

# STATISTIQUES

---

## I) TYPE DE CARACTÈRE ÉTUDIÉ

Exemple	Caractère (ou variable statistique)	Individus	Population	Modalités (ou valeurs)	Type de caractère
Couleur des yeux des élèves de la classe	La couleur des yeux	Chaque élève	La classe	Bleus, verts, marrons, ...	Qualitatif
Décimales de $\pi$	Valeur des décimales	Chaque décimale	Les décimales de $\pi$	Entiers de 0 à 9	Quantitatif discret
Durée de vie d'une pile	Durée de vie	Chaque pile	L'ensemble des piles fabriquées	Nombre positif quelconque	Quantitatif continu
Nombre d'enfants par familles	Le nombre d'enfants par familles	Chaque famille française	Ensemble des familles françaises	Entiers de 0 à 15	Quantitatif discret
Population des villes françaises	Le nombre d'habitants	Chaque ville	L'ensemble des villes françaises	Entier positif quelconque	Quantitatif continu

### Remarques :

- Dans le cas d'un caractère quantitatif discret, les données peuvent être « en vrac » (cf 7 p311) ou en tableau d'effectifs ou de fréquences (cf 8 p311)
- Dans le cas d'un caractère quantitatif continu, les données peuvent être aussi « en vrac » mais elles sont le plus souvent réparties par classes (cf 3 p311)

## II) MOYENNE ET ÉCART-TYPE

### 1) Exemple :

Soit la série statistique ci-dessous :

Durée (h)	0	1	2	3	4
effectifs	1	2	1	4	2

#### a) Moyenne

La moyenne est :  $\bar{x} = \frac{0+1+1+2+3+3+3+3+4+4}{10} = \frac{24}{10} = 2,4 \text{ h}$

ou encore :  $\bar{x} = \frac{1 \times 2 + 2 + 3 \times 4 + 4 \times 2}{10} = \frac{24}{10} = 2,4 \text{ h}$  (plus efficace !)

#### b) Écart-type

Pour estimer l'hétérogénéité de la série, on décide de calculer la moyenne des écarts à la moyenne :

$$m_{\text{écarts}} = \frac{-2,4 - 1,4 \times 2 - 0,4 + 0,6 \times 4 + 1,6 \times 2}{10} = 0 \text{ h}$$

On réalise que, quelle que soit la série étudiée, on obtiendra toujours un résultat nul car les écarts à la moyenne négatifs vont toujours « neutraliser » les écarts positifs.

Pour éviter ce phénomène on décide de mettre les écarts au carré pour qu'ils soient tous positifs. On appelle alors « écart-type » le résultat obtenu :

$$\sigma = \sqrt{\frac{(-2,4)^2 + (-1,4)^2 \times 2 + (-0,4)^2 + 0,6^2 \times 4 + 1,6^2 \times 2}{10}} = \sqrt{\frac{16,4}{10}} \approx 1,28 \text{ h}$$

**Remarque :** la moyenne et l'écart-type s'expriment dans la même unité que les valeurs de la série.

## 2) Cas général :

Soit la série statistique ci-dessous :

valeurs	$x_1$	$x_2$	...	$x_p$
effectifs	$n_1$	$n_2$	...	$n_p$

On a :

$$\bar{x} = \frac{n_1 x_1 + n_2 x_2 + \dots + n_p x_p}{n_1 + n_2 + \dots + n_p}$$

$$\sigma = \sqrt{\frac{n_1 (x_1 - \bar{x})^2 + n_2 (x_2 - \bar{x})^2 + \dots + n_p (x_p - \bar{x})^2}{n_1 + n_2 + \dots + n_p}}$$

## 3) Linéarité de la moyenne :

**Ex :** Soit la série : 10, 12, 14.

$$\bar{x}_1 =$$

$$\sigma =$$

Ajoutons 2 : la nouvelle série est : 12, 14, 16.

$$\bar{x}_2 =$$

$$\sigma =$$

Divisons par 2 : la nouvelle série est : 6, 7, 8.

$$\bar{x}_3 =$$

$$\sigma =$$

**Cas général :** Soit  $a$  un réel quelconque :

- Si l'on ajoute  $a$  à toutes les données, la moyenne augmente de  $a$  mais l'écart-type ne change pas.
- Si on multiplie toutes les données par  $a$ , la moyenne et l'écart-type sont multipliés par  $a$ .

p319 : 41, 43

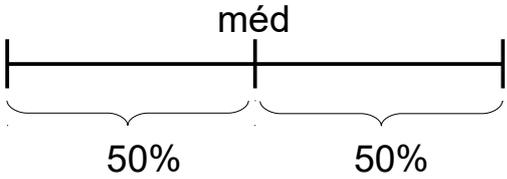
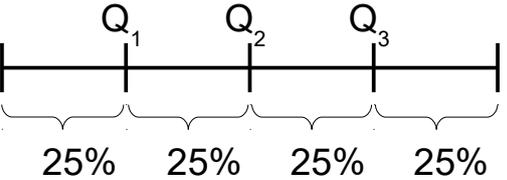
p324 : 69, 70, 72

p325 : 75, 78, 80

### III) MÉDIANE ET QUARTILES

#### 4) Déterminer la médiane et les quartiles

Soit une série rangée par ordre croissant dont l'effectif total est  $n$ .

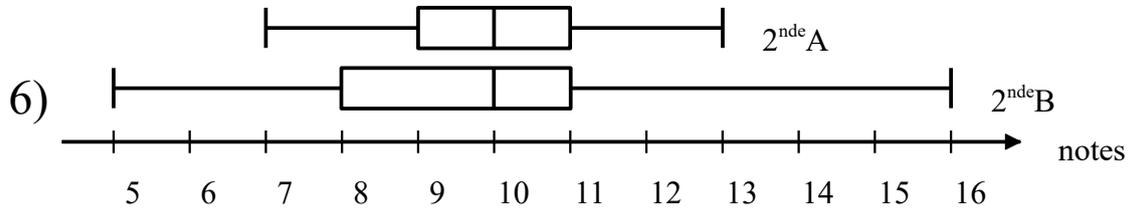
Paramètre statistique :	Pour déterminer le rang :
<p>La <b>médiane</b> partage la série en deux moitiés. C'est la valeur « centrale » de la série.</p>  <p style="text-align: center;">méd</p>	<p>On calcule <math>\frac{n+1}{2}</math> :</p> <ul style="list-style-type: none"><li>● Si le résultat est entier alors, la médiane est la valeur de rang <math>\frac{n+1}{2}</math></li><li>● Sinon, la médiane est la demi-somme des deux valeurs dont les rangs entourent <math>\frac{n+1}{2}</math></li></ul>
<p>Les <b>quartiles</b> partagent la série en 4. <math>Q_1</math> est la 1<sup>ère</sup> valeur telle que 25% des valeurs lui soient inférieures ou égales <math>Q_3</math> est la 1<sup>ère</sup> valeur telle que 75% des valeurs lui soient inférieures ou égales</p>  <p style="text-align: center;"><math>Q_1</math>   <math>Q_2</math>   <math>Q_3</math></p>	<ul style="list-style-type: none"><li>● <math>Q_1</math> est la 1<sup>ère</sup> valeur dont le rang est supérieur ou égal à <math>\frac{n}{4}</math></li><li>● <math>Q_3</math> est la 1<sup>ère</sup> valeur dont le rang est supérieur ou égal à <math>\frac{3n}{4}</math></li></ul>

Le nombre  $Q_3 - Q_1$  s'appelle l'**écart interquartile** et donne une estimation de l'hétérogénéité de la série.

## 5) Diagrammes en boîtes :

Les diagrammes en boîtes (ou boîtes à moustaches) permettent de visualiser la médiane, les quartiles et les extrêmes d'une série. Ils sont notamment commodes pour comparer plusieurs séries :

**Ex :** Deux classes comparent leurs résultats au dernier devoir.



On remarque que les deux classes ont la même médiane et ont donc globalement le même niveau.

En revanche, l'écart interquartile et l'étendue sont plus petits pour la 2<sup>nde</sup> A qui est donc plus homogène que la 2<sup>nde</sup> B.

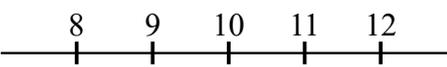
p324 : 74  
p325 : 77, 79  
p328 : 89

# IV) COMPARER LES INDICATEURS

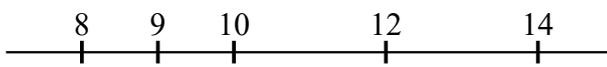
## 1) Indicateur de position ou de dispersion ?

<u>Indicateurs de position</u> (proposent une valeur centrale de la série)		<u>Indicateurs de dispersion</u> (la série est-elle regroupée autour de son centre ?)
<ul style="list-style-type: none"><li>● (mode)</li><li>● moyenne</li><li>● médiane</li></ul>	<ul style="list-style-type: none"><li>● quartiles</li></ul>	<ul style="list-style-type: none"><li>● (étendue)</li><li>● écart type</li><li>● écart interquartile</li></ul>

## 2) Interpréter un écart entre la médiane et la moyenne

- Soit la série suivante 

Ici, la moyenne et la médiane sont identiques : La série est bien « centrée »

- Soit la nouvelle série 

Ici, la moyenne est plus importante que la médiane : La série est plus « étalée à droite »

p333 : bilan 3

algo

p325 : 76, 81

p327 : 85

p329 : TP (erreur énoncé : partie A, ligne 9 algo)

## V) DANS LES EXERCICES

### 1) Caractère quantitatif avec données « en vrac »

**Enoncé :** Des amis ont mesuré leurs pieds en cm :

24,2 ; 26,2 ; 24,7 ; 23 ; 28 ; 24,7 ; 23,1 ; 26,5 ; 29,1 ; 25.

Déterminer la taille moyenne et l'écart-type, puis la taille médiane et les 1<sup>er</sup> et 3<sup>ème</sup> quartiles de cette série.

#### Rédaction :

- Calcul de la moyenne et de l'écart-type :

$$\bar{x} = \frac{24,2 + 26,2 + 24,7 + \dots + 25}{10} = 25,45 \text{ cm}$$

$$\sigma = \sqrt{\frac{(24,2 - \bar{x})^2 + (26,2 - \bar{x})^2 + (24,7 - \bar{x})^2 + \dots + (25 - \bar{x})^2}{10}} \approx 1,895 \text{ cm}$$

- Calcul de la médiane et des quartiles :

Rangeons la série par ordre croissant :

23 ; 23,1 ; 24,2 ; 24,7 ; 24,7 ; 25 ; 26,2 ; 26,5 ; 28 ; 29,1

L'effectif total est 10.

$\frac{10+1}{2} = 5,5$  donc la médiane est la demi-somme des 5<sup>èmes</sup> et 6<sup>èmes</sup> termes

$$\text{de la série : Méd} = \frac{24,7 + 25}{2} = 24,85 \text{ cm.}$$

$\frac{10}{4} = 2,5$  donc  $Q_1$  est le 3<sup>ème</sup> terme de la série :  $Q_1 = 24,2 \text{ cm.}$

$\frac{3 \times 10}{4} = 7,5$  donc  $Q_3$  est le 8<sup>ème</sup> terme de la série :  $Q_3 = 26,5 \text{ cm.}$

#### Attention :

- En fin de calcul, toujours préciser l'unité.
- En calculant l'écart-type, éviter les arrondis d'arrondis ! (d'où l'intérêt du mode statistiques de la calculatrice)
- Un rang est toujours entier : Non, la médiane n'a pas pour rang  $\frac{10+1}{2}$  !

## 2) Caractère quantitatif discret avec données en tableau d'effectifs

**Énoncé :** Les élèves d'une classe comparent le nombre de coups de téléphone qu'ils ont reçus aujourd'hui :

Nbre de coups de téléphone	0	1	2	3	4	5	6	7	8
Nbre d'élèves	7	5	2	4	8	1	3	0	1
ECC	7	12	14	18	26	27	30	30	31

Déterminer le nombre de coups de tél moyen, l'écart-type, le nombre de coups de tél médian, ainsi que les 1<sup>er</sup> et 3<sup>ème</sup> quartiles de cette série.

### Rédaction :

- Calcul de la moyenne et de l'écart-type :

$$\bar{x} = \frac{7 \times 0 + 5 \times 1 + 2 \times 2 + \dots + 8}{7 + 5 + 2 + \dots + 1} \approx 2,7 \text{ coups de téléphone.}$$

$$\sigma = \sqrt{\frac{7(0 - \bar{x})^2 + 5(1 - \bar{x})^2 + 2(2 - \bar{x})^2 + \dots + (8 - \bar{x})^2}{7 + 5 + 2 + \dots + 1}} \approx 2,17 \text{ coups de tél.}$$

- Calcul de la médiane et des quartiles :

L'effectif total est 31.

$$\frac{31 + 1}{2} = 16 \text{ donc la médiane est le } 16^{\text{ème}} \text{ terme de la série :}$$

$$\text{Méd} = 3 \text{ coups de tél.}$$

$$\frac{31}{4} = 7,75 \text{ donc } Q_1 \text{ est le } 8^{\text{ème}} \text{ terme de la série :}$$

$$Q_1 = 1 \text{ coup de tél.}$$

$$\frac{3 \times 31}{4} = 23,25 \text{ donc } Q_3 \text{ est le } 24^{\text{ème}} \text{ terme de la série :}$$

$$Q_3 = 4 \text{ coups de tél.}$$

### 3) Caractère quantitatif continu avec données réparties par classes

**Énoncé :** On s'intéresse à l'âge des fans de Bob l'éponge :

Age (ans)	[0 ; 10[	[10 ; 15[	[15 ; 20[	[20 ; 25[	[25 ; 40[	[40 ; 60[
Fréquence (%)	8,1	56	18,3	7,3	6,3	4
FCC	8,1	64,1	82,4	89,7	96	100

Déterminer une approximation de la moyenne et de l'écart-type de cette série, puis une approximation de la médiane ainsi que des 1<sup>er</sup> et 3<sup>ème</sup> quartiles.

**Attention :**

Nous ne connaissons pas ici les âges des fans de Bob l'éponge. Nous savons seulement comment ces âges sont répartis dans les différentes classes. Ne disposant donc pas des données initiales, nous ne pouvons calculer moyenne, écart-type, médiane et quartiles. En revanche, nous pouvons en chercher des approximations.

**Rédaction :**

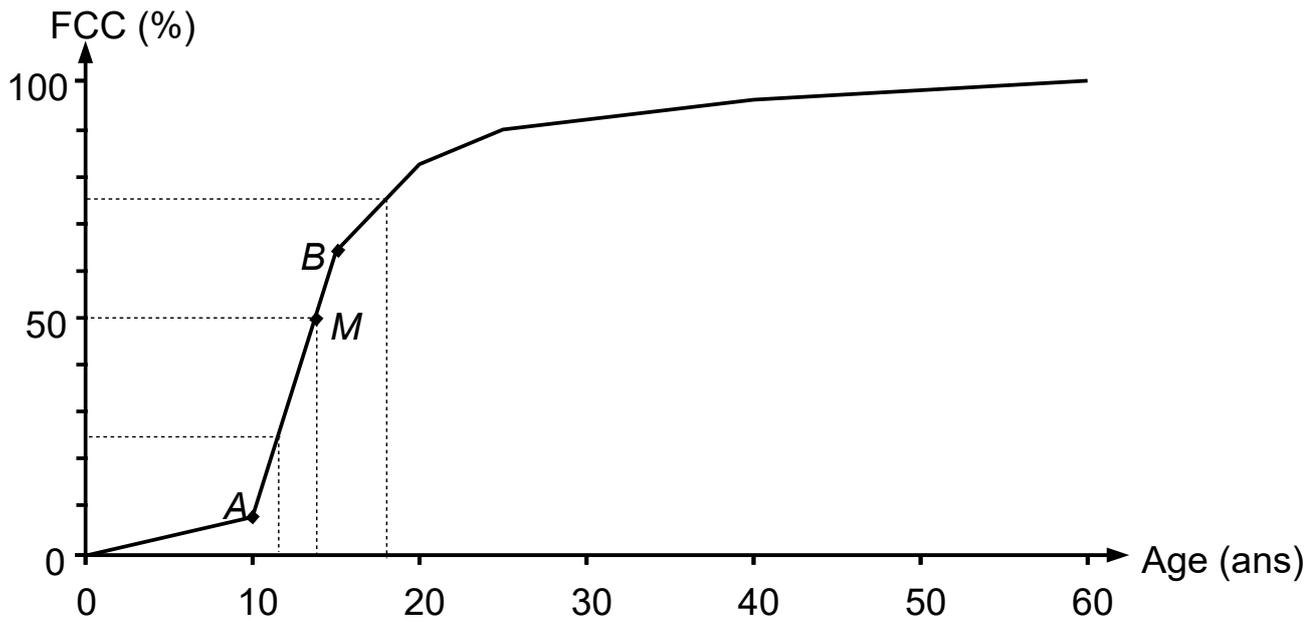
● Approximation de la moyenne et de l'écart-type :

Remplaçons chaque classe par son milieu

$$\bar{x} \approx \frac{8,1 \times 5 + 56 \times 12,5 + 18,3 \times 17,5 + \dots + 4 \times 50}{100} \approx 16,3 \text{ ans}$$

$$\sigma \approx \sqrt{\frac{8,1(5 - \bar{x})^2 + 56(12,5 - \bar{x})^2 + 18,3(17,5 - \bar{x})^2 + \dots + 4(50 - \bar{x})^2}{100}} \approx 9,1 \text{ ans}$$

- Approximation graphique de la médiane et des quartiles :  
Traçons le polygone des fréquences cumulées croissantes :



Le point de la courbe des FCC d'ordonnée 25 a pour abscisse environ 11,5  
donc :  $Q_1 \approx 11,5$  ans.

Le point de la courbe des FCC d'ordonnée 50 a pour abscisse environ 13,7  
donc : Méd  $\approx 13,7$  ans.

Le point de la courbe des FCC d'ordonnée 75 a pour abscisse environ 18  
donc :  $Q_3 \approx 18$  ans.