

ECHANTILLONNAGE

Je réfléchis

On tire 5 fois de suite une pièce de monnaie. A chaque lancer, les deux issues possibles sont

Le résultat « PFPPF » constitue un échantillon de taille pour cette expérience aléatoire à ... issues.

Si on choisit Pile comme succès, le nombre de succès obtenus pour cet échantillon est

Donc la fréquence observée de succès est :

Je retiens

Soit un entier naturel n non nul.

On répète n fois l'expérience aléatoire, c'est-à-dire sans que les résultats obtenus au fur et à mesure n'influencent les résultats suivants.

La liste des résultats obtenus est appelée

La **fréquence de succès observée** f sur un échantillon de taille n comportant k succès se calcule par la formule :

J'utilise mes connaissances

- 1) A l'aide de la calculatrice, simuler 20 échantillons de 100 lancers d'un dé équilibré.

Reporter le nombre de 6 obtenus ci-dessous et calculer la fréquence observée :

Echantillon	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Nb de 6																				
f																				

- 2) Quelle est la probabilité théorique p d'obtention d'un 6 ?
3) Quel est l'écart maximal entre p et la fréquence observée f ?

Loi des grands nombres

Une urne contient des boules, dont certaines sont rouges. Soit un entier naturel n non nul. On prélève un échantillon de taille n grande. Préciser, dans chaque cas, autour de quelle valeur, la fréquence observée de boules rouges pourrait se trouver.

- 1) L'urne contient 500 boules, dont 50 rouges.
2) L'urne contient 50 boules rouges, 10 vertes et 150 bleues.

TD PYTHON : ESTIMATION

On souhaite observer à l'aide d'un programme Python le phénomène de fluctuation d'échantillonnage et en apprécier l'étendue.

On reprend l'exemple du lancer d'un dé équilibré, on s'intéresse à la sortie de la face 6.

- 1) On considère la fonction échantillon (p, n) ci-dessous qui prend en argument la probabilité de réussite p et la taille n de l'échantillon simulé. Que renvoie-t-elle ?

```
from random import *
def echantillon(p,n):
    succes = 0
    for i in range(n):
        x = random()
        if x < p:
            succes = succes + 1
    f = succes / n
    return f
```

- 2) On souhaite observer graphiquement la fluctuation d'échantillonnage pour N échantillons de taille n et visualiser ceux pour lesquels l'écart entre la fréquence f et la probabilité p est inférieur ou égal à $\frac{1}{\sqrt{n}}$.

On propose pour cela la fonction graphique(p, n, N) ci-dessous :

```
from pylab import *
def graphique(p,n,N):
    for i in range(N):
        f = echantillon(p,n)
        if abs(f-p)<=1/sqrt(n):
            plot(i,f,'+',color='green')
        else:
            plot(i,f,'+',color='red')
    plot([0,N],[p,p],color='black')
    axis([0,N,0,1])
    show()
```

- a) L'exécuter pour $n = 100$ et $N = 1000$.
b) Que permet de visualiser le graphique obtenu ?
- 3) En modifiant la fonction graphique(p, n, N), créer une fonction écart(p, n, N) qui renvoie la proportion des cas où l'écart entre f et p est inférieur ou égal à lors de la simulation de N échantillons de taille n .

Je retiens

Lorsqu'on approche la probabilité par la fréquence l'erreur commise est

Un résultat mathématique permet d'affirmer que dans une très grande majorité des cas, on a :

Pour aller plus loin :

On lance 2 pièces de monnaie équilibrées, un élève affirme que la fréquence de « obtenir 2 côtés faces » deviendra proche de $\frac{1}{3}$ au bout d'un grand nombre de lancers. Qu'en penser ?