Plaquette 1 – Fréquences observées et fluctuation

Échantillonnage et fluctuation — Seconde

Corrections détaillées

Boîte à outils

Fréquence observée. Sur un échantillon de taille n, si un caractère apparaît k fois :

$$f = \frac{k}{n} \qquad \text{avec } 0 \leq f \leq 1$$

On l'écrit aussi en pourcentage : $f = 0{,}34$ signifie 34 %. L'écart entre f et p se lit en **points** : de 0,46 à 0,5, il y a 4 points.

Fluctuation d'échantillonnage. Deux échantillons de même taille, tirés dans la même population ou par la même expérience aléatoire, donnent en général des fréquences différentes. Ce n'est pas une erreur : c'est l'effet du hasard.

Loi des grands nombres (version vulgarisée). Lorsque n est grand, sauf exception, la fréquence observée est proche de la probabilité. Elle ne dit rien sur un petit échantillon, et elle ne dit pas que f finit par être **égale** à p.

Ce que montrent les simulations. Plus les échantillons sont grands, moins leurs fréquences sont dispersées autour de p : leur étendue et leur écart type diminuent quand n augmente.

Modèle et réalité. La probabilité est une hypothèse du modèle (pièce « équilibrée », tirage « au hasard ») ; la fréquence est une mesure faite sur la réalité.

Correction de l'exercice 1 – Calculer une fréquence observée

La fréquence observée est le nombre de « face » divisé par le nombre de lancers :

$$f = \frac{46}{100} = 0{,}46$$

soit 46 %. L'écart avec la probabilité est $|0{,}46 - 0{,}5| = 0{,}04$, c'est-à-dire 4 points de pourcentage.

Cet écart ne signifie pas que la pièce est truquée : sur 100 lancers d'une pièce équilibrée, on n'obtient presque jamais exactement 50 « face ». La fréquence fluctue autour de 0,5 d'une série de lancers à l'autre.

Réponse. $f = 0{,}46$; écart de 0,04, soit 4 points.

Correction de l'exercice 2 – Fréquence, effectif et pourcentage

a) La fréquence est l'effectif du caractère divisé par la taille de l'échantillon :
$f = \frac{85}{250} = 0{,}34$, soit 34 %.

b) On repart de la même relation, $f = \frac{k}{n}$, mais c'est cette fois l'effectif k qui est inconnu. On multiplie la fréquence par la taille de l'échantillon :

$$k = f \times n = 0{,}315 \times 400 = 126$$

Ainsi 126 élèves interrogés sur 400 pratiquent un sport en club. **Contrôle** : $\frac{126}{400} = 0{,}315$.

Réponse. a) 0,34, soit 34 % ; b) 126 élèves.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

Correction de l’exercice 3 – Probabilité ou fréquence ?

a) $\frac{1}{6}$ est une **probabilité** : elle découle de l’hypothèse « le dé est équilibré » (six faces de même chance), sans aucun lancer.

b) $\frac{13}{60}$ est une **fréquence observée** : elle résulte de 60 lancers réels et changerait si l’on recommençait.

c) 0,3 est une **probabilité** : le tirage « au hasard » donne la même chance à chacune des 10 boules, et 3 sont rouges, d’où $\frac{3}{10}$.

d) 0,42 est une **fréquence observée** sur un échantillon de 500 personnes ; la proportion dans toute la population reste inconnue.

En résumé, une probabilité vient d’un modèle et ne fluctue pas, une fréquence vient d’une observation et fluctue.

Réponse. a) probabilité ; b) fréquence ; c) probabilité ; d) fréquence.

Correction de l’exercice 4 – Deux échantillons de tailles différentes

Le tirage se fait au hasard et la moitié des boules sont blanches : la probabilité de tirer une blanche est $p = 0{,}5$.

Pour le premier échantillon, $f_1 = \frac{21}{50} = 0{,}42$, soit un écart de 0,08 avec p. Pour le second, $f_2 = \frac{243}{500} = 0{,}486$, soit un écart de 0,014.

Aucune des deux fréquences n’est égale à 0,5 : c’est la fluctuation d’échantillonnage. Mais la fréquence du grand échantillon est bien plus proche de p, comme l’annonce la loi des grands nombres : quand n est grand, la fréquence observée est, sauf exception, proche de la probabilité.

Réponse. $f_1 = 0{,}42$ (écart 0,08) et $f_2 = 0{,}486$ (écart 0,014).

Correction de l’exercice 5 – Vingt échantillons de taille 50

a) En parcourant la série, la plus petite fréquence est 0,38 et la plus grande 0,62. L’étendue est leur différence :

$$0{,}62 - 0{,}38 = 0{,}24$$

b) Seules deux fréquences sortent de l’intervalle $[0{,}4\,;0{,}6]$: 0,38 (en dessous) et 0,62 (au-dessus). Les valeurs 0,4 et 0,6, qui apparaissent deux fois chacune, sont incluses. Il reste donc $20 - 2 = 18$ échantillons, soit $\frac{18}{20} = 90\,\%$ des échantillons.

Sur le graphique, les points se répartissent de part et d’autre de la droite $f = 0{,}5$, sans jamais tomber exactement dessus de façon systématique : avec 50 lancers, la fréquence fluctue nettement.

Réponse. a) 0,38 et 0,62, étendue 0,24 ; b) 18 échantillons, soit 90 %.

Correction de l’exercice 6 – Même expérience, échantillons de taille 500

a) L’étendue est la différence entre la plus grande et la plus petite fréquence :

$$0{,}562 - 0{,}468 = 0{,}094$$

Elle est environ 2,5 fois plus petite que pour les échantillons de taille 50, car $0{,}24 \div 0{,}094 \approx 2{,}55$. Sur le graphique, à la même échelle, les points sont beaucoup plus resserrés autour de la droite $f = 0{,}5$.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s’expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

b) Plus les échantillons sont grands, moins leurs fréquences sont dispersées : la fluctuation d'échantillonnage diminue quand n augmente. C'est ce que dit la loi des grands nombres, sous sa forme vulgarisée : pour n grand, sauf exception, la fréquence observée est proche de la probabilité 0,5.

Réponse. a) Étendue 0,094, contre 0,24 ; b) la fluctuation diminue quand la taille augmente.

Correction de l'exercice 7 – La fréquence d'un « 6 » au fil des lancers

a) À chaque moment, on divise le nombre de « 6 » par le nombre de lancers effectués :

Nombre de lancers	10	100	500	1 000	5 000
Fréquence de « 6 »	0	0,14	0,164	0,157	0,162

Par exemple $\dfrac{809}{5\,000} = 0{,}1618 \approx 0{,}162$.

b) Le dé étant équilibré, la probabilité d'obtenir « 6 » est

$$p = \frac{1}{6} \approx 0{,}167$$

Après 10 lancers, la fréquence est nulle, très loin de p : sur si peu de lancers, tout peut arriver. Ensuite, les fréquences se rapprochent de 0,167 en oscillant (0,164 puis 0,157 puis 0,162). C'est la loi des grands nombres : pour un grand nombre de lancers, la fréquence est, sauf exception, proche de la probabilité.

Réponse. a) 0 ; 0,14 ; 0,164 ; 0,157 ; 0,162 ; b) vers $\dfrac{1}{6} \approx 0{,}167$.

Correction de l'exercice 8 – Comprendre la loi des grands nombres

a) **Faux.** La pièce n'a pas de mémoire : à chaque lancer, la probabilité d'obtenir « face » reste 0,5. La loi des grands nombres ne « compense » pas les écarts passés ; elle dit seulement que, sur un très grand nombre de lancers, ces quelques lancers pèsent peu dans la fréquence.

b) **Faux.** La fréquence est proche de 0,5, mais rien n'oblige l'effectif à valoir exactement 5 000 ; un résultat comme 4 968 ou 5 031 est tout à fait normal.

c) **Vrai.** C'est l'énoncé même de la loi des grands nombres : pour n grand, sauf exception, la fréquence observée est proche de la probabilité.

d) **Faux.** La loi des grands nombres ne dit rien sur 4 lancers. Obtenir quatre « face » est rare mais possible : cela arrive en moyenne une fois sur 16 séries de quatre lancers.

Réponse. a) faux ; b) faux ; c) vrai ; d) faux.

Correction de l'exercice 9 – Mille échantillons de taille 100

a) Les classes comprises dans $[0{,}4\,;0{,}6[$ sont les quatre classes centrales. On additionne leurs effectifs :

$$119 + 319 + 351 + 169 = 958$$

soit $\dfrac{958}{1\,000} = 95{,}8\,\%$ des échantillons.

b) Seules la première classe (1 échantillon) et la dernière (2 échantillons) sont concernées : $1 + 2 = 3$ échantillons sur 1 000, soit 0,3 %.

Avec 100 lancers d'une pièce équilibrée, la fréquence de « face » tombe donc presque toujours entre 0,4 et 0,6 ; un écart de plus de 0,15 avec 0,5 n'arrive qu'environ 3 fois sur mille. L'histogramme a une forme « en cloche » centrée sur 0,5.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

Réponse. a) 95,8 % ; b) 0,3 % (3 échantillons).

Correction de l’exercice 10 – Un résultat surprenant ?

La fréquence obtenue par Léa est $f = \dfrac{68}{100} = 0{,}68$. Elle appartient à la classe $[0{,}65\,;0{,}70[$.

Dans la simulation, seuls 2 échantillons sur 1 000 sont tombés dans cette classe, et aucun au-delà :

$$\frac{2}{1\,000} = 0{,}2\,\%$$

Pour une pièce équilibrée, une fréquence aussi éloignée de 0,5 est donc très rare. Cela ne prouve pas que la pièce est truquée (le hasard peut produire un résultat rare), mais cela invite à douter du modèle « pièce équilibrée » et à recommencer l’expérience avec beaucoup plus de lancers pour en avoir le cœur net.

Réponse. $f = 0{,}68$: résultat très rare (environ 0,2 % des échantillons simulés) ; on peut douter de l’équilibre de la pièce.

Correction de l’exercice 11 – Six cents lancers d’un dé

a) Pour un dé équilibré, chaque face a la probabilité $\dfrac{1}{6}$. Sur 600 lancers, on attend environ $600 \times \dfrac{1}{6} = 100$ apparitions de chaque face.

b) On divise chaque effectif par 600 :

Face	1	2	3	4	5	6
Fréquence	0,147	0,197	0,18	0,18	0,137	0,16

La probabilité vaut $\dfrac{1}{6} \approx 0{,}167$. Le plus simple est de comparer les effectifs aux 100 apparitions attendues : les écarts sont 12, 18, 8, 8, 18 et 4. Les faces 2 (trop souvent) et 5 (pas assez) s’écartent le plus, de 18 apparitions chacune, soit en fréquence

$$\frac{18}{600} = 0{,}03$$

Ces écarts de 3 points restent ordinaires sur 600 lancers : ils ne suffisent pas à affirmer que le dé est mal équilibré.

Réponse. a) Environ 100 fois ; b) 0,147 ; 0,197 ; 0,18 ; 0,18 ; 0,137 ; 0,16. Les faces 2 et 5 s’écartent le plus (18 apparitions, soit 0,03).

Correction de l’exercice 12 – Modèle et réalité

a) Le tirage se fait « au hasard » : on suppose que chaque jeton a la même chance d’être tiré. Cette équiprobabilité est une **hypothèse du modèle**, qu’on ne démontre pas. On en déduit :

$$p = \frac{12}{40} = 0{,}3$$

b) La fréquence observée est $f = \dfrac{29}{80} = 0{,}3625$.

Tom confond la fréquence, mesurée sur ses 80 tirages, avec la probabilité, fixée par la composition de l’urne. S’il recommençait 80 tirages, il obtiendrait sans doute une autre fréquence. Son résultat s’écarte de 0,0625 de la probabilité 0,3, ce qui reste une fluctuation ordinaire pour un échantillon de cette taille.

Réponse. a) $p = 0{,}3$, sous l’hypothèse d’équiprobabilité des jetons ; b) $f = 0{,}3625$: c’est

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s’expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

une fréquence, pas la probabilité.

Correction de l’exercice 13 – Lire une fonction de simulation

a) La fonction *random()* renvoie un nombre décimal au hasard dans [0 ; 1[. La condition *random() < 0.5* est vraie dans la moitié des cas : elle simule un lancer de pièce équilibrée, « vrai » correspondant à « face ». La variable *succes* compte les « face » : elle part de 0 et augmente de 1 à chaque « face ».

b) La boucle répète l’expérience 1 000 fois. À la fin, la fonction renvoie le nombre de « face » divisé par 1 000, c’est-à-dire la **fréquence observée** de « face » sur 1 000 lancers simulés, et non leur nombre.

c) Chaque appel effectue de nouveaux tirages au hasard : les deux séries de 1 000 lancers sont différentes, donc leurs fréquences aussi. C’est la fluctuation d’échantillonnage. Pour $n = 1\,000$, les valeurs renvoyées restent toutefois, sauf exception, proches de 0,5.

Réponse. a) Un lancer de pièce équilibrée ; *succes* compte les « face » ; b) la fréquence de « face » sur 1 000 lancers ; c) les tirages changent à chaque appel (fluctuation).

Correction de l’exercice 14 – Petits et grands échantillons

a) Quelle que soit la taille, la moyenne des fréquences est très proche de 0,3 : en moyenne, les échantillons ne se trompent pas ; ils sont simplement plus ou moins dispersés autour de la probabilité.

b) L’écart type mesure la dispersion des fréquences. D’une colonne à la suivante, la taille est multipliée par 4 et l’écart type est à peu près divisé par 2 :

$$0{,}085 \div 0{,}045 \approx 1{,}9 \qquad 0{,}045 \div 0{,}025 = 1{,}8 \qquad 0{,}025 \div 0{,}012 \approx 2{,}1$$

Des échantillons plus grands donnent des fréquences plus groupées autour de 0,3.

c) Il faut un écart type inférieur à 0,03. C’est le cas pour $n = 400$ (0,025) et $n = 1\,600$, mais pas pour $n = 100$ (0,045). La plus petite taille du tableau qui convient est donc $n = 400$.

Réponse. a) Toujours proches de 0,3 ; b) l’écart type est à peu près divisé par 2 ; c) $n = 400$.

Correction de l’exercice 15 – Synthèse : la roue de la kermesse

a) Les quatre secteurs ont la même chance d’être désignés et un seul est gagnant :

$$p = \frac{1}{4} = 0{,}25$$

b) Le premier jour, $f_1 = \frac{14}{40} = 0{,}35$. L’écart avec p est de 0,1, soit 10 points. Sur seulement 40 parties, un tel écart peut parfaitement venir du hasard : c’est la fluctuation d’échantillonnage. Une seule journée ne permet pas de conclure que la roue est déréglée.

c) Sur le mois, $f_2 = \frac{489}{2\,000} = 0{,}2445$, soit un écart d’environ 0,006 avec p.

d) La fréquence du mois, calculée sur 2 000 parties, est la plus fiable : d’après la loi des grands nombres, sur un grand nombre de parties la fréquence est, sauf exception, proche de la probabilité. Elle est ici très proche de 0,25 : rien n’indique que la roue soit mal équilibrée, et l’écart du premier jour était une fluctuation ordinaire.

Réponse. a) $p = 0{,}25$; b) $f_1 = 0{,}35$, écart dû à la fluctuation sur 40 parties ; c) $f_2 = 0{,}2445$; d) la fréquence du mois, car n est grand.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s’expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).