $n_i^{\nearrow}$ |
|---|---|---|---|---|
| 0 | 20 | 0.1 | 0.1 | 20 |
| 1 | 65 | 0.325 | 0.425 | 85 |
| 2 | ← 70← | 0.35 | 0.775 | 155 |
| 3 | 30 | 0.15 | 0.925 | 185 |
| 4 | 10 | 0.05 | 0.975 | 195 |
| 5 | 5 | 0.02 | 1 | 200 |
| $\sum$ | 200 | 1 | / | / |

→ $N/_2$ $ou$ 0.5

N=200, $N/_2$= 100 (Le rang de la médiane est 100 ou 0.5) : la valeur des $n_i^{\nearrow}$ qui est immédiatement supérieure à 100 est 155, et la fréquence cumulée croissante "$f_i^{\nearrow}$" immédiatement supérieure à 0.5 est 0.775, donc $\mathrm{M_e} = 2$ erreurs.

**Remarque :** Dans un tableau statistique, on considère que N est très grand et qu'il est fort possible de tomber sur des valeurs K et K+1 qui soient identiques, donc le rang de la médiane est N/2 ou 0.5 dans les cas de N pair et de N impair.

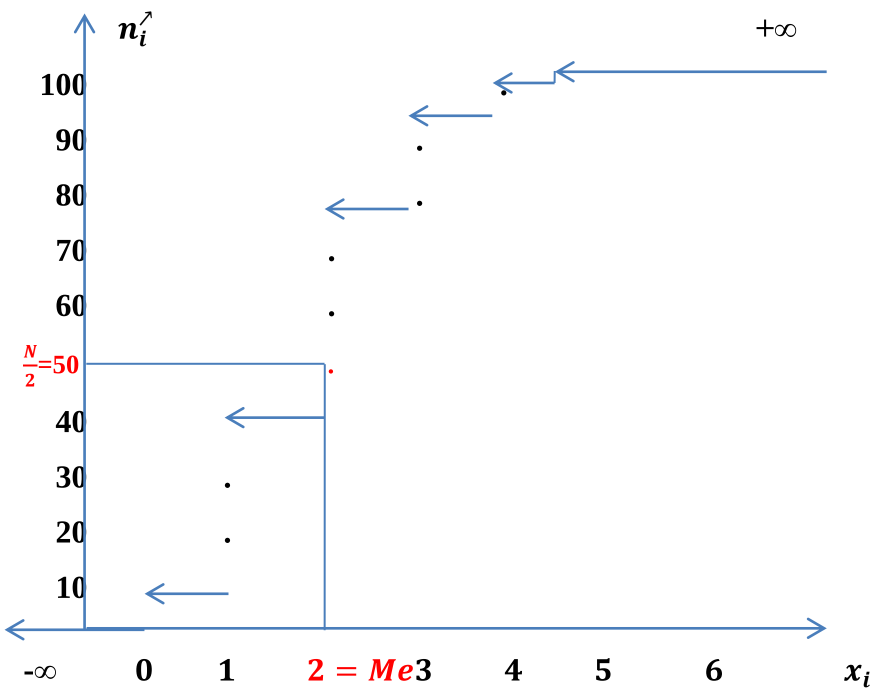
### 2.1.3-Détermination de la médiane par la méthode graphique



Dans le cas d'une VSD, la fonction cumulative {xi ; Ni} ou {xi ; Fi} est représentée graphiquement par la courbe cumulative en escalier. La médiane se déterminant par les effectifs cumulés, elle est donc logiquement déterminée graphiquement par la courbe représentant ces derniers. On peut illustrer la méthode de détermination à partir de la courbe cumulative de l'exemple c1 précédent.

Sur l'axe vertical (axe des Ni ou Fi), repérer la valeur correspondant à N/2. On projette, ensuite, ce point à l'horizontal parallèlement à l'axe des Xi, sur la courbe en escaliers. La première contremarche sur laquelle on tombe est directement projetée par une ligne verticale sur l'axe des Xi, la valeur sur laquelle on tombe est la médiane.

**NB/-** Si on tombe sur un palier au lieu d'une contremarche, cela voudrait dire que nous avons à faire un intervalle médian délimité par deux valeurs différentes, ce qui n'est pas le cas de l'exemple 1c.




## 2.2- Variable statistique continue

Dans ce cas, les données sont présentées sous forme de classes, la modalité médiane appartient forcément à une classe, appelée intervalle ou classe médiane, à partir de laquelle il faudrait la déterminer.

Comme pour la VSD, la détermination de la médiane pour une VSC se fait également de deux manières possibles : algébrique et graphique.

### 2.2.1-De manière algébrique

La méthode se déroule suivant les trois étapes ci-après :

- Repérer dans la colonne des Ni ou Fi, la valeur ou la position N/2 ou Fi = 0.5 (soit 50%), notée « $Th_2$ ».
- Déterminer la classe correspondant à cette position, c'est-à-dire la classe médiane.
- Appliquer la formule de la médiane suivante :

$$M_e = x_0 + a_i \frac{N/2 - N_{me-1}}{n_{me}}$$

Ou bien :

$$M_e = x_0 + a_i \frac{0.5 - F_{me-1}}{f_{me}}$$

Avec :

- $X_0$= borne inférieure de la classe médiane.
- a = amplitude de la classe médiane.
- $Th_i = \frac{N}{2} ou\ 0,5$.
- $N_{me}$ ou $F_{me}$ = effectif ou fréquence cumulé correspondant à la classe médiane.
- $N_{me-1}$ ou $F_{me-1}$ = effectif ou fréquence cumulé correspondant à la classe avant la classe médiane.
- $n_{me}$ ou $f_{me}$ effectif absolu ou fréquence relative correspondant à la classe médiane.

**Exemple 01 :** Déterminer la médiane de la distribution suivante :

| Classes | Effectifs |
|---|---|
| 15 – 25 | 26 |
| 25 – 30 | 33 |
| 30 – 40 | 64 |
| 40 – 50 | 7 |
| 50 – 65 | 10 |

Pour déterminer la médiane, on a besoin des effectifs ou fréquences cumulées. On construit la colonne Ni. Le tableau complet nécessaire sera comme suit (on ajoutera d'emblée la colonne $\mathbf{N_i^{\searrow}}$ décroissantes pour les besoin du graphique) :

$Th_2 = N/2 =$ 140/2 = 70, cette valeur se trouve entre les valeurs 59 et 123, c'est-à-dire entre deux lignes, dans la colonne des Ni, ce qui implique qu'on prendra

| Classes | $X_i$ | $n_i$ | $N_i$ | $N_i^{\searrow}$ | $f_i$ | $f_i^{\nearrow}$ | $f_i^{\searrow}$ |
|---|---|---|---|---|---|---|---|
| [15 - 25[ | 20 | 26 | 26 | 140 | 0.19 | 0.19 | 1 |
| [25 – 30[ | 27,5 | 33 | 59 | 114 | 0.24 | 0.43 | 0.81 |
| [30 – 40[ | 35 | 64 | 123 | 81 | 0.46 | 0.89 | 0.57 |
| [40 – 50[ | 45 | 7 | 130 | 17 | 0.05 | 0.94 | 0.11 |
| [50 - 65[ | 57,5 | 10 | 140 | 10 | 0.07 | 1 | 0.05 |
| **Total** | - | 140 | - | - | | | |

directement la ligne du bas, soit celle correspondant à Ni = 123, la classe correspondante, soit [30 - 40[, est la classe médiane, → Me ∊ [30 - 40[.

$$M_e = x_0 + a_i \frac{N/2 - N_{me-1}}{n_{me}} = 30 + 10\left[\frac{70-59}{64}\right] = 31{,}72$$

→Me = 31,72

Ou bien :

$$M_e = x_0 + a_i \frac{0.5 - F_{me-1}}{f_{me}} = 30 + 10\,\frac{0.5-0.43}{0.46}$$

$\rightarrow M_e = 31.72$

**Exemple 02 : Le** tableau suivant donne le nombre de litres de carburant consommé par 80 véhicules d'une entreprise « X ».

| **Nombre de litres** | $n_i$ | $f_i$ | $F_i$ | $N_i$ |
|---|---|---|---|---|
| [50-100[ | 16 | 0.2 | 0.2 | 16 |
| [100-200[ | 48 | 0.6 | 0.8 | 64 |
| [200-250[ | 08 | 0.1 | 0.9 | 72 |
| [250-500[ | 04 | 0.05 | 0.95 | 76 |
| [500-600[ | 04 | 0.05 | 1 | 80 |
| $\sum$ | **80** | / | / | / |

N/2=80/2=40, 64 est l'effectif cumulé croissant qui est immédiatement supérieur à 40, donc la classe médiane est [100-200[.

$$M_e = x_0 + a_i \frac{N/2 - N_{me-1}}{n_{me}} = 100 + [\,100\left[\frac{40-16}{48}\right] = 150$$

→Me = 150

Ou bien :

$$M_e = x_0 + a_i \frac{0.5 - F_{me-1}}{f_{me}} = 100 + [100\,\frac{0.5-0.2}{0.6}$$

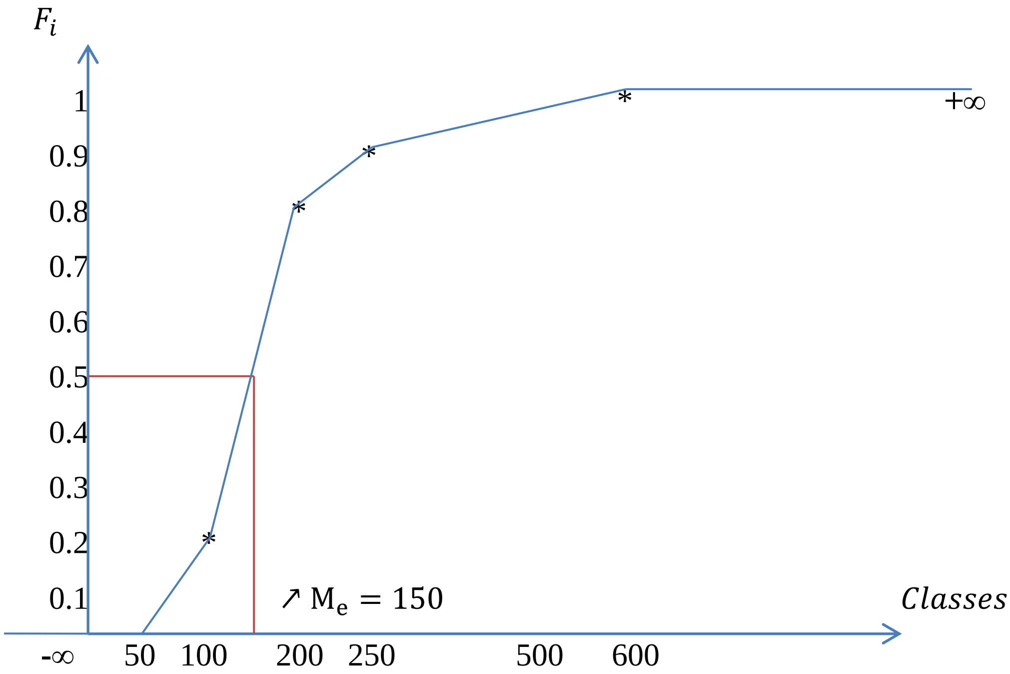
$\rightarrow M_e = 150$

**2.2.2-De manière graphique :**

A partir des deux courbes cumulatives en « S », on détermine la médiane rn projetant le point de rencontre des deux courbes sur l'axe horizontal. La valeur correspondante est la médiane.

**Remarque :**

Logiquement le point où se rencontrent les deux courbes est le point correspondant sur l'axe vertical à Ni = N/2 ou Fi = 0,5 (50%). Il a donc pour coordonnées (xi ; Ni) = (Me ; N/2).




**Remarque :**

- La médiane satisfait assez bien les conditions de YULE à l'exception de la 2éme et 6éme ; elle ne se prête pas aux calculs algébriques et ne dépend pas de toutes les valeurs.
- Elle n'est pas affectée par les valeurs extrêmes, puisqu'elle ne dépend des valeurs que par leur position (leurs effectifs cumulés).

**Notes :**

- Le mode intéresse la fonction de distribution ou la distribution des fréquences, soit les couples {xi ; ni} ou {xi ; fi}, donc graphiquement il se détermine par le diagramme en bâtons pour la VSD et l'histogramme pour la VSC.
- La médiane intéresse la fonction cumulative ou la fonction de répartition, soit les couples {xi ; Ni} ou {xi ; Fi}, donc logiquement elle se détermine par la courbe cumulative en escaliers pour la VSD et les deux courbes cumulatives en « S » pour la VSC.

### 2.3-Géneralisation de la médiane : les quantiles

En suivant la même logique de définition que celle de la médiane, on peut déterminer d'autres paramètres de position qui permettent de partager une série statistique, non seulement en deux groupes de mêmes effectifs, mais aussi en quatre, dix, cents,… ; groupes de mêmes effectifs. On appelle ces paramètres les « quantiles ».

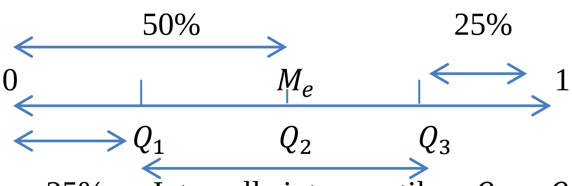
Nous définissons dans ce cours les trois quantiles les plus usités, à savoir ; les *quartiles,* les *déciles et* les *centiles.*

**2.3.1-Les quartiles :**

Notés « $Q_i$», ce sont des paramètres de position et aussi des modalités xi qui, par leurs positions (notées « Thi », partagent la série statistique en quatre groupes de mêmes effectifs, soit 25% chacun. Ils sont donc au nombre de trois : $Q_1; Q_2; Q_3$.

- $Q_1$*, appelé premier quartile, est la modalité de la série par rapport* à *laquelle* 25% *des modalités y sont inférieures* (à $Q_1$) et 75% restantes y sont supérieures (à $Q_1$).
- $Q_2$, appelé aussi deuxième quartile, est la modalité xi par rapport à laquelle 50% des modalités de la série y sont inférieures et 50% restantes y sont supérieures. Ce n'est donc rien d'autre que la médiane que nous avons déjà définie ($Q_2 = Me$).
- $Q_3$, appelé également troisième quartile, est la modalité de la série par rapport à laquelle 75% des modalités y sont inférieures et 25% restantes y sont supérieures. Il est donc le contraire de $Q_1$.

L'intervalle [$Q_1; Q_3$[, s'appelle ***<u>l'intervalle interquartile</u>*** et contient 50% des modalités centrales. On peut schématiser ces quartiles comme suit :




25% Intervalle interquartile = $Q_3 - Q_1 = 50\%$ *des modalités centrales.*

**<u>Formule</u> :** elle est très proche de celle de la médiane :

$$Q_i = x_0 + a_i \frac{{}^{iN}/_{4} - N_{Q_{i-1}}}{n_{Q_i}}$$

Avec :

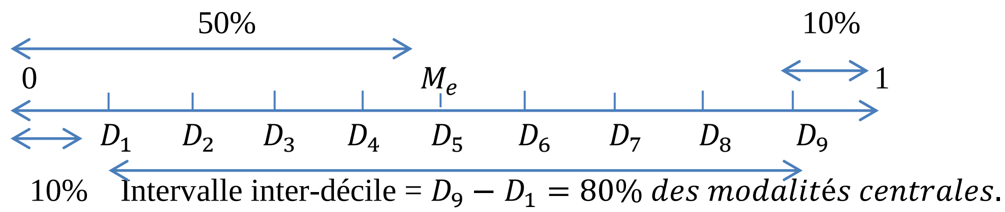
THi = $\frac{iN}{4}$ → C'est l'effectif cumulé croissant repéré dans la colonne Ni du tableau.

**2.3.2- Les déciles**

Notés « $D_i$ », ce sont des paramètres de position et aussi des modalités xi qui partagent la série statistique en dix groupes de même effectif, soit 10% chacun. Ils sont donc au nombre de neuf (9) : $D_1$ ; $D_2$ …. … .. ; $D_9$.

- $D_1$; appelé premier décile, est la modalité xi par rapport à laquelle 90% des valeurs y sont supérieures (à $D_1$), et les 10% restantes y sont inférieures.
- $D_5$; appelé aussi cinquième décile, est la modalité xi par rapport à laquelle 50% des valeurs y sont supérieures (à $D_5$), et les 50% restantes y sont inférieures. Ce n'est donc rien d'autre que la médiane : $\mathbf{D_5 = M_e}$.
- $D_9$ ; appelé neuvième décile, est le contraire de$D_1$. C'est la modalité xi par rapport à laquelle 90% des modalités y sont inférieures, et les 10% restantes y sont supérieures.

L'intervalle [$D_1$ ; $D_9$ [s'appelle intervalle inter-décile. Il contient 80% des modalités centrales. On peut schématiser ces déciles comme dans la page suivante:




**Formule :**

$$D_i = x_0 + a_i \frac{iN/_{10} - N_{D_{i-1}}}{n_{D_i}}$$

Avec :

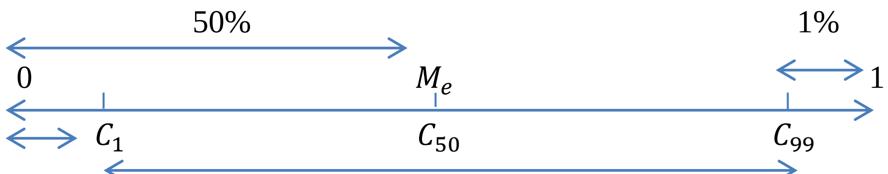
THi = $\frac{iN}{10}$ → C'est l'effectif cumulé croissant repéré dans la colonne Ni du tableau.

### 2.3.3- Les centiles :

Notés « $C_i$ », ce sont des paramètres de position et aussi des modalités xi qui partagent la série statistique en cent groupes de même effectif, soit 1% chacun. Ils sont donc au nombre de 99 : $C_1$ ; $C_2$ ; $C_3$ ; $C_{99}$.

- $C_1$; appelé premier centile, est la modalité xi de la série par rapport à laquelle 1% des valeurs y sont inférieures (à $C_1$), et les 99% restantes y sont supérieures.
- $C_{50}$; appelé aussi cinquantième centile, est la modalité de la série xi par rapport à laquelle 50% des valeurs y sont supérieures (à$C_{50}$), et les 50% restantes y sont inférieures. Ce n'est donc rien d'autre que la médiane : $\mathbf{C_{50} = M_e}$.
- $C_{99}$ ; appelé quatre-vingt-dix neuvième décile, est le contraire de $C_1$. C'est la modalité xi par rapport à laquelle 99% des modalités y sont inférieures, et les 1% restantes y sont supérieures.

L'intervalle [$C_1$ ; $C_{99}$ [s'appelle intervalle inter-centile. Il contient 98% des modalités centrales. On peut schématiser ces centiles comme suit :




1% Intervalle inter-centile → $C_{99} - C_1 = 98\%$ *des modalités centrales*.

**Formule :**

$$C_i = x_0 + a_i \frac{iN/_{100} - N_{C-1}}{n_{C_i}}$$

Avec :

THi = $\frac{iN}{100}$ → C'est l'effectif cumulé croissant repéré dans la colonne Ni du tableau.

**Rappel :**

- $M_e = Q_2 = D_5 = C_{50}$.
- $Q_3 = C_{75}$

**Section III : La moyenne arithmétique ($\overline{X}$)**

**III.1-Définition**

La moyenne arithmétique, notée ($\overline{X}$) d'une variable statistique est égale à la somme des valeurs prises par cette variable, divisée par le nombre d'observations.

**1.1-La moyenne arithmétique simple :**

($\overline{X}$) est dite simple lorsqu'à chaque valeur xi ne correspond qu'une seule observation (toutes les valeurs se répètent une seule fois).

Soient les valeurs suivantes : $\{x_1, x_{2,} \dots \dots .., x_k\}$.

$$(\overline{X}) = \frac{x_1 + x_2 \dots\dots + x_k}{N}$$

$$(\overline{X}) = \frac{1}{N} \sum x_i \ (i \ varie \ de \ 1 \ à \ k)$$

**Exemple :** Soit le nombre d'appareils vendus par cinq magasins au cours d'une semaine {7, 8, 9, 11, 14}

$\overline{X}$=1/5 (7+8+9+11+14)= 9.8 appareils.

### 1.2-La moyenne arithmétique pondérée

On dit qu'une moyenne arithmétique est pondérée lorsqu'à chaque valeur xi peut correspondre plusieurs observations.

$$(\overline{X}) = \frac{x_1 n_1 + x_2 n_2 \ldots\ldots + x_k n_k}{N}$$

$$(\overline{X}) = \frac{1}{N}\sum x_i n_i \ (i \ varie \ de\ 1 \ à \ k)$$

$$(\overline{X}) = \sum x_i \ f_i$$

$$(\overline{X}) = x_1 \ f_1 + x_2 \ f_2 + \ldots\ldots\ldots\ldots + x_k \ f_k$$

**Exemple :** Reprenons l'exemple précédent (08 magasins) : {7, 7 8, 9, 10, 10, 11, 14}

$\overline{X} = \frac{(7\times2)+(8\times1)+(9\times1)+(10\times2)+(11\times1)+(14\times1)}{8}$ = 9.5 appareil

### III.2-Cas de la variable statistique discrète

**Exemple :** soit le nombre d'erreurs d'assemblage sur un ensemble d'appareils

| Nombre d'erreurs | Nombre d'appareils | $x_i\, n_i$ | $f_i$ | $x_i\, f_i$ |
|---|---|---|---|---|
| 0 | 20 | 00 | 0.053 | 00 |
| 1 | 65 | 65 | 0.171 | 0.171 |
| 2 | 140 | 280 | 0.368 | 0.736 |
| 3 | 90 | 270 | 0.237 | 0.711 |
| 4 | 40 | 160 | 0.105 | 0.42 |
| 5 | 25 | 125 | 0.066 | 0.33 |
| $\sum$ | 380 | 900 | 1 | 2.368 |

$$\overline{X} = \frac{1}{N}\sum xini$$

$$\rightarrow \overline{X} = \frac{900}{380} = 2.368 \ erreur$$

Ou bien :

$$\overline{X} = \sum x_i \ f_i$$

$$\rightarrow \overline{X} = 2.368 \ erreur$$

### III.3-Cas de la variable statistique continue

Les modalités sont regroupées en classes et par convention, c'est le centre de classe qui est considérée comme valeur de la variable$x_i$.

**Exemple 01 :** Soit la distribution des employés d'une entreprise « X » en fonction de leur prime de rendement.

| Primes (DA) | $n_i$ | $x_i$ | $x_i n_i$ |
|---|---|---|---|
| [800-1000[ | 26 | 900 | 23 400 |
| [1000-1100[ | 33 | 1050 | 34 650 |
| [1100-1200[ | 64 | 1150 | 73 600 |
| [1200-1300[ | 07 | 1250 | 8 750 |
| [1300-1500[ | 10 | 1400 | 14 000 |
| $\sum$ | 140 | / | 154 400 |

$$\overline{X} = \frac{1}{N}\sum xini$$

$$\rightarrow \overline{X} = \frac{15\ 400}{140}$$

$\rightarrow \overline{X}$ = 1102.86 DA

## III.4-Les propriétés de la moyenne arithmétique

La moyenne arithmétique a les caractéristiques suivantes :

**4.1-La somme des écarts à la moyenne est nulle :** $\sum (x_i - \bar{x})n_i = 0$

**4.2-La somme des carrés des écarts à la moyenne arithmétique est minimale :** Ça se traduit par : $\sum n_i\,(x_i - a)^2$est minimale pour a=$\overline{X}$.

**4.3-Lorsrsque on ajoute ou l'on soustrait une quantité constante « a » aux valeurs de la variable$x_i$, la moyenne arithmétique de la série augmente ou diminue de la même quantité par la constante « a ».**

$x_i^{`} = x_i + a \ \rightarrow \overline{x_i^{`}} = \overline{x_i} + a$

$x_i^{`} = x_i - a \ \rightarrow \overline{x_i^{`}} = \overline{x_i} - a$

**4.4-**Si on mélange deux populations de tailles respectives $N_1$ $et$ $N_2$, de moyennes respectives $\overline{X}_1$ et $\overline{X}_2$, la moyenne arithmétique de l'ensemble, notée $\overline{X}_t$ est : $\overline{\mathrm{X}}_\mathrm{t} = \frac{\mathrm{N_1\overline{X}_1 + N_2\overline{X}_2}}{\mathrm{N_1+N_2}} \rightarrow \overline{\mathrm{X}}_\mathrm{t} = f_1\overline{\mathrm{X}}_1 + f_2\overline{\mathrm{X}}_2$

**4.5-La méthode de changement de variable**

De ces deux dernières propriétés découle la méthode de calcul de la moyenne arithmétique par changement de variable. Cette méthode permet de simplifier les calculs de $\bar{X}$.

On pose $x_i^{`} = \frac{x_i - x_0}{a_i}$ avec $x_0$ centre de la classe modale et $a_i$ amplitude de la classe modale. Ça donne $\overline{x_i^{`}} = \frac{\overline{x_i} - x_0}{a_i} \rightarrow \bar{X} = a\overline{X^{`}} + x_0$

**Exemple :** soit le tableau suivant

| Primes (DA) | $n_i$ | $x_i$ | $x_i^{`}$ | $x_i^{`}\mathrm{n_i}$ |
|---|---|---|---|---|
| 800-1000 | 26 | 900 | -2.5 | -65 |
| 1000-1100 | 33 | 1050 | -1 | -33 |
| 1100-1200 | 64 | 1150 | 0 | 0 |
| 1200-1300 | 07 | 1250 | 1 | 7 |
| 1300-1500 | 10 | 1400 | 2.5 | 25 |
| $\sum$ | 140 | / | / | -66 |

La classe modale : 1100-1200, donc a=100 et $x_0 = 1150$

$$\rightarrow \overline{X^{`}} = \frac{\sum x_i^{`} n_i}{N} - \frac{66}{140} = -0.47$$

$\rightarrow \bar{X} = a\overline{X^{`}} + x_0$=100(-0.47) + 1150 = 1102, 86 DA

### III.5-Généralisation de la moyenne : calcul d'autres types de moyennes

#### 5.1-La moyenne géométrique

#### 5.1.1-Rappels sur les taux de croissance et les coefficients multiplicateurs

Une quantité (ou valeur) passe de $V_0$ (valeur initiale ou de départ) à $V_t$ (valeur finale ou d'arrivée), on dit que sa variation absolue est $V_t - V_0$ et sa variation relative est $\frac{V_t - V_0}{V_0} = \frac{V_t}{V_0} - 1$. Cette variation s'appelle taux de croissance.

- Si $\frac{V_t}{V_0} - 1$ donne une valeur positive, donc il s'agit d'une hausse.
- $\frac{V_t}{V_0} - 1$ donne une valeur négative, donc il s'agit d'une baisse.

**Exemple 01 :** Un capital passe de 2000 DA à 3500 DA entre les dates 0 et t.

- Sa variation absolue est de 1500 DA ($V_t - V_0$ = 3500-2000).
- Sa variation relative est de 75% ( $\frac{V_t - V_0}{V_0} = \frac{3500-2000}{2000} \times 100$)

**Exemple 02 :** Un capital passe de 2000 DA à 1500 DA entre les dates 0 et t.

- Sa variation absolue est de - 500 DA ($V_t - V_0$ = 2000-1500).
- Sa variation relative est de -25% ( $\frac{V_t - V_0}{V_0} = \frac{2000-1500}{2000} \times 100$)

Le ratio $\frac{V_t}{V_0}$ s'appelle un coefficient multiplicateur. Taux de croissance = $\frac{V_t}{V_0} - 1$

Et $\frac{V_t}{V_0} = taux\ de\ croissance + 1$

Reprenons l'exemple précédent :

- $\frac{V_t}{V_0} = \frac{3500}{2000} = 1{,}5$: la valeur d'origine a été multipliée par 1,5
- $\frac{V_t}{V_0}\frac{1500}{2000} = 0{,}75$ : la valeur d'origine est multipliée par 0,75.

#### 5.1.2-Définition et application de la moyenne géométrique

C'est la moyenne applicable à des mesures de grandeurs dont la croissance est géométrique ou exponentielle.

La moyenne géométrique G de la série de valeurs : $x_1$, $x_2$,..., $x_k$ dont les modalités sont strictement toutes positives, est définie ainsi :

G= $\sqrt[N]{{x_1}^{n_1} \times {x_2}^{n_2} \times \ldots \times {x_k}^{n_k}} = \sqrt[N]{\pi\, {x_i}^{n_i}}$

**Exemple 01:** soit les nombres suivants: 1, 2, 4, 4, 5, 6, 6 →G=$\sqrt[7]{1 \times 2 \times 4^2 \times 5 \times 6^2} = \sqrt[7]{5760}$ =3,445

La moyenne géométrique est employée dans le calcul des taux de croissance moyens, ou des moyennes des coefficients multiplicateurs.

**Exemple 02 :** Supposons que pendant une décennie, les salaires ont cru de 50% et que pendant la décennie suivante, ils ont cru de 30%.

G=$\sqrt[20]{(1+0,5)^{10} \times (1+0.3)^{10}} = \sqrt[20]{1,5^{10} \times 1,3^{10}} = 1.39645$

Tm= 1,39645-1 = 0,39645 soit une augmentation de 39,645%

**Exemple 03 :** le prix d'un produit a cru de 15% au mois d'octobre, puis a baissé de 8% au mois de novembre et a continué de baissé en enregistrant le même taux au mois de décembre. Soit$\mathrm{T_m}$ ce taux moyen. $\mathrm{T_m} = G - 1$, avec G la moyenne géométrique des coefficients multiplicateurs.

G=$\sqrt[3]{(1+0,15) \times (1-0.08)^2} = \sqrt[3]{1,15 \times 0.92^2} = 0,991$

Tm= 0,991-1 = - 0,009 soit une diminution de de 0,9%

**5.2-La moyenne harmonique**

**5.2.2-Définition**

La moyenne harmonique d'une série statistique, notée H, est égale à l'inverse de la moyenne arithmétique des inverses des valeurs observés.

H= $\frac{\mathrm{N}}{\sum \frac{1}{\mathrm{x_i}}} = \frac{n_1+n_2+\cdots+n_k}{\frac{1}{\mathrm{x}_1}+\frac{1}{x_2}+\cdots+\frac{1}{n_k}}$ $(Moyenne\ harmonique\ simple)$ Ou H $= \frac{\mathrm{N}}{\sum \frac{\mathrm{n_i}}{\mathrm{x_i}}} = \frac{n_1+n_2+\cdots+n_k}{\frac{\mathrm{n}_1}{\mathrm{x}_1}+\frac{n_2}{x_2}+\cdots+\frac{x_k}{n_k}}$ (moyenne harmonique pondérée).

**Exemple 1 :** On achète des dollars une première fois pour 200 €au cours de 1,22 € le dollar, une seconde fois pour 100 € au cours de 0,98 € le dollar. Le cours moyen du dollar pour l'ensemble de ces deux opérations est H $= \frac{\mathrm{N}}{\sum \frac{\mathrm{n_i}}{\mathrm{x_i}}} = \frac{n_1+n_2}{\frac{\mathrm{n}_1}{\mathrm{x}_1}+\frac{n_2}{x_2}} = \frac{200+100}{\frac{200}{1,22}+\frac{100}{0,98}} = \frac{300}{163,94+102,04} = 1,128€$

**Exemple 02 :** Un atelier de reproduction a photocopié 1000 documents à la vitesse de 800 documents à l'heure et un autre atelier a photocopié 2000 documents à la vitesse de 700 documents à l'heure. Quelle est la vitesse moyenne pour photocopier les 3000 documents ?

Vitesse moyenne =$\frac{Nombre\ de\ documents\ total}{nombre\ heures} = \frac{Q}{T} = \frac{Q_1+Q_2}{T_1+T_2} = \frac{Q_1+Q_2}{\frac{Q_1}{V_1}+\frac{Q_2}{V_2}} = \frac{1000+2000}{\frac{1000}{800}+\frac{2000}{700}}$=730,46 documents à l'heure. On retrouve donc la formule de la moyenne harmonique.

**Conclusion :**

Le chapitre précèdent nous a permis de comprendre comment faire la synthèse des données statistiques et comment les présenter de manière lisible et compréhensible par un large éventail de lecteurs. Cependant, la synthèse et la présentation des données statistiques ne s'arrête pas uniquement à ce qui est développé dans le précèdent chapitre.

La présentation graphique a permis une première synthèse des informations contenues dans les tableaux statistiques mais pour approfondir l'analyse et rendre l'interprétation plus objective des informations recueillies, on a établi des mesures typiques qui résument la distribution statistique. Ainsi, l'ensemble des observations est réduit à une seule valeur numérique. Ces mesures sont appelées paramètres de tendance centrale ou paramètres de position.