

**Objectifs :** Caractéristiques de dispersion : variance, écart-type. Diagramme en boîte.

Utiliser de façon appropriée les deux couples usuels qui permettent de résumer une série statistique : (moyenne, écart-type) et (médiane, écart interquartile).

Étudier une série statistique ou mener une comparaison pertinente de deux séries statistiques à l'aide d'un logiciel ou d'une calculatrice.

### I- Médiane, quartiles et diagramme en boîte

On se donne une série statistique :

|            |       |       |     |       |
|------------|-------|-------|-----|-------|
| Valeur     | $x_1$ | $x_2$ | ... | $x_p$ |
| Effectif   | $n_1$ | $n_2$ | ... | $n_p$ |
| Fréquences | $f_1$ | $f_2$ | ... | $f_p$ |

$N$  est l'effectif total ;  $N = n_1 + n_2 + \dots + n_p$ .

#### La médiane

**Définition :** Pour une série ordonnée, la **médiane** d'une série statistique est la valeur du caractère qui partage cette série en deux groupes de même effectif.

**Méthode :** Si la série contient  $N$  valeurs rangées dans l'ordre croissant :

- si  $N$  est impair, on prend la  $\frac{N+1}{2}$  ème valeur pour médiane.

- si  $N$  est pair, on prend pour médiane la moyenne entre la  $\frac{N}{2}$  ème et la  $\frac{N}{2}+1$  ème valeur.

**Exemples :** Avec un effectif total **impair** :  $x_1 \leq x_2 \leq x_3 \leq x_4 \leq x_5$ , la médiane est  $x_3$ .

Avec un effectif total **pair** :  $x_1 \leq x_2 \leq x_3 \leq x_4 \leq x_5 \leq x_6$ , la médiane est  $\frac{x_3 + x_4}{2}$

#### Les quartiles

**Définitions :** La liste des  $N$  données est rangée par ordre croissant.

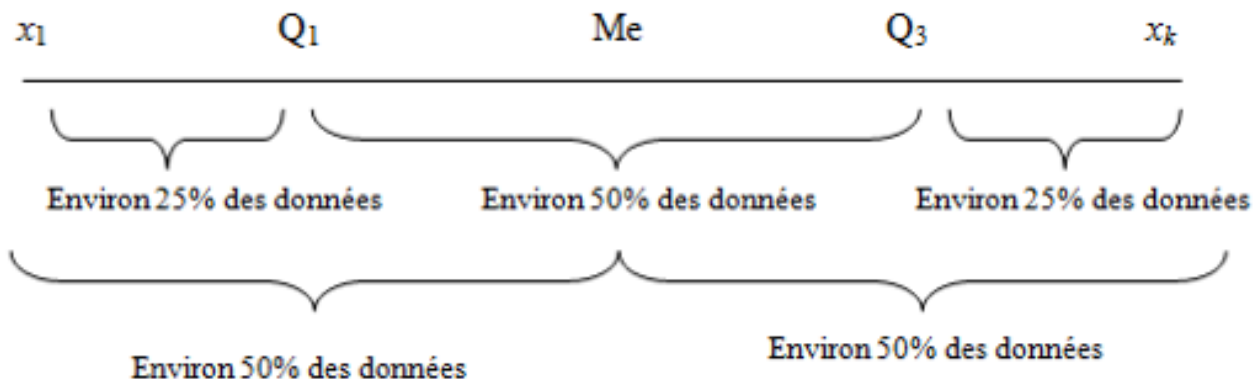
**Le premier quartile  $Q_1$**  est la plus petite donnée de la liste telle qu'au moins 25% des données soient inférieures ou égales à  $Q_1$ .

**Le troisième quartile  $Q_3$**  est la plus petite donnée de la liste telle qu'au moins 75% des données soient inférieures ou égales à  $Q_3$ .

**Méthode :** Pour  $Q_1$ , on calcule  $N/4$ , puis on détermine le premier entier  $p$  supérieur ou égal à  $N/4$ . Cet entier  $p$  est le rang de  $Q_1$ . Pour  $Q_3$ , on fait de même avec  $3N/4$

**Exemple :** Pour  $N=15$ , on a  $N/4=3,75$  et  $3N/4 = 11,25$ . Donc  $Q_1$  est la quatrième valeur de la série et  $Q_3$  est la douzième valeur.

**Écart interquartile :** c'est la différence  $Q_3 - Q_1$



## Diagramme en boîte

On peut représenter ces données sous forme de **diagramme en boîte** ou **boîte à moustaches**.

Sur ce diagramme apparaissent la valeur minimale,  $Q_1$ , Me,  $Q_3$  et la valeur maximale.

De cette boîte s'étirent deux moustaches (représentées par des traits) jusqu'au minimum et au maximum.



## II- Moyenne, Variance et écart-type

### La moyenne

**Définition** : La **moyenne** de cette série est le nombre réel, noté  $\bar{x}$ , tel que :

$$\bar{x} = \frac{n_1 x_1 + n_2 x_2 + \dots + n_p x_p}{N} \quad \text{où } N \text{ est l'effectif total ; } N = n_1 + n_2 + \dots + n_p.$$

On note souvent  $N = \sum_{i=1}^p n_i$  (somme des  $n_i$  de  $i = 1$  à  $p$ ) ;  $\bar{x} = \frac{1}{N} \sum_{i=1}^p n_i x_i = \sum_{i=1}^p f_i x_i$

Remarque : si les effectifs représentent des coefficients, on l'appelle moyenne pondérée

### La variance

Définition : c'est la moyenne des carrés des écarts à la moyenne

$$V = \frac{n_1(x_1 - \bar{x})^2 + n_2(x_2 - \bar{x})^2 + \dots + n_p(x_p - \bar{x})^2}{N} = \frac{1}{N} \sum_{i=1}^p n_i(x_i - \bar{x})^2$$

ROC : Démontrer que  $V = \left( \frac{1}{N} \sum_{i=1}^p n_i x_i^2 \right) - \bar{x}^2$

**L'écart-type**  $\sigma = \sqrt{V}$

L'avantage de l'écart-type est de s'exprimer dans la même unité que les données  $x_i$

## III- Résumé d'une série statistique

On résume souvent une série statistique par un paramètre de tendance centrale associé à un paramètre de dispersion. Deux choix sont couramment proposés : le couple **{moyenne - écart type}** qui a l'inconvénient d'associer deux paramètres sensibles aux valeurs extrêmes et le couple **{médiane - écart interquartile}** qui n'a pas ce défaut mais dont la détermination est moins pratique.

**Exercice 1 :** Pour chaque série de points marqués par des joueurs de basket-ball, indiquer par lecture directe la médiane, les premier et troisième quartiles.

- Points marqués par Tony : 8    10    12    15    20    22    27    35    38    38    45
- Points marqués par Parker : 11    12    12    18    20    25    30    30    33    38    41

Résumer les séries par le couple (médiane, écart interquartile) est-il judicieux pour comparer ces joueurs ?

Déterminer la moyenne et l'écart-type de chaque série.

Quel joueur, l'entraîneur préférera-t-il sélectionner pour le prochain match ? Pourquoi ?

### Exercice 2 :

1) Voici la répartition des notes de mathématiques de Mickaël.

|           |   |    |    |    |
|-----------|---|----|----|----|
| Notes     | 5 | 11 | 12 | 18 |
| Effectifs | 4 | 3  | 1  | 3  |

Déterminer la moyenne et l'écart type de Mickaël à l'aide de la calculatrice.

2) Voici les notes obtenues par Jackson :

|           |     |   |   |    |    |    |    |
|-----------|-----|---|---|----|----|----|----|
| Notes     | 7.5 | 8 | 9 | 10 | 12 | 13 | 14 |
| Effectifs | 1   | 1 | 2 | 4  | 5  | 3  | 1  |

Déterminer la moyenne, puis l'écart-type de Jackson à l'aide de la calculatrice.

3) Que peut-on conclure sur le niveau et les résultats de ces deux élèves ?