

# Chapitre 6 : Statistiques

## I) Paramètres de position

### 1) Moyenne :

**Définition 1 :** soit  $x_1, x_2, \dots, x_k$  les valeurs d'une série statistique et  $n_1, n_2, \dots, n_k$  les effectifs associés. La **moyenne** de la série statistique, notée  $\bar{x}$  est donnée par :

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \dots + n_kx_k}{n_1 + n_2 + \dots + n_k} = \frac{\sum_{i=1}^k n_i x_i}{\sum_{i=1}^k n_i}$$

Exemple : On s'intéresse au nombre de frères et sœurs des élèves de la 1<sup>ère</sup>S1 du LPT :

Nombre de frères et sœurs ( $x_i$ )						
Effectifs ( $n_i$ )						

$\bar{x} = \dots\dots\dots$

Rappel : La fréquence de chaque valeur est donnée par  $f_i = \dots\dots$  où  $N = \sum_{i=1}^k n_i$  est l'effectif total.

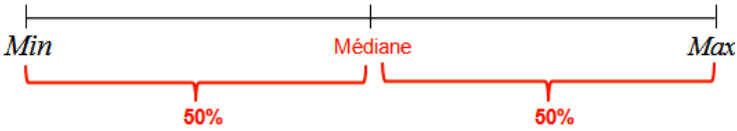
**Propriété 1 :** soit  $x_1, x_2, \dots, x_k$  les valeurs d'une série statistique et  $f_1, f_2, \dots, f_k$  les fréquences associés. La **moyenne** de la série statistique est donnée par :

$$\bar{x} = \sum_{i=1}^k f_i x_i .$$

Remarque : Pour une série regroupée en classes (caractère  $\dots\dots\dots$ ), on calcule une valeur approchée de la moyenne en prenant les  $\dots\dots\dots$  des intervalles.

### 2) Médiane :

**Définition 2 :** La **médiane** d'une série statistique, notée  $M_e$ , est telle que la moitié (50%) au moins des valeurs du caractère sont inférieures ou égales à  $M_e$  et la moitié (50%) au moins ont une valeur supérieure à  $M_e$ .

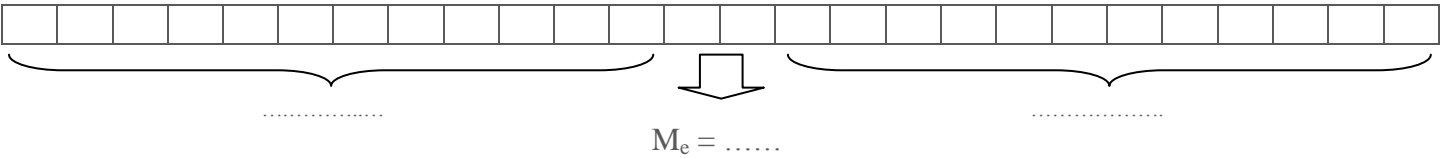


Conséquence :

- ✓ Si l'effectif total est pair,  $N = 2p$  : la médiane est comprise entre la  $\dots\dots\dots$  et  $\dots\dots\dots$  valeur.
- ✓ Si l'effectif total est impair,  $N = 2p+1$  : la médiane est la  $\dots\dots\dots$  valeur.

Remarque : La médiane, contrairement à la moyenne, n'est pas influencée par les  $\dots\dots\dots$

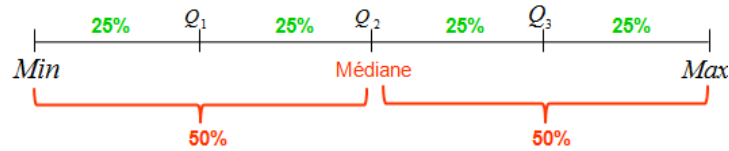
Exemple : L'effectif total est de  $N = \dots\dots$  donc la médiane est comprise entre la  $\dots\dots$  et la  $\dots\dots$  valeur



### 3) Quartiles et déciles :

#### Définition 3 :

- Le **premier quartile**  $Q_1$  d'une série statistique est la plus petite valeur de cette série telle qu'au moins 25% des données soient inférieures ou égales à  $Q_1$ .
- Le **troisième quartile**  $Q_3$  d'une série statistique est la plus petite valeur de cette série telle qu'au moins 75% des données soient inférieures ou égales à  $Q_3$ .



#### Conséquence :

- ✓ Le premier quartile  $Q_1$  est la valeur  $x_i$  dont l'indice  $i$  est le plus petit entier supérieur ou égal à .....
- ✓ Le troisième quartile  $Q_3$  est la valeur  $x_i$  dont l'indice  $i$  est le plus petit entier supérieur ou égal à .....

Exemple :  $\frac{N}{4} = \dots$  donc le premier quartile est la ..... valeur :  $Q_1 = \dots$

$\frac{3N}{4} = \dots$  donc le troisième quartile est la ..... valeur :  $Q_3 = \dots$

#### Définition 4 :

- Le **premier décile**  $d_1$  d'une série statistique est la plus petite valeur de cette série telle qu'au moins 10% des données soient inférieures ou égales à  $d_1$ .
- Le **neuvième décile**  $d_9$  d'une série statistique est la plus petite valeur de cette série telle qu'au moins 90% des données soient inférieures ou égales à  $d_9$ .

#### Conséquence :

- ✓ Le premier décile  $d_1$  est la valeur  $x_i$  dont l'indice  $i$  est le plus petit entier supérieur ou égal à .....
- ✓ Le neuvième décile  $d_9$  est la valeur  $x_i$  dont l'indice  $i$  est le plus petit entier supérieur ou égal à .....

Exemple :  $\frac{N}{10} = \dots$  donc le premier décile est la ..... valeur :  $d_1 = \dots$

$\frac{9N}{10} = \dots$  donc le neuvième décile est la ..... valeur :  $d_9 = \dots$

## II) Paramètres de dispersion

### 1) Etendue :

**Définition 5 :** L'**étendue** d'une série statistique est la différence entre la valeur maximum et la valeur minimum du caractère étudié.

Exemple :  $e = \dots - \dots = \dots$

### 2) Intervalle et écart interquartiles :

#### Définition 6 :

- L'intervalle  $[Q_1; Q_3]$  est l'**intervalle interquartile**.
- L'**écart interquartile** d'une série statistique est la différence  $Q_3 - Q_1$ .

Exemple : L'intervalle interquartile est : ..... et l'écart interquartile est : .....

### 3) Variance et écart type :

**Définition 7 :** soit  $x_1, x_2, \dots, x_k$  les valeurs d'une série statistique et  $n_1, n_2, \dots, n_k$  les effectifs associés. La **variance** de la série statistique, notée  $V$  est donnée par :

$$V = \frac{n_1(x_1 - \bar{x})^2 + n_2(x_2 - \bar{x})^2 + \dots + n_k(x_k - \bar{x})^2}{n_1 + n_2 + \dots + n_k} = \frac{\sum_{i=1}^k n_i(x_i - \bar{x})^2}{\sum_{i=1}^k n_i}$$

Autre écriture :  $V = \dots\dots\dots$

$V = \dots\dots\dots$

$V = \dots\dots\dots$

Exemple :  $V = \dots\dots\dots$

**Définition 8 :** On considère une série statistique de variance  $V$  .

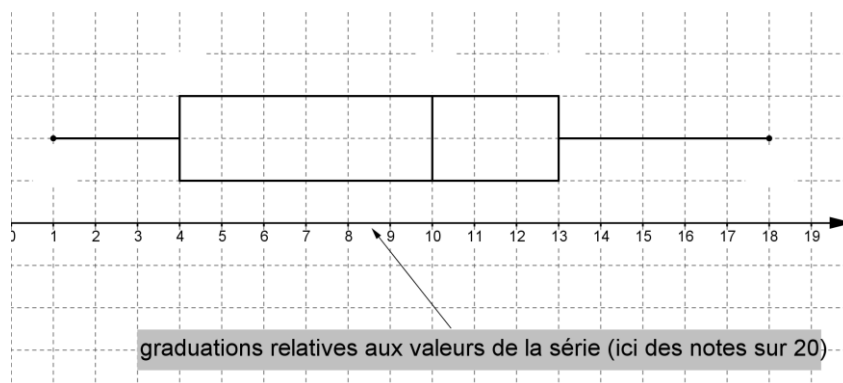
L'**écart type** de cette série est le réel noté  $\sigma$  tel que  $\sigma = \sqrt{V}$  .

Exemple :  $\sigma = \dots\dots\dots$

Remarque : L'écart type mesure la ..... de la série autour de sa moyenne. Plus l'écart type est faible, plus la population est ..... ; plus il est élevé, plus la population est .....

### III) Diagramme en boîte

Le diagramme en boîte ou « boîte à moustache » d'une série statistique présente son **minimum**, son **premier quartile**, sa **médiane**, son **troisième quartile** et son **maximum**.



Remarque : on trouve aussi des diagrammes en boîtes donnant en plus le premier et le neuvième décile...

