AIX-MARSEILLE UNIVERSITÉ

Probabilités M1

Author :
Rémi RHODES

Master 1 :
LUMINY

Table des matières

Chapitre 1

Introduction

Objectifs :
— Variables aléatoires
— Fonction génératrice, fonction caractéristique.
— Lois des grands nombres.
— Théorème central limite.
Outils :
— espaces mesurables
— mesure de Lebesgue sur $\mathbb{R}^n$
— integration de Lebesgue (mesures qcq)
— inégalité de Jensen
— espaces L^p
— Fatou, CV monotone et dominée
— derivation sous le signe somme
— classe monotone
— Fubini

Chapitre 2

Variables aléatoires

2.1 Espace probabilisé et variables aléatoires

Définition 2.1.1. *Un espace probabilisé* $(\Omega, \mathcal{T}, \mathbb{P})$ *est la donnée*
- *d'un ensemble* Ω *appelé univers, dont les éléments sont appelés résultats,*
- *d'une tribu* $\mathcal{T}$ *de parties de* Ω*, dont les éléments sont appelés évènements,*
- *d'une mesure* $\mathbb{P}$ *définie sur la tribu* $\mathcal{T}$*, qui satisfait* $\mathbb{P}(\Omega) = 1$.

Une propriété des points $\omega \in \Omega$ sera dite "vraie presque sûrement" (en abrégé p.s.) si l'ensemble des ω pour lesquels elle est fausse est inclus dans un ensemble A vérifiant $\mathbb{P}(A) = 0$.

2.2 Variables aléatoires

Définition 2.2.1. *Soit* $(\Omega, \mathcal{T}, \mathbb{P})$ *un espace probabilisé et* $(F, \mathcal{F})$ *un espace mesurable. Une variable aléatoire* $X : \Omega \to F$ *est une application mesurable, i.e.* $\forall A \in \mathcal{F}$, $X^{-1}(A) \in \mathcal{T}$.

Si $(F, \mathcal{F}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, où $\mathcal{B}(\mathbb{R})$ est la tribu des boréliens, on parle de variable aléatoire réelle et si $(F, \mathcal{F}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, on parle de vecteur aléatoire.

Si F est un ensemble dénombrable et $\mathcal{F}$ la tribu de toutes ses parties, on parle de variable aléatoire discrète.

Définition 2.2.2 (Loi d'une variable aléatoire)**.** *Soit* $(\Omega, \mathcal{T}, \mathbb{P})$ *un espace probabilisé,* $(F, \mathcal{F})$ *un espace mesurable et* $X : \Omega \to F$ *une variable aléatoire. La loi de* X *est la mesure de probabilité* $\mathbb{P}_X$ *définie sur* $(F, \mathcal{F})$ *par*

$$\forall A \in \mathcal{F}, \quad \mathbb{P}_X(A) = \mathbb{P}(X^{-1}(A)).$$

Cette définition est générale. Dans ce cours, on se restreindra au cas où F est une sous-ensemble de $\mathbb{R}^d$ muni de la tribu induite $\mathcal{F}$ par $\mathcal{B}(\mathbb{R}^d)$ sur F.

Exercice 1. *Soit* $(\Omega, \mathcal{T}, \mathbb{P})$ *un espace probabilisé et* $(F, \mathcal{F})$ *un espace mesurable. Soit* $X : \Omega \to F$ *une application.*

1. *Montrer que* $\{A \in \mathcal{F}; X^{-1}(A) \in \mathcal{T}$ *est une tribu*$\}$.

2. *En déduire : pour que X soit une v.a. il suffit que $X^{-1}(C) \in \mathcal{T}$ pour C appartenant à une famille d'ensemble engendrant la tribu $\mathcal{F}$.*
3. *En déduire que X est une v.a. réelle ssi $X^{-1}(]-\infty, x])$ pour tou $x \in \mathbb{R}$.*

Exercice 2. *On suppose $(\Omega, \mathcal{T}) = (F, \mathcal{F}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Donner la loi de X dans les cas suivants*

1. $\mathbb{P}(dx) = \mathbf{1}_{[0,1]}\, dx$ *et* $X(x) = \mathbf{1}_{[1-p,\infty[}(x)$ *où* $p \in]0,1[$.
2. $\mathbb{P}(dx) = \mathbf{1}_{[0,1]}\, dx$ *et* $X(x) = -\ln x$
3. $\mathbb{P}(dx) = \mathbf{1}_{[0,+\infty[}\lambda e^{-\lambda x}\, dx$ *(avec $\lambda > 0$) et $X(x) = [x]$ où $[x]$ désigne la partie entière de x.*

Théorème 2.2.3 (Formule de transfert)**.** *Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé et $X : \Omega \to \mathbb{R}^d$ un vecteur aléatoire. Pour toute fonction borélienne positive sur $\mathbb{R}^d$ on a*

$$\int_\Omega f(X(\omega))\mathbb{P}(d\omega) = \int_{\mathbb{R}^d} f(x)\mathbb{P}_X(dx).$$

On note $\mathbb{E}[f(X)]$ cette valeur commune.

Démonstration. La formule est immédiate lorsque f est l'indicatrice de $A \in \mathcal{B}(\mathbb{R}^d)$, donc aussi pour une combinaison linéaire de telles indicatrices. En utilisant le fait que toute fonction mesurable est limite croissante de fonctions étagées, le théorème de convergence monotone permet de conclure. □

Définition 2.2.4. *Le vecteur aléatoire $X : \Omega \to \mathbb{R}^d$ est dit intégrable ssi $\mathbb{E}[|X|] < \infty$. Dans ce cas, on appelle espérance de X la quantité $\mathbb{E}[X]$.*
Le vecteur aléatoire $X : \Omega \to \mathbb{R}^d$ est dit de puissance p-ième ($p > 0$) intégrable (ou admet des moments d'odre p) ssi $\mathbb{E}[|X|^p] < \infty$.
Dans le cas $p = 2$ et si $d = 1$ on peut définir alors la variance de X par $\mathbb{E}[X^2] - \mathbb{E}[X]^2$.
Dans le cas $p = 2$ et si $d > 1$ on peut définir alors la matrice de covariance de X par

$$\mathbb{E}[XX^T] - \mathbb{E}[X]\mathbb{E}[X^T].$$

Exercice 3. *On dit que la variable aléatoire réelle X suit une loi de Poisson de paramètre $\lambda > 0$ ssi*

$$\mathbb{P}_X = \sum_{k\in\mathbb{N}} e^{-\lambda}\frac{\lambda^n}{n!}\delta_k.$$

Montrer que X admet des moments de tout ordre. Calculer l'espérance et la variance de X.

Exercice 4. *Soit X une variable aléatoire telle que $0 \leq X \leq 1$. Montrer que $\mathrm{Var}(X) \leq 1/4$. A quelle condition a-t-on égalité ?*

Définition 2.2.5. *On dit que le vecteur aléatoire $X : \Omega \to \mathbb{R}^d$ est un vecteur à densité ssi il existe une fonction positive borélienne f telle que*

$$\mathbb{P}_X(dx) = f(x)\, dx.$$

On a les exemples importants suivants :

— Loi de Gauss réelle : $X : \Omega \to \mathbb{R}$ suit une loi normale (ou gaussienne) de paramètres $m \in \mathbb{R}, \sigma > 0$ ssi $\mathbb{P}_X(dx) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} dx$.
— Loi exponentielle : $X : \Omega \to \mathbb{R}$ suit une loi exponentielle de paramètre $\lambda > 0$ ssi $\mathbb{P}_X(dx) = \lambda e^{-\lambda x} \mathbf{1}_{x>0} dx$.
— Loi de Cauchy : $X : \Omega \to \mathbb{R}$ suit une loi de Cauchy de paramètres $a > 0, x_0 \in \mathbb{R}$ ssi $\mathbb{P}_X(dx) = \frac{a}{\pi(a^2+(x-x_0)^2)} dx$.
— Loi uniforme : $X : \Omega \to \mathbb{R}$ suit une loi uniforme sur $[a,b]$ ($a < b$) ssi $\mathbb{P}_X(dx) = \frac{1}{b-a} \mathbf{1}_{x\in[a,b]} dx$.

Proposition 2.2.6. *Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé et $X : \Omega \to \mathbb{R}^d$ un vecteur aléatoire. La loi de X est caractérisé par l'une des propriétés suivantes*

1. *pour toute fonction f borélienne positive $\int_\Omega f(X(\omega))\mathbb{P}(d\omega) = \int_{\mathbb{R}^d} f(x)\mathbb{P}_X(dx)$*
2. *pour toute fonction f continue bornée $\int_\Omega f(X(\omega))\mathbb{P}(d\omega) = \int_{\mathbb{R}^d} f(x)\mathbb{P}_X(dx)$*
3. *pour toute fonction f continue à support compact $\int_\Omega f(X(\omega))\mathbb{P}(d\omega) = \int_{\mathbb{R}^d} f(x)\mathbb{P}_X(dx)$*
4. *(si $d = 1$) pour tout $x \in \mathbb{R}$, $\mathbb{P}(X \leq x) = P_X(]-\infty, x])$.*

Démonstration. Nous donnons la preuve dans le cas $d = 1$.

On a 1) ⇒ 2) car les fonctions continues sont boréliennes et en écrivant $f = f_+ - f_-$ où f_+, f_- sont les parties positives/négatives de f.
2) ⇒ 3) évident.
3) ⇒ 4). Soient $a < b$. Pour $n \in \mathbb{N}^*$ on définit la fonction f_n continue à support compact sur $\mathbb{R}$ par $f(x) = 0$ si $x > b+1/n$ ou $x < a$, $f(x) = 1$ sur $[a+1/n, b]$ et affine sur $[a, a+1/n]$ et $[b, b+1/n]$. Clairement $f_n(x)$ converge simplement vers $\mathbf{1}_{]a,b]}$ tout en étant dominée par la fonction 1. Donc le théorème de convergence dominée implique

$$\lim_{n\to\infty} \mathbb{E}[f_n(X)] = \mathbb{P}(a < X \leq b) \quad \text{et} \quad \lim_{n\to\infty} \int_{\mathbb{R}} f_n(x)\mathbb{P}_X(dx) = \int_{\mathbb{R}} \mathbf{1}_{]a,b]}\mathbb{P}_X(dx).$$

Comme $\mathbb{E}[f_n(X)] = \int_{\mathbb{R}} f_n(x)\mathbb{P}_X(dx)$ ceci entraîne $\mathbb{P}(a < X \leq b) = \int_{\mathbb{R}} \mathbf{1}_{]a,b]}\mathbb{P}_X(dx)$.

Il faut ensuite montrer qu'"il n'y a pas de masse à l'infini". Pour cela, on observe que le théorème de convergence dominée entraîne que

$$\lim_{R\to\infty} \mathbb{P}(|X| \geq R) + \int_{\mathbb{R}} \mathbf{1}_{]-R,R[^c}\mathbb{P}_X(dx) = 0.$$

Pour tout $x > 0$, et pour $R > |x|$ on écrit alors

$$\begin{aligned} |\mathbb{P}(X \leq x) - P_X(]-\infty, x])| \leq & \mathbb{P}(|X| > R) + \int_{\mathbb{R}} \mathbf{1}_{]-R,R[^c}\mathbb{P}_X(dx) + |\mathbb{P}(R < X \leq x) - P_X(]-R, x])| \\ \leq & \mathbb{P}(|X| > R) + \int_{\mathbb{R}} \mathbf{1}_{]-R,R[^c}\mathbb{P}_X(dx). \end{aligned}$$

Cette dernière quantité peut être rendue arbitrairement petite en prenant R assez grand.
4) ⇒ 1) par le théorème de classe monotone, $\mathbb{P} \circ X^{-1} = \mathbb{P}_X$. □

Exercice 5. *Soit $a > 0$. Rappelons que l'on définit la fonction Γ*

$$\Gamma(a) = \int_0^\infty t^{a-1}e^{-t}\, dt.$$

1. *Montrer que pour $\lambda > 0$ et $a > 0$ la fonction*
$$\frac{1}{\Gamma(a)}\lambda^a e^{-\lambda t} t^{a-1} \mathbf{1}_{x>0}$$
définit une densité de probabilité sur $\mathbb{R}$, notée $\Gamma(\lambda, a)$.
2. *Si X suit une loi normale $\mathcal{N}(0,1)$. Montrer que X^2 suit une loi $\Gamma(\frac{1}{2}, \frac{1}{2})$ et retrouver le fait que $\Gamma(1/2) = \sqrt{\pi}$.*

2.3 Indépendance

Définition 2.3.1. *Indépendance d'évènements Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé. Deux évènements $A, B \in \mathcal{T}$ sont dits indépendants si*
$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$
Soit $(A_i)_{i \in I}$ une famille d'évènements. Ces évènements sont dits indépendants dans leur ensemble si pour tout sous-ensemble fini $S \subset I$
$$\mathbb{P}\Big(\bigcap_{i \in S} A_i\Big) = \prod_{i \in S} \mathbb{P}(A_i).$$

Définition 2.3.2. *Indépendance de variables aléatoires Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé et $(F, \mathcal{F})$ un espace mesuré. Deux variables aléatoires $X, Y : \Omega \to F$ sont dites indépendantes si pour tous évènements $A, B \in \mathcal{F}$*
$$\mathbb{P}(\{X \in A\} \cap \{Y \in B\}) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B).$$
Soit $X_i : \Omega \to F$ $(i \in I)$ une famille de variables aléatoires. Elles sont dits indépendantes entre elles si pour tout sous-ensemble fini $S \subset I$ et toute famille $(A_i)_{i \in S}$ d'évènements de $\mathcal{F}$
$$\mathbb{P}\Big(\bigcap_{i \in S} \{X_i \in A_i\}\Big) = \prod_{i \in S} \mathbb{P}(X_i \in A_i).$$

Proposition 2.3.3. *Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé et $(F, \mathcal{F})$ un espace mesuré. Soit $X_i : \Omega \to F$ $(i \in I)$ une famille de variables aléatoires indépendantes entre elles. Alors pour tout sous-ensemble fini $S \subset I$ et toute famille $f_i : F \to \mathbb{R}$ $(i \in S)$ de fonctions mesurables telles que $f_i(X_i)$ est intégrable on a*
$$\mathbb{E}[\prod_{i \in S} f_i(X_i)] = \prod_{i \in S} \mathbb{E}[f_i(X_i)].$$

A noter que la réciproque de cette proposition est évidente en prenant $f_i = \mathbf{1}_{A_i}$.

Démonstration. Si les f_i sont des fonctions indicatrices $f_i = \mathbf{1}_{A_i}$ alors
$$\begin{aligned}\mathbb{E}[\prod_{i \in S} f_i(X_i)] &= \mathbb{E}[\prod_{i \in S} \mathbf{1}_{X_i \in A_i}] \\ &= \prod_{i \in S} \mathbb{P}(X_i \in A_i) \\ &= \prod_{i \in S} \mathbb{E}[f_i(X_i)].\end{aligned}$$

On procède alors comme pour la formule de transfert : on vérifie la formule pour les fonctions étagées par linéarité, puis pour $f_i \geq 0$ $(i \in S)$ en les approchant par des fonctions étagées de manière croissante et enfin pour $(f_i)_{i \in S}$ intégrables en les écrivant comme des différences de fonctions positives intégrables. □

Proposition 2.3.4. *Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé et $(F, \mathcal{F})$ un espace mesuré. Soit $X_i : \Omega \to F$ $(i \in I)$ une famille de variables aléatoires indépendantes entre elles. Alors pour tout sous-ensemble fini $S = \{i_1, \ldots, i_n\} \subset I$ on a*

$$\mathbb{P}_{(X_{i_1}, \ldots, X_{i_n})}(dx_1, \ldots, dx_n) = \mathbb{P}_{X_{i_1}}(dx_1) \otimes \cdots \otimes \mathbb{P}_{X_{i_n}}(dx_n).$$

Démonstration. Soit

$$\mathcal{C} = \{A \in \mathcal{F}^{\otimes S}; \mathbb{P}_{(X_{i_1}, \ldots, X_{i_n})}(A) = \int_{F^S} \mathbf{1}_A(x_1, \ldots, x_n) \mathbb{P}_{X_{i_1}}(dx_1) \otimes \cdots \otimes \mathbb{P}_{X_{i_n}}(dx_n)\}.$$

$\mathcal{C}$ est une classe monotone qui contient la classe des pavés $A_1 \times \ldots A_n \in \mathcal{F}^{\otimes S}$ qui est stable par intersections finies, et dont la tribu engendrée coincide avec $\mathcal{F}^{\otimes S}$. □

Exercice 6. ***Loi Gamma.*** *Rappel : la loi $\gamma(p, \theta)$ $(p, \theta > 0)$ est une loi de probabilité à densité sur $\mathbb{R}_+$ donnée par*

$$\mathbf{1}_{x>0} \frac{\theta^p}{\Gamma(p)} e^{-\theta x} x^{p-1},$$

où Γ est la fonction Gamma usuelle $\Gamma(p) = \int_0^\infty x^{p-1} e^{-x}\, dx$. Cette fonction vérifie les propriétés suivantes pour tout $p > 0$: $p\Gamma(p) = \Gamma(p+1)$, $\Gamma(p) = (p-1)!$ si $p \in \mathbb{N}$ et $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

On rappelle également la formule de changement de variables dans $\mathbb{R}^2$. Si $\psi : U \subset \mathbb{R}^2 \to \psi(U) \subset \mathbb{R}^2$ est un C^1-difféomorphisme alors pour toute fonction continue $f : \psi(U) \to \mathbb{R}$

$$\int_U f(\psi(x, y))\, |J\psi(x, y)|\, dxdy = \int_{\psi(U)} f(x, y)\, dxdy$$

où $J\psi$ est la matrice jacobienne des dérivées partielles de ψ.

Soient X, Y deux variables aléatoires indépendantes suivant respectivement des lois $\gamma(p, \theta)$ et $\gamma(q, \theta)$ (même paramètre θ).

1. *Trouver la densité de la loi jointe du couple $X + Y$ et $\frac{X}{X+Y}$ et en déduire l'indépendance de ces variables aléatoires.*
   Indication : Penser à faire le changement de variables suivant : $x + y = u$, $x/(x + y) = v$.
2. *Quelles sont les lois marginales de $X + Y$ et $\frac{X}{X+Y}$?*

Exercice 7. *Soient $X_1, \ldots, X_d$ des variables aléatoires gaussiennes indépendantes de loi $\mathcal{N}(0, 1)$. La loi du chi-deux à d degrés de liberté, notée χ^2_d, est définie comme la loi de la variable $Y = X_1^2 + \cdots + X_d^2$.*

1. *Montrer que χ^2_2 est la loi exponentielle de paramètre $1/2$.*
2. *Plus généralement, montrer que la densité de la loi χ^2_d est donnée par $\frac{2^{-\frac{d}{2}}}{\Gamma(d/2)} x^{\frac{d}{2}-1} e^{-\frac{x}{2}} \perp\!\!\!\perp_{x>0}$ et vérifier que c'est aussi la ldensité de la loi Gamma $\Gamma(\frac{d}{2}, \frac{1}{2})$.*

Exercice 8. *Soit $X_1, \ldots, X_n$ des lois exponentielles indépendantes de paramètre λ. Calculer la loi de $X_1 + \cdots + X_n$*

Exercice 9. *Soit X, Y deux variables aléatoires réelles à densité et indépendantes. Calculer la probabilité $\mathbb{P}(X = Y)$.*

2.4 Inégalités

Théorème 2.4.1. *Inégalité de Markov Soit* $(\Omega, \mathcal{T}, \mathbb{P})$ *un espace probabilisé et* $X : \Omega \to \mathbb{R}$ *une variable aléatoire positive. Alors pour tout* $r > 0$

$$\mathbb{P}(X > r) \leq \frac{1}{r}\mathbb{E}[X].$$

Démonstration. On a $r\mathbf{1}_{r<X} \leq X$. Et on obtient le résultat en prenant l'espérance. □

Théorème 2.4.2. *Inégalité de Bienaymé-Tchebichev Soit* $(\Omega, \mathcal{T}, \mathbb{P})$ *un espace probabilisé et* $X : \Omega \to \mathbb{R}$ *une variable aléatoire de carré intégrable. Alors pour tout* $r > 0$

$$\mathbb{P}(|X - \mathbb{E}[X]| > r) \leq \frac{1}{r^2}\mathrm{Var}(X).$$

Démonstration. On a

$$\mathbb{P}(|X - \mathbb{E}[X]| > r) = \mathbb{P}(|X - \mathbb{E}[X]|^2 > r^2).$$

On conclut en appliquant l'inégalité de Markov et en remarquant que $\mathbb{E}[|X - \mathbb{E}[X]|^2] = \mathrm{Var}(X)$. □

Exercice 10. *Pour étudier les particules émises par une substance radioactive, on dispose d'un détecteur. On note* X *la variable aléatoire représentant le nombre de particules qui atteignent le détecteur pendant un intervalle de temps* Δt*. Le nombre maximal de particules que le détecteur peut compter pendant un intervalle de temps* Δt *est de* 10^3*.*

1. *On suppose que* X *a pour espérance* 10^2*. Donner une majoration de la probabilité que* X *dépasse* 10^3*.*
2. *On suppose de plus que* X *a une variance égale à* 10^2*. Donner une majoration plus fine de la probabilité que* X *dépasse* 10^3*.*
3. *On suppose maintenant que* X *suit une loi de Poisson de paramètre* m*. Soit* $a > 1$*.*
   (a) *Montrer que pour tout* $\theta > 0$, $\mathbb{P}(X > am) \leq \mathbb{E}[e^{\theta(X-am)}]$*.*
   (b) *En déduire que* $\mathbb{P}(X > am) \leq e^{-m(1-a+a\ln(a))}$
   (c) *Donner un majorant de la probabilité que* X *dépasse* 10^3 *si* X *suit une loi de Poisson d'espérance* 10^2*.*

Chapitre 3

Convergence de suites de variables aléatoires

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé. Pour $p \in [1, +\infty[$, la norme L^p de la variable aléatoire $X : \Omega \to \mathbb{R}$ est définie par

$$\|X\|_p := \Big(\int |X|^p \, d\mathbb{P} \Big)^{1/p}.$$

La norme L^∞ de Y est définie par

$$\|X\|_\infty := \inf\{C > 0 \,|\, \exists \Omega' \text{ tel que } \mathbb{P}(\Omega') = 1 \text{ et } |X(\omega)| \leq C \text{ pour tout } \omega' \in \Omega'\}$$

Définition 3.0.1. *Soient $(X_n)_n$ et X des variables aléatoires définies sur $(\Omega, \mathcal{T}, \mathbb{P})$ et $p \in [1, +\infty]$.*

— *La suite $(X_n)_n$ converge vers X dans L^p ssi*

$$\|X_n - X\|_p \to 0, \quad n \to \infty.$$

— *La suite $(X_n)_n$ converge vers X en probabilité ssi*

$$\forall \epsilon > 0, \quad \mathbb{P}(|X_n - X| > \epsilon) \to 0, \quad n \to \infty.$$

— *La suite $(X_n)_n$ converge vers X presque sûrement ssi*

$$\exists \Omega' \subset \Omega, \mathbb{P}(\Omega') = 1, \forall \omega \in \Omega', X_n(\omega) \to X(\omega), \quad \textit{quand } n \to \infty.$$

— *La suite $(X_n)_n$ converge en loi vers X ssi pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ continue et bornée*

$$\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)], \quad \textit{quand } n \to \infty.$$

3.1 Relations entre convergences

Proposition 3.1.1. *Soient $1 \leq p \leq q \leq \infty$. On a les implications*

$$\textit{Conv } L^\infty \Rightarrow \textit{ Conv } L^q \Rightarrow \textit{ Conv } L^p \Rightarrow \textit{ Conv } L^1 \Rightarrow \textit{ Conv en proba} \Rightarrow \textit{Conv en loi}$$

$$\textit{Conv } L^\infty \Rightarrow \textit{ Conv p.s.} \Rightarrow \textit{ Conv en proba} \Rightarrow \textit{ Conv p.s. d'une sous-suite}$$

Proof. 1) Conv $Lp \Rightarrow$ Conv L^q. Cela résulte de l'inégalité de Jensen : pour tout $1 \leq p \leq q$ et $f \geq 0$

$$\int X^p \, d\mathbb{P} = \int (X^q)^{p/q} \, d\mathbb{P} \leq \Big(\int X^q \, d\mathbb{P} \Big)^{p/q}.$$

Si $q = \infty$, cela résulte juste de la positivité de l'espérance et du fait que p.s. $X^p \leq \|X\|_\infty^p$.

2) Conv $L^1 \Rightarrow$ Conv en proba. En utlilisant Markov

$$\mathbb{P}(|X_n - X| > \epsilon) \leq \epsilon^{-1}\mathbb{E}|X_n - X| = \epsilon^{-1}\|X_n - X\|_1 \to 0, \quad n \to \infty.$$

3) Conv $L^\infty \Rightarrow$ Conv p.s. Il existe $\Omega' \subset \Omega$ tel que $\mathbb{P}(\Omega') = 1$ et

$$\sup_{\omega \in \Omega'} |X_n(\omega) - X(\omega)| \to 0, \quad n \to \infty.$$

On en déduit la convergence uniforme (donc simple) sur Ω'.

4) Conv en proba $\Rightarrow$ Conv p.s d'une sous-suite : comme $\forall k \in \mathbb{N}$, $\mathbb{P}(|X_n - X| > 1/k) \to 0$ quand $n \to \infty$ on en déduit que l'on peut trouver $(n_k)_k$ sous-suite telle que

$$\mathbb{P}(|X_{n_k} - X| > 1/k) \leq 2^{-k}.$$

Posons $A_k := \{|X_{n_k} - X| > 1/k\}$ et $\Omega' := \liminf A_k^c$. On a convergence simple de la suite $(X_{n_k})_k$ vers X sur Ω'. De plus par convergence monotone

$$\mathbb{P}((\Omega')^c) = \lim_{k_0 \to \infty} \mathbb{P}\Big(\bigcup_{k \geq k_0} A_k \Big) \leq \lim_{k_0 \to \infty} \sum_{k \geq k_0} \mathbb{P}\big(A_k\big) \leq \lim_{k_0 \to \infty} \sum_{k \geq k_0} 2^{-k} = 0.$$

5) Conv en proba $\Rightarrow$ Conv en loi. Soit f continue et bornée. Comme $(X_n)_n$ cv en proba, de toute suite extraite, on peut extraire une nouvelle sous-suite $(X_{n_k})_k$ convergeant p.s. vers X. Le théorème de cv dominée et le fait que f soit continue assure que

$$\lim_{k \to \infty} \int f(X_{n_k}) \, d\mathbb{P} = \int f(X) \, d\mathbb{P}.$$

Ceci assure que la suite $(\int f(X_n) \, d\mathbb{P})_n$ converge vers $\int f(X) \, d\mathbb{P}$. □

3.2 Tension des variables aléatoires

On va énoncer un résultat de compacité pour la convergence en loi. Tout d'abord

Définition 3.2.1. *Une suite de v.a. $(X_n)_n$ est dite tendue si et seulement si pour tout $\epsilon > 0$ il existe $M > 0$ tel que*

$$\sup_n \mathbb{P}(|X_n| > M) \leq \epsilon.$$

La notion de tension est très utile pour étudier la convergence en loi. Un premier exemple d'application sont les équivalences suivantes

Proposition 3.2.2. *Se valent :*

1. *$(X_n)_n$ converge en loi vers X*
2. *pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ continue et tendant vers 0 à l'infini*
$$\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)], \quad \textit{quand } n \to \infty.$$
3. *pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ continue à support compact*
$$\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)], \quad \textit{quand } n \to \infty.$$
4. *pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ infiniment différentiable à support compact*
$$\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)], \quad \textit{quand } n \to \infty.$$

Démonstration. Il est clair que $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4$. Montrons $4 \Rightarrow 1$. La première étape est de montrer que la suite $(X_n)_n$ est tendue. Soit $\epsilon > 0$. On peut trouver $R > 0$ tel que
$$\mathbb{P}(|X| > R-1) \leq \epsilon/2.$$
Soit φ une fonction de classe C^∞ telle que $\mathbf{1}_{|x|\leq R} \geq \varphi(x) \geq \mathbf{1}_{|x|\leq R-1}$. Comme φ est à support compact : $\exists N \in \mathbb{N}, \forall n \geq N$,
$$|\mathbb{E}[\varphi(X_n)] - \mathbb{E}[\varphi(X)]| \leq \epsilon/2.$$
Pour $n \geq N$, comme $\mathbf{1}_{|x|\leq R} \geq \varphi(x) \geq \mathbf{1}_{|x|\leq R-1}$
$$\mathbb{P}(|X_n| \leq R) \geq \mathbb{E}[\varphi(X_n)] \geq \mathbb{E}[\varphi(X)] - \epsilon/2 \geq \mathbb{P}(|X| > R-1) - \epsilon/2 \geq 1-\epsilon.$$
Soit R' tel que pour $i = 1, \ldots, N-1$, on ait $\mathbb{P}(|X_i| > R) \leq \epsilon$. Finalement, pour $\bar{R} = \max(R, R')$ on a
$$\max_n \mathbb{P}\big(|X_n| > \bar{R}\big) \leq \epsilon.$$

Soit f une fonction C^∞ à support compact et $\epsilon > 0$. Comme la suite $(X_n)_n$ est tendue, il existe $R > 0$ tel que $\sup_n \mathbb{P}(|X_n| > R) \leq \epsilon/(2\|f\|_\infty)$. Soit φ une fonction C^∞ positive sur $\mathbb{R}$, valant 0 sur $[-R, R]$ et 1 sur l'ensemble $[-R-1, R+1]^c$.

Alors on peut écrire
$$\mathbb{E}[f(X_n)] = \mathbb{E}[f(X_n)\varphi(X_n)] + \mathbb{E}[f(X_n)(1-\varphi(X_n))].$$
La fonction $(1-\varphi)$ étant à support compact, on a la convergence $\mathbb{E}[f(X_n)(1-\varphi(X_n))] \to \mathbb{E}[f(X)(1-\varphi(X))]$ lorsque $n \to \infty$. Donc pour $n \geq N$
$$\Big|\mathbb{E}[f(X_n)(1-\varphi(X_n))] - \mathbb{E}[f(X)(1-\varphi(X))]\Big| \leq \epsilon/2.$$
De plus
$$|\mathbb{E}[f(X_n)\varphi(X_n)]| \leq \|f\|_\infty \mathbb{P}(|X_n| > R) \leq \epsilon/2.$$
Ceci complète la preuve. □

Exercice 11. *Soit $(X_n)_n$ une suite de v.a. telle que pour tout n, $X_n \sim$ Bin(n, p_n) où $np_n \to \lambda \in]0, +\infty[$ quand $n \to \infty$. Montrer que la suite $(X_n)_n$ converge en loi vers une loi de Poisson de paramètre λ.*

Exercice 12. *Soit $(U_i)_{i\geq 1}$ une suite de variables aléatoires uniformes sur l'intervalle $[0,1]$. Montrer que la suite $(n\min_{1\leq i\leq n} U_i)_n$ converge en loi lorsque $n\to\infty$ vers une loi exponentielle de paramètre 1.*

Exercice 13. *Soit $(X_n)_n$ une suite de variables aléatoires convergeant en loi vers Y et $f:\mathbb{R}\to\mathbb{R}$ une fonction continue. Montrer que la suite $(f(X_n))_n$ converge en loi vers $f(Y)$.*

On a le résultat suivant

Théorème 3.2.3 (Théorème de Prokhorov)**.** *Soit $(X_n)_n$ une suite tendue de v.a. à valeurs dans $\mathbb{R}^d$. Alors il existe une v.a. X sur $\mathbb{R}^d$ et une suite extraite $(\psi(n))_n$ telle que $(X_{\psi(n)})_n$ converge en loi vers X. Inversement, si la suite $(X_n)_n$ converge en loi vers X alors cette suite est tendue.*

Remarque 3.2.4. *Ce théorème peut s'étendre (avec la même preuve) au cas d'espaces métrisables.*

Démonstration. Soit $(P_n)_n$ les lois respectives des variables aléatoires $(X_n)_n$ sur $\mathbb{R}^d$. Tout d'abord, on peut choisir une suite $(M_i)_{i\geq 1}$ strictement croissante et convergeant vers l'infini telle que $P_n(\bar{B}(0,M_i)) = \mathbb{P}(|X_n|\leq M_i)\geq 1-\frac{1}{i}$ pour tous n et i. Introduisons $\mathcal{Q}$ la famille dénombrable des boules ouvertes à centre et rayon rationnels et $\mathcal{H}$ la famille (dénombrable) des unions finies d'éléments de la forme $\bar{B}$ avec $B\in\mathcal{Q}$. Par le procédé d'extraction diagonale, on peut extraire une sous-suite notée $(P_{n'})_{n'}$ telle que $P_{n'}(H)$ converge en $n'\to\infty$ pour tout $H\in\mathcal{H}$. Notons

$$T(H) = \lim_{n'\to\infty} P_{n'}(H).$$

Pour tout ouvert O on définit alors

$$T'(O) = \sup_{H\subset O, H\in\mathcal{H}} T(H).$$

On veut montrer qu'il existe une mesure de proba sur $\mathbb{R}^d$ muni de sa tribu borélienne telle que $P(O) = T'(O)$ pour tout ouvert O. Posons pour tout ensemble arbitraire E

$$P(E) = \inf_{E\subset O} T'(O).$$

• **$\mathbb{P}$ définit une mesure de proba sur la tribu des boréliens :** 1) Notons tout d'abord que T vérifie trivialement les propriétés suivantes

$$\begin{aligned} T(H_1) &\leq T(H_2) && \text{si } H_1\subset H_2\\ T(H_1\cup H_2) &= T(H_1)+T(H_2) && \text{si } H_1\cap H_2=\emptyset\\ T(H_1\cup H_2) &\leq T(H_1)+T(H_2). \end{aligned}$$

2) Montrons que pour tout compact F et ouvert 0, il existe $H\in\mathcal{H}$ tel que $F\subset H\subset O$. En effet pour tout $x\in F$ il existe $B_x\in\mathcal{Q}$ tel que $x\in B_x\subset\bar{B}_x\subset O$. La famille $(B_x)_{x\in F}$ est un recouvrement ouvert du compact F donc on peut extraire un recouvrement fini $(B_{x_i})_i$. On peut conclure en prenant $H=\bigcup_i \bar{B}_{x_i}$.

3) Montrons ensuite que T' est sous-additive sur les ouverts. Soient donc 2 ouverts O_1 et O_2, et $H \in \mathcal{H}$ tel que $H \subset O_1 \cup O_2$. On définit

$$F_1 = \{x \in H; \mathrm{dist}(x, O_1^c) \geq \mathrm{dist}(x, O_2^c)\}$$

$$F_2 = \{x \in H; \mathrm{dist}(x, O_2^c) \geq \mathrm{dist}(x, O_2^c)\}$$

de sorte que $H = F_1 \cup F_2$ et $F_1 \subset O_1$, $F_2 \subset O_2$. Comme F_1, F_2 sont compacts (fermés et inclus dans H) le point 2 assure qu'on peut trouver $H_1, H_2 \in \mathcal{H}$ tels que $F_1 \subset H_1 \subset O_1$, $F_2 \subset H_2 \subset O_2$. D'où

$$T(H) \leq T(H_1 \cup H_2) \leq T(H_1) + T(H_2) \leq T'(O_1) + T'(O_2).$$

Puis $T'(O) \leq T'(O_1) + T'(O_2)$.

4) Montrons ensuite que T' est sous-additive sur les réunions dénombrables d'ouverts. Soit $h \in \mathcal{H}$ tel que $H \subset \bigcup_n O_n$. Comme H est compact on peut extraire un sous-recouvrement fini et donc $H \subset \bigcup_{n \leq N} O_n$ pour un certain $N \in \mathbb{N}$. Donc

$$T(H) \leq T'(\bigcup_{n \leq N} O_n) \leq \sum_{n \leq N} T'(O_n) \leq \sum_n T'(O_n),$$

ce qui implique

$$T'(\bigcup_n O_n) \leq \sum_n T'(O_n).$$

5) P est une mesure extérieure. En effet P est clairement monotone et $P(\emptyset) = 0$. Il reste à montrer que P est dénombrablement sous-additive. Soit $(E_n)_n$ une suite de sous-ensembles de $\mathbb{R}^d$. Choisissons pour tout n un ouvert O_n arbitraire tel que $E_n \subset O_n$. On a alors $\bigcup_n E_n \subset \bigcup_n O_n$ de sorte que

$$P(\bigcup_n E_n) \leq T'(\bigcup_n O_n) \leq \sum_n T'(O_n)$$

Les $(O_n)_n$ étant arbitraires on en déduit

$$P(\bigcup_n E_n) \leq \sum_n P(E_n).$$

6) Rappelons alors que si l'on dispose d'une mesure extérieure P , celle-ci définit une mesure sur la tribu $\mathcal{A}$ formé des parties A satisfaisant

$$\forall C \text{ partie de } \mathbb{R}^d, \quad P(C) = P(A \cap C) + P(A^c \cap C).$$

Il reste donc juste à montrer que cette tribu contient les boréliens. Soient F fermé et O ouvert. Soient $H_1, H_2 \in \mathcal{H}$ tels que

$$\begin{array}{lcl} H_1 \subset F^c \cap O & \text{and} & T(H_1) > T'(F^c \cap O) - \epsilon, \\ H_2 \subset H_1^c \cap O & \text{and} & T(H_2) > T'(H_1^c \cap O) - \epsilon. \end{array}$$

Comme H_1, H_2 sont disjoints et contenus dans O

$$\begin{aligned} T'(O) \geq T(H_1 \cup H_2) = T(H_1) + T(H_2) &> T'(F^c \cap O) + T'(H_1^c \cap O) - 2\epsilon \\ &\geq P(F^c \cap O) + P(F \cap O) - 2\epsilon. \end{aligned}$$

ϵ étant arbitraire on en déduit $T'(O) \geq P(F^c \cap O) + P(F \cap O)$. Soit maintenant C une partie arbitraire de $\mathbb{R}^d$, F fermé et O un ouvert contenant C. On en déduit

$$T'(O) \geq P(F^c \cap C) + P(F \cap C)$$

O contenant C étant arbitraire, on en déduit

$$P(C) \geq P(F^c \cap C) + P(F \cap C).$$

Donc les fermés sont mesurables et P est une mesure sur la tribu des boréliens. P est une proba car pour tous i il existe $H_i \in \mathcal{H}$ tel que $B(0, M_i) \subset H_i$, d'où

$$P(\mathbb{R}^d) = T'(\mathbb{R}^d) \geq T(H_i) = \lim_{n'} P_{n'}(H_i) \geq \lim_{n'} P_{n'}(B(0, M_i)) \geq 1 - 1/i$$

pour tout i.

- **Convergence en loi** Observant que pour tout ouvert O et tout $H \subset O$

$$T(H) = \lim_{n'} Pn'(O) \leq \liminf_{n'} P_{n'}(O)$$

et optimisant en H

$$P(O) = T'(O) = \sup_{H \subset O, H \in \mathcal{H}} T(H) \leq \liminf_{n'} P_{n'}(O).$$

Le théorème de Portmanteau permet de conclure. □

Proposition 3.2.5. *Soit $(X_n)_n$ une suite de v.a. qui converge en probabilité vers X. Alors $(X_n)_n$ converge en loi vers X.*

Démonstration. Montrons l'item 4 de la proposition 3.2.2. Soit f une fonction C^∞ à support compact. En particulier, f est bornée et K-Lipschitzienne. D'où

$$\begin{aligned}\big|\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)]\big| \leq & \mathbb{E}\big|f(X_n) - f(X)\big| \\ \leq & \mathbb{E}\big[\big|f(X_n) - f(X)\big|\mathbf{1}_{|X_n - X| \leq \epsilon}\big] + \mathbb{E}\big[\big|f(X_n) - f(X)\big|\mathbf{1}_{|X_n - X| > \epsilon}\big] \\ \leq & 2K\epsilon + 2\|f\|_\infty \mathbb{P}(|X_n - X| > \epsilon).\end{aligned}$$

Il suffit alors de prendre le $\limsup_n$ et de noter que ϵ peut être choisi arbitrairement petit. □

3.3 Convergence en loi et fonction de répartition

Proposition 3.3.1. *Soit $(X_n)_n$ une suite de variables aléatoires réelles et X une variable aléatoire réelle. Se valent*

1. *$(X_n)_n$ converge en loi vers X.*
2. *Les fonctions de répartition de la suite $(X_n)_n$ convergent vers la fonction de répartition de X en tout point $x \in \mathbb{R}$ tel que $\mathbb{P}(X = x) = 0$:*

$$F_{X_n}(x) \xrightarrow[n \to \infty]{} F_X(x) \quad si \quad \mathbb{P}(X = x) = 0.$$

Démonstration. Il s'agit d'approcher $\mathbf{1}_{]-\infty,x]}$ par des fonctions continues bornées. Soit h_m la fonction continue bornée, affine par morceaux telle que

— $h_m = 1$ sur $]-\infty, x - \frac{1}{m}]$
— $h_m = 0$ sur $[x, +\infty[$
— la pente de h_m vaut $-m$ sur $[x - \frac{1}{m}, x]$.

Soit g_m la fonction continue bornée, affine par morceaux telle que

— $g_m = 1$ sur $]-\infty, x]$
— $g_m = 0$ sur $[x + \frac{1}{m}, +\infty[$
— la pente de g_m vaut $-m$ sur $[x, x + \frac{1}{m}]$.

On a alors

$$0 \leq \mathbb{E}[g_m(X)] - \mathbb{E}[h_m(X)] \leq \mathbb{P}(X \in [x - \frac{1}{m}, x + \frac{1}{m}])$$

et cette dernière quantité tend vers $\mathbb{P}(X = x) = 0$ lorsque $m \to \infty$. Choisissons donc m tel que $\mathbb{E}[g_m(X)] - \mathbb{E}[h_m(X)] \leq \epsilon/2$.

Comme $(X_n)_n$ converge en loi vers X, on a pour n assez grand

$$\begin{aligned} F_X(x) - \epsilon \leq \mathbb{E}[h_m(X)] - \epsilon/2 \\ &\leq \mathbb{E}[h_m(X_n)] \leq F_{X_n}(x) \leq \mathbb{E}[g_m(X_n)] \\ &\qquad \leq \mathbb{E}[g_m(X)] + \epsilon/2 \leq F_X(x) + \epsilon. \end{aligned}$$

Réciproquement, Notons C l'ensemble des points $x \in \mathbb{R}$ tels que F n'est pas continue en x. Comme F est croissante, cet ensemble C est au plus dénombrable, donc dense dans $\mathbb{R}$. Soit $\mathcal{E}$ l'ensemble des fonctions étagées dont les points de discontinuité appartiennent à $\mathbb{R} \setminus C$

$$\mathcal{E} := \{f = \sum_{i=1}^{N} \lambda_i \mathbf{1}_{]a_i,b_i]} \mid N \in \mathbb{N}^*, \lambda_i \in \mathbb{R} \text{ et } a_i, b_i \in \mathbb{R} \setminus C\}.$$

Montrons que cet ensemble est dense dans $C_c(\mathbb{R})$ pour la norme uniforme. Soit $g \in C_c(\mathbb{R})$. Comme g est à support compact et continue, g est uniformément continue sur $\mathbb{R}$. Soit $\epsilon > 0$. Le théorème de Heine implique l'existence de $\delta > 0$ tel que

$$|g(x) - g(x')| \leq \epsilon$$

pour tout $|x - x'| \leq \delta$. Pour tout $k \in \mathbb{Z}$, choisissons α_k dans l'ensemble $]k\frac{\delta}{2}, (k+1)\frac{\delta}{2}] \cap (\mathbb{R} \setminus C)$. Notons que la suite $(\alpha_k)_k$ est croissante et vérifie $|\alpha_k - \alpha_{k-1}| \leq \delta$. Définissons une fonction f dans $\mathcal{E}$ par les relations

$$f(x) = g(\alpha_k) \quad \text{pour} \quad x \in]\alpha_{k-1}, \alpha_k].$$

Ainsi pour tout $x \in \mathbb{R}$, il existe k tel que $x \in]\alpha_{k-1}, \alpha_k]$. Ainsi $|g(x) - f(x)| = |g(x) - g(\alpha_k)| \leq \epsilon$ car $|x - \alpha_k| \leq \delta$.

Finalement pour $f \in \mathcal{E}$ on a pour $n \to \infty$

$$\mathbb{E}[f(X_n)] = \sum_{i=1}^{N} \lambda_i (F_{X_n}(b_i) - F_{X_n}(a_i)) \to \sum_{i=1}^{N} \lambda_i (F_X(b_i) - F_X(a_i)) = \mathbb{E}[f(X)]$$

car les $y = a_i, b_i$ vérifie $\mathbb{P}(X = y) = 0$. Finalement si g est dans $C_c(\mathbb{R})$ il existe une suite $(f_k)_k$ de $\mathcal{E}$ qui converge vers g pour la norme uniforme (théorème de Heine). On a donc pour tous n, k

$$\begin{aligned} \big|\mathbb{E}[g(X_n)] - \mathbb{E}[g(X)]\big| &\leq \big|\mathbb{E}[g(X_n)] - \mathbb{E}[f_k(X_n)]\big| + \big|\mathbb{E}[f_k(X_n)] - \mathbb{E}[f_k(X)]\big| + \big|\mathbb{E}[f_k(X)] - \mathbb{E}[g(X)]\big| \\ &\leq 2\epsilon + \big|\mathbb{E}[f_k(X_n)] - \mathbb{E}[f_k(X)]\big|. \end{aligned}$$

En prenant la limite sup
$$\limsup_n \left|\mathbb{E}[g(X_n)] - \mathbb{E}[g(X)]\right| \leq 2\epsilon$$
et on conclut en choisissant ϵ arbitrairement petit. □

Exercice 14. Points de discontinuité des fonctions croissantes. *Soit f une fonction croissante sur $\mathbb{R}$.*

1. *Montrer qu'en tout point $x \in \mathbb{R}$, f admet une limite à gauche $f_g(x)$ et une limite à droite $f_d(x)$ qui satisfont*
$$f_g(x) \leq f(x) \leq f_d(x)$$
et que f est continue en x si et seulement si $f_g(x) = f(x) = f_d(x)$.
2. *Soient $a < b$ et*
$$C_{k,a,b} = \{x \in]a,b[\mid f_g(x) < f_d(x)\}.$$
Montrer que $\#C_{k,a,b} \leq k(f(b) - f(a))$.
3. *En déduire que le nombre de discontinuités de f sur $]a,b[$ est au plus dénombrable.*
4. *Conclure que f a au plus un nombre dénombrable de discontinuités sur $\mathbb{R}$.*

3.4 Borel-Cantelli

On rappelle que si $(A_n)_n$ est une suite d'ensemble on définit
$$\limsup_n A_n = \bigcap_{N\in\mathbb{N}} \bigcup_{n\geq N} A_n \quad \text{et} \quad \liminf_n A_n = \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} A_n.$$

Exercice 15. *Montrer que le $\limsup_n A_n$ correspond à l'évènement "une infinité d'A_n est réalisée" et que $\liminf_n A_n$ correspond à "tous les A_n sont réalisés sauf peut-être un nombre fini".*

Lemme 3.4.1 (Borel-Cantelli)**.** *Soit $(A_n)_n$ une suite d'évènements de l'espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$.*

1. *si $\sum_n \mathbb{P}(A_n) < +\infty$ alors*
$$\mathbb{P}\Big(\limsup_n A_n\Big) = 0.$$
2. *si les $(A_n)_n$ sont indépendants et si $\sum_n \mathbb{P}(A_n) = +\infty$ alors*
$$\mathbb{P}\Big(\limsup_n A_n\Big) = 1.$$

Démonstration. 1.) Soit $B_N = \bigcup_{n\geq N} A_n$. C'est une suite décroissante d'ensemble pour l'inclusion. Donc
$$\mathbb{P}(\limsup_n A_n) = \lim_{N\to\infty} \mathbb{P}(B_N).$$
Or $\mathbb{P}(B_N) \leq \sum_{n\geq N} \mathbb{P}(A_n)$ qui est le reste d'une série convergente. Donc $\lim_{N\to\infty} \mathbb{P}(B_N) = 0$.

2.) On va montrer que le complémentaire a proba 0. Or
$$\mathbb{P}\Big((\limsup_n A_n)^c\Big) = \mathbb{P}\Big(\liminf_n A_n^c\Big) \leq \sum_N \mathbb{P}\Big(\bigcap_{n\geq N} A_n^c\Big).$$

Il faut donc montrer que chacun des termes de la série est nul. Or pour $k \geq N$

$$\mathbb{P}\Big(\bigcap_{n\geq N} A_n^c\Big) \leq \mathbb{P}\Big(\bigcap_{N\leq n\leq k} A_n^c\Big) = \prod_{N\leq n\leq k} \mathbb{P}(A_n^c) = \prod_{N\leq n\leq k} (1-\mathbb{P}(A_n))$$

En utilisant l'inégalité $1-u \leq e^{-u}$ valable pour $u \geq 0$ on en déduit

$$\prod_{N\leq n\leq k} (1-\mathbb{P}(A_n)) \leq \exp\Big(-\sum_{N\leq n\leq k} \mathbb{P}(A_n)\Big) \xrightarrow[k\to\infty]{} 0.$$

D'où le résultat. $\square$

Exercice 16. *Soit $(Y_n)_n$ une suite de v.a.. Si pour tout $\epsilon > 0$*

$$\sum_n \mathbb{P}(|Y_n| > \epsilon) < +\infty.$$

Montrer que la suite $(Y_n)_n$ converge presque sûrement vers 0.

Exercice 17. *Un singe s'amuse à taper des lettres sur une machine à écrire, sans jamais s'arrêter. On modèlise cette expérience par une suite de variables aléatoires indépendantes $(X_n)_{n\geq 0}$ prenant leurs valeurs dans l'ensemble des lettres de l'alphabet. Montrer que, presque sûrement, le singe va reproduire à un moment donné l'oeuvre complète de Victor Hugo "Les misérables".*

Chapitre 4

Fonction génératrice, fonction caractéristique

Dans ce chapitre, on va développer certains outils d'analyse fonctionnelle pour obtenir des critères plus manipulables reliés à la convergence en loi. L'idée sous-jacente est que l'espace des fonctions continues bornées utilisé dans la définition de la convergence en loi peut-être caractérisé par des sous-ensembles denses pour lesquels les calculs d'espérance sont plus explicites.

4.1 Fonction génératrices

On considère ici le cas particulier des v.a. à valeurs dans l'ensemble $\mathbb{N}$ des entiers naturels. Ces v.a. interviennent souvent dans les applications. Il est donc important de disposer d'une méthode générale qui facilite le calcul de leur loi de probabilité et de leurs moments. Une idée en ce sens consiste à associer à toute v.a. X à valeurs dans $\mathbb{N}$, une série entière qui contient tous les renseignements concernant la loi de probabilité de X et qui a l'avantage d'être bien adaptée à l'opération consistant à additionner des v.a. indépendantes.

Définition 4.1.1. *Fonction génératrice Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace de probabilité et soit $X : \Omega \to \mathbb{N}$ une variable aléatoire discrète. On appelle fonction génératrice de X la fonction*

$$G_X(t) = \sum_{n=0}^{\infty} \mathbb{P}(X = n)t^n.$$

définie pour les t où la série converge.

La première question qui se pose est le domaine de définition de cette fonction.

Proposition 4.1.2. *Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace de probabilité et soit $X : \Omega \to \mathbb{N}$ une variable aléatoire discrète. On considère $G_X(t) = \sum_{n=0}^{\infty} \mathbb{P}(X = n)t^n$. On a*

1. *cette série est absolument convergente pour $t \in [-1, 1]$.*
2. $G_X(1) = 1$
3. *deux variables aléatoires $X, Y : \Omega \to \mathbb{N}$ ont même loi si et seulement si $G_X(t) = G_Y(t)$ pour tout $t \in [0, \delta]$ pour un certain $\delta > 0$.*

Démonstration. 1) résulte du fait que $0 \leq |\mathbb{P}(X = n)t^n| \leq 1$ pour tout $n \in \mathbb{N}, t \in [-1, 1]$ et que la série $\sum_n \mathbb{P}(X = n)$ converge .

2) est évident.

Si X et Y ont même loi il est clair que $G_X(t) = G_Y(t)$ pout $t \in [0, 1]$. Réciproquement, on a pour $t \in [0, 1[$ et $n \in \mathbb{N}$

$$G_X(t) = \sum_{k=0}^{n} \mathbb{P}(X = k)t^k + R_k(t)$$

où

$$R_k(t) = \sum_{k=n+1}^{\infty} \mathbb{P}(X = k)t^k.$$

Notons que

$$|R_k(t)| \leq \sum_{k=n+1}^{\infty} t^k = \frac{t^{n+1}}{1-t} = o(t^n).$$

Donc G_X admet un DL en 0 à l'ordre n pour tout n. De même pour G_Y. Par unicité de la partie régulière d'un DL on en déduit que $\forall k\ \mathbb{P}(X = k) = \mathbb{P}(Y = k)$. Donc X et Y ont même loi. □

Dans l'exemple où X, à valeurs dans $\mathbb{N}$, suit la loi de Poisson de paramètre λ, on a

$$G_X(t) = \sum_{n=0}^{\infty} \mathbb{P}(X = n)t^n = e^{-\lambda}\frac{\lambda^k}{k!}t^n = e^{\lambda(t-1)},$$

pour tout $t \in \mathbb{R}_+$.

Prenons ensuite l'exemple où X suit une loi binomiale $B(n, p)$. On a alors

$$G_X(t) = \sum_{k=0}^{n} \binom{n}{k} p^k(1-p)^{n-k}t^k = (pt + 1 - p)^n.$$

On admet les propriétés suivantes.

Théorème 4.1.3. *Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace de probabilité et soit E un ensemble. Soit $X : \Omega \to \mathbb{N}$ une variable aléatoire discrète. Nous notons G_X la fonction génératrice de X.*

1. *Si X admet une espérance, alors $G_X(t)$ admet une dérivée à gauche $G'_X(1)$ en $t = 1$, et l'on a*
$$\mathbb{E}(X) = G'_X(1).$$
2. *Si X admet un moment d'ordre 2, alors $G_X(t)$ admet une dérivée seconde à gauche $G''_X(t)$ en $t = 1$, et l'on a*
$$\mathrm{Var}(X) = G''_X(1) + G'_X(1) - G'_X(1)^2.$$
3. *inversement si $G_X(t)$ est une fois (resp. 2 fois) dérivable en $t = 1$ alors X est intégrable et $\mathbb{E}(X) = G'_X(1)$ (resp. 2 fois dérivable et $\mathrm{Var}(X) = G''_X(1) + G'_X(1) -$*

$G'_X(1)^2$).

Démonstration. résulte du théorème de dérivation des séries. □

Proposition 4.1.4. *Soit $(\Omega, \mathcal{E}, \mathbb{P})$ un espace de probabilité et soit E un ensemble. Soient $X : \Omega \to \mathbb{N}$ et $Y : \Omega \to \mathbb{N}$ deux variables aléatoires discrètes de fonctions génératrices G_X et G_Y. Si X et Y sont indépendantes alors*

$$\forall t \in]-1, 1[, \; G_{X+Y}(t) = G_X(t) G_Y(t).$$

Démonstration. On a $G_{X+Y}(t) = \mathbb{E}[t^{X+Y}] = \mathbb{E}[t^X]\mathbb{E}[t^Y]$ par indépendance. □

Exercice 18. *Soit $(X_i)_{1 \leq i \leq}$ une suite de v.a. binomiales de paramètres respectifs (n_i, p). Montrer que $X_1 + \cdots + X_n$ suit la loi binomiale $\mathcal{B}(\sum_i n_i, p)$.*

Exercice 19. *Soit $(X_i)_{1 \leq i \leq}$ une suite de v.a. de Poisson de paramètres respectifs λ_i. Montrer que $X_1 + \cdots + X_n$ suit la loi de Poisson $\mathcal{P}(\sum_i \lambda_i)$.*

Exercice 20. *Soit $(X_n)_{n \geq 0}$ une suite de v.a. i.i.d. à valeurs dans $\mathbb{N}$ de fonction génératrice G. Soit N une v.a. à valeurs dans $\mathbb{N}$ de fonction génératrice G indépendante de la suite $(X_n)_{n \geq 0}$.*

1. *Quelle est la fonction génératrice de la variable aléatoire $S = \sum_{k=0}^{N} X_k$. ?*
2. *On suppose que X_0 et N admettent des moments d'ordre 2. Donner une expression pour $\mathbb{E}[S]$ et $\mathrm{Var}(S)$.*

4.2 Fonction caractéristique

4.2.1 Fonction caractéristique : définitions et premières propriétés

Définition 4.2.1. *La fonction caractéristique d'une v.a. $X : \Omega \to \mathbb{R}$ est définie par*

$$\phi_X(t) = \mathbb{E}[e^{itX}] = \int_{\mathbb{R}} e^{itu} \, \mathbb{P}_X(du).$$

La transformée de Fourier d'une mesure de probabilité μ définie sur la tribu des boréliens de $\mathbb{R}$ est définie par

$$\hat{\mu}(t) = \int_{\mathbb{R}} e^{itu} \mu(du).$$

On a donc l'égalité $\phi_X(t) = \hat{\mathbb{P}}_X(t)$.

Proposition 4.2.2. *On a :*

- *$|\phi_X(t)| \leq 1$ pour tout $t \in \mathbb{R}$*
- *$\phi_X(0) = 1$*
- *$t \mapsto \phi_X(t)$ est continue sur $\mathbb{R}$*
- *si X est intégrable alors $t \mapsto \phi_X(t)$ est de classe C^1 et $\phi'_X(0) = i\mathbb{E}[X]$*
- *si X est de carré intégrable alors $t \mapsto \phi_X(t)$ est de classe C^2 et $\phi''_X(0) = -\mathbb{E}[X^2]$*

Démonstration. La continuité et dérivabilité découlent des théorèmes de continuité et dérivabilité sous le signe $\int$. Par exemple, si X est intégrable, on a la majoration

$$\left|\frac{\partial}{\partial t}e^{itX}\right| = \left|iXe^{itX}\right| \leq |X|$$

ce qui implique

$$\phi'_X(t) = \mathbb{E}[iXe^{itX}].$$

□

Proposition 4.2.3. *On a : Si $X_1, \ldots, X_n$ sont des variables aléatoires indépendantes alors*

$$\phi_{X_1+\cdots+X_n}(t) = \phi_{X_1}(t) \times \cdots \times \phi_{X_n}(t).$$

Démonstration. On a par définition

$$\phi_{X_1+\cdots+X_n}(t) = \mathbb{E}[e^{it(X_1+\cdots+X_n)}] = \mathbb{E}\Big[\prod_{k=1}^{n} e^{itX_k}\Big].$$

Par indépendance, cette dernière quantité vaut $\prod_{k=1}^n \mathbb{E}\left[e^{itX_k}\right]$, ce qui correspond exactement au résultat annoncé. □

4.2.2 Fonction caractéristique et loi des variables aléatoires

Théorème 4.2.4. *Deux variables aléatoires qui ont même fonction caractéristique ont même loi, c'est-à-dire*

$$\phi_X = \phi_Y \Rightarrow \mathbb{P}_X = \mathbb{P}_Y.$$

Démonstration. Remarquons que si $f \in L^1(\mathbb{R})$ alors

$$\int_{\mathbb{R}} f(t)\phi_X(t)\,dt = \mathbb{E}[\hat{f}(X)]$$

où $\hat{f}$ désigne la transformée de Fourier de f :

$$\hat{f}(\lambda) = \int f(u)e^{i\lambda u}\,du.$$

Ainsi, si $\phi_X = \phi_Y$ alors pour toute fonction $f \in L^1(\mathbb{R})$

$$\mathbb{E}[\hat{f}(X)] = \mathbb{E}[\hat{f}(Y)].$$

On utilise alors le lemme suivant

Lemme 4.2.5. *L'image de l'espace $L^1(\mathbb{R})$ par la transformée de Fourier*

$$f \mapsto \hat{f}(\lambda) = \int f(u)e^{i\lambda u}\,du$$

est dense dans $C_0(\mathbb{R})$ muni de la norme uniforme.

On en déduit donc que $\mathbb{E}[f(X)] = \mathbb{E}[f(Y)]$ pour $f \in C_0(\mathbb{R})$, ce qui conclut la preuve. □

Preuve du lemme (admise). La transformée de Fourier préserve l'espace de Schwartz $\mathcal{S}$, qui est contenu dans $L^1(\mathbb{R})$ et dense dans l'espace $C_0(\mathbb{R})$ pour la norme uniforme. □

Théorème 4.2.6. *Si ϕ_X est la fonction caractéristique d'une v.a. X et est intégrable, i.e. $\int_{\mathbb{R}} |\phi_X(t)|\, dt < \infty$ alors X admet une densité f_X par rapport à la mesure de Lebesgue donnée par*

$$f_X(x) = \frac{1}{2\pi} \int_R e^{-ixu} \phi_X(u)\, du.$$

Preuve (admise). utilise la transformation de Fourier inverse. □

Théorème 4.2.7 (Théorème de Lévy)**.** *Soit $(X_n)_n$ une suite de variables aléatoires réelles. La suite $(X_n)_n$ converge en loi si et seulement si les fonctions caractéristiques $(\phi_n)_n$ converge simplement vers une fonction $\phi : \mathbb{R} \to \mathbb{R}$ continue en 0. Si c'est le cas, il existe une variable aléatoire réelle X telle que $\phi = \phi_X$ et la suite $(X_n)_n$ converge en loi vers X.*

Démonstration. $\Rightarrow$ Si la suite $(X_n)_n$ converge en loi alors il existe une v.a. X telle que $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$ quand $n \to \infty$ pour toute $f \in C_b(\mathbb{R})$. En particulier on peut prendre $f(u) = e^{itu}$.

$\Leftarrow$ Montrons tout d'abord que la suite $(X_n)_n$ est tendue en utilisant le théorème 3.2.3 de Prokhorov. En utilisant le théorème de Fubini, on obtient pour $u > 0$

$$\frac{1}{u} \int_{-u}^{u} (1 - \phi_n(r))\, dr = \mathbb{E}\Big[\frac{1}{u} \int_{-u}^{u} (1 - e^{irX_n})\, dr\Big] = 2\mathbb{E}\Big[1 - \frac{\sin(uX_n)}{uX_n}\Big].$$

Soit $K = 2 \min_{|y| \geq 1} \left(1 - \frac{\sin y}{y}\right) > 0$. De l'inégalité

$$\forall y \in \mathbb{R}, \quad \mathbf{1}_{|y| \geq 1} \leq \frac{2}{K}\big(1 - \frac{\sin y}{y}\big)$$

on déduit

$$\begin{aligned}
\mathbb{P}(|X_n| > \tfrac{1}{u}) = & \int_{\mathbb{R}} \mathbf{1}_{|ux| \geq 1}\, \mathbb{P}_{X_n}(dx) \\
\leq & \frac{1}{K} \int_{\mathbb{R}} 2\big(1 - \frac{\sin(ux)}{ux}\big) \mathbb{P}_{X_n}(dx) \\
= & \frac{1}{Ku} \int_{-u}^{u} (1 - \phi_n(r))\, dr
\end{aligned}$$

Concluons pour la tension de la suite $(X_n)_n$. Soit $\epsilon > 0$. Comme la suite $(\phi_n)_n$ converge simplement vers ϕ, il existe N tel que $\forall n \geq N$

$$\Big| \frac{1}{Ku} \int_{-u}^{u} (1 - \phi_n(r))\, dr - \frac{1}{Ku} \int_{-u}^{u} (1 - \phi(r))\, dr \Big| \leq \epsilon/2.$$

De plus par continuité de ϕ en 0, il existe u assez petit tel que

$$\Big|\frac{1}{Ku}\int_{-u}^{u}(1-\phi(r))\,dr\Big| \leq \epsilon/2.$$

Donc $n \geq N$ on a

$$\mathbb{P}(|X_n| > \tfrac{1}{u}) \leq \frac{1}{Ku}\int_{-u}^{u}(1-\phi_n(r))\,dr \leq \epsilon.$$

Ceci montre la tension. Le théorème 3.2.3 de Prokhorov permet de conclure à l'existence d'une suite extraite de $(X_n)_n$ convergeant en loi vers une v.a. notée X. Cette convergence nous permet d'identifier $\phi = \phi_X$ par convergence simple des $(\phi_n)_n$.

Il reste à montrer la convergence en loi. Soit $f \in L^1(\mathbb{R})$. On a

$$\mathbb{E}[\hat{f}(X_n)] = \int_{\mathbb{R}} f(t)\phi_n(t)\,dt \xrightarrow[n\to\infty]{} \int_{\mathbb{R}} f(t)\phi_X(t)\,dt = \mathbb{E}[\hat{f}(X)].$$

On peut alors facilement conclure en utilisant le lemme 4.2.5. $\square$

4.2.3 Fonctions caractéristiques des vecteurs aléatoires

Dans cette partie on notera $X = (X_1, \dots, X_n)$ un vecteur aléatoire de $\mathbb{R}^n$ et $T = (t_1, \dots, t_n)$ un élément de $\mathbb{R}^n$.

Proposition 4.2.8. *Soient $X_1, \dots, X_n$ des variables aléatoires. Celles-ci sont indépendantes si et seulement si*

$$\phi_X(T) = \phi_{X_1}(t_1) \times \cdots \times \phi_{X_n}(t_n).$$

Définition 4.2.9. *Soit X un vecteur aléatoire de $\mathbb{R}^n$. On appelle fonction caractéristique du vecteur X la fonction ϕ_X définie sur $\mathbb{R}^n$ à valeurs dans $\mathbb{C}$ par*

$$\forall T \in \mathbb{R}^n, \quad \phi_X(T) = \mathbb{E}[e^{i\langle T,X\rangle}]$$

où $\langle T, X\rangle$ est le produit scalire sur $\mathbb{R}^n$: $\langle T, X\rangle = t_1X_1 + \cdots + t_nX_n$.

A noter qu'il résulte de la définition la relation suivante

$$\phi_X(0, \dots, 0, t_i, 0, \dots, 0) = \phi_{X_i}(t_i).$$

On retrouve les propriétés élémentaires des fonctions caractéristiques

Proposition 4.2.10. *On a :*
- *$|\phi_X(T)| \leq 1$ pour tout $T \in \mathbb{R}^n$*
- *$\phi_X(0) = 1$*
- *$T \mapsto \phi_X(T)$ est continue sur $\mathbb{R}^n$*
- *si X est intégrable alors $T \mapsto \phi_X(T)$ est de classe C^1 et $\nabla\phi_X(0) = i\mathbb{E}[X]$*
- *si X est de carré intégrable alors $T \mapsto \phi_X(T)$ est de classe C^2 et $\partial^2_{t_it_j}\phi_X(0) = -\mathbb{E}[X_iX_j]$*

On obtient également un critère pour avoir un vecteur X à densité

Théorème 4.2.11. *Soit X un vecteur aléatoire de $\mathbb{R}^n$ de fonction caractéristique ϕ_X intégrable, i.e. $\int_{\mathbb{R}^n} |\phi_X(T)|\, dT < \infty$ alors X admet une densité f_X par rapport à la mesure de Lebesgue sur $\mathbb{R}^n$ donnée par*

$$f_X(x_1, \ldots, x_n) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-i\sum_{k=1}^n x_k t_k} \phi_X(t_1, \ldots, t_n)\, dt_1 \ldots dt_n.$$

Le théorème 4.2.7 de Lévy pour la convergence en loi reste vrai ainsi que le théorème 4.2.4 sur le fait que la fonction caractéristique caractérise la loi d'un vecteur aléatoire. Ceci nous permet de prouver

Proposition 4.2.12. *Soit $X = (X_1, \ldots, X_n)$ un vecteur aléatoire de $\mathbb{R}^n$. Alors $X_1, \ldots, X_n$ sont indépendantes si et seulement si*

$$\forall T = (t_1, \ldots, t_n) \in \mathbb{R}^n, \quad \phi_X(T) = \phi_{X_1}(t_1) \times \ldots \phi_{X_n}(t_n).$$

Démonstration. Une implication est évidente. Pour l'autre, on remarque que si l'on a égalité des fonctions caractéristiques, alors ϕ_X est égale à la fonction caractéristique du même vecteur aléatoire mais à composantes indépendantes. Comme la fonction caractéristique caractérise la loi, ceci entraîne que X a ses composantes indépendantes. □

4.2.4 Exercices

Exercice 21. *Calculer la fonction caractéristique de la v.a. X dans les cas suivants :*

1. *X suit une loi normale $\mathcal{N}(m, \sigma^2)$*
2. *X est une variable discrète de loi uniforme sur $\{1, \ldots, n\}$.*
3. *X suit une loi uniforme sur l'intervalle $[a, b]$ avec $a < b$.*
4. *X suit une loi exponentielle de paramètre $\lambda > 0$.*

Exercice 22. *Soient X, Y deux variables aléatoires indépendantes à densité respectivement donnée par f_X et f_Y. Montrer que la somme $X + Y$ a une densité donnée par*

$$f_{X+Y}(x) = \int_{\mathbb{R}} f_X(x - y) f_Y(y)\, dy.$$

Exercice 23. *Soit X suivant une loi de Cauchy $\mathcal{C}(x_0, a)$, i.e. avec densité sur $\mathbb{R}$ donnée par $\frac{1}{\pi} \frac{a}{a^2 + (x - x_0)^2}$.*

1. *Montrer que si $Y \sim \mathcal{C}(0, 1)$ alors $aY + x_0$ suit une loi $\mathcal{C}(x_0, a)$. En déduire que*

   $$\phi_X(t) = e^{ix_0} \phi_Y(at).$$

   On est donc ramené à calculer $\phi_Y(t)$ où $Y \sim \mathcal{C}(0, 1)$.
2. *Soit U, V des variables aléatoires indépendantes de loi exponentielles de paramètre 1. Quelle est la loi de $U - V$?*
3. *En déduire la fonction caractéristique de $U - V$.*
4. *Conclure.*

Exercice 24. *Soit X suivant une loi de Cauchy $\mathcal{C}(0,1)$. On se propose de calculer sa fonction caractéristique sans utiliser le théorème d'inversion de la transformée de Fourier.*

1. *Montrer que* $\phi_Y(t) = \frac{2}{\pi}\int_0^\infty \frac{\cos(tu)}{u^2+1}\,du$.

   Soit $\theta > 0$, *on pose* $\phi_\theta(t) = \frac{2}{\pi}\int_0^\theta \frac{\cos(tu)}{u^2+1}\,du$.
2. *Calculer* $\phi_\theta(0)$ *et* $\phi'_\theta(0)$.
3. *Montrer que* ϕ_θ *satisfait l'équation différentielle*

$$\phi''_\theta(t) - \phi_\theta(t) = -\frac{2}{\pi}\frac{\sin(t\theta)}{t}.$$

4. *Résoudre et conclure* $\phi_Y(t) = e^{-|t|}$.

Chapitre 5

Lois des grands nombres.

On va s'intéresser au comportement asymptotique d'une suite de variables aléatoires. On se donne un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et pour chaque entier $n \in \mathbb{N}$ une variable aléatoire $X_n : \Omega \to \mathbb{R}$.

Définition 5.0.1. *La suite de v.a. $(X_n)_n$ est dite identiquement distribuée si tous les X_n ont même loi, i.e. pour tout borélien $A \subset \mathbb{R}$ et $\forall n, m$*

$$\mathbb{P}(X_n \in A) = \mathbb{P}(X_m \in A).$$

Il est équivalent de dire que pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ borélienne ou intégrable par rapport à $\mathbb{P}_{X_1}$ on a

$$\mathbb{E}[f(X_n)] = \mathbb{E}[f(X_m)].$$

En particulier, $\mathbb{E}[X_n] = \mathbb{E}[X_m]$ si les $(X_n)_n$ sont intégrables et $\mathbb{E}[X_n^2] = \mathbb{E}[X_m^2]$ si les $(X_n)_n$ sont de carrés intégrables.

$(\Omega, \mathcal{T}, \mathbb{P})$

5.1 Loi faible des grands nombres

Soit $(X_n)_n$ une suite de v.a. indépendantes entre elles et identiquement distribuées (v.a. i.i.d.). On pose

$$S_n := \sum_{i=1}^{n} X_i = X_1 + \cdots + X_n.$$

Pour $\omega \in \Omega$, la quantité $S_n(\omega)/n$ est la moyenne empirique calculée sur l'échantillon donné par le résultat $\omega \in \Omega$. On cherche à étudier le comportement asymptotique de la moyenne S_n/n.

Théorème 5.1.1. Loi faible des grands nombres. *Soit $(X_n)_n$ une suite de v.a. indépendantes entre elles, identiquement distribuées et de carrés intégrables. Alors pour tout $\epsilon > 0$*

$$\lim_{n\to\infty} \mathbb{P}\Big(\Big|\frac{S_n}{n} - \mathbb{E}[X_1]\Big| > \epsilon\Big) = 0.$$

Démonstration. La preuve repose sur les 2 identités suivantes valables pour des v.a. i.i.d.

$$\mathbb{E}[S_n] = \mathbb{E}[X_1 + \cdots + X_n] = n\mathbb{E}[X_1] \tag{5.1}$$
$$\mathrm{Var}(S_n) = n\mathrm{Var}(X_1). \tag{5.2}$$

La première identité résulte de la linéarité de l'espérance alors que la 2e se montre en développant la variance

$$\mathrm{Var}(S_n) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j).$$

Les termes de covariances sont nuls par indépendance, ce qui prouve la relation. En appliquant ces 2 relations on obtient

$$\mathbb{E}[S_n/n] = \mathbb{E}[X_1] \quad \text{et} \quad \mathrm{Var}(S_n/n) = \mathrm{Var}(X_1)/n.$$

En appliquant alors l'inégalité de Bienaymé-Tchebichev

$$\mathbb{P}\big(|\frac{S_n}{n} - \mathbb{E}[X_1]| > \epsilon\big) \leq \epsilon^{-2}\mathrm{Var}(S_n/n) = \epsilon^{-2}\mathrm{Var}(X_1)/n.$$

□

Exercice 25. *Soit $(X_n)_n$ une suite de variables aléatoires identiquement distribuées et de carrés intégrables telle que*

$$\forall n, m \geq 1, \quad \mathrm{Cov}(X_n, X_m) = f(n-m)$$

où $f : \mathbb{Z} \to \mathbb{R}$ est une fonction paire satisfaisant $\lim_{n\to\infty} = f(n) = 0$. On pose

$$S_n = X_1 + \cdots + X_n.$$

1. *Montrer que*
$$\mathbb{E}[S_n^2] = nf(0) + 2\sum_{k=1}^{n}(n-k)f(k).$$
2. *En déduire que*
$$\frac{S_n}{n} \to \mathbb{E}[X_1], \quad n \to \infty \quad \textit{en proba}.$$

5.2 Loi forte des grands nombres

La loi faible des grands nombres repose sur une preuve élémentaire. On va maintenant aborder un résultat plus fort, également plus technique à établir.

Théorème 5.2.1. Loi forte des grands nombres. *Soit $(X_n)_n$ une suite de v.a. indépendantes entre elles, identiquement distribuées et intégrables. Alors la convergence suivante a lieu presque sûrement et dans L^1*

$$\lim_{n\to\infty} \frac{S_n}{n} = \mathbb{E}[X_1].$$

Exemple 5.2.2. Pile ou face. *On lance une pièce de monnaie équilibrée un grand nombre de fois de manière indépendante. Pour modéliser ces épreuves, on commence par définir la probabilité* $\tilde{P}$ *sur* $\{0,1\}$ *(muni de la tribu* $\tilde{\mathcal{T}}$ *de toutes les sous-parties) par* $\tilde{P}(\{0\}) = \tilde{P}(\{1\}) = \frac{1}{2}$. *On pose ensuite*

$$\Omega = \{0,1\}^{\mathbb{N}}, \quad \mathcal{T} = \tilde{\mathcal{T}}^{\otimes \mathbb{N}}, \quad \mathbb{P} = \tilde{\mathbb{P}}^{\otimes \mathbb{N}}.$$

On définit alors une suite de v.a. par

$$\forall n \in \mathbb{N}, \forall \omega \in \Omega, \quad X_n(\omega) = \omega_n \in \{0,1\},$$

qui représentent le résultat du lancer de pièce au n-ième jet. Alors S_n/n *représente la fréquence d'apparition de "pile" au cours des n premiers lancers. Les variables aléatoires* $(X_n)_n$ *sont indépendantes et identiquement distribuées suivant la loi de Bernoulli de paramètre* $\mathcal{B}(1/2)$. *Ainsi la loi des grands nombres nous dit que* $S_n/n \to 1/2$ *quand* $n \to \infty$ $\mathbb{P}$*-presque sûrement.*

Démonstration du Théorème 5.2.1. Tout d'abord, on remarque que, quitte à remplacer X_n par $X_n - \mathbb{E}[X_n]$, on peut supposer que les $(X_n)_n$ sont centrés, i.e. $\mathbb{E}[X_n] = 0$.

Nous allons démontrer le résultat en utilisant les lemmes suivants.

Lemme 5.2.3 (Lemme de troncature de Kolmogorov)**.** *Soient* $(X_n)_n$ *des v.a. i.i.d. intégrables de loi commune* X. *Soit* $\mu = \mathbb{E}[X]$. *On définit*

$$Y_n = X_n \mathbf{1}_{|X_n| \leq n}.$$

Alors

i) $\mathbb{E}[Y_n] \to \mu$ *quand* $n \to \infty$

ii) $\mathbb{P}(X_n = Y_n$ *pour tous n sauf un nombre fini d'indices*$) = 1$

iii) $\sum_n \mathrm{Var}(Y_n)/n^2 < +\infty$

Lemme 5.2.4 (Lemme de Kronecker)**.** *Soient* $(b_n)_n$ *une suite de réels strictement positifs croissante et tendant vers l'infini. Soit* $(x_n)_n$ *une suite de réels et définissons* $s_n = x_1 + \cdots + x_n$. *On a alors :*

$$\sum_n \frac{x_n}{b_n} \text{ converge } \Rightarrow \frac{s_n}{b_n} \to 0 \quad \text{quand } n \to \infty.$$

Lemme 5.2.5 (Inégalité de Kolmogorov)**.** *Soient* $X_1, \ldots, X_n$ *des v.a. indépendantes centrées. On pose* $S_n = X_1 + \cdots + X_n$ *et* σ_n^2 *la variance de* S_n. *On a l'inégalité suivante :*

$$\mathbb{P}(\exists k \in \{1, \ldots, n\}, |S_k| \geq t\sigma_n) \leq t^{-2}.$$

Admettons ces 2 lemmes pour démontrer la loi des grands nombres. D'après ii, on a presque sûrement

$$\lim_{n \to \infty} \Big(\frac{1}{n} \sum_{k=1}^n Y_k - \frac{1}{n} \sum_{k=1}^n X_k \Big) = 0.$$

Il reste donc à montrer que

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^n Y_k = \mu.$$

Posons $W_n = Y_n - \mathbb{E}[Y_n]$ de sorte que

$$\frac{1}{n}\sum_{k=1}^{n} Y_k = \frac{1}{n}\sum_{k=1}^{n} W_k + \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}[Y_k]$$

D'après i, $\mathbb{E}[Y_n] \to \mu$ ce qui implique $\frac{1}{n}\sum_{k=1}^{n} \mathbb{E}[Y_k] \to \mu$ d'après le lemme de Césaro.

Soit $\tilde{W}_n = \frac{W_n}{n}$ et $\tilde{S}_n = \sum_{k=1}^{n} \tilde{W}_k$. En utilisant le lemme de Kronecker, on a

$$\tilde{S} = \sum_n \frac{W_n}{n} \text{ cv } \Rightarrow \frac{1}{n}\sum_{k=1}^{n} W_k \text{ cv}.$$

Il suffit donc d'étudier la convergence presque sûre de $\tilde{S}_n$. Pour cela il suffit de montrer que $\forall \epsilon > 0$

$$\lim_{n_0\to\infty} \mathbb{P}\Big(\exists m, n \geq n_0, |\tilde{S}_m - \tilde{S}_n| > \epsilon\Big) = 0. \tag{5.3}$$

En effet, en utilisant le critère de Cauchy, on peut majorer la probabilité de l'ensemble sur lequel $\tilde{S}_n$ diverge par

$$\mathbb{P}\Big(\exists \epsilon > 0, \forall n_0, \exists m, n \geq n_0, |\tilde{S}_m - \tilde{S}_n| > \epsilon\Big) \leq \sum_{\epsilon \in Q^+} \mathbb{P}\Big(\forall n_0, \exists m, n \geq n_0, |\tilde{S}_m - \tilde{S}_n| > \epsilon\Big).$$

Le membre de droite est nul par convergence monotone.

Pour montrer (5.4), il suffit de montrer par l'inégalité triangulaire que

$$\lim_{n_0\to\infty}\lim_{m\to\infty} \mathbb{P}\Big(\exists k \in \{n_0, \dots, m\}, |\tilde{S}_k - \tilde{S}_{n_0}| > \epsilon/2\Big) = 0. \tag{5.4}$$

L'inégalité de Kolmogorov donne alors

$$\mathbb{P}\Big(\exists k \in \{n_0, \dots, m\}, |\tilde{S}_k - \tilde{S}_{n_0}| > \epsilon/2\Big) \leq 4\epsilon^{-2}\mathrm{Var}(\tilde{W}_{n_0+1} + \cdots + \tilde{W}_m).$$

En raison de l'indépendance des $\tilde{W}_n$ le membre de droite s'écrit

$$\sum_{k=n_0+1}^{m} k^{-2}\mathrm{Var}(W_k) = \sum_{k=n_0+1}^{m} k^{-2}\mathrm{Var}(Y_k),$$

donc si l'on prend d'abord la limite $m \to \infty$ puis $n_0 \to \infty$, on obtient le résultat. □

Preuve du lemme 5.2.3. Tout d'abord, $\mathbb{E}[Y_n] = \mathbb{E}[X\mathbf{1}_{|X|\leq n}] \to \mu$ par convergence dominée. Ensuite, posons $A_n = \{X_n = Y_n\}$ pour obtenir

$$\sum_n \mathbb{P}(A_n^c) = \sum_n \mathbb{P}(|X| \geq n) \leq \mathbb{E}\big[\sum_n \mathbf{1}_{n\leq|X|}\big] \leq \mathbb{E}\big[\int_0^\infty \mathbf{1}_{t\leq|X|}\,dt\big] = \mathbb{E}[|X|].$$

Comme l'espérance $\mathbb{E}[|X|]$ est finie, le lemme de Borel-Cantelli assure que

$$\mathbb{P}\Big(\limsup_n A_n^c\Big) = 0,$$

ce qui donne ii. Finalement

$$\sum_n \mathrm{Var}(Y_n)/n^2 \leq \sum_n \mathbb{E}[X^2 \mathbf{1}_{|X|\leq n}]/n^2 = \mathbb{E}[X^2 f(|X|)]$$

où

$$f(z) = \sum_{n\geq z} \frac{1}{n^2}.$$

Une comparaison série-intégrale donne rapidement une borne $f(z) \leq C \min(1, 1/z)$, ce qui assure que $\mathbb{E}[X^2 f(|X|)] < +\infty$. □

Preuve du lemme 5.2.4. Rappelons tout d'abord le lemme de Césaro généralisé qui affirme que si $(u_n)_n$ est une suite convergente vers u_∞ et si $(b_n)_n$ est comme dans le lemme alors

$$\frac{1}{b_n} \sum_{k=1}^{n} (b_k - b_{k-1}) u_k \to u_\infty.$$

Posons $u_n = \sum_{k=1}^n \frac{x_k}{b_k}$, qui converge. On a $u_n - u_{n-1} = \frac{x_n}{b_n}$. Remarquons alors que

$$s_n = \sum_{k=1}^{n} b_k (u_k - u_{k-1}) = b_n u_n - b_1 u_0 - \sum_{k=1}^{n-1} u_k (b_{k+1} - b_k).$$

En utilisant le lemme de Césaro généralisé, on en déduit que

$$\frac{s_n}{b_n} \to u_\infty - u_\infty = 0.$$

□

Preuve du lemme 5.2.5. Pour $\nu \in \{1, \ldots, n\}$ on pose $Z_\mu = \mathbf{1}_{\{S_\nu \geq t\sigma_n\}} \mathbf{1}_{\{S_k < t\sigma_n \,|\, k=1,\ldots,\nu-1\}}$ et l'on remarque que

$$Z_1 + \cdots + Z_n = \mathbf{1}_{\{\exists k \in \{1,\ldots,n\}, |S_k| \geq t\sigma_n\}}.$$

En multipliant par S_n^2 et en prenant l'espérance on en déduit que

$$\sum_{k=1}^{n} \mathbb{E}[Z_k S_n^2] \leq \mathbb{E}[S_n^2] = s_n^2.$$

Maintenant, si l'on pose $U_k = S_n - S_k$, alors U_k est indépendante de $Z_k S_k$ et donc

$$\mathbb{E}[Z_k S_n^2] = \mathbb{E}[Z_k (U_k + S_k)^2] = \mathbb{E}[Z_k U_k^2] + \mathbb{E}[Z_k S_k^2] \geq \mathbb{E}[Z_k S_k^2].$$

Or, $Z_k S_k^2 \geq Z_k t^2 \sigma_n^2$ donc

$$\begin{aligned}
\mathbb{P}(\exists k \in \{1, \ldots, n\}, |S_k| \geq t\sigma_n) \leq& \mathbb{E}[Z_1 + \cdots + Z_n] \\
\leq& (\sigma_n t)^{-2} \sum_{k=1}^{n} \mathbb{E}[Z_k S_k^2] \\
\leq& (\sigma_n t)^{-2} \sum_{k=1}^{n} \mathbb{E}[Z_k S_n^2] \\
\leq& (\sigma_n t)^{-2} \mathbb{E}[S_n^2] = t^{-2},
\end{aligned}$$

ce qui conclut la preuve. □

Corollaire 5.2.6. *Soit $(X_n)_n$ une suite de v.a. indépendantes entre elles et identiquement distribuées et soit A un borélien de $\mathbb{R}$. Alors*

$$\lim_{n\to\infty} \frac{1}{n}\#\big\{i \leq n \,|\, X_i \in A\big\} = \mathbb{P}(X_1 \in A).$$

Exercice 26. *Soit $(X_n)_n$ une suite de v.a. i.i.d. possédant des moments d'ordre 4 finis. On veut établir la loi des grands nombres presque sûrement et dans L^4.*

1. *Montrer qu'il suffit de considérer le cas de variables aléatoires centrées. On supposera dans la suite que c'est le cas.*
2. *Montrer qu'il existe une constante $C > 0$ telle que $\mathbb{E}[S_n^4] \leq Cn^2$.*
3. *En déduire que $\sum_n \mathbb{E}[(S_n/n)^4] < +\infty$ et conclure.*
4. *En déduire que si $(X_n)_n$ une suite de v.a. i.i.d. intégrables alors $S_n/n \to \mathbb{E}[X_1]$ dans L^1.*

Exercice 27. *Soit $(X_n)_n$ une suite de v.a. i.i.d. suivant la loi de Cauchy $\mathcal{C}(0,1)$.*

1. *Montrer que ces variables ne sont pas intégrables.*
2. *Calculer la loi de $(X_1 + \cdots + X_n)/n$.*
3. *En déduire que $(X_1+\cdots+X_n)/n$ ne converge pas presque sûrement vers une limite $a \in \mathbb{R}$..*

Chapitre 6

Théorème central limite.

Théorème 6.0.1 (TCL). *Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé, $(X_n)_n$ une suite de v.a. i.i.d. de carrés intégrables et de variance non nulle. On pose $S_n = \sum_{k=1}^n X_k$ et $\sigma^2 = \mathrm{Var}(X_1)$. Alors la suite de variables aléatoires*

$$\frac{1}{\sigma\sqrt{n}}\Big(S_n - n\mathbb{E}[X_1]\Big)$$

converge en loi vers une loi normale centrée réduite. En particulier pour tout intervalle $[a,b] \subset \mathbb{R}$

$$\mathbb{P}\Big(a \leq \frac{1}{\sigma\sqrt{n}}\big(S_n - n\mathbb{E}[X_1]\big) \leq b\Big) \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}\,dx.$$

L'énoncé ci-dessus indique que pour n grand la probabilité que $\frac{S_n}{n}$ soit dans l'intervalle

$$\Big[\mathbb{E}[X_1] - t\frac{\sigma}{\sqrt{n}}, \mathbb{E}[X_1] + t\frac{\sigma}{\sqrt{n}}\Big]$$

est proche de $\frac{1}{\sqrt{2\pi}}\int_{-t}^t e^{-\frac{x^2}{2}}\,dx$. Pour $t = 1,96$ cette proba vaut environ 95%.

Démonstration. Quitte à remplacer X_k par $X_k - \mathbb{E}[X_k]$ on peut supposer les $(X_n)_n$ centrées et, quitte aussi à diviser par σ, on peut aussi supposer que $\mathrm{Var}(X_1) = 1$. Donc on veut montrer que $S_n/\sqrt{n}$ converge vers la loi normale. On va pour cela passer par les fonctions caractéristiques. Calculons

$$\begin{aligned}
\phi_{S_n/\sqrt{n}}(t) =& \mathbb{E}[e^{itS_n/\sqrt{n}}] \\
=& \mathbb{E}[e^{it/\sqrt{n}(X_1+\cdots+X_n)}] \\
=& \mathbb{E}[\prod_{k=1}^n e^{it/\sqrt{n}X_k}] \\
=& \prod_{k=1}^n \mathbb{E}[e^{it/\sqrt{n}X_k}] \\
=& \mathbb{E}[e^{it/\sqrt{n}X_1}]^n
\end{aligned}$$

Ci-dessus on a utilisé tour à tour l'indépendance et l'équi-distribution des $(X_n)_n$. Comme X_1 est de carré intégrables on a le DL

$$\phi_{X_1}(t) = 1 - \frac{t^2}{2n} + o(\frac{t^2}{2n}).$$

On utilise alors le lemme suivant

Lemme 6.0.2. *Soit $(z_n)_n$ une suite de nombres complexes convergeant vers z. Alors*

$$(1 + \frac{z_n}{n})^n \to e^z.$$

Ce lemme permet de montrer que $\phi_{S_n/\sqrt{n}}(t) \to e^{-t^2/2}$ lorsque $n \to \infty$, qui est la fonction caractéristique d'une loi gaussienne centrée réduite, d'où le résultat. □

Preuve du lemme. En développant en série

$$e^z - (1 + \frac{z}{n})^n = \sum_{k=0}^{\infty} a_n^k z^k$$

avec

$$a_n^k = \frac{1}{k!}\Big(1 - \frac{n(n-1)\dots(n-k+1)}{n^k}\delta_{k\leq n}\Big).$$

D'où

$$\begin{aligned}\big|e^z - (1 + \frac{z}{n})^n\big| \leq &\sum_{k=0}^{\infty} a_n^k |z|^k \\ =&e^{|z|} - (1 + \frac{|z|}{n})^n \\ \to&0\end{aligned}$$

pour $n \to \infty$. Cette dernière convergence s'obtient par exemple en écrivant $(1+\frac{|z|}{n})^n = e^{n\ln(1+\frac{|z|}{n})}$ et en faisant un DL du $\ln$. □

6.1 Exercices

Exercice 28. *On cherche à estimer*

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

où f est une fonction intégrable sur $[0,1]$ pour la mesure de Lebesgue. On considère une suite de v.a. i.i.d. $(U_n)_n$ de loi uniforme sur $[0,1]$ et on pose

$$I_n = \frac{1}{n}\sum_{k=1}^{n} f(U_k).$$

1. *Montrer que la suite $(I_n)_n$ converge presque sûrement. Quelle est sa limite ?*

2. *On pose* $\sigma^2 = \int_{[0,1]} f^2(x)\,dx - I^2$. *On suppose* n *grand. Donner un intervalle de confiance à 95% pour* I *en fonction de* n, I_n, σ.

Exercice 29. *Utiliser le TCL pour déterminer la limite quand* $n \to \infty$ *de*

$$e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!}$$

Exercice 30. ***Erreurs d'arrondis*** *Dans un programme de calcul, l'opérateur décide d'utiliser J chiffres significatifs après la virgule et d'arrondir tous les résultats d'opérations à cette configuration (donc à* $10^{-J}/2$ *près). On suppose qu'il effectue* 10^6 *opérations successives, que les erreurs commises sur chacune sont indépendantes, de loi uniforme sur* $[-10^{-J}/2, 10^{-J}/2]$*, et que l'erreur sur le résultat final est la somme des erreurs commises sur chaque opération. Calculer la probabilité pour que l'erreur finale soit inférieure ou égale à* $10^{-J+3}/2$.

Annexe A

Corrections des exercices

A.1 Variables aléatoires

Exercice ??.

A.2 Chapitre Variables aléatoires

Exercice ??.

A.3 Convergence de suites de variables aléatoires

Exercice 16. Soit $A_n^\epsilon = \{|Y_n| > \epsilon\}$. Le lemme de Borel-Cantelli assure que

$$\mathbb{P}(\limsup_n A_n^\epsilon) = 0.$$

Posons $C^\epsilon = (\limsup_n A_n^\epsilon)^c$ pour chaque $\epsilon > 0$ et

$$S = \bigcap_{p \in \mathbb{N}^*} C^{1/p}.$$

Donc $\mathbb{P}(S) = 1$. Or tout $\omega \in S$ vérifie

$$\forall p \in \mathbb{N}^*, \exists N, \forall n \geq N, \quad |Y_n| \leq 1/p,$$

ce qui traduit la convergence simple de $(Y_n)_n$ converge vers 0 sur S.

A.4 Fonctions génératrices, fonctions caractéristiques

Exercice 21. Calculer la fonction caractéristique de la v.a. X dans les cas suivants :

1. X suit une loi normale $\mathcal{N}(m, \sigma^2)$
2. X est une variable discrète de loi uniforme sur $\{1, \dots, n\}$.
3. X suit une loi uniforme sur l'intervalle $[a, b]$ avec $a < b$.
4. X suit une loi exponentielle de paramètre $\lambda > 0$.

Exercice 24.

1. Résulte d'un simple changement de variables.
2. on a $\phi_\theta(0) = \frac{2}{\pi} \arctan \theta$ et $\phi'_\theta(0) = 0$.
3. Le théorème de dérivation sous le signe intégral mène directement à l'équation différentielle demandée.
4. Pour résoudre, on utilise la méthode de variation des constantes. L'équation homogène a pour solutions
$$y_0(t) = A \cosh t + B \sinh t.$$
La variation des constantes amène à chercher 2 fonctions A, B telles que
$$\phi_\theta(t) = A(t) \cosh t + B(t) \sinh t.$$
En dérivant
$$\phi'_\theta(t) = A'(t) \cosh t + B'(t) \sinh t + A(t) \sinh t + B(t) \cosh t.$$
On impose la condition
$$A'(t) \cosh t + B'(t) \sinh t = 0$$
pour obtenir
$$\phi'_\theta(t) = A(t) \sinh t + B(t) \cosh t,$$
d'où
$$\phi''_\theta(t) = A'(t) \sinh t + B'(t) \cosh t + A(t) \cosh t + B(t) \sinh t,$$
et
$$\phi''_\theta(t) - \phi_\theta(t) = A'(t) \sinh t + B'(t) \cosh t.$$
On a donc un système
$$A'(t) \cosh t + B'(t) \sinh t = 0 \quad \text{et} \quad A'(t) \sinh t + B'(t) \cosh t = -\frac{2}{\pi} \frac{\sin(\theta t)}{t},$$
et l'on trouve
$$A'(t) = \frac{2}{\pi} \frac{\sin(\theta t)}{t} \sinh(t) \quad \text{et} \quad B'(t) = -\frac{2}{\pi} \frac{\sin(\theta t)}{t} \cosh(t).$$
Donc
$$A(t) = \frac{2}{\pi} \int_0^t \frac{\sin(\theta u)}{u} \sinh(u)\, du + a \quad \text{et} \quad B(t) = -\frac{2}{\pi} \int_0^t \frac{\sin(\theta u)}{u} \cosh(u)\, du + b.$$
Ceci donne
$$\phi_\theta(t) = \frac{2}{\pi} \Big(\cosh t \int_0^t \frac{\sin(\theta u)}{u} \sinh(u)\, du - \sinh t \int_0^t \frac{\sin(\theta u)}{u} \cosh(u)\, du \Big) + a \cosh t + b \sinh t.$$

En utilisant les conditions initiales on trouve $a = \frac{2}{\pi}\arctan\theta$ et $b = 0$.
Il reste alors à démontrer le résultat suivant : si $f : \mathbb{R}_+ \to \mathbb{R}$ est telle qu'il existe g de classe C^1 et $\lambda \in \mathbb{R}$ t.q. $f(x) = \lambda x + x^2 g(x)$ alors pour tout $t > 0$

$$\lim_{\theta\to\infty} \int_0^{t\theta} \frac{\sin u}{u} f(u/\theta)\, du = 0.$$

Grâce à lui, on peut passer à la limite quand $\theta \to \infty$ pour obtenir pour $t \geq 0$

$$\phi_Y(t) = \cosh t - \sinh(t) = e^{-t}.$$

□

A.5 Loi des grands nombres

Exercice 26. Supposons que les v.a. $(X_n)_n$ admettent des moments d'ordre 4. L'inégalité de Markov entraîne alors

$$\mathbb{P}(|S_n/n| > \epsilon) \leq (n\epsilon)^{-4}\mathbb{E}[S_n^4].$$

Le point crucial est d'obtenir une estimation de $\mathbb{E}[S_n^4]$. En développant

$$\mathbb{E}[S_n^4] = \sum_{i_1,i_2,i_3,i_4} \mathbb{E}[X_{i_1}X_{i_2}X_{i_3}X_{i_4}].$$

Il y a plusieurs cas possibles

- tous les indices égaux : l'espérance vaut alors $\mathbb{E}[X_1^4]$ et il y a n termes de ce type. Contribution dans la somme $n\mathbb{E}[X_1^4]$.
- au moins un indice est différent de tous les autres : par indépendance, l'espérance vaut alors 0.
- 2 paires d'indices égaux : l'espérance vaut alors $\mathbb{E}[X_1^2]^2$ et il y a $3n(n-1)$ termes de ce type.

Au final on voit que l'on peut borner

$$\mathbb{E}[S_n^4] \leq n\mathbb{E}[X_1^4] + 3n^2\mathbb{E}[X_1^2]^2 \quad \text{d'où} \quad \mathbb{P}(|S_n/n| > \epsilon) \leq C\epsilon^{-4}n^{-2}.$$

Comme la série $\sum_n \frac{1}{n^2}$ cv on peut appliquer le lemme pour la convergence presque sûre. L'estimée ci-dessus montre également la convergence dans L^4, donc dans L^1. □

Exercice 27. Soit $(X_n)_n$ une suite de v.a. i.i.d. suivant la loi de Cauchy $\mathcal{C}(0, 1)$.

1. Utiliser le critère de Riemann.
2. En utilisant la fonction caractéristique de la loi de Cauchy, on montre que $(X_1+\cdots+X_n)/n$ suit encore une loi de Cauchy $\mathcal{C}(0, 1)$.
3. Si S_n/n converge p.s. vers a alors il converge également en loi vers a. Impossible vu la Q précédente.

A.6 Théorème de la limite centrale

Exercice ??.

Bibliographie

[1] Billingsley P. : Probability and measure. John Wiley & Sons Inc., New York, third edition.