Chapitre 8 19 mai 2026

Concentration et loi des grands nombres

Introduction

Objectifs du chapitre et Capacités attendues :

— Comprendre et appliquer l'inégalité de Bienaymé-Tchebychev
— Maîtriser l'inégalité de concentration pour les moyennes d'échantillons
— Énoncer et interpréter la loi des grands nombres
— Utiliser ces outils pour majorer ou minorer des probabilités
— Appliquer ces concepts à des situations concrètes

Annexes historiques

Histoire des probabilités :

— **Jacob Bernoulli** (1654-1705) : Premier énoncé de la loi des grands nombres dans "Ars Conjectandi"
— **Irinée-Jules Bienaymé** (1796-1878) : Statisticien français qui a généralisé les inégalités
— **Pafnoutti Tchebychev** (1821-1894) : Mathématicien russe qui a formalisé l'inégalité
— **Andreï Markov** (1856-1922) : Élève de Tchebychev, a développé l'inégalité de Markov

1 Inégalité de Bienaymé-Tchebychev

Définition 1.1: Variable aléatoire positive ou nulle

Une variable aléatoire X est dite positive ou nulle lorsque toutes les valeurs prises par celle-ci sont des réels positifs ou nuls.

$$\forall \omega \in \Omega, \quad X(\omega) \geq 0$$

Théorème 1.1: Inégalité de Markov

Soit X une variable aléatoire réelle **positive ou nulle** d'espérance $E(X)$. Alors, pour tout réel $a > 0$:

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

Démonstration. Soit X une variable aléatoire réelle positive ou nulle dont on note x_i les valeurs possibles.

Étape 1 : Décomposition de l'espérance :

$$E(X) = \sum_i x_i \cdot P(X = x_i)$$

Étape 2 : Séparation des termes : On sépare la somme en deux parties :

$$E(X) = \sum_{x_i < a} x_i \cdot P(X = x_i) + \sum_{x_i \geq a} x_i \cdot P(X = x_i)$$

Étape 3 : Minoration des termes : Comme X est positive ou nulle, tous les termes sont positifs. En particulier :

$$\sum_{x_i < a} x_i \cdot P(X = x_i) \geq 0$$

Donc :

$$E(X) \geq \sum_{x_i \geq a} x_i \cdot P(X = x_i)$$

Étape 4 : Utilisation de la contrainte $x_i \geq a$:
Pour tout $x_i \geq a$, on a $x_i \geq a$, donc :

$$x_i \cdot P(X = x_i) \geq a \cdot P(X = x_i)$$

Par conséquent :

$$\sum_{x_i \geq a} x_i \cdot P(X = x_i) \geq \sum_{x_i \geq a} a \cdot P(X = x_i)$$

Étape 5 : Factorisation et probabilité totale

$$\sum_{x_i \geq a} a \cdot P(X = x_i) = a \cdot \sum_{x_i \geq a} P(X = x_i) = a \cdot P(X \geq a)$$

Étape 6 : Conclusion En combinant les inégalités :

$$E(X) \geq a \cdot P(X \geq a)$$

Comme $a > 0$, on peut diviser par a :

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

□

Remarque
L'inégalité de Markov donne une borne supérieure pour la probabilité qu'une variable aléatoire positive dépasse un certain seuil. Elle est particulièrement utile quand on ne connaît pas la loi exacte de la variable, mais seulement son espérance.

Théorème 1.2: Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire d'espérance $E(X) = \mu$ et de variance $V(X)$. Pour tout réel $\delta > 0$:

$$P(|X-\mu| \geq \delta) \leq \frac{V(X)}{\delta^2}$$

$$P(|X-\mu| < \delta) \geq 1 - \frac{V(X)}{\delta^2}$$

Démonstration : **Étape 1 : Variable auxiliaire :** Considérons la variable aléatoire $Y = (X-\mu)^2$. Cette variable est positive ou nulle car c'est un carré.

Étape 2 : Espérance de Y

$$E(Y) = E[(X-\mu)^2] = V(X)$$

Étape 3 : Lien avec l'événement d'intérêt : Remarquons que :

$$|X-\mu| \geq \delta \iff (X-\mu)^2 \geq \delta^2$$

Donc :

$$P(|X-\mu| \geq \delta) = P((X-\mu)^2 \geq \delta^2)$$

Étape 4 : Application de l'inégalité de Markov : Appliquons l'inégalité de Markov à Y avec le seuil δ^2 :

$$P(Y \geq \delta^2) \leq \frac{E(Y)}{\delta^2} = \frac{V(X)}{\delta^2}$$

Étape 5 : Conclusion de la première inégalité :

$$P(|X-\mu| \geq \delta) \leq \frac{V(X)}{\delta^2}$$

Étape 6 : Seconde inégalité : Les événements $|X-\mu| \geq \delta$ et $|X-\mu| < \delta$ sont complémentaires, donc :

$$P(|X-\mu| \geq \delta) \leq \frac{V(X)}{\delta^2}$$

$$-P(|X-\mu| \geq \delta) \geq -\frac{V(X)}{\delta^2}$$

$$1 - P(|X-\mu| \geq \delta) \geq 1 - \frac{V(X)}{\delta^2}$$

$$P(|X-\mu| < \delta) \geq 1 - \frac{V(X)}{\delta^2}$$

□

Remarque
L'inégalité de Bienaymé-Tchebychev montre que plus la variance est petite, plus la variable est concentrée autour de son espérance. Elle permet de quantifier cette concentration sans connaître la loi exacte de la variable.

Exemple (Interprétation géométrique)
Si $V(X)$ est petite, alors X est très concentrée autour de μ.
Si $V(X)$ est grande, alors X peut s'éloigner significativement de μ.

2 Inégalité de concentration

Définition 2.1: Échantillon et moyenne

Soit $(X_1, X_2, \dots, X_n)$ un échantillon de variables aléatoires indépendantes et identiquement distribuées d'espérance μ et de variance V. La moyenne de l'échantillon est :

$$M_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

On dit que les variables sont :

- **Indépendantes** : la réalisation d'une variable n'influence pas les autres
- **Identiquement distribuées** : toutes suivent la même loi de probabilité

Théorème 2.1: Inégalité de concentration

Soit M_n la moyenne d'un échantillon de taille n. Pour tout $\delta > 0$:

$$P(|M_n - \mu| \geq \delta) \leq \frac{V}{n\delta^2}$$

$$P(|M_n - \mu| < \delta) \geq 1 - \frac{V}{n\delta^2}$$

Démonstration : **Étape 1 : Espérance de la moyenne** Par linéarité de l'espérance :

$$E(M_n) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n} \cdot n\mu = \mu$$

Étape 2 : Variance de la moyenne Comme les variables sont indépendantes :

$$V(M_n) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{1}{n^2} \cdot nV = \frac{V}{n}$$

Étape 3 : Application de Bienaymé-Tchebychev Appliquons l'inégalité de Bienaymé-Tchebychev à M_n :

$$P(|M_n - E(M_n)| \geq \delta) \leq \frac{V(M_n)}{\delta^2}$$

Étape 4 : Substitution En remplaçant $E(M_n) = \mu$ et $V(M_n) = \frac{V}{n}$:

$$P(|M_n - \mu| \geq \delta) \leq \frac{V}{n\delta^2}$$

Étape 5 : Inégalité complémentaire

$$P(|M_n - \mu| \geq \delta) \leq \frac{V}{n\delta^2}$$

$$1 - P(|M_n - \mu| \geq \delta) \geq 1 - \frac{V}{n\delta^2}$$

$$P(|M_n - \mu| < \delta) \geq 1 - \frac{V}{n\delta^2}$$

□

Remarque
L'inégalité de concentration montre que la moyenne d'un échantillon se concentre autour de l'espérance théorique quand la taille de l'échantillon augmente. Le terme $\frac{1}{n}$ dans la variance explique pourquoi la concentration s'améliore avec n.

3 Loi des grands nombres

Théorème 3.1: Loi faible des grands nombres

Soit $(X_1, X_2, \ldots, X_n)$ un échantillon de variables aléatoires d'espérance μ et M_n leur moyenne. Pour tout $\delta > 0$:

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

Démonstration : **Étape 1 : Inégalité de concentration :** D'après l'inégalité de concentration, on a :

$$P(|M_n - \mu| \geq \delta) \leq \frac{V}{n\delta^2}$$

Étape 2 : Encadrement : Comme toute probabilité est positive :

$$0 \leq P(|M_n - \mu| \geq \delta) \leq \frac{V}{n\delta^2}$$

Étape 3 : Passage à la limite : Quand $n \to +\infty$, le majorant tend vers 0 :

$$\lim_{n\to+\infty} \frac{V}{n\delta^2} = 0$$

Étape 4 : Théorème des gendarmes : Par le théorème des gendarmes :

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

Étape 5 : Interprétation : Ceci signifie que pour n suffisamment grand, la moyenne empirique M_n est arbitrairement proche de l'espérance théorique μ avec une probabilité arbitrairement proche de 1. □

Remarque
La loi des grands nombres justifie l'approche fréquentiste des probabilités : la fréquence relative d'un événement converge vers sa probabilité théorique quand le nombre d'expériences augmente.

Exemple (Convergence en pratique)
Pour un dé équilibré :

- Avec 10 lancers, la moyenne peut être entre 2,5 et 4,5
- Avec 100 lancers, la moyenne est généralement entre 3,2 et 3,8
- Avec 1000 lancers, la moyenne est presque toujours entre 3,4 et 3,6
- Avec 10000 lancers, la moyenne est extrêmement proche de 3,5

4 Applications avancées et contre-exemples

Exemple (Cas pathologique)
Considérons une variable aléatoire qui prend la valeur n^2 avec probabilité $\frac{1}{n^2}$ et 0 avec probabilité $1 - \frac{1}{n^2}$.
Son espérance est 1, mais elle peut prendre des valeurs arbitrairement grandes avec probabilité faible. L'inégalité de Markov donne une borne qui peut être très pessimiste.

Remarque (Optimalité des inégalités)
Les inégalités de Markov et Bienaymé-Tchebychev sont optimales dans le sens où il existe des variables aléatoires pour lesquelles l'égalité est atteinte (ou approchée arbitrairement).

5 Applications et interprétations pratiques

Propriété 5.1: Interprétation fréquentiste

La loi des grands nombres justifie l'approche fréquentiste des probabilités :

$$\lim_{n\to+\infty} \frac{\text{Nombre de succès}}{n} = p$$

où p est la probabilité théorique du succès.

Exemple
Dans un casino, la roulette a une probabilité de 18/37 pour un numéro simple. Sur 1000 parties, la fréquence de gain peut être de 17,8%, sur 10 000 parties de 18,2%, et sur 100 000 parties elle se rapproche de 18/37 ≈ 48,65%.

6 Cas des lois binomiales

Propriété 6.1: Application aux fréquences binomiales

Soit $X \hookrightarrow B(n,p)$ et $F_n = \frac{X}{n}$ la fréquence de succès.

$$P(|F_n - p| \geq \delta) \leq \frac{p(1-p)}{n\delta^2} \leq \frac{1}{4n\delta^2}$$

car $p(1-p) \leq \frac{1}{4}$ pour tout $p \in [0,1]$.

Démonstration. $X = X_1 + \cdots + X_n$ où $X_i \hookrightarrow \mathcal{B}(p)$, donc $F_n = \frac{X_1+\cdots+X_n}{n}$.
L'inégalité de concentration donne le résultat. □