MATHÉMATIQUES

Maîtriser les probabilités et la statistique

Une fiche d'exercices pour dénombrer les issues, distinguer probabilité et fréquence, comprendre la moyenne

Fiche de cours : Probabilités et statistiques.

① Pourquoi cette fiche ?

Le hasard est l'un des sujets où l'intuition trompe le plus souvent. Spontanément, on tend à juger toutes les issues d'une expérience « aussi probables », à confondre une fréquence observée sur quelques essais avec une probabilité, à juger une suite de résultats « plus probable » parce qu'elle ressemble au désordre que l'on imagine. Quant à la moyenne, c'est un nombre dont le sens reste flou pour beaucoup : ce n'est ni la valeur la plus fréquente, ni nécessairement une valeur observée.

Cette fiche aborde les probabilités et la statistique en démasquant ces intuitions trompeuses. Elle s'appuie sur deux idées clés : **dénombrer explicitement** les issues d'une expérience pour ne pas se contenter d'une équiprobabilité supposée ; **simuler ou expérimenter** pour observer la stabilisation des fréquences vers la probabilité théorique. Pour la moyenne, elle propose une approche par redistribution (« la valeur qui égalise ») plus robuste que la simple formule.

② Expérience aléatoire et issues

Rappel : vocabulaire de base

Une **expérience aléatoire** est une expérience dont on ne peut prévoir le résultat avec certitude, mais dont on connaît tous les résultats possibles. Chacun de ces résultats possibles est appelé une **issue**. L'ensemble des issues est appelé l'**univers** de l'expérience. Un **événement** est un sous-ensemble de l'univers (une « partie » des issues). On dit qu'un événement est **réalisé** si l'issue de l'expérience appartient à cet événement.
Probabilité d'un événement. La probabilité d'un événement A est un nombre P(A) compris entre 0 et 1. Si toutes les issues sont équiprobables, on a :

$$\mathrm{P}(\mathrm{A}) = \frac{\text{nombre d'issues qui réalisent A}}{\text{nombre total d'issues}}.$$

Exercice 1 *Lister les issues* Pour chaque expérience aléatoire, lister toutes les issues possibles, puis indiquer combien il y en a.

a) On lance une pièce de monnaie.

b) On lance un dé à six faces.

c) On tire une carte d'un jeu de 32 cartes.

d) On lance deux pièces simultanément (et l'on note le résultat de chacune).

③ Équiprobabilité : quand est-elle légitime ?

Rappel : le piège de l'équiprobabilité abusive

On appelle **équiprobabilité** la situation où toutes les issues d'une expérience ont la même probabilité. C'est un cas particulier, qu'il faut justifier (par la symétrie de l'objet, par exemple : dé non truqué, pièce équilibrée, urne avec boules indiscernables).
Erreur fréquente. L'élève suppose souvent l'équiprobabilité par défaut, même quand elle ne tient pas. Par exemple, il pense que « obtenir 7 ou 12 comme somme de deux dés » sont des événements équiprobables. Or il y a **plusieurs façons** d'obtenir 7 (six couples : (1,6), (2,5), (3,4), (4,3), (5,2), (6,1)) mais une seule façon d'obtenir 12 (un seul couple : (6,6)). Les deux événements ne sont pas équiprobables.
Méthode sûre. Avant de calculer une probabilité, dénombrer explicitement les issues : faire un arbre, un tableau à double entrée, ou lister tous les couples.

Exercice 2 *Somme de deux dés* On lance deux dés équilibrés à six faces et l'on note la somme des résultats.

a) Combien y a-t-il d'issues possibles pour le couple (résultat du premier dé, résultat du second dé) ?

b) Compléter un tableau à double entrée donnant la somme pour chaque couple.

c) Combien d'issues réalisent l'événement « la somme vaut 7 » ? Quelle est la probabilité de cet événement ?

d) Même question pour les sommes 2, 6, 9, 12.

Exercice 3 *Arbre et issues* On tire successivement et avec remise deux boules d'une urne contenant 3 boules rouges et 2 boules bleues.

a) Construire un arbre des issues possibles pour la couleur de la première boule, puis de la seconde.

b) En supposant que les boules sont indiscernables au toucher, donner la probabilité de tirer une boule rouge à chaque tirage.

c) En déduire la probabilité de tirer deux boules rouges.

d) Quelle est la probabilité d'avoir au moins une boule bleue ?

④ Probabilité et fréquence

Rappel : la loi des grands nombres

La **fréquence** d'un événement, calculée sur n répétitions d'une expérience aléatoire, est le rapport

$$f = \frac{\text{nombre de fois où l'événement est réalisé}}{n}.$$

Sur peu d'essais, la fréquence peut s'écarter beaucoup de la probabilité théorique. Mais lorsqu'on multiplie les essais, la fréquence **tend à se stabiliser** vers la probabilité. C'est ce qu'on appelle la **loi des grands nombres**.
Conséquence pratique. Un événement de probabilité 0,01 n'est pas « impossible » : en moyenne, il se produit une fois sur cent. À la cent-millième tentative, il se sera produit environ mille fois. Réduire les petites probabilités à zéro est une erreur fréquente, qui sous-estime systématiquement les événements rares.

Exercice 4 *Comparer fréquences observées et probabilité* Léa lance 20 fois un dé équilibré et obtient les résultats suivants : 3, 1, 6, 4, 2, 5, 6, 3, 1, 4, 2, 5, 3, 6, 1, 2, 4, 5, 3, 6.

a) Calculer la fréquence d'apparition de chaque face.

b) Comparer ces fréquences avec la probabilité théorique de chaque face.

c) Si Léa avait lancé 1 000 fois le dé, peut-on prévoir avec certitude le nombre exact d'apparitions de la face 6 ? Et son ordre de grandeur ?

⑤ Démasquer le biais de représentativité

Rappel : le hasard ne ressemble pas toujours au hasard

Spontanément, on juge une suite de résultats « plus probable » quand elle ressemble à du hasard, c'est-à-dire qu'elle alterne ou ne montre pas de régularité apparente. Mais cette intuition est trompeuse.
Si l'on lance une pièce six fois, les deux suites suivantes sont **exactement** aussi probables :

PFPPFP et PPPPPP.

Chacune a une probabilité $\left(\frac{1}{2}\right)^6 = \frac{1}{64}$. Le fait que la première « ressemble » au hasard et que la seconde paraisse « ordonnée » n'a rien à voir avec leur probabilité. C'est ce qu'on appelle le **biais de représentativité**.

Exercice 5 *Vrai ou faux* Pour chaque affirmation, dire si elle est vraie ou fausse, en justifiant.

a) Si on lance 5 fois une pièce équilibrée et qu'on obtient 5 piles d'affilée, le sixième lancer a plus de chances d'être face que pile.

b) La suite PFFPFP est plus probable que la suite PPPPPP en six lancers.

c) Sur cinquante lancers, on obtient nécessairement environ 25 piles et 25 faces.

d) Plus on multiplie les essais, plus la fréquence de pile s'approche de $\frac{1}{2}$.

⑥ Comprendre la moyenne

Rappel : la moyenne, ou « valeur qui égalise »

La moyenne d'une série de valeurs est définie par :

$$\overline{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}.$$

Cette formule cache une idée intuitive plus parlante : la moyenne est la valeur qui, si on la donnait à tout le monde, donnerait le même **total**. C'est la valeur qui « égalise » quand on redistribue.
Erreurs fréquentes.

- Confondre la moyenne avec le **mode** (la valeur la plus fréquente). Les deux notions coïncident dans certaines séries (par exemple si toutes les valeurs sont égales), mais pas en général.

- Croire que la moyenne doit être une valeur effectivement présente dans la série. Or la moyenne de la série (2 ; 4 ; 6) est 4, qui est dans la série, mais celle de (1 ; 2 ; 4) est $\frac{7}{3}$, qui n'y est pas.

Exercice 6 *Calculer une moyenne*

a) Calculer la moyenne de la série (8 ; 10 ; 12 ; 14 ; 16).

b) Calculer la moyenne de la série (2 ; 5 ; 5 ; 7 ; 11).

c) Une élève a obtenu 12, 14, 9, 16 aux quatre devoirs du trimestre, comptant respectivement coefficients 1, 2, 1, 3. Calculer sa moyenne pondérée.

Exercice 7 *Démasquer la confusion moyenne/mode* Voici la liste des notes obtenues par les 10 élèves d'un groupe : 5, 8, 8, 8, 10, 12, 13, 14, 17, 18.

a) Calculer la moyenne de cette série.

b) Quelle est la valeur la plus fréquente (mode) de la série ?

c) La moyenne est-elle égale à la valeur la plus fréquente ?

Exercice 8 *Effet d'une valeur extrême* Une classe de 10 élèves a obtenu les notes suivantes au dernier devoir : 11, 12, 12, 13, 13, 13, 14, 14, 15, 15. Calculer la moyenne. On apprend qu'il faut ajouter une onzième note : 0 (un élève absent à qui la note 0 a été attribuée).

a) Calculer la nouvelle moyenne.

b) De combien la moyenne a-t-elle diminué ? Comment expliquer cet effet ?

7 Pour s'auto-évaluer

Rappel : cinq questions à se poser

Avant et pendant un calcul de probabilité ou de moyenne, prendre l'habitude de se poser ces cinq questions.

- Ai-je dénombré explicitement toutes les issues, ou suis-je en train de supposer qu'elles sont équiprobables sans le justifier ?
- Suis-je en train de confondre la fréquence (observée sur des essais) et la probabilité (valeur théorique) ?
- Une suite de résultats « ordonnée » me semble-t-elle moins probable qu'une suite « en désordre » ? C'est une intuition trompeuse : deux suites de longueur fixe sont équiprobables.

- Pour une moyenne, suis-je en train de la confondre avec le mode (la valeur la plus fréquente) ? Une valeur extrême peut-elle fausser la moyenne ?
- Mon résultat est-il cohérent ? Une probabilité est toujours dans $[0\ ;\ 1]$; une moyenne est toujours comprise entre la plus petite et la plus grande valeur de la série.

Solutions des exercices

Corrigé de l'exercice 1.

a) Issues : pile, face. Il y en a 2.

b) Issues : 1, 2, 3, 4, 5, 6. Il y en a 6.

c) 32 issues (chaque carte est une issue distincte).

d) Issues : (P, P), (P, F), (F, P), (F, F). Il y en a 4. *Remarque :* les issues (P, F) et (F, P) sont distinctes parce qu'on note séparément le résultat de chaque pièce.

Corrigé de l'exercice 2.

a) $6 \times 6 = 36$ issues.

b) Les sommes vont de 2 (couple (1,1)) à 12 (couple (6,6)). Voici le tableau :

+	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

c) Les couples donnant 7 sont (1,6), (2,5), (3,4), (4,3), (5,2), (6,1) : 6 couples. Probabilité : $\frac{6}{36} = \frac{1}{6}$.

d) Somme 2 : 1 couple, probabilité $\frac{1}{36}$. Somme 6 : 5 couples, probabilité $\frac{5}{36}$. Somme 9 : 4 couples, probabilité $\frac{4}{36} = \frac{1}{9}$. Somme 12 : 1 couple, probabilité $\frac{1}{36}$.

Conclusion. Les sommes ne sont **pas équiprobables** : 7 est six fois plus probable que 12. L'erreur consistant à supposer une équiprobabilité conduirait à conclure que P(somme = 7) = P(somme = 12), ce qui est faux.

Corrigé de l'exercice 3.

a) L'arbre comporte deux niveaux : au premier, deux branches (R, B) ; à chaque branche correspondent à nouveau deux branches au second niveau (R, B). Quatre chemins possibles : RR, RB, BR, BB.

b) À chaque tirage, $P(R) = \frac{3}{5}$ et $P(B) = \frac{2}{5}$.

c) $P(RR) = \frac{3}{5} \times \frac{3}{5} = \frac{9}{25}$.

d) L'événement contraire de « au moins une bleue » est « aucune bleue », c'est-à-dire deux rouges. $P(\text{au moins une bleue}) = 1 - P(RR) = 1 - \frac{9}{25} = \frac{16}{25}$.

Corrigé de l'exercice 4.

a) Comptage : $1 \to 3$ fois ; $2 \to 3$; $3 \to 4$; $4 \to 3$; $5 \to 3$; $6 \to 4$. Fréquences : $\dfrac{3}{20} = 0{,}15$ pour les faces 1, 2, 4, 5 ; $\dfrac{4}{20} = 0{,}20$ pour 3 et 6.

b) Probabilité théorique de chaque face : $\dfrac{1}{6} \approx 0{,}167$. Les fréquences observées s'écartent de cette valeur, mais elles en sont assez proches. Sur 20 lancers, l'écart est normal.

c) On ne peut pas prévoir le nombre exact (le hasard reste). Mais l'ordre de grandeur est prévisible : environ $\dfrac{1\,000}{6} \approx 167$ apparitions de la face 6. Sur 1 000 lancers, l'écart relatif à la probabilité théorique sera bien plus petit que sur 20 : c'est la loi des grands nombres.

Corrigé de l'exercice 5.

a) **Faux.** La pièce « n'a pas de mémoire ». À chaque lancer, $P(P) = P(F) = \dfrac{1}{2}$, indépendamment des lancers précédents. C'est une autre forme de biais (parfois appelé « erreur du joueur »).

b) **Faux.** Toute suite donnée de six lancers a pour probabilité $\dfrac{1}{64}$. Les deux suites sont équiprobables.

c) **Faux.** Sur cinquante lancers, on n'obtient pas nécessairement exactement 25 / 25. La répartition est aléatoire ; on attend en moyenne 25 / 25, mais l'écart 20 / 30 est tout à fait plausible.

d) **Vrai.** C'est la loi des grands nombres.

Corrigé de l'exercice 6.

a) Moyenne $= \dfrac{8 + 10 + 12 + 14 + 16}{5} = \dfrac{60}{5} = 12$.

b) Moyenne $= \dfrac{2 + 5 + 5 + 7 + 11}{5} = \dfrac{30}{5} = 6$. *Remarque : la valeur 6 n'apparaît pas dans la série, ce qui est normal pour une moyenne.*

c) Moyenne $= \dfrac{1 \times 12 + 2 \times 14 + 1 \times 9 + 3 \times 16}{1 + 2 + 1 + 3} = \dfrac{12 + 28 + 9 + 48}{7} = \dfrac{97}{7} \approx 13{,}9$.

Corrigé de l'exercice 7.

a) Moyenne $= \dfrac{5 + 8 + 8 + 8 + 10 + 12 + 13 + 14 + 17 + 18}{10} = \dfrac{113}{10} = 11{,}3$.

b) La valeur la plus fréquente est 8 (présente trois fois).

c) Non, la moyenne (11,3) et le mode (8) sont **distincts**. Confondre les deux est une erreur fréquente.

Corrigé de l'exercice 8.

a) Moyenne initiale $= \dfrac{132}{10} = 13{,}2$. Nouvelle moyenne avec 0 : $\dfrac{132 + 0}{11} = \dfrac{132}{11} = 12$.

b) La moyenne a diminué de 1,2 point. La valeur extrême 0, très éloignée des autres notes (qui se situent entre 11 et 15), tire fortement la moyenne vers le bas. La moyenne est sensible aux valeurs extrêmes : c'est l'une de ses limites.