

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 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} .

1.1 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 :

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

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 raisonne par disjonction des cas en décomposant l'ensemble image $X(\Omega)$ en deux sous-ensembles disjoints X_1 et X_2 tels que $X(\omega) = X_1 \cup X_2$.

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

.....
.....
.....

1.2 Inégalité de Bienaymé-Tchebychev



Propriété 2 Inégalité de Bienaymé-Tchebychev

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

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

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 l'inégalité de Bienaymé-Tchebychev (Capacité 5 p. 407)

Métropole juin 2024 jour 1.

Lorsqu'une commande de téléviseur est passée par un client, on considère que le temps de livraison du téléviseur est modélisé par une variable aléatoire T égale à la somme de deux variables aléatoires T_1 et T_2 . La variable aléatoire T_1 modélise le nombre entier de jours pour l'acheminement du téléviseur depuis un entrepôt de stockage vers une plateforme de distribution. La variable aléatoire T_2 modélise le nombre entier de jours pour l'acheminement du téléviseur depuis cette plateforme jusqu'au domicile du client.

On admet que les variables aléatoires T_1 et T_2 sont indépendantes, et on donne :

- L'espérance $E(T_1) = 4$ et la variance $V(T_1) = 2$;
- L'espérance $E(T_2) = 3$ et la variance $V(T_2) = 1$.

1. Déterminer l'espérance $E(T)$ et la variance $V(T)$ de la variable aléatoire T .
2. On note A l'événement « Le client reçoit sa commande entre 5 et 9 jours après sa commande, bornes comprises ». Exprimer A en fonction de T et de $E(T)$.

3. Sélectionner la ou les expressions correctes de l'événement contraire \bar{A} :

- $\bar{A} = |T - E(T) > 2|$
- $\bar{A} = |T - E(T) \geq 2|$
- $\bar{A} = |T - E(T) \geq 3|$

4. Majorer $\mathbb{P}(\bar{A})$ avec l'inégalité de *Bienayme-Tchebychev*.

5. En déduire une minoration de $\mathbb{P}(A)$.

Capacité 2 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 $k \geq 1$, on a :

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

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 $\mathbb{P}(|X - \mu| \geq k\sigma) \leq 0,1$.

b. Interpréter la valeur de k obtenue.

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

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

2 Inégalité de concentration et loi faible des grands nombres

2.1 Retour sur les sommes de variables aléatoires

Capacité 3

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 $\mathbb{P}(X = 0) = p$ et $\mathbb{P}(X = 1) = 1 - p$.

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 :

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

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 .

2.2 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 $M_n = \frac{X_1 + X_2 + \dots + X_n}{n}$ 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** :

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

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 :

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

Démonstration

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Capacité 4 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 $\mathbb{P}(|M_n - 0,5| \geq 0,1) \leq \frac{25}{n}$.
3. Déterminer une valeur de n telle que $\mathbb{P}(|M_n - 0,5| \geq 0,1) \leq 0,05$.

On dit que l'on obtient pour M_n une précision de 0,1 avec un risque de 0,05.

2.3 Loi faible des grands nombres de Bernoulli

Propriété 4

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

Pour tout réel $\delta > 0$, on a :

$$\lim_{n \rightarrow +\infty} \mathbb{P}(|M_n - \mu| \geq \delta) = 0$$

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 5 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 $\mathbb{P}(|M_n - p| \geq \varepsilon) \leq \frac{1}{4n\varepsilon^2}$.
2. En déduire une méthode pour estimer la probabilité p si elle est inconnue.

Algorithmique 1 Déterminer une fréquence d'échantillon

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

1. Démontrer que $\mathbb{P}(|S_n - np| \geq \sqrt{n}) \leq p(1-p)$ 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 .

```
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 = s + 1
    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 .

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

```
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 abs(val - n * p) >= sqrt(n):
            b = b + 1
    return b / m
```

5. Le graphique ci-dessous représente un échantillon de dix valeurs de `frequence_fluctuation_echantillon(100, 0.4, 1000)`.
 Comparer les fréquences obtenues avec la majoration de $\mathbb{P}(|S_n - np| \geq \sqrt{n})$ obtenue en question 1.

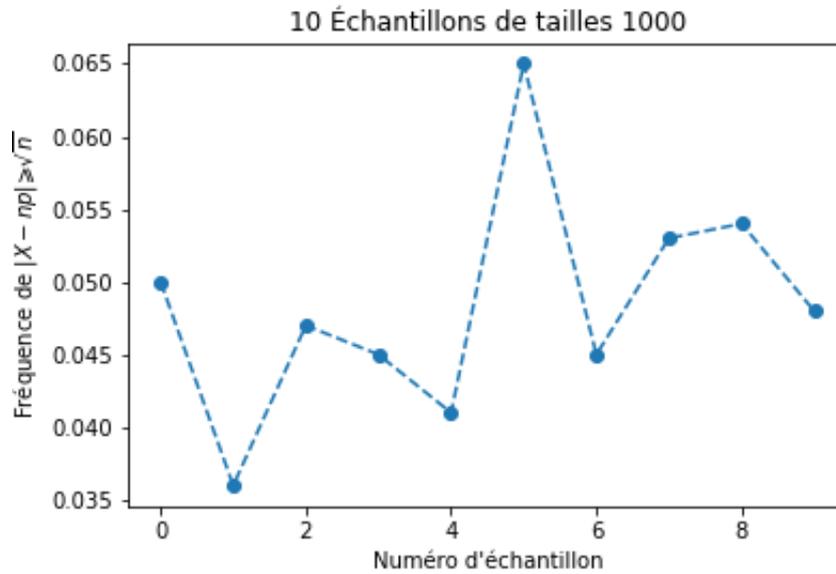


Table des matières

1	Inégalités de Markov et de Bienaymé-Tchebychev	1
1.1	Inégalité de Markov	1
1.2	Inégalité de Bienaymé-Tchebychev	2
2	Inégalité de concentration et loi faible des grands nombres	3
2.1	Retour sur les sommes de variables aléatoires	3
2.2	Inégalité de concentration	4
2.3	Loi faible des grands nombres de Bernoulli	5