

Corrigé

Échantillonnage

Observer la fluctuation d'échantillonnage

Exercice 1. Fluctuation d'échantillonnage

Dans une classe, 30 élèves lancent chacun 50 fois une pièce équilibrée et notent la fréquence de « Pile ». Voici quelques résultats : 0,46 ; 0,58 ; 0,50 ; 0,38 ; 0,62 ; 0,54 ; 0,44 ; 0,70.

1. Pourquoi les élèves n'obtiennent-ils pas tous la même fréquence?
Chaque échantillon est obtenu au hasard : c'est la fluctuation d'échantillonnage.
2. Combien de « Pile » a obtenu l'élève dont la fréquence est 0,38?
 $0,38 \times 50 = 19$ « Pile ».
3. Calculer la moyenne et l'étendue des huit fréquences citées.
Somme : 4,22 ; moyenne $\frac{4,22}{8} \approx 0,53$. Étendue : $0,70 - 0,38 = 0,32$.
4. Calculer $\frac{1}{\sqrt{50}}$ arrondi au centième puis l'intervalle $\left[0,5 - \frac{1}{\sqrt{50}} ; 0,5 + \frac{1}{\sqrt{50}}\right]$.
 $\frac{1}{\sqrt{50}} \approx 0,14$. L'intervalle est environ $[0,36 ; 0,64]$.
5. Quelles fréquences parmi celles citées sont en dehors de cet intervalle ? Est-ce surprenant ?
Seule 0,70 est en dehors. Ce n'est pas surprenant : environ 5 % des échantillons (donc 1 ou 2 élèves sur 30) peuvent sortir de l'intervalle.
6. Les 30 élèves regroupent leurs résultats. Combien de lancers cela représente-t-il ?
 $30 \times 50 = 1\,500$ lancers.
7. Ils obtiennent 768 « Pile » sur 1 500 lancers. Calculer la fréquence obtenue et expliquer pourquoi elle est plus proche de 0,5 que la plupart des fréquences individuelles.
 $f = \frac{768}{1500} = 0,512$. L'échantillon est beaucoup plus grand, donc la fluctuation est plus faible (loi des grands nombres).

Exercice 2. Taille de l'échantillon

On a simulé, pour différentes tailles n , 100 échantillons de lancers d'une pièce équilibrée. Le tableau indique la plus petite et la plus grande fréquence de « Pile » obtenues.

n	25	100	400	2 500
Fréquence minimale	0,28	0,39	0,44	0,478
Fréquence maximale	0,72	0,62	0,55	0,524

1. Calculer l'étendue des fréquences pour chaque taille d'échantillon.

0,44 ; 0,23 ; 0,11 ; 0,046.

2. Comment évolue la fluctuation lorsque n augmente ?

Elle diminue : les fréquences se resserrent autour de 0,5.

3. Calculer $\frac{2}{\sqrt{n}}$ pour chaque valeur de n et comparer avec les étendues.

0,4 ; 0,2 ; 0,1 ; 0,04 : c'est proche des étendues observées (l'intervalle de fluctuation a pour amplitude $\frac{2}{\sqrt{n}}$).

4. Par combien faut-il multiplier n pour diviser l'amplitude de l'intervalle par 5 ?

$$\frac{2}{\sqrt{kn}} = \frac{1}{5} \times \frac{2}{\sqrt{n}} \iff \sqrt{k} = 5 \iff k = 25.$$

Déterminer un intervalle de fluctuation et prendre une décision

Exercice 3. Intervalle de fluctuation

Pour chaque situation, déterminer l'intervalle $\left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}} \right]$.

1. $p = 0,3$ et $n = 100$.

$$\frac{1}{\sqrt{100}} = 0,1 : [0,2 ; 0,4].$$

2. $p = 0,5$ et $n = 2 500$.

$$\frac{1}{\sqrt{2500}} = 0,02 : [0,48 ; 0,52].$$

3. $p = 0,12$ et $n = 400$.

$$\frac{1}{\sqrt{400}} = 0,05 : [0,07 ; 0,17].$$

4. $p = 0,65$ et $n = 900$ (arrondir au millième).

$$\frac{1}{\sqrt{900}} \approx 0,033 : [0,617 ; 0,683].$$

5. $p = 0,25$ et $n = 64$.

$$\frac{1}{\sqrt{64}} = 0,125 : [0,125 ; 0,375].$$

6. Que se passe-t-il quand la taille n de l'échantillon est multipliée par 4?

$$\frac{1}{\sqrt{4n}} = \frac{1}{2} \times \frac{1}{\sqrt{n}} : \text{l'amplitude de l'intervalle est divisée par 2.}$$

7. Quelle taille d'échantillon faut-il pour que l'intervalle ait une amplitude de 0,05?

$$\frac{2}{\sqrt{n}} = 0,05 \iff \sqrt{n} = 40 \iff n = 1\,600.$$

Exercice 4. Prise de décision

1. Un fabricant affirme que 8 % de ses ampoules sont défectueuses. Sur un lot de 625 ampoules, on en trouve 70 défectueuses. Le lot est-il conforme à l'affirmation du fabricant?

$f = \frac{70}{625} = 0,112$. $\frac{1}{\sqrt{625}} = 0,04$: intervalle $[0,04 ; 0,12]$. f est dans l'intervalle : rien ne permet de remettre en cause l'affirmation.

2. À partir de combien d'ampoules défectueuses sur ce lot de 625 aurait-on remis en cause l'affirmation?

Il faut $f > 0,12$, soit plus de $0,12 \times 625 = 75$ ampoules : à partir de 76 ampoules défectueuses.

3. Dans une commune, on a observé 1 080 naissances de garçons sur 2 000 naissances. La probabilité qu'un nouveau-né soit un garçon est environ 0,512. Cette commune est-elle atypique?

$f = \frac{1080}{2000} = 0,54$. $\frac{1}{\sqrt{2000}} \approx 0,022$: intervalle environ $[0,490 ; 0,534]$. $f = 0,54$ est en dehors : au risque de 5 %, la commune semble atypique.

4. On lance 900 fois un dé et on obtient 180 fois le 6. Peut-on penser que le dé est truqué?

$p = \frac{1}{6} \approx 0,167$, $\frac{1}{\sqrt{900}} \approx 0,033$: intervalle environ $[0,133 ; 0,200]$. $f = \frac{180}{900} = 0,2$ est à la limite de l'intervalle : on ne peut pas conclure avec certitude, il faudrait plus de lancers.

5. On relance ce dé 3 600 fois et on obtient 720 fois le 6. Que peut-on conclure?
 $\frac{1}{\sqrt{3600}} \approx 0,017$: intervalle environ $[0,150 ; 0,184]$. $f = \frac{720}{3600} = 0,2$ est en dehors : au risque de 5 %, on peut penser que le dé est truqué.
6. Dans une entreprise de 100 salariés, il y a 32 femmes alors que la population active compte 48 % de femmes. Peut-on soupçonner une discrimination à l'embauche?
 $\frac{1}{\sqrt{100}} = 0,1$: intervalle $[0,38 ; 0,58]$. $f = 0,32$ est en dehors : au risque de 5 %, la situation n'est pas due au seul hasard et mérite d'être étudiée.

Estimer une proportion

Exercice 5. Sondage

Un sondage réalisé auprès de 1 600 personnes indique que 38 % d'entre elles utilisent les transports en commun chaque jour.

- Combien de personnes interrogées utilisent les transports en commun chaque jour?
 $0,38 \times 1\,600 = 608$ personnes.
- Donner un intervalle dans lequel se trouve, avec une confiance de 95 %, la proportion réelle dans la population.
 $\frac{1}{\sqrt{1600}} = 0,025$: $[0,355 ; 0,405]$, soit entre 35,5 % et 40,5 %.
- Une ville compte 250 000 habitants. Entre combien et combien d'habitants utilisent vraisemblablement les transports en commun chaque jour?
Entre $0,355 \times 250\,000 = 88\,750$ et $0,405 \times 250\,000 = 101\,250$ habitants.
- Combien aurait-il fallu interroger de personnes pour obtenir un intervalle d'amplitude 0,02 ?
On veut $\frac{2}{\sqrt{n}} = 0,02$, soit $\sqrt{n} = 100$, soit $n = 10\,000$ personnes.

Exercice 6. Deux candidats

Avant une élection, un sondage auprès de 400 électeurs donne 53 % des voix au candidat A et 47 % au candidat B.

1. Donner l'intervalle de confiance au niveau 95 % de la proportion d'électeurs votant pour A.

$$\frac{1}{\sqrt{400}} = 0,05 : [0,48 ; 0,58].$$

2. Peut-on affirmer que A sera élu?

Non : l'intervalle contient des valeurs inférieures à 0,5, A pourrait être minoritaire.

3. Un second sondage auprès de 2 500 électeurs donne encore 53 % pour A. Que peut-on dire maintenant?

$\frac{1}{\sqrt{2500}} = 0,02 : [0,51 ; 0,55]$. Toutes les valeurs sont supérieures à 0,5 : avec une confiance de 95 %, A est majoritaire.

4. Quel est le nombre minimal de personnes à interroger pour que l'intervalle de confiance centré sur 0,53 ne contienne pas 0,5?

Il faut $\frac{1}{\sqrt{n}} < 0,03 \iff \sqrt{n} > \frac{1}{0,03} \approx 33,3 \iff n > 1\,111,1$: au moins 1 112 personnes.

Simuler une expérience avec Python

Exercice 7. Simulation en Python

On considère la fonction suivante :

```
from random import randint
def experience(n):
    compteur = 0
    for i in range(n):
        if randint(1, 6) == 6:
            compteur = compteur + 1
    return compteur / n
```

1. Que simule `randint(1, 6)`? Que renvoie la fonction `experience(n)`?

`randint(1, 6)` simule le lancer d'un dé équilibré. La fonction renvoie la fréquence des 6 obtenus sur n lancers.

2. Quel est le rôle de la variable compteur? Combien de fois la boucle est-elle exécutée?
compteur compte le nombre de 6 obtenus; la boucle est exécutée n fois.
3. Autour de quelle valeur devraient se situer les résultats de expérience (10000)? Pourquoi?
Autour de $\frac{1}{6} \approx 0,167$ d'après la loi des grands nombres (écart inférieur à 0,01 dans 95 % des cas).
4. Modifier la fonction pour qu'elle simule le lancer d'une pièce et renvoie la fréquence de « Pile ».
On remplace le test par `if randint(0, 1) == 1:` (en codant Pile par 1).
5. Modifier la fonction pour qu'elle renvoie la fréquence des résultats pairs d'un dé.
On remplace le test par `if randint(1, 6) % 2 == 0:`. La fréquence doit être proche de 0,5.
6. Écrire une fonction `echantillons(N, n)` qui renvoie la liste des fréquences de 6 obtenues pour N échantillons de taille n .

```
def echantillons(N, n):
    L = []
    for k in range(N):
        L.append(experience(n))
    return L
```
7. Compléter cette fonction pour qu'elle renvoie la proportion d'échantillons dont la fréquence est dans l'intervalle $\left[\frac{1}{6} - \frac{1}{\sqrt{n}}; \frac{1}{6} + \frac{1}{\sqrt{n}}\right]$. Quelle valeur doit-on observer environ?
On ajoute `from math import sqrt`, puis on compte les fréquences f de L telles que `abs(f - 1/6) <= 1/sqrt(n)` et on divise par N . On doit trouver environ 0,95 ou plus.