Les statistiques décrivent ce qui s'est déjà produit : on organise, résume et représente des données (notes d'une classe, durées de trajet, tailles). Les probabilités anticipent ce qui pourrait se produire : on mesure la chance qu'un événement survienne dans une expérience aléatoire (lancer de dé, tirage d'une carte). Les difficultés principales viennent de confusions durables : confondre moyenne et médiane, croire que toutes les issues sont également probables (somme 7 vs somme 12 avec deux dés), traiter un événement comme une certitude parce que sa probabilité est élevée.

PARTIE I. STATISTIQUES

Vocabulaire (5e–3e)

Série statistique : liste de valeurs numériques collectées sur un ensemble d'individus (les notes des 25 élèves d'une classe, les tailles de 30 nouveau-nés).

Effectif d'une valeur : nombre d'individus ayant cette valeur. Dans la série (12 ; 8 ; 12 ; 15 ; 12), l'effectif de la valeur 12 est 3.

Effectif total : nombre total d'individus de la série (noté N). Dans l'exemple ci-dessus, $N = 5$.

Fréquence d'une valeur : rapport de l'effectif à l'effectif total, $f = \dfrac{\text{effectif}}{N}$. Elle est toujours comprise entre 0 et 1, ou exprimée en pourcentage entre 0 % et 100 %.

Distribution des fréquences : la somme de toutes les fréquences d'une série vaut toujours 1 (ou 100 %).

Exemples On relève les notes de douze élèves à un contrôle :

12 ; 8 ; 15 ; 12 ; 10 ; 14 ; 8 ; 12 ; 17 ; 10 ; 14 ; 8.

Note	8	10	12	14	15	17
Effectif	3	2	3	2	1	1
Fréquence	$\frac{3}{12}$	$\frac{2}{12}$	$\frac{3}{12}$	$\frac{2}{12}$	$\frac{1}{12}$	$\frac{1}{12}$

Vérification : somme des effectifs = $12 = N$ ✓ somme des fréquences = $\dfrac{12}{12} = 1$ ✓.

Moyenne d'une série (5e–4e)

La **moyenne** d'une série, notée $\bar{x}$, est la somme de toutes les valeurs divisée par l'effectif total :

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

Moyenne pondérée : quand les valeurs se répètent, on utilise les effectifs. Si la valeur x_i a pour effectif n_i et que $N = n_1 + n_2 + \cdots + n_p$:

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \cdots + n_px_p}{N}$$

Exemples Reprenons la série des douze notes. On appelle A la moyenne de cette série :

$$A = \frac{3 \times 8 + 2 \times 10 + 3 \times 12 + 2 \times 14 + 1 \times 15 + 1 \times 17}{12}$$
$$A = \frac{24 + 20 + 36 + 28 + 15 + 17}{12}$$
$$A = \frac{140}{12} \approx 11{,}67$$

La moyenne de la classe est environ $11{,}67$.

Interprétation : la moyenne est la « valeur d'équilibre » : si l'on redistribuait équitablement toutes les valeurs, chacun aurait la moyenne. Attention : la moyenne peut ne correspondre à aucune valeur réelle de la série (ici, aucun élève n'a eu $11{,}67$).

Médiane d'une série (4e–3e)

La **médiane** partage la série ordonnée en deux moitiés de même effectif : la moitié des valeurs lui sont inférieures ou égales, l'autre moitié lui sont supérieures ou égales.

Méthode :

1. On range la série dans l'ordre **croissant**.
2. Si l'effectif total N est **impair**, la médiane est la valeur du milieu (la valeur de rang $\frac{N+1}{2}$).
3. Si N est **pair**, la médiane est la demi-somme des deux valeurs centrales (rangs $\frac{N}{2}$ et $\frac{N}{2} + 1$).

Exemples

- Série B à sept valeurs : $(5\,;\,8\,;\,9\,;\,12\,;\,14\,;\,14\,;\,20)$ (déjà triée). $N = 7$ impair, la médiane est la valeur de rang 4 : $\text{Med}(B) = 12$.
- Série C à huit valeurs : $(5\,;\,8\,;\,9\,;\,12\,;\,14\,;\,14\,;\,18\,;\,20)$. $N = 8$ pair, la médiane est la demi-somme des valeurs de rangs 4 et 5 :

  $$\text{Med}(C) = \frac{12 + 14}{2} = 13$$

- Reprenons les douze notes. Triées : $(8\,;\,8\,;\,8\,;\,10\,;\,10\,;\,12\,;\,12\,;\,12\,;\,14\,;\,14\,;\,15\,;\,17)$. $N = 12$, les valeurs de rangs 6 et 7 sont 12 et 12 : la médiane vaut 12.

*Moyenne et médiane ne coïncident pas en général. Quelques valeurs très grandes (ou très petites) influencent beaucoup la moyenne, mais peu la médiane. C'est pourquoi, pour décrire les salaires d'un pays, la **médiane** est souvent plus représentative que la moyenne : elle résiste aux valeurs extrêmes.*

Exercice 1 ***Moyenne et médiane (5e–3e)*** On relève le nombre de livres lus par huit élèves sur l'année :

$$A = (3\,;\,7\,;\,5\,;\,2\,;\,8\,;\,5\,;\,4\,;\,6).$$

1. Calculer la moyenne et la médiane de A.
2. Un neuvième élève a lu 30 livres. Calculer la nouvelle moyenne et la nouvelle médiane de A.
3. Quel indicateur a le plus changé ? Expliquer.

Étendue et quartiles (3^e–2^{de})

Les indicateurs de position (moyenne, médiane) résument la série par un seul nombre. Les indicateurs de **dispersion** décrivent, eux, comment les valeurs sont réparties autour de ce centre.

Étendue : différence entre la plus grande et la plus petite valeur de la série.

$$\text{Étendue} = \max - \min$$

Premier quartile Q_1 **:** plus petite valeur de la série triée telle qu'au moins 25 % des valeurs lui sont inférieures ou égales.

Troisième quartile Q_3 **:** plus petite valeur telle qu'au moins 75 % des valeurs lui sont inférieures ou égales.

Écart interquartile : $Q_3 - Q_1$. Il mesure la dispersion de la « moitié centrale » de la série.

Méthode pratique : on trie la série, on calcule $\dfrac{N}{4}$ pour Q_1 et $\dfrac{3N}{4}$ pour Q_3. Si le résultat est entier, on prend la valeur de ce rang ; sinon, on arrondit au rang entier supérieur.

Exemples Série triée : $(5\,;\,7\,;\,8\,;\,10\,;\,11\,;\,12\,;\,14\,;\,16\,;\,18\,;\,20\,;\,22\,;\,25)$. $N = 12$.

- Étendue : $25 - 5 = 20$.
- $\dfrac{N}{4} = 3$ entier : Q_1 est la valeur de rang 3, soit $Q_1 = 8$.
- $\dfrac{3N}{4} = 9$ entier : Q_3 est la valeur de rang 9, soit $Q_3 = 18$.
- Écart interquartile : $Q_3 - Q_1 = 18 - 8 = 10$.

Les quartiles ne sont pas sensibles aux valeurs extrêmes. En décrivant la moitié centrale des données, l'écart interquartile complète la médiane de la même manière que l'écart-type (non traité ici) complète la moyenne.

Exercice 2 ***Indicateurs complets (3^e–2^{de})*** Voici les durées (en minutes) du trajet domicile-collège pour vingt élèves :

$$B = (5\,;\,8\,;\,10\,;\,12\,;\,12\,;\,15\,;\,15\,;\,18\,;\,20\,;\,20\,;\,22\,;\,25\,;\,25\,;\,28\,;\,30\,;\,30\,;\,35\,;\,40\,;\,45\,;\,60).$$

La série est déjà triée.

1. Calculer la moyenne, la médiane, l'étendue et les deux quartiles.
2. Calculer l'écart interquartile.
3. Interpréter : la moitié des élèves ont un trajet de quelle durée maximale ?

PARTIE II. PROBABILITÉS

Vocabulaire des probabilités (4^e–3^e)

Expérience aléatoire : expérience dont on ne peut pas prévoir le résultat à l'avance, mais dont on connaît tous les résultats possibles. Exemple : lancer un dé cubique équilibré.

Issue : résultat possible de l'expérience. Pour un dé : les issues sont 1, 2, 3, 4, 5, 6.

Univers : ensemble de toutes les issues, noté Ω. Pour un dé : $\Omega = \{1\,;\,2\,;\,3\,;\,4\,;\,5\,;\,6\}$.

Événement : ensemble d'issues regroupées par une condition commune. Pour un dé, « obtenir un nombre pair » est l'événement $\{2\,;\,4\,;\,6\}$.

Événement élémentaire : événement ne contenant qu'une seule issue.

Événement certain : événement égal à l'univers (il se réalise toujours).

Événement impossible : événement vide (il ne se réalise jamais).

Probabilité dans un cas d'équiprobabilité (4^e–3^e)

Dans une expérience où toutes les issues ont la même chance de se produire (on parle d'**équiprobabilité**) :

$$P(\text{événement}) = \frac{\text{nombre d'issues favorables}}{\text{nombre total d'issues}}$$

Propriétés à connaître :

- $0 \leqslant P(A) \leqslant 1$ pour tout événement A ;
- $P(\Omega) = 1$ (événement certain) ;
- $P(\emptyset) = 0$ (événement impossible) ;
- si A et B sont **incompatibles** (ils n'ont aucune issue commune), $P(A \text{ ou } B) = P(A) + P(B)$.

Exemples On lance un dé cubique équilibré. Calculons quelques probabilités. On note A, B, C les événements suivants :

- A : « obtenir un nombre pair » $= \{2\,;\,4\,;\,6\}$. $P(A) = \frac{3}{6} = \frac{1}{2}$.
- B : « obtenir un multiple de 3 » $= \{3\,;\,6\}$. $P(B) = \frac{2}{6} = \frac{1}{3}$.
- C : « obtenir un nombre supérieur à 4 » $= \{5\,;\,6\}$. $P(C) = \frac{2}{6} = \frac{1}{3}$.

Attention : l'équiprobabilité n'est pas automatique. Dans « lancer deux dés et noter la somme », les résultats 2, 3, *…,* 12 *ne sont* ***pas*** *équiprobables (voir plus loin). Avant d'appliquer la formule, il faut vérifier que les issues choisies sont effectivement équiprobables.*

Événement contraire et opérations sur les événements (3^e–2^{de})

Événement contraire de A **:** noté $\bar{A}$. Il est constitué de toutes les issues qui ne sont pas dans A. Autrement dit, $\bar{A}$ se réalise lorsque A ne se réalise pas.

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

Intersection $A \cap B$ **:** événement qui se réalise quand A **et** B se réalisent simultanément.
Union $A \cup B$ **:** événement qui se réalise quand A **ou** B se réalise (au moins l'un des deux).
Formule générale : $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. Si A et B sont incompatibles ($A \cap B = \emptyset$), cette formule se simplifie en $P(A \cup B) = P(A) + P(B)$.

Exemples On reprend le dé cubique et les événements A (pair) et B (multiple de 3) définis plus haut.

- Contraire de A : $\bar{A} = \{1\,;\,3\,;\,5\}$ (« impair »). $P(\bar{A}) = \frac{3}{6} = \frac{1}{2} = 1 - P(A)$. ✓
- Intersection : $A \cap B = \{6\}$ (pair **et** multiple de 3). $P(A \cap B) = \frac{1}{6}$.
- Union : $A \cup B = \{2\,;\,3\,;\,4\,;\,6\}$ (pair **ou** multiple de 3). $P(A \cup B) = \frac{4}{6} = \frac{2}{3}$.
  On vérifie la formule : $P(A) + P(B) - P(A \cap B) = \frac{1}{2} + \frac{1}{3} - \frac{1}{6} = \frac{3}{6} + \frac{2}{6} - \frac{1}{6} = \frac{4}{6} = \frac{2}{3}$. ✓

Exercice 3 ***Une urne (4^e–3^e)*** Une urne contient 4 boules rouges, 3 boules vertes et 5 boules bleues, toutes indiscernables au toucher. On tire une boule au hasard. Calculer :

1. la probabilité de R : « tirer une boule rouge » ;
2. la probabilité de $\bar{V}$: « ne pas tirer une boule verte » ;

3. la probabilité de $R \cup B$: « tirer une boule rouge ou bleue ».

Expériences à deux étapes : arbre et tableau (3^e–2^de)

Quand une expérience se décompose en plusieurs étapes (lancer deux dés, tirer deux cartes, jouer deux parties), on peut énumérer les issues avec deux outils :

- un **arbre de probabilités**, utile quand les étapes sont différentes ou dépendantes ;
- un **tableau à double entrée**, bien adapté quand les deux étapes sont similaires (deux dés, deux pièces).

Exemple : somme de deux dés. On lance deux dés équilibrés et on note la somme. Le tableau à double entrée des $6 \times 6 = 36$ issues équiprobables :

+	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

On compte les issues favorables à chaque somme :

- $P(\text{somme} = 2) = \dfrac{1}{36}$ (une seule case : $1 + 1$) ;
- $P(\text{somme} = 7) = \dfrac{6}{36} = \dfrac{1}{6}$ (six cases sur la diagonale) ;
- $P(\text{somme} = 12) = \dfrac{1}{36}$ (une seule case : $6 + 6$).

Les sommes 2, 7 et 12 ne sont pas équiprobables : la somme 7 est six fois plus probable que la somme 12. L'intuition qui veut que « tous les résultats soient également probables » est ici clairement fausse, car les sommes ne sont pas les issues élémentaires : les vraies issues sont les couples $(d_1\,;\,d_2)$, au nombre de 36.

Exercice 4 *Deux dés (3^e–2^de)* On lance deux dés cubiques équilibrés et on note la somme.

1. Déterminer la probabilité d'obtenir la somme 10.
2. Déterminer la probabilité d'obtenir une somme au moins égale à 10.
3. Est-il plus probable d'obtenir une somme égale à 7 ou une somme paire ?
4. Soit S l'événement « la somme est strictement inférieure à 5 ». Calculer $P(S)$ puis $P(\bar{S})$.

Exercice 5 *Arbre de probabilités (3^e–2^de)* On dispose de deux urnes :

- l'urne U_1 contient 3 boules rouges et 2 boules bleues ;
- l'urne U_2 contient 1 boule rouge et 4 boules bleues.

On choisit d'abord une urne au hasard (équiprobabilité), puis on tire une boule dans l'urne choisie.

1. Représenter la situation par un arbre de probabilités. Vérifier que la somme des probabilités des branches issues de chaque nœud vaut 1.
2. Calculer la probabilité de tirer une boule rouge.

3. Calculer la probabilité de l'événement « avoir choisi l'urne U_1 et tiré une boule bleue ».

Erreurs classiques en probabilités et statistiques

Erreur	Exemple faux	Correction
Confondre moyenne et médiane	« La moyenne des salaires est 2000 €, donc la moitié des salariés gagnent 2000 € ou moins »	La moyenne ne donne pas la répartition : c'est la **médiane** qui partage la série en deux moitiés
Équiprobabilité abusive des sommes	« Avec deux dés, les sommes 7 et 12 sont aussi probables, il y a 11 sommes possibles »	Les sommes ne sont pas équiprobables : il y a 6 façons d'obtenir 7 contre 1 seule pour 12
Additionner $P(A)$ et $P(B)$ quand A et B ne sont pas incompatibles	$P(\text{pair}) + P(\text{multiple de } 3) = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}$	Il faut soustraire $P(A \cap B)$: $\frac{1}{2} + \frac{1}{3} - \frac{1}{6} = \frac{2}{3}$
Oublier le contexte d'équiprobabilité	Appliquer $P = \frac{\text{favorables}}{\text{total}}$ sans vérifier que les issues sont équiprobables	La formule n'est valable que lorsque **toutes les issues de l'univers sont également probables**
Confondre fréquence observée et probabilité théorique	Après 10 lancers d'une pièce donnant 7 Pile, conclure $P(\text{Pile}) = 0{,}7$	La probabilité théorique est $0{,}5$; la fréquence observée ne la rejoint qu'après un grand nombre de lancers (loi des grands nombres)
Valeur extrême qui « casse » la moyenne	Ignorer l'impact d'une valeur aberrante sur la moyenne	Vérifier l'étendue : si une valeur est très éloignée des autres, la **médiane** est plus représentative que la moyenne

Exercice 6 *QCM (4^e–2^{de})* Pour chaque question, une seule réponse est correcte.

1. La série $(3\,;\,7\,;\,5\,;\,3\,;\,8\,;\,4)$ a pour moyenne :

   A. 5 **B.** 5,5 **C.** 4,5 **D.** 6

2. Dans la même série, la médiane vaut :

   A. 3 **B.** 4 **C.** 4,5 **D.** 5

3. On tire une carte au hasard dans un jeu de 32 cartes. La probabilité d'obtenir un cœur est :

   A. $\frac{1}{2}$ **B.** $\frac{1}{4}$ **C.** $\frac{1}{8}$ **D.** $\frac{1}{32}$

4. Un dé équilibré à six faces est lancé. La probabilité d'obtenir un nombre **pair ou supérieur à** 3 est :

   A. $\frac{3}{6}$ **B.** $\frac{4}{6}$ **C.** $\frac{6}{6}$ **D.** $\frac{7}{6}$

5. On lance deux dés équilibrés et on note la somme. L'événement le plus probable est la somme égale à :

   A. 2 **B.** 6 **C.** 7 **D.** 12

6. Un sondage indique que $P(\text{vélo}) = 0{,}3$ et $P(\text{transport en commun}) = 0{,}5$. Sachant qu'aucun élève n'utilise les deux, la probabilité qu'un élève n'utilise ni vélo ni transport en commun est :

   A. 0,2 **B.** 0,5 **C.** 0,8 **D.** 1

Exercice 7 ***Vrai ou faux ?*** Pour chaque affirmation, dire si elle est vraie ou fausse et justifier :

A. La moyenne d'une série est toujours l'une des valeurs de la série.

B. La médiane d'une série de 10 valeurs est toujours la cinquième valeur après tri.

C. L'étendue mesure la dispersion autour de la moyenne.

D. Une probabilité peut valoir $1{,}2$ si l'événement est très probable.

E. Si $P(A) = 0{,}4$, alors $P(\bar{A}) = 0{,}6$.

F. Avec deux pièces équilibrées, l'événement « obtenir au moins un Pile » a pour probabilité $\frac{1}{2}$.

G. Dans une urne avec 5 boules rouges et 5 boules bleues, la probabilité de tirer une rouge est $\frac{1}{2}$, donc sur 10 tirages on obtient toujours exactement 5 rouges.

H. $P(A \cup B) = P(A) + P(B)$ est vrai pour tous événements A et B.

Tu bloques sur un exercice ? Consulte la fiche ***Que faire quand je bloque ?***
Tu veux retenir durablement ces méthodes ? Consulte la fiche ***Mémorisation et sciences cognitives****.*
Tu fais souvent les mêmes erreurs ? Remplis ton ***Carnet d'erreurs****.*