Histoire 1

La parution de *l'Ars Conjectandi* de Jacques Bernoulli (1713), reprenant notamment d'anciens travaux de Huygens, marque une rupture dans l'histoire des probabilités. On y trouve la première étude de la distribution binomiale, introduite dans le cadre d'un tirage sans remise pour un modèle d'urne.
Un résultat majeur de cet ouvrage est son « *théorème d'or* », la **loi des grands nombres**, qui relie fréquences et probabilité, valide le principe de l'échantillonnage et est le premier exemple de « théorème limite » en théorie des probabilités. Le mathématicien français Bienaymé (en 1853, publication en 1867) et le mathématicien russe Tchebychev (en 1867) démontrent l'inégalité qui porte leur nom, en parlant de fréquences d'échantillons plutôt que de variables aléatoires. Ils fournissent ainsi la possibilité d'une démonstration plus simple de la **loi des grands nombres**.

1 Un problème historique

Histoire 2 ***Problème du chevalier de Méré***

Un problème historique dû au Chevalier de Méré est rapporté dans la correspondance entre Pascal et Fermat. Grand joueur, le chevalier de Méré s'intéressait aux jeux de hasard sur lesquels il misait de l'argent. À l'issue de nombreuses parties, il avait constaté avoir plus d'une chance sur deux d'obtenir au moins une fois un six en lançant quatre fois un dé à six faces et moins d'une chance sur deux d'obtenir au moins un double-six en lançant 24 fois deux dés. Ce résultat lui semblait en contradiction avec l'égalité des rapports $\frac{24}{36}$ et $\frac{4}{6}$ du nombre de lancers au nombre de faces.

1. Calculer la probabilité d'obtenir au moins un six à l'issue de 4 lancers d'un dé.
2. Calculer la probabilité d'obtenir au moins un double-six à l'issue de 24 lancers de deux dés.
3. A-t-on plus ou moins d'une chance sur deux d'obtenir au moins un six en lançant quatre fois un dé à six faces ?
4. A-t-on plus ou moins d'une chance sur deux d'obtenir au moins un double six en lançant vingt-quatre fois deux dés à six faces ?
5. Le texte ci-dessous reproduit l'extrait d'une lettre adressée par Fermat à Pascal en 1654.

> *Monsieur,*
> *Je n'ai pas eu le temps de vous envoyer la démonstration d'une difficulté qui étonnait fort M. de Méré. Il me disait donc qu'il avait trouvé fausseté dans les nombres par cette raison : si on entreprend de faire un six avec un dé, il y a avantage de l'entreprendre en 4, comme de 671 à 625. Si on entreprend de faire un double six avec deux dés, il y a désavantage de l'entreprendre en 24. Et néanmoins 24 est à 36 (qui est le nombre des faces de deux dés) comme 4 à 6 (qui est le nombre des faces d'un dé). Voilà quel était son grand scandale qui lui faisait dire hautement que les propositions n'étaient pas constantes et que l'arithmétique se démentait : mais vous en verrez bien aisément la raison par les principes où vous êtes.*

a. Commenter l'explication apportée par Fermat sur la probabilité d'obtenir au moins un six en lançant quatre fois un dé à six faces.

b. Quelle est l'erreur de raisonnement commise par le Chevalier de Méré ?

2 Inégalités de Markov et de Bienaymé-Tchebychev

Dans tout ce chapitre, on considère des variables aléatoires réelles définies sur un univers Ω fini muni d'une loi de probabilité $\mathbb{P}$.

2.1 Inégalité de Markov

Propriété 1 *Inégalité de Markov*

Si X est une variable aléatoire à valeurs positives et soit a un réel strictement positif, alors :

$$\mathbb{P}(X \geqslant a) \leqslant \frac{\mathbb{E}(X)}{a}$$

Interprétation : La probabilité que X prenne des valeurs plus grandes que a est d'autant plus petite que a est grand.

Démonstration

Soit Ω l'univers fini sur lequel est défini la variable aléatoire X.
On décompose l'ensemble image $X(\Omega)$ en deux sous-ensembles disjoints X_1 et X_2 tels que $X(\omega) = X_1 \cup X_2$.

- $X_1 = \{x \in X(\Omega),\ x \geqslant a\}$
- $X_2 = \{x \in X(\Omega),\ x < a\}$

....

....

....

....

....

....

....

....

2.2 Inégalité de Bienaymé-Tchebychev

Propriété 2 ***Inégalité de Bienaymé-Tchebychev***

Si X est une variable aléatoire d'espérance μ et de variance V alors :

$$\text{pour tout réel } \delta > 0, \quad \mathbb{P}\left(\left|X-\mu\right| \geqslant \delta\right) \leqslant \frac{V}{\delta^2}$$

Interprétation : L'inégalité de Bienaymé-Tchebychev met en évidence l'écart-type comme mesure de dispersion des valeurs d'une variable aléatoire autour de son espérance. La probabilité que les valeurs d'une variable aléatoire s'écartent de son espérance sont d'autant plus petites que sa variance et donc son écart-type, sont petits.

Démonstration

....

....

....

....

....

....

Capacité 1 ***Appliquer les inégalités de Markov et de Bienaymé-Tchebychev (Capacité 5 p. 407)***

Le nombre de croissants vendus dans une boulangerie en une journée est une variable aléatoire d'espérance 50 et de variance 25.

1. Déterminer un majorant de la probabilité que la vente du jour dépasse 75 par l'inégalité de Markov.
2. Déterminer un majorant de la probabilité que la vente du jour dépasse 75 par l'inégalité de Bienaymé-Tchebychev.
3. Comparer les deux majorations obtenues.

2.3 Limites de l'inégalité de Bienaymé-Tchebychev

Capacité 2 ***Inégalité de Bienaymé-Tchebychev et intervalle de fluctuation***

Au basket, Céline réussit son premier lancer franc dans 80% des cas.
À l'entraînement, sur une série de 100 lancers, on note X le nombre de paniers réussis par Céline.

1. **a.** Déterminer la loi de probabilité de la variable aléatoire X.
   b. Déterminer l'espérance μ et l'écart-type σ de X.
   c. Démontrer à l'aide de l'inégalité de Bienaymé-Tchebychev que :

$$\mathbb{P}\left(\left|X-\mu\right| \geqslant 2\sigma\right) \leqslant 0,25$$

2. **a.** Compléter la fonction Python ci-dessous pour qu'elle simule une réalisation de la variable aléatoire X.

```
from random import random

def X():
    """Simule la réalisation d'une va de loi B(100, 0.8)"""
    s = 0
    for k in range(100):
        if random() ......... 0.8: #succès
            s = ..........
    return s
```

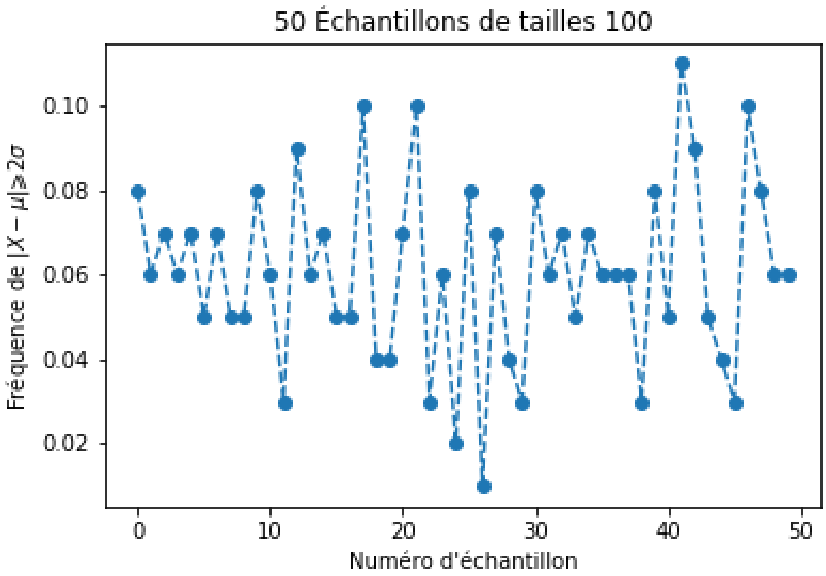
b. Compléter la fonction Python ci-dessous pour qu'elle renvoie un échantillon de n réalisations de la variable aléatoire X.

```
def echantillon(n):
    """Renvoie une liste de n réalisations de X"""
    return [............ for k in range(n)]
```

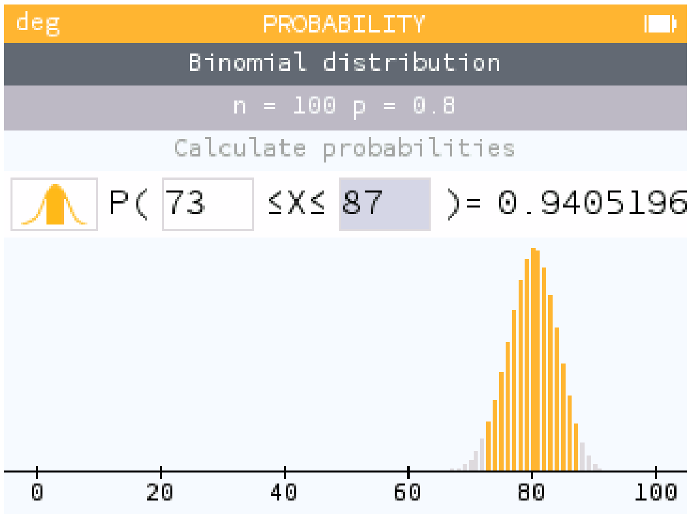
c. Compléter la fonction Python ci-dessous pour qu'elle renvoie la fréquence de réalisations de la variable aléatoire X telles que l'écart avec son espérance μ est supérieur à deux fois son écart-type σ.

```
def frequence_fluctuation_2sigma(n):
    """Mesure la fréquence de réalisations de la va X
    d'espérance mu et d'écart-type sigma
    telles que |X-mu|>= 2* sigma sur un échantillon
    de taille n"""
    sigma = ..........
    mu = ..........
    compteur = 0
    echant = echantillon(n)
    for val in echant:
        if abs(............) >= ..............:
            compteur = ..............
    return compteur / n
```

d. Le graphique ci-dessous représente un échantillon de valeurs de `frequence_fluctuation_2sigma(100)`.
Comparer avec le majorant donné par l'inégalité de Bienaymé-Tchebychev.



e. Démontrer que $\mathbb{P}(|X-\mu| \geqslant 2\sigma) = 1 - \mathbb{P}(73 \leqslant X \leqslant 87)$.
$[73\,;87]$ est un intervalle de fluctuation de X au seuil de 94%. Justifier cette dénomination.



Capacité 3 *Inégalité de Bienaymé-Tchebychev et probabilité des écarts à l'espérance*

Soit X une variable aléatoire d'espérance μ et d'écart-type σ.

1. **a.** Démontrer que pour tout entier naturel $k \geqslant 1$, on a :

$$\mathbb{P}(|X-\mu| \geqslant k\sigma) \leqslant \frac{1}{k^2}$$

b. Commenter cette phrase du programme officiel :

L'inégalité de Bienaymé-Tchebychev montre que des écarts de X à μ de quelques σ deviennent improbables.

2. **a.** Déterminer une valeur de k pour laquelle $\mathbb{P}(|X-\mu| \geqslant k\sigma) \leqslant 0,1$.

b. Interpréter la valeur de k obtenue.

3. Démontrer que pour tout réel $\delta \geqslant 1$, on a :

$$\mathbb{P}\left(\mu - \delta\sigma < X < \mu + \delta\sigma\right) \geqslant \frac{\delta^2 - 1}{\delta^2}$$

3 Inégalité de concentration et loi faible des grands nombres

3.1 Retour sur les sommes de variables aléatoires

Capacité 4

On rappelle qu'une variable aléatoire X définie sur un univers fini Ω muni d'une loi de probabilité $\mathbb{P}$ suit une loi de Bernoulli de paramètre p si elle prend comme valeurs 1 et 0 avec les probabilités $\mathbb{P}\left(X = 0\right) = p$ et $\mathbb{P}\left(X = 1\right) = 1 - p$.
Soit n un entier strictement positif, $X_1, X_2, \dots, X_n$ désignent n variables aléatoires indépendantes de même loi de Bernoulli de paramètre p.
On définit les variables aléatoires :

$$S_n = X_1 + X_2 + \ldots + X_n = \sum_{k=1}^{n} X_k \quad \text{et} \quad M_n = \frac{S_n}{n}$$

1. Déterminer l'espérance et la variance de X_1.

2. **a.** Que représente la variable S_n ? Quelle est sa loi de probabilité ?

b. Déterminer l'espérance et la variance de la variable aléatoire S_n.

3. **a.** Que représente la variable M_n ? Quelle est sa loi de probabilité ?

b. Déterminer l'espérance et la variance de la variable aléatoire M_n.

3.2 Inégalité de concentration

Propriété 3 *Inégalité de concentration*

Soit n un entier strictement positif et soit , $X_1, X_2, \dots, X_n$, n variables aléatoires indépendantes, suivant une même loi de probabilité d'espérance μ et de variance V.
On note $M_n = \dfrac{X_1 + X_2 + \ldots + X_n}{n}$ la variable aléatoire moyenne sur l'échantillon de taille n constitué par le n-uplet $(X_1, X_2, \dots, X_n)$.
M_n vérifie **l'inégalité de concentration** :

$$\text{pour tout réel } \delta > 0, \quad \mathbb{P}\left(\left| M_n - \mu \right| \geqslant \delta\right) \leqslant \frac{V}{n\delta^2}$$

Démonstration

....

....

....

....

....

....

....

....

Corollaire ***Inégalité de concentration pour une variable de Bernoulli***

Soit n un entier strictement positif et un échantillon de taille n d'une variable aléatoire suivant une loi de Bernoulli de paramètre $p > 0$, d'espérance p et de variance $p(1-p)$.
La variable aléatoire moyenne M_n sur cet échantillon de taille n vérifie l'inégalité de concentration :

$$\text{pour tout réel } \delta > 0, \quad \mathbb{P}\left(\left| M_n - p \right| \geqslant \delta\right) \leqslant \frac{p(1-p)}{n\delta^2}$$

Démonstration

....

....

....

....

....

Capacité 5 ***Appliquer l'inégalité de Bienaymé-Tchebychev pour définir une taille d'échantillon, en fonction de la précision et du risque choisi***

On lance une pièce équilibrée n fois de suite avec n un entier strictement positif.
La variable aléatoire M_n donne la proportion de Pile obtenus au cours des n lancers.

1. Déterminer l'espérance et la variance de M_n.
2. Justifier que $\mathbb{P}\left(\left| M_n - 0,5 \right| \geqslant 0,1\right) \leqslant \dfrac{25}{n}$.
3. Déterminer une valeur de n telle que $\mathbb{P}\left(\left| M_n - 0,5 \right| \geqslant 0,1\right) \leqslant 0,05$.
   On dit que l'on obtient pour M_n une précision de $0,1$ avec un risque de $0,05$.

3.3 Loi faible des grands nombres de Bernoulli

Propriété 4

Soit n un entier strictement positif et soit , $X_1, X_2, \ldots, X_n$, n variables aléatoires indépendantes, suivant une même loi de probabilité d'espérance μ et de variance V.

On note $M_n = \dfrac{X_1 + X_2 + \ldots + X_n}{n}$ la variable aléatoire moyenne sur l'échantillon de taille n constitué par le n-uplet $(X_1, X_2, \ldots, X_n)$.

Pour tout réel $\delta > 0$, on a :

$$\lim_{n \to +\infty} \mathbb{P}\left(\left| M_n - \mu \right| \geqslant \delta\right) = 0$$

Interprétation : L'écart entre la moyenne d'un échantillon d'une variable aléatoire et l'espérance de cette variable ne dépasse une valeur donnée à l'avance qu'avec une probabilité qui tend vers zéro quand la taille de l'échantillon tend vers l'infini.

Démonstration ***Au programme***

....

....

....

....

....

....

....

....

Capacité 6 ***Utiliser la loi faible des grands nombres pour estimer une probabilité***

Soit X une variable aléatoire suivant une loi de Bernoulli de paramètre $p > 0$. Pour tout entier $n > 0$, on note M_n la variable aléatoire moyenne sur un échantillon de taille n de la variable aléatoire X.

1. Démontrer que pour tout réel $\varepsilon > 0$ et tout entier $n > 0$, on a $\mathbb{P}\left(\left| M_n - p \right| \geqslant \varepsilon\right) \leqslant \dfrac{1}{4n\varepsilon^2}$.
2. En déduire une méthode pour estimer la probabilité p si elle est inconnue.

Capacité 7 ***Estimer une espérance avec la loi faible des grands nombres***

On veut estimer l'espérance d'une variable aléatoire qui renvoie le rang du premier 6 lors d'une série de lancers d'un dé à 6 faces équilibré.

1. Compléter les fonctions Python ci-dessous en respectant la spécification donnée en documentation.

```python
from random import randint

def premier6():
```

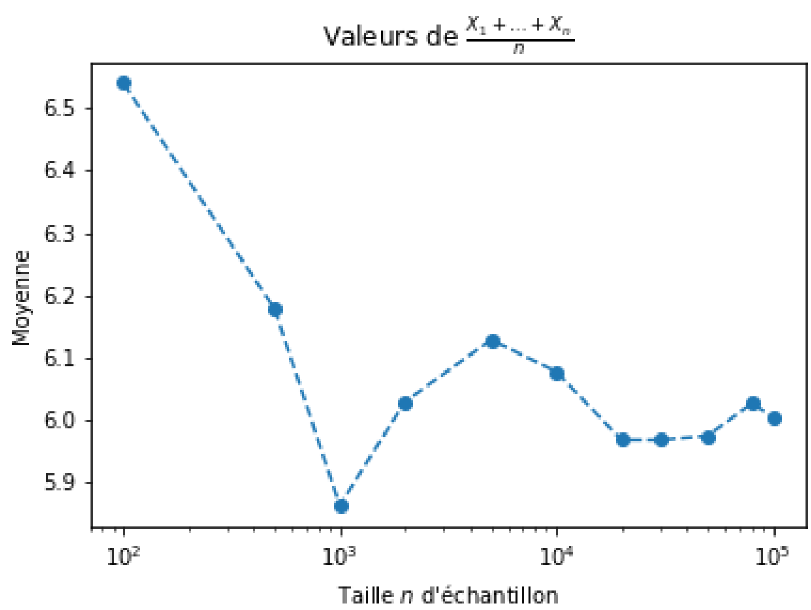
```python
    """Renvoie le rang du premier 6
    lors d'une série de lancers de dés à 6 faces équilibrés"""
    n = 1
    while randint(1, 6) .....:
        n = .....
    return n

def echantillon_premier6(n):
    """Renvoie un échantillon de n réalisations
    de la va premier6"""
    return [.............. for k in range(n)]

def moyenne(liste):
    """Renvoie la moyenne d'une liste de nombre"""
    s = 0
    for ................:
        s = ............
    return .............

def moyenne_echantillon_premier6(n):
    """Renvoie la moyenne sur un échantillon
    de taille n de la va premier6"""
    return ......................
```

2. On a représenté sur le graphique ci-dessous des valeurs de `moyenne_echantillon_premier6(n)` pour des valeurs de `n` croissantes. Quelle conjecture peut-on faire sur l'espérance de la variable aléatoire renvoyant le rang du premier 6 ? Voir le DM sur la loi géométrique !



Algorithmique 1 *Déterminer une fréquence d'échantillon*

S_n est une variable aléatoire qui suit une loi binomiale $\mathscr{B}(n\,;\,p)$ avec n entier naturel, $n \geqslant 1$ et $p \in]0\,;\,1[$.

1. Démontrer que $\mathbb{P}\left(\left|S_n - np\right| \geqslant \sqrt{n}\right) \leqslant p(1-p)$ avec l'inégalité de Bienaymé-Tchebychev.

2. Compléter la fonction Python `S(n,p)` ci-dessous pour qu'elle renvoie une réalisation de la variable aléatoire S_n.

```python
from random import random

def S(n, p):
    """Renvoie une simulation d'une va de loi B(n,p)"""
    s = 0
    for k in range(n):
        if random() ....... p:
            s = .........
    return s
```

3. Compléter la fonction Python `echantillon_S(n,p,m)` ci-dessous pour qu'elle renvoie un échantillon de taille m de réalisations de la variable aléatoire S_n.

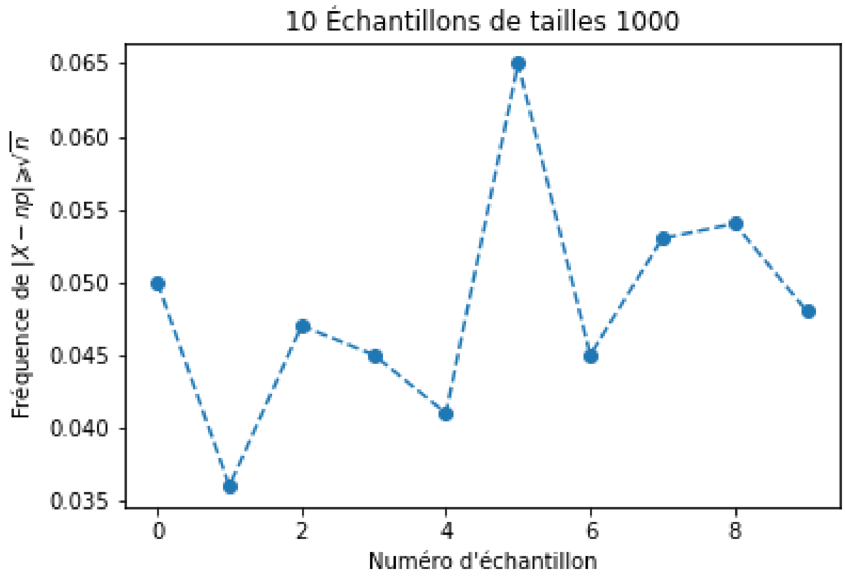
```python
def echantillon_S(n,p,m):
    """Renvoie un échantillon de taille m
    d'une va S de loi B(n,p)"""
    return [.......... for k in range(m)]
```

4. Compléter la fonction Python `frequence_fluctuation_echantillon(n, p, m)` ci-dessous pour qu'elle renvoie la fréquence de réalisations de la variable aléatoire S_n à une distance de l'espérance $\mathbb{E}(S_n) = np$ supérieure ou égale à $\sqrt{n}$. sur un échantillon de taille m de réalisations de la variable aléatoire S_n.

```python
def frequence_fluctuation_echantillon(n, p, m):
    """Renvoie la fréquence de réalisations telles que
    |S(n,p) - n * p| >= sqrt(n)
    sur un échantillon de taille m d'une va S de loi B(n,p)"""
    b = 0
    echant = echantillon_S(n,p,m)
    for val in echant:
        if ....................:
            b = b + 1
    return b / m
```

5. Le graphique ci-dessous représente un échantillon de dix valeurs de `frequence_fluctuation_echantil` `0.4, 1000)`.

   Comparer les fréquences obtenues avec la majoration de $\mathbb{P}\left(\left|S_n - np\right| \geqslant \sqrt{n}\right)$ obtenue en question **1.**



4 Applications

4.1 Recherche d'une condition suffisante sur la taille d'un échantillon (Capacité 6 p. 407)

Histoire 3 ***Retour sur le problème du chevalier de Méré***

1. En appliquant l'inégalité de Bienaymé-Tchebychev, déterminer au bout de combien de répétitions d'un lancer de quatre dés, la fréquence d'apparition d'au moins un six est supérieure ou égale à $\frac{1}{2}$ avec une probabilité supérieure ou égale à $0,95$.

2. En appliquant l'inégalité de Bienaymé-Tchebychev, déterminer au bout de combien de répétitions de vint-quatre lancers de deux dés, la fréquence d'apparition d'au moins un double six est inférieure ou égale à $\frac{1}{2}$ avec une probabilité supérieure ou égale à $0,95$.

4.2 Étude d'une marche aléatoire

Algorithmique 2 ***Manuel Indice, TP 2 page 423***

Table des matières