

## STATISTIQUES EN PREMIÈRE ES ET L

### I - RAPPELS

#### DÉFINITIONS

Les statistiques permettent d'étudier un **caractère** d'une **population**.

Le nombre d'éléments de la population s'appelle l'**effectif global** (ou l'**effectif total**).

Pour une valeur de caractère donnée, l'**effectif** est le nombre d'éléments correspondant à cette valeur.

Une **série statistique** est un tableau donnant les effectifs pour chacune des valeurs possibles du caractère.

#### EXEMPLE

On effectue une étude portant sur l'âge des élèves d'un lycée.

- le **caractère** étudié est l'âge
- la **population** est l'ensemble des élèves du lycée
- l'**effectif global** est le nombre d'élèves du lycée
- le tableau ci-dessous est la **série statistique** pour ce caractère dans un lycée donné :

âges (en années)	14	15	16	17	18	19	20
effectifs	3	22	65	82	59	35	2

### II - MÉDIANE - QUARTILES - DÉCILES

#### DÉFINITION

La **médiane** d'une série statistique est la valeur du caractère qui partage la population en deux classes de même effectif.

#### REMARQUE

En pratique pour trouver la médiane d'une série statistique d'effectif global  $n$  :

- On ordonne les valeurs du caractère dans l'ordre croissant.
- Si  $n$  est pair, la médiane sera la moyenne des valeurs du terme de rang  $\frac{n}{2}$  et du terme de rang  $\frac{n}{2} + 1$ .
- Si  $n$  est impair, la médiane sera la valeur du terme de rang  $\frac{n+1}{2}$ .
- Lorsque l'effectif global est élevé, il est souvent utile de calculer les effectifs cumulés pour trouver cette valeur.

**EXEMPLE**

On lance 10 fois un dé à six faces. Les résultats obtenus sont : 1;5;6;6;3;2;3;1;4;1

On trie ces valeurs par ordre croissant : 1;1;1;2;3;3;3;4;5;6;6

$n = 10$  étant pair on effectue la moyenne du cinquième et du sixième terme (3 et 3) et la médiane est donc 3.

**DÉFINITIONS**

- Le **premier quartile**  $Q_1$  d'une série statistique est la plus petite valeur des termes de la série pour laquelle au moins un quart des données sont inférieures ou égales à  $Q_1$ .
- Le **troisième quartile**  $Q_3$  d'une série statistique est la plus petite valeur des termes de la série pour laquelle au moins trois quarts des données sont inférieures ou égales à  $Q_3$ .
- Le **premier décile**  $D_1$  d'une série statistique est la plus petite valeur des termes de la série pour laquelle au moins 10% des données sont inférieures ou égales à  $D_1$ .
- Le **neuvième décile**  $D_9$  d'une série statistique est la plus petite valeur des termes de la série pour laquelle au moins 90% des données sont inférieures ou égales à  $D_9$ .

**DÉFINITION**

L'**écart interquartile** est la différence entre le troisième et le premier quartile  $Q_3 - Q_1$ .

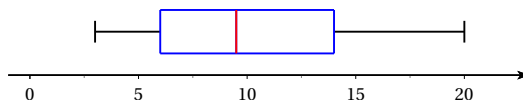
**REMARQUE**

L'écart interquartile mesure la dispersion autour de la médiane.

**III - DIAGRAMME EN BOÎTE**

On peut résumer un certain nombre d'informations relatives à une série statistique grâce à un **diagramme en boîte** (aussi appelé *boîte à moustache*) qui fait apparaître (voir figure ci-dessus) :

- les valeurs minimum et maximum
- le premier et le troisième quartile ( $Q_1$  et  $Q_3$ )
- la médiane

**EXEMPLE**

La figure ci-dessus représente une série statistique de valeurs extrêmes 3 et 20, de premier quartile 6, de troisième quartile 14 et de médiane 9,5.