

ÉCHANTILLONNAGE E01

EXERCICE N°1 (Le corrigé)

L'étiquetage d'un sac de 100 kg de café précise qu'il contient 30 % de Robusta et 70 % d'Arabica. On prélève une poignée de 153 grains et on obtient 45 grains de Robusta.

1) Indiquer la population étudiée, le caractère étudié, sa proportion théorique et la taille de l'échantillon.

La population est l'ensemble des grains café du sac.

Le caractère étudié est le fait d'être Robusta ou non.

Et pourquoi ce ne serait pas la nature du grain : Robusta ou Arabica ?

Ici, on se laisse guider par l'énoncé qui met l'accent sur Robusta.

La proportion théorique est alors 0,3.

30 % c'est 30 sur 100 qui vaut 0,3

La taille de l'échantillon est de 153.

2) Calculer la fréquence du caractère dans cet échantillon.

$$\frac{45}{153} \approx 0,2941 \text{ soit environ } 29,41 \%$$

ÉCHANTILLONNAGE E01

EXERCICE N°2 (Le corrigé)

On lance un dé équilibré et on observe si on obtient un résultat pair ou non.

1) Compléter la fonction python ci-dessous pour qu'elle simule cette expérience aléatoire et renvoie la fréquence de l'évènement « Obtenir un résultat pair » dans un échantillon de taille 5000.

```
from random import random
def de_pair():
    nombre = 0
    for compteur in range(5000):
        if random() < 0.5 :
            nombre = nombre + 1
    return nombre/5000
```

```
from random import random
def de_pair():
    nombre = 0
    for compteur in range(...):
        if random() < ... :
            nombre = nombre + 1
    return ...
```

2) Exécuter 10 fois cette fonction et donner le plus petit intervalle auquel appartiennent tous les résultats renvoyés par la fonction.

Dans la console, on peut taper l'instruction au premier prompt (c'est le symbole >>>) : (arrivé aux « deux points » on appuie sur la touche « entrer » et on termine de taper la seconde ligne de l'instruction puis on appuie deux fois sur la touche « entrer ».

```
>>> for i in range(10):
    print(i+1, " ", de_pair())

1 )  0.505
2 )  0.504
3 )  0.506
4 )  0.4924
5 )  0.5042
6 )  0.5048
7 )  0.4996
8 )  0.4924
9 )  0.495
10 )  0.5036
>>> |
```

sinon, on tape 10 fois de suite l'instruction : `>>> de_pair()` et on note le résultat à chaque fois...

Vous n'aurez pas les mêmes résultats que ce qui est affiché et c'est normal puisqu'on utilise la fonction « random » du module « random » .

Après simulation, je trouve l'intervalle : [0,4924 ; 0,506]

Bien sûr, l'intervalle dépendra de vos résultats...

ÉCHANTILLONNAGE E01

EXERCICE N°3 (Le corrigé)

1) Que permet d'estimer l'algorithme ci-dessous ?

```
Somme ← 0
Pour compteur allant de 1 à 1000
  A ← nombre entier aléatoire entre 1 et 6
  Si A = 6
    Somme ← Somme + 1
Fréquence ← Somme / 1000
Afficher Fréquence
```

Cet algorithme permet de la probabilité d'obtenir un 6 avec un dé à 6 faces.

Il simule 1000 fois le lancer d'un dé à 6 faces (bien équilibré), compte le nombre de 6 puis le divise par 1000 et affiche le résultat. On obtient la fréquence d'apparition du 6 et on sait que la fréquence « tend à s'approcher de la probabilité » : c'est la loi des grands nombres (propriété n°1 de ce [cours](#) (page 2))

2) Comment pourrait-on améliorer la précision de cette estimation ?

On peut améliorer l'estimation en augmentant le nombre de tirage, en remplaçant 1000 par 10000 (ou tout nombre bien plus grand que 1000) partout dans l'algorithme.
Pour les curieux : voir [wikipédia](#))

ÉCHANTILLONNAGE E01

EXERCICE N°4 Des maths en justice ! (Le corrigé)

En 1976, un accusé d'origine mexicaine condamné pour différents crimes au Texas attaqua le jugement sous le motif que la désignation des jurés dans l'état du Texas était discriminatoire pour les Américains d'origine mexicaine.

Son argument était que ceux-ci n'étaient pas suffisamment représentés dans les jurys populaires.

À cette époque, au Texas, 79,1 % de la population était mexico-américaine, et parmi les 870 personnes convoquées en tant que « grands jurés » (sur une période de 11 ans), 339 étaient d'origine mexicaine.

La constitution des jurys pouvait-elle être considérée comme impartiale ?

Je me suis demandé si je devais laisser cet exercice, puisque son but est de susciter le débat...en classe...

J'explique rapidement :

On calcule la proportion $\frac{339}{870} \approx 0,39$ (soit 39%) et on constate qu'on est bien en dessous des 79,1 %...

Et là bien sûr : « C'était pas impartial, c'est honteux ! »

Et évidemment, si on s'arrête là c'est normal...

Sauf que, après vérification (car c'est une histoire vraie), il se trouve que l'argument de l'accusé a été rejeté.

« C'est normal puisque c'est pas impartial ! »

Et en fait...

Sur la période de 11 ans la proportion d'Américains d'origine mexicaine a évolué et il se trouve qu'il y a des conditions pour être jury :

- être âgé de plus de 18 ans ;
- posséder la nationalité américaine ;
- être un résident du district où a lieu le procès depuis un an minimum ;
- parler anglais et savoir le lire et l'écrire correctement.

L'accusation a d'ailleurs utilisé cet argument en s'appuyant sur le fait que 56 % répondaient aux critères. Ce fut sans succès car l'écart reste encore trop important...

Au final, voici ce qu'a écrit la cour suprême :

« étant donné les nombreuses facettes de la motivation humaine, il serait peu approprié de prendre comme loi établie que des humains appartenant à un groupe ne pratiqueront pas de discrimination à l'égard des membres d'un autre groupe »

Ce qu'il faut retenir (au delà du fait que « la justice, c'est compliqué ! ») :

Le fait qu'un calcul soit vrai n'est pas un argument tant que l'on a pas vérifié qu'il correspond bien à ce que l'on veut démontrer.

ÉCHANTILLONNAGE E01

EXERCICE N°1

L'étiquetage d'un sac de 100 kg de café précise qu'il contient 30 % de Robusta et 70 % d'Arabica. On prélève une poignée de 153 grains et on obtient 45 grains de Robusta.

- 1) Indiquer la population étudiée, le caractère étudié, sa proportion théorique et la taille de l'échantillon.
- 2) Calculer la fréquence du caractère dans cet échantillon.

EXERCICE N°2

On lance un dé équilibré et on observe si on obtient un résultat pair ou non.

- 1) Compléter la fonction python ci-dessous pour qu'elle simule cette expérience aléatoire et renvoie la fréquence de l'évènement « Obtenir un résultat pair » dans un échantillon de taille 5000.

```
from random import random
def de_pair():
    nombre = 0
    for compteur in range(...):
        if random() < ... :
            nombre = nombre + 1
    return ...
```

- 2) Exécuter 10 fois cette fonction et donner le plus petit intervalle auquel appartiennent tous les résultats renvoyés par la fonction.

EXERCICE N°3

- 1) Que permet d'estimer l'algorithme ci-dessous ?

```
Somme ← 0
Pour compteur allant de 1 à 1000
    A ← nombre entier aléatoire entre 1 et 6
    Si A = 6
        Somme ← Somme + 1
Fréquence ← Somme / 1000
Afficher Fréquence
```

- 2) Comment pourrait-on améliorer la précision de cette estimation ?

EXERCICE N°4 Des maths en justice !

En 1976, un accusé d'origine mexicaine condamné pour différents crimes au Texas attaqua le jugement sous le motif que la désignation des jurés dans l'état du Texas était discriminatoire pour les Américains d'origine mexicaine.

Son argument était que ceux-ci n'étaient pas suffisamment représentés dans les jurys populaires.

À cette époque, au Texas, 79,1 % de la population était mexico-américaine, et parmi les 870 personnes convoquées en tant que « grands jurés » (sur une période de 11 ans), 339 étaient d'origine mexicaine.

La constitution des jurys pouvait-elle être considérée comme impartiale ?