I. Echantillonnage

Alice propose un jeu d'argent à Bob.

Elle dispose d'une pièce de monnaie, elle lui propose de la lancer 500 fois de sorte qu'à chaque fois qu'elle obtient pile Bob donne un euro à Alice et au contraire à chaque fois qu'elle obtient face Alice donne un euro à Bob.

Bob accepte cette proposition qui lui parait équitable.

Alice joue et a obtenu 300 Fois pile et 200 fois Face.

Au final Alice a reçu 300 euros de Bob et a donné 200 euros ce qui lui vaut un gain total de 100 euros. ***BDevant l'extravagance de la somme a donné, Bob met en doute le fait que la pièce soit bien équilibrée***

En fait, on ne peut jamais être sûr de rien, mais d'un point de vue statistique, on peut avoir une idée de la réponse, avec une certaine marge d'erreur, grâce aux simulations.

Autrement on va reproduire ladite expérience un grand nombre de fois en supposant que la pièce soit bien équilibrée pour vérifier si ce qui est arrivé à Bob est si rare que ça.

Algorithme 1 :

Traitement :
$n = 500$
Npile=0
Pour i allant de 1 à n **Faire**
    hasard prend la valeur 0 ou 1.
    Si (hasard=0) **Alors**
        Npile=Npile+1
    Fin Si
Fin Pour
Afficher Npile.

En effectuant la simulation suivante, par exemple en Python

codingground SIMPLY EASY CODING | Execute Python Online (Python v2.7.13)

Execute | Share | main.py | STDIN

```
# Hello World program in Python
from random import randint

print "Hello World!\n"



def simulation(n):
    Npile=0
    for i in range(n):
        hasard=randint(0,1)
        if hasard==0:
            Npile=Npile+1
    return(Npile)

L=[]
for i in range(1000):
    L.append(simulation(500))
L.sort()
print(L)
```

on constate qu'effectivement le cas présenté au départ ne se produit quasiment jamais ; il y avait très peu de chance que Bob perde autant ; cela pousse à soupçonner une entourloupe d'autant plus.

Définition 1.

Un échantillon (aléatoire) de taille n est la liste des résultats obtenus par n répétitions indépendantes d'une même expérience aléatoire.

Exemple :

Ici notre échantillon est de taille $n = 500$ et comporte 300 Pile et 200 Face. Nous avons observé dans cet exemple une fréquence de « Pile »qui vaut $f_p = \dfrac{300}{500}$ et une fréquence de « Face »qui vaut $f_f = \dfrac{200}{500}$. On se demande s'il s'agit du hasard et qu'il est raisonnable d'estimer que la probabilité d'obtenir pile avec cette pièce de monnaie est $p = \dfrac{1}{2}$. Le résultat suivant que l'on admet permet de décider :

Théorème 1.

On considère une expérience aléatoire à deux issues possibles (succès ou échec) dont **on connait** la probabilité de succès p. On produit un échantillon de taille n et on observe la fréquence f de succès.
Si p est compris entre $0,2$ et $0,8$ et si n est supérieur ou égal à 25, alors dans 95% des cas au moins, f appartient à l'intervalle :

$$\left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}}\right]$$

que l'on appelle **intervalle de fluctuation** au seuil 95% (ou seuil $0,95$).

Remarque : Réciproquement dans les mêmes conditions on a dans 95% des cas au moins, p appartient à l'intervalle :

$$\left[f - \frac{1}{\sqrt{n}} ; f + \frac{1}{\sqrt{n}}\right]$$

On utilise ce nouvel intervalle, appelé **intervalle de confiance** de p lorsqu'on ne connait pas p.

Remarque : En mathématiques, un intervalle de fluctuation permet de détecter un écart important par rapport à la valeur théorique pour une grandeur établie sur un échantillon. C'est un intervalle dans lequel la grandeur observée est censée se trouver avec une forte probabilité (souvent de l'ordre de 95 %).
Le fait d'obtenir une valeur en dehors de cet intervalle s'interprète alors en mettant en cause la représentativité de l'échantillon ou la valeur théorique. À l'inverse, le fait que la moyenne soit comprise dans l'intervalle n'est pas une garantie de la validité de l'échantillon ou du modèle.

Exemple :

Ici donc notre échantillon est de taille 500, et la fréquence de pile que l'on a observé vaut $\dfrac{300}{500} = 0.600$.

Les bornes de l'intervalle de fluctuation sont : $0,5 - \dfrac{1}{\sqrt{[500}} \simeq 0,455$ et $0,5 + \dfrac{1}{\sqrt{500}} \simeq 0,544$

Le théorème précédent assure que si on effectue la même expérience avec une pièce qui est parfaitement équilibrée alors dans au moins 95% on observa une fréquence de pile comprise $0,455$ et $0,544$.
Notre observation est donc « rare »pour une pièce bien équilibrée, puisque $0,600$ n'est pas compris entre $0,455$ et $0,544$, par conséquent on a toutes les raison de refuser l'hypothèse « la pièce est bien équilibrée. »