Plaquette 4 – Simuler avec Python : fonctions et séries de fréquences

Échantillonnage et fluctuation — Seconde

Corrections détaillées

Boîte à outils

Une fonction par niveau.

— *experience()* : renvoie le résultat d'**une** expérience (1 pour un succès, 0 sinon ; ou un nombre, une chaîne, un booléen).
— *frequence(n)* : répète n fois l'expérience et renvoie la fréquence de succès.
— une fonction qui appelle *frequence(n)* k fois fournit une **série de k fréquences**.

Moyenne et écart type au fil de la boucle. On accumule la somme s des valeurs et la somme s_2 de leurs carrés, puis :

$$\overline{x} = \frac{s}{k} \qquad \sigma = \sqrt{\frac{s_2}{k} - \overline{x}^2}$$

(la variance est la moyenne des carrés moins le carré de la moyenne). Avec *from math import sqrt*.

Pièges fréquents. Un *return* placé dans la boucle arrête la fonction au premier tour ; un compteur initialisé dans la boucle est remis à zéro à chaque tour ; *randint(a, b)* inclut ses deux bornes.

Correction de l'exercice 1 – Expérience et répétition

a) *experience()* lance un dé ; si le résultat vaut 5 ou 6, elle renvoie 1 (succès), sinon 0. Le second *return* n'est atteint que lorsque le premier n'a pas été exécuté. Deux faces sur six donnent un succès :

$$P(\text{succès}) = \frac{2}{6} = \frac{1}{3}$$

b) *total* additionne les valeurs renvoyées, c'est-à-dire compte les succès, puisque chaque succès ajoute 1 et chaque échec 0. La fonction renvoie donc la fréquence de succès sur n expériences. Pour $n = 10\,000$, d'après la loi des grands nombres, elle sera, sauf exception, proche de $\frac{1}{3} \approx 0{,}333$.

Réponse. a) 1 si le dé donne 5 ou 6, avec la probabilité $\frac{1}{3}$; b) la fréquence de succès, proche de 0,333.

Correction de l'exercice 2 – La moyenne d'une série de lancers

a) La variable s accumule les faces : 2, puis 8, 11 et enfin 16. La fonction renvoie

$$\frac{16}{4} = 4$$

b) Les moyennes obtenues s'approchent de 3,5 quand le nombre de lancers augmente : l'écart vaut 0,08 pour 100 lancers, environ 0,02 pour 1 000 lancers et moins de 0,01 pour 10 000 lancers. Comme les fréquences de chaque face se rapprochent de $\frac{1}{6}$, la moyenne des faces se rapproche de la moyenne $\frac{1+2+3+4+5+6}{6} = 3{,}5$. Avec 4 lancers, en revanche, la moyenne (4) peut s'en écarter nettement.

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) 4 ; b) les moyennes se rapprochent de 3,5 quand n augmente.

Correction de l’exercice 3 – Moyenne d’une série de fréquences

a) La boucle appelle 20 fois *frequence(50)*, et chaque appel simule 50 lancers : au total $20 \times 50 = 1\,000$ lancers.

b) s accumule les quatre fréquences :

$$\frac{0{,}42 + 0{,}56 + 0{,}5 + 0{,}48}{4} = \frac{1{,}96}{4} = 0{,}49$$

c) Les écarts à 0,5 sont tantôt positifs, tantôt négatifs, et se compensent en partie dans la moyenne. D’ailleurs, quand les échantillons ont tous la même taille, la moyenne des quatre fréquences est la fréquence de « face » sur l’ensemble des 200 lancers : elle repose sur un échantillon quatre fois plus grand, donc fluctue moins.

Réponse. a) 1 000 lancers ; b) 0,49 ; c) les écarts se compensent : c’est la fréquence sur 200 lancers.

Correction de l’exercice 4 – Écart type d’une série de fréquences

a) s est la somme des k fréquences et *s2* la somme de leurs carrés : chaque fréquence f est ajoutée à s, et son carré $f * f$ à *s2*.

b) On obtient $s = 0{,}4 + 0{,}5 + 0{,}6 + 0{,}5 = 2$ et $s_2 = 0{,}16 + 0{,}25 + 0{,}36 + 0{,}25 = 1{,}02$. La moyenne vaut $m = \frac{2}{4} = 0{,}5$, puis

$$\sigma = \sqrt{\frac{1{,}02}{4} - 0{,}5^2} = \sqrt{0{,}255 - 0{,}25} = \sqrt{0{,}005} \approx 0{,}071$$

Contrôle par la définition : les écarts à la moyenne sont $-0{,}1$, 0, 0,1, 0, dont la moyenne des carrés vaut $\frac{0{,}01 + 0 + 0{,}01 + 0}{4} = 0{,}005$.

c) L’écart type mesure la dispersion des fréquences autour de leur moyenne : plus il est petit, plus les échantillons donnent des résultats proches les uns des autres.

Réponse. a) La somme des fréquences et la somme de leurs carrés ; b) $\sigma = \sqrt{0{,}005} \approx 0{,}071$; c) la dispersion des fréquences.

Correction de l’exercice 5 – Adapter une fonction à une autre expérience

a) On compare la chaîne renvoyée à "*face*" et on compte les égalités :

```
def frequence_face(n):
    c = 0
    for i in range(n):
        if lancer() == "face":
            c = c + 1
    return c / n
```

b) Il suffit de changer le seuil : *if random() < 0.6:*. La fonction renvoie alors « face » avec la probabilité 0,6.

c) On numérote les 10 boules : les numéros 1 à 3 sont rouges, 4 à 8 vertes, 9 et 10 jaunes. Le tirage d’un numéro avec *randint(1, 10)* est équiprobable :

```
from random import randint

def tirage():
    b = randint(1, 10)
    if b <= 3:
```

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).

```
        return "rouge"
    if b <= 8:
        return "verte"
    return "jaune"
```

On a bien $P(\text{rouge}) = \frac{3}{10}$, $P(\text{verte}) = \frac{5}{10}$ et $P(\text{jaune}) = \frac{2}{10}$.

Réponse. a) Compter les *lancer() == "face"* puis diviser par *n* ; b) *random() < 0.6* ; c) *randint(1, 10)* : 1–3 rouge, 4–8 verte, 9–10 jaune.

Correction de l’exercice 6 – Combien de lancers pour un 6, en moyenne ?

a) La fonction joue k parties, additionne dans s le nombre de lancers de chacune, puis divise par k : elle renvoie le nombre moyen de lancers nécessaires pour obtenir un « 6 », sur k parties simulées.

b) Plus le nombre de parties est grand, plus les moyennes semblent se stabiliser autour de 6 : on conjecture qu’il faut en moyenne 6 lancers pour obtenir un « 6 ». C’est cohérent avec la probabilité $\frac{1}{6}$: sur un grand nombre de lancers, un lancer sur six environ donne un « 6 ».

c) C’est faux. Le dé n’a pas de mémoire : au sixième lancer, la probabilité d’obtenir « 6 » vaut toujours $\frac{1}{6}$. La moyenne de 6 lancers cache une forte dispersion : certaines parties se terminent au premier lancer, d’autres demandent 15 ou 20 lancers.

Réponse. a) Le nombre moyen de lancers sur k parties ; b) environ 6 ; c) faux, le dé n’a pas de mémoire.

Correction de l’exercice 7 – Le plus grand de deux dés

a) On tire deux entiers et on renvoie le plus grand :

```
from random import randint

def maximum deux des():
    return max(randint(1, 6), randint(1, 6))
```

b) Le maximum vaut 6 dès qu’au moins un dé donne 6. Dans le tableau, la ligne du 6 du premier dé compte 6 issues et la colonne du 6 du second dé aussi, la case $(6\,;6)$ étant commune :

$$P(\max = 6) = \frac{6 + 6 - 1}{36} = \frac{11}{36} \approx 0{,}306$$

c) La fréquence observée vaut $\frac{311}{1\,000} = 0{,}311$, très proche de 0,306 : l’écart de 0,005 est une fluctuation normale pour 1 000 simulations.

Réponse. a) *return max(randint(1, 6), randint(1, 6))* ; b) $\frac{11}{36} \approx 0{,}306$; c) 0,311, très proche.

Correction de l’exercice 8 – Une fonction qui renvoie un booléen

a) $a == b$ est une comparaison : la fonction renvoie un booléen, *True* lorsque les deux dés donnent le même résultat (un « double »), *False* sinon.

b) Parmi les 36 issues équiprobables, les doubles sont $(1,1)$, $(2,2)$, …, $(6,6)$, soit 6 issues : $P = \frac{6}{36} = \frac{1}{6}$.

c) Un booléen peut servir directement de condition :

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).

```
def frequence_doubles(n):
    c = 0
    for i in range(n):
        if double():
            c = c + 1
    return c / n
```

La fréquence observée vaut $\frac{152}{1\,000} = 0{,}152$, à comparer à

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

L'écart, 0,015, reste du même ordre que les fluctuations observées sur 1 000 répétitions.

Réponse. a) Un booléen, *True* pour un double ; b) $\frac{1}{6}$; c) compter les *True* ; 0,152, proche de 0,167.

Correction de l'exercice 9 – Exécuter une partie à la main

a) On suit les variables :

Lancer	*random()*	« face » ?	*gain* après le tour
1	0,8	non	5
2	0,6	non	2
3	0,3	oui	—

Au troisième lancer, la condition *random() < 0.5* est vraie : la fonction renvoie la valeur de *gain* à ce moment, soit 2 €, ce qui correspond bien à « face au troisième lancer ».

b) La pièce donne « pile » quatre fois : *gain* passe successivement à 5, 2, 1 puis reste à 1, et *lancers* atteint 4. La condition *lancers < 4* devient fausse, la boucle s'arrête et la fonction renvoie 0 : le joueur ne gagne rien.

Réponse. a) La fonction renvoie 2 (face au troisième lancer) ; b) 0.

Correction de l'exercice 10 – Un return mal placé

a) L'instruction *return* est indentée dans la boucle : elle est exécutée dès le premier tour. Une instruction *return* termine immédiatement la fonction, donc un seul lancer est simulé. Après ce lancer, *c* vaut 0 ou 1, et la fonction renvoie

$$\frac{0}{n} = 0 \qquad \text{ou} \qquad \frac{1}{n}$$

b) Il faut sortir le *return* de la boucle, en le plaçant au même niveau que *for* :

```
def frequence_pile(n):
    c = 0
    for i in range(n):
        if random() < 0.5:
            c = c + 1
    return c / n
```

La boucle s'exécute alors *n* fois avant que la fréquence ne soit renvoyée.

Réponse. a) Le *return* dans la boucle arrête la fonction au premier tour ; b) placer *return c / n* après la boucle.

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 11 – Gain moyen d’un jeu

a) La fonction additionne les gains de n parties et divise par n : elle renvoie le gain moyen par partie sur n parties simulées.

b) Si la fréquence des « 6 » sur les 10 000 parties est f, le gain total vaut $5 \times 10\,000 f$ et le gain moyen $5f$. On a donc $5f = 0{,}832$, soit $f = 0{,}1664$, très proche de $\frac{1}{6}$: on retrouve bien

$$5 \times \frac{1}{6} \approx 0{,}83$$

c) Le joueur gagne en moyenne environ 0,83 € par partie mais paie 1 € pour jouer : il perd en moyenne environ 0,17 € par partie. Sur un grand nombre de parties, le jeu est désavantageux pour le joueur, et rapporte à l’organisateur.

Réponse. a) Le gain moyen par partie ; b) $0{,}832 = 5f$ avec $f \approx \frac{1}{6}$; c) non : perte moyenne d’environ 0,17 € par partie.

Correction de l’exercice 12 – Proportion d’échantillons proches de la probabilité

a) $|f - 0{,}5|$ est la distance entre la fréquence observée et la probabilité 0,5. La condition est vraie quand cette distance est au plus 0,1, c’est-à-dire quand

$$0{,}4 \leq f \leq 0{,}6$$

La fonction compte les échantillons « proches » de 0,5 et renvoie leur proportion parmi les k échantillons.

b) Avec 25 lancers, 79 % des échantillons seulement ont une fréquence entre 0,4 et 0,6 : environ un sur cinq s’en écarte davantage. Avec 100 lancers, c’est le cas de 95,2 % d’entre eux, et avec 400 lancers de tous les échantillons simulés. Plus l’échantillon est grand, plus la fréquence est, sauf exception, proche de la probabilité : c’est la loi des grands nombres observée par simulation.

Réponse. a) $0{,}4 \leq f \leq 0{,}6$; b) la proportion d’échantillons proches de 0,5 augmente avec n : 79 %, 95,2 %, 100 %.

Correction de l’exercice 13 – Simuler un tirage avec remise

a) Chaque tirage avec remise se fait dans la même urne : une boule est rouge avec la probabilité $\frac{4}{10} = 0{,}4$. On simule les deux tirages et on teste les deux conditions :

```
from random import random

def deux_rouges():
    b1 = random() < 0.4
    b2 = random() < 0.4
    return b1 and b2
```

b) Sur 5 000 simulations, on peut s’attendre à environ

$$5\,000 \times 0{,}16 = 800$$

réalisations de l’événement.

c) Sur 50 simulations, on attend environ $50 \times 0{,}16 = 8$ réalisations ; l’élève en a obtenu 14, soit une fréquence de 0,28. Sur un si petit nombre d’essais, la fluctuation est forte, et cela ne suffit pas à rejeter la probabilité admise. Il faut relancer la simulation avec un grand nombre de tirages, par exemple 10 000, pour comparer sérieusement la fréquence à 0,16.

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) *return (random() < 0.4) and (random() < 0.4)* ; b) environ 800 ; c) 50 essais, c'est trop peu pour conclure.

Correction de l'exercice 14 – Lire une fonction mystère

a) On lance deux dés ; l'événement est « la somme est au moins 10 **ou** les deux dés donnent le même résultat ». La fonction renvoie la fréquence de cet événement sur n lancers.

b) Les sommes au moins égales à 10 correspondent aux issues $(4,6)$, $(5,5)$, $(6,4)$, $(5,6)$, $(6,5)$, $(6,6)$: 6 issues. Les doubles sont $(1,1)$, …, $(6,6)$: 6 issues. Les doubles $(5,5)$ et $(6,6)$ appartiennent aux deux listes et ne doivent être comptés qu'une fois :

$$P = \frac{6+6-2}{36} = \frac{10}{36} = \frac{5}{18} \approx 0{,}278$$

c) Avec 100 000 lancers, d'après la loi des grands nombres, la fréquence renvoyée sera, sauf exception, très proche de 0,278.

Réponse. a) « somme au moins 10 ou double » ; b) $\frac{5}{18} \approx 0{,}278$; c) environ 0,278.

Correction de l'exercice 15 – Synthèse : les paris du chevalier de Méré

a) On lance au plus 4 fois le dé et l'on s'arrête dès qu'un « 6 » sort :

```
from random import randint

def jeu1():
    for i in range(4):
        if randint(1, 6) == 6:
            return True
    return False
```

b) Un double six demande que les **deux** dés donnent 6 : la condition est *randint(1, 6) == 6 and randint(1, 6) == 6*.

c) Sur 10 000 puis 100 000 parties, la fréquence de gain du jeu 1 dépasse 0,5, tandis que celle du jeu 2 passe sous 0,5 : le jeu 1 semble favorable au joueur, le jeu 2 défavorable. Sur 1 000 parties, les deux fréquences étaient au-dessus de 0,5 (0,510 et 0,507) : l'écart entre les deux jeux, quelques centièmes, était noyé dans la fluctuation d'échantillonnage. Il faut beaucoup de parties pour qu'il apparaisse.

d) Sur 100 000 parties, les fréquences s'écartent des probabilités admises de

$$|0{,}517 - 0{,}518| = 0{,}001 \qquad |0{,}490 - 0{,}491| = 0{,}001$$

Les simulations sont donc tout à fait cohérentes avec ces valeurs, comme le prévoit la loi des grands nombres. Le chevalier de Méré, qui gagnait au jeu 1, perdait à la longue au jeu 2.

Réponse. a) Boucle de 4 lancers, *return True* dès qu'un 6 sort ; b) *randint(1, 6) == 6 and randint(1, 6) == 6* ; c) le jeu 1 ; l'écart est plus petit que la fluctuation sur 1 000 parties ; d) oui, écarts de 0,001.

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).