

Statistiques et initiation aux probabilités

5^e

Programme du cycle 4 — BO spécial n°11 du 26 novembre 2015

Deux chapitres en un, et une même question : **que peut-on dire d'un ensemble de nombres ?** Les statistiques décrivent ce qui a été observé, les probabilités anticipent ce qui pourrait l'être. Les premières regardent en arrière, les secondes en avant.

1 Décrire une série

DÉFINITION

La **moyenne** est la somme des valeurs divisée par leur nombre : c'est le point d'équilibre. La **médiane** est la valeur qui partage la série rangée en deux moitiés. Les deux répondent à « quelle est la valeur typique ? » — et donnent parfois des réponses très différentes.

$$\bar{x} = \frac{\text{somme des valeurs}}{\text{effectif}}$$

la médiane est une position dans la série rangée, pas un calcul

effectif impair — 7 valeurs



la 4^e valeur : médiane = 8

effectif pair — 6 valeurs



demi-somme des deux du milieu : $(10+14)/2 = 12$

et toujours : ranger la série AVANT de chercher quoi que ce soit
la médiane n'est pas forcément une valeur de la série — ici 12 n'y figure pas
Effectif impair : la valeur du milieu. Effectif pair : la demi-somme des deux.

LE PIÈGE QUI COÛTE LE PLUS DE POINTS

Chercher la médiane sans avoir rangé la série. — la médiane est une **position**, pas un calcul. Prendre la valeur centrale d'une liste en désordre donne un nombre qui ne signifie rien — et l'erreur est invisible, puisque le résultat ressemble à une réponse.

Le réflexe : réécrire la série rangée en entier sur la copie. Deux minutes qui sécurisent toutes les questions suivantes.

L'étendue — la différence entre la plus grande et la plus petite valeur — complète ces deux indicateurs. Deux séries de même moyenne peuvent avoir des étendues très différentes, et c'est cette différence qui distingue une classe homogène d'une classe écartée.

DÉFINITION

La **fréquence** d'une valeur est son effectif rapporté à l'effectif total. C'est ce qui permet de comparer deux groupes de tailles différentes : 12 réussites sur 20 et 21 sur 35 donnent la même fréquence, 0,6 — soit 60 % — alors que les effectifs, eux, ne se ressemblent pas.

$$f = \frac{\text{effectif de la valeur}}{\text{effectif total}}$$

LIRE UN TABLEAU D'EFFECTIFS

- Repérer l'**effectif total** : c'est la somme de la ligne des effectifs, jamais le nombre de colonnes.
- Pour la moyenne, multiplier chaque valeur par son effectif avant d'additionner.
additionner les valeurs sans les pondérer revient à traiter une note obtenue par douze élèves comme une note obtenue par un seul.
- Pour la médiane, cumuler les effectifs jusqu'à dépasser la moitié du total : la valeur atteinte à ce moment est la médiane.

EXEMPLE

Notes : 8 (×3 élèves), 12 (×5), 15 (×4). Moyenne et médiane ?

- ① L'effectif total vaut 12. La moyenne est pondérée :

$$\frac{3 \times 8 + 5 \times 12 + 4 \times 15}{12}$$

- ② Soit :

$$\frac{24 + 60 + 60}{12} = 12$$

- ③ Pour la médiane, on cumule : 3 élèves à 8, puis 8 élèves cumulés à 12. La 6^e et la 7^e note valent donc 12.

moyenne 12 et médiane 12

EXEMPLE

Dans une classe de 25 élèves, 10 viennent en bus, 6 à vélo, 7 à pied et 2 en voiture. Quelles sont les fréquences ?

- ① Le bus.

$$\frac{10}{25} = 0,4$$

- ② Le vélo.

$$\frac{6}{25} = 0,24$$

- ③ La marche, puis la voiture.

$$\frac{7}{25} = 0,28$$

- ④ Et le dernier mode.

$$\frac{2}{25} = 0,08$$

- ⑤ Contrôle indispensable : la somme des fréquences.

$$0,4 + 0,24 + 0,28 + 0,08 = 1$$

elle vaut toujours 1, puisqu'on a réparti la totalité de l'effectif.

40 %, 24 %, 28 % et 8 % — et 40 + 24 + 28 + 8 = 100

LE PIÈGE QUI COÛTE LE PLUS DE POINTS

Comparer deux groupes en comparant leurs effectifs. — « 10 élèves prennent le bus ici, 12 là-bas » ne dit rien si les deux classes n'ont pas le même nombre d'élèves. Un effectif ne se compare qu'à effectif total égal.

Le réflexe : dès que les deux totaux diffèrent, passer aux **fréquences** : 10 sur 25 fait 40 %, 12 sur 40 fait 30 % — et le classement s'inverse.

2 Le vocabulaire des probabilités

DÉFINITION

Une **expérience aléatoire** a plusieurs résultats possibles et on ne sait pas lequel sortira : lancer un dé, tirer une carte. Chaque résultat possible est une **issue**. Un **événement** regroupe une ou plusieurs issues : « obtenir un nombre pair » réunit trois issues sur six.

RÈGLE

Quand toutes les issues ont la même chance de se produire — on dit qu'il y a **équiprobabilité** — la probabilité d'un événement se compte.

$$P(A) = \frac{\text{nombre d'issues favorables}}{\text{nombre d'issues possibles}}$$

toute probabilité vit entre 0 et 1 — un résultat hors bornes est une erreur



la somme des probabilités de toutes les issues vaut toujours 1

Une probabilité se lit comme une position sur cette échelle.

RÈGLE

La probabilité de l'événement **contraire** complète à 1. C'est souvent le chemin le plus court : « au moins un » se calcule presque toujours par « aucun ».

$$P(\overline{A}) = 1 - P(A)$$

CALCULER UNE PROBABILITÉ EN SITUATION D'ÉQUIPROBABILITÉ

- ① **Décrire l'expérience** et lister toutes les issues possibles. Un dé à six faces : six issues ; une pièce : deux.
- ② **Vérifier l'équiprobabilité** — dé équilibré, pièce non truquée, boules indiscernables au toucher. Sans elle, la formule ne s'applique pas.
l'énoncé le précise toujours ; s'il ne le précise pas, il faut le dire.
- ③ **Compter les issues favorables** à l'événement, puis diviser par le nombre total d'issues.
- ④ **Contrôler** : le résultat est entre 0 et 1. Et si l'événement contraire est plus simple à compter, on calcule celui-là puis on complète à 1.

LE PIÈGE QUI COÛTE LE PLUS DE POINTS

Annoncer une probabilité de 1,5 ou de $-0,2$. — aucune probabilité ne sort de l'intervalle $[0; 1]$. Un tel résultat signale toujours une erreur : dénominateur oublié, addition à la place d'une division, ou issues comptées deux fois.

Le réflexe : vérifier que le numérateur ne dépasse jamais le dénominateur. Il ne peut pas y avoir plus de cas favorables que de cas possibles.

Un dernier outil, qu'on rencontrera en 4^e et en 3^e : le **diagramme en boîte et à moustaches**. La boîte y contient la moitié centrale des valeurs, un trait marque la médiane, et les moustaches vont jusqu'aux extrêmes. On peut déjà en lire l'idée : ce diagramme ne montre pas les valeurs, il montre comment elles se répartissent — et deux boîtes côte à côte comparent deux séries sans un seul calcul.

À VÉRIFIER

- Ai-je **rangé** la série avant de chercher la médiane ?
- Pour comparer deux groupes de tailles différentes, suis-je passé aux **fréquences** ?
- Chaque probabilité annoncée est-elle comprise entre 0 et 1 ?
- La somme des probabilités de toutes les issues vaut-elle bien 1 ?

À RETENIR

- **Ranger la série** avant de chercher la médiane — sinon tout est faux.
- Effectif **impair** : la valeur du milieu. **Pair** : la demi-somme des deux du milieu.
- La médiane n'appartient pas forcément à la série.
- **Issue** = un résultat ; **événement** = un groupe d'issues.
- $P(A) = \frac{\text{favorables}}{\text{possibles}}$, et toute probabilité vit **entre 0 et 1**.
- $P(\overline{A}) = 1 - P(A)$: « ne pas » appelle presque toujours le contraire.
- Même moyenne ne veut pas dire mêmes séries : regarder l'**étendue**.