

Seconde Générale et Technologique
Mathématiques | Chapitre 11 : Echantillonnage
Enoncés des exercices

Les exercices sont classés en trois niveaux de difficulté :

- ★ Exercices d'application : comprendre les notions essentielles du cours
- ★★ Exercices d'entraînement : prendre les bons réflexes
- ★★★ Exercices d'approfondissement : aller plus loin

Exercices gratuits	Exercices sur abonnement*
★ 1 ★★ 4 ★★★	★ 2 – 3 ★★ 5 – 6 ★★★ 7 – 8 – 9

EXERCICES GRATUITS

Exercice 1 ★

Echantillonnage

On lance 300 fois un dé cubique numérotés de 1 à 6 et on a obtenu les fréquences de sorties des numéros 1 ; 2 ; 3 ; 4 ; 5 ; 6 respectivement f_1 ; f_2 ; f_3 ; f_4 ; f_5 ; f_6 suivantes :

Face i	1	2	3	4	5	6
Fréquence f_i	0,20	0,11	0,18	0,20	0,17	0,14

Cet échantillon ayant fourni des fréquences variant parfois pratiquement du simple au double on se demande si on peut penser que le dé est truqué (mal équilibré).

1. Si l'on suppose le dé parfaitement équilibré, quelle est la probabilité p doit-on associer à chaque face ?
2. Calculer, à 10^{-3} près, les bornes de l'intervalle de fluctuation au seuil de 95 % :
 $I = \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}} \right]$. Rappeler sa signification.
3. Pensez-vous que les écarts entre les fréquences f_1 ; f_2 ; f_3 ; f_4 ; f_5 ; f_6 observées et la probabilité théorique p s'expliquent par la seule fluctuation d'échantillonnage due au hasard ? Êtes-vous sûr d'avoir raison ?
4. Même question si l'on suppose que ces mêmes fréquences observées sont issues d'un échantillon de taille 500.

Exercice 4 ★★

Fluctuation d'échantillonnage

Dans une compétition sportive nationale où le taux de réussite est de $p = 0.55$, deux centres de formation, A et B, participent avec leurs meilleurs athlètes. L'objectif de cette compétition est de mesurer le taux de réussite des athlètes dans leurs performances.

Dans le centre A, composé de 100 athlètes, 46 d'entre eux ont réussi à atteindre les objectifs fixés lors de la compétition. Dans le centre B, qui compte 2500 athlètes, 1275 d'entre eux ont réussi à atteindre les mêmes objectifs.

1. Quelle est la fréquence f_1 correspondant au pourcentage de réussite du centre A ?
2. Quelle est la fréquence f_2 correspondant au pourcentage de réussite du centre B ?
3. Lequel de ces deux centres a le mieux réussi le concours ?
4. On assimile le centre A à un échantillon de taille $n = 100$. Donner l'intervalle de fluctuation au seuil de 95 %. f_1 appartient-il à cet intervalle ?
5. On assimile le centre B à un échantillon de taille $n = 2500$. Donner l'intervalle de fluctuation au seuil de 95 %. f_2 appartient-il à cet intervalle ?
6. Conclure sur le centre qui est le plus représentatif du résultat national à ce concours.

EXERCICES SUR ABONNEMENT

Exercice 2 * ★

Intervalle de fluctuation

Une classe de 50 élèves est choisie pour participer à un sondage sur leurs préférences en matière de sport. Parmi ces 50 élèves, 30 déclarent pratiquer régulièrement un sport et 20 ne pratiquent pas de sport régulièrement.

1. Quelle est la proportion d'élèves de la classe qui pratiquent régulièrement un sport ?
2. Calculer l'intervalle de fluctuation à 95% pour estimer la proportion d'élèves de l'ensemble de l'école qui pratiquent régulièrement un sport.
3. En utilisant l'intervalle de fluctuation calculé à la question précédente, peut-on affirmer avec certitude que plus de la moitié des élèves de l'école pratiquent régulièrement un sport ?

Exercice 3 * ★

Echantillonnage

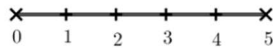
Construire un algorithme de simulation d'expérience.

1. Avec des boules numérotées

Un jeu est constitué de boules numérotées de 1 à 5.

On gagne le jeu si on tire au hasard la boule 1 ou 2, on perd sinon. On veut construire un programme modélisant la situation en effectuant un tirage au hasard :

- (a) Sur le graphique ci-dessous, si on considère que le tirage au hasard est un point du segment ci-dessous, colorier la partie correspondant à une partie gagnée puis celle représentant une partie perdue.



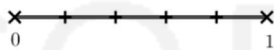
- (b) *boule* est un nombre aléatoire compris entre 0 et 5.

Compléter le programme ci-dessous pour représenter la situation du jeu.

```
import random
boule_au_hasard = 5*random.random()

if .....<= boule_au_hasard < ..... :
    print("gagne")
if .... <= boule_au_hasard < ..... :
    print("perdu")
```

2. on veut effectuer une simulation pour avoir une probabilité de gagner de $\frac{2}{5}$ et de perdre de $\frac{3}{5}$.
(*Nombre – aleatoire est un nombre aléatoire compris entre 0 et 1*) :



```
import random
Nombre_aleatoire = random.random()

if ..... <= Nombre_aleatoire < ..... :
    print("gagne")
if .... <= Nombre_aleatoire < ..... :
    print("perdu")
```

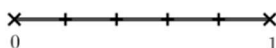
3. Avec des boules de couleurs

Dans un jeu télévisé, les candidats doivent prendre au hasard une boule dans une urne.

L'urne est composée de 3 boules blanches et de 2 boules noires.

On souhaite programmer cette simulation à l'aide d'un programme écrit en python. En s'inspirant des questions précédentes, compléter le programme ci-dessous pour qu'il simule la situation.

(*Nombre – aleatoire est un nombre aléatoire compris entre 0 et 1*) :



```
import random
Nombre_aleatoire = random.random()

if ..... <= Nombre_aleatoire < ..... :
    print("boule noire")
if .... <= Nombre_aleatoire < ..... :
    print("boule blanche")
```

Exercice 5 * ★★

Echantillonnage

Validité d'un sondage.

Pour une élection entre deux candidats *Alpha* et *Beta*, On a effectué un sondage sur une population de 100 personnes.

Les résultats du sondages donnent 49,79% d'intentions de votes pour *Alpha*.

Celui-ci prétend que l'écart est trop faible pour être certain de la défaite.

Êtes-vous d'accord ?

Exercice 6 * ★★

Un intervalle de confiance

En vue d'une élection, un institut de sondage veut estimer la proportion d'électeurs favorables au candidat sortant. Pour cela, l'institut procède à un sondage aléatoire de taille 3 600 et obtient 1 872 intentions de vote pour le candidat sortant.

- 1.a) Calculer la fréquence f des intentions de vote en faveur du candidat sortant sur ce sondage.
- b) Peut-on affirmer en supposant que les électeurs maintiennent leur choix le jour du vote, que le candidat sortant sera élu ? Pourquoi ?
2. a) Si f est la fréquence du caractère étudié sur un échantillon de taille 3 600 et p la proportion du caractère dans la population, rappeler quel est l'intervalle de fluctuation auquel f appartient, au seuil de 95%.
- b) En déduire un intervalle de centre f auquel p appartient, au niveau de confiance 0,95.
- c) Peut-on en déduire que le candidat sortant a de très bonnes chances d'être élu ?
3. Si la taille de l'échantillon du sondage avait été 1 500, pour une fréquence obtenue de 0,52 , étudiez si les conclusions auraient été les mêmes.

Exercice 7* ★★★

Echantillonnage

Simulation de tirage, conformité d'un résultat

Dans un jeu, on doit tirer une boule au hasard parmi 10 boules. Ces boules sont réparties de la manière suivante :

- 5 boules rouges.
- 3 boules bleues.
- 2 boules vertes.

Malheureusement, des boules ont été égarées et le jeu est inutilisable si on ne procède à une simulation d'échantillon.

1. Compléter les pointillés dans le programme en python ci-dessous pour qu'il simule un tirage de boule traduisant la situation (*random.random()* est un nombre aléatoire compris entre 0 et 1).

```
import random
boule_au_hasard = random.random()

if boule_au_hasard < .... :
    print("boule rouge")
if .... <= boule_au_hasard < .... :
    print("boule bleue")
if ... <= boule_au_hasard:
    print("boule verte")
```

2. Gêralt et Jaskier font une partie en utilisant ce programme.
Au bout de 150 lancers, Jaskier se rend compte qu'il a obtenu 35 fois la boule bleue alors qu'il s'estime en droit de l'obtenir 45 fois et émet des réserves sur la probité du programme utilisé.
 - (a) Expliquer le raisonnement de Jaskier.
 - (b) En utilisant l'estimation d'un échantillon infirmer ou confirmer ce raisonnement.

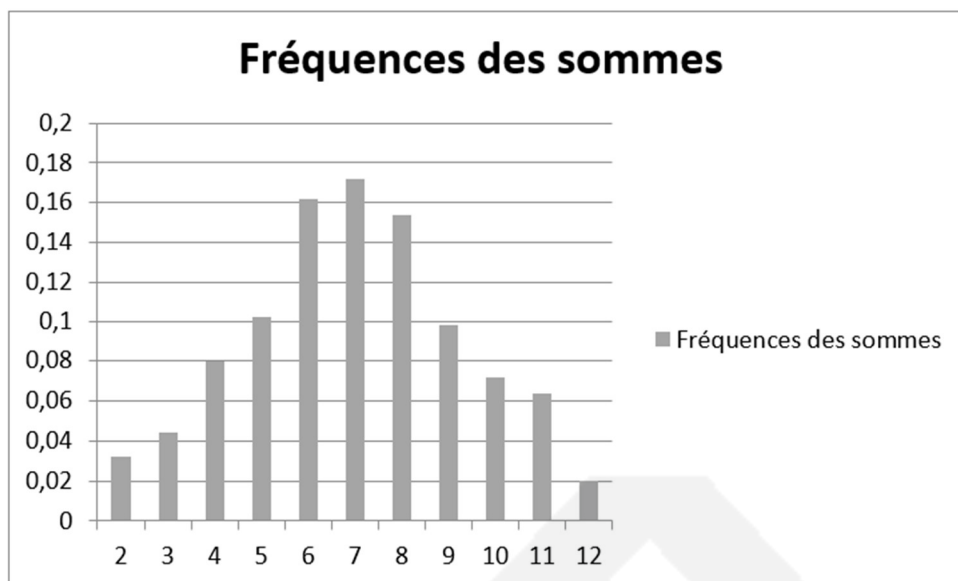
Exercice 8 * ★★★

Valider un modèle d'une expérience aléatoire, puis décider à partir d'un échantillon

Une expérience aléatoire consiste à lancer deux dés cubiques numérotés de 1 à 6.

On s'intéresse à la somme des chiffres obtenus sur les faces supérieures des 2 dés.

1. Quelles sont les sommes possibles ?
2. Quelle est la somme la plus probable ? justifier.
3. Si l'on suppose les deux dés parfaitement équilibrés, quelle probabilité p_i doit-on associer à chaque somme ?
4. On lance 500 fois les deux dés cubiques numérotés de 1 à 6 et on a obtenu les fréquences f_i de sortie des sommes S_i suivantes représentées d'abord par un diagramme en bâtons, puis par une table de valeurs donnant les sommes possibles, les occurrences par somme et les fréquences obtenus sous Excel .



Sommes possibles	<u>ni</u>	<u>fi</u>
2	12	0,024
3	26	0,052
4	40	0,08
5	57	0,114
6	58	0,116
7	89	0,178
8	73	0,146
9	69	0,138
10	37	0,074
11	33	0,066

12	6	0,012
----	---	-------

a) Calculer, à 10^{-3} près, les bornes de l'intervalle de fluctuation au seuil de 95 % :

$I = \left[p - \frac{1}{\sqrt{n}}; p + \frac{1}{\sqrt{n}} \right]$ pour la fréquence f_6 de la somme $S_6 = 7$. Rappeler sa

signification.

b) Pensez-vous que l'écart entre la fréquence

f_6 observées et la probabilité théorique p_6 correspondante s'explique par la

seule fluctuation d'échantillonnage due au hasard ? Êtes-vous sûr d'avoir raison ?

c) Même question si l'on suppose que ces mêmes fréquences observées sont

issues d'un échantillon de taille 15000.

Exercice 9 * ★★★★★

Autour de l'intervalle de confiance

Un institut de sondage souhaite évaluer l'opinion des électeurs sur une proposition de loi. La proportion de votants en faveur de la loi est estimée à 62%.

1. Pour déterminer si la proposition de loi a un impact significatif, l'institut réalise deux échantillons et compare les résultats :
 - a) Sur un échantillon de 300 électeurs, 72% se prononcent en faveur de la loi.
 - b) Sur un échantillon de 800 électeurs, 67% se prononcent en faveur de la loi.
2. Déterminer la taille minimale de l'échantillon n nécessaire pour que l'intervalle de confiance soit de 95% de largeur maximale de 0,1. Autrement dit, la différence entre les bornes supérieure et inférieure doit être de 0,1.