Intégration et Probabilités

Cours de Adrien Kassel
Notes de Alexis Marchand

ENS de Lyon
S2 2017-2018
Niveau L3

Table des matières

1 Espaces de probabilité et variables aléatoires

1.1 Rappels de théorie de la mesure – espaces mesurables et espaces mesurés

Définition 1.1.1 (Espace mesurable)**.** *Un* espace mesurable $(E, \mathcal{E})$ *est la donnée d'un ensemble E et d'un $\mathcal{E} \subset \mathcal{P}(E)$ t.q.*

(i) $E \in \mathcal{E}$.

(ii) $\forall A \in \mathcal{E},\ E \backslash A \in \mathcal{E}$.

(iii) $\forall (A_n)_{n \in \mathbb{N}} \in \mathcal{E}^{\mathbb{N}},\ \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{E}$.

*On dit que $\mathcal{E}$ est une σ-*algèbre *(ou* tribu*) sur E. Les éléments de $\mathcal{E}$ sont appelés* ensembles mesurables*.*

Définition 1.1.2 (Tribu engendrée)**.** *Soit E un ensemble et $C \subset \mathcal{P}(E)$. La tribu :*

$$\sigma(C) = \bigcap_{\substack{\mathcal{A} \text{ tribu sur } E \\ C \subset \mathcal{A}}} \mathcal{A}$$

est appelée tribu engendrée *par C. C'est la plus petite tribu sur E qui contient C.*

Définition 1.1.3 (Tribu borélienne)**.** *Soit X un espace topologique et $\mathcal{T}$ la topologie de X. La tribu $\sigma(\mathcal{T})$ est appelée* tribu borélienne *de X et notée* $\mathrm{Bor}(X)$*.*

Définition 1.1.4 (Classe monotone)**.** *Soit E un ensemble. On appelle* classe monotone *sur E tout $\mathcal{M} \subset \mathcal{P}(E)$ t.q.*

(i) $E \in \mathcal{M}$.

(ii) $\forall (A, B) \in \mathcal{M}^2,\ A \subset B \Longrightarrow (B \backslash A) \in \mathcal{M}$.

(iii) $\forall (A_n)_{n \in \mathbb{N}} \in \mathcal{M}^{\mathbb{N}},\ (A_n)_{n \in \mathbb{N}}$ *croissante* $\Longrightarrow \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{M}$.

Remarque 1.1.5. *Toute tribu est une classe monotone.*

Définition 1.1.6 (Classe monotone engendrée)**.** *Soit E un ensemble et $C \subset \mathcal{P}(E)$. La classe monotone :*

$$\mathcal{M}(C) = \bigcap_{\substack{\mathfrak{M} \text{ classe monotone sur } E \\ C \subset \mathfrak{M}}} \mathfrak{M}$$

est appelée classe monotone engendrée *par C. C'est la plus petite classe monotone sur E qui contient C.*

Théorème 1.1.7 (Lemme de classe monotone)**.** *Soit E un ensemble et $C \subset \mathcal{P}(E)$. On suppose que C est stable par intersections finies. Alors $\mathcal{M}(C) = \sigma(C)$.*

Définition 1.1.8 (Mesure positive)**.** *Soit $(E, \mathcal{A})$ un espace mesurable. On appelle* mesure positive *sur $(E, \mathcal{A})$ toute application $\mu : \mathcal{A} \to [0, +\infty]$ vérifiant :*

(i) $\mu(\varnothing) = 0$.

(ii) *Pour toute famille $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{A}$ deux à deux disjoints, on a $\mu\left(\bigsqcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n)$.*

On dit alors que $(E, \mathcal{A}, \mu)$ est un espace mesuré.

Proposition 1.1.9. *Soit $(E, \mathcal{A}, \mu)$ un espace mesuré.*

(i) $\forall (A, B) \in \mathcal{A}^2,\ A \subset B \Longrightarrow \mu(A) \leqslant \mu(B)$.

(ii) $\forall (A, B) \in \mathcal{A}^2,\ \mu(A) + \mu(B) = \mu(A \cup B) + \mu(A \cap B)$.

(iii) *Soit $(A_n)_{n\in\mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ une suite croissante pour l'inclusion. Alors :*

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to+\infty} \mu(A_n).$$

(iv) *Soit $(A_n)_{n\in\mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ une suite décroissante pour l'inclusion. Alors :*

$$\mu(A_0) < +\infty \Longrightarrow \mu\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\to+\infty} \mu(A_n).$$

(v) $\forall (A_n)_{n\in\mathbb{N}} \in \mathcal{A}^{\mathbb{N}},\ \mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leqslant \sum_{n\in\mathbb{N}} \mu(A_n)$.

Proposition 1.1.10. *Soit $(E, \mathcal{A})$ un espace mesurable. Soit μ et ν deux mesures sur $(E, \mathcal{A})$. On suppose qu'il existe une classe $C \subset \mathcal{A}$ stable par intersections finies t.q. $\sigma(C) = \mathcal{A}$ et $\forall A \in C, \mu(A) = \nu(A)$.*

(i) *Si $\mu(E) = \nu(E) < +\infty$, alors $\mu = \nu$.*

(ii) *S'il existe une suite $(E_n)_{n\in\mathbb{N}} \in C^{\mathbb{N}}$ croissante pour l'inclusion t.q. $E = \bigcup_{n\in\mathbb{N}} E_n$ et $\forall n \in \mathbb{N},\ \mu(E_n) = \nu(E_n) < +\infty$, alors $\mu = \nu$.*

Démonstration. Utiliser le lemme de classe monotone (théorème 1.1.7). □

Définition 1.1.11. *Soit $(E, \mathcal{A}, \mu)$ un espace mesuré.*

(i) *On dit que μ est une* mesure de probabilité *lorsque $\mu(E) = 1$.*

(ii) *On dit que μ est* σ-finie *lorsqu'il existe une suite $(E_n)_{n\in\mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ croissante pour l'inclusion t.q. $E = \bigcup_{n\in\mathbb{N}} E_n$ et $\forall n \in \mathbb{N},\ \mu(E_n) < +\infty$.*

(iii) *Un élément $x \in E$ est dit un* atome *de μ lorsque $\{x\} \in \mathcal{A}$ et $\mu(\{x\}) > 0$.*

(iv) *On dit que μ est* diffuse *lorsque μ n'a pas d'atome.*

1.2 Espaces de probabilité

Définition 1.2.1 (Espace de probabilité)**.** *Un* espace de probabilité *est un espace mesuré $(\Omega, \mathcal{A}, \mathbb{P})$, où $\mathbb{P}$ est une mesure de probabilité.*

(i) *Ω est l'espace des éventualités, il est appelé l'*univers.

(ii) *$\mathcal{A}$ est l'ensemble des* événements.

Remarque 1.2.2. *Dans une situation donnée, l'espace de probabilité n'est pas unique ni canonique. Souvent, il sera omis. Les objets importants sont les* variables aléatoires.

Exemple 1.2.3. *Il existe une unique mesure de probabilité $\mathbb{P}$ sur $\left(\{0,1\}^{\mathbb{N}}, \mathrm{Bor}\left(\{0,1\}^{\mathbb{N}}\right)\right)$ t.q.*

$$\forall \ell \in \mathbb{N},\ \forall s \in \{0,1\}^{\ell},\ \mathbb{P}(C_s) = \frac{1}{2^{\ell}},$$

où $C_s = \left\{\omega \in \{0,1\}^{\mathbb{N}},\ \forall i \in [\![0, \ell-1]\!],\ \omega_i = s_i\right\}$ pour $\ell \in \mathbb{N}$ et $s = (s_0, \ldots, s_{\ell-1}) \in \{0,1\}^{\ell}$.

1.3 Rappels de théorie de la mesure – fonctions mesurables

Définition 1.3.1 (Fonction mesurable)**.** *Soit* $(E, \mathcal{E})$ *et* $(F, \mathcal{F})$ *deux espaces mesurables. Une application* $f : E \to F$ *est dite* mesurable *lorsque :*

$$\forall A \in \mathcal{F},\ f^{-1}(A) \in \mathcal{E}.$$

Proposition 1.3.2.

(i) *Une composée de fonctions mesurables est mesurable.*

(ii) *Si* $(E, \mathcal{E})$ *et* $(F, \mathcal{F})$ *sont deux espaces mesurables,* $C \subset \mathcal{P}(F)$ *est t.q.* $\sigma(C) = \mathcal{F}$*, alors une application* $f : E \to F$ *est mesurable dès que* $\forall A \in C,\ f^{-1}(A) \in \mathcal{E}$*.*

(iii) *Si* X *et* Y *sont deux espaces topologiques munis de leurs tribus boréliennes respectives, alors toute application continue* $f : X \to Y$ *est mesurable.*

Proposition 1.3.3. *Soit* $(E, \mathcal{E})$ *un espace mesurable.*

(i) *Si* $f, g : E \to \mathbb{R}$ *sont mesurables, alors* $(f+g)$*,* $\max(f, g)$*,* $\min(f, g)$*,* f^+ *et* f^- *sont mesurables.*

(ii) *Si* $(f_n)_{n\in\mathbb{N}} \in \left(\mathbb{R}^E\right)^{\mathbb{N}}$ *est une suite de fonctions mesurables, alors* $\sup f_n$*,* $\inf f_n$*,* $\limsup_{n\to+\infty} f_n$ *et* $\liminf_{n\to+\infty} f_n$ *sont mesurables.*

(iii) *Si* $(f_n)_{n\in\mathbb{N}} \in \left(\mathbb{R}^E\right)^{\mathbb{N}}$ *est une suite de fonctions mesurables convergeant simplement vers* $f : E \to \mathbb{R}$*, alors* f *est mesurable.*

(iv) *Si* $(f_n)_{n\in\mathbb{N}} \in \left(\mathbb{R}^E\right)^{\mathbb{N}}$ *est une suite de fonctions mesurables, alors* $\left\{x \in E,\ (f_n(x))_{n\in\mathbb{N}} \text{ converge}\right\}$ *est mesurable.*

Définition 1.3.4 (Mesure image)**.** *Soit* $(E, \mathcal{E})$ *et* $(F, \mathcal{F})$ *deux espaces mesurables,* $f : E \to F$ *une fonction mesurable. Soit* μ *une mesure positive sur* $(E, \mathcal{E})$*. La* mesure image *de* μ *par* f*, notée* $f(\mu)$ *ou* $f_*\mu$*, est définie par :*

$$f_*\mu : B \in \mathcal{F} \longmapsto \mu\left(f^{-1}(B)\right) \in [0, +\infty].$$

C'est bien une mesure positive.

Définition 1.3.5 (Fonction étagée)**.** *Soit* $(E, \mathcal{E})$ *un espace mesurable. On appelle* fonction étagée *sur* E *toute application* $\varphi : E \to \mathbb{R}$ *mesurable t.q.* $\varphi(E)$ *est fini.*

Proposition 1.3.6. *Soit* $(E, \mathcal{E})$ *un espace mesurable.*

(i) *Toute fonction mesurable positive* $E \to [0, +\infty[$ *est limite simple croissante de fonctions étagées positives.*

(ii) *Toute fonction mesurable* $E \to \mathbb{R}$ *est limite simple de fonctions étagées.*

1.4 Variables aléatoires

Définition 1.4.1 (Variable aléatoire)**.** *Soit* $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace de probabilité et* $(E, \mathcal{E})$ *un espace mesurable. On appelle* variable aléatoire *à valeurs dans* E *toute application mesurable* $X : \Omega \to E$*.*

Définition 1.4.2 (Loi d'une variable aléatoire)**.** *Soit* $(\Omega, \mathcal{A}, \mathbb{P})$ *un espace de probabilité,* $(E, \mathcal{E})$ *un espace mesurable et* $X : \Omega \to E$ *une variable aléatoire. On appelle* loi *de* X*, notée* $\mathbb{P}_X$*, la mesure image de* $\mathbb{P}$ *par* X*. Autrement dit :*

$$\forall B \in \mathcal{E},\ \mathbb{P}_X(B) = \mathbb{P}\left(X^{-1}(B)\right).$$

C'est une mesure de probabilités sur $(E, \mathcal{E})$*. Pour* $B \in \mathcal{E}$*, on notera souvent* $\mathbb{P}(X \in B)$ *plutôt que* $\mathbb{P}_X(B)$*.*

Remarque 1.4.3. *Soit* $(E, \mathcal{E}, \mu)$ *un espace de probabilité. Alors il existe un espace de probabilité* $(\Omega, \mathcal{A}, \mathbb{P})$ *et une variable aléatoire* $X : \Omega \to E$ *t.q.* $\mathbb{P}_X = \mu$. *En effet, il suffit de prendre* $(\Omega, \mathcal{A}, \mathbb{P}) = (E, \mathcal{E}, \mu)$ *et* $X = id_E$. *On dit que* X *est la* variable aléatoire canonique *de loi* μ.

Définition 1.4.4 (Loi discrète, loi à densité)**.** *Soit* X *une variable aléatoire à valeurs dans un espace mesurable* $(E, \mathcal{E})$.

(i) *On dit que* X *est* discrète *lorsque* E *est dénombrable et* $\mathcal{E} = \mathcal{P}(E)$. *On a alors :*

$$\forall A \in \mathcal{P}(E),\ \mathbb{P}(X \in A) = \sum_{a \in A} \mathbb{P}(X = a).$$

(ii) *On dit que* X *est* à densité *lorsque* $(E, \mathcal{E}) = (\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$, *et* $\mathbb{P}_X$ *admet une densité* f *par rapport à la mesure de Lebesgue* λ. *On a alors :*

$$\forall A \in \mathrm{Bor}(\mathbb{R}),\ \mathbb{P}(X \in A) = \int_A f \ \mathrm{d}\lambda.$$

Exemple 1.4.5 (Lois discrètes)**.**

(i) Loi uniforme $\mathcal{U}(E)$ *sur* E, *avec* E *fini non vide :* $\forall A \in \mathcal{P}(E),\ \mathbb{P}(X \in A) = \frac{|A|}{|E|}$.

(ii) Loi de Bernoulli $\mathcal{B}(p)$ *sur* $\{0,1\}$, *avec* $p \in [0,1]$ *:* $\mathbb{P}(X = 1) = p$ *et* $\mathbb{P}(X = 0) = 1 - p$.

(iii) Loi binomiale $\mathcal{B}(n,p)$ *sur* $\{0, \ldots, n\}$, *avec* $p \in [0,1]$ *et* $n \in \mathbb{N}^*$ *:*

$$\forall k \in \{0, \ldots, n\},\ \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

(iv) Loi géométrique $\mathcal{G}(p)$ *sur* $\mathbb{N}^*$, *avec* $p \in]0,1[$ *:*

$$\forall k \in \mathbb{N}^*,\ \mathbb{P}(X = k) = (1-p)^{k-1} p.$$

(v) Loi de Poisson $\mathcal{P}(\lambda)$ *sur* $\mathbb{N}$, *avec* $\lambda \in \mathbb{R}_+^*$ *:*

$$\forall k \in \mathbb{N},\ \mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

Exemple 1.4.6 (Lois à densité)**.**

(i) Loi uniforme $\mathcal{U}([a,b])$, *avec* $a < b$ *:* $\forall x \in \mathbb{R},\ f(x) = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x)$.

(ii) Loi exponentielle $\mathrm{Exp}(\lambda)$, *avec* $\lambda \in \mathbb{R}_+^*$ *:*

$$\forall x \in \mathbb{R},\ f(x) = \lambda e^{-\lambda x} \mathbb{1}_{[0,+\infty[}(x).$$

(iii) Loi normale *(ou* gaussienne*)* $\mathcal{N}(m, \sigma^2)$, *avec* $m \in \mathbb{R}$ *et* $\sigma \in \mathbb{R}_+^*$ *:*

$$\forall x \in \mathbb{R},\ f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right).$$

(iv) Loi de Cauchy $\mathcal{C}(c)$, *avec* $c \in \mathbb{R}_+^*$ *:*

$$\forall x \in \mathbb{R},\ f(x) = \frac{c}{\pi(x^2 + c^2)}.$$

1.5 Espérance d'une variable aléatoire

Définition 1.5.1 (Espérance)**.** *Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité, $X : \Omega \to \mathbb{R}$ une variable aléatoire, avec $d \in \mathbb{N}$. L'*espérance *de X est définie sous réserve d'existence par :*

$$\mathbb{E}(X) = \int_{\Omega} X \, \mathrm{d}\mathbb{P}.$$

Si $X \geqslant 0$, $\mathbb{E}(X)$ est toujours définie dans $[0, +\infty]$. Sinon, $\mathbb{E}(X)$ est définie dès que $\mathbb{E}(|X|) < +\infty$.

Remarque 1.5.2. *L'espérance hérite des propriétés de l'intégrale ; en particulier, l'espérance est linéaire.*

Proposition 1.5.3 (Formule de transfert)**.** *Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité, $(E, \mathcal{E})$ un espace mesurable, $X : \Omega \to E$ une variable aléatoire et $f : E \to \mathbb{R}$ une fonction intégrable (par rapport à la mesure $\mathbb{P}_X$). Alors $f(X)$ est intégrable (par rapport à la mesure $\mathbb{P}$), et :*

$$\mathbb{E}\left(f(X)\right) = \int_E f \, \mathrm{d}\mathbb{P}_X.$$

Proposition 1.5.4. *Soit X une variable aléatoire à valeurs dans un espace mesurable $(E, \mathcal{E})$. Alors la loi de X est déterminée par les $\mathbb{E}\left(f(X)\right)$, où f parcourt l'ensemble des fonctions réelles mesurables bornées.*

Démonstration. $\forall A \in \mathcal{E}$, $\mathbb{P}(X \in A) = \mathbb{E}\left(\mathbb{1}_A(X)\right)$. □

Vocabulaire 1.5.5 (Lois marginales)**.** *Si $X = (X_1, \ldots, X_d)$ est une variable aléatoire à valeurs dans $\mathbb{R}^d$, alors la loi de X_i est appelée i-ième* loi marginale *de X.*

Proposition 1.5.6. *Soit $X = (X_1, \ldots, X_d)$ une variable aléatoire à valeurs dans $\mathbb{R}^d$. Si X admet une densité par rapport à la mesure de Lebesgue λ_d sur $\mathbb{R}^d$, alors chaque X_i admet une densité par rapport à la mesure de Lebesgue λ sur $\mathbb{R}$.*

Remarque 1.5.7. *Les lois marginales d'une variable aléatoire X sont déterminées par la loi de X, mais la réciproque est fausse.*

Exemple 1.5.8. *Soit Y une variable aléatoire à densité f sur $\mathbb{R}$. On définit $g : (x_1, x_2) \in \mathbb{R}^2 \longmapsto f(x_1) f(x_2)$. On pose X une variable aléatoire sur $\mathbb{R}^2$ de densité g et $X' = (Y, Y)$. Alors X et X' sont des variables aléatoires sur $\mathbb{R}^2$ possédant les mêmes lois marginales, mais si $\Delta = \{(x, x),\ x \in \mathbb{R}\}$, alors $\mathbb{P}(X \in \Delta) = 0$ et $\mathbb{P}(X' \in \Delta) = 1$.*

1.6 Moments d'une variable aléatoire

Définition 1.6.1 (Moments)**.** *Soit X une variable aléatoire réelle et $p \in \mathbb{N}^*$. On dit que X admet un* moment d'ordre *p lorsque $\mathbb{E}\left(|X|^p\right) < +\infty$, et on appelle alors moment d'ordre p de X la quantité $\mathbb{E}\left(X^p\right)$.*

Exemple 1.6.2. *Soit X une variable aléatoire réelle à densité f. Si X admet un moment d'ordre p, alors selon la formule de transfert :*

$$\mathbb{E}\left(X^p\right) = \int_{\mathbb{R}} x^p f(x) \, \mathrm{d}x.$$

Notation 1.6.3. *Si $(\Omega, \mathcal{A}, \mathbb{P})$ est un espace de probabilité et $p \in [1, +\infty[$, on notera $L^p(\Omega, \mathcal{A}, \mathbb{P})$ (ou simplement L^p) l'espace des variables aléatoires réelles de puissance p-ième intégrable, quotienté par la relation d'égalité presque-partout, et muni de $\|\cdot\|_p$.*

Proposition 1.6.4 (Inégalité de Hölder)**.** *Soit X et Y deux variables aléatoires réelles, $(p, q) \in]1, +\infty[^2$ avec $1 = \frac{1}{p} + \frac{1}{q}$. Alors :*

$$\mathbb{E}\left(|XY|\right) \leqslant \mathbb{E}\left(|X|^p\right)^{\frac{1}{p}} \mathbb{E}\left(|Y|^q\right)^{\frac{1}{q}}.$$

Corollaire 1.6.5 (Inégalité de Cauchy-Schwarz)**.** *Soit X et Y deux variables aléatoires réelles. Alors :*

$$\mathbb{E}\left(|XY|\right)^2 \leqslant \mathbb{E}\left(|X|^2\right)\mathbb{E}\left(|Y|^2\right).$$

Proposition 1.6.6. *Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité. Soit $(p,q) \in [1, +\infty[^2$. Alors :*

$$p \leqslant q \Longrightarrow L^p\left(\Omega, \mathcal{A}, \mathbb{P}\right) \supset L^q\left(\Omega, \mathcal{A}, \mathbb{P}\right).$$

Exemple 1.6.7. *La loi normale a des moments de tout ordre, mais la loi de Cauchy n'a aucun moment.*

Remarque 1.6.8. *Les moments permettent d'estimer la "queue de distribution" d'une variable aléatoire X, i.e. la fonction $x \longmapsto \mathbb{P}\left(X \geqslant x\right)$*

Proposition 1.6.9 (Inégalité de Markov)**.** *Soit X une variable aléatoire réelle positive. Soit $p \in \mathbb{N}^*$ t.q. X admet un moment d'ordre p. Alors :*

$$\forall x \in \mathbb{R},\ \mathbb{P}\left(X \geqslant x\right) \leqslant \frac{\mathbb{E}\left(X^p\right)}{x^p}.$$

Proposition 1.6.10 (Inégalité de Chernoff)**.** *Soit X une variable aléatoire réelle. Alors :*

$$\forall \lambda \in \mathbb{R},\ \forall x \in \mathbb{R},\ \mathbb{P}\left(X \geqslant x\right) \leqslant e^{-\lambda x}\mathbb{E}\left(e^{\lambda X}\right).$$

1.7 Variance et covariance

Définition 1.7.1 (Variance)**.** *Soit X une variable aléatoire réelle admettant un moment d'ordre 2. On définit la* variance *de X par :*

$$\mathrm{Var}(X) = \mathbb{E}\left(\left(X - \mathbb{E}(X)\right)^2\right).$$

*On définit de plus l'*écart-type *de X par $\sigma_X = \sqrt{\mathrm{Var}(X)}$.*

Proposition 1.7.2 (Inégalité de Bienaymé-Tchebychev)**.** *Soit X une variable aléatoire réelle admettant un moment d'ordre 2. Alors :*

$$\forall a \in \mathbb{R}_+^*,\ \mathbb{P}\left(|X - \mathbb{E}(X)| \geqslant a\right) \leqslant \frac{\mathrm{Var}(X)}{a^2}.$$

Définition 1.7.3 (Covariance)**.** *Soit X et Y deux variables aléatoires réelles admettant des moments d'ordre 2. On définit la* covariance *de X et Y par :*

$$\mathrm{Cov}(X,Y) = \mathbb{E}\left(\left(X - \mathbb{E}(X)\right)\left(Y - \mathbb{E}(Y)\right)\right).$$

On définit de plus la corrélation *de X et Y par $\mathrm{corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathbb{E}\left(|X|^2\right)\mathbb{E}\left(|Y|^2\right)}}$.*

Définition 1.7.4 (Matrice des covariances)**.** *Soit $X = (X_1, \ldots, X_d)$ une variable aléatoire à valeurs dans $\mathbb{R}^d$. On définit la* matrice des covariances *de X par :*

$$K_X = \left(\mathrm{Cov}\left(X_i, X_j\right)\right)_{1 \leqslant i,j \leqslant d} \in \mathbb{M}_d\left(\mathbb{R}\right).$$

Proposition 1.7.5. *Soit X une variable aléatoire à valeurs dans $\mathbb{R}^d$.*

(i) *La matrice des covariances K_X est symétrique positive.*

(ii) *Si $A \in \mathbb{M}_d\left(\mathbb{R}\right)$, alors $K_{AX} = AK_X{}^tA$.*

1.8 Complément – théorème de Stieltjes

Définition 1.8.1 (Fonction de répartition d'une mesure)**.** *Soit μ une mesure sur* $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$*. On définit la* fonction de répartition de μ *par :*

$$F_\mu : x \in \mathbb{R} \longmapsto \mu\left(]-\infty, x]\right).$$

Théorème 1.8.2 (Théorème de Stieltjes)**.**

(i) *Soit μ une mesure finie sur* $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$*. Alors F_μ est croissante, bornée, continue à droite, et* $\lim_{-\infty} F_\mu = 0$.

(ii) *Soit $F : \mathbb{R} \to \mathbb{R}_+$ une fonction croissante, bornée, continue à droite, et t.q.* $\lim_{-\infty} F = 0$. *Alors il existe une unique mesure finie μ sur* $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$ *t.q. $F_\mu = F$. Pour $f : \mathbb{R} \to [0, +\infty]$ mesurable, on appelle alors* intégrale de Stieltjes *de f par rapport à F l'intégrale* $\int_{\mathbb{R}} f \,\mathrm{d}\mu$.

Démonstration. (ii) *Unicité.* Appliquer la proposition 1.1.10. *Existence.* Appliquer le théorème de Carathéodory. ☐

Remarque 1.8.3. *On peut étendre la construction donnée par le théorème de Stieltjes au cas d'une mesure μ σ-finie, en posant :*

$$F_\mu : x \in \mathbb{R} \longmapsto \begin{cases} \mu\left(]0, x]\right) & \text{si } x \geqslant 0 \\ -\mu\left(]x, 0]\right) & \text{si } x < 0 \end{cases}.$$

Par exemple, la mesure de Lebesgue correspond à la fonction $F : x \in \mathbb{R} \longmapsto x$.

1.9 Fonctions associées à une variable aléatoire

1.9.1 Fonction de répartition

Définition 1.9.1 (Fonction de répartition)**.** *Soit X une variable aléatoire réelle. On définit la* fonction de répartition *de X par :*

$$F_X : x \in \mathbb{R} \longmapsto \mathbb{P}\left(X \leqslant x\right).$$

Proposition 1.9.2. *Soit X une variable aléatoire réelle.*

(i) *F_X est croissante, bornée, continue à droite, et* $\lim_{-\infty} F_X = 0$, $\lim_{+\infty} F_X = 1$.

(ii) *Pour $x \in \mathbb{R}$, F_X est continue en x* ssi $\mathbb{P}(X = x) = 0$.

(iii) *F_X caractérise la loi de X.*

Exemple 1.9.3. *Si X suit la loi de Cauchy $\mathcal{C}(c)$, alors $\forall x \in \mathbb{R}$, $F_X(x) = \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{x}{c}\right)$.*

Proposition 1.9.4. *Soit μ une mesure de probabilité sur* $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$ *et F_μ sa fonction de répartition. On pose :*

$$G : u \in [0, 1[\longmapsto \inf\left\{x \in \mathbb{R},\ F_\mu(x) \geqslant u\right\}.$$

Si U est une variable aléatoire de loi $\mathcal{U}\left([0, 1]\right)$, alors $G(U)$ a pour loi μ.

1.9.2 Série génératrice

Définition 1.9.5 (Série génératrice)**.** *Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. On définit la* série génératrice *de X par :*

$$G_X : z \in \overline{D(0,1)} \longmapsto \mathbb{E}\left(z^X\right) = \sum_{n=0}^{\infty} \mathbb{P}\left(X = n\right) z^n.$$

Proposition 1.9.6. *Soit X une variable aléatoire à valeurs dans $\mathbb{N}$.*

(i) *G_X est holomorphe sur $D(0,1)$ et continue sur $\overline{D(0,1)}$.*

(ii) *G_X caractérise la loi de X : $\forall n \in \mathbb{N}$, $\mathbb{P}(X = n) = \frac{G_X^{(n)}(0)}{n!}$.*

(iii) *Soit $p \in \mathbb{N}$ t.q. X admet un moment d'ordre p. Alors g_X est p fois dérivable en 1 et :*

$$g_X^{(p)}(1) = \mathbb{E}\left(X(X-1)\cdots(X-p+1)\right).$$

1.9.3 Fonction caractéristique

Définition 1.9.7 (Fonction caractéristique)**.** *Soit X une variable aléatoire à valeurs dans $\mathbb{R}^d$. On définit la* fonction caractéristique *de X par :*

$$\Phi_X : \xi \in \mathbb{R}^d \longmapsto \mathbb{E}\left(e^{i\langle \xi | X\rangle}\right),$$

où $\langle \cdot \mid \cdot \rangle$ est le produit scalaire canonique sur $\mathbb{R}^d$.

Définition 1.9.8 (Transformée de Fourier d'une mesure)**.** *Soit μ une mesure sur $\left(\mathbb{R}^d, \mathrm{Bor}\left(\mathbb{R}^d\right)\right)$. On définit la* transformée de Fourier *de μ par :*

$$\hat{\mu} : \xi \in \mathbb{R}^d \longmapsto \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-i\langle \xi | \cdot \rangle} \, \mathrm{d}\mu.$$

Lemme 1.9.9. *Soit $\sigma \in \mathbb{R}_+^*$. On considère :*

$$g_\sigma : x \in \mathbb{R} \longmapsto \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right).$$

Soit μ_σ la mesure sur $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$ de densité g_σ. Alors :

$$\forall \xi \in \mathbb{R},\ \hat{\mu}_\sigma(\xi) = \frac{1}{\sigma} g_{1/\sigma}(\xi).$$

Théorème 1.9.10. *La transformée de Fourier est injective sur l'ensemble des mesures boréliennes sur $\mathbb{R}^d$.*

Démonstration. On se place dans le cas où $d = 1$. Soit ν une mesure sur $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$. Pour $\sigma \in \mathbb{R}_+^*$, on pose :

$$f_\sigma : x \in \mathbb{R} \longmapsto (g_\sigma * \mu)(x) = \int_{\mathbb{R}} g_\sigma(x - \cdot) \, \mathrm{d}\nu,$$

avec les notations du lemme 1.9.9. Et on note ν_σ la mesure de densité f_σ sur $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$. En utilisant le lemme 1.9.9, on montre que $\hat{\nu}$ détermine ν_σ pour tout $\sigma \in \mathbb{R}_+^*$. Comme la famille $(g_\sigma)_{\sigma \in \mathbb{R}_+^*}$ est une approximation de la masse de Dirac δ_0 en 0 (i.e. $\forall \sigma \in \mathbb{R}_+^*$, $\int_{\mathbb{R}} g_\sigma \, \mathrm{d}\lambda = 1$ et $\forall \varepsilon > 0$, $\int_{\mathbb{R}\setminus[-\varepsilon,+\varepsilon]} g_\sigma \, \mathrm{d}\lambda \xrightarrow[\sigma \to 0]{} 0$). On en déduit que $\forall \varphi \in \mathcal{C}_c^0(\mathbb{R})$, $\int \varphi \, \mathrm{d}\nu_\sigma \xrightarrow[\sigma \to 0]{} \int \varphi \, \mathrm{d}\nu$. Par densité de $\mathcal{C}_c^0(\mathbb{R})$ dans $L^1(\mathbb{R})$, on en déduit que $\hat{\nu}$ détermine ν. □

Corollaire 1.9.11. *Si X est une variable aléatoire à valeurs dans $\mathbb{R}^d$, Φ_X caractérise la loi de X.*

Démonstration. C'est une conséquence du théorème 1.9.10 car $\forall \xi \in \mathbb{R}^d,\ \Phi_X(\xi) = (2\pi)^{d/2}\hat{\mathbb{P}}_X(-\xi)$. □

Proposition 1.9.12. *Soit X une variable aléatoire réelle. Soit $p \in \mathbb{N}$ t.q. X admet un moment d'ordre p. Alors Φ_X est de classe $\mathcal{C}^p$ sur $\mathbb{R}$ et :*

$$\mathbb{E}(X^p) = (-i)^p \Phi_X^{(p)}(0).$$

1.9.4 Transformée de Laplace

Définition 1.9.13 (Transformée de Laplace)**.** *Soit X une variable aléatoire réelle positive. On définit la* transformée de Laplace *de X par :*

$$L_X : \lambda \in \mathbb{R} \longmapsto \mathbb{E}\left(e^{-\lambda X}\right).$$

Proposition 1.9.14. *Soit X une variable aléatoire réelle positive.*

(i) *L_X est $\mathcal{C}^\infty$ sur $]0, +\infty[$ et $\mathcal{C}^0$ en 0.*

(ii) L_X *caractérise la loi de* X.

(iii) *Soit* $p \in \mathbb{N}$ *t.q.* X *admet un moment d'ordre* p. *Alors* L_X *est* p *fois dérivable en* 0 *et :*

$$\mathbb{E}\left(X^p\right) = (-1)^p L_X^{(p)}(0).$$

Démonstration. (ii) On note $\mathcal{A} = \operatorname{Vect}\left(\left\{\left(x \in [0,+\infty] \longmapsto e^{-\lambda x}\right),\ \lambda \in \mathbb{R}_+\right\}\right)$. Soit X et X' deux variables aléatoires réelles t.q. $L_X = L_{X'}$. On a alors aisément $\forall f \in \mathcal{A}$, $\mathbb{E}\left(f(X)\right) = \mathbb{E}\left(f\left(X'\right)\right)$. Or $\mathcal{A}$ est une algèbre de fonctions continues sur le compact $[0,+\infty]$, et $\mathcal{A}$ sépare les points (i.e. $\forall x \neq y$, $\exists f \in \mathcal{A}$, $f(x) \neq f(y)$). Selon le théorème de Stone-Weierstraß, $\mathcal{A}$ est dense dans $\left(\mathcal{C}^0\left([0,+\infty]\right), \|\cdot\|_\infty\right)$. Par convergence dominée, on en déduit que $\mathbb{E}\left(f(X)\right) = \mathbb{E}\left(f\left(X'\right)\right)$ pour tout f mesurable bornée. En appliquant cela à des fonctions caractéristiques, on en déduit que $\mathbb{P}_X = \mathbb{P}_{X'}$. □

Remarque 1.9.15. *On peut étendre le domaine de définition de* L_X *à* $\{z \in \mathbb{C},\ \Re(z) \geqslant 0\}$. *On obtient une fonction holomorphe sur* $\{z \in \mathbb{C},\ \Re(z) > 0\}$. *Pour déterminer la loi de* X, *il suffit donc de connaître* L_X *sur un sous-ensemble de* $\mathbb{R}_+^*$ *contenant un point d'accumulation.*

Proposition 1.9.16. *Soit* X *une variable aléatoire à valeurs réelles positives. S'il existe* $\varepsilon > 0$ *t.q.* $\mathbb{E}\left(e^{\varepsilon X}\right) < +\infty$, *on dit que* X *admet des* moments exponentiels. *Dans ce cas,* X *admet tous ses moments,* L_X *est holomorphe sur* $\{z \in \mathbb{C},\ \Re(z) > -\varepsilon\}$ *et :*

$$\forall \lambda \in \left]-\varepsilon,+\varepsilon\right[,\ L_X(\lambda) = \sum_{k \in \mathbb{N}} \mathbb{E}\left(X^k\right) \frac{(-\lambda)^k}{k!}.$$

Cette série s'appelle la fonction génératrice des moments.

Corollaire 1.9.17. *Si une variable aléatoire* X *à valeurs réelles positives admet des moments exponentiels, alors la suite* $\left(\mathbb{E}\left(X^k\right)\right)_{k \in \mathbb{N}}$ *caractérise la loi de* X.

2 Indépendance

2.1 Événements indépendants

Notation 2.1.1. *Dans toute la suite, on fixe un espace de probabilité* $(\Omega, \mathcal{A}, \mathbb{P})$.

Définition 2.1.2 (Événements indépendants)**.**

(i) *On dit que deux événements* A *et* B *sont* indépendants, *et on note* $A \perp B$, *lorsque :*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

(ii) *On dit que* n *événements* $A_1, \ldots, A_n$ *sont* indépendants *(dans leur ensemble) lorsque :*

$$\forall J \in \mathcal{P}\left(\{1,\ldots,n\}\right),\ \mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}\left(A_j\right).$$

(iii) *On dit qu'une famille d'événements* $\left(A_j\right)_{j \in J}$ *est* indépendante *lorsque :*

$$\forall J \in \mathcal{P}_{\mathrm{f}}(I),\ \mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}\left(A_j\right),$$

où $\mathcal{P}_{\mathrm{f}}(I)$ *est l'ensemble des parties finies de* I.

Remarque 2.1.3. *Si* n *événements sont indépendants (dans leur ensemble), alors ils sont deux à deux indépendants ; mais la réciproque est fausse.*

2.2 Sous-tribus indépendantes

Définition 2.2.1 (Sous-tribus indépendantes)**.**

(i) *On dit que n sous-tribus $\mathcal{B}_1, \dots, \mathcal{B}_n$ de $\mathcal{A}$ sont indépendantes lorsque :*

$$\forall (A_1, \dots, A_n) \in \mathcal{B}_1 \times \cdots \times \mathcal{B}_n,\ \mathbb{P}\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n \mathbb{P}(A_i),$$

ou de manière équivalente, lorsque $\forall (A_1, \dots, A_n) \in \mathcal{B}_1 \times \cdots \times \mathcal{B}_n$, les événements $A_1, \dots, A_n$ sont indépendants.

(ii) *On dit qu'une famille $(\mathcal{B}_i)_{i \in I}$ de sous-tribus de $\mathcal{A}$ est indépendante lorsque toute sous-famille finie de $(\mathcal{B}_i)_{i \in I}$ est indépendante.*

Remarque 2.2.2. *Si A est un événement, notons que $\sigma(\{A\}) = \left\{\varnothing, A, \complement A, \Omega\right\}$. Si $(A_i)_{i \in I}$ est une famille d'événements, alors cette famille est indépendante* ssi *la famille de tribus $(\sigma(\{A_i\}))_{i \in I}$ est indépendante.*

Lemme 2.2.3. *Soit $\mathcal{B}_1, \dots, \mathcal{B}_n$ des sous-tribus de $\mathcal{A}$. Pour $i \in \{1, \dots, n\}$, soit $\mathcal{C}_i \subset \mathcal{B}_i$ t.q. $\sigma(\mathcal{C}_i) = \mathcal{B}_i$, $\Omega \in \mathcal{C}_i$ et $\mathcal{C}_i$ est stable par intersections finies. On suppose que :*

$$\forall (C_1, \dots, C_n) \in \mathcal{C}_1 \times \cdots \times \mathcal{C}_n,\ \mathbb{P}\left(\bigcap_{i=1}^n C_i\right) = \prod_{i=1}^n \mathbb{P}(C_i).$$

Alors $\mathcal{B}_1, \dots, \mathcal{B}_n$ sont indépendantes.

Démonstration. On fixe d'abord $(C_2, \dots, C_n) \in \mathcal{C}_2 \times \cdots \times \mathcal{C}_n$. On définit :

$$\mathfrak{M}_1 = \{A \in \mathcal{B}_1,\ \mathbb{P}(A \cap C_2 \cap \cdots \cap C_n) = \mathbb{P}(A)\mathbb{P}(C_2) \cdots \mathbb{P}(C_n)\}.$$

On montre que $\mathfrak{M}_1$ est une classe monotone stable par intersections finies et contenant $\mathcal{C}_1$. Selon le lemme de classe monotone (théorème 1.1.7), $\mathfrak{M}_1 = \mathcal{B}_1$. Puis on raisonne par récurrence. □

Notation 2.2.4. *Si $(\mathcal{B}_i)_{i \in I}$ est une famille de sous-tribus de $\mathcal{A}$, on note :*

$$\bigvee_{i \in I} \mathcal{B}_i = \sigma\left(\bigcup_{i \in I} \mathcal{B}_i\right).$$

Lemme 2.2.5 (Lemme de regroupement par paquets)**.** *Soit $(\mathcal{B}_i)_{i \in I}$ une famille de sous-tribus indépendantes de $\mathcal{A}$. Soit $(I_\lambda)_{\lambda \in \Lambda}$ une partition de I. Alors les sous-tribus $\left(\bigvee_{i \in I_\lambda} \mathcal{B}_i\right)_{\lambda \in \Lambda}$ sont indépendantes.*

2.3 Variables aléatoires indépendantes

Définition 2.3.1 (Tribu engendrée par une variable aléatoire)**.** *Soit $(E, \mathcal{E})$ un espace mesurable et $X : \Omega \to E$ une variable aléatoire. On appelle* tribu engendrée *par X la tribu :*

$$\sigma(X) = \left\{X^{-1}(B),\ B \in \mathcal{E}\right\}.$$

C'est la plus petite sous-tribu de $\mathcal{A}$ qui rende X mesurable. De même, si $(X_i)_{i \in I}$ est une famille de variables aléatoires, la tribu engendrée par les $(X_i)_{i \in I}$ est la plus petite tribu de $\mathcal{A}$ qui rende tous les $(X_i)_{i \in I}$ mesurables :

$$\sigma(X_i,\ i \in I) = \bigvee_{i \in I} \sigma(X_i).$$

Proposition 2.3.2. *Soit $(E, \mathcal{E})$ un espace mesurable. Soit $X : \Omega \to E$ et $Y : \Omega \to \mathbb{R}$ deux variables aléatoires. Alors Y est $\sigma(X)$-mesurable* ssi *il existe une fonction $f : E \to \mathbb{R}$ mesurable t.q.*

$$Y = f(X).$$

Démonstration. ($\Leftarrow$) Clair. ($\Rightarrow$) Le montrer d'abord pour Y étagée, puis passer à la limite. $\square$

Définition 2.3.3 (Variables aléatoires indépendantes)**.** *Soit $((E_i, \mathcal{E}_i))_{i\in I}$ une famille d'espaces mesurables et $(X_i)_{i\in I} \in \prod_{i\in I} {E_i}^{\Omega}$ une famille de variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbb{P})$. On dit que les variables aléatoires $(X_i)_{i\in I}$ sont* indépendantes *lorsque les tribus $(\sigma(X_i))_{i\in I}$ sont indépendantes, ou de manière équivalente lorsque :*

$$\forall J \in \mathcal{P}_{\mathrm{f}}(I),\ \forall (A_j)_{j\in J} \in \prod_{j\in J} \mathcal{E}_j,\ \mathbb{P}\left(\bigcap_{j\in J} X_j^{-1}(A_j)\right) = \prod_{j\in J} \mathbb{P}\left(X_j^{-1}(A_j)\right),$$

où $\mathcal{P}_{\mathrm{f}}(I)$ est l'ensemble des parties finies de I.

Remarque 2.3.4. *Soit $(\mathcal{B}_i)_{i\in I}$ une famille de sous-tribus indépendantes de $\mathcal{A}$ et soit $(X_i)_{i\in I}$ une famille de variables aléatoires sur Ω. Si X_i est $\mathcal{B}_i$-mesurable pour tout $i \in I$, alors les $(X_i)_{i\in I}$ sont indépendantes.*

2.4 Rappels sur les produits de mesures

Définition 2.4.1 (Tribu produit)**.** *Soit $((E_i, \mathcal{E}_i))_{i\in I}$ une famille d'espaces mesurables. La tribu produit $\bigotimes_{i\in I} \mathcal{E}_i$ est la plus petite tribu sur $\prod_{i\in I} E_i$ t.q. pour tout $i \in I$, la projection $\pi_i : \prod_{j\in I} E_j \longrightarrow E_i$ est mesurable.*

Exemple 2.4.2. *Si X et Y sont deux espaces métriques séparables, alors :*

$$\mathrm{Bor}(X \times Y) = \mathrm{Bor}(X) \otimes \mathrm{Bor}(Y).$$

Proposition 2.4.3. *Soit $(E, \mathcal{E}, \mu)$ et $(F, \mathcal{F}, \nu)$ deux espaces mesurés σ-finis. Alors il existe une unique mesure notée $\mu \otimes \nu$ sur $(E \times F, \mathcal{E} \otimes \mathcal{F})$ t.q.*

$$\forall (A, B) \in \mathcal{E} \times \mathcal{F},\ (\mu \otimes \nu)(A \times B) = \mu(A)\nu(B).$$

Remarque 2.4.4. *On peut de même définir une mesure produit sur un produit fini d'espaces. Le produit de tribus et le produit de mesures sont alors associatifs.*

2.5 Caractérisation de l'indépendance en termes de lois

Proposition 2.5.1. *Soit $(X_i)_{1\leqslant i\leqslant k}$ une famille finie de variables aléatoires sur Ω. Alors les $(X_i)_{1\leqslant i\leqslant k}$ sont indépendantes* ssi

$$\mathbb{P}_{(X_1,\ldots,X_k)} = \mathbb{P}_{X_1} \otimes \cdots \otimes \mathbb{P}_{X_k}.$$

Exemple 2.5.2. *Soit U et V deux variables aléatoires indépendantes t.q. $U \sim \mathrm{Exp}(1)$ et $V \sim \mathcal{U}([0,1])$. Alors les variables aléatoires $X = \sqrt{U}\cos(2\pi V)$ et $Y = \sqrt{U}\sin(2\pi V)$ sont indépendantes.*

Corollaire 2.5.3. *Soit $((E_i, \mathcal{E}_i))_{i\in I}$ une famille d'espaces mesurables, $(X_i)_{i\in I} \in \prod_{i\in I} {E_i}^{\Omega}$ une famille de variables aléatoires indépendantes et $(f_i)_{i\in I} \in \prod_{i\in I} \mathbb{R}^{E_i}$ une famille de fonctions intégrables (f_i est intégrable pour la mesure $\mathbb{P}_{X_i}$, i.e. $f_i \circ X_i$ est intégrable pour la mesure $\mathbb{P}$). Alors :*

$$\mathbb{E}\left(\prod_{i\in I} f_i(X_i)\right) = \prod_{i\in I} \mathbb{E}(f(X_i)).$$

En particulier, si les $(X_i)_{i\in I}$ sont réelles et d'espérance finie, alors :

$$\mathbb{E}\left(\prod_{i\in X} X_i\right) = \prod_{i\in I} \mathbb{E}(X_i).$$

Définition 2.5.4 (Variables aléatoires décorrélées)**.** *Deux variables aléatoires réelles X et Y admettant des moments d'ordre 2 sont dites* décorrélées *lorsque* $\operatorname{Cov}(X, Y) = 0$.

Corollaire 2.5.5. *Deux variables aléatoires indépendantes sont décorrélées.*

Exemple 2.5.6. *Soit X et E deux variables aléatoires indépendantes t.q. $X \sim \mathcal{N}(0,1)$ et $E \sim \mathcal{U}(\{-1,1\})$. Alors $\operatorname{Cov}(X, EX) = 0$, donc X et EX sont décorrélées. Pourtant, le support de la loi du couple (X, EX) est de mesure de Lebesgue nulle dans $\mathbb{R}^2$, donc (X, EX) n'est pas à densité, donc X et EX ne sont pas indépendantes (car X et EX sont toutes deux à densité).*

Proposition 2.5.7. *Soit X une variable aléatoire réelle.*

(i) *Soit $\mathcal{B}$ une sous-tribu de $\mathcal{A}$ t.q.*

$$\forall B \in \mathcal{B},\ \mathbb{P}(B) \in \{0,1\}.$$

Si X est $\mathcal{B}$-mesurable, alors X est presque-sûrement constante.

(ii) *Soit $f : \mathbb{R} \to \mathbb{R}$ mesurable t.q. $f(X)$ et X sont indépendantes. Alors $f(X)$ est presque-sûrement constante.*

Démonstration. (i) Soit $F_X : x \longmapsto \mathbb{P}(X \leqslant x)$ la fonction de répartition de x. Alors $F_X(\mathbb{R}) \subset \{0,1\}$. Or F_X est croissante, continue à droite, et $\lim_{-\infty} F_X = 0$, $\lim_{+\infty} F_X = 1$. Donc il existe $x_0 \in \mathbb{R}$ t.q. $F_X = \mathbb{1}_{[x_0,+\infty[}$. Ainsi, $\mathbb{P}(X = x_0) = 1$. (ii) Posons $Y = f(X)$. f étant mesurable, on a $\sigma(Y) \subset \sigma(X)$. Or X et Y sont indépendantes, donc $\sigma(X)$ et $\sigma(Y)$ sont indépendantes, donc $\sigma(Y)$ et $\sigma(Y)$ sont indépendantes. D'où :

$$\forall A \in \sigma(Y),\ \mathbb{P}(A) = \mathbb{P}(A \cap A) = \mathbb{P}(A)^2.$$

Ainsi $\forall A \in \sigma(Y)$, $\mathbb{P}(A) \in \{0,1\}$. D'après le (i), puisque Y est $\sigma(Y)$-mesurable, Y est presque-sûrement constante. □

Proposition 2.5.8. *Soit $X = (X_1, \ldots, X_d)$ une variable aléatoire à valeurs dans $\mathbb{R}^d$, à densité f_X. On suppose qu'il existe des fonctions $(f_1, \ldots, f_d) \in (\mathbb{R}_+)^{\mathbb{R}}$ mesurables t.q.*

$$\forall (x_1, \ldots, x_d) \in \mathbb{R}^d,\ f_X(x_1, \ldots, x_d) = \prod_{i=1}^{d} f_i(x_i).$$

Alors les $(X_i)_{1 \leqslant i \leqslant d}$ sont indépendantes et pour tout $i \in \{1, \ldots, d\}$, il existe une constante $c_i > 0$ t.q. X_i est à densité $c_i f_i$.

2.6 Existence de suites de variables aléatoires réelles indépendantes

Lemme 2.6.1. *Il existe une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires indépendantes de loi $\mathcal{U}(\{0,1\})$.*

Démonstration. Cela revient à prouver l'existence de la mesure d'équiprobabilité sur $\{0,1\}^{\mathbb{N}}$, qui est une conséquence du théorème de Carathéodory. □

Théorème 2.6.2. *Soit $(\mu_n)_{n\in\mathbb{N}}$ une suite de mesures de probabilité sur* $(\mathbb{R}, \operatorname{Bor}(\mathbb{R}))$. *Alors il existe une suite de variables aléatoires indépendantes $(Y_n)_{n\in\mathbb{N}}$ t.q. $\forall n \in \mathbb{N}$, $\mathbb{P}_{Y_n} = \mu_n$.*

Démonstration. On se donne d'abord une suite $(X_k)_{k\in\mathbb{N}}$ de variables aléatoires indépendantes de loi $\mathcal{U}(\{0,1\})$. Soit $\varphi : \mathbb{N}^2 \to \mathbb{N}$ une bijection. Pour $(m,n) \in \mathbb{N}^2$, on pose :

$$Z_{m,n} = X_{\varphi(m,n)}.$$

Alors les variables aléatoires $(Z_{m,n})_{(m,n)\in\mathbb{N}^2}$ sont indépendantes. On pose ensuite, pour $n \in \mathbb{N}$:

$$U_n = \sum_{m=0}^{\infty} \frac{1}{2^{m+1}} Z_{m,n}.$$

Selon le lemme de regroupement par paquets (lemme 2.2.5), les variables aléatoires $(U_n)_{n\in\mathbb{N}}$ sont indépendantes. De plus, $\forall n \in \mathbb{N}$, $U_n \sim \mathcal{U}([0,1])$. Pour $n \in \mathbb{N}$, on définit enfin $F_n : x \in \mathbb{R} \longmapsto \mu_n(]-\infty, x])$ la fonction de répartition de μ_n ; et on pose :

$$G_n : u \in]0,1[\longmapsto \inf\{x \in \mathbb{R},\ F_n(x) \geqslant u\}.$$

Selon la proposition 1.9.4, $Y_n = G_n(U_n)$ a pour loi μ_n ; et les $(Y_n)_{n\in\mathbb{N}}$ sont indépendantes. □

Remarque 2.6.3. *Si $((E_n, \mathcal{E}_n, \mu_n))_{n\in\mathbb{N}}$ est une suite d'espaces de probabilité, alors il existe une mesure de probabilité μ sur $(\prod_{n\in\mathbb{N}} E_n, \bigotimes_{n\in\mathbb{N}} \mathcal{E}_n)$ t.q.*

$$\forall F \in \mathcal{P}_{\mathrm{f}}(\mathbb{N}),\ \forall (A_n)_{n\in F} \in \prod_{n\in F} \mathcal{E}_n,\ \mu\left(\left(\prod_{n\in F} A_n\right) \times \left(\prod_{m\in\mathbb{N}\setminus F} E_m\right)\right) = \prod_{n\in F} \mu_n(A_n).$$

2.7 Sommes de variables aléatoires réelles indépendantes

Définition 2.7.1 (Convolution de deux mesures)**.** *Soit μ et ν deux mesures sur* $(\mathbb{R}, \mathrm{Bor}(\mathbb{R}))$*. On définit le* produit de convolution $\mu * \nu$ *par :*

$$\mu * \nu = S_*(\mu \otimes \nu),$$

où $S : (x,y) \in \mathbb{R}^2 \longmapsto x + y \in \mathbb{R}$. Autrement dit $\forall A \in \mathrm{Bor}(\mathbb{R})$, $(\mu * \nu)(A) = (\mu \otimes \nu)(S^{-1}(A))$.

Proposition 2.7.2. *Si $X_1, \dots, X_k$ sont des variables aléatoires réelles indépendantes, alors la loi de la somme est donnée par :*

$$\mathbb{P}_{X_1+\dots+X_k} = \mathbb{P}_{X_1} * \dots * \mathbb{P}_{X_k}.$$

Proposition 2.7.3. *Si $X_1, \dots, X_k$ sont des variables aléatoires réelles indépendantes, alors la fonction caractéristique de la somme est donnée par :*

$$\Phi_{X_1+\dots+X_k} = \Phi_{X_1} \cdots \Phi_{X_k}.$$

Proposition 2.7.4. *Si $X_1, \dots, X_k$ sont des variables aléatoires réelles admettant des moments d'ordre 2, alors :*

$$\mathrm{Var}\left(\sum_{i=1}^k X_i\right) = \sum_{i=1}^k \mathrm{Var}(X_i) + 2 \sum_{1\leqslant i<j\leqslant k} \mathrm{Cov}(X_i, X_j).$$

En particulier, si les $(X_i)_{1\leqslant i\leqslant k}$ sont deux à deux décorrélées (donc si elles sont indépendantes), alors $\mathrm{Var}\left(\sum_{i=1}^k X_i\right) = \sum_{i=1}^k \mathrm{Var}(X_i)$.

2.8 Loi faible des grands nombres

Théorème 2.8.1 (Loi faible des grands nombres dans L^2)**.** *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles indépendantes et suivant toutes la loi d'une variable aléatoire X. On suppose que X admet un moment d'ordre 2. Alors :*

$$\frac{1}{n}\sum_{k=1}^n X_k \xrightarrow[n\to+\infty]{L^2} \mathbb{E}(X).$$

Exemple 2.8.2 (Loi des événements rares)**.** *Pour $n \in \mathbb{N}$, soit Y_n une variable aléatoire de loi $\mathcal{B}(n, p_n)$, avec $p_n \in [0,1]$. On suppose que $np_n \xrightarrow[n\to+\infty]{} \lambda \in \mathbb{R}_+^*$. Alors $(Y_n)_{n\in\mathbb{N}}$ converge en loi vers une variable aléatoire de Poisson de paramètre λ.*

2.9 Autres caractérisations de l'indépendance

Proposition 2.9.1. *Soit* $((E_i, \mathcal{E}_i))_{1 \leqslant i \leqslant k}$ *une famille d'espaces mesurables et* $(X_i)_{1 \leqslant i \leqslant k} \in \prod_{1 \leqslant i \leqslant k} {E_i}^{\Omega}$ *une famille de variables aléatoires discrètes sur* Ω*. Alors les* $(X_i)_{1 \leqslant i \leqslant k}$ *sont indépendantes* ssi

$$\forall (x_1, \ldots, x_k) \in E_1 \times \cdots \times E_k,\ \mathbb{P}\left(\bigcap_{i=1}^{k} (X_i = x_i)\right) = \prod_{i=1}^{k} \mathbb{P}(X_i = x_i).$$

Proposition 2.9.2. *Soit* $(X_i)_{1 \leqslant i \leqslant k} \in \prod_{1 \leqslant i \leqslant k} {E_i}^{\Omega}$ *une famille de variables aléatoires réelles sur* Ω*. Pour* $i \in \{1, \ldots, n\}$*, on note* F_{X_i} *la fonction de répartition de* X_i *(c.f. définition 1.9.1). Alors les* $(X_i)_{1 \leqslant i \leqslant k}$ *sont indépendantes* ssi

$$\forall (x_1, \ldots, x_k) \in \mathbb{R}^k,\ \mathbb{P}\left(\bigcap_{i=1}^{k} (X_i \leqslant x_i)\right) = \prod_{i=1}^{k} F_{X_i}(x_i).$$

Proposition 2.9.3. *Soit* $(X_i)_{1 \leqslant i \leqslant k} \in \prod_{1 \leqslant i \leqslant k} {E_i}^{\Omega}$ *une famille de variables aléatoires réelles sur* Ω*. Pour* $i \in \{1, \ldots, n\}$*, on note* Φ_{X_i} *la fonction caractéristique de* X_i *(c.f. définition 1.9.7). Alors les* $(X_i)_{1 \leqslant i \leqslant k}$ *sont indépendantes* ssi

$$\Phi_{(X_1, \ldots, X_k)} = \Phi_{X_1} \cdots \Phi_{X_k}.$$

2.10 Lemme de Borel-Cantelli et loi du zéro-un de Kolmogorov

Définition 2.10.1 (lim sup et lim inf)**.** *Soit* $(A_n)_{n \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$ *une suite d'événements. On définit :*

$$\limsup_{n \to +\infty} A_n = \bigcap_{n \in \mathbb{N}} \bigcup_{k \geqslant n} A_k \in \mathcal{A} \qquad \textit{et} \qquad \liminf_{n \to +\infty} A_n \in \mathcal{A} = \bigcup_{n \in \mathbb{N}} \bigcap_{k \geqslant n} A_k \in \mathcal{A}.$$

$\limsup_{n \to +\infty} A_n$ *est l'événement "une infinité de* A_n *sont réalisés" et* $\liminf_{n \to +\infty} A_n$ *est l'événement "tous les* A_n *sont réalisés à partir d'un certain rang".*

Proposition 2.10.2. *Soit* $(A_n)_{n \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$*. Alors :*

(i) $\liminf_{n \to +\infty} A_n \subset \limsup_{n \to +\infty} A_n$.

(ii) $\complement(\liminf_{n \to +\infty} A_n) = \limsup_{n \to +\infty} \complement A_n$.

Théorème 2.10.3 (Lemme de Borel-Cantelli)**.** *Soit* $(A_n)_{n \in \mathbb{N}} \in \mathcal{A}^{\mathbb{N}}$*.*

(i) *Si* $\sum_{n \in \mathbb{N}} \mathbb{P}(A_n) < +\infty$*, alors* $\mathbb{P}\left(\limsup_{n \to +\infty} A_n\right) = 0$.

(ii) *Si* $\sum_{n \in \mathbb{N}} \mathbb{P}(A_n) = +\infty$ *et les* $(A_n)_{n \in \mathbb{N}}$ *sont indépendants, alors* $\mathbb{P}\left(\limsup_{n \to +\infty} A_n\right) = 1$.

Définition 2.10.4 (Tribu asymptotique)**.** *Soit* $(\mathcal{B}_n)_{n \in \mathbb{N}}$ *une suite de sous-tribus de* $\mathcal{A}$*. On pose :*

$$\mathcal{A}_\infty = \bigcap_{n \in \mathbb{N}} \bigvee_{k \geqslant n} \mathcal{B}_k.$$

$\mathcal{A}_\infty$ *est la* tribu asymptotique *des* $(\mathcal{B}_n)_{n \in \mathbb{N}}$*. Ses événements sont appelés* événements asymptotiques.

Remarque 2.10.5. *La tribu asymptotique des* $(\mathcal{B}_n)_{n \in \mathbb{N}}$ *est l'information contenue dans la suite* $(\mathcal{B}_n)_{n \in \mathbb{N}}$ *qui ne dépend d'aucune sous-famille finie.*

Théorème 2.10.6 (Loi du zéro-un de Kolmogorov)**.** *Soit* $(\mathcal{B}_n)_{n \in \mathbb{N}}$ *une suite de sous-tribus de* $\mathcal{A}$*. Si les* $(\mathcal{B}_n)_{n \in \mathbb{N}}$ *sont indépendantes, alors la tribu asymptotique* $\mathcal{A}_\infty$ *des* $(\mathcal{B}_n)_{n \in \mathbb{N}}$ *est triviale, i.e.*

$$\forall A \in \mathcal{A}_\infty,\ \mathbb{P}(A) \in \{0, 1\}.$$

Démonstration. Soit $n \in \mathbb{N}$. Alors les tribus $\mathcal{B}_0, \ldots, \mathcal{B}_n, \mathcal{A}_{n+1}$ sont indépendantes par regroupement par paquets (c.f. lemme 2.2.5). Comme $\mathcal{A}_\infty \subset \mathcal{A}_{n+1}$, les tribus $\mathcal{B}_0, \ldots, \mathcal{B}_n, \mathcal{A}_\infty$ sont indépendantes, et ceci est vrai pour tout $n \in \mathbb{N}$. À nouveau par regroupement par paquets, $\mathcal{A}_\infty$ est indépendante de $\bigvee_{n\in\mathbb{N}} \mathcal{B}_n$. Or $\mathcal{A}_\infty \subset \bigvee_{n\in\mathbb{N}} \mathcal{B}_n$, donc $\mathcal{A}_\infty$ est indépendante de $\mathcal{A}_\infty$. Ainsi :

$$\forall A \in \mathcal{A}_\infty,\ \mathbb{P}(A) = \mathbb{P}(A \cap A) = \mathbb{P}(A)^2,$$

d'où le résultat. $\square$

Corollaire 2.10.7 (Loi du zéro-un de Borel)**.** *Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements indépendants. Alors $\limsup_{n\to+\infty} A_n$ est un événement asymptotique et on a la dichotomie suivante :*

(i) *Si $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) < +\infty$, alors $\mathbb{P}\left(\limsup_{n\to+\infty} A_n\right) = 0$.*

(ii) *Si $\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = +\infty$, alors $\mathbb{P}\left(\limsup_{n\to+\infty} A_n\right) = 1$.*

Exemple 2.10.8. *Quelques applications :*

(i) *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes de loi $\mathcal{B}\left(\frac{1}{2}\right)$. Pour $n \in \mathbb{N}$, on pose $R_n = \max\{\ell \in \mathbb{N},\ X_n = \cdots = X_{n+\ell-1} = 1)$, puis $M_n = \max(R_0, \ldots, R_n)$. Alors :*

$$\mathbb{P}\left(\frac{M_n}{\log_2 n} \xrightarrow[n\to+\infty]{} 1\right) = 1.$$

(ii) *Il n'existe pas de mesure $\mathbb{P}$ sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ t.q.*

$$\forall k \in \mathbb{N}^*,\ \mathbb{P}(k\mathbb{N}) = \frac{1}{k}.$$

(iii) *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes de loi $\mathcal{B}\left(\frac{1}{2}\right)$. Alors pour tout $k \in \mathbb{N}^*$ et pour tout $\varepsilon \in \{0,1\}^k$, le motif $(\varepsilon_1, \ldots, \varepsilon_k)$ apparaît presque-sûrement une infinité de fois dans la suite $(X_n)_{n\in\mathbb{N}}$.*

(iv) *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes de même loi. On pose, pour $n \in \mathbb{N}$, $S_n = X_0 + \cdots + X_n$. Alors les événements $\left(\limsup_{n\to+\infty} S_n = +\infty\right)$ et $(\liminf_{n\to+\infty} S_n = -\infty)$ sont de probabilité 0 ou 1. De plus, si les $(X_n)_{n\in\mathbb{N}}$ sont de loi $\mathcal{U}(\{-1,1\})$, alors :*

$$\mathbb{P}\left(\liminf_{n\to+\infty} S_n = -\infty \text{ et } \limsup_{n\to+\infty} S_n = +\infty\right) = 1.$$

On dit que $(S_n)_{n\in\mathbb{N}}$ oscille.

3 Convergence de variables aléatoires

3.1 Convergence presque-sûre et convergence L^p

Notation 3.1.1. *Dans toute la suite, on fixe un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$.*

Définition 3.1.2 (Convergence presque-sûre et convergence L^p)**.** *Soit $(X_n)_{n\in\mathbb{N}}$ et X des variables aléatoires réelles définies sur Ω.*

(i) *On dit que $X_n \xrightarrow[n\to+\infty]{\text{p.s.}} X$ ($(X_n)_{n\in\mathbb{N}}$ converge* presque-sûrement *vers X) lorsque :*

$$\mathbb{P}\left(X_n \xrightarrow[n\to+\infty]{} X\right) = 1.$$

Cela est équivalent à dire que $X_n \xrightarrow[n\to+\infty]{} X$ simplement presque-partout.

(ii) *Pour $p \geqslant 1$, si les $(X_n)_{n\in\mathbb{N}}$ et X sont dans L^p, on dit que $X_n \xrightarrow[n\to+\infty]{L^p} X$ lorsque :*

$$\mathbb{E}\left(|X_n - X|^p\right) \xrightarrow[n\to+\infty]{} 0.$$

Proposition 3.1.3. *Soit $(X_n)_{n\in\mathbb{N}}$ et X des variables aléatoires réelles définies sur Ω. Soit $p \geqslant 1$.*

(i) *Si $X_n \xrightarrow[n\to+\infty]{\text{p.s.}} X$ et si $\exists Y \in L^p$, $\forall n \in \mathbb{N}$, $|X_n| \leqslant Y$, alors $X_n \xrightarrow[n\to+\infty]{L^p} X$.*

(ii) *Si $X_n \xrightarrow[n\to+\infty]{L^p} X$, alors il existe une extractrice φ t.q. $X_{\varphi(n)} \xrightarrow[n\to+\infty]{\text{p.s.}} X$.*

3.2 Loi forte des grands nombres

Lemme 3.2.1. *Soit X une variable aléatoire réelle positive. Alors :*

$$\forall p \in \mathbb{R}_+^*,\ \mathbb{E}\left(X^p\right) = \int_0^\infty pt^{p-1}\mathbb{P}\left(X > t\right)\,\mathrm{d}t.$$

Théorème 3.2.2 (Loi forte des grands nombres)**.** *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles indépendantes et suivant toutes la loi d'une variable aléatoire X. On suppose que X admet un moment d'ordre 1. Alors :*

$$\frac{1}{n}\sum_{k=1}^n X_k \xrightarrow[n\to+\infty]{\text{p.s.}} \mathbb{E}(X).$$

Démonstration (Première méthode)**.** Quitte à décomposer X en $X = X^+ - X^-$, on peut supposer que $X \geqslant 0$. *Étape 1.* Pour $n \in \mathbb{N}^*$, posons :

$$Y_n = X_n \mathbb{1}\left(X_n \leqslant n\right).$$

Montrons que presque-sûrement, pour n assez grand, $X_n = Y_n$ (i.e. $\mathbb{P}\left(\liminf_{n\to+\infty}\left(X_n = Y_n\right)\right) = 1$). Pour cela, notons que :

$$\sum_{n\in\mathbb{N}^*}\mathbb{P}\left(X_n \neq Y_n\right) = \sum_{n\in\mathbb{N}^*}\mathbb{P}\left(X > n\right) \leqslant \int_0^\infty \mathbb{P}\left(X > t\right)\,\mathrm{d}t = \mathbb{E}(X) < +\infty.$$

Selon le lemme de Borel-Cantelli (théorème 2.10.3) :

$$\mathbb{P}\left(\liminf_{n\to+\infty}\left(X_n = Y_n\right)\right) = 1 - \mathbb{P}\left(\limsup_{n\to+\infty}\left(X_n \neq Y_n\right)\right) = 1.$$

Étape 2. Il suffit maintenant de prouver que $\frac{T_n}{n} \xrightarrow[n\to+\infty]{\text{p.s.}} \mathbb{E}(X)$, avec $T_n = \sum_{k=1}^n Y_k$. Soit $\alpha > 1$. Posons $k(n) = \lfloor \alpha^n \rfloor$. Alors, pour $\varepsilon > 0$:

$$\begin{aligned}
\sum_{n\in\mathbb{N}^*}\mathbb{P}\left(\left|T_{k(n)} - \mathbb{E}\left(T_{k(n)}\right)\right| > \varepsilon k(n)\right) &\leqslant \frac{1}{\varepsilon^2}\sum_{n\in\mathbb{N}^*}\frac{1}{k(n)^2}\operatorname{Var}\left(T_{k(n)}\right) \leqslant \frac{1}{\varepsilon^2}\sum_{n\in\mathbb{N}^*}\frac{1}{k(n)^2}\sum_{1\leqslant m\leqslant k(n)}\operatorname{Var}\left(Y_m\right) \\
&= \frac{1}{\varepsilon^2}\sum_{m\in\mathbb{N}^*}\operatorname{Var}\left(Y_m\right)\sum_{\substack{n\in\mathbb{N}^* \\ k(n)\geqslant m}}\frac{1}{k(n)^2} = \frac{1}{\varepsilon^2}\sum_{m\in\mathbb{N}^*}\operatorname{Var}\left(Y_m\right)\sum_{n\in\mathbb{N}^*}\frac{4}{\alpha^{2n}} \\
&\leqslant \frac{4}{\varepsilon^2\left(1-\alpha^{-2}\right)}\sum_{m\in\mathbb{N}^*}\frac{\operatorname{Var}\left(Y_m\right)}{m^2} \leqslant \frac{4}{\varepsilon^2\left(1-\alpha^{-2}\right)}\sum_{m\in\mathbb{N}^*}\frac{1}{m^2}\mathbb{E}\left(Y_m^2\right) \\
&= \frac{8}{\varepsilon^2\left(1-\alpha^{-2}\right)}\sum_{m\in\mathbb{N}^*}\frac{1}{m^2}\int_0^\infty t\mathbb{1}_{[0,m]}(t)\mathbb{P}\left(X > t\right)\,\mathrm{d}t \\
&= \frac{8}{\varepsilon^2\left(1-\alpha^{-2}\right)}\int_0^\infty\left(\sum_{m=\lceil t\rceil}^\infty\frac{1}{m^2}\right)t\mathbb{P}\left(X > t\right)\,\mathrm{d}t.
\end{aligned}$$

En utilisant le fait que $\sum_{m=\lceil t\rceil}^{\infty} \frac{1}{m^2} \sim \frac{1}{t}$, on obtient l'existence d'une constante K t.q.

$$\sum_{n\in\mathbb{N}^*} \mathbb{P}\left(\left|T_{k(n)} - \mathbb{E}\left(T_{k(n)}\right)\right| > \varepsilon k(n)\right) \leqslant K\mathbb{E}(X) < +\infty.$$

Par Borel-Cantelli, on obtient :

$$\forall \varepsilon > 0,\ \mathbb{P}\left(\liminf_{n\to+\infty}\left(\frac{\left|T_{k(n)} - \mathbb{E}\left(T_{k(n)}\right)\right|}{k(n)} \leqslant \varepsilon\right)\right) = 1.$$

On en déduit donc que $\frac{T_{k(n)} - \mathbb{E}\left(T_{k(n)}\right)}{k(n)} \xrightarrow[n\to+\infty]{\text{p.s.}} 0$. Or, par convergence dominée, $\mathbb{E}\left(Y_n\right) \xrightarrow[n\to+\infty]{} \mathbb{E}(X)$, d'où :

$$\frac{T_{k(n)}}{k(n)} \xrightarrow[n\to+\infty]{\text{p.s.}} \mathbb{E}(X).$$

Étape 3. Soit maintenant $n \in \mathbb{N}^*$ et $m \in \mathbb{N}^*$ t.q. $k(m) \leqslant n \leqslant k(m+1)$. Alors :

$$\frac{T_{k(m)}}{k(m)} \cdot \frac{k(m)}{k(m+1)} \leqslant \frac{T_n}{n} \leqslant \frac{T_{k(m+1)}}{k(m+1)} \cdot \frac{k(m+1)}{k(m)}.$$

On en déduit :

$$\frac{1}{\alpha}\mathbb{E}(X) \leqslant \liminf_{n\to+\infty} \frac{T_n}{n} \leqslant \limsup_{n\to+\infty} \frac{T_n}{n} \leqslant \alpha\mathbb{E}(X).$$

En faisant tendre $\alpha \to 1$ (en prenant un nombre dénombrable de valeurs), on obtient $\frac{T_n}{n} \xrightarrow[n\to+\infty]{\text{p.s.}} \mathbb{E}(X)$. □

Démonstration (Seconde méthode)**.** Pour $n \in \mathbb{N}$, posons $S_n = \sum_{k=1}^{n} X_k$. Pour $a > \mathbb{E}(X)$, on considère :

$$M = \sup_{n\in\mathbb{N}} \left(S_n - na\right) \in [0, +\infty].$$

Si on montre que M est presque-sûrement finie pour tout $a > \mathbb{E}(X)$, on aura $\forall n \in \mathbb{N},\ S_n \leqslant M + na$ donc :

$$\forall a > \mathbb{E}(X),\ \limsup_{n\to+\infty} \frac{S_n}{n} \leqslant a.$$

En faisant tendre $a \to \mathbb{E}(X)$ (en prenant un nombre dénombrable de valeurs), $\limsup_{n\to+\infty} \frac{S_n}{n} \leqslant \mathbb{E}(X)$ presque-sûrement, puis en appliquant le même argument à $\left(-X_n\right)_{n\in\mathbb{N}^*}$, $\liminf_{n\to+\infty} \frac{S_n}{n} \geqslant \mathbb{E}(X)$ presque-sûrement, d'où le résultat. Reste donc à prouver que pour tout $a > \mathbb{E}(X)$, M est presque-sûrement finie. Notons que :

$$\forall k \in \mathbb{N},\ (M < +\infty) = \left(\sup_{n>k}\left(X_{k+1} + \cdots + X_n\right) - (n-k)a < +\infty\right) \in \bigvee_{j>k} \sigma\left(X_j\right).$$

Donc $(M < +\infty) \in \bigcap_{k\in\mathbb{N}} \bigvee_{j>k} \sigma\left(X_j\right)$. C'est donc un événement asymptotique en les $\left(X_n\right)_{n\in\mathbb{N}^*}$. Comme les $\left(X_n\right)_{n\in\mathbb{N}^*}$, la loi du zéro-un de Kolmogorov (théorème 2.10.6) donne $\mathbb{P}\left(M < +\infty\right) \in \{0, 1\}$. Supposons par l'absurde que $\mathbb{P}\left(M < +\infty\right) = 0$, i.e. $\mathbb{P}\left(M = +\infty\right) = 1$. Pour $k \in \mathbb{N}$, on pose :

$$M_k = \sup_{0\leqslant n\leqslant k} \left(S_n - nk\right) \qquad \text{et} \qquad M'_k = \sup_{0\leqslant n\leqslant k} \left(S_{n+1} - X_1 - nk\right).$$

Alors M_k et M'_k ont même loi. On a $M = \lim_{k\to+\infty} M_k$; posons $M' = \lim_{k\to+\infty} M'_k$. Alors M et M' ont même loi. Par ailleurs :

$$\forall k \in \mathbb{N},\ M_{k+1} = \max\left(0, M'_k + X_1 - a\right) = M'_k - \min\left(M'_k, a - X_1\right).$$

D'où $\mathbb{E}\left(\min\left(M'_k, a - X_1\right)\right) = \mathbb{E}\left(M'_k\right) - \mathbb{E}\left(M_{k+1}\right) \leqslant 0$. Or $M'_k \geqslant 0$, donc on a la domination $\left|\min\left(M'_k, a - X_1\right)\right| \leqslant |a - X_1|$. Par convergence dominée :

$$\mathbb{E}\left(\min\left(M', a - X_1\right)\right) = \lim_{k\to+\infty} \mathbb{E}\left(\min\left(M'_k, a - X_1\right)\right) \leqslant 0.$$

Mais $M' = +\infty$ presque-sûrement, donc $\mathbb{E}\left(\min\left(M', a - X_1\right)\right) = \mathbb{E}\left(a - X_1\right) > 0$ comme $a > \mathbb{E}(X)$. C'est absurde. □

Corollaire 3.2.3. *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles indépendantes et suivant toutes la loi d'une variable aléatoire X. On suppose que X admet un moment d'ordre 1. Alors :*

$$\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow[n\to+\infty]{L^1} \mathbb{E}(X).$$

Remarque 3.2.4. *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles positives indépendantes et suivant toutes la loi d'une variable aléatoire X t.q. $\mathbb{E}(X)=+\infty$. Alors $\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow[n\to+\infty]{\text{p.s.}} +\infty$.*

Exemple 3.2.5. *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes suivant la loi $\mathcal{C}(c)$, avec $c\in\mathbb{R}_+^*$ (c.f. exemple 1.4.6). Le calcul des fonctions caractéristiques montre que, pour tout $n\in\mathbb{N}^*$, $\frac{X_1+\cdots+X_n}{n}$ suit la loi $\mathcal{C}(c)$. Dans ce cas, la loi des grands nombres ne s'applique pas (car la loi de Cauchy n'admet pas de moment d'ordre 1).*

3.3 Convergence en loi

Notation 3.3.1. *On notera $\mathcal{C}_b^0\left(\mathbb{R}^d\right)$ l'espace des fonctions continues bornées $\mathbb{R}^d\to\mathbb{R}$, muni de $\|\cdot\|_\infty$. On notera de plus $\mathcal{C}_c^0\left(\mathbb{R}^d\right)\subset\mathcal{C}_b^0\left(\mathbb{R}^d\right)$ le sous-espace des fonctions continues à support compact.*

Définition 3.3.2 (Convergence étroite)**.** *Soit $(\mu_n)_{n\in\mathbb{N}}$ et μ des mesures de probabilité sur $\mathbb{R}^d$. On dit que $\mu_n \xrightarrow[n\to+\infty]{\text{e}} \mu$ ($(\mu_n)_{n\in\mathbb{N}}$* converge étroitement *vers μ) lorsque :*

$$\forall\varphi\in\mathcal{C}_b^0\left(\mathbb{R}^d\right),\ \int_{\mathbb{R}^d}\varphi\ \mathrm{d}\mu_n \xrightarrow[n\to+\infty]{} \int_{\mathbb{R}^d}\varphi\ \mathrm{d}\mu.$$

Définition 3.3.3 (Convergence en loi)**.** *Soit $(X_n)_{n\in\mathbb{N}}$ et X des variables aléatoires à valeurs dans $\mathbb{R}^d$. On dit que $X_n \xrightarrow[n\to+\infty]{\text{loi}} X$ ($(X_n)_{n\in\mathbb{N}}$* converge *vers X* en loi*) lorsque $\mathbb{P}_{X_n} \xrightarrow[n\to+\infty]{\text{e}} \mathbb{P}_X$, i.e.*

$$\forall\varphi\in\mathcal{C}_b^0\left(\mathbb{R}^d\right),\ \mathbb{E}\left(\varphi\left(X_n\right)\right) \xrightarrow[n\to+\infty]{} \mathbb{E}\left(\varphi(X)\right).$$

Remarque 3.3.4.

(i) *On peut parler de convergence en loi de variables aléatoires définies sur des espaces distincts.*

(ii) *La variable aléatoire "limite", en cas de convergence en loi, n'est pas unique.*

Proposition 3.3.5. *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires à valeurs dans $\mathbb{N}$, et X une variable aléatoire à valeurs dans $\mathbb{N}$. S'équivalent :*

(i) $X_n \xrightarrow[n\to+\infty]{\text{loi}} X$.

(ii) $\forall k\in\mathbb{N}$, $\mathbb{P}\left(X_n=k\right) \xrightarrow[n\to+\infty]{} \mathbb{P}\left(X=k\right)$.

Exemple 3.3.6.

(i) *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles t.q. pour tout $n\in\mathbb{N}$, X_n admet une densité p_n. On suppose que $(p_n)_{n\in\mathbb{N}}$ converge presque-partout et que $\exists q\in L^1, \forall n\in\mathbb{N}, |p_n|\leqslant q$. Alors $(X_n)_{n\in\mathbb{N}}$ converge en loi.*

(ii) *Pour $n\in\mathbb{N}^*$, soit $X_n\sim\mathcal{U}\left(\left\{0,\frac{1}{n},\frac{2}{n},\ldots,1\right\}\right)$. Alors $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers $\mathcal{U}\left([0,1]\right)$.*

(iii) *Pour $n\in\mathbb{N}$, soit $X_n\sim\mathcal{N}\left(0,\sigma_n^2\right)$, où $(\sigma_n)_{n\in\mathbb{N}}\in\left(\mathbb{R}_+^*\right)^{\mathbb{N}}$ est t.q. $\sigma_n \xrightarrow[n\to+\infty]{} 0$. Alors $(X_n)_{n\in\mathbb{N}}$ converge en loi vers la mesure de Dirac en 0.*

3.4 Convergence des mesures empiriques

Définition 3.4.1 (Mesure empirique)**.** *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires à valeurs dans $\mathbb{R}^d$. Pour $\omega \in \Omega$ et $n \in \mathbb{N}^*$, on définit la* mesure empirique *de l'échantillon $(X_1(\omega), \dots, X_n(\omega))$ par :*

$$\mu_{n,\omega} = \frac{1}{n} \sum_{k=1}^{n} \delta_{X_k(\omega)},$$

où δ_x est la mesure de Dirac en x pour tout $x \in \mathbb{R}^d$.

Lemme 3.4.2. *Soit $(\mu_n)_{n\in\mathbb{N}}$ et μ des mesures de probabilité sur $\mathbb{R}^d$. Soit $H \subset \mathcal{C}_b^0\left(\mathbb{R}^d\right)$ t.q. $\overline{H} \supset \mathcal{C}_c^0\left(\mathbb{R}^d\right)$ (pour $\|\cdot\|_\infty$). S'équivalent :*

(i) $\mu_n \xrightarrow[n\to+\infty]{\text{e}} \mu$.

(ii) $\forall f \in \mathcal{C}_c^0\left(\mathbb{R}^d\right),\ \int_{\mathbb{R}^d} f\ \mathrm{d}\mu_n \xrightarrow[n\to+\infty]{} \int_{\mathbb{R}^d} f\ \mathrm{d}\mu$.

(iii) $\forall f \in H,\ \int_{\mathbb{R}^d} f\ \mathrm{d}\mu_n \xrightarrow[n\to+\infty]{} \int_{\mathbb{R}^d} f\ \mathrm{d}\mu$.

Théorème 3.4.3 (Théorème de Glivenko-Cantelli)**.** *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes à valeurs dans $\mathbb{R}^d$ et suivant toutes la loi d'une variable aléatoire X. Alors, pour presque tout $\omega \in \Omega$:*

$$\mu_{n,\omega} \xrightarrow[n\to+\infty]{\text{e}} \mathbb{P}_X.$$

Démonstration. On veut montrer que, pour presque tout $\omega \in \Omega$:

$$\forall f \in \mathcal{C}_b^0\left(\mathbb{R}^d\right),\ \underbrace{\int_{\mathbb{R}^d} f\ \mathrm{d}\mu_{n,\omega}}_{\frac{1}{n}\sum_{i=1}^n f(X_i(\omega))} \xrightarrow[n\to+\infty]{} \underbrace{\int_{\mathbb{R}^d} f\ \mathrm{d}\mathbb{P}_X}_{\mathbb{E}(f(X))}.$$

Or, la loi forte des grands nombres (théorème 3.2.2) donne :

$$\forall f \in \mathcal{C}_b^0\left(\mathbb{R}^d\right),\ \frac{1}{n}\sum_{i=1}^{n} f\left(X_i\right) \xrightarrow[n\to+\infty]{\text{p.s.}} \mathbb{E}\left(f(X)\right).$$

On se donne maintenant H un sous-ensemble dénombrable et dense de $\mathcal{C}_c^0\left(\mathbb{R}^d\right)$. On a alors $\forall f \in H,\ \frac{1}{n}\sum_{i=1}^n f\left(X_i\right) \xrightarrow[n\to+\infty]{\text{p.s.}} \mathbb{E}\left(f(X)\right)$. H étant dénombrable, on en déduit que pour presque tout $\omega \in \Omega$: $\forall f \in H,\ \frac{1}{n}\sum_{i=1}^n f\left(X_i(\omega)\right) \xrightarrow[n\to+\infty]{} \mathbb{E}\left(f(X)\right)$. Le lemme 3.4.2 permet maintenant de conclure. □

Théorème 3.4.4 (Théorème central limite)**.** *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles indépendantes et suivant toutes la loi d'une variable aléatoire X. On suppose que X admet un moment d'ordre 2. Notons $\sigma = \sqrt{\mathrm{Var}(X)}$, et posons $S_n = X_1 + \cdots + X_n$ pour $n \in \mathbb{N}^*$. Alors :*

$$\frac{S_n - n\mathbb{E}(X)}{\sigma\sqrt{n}} \xrightarrow[n\to+\infty]{\text{loi}} \mathcal{N}(0,1).$$

Démonstration. Quitte à remplacer X par $\frac{1}{\sigma}\left(X - \mathbb{E}(X)\right)$, on suppose que $\mathbb{E}(X) = 0$ et $\sigma = 1$. On calcule alors les fonctions caractéristiques :

$$\forall \xi \in \mathbb{R},\ \Phi_{S_n/\sqrt{n}}(\xi) = \Phi_X\left(\frac{\xi}{\sqrt{n}}\right)^n.$$

Or X admet un moment d'ordre 2 donc Φ_X est de classe $\mathcal{C}^2$, d'où, à $\xi \in \mathbb{R}$ fixé :

$$\Phi_X\left(\frac{\xi}{\sqrt{n}}\right) = 1 - \frac{\xi^2}{2n} + o\left(\frac{1}{n}\right).$$

Or il existe $n_0 \in \mathbb{N}$ t.q. $\forall n \geqslant n_0$, $\left|\Phi_X\left(\frac{\xi}{\sqrt{n}}\right) - 1\right| < 1$. Si on note $\operatorname{Log} : \mathbb{C}\backslash\mathbb{R}_- \to \mathbb{C}$ la détermination principale du logarithme, on a alors :

$$\forall \xi \in \mathbb{R},\ \Phi_{S_n/\sqrt{n}}(\xi) = \exp\left(n \operatorname{Log}\left(\Phi_X\left(\frac{\xi}{\sqrt{n}}\right)\right)\right) \xrightarrow[n\to+\infty]{} \exp\left(-\frac{\xi^2}{2}\right) = \Phi_Y(\xi),$$

où $Y \sim \mathcal{N}(0,1)$. D'après le théorème 3.5.1, on a bien $\frac{S_n}{\sqrt{n}} \xrightarrow[n\to+\infty]{\text{loi}} Y$. □

3.5 Théorème de Lévy

Théorème 3.5.1 (Théorème de Lévy)**.**

(i) *Soit $(\mu_n)_{n\in\mathbb{N}}$ et μ des mesures de probabilité sur $\mathbb{R}^d$. Alors :*

$$\left(\mu_n \xrightarrow[n\to+\infty]{\text{e}} \mu\right) \Longleftrightarrow \left(\forall \xi \in \mathbb{R}^d,\ \hat{\mu}_n(\xi) \xrightarrow[n\to+\infty]{} \hat{\mu}(\xi)\right),$$

où $\hat{\mu}$ désigne la transformée de Fourier de μ (c.f. définition 1.9.8).

(ii) *Soit $(X_n)_{n\in\mathbb{N}}$ et X des variables aléatoires à valeurs dans $\mathbb{R}^d$. Alors :*

$$\left(X_n \xrightarrow[n\to+\infty]{\text{loi}} X\right) \Longleftrightarrow \left(\forall \xi \in \mathbb{R}^d,\ \Phi_{X_n}(\xi) \xrightarrow[n\to+\infty]{} \Phi_X(\xi)\right).$$

Démonstration. (i) Il est clair que si $\mu_n \xrightarrow[n\to+\infty]{\text{e}} \mu$, alors $\hat{\mu}_n \xrightarrow[n\to+\infty]{} \hat{\mu}$ simplement, car pour tout $\xi \in \mathbb{R}^d$, $x \longmapsto \frac{e^{i\langle\xi|x\rangle}}{(2\pi)^{d/2}}$ est une fonction continue bornée. Réciproquement, supposons que $\hat{\mu}_n \xrightarrow[n\to+\infty]{} \hat{\mu}$ simplement. On se place dans le cas où $d = 1$. Pour $\sigma \in \mathbb{R}_+^*$, on pose $g_\sigma : x \in \mathbb{R} \longmapsto \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right)$. On note :

$$H = \left\{g_\sigma * f,\ \sigma \in \mathbb{R}_+^*,\ f \in \mathcal{C}_c^0(\mathbb{R})\right\} \subset \mathcal{C}_b^0(\mathbb{R}).$$

Pour tout $f \in \mathcal{C}_c^0(\mathbb{R})$, on a $g_\sigma * f \xrightarrow[\sigma\to 0]{} f$ uniformément, ce qui prouve que $\overline{H} \supset \mathcal{C}_c^0(\mathbb{R})$. Selon le lemme 3.4.2, il suffit de prouver que $\forall \varphi \in H$, $\int_{\mathbb{R}} \varphi \,\mathrm{d}\mu_n \xrightarrow[n\to+\infty]{} \int_{\mathbb{R}} \varphi \,\mathrm{d}\mu$. Pour cela, notons que si $\sigma \in \mathbb{R}_+^*$ et $f \in \mathcal{C}_c^0(\mathbb{R})$, on a pour toute mesure ν, selon le lemme 1.9.9 :

$$\int_{\mathbb{R}} g_\sigma * f \,\mathrm{d}\nu = \frac{1}{\sigma\sqrt{2\pi}} \int_{\mathbb{R}} f(x) \left(\int_{\mathbb{R}} e^{-i\xi x} g_{1/\sigma}(\xi) \hat{\nu}(-\xi) \,\mathrm{d}\xi\right) \mathrm{d}x.$$

Comme $\hat{\mu}_n \xrightarrow[n\to+\infty]{} \hat{\mu}$ simplement, on en déduit par convergence dominée que $\int_{\mathbb{R}} g_\sigma * f \,\mathrm{d}\mu_n \xrightarrow[n\to+\infty]{} \int_{\mathbb{R}} g_\sigma * f \,\mathrm{d}\mu$, d'où le résultat. (ii) C'est un corollaire direct de (i). □

3.6 Autres caractérisations de la convergence en loi

Théorème 3.6.1 (Théorème porte-manteau)**.** *Soit $(\mu_n)_{n\in\mathbb{N}}$ et μ des mesures de probabilité sur $\mathbb{R}^d$. Sont équivalentes :*

(i) $\mu_n \xrightarrow[n\to+\infty]{\text{e}} \mu$.

(ii) $\forall G$ *ouvert de* $\mathbb{R}^d$, $\liminf_{n\to+\infty} \mu_n(G) \geqslant \mu(G)$.

(iii) $\forall F$ *fermé de* $\mathbb{R}^d$, $\limsup_{n\to+\infty} \mu_n(F) \leqslant \mu(F)$.

(iv) $\forall B \in \operatorname{Bor}\left(\mathbb{R}^d\right)$, $\mu(\partial B) = 0 \Longrightarrow \mu_n(B) \xrightarrow[n\to+\infty]{} \mu(B)$.

Démonstration. (i) $\Rightarrow$ (ii) Soit G un ouvert de $\mathbb{R}^d$. Soit $(\varphi_k)_{k\in\mathbb{N}}$ une suite croissante de fonctions continues positives convergeant simplement vers $\mathbb{1}_G$, par exemple $\varphi_k : x \longmapsto \min\left(1, k\cdot d\left(x, \mathbb{R}^d \backslash G\right)\right)$. Ainsi :

$$\begin{aligned}\liminf_{n\to+\infty} \mu_n(G) &= \liminf_{n\to+\infty} \int_{\mathbb{R}} \sup_{k\in\mathbb{N}} \varphi_k \, \mathrm{d}\mu_n = \liminf_{n\to+\infty} \sup_{k\in\mathbb{N}} \int_{\mathbb{R}} \varphi_k \, \mathrm{d}\mu_n \\ &\geqslant \sup_{k\in\mathbb{N}} \liminf_{n\to+\infty} \int_{\mathbb{R}} \varphi_k \, \mathrm{d}\mu_n = \sup_{k\in\mathbb{N}} \int_{\mathbb{R}} \varphi_k \, \mathrm{d}\mu = \int_{\mathbb{R}} \sup_{k\in\mathbb{N}} \varphi_k \, \mathrm{d}\mu = \mu(G).\end{aligned}$$

(ii) $\Leftrightarrow$ (iii) Il suffit de passer au complémentaire. (ii) et (iii) $\Rightarrow$ (iv) Soit $B \in \mathrm{Bor}\left(\mathbb{R}^d\right)$ t.q. $\mu(\partial B) = 0$. Alors $\mu(B) = \mu\left(\overline{B}\right) = \mu\left(\mathring{B}\right)$, d'où :

$$\mu(B) = \mu\left(\mathring{B}\right) \leqslant \liminf_{n\to+\infty} \mu_n\left(\mathring{B}\right) \leqslant \liminf_{n\to+\infty} \mu_n(B) \leqslant \limsup_{n\to+\infty} \mu_n(B) \leqslant \limsup_{n\to+\infty} \mu_n\left(\overline{B}\right) \leqslant \mu\left(\overline{B}\right) = \mu(B).$$

(iv) $\Rightarrow$ (i) Soit $\varphi \in \mathcal{C}_b^0\left(\mathbb{R}^d\right)$. Quitte à écrire $\varphi = \varphi^+ - \varphi^-$, on peut supposer que $\varphi \geqslant 0$. Soit $K > 0$ t.q. $\forall x \in \mathbb{R}^d,\ 0 \leqslant \varphi(x) \leqslant K$. On a :

$$\int_{\mathbb{R}^d} \varphi \, \mathrm{d}\mu = \int_0^K \mu\left(\underbrace{\varphi^{-1}\left([t, +\infty[\right)}_{A_t}\right) \mathrm{d}t.$$

Et on a la même égalité en remplaçant μ par μ_n. Or $\partial A_t \subset \varphi^{-1}(\{t\})$ pour tout $t \in [0, K]$. Et $\varphi^{-1}(\{t\})$ est de mesure strictement positive (pour μ) pour au plus un nombre dénombrable de valeurs de t. On en déduit que pour presque tout $t \in [0, K]$ (au sens de la mesure de Lebesgue), $\mu(\partial A_t) = 0$. Par hypothèse, on a donc $\mu_n(A_t) \xrightarrow[n\to+\infty]{} \mu(A_t)$ pour presque tout $t \in [0, K]$. Par convergence dominée,il vient $\int_{\mathbb{R}^d} \varphi \, \mathrm{d}\mu_n \xrightarrow[n\to+\infty]{} \int_{\mathbb{R}^d} \varphi \, \mathrm{d}\mu$. □

Corollaire 3.6.2. *Soit $(X_n)_{n\in\mathbb{N}}$ et X des variables aléatoires réelles. Alors :*

$$\left(X_n \xrightarrow[n\to+\infty]{\text{loi}} X\right) \Longleftrightarrow \left(\forall x \in \mathbb{R},\ F_X\ \mathcal{C}^0 \text{ en } x \Longrightarrow F_{X_n}(x) \xrightarrow[n\to+\infty]{} F_X(x)\right).$$

Théorème 3.6.3. *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles indépendantes et suivant toutes la loi d'une variable aléatoire X. On suppose que X admet un moment d'ordre 2. Notons $\sigma = \sqrt{\mathrm{Var}(X)}$, et posons $S_n = X_1 + \cdots + X_n$ pour $n \in \mathbb{N}^*$. Alors :*

$$\forall (a, b) \in \mathbb{R}^2,\ a \leqslant b \Longrightarrow \mathbb{P}\left(a \leqslant \frac{S_n - n\mathbb{E}(X)}{\sigma\sqrt{n}} \leqslant b\right) \xrightarrow[n\to+\infty]{} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} \, \mathrm{d}x.$$

3.7 Convergence en probabilité

Définition 3.7.1 (Convergence en probabilité)**.** *Soit $(X_n)_{n\in\mathbb{N}}$ et X des variables aléatoires réelles définies sur Ω. On dit que $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$ ($(X_n)_{n\in\mathbb{N}}$ converge* en probabilité *vers X) lorsque :*

$$\forall \varepsilon > 0,\ \mathbb{P}\left(|X_n - X| > \varepsilon\right) \xrightarrow[n\to+\infty]{} 0.$$

Proposition 3.7.2. *On note $\mathcal{L}^0$ l'espace des variables aléatoires $\Omega \to \mathbb{R}$ et L^0 le quotient de $\mathcal{L}^0$ par la relation d'égalité presque-partout. On définit :*

$$d : (X, Y) \in \left(L^0\right)^2 \longmapsto \mathbb{E}\left(\min\left(1, |X - Y|\right)\right).$$

Alors :

(i) *d est une distance sur L^0.*

(ii) *Soit* $(X_n)_{n\in\mathbb{N}}$ *et* X *des variables aléatoires réelles définies sur* Ω*. Alors :*

$$\left(X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X\right) \iff \left(d\left(X_n, X\right) \xrightarrow[n\to+\infty]{} 0\right).$$

(iii) *L'espace métrique* (L^0, d) *est complet.*

Proposition 3.7.3. *Soit* $(X_n)_{n\in\mathbb{N}}$ *et* X *des variables aléatoires réelles définies sur* Ω*. Si* $X_n \xrightarrow[n\to+\infty]{\mathbb{P}} X$*, alors on peut extraire de* $(X_n)_{n\in\mathbb{N}}$ *une sous-suite convergeant presque-sûrement vers* X*.*

4 Transformée de Fourier

4.1 Définitions et premières propriétés

Définition 4.1.1 (Transformée de Fourier d'une fonction L^1)**.** *Soit* $f \in L^1\left(\mathbb{R}^d\right)$*. La* transformée de Fourier *de* f *est définie par :*

$$\hat{f} : \xi \in \mathbb{R}^d \longmapsto \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-i\langle \xi | x\rangle} f(x)\, \mathrm{d}x.$$

Notation 4.1.2. *Soit* $y \in \mathbb{R}^d$*.*

(i) *On définit un opérateur de* translation $\tau_y : L^1\left(\mathbb{R}^d\right) \to L^1\left(\mathbb{R}^d\right)$ *par :*

$$\forall f \in L^1\left(\mathbb{R}^d\right),\ \forall x \in \mathbb{R}^d,\ \tau_y f(x) = f(x-y).$$

(ii) *On définit un opérateur de* modulation $e_y : L^1\left(\mathbb{R}^d\right) \to L^1\left(\mathbb{R}^d\right)$ *par :*

$$\forall f \in L^1\left(\mathbb{R}^d\right),\ \forall x \in \mathbb{R}^d,\ e_y f(x) = e^{i\langle y|x\rangle} f(x).$$

(iii) *On définit un opérateur de* réflexion $R : L^1\left(\mathbb{R}^d\right) \to L^1\left(\mathbb{R}^d\right)$ *par :*

$$\forall f \in L^1\left(\mathbb{R}^d\right),\ \forall x \in \mathbb{R}^d,\ Rf(x) = f(-x).$$

Proposition 4.1.3.

(i) *On a :*

$$\forall f \in L^1\left(\mathbb{R}^d\right),\ \widehat{\tau_y f} = e_{-y}\hat{f} \qquad et \qquad \widehat{e_y f} = \tau_f \hat{f}.$$

(ii) *Soit* $a \in \mathbb{R}^*$*. On pose* $g : x \in \mathbb{R}^d \longmapsto f\left(\frac{x}{a}\right)$*. Alors :*

$$\forall \xi \in \mathbb{R}^d,\ \hat{g}(\xi) = |a|^d \hat{f}(a\xi).$$

(iii) *On a :*

$$\forall f \in L^1\left(\mathbb{R}^d\right),\ \overline{\hat{f}} = R\hat{\overline{f}}.$$

Proposition 4.1.4. *Soit* $f \in L^1\left(\mathbb{R}^d\right)$*. Alors* $\hat{f}$ *est uniformément continue et :*

$$\hat{f}(\xi) \xrightarrow[\|\xi\|\to+\infty]{} 0.$$

De plus, $\left\|\hat{f}\right\|_\infty \leqslant \frac{1}{(2\pi)^{d/2}} \|f\|_1$*. En particulier, l'application :*

$$\left|\begin{array}{c} L^1\left(\mathbb{R}^d\right) \longrightarrow L^\infty\left(\mathbb{R}^d\right) \\ f \longmapsto \hat{f} \end{array}\right.$$

est linéaire continue.

Proposition 4.1.5 (Formule de réciprocité)**.** *Soit* $(f, g) \in L^1\left(\mathbb{R}^d\right)^2$*. Alors* $\hat{f}g$ *et* $f\hat{g}$ *sont intégrables et :*

$$\int_{\mathbb{R}^d} \hat{f} g = \int_{\mathbb{R}^d} f\hat{g}.$$

4.2 Propriétés de régularités

Notation 4.2.1 (Multi-indices)**.** *Soit $\alpha \in \mathbb{N}^d$.*

(i) *On note $|\alpha| = \sum_{i=1}^d \alpha_i$.*

(ii) *Pour $x \in \mathbb{R}^d$, on note $x^\alpha = x_1^{\alpha_1} \cdots x_d^{\alpha_d}$.*

Notation 4.2.2. *Soit F un espace de Banach, U un ouvert de $\mathbb{R}^m$, $f : U \to F$, $a \in U$. Supposons f k fois différentiable en a. Pour $\alpha \in \mathbb{N}^m$ t.q. $|\alpha| \leqslant k$, on pose :*

$$\partial^\alpha f(a) = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \left(\frac{\partial}{\partial x_2}\right)^{\alpha_2} \cdots \left(\frac{\partial}{\partial x_m}\right)^{\alpha_m} f(a).$$

Proposition 4.2.3. *Soit $f \in L^1\left(\mathbb{R}^d\right)$. On suppose qu'il existe un entier $k \in \mathbb{N}^*$ t.q. $x \longmapsto x^k f(x)$ est intégrable. Alors $\hat{f} \in \mathcal{C}^k\left(\mathbb{R}^d\right)$ et pour tout multi-indice $\alpha \in \mathbb{N}^d$ t.q. $|\alpha| \leqslant k$, on a :*

$$\partial^\alpha \hat{f}(\xi) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-i\langle \xi | x \rangle} (-ix)^\alpha f(x) \, \mathrm{d}x.$$

Proposition 4.2.4. *Soit $f \in L^1\left(\mathbb{R}^d\right)$. On suppose qu'il existe un entier $k \in \mathbb{N}^*$ t.q. $f \in \mathcal{C}^k\left(\mathbb{R}^d\right)$ et $\forall \alpha \in \mathbb{N}^d,\ |\alpha| \leqslant k \Longrightarrow \partial^\alpha f \in L^1\left(\mathbb{R}^d\right)$. Alors, pour tout $\alpha \in \mathbb{N}^d$ t.q. $|\alpha| \leqslant k$, on a :*

$$\forall \xi \in \mathbb{R}^d,\ \widehat{\partial^\alpha f}(\xi) = (i\xi)^\alpha \hat{f}(\xi).$$

Corollaire 4.2.5. *Soit $f \in L^1\left(\mathbb{R}^d\right)$. On suppose qu'il existe un entier $k \in \mathbb{N}^*$ t.q. $f \in \mathcal{C}^k\left(\mathbb{R}^d\right)$ et $\forall \alpha \in \mathbb{N}^d,\ |\alpha| \leqslant k \Longrightarrow \partial^\alpha f \in L^1\left(\mathbb{R}^d\right)$. Alors :*

$$\hat{f}(\xi) = o_{+\infty}\left(\frac{1}{\|\xi\|^k}\right).$$

4.3 Convolution

Proposition 4.3.1. *Soit $(f, g) \in L^1\left(\mathbb{R}^d\right)^2$. Alors :*

$$\widehat{f * g} = (2\pi)^{d/2} \hat{f} \cdot \hat{g},$$

avec $\forall x \in \mathbb{R}^d,\ (f * g)(x) = \int_{\mathbb{R}^d} f(x-y) g(y) \, \mathrm{d}y$.

Exemple 4.3.2. *Pour $\sigma \in \mathbb{R}_+^*$, on pose :*

$$g_\sigma : x \in \mathbb{R}^d \longmapsto \frac{1}{\sigma^d} \exp\left(-\left(\frac{\|x\|}{\sigma}\right)^2\right).$$

Alors :

$$\hat{g}_\sigma = \frac{1}{\sigma^d} g_{1/\sigma}.$$

4.4 Formule d'inversion

Théorème 4.4.1. *Soit $f \in L^1\left(\mathbb{R}^d\right)$ t.q. $\hat{f} \in L^1\left(\mathbb{R}^d\right)$. Alors :*

$$\widehat{\left(\hat{f}\right)} = Rf,$$

où R est l'opérateur défini dans la notation 4.1.2.

Corollaire 4.4.2. *L'application :*

$$\left| \begin{array}{c} L^1\left(\mathbb{R}^d\right) \longrightarrow L^\infty\left(\mathbb{R}^d\right) \\ f \longmapsto \hat{f} \end{array} \right.$$

est linéaire continue et injective.

4.5 Transformée de Fourier dans L^2

Proposition 4.5.1. *Soit $f \in \mathcal{C}_c^\infty\left(\mathbb{R}^d\right) \subset L^1\left(\mathbb{R}^d\right) \cap L^2\left(\mathbb{R}^d\right)$. Alors $\hat{f} \in L^2\left(\mathbb{R}^d\right)$ et :*

$$\left\|\hat{f}\right\|_2 = \|f\|_2 .$$

Définition 4.5.2 (Transformée de Fourier d'une fonction L^2)**.** *L'application :*

$$\left| \begin{aligned} \mathcal{C}_c^\infty\left(\mathbb{R}^d\right) &\longrightarrow L^2\left(\mathbb{R}^d\right) \\ f &\longmapsto \hat{f} \end{aligned} \right.$$

est une isométrie (peut-être pas surjective) donc une application uniformément continue. Comme $\mathcal{C}_c^\infty\left(\mathbb{R}^d\right)$ est dense dans $L^2\left(\mathbb{R}^d\right)$ et comme $L^2\left(\mathbb{R}^d\right)$ est complet, cette application admet un unique prolongement uniformément continue $\mathcal{F} : L^2\left(\mathbb{R}^d\right) \to L^2\left(\mathbb{R}^d\right)$. On dit que $\mathcal{F}$ est la transformée de Fourier dans L^2 *(ou la transformée de Fourier-Plancherel).*

Théorème 4.5.3.

(i) $\forall f \in L^1\left(\mathbb{R}^d\right) \cap L^2\left(\mathbb{R}^d\right)$, $\mathcal{F}f = \hat{f}$.

(ii) $\mathcal{F} : L^2\left(\mathbb{R}^d\right) \to L^2\left(\mathbb{R}^d\right)$ *est une isométrie bijective.*

Définition 4.5.4 (Écart-type)**.** *Soit $f \in L^2(\mathbb{R}) \setminus \{0\}$. Alors la fonction $\left(\frac{|f|}{\|f\|_2}\right)^2$ définit une densité de probabilité sur $\mathbb{R}$. On note $\sigma(f) \in [0, +\infty]$ l'écart-type associé.*

Théorème 4.5.5 (Principe d'incertitude de Heisenberg)**.** *Soit $f \in L^2(\mathbb{R}) \setminus \{0\}$. Alors :*

$$\sigma(f)\sigma\left(\mathcal{F}f\right) \geqslant \frac{1}{2},$$

avec égalité ssi *f est une translation et modulation d'une gaussienne.*

Démonstration. Par densité, on suppose que $f \in \mathcal{C}_c^\infty(\mathbb{R})$. Ainsi, $\mathcal{F}f = \hat{f}$. On pose X et Y des variables aléatoires réelles de densités respectives $\left(\frac{|f|}{\|f\|_2}\right)^2$ et $\left(\frac{|\hat{f}|}{\|\hat{f}\|_2}\right)^2$. On peut supposer que $\|f\|_2 = 1$ (donc $\left\|\hat{f}\right\|_2 = 1$) et $\mathbb{E}(X) = \mathbb{E}(Y) = 0$. On a :

$$\sigma\left(\hat{f}\right)^2 = \int_{\mathbb{R}} \omega^2 \left|\hat{f}(\omega)\right|^2 \mathrm{d}\omega = \int_{\mathbb{R}} \left|i\omega\hat{f}(\omega)\right|^2 \mathrm{d}\omega = \int_{\mathbb{R}} \left|\widehat{f'}(\omega)\right|^2 \mathrm{d}\omega = \int_{\mathbb{R}} |f'(t)|^2 \mathrm{d}t.$$

En intégrant par parties, puis en utilisant l'inégalité de Cauchy-Schwarz, on a donc :

$$1 = \int_{\mathbb{R}} |f(t)|^2 \mathrm{d}t = -2\Re\left(\int_{\mathbb{R}} tf(t)\overline{f'(t)}\,\mathrm{d}t\right) \leqslant 2\left(\int_{\mathbb{R}} t^2 |f(t)|^2 \mathrm{d}t\right)^{1/2} \left(\int_{\mathbb{R}} |f'(t)|^2 \mathrm{d}t\right)^{1/2} = 2\sigma(f)\sigma\left(\hat{f}\right).$$

□

4.6 Formule sommatoire de Poisson

Théorème 4.6.1 (Formule sommatoire de Poisson)**.** *Soit $F \in \mathcal{C}^1(\mathbb{R})$. On suppose que :*

(i) $\sum_{n\in\mathbb{Z}} \left|\hat{F}(2\pi n)\right| < +\infty$.

(ii) $\exists M > 0$, $\exists \alpha > 1$, $\forall x \in \mathbb{R}$, $|F(x)| \leqslant \frac{M}{(1+|x|)^\alpha}$.

Alors :

$$\sum_{n\in\mathbb{Z}} \hat{F}(2\pi n) = \frac{1}{\sqrt{2\pi}} \sum_{n\in\mathbb{N}} F(n).$$

Démonstration. Considérer $f : x \in \mathbb{R} \longmapsto \sum_{n\in\mathbb{Z}} F(x+n)$. Montrer que f est périodique et suffisamment régulière et lui appliquer la théorie des séries de Fourier. □

4.7 Équation de la chaleur

Remarque 4.7.1. *Soit $f \in L^1\left(\mathbb{R}^d\right)$. L'*équation de la chaleur *s'écrit :*

$$\begin{cases} \partial_t u = \frac{1}{2}\Delta u \\ u(0,\cdot) = f \end{cases},$$

d'inconnue $u : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ deux fois différentiable. On peut chercher une solution de cette équation en "passant en Fourier en x". Sous hypothèse de régularité, si u est solution de l'équation de la chaleur, alors $\hat{u}$ est solution de l'équation différentielle ordinaire suivante :

$$\begin{cases} \partial_t \hat{u} = -\frac{\|\xi\|^2}{2}\hat{u} \\ \hat{u}(0,\cdot) = f \end{cases}.$$

Ainsi, $\forall (t,\xi) \in \mathbb{R}_+ \times \mathbb{R}^d,\ \hat{u}(t,\xi) = \exp\left(-\frac{t}{2}\|\xi\|\right)^2 \hat{f}(\xi) = K\hat{g}_{\sqrt{t}}(\xi)\hat{f}(\xi)$, où $K \in \mathbb{R}_+^*$ est une constante et g_σ est la gaussienne d'écart-type σ (c.f. lemme 1.9.9). Il vient :

$$\forall (t,\xi) \in \mathbb{R}_+ \times \mathbb{R}^d,\ u(t,\xi) = Kg_{\sqrt{t}} * Rf.$$

On vérifie a posteriori *qu'on a bien une solution de l'équation de la chaleur.*

5 Processus de branchement

5.1 Arbres de Galton-Watson

Définition 5.1.1 (Arbre de Galton-Watson)**.** *Soit μ une loi de probabilité sur $\mathbb{N}$, $(X_{n,i})_{(n,i)\in(\mathbb{N}^*)^2}$ une famille de variables aléatoires indépendantes de loi μ. L'*arbre de Galton-Watson *de loi de reproduction μ est défini par la suite $(Z_n)_{n\in\mathbb{N}}$ définie par récurrence par :*

$$Z_0 = 1 \qquad \text{et} \qquad \forall n \in \mathbb{N},\ Z_{n+1} = \sum_{i=1}^{Z_n} X_{n+1,i}.$$

Il s'interprète comme suit : à l'étape n, on a n individus, chacun se reproduit et donne $X_{n,i}$ descendants, et Z_n représente le nombre d'individus de la n-ième génération.

Définition 5.1.2. *Soit μ une loi de probabilité sur $\mathbb{N}$. On considère l'arbre de Galton-Watson $(Z_n)_{n\in\mathbb{N}}$ de loi de reproduction μ.*

(i) *On dit que l'arbre est* infini*, ou qu'il y a* survie*, lorsque $\forall n \in \mathbb{N},\ Z_n \neq 0$.*

(ii) *Dans le cas contraire, on dit que l'arbre est* fini *ou qu'il y a* extinction.

Remarque 5.1.3. *Soit μ une loi de probabilité sur $\mathbb{N}$. On considère l'arbre de Galton-Watson de loi de reproduction μ.*

(i) *Si $\mu(\{0\}) = 0$, l'arbre est presque-sûrement infini.*

(ii) *Si $\mu(\{0\}) > 0$, il y a une probabilité non nulle qu'il y ait extinction.*

Théorème 5.1.4. *Soit μ une loi de probabilité sur $\mathbb{N}$ t.q. $\mu(\{1\}) \neq 1$. On note $m_\mu = \sum_{k\in\mathbb{N}} k\mu(\{k\})$ la moyenne de μ et on considère l'arbre de Galton-Watson de loi de reproduction μ.*

(i) *Si $m_\mu \leqslant 1$, il y a extinction presque-sûrement.*

(ii) *Si $m_\mu > 1$, il y a une probabilité non nulle de survie.*

On parle de phase sous-critique *(resp.* phase sur-critique*) lorsque $m_\mu < 1$ (resp. $m_\mu > 1$). On parle de* phase critique *lorsque $m_\mu = 1$.*

Démonstration. Soit X une variable aléatoire de loi μ. Soit $n \in \mathbb{N}^*$ Calculons la série génératrice (c.f. définition 1.9.5) de Z_n en fonction de celle de X :

$$\begin{aligned}
\forall s \in [0,1],\ G_{Z_n}(s) &= \sum_{y\in\mathbb{N}} \mathbb{P}\left(Z_n = y\right) s^y = \sum_{y\in\mathbb{N}}\sum_{x\in\mathbb{N}} \mathbb{P}\left(Z_n = y \mid Z_{n-1} = x\right)\mathbb{P}\left(Z_{n-1} = x\right) s^y \\
&= \sum_{y\in\mathbb{N}}\sum_{x\in\mathbb{N}} \mathbb{P}\left(\sum_{i=1}^{x} X_{n,i} = y\right)\mathbb{P}\left(Z_{n-1} = x\right) s^y \\
&= \sum_{x\in\mathbb{N}} \mathbb{P}\left(Z_{n-1} = x\right)\sum_{y\in\mathbb{N}} \mathbb{P}\left(\sum_{i=1}^{x} X_{n,i} = y\right) s^y \\
&= \sum_{x\in\mathbb{N}} \mathbb{P}\left(Z_{n-1} = x\right) G_{\sum_{i=1}^{x} X_{n,i}}(s) = \sum_{x\in\mathbb{N}} \mathbb{P}\left(Z_{n-1} = x\right)\left(G_X(s)\right)^x \\
&= \left(G_{Z_{n-1}} \circ G_X\right)(s).
\end{aligned}$$

Par récurrence, on en déduit que $\forall n \in \mathbb{N},\ G_{Z_n} = \left(G_X\right)^n = \underbrace{G_X \circ \cdots \circ G_X}_{n \text{ fois}}$. On s'intéresse maintenant à l'événement d'extinction : $E = \bigcup_{n\in\mathbb{N}}\left(Z_n = 0\right)$. Par réunion croissante, on a :

$$\mathbb{P}(E) = \lim_{n\to+\infty} \mathbb{P}\left(Z_n = 0\right) = \lim_{n\to+\infty} G_{Z_n}(0) = \lim_{n\to+\infty} \left(G_X\right)^n(0).$$

On se ramène donc à une étude de système dynamique : on montre que si $m_\mu = G_X'(1) \leqslant 1$, alors G_X a un unique point fixe (qui est 1), sinon G_X a exactement deux points fixes (un en 1 et un dans $]0,1[$). □

Remarque 5.1.5. *Le calcul précédent montre que* $\mathbb{E}\left(Z_n\right) = G_{Z_n}'(1) = \left(G_X'(1)\right)^n = m_\mu^n$.

5.2 Phases sous-critique et critique pour les arbres de Galton-Watson

Proposition 5.2.1. *Soit μ une loi de probabilité sur $\mathbb{N}$ t.q. $m_\mu = \sum_{k\in\mathbb{N}} k\mu\left(\{k\}\right) < 1$. On considère l'arbre de Galton-Watson $\left(Z_n\right)_{n\in\mathbb{N}}$ de loi de reproduction μ. Alors :*

$$\forall n \in \mathbb{N},\ \mathbb{P}\left(Z_n > 0\right) \leqslant m_\mu^n.$$

Démonstration. Noter que :

$$\mathbb{P}\left(Z_n > 0\right) = 1-\left(G_X\right)^n(0) = G_X(1) - G_X\left(\left(G_X\right)^{n-1}(0)\right) \leqslant m_\mu\left(1 - \left(G_X\right)^{n-1}(0)\right) = m_\mu \mathbb{P}\left(Z_{n-1} > 0\right).$$

□

Proposition 5.2.2. *Soit μ une loi de probabilité sur $\mathbb{N}$ t.q. $\mu\left(\{1\}\right) \neq 1$ et $m_\mu = \sum_{k\in\mathbb{N}} k\mu\left(\{k\}\right) = 1$. On suppose que μ admet un moment d'ordre 2 et que $\sigma_\mu^2 = \sum_{k\in\mathbb{N}} k(k-1)\mu\left(\{k\}\right) \neq 0$. On considère l'arbre de Galton-Watson $\left(Z_n\right)_{n\in\mathbb{N}}$ de loi de reproduction μ. Alors :*

$$\mathbb{P}\left(Z_n > 0\right) \sim \frac{2}{n\sigma_\mu^2}.$$

Démonstration. Noter que $G_X(s) = 1 - (1-s) + \frac{(1-s)^2}{2}\sigma_\mu^2 + o_1\left((1-s)^2\right)$. Ainsi :

$$\frac{1}{1 - G_X(s)} - \frac{1}{1-s} \underset{1}{\sim} \frac{\sigma_\mu^2}{2}.$$

Comme $\left(G_X\right)^n(0) \xrightarrow[n\to+\infty]{} 1$, il vient $\frac{1}{1-\left(G_X\right)^n(0)} - \frac{1}{1-\left(G_X\right)^{n-1}(0)} \sim \frac{\sigma_\mu^2}{2}$. On obtient le résultat en sommant. □

5.3 Phase sur-critique pour les arbres de Galton-Watson

5.3.1 Conditionnement à l'extinction

Lemme 5.3.1. *Soit μ une loi de probabilité sur $\mathbb{N}$. On considère l'arbre de Galton-Watson $(Z_n)_{n\in\mathbb{N}}$ de loi de reproduction μ et on note $T = \sum_{n\in\mathbb{N}} Z_n$. Si $(X_n)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires indépendantes de loi μ et si $S_n = X_1 + \cdots + X_n - (n-1)$ pour $n \in \mathbb{N}^*$, alors T a la même loi que :*

$$T' = \min\left\{n \in \mathbb{N}^*,\ S_n = 0\right\}.$$

Théorème 5.3.2. *Soit μ une loi de probabilité sur $\mathbb{N}$ t.q. $m_\mu = \sum_{k\in\mathbb{N}} k\mu(\{k\}) > 1$. On note $q > 0$ la probabilité d'extinction. Alors le processus conditionné à l'extinction est un arbre de Galton-Watson dont la loi de reproduction μ_q est donnée par :*

$$\forall k \in \mathbb{N},\ \mu_q(\{k\}) = \mu(\{k\})\, q^{k-1}.$$

De plus, ce processus est sous-critique.

Exemple 5.3.3. *On suppose que μ est une loi de Poisson de paramètre $\lambda > 1$ (c.f. exemple 1.4.5). On se donne $\lambda' < 1$ t.q. $\lambda e^{-\lambda} = \lambda' e^{-\lambda'}$. Alors un arbre de Galton-Watson de loi de reproduction $\mathcal{P}(\lambda)$ conditionné à l'extinction est un arbre de Galton-Watson de loi de reproduction $\mathcal{P}(\lambda')$.*

5.3.2 Population totale

Théorème 5.3.4. *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variable aléatoires réelles indépendantes suivant toutes la loi d'une variable aléatoire X. On suppose X intégrable. Pour $a \in \mathbb{R}$, on note :*

$$I(a) = \begin{cases} \sup_{t\geqslant 0}\left(ta - \ln \mathbb{E}\left(e^{tX}\right)\right) & si\ a > \mathbb{E}(X) \\ \sup_{t\leqslant 0}\left(ta - \ln \mathbb{E}\left(e^{tX}\right)\right) & si\ a \leqslant \mathbb{E}(X) \end{cases}.$$

Alors :

(i) $\forall a \geqslant \mathbb{E}(X),\ \forall n \in \mathbb{N}^*,\ \mathbb{P}(X_1 + \cdots + X_n \geqslant na) \leqslant e^{-nI(a)}$.

(ii) $\forall a \leqslant \mathbb{E}(X),\ \forall n \in \mathbb{N}^*,\ \mathbb{P}(X_1 + \cdots + X_n \leqslant na) \leqslant e^{-nI(a)}$.

Démonstration. On se place dans le cas où $a \geqslant \mathbb{E}(X)$ (sinon, on remplace X par $-X$). Alors :

$$\begin{aligned}\forall t \in \mathbb{R}_+,\ \mathbb{P}(X_1 + \cdots + X_n \geqslant na) &= \mathbb{P}\left(\exp\left(t\sum_{i=1}^n X_i\right) \geqslant \exp(tna)\right) \\ &\leqslant e^{-tna}\mathbb{E}\left(\exp\left(t\sum_{i=1}^n X_i\right)\right) = \exp\left(-n\left(ta + \ln \mathbb{E}\left(e^{tX}\right)\right)\right).\end{aligned}$$

D'où le résultat en passant au sup sur t. □

Théorème 5.3.5. *Soit μ une loi de probabilité sur $\mathbb{N}$ t.q. $m_\mu = \sum_{k\in\mathbb{N}} k\mu(\{k\}) > 1$. On considère l'arbre de Galton-Watson $(Z_n)_{n\in\mathbb{N}}$ de loi de reproduction μ et on note $T = \sum_{n\in\mathbb{N}} Z_n$. Alors :*

$$\forall k \in \mathbb{N},\ \mathbb{P}(k \leqslant T < +\infty) \leqslant \frac{e^{-kI}}{1 - e^{-I}},$$

où $I = \sup_{t\in\mathbb{R}_-}\left(t - \ln \mathbb{E}\left(e^{tX}\right)\right) > 0$.

Démonstration. Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes de loi μ. Pour $n \in \mathbb{N}^*$, on pose $S_n = X_1 + \cdots + X_n - (n-1)$ et :

$$T' = \min\left\{n \in \mathbb{N}^*,\ S_n = 0\right\}.$$

Alors T et T' ont même loi selon le lemme 5.3.1. Et on a :

$$\forall k \in \mathbb{N},\ \mathbb{P}(k \leqslant T < +\infty) = \sum_{s=k}^{\infty} \mathbb{P}(T' = s) \leqslant \sum_{s=k}^{\infty} \mathbb{P}(X_1 + \cdots + X_s \leqslant s).$$

On applique le théorème 5.3.4 pour obtenir $\mathbb{P}(k \leqslant T < +\infty) \leqslant \sum_{s=k}^{\infty} e^{-sI}$. Après avoir montré que $I \neq 0$ (par exemple en observant la dérivée de $t \longmapsto t - \ln \mathbb{E}\left(e^{tX}\right)$ en 0), on en déduit l'inégalité souhaitée. □

5.4 Temps d'arrêt et population totale d'un arbre de Galton-Watson

Théorème 5.4.1. *Soit $(Y_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes et identiquement distribuées à valeurs dans $\mathbb{N}\cup\{-1\}$. On fixe $k\in\mathbb{N}$ et on considère, pour $n\in\mathbb{N}^*$:*

$$S_n = k + Y_1 + \cdots + Y_n \qquad \text{et} \quad H_0 = \min\{n\in\mathbb{N}^*,\ S_n = 0\}.$$

Alors :

$$\forall n\in\mathbb{N}^*,\ \mathbb{P}(H_0 = n) = \frac{k}{n}\mathbb{P}(S_n = 0).$$

Autrement dit, $\forall n\in\mathbb{N}^*$, $\mathbb{P}(H_0 = n \mid S_n = 0) = \frac{k}{n}$.

Démonstration. On raisonne par récurrence sur n. Pour $n=1$, le résultat est clair. Pour l'hérédité, on conditionne selon la valeur de Y_1. □

Théorème 5.4.2. *Soit μ une loi de probabilité sur $\mathbb{N}$. On considère l'arbre de Galton-Watson $(Z_n)_{n\in\mathbb{N}}$ de loi de reproduction μ et on note $T=\sum_{n\in\mathbb{N}} Z_n$. Si $(X_n)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires indépendantes de loi μ, alors :*

$$\forall n\in\mathbb{N}^*,\ \mathbb{P}(T=n) = \frac{1}{n}\mathbb{P}(X_1+\cdots+X_n = n-1).$$

Démonstration. Appliquer le lemme 5.3.1 puis le théorème 5.4.1. □

6 Marches aléatoires

6.1 Généralités

Définition 6.1.1 (Marche aléatoire)**.** *Soit μ une loi de probabilité sur $\mathbb{Z}^d$, et soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes de loi μ. On appelle* marche aléatoire *de loi de saut μ la suite $(S_n)_{n\in\mathbb{N}}$ définie par :*

$$\forall n\in\mathbb{N},\ S_n = \sum_{k=1}^{n} X_k \in \mathbb{Z}^d.$$

Définition 6.1.2 (Récurrence et transience)**.** *Soit μ une loi de probabilité sur $\mathbb{Z}^d$. On considère la marche aléatoire $(S_n)_{n\in\mathbb{N}}$ de loi de saut μ. Pour $x\in\mathbb{Z}^d$, on note :*

$$\mathcal{R}_x = \limsup_{n\to+\infty}(S_n = x) = \bigcap_{N\in\mathbb{N}}\bigcup_{n\geqslant N}(S_n = x).$$

Autrement dit, $\mathcal{R}_x$ est l'événement "la marche aléatoire passe par x une infinité de fois".

(i) *Si $\mathbb{P}(\mathcal{R}_0)=1$, on dit que la marche aléatoire est* récurrente*, et on a alors $\mathbb{P}\left(\bigcap_{x\in\mathbb{Z}^d}\mathcal{R}_x\right)=1$.*

(ii) *Si $\mathbb{P}\left(\complement\mathcal{R}_0\right)=1$, on dit que la marche aléatoire est* transiente*, et on a alors $\mathbb{P}\left(\bigcap_{x\in\mathbb{Z}^d}\complement\mathcal{R}_x\right)=1$.*

6.2 Principe de dichotomie

Notation 6.2.1. *Soit $(S_n)_{n\in\mathbb{N}}$ une marche aléatoire sur $\mathbb{Z}^d$ de loi de saut μ.*

- *On définit une fonction $g : x\in\mathbb{Z}^d \longmapsto \sum_{n\in\mathbb{N}}\mathbb{P}(S_n = x)\in[0,+\infty]$.*
- *On définit une suite $\left(H_0^{(k)}\right)_{k\in\mathbb{N}}$ de variables aléatoires par $H_0^{(0)}=0$ et $\forall k\in\mathbb{N}$, $H_0^{(k+1)} = \inf\left\{n > H_0^{(k)},\ S_n = 0\right\}$.*
- *On note de plus $H_0 = H_0^{(1)}$.*

Lemme 6.2.2. *Soit $(S_n)_{n\in\mathbb{N}}$ une marche aléatoire sur $\mathbb{Z}^d$ de loi de saut μ. Alors :*

$$\forall k\in\mathbb{N},\ \mathbb{P}\left(H_0^{(k)} < +\infty\right) = \mathbb{P}(H_0 < +\infty)^k.$$

Démonstration. Pour $r \in \mathbb{N}$ et $m \in \mathbb{N}$, on pose $S_m^{(r)} = \sum_{k=r+1}^{r+m} X_k$. Ainsi, pour $k \in \mathbb{N}$:

$$\begin{aligned}
\mathbb{P}\left(H_0^{(k+1)} < +\infty\right) &= \sum_{r=1}^{\infty} \mathbb{P}\left(H_0^{(k)} = r, H_0^{(k+1)} < +\infty\right) = \sum_{r=1}^{\infty} \mathbb{P}\left(\left(H_0^{(k)} = r\right) \cap \bigcup_{m \in \mathbb{N}^*} \left(S_m^{(r)} = 0\right)\right) \\
&= \sum_{r=1}^{\infty} \mathbb{P}\left(H_0^{(k)} = r\right) \mathbb{P}\left(\bigcup_{m \in \mathbb{N}^*} \left(S_m^{(r)} = 0\right)\right) = \sum_{r=1}^{\infty} \mathbb{P}\left(H_0^{(k)} = r\right) \mathbb{P}\left(H_0 < +\infty\right) \\
&= \mathbb{P}\left(H_0^{(k)} < +\infty\right) \cdot \mathbb{P}\left(H_0 < +\infty\right).
\end{aligned}$$

On en déduit le résultat par récurrence. $\square$

Proposition 6.2.3. *Soit $(S_n)_{n\in\mathbb{N}}$ une marche aléatoire sur $\mathbb{Z}^d$ de loi de saut μ. Alors :*

$$g(0) = \frac{1}{1 - \mathbb{P}\left(H_0 < +\infty\right)}.$$

Démonstration. Notons qu'on a une bijection entre $\{n \in \mathbb{N},\ S_n = 0\}$ et $\left\{k \in \mathbb{N},\ H_0^{(k)} < +\infty\right\}$. Donc :

$$g(0) = \mathbb{E}\left(\sum_{n\in\mathbb{N}} \mathbb{1}\left(S_n = 0\right)\right) = \mathbb{E}\left(\sum_{k\in\mathbb{N}} \mathbb{1}\left(H_0^{(k)} < +\infty\right)\right) = \sum_{k\in\mathbb{N}} \mathbb{P}\left(H_0 < +\infty\right)^k = \frac{1}{1 - \mathbb{P}\left(H_0 < +\infty\right)}.$$

$\square$

Théorème 6.2.4. *Soit $(S_n)_{n\in\mathbb{N}}$ une marche aléatoire sur $\mathbb{Z}^d$ de loi de saut μ.*

(i) *Si $\mathbb{P}\left(H_0 < +\infty\right) < 1$ (i.e. $g(0) < +\infty$), alors la marche aléatoire est transiente.*

(ii) *Si $\mathbb{P}\left(H_0 < +\infty\right) = 1$ (i.e. $g(0) = +\infty$), alors la marche aléatoire est récurrente.*

En particulier, $\mathbb{P}\left(\mathcal{R}_0\right) \in \{0, 1\}$.

Démonstration. (i) Supposons $\mathbb{P}\left(H_0 < +\infty\right) < 1$. Alors $\mathbb{E}\left(\sum_{n\in\mathbb{N}} \mathbb{1}\left(S_n = 0\right)\right) = g(0) < +\infty$ donc $\sum_{n\in\mathbb{N}} \mathbb{1}\left(S_n = 0\right)$ est presque-sûrement finie, et la marche aléatoire est transiente. (ii) Supposons $\mathbb{P}\left(H_0 < +\infty\right) = 1$. Alors $\forall k \in \mathbb{N}$, $\mathbb{P}\left(H_0^{(k)} < +\infty\right) = \mathbb{P}\left(H_0 < +\infty\right)^k = 1$. Donc, par intersection décroissante :

$$\mathbb{P}\left(\mathcal{R}_0\right) = \mathbb{P}\left(\bigcap_{k\in\mathbb{N}} \left(H_0^{(k)} < +\infty\right)\right) = \lim_{k\to+\infty} \mathbb{P}\left(H_0^{(k)} < +\infty\right) = 1.$$

Donc la marche aléatoire est récurrente. $\square$

6.3 Marche aléatoire simple

Définition 6.3.1 (Marche aléatoire simple)**.** *On appelle* marche aléatoire simple *sur $\mathbb{Z}^d$ la marche aléatoire dont la loi de saut μ est donnée par :*

$$\forall i \in \{1, \dots, d\},\ \mu\left(\{e_i\}\right) = \mu\left(\{-e_i\}\right) = \frac{1}{2d},$$

où $(e_1, \dots, e_d)$ est la base canonique de $\mathbb{R}^d$.

Lemme 6.3.2. *La marche aléatoire simple sur $\mathbb{Z}^d$ est récurrente* ssi *:*

$$\int_{[-\pi,\pi]^d} \Re\left(\frac{1}{1 - t\Phi_\mu(\xi)}\right) \mathrm{d}\xi \xrightarrow[t\to 1^-]{} +\infty,$$

où $\Phi_\mu : \xi \in \mathbb{R}^d \longmapsto \int_{\mathbb{R}^d} e^{i\langle\xi|\cdot\rangle}\,\mathrm{d}\mu = \sum_{x\in\mathbb{Z}^d} \mu\left(\{x\}\right) e^{i\langle\xi|x\rangle}$ est la fonction caractéristique de μ.

Démonstration. Pour $t \in [0,1]$, posons $g_t(0) = \sum_{n\in\mathbb{N}} t^n \mathbb{P}(S_n = 0)$. Ainsi $g(0) = \lim t \to 1^- g_t(0)$ par convergence monotone. Soit $n \in \mathbb{N}$. En considérant la mesure μ^{*n} (produit de convolution itéré n fois), on a :

$$\int_{[-\pi,\pi]^d} \Phi_\mu(\xi)^n \,\mathrm{d}\xi = \int_{[-\pi,\pi]^d} \Phi_{\mu^{*n}}(\xi) \,\mathrm{d}\xi = \sum_{x\in\mathbb{Z}^d} \mu^{*n}(\{x\}) \int_{[-\pi,\pi]^d} e^{i\langle \xi | x\rangle} \,\mathrm{d}\xi$$
$$= (2\pi)^d \mu^{*n}(\{0\}) = (2\pi)^d \mathbb{P}(S_n = 0).$$

On en déduit par convergence dominée que $\forall t \in [0,1[,\ g_t(0) = \frac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d} \frac{1}{1-t\Phi_\mu(\xi)} \,\mathrm{d}\xi$, puis on passe à la partie réelle, on fait tendre $t \to 1^-$ et on applique le théorème 6.2.4. □

Théorème 6.3.3 (Théorème de Pólya)**.**

(i) *Si $d \leqslant 2$, la marche aléatoire simple sur $\mathbb{Z}^d$ est récurrente.*

(ii) *Si $d \geqslant 3$, la marche aléatoire simple sur $\mathbb{Z}^d$ est transiente.*

Démonstration. *Cas 1* : $d = 1$. On a :

$$\forall m \in \mathbb{N},\ \mathbb{P}(S_{2m} = 0) = \binom{2m}{m}\left(\frac{1}{2}\right)^{2m} \sim \frac{1}{\sqrt{\pi m}}.$$

Il vient $g(0) = \sum_{n\in\mathbb{N}} \mathbb{P}(S_n = 0) = +\infty$, donc la marche aléatoire est récurrente selon le théorème 6.2.4. *Cas 2* : $d = 2$. On pose $\varepsilon_1 = \frac{\sqrt{2}}{2}(1,1) \in \mathbb{Z}^2$, $\varepsilon_2 = \frac{\sqrt{2}}{2}(-1,1) \in \mathbb{Z}^2$. Alors $(\varepsilon_1, \varepsilon_2)$ est une base de $\mathbb{R}^2$ dont on note $(\varepsilon_1^*, \varepsilon_2^*)$ la base duale. Pour $n \in \mathbb{N}$, on note $S_n^{(1)} = \varepsilon_1^*(S_n)$ et $S_n^{(2)} = \varepsilon_2^*(S_n)$. Alors $\left(S_n^{(1)}\right)_{n\in\mathbb{N}}$ et $\left(S_n^{(2)}\right)_{n\in\mathbb{N}}$ sont des marches aléatoires simples indépendantes sur $\mathbb{Z}$. Donc :

$$\forall m \in \mathbb{N},\ \mathbb{P}(S_{2m} = 0) = \mathbb{P}\left(S_{2m}^{(1)} = 0\right)\mathbb{P}\left(S_{2m}^{(2)} = 0\right) \sim \frac{1}{\pi m}.$$

Il vient $g(0) = +\infty$, donc la marche aléatoire est récurrente selon le théorème 6.2.4. *Cas 3* : $d \geqslant 3$. On a :

$$\forall \xi \in \mathbb{R}^d,\ \Phi_\mu(\xi) = \frac{1}{d}\sum_{k=1}^{d} \cos(\xi_k).$$

Donc $\frac{1}{1-\Phi_\mu(\xi)} \underset{0}{\sim} \frac{2d}{\|\xi\|^2}$. On conclut alors avec le lemme 6.3.2. □

6.4 Exemples de résultats plus généraux en dimension 1

Théorème 6.4.1. *Soit $(S_n)_{n\in\mathbb{N}}$ une marche aléatoire sur $\mathbb{Z}$ de loi de saut μ. On suppose que $\sum_{k\in\mathbb{Z}} |k|\,\mu(\{k\}) < +\infty$. Alors la marche aléatoire est récurrente* ssi $\sum_{k\in\mathbb{Z}} k\mu(\{k\}) = 0$.

Notation 6.4.2. *Soit $(S_n)_{n\in\mathbb{N}}$ une marche aléatoire sur $\mathbb{Z}$ de loi de saut μ. Pour $n \in \mathbb{N}^*$, on note $\overline{X}_n = \frac{S_n}{n}$.*

Définition 6.4.3 (Entropie et pression)**.** *Soit $(S_n)_{n\in\mathbb{N}}$ une marche aléatoire sur $\mathbb{Z}$ de loi de saut μ. On définit :*

(i) *L'*entropie $s : x \in \mathbb{R} \longmapsto \sup_{n\in\mathbb{N}^*}\left(\frac{1}{n}\ln \mathbb{P}\left(\overline{X}_n \geqslant x\right)\right)$.

(ii) *La* pression $p : \lambda \in \mathbb{R}_+ \longmapsto \ln \mathbb{E}\left(e^{\lambda X_1}\right)$.

Lemme 6.4.4. *Soit $(S_n)_{n\in\mathbb{N}}$ une marche aléatoire sur $\mathbb{Z}$ de loi de saut μ. Alors :*

$$\forall x \in \mathbb{R},\ \forall n \in \mathbb{N}^*,\ \mathbb{P}(X_1 \geqslant x)^n \leqslant \mathbb{P}\left(\overline{X}_n \geqslant x\right) \leqslant e^{ns(x)}.$$

Théorème 6.4.5. *Soit $(S_n)_{n\in\mathbb{N}}$ une marche aléatoire sur $\mathbb{Z}$ de loi de saut μ. Alors :*

$$\forall \lambda \in \mathbb{R}_+,\ p(\lambda) = \sup_{u\in\mathbb{R}} (\lambda u + s(u)).$$

Proposition 6.4.6. *Soit* $(S_n)_{n\in\mathbb{N}}$ *une marche aléatoire sur* $\mathbb{Z}$ *de loi de saut* μ. *Alors pour tout* $x \in \mathbb{R}$, *la suite* $\left(-\ln \mathbb{P}\left(\overline{X}_n \geqslant x\right)\right)_{n\in\mathbb{N}^*}$ *est sous-additive et la suite* $\left(\frac{1}{n}\ln \mathbb{P}\left(\overline{X}_n \geqslant x\right)\right)_{n\in\mathbb{N}^*}$ *converge dans* $[-\infty, 0]$ *vers* $s(x)$.

Proposition 6.4.7. *Soit* $(S_n)_{n\in\mathbb{N}}$ *une marche aléatoire sur* $\mathbb{Z}$ *de loi de saut* μ. *Alors l'entropie* s *est concave.*

Théorème 6.4.8 (Théorème de Cramér)**.** *Soit* $(S_n)_{n\in\mathbb{N}}$ *une marche aléatoire sur* $\mathbb{Z}$ *de loi de saut* μ. *Alors pour tout* $x \in \mathbb{R}$, *la suite* $\left(\frac{1}{n}\ln \mathbb{P}\left(\overline{X}_n \geqslant x\right)\right)_{n\in\mathbb{N}^*}$ *converge dans* $[-\infty, 0]$ *et :*

$$\frac{1}{n}\ln \mathbb{P}\left(\overline{X}_n \geqslant x\right) \xrightarrow[n\to+\infty]{} \inf_{\lambda\in\mathbb{R}_+}\left(\ln \mathbb{E}\left(e^{\lambda X_1}\right) - \lambda x\right).$$