Statistiques descriptives

Fiche méthode

(Maths 2nde · Statistiques et probabilités · Statistique descriptive)

La démarche pour étudier une série statistique

1. Je regarde comment la série est donnée : liste brute, tableau d'effectifs, indicateurs déjà calculés, classes.

2. Je range et je cumule : série ordonnée ou effectifs cumulés croissants, et l'effectif total N.

3. Je calcule l'indicateur demandé avec sa règle : somme ÷ N, rangs, écarts au carré, centres ou moyennes des classes.

4. J'interprète par une phrase avec l'unité, en écrivant « ≈ » pour une estimation.

Le réflexe — un indicateur se justifie : rangs calculés pour la médiane et les quartiles, somme et effectif pour la moyenne, étapes de la variance pour l'écart type.

Cas 1 · Moyenne, groupes et linéarité

Réflexe — je repasse par les sommes N × moyenne ; pour y = ax + b, la moyenne devient a × moyenne + b.

Exemple. Questions types : « Calculer la moyenne de l'ensemble », « Déterminer la nouvelle moyenne ». Exemple : 24 élèves ont 9,5 de moyenne, 16 autres ont 12. Moyenne des 40 ?

Sommes : $24 \times 9{,}5 = 228$ et $16 \times 12 = 192$.

$\overline{x} = \dfrac{228 + 192}{40} = \dfrac{420}{40} = 10{,}5$.

La demi-somme 10,75 serait fausse : les groupes n'ont pas le même effectif.

Attention — ajouter une valeur supérieure à la moyenne la fait monter, inférieure la fait baisser.

Cas 2 · Médiane et quartiles

Réflexe — je range, je compte N, je calcule les rangs, puis je lis les valeurs.

Exemple. Questions types : « Déterminer la médiane, les quartiles, l'écart interquartile ». Exemple : 3 ; 5 ; 5 ; 8 ; 9 ; 11 ; 14 ; 17 ; 20 ; 24.

$N = 10$ est pair : $Me = \dfrac{9 + 11}{2} = 10$ (rangs 5 et 6).

$\dfrac{10}{4} = 2{,}5$, arrondi à 3 : $Q_1 = 5$. $\dfrac{30}{4} = 7{,}5$, arrondi à 8 : $Q_3 = 17$.

Écart interquartile : $17 - 5 = 12$.

Attention — on range la série avant tout, en gardant les valeurs répétées.

Cas 3 · Écart type

Réflexe — moyenne, écarts, carrés × effectifs, somme ÷ N, puis racine carrée.

Exemple. Questions types : « Calculer la variance puis l'écart type ». Exemple : nombre d'animaux par foyer, valeurs 0 ; 1 ; 2 ; 3, effectifs 4 ; 7 ; 6 ; 3.

$N = 20$ et $\overline{x} = \dfrac{0 + 7 + 12 + 9}{20} = \dfrac{28}{20} = 1{,}4$.

$V = \dfrac{4(0 - 1{,}4)^2 + 7(1 - 1{,}4)^2 + 6(2 - 1{,}4)^2 + 3(3 - 1{,}4)^2}{20} = \dfrac{7{,}84 + 1{,}12 + 2{,}16 + 7{,}68}{20} = \dfrac{18{,}8}{20} = 0{,}94$.

$\sigma = \sqrt{0{,}94} \approx 0{,}97$ animal.

Attention — la variance n'est pas l'écart type : il reste la racine carrée.

Cas 4 · Comparer deux séries

Réflexe — positions, puis dispersions, puis valeur extrême, puis une phrase de conclusion.

Exemple. Questions types : « Laquelle est la plus régulière ? Justifier ». Exemple : deux machines produisent des pièces de longueur moyenne 40 mm ; écarts types 1,2 mm (A) et 4,5 mm (B).

Les moyennes sont égales : on compare les dispersions.

$1{,}2 < 4{,}5$: les pièces de A s'écartent beaucoup moins de 40 mm.

La machine A est la plus régulière : on la retient pour des pièces de précision.

Attention — on compare moyenne avec écart type, médiane avec écart interquartile, jamais en croisant.

Cas 5 · Série en classes : moyenne

Réflexe — moyennes des classes connues : calcul exact ; sinon centres et « ≈ ».

Exemple. Questions types : « Calculer une valeur approchée de la moyenne ». Exemple : durée d'un trajet en bus, $[10\,;14[$: 7 ; $[14\,;18[$: 18 ; $[18\,;22[$: 11 ; $[22\,;26[$: 4.

$N = 40$; centres : 12 ; 16 ; 20 ; 24.

$$\overline{x} \approx \frac{7\times 12 + 18\times 16 + 11\times 20 + 4\times 24}{40} = \frac{84 + 288 + 220 + 96}{40} = \frac{688}{40} = 17{,}2 \text{ min.}$$

C'est une estimation : chaque durée est remplacée par le centre de sa classe.

Attention — le centre de [a ; b[est (a + b) ÷ 2, pas l'amplitude.

Cas 6 · Classe médiane

Réflexe — effectifs cumulés, première classe qui atteint N ÷ 2, proportionnalité dans la classe.

Exemple. Questions types : « Déterminer la classe médiane, estimer la médiane ». Exemple : temps d'attente, $N = 40$: $[0\,;5[$: 5 ; $[5\,;10[$: 9 ; $[10\,;15[$: 16 ; $[15\,;20[$: 10.

Effectifs cumulés : 5 ; 14 ; 30 ; 40. $\dfrac{N}{2} = 20$: le cumul passe de 14 à 30 dans $[10\,;15[$, la classe médiane.

$$Me \approx 10 + 5 \times \frac{20 - 14}{16} = 10 + 5 \times 0{,}375 = 11{,}875, \text{ soit environ } 11{,}9 \text{ min.}$$

Attention — la classe médiane n'est pas la classe de plus grand effectif.

Les pièges

- La médiane se lit sur les rangs de la série rangée, jamais comme la demi-somme du minimum et du maximum.
- La moyenne de deux moyennes n'est pas la moyenne de l'ensemble si les effectifs diffèrent.
- La moyenne des écarts à la moyenne vaut toujours 0 : c'est pour cela qu'on élève au carré.
- Dans une série en classes, la moyenne calculée avec les centres est une valeur approchée.
- Les points du polygone des fréquences cumulées se placent aux bornes supérieures des classes.

Ce document est la propriété intellectuelle de Ben, professeur particulier (benprofparticulier.fr), protégée par le Code de la propriété intellectuelle (art. L. 111-1 et L. 122-4). Il est mis à disposition pour un usage personnel de révision. Toute reproduction, diffusion, mise en ligne ou revente, totale ou partielle, sans autorisation écrite est interdite et constitue une contrefaçon (art. L. 335-2).