Concentration et loi des grands nombres
Problème de recherche : Bienaymé-Tchebychev est-elle pessimiste ?

On lance n fois une pièce équilibrée et on note F_n la fréquence de pile obtenue. On s'intéresse à la probabilité $P(|F_n - 0{,}5| \geqslant 0{,}05)$, c'est-à-dire au risque que la fréquence s'écarte de plus de 5 points de 0,5.

La mission

1. Donner la majoration de $P(|F_n - 0{,}5| \geqslant 0{,}05)$ fournie par l'inégalité de concentration, pour $n = 100$, $n = 400$ et $n = 1\,600$.
2. Écrire un programme Python qui simule 10 000 échantillons de taille n et estime cette probabilité. L'utiliser pour $n = 100$, 400 et 1 600, et comparer avec les majorations.
3. Quelle taille n l'inégalité impose-t-elle pour garantir un risque d'au plus 5 % ? D'après les simulations, quelle taille suffit réellement ?
4. Comment ces résultats illustrent-ils la loi des grands nombres ?

Coups de pouce

Bloqué ? Des indices progressifs sont disponibles en ligne. N'en ouvrir qu'un à la fois, et seulement après avoir cherché !

Éléments de correction

1. Les majorations

Avec $p = 0{,}5$: $P(|F_n - 0{,}5| \geqslant 0{,}05) \leqslant \frac{0{,}25}{n \times 0{,}05^2} = \frac{100}{n}$. Pour $n = 100$: 1 (aucune information) ; $n = 400$: 0,25 ; $n = 1\,600$: 0,0625.

2. Simulation

```
from random import random
def frequence(n):
    return sum(1 for _ in range(n) if random() < 0.5) / n
def risque(n, N=10000):
    return sum(1 for _ in range(N) if abs(frequence(n) - 0.5) >= 0.05) / N
for n in [100, 400, 1600]:
    print(n, risque(n))
```

n	100	400	1 600
majoration	1	0,25	0,0625
simulation (environ)	0,37	0,05	presque 0

Les valeurs simulées (proches des probabilités exactes 0,368 ; 0,051 ; 0,000 07) sont bien inférieures aux majorations : celles-ci sont vraies, mais souvent très larges.

3. La taille nécessaire

L'inégalité impose $\frac{100}{n} \leqslant 0{,}05$, soit $n \geqslant 2\,000$. Les simulations montrent qu'avec $n = 400$ le risque est déjà d'environ 5 % : en pratique, une taille cinq fois plus petite suffit. L'inégalité a l'avantage d'être **garantie** pour toute loi, au prix d'une prudence excessive.

4. Loi des grands nombres

Pour tout écart $\delta > 0$ fixé, la majoration $\frac{p(1-p)}{n\delta^2}$ tend vers 0 quand n augmente : la probabilité que F_n s'écarte de p de plus de δ tend vers 0. C'est la loi des grands nombres, que les simulations rendent visible : plus l'échantillon est grand, plus les fréquences se concentrent autour de 0,5.

Enrichissement historique : l'inégalité porte les noms d'Irénée-Jules Bienaymé (1853) et de Pafnouti Tchebychev (1867).

Ce qu'on a utilisé du chapitre

- Inégalité de concentration ; loi des grands nombres.
- Simuler des échantillons avec Python et interpréter.