Echantillonnage

L'échantillonnage est un domaine assez récent des mathématiques, entre les statistiques et les probabilités.
Ce domaine intervient dans les situations où l'on veut étudier certaines propriétés d'une population de trop grand effectif pour être observée de façon exhaustive.
On prélève alors un **échantillon** de cette population, échantillon à partir duquel on souhaite obtenir des informations sur la population totale avec un certain degré de précision.
Un exemple omniprésent dans les média est la réalisation de sondages d'opinion.

I Notion d'échantillon

Définition 1
Un échantillon de taille n est constitué des résultats de n répétitions indépendantes de la même expérience sur l'ensemble des personnes ou objets sur lequel porte l'étude statistique (la population).

Exemple 1
- Sur l'ensemble des clés usb produites par une entreprise en un mois, on en prélève 150. On dit que cet ensemble de 150 clés usb constituent un **échantillon de taille 150** de la population de toutes les clés usb produites en un mois.
- On s'intéresse aux intentions de vote lors d'une élection. On sonde 1000 personnes en leur demandant leur intention de vote. L'ensemble des 1000 personnes constitue ....
....
- On lance une pièce de monnaie 50 fois de suite et on note les résultats obtenus. L'ensemble des 50 lancers constitue ....

II Fluctuation d'échantillonnage

Dans une population, on s'intéresse à l'apparition d'un certain caractère. On note p la proportion d'individus présentant ce caractère dans la population totale.
On prélève un échantillon dans la population et on note f la fréquence d'apparition du caractère observée dans cet échantillon.
En étudiant plusieurs échantillons prélevés dans la même population, on constate que la fréquence observée fluctue autour de la proportion p. Ce phénomène, dû au hasard dans la constitution des échantillons est appelé **fluctuation d'échantillonnage**

Exemple 2
On lance 50 fois un dé équilibré à 6 faces. La probabilité de tomber sur le 1, le 2, le 3 le 4 et le 6 est de 1/6
On note les fréquences d'apparition des 6 faces dans un tableau :

Face du dé	1	2	3	4	5	6
Fréquence						

III Loi des grands nombres

Propriété 1
Lorsque n devient grand, sauf exception, la fréquence observée est proche de la probabilité.

Exemple 3
A l'aide de python, on simule un programme permettant d'afficher en sortie la fréquence de 1 ou de 6 en lançant 10 dés, ou 100, ou 1000 etc. (échantillon de taille 10, puis 100, puis 1000 etc.)

```python
from random import*

def dé(n):
    s=0
    for k in range(n):
        r=randint(1,6)
        if r==1 or r==6:
            s=s+1
    return(s/n)
```

On éxecute le programme pour des valeurs de n de plus en plus grandes. Et voici les résultats obtenus à l'aide d'un logiciel :

```
>>> dé(10)
0.2
>>> dé(100)
0.32
>>> dé(1000)
0.328
>>> dé(5000)
0.3372
>>> dé(100000)
0.33353
```

On constate que plus n devient grand, plus ....

IV Estimer une proportion

Définition 2
Dans une population la proportion p d'individus présentant un certain caractère est inconnue.
On prélève dans cette population un échantillon aléatoire de taille n et on note f la fréquence d'apparition du caractère dans l'échantillon.
La fréquence observée est appelée ***une estimation*** *de la proportion p*

Propriété 2
Pour un n assez grand, f donne une bonne estimation de p dans environ 95% des cas.

Exercice 1
Jean a acheté un dé à 6 faces truqué, c'est à dire que toutes les faces n'ont pas la même probabilité d'apparaître. Mais il ne connait pas les probabilités réelles d'apparition de chaque face. Il effectue 500 lancers et obtient les résultats ci-dessous :

Face	1	2	3	4	5	6
Effectif	42	82	80	90	88	118

Donner une estimation de la probabilité d'obtenir le nombre 6 avec ce dé.