

## 1 Caractéristiques de position

### RAPPELS

- La moyenne  $\bar{x}$  peut se calculer à partir :

– des effectifs :  $\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_px_p}{N}$

avec  $N = n_1 + n_2 + \dots + n_p$ ;

– des fréquences :  $\bar{x} = f_1x_1 + f_2x_2 + \dots + f_px_p$ .

- La médiane  $Me$  est le nombre tel que :

— si  $N$  est impair,  $N = 2k + 1$ , alors  $Me = x_{k+1}$  ;

— si  $N$  est pair,  $N = 2k$ , alors  $Me = \frac{x_k + x_{k+1}}{2}$ .

- Les quartiles  $Q_1$  et  $Q_3$  sont les valeurs que l'on calcule de la manière suivante :

$Q_1$  (resp.  $Q_3$ ) est la valeur dont le rang est

le premier entier supérieur ou égal à  $\frac{N}{4}$

(resp.  $\frac{3N}{4}$ ).

#### Définition 1

Le premier décile, noté  $D_1$  (respectivement le 9<sup>e</sup> décile, noté  $D_9$ ), est la plus petite valeur d'une série statistique telle qu'au moins 10 % (respectivement au moins 90 %) des données soient inférieures ou égales à ce nombre.

## 2 Caractéristiques de dispersion

Voir le mémento pour les définitions de l'étendue et de l'écart interquartile.

#### Définition 2

Étant donnée une série statistique de modalités  $(x_i)$ , d'effectif total  $N$

et de moyenne  $\bar{x}$ , on appelle variance le réel :  $V = \frac{1}{N} \sum_{i=1}^n n_i (x_i - \bar{x})^2$ .

**REMARQUE :**  $V$  est la moyenne des carrés des écarts à la moyenne des valeurs  $x_i$  de la série.

#### Propriété 1

La variance est égale à la différence entre la moyenne des carrés et le carré

de la moyenne :  $V = \frac{1}{N} \sum_{i=1}^n n_i x_i^2 - \bar{x}^2 = \overline{x^2} - \bar{x}^2$ .

### Définition 3

L'écart type est la racine carrée de la variance :  $\sigma = \sqrt{V}$ .

**EXEMPLE :** On étudie les notes d'une classe de 29 élèves de première S.

Notes	0,5	1	2	2,5	4	5	5,5	6,5	7,5	8	8,5	9	9,5	11	12	12,5	13,5	17,5	18,5
Effectif	1	3	1	2	1	1	2	1	1	3	1	2	2	2	1	1	2	1	1

On calcule la moyenne  $\bar{x} \approx 7,7$  et alors la variance vaut :  $V \approx 22,7$  et l'écart type :  $\sigma \approx 4,8$ .

#### Note :

Le diagramme en boîte a été inventé en 1977 par John Tukey.

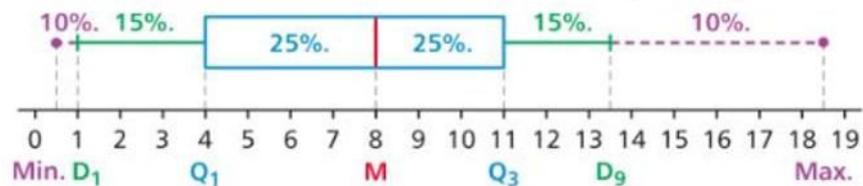
## 3 Diagramme en boîte

### Définition 4

On appelle **diagramme en boîte** (ou boîte à moustaches), 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  $D_9$ .

- REMARQUES :**
- On prend parfois pour extrémités des moustaches les valeurs extrêmes de la série.
  - Par convention, la largeur du rectangle est souvent prise proportionnelle à  $\sqrt{N}$ .
  - Le diagramme en boîte permet de comparer les caractéristiques de deux séries (voir savoir-faire 4 p. 234).

**EXEMPLE :** Avec l'exemple précédent on a :  $D_1 = 1$ ,  $D_9 = 13,5$ ,  $Q_1 = 4$ ,  $Q_3 = 11$  et  $M = 8$ .



## Résumé d'une série statistique

Résumer une série, c'est indiquer la **répartition des données** en utilisant différents indicateurs. Deux questions peuvent alors être posées :

- Autour de quelle valeur centrale les données sont-elles réparties ?
- Quelle est l'importance de la dispersion des données autour de cette valeur centrale ?

On utilise habituellement un paramètre de position indiquant une tendance centrale et un paramètre de dispersion.

Paramètre de tendance centrale	Paramètre de dispersion	Propriété
médiane : $Me$	écart interquartile : $E_i = Q_3 - Q_1$	peu sensible aux valeurs extrêmes
moyenne : $\bar{x}$	écart-type : $s$	sensible aux valeurs extrêmes