# LOIS DES GRANDS NOMBRES

**Rédigé par** :

Hippolyte SODJINOU

...

*Adresse mail* : **sodjinouhippolyte@gmail.com**

Année académique 2023-2024

# Table des matières

## Introduction

La loi des grands nombres est un énoncé central du calcul des probabilités, qui en particulier fait asymptotiquement émerger le déterminisme au sein d’un modèle désordonné (aléatoire). Elle illustre une intuition naturelle et commune : lancer une pièce de monnaie un grand nombre de fois, la proportion ou moyenne « empirique » (rapport entre le nombre de pile et le nombre total de lancers) tend à s’approcher de la moyenne « théorique » de la pièce ( 1 2 si la pièce est équilibrée). Sa formalisation, et la démonstration de l’énoncé le plus complet, ont toutefois mis du temps historiquement à s’affirmer. Il existe des versions innombrables de la loi des grands nombres, et cette leçon ne concerne essentiellement que la forme la plus traditionnelle pour des sommes de variables aléatoires indépendantes équidistribuées, correspondant à une expérience répétée.

# 1 Notion de variables aléatoires

## 1.1 Tribu

**Définition 1.1**. Soit Ω un ensemble non vide. On appelle tribu sur Ω, une famille $\mathcal{F}$ de parties de Ω vérifiant :

- $\Omega \in \mathcal{F}$.
- $\forall A \in \mathcal{F}, A^C \in \mathcal{F}$ (stabilité par passage au complémentaire),
- $((A_n)_{n\in N^*} \subset \mathcal{F}) \Longrightarrow (\bigcup_{n\in N^*} An) \in \mathcal{F}$ (stabilité par réunion dénombrable).

**Exemple 1.1**. Sont des tribus de parties de Ω,

- $\mathcal{F}_0 = \{\emptyset, \Omega\}$ .
- $\mathcal{F}_1 = P(\Omega)$.
- Pour toute partie non vide A de Ω, $\mathcal{F} = \{\emptyset, A, A^C, \Omega\}$

**Proposition 1.1**. Soit $\mathcal{F}$ une tribu sur un ensemble Ω; on a :

- $\mathcal{F}$ est stable par réunion finie.
- $\mathcal{F}$ est stable par intersection dénombrable ou finie.
- Si A et B sont des éléments de $\mathcal{F}$, $A \setminus B$ et $A \Delta B$ sont aussi des éléments de $\mathcal{F}$.

**Définition 1.2.** On appelle espace mesurable, un couple $(\Omega, \mathcal{F})$, où Ω est un ensemble non vide et $\mathcal{F}$ une tribu de parties de Ω.

**Définition 1.3.**

1. Soit Ω un ensemble fini et $\mathcal{F}$ une tribu de parties de Ω. Le couple $(\Omega, \mathcal{F})$ est appelé espace probabilisable.

2. On appelle probabilité définie sur l'espace probabilisable $(\Omega, \mathcal{F})$, toute application *p* définie de $\mathcal{F}$ dans [0, 1] qui vérifie les propriétés suivantes :

   a) $p(\Omega) = 1$.
   b) Si *A* et *B* deux évènements incompatibles c'est-à-dire $A \cap B=\emptyset$, on a : $p(A \cup B)=p(A) + p(B)$.

Le triplet $(\Omega, \mathcal{F}, p)$ est appelé espace probabilisé. Pour tout évènement *A*, le réel *p(A)* est appelé probabilité de l'évènement *A*.

**Proposition 1.2.** Soit $(\Omega, \mathcal{F}, p)$ un espace probabilisé fini, $A$ et $B$ deux évènements de $\Omega$. On a :

1. $p(\emptyset) = 0$ .
2. $p(\bar{A}) = 1 - p(A)$.
3. $p(A \setminus B) = p(A) - p(A \cap B)$.
4. *Si* $A \subset B$, *alors on a :* $p(A) \leq p(B)$.
5. $p(A \cup B) = p(A) + p(B) - p(A \cap B)$.

## 1.2 Variable aléatoire réelle

On désigne par $\mathcal{P}(R)$ l'ensemble des parties de $R$.

**Définition 2.1.** Soit $(\Omega, \mathcal{F})$ un espace probabilisable ; on appelle variable aléatoire réelle (v.a.r.) définie sur $(\Omega, \mathcal{F})$ toute application :

$$X : (\Omega, \mathcal{F}) \longrightarrow (R, \mathcal{B})$$

telle que $X^{-1}(\mathcal{B}) \subset \mathcal{F}$. C'est-à -dire que $X^{-1}(\mathcal{B}) \in \mathcal{F}$ pour tout $B \in \mathcal{B}$. L'ensemble $H = X(\Omega)$ est appelé ensemble fondamental de la variable aléatoire réelle X.

$\mathcal{B}$ est la tribu engendrée par l'ensemble des parties de $\mathcal{R}$ (tribu borélienne).

**Remarque.** Lorsque $\mathcal{F} = P(\Omega)$, alors toute application X définie de $\Omega$ dans $R$ est une variable aléatoire réelle.

### 1.2.1 Fonction de répartition

**Définition 2.2 :**

Soit X une v.a.r. définie sur un espace probabilisé $(\Omega, \mathcal{F}, p)$. L'application F définie sur $R$, à valeurs dans [0, 1], par :

$$F : R \longrightarrow [0, 1]$$

$$x \mapsto F(x) := p(X < x)$$

est appelée la fonction de répartition de X.

**Remarque :** On a :

$$\lim_{x \to -\infty} F(x) = 0$$

$$\lim_{x \to +\infty} F(x) = 1$$

**Théorème .** La fonction de répartition d'une v.a.r. X est croissante et continue à gauche en chacun de ses points de discontinuité. En chaque point $x_0$ de discontinuité, le saut de la fonction de répartition F, c'est-à-dire $F(x_0 + 0) - F(x_0)$ est la probabilité que X prenne la valeur $x_0$. On a alors :

$$F(x_0 + 0) - F(x_0) = p(X = x_0) = P_X(x_0).$$

**Théorème.** Soit F une fonction définie sur $R$ à valeurs dans [0, 1] vérifiant les propriétés suivantes :

- F est croissante ;
- $\lim_{x\to-\infty} F(x) = 0$ et $\lim_{x\to-\infty} F(x) = 0$;
- F est continue à gauche en chacun de ses points de continuité. Alors F est la fonction de répartition d'une variable aléatoire réelle.

## 1.3 Variables aléatoires discrètes

**Définition .** Une v.a.r. est dite discrète si son ensemble fondamental $\mathcal{H} = X(\Omega)$ est un sous ensemble dénombrable de $R$. Dans ce cas, on notera $\{ x_n, n \in N^* \}$ les éléments de $\mathcal{H}$ rangés par ordre de valeurs croissantes.

**Exemple** :

On considère X(Ω) = {1, 2, 3, 4, 5, 6} et Y (Ω) = {0, 1, 2, 3, 4, 5}. On peut donc dire que les v.a.r. X et Y sont discrètes finies.

### 1.3.1 Loi de probabilité d'une v.a.r. discrète finie

Soit X une v.a.r. sur l'espace probabilisé $(\Omega, \mathcal{F}, p)$ telle que $X(\Omega) = \{x_1, x_2, \cdots, x_n\}$ avec $x_1 < x_2 < \cdots < x_n$. On appelle loi de probabilité de X, l'ensemble des couples $(x_i , p_i)_{1\leq i\leq n}$ où $x_i \in X(\Omega)$ et $p_i = p(X = x_i)$.

#### *1.3.1.1 Indépendance de deux v.a.r.*

Soient X et Y deux v.a.r. définies sur un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), p)$. On dit que X et Y sont indépendantes pour la probabilité $p$ si pour tout $(x, y) \in X(\Omega)\times Y(\Omega)$, on a :

$$p[(X = x) \cap (Y = y)] = p(X = x) \cdot p(Y = y).$$

### 1.3.2 Moments d'une v.a.r. discrète finie

#### *1.3.2.1 Moments d'ordre k avec k $\in N^*$*

Soient X une v.a.r. de loi $\{((x_i, p_i) / 1 \leq i \leq n)\}$, k un entier naturel non nul. On appelle moment d'ordre k de X, le nombre réel noté $m_k(X)$ défini par :

$$m_k(X)= \sum_{i=1}^{n} (x_i)^k p_i .$$

#### 1.3.2.2 Espérance mathématique

**Définition .** L'espérance mathématique d'une v.a.r. X est égale au moment d'ordre 1 de X. On le note E(X) et on a :

$$E(X) = \sum_{i=1}^{n} x_i p_i$$

**Proposition.** Soient X et Y deux v.a.r. sur un espace probabilisé fini $(\Omega, \mathcal{F}, p)$, λ et μ deux nombres réels. On a :

$$E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y).$$

En particulier, on a :

$$E(\lambda X + \mu) = \lambda E(X) + \mu.$$

### 1.3.3 Variance d'une v.a.r.

**Définition.** On appelle variance d'une v.a.r. X, le nombre réel positif ou nul noté V (X) défini par :

$$V(X) = \sum_{i=1}^{n} ( (x_i - E(X))^2 p_i = E(( X - E(X) )^2)$$

L'écart type de X est défini par :

$$\sigma(X) = \sqrt{V(X)}$$

**Proposition.** Soient X une v.a.r. sur un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), p)$, λ et μ deux nombres réels. On a :

$- V(X) = E(X^2) - (E(X))^2.$

$- V(\lambda X + \mu) = \lambda^2 V(X)$

$- \sigma(\lambda X + \mu) = |\lambda|\, \sigma(X).$

## 1.4 Variable aléatoire absolument continue

**Définition .** Une v.a.r. X est dite continue si sa fonction de répartition est continue à droite, donc continue. Il en résulte que pour tout nombre réel x, on a ;

$$P_X(\{x\}) = 0$$

**Remarque**

Soient deux nombres réels a et b tels que a $< b$. Alors, on a :

$$P_X([a,b]) = P_X(]a,b]) = P_X([a,b[) = P_X(]a,b[) = 0$$

### 1.4.1 Densité et fonction de répartition

**Définition** . On dit qu'une v.a.r. X admet une densité de probabilité $f$ lorsque sa fonction de répartition peut s'écrire sous la forme :

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

où $f$ est une fonction positive à valeurs réelles ayant un nombre fini de points de discontinuité et telle que :

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

*De manière équivalente, une variable aléatoire X est dite à densité, si sa fonction de répartition est continue sur R et de classe $C^1(R)$ sauf peut-être en un nombre fini de points.*

**Définition.** Une fonction réelle $f$ est une densité de probabilité si :

$-f \text{ est continue sur } R \text{ sauf peut} - \text{être en un nombre fini de points}$ ;

$- f(x) \geq 0 \text{ pour tout } x \in R$ ;

$- \int_{-\infty}^{+\infty} f(x)dx = 1.$

**Proposition :**

Soit X une variable aléatoire réelle admettant une densité de probabilité $f$ ; on a :

$-$ pour tout a $\in$ R, $p(\mathrm{X} = \mathrm{a}) = 0.$

$-$ pour tous réels a et b tels que a $< b$,

$$p(\mathrm{a} < X < b) = p(\mathrm{a} \leq \mathrm{X} < b) = p(\mathrm{a} < X \leq b) = p(\mathrm{a} \leq \mathrm{X} \leq \mathrm{b}) = \int_{\mathrm{a}}^{\mathrm{b}} \mathrm{f(t)}\ \mathrm{dt}.$$

− pour tout réel a,

$p(X < a) = p(X \leq a) = \int_{-\infty}^{a} f(t)dt.$

$-$ *pour tout réel b,*

$- p(X > b) = p(X \geq b) = \int_{b}^{+\infty} f(t)dt.$

### 1.4.2 Espérance mathématique

**Définition.** Soit $X$ une v.a.r. de densité de probabilité $f$. Si l'intégrale $\int_{-\infty}^{+\infty} xf(x)dx$ converge, X admet une espérance mathématique E(X) et

$$\mathrm{E(X)} = \int_{-\infty}^{+\infty} xf(x)dx.$$

**Proposition**

Soient X et $Y$ deux v.a.r. à densité de probabilité admettant chacune une espérance mathématique, $a$ et $b$ deux nombres réels. Alors la v.a.r $aX + bY$ admet une espérance mathématique et

$$E(aX + bY) = aE(X) + bE(Y).$$

En particulier :

$$E(aX + b) = aE(X) + b.$$

### 1.4.3 Variance et écart-type

**Théorème et définition :**

− Soit $X$ une v.a.r. à densité $f$. Si l'intégrale $\int_{-\infty}^{+\infty} x^2 f(x)dx$ converge, alors la v.a.r. $X^2$ admet une espérance mathématique et

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)dx\,.$$

$E(X^2)$ est appelée moment d'ordre deux de la v.a.r. $X$.

− Si $X$ admet une espérance mathématique et $E(X^2)$ existe, on appelle variance de $X$ le nombre réel

$$V(X) = E\left[\left(X - E(X)\right)^2\right].$$

$-$ Si $X$ admet une variance, on appelle écart-type de $X$ le nombre réel $\sigma(X) = \sqrt{V(X)}$

**Remarque**

Si $E(X)$ n'existe pas, $V\ (X)$ n'existe pas.

**Proposition**

$-$ Soit $X$ une v.a.r. de densité $f$, admettant une espérance mathématique $E(X)$.

Alors, $V\ (X)$ existe si et seulement si $\int_{-\infty}^{+\infty}\big(X - E(X)\big)^2 f(x)dx$ converge. On a alors :

$$V\ (X) = \int_{-\infty}^{+\infty}\big(X - E(X)\big)^2 f(x)dx.$$

$-$ Soit $X$ une v.a.r. de densité $f$, admettant une espérance mathématique $E(X)$. Alors, $V\ (X)$ existe si et seulement si $E(X^2)$ existe et on a :

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

**Définition**

$-$ Si $\sigma(X)\ =\ 1$, on dit que la v.a.r. $X$ est réduite.

$-$ Si $E(X)\ =\ 0$, on dit que la v.a.r. $X$ est centrée.

$-$ $\frac{X-E(X)}{\sigma(X)}$ est appelée v.a.r. centrée réduite associée à $X$.

### 1.4.4 Variables aléatoires à densité indépendantes

**Définition.** Soient $X$ et $Y$ deux v.a.r. à densité ; on dit que $X$ et $Y$ sont indépendantes lorsque pour tous $x$ et $y$ réels,

$$p\ [(X\ \leq\ x)\ \cap\ (Y\ \leq\ y)]\ =\ p\ (X\ \leq\ x)\, p\ (Y\ \leq\ y.)$$

**Proposition :**

Si $X$ et $Y$ sont deux v.a.r. indépendantes, à densité, admettant chacune une espérance mathématique, alors $XY$ est une v.a.r. à densité, admet une espérance mathématique et :

$$E(XY\ )\ =\ E(X)E(Y)$$

**Proposition :**

Si $X$ et $Y$ sont deux v.a.r. indépendantes, à densités respectives $f$ et $g$. $Z\ =\ X\ +\ Y$ admet pour densité la fonction $h$ définie pour tout $x\ \in\ R$ par :

$$h(x) = \int_{-\infty}^{+\infty} f(u)g(x-u)du = \int_{-\infty}^{+\infty} g(v)f(x-v)dv$$

**Proposition :**

Si $X$ et $Y$ sont deux v.a.r.à densité, indépendantes et admettant chacune une variance, alors $X + Y$ admet une variance et

$$V(X + Y) = V(X) + V(Y).$$

# 2 Propriétés générales

## 2.1 Variable aléatoire centrée réduite

**Définition**

− Une v.a.r. $X$ est dite centrée si son espérance mathématique est nulle.

− Une v.a.r. $X$ est dite réduite si son écart type est égal à 1.

− Une v.a.r. $X$ est dite centrée réduite si son espérance mathématique est nulle et son écart type est égal à 1.

**Théorème-Définition**

Soit $X$ une variable aléatoire d'espérance $E(X)$ et d'écart type $\sigma(X) \neq 0$. Alors la v.a.r. $Z = \frac{X-E(X))}{\sigma(X}$ est une variable aléatoire centrée réduite. C'est la variable aléatoire centrée réduite associée à la v.a.r. $X$.

**Fonction caractéristique**

Une autre fonction très utile est la fonction caractéristique $\varphi_X$ de la variable aléatoire réelle $X$.

**Définition.** Soit $X$ une variable aléatoire. La fonction caractéristique $\varphi_X$ de la variable aléatoire réelle $X$ est définie par :

$$\varphi_X(t) = E(e^{itX})$$

où $i$ est le nombre imaginaire vérifiant $i^2 = -1$.

## 2.2 Inégalités intéressante

### 2.2.1 Inégalité de Markov

**Proposition :**

Soit $X$ une variable aléatoire réelle sur un espace probabilisé $(\Omega,\ \mathcal{F},\ p)$, $g$ une fonction borélienne définie de $R$ dans $]0; +\infty[$ . Alors, si $E\ [g\ (X)]$ existe, pour tout $k\ \in\ R_+^*$ , on a

$$p\ (g\ (X)\ \geq\ k)\ \leq\ \frac{E\ [g\ (X)]}{k}$$

### 2.2.2 Inégalité de Chebyshev

**Proposition :**

Soit $X$ une v.a.r. sur un espace probabilisé $(\Omega,\ \mathcal{F},\ \mathrm{p})$. Alors, si $V\ (X)$ existe, pour tout réel $a$ strictement positif, on a :

$$p\ (|X\ -\ E(X)| \geq\ a) \leq \frac{V\ (X)}{a^2}\ .$$

**Remarque**. L'inégalité de **Chebyshev** ne requiert que la connaissance de l'espérance et de la variance de $X$. Elle donne une borne supérieure quant à la probabilité que la variable aléatoire réelle $X$ prenne des valeurs comprises à l'intérieur de l'intervalle $[E(X)\ -\ a;\ E(X)\ +\ a]$ . En pratique, l'inégalité de **Chebyshev** est très utile si l'on veut obtenir une estimation des ordres de grandeurs pour des intervalles de confiance symétriques autour de la moyenne lorsque l'on ne dispose que des valeurs de $E(X)$ et de $V\ (X)$.

# 3 Modes de convergences

## 3.1 Convergence en probabilités

**Définition.** On dit qu'une suite $(\mathrm{X_n})_\mathrm{n} \in \mathrm{N}$ converge en probabilité vers une v.a.r. $X$ et on note $X_n \xrightarrow{P} X$ si

$$\forall\ \varepsilon >\ 0,\ \lim_{n\to+\infty} P\ (|X_n\ -\ X|\ \geq\ \varepsilon)\ =\ 0.$$

**Exemple :** Soit $(\mathrm{X_n})_\mathrm{n} \in \mathrm{N}$ une suite de variables aléatoires réelles définies sur un même espace probabilisé. On suppose que chaque $X_n$ suit la loi de Poisson de paramètre $\frac{1}{n}$. Montrer que la suite $(\mathrm{X_n})_\mathrm{n} \in \mathrm{N}$ converge en probabilité vers la variable nulle.

**Théorème :** (Critère de convergence en probabilité)

Soit $(\mathrm{X_n})_\mathrm{n} \in \mathrm{N}$ une suite de variables aléatoires ; s'il existe $r\ >\ 0$ tel que la suite de terme général $E\ (\ |X_n|^r)$ tende vers 0, alors :

$$X_n \overset{P}{\rightarrow} 0$$

**Preuve**.

D'après le théorème de Markov, on a :

$$P\,(|X_n\,| > \;\varepsilon) \leq \frac{E\,(\,|X_n|^r)}{\varepsilon^r} \quad \rightarrow \quad 0.$$

## 3.2 Convergence en moyenne d'ordre α

**Définition.** Soit un espace probabilisé (Ω, $\mathcal{B}$, P), une suite $(X_n)_n \in N$ de v.a.r. et une v.a.r. $X$ définies sur (Ω, $\mathcal{B}$, P). On dit que la suite $(X_n)_n \in N$ converge en moyenne d'ordre α (avec α > 0) vers la v.a.r. $X$ et on note $X_n \overset{L^\alpha}{\rightarrow} X$ si :

$$\lim_{n\to+\infty} E(\,|X_n \;-\; X|^\alpha) \;=\; 0.$$

En pratique, α vaut souvent 2. On parle alors de convergence en **moyenne quadratique**.

## 3.3 Convergence presque sûre

**Définition.** Soit un espace probabilisé (Ω, $\mathcal{B}$, P). On dit qu'une suite $(X_n)_n \in N$ converge presque sûrement vers une v.a.r. $X$ et on note $X_n \overset{p.s}{\rightarrow} X$ si :

$$\exists N \;\subset\; \Omega, avec\; P(N) \;=\; 0, \omega \;\in\; \Omega \;-\; N \Rightarrow \;(X_N\,(\omega) \;\rightarrow\; X(\omega)).$$

Un ensemble $N$ tel que $P(N) \;=\; 0$ est dit de mesure nulle pour la probabilité $P$.

### 3.3.1 Critère de convergence presque sûre

Le critère suivant est souvent utilisé pour démontrer la convergence presque sûre.

**Propriété :**

Si pour tout λ positif, la série des probabilités $P(|X_n \;-\; X| \;>\; \lambda)$ est de somme finie, alors la suite $(X_n)_n \in N$ converge presque sûrement vers la v.a.r. $X$ . Cela équivaut à :

$$\left[\forall\lambda \;>\; 0, \sum_{n=1}^{+\infty} P(|X_n \;-\; X| \;>\; \lambda) \;<\; \infty\right] \Rightarrow \;(X_n \overset{p.s}{\rightarrow} X)$$

# 4 Lois des Grands Nombres (LGN)

## 4.1 Loi faible des grands nombres

**Théorème : (Première forme)**

Soit $(X_n)_n \in N$ une suite de v.a.r. discrètes définies sur un même espace probabilisé. On suppose que ces variables admettent toutes la même espérance mathématique m et la même variance $\sigma^2$, et sont deux à deux non corrélées. On pose ;

$$\overline{X_n} = \frac{X_1 + X_2 + X_3 + \ldots + X_n}{n}$$

Alors, pour tout $\varepsilon > 0$, on a ;

$$\lim_{n \to +\infty} P\left(|\overline{X_n} - m| \geq \varepsilon\right) = 0.$$

Autrement dit, la suite ( $\overline{X_n}$ ) converge en probabilité vers la variable aléatoire $m$.

**Preuve.**

Il suffit d'appliquer l'inégalité de Bienaymé-Chebyshev à la variable $\overline{X_n}$ en remarquant que

$$E(\overline{X_n}) = m \; et \; V\left(\overline{X_n}\right) = n\sigma^2.$$

**Remarque**

Les hypothèses du théorème précédent sont remplies en particulier lorsque toutes les variables $(X_n)_n$ sont mutuellement indépendantes, de même loi et admettent une espérance et une variance.

**Théorème de Bernoulli :**

Soit $(X_n)_n$ une suite de v.a.r. indépendantes suivant la loi de Bernouilli de paramètre p. On pose :

$$\overline{X_n} = \frac{X_1 + X_2 + X_3 + \ldots + X_n}{n}$$

Alors, pour tout $\varepsilon > 0$, on a

$$\lim_{n \to +\infty} P\left(|\overline{X_n} - p| \geq \varepsilon\right) = 0.$$

Autrement dit, la suite ( $\overline{X_n}$ ) converge en probabilité vers la variable aléatoire $p$.

**Preuve.** Il s'agit d'un cas particulier du théorème précédent.

**Théorème :( Deuxième forme)**

Soit $(\mathrm{X_n})_\mathrm{n}$ une suite de v.a.r. indépendantes, de même loi, et variance finie. Alors, la variable aléatoire réelle $\frac{1}{n}\sum_{i=1}^{n} X_i$ converge en moyenne quadratique vers l'espérance mathématique $m$ commune aux variables aléatoires. En d'autres termes, on a :

$$\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i \overset{L^2}{\rightarrow} m$$

## 4.2 Loi forte des grands nombres (LFGN):

**Théorème :**

Soit $(\mathrm{X_n})_\mathrm{n}$ une suite de v.a.r. indépendantes, de même loi, et avec une espérance mathématique finie $m$. Alors, la variable aléatoire réelle $\frac{1}{n}\sum_{i=1}^{n} X_i$ converge presque sûrement vers l'espérance mathématique $m$ commune aux variables aléatoires. En d'autres termes, on a :

$$\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i \overset{p^s}{\rightarrow} m$$

## 4.3 LFGN sans équidistribution

**Théorème : (Kolmogorow)**

Soit $(\mathrm{X_k})_\mathrm{k}$ une suite de variables aléatoires indépendantes vérifiant :

a) pour tout $k \geq 1, E\,({X_k}^2) < +\infty$ ;

b) il existe une suite $(a_k)$ de réels strictement positifs qui tend en croissant vers $+\infty$ telle que

$$\sum_{k=1}^{n} \frac{Var(X_k)}{{a_k}^2} < +\infty.$$

Alors $(S_n - E(S_n))/a_n$ converge presque sûrement vers 0. Si de plus $\mathrm{a_n^{-1}}\,\mathrm{E}(S_n) \rightarrow m$, $S_n/a_n$ converge p.s. vers $m$.

**Théorème : (**Inégalité maximale de Kolmogorov**)**

Si les $X_k$ sont indépendantes, centrées et de carrés intégrables

$$\forall\, t > 0, P(\, max_{1\leq k\leq n}\ |S_k|) \geq t \leq \frac{1}{t^2} Var\,(S_n)$$

# 5 Applications

## 5.1 La méthode de Monte Carlo

La loi des grands nombres fournit une méthode de calcul approché d'intégrales, intéressante lorsque la fonction à intégrer est très irrégulière ou lorsque la dimension de l'espace est élevée. Supposons que l'on veuille effectuer un calcul approché de

$$I = \int_{[0,1]^d} f(x)dx.$$

où $f$ est Lebesgue intégrable sur $[0,1]^d$. Soit $(U_i)i \geq 1$, une suite de variables aléatoires indépendantes de même loi uniforme sur [0,1]. On déduit facilement de la LFGN que :

$$\boldsymbol{\frac{1}{n}\sum_{i=1}^{n} f(U_{(k-1)d+1},\ U_{(k-1)d+2,},\ldots,U_{kd}) \overset{p.s}{\rightarrow} E\big(f(U_1,U_2,\ldots,U_d)\big)} = \int_{[0,1]^d} f(x)dx.$$

## 5.2 Estimation de paramètres

La LFGN permet de définir des estimateurs convergents de paramètres d'une loi inconnue µ (ou partiellement inconnue, par exemple on sait qu'il s'agit d'une loi de Poisson paramètre α dont on ignore la valeur).

Pour cela on utilise une suite d'observations indépendantes $X_1(\omega),\ldots,X_n(\omega)$ où les $X_i$ sont i.i.d. de même loi µ. On souhaite estimer un paramètre $\theta$ de la forme $\theta = \int_R H\,d\mu$. L'idée est de remplacer la mesure déterministe mais inconnue µ par la mesure aléatoire $\mu_n$ calculable à partir des observations :

$$\mu_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}$$

Cette mesure est appelée mesure empirique. La fonction de répartition empirique est simplement sa fonction de répartition :

$$F_n(x) = \mu_n(]-\infty, x]).$$

On propose d'estimer $\theta$ par

$$\hat{\theta}_n = \int_R H\,d\mu_n = \frac{1}{n}\sum_{i=1}^{n} H(X_i)$$

La définition de $\theta$ suppose implicitement que $H$ est µ intégrable. Cette intégrabilité s'écrit encore $E(|H(X_1)|) < +\infty$. Ainsi par la loi forte des grands nombres,

$$\hat{\theta}_n \overset{p.s}{\rightarrow} E(\mathrm{H}(\mathrm{X}_1)) = \theta = \int_R H \, d\mu.$$

On dit que $\hat{\theta}_n$ est un estimateur fortement consistant de $\theta$. Il est aussi sans biais puisque $E(\hat{\theta}_n) = \theta$.

Cette méthode permet notamment d'estimer les moments de μ : en prenant $H(x) = x^r$ , $\theta = E(X_1^r) = \int_R x^r \mu(dx)$. Le cas $r = 1$ revêt une importance particulière. L'estimateur $\hat{\theta}_n$ est alors simplement la moyenne empirique

$$\overline{\boldsymbol{X}}_{\boldsymbol{n}} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

## Conclusion

Les lois des grands nombres jouent un rôle fondamental dans la théorie des probabilités en offrant un cadre qui relie le comportement empirique des suites de variables aléatoires à leurs propriétés théoriques. La distinction entre la loi faible et la loi forte des grands nombres permet de comprendre comment la moyenne empirique converge vers l'espérance mathématique à différentes échelles de rigueur. Leurs applications variées, comme la méthode de Monte Carlo et l'estimation de paramètres, illustrent l'importance de ces résultats dans des domaines aussi divers que la physique et les statistiques.

Ainsi, que ce soit dans l'analyse des processus stochastiques ou dans la simulation numérique, les lois des grands nombres offrent des outils puissants pour appréhender et modéliser des phénomènes aléatoires, consolidant le lien entre le hasard et le déterminisme.

## Bibliographie

OGOUYANDJOU, C. (2022). *COURS DE THEORIE DES PROBABILIT´ES.* IMSP,Benin.

Stout, W. F. (1974). *Almost sure convergence.* Royaume-Uni: Academic Press.

Suquet, C. (2003). *Lois des grands nombres.* Lille: Universit´e des Sciences et Technologies de Lille.

Toulouse, P. S. (1999). *Thèmes de Probabilités et Statistique.* Dunod.