

1 Médiane et quartiles d'une série statistique

1.1 Médiane

Définition 1. La valeur médiane partage les valeurs d'une série statistique en deux groupes de même effectif :

- les valeurs inférieures ou égales à la valeur médiane,
- les valeurs supérieures ou égales à la valeur médiane.

1.2 Quartiles

1.2.1 Définition

Définition 2. La médiane Me et les quartiles Q_1 et Q_3 partagent la population d'une série statistique suivant le principe suivant :

25% des valeurs d'une série statistique sont inférieures ou égale au 1^{er} quartile

25% des valeurs d'une série statistique sont supérieures ou égale au 3^{eme} quartile

1.2.2 Écart inter quartile et étendue

Définition 3. Soit une série statistique. On appelle écart interquartile la différence entre les quartiles Q_3 et Q_1 :

$$E = Q_3 - Q_1$$

Il mesure la dispersion des 50% des valeurs qui entourent la médiane.

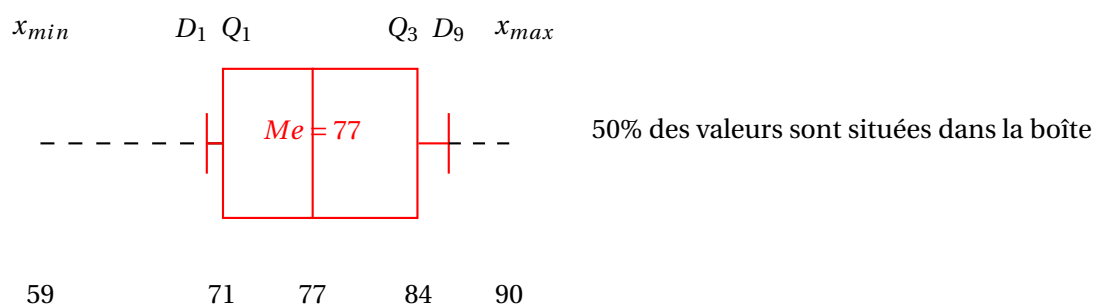
Définition 4. x_{min} et x_{max} étant respectivement la plus petite et la plus grande valeur d'une série statistique, l'étendue est la différence des valeurs extrêmes : $e = x_{max} - x_{min}$

1.2.3 Représentation par un diagramme en boîte

Définition 5. On appelle diagramme en boîte, un diagramme constitué d'une boîte rectangulaire délimitée par les quartiles Q_1 et Q_3 , indiquant la médiane M , et de moustaches qui relient les quartiles aux déciles D_1 et D_9 .

Exemple :

Rythme cardiaque	59	63	70	71	73	77	79	80	84	86	88	90
Effectif	2	1	5	2	5	4	2	2	4	5	1	1



2 Moyenne et variance

2.1 Moyenne

Définition 6. Si on note (x_i) ($1 \leq i \leq p$) les valeurs d'une série statistique, et n_i l'effectif relatif à la valeur x_i , N l'effectif total ($N = \sum_{i=1}^p n_i$), alors la moyenne, notée \bar{x} vaut

$$\bar{x} = \frac{\sum_{i=1}^p x_i n_i}{N}.$$

2.2 Variance et écart type

Définition 7. Soit (x_i, n_i) une série statistique d'effectif total N ($1 \leq i \leq p$).

- La variance de cette série est le nombre : $V = \frac{\sum_{i=1}^p n_i (x_i - \bar{x})^2}{N} = \frac{n_1(x_1 - \bar{x})^2 + n_2(x_2 - \bar{x})^2 + \dots + n_p(x_p - \bar{x})^2}{N}$
- L'écart type de cette série est le nombre : $s = \sqrt{V}$

Remarque : On définit la fréquence f_i relative à la valeur x_i par $f_i = \frac{n_i}{N}$, alors on peut aussi définir la variance par : $V = \sum_{i=1}^p f_i (x_i - \bar{x})^2$

Théorème 1. La variance de la série statistique (x_i, n_i) ($1 \leq i \leq p$) d'effectif N est :

$$V = \frac{\sum_{i=1}^p n_i x_i^2}{N} - \bar{x}^2$$

Démonstration :

$$\begin{aligned} V &= \frac{n_1(x_1 - \bar{x})^2 + n_2(x_2 - \bar{x})^2 + \dots + n_p(x_p - \bar{x})^2}{N} \\ &= \frac{n_1(x_1^2 - 2x_1\bar{x} + \bar{x}^2) + n_2(x_2^2 - 2x_2\bar{x} + \bar{x}^2) + \dots + n_p(x_p^2 - 2x_p\bar{x} + \bar{x}^2)}{N} \\ &= \frac{(n_1x_1^2 + n_2x_2^2 + \dots + n_px_p^2) - (2n_1x_1\bar{x} + 2n_2x_2\bar{x} + \dots + 2n_px_p\bar{x}) + (n_1\bar{x}^2 + n_2\bar{x}^2 + \dots + n_p\bar{x}^2)}{N} \\ &= \frac{(n_1x_1^2 + n_2x_2^2 + \dots + n_px_p^2) - 2\bar{x}(n_1x_1 + n_2x_2 + \dots + n_px_p) + \bar{x}^2(n_1 + n_2 + \dots + n_p)}{N} \\ &= \frac{(n_1x_1^2 + n_2x_2^2 + \dots + n_px_p^2)}{N} - 2\bar{x} \frac{(n_1x_1 + n_2x_2 + \dots + n_px_p)}{N} + \bar{x}^2 \frac{(n_1 + n_2 + \dots + n_p)}{N} \\ &= \frac{n_1x_1^2 + n_2x_2^2 + \dots + n_px_p^2}{N} - 2\bar{x} \cdot \bar{x} + \bar{x}^2 \\ &= \frac{n_1x_1^2 + n_2x_2^2 + \dots + n_px_p^2}{N} - \bar{x}^2 \end{aligned}$$



3 EXERCICES : Les exercices de base

1. Un élève a obtenu les notes suivantes aux quatre épreuves d'un examen :

Matière	Histoire	Français	Anglais	Mathématiques
note	12	14	11	14
coefficient		4	2	

Il ne se souvient plus du coefficient de la note de Mathématiques et de la note d'Histoire excepté que le coefficient de Mathématiques est supérieur de 1 à celui d'Histoire. Il pense que sa moyenne est 13, est-il possible de retrouver le coefficient de Mathématiques et d'Histoire ?

2. Soient deux séries de nombres :

Série 1 :	0	5	9	6	11	14	3	15	16	10	11
Série 2 :	12	8	4	11	5	14	14	11	8	3	18

- (a) Les deux séries suivantes ont-elles la même médiane ?
- (b) Si on retire à chaque série la note la plus haute et la note la plus basse, ont-elles la même médiane ?
3. Construire une série de 8 notes comprises entre 2 et 18, dont la médiane est 10 et la moyenne 10 et dont l'étendue soit 10. Y a-t-il plusieurs choix possibles
4. Une entreprise MATHIX emploie différents types de personnels. Le tableau suivant donne la répartition des salaires mensuels en centaines d'euros.

Salaire mensuel net(en centaines d'euros)	10	12	15	25	30
Nombres de personnes	5	8	24	13	2

- (a) Quel est l'effectif total ? Quelle est l'étendue ?
- (b) Donner le tableau des effectifs cumulés croissants et fréquences cumulées croissantes.
- (c) Pour cette entreprise donner le salaire médian et le salaire moyen.
- (d) Donner la représentation en boîte .

4 EXERCICES : Les exercices de base (corrigés)

1. Soit n le coefficient de Mathématiques et donc $n - 1$ celui d'Histoire.

On a alors :

Matière	Histoire	Français	Anglais	Mathématiques
note	12	14	11	14
coefficient	$n - 1$	4	2	n

on doit résoudre $13 = \frac{12n - 12 + 14 \times 4 + 11 \times 2 + 14n}{5 + 2n}$

soit $26n + 66 = 26n + 65$ donc $66 = 65$ ce qui est impossible.

2. (a) On classe les deux séries par ordre croissant :

Série 1 :	0	3	5	6	9	10	11	11	14	15	16
Série 2 :	3	4	5	8	8	11	11	12	14	14	18

Médiane de la série 1 : 10 Médiane de la série 2 : 11

Les deux séries n'ont pas la même médiane.

- (b) Si on retire à chaque série la note la plus haute et la note la plus basse, on obtient de

nouvelles séries :

Série 1 :	3	5	6	9	10	11	11	14	15
Série 2 :	4	5	8	8	11	11	12	14	14

Médiane de la série 1 bis : 10 Médiane de la série 2 bis : 11

Les deux médianes ne changent pas.

3. Par exemple : 4 9 9 10 10 11 13 14

Il y a plusieurs choix.

4. On a :

Salaire mensuel net(en centaines d'euros)	10	12	15	25	30
Nombres de personnes	5	8	24	13	2

- (a) L'effectif total est de 52 personnes et l'étendue est 2000 € ($30 - 10 = 20$ centaines d'euros).

- (b) Le tableau des effectifs cumulés et fréquences cumulées croissantes est :

Salaire mensuel net(en centaines d'euros)	10	12	15	25	30
Effectif cumulé croissant	5	13	37	50	52
Fréquence cumulée croissante (en %)	9,62%	25%	71,15%	96,15%	100%

- i. Le salaire médian correspondant à la valeur des 26^e et 27^e salaires soit 1500 €

- ii. Le salaire moyen est

$$\frac{5 \times 1000 + 8 \times 1200 + 24 \times 1500 + 13 \times 2500 + 2 \times 3000}{5 + 8 + 24 + 13 + 2} = \frac{89100}{52} \approx 1713$$

soit un salaire moyen de 1713 €

- (c) La représentation en boîte est :

