Variables aléatoires
Problème de recherche : l'échantillon suspect

Un fabricant de cartes électroniques achète des condensateurs. Son fournisseur garantit que 5 % au plus des condensateurs sont défectueux. À la réception d'un lot, un technicien teste un échantillon de 50 condensateurs, prélevés au hasard dans un très grand stock. Il trouve 6 condensateurs défectueux, soit une fréquence de 12 %.

On suppose d'abord que la garantie est respectée avec exactement 5 % de défectueux : chaque condensateur testé est défectueux avec la probabilité 0,05, indépendamment des autres.

La mission

1. Écrire une fonction Python `echantillon()` qui simule le test de 50 condensateurs et renvoie le nombre de défectueux. L'exécuter plusieurs fois : les résultats sont-ils toujours les mêmes ?
2. Simuler 1 000 échantillons. Dans quelle proportion d'échantillons trouve-t-on au moins 6 défectueux ?
3. Déterminer, à l'aide de la simulation, un nombre s tel qu'environ 95 % des échantillons contiennent au plus s défectueux.
4. Le technicien doit-il mettre en doute la garantie du fournisseur ? Peut-il en être certain ?

Pour aller plus loin

Le technicien teste un deuxième échantillon de 50 et trouve de nouveau 6 défectueux. Qu'en penser ?

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. La fonction

```
from random import random
def echantillon():
    d = 0
    for i in range(50):
        if random() < 0.05:
            d = d + 1
    return d
```

Les résultats changent d'une exécution à l'autre (souvent 1, 2, 3 ou 4) : c'est la fluctuation d'échantillonnage.

2. Mille échantillons

```
c = 0
for k in range(1000):
    if echantillon() >= 6:
        c = c + 1
print(c / 1000)
```

On obtient des valeurs proches de 0,04 : environ 4 % des échantillons contiennent au moins 6 défectueux. (Le calcul exact, hors programme, donne 3,8 %.)

3. Le seuil à 95 %

En comptant les échantillons avec au plus 4, puis au plus 5 défectueux, on trouve environ 90 % puis environ 96 %. Donc $s = 5$: environ 95 % des échantillons contiennent au plus 5 défectueux, soit une fréquence d'au plus 10 %.

4. La décision

Si la garantie était respectée, un échantillon avec 6 défectueux ou plus n'arriverait que dans environ 4 % des cas. C'est rare : le technicien a de bonnes raisons de douter de la garantie et de réclamer un contrôle plus poussé.

Il ne peut pas en être certain : ce résultat reste possible, environ une fois sur vingt-cinq, avec un fournisseur honnête.

Pour aller plus loin

Deux échantillons indépendants avec au moins 6 défectueux : la probabilité est d'environ $0{,}04 \times 0{,}04 = 0{,}0016$ si la garantie est respectée. C'est très improbable : le doute devient presque une certitude.

Ce qu'on a utilisé du chapitre

- Simulation d'une répétition d'épreuves de Bernoulli avec Python.
- Fluctuation d'échantillonnage, fréquence observée.
- Indépendance et produit des probabilités.