

1 Rappels de seconde

Les statistiques descriptives sont un outil mathématique permettant de résumer beaucoup de données afin de pouvoir mieux les comprendre et les comparer. Par exemple, la note moyenne permet d'éviter d'avoir à regarder toutes les notes.

Dans tout le cours, on travaillera sur l'exemple donné par la série statistique suivante :

12; 12; 5; 5; 14; 8; 14; 5; 12; 12; 20; 4; 4; 5; 12; 14; 20; 17; 20; 14; 12; 12; 14; 12; 20; 4

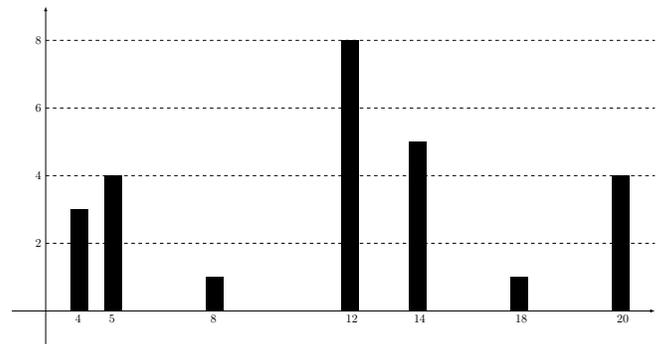
1.1 Tableaux d'effectifs

Vocabulaire :

- les nombres présents dans la série statistiques sont les *valeurs* (parfois aussi appelées *caractères*).
- Le nombre de fois où elles apparaissent est l'*effectif* associé à cette valeur.
- Le nombre de termes inférieurs ou égaux à une valeur donnée est l'*effectif cumulé croissant* associé à cette valeur.

On représente ainsi les séries statistiques souvent sous forme de tableau (et parfois d'histogramme). Ainsi, dans l'exemple :

Valeur	4	5	8	12	14	17	20
Effectifs	3	4	1	8	5	1	4
E.C.C.	3	7	8	16	21	22	26



Ici, l'effectif total est 26.

1.2 Moyenne

La moyenne (arithmétique), notée en général \bar{x} , est la valeur par laquelle on pourrait remplacer tous les termes en conservant le total. Elle se calcule par :

$$\bar{x} = \frac{\text{VALEUR} \times \text{EFFECTIF} + \dots + \text{VALEUR} \times \text{EFFECTIF}}{\text{EFFECTIF TOTAL}}$$

Exemple : Dans l'exemple, on a :

$$\bar{x} = \frac{4 \times 3 + 5 \times 4 + 8 \times 1 + 12 \times 8 + 14 \times 5 + 17 \times 1 + 20 \times 4}{26} = \frac{303}{26}$$

Remarques :

- La moyenne est forcément comprise entre le plus petit terme et le plus grand terme. Ceci permet une vérification rapide (erreur fréquente : oubli des parenthèses)
- Sauf indication contraire de l'énoncé, il faut garder la valeur exacte (sous forme de fraction simplifiée)
- La moyenne a l'inconvénient d'être potentiellement modifiée de manière forte par des exceptions très grandes ou très petites.

1.3 Médiane, quartiles

Définitions

- La *médiane* est le plus petit terme M tel que au moins la moitié des termes soient inférieurs ou égaux à M ¹
- Le *premier quartile* est le plus petit terme Q_1 tel que au moins un quart des termes soient inférieurs ou égaux à Q_1
- Le *troisième quartile* est le plus petit terme Q_3 tel que au moins trois quarts des termes soient inférieurs ou égaux à Q_3

Calcul : Il se fait à partir d'un tableau d'effectifs cumulés croissants. On calcule :

- Pour le premier quartile, $\frac{\text{EFFECTIF TOTAL}}{4}$
- Pour le troisième quartile, $\frac{3 \times \text{EFFECTIF TOTAL}}{4}$
- Pour la médiane, $\frac{\text{EFFECTIF TOTAL}}{2}$

1. Sauf, dans des cas particuliers, où la médiane est le milieu entre deux termes

Puis, on cherche dans la ligne des effectifs cumulés croissants la première case qui arrive ou dépasse le résultat du calcul.

Le premier quartile, le troisième quartile ou la médiane² sont la valeur correspondante.

Exemple :

Valeur	4	5	8	12	14	17	20
Effectifs	3	4	1	8	5	1	4
E.C.C.	3	7	8	16	21	22	26

— Pour le premier quartile, on calcule : $\frac{26}{4} = 6,5$

On cherche donc le 7^e terme. La première case qui dépasse ceci dans la ligne des E.C.C. correspond à une valeur de 5.

Ainsi, $Q_1 = 5$.

— Pour la médiane, on calcule : $\frac{26}{2} = 13$

On cherche donc le 13^e terme (en fait, le milieu entre le 13^e et le 14^e). La première case qui dépasse ceci dans la ligne des E.C.C. correspond à une valeur de 12.

Ainsi, $M = 12$.

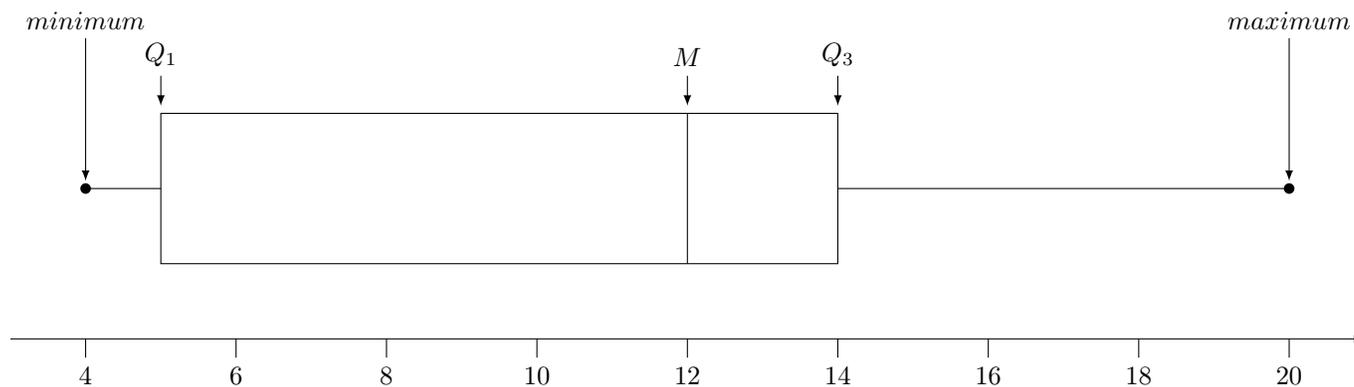
— Pour le troisième quartile, on calcule $\frac{3 \times 26}{4} = 19,5$

On cherche donc le 20^e terme. La première case qui dépasse ceci dans la ligne des E.C.C. correspond à une valeur de 14.

Ainsi, $Q_3 = 14$.

1.4 Diagramme en boîtes

Le diagramme en boîtes, aussi appelé diagramme boîte à moustaches, est la représentation, au-dessus d'une droite graduée, des valeurs calculées auparavant :

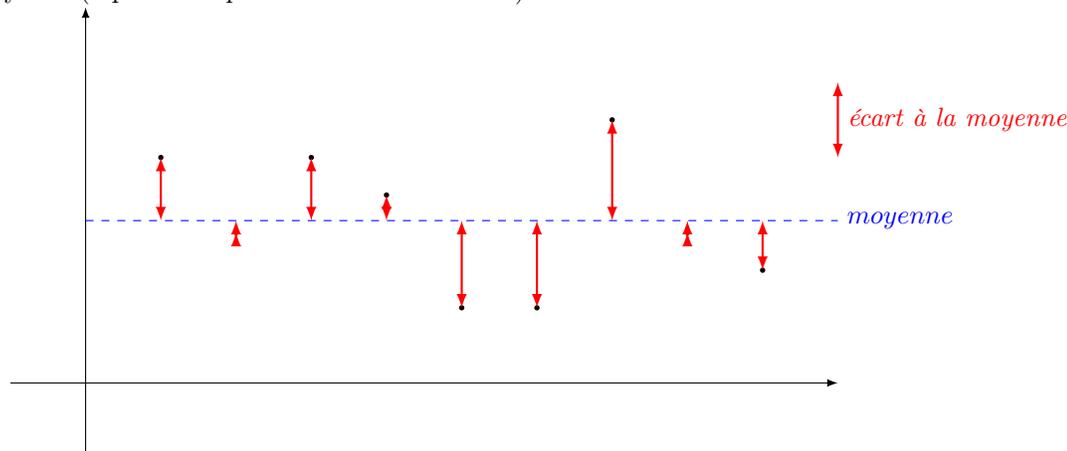


Il permet surtout de comparer plusieurs séries statistiques.

2 Variance et écart-type

2.1 Définition

Définition : on considère une série statistique. La variance, notée V , est définie comme la moyenne des carrés des écarts à la moyenne (représentés par des flèches ci-dessous)



C'est une donnée qui détermine l'étalement de la série statistique : plus elle est grande, plus la série est étalée ; plus elle est petite, plus la série est resserrée.

2. Sauf si on tombe pile, pour la médiane, auquel cas on prend le milieu entre cette valeur et la valeur suivante

L'écart-type, noté σ , est donné par :

$$\sigma = \sqrt{V}$$

Remarque : l'écart-type a la même unité que les données.

2.2 Calculs

On peut calculer la variance à partir de la définition, mais ceci donne souvent des calculs longs. On se sert, en général, de :

$$V = \overline{x^2} - \bar{x}^2$$

(la moyenne des carrés moins le carré de la moyenne)

Exemple

Valeurs x	4	5	8	12	14	17	20
x^2	16	25	64	144	196	289	400
Effectifs	3	4	1	8	5	1	4

Nous avons déjà calculé la moyenne : $\bar{x} = \frac{303}{26}$.

De même, la moyenne des carrés vaut : $\overline{x^2} = \frac{16 \times 3 + 25 \times 4 + 64 \times 1 + 144 \times 8 + 196 \times 5 + 289 \times 1 + 400 \times 4}{26} = \frac{4233}{26}$

Ainsi, la variance vaut :

$$\begin{aligned} V &= \overline{x^2} - \bar{x}^2 \\ &= \frac{4233}{26} - \left(\frac{303}{26}\right)^2 \\ &= \frac{18249}{676} \\ &\approx 27 \end{aligned}$$

L'écart-type vaut donc :

$$\begin{aligned} \sigma &= \sqrt{V} \\ &= \sqrt{\frac{18249}{676}} \\ &\approx 5,2 \end{aligned}$$

Remarque : L'écart-type sera toujours plus petit que l'étendue (maximum - minimum), ce qui donne une première vérification.

3 Calculatrice (TI)

3.1 Rentrer les données

Pour rentrer les données dans la calculatrice, on doit se rendre dans l'écran des listes : **stats**, puis 1 : Edit. On obtient l'écran ci-dessous.

L1	L2	L3
-----	-----	-----

L1(1) =

On rentre, sous L1, les valeurs ; puis sous L2 les effectifs.

Ainsi, pour l'exemple qu'on a eu tout le cours, on aura :

L1	L2	L3
4	3	-----
5	4	-----
8	1	-----
12	8	-----
14	5	-----
17	1	-----
20	4	-----

L3(1) =

3.2 Stats-1Var

- On revient à l'écran de calcul en faisant **2nde** **mode** (Quitter).
- On appuie sur **stats**, puis, dans le menu CALC, on choisit 1 : Stats-1Var.
- On indique quelles listes considérer en rentrant L_1 et L_2 : **2nde** **1** puis **,** puis **2nde** **2**.

— La calculatrice affiche :

(Utiliser les flèches pour se déplacer).

Le tableau ci-dessous explique les données affichées :

\bar{x}	Moyenne
Σx	Somme de tous les termes
Σx^2	Somme des carrés des termes
Sx	A oublier
σx	Ecart-type
n	Effectif total
$\min X$	Minimum (valeur la plus petite)
Q_1	Version spéciale du premier quartile
Med	Médiane
Q_3	Version spéciale du troisième quartile
$\max X$	Maximum (valeur la plus grande)

Remarques :

- Les quartiles sont calculés d'une manière différente de la méthode du cours. En général, la calculatrice affichera des valeurs plus proches de la médiane. *Il ne faut donc pas se servir de ces informations !*
- Σx et Σx^2 sont utiles pour calculer les valeurs exactes ! En effet, il suffit de les diviser par l'effectif total pour obtenir \bar{x} et $\overline{x^2}$.