Université A- mira de Bejaia
Faculté des sciences économiques, commerciales et des sciences de gestion
Département SEGC(LMD)
Module stat I
Enseignante : Dr. BERRAH

Chapitre 3 : Les paramètres de tendance centrale

Les paramètres de tendance centrale, également appelés "mesures de tendance centrale", sont des valeurs uniques qui représentent le mieux un ensemble de données. Ils sont exprimés dans la même unité que les valeurs des variables. L'appellation "tendance centrale" vient du fait que ces paramètres donnent une idée de ce qui se passe au centre d'une distribution, d'un ensemble de données. On distingue trois mesures de tendance centrale principales : le mode, la médiane et la moyenne arithmétique. Chacun de ces paramètres offre une perspective différente sur la distribution des données, ce qui les rend complémentaires dans la description et l'analyse d'un ensemble de données.

1. Le Mode

Le mode, noté $\boldsymbol{Mo}$ d'une série statistique est la valeur du caractère la plus fréquente ou dominante dans l'échantillon. Autrement dit, c'est la valeur de x_i qui a la fréquence (absolue ou relative) la plus grande.

Lorsqu'une distribution présente plusieurs modes, on la qualifie de "multimodale" (bimodale pour deux modes, trimodale pour trois modes, etc.).

La méthode de détermination du mode d'une série statistique varie en fonction de la nature de la variable étudiée :

1.1. Cas d'une variable discrète : Lorsque la variable statistique est discrète, le mode est la valeur de la variable x_i qui possède l'effectif le plus élevé. Il est donc directement identifiable dans une série statistique.

Graphiquement, le mode correspond à la valeur de x_i représentée par le bâton le plus haut dans un diagramme en bâtons. Sa valeur est déterminée par l'axe des abscisses.

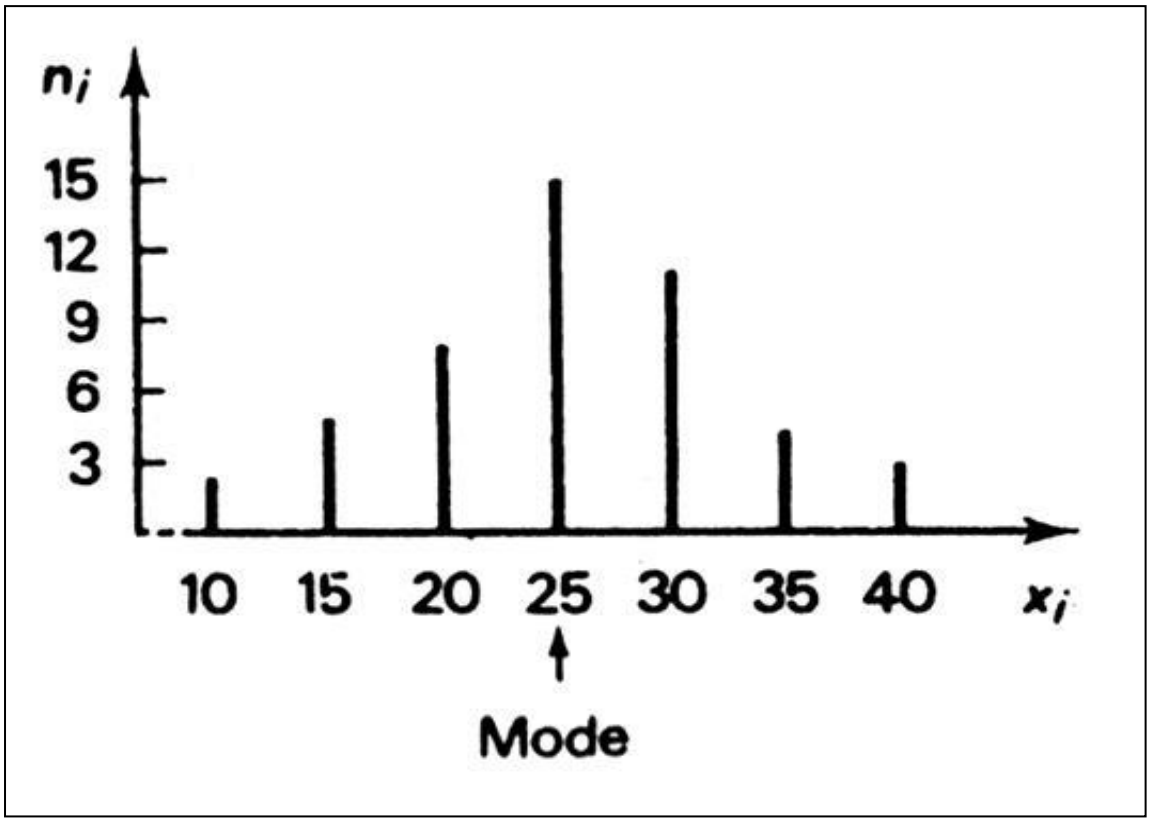
Exemple1 : Soit le tableau suivant.

x_i	10	15	20	**25**	30	35	40
n_i	2	3	8	17	13	5	3

Le mode : $\mathbf{M_O = 25,}$ car il correspond à la valeur maximale des effectifs (maxi n_i =17).

Dans ce cas le mode signifie que la majorité des individus ont le caractère xi= 25.

Le mode graphiquement correspond au bâton le plus élevé, voir le graphe suivant.



Le mode graphiquement correspond à la valeur 25 sur l'axe des abscisses.

1.2. Cas d'une variable continue : Dans le cas d'une variable continue, les valeurs prises par la variable sont regroupées en classes. On parle alors de classe modale, qui est la classe ayant l'effectif le plus élevé. Le mode est ensuite calculé à l'aide de la formule suivante :

$$Mo = X_{\min} + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times a_i$$

Où :

Mo : La valeur du mode.

$X_{\min}$: La borne inférieure de la classe modale.

Δ_1 : La différence entre l'effectif de la classe modale et l'effectif de la classe précédente.

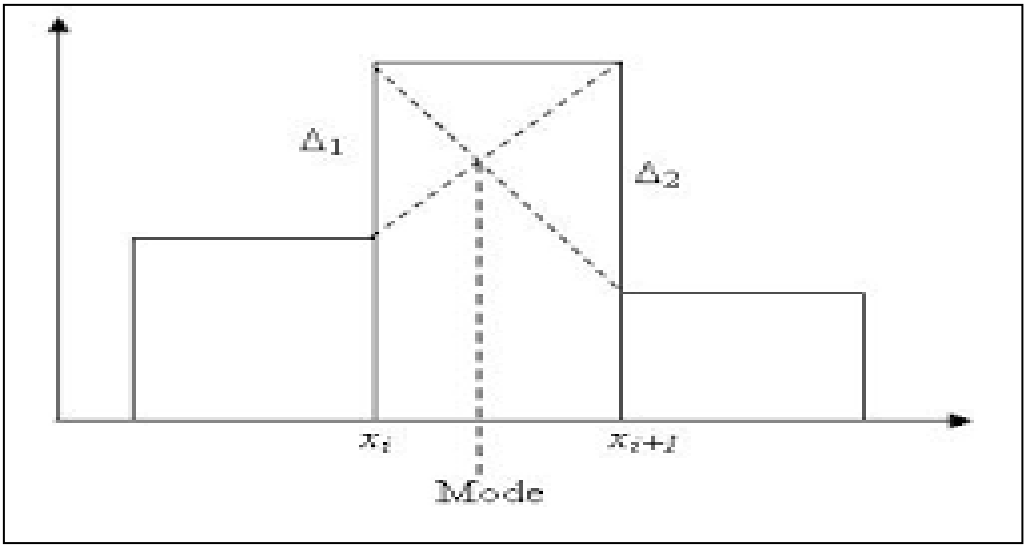
Δ_2 : La différence entre l'effectif de la classe modale et l'effectif de la classe suivante.

a_i : L'amplitude de la classe modale.

Graphiquement, la valeur du mode est déterminée à partir d'un histogramme. Pour ce faire :

1. **Identifier la classe modale :** Repérer le rectangle le plus haut, qui correspond à la classe modale (celle ayant l'effectif le plus grand).

2. **Appliquer la méthode des diagonales :** Tracer les diagonales des rectangles situés à gauche et à droite de la classe modale, comme illustré dans le schéma ci-après.
3. Le point d'intersection des diagonales indique la valeur du mode sur l'axe des abscisses.



Remarque importante :

Lorsque les classes ont des amplitudes inégales, il est nécessaire de corriger les effectifs avant de déterminer le mode. On calcule les **effectifs corrigés** (n_ic) en divisant l'effectif de chaque classe (n_i) par son amplitude (a_i) : $nic = \frac{ni}{ai}$

On utilise ensuite ces effectifs corrigés pour calculer le mode, en appliquant les mêmes procédures que précédemment.

Pour la représentation graphique (histogramme), on remplace les effectifs (n_i) par les effectifs corrigés (n_ic) sur l'axe des ordonnées. On poursuit ensuite les mêmes procédures précédentes pour déterminer le mode graphiquement, en utilisant la méthode des diagonales.

- **Exemple (cas des amplitudes égales) :** Soit le tableau suivant. Calculer le mode et le déterminer graphiquement.

x_i	**ni**
[100-110[	8
[110-120[	22
[120-130[	38
[130-140[	12
[140-150[	6
Total	**86**

Classe modale

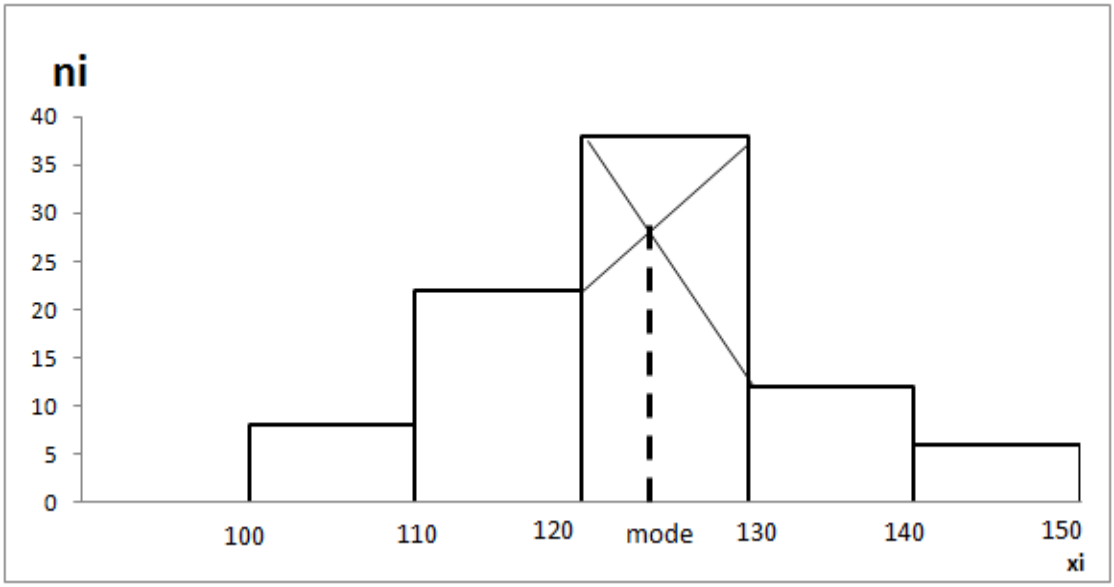
Le plus grand effectif (*ni*)

Solution

- On détermine tout d'abord la classe modale, à travers le tableau on remarque que l'effectif le plus grand est 38, donc la classe modale est [120-130[.
- Maintenant on applique la formule suivante pour calculer le mode :

$$Mo = X_{\min} + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times a_i = 120 + \frac{38-22}{(38-22)+(38-12)} \times 10 = 123{,}81$$

Graphiquement, le mode est déterminé sur l'histogramme ci-après :



Exemple 2 (cas des amplitudes inégales) : soit le tableau suivant.

Classes	n_i	a_i	n_i corrigés (= ni/ai)
[0 - 1[	35	1	**35**
[1 - 2[	30	1	30
[2 - 3[	25	1	25
[3 - 5[	54	2	27
[5 - 10[	100	5	20
[10-20[	80	10	8
[20-22[	6	2	3
total	330	-	-

Classe modale

Le plus grand effectif corrigé (nic)

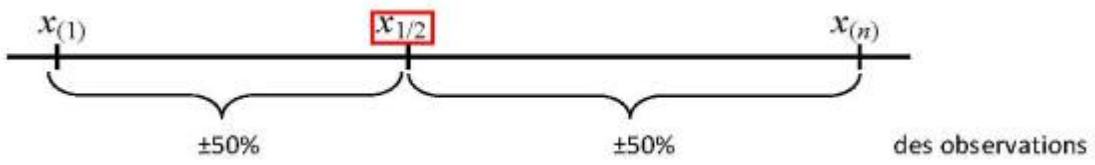
Solution : Comme les amplitudes (a_i) sont inégales donc la classe modale correspond à l'effectif corrigé (nic) le plus grand, qui est égal à 35. D'où la classe modale est : [0 - 1[

Donc on applique la formule pour calculer le mode :

$$Mo = X_{\min} + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times a_i \Rightarrow Mo = 0 + \frac{(35-0)}{(35-0)+(35-30)} \times 1 = 0{,}875 \text{ HA}$$

2. La Médiane (notée Mé)

La médiane (Me) d'une série statistique est la valeur du caractère (x_i) qui partage la série statistique en deux groupes de taille égale, à condition que les valeurs de la variable statistique aient été préalablement classées par ordre croissant.



Dans une distribution statistique, le calcul de la médiane s'effectue à partir de la colonne des effectifs cumulés croissants (ou des fréquences relatives cumulées croissantes). La méthode à suivre dépend de la nature de la variable statistique étudiée :

2.1. Cas d'une variable discrète : Dans ce cas il faut voir le nombre total des effectifs (N), deux cas se présentent[1] :

- **Si le nombre d'observations (N) est impair :**

  Dans ce cas, la médiane : $M_e = x_{\frac{n+1}{2}}$

- **Si le nombre d'observations (N) est pair :**

  Dans ce cas, la médiane : $M_e = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$

Graphiquement la médiane (Mé) est déterminée sur le diagramme cumulatif croissant des effectifs (ou des fréquences relatives), c'est la valeur de xi située sur l'axe des abscisses à laquelle est associée à la valeur N/2 (ou 0,5) sur l'axe des ordonnées.

Exemple

Le tableau ci-dessous présente la distribution observée des âges des répondants à une enquête :

[1] Cette application correspond au sériés pondérées (les plus utilisées dans la pratique)

Age du répondant (xi)	Effectif (ni)	$n_i \uparrow$
17	18	18
18	85	103
19	99	202
20	52	254
21	31	285
22	15	300
Total	300	-

- Calculer la médiane et la déterminer graphiquement.

Solution : N=300 est pair, donc la médiane se calcule par la formule suivante :

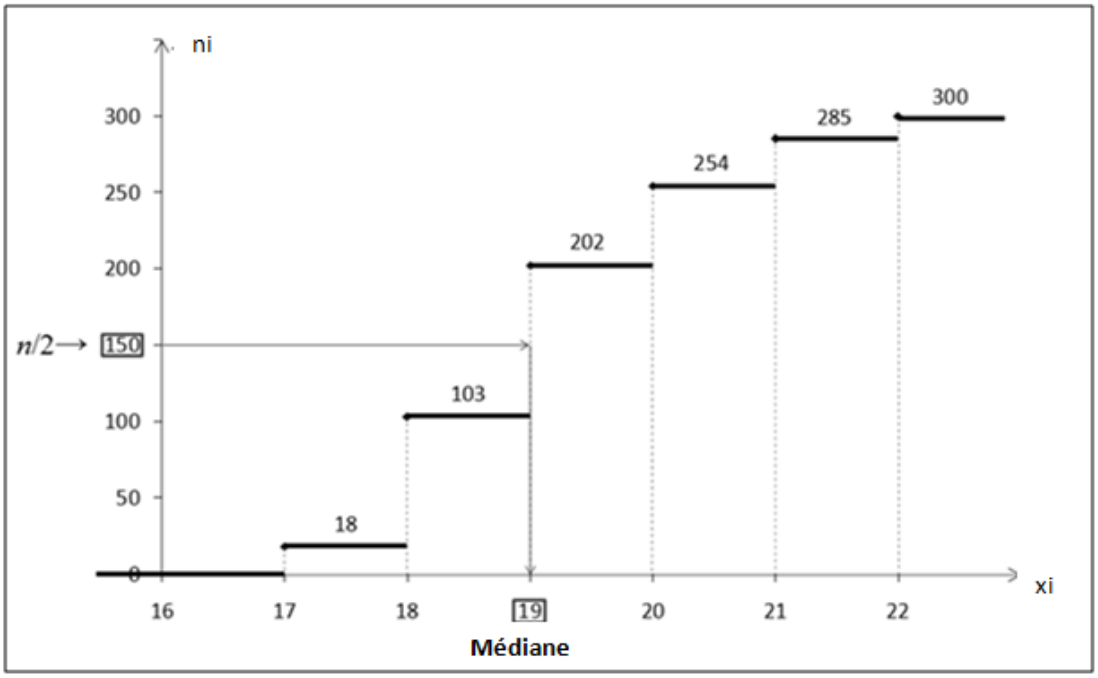
$$M_e = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} \quad \Rightarrow \quad M_e = \frac{x_{\frac{300}{2}} + x_{\frac{300}{2}+1}}{2} = \frac{x_{150} + x_{151}}{2}$$

On remarque que les valeurs x_{150} et x_{151} sont situées dans la colonne des effectifs cumulée croissantes ($n_i \uparrow$) entre les valeurs 103 et 202, cette dernière valeur correspond à la valeur 19 sur la première colonne (xi).

Donc la médiane : $M_e = \frac{19+19}{2} = 19.$

Cette valeur signifie que 50% des répondants sont âgés de moins de 19 ans.

Graphiquement la médiane se détermine sur le diagramme en escalier, voir le schéma suivant.



2.2. Cas d'une variable continue

Dans ce cas, pour calculer la médiane, il faut déterminer au préalable la classe médiane, c'est la classe qui correspond à la valeur N/2 sur la colonne $N\uparrow$ ou bien à la valeur 0.5 sur la colonne $F\uparrow$. Puis on applique la formule suivante pour calculer la valeur de la médiane :

$$Mé = X_{\min} + \frac{N/2 - n_{i\,précédent}\uparrow}{n_{i\,Me}} \times a_i$$ Sachant :

$Mé$: La valeur de la médiane.

$X_{\min}$: La borne inférieure de la classe médiane.

$n_{i\,précédent}\uparrow$: L'effectif cumulé croissant de la classe située avant la classe médiane.

$n_{i\,Me}$: L'effectif de la classe médiane.

a_i : L'amplitude de la classe médiane.

Exemple : Soit le tableau suivant, calculer la médiane.

xi	n_i	$n_i\uparrow$	$n_i\downarrow$
[1000 - 1500[	6	6	65
[1500 - 2000[	12	18	59
[2000 - 2500[	25	43	47
[2500 - 3000[	17	60	22
[3000- 3500[	5	65	5
Total	65	-	-

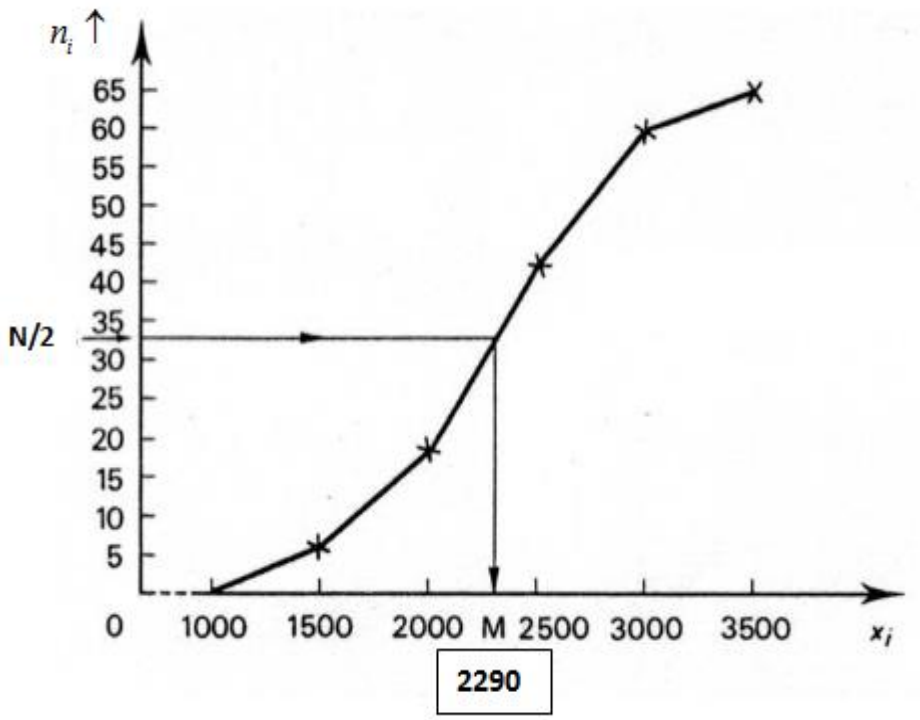
Solution

Dans ce cas, on a N/2 = 65/2= 32.5.

Cette valeur est située dans la colonne des effectifs cumulés croissants entre les valeurs 18 et 43, cette dernière correspond à la classe [2000-2500[.

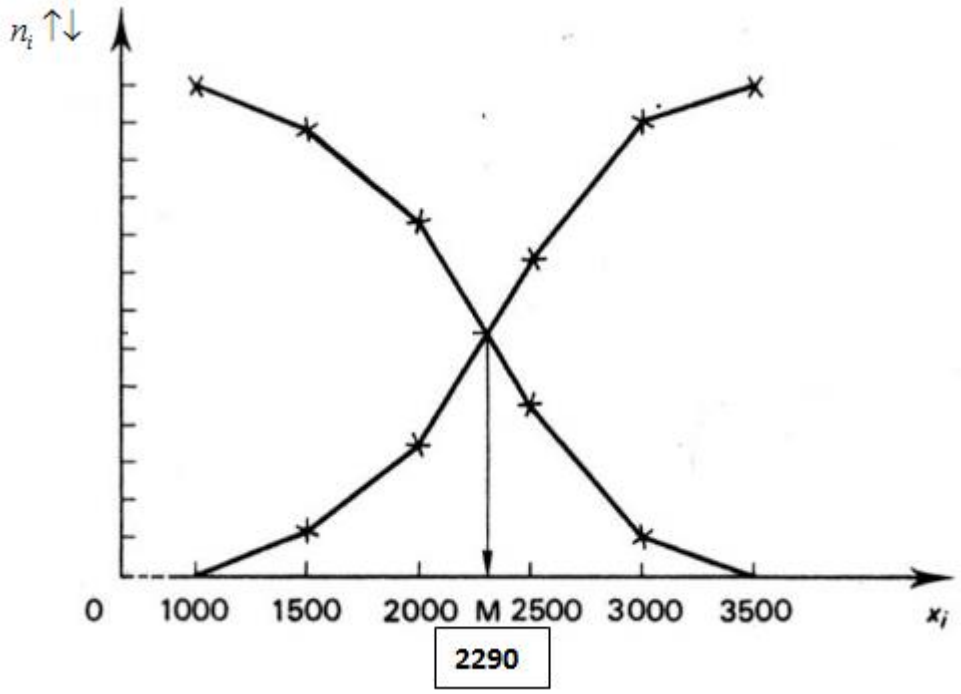
Donc $Mé \in [2000; 2500[$

$$\Rightarrow Mé = X_{\min} + \frac{N/2 - N_{i\,précédent}^{\uparrow}}{ni_{Mé}} \times ai = 2000 + \frac{32{,}5 - 18}{25} \times 500 = 2290$$



Graphiquement la valeur de la médiane : Mé = 2290, qui correspond au rang N/2 sur l'axe des effectifs cumulés croissants.

On peut aussi obtenir la valeur de la médiane en traçant sur le même graphique les deux courbes des effectifs cumulés croissants et décroissants. Le point d'intersection donne la valeur de la médiane sur l'axe des abscisses. Voir le schéma suivant :



Remarque : Dans le cas où nous utilisons les fréquences relatives cumulées croissantes, le calcul de la médiane se fait par la formule suivante : $Mé = X_{\min} + \frac{0,5 - F_i^{\uparrow}{}_{précédent}}{fi_{Mé}} \times ai$

Exemple : Soit le tableau suivant, calculer la médiane.

xi	n_i	f_i	$f_i \uparrow$
[3 - 4[	12	0,024	0.024
[4 - 6[	30	0,061	0.085
[6 - 8[	120	0,246	0.331
[8 - 9[	210	0,431	0,762
[9- 10[	90	0,184	0,946
[10- 12[	25	0,051	≈ 1
Total	487	≈ 1	-

Solution : Pour nos calculs, nous avons retenu les fréquences relatives cumulées croissantes.

- Mé $(\alpha = 0,5)$, cette valeur est située dans la colonne des fréquences relatives cumulées croissantes ($f_i \uparrow$) entre les valeurs 0,331 et 0,762, cette dernière correspond à la classe $[8;9[$. Donc $Mé \in [8;9[\Rightarrow Mé = X_{\min} + \frac{\alpha - F_i^{\uparrow}{}_{précédent}}{fi_{Mé}} \times ai = 8 + \frac{0,5 - 0,331}{0,431} \times 1 = 8,39$

3. La Moyenne arithmétique

La moyenne arithmétique d'une distribution statistique pour une variable X, est égale au rapport de la somme des valeurs prises par cette variable sur le nombre d'observations. On la note généralement $\overline{X}$ et se lit « X barre »

Elle peut être simple ou pondérée :

- **La moyenne arithmétique simple** : on dit la moyenne arithmétique est simple lorsque les valeurs prises par la variable X n'apparaissent qu'une seule fois. Dans ce cas chaque effectif ni est égal à 1.

$$\overline{X} = \frac{x1 + x2 + ... + xk}{n1 + n2 + ... + nk} \quad \Rightarrow \quad \overline{X} = \frac{\sum_{i=1}^{k} X_i}{N}$$

Exemple : 8 étudiants ont obtenu les notes suivantes : 3, 5, 7, 9, 10, 11, 12, 18.

La note moyenne est : $\overline{X} = \frac{\sum_{i=1}^{n} x_i}{N} = \frac{3+5+7+9+10+11+12+18}{8} = 9.37$

- **La moyenne arithmétique pondérée :** on dit la moyenne arithmétique est pondérée lorsque les valeurs prises par la variable X peuvent correspondre à plusieurs valeurs d'effectifs ni.

$$\overline{X} = \frac{n1 * x1 + n2 * x2 + ... + nk * xk}{n1 + n2 + ... + nk} \quad \Rightarrow \quad \overline{X} = \frac{\sum_{i=1}^{k} n_i x_i}{N}$$

Ou, en travaillant avec des fréquences relatives au lieu des effectifs :

$$\overline{X} = \sum_{i=1}^{n} f_i x_i \ \ avec \ \ f_i = \frac{n_i}{N} \quad et \quad \sum_{i}^{n} f_i = 1$$

Exemple : 8 étudiants ont obtenu les notes suivantes : 3, 3, 3, 5, 9, 9, 11, 11.

La note moyenne est :

$$\overline{X} = \frac{\sum_{i=1}^{n} n_i x_i}{N} = \frac{n1 * x1 + n2 * x2 + ... + nk * xk}{n1 + n2 + ... + nk} = \frac{3*3 + 1*5 + 2*9 + 2*11}{8} = 6.75$$

Comment calculer la moyenne arithmétique ?

3.1. Cas d'une variable discrète : Dans le cas d'une variable discrète, le calcul de la moyenne arithmétique se fait directement à partir les modalités de la variable étudiée. On ajoute au tableau statistique une nouvelle colonne contenant les produits des valeurs de la variable et celles des effectifs ($n_i \cdot x_i$). Puis on utilise la formule suivante pour calculer la moyenne arithmétique : $\overline{X} = \frac{\sum_{i=1}^{n} n_i \cdot x_i}{\sum n_i}$

Exemple : Soit le tableau suivant.

X_i	n_i	$n_i \cdot x_i$
18	6	108
19	10	190
20	4	80
21	2	42
22	2	44
Total	24	464

$$\overline{X} = \frac{\sum_{i=1}^{7} n_i \cdot x_i}{N} = \frac{464}{24} = 19{,}33$$

3.2. Cas d'une variable continue : Dans ce cas, les modalités de la variable sont alors des classes et par conséquent on utilise dans les calculs de la moyenne les centres de classes. Dans ce cas on ajoute au tableau statistique deux nouvelles colonnes, l'une relative aux centres de classe c_i et l'autre colonne contenant les produits des valeurs de la variable et celles des effectifs ($n_i \times c_i$). Puis on utilise la formule suivante pour calculer la moyenne arithmétique :

$$\overline{X} = \frac{\sum_i n_i \times c_i}{N}$$

Exemple : Soit la distribution d'un échantillon de 24 étudiants selon leur taille.

X	n_i	c_i (centre de classe)	$n_i \times c_i$
[1.5 ; 1.6 [	6	1.55	9.3
[1.6 ; 1.7 [	7	1.65	11.55
[1.7 ; 1.8 [	8	1.75	14
[1.8 ; 1.9 [	2	1.85	3.7
[1.9 ; 2 [	1	1.95	1.95
Total	24	-	40.5

D'où la moyenne arithmétique : $\overline{X} = \frac{\sum_i n_i \times c_i}{N} = \frac{40,5}{24} \approx 1,687$

Propriétés de la moyenne arithmétique :

- Si on définit une nouvelle variable Z telle que : $Z_i = a\,\chi_i + b$, sachant a et b sont des constantes réelles, alors : $\overline{Z} = a\,\overline{X} + b$

  Cela signifie que si on multiplie toutes les valeurs de la variable par une constante a, la moyenne sera elle aussi multipliée par a, et, si on ajoute à toutes les valeurs de la variable une constante b, la moyenne augmentera elle aussi d'une valeur égale à b.

La même chose lorsque : $Z_i = a\,\chi_i - b \quad alors \quad \overline{Z} = a\,\overline{X} - b$

4. Généralisation de la notion de moyenne : les autres types de moyennes

Outre la moyenne arithmétique, il existe d'autres types de moyennes, notamment la moyenne géométrique, la moyenne harmonique et la moyenne quadratique, qui reposent sur des principes différents.

4.1. La moyenne géométrique (notée G) : La moyenne géométrique d'une distribution statistique pour une variable X est égale à la racine n-ième du produit des n valeurs positives de cette variable. . On la note généralement : G_X. Elle est utilisée principalement utilisée dans le calcul de taux de croissance moyens. Sa formule de calcul est donnée dans les deux cas suivants :

- **Cas de la moyenne géométrique simple :**

$$G = \sqrt[N]{X_1 \times X_2 \times X_3 \times ... \times X_k} = \sqrt[N]{\prod_{i=1}^{k} X_i} = \left(\prod_{i=1}^{k} X_i \right)^{\frac{1}{N}} .$$

- **Cas de la moyenne géométrique pondérée :** Dans ce cas, la formule de la moyenne géométrique devient :

$$G = \sqrt[N]{X_1^{n_1} \times X_2^{n_2} \times .... \times X_k^{n_k}}$$

Remarque : le taux de croissance moyen Tx : est égal : $Tx = (G-1) \times 100$

Exemple :

Une société a vu son bénéfice augmenter ces trois dernières années de 10 % la première année, 21% la deuxième année et 3% la troisième année.

- Calculer la moyenne géométrique.
- Calculer le taux de croissance moyen de bénéfice.

Solution

- La moyenne géométrique : $G = \sqrt[N]{X_1^{n_1} \times X_2^{n_2} \times .... \times X_k^{n_k}}$

On a : $X_i = 1 + t_i$, sachant t_i est le taux de croissance. Donc $G = \sqrt[3]{{}_1^{n_1} \times X_2^{n_2} \times .... \times X_k^{n_k}}$

$$= \sqrt[3]{(1{,}1)^1 \times (1{,}21)^1 \times (1{,}02)^1} = 1{,}1071$$

- Le taux de croissance moyen du bénéfice :

Au cours de cette période, le taux moyen de croissance Tx est : $Tx = (G-1) \times 100 = (1{,}1071 - 1) \times 100 = 10{,}72\%$

4-2- La moyenne harmonique (notée H) : La moyenne harmonique d'une distribution statistique pour une variable X est égale à l'inverse de la moyenne arithmétique des inverses des valeurs de cette variable. On la note généralement H_x. Elle est principalement utilisée dans le calcul de moyennes de rapports, notamment de vitesses moyennes et d'autres indices statistiques (km/h, DA/minute, quintaux/hectare, etc.). Sa formule de calcul est donnée dans les deux cas suivants :

- **Cas de la moyenne harmonique simple :** $H = \dfrac{N}{\sum_i \dfrac{1}{X_i}} = \dfrac{N}{\dfrac{1}{X_1} + \dfrac{1}{X_2} + ... + \dfrac{1}{X_n}}$

- **Cas de la moyenne harmonique pondérée :** $H = \dfrac{\sum n_i}{\sum_i \dfrac{n_i}{X_i}} = \dfrac{N}{\dfrac{n_1}{X_1} + \dfrac{n_2}{X_2} + \cdots + \dfrac{n_k}{X_k}}$

Exemple : Une voiture roule 1 heure à la vitesse de 80 km/h et parcours un tronçon de 60 km, de la suite de son trajet, à une vitesse de 120 km/h. Le trajet de retour est accompli avec une vitesse de 100km/h.

Calculer la vitesse moyenne de cette voiture.

Solution : La vitesse moyenne de cette voiture est une moyenne harmonique :

1- Le temps =1h et la distance (n1) = 80 km $\Rightarrow$ la vitesse (x1) = 80km/h.

2- La distance (n2) = 60 km et la vitesse (x2) = 120 km/h.

3- La distance (n3) = 80 +60 = 140 KM et la vitesse (x3) = 100km/h.

$$H = \frac{N}{\dfrac{n_1}{x_1} + \dfrac{n_2}{x_2} + \dfrac{n_3}{x_3}} \quad \text{Donc} \quad H = \frac{80 + 60 + 140}{\dfrac{80}{80} + \dfrac{60}{120} + \dfrac{140}{100}} = 96{,}55 km/h$$

La vitesse moyenne est 96,55 km/h.

4-3- La moyenne quadratique (notée Q) : La moyenne quadratique d'une distribution statistique pour une variable X est égale à la racine carrée de la moyenne arithmétique des carrés des valeurs (x_i). On la note généralement Q_x. Bien que son utilisation soit moins fréquente que d'autres types de moyennes, elle trouve son application principale dans les mesures de dispersion, notamment pour le calcul de l'écart-type. Sa formule de calcul est donnée dans les deux cas suivants :

- **Cas de la moyenne quadratique simple :** $Q = \sqrt{\dfrac{1}{N}\sum_{i=1}^{n} X_i^2}$

- **Cas de la moyenne quadratique pondérée :** $Q = \sqrt{\dfrac{1}{N}\sum_{i=1}^{k} n_i \cdot X_i^2}$

Remarque :

- Dans le cas d'une variable continue, les valeurs sont regroupées en classes. Pour calculer les différents types de moyennes (arithmétique, géométrique, harmonique,

quadratique), on utilise le **centre de classe** comme valeur représentative de chaque classe.

- $H \prec G \prec \overline{X} \prec Q$: cette relation est pratiquement toujours vérifiée.

Nous pouvons vérifier cette relation en traitant l'exemple suivant

Exemple : soit le tableau suivant

x_i	1	2	3	4
n_i	5	15	10	10

- Calculer les valeurs des moyennes arithmétique, harmonique, géométrique et quadratique.

Solution :

- La moyenne arithmétique : $\overline{X} = \frac{1\times 5 + 2\times 15 + 3\times 10 + 4\times 10}{40} = 2{,}62$
- La moyenne géométrique : $G = \sqrt[40]{1^5 \times 2^{15} \times 3^{10} \times 4^{10}} = 2{,}41$
- La moyenne harmonique : $H = \frac{40}{\frac{5}{1} + \frac{15}{2} + \frac{10}{3} + \frac{10}{4}} = 2{,}18$
- La moyenne quadratique : $Q = \sqrt{\frac{1^2 \times 5 + 2^2 \times 15 + 3^2 \times 10 + 4^2 \times 10}{40}} = 2{,}80$

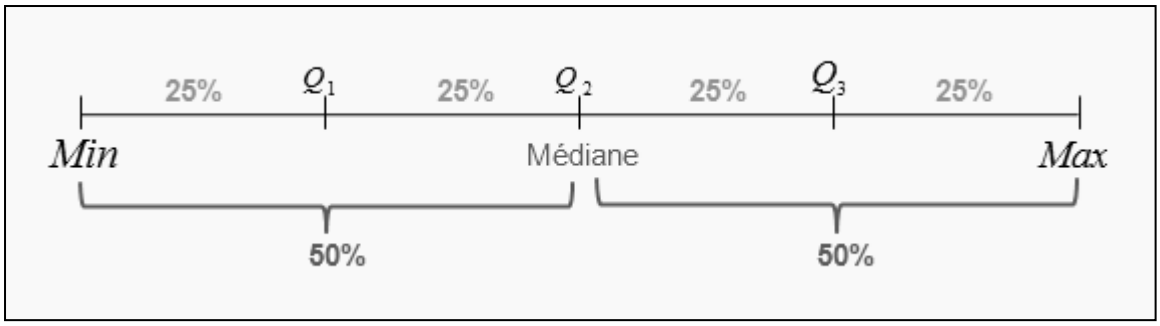
On remarque : $H \prec G \prec \overline{X} \prec Q$

5- Généralisation de la notion de la médiane : Les quantiles

Les quantiles constituent une généralisation de la notion de médiane, qui en représente un cas particulier. Ce sont des valeurs du caractère (x_i) qui partagent la série statistique en sous-groupes de même taille. Ils sont déterminés de manière similaire à la médiane, à partir des effectifs cumulés croissants (ou des fréquences relatives cumulées croissantes).

On distingue trois types de quantiles : les quartiles, les déciles et les centiles : Ils divisent la série statistique en cent groupes égaux.

5.1. Les Quartiles : ce sont les valeurs de xi qui partagent la série statistique en quatre sous ensembles égaux. Ils sont de trois nombre : Q1, Q2 et Q3 et ils correspondent respectivement sur la série statistique aux valeurs : 1/4 (ou 25%), 2/4(ou 50%) et 3/4 (ou 75%), voir la figure suivante.



- le 1er quartile Q1$\left(\alpha = \frac{1}{4}\right)$: 25% des observations lui soient inférieures.
- le 2e quartile Q2 $\left(\alpha = \frac{2}{4}\right)$: est la médiane, telle que 50% des observations lui soient inférieures.
- le 3e quartile Q3 $\left(\alpha = \frac{3}{4}\right)$: telle que 75% des observations lui soient sont inférieures.

5-2-Les Déciles : Les déciles, sont au nombre de neuf (D1, D2,…., D9), qui divisent les effectifs de la série en 10 parties égales.

- Le 1er décile D1$\left(\alpha = \frac{1}{10}\right)$: est la valeur de la variable telle que 10% des observations lui soient inférieures.
- ……..
- Le 9e décile D9$\left(\alpha = \frac{9}{10}\right)$: est la valeur de la variable telle que 90% des observations lui soient inférieures.

Le 5e décile est égal la médiane. $D_5\left(\alpha = \frac{5}{10}\right)$

5-3-Les centiles : Les centiles, sont au nombre de 99 (C1, C2,…., D99), qui divisent les effectifs de la série en 100 parties égales.

- Le 1er centile C1$\left(\alpha = \frac{1}{100}\right)$: est la valeur de la variable telle que 1% des observations lui soient inférieures.
- ……..
- Le 99e centile C99$\left(\alpha = \frac{99}{100}\right)$: est la valeur de la variable telle que 99% des observations lui soient inférieures.

Le 50e centile $C_{50}\left(\alpha = \frac{50}{100}\right)$ correspond à la médiane. Donc : $Mé = Q_2 = D_5 = C_{50}$

Comment calculer les quantiles ?

Les quantiles, qu'ils soient dans le cas continu ou discret, se calculent sur le même principe que la médiane à partir des effectifs ou des fréquences relatives cumulées ; on détermine l'effectif ou la fréquence relative cumulée correspondant au quantile envisagé de paramètre α . La suite des calculs est identique à celle de la médiane.

Exemples d'illustration sur le calcul des quantiles

Exemple1 (Cas d'une variable discrète) : Soit la distribution statistique suivante qui indique la répartition du nombre d'enfants par ménage (voir le tableau ci-après).

Calculer les quantiles suivants : Q1, Q2, D5, D8, C50 et C95.

Nombre d'enfants (xi)	Nombre de ménage (ni)	$n_i \uparrow$
0	20	20
1	65	85
2	70	155
3	30	185
4	10	195
5	5	200
Total	200	-

Solution :

- Q1$\left(\alpha = \frac{1}{4},\ N = 200\right)$: $Q_1 = \frac{x_{n\alpha} + x_{n\alpha+1}}{2} = \frac{x_{50} + x_{51}}{2} = \frac{1+1}{2} = 1 \text{ enfants}$

Cette valeur indique que 25% de ménages ont au plus d'un enfant et 75% ont plus d'un enfant.

- Q2$\left(\alpha = \frac{1}{2},\ N = 200\right)$: $Q_2 = \frac{x_{n\alpha} + x_{n\alpha+1}}{2} = \frac{x_{100} + x_{101}}{2} = \frac{2+2}{2} = 2$ enfants

Cette valeur indique que 50% de ménages ont au plus de deux enfants et l'autre 50% de ménages ont plus de deux enfants.

- D5$\left(\alpha = \frac{5}{10},\ N = 200\right)$: $D_5 = \frac{x_{n\alpha} + x_{n\alpha+1}}{2} = \frac{x_{100} + x_{101}}{2} = \frac{2+2}{2} = 2$ enfants

- D8$\left(\alpha = \frac{8}{10},\ N = 200\right)$: $D_8 = \frac{x_{n\alpha} + x_{n\alpha+1}}{2} = \frac{x_{160} + x_{161}}{2} = \frac{3+3}{2} = 3$ enfants

Cette valeur indique que 80% de ménages ont au plus de 3 enfants et 20% de ménages ont plus de 3 enfants.

- C50$\left(\alpha = \frac{50}{100},\ N = 200\right)$: $C_{50} = \frac{x_{n\alpha} + x_{n\alpha+1}}{2} = \frac{x_{100} + x_{101}}{2} = \frac{2+2}{2} = 2$ enfants

- C95$\left(\alpha = \frac{95}{100},\ N = 200\right)$: $C_{50} = \frac{x_{n\alpha} + x_{n\alpha+1}}{2} = \frac{x_{190} + x_{191}}{2} = \frac{4+4}{2} = 4$ enfants

Cette valeur indique que 95% de ménages ont au plus de 4 enfants et 5% de ménages ont plus de 4 enfants.

Exemple 2 (Cas d'une variable continue) : Soit le tableau suivant, calculer Q1, D4 et C85

xi	n_i	f_i	$f_i \uparrow$
[3 - 4[	12	0,024	0.024
[4 - 6[	30	0,061	0.085
[6 - 8[	120	0,246	0.331
[8 - 9[	210	0,431	0,762
[9- 10[	90	0,184	0,946
[10- 12[	25	0,051	≈1
Total	487	≈1	-

Solution : Pour nos calculs, nous avons retenu les fréquences relatives cumulées croissantes.

- Q1$(\alpha = 0{,}25)$, cette valeur est située dans la colonne des fréquences relatives cumulées croissantes ($f_i \uparrow$) entre les valeurs 0,085 et 0,331, cette dernière correspond à la classe

$[6;8[$. Donc $Q1 \in [6;8[\Rightarrow Q_1 = X_{\min} + \frac{\alpha - F_i^{\uparrow}{}_{précédent}}{fi_{Q1}} \times ai = 6 + \frac{0{,}25 - 0{,}085}{0{,}246} \times 2 = 7{,}34$

- D4$(\alpha = 0{,}4:)$: cette valeur est située dans la colonne des fréquences relatives cumulées croissantes ($f_i \uparrow$) entre les valeurs 0,331 et 0, 762, cette dernière correspond à la classe

$[8;9[$. Donc $D4 \in [8\ ;\ 9[\Rightarrow D4 = X_{\min} + \frac{\alpha - F_i^{\uparrow}{}_{précédent}}{fi_{D4}} \times ai = 8 + \frac{0{,}4 - 0{,}331}{0{,}431} \times 1 = 8{,}16$

- C85$(\alpha = 0{,}85,)$: cette valeur est située dans la colonne des fréquences relatives cumulées croissantes ($f_i \uparrow$) entre les valeurs 0, 762 et 0,946, cette dernière correspond à la classe

$[9;10[$.Donc $C85 \in [9;10[\Rightarrow C85 = X_{\min} + \frac{\alpha - F_i^{\uparrow}{}_{précédent}}{fi_{C85}} \times ai = 9 + \frac{0{,}85 - 0{,}762}{0{,}184} \times 1 = 9{,}48$

Exercices

Exercice 1 : Le tableau suivant indique la distribution de 100 logements selon le nombre de pièces dans un quartier.

Nombre de pièce (xi)	1	2	3	4	5	6
Nombre de logements (ni)	5	10	20	30	25	10

1- Quelle est la population statistique étudiée ?
2- Quel est le caractère étudié de cette distribution ? préciser sa nature.
3- Donner le mode de cette série et déterminer le graphiquement. Quelle est sa signification ?
4- Tracer le diagramme cumulatif des fréquences relatives.
5- Calculer la médiane et déterminer sa valeur graphiquement.
6- Calculer les quantiles ; Q1, Q2, Q3, D5, C50 et C75.
7- Calculer la moyenne arithmétique.

Exercice 2 : On considère une population de 1000 individus répartis en fonction de leur âge :

Age	[0 - 10[	[10 - 15[	[15 - 20[	[20- 30[	[30 - 40[	[40 - 60[	[60 - 80[
Nombre	120	100	140	200	180	160	100

1- Déterminer la population statistique, le caractère et sa nature.
2- Déterminer le mode graphiquement et calculer sa valeur.
3- Tracer les courbes des effectifs cumulées (croissants et décroissants)
4- Déterminer la médiane graphiquement et calculer sa valeur. Quelle est sa signification ?
5- Calculer les quantiles ; Q1, Q2, Q3, D5, C50 et C75.
6- Calculer le moyen âge de ces individus.
7- Quel est le nombre d'individus ayant l'âge moins de 20 ans ?
8- Quel est le nombre d'individus ayant l'âge au moins 15 ans ?

Exercice 3 : Les taux de croissance annuel du PIB d'un pays au cours des années 2000-2003 se présentent dans le tableau suivant :

Année	2000	2001	2002	2003
Taux	7,2	6,3	7	4,3

- Quel est le taux de croissance annuel moyen du PIB au cours de ces quatre années ?