

Statistiques

1. Couple médiane. Intervalle interquartile

1.1. La médiane

Définition:

La médiane d'une série statistique est la valeur du caractère qui partage la population en deux effectifs égaux. Il y a donc autant de valeurs supérieures à la médiane que de valeurs inférieures.

Méthode :

Il faut ranger les valeurs par ordre croissant et diviser par deux l'effectif total. On appelle N ce nombre. Si N est nombre entier, la médiane est une valeur comprise entre les valeurs de rang N et de rang N+1.

Si N n'est pas un entier, la médiane est la première valeur de rang supérieure à N.

Application élémentaire :

Déterminer la médiane de la série 3 ; 7 ; 5 ; 6 ; 9

On range la série par ordre croissant : 3 ; 5 ; 6 ; 7 ; 9

La médiane est 6

Déterminer la médiane de la série 5 ; 8 ; 7 ; 5 ; 9 ; 2

On range la série par ordre croissant : 2 ; 5 ; 5 ; 7 ; 8 ; 9

La médiane est entre 5 et 7 car l'effectif est pair. La médiane est 6

Exemple 1 :

Un boulanger teste les masses (en grammes) de 30 baguettes qu'il vient de fabriquer, il obtient les résultats suivants :

235;235;237;238;238;239;239;239;240;241;241;243;245;247;247;249;250;205;250;250;250;251;251;253;253;255;255;255;257;260;

Déterminer la médiane de cette série :

$$30/2 = 15$$

La médiane est entre la 15e et la 16e valeur :

$$Me=248g$$

Exemple 2 :

Le tableau ci-dessous indique la durée (en minutes) de connexion internet par jour de 43 familles interrogées

Durée en minutes	40	60	80	120	180	200	240	300
Effectif	2	9	11	7	5	2	4	3

Déterminer la médiane de cette série :

$$43/2 = 21,5$$

La médiane est la 22eme valeur :

Durée en minutes	40	60	80	120	180	200	240	300
Effectif	2	9	11	7	5	2	4	3
Effectifs cumulés croissants	2	11	22	29	34	36	40	43

Me= 80 minutes

Remarque :

La médiane est un caractère de **position** d'une série statistique.

Il ne mesure pas **la dispersion** des valeurs mais donne une **valeur centrale** à la série.

1.2. Les quartiles

Définition:

On considère une série dont les données sont rangées dans l'ordre croissant :

Les quartiles sont des données de la série qui la partage en quatre parties à peu près de même effectif.

• Le premier quartile noté Q_1 , de la série ordonnée est la plus petite valeur de la série telle que 25% des valeurs soient inférieurs ou égales à Q_1

• Le troisième quartile noté Q_3 , de la série ordonnée est la plus petite valeur de la série telle que 75% des valeurs soient inférieurs ou égales à Q_3

Exemple 1 :

Dans l'exemple 1 précédent portant sur les masses des baguettes le quart de l'effectif étant 30

$\frac{30}{4}=7,5$. On arrondit à l'entier supérieur 8 : Q_1 est la donnée de rang 8 soit $Q_1=239g$

et $3 \times \frac{30}{4}=22,5$ On arrondit à l'entier supérieur 23. Q_3 est la donnée de rang 23 $Q_3=251g$

Exemple 2 :

Dans l'exemple 2 précédent portant sur la durée de connexion internet le quart de l'effectif étant

$\frac{43}{4}=10,75$ On arrondit à l'entier supérieur 11. Q_1 est la donnée de rang 11 soit $Q_1 = 60$ min

et $3 \times \frac{43}{4}=32,25$ On arrondit à l'entier supérieur 33. Q_3 est la donnée de rang 33 soit $Q_3 = 180$ min

1.3. L'écart interquartile

Définition :

L'écart interquartile est égal à la différence $Q_3 - Q_1$

Exemples :

Dans l'exemple 1 :

$Q_3 - Q_1 = 251 - 239 = 12$ L'écart interquartile est 12 g

Dans l'exemple 2 :

$Q_3 - Q_1 = 180 - 60 = 120$ L'écart interquartile est 120 minutes

Remarques :

L'écart interquartile est un paramètre de **dispersion** de la série.

La médiane et les quartiles ne sont pas influencés par les valeurs extrêmes de la série. Ils dépendent seulement du nombre de valeurs de la série.

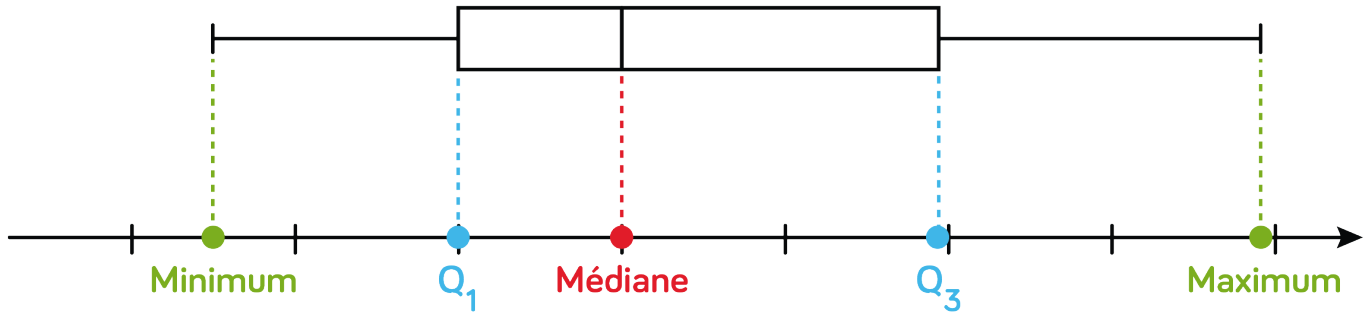
Ils forment un couple d'indicateur (position-dispersion) qui peut permettre d'analyser une série.

2. Diagramme en boite

Une série statistique peut être représentée par un diagramme appelé « boîte à moustache »

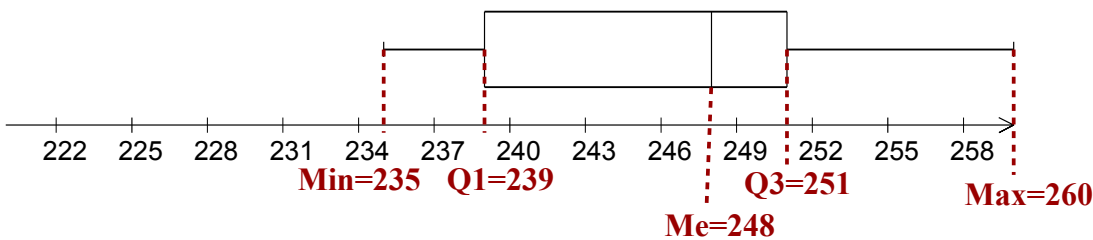
2.1. Définition .

On appelle diagramme en boîte ou boîte à moustache d'une série , la représentation graphique ci-dessous. Elle est composée de deux rectangles et de deux segments dont les longueurs correspondent aux paramètres de la série, représentés sur un axe gradué



Les boîtes à moustaches sont un moyen simple pour comparer un même caractère sur plusieurs séries statistiques.

Exemple 1 :



Exemple 2 :

3. Couple moyenne écart type

3.1. La moyenne

Définition basique de la moyenne

La moyenne d'une série statistique, notée \bar{x} , est égale à $\bar{x} = \frac{\text{somme totale des valeurs prises par le caractère}}{\text{effectif total}}$

Définition de la moyenne en notation statistique :

Une série statistique avec p valeurs de son caractères peut se résumer avec un tableau des effectifs :

x_i	x_1	x_2	x_3	x_4	...	x_p
n_i	n_1	n_2	n_3	n_4	...	n_p

La moyenne de la série vaut :

$$\bar{x} = \frac{n_1 \times x_1 + n_2 \times x_2 + n_3 \times x_3 + \dots + n_p \times x_p}{n_1 + n_2 + n_3 + \dots + n_p}$$

Une série statistique avec p valeurs de son caractères peut se résumer avec un tableau des effectifs :

x_i	x_1	x_2	x_3	x_4	...	x_p
f_i	f_1	f_2	f_3	f_4	...	f_p

Ou si on connaît les fréquences de la série à la place de effectifs : $\bar{x} = f_1 \times x_1 + f_2 \times x_2 + f_3 \times x_3 + \dots + f_p \times x_p$

Exemple 1:

Soit la série statistique répertoriant la taille en mètres de 100 requins blancs

taille(en m)	1,5	2	2,5	3	3,5	4	4,5
Effectif	8	10	25	32	19	4	2

La taille moyenne est :

$$\bar{x} = \frac{1,5 \times 8 + 2 \times 10 + 2,5 \times 25 + 3,5 \times 19 + 4 \times 4 + 4,5 \times 2}{8 + 10 + 25 + 32 + 19 + 4 + 2} = \frac{282}{100} = 2,82$$

La taille moyenne des requins blancs étudiés est de 21,82 m

Exemple 2 :

On étudie dans une maternité la taille de 50 nouveaux nés. Les résultats sont donnés en fréquence.

Taille en cm	47	48	49	50	51	52
Fréquence	0,1	0,16	0,24	0,3	0,18	0,02

On sait que $\bar{x} = f_1 \times x_1 + f_2 \times x_2 + f_3 \times x_3 + \dots + f_p \times x_p$

$$\bar{x} = 0,1 \times 47 + 0,16 \times 48 + 0,24 \times 49 + 0,3 \times 50 + 0,18 \times 51 + 0,02 \times 52 = 49,36$$

La taille moyenne des bébés de cette maternité est 49,36 cm

Exemple 3 :

Un supermarché a relevé les dépenses (en €) de ses clients en 2 heures un jour donné, les résultats sont rassemblés dans le tableau suivant :

Dépenses(en €)	[0;30[[30;60[[60;100[[100;120[
Milieu de classe	15	45	80	110
Effectif	12	25	42	67

Pour calculer la moyenne on détermine les milieux des classes de la distribution puis on effectue le calcul :

$$\bar{x} = \frac{15 \times 12 + 45 \times 25 + 80 \times 42 + 110 \times 67}{12 + 25 + 42 + 67} = 82,43$$

La dépense moyenne par client est de 82,43 € dans ce supermarché

Remarques sur la moyenne :

La moyenne est un **indicateur de position** : elle donne une tendance centrale de la série.

Exemple : Voici les notes de deux élèves :

élève 1 : 11 ; 10 ; 9 ; 10 $\bar{x}=10$

élève 2 : 2 ; 18 ; 3 ; 17 $\bar{x}=10$

Elle n'indique rien quant à la dispersion des valeurs de la série autour de la moyenne...

3.2. Écart-type

Définition :

L'écart-type d'une série statistique est un nombre σ défini par

$$\sigma = \sqrt{\frac{n_1(x_1 - \bar{x})^2 + n_2(x_2 - \bar{x})^2 + \dots + n_p(x_p - \bar{x})^2}{n_1 + n_2 + \dots + n_p}} \quad (\text{à ne pas connaître par cœur, ouf!})$$

Il mesure la **dispersion** des valeurs autour de la moyenne. On le détermine uniquement à l'aide de la calculatrice. Il possède la même unité que les valeurs de la série.

Associé à la moyenne, il peut permettre une analyse plus fine d'une série statistique.

4. Utilisation de la calculatrice :

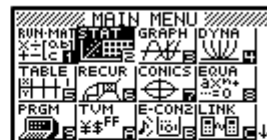
Déterminer les paramètres de la série statistique ci-contre :

Valeurs	0	2	3	5	8
Effectifs	16	12	28	32	21

Accès au mode statistique

Touche **MENU** icône  appuyer sur **EXE**.

→ Si les listes ne sont pas vides les effacer.
Voir paragraphe « Effacement des données ».



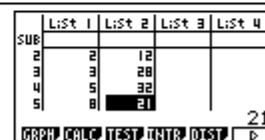
Entrée des données

Mettre les valeurs dans une liste, par exemple **List 1**.

Touche **EXE** pour passer à la ligne suivante.

Mettre les effectifs dans une autre liste, par exemple **List 2**.

→ Il est possible de se déplacer dans les listes à l'aide des flèches.



SUB	List 1	List 2	List 3	List 4
2	2	12		
3	3	28		
4	5	32		
5	8	21		

Affichage des résultats




Touche **MENU**, choisir **STAT**
Sélectionner **CALC** (Touche **F2**) puis **SET** (touche **F6**).

1 Var Xlist :List 1 (touche **F1**)

1 Var Freq :List 2 (touche **F2** et entrer 2)

Appuyer sur **EXE** puis choisir **1 Var** (touche **F1**).

On peut lire : la moyenne \bar{x}
la somme des données $\sum x$
la somme des carrés des données $\sum x^2$
l'écart type σx
l'effectif total n

Flèche  pour faire défiler la suite des résultats.

On peut lire : la valeur minimum **minX**
le 1^{er} quartile **Q1**
la médiane **Med**
le 3^{ème} quartile **Q3**
la valeur maximum **maxX**
le mode **Mod**

List 1	List 2	List 3	List 4
2	12		
3	28		
4	32		
5	21		

```
1Var XList :List1
1Var Freq  :List2
Sélectionner No:liste
List[1~26]: 2
1 LIST
```

```
1Var XList :List1
1Var Freq  :List2
2Var XList :List1
2Var Freq  :List2
2Var XList :List2
2Var Freq  :1
1 LIST
```


```
1 variable
x̄ =4
Σx =436
Σx² =2444
σx =2.53417014
sx =2.54587538
n =109
```

```
1 variable
minX =0
Q1 =2
Med =3
Q3 =5
maxX =8
Mod =5
```

```
Mod:n=1
Mod:F=32
```

Effacement des données

Instruction **QUIT** (touches **SHIFT** **EXIT**). Placer le curseur sur le nom de la liste à effacer, par exemple L2.

Taper  (touche **F6**), puis

DEL-A (touche **F4**).

Confirmer par **OUI** (touche **F1**).

SUB	List 1	List 2	List 3	List 4
1	0	16		
2	2	12		
3	3	28		
4	5	32		

SUB	List 1	List 2	List 3	List 4
1	0			
2	2			
3	3			
4	5			

(Document de l'IREM de Lyon)