BURHAN

Statistique

3ème Année Collège · Mathématiques

Version en ligne : burhan.ma/3ac/mathematiques/statistique/cours

Chapitre	Semestre
Statistique	Semestre 2

Pour commencer

Voici une situation fictive. Dans une ligue imaginaire de handball, deux équipes ont joué cinq matchs.

Les Faucons ont marqué 21, 24, 22, 25 et 23 buts. Les Lynx ont marqué 12, 38, 15, 36 et 14 buts.

Devine : Qui marque le plus ?

Sans calculer, fais une prédiction. Quelle équipe marque le plus de buts par match, en moyenne ?

A. Les Faucons, car ils marquent toujours plus de 20 buts.

Faux.

Faux. Les Faucons sont réguliers, mais régulier ne veut pas dire meilleur en moyenne. Leur total est de 115 buts, comme celui des Lynx.

B. Les Lynx, car ils ont réussi deux très gros scores.

Faux.

Faux. Les Lynx ont aussi trois petits scores, qui compensent les deux gros. Leur total est de 115 buts, comme celui des Faucons.

C. Les deux équipes, à égalité.

Juste.

Juste. Chaque équipe a marqué 115 buts en cinq matchs, soit 23 buts par match en moyenne.

Correction

Total des Faucons, en buts :

$$\begin{aligned} &21 + 24 + 22 + 25 + 23 \\ &= 115 \end{aligned}$$

Total des Lynx, en buts :

$$\begin{aligned} &12 + 38 + 15 + 36 + 14 \\ &= 115 \end{aligned}$$

Chaque équipe a donc marqué 115 buts en cinq matchs. En moyenne, chacune marque $\frac{115}{5} = 23$ buts par match.

Pourtant, les Faucons et les Lynx ne jouent pas du tout de la même façon. Pour décrire ces différences, tu vas apprendre à résumer une série par plusieurs nombres. Tu retrouveras ces deux équipes à la fin du chapitre.

Rappels : effectifs, fréquences, effectifs cumulés

Effectif et fréquence

Rappel de 1re année : on étudie un **caractère**, comme la pointure, sur une **population**, comme les paires vendues par un magasin. Chaque élément de la population est un **individu**.

Chaque individu prend une **valeur** du caractère. La liste des valeurs relevées est une **série statistique**. Le nombre d'individus qui prennent une valeur s'appelle l'**effectif** de cette valeur.

À toi : Une fréquence

Un magasin imaginaire a noté les pointures des paires de chaussures vendues un samedi.

Pointure	37	38	39	40	41
Effectif	6	10	12	8	4

Rappel de 1re année : la **fréquence** d'une valeur est la part de la population qui la prend. On l'obtient en divisant son effectif par le nombre total d'individus. Quelle est la fréquence de la pointure 38 ?

A. $\frac{10}{40}$, soit $0{,}25$
Juste.
Juste. L'effectif total est $6 + 10 + 12 + 8 + 4 = 40$ paires. La fréquence est l'effectif de la valeur divisé par ce total.

B. $\frac{10}{5}$, soit 2
Faux.
Faux. Tu as divisé par le nombre de pointures différentes. Une fréquence ne dépasse jamais 1 : divise par le nombre total de paires.

C. $\frac{10}{100}$, soit $0{,}1$
Faux.
Faux. On divise par l'effectif total, ici 40, et non par 100. Le nombre 100 sert seulement à passer au pourcentage.

D. $\frac{40}{10}$, soit 4
Faux.
Faux. Le quotient est à l'envers : l'effectif de la valeur se place au numérateur.

Correction

L'effectif total est $6 + 10 + 12 + 8 + 4 = 40$ paires.

La fréquence de la pointure 38 est donc $\frac{10}{40} = 0{,}25$, soit 25 pour cent.

Définition : Effectif et fréquence

On étudie un **caractère** sur une **population** : chaque individu de la population prend une **valeur** du caractère. L'ensemble des valeurs relevées forme une **série statistique**.

- L'**effectif** d'une valeur est le nombre d'individus qui prennent cette valeur.
- L'**effectif total** N est le nombre d'individus de la population : c'est la somme des effectifs de toutes les valeurs.
- La **fréquence** d'une valeur d'effectif n est le quotient $f = \frac{n}{N}$.

Une fréquence est un nombre compris entre 0 et 1. On l'écrit sous forme de fraction, de nombre décimal ou de pourcentage : la fréquence $\frac{1}{4} = 0{,}25$ s'écrit aussi 25 pour cent. La somme des fréquences de toutes les valeurs est égale à 1.

En clair : la fréquence dit quelle **part** de la population prend une valeur. Une fréquence de $0{,}5$ veut dire : la moitié de la population.

Le tableau complet

Exemple : Le tableau des pointures

On complète le tableau du magasin imaginaire, avec $N = 40$.

Pointure	37	38	39	40	41
Effectif	6	10	12	8	4
Fréquence (fraction)	$\frac{6}{40}$	$\frac{10}{40}$	$\frac{12}{40}$	$\frac{8}{40}$	$\frac{4}{40}$
Fréquence (décimal)	$0{,}15$	$0{,}25$	$0{,}3$	$0{,}2$	$0{,}1$
Pour cent	15	25	30	20	10

Pour le pourcentage, on multiplie la fréquence par 100.

Contrôle : les cinq fréquences ont bien pour somme 1.

$$0{,}15 + 0{,}25 + 0{,}3 + 0{,}2 + 0{,}1 = 1$$

Trouve l'erreur : Un tableau à vérifier

Un parking imaginaire a compté ses voitures pendant 25 jours. Un tableau donne les fréquences $0{,}2$; $0{,}36$; $0{,}28$ et $0{,}12$ pour les quatre valeurs relevées. Ce tableau est-il juste ?

A. Oui, car chaque fréquence est comprise entre 0 et 1.

Faux.

Faux. C'est nécessaire, mais pas suffisant : il faut aussi contrôler la somme des fréquences.

B. Non, car la somme des fréquences ne vaut pas 1.

Juste.

Juste. $0{,}2 + 0{,}36 + 0{,}28 + 0{,}12 = 0{,}96$, et non 1. Une fréquence est fausse, ou un effectif a été oublié.

C. On ne peut rien dire sans les effectifs.

Faux.

Faux. Le contrôle de la somme se fait sur les fréquences seules : leur somme doit valoir 1.

Correction

La somme des fréquences vaut :

$$\begin{aligned} &0{,}2 + 0{,}36 + 0{,}28 + 0{,}12 \\ &= 0{,}96 \end{aligned}$$

Elle devrait valoir 1. En effectifs, cela fait $25 \times 0{,}96 = 24$ jours seulement sur 25.

Le tableau est donc faux.

Attention : Diviser par l'effectif total

Une fréquence se calcule avec l'**effectif total**, pas avec le nombre de valeurs différentes. Série fictive : 50 relevés, cinq valeurs différentes, une valeur d'effectif 20.

✗ $\dfrac{20}{5} = 4$: impossible, une fréquence ne dépasse pas 1.

✓ $\dfrac{20}{50} = 0{,}4$, soit 40 pour cent.

Les effectifs cumulés

Rappel de 2e année : quand les valeurs sont des nombres, on les range dans l'ordre croissant. On peut alors **cumuler** les effectifs, c'est-à-dire les ajouter les uns aux autres.

Devine : Au plus 39

Reprends le tableau des pointures du magasin imaginaire. Combien de paires vendues ont une pointure inférieure ou égale à 39 ?

A. 12

Faux.

Faux. 12 est l'effectif de la pointure 39 seule. Les pointures 37 et 38 sont aussi inférieures à 39.

B. 22

Faux.

Faux. Tu as ajouté 10 et 12, mais oublié les 6 paires de pointure 37.

C. 28

Juste.

Juste. On ajoute les effectifs de 37, 38 et 39 : $6 + 10 + 12 = 28$.

D. 24

Faux.

Faux. $12 + 8 + 4 = 24$ compte les pointures supérieures ou égales à 39.

Correction

Les pointures inférieures ou égales à 39 sont 37, 38 et 39. On ajoute leurs effectifs :
$6 + 10 + 12 = 28$ paires.

Définition : Effectifs cumulés et fréquences cumulées

Les valeurs du caractère sont des nombres, rangés dans l'ordre croissant.

- L'**effectif cumulé croissant** d'une valeur est la somme de son effectif et des effectifs de toutes les valeurs plus petites qu'elle.
- La **fréquence cumulée croissante** d'une valeur est la somme de sa fréquence et des fréquences de toutes les valeurs plus petites qu'elle.

L'effectif cumulé croissant d'une valeur est le nombre d'individus dont la valeur est inférieure ou égale à elle. Celui de la plus grande valeur est égal à l'effectif total N.

Exemple : Les cumuls des pointures

Pointure	37	38	39	40	41
Effectif	6	10	12	8	4
Effectif cumulé croissant	6	16	28	36	40
Fréquence cumulée croissante	$0{,}15$	$0{,}4$	$0{,}7$	$0{,}9$	1

Chaque cumul est le précédent augmenté de l'effectif : $16 = 6 + 10$, puis $28 = 16 + 12$. Le dernier cumul est 40, l'effectif total.

Lecture : 36 paires ont une pointure inférieure ou égale à 40.

Méthode et entraînement

Méthode : Remplir un tableau statistique

1. **Additionner** tous les effectifs : on obtient l'effectif total N.
2. **Diviser** chaque effectif par N : on obtient les fréquences. Multiplie par 100 pour avoir un pourcentage.
3. **Vérifier** que la somme des fréquences fait 1.
4. **Cumuler** les effectifs dans l'ordre croissant des valeurs. Le dernier cumul doit être N.

À toi : Une bibliothèque imaginaire

Une bibliothèque imaginaire a noté chaque jour le nombre de dictionnaires empruntés.

Dictionnaires empruntés	0	1	2	3
Nombre de jours	3	7	6	4

Donne la fréquence de la valeur 2 en pourcentage, puis l'effectif cumulé croissant de la valeur 2.

Indice 1
L'outil : l'effectif total N, puis le quotient de l'effectif par N.
Indice 2
Le calcul : $N = 3 + 7 + 6 + 4 = 20$, et la fréquence de la valeur 2 est $\frac{6}{20}$.
Indice 3
Point délicat : l'effectif cumulé de la valeur 2 compte aussi les jours à 0 et à 1 dictionnaire.

Réponse
Effectif total. On additionne les effectifs : $N = 3 + 7 + 6 + 4 = 20$ jours.

Fréquence. La fréquence de la valeur 2 est $\frac{6}{20} = 0{,}3$, soit 30 pour cent.

Cumul. L'effectif cumulé croissant de la valeur 2 ajoute les jours à 0, 1 et 2 dictionnaires : $3 + 7 + 6 = 16$.

Conclusion. La valeur 2 a pour fréquence 30 pour cent, et pendant 16 jours, au plus 2 dictionnaires ont été empruntés.

La moyenne arithmétique

Découvrir la moyenne

À toi : Quatre jours de course

Une coureuse imaginaire a parcouru 5 km, 7 km, 6 km et 10 km en quatre jours. Quelle distance a-t-elle parcourue en moyenne par jour ?

A. $7{,}5$ km

Faux.

Faux. $7{,}5$ est à mi-chemin entre 5 et 10, la plus petite et la plus grande distance. Les distances 6 et 7 comptent aussi.

B. 28 km

Faux.

Faux. 28 km est la distance totale. Il reste à la partager entre les quatre jours.

C. 7 km

Juste.

Juste. La distance totale est 28 km, et $\frac{28}{4} = 7$.

D. $6{,}5$ km

Faux.

Faux. $6{,}5$ est le milieu de 6 et 7, les deux distances du milieu une fois rangées. Ce n'est pas la moyenne.

Correction

La distance totale est $5 + 7 + 6 + 10 = 28$ km. On la partage en parts égales entre les quatre jours :

$$\frac{28}{4} = 7$$

En moyenne, la coureuse a parcouru 7 km par jour.

Rappel de 2e année : la **moyenne arithmétique** partage le total en parts égales entre tous les individus. Quand une valeur se répète, il faut en tenir compte.

Devine : Dix jours à vélo

Un cycliste imaginaire a roulé 20 km par jour pendant 7 jours, puis 30 km par jour pendant 3 jours. Quelle est sa distance moyenne par jour ?

A. 25 km

Faux.

Faux. $\frac{20 + 30}{2} = 25$ compte chaque distance une seule fois. Or 20 km revient 7 jours sur 10.

B. 23 km

Juste.

Juste. Le total est $7 \times 20 + 3 \times 30 = 230$ km, sur 10 jours.

C. 230 km

Faux.

Faux. 230 km est la distance totale des 10 jours. Il reste à diviser par 10.

D. 115 km

Faux.
Faux. Tu as divisé 230 par 2, le nombre de distances différentes. On divise par le nombre de jours, 10.

Correction
Distance totale. Sept jours à 20 km, puis trois jours à 30 km :

$$\begin{aligned} &7 \times 20 + 3 \times 30 \\ &= 140 + 90 \\ &= 230 \end{aligned}$$

Nombre de jours. Le cycliste a roulé $7 + 3 = 10$ jours.

Conclusion. Sa distance moyenne est $\dfrac{230}{10} = 23$ km par jour.

Définition et méthode

Définition : Moyenne arithmétique

Une série statistique a pour valeurs x_1, x_2, ..., x_p, d'effectifs respectifs n_1, n_2, ..., n_p. Son effectif total est $N = n_1 + n_2 + \cdots + n_p$. La **moyenne arithmétique** de la série est le nombre

$$m = \frac{n_1 \times x_1 + n_2 \times x_2 + \cdots + n_p \times x_p}{N}.$$

Quand chaque valeur n'est relevée qu'une fois, la moyenne est la somme des valeurs divisée par leur nombre.

En clair : une valeur d'effectif 3 entre trois fois dans le total. On additionne les produits « effectif $\times$ valeur », puis on divise par l'effectif total N.

Exemple : La ligne des produits

Une équipe imaginaire de football a joué 20 matchs. Voici le nombre de buts marqués par match.

Buts	0	1	2	3	4	Total
Effectif	3	6	5	4	2	20
Effectif $\times$ valeur	0	6	10	12	8	36

La somme des produits vaut 36, et $N = 20$. La moyenne est donc :

$$m = \frac{36}{20} = 1{,}8$$

L'équipe marque en moyenne $1{,}8$ but par match. Contrôle : $1{,}8$ est bien entre 0 et 4.

Méthode : Calculer une moyenne à la main

1. **Multiplier** chaque valeur par son effectif.
2. **Additionner** tous ces produits.
3. **Diviser** cette somme par l'effectif total N, et non par le nombre de valeurs différentes.
4. **Contrôler** : le résultat doit rester compris entre la plus petite et la plus grande valeur de la série.

Attention : Les effectifs comptent

Une valeur répétée compte plusieurs fois. Série fictive : la valeur 10 a l'effectif 4, la valeur 25 a l'effectif 1.

✗ $\dfrac{10+25}{2} = 17{,}5$.

✓ Chaque valeur est multipliée par son effectif :

$$m = \frac{4\times 10 + 1\times 25}{5} = \frac{65}{5} = 13$$

Avec les fréquences

Reprends l'équipe imaginaire de football. Les fréquences des valeurs 0, 1, 2, 3 et 4 sont $0{,}15$; $0{,}3$; $0{,}25$; $0{,}2$ et $0{,}1$. Multiplie chaque valeur par sa fréquence, puis additionne :

$$\begin{aligned} &0{,}15\times 0 + 0{,}3\times 1 \\ &\quad + 0{,}25\times 2 + 0{,}2\times 3 \\ &\quad + 0{,}1\times 4 = 1{,}8. \end{aligned}$$

C'est exactement la moyenne.

Pourquoi est-ce vrai ? Ce n'est pas un hasard propre à cette équipe. La propriété suivante l'affirme pour toute série, et sa preuve est courte.

Propriété : Moyenne et fréquences

Avec les notations de la moyenne arithmétique, notons f_1, f_2, …, f_p les fréquences des valeurs x_1, x_2, …, x_p. Alors

$$m = f_1\times x_1 + f_2\times x_2 + \cdots + f_p\times x_p.$$

Démonstration

Étape 1. D'après la définition de la moyenne arithmétique, $m = \dfrac{n_1\times x_1 + n_2\times x_2 + \cdots + n_p\times x_p}{N}$.

Étape 2. Diviser une somme par N, c'est diviser chacun de ses termes par N. Donc $m = \frac{n_1 \times x_1}{N} + \frac{n_2 \times x_2}{N} + \cdots + \frac{n_p \times x_p}{N}$.

Étape 3. Diviser un produit par N, c'est diviser un seul de ses facteurs par N : $\frac{n_1 \times x_1}{N} = \frac{n_1}{N} \times x_1$. Or $\frac{n_1}{N} = f_1$, par la définition de la fréquence. Donc $\frac{n_1 \times x_1}{N} = f_1 \times x_1$. Il en va de même pour les autres termes.

Étape 4. Donc $m = f_1 \times x_1 + f_2 \times x_2 + \cdots + f_p \times x_p$.

À retenir : Ce qu'il faut retenir de la preuve

Toute l'idée tient en deux gestes. On partage la division par N entre tous les termes de la somme. Dans chaque terme, seul l'effectif est divisé par N : le quotient $\frac{n_1}{N}$ est la fréquence f_1.

Vérifie : Le cœur de la preuve

Dans la preuve, quel argument fait passer des effectifs aux fréquences ?

A. On multiplie la somme par N, puis on simplifie chaque terme.

Faux.

Faux. La preuve ne multiplie jamais par N : elle part d'une division par N.

B. La somme des fréquences vaut 1, donc N disparaît du calcul.

Faux.

Faux. La somme des fréquences vaut bien 1, mais la preuve n'utilise pas ce fait.

C. On remplace N par le nombre de valeurs différentes.

Faux.

Faux. La preuve garde partout le même effectif total N. Diviser par le nombre de valeurs différentes est une erreur de calcul de moyenne.

D. On divise chaque terme par N, en ne divisant que son effectif : $\frac{n_1}{N} = f_1$.

Juste.

Juste. Une somme divisée par N : chacun de ses termes est divisé par N. Un produit divisé par N : un seul de ses facteurs l'est, ici l'effectif. Ainsi $\frac{n_1 \times x_1}{N} = \frac{n_1}{N} \times x_1 = f_1 \times x_1$.

Correction

La preuve part de la définition de m. Elle divise chaque terme de la somme par N.

Dans le terme $n_1 \times x_1$, elle ne divise qu'un facteur, l'effectif n_1 : $\frac{n_1 \times x_1}{N} = \frac{n_1}{N} \times x_1$.

Enfin, $\frac{n_1}{N} = f_1$, par la définition de la fréquence, et de même pour chaque terme.

S'entraîner et rédiger

Exemple : Une moyenne avec les fréquences, et son hypothèse

Une tirelire fictive contient 20 pièces : 10 pièces de 1 dirham, 6 de 2 dirhams et 4 de 5 dirhams. Avec $N = 20$, les fréquences sont $0{,}5$; $0{,}3$ et $0{,}2$. La valeur moyenne d'une pièce est

$$\begin{aligned} m &= 0{,}5 \times 1 + 0{,}3 \times 2 \\ &\quad + 0{,}2 \times 5 \\ &= 0{,}5 + 0{,}6 + 1 = 2{,}1. \end{aligned}$$

Une pièce vaut donc $2{,}1$ dirhams en moyenne.

Contrôle avec les effectifs :

$$\begin{aligned} &\frac{10 \times 1 + 6 \times 2 + 4 \times 5}{20} \\ &= \frac{42}{20} = 2{,}1. \end{aligned}$$

L'hypothèse compte : toutes les fréquences doivent être calculées avec le même effectif total N. Calculée parmi les 10 pièces de 2 et de 5 dirhams seulement, la fréquence des pièces de 5 dirhams devient $\frac{4}{10} = 0{,}4$. Le calcul donne alors $0{,}5 \times 1 + 0{,}3 \times 2 + 0{,}4 \times 5 = 3{,}1$.

Ce résultat est faux : la moyenne vaut $2{,}1$. Signe d'alerte : ces fréquences ont pour somme $0{,}5 + 0{,}3 + 0{,}4 = 1{,}2$, et non 1.

À toi : Dix jours de météo

Dans une ville imaginaire, on a relevé la température maximale pendant 10 jours : 18 degrés pendant 2 jours, 20 degrés pendant 5 jours, 23 degrés pendant 3 jours. Calcule la température moyenne avec les effectifs, puis avec les fréquences.

Indice 1
L'outil : la ligne des produits « effectif $\times$ valeur », puis la division par N.
Indice 2
Le calcul : $2 \times 18 + 5 \times 20 + 3 \times 23 = 36 + 100 + 69 = 205$.
Indice 3
Point délicat : divise par $N = 10$ jours, et non par les 3 valeurs. Les fréquences sont $\frac{2}{10}$, $\frac{5}{10}$ et $\frac{3}{10}$.

Réponse
Avec les effectifs. La somme des produits « effectif $\times$ valeur » vaut 205, et $N = 10$:

$$m = \frac{205}{10} = 20{,}5$$

Avec les fréquences. Les fréquences sont $0{,}2$; $0{,}5$ et $0{,}3$:

$$\begin{aligned} m &= 0{,}2 \times 18 \\ &\quad + 0{,}5 \times 20 \\ &\quad + 0{,}3 \times 23 \\ &= 3{,}6 + 10 + 6{,}9 \\ &= 20{,}5 \end{aligned}$$

Conclusion. La température moyenne est de $20{,}5$ degrés. Les deux calculs donnent le même résultat, comme l'annonce la propriété.

Remarque : Coin examen : rédiger le calcul d'une moyenne

Niveau examen régional : une rédaction modèle écrit chaque étape du calcul, puis conclut par une phrase avec l'unité. Exemple fictif : à un guichet imaginaire, 8 personnes ont attendu 5 minutes, 7 personnes 10 minutes et 5 personnes 15 minutes.

« L'effectif total est $N = 8 + 7 + 5 = 20$. La moyenne est

$$\begin{aligned} m &= \frac{8 \times 5 + 7 \times 10 + 5 \times 15}{20} \\ &= \frac{185}{20} = 9{,}25. \end{aligned}$$

La durée d'attente moyenne est de $9{,}25$ minutes. »

Le mode et la valeur médiane

Le mode

Devine : Le nombre de pains le plus courant

Une boulangerie imaginaire a noté combien de pains chaque client a achetés un matin.

Pains achetés	1	2	3	4
Nombre de clients	9	14	5	2

Quel nombre de pains est le plus courant ?

A. 14

Faux.

Faux. 14 est un nombre de clients, c'est-à-dire un effectif. La question porte sur un nombre de pains.

B. 2

Juste.

Juste. 14 clients sur 30 ont acheté 2 pains : c'est la valeur de plus grand effectif.

C. 4

Faux.

Faux. 4 est le plus grand nombre de pains acheté, mais seulement par 2 clients.

Correction
La valeur la plus courante est celle qui a le plus grand effectif. Le plus grand effectif du tableau est 14 ; il correspond à 2 pains. La réponse est donc 2 pains.

Définition : Mode

Le **mode** d'une série statistique est une valeur du caractère qui a le plus grand effectif. Une série peut avoir plusieurs modes : c'est le cas quand plusieurs valeurs ont ce plus grand effectif.

Exemple : Deux modes

Un magasin imaginaire a vendu des vélos de quatre couleurs : 7 rouges, 7 bleus, 4 noirs et 2 verts. Le plus grand effectif est 7, atteint deux fois. Cette série a donc deux modes : rouge et bleu.

Ici, le caractère n'est pas un nombre, et le mode a quand même un sens.

Attention : Une valeur, pas un effectif

Le mode est une **valeur** du caractère. Série fictive : des durées de trajet, en minutes ; la durée 45 a le plus grand effectif, 9.

✗ « Le mode est 9. »

✓ « Le mode est 45 minutes. »

Le mode n'est pas non plus forcément la plus grande valeur.

Le milieu d'une série

À toi : Sept trajets

Voici sept durées fictives de trajet, en minutes, rangées dans l'ordre croissant : 12, 15, 15, 18, 20, 26, 40. Quelle durée partage cette liste en deux groupes de même effectif ?

A. 26

Faux.
Faux. 26 est à mi-chemin entre 12 et 40, la durée la plus courte et la plus longue. Il y a 5 durées avant 26 et une seule après.

B. 15

Faux.
Faux. 15 est la durée qui revient le plus, c'est-à-dire le mode. Elle ne coupe pas la liste en deux.

C. 18

Juste.

Juste. 18 est la quatrième durée : trois durées avant, trois après.

Correction

La liste compte 7 durées. La quatrième, 18, laisse trois durées avant elle (12, 15, 15) et trois après (20, 26, 40).

Devine : Huit trajets

On ajoute une huitième durée, 22 minutes. La liste rangée devient 12, 15, 15, 18, 20, 22, 26, 40.

Le **rang** d'une durée est sa place dans cette liste rangée : rang 1 pour la plus courte, rang 8 pour la plus longue. Quelle valeur partage maintenant la liste en deux groupes de quatre ?

A. 19

Juste.

Juste. Les deux durées du milieu sont 18 et 20 ; on prend leur moyenne, $\frac{18+20}{2} = 19$.

B. 18

Faux.

Faux. 18 est la quatrième durée : il reste quatre durées après elle, mais seulement trois avant.

C. 20

Faux.

Faux. 20 est la cinquième durée : quatre durées avant elle, mais trois seulement après.

D. Aucune : une liste de huit valeurs n'a pas de milieu.

Faux.

Faux. Aucune durée de la liste ne convient, mais un nombre placé entre 18 et 20 partage bien la liste en deux groupes de quatre.

Correction

Avec 8 durées, le milieu tombe entre la durée de rang 4, 18, et celle de rang 5, 20. On convient de prendre leur moyenne :

$$\frac{18+20}{2} = 19$$

Quatre durées sont inférieures à 19, quatre lui sont supérieures.

Définition : Valeur médiane

On range les N valeurs d'une série statistique numérique dans l'ordre croissant, chaque valeur étant écrite autant de fois que son effectif. La **valeur médiane** (on dit aussi la **médiane**) partage cette liste en deux groupes de même effectif :

- si N est impair, la médiane est la valeur de rang $\frac{N+1}{2}$;
- si N est pair, la médiane est la moyenne des deux valeurs de rangs $\frac{N}{2}$ et $\frac{N}{2}+1$.

Ainsi, au moins la moitié des valeurs de la série sont inférieures ou égales à la médiane, et au moins la moitié sont supérieures ou égales à elle.

En clair : on range toutes les valeurs, puis on cherche le milieu de la liste. Avec un nombre impair de valeurs, une valeur est au milieu. Avec un nombre pair, le milieu tombe entre deux valeurs : on prend leur moyenne.

Exemple : Un effectif total impair

Une station météo imaginaire a relevé la vitesse du vent, en km/h, pendant 9 jours : 14, 8, 20, 11, 8, 25, 16, 11, 9.

Liste rangée : 8, 8, 9, 11, 11, 14, 16, 20, 25.

Ici $N = 9$ est impair : le rang du milieu est $\frac{9+1}{2} = 5$. La cinquième valeur est 11 : la médiane vaut 11 km/h.

Effectif pair et tableau

Exemple : Un effectif total pair

Une athlète imaginaire a réussi 6 sauts, en centimètres : 412, 398, 405, 420, 401, 415.

Liste rangée : 398, 401, 405, 412, 415, 420.

Ici $N = 6$ est pair : on prend les valeurs de rangs 3 et 4, soit 405 et 412. La médiane est leur moyenne :

$$\frac{405+412}{2} = 408{,}5$$

La médiane des sauts vaut $408{,}5$ cm.

Quand la série est donnée par un tableau, écrire toute la liste devient long. Les effectifs cumulés croissants donnent directement la place de chaque valeur dans la liste rangée.

Méthode : Trouver la médiane avec un tableau

1. **Calculer** N et les effectifs cumulés croissants.
2. **Situer** le milieu : un rang, $\frac{N+1}{2}$, quand N est impair ; deux rangs, $\frac{N}{2}$ et $\frac{N}{2}+1$, quand N est pair.

3. **Lire** la valeur de chaque rang : la première valeur dont l'effectif cumulé atteint ou dépasse ce rang.
4. **Conclure** : cette valeur, ou la moyenne des deux valeurs lues.

Exemple : Un club de lecture imaginaire

Un club de lecture imaginaire compte 25 membres. On a noté le nombre de livres lus par chacun en un mois.

Livres lus	0	1	2	3	4
Effectif	2	9	5	6	3
Effectif cumulé croissant	2	11	16	22	25

Ici $N = 25$ est impair : le rang du milieu est $\frac{25+1}{2} = 13$.

Le cumul 11 n'atteint pas ce rang. Le premier cumul qui atteint 13 est 16, celui de la valeur 2.

La médiane vaut 2 livres, alors que le mode vaut 1 livre.

Vérifier et s'entraîner

Vérifie : Un membre de plus

Un nouveau membre rejoint le club de lecture imaginaire. Il a lu 4 livres ce mois-ci. Que devient la médiane ?

A. Elle reste égale à 2.

Juste.

Juste. $N = 26$: rangs 13 et 14. Le cumul de la valeur 2 est 16 : les deux rangs portent la valeur 2 .

B. Elle passe à $2{,}5$.

Faux.

Faux. Le milieu tombe bien entre les rangs 13 et 14, mais ces deux rangs portent la même valeur, 2.

C. Elle passe à 3.

Faux.

Faux. Une seule valeur ajoutée en haut de la liste ne déplace le milieu que d'un demi-rang.

Correction

Avec le nouveau membre, $N = 26$ est pair : on cherche les rangs 13 et 14.

Les cumuls deviennent $2, 11, 16, 22, 26$. Les rangs 12 à 16 portent la valeur 2, donc les rangs 13 et 14 aussi.

La médiane reste $\frac{2+2}{2} = 2$.

Attention : Le milieu de la liste, pas du tableau

La médiane se cherche dans la liste rangée de **toutes** les valeurs. Série fictive : valeurs $2, 4, 6$, d'effectifs $1, 1, 7$.

✗ « La médiane est 4, la valeur du milieu du tableau. »

✓ $N = 9$, rang du milieu 5. Les cumuls sont $1, 2, 9$: la médiane est 6.

À toi : Des mois de pluie

Une ville imaginaire a compté les jours de pluie de chaque mois, pendant 20 mois.

Jours de pluie	2	4	5	7	9
Nombre de mois	3	7	4	4	2

Donne le mode, puis la médiane de cette série.

Indice 1
L'outil : le mode se lit sur le plus grand effectif ; la médiane, sur les effectifs cumulés croissants.
Indice 2
Le calcul : les cumuls sont $3, 10, 14, 18, 20$; $N = 20$ est pair, donc on cherche les rangs 10 et 11.
Indice 3
Point délicat : le rang 10 et le rang 11 ne portent pas la même valeur.

Réponse
Mode. Le plus grand effectif est 7 : le mode est 4 jours de pluie.

Rangs. Ici $N = 20$ est pair : on cherche les rangs 10 et 11. Le cumul de la valeur 4 est 10 : le rang 10 porte la valeur 4. Le rang 11 porte la valeur suivante, 5.

Médiane. On prend la moyenne de ces deux valeurs :

$$\frac{4+5}{2} = 4{,}5$$

Conclusion. Le mode vaut 4 jours de pluie et la médiane $4{,}5$ jours.

Diagramme en bâtons et diagramme circulaire

Le diagramme en bâtons

Devine : La hauteur d'un bâton

Un cinéma imaginaire a noté, pendant 24 jours, le nombre de séances complètes par jour.

Séances complètes	0	1	2	3	4
Nombre de jours	3	5	8	6	2

On va tracer un bâton au-dessus de chaque valeur. De quoi dépendra la hauteur d'un bâton ?

A. De la valeur elle-même : plus elle est grande, plus le bâton est haut.

Faux.

Faux. Le bâton de la valeur 4 serait alors le plus haut. Or cette valeur n'a que 2 jours : la hauteur ne suit pas la valeur.

B. De la place de la valeur dans le tableau : le premier bâton est le plus haut.

Faux.

Faux. L'ordre des valeurs ne change pas les hauteurs. La valeur 0, placée en premier, n'a que 3 jours.

C. De l'effectif de la valeur : plus il est grand, plus le bâton est haut.

Juste.

Juste. Le bâton le plus haut sera donc celui de la valeur 2, d'effectif 8.

Correction

La hauteur d'un bâton suit l'effectif de sa valeur. Le plus grand effectif est 8, celui de la valeur 2 : son bâton sera le plus haut. Vérifie-le sur le diagramme qui suit.

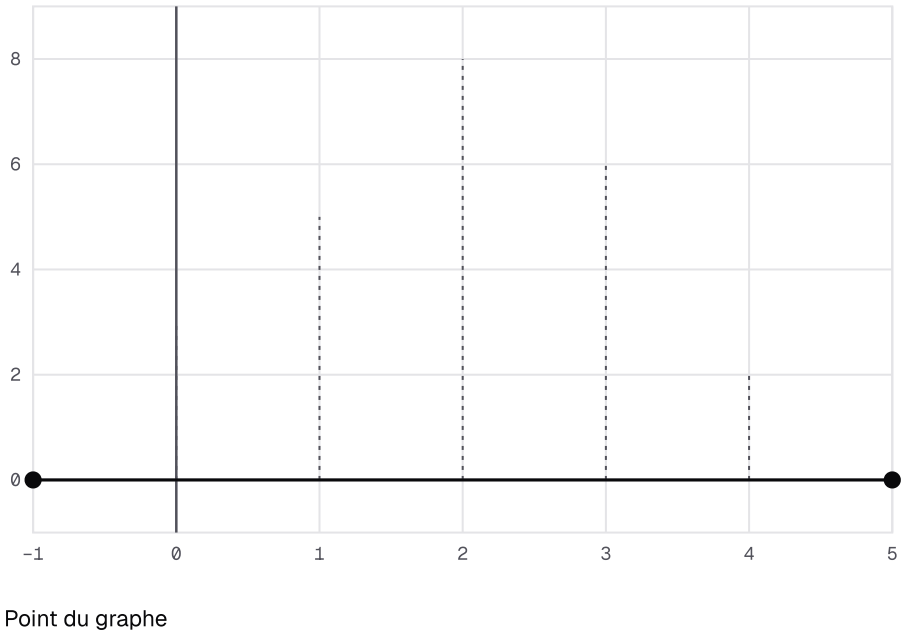


Diagramme en bâtons des séances complètes du cinéma imaginaire, sur 24 jours.

Définition : Diagramme en bâtons et diagramme circulaire

- Dans un **diagramme en bâtons**, on place les valeurs du caractère sur un axe horizontal. Au-dessus de chaque valeur, on trace un bâton vertical dont la hauteur est proportionnelle à l'effectif de cette valeur (ou à sa fréquence).
- Dans un **diagramme circulaire**, on partage un disque en secteurs, un secteur par valeur. La mesure de l'angle de chaque secteur est proportionnelle à l'effectif de la valeur : un secteur de fréquence f a un angle de mesure $f \times 360^\circ$.

En clair : dans les deux diagrammes, la taille de chaque élément suit l'effectif. C'est la hauteur pour un bâton, l'angle pour un secteur.

Vérifie : Lire le diagramme

D'après le diagramme en bâtons, pendant combien de jours le cinéma a-t-il eu au moins 3 séances complètes ?

A. 6 jours
Faux.
Faux. 6 est la hauteur du bâton de la valeur 3 seule. « Au moins 3 » compte aussi la valeur 4.

B. 8 jours
Juste.
Juste. On ajoute les hauteurs des bâtons de 3 et de 4 : $6 + 2 = 8$.

C. 7 jours
Faux.
Faux. $3 + 4 = 7$ ajoute les valeurs lues sur l'axe horizontal. Il faut ajouter les hauteurs des bâtons.

Correction
« Au moins 3 » signifie 3 ou 4 séances. Les bâtons de ces valeurs mesurent 6 et 2. Le cinéma a donc eu au moins 3 séances complètes pendant $6 + 2 = 8$ jours.

Le diagramme circulaire

Rappel de 1re année : un diagramme circulaire montre la part de chaque valeur dans l'ensemble. Un tour complet mesure 360°, et un angle se mesure au rapporteur. Ici, c'est toi qui construis le diagramme, sur ton cahier.

À toi : L'angle d'un secteur

Une série fictive a pour effectif total $N = 40$. Une de ses valeurs a l'effectif 10. Quel angle mesure le secteur de cette valeur ?

A. $10°$

Faux.

Faux. Un effectif n'est pas un angle. Le disque entier, $360°$, représente les 40 individus.

B. $25°$

Faux.

Faux. 25 est le pourcentage de cette valeur. Un pourcentage ne se lit pas en degrés.

C. $36°$

Faux.

Faux. $\frac{10}{100} \times 360 = 36$: tu as divisé par 100 au lieu de 40.

D. $90°$

Juste.

Juste. La fréquence est $\frac{10}{40} = 0{,}25$, et $0{,}25 \times 360° = 90°$.

Correction

La fréquence de la valeur est $\frac{10}{40} = 0{,}25$. L'angle de son secteur est $0{,}25 \times 360° = 90°$: un quart du disque.

Exemple : Les angles du cinéma imaginaire

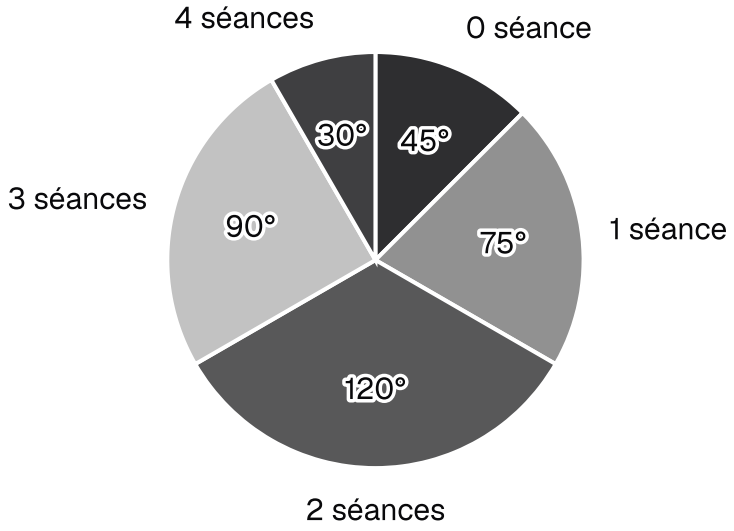
Pour les 24 jours du cinéma, une valeur d'effectif n a un secteur d'angle

$$\frac{n}{24} \times 360° = n \times 15°$$

Séances complètes	0	1	2	3	4	Total
Nombre de jours	3	5	8	6	2	24
Angle	$45°$	$75°$	$120°$	$90°$	$30°$	$360°$

Sur ton cahier, trace un disque et un rayon de départ. Place au rapporteur les secteurs de $45°$, $75°$, $120°$, $90°$ et $30°$, l'un après l'autre.

Le secteur de la valeur 2 occupe le tiers du disque.



Le diagramme circulaire des 24 jours du cinéma : chaque secteur a pour angle son nombre de jours multiplié par 15 degrés.

Méthode : Tracer et lire les deux diagrammes

1. **Bâtons** : gradue l'axe vertical, puis trace au-dessus de chaque valeur un bâton haut comme son effectif.
2. **Angles** : pour chaque valeur, calcule $\frac{n}{N} \times 360°$. Vérifie que leur somme fait $360°$.
3. **Secteurs** : trace-les au rapporteur, l'un à la suite de l'autre.
4. **Lecture inverse** : un secteur d'angle a degrés correspond à l'effectif $\frac{a}{360} \times N$.

Lire dans les deux sens

Vérifie : Retrouver un effectif

Le diagramme circulaire d'une série fictive d'effectif total $N = 60$ a un secteur de $72°$. Quel est l'effectif de la valeur de ce secteur ?

A. 12

Juste.

Juste. $\frac{72}{360} = 0{,}2$, et $0{,}2 \times 60 = 12$.

B. 20

Faux.

Faux. $\frac{72}{360} \times 100 = 20$ donne le pourcentage, pas l'effectif.

C. 72

Faux.

Faux. 72 est la mesure de l'angle, en degrés. Il faut la convertir en effectif.

Correction

La fréquence est $\frac{72}{360} = 0{,}2$. L'effectif est donc $0{,}2 \times 60 = 12$.

Contrôle : $\frac{12}{60} \times 360° = 72°$.

Voici un second diagramme en bâtons, pour une autre série fictive. Un péage imaginaire a compté les personnes présentes dans chacune de 30 voitures.

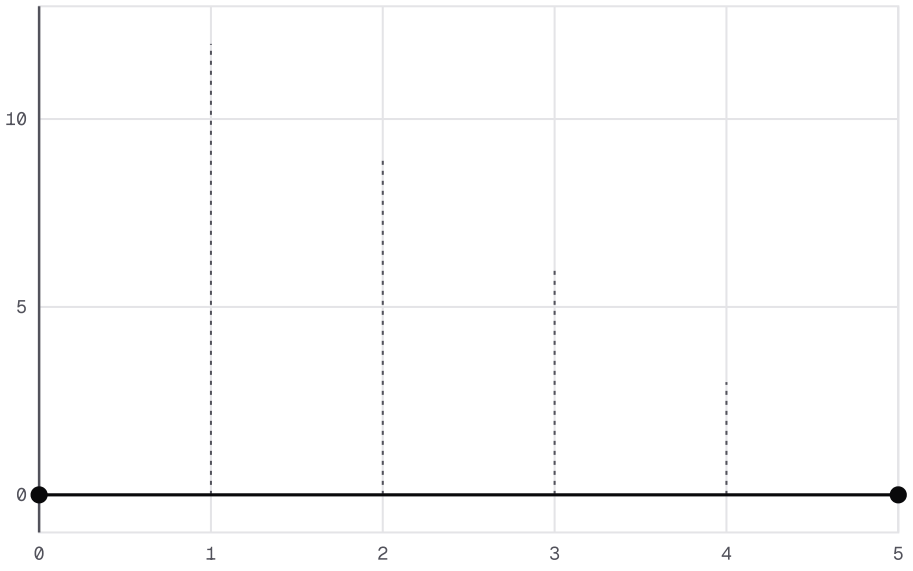


Diagramme en bâtons : nombre de personnes par voiture à un péage imaginaire, pour 30 voitures.

À toi : Du bâton au secteur

Lis sur ce diagramme l'effectif de chaque valeur. Calcule ensuite les angles du diagramme circulaire de la même série.

Indice 1

L'outil : la hauteur de chaque bâton donne l'effectif ; l'angle vaut $\frac{n}{N} \times 360°$.

Indice 2

Le calcul : $N = 12 + 9 + 6 + 3 = 30$, et $\frac{360}{30} = 12$: chaque voiture compte pour $12°$.

Indice 3

Point délicat : contrôle que la somme des angles fait bien $360°$.

Réponse

Effectifs. Les hauteurs des bâtons donnent 12, 9, 6 et 3 voitures pour 1, 2, 3 et 4 personnes.

Angles. Chaque voiture compte pour $12°$. Les angles sont $12 \times 12° = 144°$, $9 \times 12° = 108°$, $6 \times 12° = 72°$ et $3 \times 12° = 36°$.

Contrôle. La somme des angles vaut bien $144 + 108 + 72 + 36 = 360$ degrés.

Comparer deux séries statistiques

Deux lignes de bus

Situation fictive : une ville imaginaire a deux lignes de bus, A et B. On a chronométré 20 trajets de la ligne A et 25 trajets de la ligne B.

Durée (minutes)	10	15	20	25	60
Trajets de la ligne A	4	10	4	2	0
Trajets de la ligne B	13	5	5	0	2

Devine : Que comparer ?

Pour la ligne A, 14 trajets durent au plus 15 minutes. Pour la ligne B, il y en a 18. Quelle ligne a la plus grande part de trajets courts ? Que dois-tu comparer pour répondre ?

A. Les effectifs 18 et 14.

Faux.

Faux. La ligne B a aussi plus de trajets en tout : 25 contre 20. Comparer 18 et 14 ne suffit pas.

B. Les fréquences $\frac{14}{20}$ et $\frac{18}{25}$.

Juste.

Juste. Les effectifs totaux sont différents : on compare des fréquences.

C. Les nombres de trajets en tout, 20 et 25.

Faux.

Faux. Le nombre total de trajets ne dit rien, seul, de la part des trajets courts.

Correction

Les effectifs totaux sont différents, donc on compare les fréquences.

- Ligne A : $\frac{14}{20} = 0{,}7$, soit 70 pour cent.
- Ligne B : $\frac{18}{25} = 0{,}72$, soit 72 pour cent.

La part de trajets courts est donc un peu plus grande pour la ligne B.

Exemple : Les paramètres de la ligne A

Ligne A, avec $N = 20$.

Moyenne : la somme des produits « effectif $\times$ valeur » vaut

$$\begin{aligned} &4 \times 10 + 10 \times 15 \\ &\quad + 4 \times 20 + 2 \times 25 \\ &= 320 \end{aligned}$$

donc $m = \dfrac{320}{20} = 16$ minutes.

- Médiane : sur les cinq colonnes du tableau, les cumuls sont 4, 14, 18, 20, 20 (la durée 60 minutes a l'effectif 0). Les rangs 10 et 11 portent tous deux la valeur 15 : la médiane vaut 15 minutes.
- Mode : 15 minutes, d'effectif 10.

La ligne B

Vérifie : La médiane de la ligne B

Pour la ligne B, $N = 25$. Quelle est la médiane des durées ?

A. 20 minutes

Faux.

Faux. 20 est la valeur placée au milieu du tableau. La médiane se cherche dans la liste rangée des 25 durées.

B. 17 minutes

Faux.

Faux. 17 minutes est la moyenne des durées de B, pas sa médiane.

C. 10 minutes

Juste.

Juste. Le rang du milieu est $\dfrac{25+1}{2} = 13$. Le premier cumul, 13, l'atteint déjà : c'est celui de la valeur 10.

D. 15 minutes

Faux.

Faux. Le rang du milieu est 13, et les 13 premières durées de la liste rangée valent toutes 10 minutes.

Correction

Ici $N = 25$ est impair : le rang du milieu est $\dfrac{25+1}{2} = 13$.

Sur les cinq colonnes, les cumuls de B sont 13, 18, 23, 23, 25 (la durée 25 minutes a l'effectif 0). Le cumul de la valeur 10 atteint déjà 13 : la médiane vaut 10 minutes.

Exemple : Moyenne et mode de la ligne B

Ligne B, avec $N = 25$.

Moyenne : la somme des produits « effectif $\times$ valeur » vaut

$$\begin{aligned} &13 \times 10 + 5 \times 15 \\ &\quad + 5 \times 20 + 2 \times 60 \\ &= 425 \end{aligned}$$

donc $m = \dfrac{425}{25} = 17$ minutes.

- Mode : 10 minutes, d'effectif 13.

Les deux trajets de 60 minutes pèsent $2 \times 60 = 120$ minutes dans la somme 425.

Méthode : Comparer deux séries, pas à pas

1. **Calculer** les mêmes paramètres pour les deux séries : moyenne, médiane, mode.
2. **Passer aux fréquences** quand les effectifs totaux diffèrent.
3. **Comparer** chaque paramètre, et regarder les deux diagrammes ensemble.
4. **Répondre** à la question posée par une phrase qui cite les paramètres.

Attention : Une valeur isolée tire la moyenne

Une **valeur isolée** est une valeur très éloignée de toutes les autres. Elle change beaucoup la moyenne et peu la médiane.

Série fictive : 6, 7, 7, 8, 42.

✗ « La moyenne vaut 14, donc les valeurs tournent autour de 14. »

✓ La moyenne vaut $\dfrac{70}{5} = 14$, mais la médiane vaut 7 : quatre valeurs sur cinq sont proches de 7.

Conclure

Exemple : Comparer avec des fréquences

Part des trajets d'au plus 15 minutes :

- ligne A : $\dfrac{14}{20} = 0{,}7$, soit 70 pour cent ;
- ligne B : $\dfrac{18}{25} = 0{,}72$, soit 72 pour cent.

Les effectifs 14 et 18 ne se comparent pas directement : les deux lignes n'ont pas le même nombre de trajets.

Vérifie : La meilleure conclusion

Pour un trajet habituel, quelle phrase compare le mieux les deux lignes, avec tous les paramètres calculés ?

A. La ligne A est plus rapide, car sa moyenne, 16 minutes, est plus petite que celle de B.

Faux.

Faux. La phrase n'utilise que la moyenne. La médiane de B, 10 minutes, est plus petite que celle de A, 15 minutes : au moins la moitié des trajets de B durent au plus 10 minutes.

B. Au moins la moitié des trajets de B durent au plus 10 minutes ; deux trajets très longs relèvent sa moyenne.

Juste.

Juste. La phrase cite la médiane et explique l'écart avec la moyenne.

C. Les deux lignes se valent, car leurs moyennes, 16 et 17 minutes, sont proches.

Faux.

Faux. Des moyennes proches peuvent cacher des séries très différentes : les médianes sont 15 et 10.

Correction

Moyennes : 16 minutes pour A, 17 pour B. Médianes : 15 pour A, 10 pour B.

Au moins la moitié des trajets de B durent au plus 10 minutes ; pour A, au moins la moitié durent au plus 15 minutes. Mais les deux trajets de 60 minutes relèvent la moyenne de B au-dessus de celle de A.

À toi, puis rédiger

À toi : Deux joueurs imaginaires

Deux joueurs imaginaires de basket comparent leurs points par match. Joueur X, sur 5 matchs : 8, 10, 12, 10, 10. Joueur Y, sur 6 matchs : 2, 4, 30, 4, 6, 8. Compare les deux joueurs avec la moyenne et la médiane.

Indice 1

L'outil : pour chaque joueur, la moyenne, puis la médiane de la liste rangée.

Indice 2

Le calcul : X totalise 50 points en 5 matchs, Y totalise 54 points en 6 matchs.

Indice 3

Point délicat : Y a joué 6 matchs, un nombre pair ; sa médiane est la moyenne de deux valeurs.

Réponse

Joueur X. Moyenne : $m = \dfrac{50}{5} = 10$. Sur la liste rangée 8, 10, 10, 10, 12, la médiane est la troisième valeur, 10.

Joueur Y. Moyenne : $m = \dfrac{54}{6} = 9$. Sur la liste rangée $2, 4, 4, 6, 8, 30$, la médiane est la moyenne des troisième et quatrième valeurs : $\dfrac{4+6}{2} = 5$.

Conclusion. X marque plus en général. La moyenne de Y, 9, est relevée par un seul match à 30 points ; sa médiane, 5, décrit mieux ses matchs habituels.

Remarque : Coin examen : rédiger une comparaison

Niveau examen régional : une rédaction modèle termine la comparaison par une phrase qui répond à la question. Pour les deux lignes de bus imaginaires :

« Ligne A : $m = 16$ minutes, médiane 15 minutes. Ligne B : $m = 17$ minutes, médiane 10 minutes.

La moyenne de B est plus grande, à cause de deux trajets de 60 minutes. Mais sa médiane est plus petite : au moins la moitié des trajets de B durent au plus 10 minutes.

Pour un trajet habituel, la ligne B est donc plus rapide. »

Synthèse : la statistique du collège

Tu sais maintenant décrire une série avec quatre outils : effectifs, fréquences, moyenne et médiane. Le mode complète la boîte à outils. Voici l'essentiel, puis un dernier entraînement.

Effectif total

$$N = n_1 + n_2 + \cdots + n_p$$

Fréquence d'une valeur

$$f = \frac{n}{N}$$

Fréquence en pourcentage

$$f \times 100$$

Moyenne avec les effectifs

$$m = \frac{n_1 \times x_1 + \cdots + n_p \times x_p}{N}$$

Moyenne avec les fréquences

$$m = f_1 \times x_1 + \cdots + f_p \times x_p$$

Rang de la médiane, effectif total impair

$$\frac{N+1}{2}$$

Rangs de la médiane, effectif total pair

$$\frac{N}{2} \text{ et } \frac{N}{2} + 1$$

Angle d'un secteur

$$f \times 360^\circ$$

À retenir : Ce qui relie les notions

- Les **effectifs** comptent ; les **fréquences** les divisent par N.
- Les effectifs cumulés croissants situent chaque rang : ils mènent à la **médiane**.
- La **moyenne** se calcule avec les effectifs, ou avec les fréquences sans connaître N.

- Le **mode** est la valeur de plus grand effectif.

À retenir : Moyenne ou médiane ?

La moyenne utilise toutes les valeurs : une valeur isolée la déplace. La médiane coupe la série rangée en deux moitiés : elle résiste aux valeurs isolées. Pour comparer deux séries, regarde les deux.

À toi : Retour à la ligue imaginaire

Reprends les deux équipes du début du chapitre. Faucons : 21, 24, 22, 25, 23 buts. Lynx : 12, 38, 15, 36, 14 buts.

Pour chaque équipe, donne l'effectif et la fréquence des matchs à plus de 20 buts, puis la moyenne et la médiane. Conclus.

Indice 1
L'outil : compte les matchs, puis divise par $N = 5$; pour la médiane, range les cinq scores.
Indice 2
Le calcul : chaque équipe totalise 115 buts ; les listes rangées sont 21, 22, 23, 24, 25 et 12, 14, 15, 36, 38.
Indice 3
Point délicat : la médiane est le troisième score de la liste rangée, pas le troisième match joué.

Réponse
Faucons. Ils dépassent 20 buts dans 5 matchs sur 5 : fréquence 1, soit 100 pour cent ; $m = \dfrac{115}{5} = 23$; médiane 23.

Lynx. Ils dépassent 20 buts dans 2 matchs sur 5 : fréquence $\dfrac{2}{5} = 0{,}4$, soit 40 pour cent ; $m = 23$; médiane 15.

Conclusion. Les deux équipes ont la même moyenne, mais les Faucons sont réguliers. Les Lynx alternent deux très gros scores et trois petits : la plupart de leurs matchs sont sous la moyenne.

Tu sais maintenant lire un tableau, calculer les paramètres d'une série et comparer deux séries avec méthode.

SOURCES

Référentiel BURHAN du chapitre « Statistique » (3ac · general · mathematiques)