

# Fluctuation d'échantillonnage

## Intervalle de fluctuation

### Test d'hypothèse

On a un dé à 4 faces et on veut savoir s'il est bien équilibré ou s'il est truqué.

On lance 100 fois ce dé.

S'il donne 25 fois le 4, on n'a pas de raison de penser qu'il est truqué.

On lance 100 fois ce dé à quatre faces et à chaque fois c'est la 4 qui sort.

Est-ce possible ?

Oui.

Est-ce probable ?

Non, la probabilité de cet événement est  $\left(\frac{1}{4}\right)^{100} \approx$

0,000 001.

Il est possible qu'un dé donne 100 fois de suite le même résultat. Mais si cela arrive, on pensera que le dé est truqué.

On affirme que le dé est truqué. On n'en est pas vraiment certain mais il y a de très grandes chances qu'il soit truqué.

Bon, là c'est facile de se faire une idée. Voyons un autre exemple.

On lance 100 fois un dé à quatre faces et le 4 sort 33 fois.

On peut attendre que le 4 sorte 25 fois et il sort 33 fois.

Peut-on penser que le dé est truqué ?

Pour un dé normal, le nombre d'apparitions du 4 lors de 100 lancers est une variable aléatoire  $X$  binomiale de paramètres 100 et 0,25.

L'espérance de cette variable aléatoire est 25. On a trouvé 33. Cela fait un écart de 8. Est-ce que cela peut arriver avec un dé normal ?

Lorsqu'on lance 100 fois un dé à quatre faces, le nombre de 4 appartient à l'intervalle  $[0 ; 100]$ .

On va partager ces résultats en deux catégories :

- la catégorie A qui contient les résultats dont l'écart à l'espérance est inférieur ou égal à 7.

- la catégorie B qui contient dont l'écart à l'espérance est supérieur ou égal à 8.

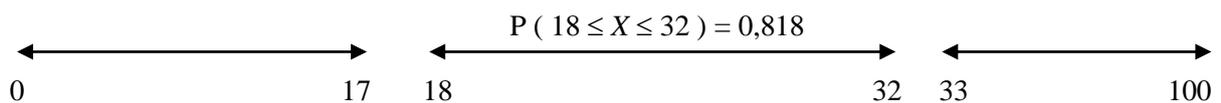
A contient les résultats appartenant à  $[18 ; 32]$ .

B contient les résultats appartenant à  $[0 ; 17] \cup [33 ; 100]$ .

On demande à la calculatrice et on trouve que

$$P(X \in A) = P(18 \leq X \leq 32) = P(X \leq 32) - P(X \leq 17) = 0,918$$

$$\text{donc } P(X \in B) = 1 - 0,918 = 0,082.$$



Qu'est-ce que cela signifie ?

On lance 100 fois un dé à quatre faces non truqué et on compte le nombre de 4.

On s'intéresse à l'écart entre 25 et ce nombre de 4.

Cet écart peut être supérieur ou égal à 8. Cela arrive avec une probabilité de 0,082, c'est-à-dire environ une fois sur 12.

Ce n'est pas très fréquent mais cela peut arriver. Si cela arrive, on ne va pas accuser le dé d'être truqué.

On se demandait si le dé est bien équilibré. On appelle cela un test d'hypothèse. Dans ce cas, on accepte l'hypothèse que le dé est bien équilibré.

Avec une probabilité de  $0,082 \approx \frac{1}{12}$ , on a accepté l'hypothèse que le dé est bien équilibré.

À partir de quel seuil rejette-t-on cette hypothèse ?

C'est un peu arbitraire mais si on atteint un écart dont la probabilité de se produire est inférieure à 0,05, on considère que le test conduit à rejeter l'hypothèse.

$$0,05 = 5 \% = \frac{1}{20}.$$

On lance 100 fois un dé à quatre faces et le 4 sort 34 fois.

L'écart entre 34 et l'espérance est 9.

La probabilité que l'écart soit 8 ou moins  $P(17 \leq X \leq 33) = P(X \leq 33) - P(X \leq 16) = 0,951$ .

La probabilité que l'écart soit 9 ou plus est  $P(X \leq 16 \text{ ou } X \geq 34) = 1 - 0,951 = 0,049$ .

Cela fait 4,9 % qui est inférieur à 5 %.

On refuse donc l'hypothèse que le dé est équilibré.

On considère que le dé est truqué.

L'intervalle de fluctuation au seuil de 95 %

$$P(18 \leq X \leq 32) = 0,918 = 91,8 \%$$

$$P(17 \leq X \leq 33) = 0,951 = 95,1 \%$$

Le seuil de 5 % est dépassé entre l'intervalle [ 18 ; 32 ] et l'intervalle [ 17 ; 33 ].

On peut aussi penser aux intervalles [ 18 ; 33 ] et l'intervalle [ 17 ; 32 ].

On parle d'intervalle de fluctuation des effectifs au seuil de 95 %.

Il faut en choisir un parmi les quatre, on choisit [ 17 ; 32 ] : pour la borne de gauche, on choisit le plus petit nombre, pour la borne de droite aussi. C'est l'habitude mais c'est tout-à-fait arbitraire, les trois autres intervalles sont bons aussi.

On a une autre habitude : présenter les résultats non avec des effectifs mais avec des fréquences.

Pour passer des effectifs aux fréquences on divise par  $n = 100$ .

L'intervalle de fluctuation des fréquences au seuil de 95 % est [ 0,17 ; 0,32 ].

Lors d'une série de 100 lancers d'un dé à quatre faces, si la fréquence des 4 est comprise entre 0,16 et 0,32, on acceptera l'hypothèse que le dé est équilibré. Sinon on rejettera cette hypothèse et on considérera que le dé est truqué.

Un intervalle de fluctuation sert à tester une hypothèse.

Exercice : on lance 1000 fois un dé à quatre faces.

1/ Le 4 apparaît 220 fois. Au seuil de 95 %, peut-on considérer que le dé est bien équilibré ?

2/ Donner l'intervalle de fluctuation des effectifs au seuil de 95 %.

3/ On fait l'hypothèse que ce dé est équilibré. Donner la règle de décision concernant cette hypothèse.

4/ Donner l'intervalle de fluctuation des fréquences au seuil de 95 %.

5/ On fait l'hypothèse que ce dé est équilibré. En utilisant les fréquences, donner la règle de décision concernant cette hypothèse.