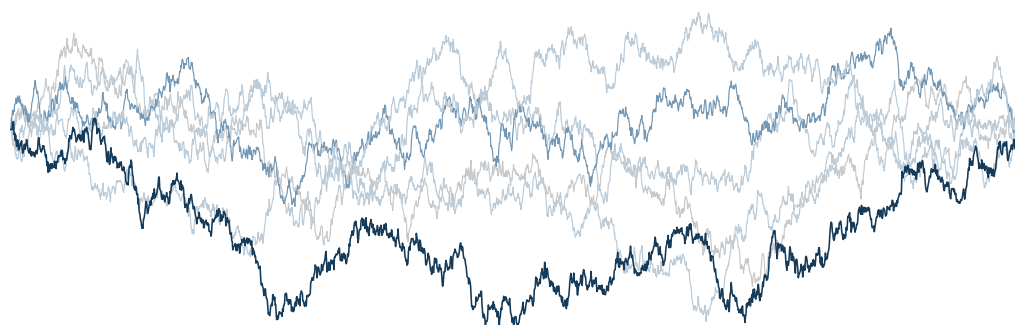


Théorèmes limites et applications

Cours, exercices et corrigés



Ce document réunit les notes du cours et les exercices et problèmes, ainsi que leurs corrigés, issus des travaux dirigés, devoirs maison et examens du cours enseigné au M2 de l'aleatoire de l'Université Paris-Saclay entre 2020 et 2025.

La [partie I](#) développe la convergence en loi, le théorème de Prokhorov, les espaces de trajectoires, le théorème de Donsker, le pont brownien et les mesures aléatoires de Poisson.

La [partie II](#) rassemble les exercices et leurs corrigés, répartis en sept chapitres thématiques. Le [sommaire](#) relie chaque thème du cours aux exercices correspondants.

Sommaire

Les huit chapitres du cours et les sept feuilles d'exercices se correspondent selon le tableau ci-dessous.

Thème	Cours	Exercices
Convergence étroite	Chapitre 1, p. 6	Feuille 1, p. 98
Modes de convergence	Chapitre 2, p. 14	Feuille 2, p. 111
Mesures sur un espace métrique	Chapitre 3, p. 18	Feuille 3, p. 125
Fonctions continues	Chapitre 4, p. 35	Feuille 4, p. 155
Donsker et pont brownien	Chapitre 5, p. 55 Chapitre 6, p. 62	Feuille 5, p. 177
Trajectoires càdlàg (Skorokhod)	Chapitre 7, p. 69	Feuille 6, p. 196
Mesures aléatoires de Poisson	Chapitre 8, p. 82	Feuille 7, p. 203

Table des matières

Sommaire	2
Partie I Cours	5
Notations	5
Références	5
1 Convergence étroite dans $\mathbb{R}^k$	6
1.1 Définition	6
1.2 Restriction des fonctions test	7
1.3 Fonctions de répartition ($k = 1$) et tension	8
1.4 Fonctions caractéristiques, théorème de Lévy	11
Interlude : séparabilité	13
2 Autres modes de convergence	14
2.1 Convergence presque sûre, en probabilité	14
2.2 Uniforme intégrabilité et convergence dans L^1	15
3 Mesures de probabilité sur un espace métrique	18
3.1 Régularité de mesures	18
3.2 Théorème de porte-manteau	18
3.3 Convergence étroite dans $\mathbb{R}^{\mathbb{N}}$	22
3.4 Tension et théorème de Prokhorov	24
3.5 Propriétés topologiques de $\mathcal{M}_1(E)$	29
3.6 Compléments de topologie	32
4 L'espace des fonctions continues sur un compact	35
4.1 Caractère polonais	35
4.2 Marginales finidimensionnelles	36
4.3 Convergence en loi dans cet espace	38
4.4 Compacité dans C	39
4.5 Tension	41
4.6 Critère de tension de Kolmogorov	43
4.7 Fonctions continues sur $\mathbb{R}_+$	46
Interlude : couplages	48
C.1 Théorème de représentation de Skorokhod	48
C.2 Distance en variation totale	51
C.3 Distance de Lévy–Prokhorov	52
C.4 Monotonie, ordre stochastique	53
5 Théorème de Donsker et applications	55
5.1 Le théorème de Donsker	55
5.2 Loi de l'arcsinus	58
5.3 Théorème de Lévy	60
6 Le pont brownien	62
6.1 Définition du pont brownien	62
6.2 Le théorème local limite	63
6.3 Convergence vers le pont brownien	66

7	Processus discontinus : l'espace de Skorokhod	69
7.1	Propriétés de base	69
7.2	Topologie de Skorokhod (J_1) sur $\mathbb{D}$	71
7.3	Convergence étroite dans $\mathbb{D}$	74
7.4	Critères de tension	75
7.5	Compléments facultatifs	78
8	Mesures aléatoires de Poisson	82
8.1	Rappels sur les lois de Poisson	82
8.2	Mesures aléatoires de Poisson	83
8.3	Points multiples	87
8.4	Formules de Campbell (additive et exponentielle)	89
8.5	Transformations de nuages poissoniens	92
8.6	Nuages de Poisson avec paramètre temporel	93
8.7	Formule de Palm	95
Partie II	Exercices	97
1	Convergence étroite dans $\mathbb{R}^k$	98
1.1	Énoncés	98
1.2	Corrigés	101
2	Différents modes de convergence (presque sûre, en probabilité, L^1)	111
2.1	Énoncés	111
2.2	Corrigés	114
3	Mesures de probabilité sur un espace métrique	125
3.1	Énoncés	125
3.2	Corrigés	133
4	L'espace des fonctions continues sur un compact	155
4.1	Énoncés	155
4.2	Corrigés	161
5	Théorème de Donsker, pont brownien	177
5.1	Énoncés	177
5.2	Corrigés	183
6	L'espace des fonctions càdlàg	196
6.1	Énoncés	196
6.2	Corrigés	198
7	Mesures aléatoires de Poisson	203
7.1	Énoncés	203
7.2	Corrigés	205

Partie I

Cours

Notations

- (E, d) : espace métrique.
- $\mathcal{B}_E$ ou $\mathcal{B}$: tribu borélienne (la plus petite qui contient les ouverts).
- $\mathcal{M}_1(E)$: mesures de probabilité sur E .
- Si $f : E \rightarrow \mathbb{R}$ est mesurable positive ou μ -intégrable, on note

$$\mu(f) = \langle \mu, f \rangle = \int_E f \, d\mu.$$

- $\mathcal{C}_b(E)$: fonctions continues bornées $E \rightarrow \mathbb{R}$.
- $\mathcal{C}_c(E)$: fonctions continues $E \rightarrow \mathbb{R}$ à support compact.
- $\mathcal{C}(I, E)$ (également noté $C(I, E)$) : fonctions continues de I dans E .
- $\mathcal{ID}(I, E)$: fonctions càdlàg de I dans E ; le domaine est omis lorsqu'il est fixé.
- $X_n \xrightarrow{\text{loi}} X$ ou $X_n \Rightarrow X$: convergence en loi. Pour des mesures, $\mu_n \Rightarrow \mu$ désigne la convergence étroite.

Références

- P. Billingsley, *Convergence of Probability Measures* (2nd ed.).
- O. Kallenberg, *Foundations of Modern Probability* (2nd ed.).
- A. Klenke, *Probability Theory : A Comprehensive Course*.
- J. Kingman, *Poisson Processes*.
- G. Last, M. Penrose, *Lectures on the Poisson Process*.
- P.-L. Méliot, [Convergence de mesures, processus de Poisson et processus de Lévy](#) (notes de cours)
- G. Miermont, [Théorèmes limites et processus de Poisson](#) (notes de cours)
- T. Lévy, [Convergence de mesures, Grandes déviations, Percolation](#). (notes de cours)

1 Convergence étroite dans $\mathbb{R}^k$

Exercices associés : feuille 1, p. 98.

Dans ce chapitre

1.1	Définition	6
1.2	Restriction des fonctions test	7
1.3	Fonctions de répartition ($k = 1$) et tension	8
1.4	Fonctions caractéristiques, théorème de Lévy	11

1.1 Définition

Définition 1.1. Soit (E, d) un espace métrique et $\mu_n, \mu \in \mathcal{M}_1(E)$. On dit que (μ_n) converge *étroitement* vers μ si

$$\forall f \in \mathcal{C}_b(E), \quad \mu_n(f) \longrightarrow \mu(f).$$

On note alors $\mu_n \Rightarrow \mu$.

Exemple 1.2. — Si $x_n \rightarrow x \in \mathbb{R}^k$, alors $\delta_{x_n} \Rightarrow \delta_x$.

$$\text{— } \mu_n = \frac{1}{n+1} \sum_{i=0}^n \delta_{i/n} \Rightarrow \text{Leb}_{[0,1]}.$$

Formulation probabiliste. Si X_n, X sont des variables aléatoires de lois respectives μ_n, μ , on dit que $X_n \xrightarrow{\text{loi}} X$ si $\mu_n \Rightarrow \mu$, c'est-à-dire

$$\forall f \in \mathcal{C}_b(E), \quad \mathbb{E}[f(X_n)] \longrightarrow \mathbb{E}[f(X)].$$

Proposition 1.3. Soit (E, d) un espace métrique et $\mu_n, \mu, \nu \in \mathcal{M}_1(E)$. Si $\mu_n \Rightarrow \mu$ et $\mu_n \Rightarrow \nu$, alors $\mu = \nu$.

Démonstration. On a alors $\mu(f) = \nu(f)$ pour tout $f \in \mathcal{C}_b(E)$. Soit F fermé et $K > 0$. On pose

$$g_{F,K}(x) = \max(1 - Kd(x, F), 0).$$

La fonction vaut 1 sur F , puis décroît linéairement en fonction de la distance à F jusqu'à s'annuler à distance $1/K$. Comme $g_{F,K}$ converge simplement vers $\mathbb{1}_F$ lorsque $K \rightarrow \infty$, par convergence dominée on en déduit que $\mu(F) = \nu(F)$, et donc $\mu = \nu$ par le théorème des classes monotones. $\square$

Dans la suite de ce chapitre, on se restreint à $\mathbb{R}^k$, muni de la distance euclidienne.

1.2 Restriction des fonctions test

Théorème 1.4 (Petit porte-manteau). Soit $\mu_n, \mu \in \mathcal{M}_1(\mathbb{R}^k)$. On a $\mu_n \Rightarrow \mu$ si et seulement si, pour toute fonction $f : \mathbb{R}^k \rightarrow \mathbb{R}_+$ mesurable bornée, continue μ -presque partout (c'est-à-dire $\mu(\{x \in \mathbb{R}^k : f \text{ continue en } x\}) = 1$), on a $\mu_n(f) \rightarrow \mu(f)$.

Voir le théorème de porte-manteau 3.2 pour une preuve dans un cadre plus général.

Proposition 1.5. Soit $\mu_n, \mu \in \mathcal{M}_1(\mathbb{R}^k)$. Soit H un ensemble de fonctions $\mathbb{R}^k \rightarrow \mathbb{R}$ mesurables bornées tel que $\mathcal{C}_c(\mathbb{R}^k) \subset \overline{H}$, l'adhérence étant prise pour la norme uniforme. Alors

$$[\forall f \in H, \mu_n(f) \rightarrow \mu(f)] \implies \mu_n \Rightarrow \mu.$$

Démonstration. Étape 1. On montre que $\mu_n(f) \rightarrow \mu(f)$ pour tout $f \in \mathcal{C}_c(\mathbb{R}^k)$ implique $\mu_n \Rightarrow \mu$. L'idée est un argument de troncature. Soit $\chi_r : \mathbb{R}^k \rightarrow [0, 1]$ continue, avec $\chi_r(x) = 1$ pour $\|x\| \leq r$ et $\chi_r(x) = 0$ pour $\|x\| \geq r+1$. Soit $f \in \mathcal{C}_b(\mathbb{R}^k)$. Alors

$$|\mu_n(f) - \mu(f)| \leq |\mu_n(f) - \mu_n(f\chi_r)| + |\mu_n(f\chi_r) - \mu(f\chi_r)| + |\mu(f\chi_r) - \mu(f)|.$$

Quand $n \rightarrow \infty$, la limite supérieure du premier terme est majorée par $\|f\|_\infty(1 - \mu(\chi_r))$, celle du deuxième terme est nulle, et le troisième terme est majoré par $\|f\|_\infty(1 - \mu(\chi_r))$.

Ainsi

$$\limsup_{n \rightarrow \infty} |\mu_n(f) - \mu(f)| \leq 2\|f\|_\infty(1 - \mu(\chi_r)).$$

Comme $\mu(\chi_r) \rightarrow 1$ quand $r \rightarrow \infty$ par convergence dominée, cela conclut l'étape 1.

Étape 2. Soit $f \in \mathcal{C}_c(\mathbb{R}^k)$. Soit $g \in H$ tel que $\|f - g\|_\infty \leq \varepsilon$. Alors

$$|\mu_n(f) - \mu(f)| \leq \underbrace{|\mu_n(f) - \mu_n(g)|}_{\leq \varepsilon} + \underbrace{|\mu_n(g) - \mu(g)|}_{\rightarrow 0 \text{ car } g \in H} + \underbrace{|\mu(g) - \mu(f)|}_{\leq \varepsilon}.$$

Donc $\mu_n(f) \rightarrow \mu(f)$. On conclut par l'étape 1. □

Application. Soit $\mu \in \mathcal{M}_1(\mathbb{R})$ et $(X_n)_{n \geq 1}$ des variables aléatoires i.i.d. de loi μ . Soit $\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$. Alors, presque sûrement, $\mu_n \Rightarrow \mu$. En particulier, toute probabilité de $\mathcal{M}_1(\mathbb{R})$ est limite étroite de mesures atomiques.

Démonstration. Soit $(f_k)_{k \geq 1}$ une suite de $\mathcal{C}_c(\mathbb{R})$ dense dans $\mathcal{C}_c(\mathbb{R})$: prendre par exemple les polynômes à coefficients rationnels multipliés par des fonctions plateau $T_{K,r}$, avec $K, r \in \mathbb{N}^*$, égales à 1 sur $[-K, K]$, nulles hors de $[-K - 1/r, K + 1/r]$ et affines sur les deux intervalles intermédiaires. On pose $H = \{f_k : k \geq 1\}$. Comme

$$\mu_n(f_k) = \frac{1}{n} \sum_{i=1}^n f_k(X_i) \xrightarrow[n \rightarrow \infty]{\text{p.s.}} \mathbb{E}[f_k(X_1)] = \mu(f_k)$$

d'après la loi des grands nombres, on a presque sûrement $\mu_n(f) \rightarrow \mu(f)$ pour tout

$f \in H$, d'où le résultat par la proposition. $\square$

1.3 Fonctions de répartition ($k = 1$) et tension

Définition 1.6. Si $\mu \in \mathcal{M}_1(\mathbb{R})$, on note $F_\mu(x) = \mu(]-\infty, x])$, pour $x \in \mathbb{R}$, sa *fonction de répartition*.

Remarque 1.7. — $\lim_{+\infty} F_\mu = 1$, $\lim_{-\infty} F_\mu = 0$, F_μ est croissante, càdlàg (continue à droite avec des limites à gauche en tout point).

— Réciproquement, si $F : \mathbb{R} \rightarrow [0, 1]$ est croissante, continue à droite et vérifie $\lim_{-\infty} F = 0$, $\lim_{+\infty} F = 1$, alors il existe $\mu \in \mathcal{M}_1(\mathbb{R})$ telle que $F = F_\mu$. En effet, on peut prendre pour μ l'image de la mesure de Lebesgue sur $]0, 1[$ par

$$x \mapsto \inf\{y \in \mathbb{R} : F(y) \geq x\}.$$

C'est la mesure de Lebesgue–Stieltjes associée à F ; voir exercices.

— μ a un atome en x si et seulement si $\Delta F_\mu(x) = F_\mu(x) - F_\mu(x-) = \mu(\{x\}) > 0$.

— μ a un nombre dénombrable d'atomes, car

$$\forall k \geq 1, \quad \#\{x \in \mathbb{R} : \mu(\{x\}) \geq 1/k\} \leq k.$$

— μ n'a pas d'atome si et seulement si F_μ est continue.

— Si μ a une densité g par rapport à la mesure de Lebesgue, alors F_μ est dérivable presque partout, de dérivée g , et $F_\mu(x) = \int_{-\infty}^x g(t) dt$.

Attention : F_μ est toujours presque partout dérivable, mais n'est pas toujours l'intégrale de sa dérivée (c'est en fait le cas si et seulement si $\mu \ll \text{Leb}$). Prendre l'escalier du diable, ou de Cantor, dont la dérivée est nulle presque partout : c'est une fonction continue croissante, présentant des plateaux à toutes les échelles.

Proposition 1.8. Soit $\mu_n, \mu \in \mathcal{M}_1(\mathbb{R})$. Alors $\mu_n \Rightarrow \mu$ si et seulement si

$$F_{\mu_n}(x) \rightarrow F_\mu(x) \quad \text{pour tout } x \in \mathbb{R} \text{ tel que } \Delta F_\mu(x) = 0.$$

Démonstration. Si $\mu_n \Rightarrow \mu$ et si $\Delta F_\mu(x) = 0$, alors $f(z) = \mathbb{1}_{\{z \leq x\}}$ est continue μ -presque partout et donc, par le petit porte-manteau, $\mu_n(f) \rightarrow \mu(f)$.

Réciproquement, on a alors $\mu_n(g) \rightarrow \mu(g)$ pour toute fonction g de la forme

$$g = \sum_{i=1}^r \alpha_i \mathbb{1}_{]x_i, y_i]},$$

avec x_i et y_i des non-atomes de μ . Or toute fonction de $\mathcal{C}_c(\mathbb{R})$ est limite uniforme de telles fonctions. On conclut avec la proposition de la sous-section 2. $\square$

Définition 1.9. On dit qu'une famille $(\mu_i)_{i \in I}$ de mesures de $\mathcal{M}_1(\mathbb{R})$ est *tendue* si

$$\forall \varepsilon > 0, \quad \exists A > 0, \quad \sup_{i \in I} \mu_i(\mathbb{R} \setminus [-A, A]) \leq \varepsilon.$$

Remarque 1.10. Une suite $(\mu_n)_{n \geq 1}$ de $\mathcal{M}_1(\mathbb{R})$ est tendue si et seulement si

$$\forall \varepsilon > 0, \quad \exists A > 0, \quad \limsup_{n \rightarrow \infty} \mu_n(\mathbb{R} \setminus [-A, A]) \leq \varepsilon.$$

Ceci provient du fait qu'une mesure $\mu \in \mathcal{M}_1(\mathbb{R})$ est tendue puisque $\lim_{A \rightarrow \infty} \mu(\mathbb{R} \setminus [-A, A]) = 0$.

Le résultat suivant est un résultat de type « compacité » et est un cas particulier du théorème de Prokhorov 3.24, présenté au chapitre 3.

Théorème 1.11. Soit $(\mu_n)_{n \geq 1}$ une suite de $\mathcal{M}_1(\mathbb{R})$. Alors $(\mu_n)_{n \geq 1}$ est tendue si et seulement si, pour toute extraction φ , il existe une extraction ψ telle que $(\mu_{\varphi \circ \psi(n)})_{n \geq 1}$ converge étroitement.

Par extraction, on veut dire sous-suite ou, plus formellement, une injection croissante $\mathbb{N}^* \rightarrow \mathbb{N}^*$.

Démonstration. Sens réciproque. Par l'absurde, on suppose qu'il existe $\varepsilon > 0$ tel que, pour tout $n \geq 1$,

$$\mu_{\varphi(n)}(\mathbb{R} \setminus [-n, n]) > \varepsilon.$$

On peut prendre φ croissante car, à $i \geq 1$ fixé, $\lim_{n \rightarrow \infty} \mu_i(\mathbb{R} \setminus [-n, n]) = 0$. Par hypothèse, soit ψ une extraction telle que $(\mu_{\varphi \circ \psi(n)})_{n \geq 1}$ converge étroitement vers $\mu \in \mathcal{M}_1(\mathbb{R})$. Alors, pour tout $x > 0$ tel que $\mu(\{x\}) = \mu(\{-x\}) = 0$, on a, pour n assez grand, $[-x, x] \subset [-\psi(n), \psi(n)]$. Donc

$$\begin{aligned} \mu([-x, x]) &= \lim_{n \rightarrow \infty} \mu_{\varphi \circ \psi(n)}([-x, x]) \\ &\leq \liminf_{n \rightarrow \infty} \mu_{\varphi \circ \psi(n)}([-\psi(n), \psi(n)]) \leq 1 - \varepsilon. \end{aligned}$$

En faisant $x \rightarrow \infty$, on obtient $\mu(\mathbb{R}) < 1$, absurde.

Sens direct. Quitte à travailler avec $\mu_{\varphi(n)}$, on suppose $\varphi = \text{id}$.

Étape 1 : identifier une limite potentielle. On pose $F_n = F_{\mu_n}$. Pour tout $q \in \mathbb{Q}$, la suite $(F_n(q))_{n \geq 1}$ est bornée. Par argument diagonal, il existe une extraction ψ telle que, pour tout $q \in \mathbb{Q}$, $(F_{\psi(n)}(q))_{n \geq 1}$ converge vers une limite notée $F_\infty(q)$. On pose alors

$$F(x) = \inf\{F_\infty(q) : q > x, q \in \mathbb{Q}\}.$$

On vérifie que F est croissante, continue à droite. Soit $\varepsilon > 0$ et $A > 0$ tel que

$$\sup_{n \geq 1} \mu_n(\mathbb{R} \setminus [-A, A]) \leq \varepsilon.$$

Soit $q \in \mathbb{Q}$, $q > A$. Alors $F_n(q) \geq 1 - \varepsilon$. En faisant $n \rightarrow \infty$, on obtient $F_\infty(q) \geq 1 - \varepsilon$, d'où $F(A) \geq 1 - \varepsilon$. On montre de même que $F(x) \leq \varepsilon$ pour $x < -A$. Donc $\lim_{-\infty} F = 0$,

$\lim_{+\infty} F = 1$. Ainsi il existe $\mu \in \mathcal{M}_1(\mathbb{R})$ telle que $F = F_\mu$: c'est la mesure de Lebesgue–Stieltjes associée à F , voir la remarque précédente.

Étape 2. On vérifie que $\mu_{\psi(n)} \Rightarrow \mu$. Soit $x \in \mathbb{R}$ tel que $\mu(\{x\}) = 0$. Alors

$$F(x) = \sup\{F_\infty(q) : q < x, q \in \mathbb{Q}\} = \inf\{F_\infty(q) : q > x, q \in \mathbb{Q}\}.$$

Pour la première égalité, on peut utiliser $F(x) = \lim_{u \uparrow x} F(u)$. Ainsi, si $q < x < q'$ avec $q, q' \in \mathbb{Q}$,

$$\begin{aligned} F_\infty(q) &= \lim_{n \rightarrow \infty} F_{\psi(n)}(q) \leq \liminf_{n \rightarrow \infty} F_{\psi(n)}(x) \\ &\leq \limsup_{n \rightarrow \infty} F_{\psi(n)}(x) \leq \limsup_{n \rightarrow \infty} F_{\psi(n)}(q') = F_\infty(q'). \end{aligned}$$

En faisant $q \uparrow x, q' \downarrow x$, on conclut que $F_{\psi(n)}(x) \rightarrow F(x)$.

□

Remarque 1.12. — Il est tentant de reformuler en disant que $\{\mu_n : n \geq 1\}$ est d'adhérence compacte dans $\mathcal{M}_1(\mathbb{R})$ muni d'une distance qui métrise la convergence étroite. Ce point sera précisé à la section 3.5.

— L'extension à $\mathbb{R}^k$ est similaire (voir la proposition 5.21 dans Kallenberg).

Définition 1.13. On dit qu'une famille $(X_i)_{i \in I}$ de variables aléatoires réelles est *tendue* si la famille de leurs lois l'est.

Exemple 1.14. Si $(X_i)_{i \in I}$ est bornée dans L^1 , c'est-à-dire s'il existe $C > 0$ tel que $\sup_{i \in I} \mathbb{E}[|X_i|] \leq C$, alors $(X_i)_{i \in I}$ est tendue. En effet,

$$\mathbb{P}(|X_i| \geq A) \leq \frac{\mathbb{E}[|X_i|]}{A} \leq \frac{C}{A}.$$

Lemme 1.15 (Sous-sous-suites pour les mesures). Soit $\mu_n, \mu \in \mathcal{M}_1(\mathbb{R})$. Alors $\mu_n \Rightarrow \mu$ si et seulement si, pour toute extraction φ , il existe une extraction ψ telle que $\mu_{\varphi \circ \psi(n)} \Rightarrow \mu$.

Démonstration. Le sens direct est immédiat. Pour la réciproque, par l'absurde, soient $\varepsilon > 0, f \in \mathcal{C}_b(\mathbb{R})$ et φ une extraction tels que, pour tout n , $|\mu_{\varphi(n)}(f) - \mu(f)| \geq \varepsilon$. Par hypothèse, il existe une extraction ψ telle que $\mu_{\varphi \circ \psi(n)} \Rightarrow \mu$. Alors $\mu_{\varphi \circ \psi(n)}(f) \rightarrow \mu(f)$: contradiction. □

Corollaire 1.16. Soit $(\mu_n)_{n \geq 1}$ une suite de $\mathcal{M}_1(\mathbb{R})$ et $\mu \in \mathcal{M}_1(\mathbb{R})$. Alors $\mu_n \Rightarrow \mu$ si et seulement si

- $(\mu_n)_{n \geq 1}$ est tendue ;
- si $\mu_{\varphi(n)} \Rightarrow \nu$ pour une extraction φ et $\nu \in \mathcal{M}_1(\mathbb{R})$, alors $\mu = \nu$.

Démonstration. Le sens direct est clair. Pour la réciproque, on raisonne par l'absurde. D'après le lemme précédent, il existe une extraction φ telle que, pour toute extraction ψ , $\mu_{\varphi \circ \psi(n)} \not\Rightarrow \mu$. Or $(\mu_{\varphi(n)})$ est tendue, donc il existe une extraction ψ et $\nu \in \mathcal{M}_1(\mathbb{R})$ telles que $\mu_{\varphi \circ \psi(n)} \Rightarrow \nu$. Donc $\nu = \mu$ par hypothèse : absurde. $\square$

En pratique, le corollaire précédent est *très utile* pour montrer que $\mu_n \Rightarrow \mu$, comme suit :

- (i) On montre que (μ_n) est tendue : *tension*.
- (ii) On prend une extraction φ et ν telles que $\mu_{\varphi(n)} \Rightarrow \nu$. On montre que $\nu = \mu$ en utilisant une propriété qui caractérise une mesure : *unicité de la limite*.

1.4 Fonctions caractéristiques, théorème de Lévy

Définition 1.17. Pour $\mu \in \mathcal{M}_1(\mathbb{R}^k)$, on note

$$\phi_\mu(u) = \int_{\mathbb{R}^k} e^{i\langle u, x \rangle} \mu(dx), \quad u \in \mathbb{R}^k,$$

la *fonction caractéristique* de μ ; $\langle u, x \rangle$ désigne le produit scalaire.

Il est clair que $\mu_n \Rightarrow \mu$ implique $\phi_{\mu_n} \rightarrow \phi_\mu$ simplement. Pour la réciproque, on a le résultat plus précis suivant.

Théorème 1.18 (Lévy). Soit $\mu_n \in \mathcal{M}_1(\mathbb{R}^k)$ telle que

- (i) pour tout $u \in \mathbb{R}^k$, $\phi_{\mu_n}(u)$ converge vers une limite notée $\phi(u)$;
- (ii) $u \mapsto \phi(u)$ est continue en 0.

Alors il existe $\mu \in \mathcal{M}_1(\mathbb{R}^k)$ telle que $\phi = \phi_\mu$ et $\mu_n \Rightarrow \mu$.

Pour la preuve, on se restreint à $k = 1$ pour simplifier.

Lemme 1.19 (Injectivité de la fonction caractéristique). Soit $\mu, \nu \in \mathcal{M}_1(\mathbb{R})$ telles que $\phi_\mu = \phi_\nu$. Alors $\mu = \nu$.

Idée de preuve. On convolue avec une gaussienne. Soit

$$g_\varepsilon(x) = \frac{1}{\sqrt{2\pi\varepsilon}} e^{-x^2/(2\varepsilon)}, \quad (\mu * g_\varepsilon)(y) = \int \mu(dx) g_\varepsilon(x - y).$$

C'est une densité de probabilité $\mathcal{C}^\infty$; on vérifie que $(\mu * g_\varepsilon)(y) dy \Rightarrow \mu$ quand $\varepsilon \rightarrow 0$. Or

$$\phi_{g_\varepsilon}(z) = e^{-\varepsilon z^2/2} = g_{1/\varepsilon}(z) \sqrt{\frac{2\pi}{\varepsilon}}, \quad g_\varepsilon(z) = \frac{1}{\sqrt{2\pi\varepsilon}} \phi_{g_{1/\varepsilon}}(z).$$

Ainsi,

$$(\mu * g_\varepsilon)(y) = \int \mu(dx) \frac{1}{\sqrt{2\pi\varepsilon}} \phi_{g_{1/\varepsilon}}(x - y)$$

$$\begin{aligned}
&= \frac{1}{\sqrt{2\pi\varepsilon}} \int \mu(dx) \int dz e^{iz(x-y)} g_{1/\varepsilon}(z) \\
&\stackrel{\text{Fubini}}{=} \frac{1}{\sqrt{2\pi\varepsilon}} \int dz g_{1/\varepsilon}(z) e^{-izy} \phi_\mu(z).
\end{aligned}$$

Ainsi $\mu * g_\varepsilon$ s'exprime en fonction de ϕ_μ et ε . On conclut en faisant $\varepsilon \rightarrow 0$. $\square$

Lemme 1.20 (Lien entre queue et fonction caractéristique). Il existe $c > 0$ tel que, pour tout $\mu \in \mathcal{M}_1(\mathbb{R})$ et tout $K > 0$,

$$\mu(\mathbb{R} \setminus [-K, K]) \leq cK \int_{-1/K}^{1/K} (1 - \operatorname{Re} \phi_\mu(u)) du.$$

Démonstration. On a

$$\begin{aligned}
K \int_{-1/K}^{1/K} (1 - \operatorname{Re} \phi_\mu(u)) du &= K \int_{-1/K}^{1/K} \left(1 - \int_{\mathbb{R}} \cos(ux) \mu(dx) \right) du \\
&\stackrel{\text{Fubini}}{=} \int_{\mathbb{R}} \mu(dx) K \int_{-1/K}^{1/K} (1 - \cos(ux)) du \\
&= 2 \int_{\mathbb{R}} \mu(dx) \left(1 - \frac{\sin(x/K)}{x/K} \right) \\
&\geq 2 \int_{\mathbb{R} \setminus [-K, K]} \mu(dx) \left(1 - \frac{\sin(x/K)}{x/K} \right) \\
&\geq 2(1 - \sin(1)) \mu(\mathbb{R} \setminus [-K, K]),
\end{aligned}$$

car $\sin(y)/y \leq \sin(1)$ pour $|y| \geq 1$. $\square$

Preuve du théorème de Lévy. Étape 1 : tension. On a

$$\mu_n(\mathbb{R} \setminus [-A, A]) \leq cA \int_{-1/A}^{1/A} (1 - \operatorname{Re} \phi_{\mu_n}(u)) du \rightarrow cA \int_{-1/A}^{1/A} (1 - \operatorname{Re} \phi(u)) du.$$

La dernière quantité tend vers 0 quand $A \rightarrow \infty$, car ϕ est continue en 0 et $\phi(0) = 1$. On en déduit que

$$\lim_{A \rightarrow \infty} \limsup_{n \rightarrow \infty} \mu_n(\mathbb{R} \setminus [-A, A]) = 0.$$

Donc $(\mu_n)_{n \geq 1}$ est tendue. En particulier, il existe une extraction φ_0 et $\mu \in \mathcal{M}_1(\mathbb{R})$ telles que $\mu_{\varphi_0(n)} \Rightarrow \mu$. Alors $\phi_{\mu_{\varphi_0(n)}} \rightarrow \phi_\mu$ simplement. Donc $\phi = \phi_\mu$, ce qui montre déjà que ϕ est la fonction caractéristique d'une mesure de probabilité.

Étape 2 : unicité de la limite. Soit φ une extraction et $\nu \in \mathcal{M}_1(\mathbb{R})$ telles que $\mu_{\varphi(n)} \Rightarrow \nu$. Alors $\phi_{\mu_{\varphi(n)}} \rightarrow \phi_\nu$ simplement. Or $\phi_{\mu_n} \rightarrow \phi = \phi_\mu$ simplement. Donc $\phi_\mu = \phi_\nu$, et $\mu = \nu$ par injectivité de la fonction caractéristique. $\square$

On retrouve une structure du raisonnement en deux étapes, qui est le credo de ce cours. Il y a essentiellement trois ingrédients :

- (i) Tension.
 - (ii) Une certaine propriété $\mathcal{P}$ passe à la limite étroite (pour le théorème de Lévy, $\mu_n \Rightarrow \mu$ implique $\phi_{\mu_n} \rightarrow \phi_\mu$ simplement).
 - (iii) Cette propriété $\mathcal{P}$ caractérise les mesures de probabilité.
- Souvent, les points 2 et 3 sont appelés « identification de la limite » ou « unicité de la limite ».

C'est aussi pour ces raisons qu'il est souvent important, dans des espaces généraux, d'avoir une « bonne » caractérisation des compacts (en vue de la tension) et une « bonne » caractérisation des mesures de probabilité (en vue de l'unicité de la limite).

Interlude : séparabilité

Dans tout le cours, la notion de *séparabilité* est cruciale. On dit qu'un espace métrique E (ou plus généralement un espace topologique) est *séparable* s'il admet une suite dénombrable dense.

Par exemple, sont séparables $(\mathcal{C}_c(\mathbb{R}), \|\cdot\|_\infty)$ et tout espace métrique compact (utiliser Borel–Lebesgue), et $(\mathcal{C}_b(\mathbb{R}), \|\cdot\|_\infty)$ ne l'est pas (exercice).

La séparabilité est une propriété très utile lorsqu'on considère des produits d'espaces. Si E, F sont des espaces métriques séparables,

$$\mathcal{B}(E \times F) = \mathcal{B}(E) \otimes \mathcal{B}(F),$$

ce qui est faux en général. En particulier, si X, Y sont des variables aléatoires à valeurs dans (E, d) séparable, alors $d(X, Y)$ est mesurable comme composée de

$$\begin{aligned} (\Omega, \mathcal{F}) &\longrightarrow (E \times E, \mathcal{B}(E) \otimes \mathcal{B}(E)), \\ \omega &\longmapsto (X(\omega), Y(\omega)), \end{aligned}$$

et de

$$\begin{aligned} (E \times E, \mathcal{B}(E \times E)) &\longrightarrow \mathbb{R}, \\ (x, y) &\longmapsto d(x, y), \end{aligned}$$

la seconde application étant mesurable car continue, et les deux tribus intermédiaires étant égales.

La séparabilité est une notion topologique, au sens où, si E et F sont homéomorphes, alors E est séparable si et seulement si F l'est. N.B. : la complétude n'est pas une notion topologique.

2 Autres modes de convergence

Exercices associés : feuille 2, p. 111.

Dans ce chapitre

2.1	Convergence presque sûre, en probabilité	14
2.2	Uniforme intégrabilité et convergence dans L^1	15

2.1 Convergence presque sûre, en probabilité

Soient $(X_n)_{n \geq 1}$ et X des variables aléatoires définies sur le même espace de probabilité, à valeurs dans (E, d) séparable.

Définition 2.1. — On dit que X_n converge *presque sûrement* (p.s.) vers X , et on note $X_n \xrightarrow{\text{p.s.}} X$, si

$$\mathbb{P}(\{\omega \in \Omega : X_n(\omega) \rightarrow X(\omega)\}) = 1.$$

— On dit que X_n converge *en probabilité* vers X , et on note $X_n \xrightarrow{\mathbb{P}} X$, si

$$\forall \varepsilon > 0, \quad \mathbb{P}(d(X_n, X) > \varepsilon) \rightarrow 0.$$

Remarque 2.2. — $\mathbb{P}(X_n \rightarrow X)$ fait intervenir la loi jointe de $(X, X_1, X_2, X_3, \dots)$, tandis que $\mathbb{P}(d(X_n, X) > \varepsilon)$ ne fait intervenir que la loi jointe de (X_n, X) .

— On verra plus tard que les convergences presque sûre et en probabilité impliquent la convergence en loi, et que la convergence en loi vers une constante est équivalente à la convergence en probabilité vers cette constante.

Proposition 2.3. On a $X_n \xrightarrow{\mathbb{P}} X$ si et seulement si de toute sous-suite $(X_{\varphi(n)})$, on peut extraire une sous-sous-suite $(X_{\varphi \circ \psi(n)})$ qui converge presque sûrement vers X . En particulier, $X_n \xrightarrow{\text{p.s.}} X$ implique $X_n \xrightarrow{\mathbb{P}} X$.

Démonstration. Vérifions tout d'abord que

$$X_n \xrightarrow{\mathbb{P}} X \iff \mathbb{E}[d(X_n, X) \wedge 1] \rightarrow 0,$$

où $\wedge$ désigne le minimum. Pour le sens réciproque, on a, pour $0 < \varepsilon < 1$,

$$\mathbb{P}(d(X_n, X) > \varepsilon) \leq \frac{1}{\varepsilon} \mathbb{E}[d(X_n, X) \wedge 1] \rightarrow 0.$$

Pour le sens direct, on écrit, pour $\varepsilon > 0$,

$$\begin{aligned} \mathbb{E}[d(X_n, X) \wedge 1] &= \mathbb{E}[(d(X_n, X) \wedge 1) \mathbb{1}_{\{d(X_n, X) \leq \varepsilon\}}] \\ &\quad + \mathbb{E}[(d(X_n, X) \wedge 1) \mathbb{1}_{\{d(X_n, X) > \varepsilon\}}] \\ &\leq \varepsilon + \mathbb{P}(d(X_n, X) > \varepsilon). \end{aligned}$$

Donc $\limsup_{n \rightarrow \infty} \mathbb{E}[d(X_n, X) \wedge 1] = 0$. Revenons à la preuve de la proposition.

Sens direct. Soit φ une extraction. Il existe alors une extraction ψ telle que, pour tout $n \geq 1$,

$$\mathbb{E}[d(X_{\varphi \circ \psi(n)}, X) \wedge 1] \leq 2^{-n}.$$

On a

$$\mathbb{E} \left[\sum_{n=1}^{\infty} d(X_{\varphi \circ \psi(n)}, X) \wedge 1 \right] = \sum_{n=1}^{\infty} \mathbb{E}[d(X_{\varphi \circ \psi(n)}, X) \wedge 1] < \infty.$$

Donc presque sûrement $\sum_{n=1}^{\infty} d(X_{\varphi \circ \psi(n)}, X) \wedge 1 < \infty$, et donc $X_{\varphi \circ \psi(n)} \xrightarrow{\text{p.s.}} X$.

Sens réciproque. Par l'absurde, si $X_n \not\xrightarrow{\mathbb{P}} X$, on trouve $\varepsilon > 0$ et une extraction φ telles que, pour tout n , $\mathbb{E}[d(X_{\varphi(n)}, X) \wedge 1] \geq \varepsilon$. Soit alors ψ une extraction telle que $X_{\varphi \circ \psi(n)} \xrightarrow{\text{p.s.}} X$. Par convergence dominée,

$$\mathbb{E}[d(X_{\varphi \circ \psi(n)}, X) \wedge 1] \rightarrow 0,$$

ce qui est absurde. □

2.2 Uniforme intégrabilité et convergence dans L^1

Dans cette partie, $(X_n)_{n \geq 1}$ et X sont des variables aléatoires réelles intégrables définies sur le même espace de probabilité.

Définition 2.4. On dit que (X_n) converge vers X dans L^1 , et on note $X_n \xrightarrow{L^1} X$, si $\mathbb{E}[|X_n - X|] \rightarrow 0$.

Le théorème de convergence dominée permet de passer de la convergence presque sûre à la convergence L^1 avec une hypothèse de domination. Aussi, le lemme de Scheffé (voir la feuille d'exercices 1) dit que

$$X_n \geq 0, X \geq 0, X_n \xrightarrow{\text{p.s.}} X, \mathbb{E}[X_n] \rightarrow \mathbb{E}[X] \implies X_n \xrightarrow{L^1} X.$$

Le concept d'uniforme intégrabilité fait le lien entre convergence en probabilité et convergence L^1 .

Si $X \in L^1$, on a $\mathbb{E}[|X| \mathbb{1}_{\{|X| \geq x\}}] \rightarrow 0$ quand $x \rightarrow \infty$, par convergence dominée. L'uniforme intégrabilité étend cela à une famille de variables aléatoires.

Théorème 2.5. Soit $(X_i)_{i \in I}$ des variables aléatoires réelles intégrables. Les trois conditions suivantes sont équivalentes :

- (1) $\sup_{i \in I} \mathbb{E}[|X_i| \mathbb{1}_{\{|X_i| \geq x\}}] \rightarrow 0$ quand $x \rightarrow \infty$.
- (2) $\sup_{i \in I} \mathbb{E}[|X_i|] < \infty$ et, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que, pour tout A mesurable,

$$\mathbb{P}(A) \leq \delta \implies \sup_{i \in I} \mathbb{E}[|X_i| \mathbb{1}_A] \leq \varepsilon.$$

(3) Il existe une fonction croissante $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ telle que

$$\frac{\varphi(x)}{x} \longrightarrow +\infty \quad (x \rightarrow \infty), \quad \sup_{i \in I} \mathbb{E}[\varphi(|X_i|)] < \infty.$$

C'est le critère de de la Vallée Poussin.

Lorsque ces conditions sont réalisées, on dit que $(X_i)_{i \in I}$ est *uniformément intégrable*, ou, en abrégé, UI.

Avant de démontrer cela, mentionnons un corollaire utile de (2).

Corollaire 2.6. Une somme de deux familles UI est UI.

Preuve du théorème. (1) $\Rightarrow$ (2). Pour la bornitude dans L^1 , on écrit, pour x assez grand,

$$\mathbb{E}[|X_i|] \leq \mathbb{E}[|X_i| \mathbb{1}_{\{|X_i| < x\}}] + \mathbb{E}[|X_i| \mathbb{1}_{\{|X_i| \geq x\}}] \leq x + \sup_{i \in I} \mathbb{E}[|X_i| \mathbb{1}_{\{|X_i| \geq x\}}] < \infty.$$

Ensuite, on fixe $\varepsilon > 0$. Par hypothèse, il existe $x_\varepsilon > 0$ tel que $\sup_{i \in I} \mathbb{E}[|X_i| \mathbb{1}_{\{|X_i| \geq x_\varepsilon\}}] \leq \varepsilon$. Alors, si $\mathbb{P}(A) \leq \varepsilon/x_\varepsilon$, pour $i \in I$ on a

$$\begin{aligned} \mathbb{E}[|X_i| \mathbb{1}_A] &= \mathbb{E}[|X_i| \mathbb{1}_A \mathbb{1}_{\{|X_i| < x_\varepsilon\}}] + \mathbb{E}[|X_i| \mathbb{1}_A \mathbb{1}_{\{|X_i| \geq x_\varepsilon\}}] \\ &\leq x_\varepsilon \mathbb{P}(A) + \varepsilon \leq 2\varepsilon. \end{aligned}$$

(2) $\Rightarrow$ (1). D'après l'inégalité de Markov, si $x^{-1} \sup_{i \in I} \mathbb{E}[|X_i|] \leq \delta$, on a, pour tout $i \in I$, $\mathbb{P}(|X_i| \geq x) \leq \delta$ et donc $\mathbb{E}[|X_i| \mathbb{1}_{\{|X_i| \geq x\}}] \leq \varepsilon$.

(3) $\Rightarrow$ (1). Soit $\varepsilon > 0$ et $x_\varepsilon > 0$ tel que $x \leq \varepsilon \varphi(x)$ pour tout $x \geq x_\varepsilon$. Alors

$$\sup_{i \in I} \mathbb{E}[|X_i| \mathbb{1}_{\{|X_i| \geq x_\varepsilon\}}] \leq \varepsilon \sup_{i \in I} \mathbb{E}[\varphi(|X_i|) \mathbb{1}_{\{|X_i| \geq x_\varepsilon\}}] \leq \varepsilon \sup_{i \in I} \mathbb{E}[\varphi(|X_i|)].$$

(1) $\Rightarrow$ (3). Soit (x_n) une suite strictement croissante tendant vers $+\infty$ telle que

$$\sup_{i \in I} \mathbb{E}[|X_i| \mathbb{1}_{\{|X_i| \geq x_n\}}] \leq 2^{-n}.$$

On pose alors $\varphi(x) = \sum_{n=1}^{\infty} (x - x_n)_+$, pour $x \geq 0$, qui est croissante. De plus,

$$\frac{\varphi(x)}{x} = \sum_{n=1}^{\infty} \left(1 - \frac{x_n}{x}\right)_+ \longrightarrow +\infty$$

par convergence monotone, et

$$\mathbb{E}[\varphi(|X_i|)] \leq \sum_{n=1}^{\infty} \mathbb{E}[|X_i| \mathbb{1}_{\{|X_i| \geq x_n\}}] \leq 1.$$

□

On applique souvent le critère (3) avec $\varphi(x) = x^2$.

Remarque 2.7. On a vu (voir la feuille d'exercices 1) que $(X_i)_{i \in I}$ est tendue si et seulement s'il existe $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ telle que $\lim_{+\infty} \varphi = +\infty$ et $\sup_{i \in I} \mathbb{E}[\varphi(|X_i|)] < \infty$. C'est donc une notion plus faible que l'uniforme intégrabilité, et

$$\boxed{\text{UI} \implies \text{tension.}}$$

Proposition 2.8. Soit $(X_n)_{n \geq 1}$ des variables aléatoires réelles intégrables et X une variable aléatoire réelle. On a $X \in L^1$ et $X_n \xrightarrow{L^1} X$ si et seulement si $X_n \xrightarrow{\mathbb{P}} X$ et $(X_n)_{n \geq 1}$ est UI.

Démonstration. Sens direct. On suppose $X_n \xrightarrow{L^1} X$. Alors $X_n \xrightarrow{\mathbb{P}} X$ par Markov. Pour montrer que $(X_n)_{n \geq 1}$ est UI, il suffit de montrer que $(X_n - X)_{n \geq 1}$ est UI, puisque la somme de deux familles UI est UI par la caractérisation (2). Pour cela, si $\varepsilon > 0$, on choisit n_0 tel que $n \geq n_0$ implique $\mathbb{E}[|X_n - X|] \leq \varepsilon$. Pour ces n , toute espérance tronquée est donc majorée par ε . On choisit ensuite x_0 tel que, pour $x \geq x_0$,

$$\max_{1 \leq i < n_0} \mathbb{E}[|X_i - X| \mathbb{1}_{\{|X_i - X| \geq x\}}] \leq \varepsilon.$$

Alors $\sup_{i \geq 1} \mathbb{E}[|X_i - X| \mathbb{1}_{\{|X_i - X| \geq x\}}] \leq \varepsilon$.

Sens réciproque. Vérifions d'abord que $X \in L^1$. Soit φ une extraction telle que $X_{\varphi(n)} \xrightarrow{\text{p.s.}} X$. Alors, d'après le lemme de Fatou,

$$\mathbb{E}[|X|] \leq \liminf_{n \rightarrow \infty} \mathbb{E}[|X_{\varphi(n)}|] < \infty$$

par UI. Ainsi, comme ci-dessus, $(X_n - X)_{n \geq 1}$ est UI. En choisissant δ comme dans la caractérisation (2), si $\mathbb{P}(|X_n - X| \geq \varepsilon) \leq \delta$, on en déduit que

$$\mathbb{E}[|X_n - X|] \leq \mathbb{E}[|X_n - X| \mathbb{1}_{\{|X_n - X| \geq \varepsilon\}}] + \mathbb{E}[|X_n - X| \mathbb{1}_{\{|X_n - X| < \varepsilon\}}] \leq 2\varepsilon.$$

□

3 Mesures de probabilité sur un espace métrique

Exercices associés : feuille 3, p. 125.

Dans ce chapitre

3.1	Régularité de mesures	18
3.2	Théorème de porte-manteau	18
3.3	Convergence étroite dans $\mathbb{R}^N$	22
3.4	Tension et théorème de Prokhorov	24
3.5	Propriétés topologiques de $\mathcal{M}_1(E)$	29
3.6	Compléments de topologie	32

3.1 Régularité de mesures

Théorème 3.1. Soit (E, d) un espace métrique et $\mu \in \mathcal{M}_1(E)$. Pour tout $A \in \mathcal{B}(E)$,

$$\mu(A) = \inf\{\mu(O) : O \text{ ouvert}, A \subset O\} = \sup\{\mu(F) : F \text{ fermé}, F \subset A\}. \quad (1)$$

Démonstration. C'est une preuve à la « théorie de la mesure ». On note

$$\mathcal{A} = \{A \in \mathcal{B}(E) : A \text{ vérifie (1)}\}.$$

Si O est ouvert, $O \in \mathcal{A}$: pour l'infimum, c'est immédiat ; pour le supremum, soit $F_n = \{x \in E : d(x, O^c) \geq 1/n\}$, fermé. Comme $O = \bigcup_{n \geq 1} F_n$, on a $\mu(O) = \lim_{n \rightarrow \infty} \mu(F_n)$.

Montrons que $\mathcal{A}$ est une tribu. On a $\emptyset \in \mathcal{A}$; le passage au complémentaire est immédiat. Soit $(A_n)_{n \geq 1}$ une suite de $\mathcal{A}$, et F_n fermés, O_n ouverts tels que $F_n \subset A_n \subset O_n$ et $\mu(O_n \setminus F_n) \leq \varepsilon/2^n$. Alors $O = \bigcup_{n \geq 1} O_n$ est un ouvert contenant $\bigcup_{n \geq 1} A_n$ tel que

$$\mu\left(O \setminus \bigcup_{n \geq 1} A_n\right) \leq \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} \leq \varepsilon.$$

Cependant, $F' = \bigcup_{n \geq 1} F_n$ n'est pas toujours fermé. Pour contourner cela, on remarque que $\bigcup_{n=1}^N F_n \uparrow F'$ quand $N \rightarrow \infty$, et donc il existe $N \geq 1$ tel que $\mu(F' \setminus \bigcup_{n=1}^N F_n) \leq \varepsilon$. On pose $F = \bigcup_{n=1}^N F_n$. On a alors

$$\mu(O \setminus F) \leq \mu(O \setminus F') + \mu(F' \setminus F) \leq 2\varepsilon.$$

□

En particulier, une mesure de probabilité est caractérisée par sa valeur sur les ouverts (ou les fermés), ce qu'on peut directement voir avec le lemme des classes monotones.

3.2 Théorème de porte-manteau

Ici, (E, d) est un espace métrique.

Théorème 3.2 (Porte-manteau / Alexandrov). Soit $(\mu_n)_{n \geq 1}$ une suite de $\mathcal{M}_1(E)$ et $\mu \in \mathcal{M}_1(E)$. Les propriétés suivantes sont équivalentes :

- (1) $\mu_n \Rightarrow \mu$.
- (2) Pour toute fonction $f : E \rightarrow \mathbb{R}$ lipschitzienne bornée, $\mu_n(f) \rightarrow \mu(f)$.
- (3) Pour tout F fermé, $\limsup_{n \rightarrow \infty} \mu_n(F) \leq \mu(F)$.
- (4) Pour tout O ouvert, $\liminf_{n \rightarrow \infty} \mu_n(O) \geq \mu(O)$.
- (5) Pour tout $A \in \mathcal{B}(E)$ tel que $\mu(\partial A) = 0$, on a $\lim_{n \rightarrow \infty} \mu_n(A) = \mu(A)$.
- (6) Pour toute fonction $f : E \rightarrow \mathbb{R}$ mesurable bornée et continue en μ -presque tout point, on a $\mu_n(f) \rightarrow \mu(f)$.

L'intuition de (3) est que la masse ne peut pas sortir des fermés.

En termes probabilistes, si $\mathcal{L}(X_n) = \mu_n$ et $\mathcal{L}(X) = \mu$,

$$\begin{aligned} X_n \xrightarrow{\text{loi}} X &\iff \mathbb{E}[f(X_n)] \rightarrow \mathbb{E}[f(X)] \text{ pour toute } f : E \rightarrow \mathbb{R} \text{ lipschitzienne bornée} \\ &\iff \limsup_{n \rightarrow \infty} \mathbb{P}(X_n \in F) \leq \mathbb{P}(X \in F) \text{ pour tout fermé } F, \end{aligned}$$

et ainsi de suite. Avant la preuve, voyons quelques applications.

Proposition 3.3. Soient X_n, X des variables aléatoires à valeurs dans (E, d) telles que $X_n \xrightarrow{\text{loi}} X$. Soit $f : E \rightarrow \mathbb{R}$ mesurable, presque sûrement continue en X . Alors $f(X_n) \xrightarrow{\text{loi}} f(X)$.

Démonstration. Soit $g : \mathbb{R} \rightarrow \mathbb{R}$ continue bornée. Alors $g \circ f : E \rightarrow \mathbb{R}$ est mesurable bornée, presque sûrement continue en X . Donc, par (6) de porte-manteau, $\mathbb{E}[g \circ f(X_n)] \rightarrow \mathbb{E}[g \circ f(X)]$, ce qui est précisément la définition de $f(X_n) \xrightarrow{\text{loi}} f(X)$. $\square$

Application. Soit (E, d) un espace métrique séparable, et X_n, Y_n des variables aléatoires à valeurs dans E telles que $X_n \Rightarrow X$ et $d(X_n, Y_n) \xrightarrow{\mathbb{P}} 0$. Alors $Y_n \Rightarrow X$.

Démonstration. On applique porte-manteau. Soit F fermé, et notons $F_\varepsilon = \{x \in E : d(x, F) \leq \varepsilon\}$, fermé. Alors

$$\mathbb{P}(Y_n \in F) \leq \mathbb{P}(X_n \in F_\varepsilon) + \mathbb{P}(d(X_n, Y_n) \geq \varepsilon).$$

Donc

$$\limsup_{n \rightarrow \infty} \mathbb{P}(Y_n \in F) \leq \limsup_{n \rightarrow \infty} \mathbb{P}(X_n \in F_\varepsilon) \leq \mathbb{P}(X \in F_\varepsilon).$$

Or $\mathbb{P}(X \in F_\varepsilon) \rightarrow \mathbb{P}(X \in F)$ quand $\varepsilon \rightarrow 0$, car F est fermé, ce qui conclut. $\square$

Corollaire 3.4 (Lemme de Slutsky). Soient (E, d_E) et $(E', d_{E'})$ séparables. Si $X_n \Rightarrow X$ dans E et $X'_n \xrightarrow{\mathbb{P}} x'$ dans E' , avec $x' \in E'$ fixé, alors

$$(X_n, X'_n) \Rightarrow (X, x').$$

Démonstration. On munit $E \times E'$ de la distance

$$d((a, a'), (b, b')) = d_E(a, b) + d_{E'}(a', b').$$

Ainsi $(X_n, x') \Rightarrow (X, x')$, par composition de $X_n \xrightarrow{\text{loi}} X$ avec $x \mapsto (x, x')$ continue. Par ailleurs,

$$d((X_n, x'), (X_n, X'_n)) = d_{E'}(X'_n, x') \xrightarrow{\mathbb{P}} 0.$$

D'où le résultat. □

Application. Dans $\mathbb{R}$, si $X_n \xrightarrow{\text{loi}} X$ et $Y_n \xrightarrow{\mathbb{P}} 1$, alors $X_n/Y_n \xrightarrow{\text{loi}} X$, en définissant arbitrairement le quotient lorsque $Y_n = 0$. En effet, $(X_n, Y_n) \xrightarrow{\text{loi}} (X, 1)$ par Slutsky, puis $g(X_n, Y_n) \xrightarrow{\text{loi}} g(X, 1)$ avec

$$g(x, y) = \begin{cases} x/y & \text{si } y \neq 0, \\ 0 & \text{sinon,} \end{cases}$$

presque sûrement continue en $(X, 1)$.

Preuve du théorème de porte-manteau. (1) $\Rightarrow$ (2). Immédiat.

(2) $\Rightarrow$ (3). Soit F fermé et $K > 0$. On pose

$$g_{F,K}(x) = \max(1 - Kd(x, F), 0).$$

Cette fonction présente un plateau à hauteur 1 sur F et s'annule à distance $1/K$ de F . Alors $g_{F,K}$ est K -lipschitzienne et vérifie $\mathbb{1}_F \leq g_{F,K} \leq 1$. Donc $\mu_n(F) \leq \mu_n(g_{F,K})$, et

$$\limsup_{n \rightarrow \infty} \mu_n(F) \leq \mu(g_{F,K}).$$

Par convergence dominée, $\mu(g_{F,K}) \rightarrow \mu(F)$ quand $K \rightarrow \infty$, d'où le résultat.

(3) $\Rightarrow$ (4). Immédiat par passage au complémentaire.

(3) et (4) $\Rightarrow$ (5). Soit $A \in \mathcal{B}(E)$ tel que $\mu(\partial A) = 0$, de sorte que $\mu(A^\circ) = \mu(A) = \mu(\overline{A})$. Alors

$$\begin{aligned} \mu(A) = \mu(A^\circ) &\leq \liminf_{n \rightarrow \infty} \mu_n(A^\circ) \leq \limsup_{n \rightarrow \infty} \mu_n(A) \\ &\leq \limsup_{n \rightarrow \infty} \mu_n(\overline{A}) \leq \mu(\overline{A}) = \mu(A). \end{aligned}$$

En insérant également $\liminf_n \mu_n(A)$ entre les deux bornes, on voit que toutes les inégalités sont des égalités, et on obtient (5).

(5) $\Rightarrow$ (6). Soit f continue μ -presque partout, bornée par M . Quitte à écrire $f = f_+ - f_-$, on peut supposer $f \geq 0$. Alors, par Fubini,

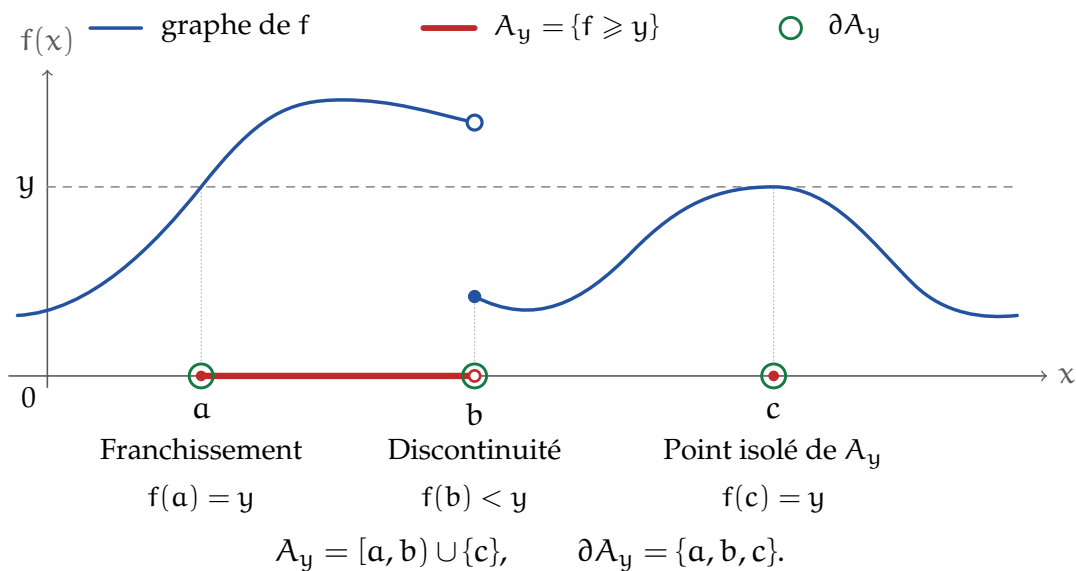
$$\begin{aligned}\mu_n(f) &= \int_E f(x) \mu_n(dx) = \int_E \int_0^\infty \mathbb{1}_{\{y \leq f(x)\}} dy \mu_n(dx) \\ &= \int_0^\infty dy \int_E \mathbb{1}_{\{y \leq f(x)\}} \mu_n(dx) = \int_0^\infty dy \mu_n(\{f \geq y\}) = \int_0^M \mu_n(A_y) dy,\end{aligned}$$

où $A_y = \{f \geq y\}$.

Pour appliquer (5), regardons ∂A_y . Soit D l'ensemble des points de discontinuité de f . Vérifions que

$$\partial A_y \subset \{x \in E : f(x) = y\} \cup D.$$

Exemple pour $E = \mathbb{R}$. La figure ci-dessous représente le graphe de f en bleu, l'ensemble $A_y = [a, b) \cup \{c\}$ en rouge sur l'axe des abscisses et sa frontière $\partial A_y = \{a, b, c\}$ en vert. En a , f franchit continûment le niveau y ; en b , elle présente un saut avec $f(b) < y$; en c , elle atteint le niveau y en un maximum local, donnant un point isolé de A_y . Ainsi, $f(a) = f(c) = y$ et $b \in D$.



En effet, soit $x_n \in A_y$ tel que $x_n \rightarrow x \in \partial A_y$. Si $x \in D$, il n'y a rien à faire. Si $x \notin D$, comme $f(x_n) \geq y$ pour tout n , on a $f(x) \geq y$. Alors, soit $f(x) = y$, soit $f(x) > y$, auquel cas $x \in A_y^\circ$, ce qui est impossible.

Or il y a au plus un nombre dénombrable de y tels que $\mu(\{f = y\}) > 0$: pour un entier $r \geq 1$, il y a au plus r valeurs de y telles que $\mu(\{f = y\}) \geq 1/r$, car μ est une probabilité. Donc, pour Lebesgue-presque tout y , $\mu(\partial A_y) = 0$ et $\mu_n(A_y) \rightarrow \mu(A_y)$. Par convergence dominée, on conclut que

$$\mu_n(f) \longrightarrow \int_0^M \mu(A_y) dy = \mu(f).$$

(6) $\Rightarrow$ (1). Immédiat. □

3.3 Convergence étroite dans $\mathbb{R}^{\mathbb{N}}$

On note $\mathbb{R}^{\mathbb{N}} = \{(x_1, x_2, \dots) : \forall i \geq 1, x_i \in \mathbb{R}\}$ l'ensemble des suites réelles.

3.3.1 Tribu sur $\mathbb{R}^{\mathbb{N}}$

Définition 3.5. La *tribu produit* sur $\mathbb{R}^{\mathbb{N}}$, notée $\mathcal{B}(\mathbb{R})^{\otimes \mathbb{N}}$, est la plus petite tribu sur $\mathbb{R}^{\mathbb{N}}$ qui rend, pour tout $k \geq 1$, les applications

$$\begin{aligned} \pi_k : \mathbb{R}^{\mathbb{N}} &\longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R})), \\ (x_i)_{i \geq 1} &\longmapsto x_k \end{aligned}$$

mesurables. Elle est parfois appelée *tribu cylindrique*.

Remarque 3.6. On vérifie que :

— Les ensembles de la forme

$$\{(x_i)_{i \geq 1} : x_{i_1} \in A_1, \dots, x_{i_k} \in A_k\} = \pi_{i_1}^{-1}(A_1) \cap \dots \cap \pi_{i_k}^{-1}(A_k),$$

pour $A_1, \dots, A_k \in \mathcal{B}(\mathbb{R})$ et $1 \leq i_1 < i_2 < \dots < i_k$, appelés *cylindres*, forment un π -système générateur de $\mathcal{B}(\mathbb{R})^{\otimes \mathbb{N}}$.

— Une fonction $(\Omega, \mathcal{F}) \rightarrow (\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R})^{\otimes \mathbb{N}})$, $\omega \mapsto (x_i(\omega))_{i \geq 1}$, est mesurable si et seulement si, pour tout $k \geq 1$, $\omega \mapsto x_k(\omega)$ est mesurable. En particulier, $(X_i)_{i \geq 1}$ est une variable aléatoire à valeurs dans $\mathbb{R}^{\mathbb{N}}$ si et seulement si chaque X_i est une variable aléatoire.

Définition 3.7. Soit $\mu \in \mathcal{M}_1(\mathbb{R}^{\mathbb{N}})$. Pour $1 \leq i_1 < \dots < i_k$, on note $\mu_{i_1, \dots, i_k}$ la mesure image de μ par l'application

$$\mathbb{R}^{\mathbb{N}} \longrightarrow \mathbb{R}^k, \quad (x_i)_{i \geq 1} \longmapsto (x_{i_1}, \dots, x_{i_k}).$$

Si $(X_i)_{i \geq 1}$ a pour loi μ , c'est la loi de $(X_{i_1}, \dots, X_{i_k})$. Les mesures $\{\mu_{i_1, \dots, i_k} : 1 \leq i_1 < \dots < i_k\}$ sont appelées *marginale de dimension k de μ* . La famille formée de toutes les marginales de dimension k , $k \geq 1$, est appelée *famille des marginales de dimension finie de μ* (fidi en abrégé).

Proposition 3.8. Une mesure $\mu \in \mathcal{M}_1(\mathbb{R}^{\mathbb{N}})$ est caractérisée par ses marginales fidi. En particulier, la loi d'une variable aléatoire $(X_i)_{i \geq 1} \in \mathbb{R}^{\mathbb{N}}$ est caractérisée par ses lois fidi.

Démonstration. Ceci provient du fait que les cylindres forment un π -système générateur de $\mathcal{B}(\mathbb{R})^{\otimes \mathbb{N}}$; deux mesures de $\mathcal{M}_1(\mathbb{R}^{\mathbb{N}})$ qui coïncident dessus sont donc égales d'après le lemme des classes monotones. $\square$

En termes probabilistes,

$$(X_i)_{i \geq 1} \stackrel{\text{loi}}{=} (Y_i)_{i \geq 1} \iff \forall k \geq 1, \forall 1 \leq i_1 < \dots < i_k, \quad (X_{i_1}, \dots, X_{i_k}) \stackrel{\text{loi}}{=} (Y_{i_1}, \dots, Y_{i_k}).$$

3.3.2 Distance sur $\mathbb{R}^{\mathbb{N}}$

Pour pouvoir parler de convergence étroite dans $\mathbb{R}^{\mathbb{N}}$, on définit une distance sur $\mathbb{R}^{\mathbb{N}}$.

Définition 3.9. Pour $(x_i)_{i \geq 1}, (y_i)_{i \geq 1} \in \mathbb{R}^{\mathbb{N}}$, on pose

$$d((x_i)_{i \geq 1}, (y_i)_{i \geq 1}) = \sum_{i=1}^{\infty} \frac{|x_i - y_i|}{1 + |x_i - y_i|} \frac{1}{2^i}.$$

Si $(x_i)_{i \geq 1}$ et $(y_i)_{i \geq 1}$ coïncident sur les K premiers termes (c'est-à-dire $x_i = y_i$ pour $i \leq K$), on a

$$d((x_i)_{i \geq 1}, (y_i)_{i \geq 1}) \leq \frac{1}{2^{K-1}}.$$

En particulier, les suites de rationnels nulles à partir d'un certain rang sont denses dans $\mathbb{R}^{\mathbb{N}}$, qui est donc séparable.

On vérifie aussi que (exercice), si $(x_i^n)_{i \geq 1} \in \mathbb{R}^{\mathbb{N}}$ et $(x_i)_{i \geq 1} \in \mathbb{R}^{\mathbb{N}}$,

$$d(x^n, x) \rightarrow 0 \iff \forall i \geq 1, \quad x_i^n \rightarrow x_i.$$

Remarque 3.10 (Facultative). La distance d métrise la topologie produit sur $\mathbb{R}^{\mathbb{N}}$, au sens où les ouverts définis par d coïncident avec la plus petite topologie (c'est-à-dire collection d'ouverts) pour laquelle toutes les projections π_k sont continues.

Lemme 3.11. La tribu borélienne $\mathcal{B}(\mathbb{R}^{\mathbb{N}})$ coïncide avec la tribu produit $\mathcal{B}(\mathbb{R})^{\otimes \mathbb{N}}$.

Démonstration. La continuité des projections π_k donne $\mathcal{B}(\mathbb{R})^{\otimes \mathbb{N}} \subset \mathcal{B}(\mathbb{R}^{\mathbb{N}})$. Pour l'autre inclusion, fixons $z \in \mathbb{R}^{\mathbb{N}}$. La fonction

$$x \mapsto d(x, z) = \lim_{K \rightarrow \infty} \sum_{i=1}^K 2^{-i} \frac{|x_i - z_i|}{1 + |x_i - z_i|}$$

est mesurable pour la tribu produit, comme limite de fonctions mesurables. Toute boule ouverte $B(z, r)$ appartient donc à cette tribu. Enfin, $\mathbb{R}^{\mathbb{N}}$ est séparable : les boules de centres dans une partie dense dénombrable et de rayons rationnels strictement positifs forment une base dénombrable d'ouverts. Tout ouvert est donc une union dénombrable de telles boules, ce qui prouve $\mathcal{B}(\mathbb{R}^{\mathbb{N}}) \subset \mathcal{B}(\mathbb{R})^{\otimes \mathbb{N}}$. $\square$

3.3.3 Convergence étroite dans $\mathbb{R}^{\mathbb{N}}$

Dans $\mathbb{R}^{\mathbb{N}}$, la convergence en loi est équivalente à la convergence des marginales fini-dimensionnelles.

Théorème 3.12. Soient $(X^n)_{n \geq 1}$ et X des variables aléatoires de $\mathbb{R}^{\mathbb{N}}$. Alors

$$X^n \xrightarrow{\text{loi}} X \iff \forall k \geq 1, \forall 1 \leq i_1 < \dots < i_k, \quad (X_{i_1}^n, \dots, X_{i_k}^n) \xrightarrow{\text{loi}} (X_{i_1}, \dots, X_{i_k}).$$

Démonstration. Le sens direct est clair par continuité de

$$\pi_{i_1, \dots, i_k} : \mathbb{R}^{\mathbb{N}} \longrightarrow \mathbb{R}^k, \quad (x_i)_{i \geq 1} \longmapsto (x_{i_1}, \dots, x_{i_k}).$$

Pour la réciproque, soit $f : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}$ lipschitzienne bornée. Soit $\varepsilon > 0$, et $\delta > 0$ tel que $d(x, y) \leq \delta$ implique $|f(x) - f(y)| \leq \varepsilon$, pour $x, y \in \mathbb{R}^{\mathbb{N}}$. On choisit alors K tel que $2^{-K} < \delta$. On écrit, en notant $T_K x = (x_1, \dots, x_K, 0, 0, \dots)$,

$$\begin{aligned} |\mathbb{E}[f(X^n)] - \mathbb{E}[f(X)]| &\leq \underbrace{\mathbb{E}[|f(X^n) - f(T_K X^n)|]}_{\leq \varepsilon} \\ &\quad + \underbrace{|\mathbb{E}[f(T_K X^n)] - \mathbb{E}[f(T_K X)]|}_{\rightarrow 0 \text{ par hypothèse}} \\ &\quad + \underbrace{\mathbb{E}[|f(T_K X) - f(X)|]}_{\leq \varepsilon}. \end{aligned}$$

Ceci conclut par porte-manteau. □

3.4 Tension et théorème de Prokhorov

Dans le cadre métrique général, la tension fait intervenir la notion de compacité. On a besoin de quelques résultats topologiques.

Définition 3.13. Soit (E, d) un espace métrique et $A \subset E$. On dit que A est *relativement compact* si $\overline{A}$ est compact.

Remarque 3.14. A est relativement compact si et seulement si de toute suite de A on peut extraire une sous-suite qui converge. En particulier, $\{x_n : n \geq 1\}$ est relativement compact si et seulement si, pour toute extraction φ , il existe une extraction ψ telle que $(x_{\varphi \circ \psi(n)})_{n \geq 1}$ converge.

Définition 3.15. Soit (E, d) un espace métrique et $A \subset E$. On dit que A est *précompact* si, pour tout $\varepsilon > 0$, A peut être recouvert par un nombre fini de boules de rayon ε .

Lemme 3.16. A est précompact si et seulement si $\overline{A}$ est précompact.

Démonstration. Le sens réciproque est clair. Pour le sens direct, si $A \subset \bigcup_{n=1}^N B(x_n, \varepsilon)$, vérifions que $\overline{A} \subset \bigcup_{n=1}^N B(x_n, 2\varepsilon)$. Soit $y \in \overline{A}$ et $y_m \rightarrow y$ avec $y_m \in A$. Soit m_0 tel que $d(y, y_{m_0}) < \varepsilon$, et soit n_1 tel que $y_{m_0} \in B(x_{n_1}, \varepsilon)$. Alors $y \in B(x_{n_1}, 2\varepsilon)$ par inégalité triangulaire. □

Avec la séparabilité, la complétude joue un rôle important dans le théorème de Prokhorov. Contrairement à la séparabilité, la complétude n'est pas une notion topologique : (E, d_1) et (E, d_2) peuvent être homéomorphes, avec (E, d_1) complet et (E, d_2) non complet. Par exemple, prendre $E = \mathbb{R}$, $d_1(x, y) = |x - y|$ et $d_2(x, y) = |\arctan(x) - \arctan(y)|$.

Dans un espace complet, la précompacité permet de caractériser les parties relativement compactes.

Exemple 3.17. Un espace métrique compact est précompact et complet.

Le résultat suivant sera très utile ; voir l'exercice pour la preuve.

Proposition 3.18. Soit (E, d) complet et $A \subset E$. Alors

$$A \text{ précompact} \iff A \text{ relativement compact.}$$

Définition 3.19. Soit (E, d) un espace métrique.

- On dit qu'une mesure $\mu \in \mathcal{M}_1(E)$ est *tendue* si, pour tout $\varepsilon > 0$, il existe K compact tel que $\mu(K) \geq 1 - \varepsilon$.
- On dit qu'une famille $(\mu_i)_{i \in I}$ de mesures est *tendue* si

$$\forall \varepsilon > 0, \quad \exists K \text{ compact, } \inf_{i \in I} \mu_i(K) \geq 1 - \varepsilon.$$

Une famille de variables aléatoires est tendue si la famille de leurs lois l'est.

Proposition 3.20. Soit $(\mu_i)_{i \in I}$ une famille tendue sur (E, d) et $f : (E, d) \rightarrow (F, d')$ continue. Alors $(f_*\mu_i)_{i \in I}$ est tendue, où $f_*\mu_i$ est la mesure image de μ_i par f .

Démonstration. Soient $\varepsilon > 0$ et K compact tels que $\inf_{i \in I} \mu_i(K) > 1 - \varepsilon$. Comme f est continue, $f(K)$ est compact. De plus $f^{-1}(f(K)) \supset K$, donc

$$\inf_{i \in I} f_*\mu_i(f(K)) = \inf_{i \in I} \mu_i(f^{-1}(f(K))) \geq \inf_{i \in I} \mu_i(K) > 1 - \varepsilon.$$

□

Définition 3.21 (Un peu simplifiée). On dit que (E, d) est *polonais* s'il est séparable et complet.

Remarque 3.22. La « vraie » définition d'un espace polonais est : un espace topologique séparable, métrisable avec une distance complète (cf. partie 4). Mais, pour simplifier dans ce cours, polonais signifie métrique, complet, séparable.

Proposition 3.23. Soit (E, d) polonais et $\mu \in \mathcal{M}_1(E)$. Alors μ est tendue.

Démonstration. Soit $(x_k)_{k \geq 1}$ une suite dense. Soit $\varepsilon > 0$. Pour tout $n \geq 1$, on choisit N_n tel que

$$\mu \left(\bigcup_{k=1}^{N_n} \bar{B}(x_k, 1/n) \right) \geq 1 - \frac{\varepsilon}{2^n}.$$

On pose alors $K = \bigcap_{n \geq 1} \bigcup_{k=1}^{N_n} \bar{B}(x_k, 1/n)$. K est fermé et précompact. Il est donc

compact. Par ailleurs,

$$\mu(K^c) \leq \sum_{n=1}^{\infty} \left(1 - \mu \left(\bigcup_{k=1}^{N_n} \bar{B}(x_k, 1/n) \right) \right) \leq \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon.$$

□

Théorème 3.24 (Prokhorov). Soit (E, d) métrique et $(\mu_n)_{n \geq 1}$ une suite de $\mathcal{M}_1(E)$.

- (i) Si $(\mu_n)_{n \geq 1}$ est tendue, alors, pour toute extraction φ , il existe une extraction ψ telle que $(\mu_{\varphi \circ \psi(n)})_{n \geq 1}$ converge étroitement.
- (ii) Si (E, d) est polonais, la réciproque est vraie.

On commence par un lemme.

Proposition 3.25. Soit (E, d_E) métrique séparable. Alors E est homéomorphe à un sous-ensemble de $(\mathbb{R}^{\mathbb{N}}, d)$.

Démonstration. Soit $(z^i)_{i \geq 1}$ une suite dense de E . On pose

$$\Phi : E \longrightarrow \mathbb{R}^{\mathbb{N}}, \quad x \longmapsto (d_E(x, z^i))_{i \geq 1}.$$

- Φ est continue car la convergence dans $\mathbb{R}^{\mathbb{N}}$ est la convergence des coordonnées.
- Supposons $\Phi(x) = \Phi(y)$. Si $x \neq y$, on choisit z^i tel que $d_E(x, z^i) < d_E(x, y)/2$. On ne peut avoir $d_E(x, z^i) = d_E(y, z^i)$, par l'inégalité triangulaire.
- Supposons $\Phi(x^n) \rightarrow \Phi(x)$. Par l'absurde, supposons $x^n \not\rightarrow x$. Il existe alors $\varepsilon > 0$ et une extraction φ tels que $d_E(x^{\varphi(n)}, x) \geq \varepsilon$ pour tout n . On choisit z^i tel que $d_E(z^i, x) < \varepsilon/2$. Alors $d_E(x^{\varphi(n)}, z^i) \geq \varepsilon/2$. Or $d_E(x^{\varphi(n)}, z^i) \rightarrow d_E(x, z^i)$: absurde.

Donc Φ est un homéomorphisme sur son image. □

Revenons à la preuve du théorème de Prokhorov.

Preuve du premier point. C'est le sens le plus utile, mais aussi le plus difficile. L'idée est de traiter successivement les cas $\mathbb{R}^d$, $\mathbb{R}^{\mathbb{N}}$, $E \subset \mathbb{R}^{\mathbb{N}}$, E .

- Cas $E = \mathbb{R}$. Il a été vu à la section 1.3 ; le cas $\mathbb{R}^d$ se traite de même (cf. feuille d'exercices ou Kallenberg, deuxième édition, proposition 5.21).
- Cas $E = \mathbb{R}^{\mathbb{N}}$. Notons

$$\widehat{\Pi}_k : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}^k, \quad (x_i)_{i \geq 1} \longmapsto (x_1, \dots, x_k),$$

qui est continue. L'image d'un compact par une fonction continue étant compacte, pour tout $k \geq 1$, $(\widehat{\Pi}_k \mu_n)_{n \geq 1}$ est tendue dans $\mathbb{R}^k$. Par procédé diagonal, pour toute extraction φ , il existe une extraction ψ telle que, si $\sigma = \varphi \circ \psi$, pour tout $k \geq 1$, $(\widehat{\Pi}_k \mu_{\sigma(n)})_{n \geq 1}$ converge étroitement vers une mesure $\nu_k \in \mathcal{M}_1(\mathbb{R}^k)$. Les $(\nu_k)_{k \geq 1}$ étant cohérentes (c'est-à-dire $\widehat{\Pi}_k \nu_{k+1} = \nu_k$), d'après le théorème

d'extension de Kolmogorov, il existe $\nu \in \mathcal{M}_1(\mathbb{R}^{\mathbb{N}})$ telle que $\widehat{\Pi}_k \nu = \nu_k$ pour tout $k \geq 1$. Il s'ensuit que $\mu_{\sigma(n)} \xrightarrow{\text{étroite}} \nu$, car dans $\mathbb{R}^{\mathbb{N}}$ la convergence étroite est équivalente à la convergence étroite des marginales finidimensionnelles.

- *Cas E σ -compact* (E est une union dénombrable de compacts). E étant séparable, il est homéomorphe à un sous-ensemble de $\mathbb{R}^{\mathbb{N}}$, qui est σ -compact (l'image d'un compact par une fonction continue est compacte), et a fortiori mesurable dans $\mathbb{R}^{\mathbb{N}}$.

La convergence étroite et la tension étant stables par composition par une fonction continue, on peut supposer que $E \subset \mathbb{R}^{\mathbb{N}}$ avec E mesurable. On vérifie que

$$\mathcal{B}(E) = \{A \in \mathcal{B}(\mathbb{R}^{\mathbb{N}}) : A \subset E\} \quad (\text{exercice}).$$

On étend μ_n sur $\mathbb{R}^{\mathbb{N}}$ en posant $\widehat{\mu}_n(A) = \mu_n(A \cap E)$ pour $A \in \mathcal{B}(\mathbb{R}^{\mathbb{N}})$; c'est-à-dire que $\widehat{\mu}_n$ est la mesure image de μ_n par l'inclusion $I : E \rightarrow \mathbb{R}^{\mathbb{N}}, x \mapsto x$, continue.

- a) Alors $(\widehat{\mu}_n)_{n \geq 1}$ est tendue comme image de $(\mu_n)_{n \geq 1}$ par I. Soit φ une extraction. D'après l'étape précédente, il existe une extraction ψ telle que, si $\sigma = \varphi \circ \psi$, $(\widehat{\mu}_{\sigma(n)})_{n \geq 1}$ converge étroitement vers $\widehat{\mu} \in \mathcal{M}_1(\mathbb{R}^{\mathbb{N}})$.
- b) Vérifions que $\widehat{\mu}(E) = 1$. Soit K_i compact de E tel que $\mu_{\sigma(n)}(K_i) \geq 1 - 1/i$ pour tout $n \geq 1$. Puisque K_i est fermé dans $\mathbb{R}^{\mathbb{N}}$ (exercice), par Portmanteau,

$$\widehat{\mu}(E) \geq \widehat{\mu}(K_i) \geq \limsup_{n \rightarrow \infty} \widehat{\mu}_{\sigma(n)}(K_i) = \limsup_{n \rightarrow \infty} \mu_{\sigma(n)}(K_i) \geq 1 - \frac{1}{i}.$$

En faisant $i \rightarrow \infty$, on obtient $\widehat{\mu}(E) = 1$. On définit alors μ sur E en posant $\mu(A) = \widehat{\mu}(A)$ pour $A \in \mathcal{B}(E)$.

- c) Vérifions que $\mu_{\sigma(n)} \xrightarrow{\text{étroite}} \mu$. On utilise Portmanteau. Soit O ouvert de E. Alors $O = E \cap \widehat{O}$, avec $\widehat{O}$ ouvert de $\mathbb{R}^{\mathbb{N}}$ (exercice), et

$$\liminf_{n \rightarrow \infty} \mu_{\sigma(n)}(O) = \liminf_{n \rightarrow \infty} \widehat{\mu}_{\sigma(n)}(\widehat{O}) \geq \widehat{\mu}(\widehat{O}) = \widehat{\mu}(\widehat{O} \cap E) = \mu(O),$$

car $\widehat{\mu}(E) = 1$.

- *Cas général*. Soit K_i compact tel que $\mu_n(K_i) \geq 1 - 1/i$ pour tout $n \geq 1$. On pose $E_0 = \bigcup_{i=1}^{\infty} K_i$, de sorte que $\mu_n(E_0) = 1$ pour tout $n \geq 1$. On applique le cas précédent à $\widetilde{\mu}_n$, définie sur E_0 par $\widetilde{\mu}_n(A) = \mu_n(A)$. En particulier, comme $\mu_n(E_0) = 1$, $\mu_n(A) = \widetilde{\mu}_n(A \cap E_0)$ pour $A \in \mathcal{B}(E)$. La famille $(\widetilde{\mu}_n)_{n \geq 1}$ est bien tendue car, pour tous $i, n \geq 1$, $\widetilde{\mu}_n(K_i) \geq 1 - 1/i$, avec K_i compact.

Soit φ une extraction; il existe une extraction ψ telle que, si $\sigma = \varphi \circ \psi$, alors $\widetilde{\mu}_{\sigma(n)} \xrightarrow{\text{étroite}} \widetilde{\mu} \in \mathcal{M}_1(E_0)$. On définit alors μ par $\mu(A) = \widetilde{\mu}(A \cap E_0)$ pour $A \in \mathcal{B}(E)$.

On conclut que $\mu_{\sigma(n)} \xrightarrow{\text{étroite}} \mu$, car $\mu_{\sigma(n)}$ et μ sont les mesures images de $\widetilde{\mu}_{\sigma(n)}$ et $\widetilde{\mu}$ par l'inclusion $I : E_0 \rightarrow E, x \mapsto x$, continue.

□

Preuve du deuxième point. On suppose (E, d) polonais et que, pour toute extraction φ , il existe une extraction ψ telle que $\mu_{\varphi \circ \psi(n)}$ converge étroitement. Fixons $\varepsilon > 0$ et une suite dense $(x_i)_{i \geq 1}$. Vérifions d'abord que

$$\forall k \geq 1, \quad \exists N_k \geq 1, \quad \forall n \geq 1, \quad \mu_n \left(\bigcup_{i=1}^{N_k} B(x_i, 1/k) \right) \geq 1 - \frac{\varepsilon}{2^k}.$$

Sinon, il existe $k \geq 1$ tel que cette propriété échoue. Pour toute mesure fixée μ_j , la masse de $\bigcup_{i=1}^{N_k} B(x_i, 1/k)$ tend vers 1 lorsque $N \rightarrow \infty$. Posons $\varphi(0) = 0$. À l'étape n , choisissons $N \geq n$ assez grand pour que tous les indices $j \leq \varphi(n-1)$ satisfassent la borne avec ce N . Puisque la propriété échoue, un indice $\varphi(n)$ viole cette borne au rang N ; il est nécessairement strictement supérieur à $\varphi(n-1)$. Par inclusion des recouvrements, il la viole aussi au rang n . Ainsi, pour tout n ,

$$\mu_{\varphi(n)} \left(\bigcup_{i=1}^n B(x_i, 1/k) \right) < 1 - \frac{\varepsilon}{2^k}.$$

Soit alors ψ une extraction telle que $\mu_{\varphi \circ \psi(n)} \xrightarrow{\text{étroite}} \mu$. Par Portmanteau, pour tout $m \geq 1$,

$$\begin{aligned} 1 - \frac{\varepsilon}{2^k} &\geq \liminf_{n \rightarrow \infty} \mu_{\varphi \circ \psi(n)} \left(\bigcup_{i=1}^{\psi(n)} B(x_i, 1/k) \right) \\ &\geq \liminf_{n \rightarrow \infty} \mu_{\varphi \circ \psi(n)} \left(\bigcup_{i=1}^m B(x_i, 1/k) \right) \geq \mu \left(\bigcup_{i=1}^m B(x_i, 1/k) \right). \end{aligned}$$

Donc, en faisant $m \rightarrow \infty$, $\mu(E) \leq 1 - \varepsilon/2^k$: absurde. On pose alors

$$K = \bigcap_{k \geq 1} \bigcup_{i=1}^{N_k} \overline{B}(x_i, 1/k).$$

Comme précédemment, pour montrer qu'une mesure de probabilité sur un espace polonais est tendue, on voit que K est compact et que $\mu_n(K) \geq 1 - \varepsilon$ pour tout $n \geq 1$. □

Corollaire 3.26. Soit (E, d) polonais, $(\mu_n)_{n \geq 1}$ une suite de $\mathcal{M}_1(E)$ et $\mu \in \mathcal{M}_1(E)$. Alors $\mu_n \xrightarrow{\text{étroite}} \mu$ si et seulement si

- $(\mu_n)_{n \geq 1}$ est tendue ;
- si φ est une extraction et $\nu \in \mathcal{M}_1(E)$ vérifie $\mu_{\varphi(n)} \xrightarrow{\text{étroite}} \nu$, alors $\mu = \nu$.

La preuve est la même que celle du corollaire 1.16 pour $E = \mathbb{R}$.

Reformulation probabiliste. Soient $(X_n)_{n \geq 1}$ et X des variables aléatoires à valeurs dans un espace polonais. Alors $X_n \xrightarrow{\text{loi}} X$ si et seulement si

- (X_n) est tendue, c'est-à-dire $\forall \varepsilon > 0, \exists K_\varepsilon$ compact, $\limsup_{n \rightarrow \infty} \mathbb{P}(X_n \notin K_\varepsilon) \leq \varepsilon$;
- si $X_{\varphi(n)} \xrightarrow{\text{loi}} Y$, alors $X \stackrel{\text{loi}}{=} Y$.

Application. Soient $(X_n)_{n \geq 1}$ et $(Y_n)_{n \geq 1}$ des variables aléatoires à valeurs dans (E, d) et (E', d') polonais. On suppose $X_n \xrightarrow{\text{loi}} X$, $Y_n \xrightarrow{\text{loi}} Y$ et $X_n \perp\!\!\!\perp Y_n$ pour tout n . Alors

$$(X_n, Y_n) \xrightarrow{\text{loi}} (\tilde{X}, \tilde{Y}), \quad \tilde{X} \stackrel{\text{loi}}{=} X, \quad \tilde{Y} \stackrel{\text{loi}}{=} Y, \quad \tilde{X} \perp\!\!\!\perp \tilde{Y}.$$

Démonstration. Étape 1 : $((X_n, Y_n))_{n \geq 1}$ est tendue. En effet, puisque $X_n \xrightarrow{\text{loi}} X$ et $Y_n \xrightarrow{\text{loi}} Y$, pour $\varepsilon > 0$, il existe des compacts K, K' de E, E' tels que $\mathbb{P}(X_n \in K) \geq 1 - \varepsilon$ et $\mathbb{P}(Y_n \in K') \geq 1 - \varepsilon$. Alors $\mathbb{P}((X_n, Y_n) \in K \times K') \geq 1 - 2\varepsilon$, avec $K \times K'$ compact de $E \times E'$.

Étape 2 : unicité de la limite. Si $(X_{\varphi(n)}, Y_{\varphi(n)}) \xrightarrow{\text{loi}} (\tilde{X}, \tilde{Y})$, alors, pour tous $f \in \mathcal{C}_b(E)$, $g \in \mathcal{C}_b(E')$,

$$\mathbb{E}[f(X_{\varphi(n)})g(Y_{\varphi(n)})] = \mathbb{E}[f(X_{\varphi(n)})]\mathbb{E}[g(Y_{\varphi(n)})].$$

En passant à la limite lorsque $n \rightarrow \infty$, on obtient

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

Comme dans la preuve de Portmanteau, on approche $\mathbb{1}_{F \times F'}$, avec F, F' fermés, et on conclut par classe monotone. $\square$

3.5 Propriétés topologiques de $\mathcal{M}_1(E)$

Soit (E, d) un espace métrique.

Définition 3.27. Pour $\mu, \nu \in \mathcal{M}_1(E)$, on pose

$$d_{LP}(\mu, \nu) = \inf\{\varepsilon > 0 : \forall A \in \mathcal{B}_E, \mu(A) \leq \nu(A^\varepsilon) + \varepsilon \text{ et } \nu(A) \leq \mu(A^\varepsilon) + \varepsilon\},$$

où $A^\varepsilon = \{x \in E : d(x, A) < \varepsilon\}$ est le ε -voisinage ouvert de A . d_{LP} est appelée *distance de Lévy-Prokhorov* sur $\mathcal{M}_1(E)$.

Proposition 3.28. d_{LP} est une distance.

Démonstration. La symétrie est immédiate. L'inégalité triangulaire provient de $(A^\varepsilon)^\eta \subset A^{\varepsilon+\eta}$. Pour la séparation, si $d_{LP}(\mu, \nu) = 0$, alors, pour tout fermé F ,

$$\mu(F) \leq \nu(F^{1/n}) + \frac{1}{n}, \quad \nu(F) \leq \mu(F^{1/n}) + \frac{1}{n}.$$

Comme $F = \bigcap_{n \geq 1} F^{1/n}$, en faisant $n \rightarrow \infty$, on en déduit $\mu(F) \leq \nu(F)$ et $\nu(F) \leq \mu(F)$. Donc $\mu = \nu$ sur les fermés, donc $\mu = \nu$. $\square$

Il peut parfois être utile d'utiliser une définition non symétrique de d_{LP} .

Lemme 3.29. On a

$$d_{LP}(\mu, \nu) = \inf\{\varepsilon > 0 : \forall A \in \mathcal{B}_E, \mu(A) \leq \nu(A^\varepsilon) + \varepsilon\}.$$

Démonstration. Soit $\varepsilon > 0$ tel que $\mu(A) \leq \nu(A^\varepsilon) + \varepsilon$ pour tout $A \in \mathcal{B}_E$. Soit $A \in \mathcal{B}_E$. On vérifie que $((A^\varepsilon)^c)^\varepsilon \subset A^c$, de sorte que

$$\nu(A) = 1 - \nu(A^c) \leq 1 - \nu(((A^\varepsilon)^c)^\varepsilon) \leq 1 - \mu((A^\varepsilon)^c) + \varepsilon = \mu(A^\varepsilon) + \varepsilon.$$

□

Lemme 3.30. Si E est séparable, pour toute probabilité $\mu \in \mathcal{M}_1(E)$ et tout $\eta > 0$, il existe un nombre fini de fonctions $f_1, \dots, f_m \in \mathcal{C}_b(E)$ telles que, pour toute $\nu \in \mathcal{M}_1(E)$,

$$\max_{1 \leq j \leq m} |\nu(f_j) - \mu(f_j)| < \eta \implies d_{LP}(\mu, \nu) \leq 3\eta.$$

Démonstration. Fixons une suite dense $(x_i)_{i \geq 1}$ dans E . Choisissons N tel que

$$\mu\left(\bigcup_{i=1}^N B(x_i, \eta)\right) > 1 - \eta.$$

Pour $I \subset \{1, \dots, N\}$ non vide, posons

$$B_I = \bigcup_{i \in I} B(x_i, \eta), \quad f_I(x) = \max\left(1 - \frac{d(x, B_I)}{\eta}, 0\right),$$

et posons $B_\emptyset = \emptyset$, $f_\emptyset = 0$. Ces 2^N fonctions sont continues, comprises entre 0 et 1. Supposons $|\nu(f_I) - \mu(f_I)| < \eta$ pour tout I . Pour un borélien A , posons

$$I_A = \{i \in \{1, \dots, N\} : A \cap B(x_i, \eta) \neq \emptyset\}.$$

Alors $A \setminus B_{I_A}$ est contenu dans le complémentaire du recouvrement et $\{f_{I_A} > 0\} \subset A^{3\eta}$. Par conséquent,

$$\begin{aligned} \mu(A) &\leq \mu(B_{I_A}) + \eta \leq \mu(f_{I_A}) + \eta \\ &< \nu(f_{I_A}) + 2\eta \leq \nu(A^{3\eta}) + 2\eta. \end{aligned}$$

La définition non symétrique de la distance de Lévy–Prokhorov, établie dans le lemme précédent, donne $d_{LP}(\mu, \nu) \leq 3\eta$. □

La distance de Lévy–Prokhorov est liée à la convergence étroite.

Proposition 3.31. Soit $(\mu_n)_{n \geq 1}$ une suite de $\mathcal{M}_1(E)$ et $\mu \in \mathcal{M}_1(E)$.

- (1) $d_{LP}(\mu_n, \mu) \rightarrow 0$ implique $\mu_n \xrightarrow{\text{étroite}} \mu$.
- (2) Si E est séparable, $\mu_n \xrightarrow{\text{étroite}} \mu$ implique $d_{LP}(\mu_n, \mu) \rightarrow 0$.

Ainsi, lorsque E est séparable, d_{LP} métrise la convergence étroite sur $\mathcal{M}_1(E)$.

Démonstration. (1) On suppose $d_{LP}(\mu_n, \mu) \rightarrow 0$. Soient F fermé, $\varepsilon > 0$. Soit n_0 tel que $n \geq n_0$ implique $\mu_n(F) \leq \mu(F^\varepsilon) + \varepsilon$. Donc $\limsup_{n \rightarrow \infty} \mu_n(F) \leq \mu(F^\varepsilon) + \varepsilon$. En faisant $\varepsilon \rightarrow 0$, on en déduit $\limsup_{n \rightarrow \infty} \mu_n(F) \leq \mu(F)$. Donc, par Portmanteau, $\mu_n \xrightarrow{\text{étroite}} \mu$.

(2) Soit E séparable et $\mu_n \xrightarrow{\text{étroite}} \mu$. Fixons $\eta > 0$ et choisissons les fonctions $f_1, \dots, f_m$ du lemme 3.30. Comme leur nombre est fini,

$$\max_{1 \leq j \leq m} |\mu_n(f_j) - \mu(f_j)| \rightarrow 0.$$

Pour n assez grand, le lemme donne donc $d_{LP}(\mu_n, \mu) \leq 3\eta$. Comme $\eta > 0$ est arbitraire, $d_{LP}(\mu_n, \mu) \rightarrow 0$. □

Corollaire 3.32 (Du théorème de Prokhorov). Si (E, d) est compact, alors $(\mathcal{M}_1(E), d_{LP})$ est compact.

Démonstration. Soit $(\mu_n)_{n \geq 1}$ une suite de $\mathcal{M}_1(E)$. Il suffit de montrer que $(\mu_n)_{n \geq 1}$ admet une valeur d'adhérence. Comme E est compact, $(\mu_n)_{n \geq 1}$ est tendue. Il existe donc une extraction φ et $\mu \in \mathcal{M}_1(E)$ telles que $\mu_{\varphi(n)} \xrightarrow{\text{étroite}} \mu$. Comme E est séparable, $\mu_{\varphi(n)} \rightarrow \mu$ dans $(\mathcal{M}_1(E), d_{LP})$. □

Remarque 3.33. Lorsque (E, d) est polonais, le théorème de Prokhorov se reformule ainsi : si $A \subset \mathcal{M}_1(E)$, la famille $\{\mu : \mu \in A\}$ est tendue si et seulement si A est relativement compact dans $(\mathcal{M}_1(E), d_{LP})$. Sous la seule hypothèse de séparabilité, l'implication de la tension vers la compacité relative reste vraie.

Théorème 3.34. Soit (E, d) un espace métrique.

- (1) Si E est séparable, alors $(\mathcal{M}_1(E), d_{LP})$ est séparable.
- (2) Si E est complet séparable, alors $(\mathcal{M}_1(E), d_{LP})$ est complet séparable.

Démonstration. (1) Soit $(x_i)_{i \geq 1}$ une suite dense de E . On montre que l'ensemble des mesures de la forme $\nu = \sum_{m \in M} r_m \delta_{x_m}$, avec $M \subset \mathbb{N}^*$ fini, $r_m \in \mathbb{Q}_+$ et $\sum_{m \in M} r_m = 1$, est dense.

Soit $\mu \in \mathcal{M}_1(E)$. Soient $\varepsilon > 0$ et $M \subset \mathbb{N}^*$, $\text{Card } M < \infty$, tels que $\mu(\bigcup_{m \in M} B(x_m, \varepsilon)) \geq 1 - \varepsilon$. En remplaçant les $B(x_m, \varepsilon)$ par des parties disjointes de

celles-ci, on écrit

$$\bigcup_{m \in M} B(x_m, \varepsilon) = \bigsqcup_{m \in M} B_m, \quad x_m \in B_m \subset B(x_m, \varepsilon).$$

On pose alors $\nu = \sum_{m \in M} r_m \delta_{x_m}$, avec les (r_m) choisis tels que $\sum_{m \in M} |r_m - \mu(B_m)| \leq \varepsilon$. Si $A \in \mathcal{B}(E)$, on pose

$$\tilde{A} = \bigcup_{\substack{m \in M \\ B_m \cap A \neq \emptyset}} B_m.$$

Alors $\mu(A) \leq \mu(\tilde{A}) + \mu(A \setminus \tilde{A}) \leq \mu(\tilde{A}) + \varepsilon$. Or $|\mu(\tilde{A}) - \nu(\tilde{A})| \leq \varepsilon$ et $\tilde{A} \subset A^{2\varepsilon}$, donc $\mu(A) \leq \nu(A^{2\varepsilon}) + 2\varepsilon$, donc $d_{LP}(\mu, \nu) \leq 2\varepsilon$.

(2) Soit $(\mu_n)_{n \geq 1}$ une suite de Cauchy pour d_{LP} . Il suffit de montrer que $(\mu_n)_{n \geq 1}$ est tendue (elle convergera alors vers cette valeur d'adhérence, par les propriétés des suites de Cauchy). Soient $\varepsilon > 0$ et une extraction φ tels que

$$\forall i \geq \varphi(n), \quad d_{LP}(\mu_i, \mu_{\varphi(n)}) < \frac{\varepsilon}{2^{n+1}}.$$

Par tension de $\mu_1, \dots, \mu_{\varphi(n)}$, il existe K_n compact tel que $\mu_i(K_n) \geq 1 - \varepsilon/2^{n+1}$ pour $i = 1, \dots, \varphi(n)$. Alors, pour $i \geq \varphi(n)$,

$$\mu_i(K_n^{\varepsilon/2^{n+1}}) \geq \mu_{\varphi(n)}(K_n) - \frac{\varepsilon}{2^{n+1}} \geq 1 - \frac{\varepsilon}{2^n}.$$

C'est aussi vrai pour $i = 1, 2, \dots, \varphi(n)$. En posant

$$K = \bigcap_{n \geq 1} \overline{K_n^{\varepsilon/2^{n+1}}},$$

on vérifie que K est compact (car précompact et fermé) et de mesure $\geq 1 - \varepsilon$. Ceci conclut. $\square$

3.6 Compléments de topologie

Quelques rappels de topologie (cf. par exemple les notes de cours *Topologie, analyse et calcul différentiel* de Frédéric Paulin, disponibles sur sa page web).

— Une *topologie* sur un ensemble E est un ensemble $\mathcal{O}$ de parties de E , appelées *ouverts*, tel que :

- (i) toute intersection finie d'éléments de $\mathcal{O}$ appartient à $\mathcal{O}$;
- (ii) toute union d'éléments de $\mathcal{O}$ appartient à $\mathcal{O}$.

Un espace topologique est un ensemble muni d'une topologie.

— Une *base d'ouverts* d'un espace topologique $(E, \mathcal{O})$ est une partie $\mathcal{B}$ de $\mathcal{O}$ telle que tout ouvert de E soit union d'éléments de $\mathcal{B}$, c'est-à-dire

$$\forall U \in \mathcal{O}, \quad \forall x \in U, \quad \exists V \in \mathcal{B}, \quad x \in V \subset U.$$

Il est possible de montrer que si $(E, \mathcal{O})$ a une base dénombrable d'ouverts, alors E est séparable. La réciproque est vraie dans un espace métrique, mais fausse en général (prendre par exemple $E = \mathbb{R}$ muni de la topologie cofinie, où les ouverts sont $\emptyset$ et les ensembles de la forme $\mathbb{R} \setminus F$ avec $F \subset \mathbb{R}$ fini).

- En général, la propriété « être une base d'ouverts d'une certaine topologie » n'est pas automatique pour un ensemble de parties d'un ensemble. On a toutefois le critère suivant : si $\mathcal{B}$ est une partie de $\mathcal{P}(E)$ telle que $\bigcup \mathcal{B} = E$ et telle que

$$\forall U, V \in \mathcal{B}, \quad \forall x \in U \cap V, \quad \exists W \in \mathcal{B}, \quad x \in W \subset U \cap V,$$

alors l'ensemble $\mathcal{O}$ des unions d'éléments de $\mathcal{B}$ est la topologie engendrée par $\mathcal{B}$. En particulier, $\mathcal{B}$ est une base d'ouverts de sa topologie engendrée (plus petite topologie qui contient $\mathcal{B}$).

- Si $(E, \mathcal{O})$ est un espace topologique, on dit qu'une suite (x_n) d'éléments de E converge vers $x \in E$ si, pour tout $U \in \mathcal{O}$ tel que $x \in U$, il existe N tel que $n \geq N \Rightarrow x_n \in U$.

Attention : deux topologies différentes peuvent avoir les mêmes suites convergentes (prendre par exemple $\mathbb{N}$ muni de la topologie discrète, qui est métrisable, τ_1 , et

$$\tau_2 = \left\{ A \subset \mathbb{N} : 1 \notin A \text{ ou } \sum_{n \in \mathbb{N} \setminus A} \frac{1}{n} < \infty \right\}.$$

- Si d est une distance sur E , l'ensemble $\mathcal{O}$ des parties U de E telles que $\forall x \in U, \exists \varepsilon > 0, B(x, \varepsilon) \subset U$, est une topologie sur E , appelée *topologie induite* par d . Un espace topologique $(E, \mathcal{O})$ est dit *métrisable* s'il existe une telle distance d .
- Les notions « être à base dénombrable d'ouverts », « séparabilité », « convergence », « être métrisable » sont invariantes par homéomorphismes.

Il est parfois plus naturel de directement définir une topologie sans référence à une métrique, qui peut être peu naturelle, peu commode à utiliser, ou rarement canonique. Par ailleurs, un espace topologique peut être complet ou non complet pour des métriques qui le métrisent.

C'est pour cette raison que la « bonne » définition d'un espace polonais est : *un espace topologique séparable, métrisable par une distance complète*. Mettons en œuvre ce qui précède dans le cadre d'une topologie sur l'espace des mesures de probabilité.

Soit E un espace topologique métrisable. Compte tenu de la définition de la convergence étroite, il est naturel de considérer la plus petite topologie sur $\mathcal{M}_1(E)$ pour laquelle les fonctions $\mu \mapsto \mu(f)$ sont continues pour $f \in \mathcal{C}_b(E)$ (c'est-à-dire la *topologie initiale* sur $\mathcal{M}_1(E)$ définie par les applications $\mu \mapsto \mu(f)$ pour $f \in \mathcal{C}_b(E)$), appelée *topologie faible-étoile* sur $\mathcal{M}_1(E)$ (aussi abrégée en faible-*).

Lemme 3.35. La famille de parties

$$B_{(\mu, \varepsilon, f_1, \dots, f_n)} = \{\nu \in \mathcal{M}_1(E) : |\mu(f_i) - \nu(f_i)| < \varepsilon \text{ pour tout } 1 \leq i \leq n\},$$

pour $\varepsilon > 0$, $f_1, \dots, f_n \in \mathcal{C}_b(E)$ et $\mu \in \mathcal{M}_1(E)$, constitue une base d'ouverts de la topologie faible-* sur $\mathcal{M}_1(E)$.

Démonstration. Ces parties sont bien des ouverts de la topologie faible-*. La topologie engendrée par cette famille rend les fonctions $\mu \mapsto \mu(f)$, pour $f \in \mathcal{C}_b(E)$, continues, donc elle contient la topologie faible-*. $\square$

Proposition 3.36. (1) Soient $\mu_n, \mu \in \mathcal{M}_1(E)$. Alors $\mu_n \xrightarrow{\text{étroite}} \mu$ si et seulement si $\mu_n \rightarrow \mu$ au sens de la topologie faible-*.

(2) Lorsque E est séparable, d_{LP} métrise la topologie faible-*.

Démonstration. (1) On suppose $\mu_n \xrightarrow{\text{étroite}} \mu$. Soit U un ouvert faible-* contenant μ . Il existe alors un ouvert de la forme $B_{(\mu, \varepsilon, f_1, \dots, f_k)}$ inclus dans U . Puisque $\mu_n \xrightarrow{\text{étroite}} \mu$, $\max_{1 \leq i \leq k} |\mu_n(f_i) - \mu(f_i)| \rightarrow 0$ et donc $\mu_n \in B_{(\mu, \varepsilon, f_1, \dots, f_k)}$ pour n assez grand.

Réciproquement, on suppose $\mu_n \rightarrow \mu$ au sens de la topologie faible-*. Alors, pour tout $f \in \mathcal{C}_b(E)$ et tout $\varepsilon > 0$, $\mu_n \in B_{(\mu, \varepsilon, f)}$ pour n assez grand, c'est-à-dire $|\mu_n(f) - \mu(f)| < \varepsilon$.

(2) Soit U un ouvert faible-* et $\mu \in U$. Raisonnons par l'absurde en supposant que, pour tout $\varepsilon > 0$, $B_{d_{LP}}(\mu, \varepsilon) \not\subset U$. Il existe alors une suite $\mu_n \in \mathcal{M}_1(E)$ telle que $\mu_n \rightarrow \mu$ pour d_{LP} et $\mu_n \notin U$ pour tout $n \geq 0$. Mais on a vu que $\mu_n \rightarrow \mu$ pour d_{LP} implique $\mu_n \xrightarrow{\text{étroite}} \mu$, et par (1) cela implique que $\mu_n \rightarrow \mu$ pour la topologie faible-* : absurde.

Réciproquement, fixons $\mu \in \mathcal{M}_1(E)$ et $\varepsilon > 0$. Choisissons $0 < \eta < \varepsilon/3$ et les fonctions $f_1, \dots, f_m$ du lemme 3.30. L'ensemble

$$V = \{\nu \in \mathcal{M}_1(E) : |\nu(f_j) - \mu(f_j)| < \eta \text{ pour tout } 1 \leq j \leq m\}$$

est un voisinage faible-* de μ , et le lemme donne

$$V \subset \{\nu : d_{LP}(\mu, \nu) \leq 3\eta\} \subset B_{d_{LP}}(\mu, \varepsilon).$$

Toute boule de centre μ est donc un voisinage faible-* de μ . Les deux topologies coïncident. $\square$

Lorsque E est un espace topologique métrisable et que $\mathcal{M}_1(E)$ est muni de la topologie faible-*, il est possible de montrer que

$$E \text{ compact} \iff \mathcal{M}_1(E) \text{ compact},$$

$$E \text{ séparable} \iff \mathcal{M}_1(E) \text{ séparable},$$

$$E \text{ polonais} \iff \mathcal{M}_1(E) \text{ polonais}.$$

4 L'espace des fonctions continues sur un compact

Exercices associés : feuille 4, p. 155.

Dans ce chapitre

4.1	Caractère polonais	35
4.2	Marginales finidimensionnelles	36
4.3	Convergence en loi dans cet espace	38
4.4	Compacité dans C	39
4.5	Tension	41
4.6	Critère de tension de Kolmogorov	43
4.7	Fonctions continues sur $\mathbb{R}_+$	46

Notation. Si K, E sont des espaces métriques avec K compact, on note $\mathcal{C}(K, E)$ l'ensemble des fonctions continues de K dans E , muni de la distance $d(f, g) = \sup_{x \in K} d_E(f(x), g(x))$.

4.1 Caractère polonais

Proposition 4.1. Soit (K, ρ) un espace métrique compact et (E, d_E) un espace polonais. Alors $(\mathcal{C}(K, E), d)$ est polonais.

Démonstration. Complétude. Si $(f_n)_{n \geq 1}$ est de Cauchy dans $\mathcal{C}(K, E)$, alors, pour tout $x \in K$, $(f_n(x))_{n \geq 1}$ est de Cauchy dans E , donc converge vers une limite notée $f(x)$. On vérifie alors que $d(f_n, f) \rightarrow 0$ et que $f \in \mathcal{C}(K, E)$.

Séparabilité. On pose, pour $m, n \geq 1$,

$$C_{m,n} = \{f \in \mathcal{C}(K, E) : \forall x, y \in K, \rho(x, y) < 1/m \Rightarrow d_E(f(x), f(y)) < 1/n\}.$$

Puisque K est compact, pour tout $m \geq 1$, il existe un ensemble fini K_m tel que $K = \bigcup_{x \in K_m} B(x, 1/(2m))$. On note $K_m = \{x_1^m, \dots, x_{k_m}^m\}$. Aussi, par séparabilité de E , pour tout $n \geq 1$, il existe un ensemble dénombrable $\mathcal{U}_n = \{U_1^n, U_2^n, \dots\}$ d'ouverts qui recouvrent E et de diamètre $< 1/n$. Pour un k_m -uplet $(i_1, \dots, i_{k_m})$ d'entiers ≥ 1 , on note

$$F_{(i_1, \dots, i_{k_m})}^{m,n} = \{f \in C_{m,n} : f(x_j^m) \in U_{i_j}^n \text{ pour tout } 1 \leq j \leq k_m\}.$$

Lorsque $F_{(i_1, \dots, i_{k_m})}^{m,n} \neq \emptyset$, on choisit une fonction $f_{(i_1, \dots, i_{k_m})}^{m,n} \in F_{(i_1, \dots, i_{k_m})}^{m,n}$. Soit $D_{m,n}$ l'ensemble de telles fonctions.

Vérifions que, pour tout $f \in C_{m,n}$, il existe $g \in D_{m,n}$ tel que $d_E(f(x_j^m), g(x_j^m)) \leq 1/n$ pour tout $1 \leq j \leq k_m$. Soit $f \in C_{m,n}$. On observe que

$$C_{m,n} = \bigcup_{i_1, \dots, i_{k_m} \geq 1} F_{(i_1, \dots, i_{k_m})}^{m,n}.$$

Choisissons $i_1, \dots, i_{k_m}$ tels que $f(x_j^m) \in U_{i_j}^n$ pour $1 \leq j \leq k_m$. En particulier $F_{(i_1, \dots, i_{k_m})}^{m,n} \neq \emptyset$. On peut donc prendre $g = f_{(i_1, \dots, i_{k_m})}^{m,n}$, qui convient, car $\text{diam}(U_i^n) < 1/n$ pour tout i .

Soit maintenant $D = \bigcup_{m,n} D_{m,n}$, dénombrable. Vérifions que D est dense dans $\mathcal{C}(K, E)$. Soient $f \in \mathcal{C}(K, E)$ et $\varepsilon > 0$. Fixons $n > 3/\varepsilon$. Puisque f est uniformément continue, il existe $m \geq 1$ tel que $f \in C_{m,n}$. Soit alors $g \in D_{m,n}$ tel que $d_E(f(x_j^m), g(x_j^m)) < 1/n$ pour tout $1 \leq j \leq k_m$.

Par inégalité triangulaire, on en déduit, pour tout $x \in K$, avec x_j^m tel que $x \in B(x_j^m, 1/(2m))$,

$$d_E(f(x), g(x)) \leq d_E(f(x), f(x_j^m)) + d_E(f(x_j^m), g(x_j^m)) + d_E(g(x_j^m), g(x)) \leq \frac{3}{n} \leq \varepsilon.$$

□

4.2 Marginales finidimensionnelles

4.2.1 L'espace des fonctions

Soit (E, d) un espace polonais et I un ensemble quelconque. L'ensemble E^I des fonctions de I dans E est naturellement muni de la tribu produit (comme pour $\mathbb{R}^{\mathbb{N}}$ à la section 3.3).

Définition 4.2. La *tribu produit* sur E^I , notée $\mathcal{B}(E)^{\otimes I}$, est la plus petite tribu sur E^I qui rend, pour tout $t \in I$, les applications

$$\pi_t : E^I \longrightarrow (E, \mathcal{B}(E)), \quad (x_i)_{i \in I} \longmapsto x_t,$$

mesurables. Elle est parfois appelée *tribu cylindrique*.

Remarque 4.3. On vérifie que :

— les ensembles de la forme

$$\{(x_i)_{i \in I} : x_{t_1} \in A_1, \dots, x_{t_k} \in A_k\} = \pi_{t_1}^{-1}(A_1) \cap \dots \cap \pi_{t_k}^{-1}(A_k),$$

pour $A_1, \dots, A_k \in \mathcal{B}(E)$ et $t_1, \dots, t_k \in I$, appelés *cylindres*, forment un π -système générateur de $\mathcal{B}(E)^{\otimes I}$;

— une fonction $(\Omega, \mathcal{A}) \rightarrow (E^I, \mathcal{B}(E)^{\otimes I})$, $\omega \mapsto (x_i(\omega))_{i \in I}$, est mesurable si et seulement si, pour tout $i \in I$, $\omega \mapsto x_i(\omega)$ est mesurable. En particulier, $(X_i)_{i \in I}$ est une variable aléatoire à valeurs dans E^I si et seulement si X_i est une variable aléatoire pour tout $i \in I$.

Définition 4.4. Soit $\mu \in \mathcal{M}_1(E^I)$. Pour $i_1, \dots, i_k \in I$, on note $\mu_{i_1, \dots, i_k}$ la mesure image de μ sur $(E^k, \mathcal{B}(E)^{\otimes k})$ par l'application

$$E^I \longrightarrow E^k, \quad (x_i)_{i \in I} \longmapsto (x_{i_1}, \dots, x_{i_k}).$$

Si $(X_i)_{i \in I}$ a pour loi μ , c'est la loi de $(X_{i_1}, \dots, X_{i_k})$. Les mesures $\{\mu_{i_1, \dots, i_k} : i_1, \dots, i_k \in I\}$ sont appelées *marginales de dimension k de μ* . La famille formée de toutes les

marginales de dimension k , $k \geq 1$, est appelée *famille des marginales de dimension finie* de μ (fidi en abrégé).

Proposition 4.5. Une mesure $\mu \in \mathcal{M}_1(E^I)$ est caractérisée par ses marginales fidi. En particulier, la loi d'une variable aléatoire $(X_i)_{i \in I} \in E^I$ est caractérisée par ses lois fidi.

Démonstration. Ceci provient du fait que les cylindres forment un π -système générateur de $\mathcal{B}(E)^{\otimes I}$, et donc deux mesures de $\mathcal{M}_1(E^I)$ qui coïncident dessus sont égales d'après le lemme des classes monotones. $\square$

L'avantage de ce point de vue est l'existence de variables aléatoires.

Théorème 4.6 (Extension de Kolmogorov). Soit E un espace polonais et I un ensemble quelconque. Pour chaque partie finie non vide $J \subset I$, on se donne une mesure de probabilité μ_J sur E^J . On suppose que cette famille est cohérente : pour toutes parties finies non vides $J \subset K \subset I$, la projection $\pi_{K,J} : E^K \rightarrow E^J$ vérifie

$$(\pi_{K,J})_* \mu_K = \mu_J.$$

Alors il existe une unique mesure de probabilité μ sur $(E^I, \mathcal{B}(E)^{\otimes I})$ dont la projection sur E^J est μ_J pour toute partie finie non vide $J \subset I$.

Ceci permet par exemple de définir le (pré)mouvement brownien comme variable aléatoire à valeurs dans $\mathbb{R}^{\mathbb{R}_+}$.

Problème : en général, E^I n'est pas polonais. Ce cadre n'est pas bien adapté à la convergence étroite !

4.2.2 L'espace des fonctions continues

Notons ici $C = \mathcal{C}(K, E)$, avec K compact, (E, d) un espace polonais, et $\mathcal{B}_C$ la tribu borélienne sur C . On note $\mathcal{P}_C$ la trace sur C de la tribu produit $\mathcal{B}(E)^{\otimes K}$, c'est-à-dire

$$\mathcal{P}_C = \{A \cap C : A \in \mathcal{B}(E)^{\otimes K}\}.$$

De manière équivalente, c'est la plus petite tribu sur C qui rend toutes les projections $\pi_t : C \rightarrow E, f \mapsto f(t)$, mesurables (exercice).

Proposition 4.7. Les tribus $\mathcal{B}_C$ et $\mathcal{P}_C$ sont les mêmes.

Démonstration. Puisque, pour tout $x \in K$, $\pi_x : C \rightarrow E$ est continue, elle est mesurable pour $\mathcal{B}_C$. Donc $\mathcal{P}_C \subset \mathcal{B}_C$.

Pour montrer que $\mathcal{B}_C \subset \mathcal{P}_C$, il y a deux étapes.

Étape 1 : les boules ouvertes de C sont mesurables pour $\mathcal{P}_C$. Pour le voir, on fixe $f_0 \in C$ et on considère $F : C \rightarrow \mathbb{R}, f \mapsto d(f, f_0)$. Puisque K est compact, il existe $D \subset K$

dénombrable dense. Par continuité,

$$d(f, f_0) = \sup_{x \in D} d(f(x), f_0(x)).$$

Puisque, pour tout $x \in D$, $(C, \mathcal{P}_C) \rightarrow \mathbb{R}$, $f \mapsto d(f(x), f_0(x))$, est mesurable (c'est la composée $\phi \circ \pi_x$ avec $\phi : E \rightarrow \mathbb{R}$, $y \mapsto d(y, f_0(x))$, continue donc mesurable), $F : (C, \mathcal{P}_C) \rightarrow \mathbb{R}$ est mesurable comme le supremum d'une famille dénombrable d'applications mesurables. Ainsi, pour tout $r > 0$, $F^{-1}([0, r]) \in \mathcal{P}_C$, c'est-à-dire $\{f \in C : d(f, f_0) < r\} \in \mathcal{P}_C$.

Étape 2. Par séparabilité de C , tout ouvert de C est union dénombrable de boules ouvertes (exercice). Cela conclut. $\square$

D'après ce qui précède, les marginales fidi $\{\mu_{t_1, \dots, t_k} : t_1, \dots, t_k \in K\}$ d'une mesure $\mu \in \mathcal{M}_1(C)$ la caractérisent.

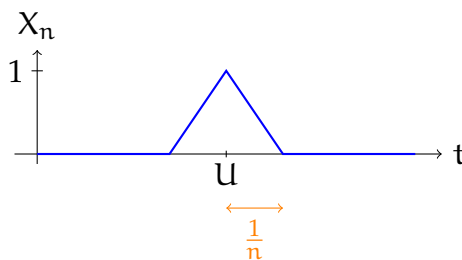
Attention : l'ensemble $\{f \in \mathbb{R}^{[0,1]} : f \text{ est continue}\}$ n'est pas mesurable par rapport à $\mathcal{B}(\mathbb{R})^{\otimes [0,1]}$ (exercice).

4.3 Convergence en loi dans cet espace

Définition 4.8. Soient $\mu_n, \mu \in \mathcal{M}_1(C)$. On dit que μ_n converge vers μ au sens des fidi si, pour tous $t_1, \dots, t_k \in K$, $(\mu_n)_{t_1, \dots, t_k} \xrightarrow{\text{étroite}} \mu_{t_1, \dots, t_k}$. On note $\mu_n \xrightarrow{\text{f.d.}} \mu$. De même, une suite $(X_n)_{n \geq 1}$ de variables aléatoires à valeurs dans C converge vers X au sens des fidi si, pour tous $t_1, \dots, t_k \in K$,

$$(X_n(t_1), \dots, X_n(t_k)) \xrightarrow{\text{loi}} (X(t_1), \dots, X(t_k)).$$

Attention ! Par continuité des projections, la convergence étroite dans C implique la convergence des fidi. Cependant, contrairement à $\mathbb{R}^{\mathbb{N}}$, la réciproque est fautive : soit U uniforme sur $[0, 1]$ et $X_n(t) = \max(0, 1 - n|U - t|)$ pour $0 \leq t \leq 1$.



Alors $X_n \xrightarrow{\text{f.d.}} 0$, mais $X_n \not\xrightarrow{\text{loi}} 0$ dans C . En effet, si c'était le cas, puisque $\sup : C \rightarrow \mathbb{R}$ est continue, on aurait $1 = \sup X_n \xrightarrow{\text{loi}} 0$: absurde.

Cependant, on a le résultat suivant.

Théorème 4.9. Soient $\mu_n, \mu \in \mathcal{M}_1(C)$.

- (1) $\mu_n \xrightarrow{\text{étroite}} \mu$ si et seulement si $\mu_n \xrightarrow{\text{f.d.}} \mu$ et $(\mu_n)_{n \geq 1}$ est tendue.
- (2) Si les fidi de (μ_n) convergent et si $(\mu_n)_{n \geq 1}$ est tendue, alors elle converge

étroitement.

Démonstration. (1) Le sens direct résulte de la continuité des projections et du théorème de Prokhorov. Réciproquement, d'après le théorème de Prokhorov, il suffit de montrer que si $\mu_{\varphi(n)} \xrightarrow{\text{étroite}} \nu$, alors $\mu = \nu$. Mais alors $\mu_{\varphi(n)} \xrightarrow{\text{f.d.}} \nu$. Donc μ et ν ont les mêmes fidi, donc sont égales. (2) Par tension, il existe une extraction φ et $\mu \in \mathcal{M}_1(C)$ telles que $\mu_{\varphi(n)} \xrightarrow{\text{étroite}} \mu$. Alors $\mu_{\varphi(n)} \xrightarrow{\text{f.d.}} \mu$, et donc $\mu_n \xrightarrow{\text{f.d.}} \mu$. Donc $\mu_n \xrightarrow{\text{étroite}} \mu$ par (1). $\square$

Ainsi, dans $\mathcal{C}([0, 1], E)$, $(X_n(t))_{0 \leq t \leq 1} \xrightarrow{\text{loi}} (X(t))_{0 \leq t \leq 1}$ si et seulement si, pour tous $0 \leq t_1 < \dots < t_k \leq 1$, on a

$$(X_n(t_1), \dots, X_n(t_k)) \xrightarrow{\text{loi}} (X(t_1), \dots, X(t_k)),$$

et $(X_n)_{n \geq 1}$ est tendue.

Remarque 4.10. Étant donnée une famille $\mathcal{F}$ de fidi cohérentes, d'après le théorème d'extension de Kolmogorov, il existe $\mu \in \mathcal{M}_1(\mathbb{R}^{[0,1]})$ dont les fidi sont $\mathcal{F}$. Mais rien ne garantit qu'il existe une telle $\mu \in \mathcal{M}_1(C)$. D'après le théorème précédent, c'est le cas si les fidi $\mathcal{F}$ sont limites fidi d'une famille tendue de $\mathcal{M}_1(C)$.

Pour appliquer le théorème précédent, il s'agit maintenant d'identifier les compacts de C .

4.4 Compacité dans C

La notion d'équicontinuité est fondamentale.

Définition 4.11. Une famille $\mathcal{F} \subset \mathcal{C}(X, Y)$ est dite *équicontinue* si

$$\forall \varepsilon > 0, \quad \exists \eta > 0, \quad \forall x, y \in X, \quad \forall f \in \mathcal{F}, \quad d_X(x, y) \leq \eta \Rightarrow d_Y(f(x), f(y)) \leq \varepsilon.$$

En notant, pour $f \in \mathcal{F}$ et $\eta > 0$,

$$\omega(f, \eta) = \sup\{d_Y(f(x), f(y)) : x, y \in X, d_X(x, y) \leq \eta\},$$

le module de continuité de f , c'est équivalent à

$$\forall \varepsilon > 0, \quad \exists \eta > 0, \quad \sup_{f \in \mathcal{F}} \omega(f, \eta) \leq \varepsilon.$$

Remarque 4.12. Pour K compact :

- pour $f \in \mathcal{C}(K, Y)$, $\lim_{h \rightarrow 0} \omega(f, h) = 0$, car f est uniformément continue ;
- pour tout $h > 0$, l'application $\mathcal{C}(K, Y) \rightarrow \mathbb{R}_+$, $f \mapsto \omega(f, h)$, est continue.

Théorème 4.13 (Arzelà–Ascoli). Soit (K, d_K) compact et (E, d_E) un espace métrique complet. Soit $\mathcal{F} \subset \mathcal{C}(K, E)$. Alors $\mathcal{F}$ est relativement compact dans $\mathcal{C}(K, E)$ si et seulement si

- $\mathcal{F}$ est équicontinue ;
- il existe $D \subset K$ dénombrable dense tel que, pour tout $x \in D$, $\mathcal{F}(x) = \{f(x) : f \in \mathcal{F}\}$ est relativement compact dans E .

Démonstration. $\Leftarrow$ Soit $(f_n)_{n \geq 1}$ une suite de $\mathcal{F}$. On trouve une sous-suite qui converge uniformément. Écrivons $D = \{a_n : n \geq 1\}$. Puisque $(f_n(a_k))_{n \geq 1}$ est dans $\overline{\mathcal{F}(a_k)}$, compact dans E , il existe par procédé diagonal une extraction φ telle que $(f_{\varphi(n)}(a_k))_{n \geq 1}$ converge pour tout $k \geq 1$. Posons $g_n(x) = f_{\varphi(n)}(x)$ pour simplifier les notations.

Vérifions que, pour tout $x \in K$, $(g_n(x))_{n \geq 1}$ converge. Comme E est complet, il suffit de montrer qu'elle est de Cauchy. Pour cela, on fixe $\varepsilon > 0$ et $\eta > 0$ tels que $\sup_{f \in \mathcal{F}} \omega(f, \eta) \leq \varepsilon$. Par compacité, il existe $L \geq 1$ tel que $K = \bigcup_{i=1}^L B(a_i, \eta)$. Pour $1 \leq i \leq L$, puisque $(g_n(a_i))_{n \geq 1}$ converge, elle est de Cauchy. Il existe donc N tel que

$$p, q \geq N, \quad 1 \leq k \leq L \quad \Rightarrow \quad d_E(g_p(a_k), g_q(a_k)) \leq \varepsilon.$$

Fixons $x \in K$. Soit $1 \leq k \leq L$ tel que $d_K(x, a_k) < \eta$. Alors, pour $p, q \geq N$,

$$\begin{aligned} d_E(g_p(x), g_q(x)) &\leq d_E(g_p(x), g_p(a_k)) + d_E(g_p(a_k), g_q(a_k)) \\ &\quad + d_E(g_q(a_k), g_q(x)) \leq 3\varepsilon. \end{aligned}$$

De plus, en notant $g(x) = \lim_{n \rightarrow \infty} g_n(x)$, en passant à la limite $q \rightarrow \infty$ dans l'inégalité précédente, on obtient

$$\forall x \in K, \quad \forall p \geq N, \quad d_E(g_p(x), g(x)) \leq 3\varepsilon.$$

La convergence est donc uniforme et $g \in \mathcal{C}(K, E)$.

$\Rightarrow$ Soit $x \in K$. Montrons que $\overline{\mathcal{F}(x)}$ est compact. L'application $\pi_x : \mathcal{C}(K, E) \rightarrow E$, $f \mapsto f(x)$, est 1-lipschitzienne donc continue. Donc $\pi_x(\overline{\mathcal{F}}) = \overline{\mathcal{F}(x)}$ est compact. Or $\mathcal{F} \subset \overline{\mathcal{F}}$, donc $\mathcal{F}(x) \subset \overline{\mathcal{F}(x)}$, donc $\overline{\mathcal{F}(x)} \subset \overline{\mathcal{F}(x)}$. Donc $\overline{\mathcal{F}(x)}$, fermé dans un compact, est compact.

Soit $\varepsilon > 0$. Par précompacité de $\mathcal{F}$, il existe $f_1, \dots, f_N \in \mathcal{F}$ tels que $\mathcal{F} \subset \bigcup_{k=1}^N B(f_k, \varepsilon)$. Par uniforme continuité, on peut trouver $\eta > 0$ tel que

$$\forall x, y \in K, \quad \forall 1 \leq k \leq N, \quad d_K(x, y) \leq \eta \Rightarrow d_E(f_k(x), f_k(y)) \leq \varepsilon.$$

Alors, pour tout $f \in \mathcal{F}$ et tous $x, y \in K$ tels que $d_K(x, y) < \eta$, soit $1 \leq k \leq N$ tel que $d(f, f_k) < \varepsilon$ dans $\mathcal{C}(K, E)$, et alors

$$d_E(f(x), f(y)) \leq d_E(f(x), f_k(x)) + d_E(f_k(x), f_k(y)) + d_E(f_k(y), f(y)) \leq 3\varepsilon.$$

Donc $\mathcal{F}$ est équicontinue. □

Remarque 4.14. Si, dans E , les boules fermées sont compactes (comme dans $\mathbb{R}^k$) et si K est un compact convexe d'un espace vectoriel normé, muni de la distance induite par la norme, alors $\mathcal{F}$ est relativement compacte si et seulement si elle est équicontinue et $\mathcal{F}(x_0)$ est relativement compact pour un $x_0 \in K$.

En effet, choisissons $\eta > 0$ tel que $\sup_{f \in \mathcal{F}} \omega(f, \eta) \leq 1$. Pour tout $x \in K$, on découpe le segment $[x_0, x]$ en au plus $1 + \lceil \text{diam}(K)/\eta \rceil$ morceaux de longueur au plus η . L'inégalité triangulaire donne une borne uniforme sur $d_E(f(x), f(x_0))$, pour $f \in \mathcal{F}$ et $x \in K$. Les ensembles $\mathcal{F}(x)$ sont donc bornés, puis relativement compacts dans E , et le théorème d'Arzelà–Ascoli s'applique.

Un cas typique d'application est dans $\mathcal{C}([0, 1], \mathbb{R}^k)$, en prenant $x_0 = 0$.

4.5 Tension

Voyons maintenant une application probabiliste.

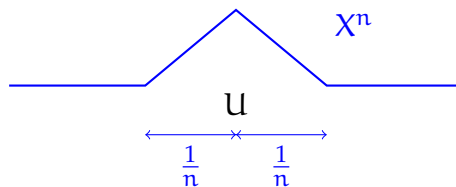
Théorème 4.15. Soit $(X^n)_{n \geq 1}$ une suite de variables aléatoires à valeurs dans $\mathcal{C}([0, 1], E)$, avec E polonais. Elle est tendue si et seulement si

- (1) pour tous $\varepsilon > 0, \eta > 0$, il existe $\delta > 0$ tel que $\limsup_{n \rightarrow \infty} \mathbb{P}(\omega(X^n, \delta) \geq \eta) \leq \varepsilon$;
- (2) pour tout $t \in [0, 1] \cap \mathbb{Q}$, $(X^n(t))_{n \geq 1}$ est tendue.

Si, dans E , les fermés bornés sont compacts, on peut remplacer (2) par (2') : $(X^n(0))_{n \geq 1}$ est tendue.

Remarque 4.16. Si $(X^n)_{n \geq 1}$ converge au sens des fidi, la condition (2) est vérifiée.

Exemple 4.17. Si U est uniforme sur $[0, 1]$ et $X^n(t) = \max(0, 1 - n|U - t|)$ pour $0 \leq t \leq 1$, la suite $(X^n)_{n \geq 1}$ n'est pas tendue.



Preuve du théorème. $\Rightarrow$ Soient $\varepsilon > 0$ et K un compact de C tel que $\mathbb{P}(X^n \notin K) \leq \varepsilon$ pour tout $n \geq 1$. Soit $\eta > 0$.

(1) Par Arzelà–Ascoli, on peut trouver δ tel que $\sup_{f \in K} \omega(f, \delta) < \eta$. On a alors $\mathbb{P}(\omega(X^n, \delta) \geq \eta) \leq \mathbb{P}(X^n \notin K) \leq \varepsilon$.

(2) Soit $t \in [0, 1]$. Il existe A compact de E tel que $f(t) \in A$ pour tout $f \in K$. Alors $\mathbb{P}(X^n(t) \notin A) \leq \mathbb{P}(X^n \notin K) \leq \varepsilon$. Donc $(X^n(t))_{n \geq 1}$ est tendue.

$\Leftarrow$ Fixons $\varepsilon > 0$. Soit $(t_i)_{i \geq 1}$ une énumération des rationnels de $[0, 1]$. On peut trouver des compacts A_i de E tels que $\mathbb{P}(X^n(t_i) \notin A_i) \leq \varepsilon/2^{i+1}$ pour tout $n \geq 1$. Pour chaque $k \geq 1$, la condition (1), appliquée avec $\eta = 2^{-k}$ et $\varepsilon/2^{k+2}$, fournit $\delta'_k > 0$ et N_k tels que

$$n \geq N_k \quad \Rightarrow \quad \mathbb{P}(\omega(X^n, \delta'_k) > 2^{-k}) \leq \frac{\varepsilon}{2^{k+1}}.$$

Pour chacun des indices $n < N_k$, la continuité des trajectoires entraîne $\omega(X^n, \delta) \rightarrow 0$ presque sûrement lorsque $\delta \downarrow 0$. En diminuant δ'_k , on obtient donc $\delta_k > 0$ tel que

$$\forall n \geq 1, \quad \mathbb{P}(\omega(X^n, \delta_k) > 2^{-k}) \leq \frac{\varepsilon}{2^{k+1}}.$$

En posant

$$K_\varepsilon = \{f \in C : \forall i \geq 1, f(t_i) \in A_i; \forall k \geq 1, \omega(f, \delta_k) \leq 2^{-k}\},$$

on voit que $\mathbb{P}(X^n \notin K_\varepsilon) \leq \varepsilon$, que K_ε est fermé ($f \mapsto \omega(f, \delta_k)$ est continue) et que K_ε est compact par Arzelà–Ascoli.

Si les fermés bornés de E sont compacts et si l'on suppose seulement (2'), choisissons un compact $A \subset E$ tel que $\sup_n \mathbb{P}(X^n(0) \notin A) \leq \varepsilon/2$. Avec les mêmes δ_k , l'ensemble

$$K'_\varepsilon = \{f \in C : f(0) \in A; \forall k \geq 1, \omega(f, \delta_k) \leq 2^{-k}\}$$

est compact par la remarque suivant Arzelà–Ascoli, et $\sup_n \mathbb{P}(X^n \notin K'_\varepsilon) \leq \varepsilon$. Cela prouve également la dernière assertion. $\square$

Remarque 4.18. La condition (1) est équivalente à : pour toute suite $\delta_k \rightarrow 0$ et toute suite (n_k) , $\omega(X^{n_k}, \delta_k) \xrightarrow{\mathbb{P}} 0$ (exercice).

Corollaire 4.19. Soient $(X^n)_{n \geq 1}$ des variables aléatoires dans C telles que

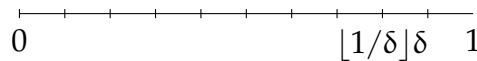
(1) pour tous $\varepsilon > 0, \eta > 0$, il existe $0 < \delta < 1$ tel que, pour tout $t \in [0, 1]$,

$$\limsup_{n \rightarrow \infty} \frac{1}{\delta} \mathbb{P} \left(\sup_{t \leq s \leq (t+\delta) \wedge 1} d(X_s^n, X_t^n) > \eta \right) \leq \varepsilon;$$

(2) pour tout $t \in [0, 1] \cap \mathbb{Q}$, $(X_t^n)_{n \geq 1}$ est tendue.

Alors $(X^n)_{n \geq 1}$ est tendue dans C .

Démonstration. On vérifie la condition (1) du théorème. Pour cela, on découpe $[0, 1]$ en intervalles de longueur δ .



Soit $f \in C$. Vérifions que

$$\{\omega(f, \delta) > 3\eta\} \subset \bigcup_{i=0}^{\lfloor 1/\delta \rfloor} \left\{ \sup_{i\delta \leq s \leq ((i+1)\delta) \wedge 1} d(f(s), f(i\delta)) > \eta \right\}.$$

On suppose que, pour tout $0 \leq i \leq \lfloor 1/\delta \rfloor$, $\sup_{i\delta \leq s \leq ((i+1)\delta) \wedge 1} d(f(s), f(i\delta)) \leq \eta$. Si $|s - t| \leq \delta$, deux cas se présentent.

— Cas 1 : s, t sont dans le même intervalle $[i\delta, ((i+1)\delta) \wedge 1]$. Alors

$$d(f(s), f(t)) \leq d(f(t), f(i\delta)) + d(f(s), f(i\delta)) \leq 2\eta.$$

— Cas 2 : s, t sont dans deux intervalles consécutifs, avec $s \in [i\delta, (i+1)\delta]$ et $t \in [(i+1)\delta, ((i+2)\delta) \wedge 1]$. Alors

$$d(f(s), f(t)) \leq d(f(s), f(i\delta)) + d(f((i+1)\delta), f(i\delta)) + d(f(t), f((i+1)\delta)) \leq 3\eta.$$

Ainsi,

$$\mathbb{P}(\omega(X^n, \delta) > 3\eta) \leq (\lfloor 1/\delta \rfloor + 1) \times \max_{t \in \{0, \delta, 2\delta, \dots, \lfloor 1/\delta \rfloor \delta\}} \mathbb{P} \left(\sup_{t \leq s \leq (t+\delta) \wedge 1} d(X_s^n, X_t^n) > \eta \right).$$

La limite supérieure est $\leq 2\varepsilon$, car $\lfloor 1/\delta \rfloor + 1 \leq 1/\delta + 1 < 2/\delta$. □

4.6 Critère de tension de Kolmogorov

Le problème des caractérisations précédentes de la tension est qu'elles font intervenir des variables aléatoires malaisées à étudier. Des critères de tension efficaces peuvent faire intervenir la loi de X^n de la façon la moins compliquée possible, par exemple à travers des marginales de dimension finie. Le critère de Kolmogorov ne fait intervenir que les marginales de dimension 2 (!).

On introduit d'abord quelques notations.

Notation.

- Soit $D_n = \{i/2^n : 0 \leq i \leq 2^n\}$ l'ensemble des dyadiques de niveau n dans $[0, 1]$, et $D = \bigcup_{n \geq 0} D_n$.
- Pour $X = (X_t)_{0 \leq t \leq 1} \in E^{[0,1]}$ et $\gamma > 0$, on note

$$M_\gamma(X) = \sup_{\substack{s, t \in D \\ s \neq t}} \frac{d(X_s, X_t)}{|s - t|^\gamma}.$$

Ainsi, si X est γ -höldérien, $M_\gamma(X) < \infty$.

Théorème 4.20 (Critère de tension de Kolmogorov). (1) Soit $X = (X_t)_{0 \leq t \leq 1}$ une variable aléatoire à valeurs dans $E^{[0,1]}$. On suppose qu'il existe $\alpha, \beta, c > 0$ tels que

$$\forall s, t \in [0, 1], \quad \mathbb{E}[d(X_s, X_t)^\alpha] \leq c|s - t|^{1+\beta}. \quad (*)$$

Alors, pour tout $0 < \gamma < \beta/\alpha$, $\mathbb{E}[M_\gamma(X)^\alpha] \leq C(\alpha, \beta, \gamma, c)$.

(2) Soient $(X^n)_{n \geq 1}$ des variables aléatoires à valeurs dans C . On suppose que

- a) pour tout $t \in [0, 1]$, $(X_t^n)_{n \geq 1}$ est tendue;
- b) il existe $\alpha, \beta, c > 0$ tels que, pour tout $n \geq 1$ et tous $s, t \in [0, 1]$, $\mathbb{E}[d(X_s^n, X_t^n)^\alpha] \leq c|s - t|^{1+\beta}$.

Alors $(X^n)_{n \geq 1}$ est tendue, et toute limite de sous-suite est presque sûrement γ -höldérienne pour tout $0 < \gamma < \beta/\alpha$.

La première partie de la preuve est déterministe. L'idée est de contrôler $d(X_s, X_t)$ « uniformément » en $s, t \in D$, par l'inégalité triangulaire le long d'un chemin de dyadiques « consécutifs ».

Notons

$$\Delta_k = \max_{0 \leq i < 2^k} d(X_{i2^{-k}}, X_{(i+1)2^{-k}}), \quad K_\gamma = \sup_{k \geq 0} 2^{k\gamma} \Delta_k.$$

Lemme 4.21. Pour tout $\gamma > 0$, $M_\gamma(X) \leq \frac{2^{\gamma+1}}{1-2^{-\gamma}} K_\gamma$.

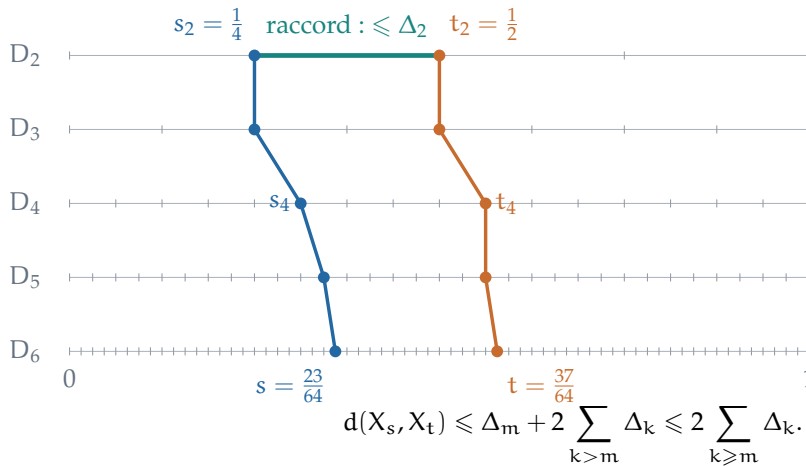
Démonstration. On commence par regrouper les couples de dyadiques suivant leur distance :

$$M_\gamma(X) \leq \sup_{m \geq 0} 2^{(m+1)\gamma} \sup_{\substack{s, t \in D \\ |s-t| \leq 2^{-m}}} d(X_s, X_t). \quad (**)$$

Le chaînage dyadique donne, pour tout $m \geq 0$,

$$\sup_{\substack{s, t \in D \\ |s-t| \leq 2^{-m}}} d(X_s, X_t) \leq 2 \sum_{k=m}^{\infty} \Delta_k. \quad (***)$$

En effet, pour $s, t \in D$, $s \leq t$ et $t - s \leq 2^{-m}$, posons $s_k = 2^{-k} \lfloor 2^k s \rfloor$ et $t_k = 2^{-k} \lfloor 2^k t \rfloor$. Les points s_m, t_m sont égaux ou voisins dans D_m . Deux approximations successives de s sont égales ou voisines dans D_{k+1} , et $s_k = s$ pour k assez grand. Ainsi $d(X_{s_m}, X_s) \leq \sum_{k>m} \Delta_k$, et de même pour t . L'inégalité triangulaire prouve (**).



$$m = 2, \quad t - s = \frac{14}{64} < 2^{-2}$$

$$s_k = 2^{-k} \lfloor 2^k s \rfloor$$

$$t_k = 2^{-k} \lfloor 2^k t \rfloor$$

Deux approximations successives sont égales ou voisines dans D_{k+1} .

Leur accroissement dans E est donc au plus Δ_{k+1} .

Les lignes colorées relient les approximations dyadiques des temps s et t . Les deux chaînes rejoignent des points voisins de D_2 et s'arrêtent ici à D_6 . Un segment vertical indique une approximation inchangée; un segment oblique, un déplacement d'une maille de la grille suivante.

Si $K_\gamma = \infty$, le résultat est immédiat. Sinon,

$$2 \sum_{k=m}^{\infty} \Delta_k \leq 2K_\gamma \sum_{k=m}^{\infty} 2^{-k\gamma} = \frac{2K_\gamma}{1-2^{-\gamma}} 2^{-m\gamma}.$$

En reportant cette borne dans (**), on obtient le résultat. $\square$

Preuve du théorème. (1) L'hypothèse (*) donne

$$\mathbb{E}[\Delta_k^\alpha] \leq \sum_{i=0}^{2^k-1} \mathbb{E}[d(X_{i2^{-k}}, X_{(i+1)2^{-k}})^\alpha] \leq 2^k c 2^{-k(1+\beta)} = c 2^{-k\beta}.$$

Pour tout $\alpha > 0$, on en déduit

$$\mathbb{E}[K_\gamma^\alpha] \leq \sum_{k \geq 0} 2^{k\alpha\gamma} \mathbb{E}[\Delta_k^\alpha] \leq c \sum_{k \geq 0} 2^{-k(\beta-\alpha\gamma)} < \infty,$$

car $\gamma < \beta/\alpha$. Le lemme précédent entraîne alors $\mathbb{E}[M_\gamma(X)^\alpha] \leq C(\alpha, \beta, \gamma, c)$.

(2) Pour la tension, il suffit de vérifier que, pour tous $\varepsilon > 0, \eta > 0$, il existe $\delta > 0$ tel que $\mathbb{P}(\omega(X^n, \delta) \geq \eta) \leq \varepsilon$.

Par définition de $M_\gamma(X^n)$, pour tous $s, t \in D$, on a $d(X_s^n, X_t^n) \leq M_\gamma(X^n)|s - t|^\gamma$. Par continuité, c'est aussi vrai pour tous $s, t \in [0, 1]$. Ainsi, si $|s - t| \leq \delta$, on a $d(X_s^n, X_t^n) \leq M_\gamma(X^n)\delta^\gamma$, de sorte que $\omega(X^n, \delta) \leq M_\gamma(X^n)\delta^\gamma$. Par l'inégalité de Markov,

$$\mathbb{P}(\omega(X^n, \delta) \geq \eta) \leq \eta^{-\alpha} \delta^{\gamma\alpha} \mathbb{E}[M_\gamma(X^n)^\alpha].$$

On en déduit le résultat par (1).

Enfin, supposons $X^{\varphi(n)} \xrightarrow{\text{loi}} X$ dans C . Pour $s, t \in [0, 1]$ et $R > 0$, l'application $f \mapsto \min\{d(f(s), f(t))^\alpha, R\}$ est continue bornée. Par convergence en loi,

$$\begin{aligned} \mathbb{E}[\min\{d(X_s, X_t)^\alpha, R\}] &= \lim_{n \rightarrow \infty} \mathbb{E}[\min\{d(X_s^{\varphi(n)}, X_t^{\varphi(n)})^\alpha, R\}] \\ &\leq c|s - t|^{1+\beta}. \end{aligned}$$

En faisant $R \uparrow \infty$, la convergence monotone montre que X vérifie (*). Le point (1) et la continuité de X donnent donc la régularité γ -höldérienne presque sûre pour chaque $0 < \gamma < \beta/\alpha$. En prenant une suite d'exposants croissant vers β/α et une intersection dénombrable, cette régularité vaut simultanément pour tous ces exposants. $\square$

Remarque 4.22. Le point (1) permet de démontrer aussi le critère de continuité de Kolmogorov : si $(X_t)_{0 \leq t \leq 1} \in E^{[0,1]}$ vérifie (*), alors, pour tout $0 < \gamma < \beta/\alpha$, il existe un processus $(\tilde{X}_t)_{0 \leq t \leq 1}$ γ -höldérien tel que, pour tout $t \in [0, 1]$, $\mathbb{P}(X_t = \tilde{X}_t) = 1$.

En effet, puisque $\mathbb{E}[M_\gamma(X)^\alpha] < \infty$, il existe A de probabilité 1 tel que, pour tout $\omega \in A$, X est γ -höldérien sur D . Pour $\omega \in A$, on peut donc prolonger X en une fonction γ -höldérienne, notée $\tilde{X}$. Pour $\omega \notin A$, on pose $\tilde{X} = 0$ (ou un point fixé de E).

Vérifions que, pour $t \in [0, 1]$, $\mathbb{P}(X_t = \tilde{X}_t) = 1$. Par (*) et l'inégalité de Markov, si

$s_n \rightarrow t$ et $s_n \in D$, $X_{s_n} \xrightarrow{\mathbb{P}} X_t$. D'autre part, par construction, $X_{s_n} \xrightarrow{\text{p.s.}} \tilde{X}_t$. Donc $X_{s_n} \xrightarrow{\mathbb{P}} \tilde{X}_t$. Donc $\mathbb{P}(X_t = \tilde{X}_t) = 1$.

Application. Soient $(X_i)_{i \geq 1}$ des variables aléatoires réelles i.i.d. telles que $\mathbb{E}[X_1] = 0$ et $\mathbb{E}[X_1^4] < \infty$. On pose $S_0 = 0$, $S_n = X_1 + \dots + X_n$ pour $n \geq 1$ et

$$S_t = (1 - \{t\})S_{[t]} + \{t\}S_{[t]+1}, \quad t \geq 0.$$

Enfin, on définit $S_t^{(n)} = S_{nt}/\sqrt{n}$ pour $0 \leq t \leq 1$, de sorte que $S^{(n)} \in C = \mathcal{C}([0, 1], \mathbb{R})$. Alors $(S^{(n)})_{n \geq 1}$ est tendue dans C .

On verra plus tard que ce résultat reste vrai sous l'hypothèse plus faible $\mathbb{E}[X_1^2] < \infty$, mais la preuve est plus délicate.

Démonstration. Posons $v = \mathbb{E}[X_1^2]$, $m_4 = \mathbb{E}[X_1^4]$ et $C = 1 + m_4 + 3v^2$. Pour $0 \leq s < t \leq 1$ et $1 \leq i \leq n$, notons

$$a_i = \text{Leb}([ns, nt] \cap [i-1, i]).$$

Par définition de l'interpolation, $S_{nt} - S_{ns} = \sum_{i=1}^n a_i X_i$, avec $0 \leq a_i \leq 1$ et $\sum_i a_i = n(t-s)$. L'indépendance et le centrage donnent

$$\begin{aligned} \mathbb{E}[|S_{nt} - S_{ns}|^4] &= m_4 \sum_i a_i^4 + 6v^2 \sum_{i < j} a_i^2 a_j^2 \\ &\leq C \left(\sum_i a_i^2 \right)^2 \leq Cn^2(t-s)^2. \end{aligned}$$

En particulier, en prenant $s = 0$ et $t = 1$,

$$\mathbb{E}[S_n^4] \leq Cn^2. \quad (\dagger)$$

Le même calcul donne, sans restriction sur les extrémités s, t ,

$$\mathbb{E}[|S_{nt} - S_{ns}|^4] \leq Cn^2|t-s|^2. \quad (\dagger\dagger)$$

Après division par n^2 , on obtient $\mathbb{E}[|S_t^{(n)} - S_s^{(n)}|^4] \leq C|t-s|^2$. Enfin, pour $s = 0$, l'indépendance et le centrage donnent aussi

$$\mathbb{E}[|S_t^{(n)}|^2] = \frac{v}{n} \sum_i a_i^2 \leq vt.$$

Les marginales sont donc tendues par l'inégalité de Markov. Le critère de Kolmogorov s'applique avec $\alpha = 4$ et $\beta = 1$, et prouve la tension de $(S^{(n)})_{n \geq 1}$. $\square$

4.7 Fonctions continues sur $\mathbb{R}_+$

En pratique, les processus aléatoires sont souvent indexés par $\mathbb{R}_+$ plutôt que par $[0, 1]$. Comme, sur l'espace $\mathcal{C}(\mathbb{R}_+, E)$, la distance uniforme $D_\infty(f, g) = \sup_{t \geq 0} d(f(t), g(t))$ peut

valoir $+\infty$, il faut faire attention. On placera toujours sur $\mathcal{C}(\mathbb{R}_+, E)$ une autre topologie, celle de la convergence uniforme sur les compacts, induite par la distance

$$D(f, g) = \sum_{n=1}^{\infty} \frac{\left(\sup_{0 \leq t \leq n} d(f(t), g(t)) \right)}{2^n} \wedge 1.$$

Remarque 4.23 (Facultative). En termes topologiques, c'est la topologie compacte-ouverte sur $\mathcal{C}(\mathbb{R}_+, E)$, dont une sous-base est donnée par les ensembles

$$\{f \in \mathcal{C}(\mathbb{R}_+, E) : f(K) \subset O\},$$

avec K compact de $\mathbb{R}_+$ et O ouvert de E .

On vérifie aisément que :

- $f_n \rightarrow f$ dans $\mathcal{C}(\mathbb{R}_+, E)$ si et seulement si $f_n \rightarrow f$ uniformément sur les compacts ;
- $\mathcal{C}(\mathbb{R}_+, E)$ est polonais ;
- les marginales finidimensionnelles caractérisent les mesures de $\mathcal{M}_1(\mathcal{C}(\mathbb{R}_+, E))$, car la tribu borélienne sur $\mathcal{C}(\mathbb{R}_+, E)$ coïncide encore avec la trace sur $\mathcal{C}(\mathbb{R}_+, E)$ de la tribu produit sur $E^{\mathbb{R}_+}$.

L'avantage important de cet espace est que la relative compacité est une propriété locale, cf. la proposition suivante (preuve laissée en exercice).

Proposition 4.24. — Une partie $\mathcal{F}$ de $\mathcal{C}(\mathbb{R}_+, E)$ est relativement compacte si et seulement si, pour tout $N \geq 1$, les parties $\mathcal{F}_N = \{f|_{[0, N]} : f \in \mathcal{F}\}$ sont relativement compactes dans $\mathcal{C}([0, N], E)$, où $f|_{[0, N]}$ est la restriction de f à $[0, N]$.

— Une suite $(X^n)_{n \geq 1}$ de variables aléatoires de $\mathcal{C}(\mathbb{R}_+, E)$ est tendue si et seulement si, pour tout $N \geq 1$, la suite des restrictions $(X^n|_{[0, N]})_{n \geq 1}$ est tendue dans $\mathcal{C}([0, N], E)$.

Ainsi, la convergence en loi dans $\mathcal{C}(\mathbb{R}_+, E)$ se ramène par restriction à la convergence en loi dans les espaces $\mathcal{C}(I, E)$, avec I un intervalle compact, et les résultats obtenus précédemment peuvent être mis en œuvre.

Interlude : couplages

Un *couplage* de deux variables aléatoires X et Y (pas forcément définies sur le même espace de probabilité) est une variable aléatoire (X', Y') telle que $X' \stackrel{\text{loi}}{=} X$ et $Y' \stackrel{\text{loi}}{=} Y$.

Exemple C.1. Soient $n \geq 1$ un entier et $0 \leq p \leq q \leq 1$. Alors, pour tout entier $0 \leq k \leq n$,

$$\mathbb{P}(\text{Bin}(n, p) \geq k) \leq \mathbb{P}(\text{Bin}(n, q) \geq k).$$

Démonstration. On construit un couplage dans lequel $\text{Bin}(n, p) \leq \text{Bin}(n, q)$ en considérant $(U_i)_{1 \leq i \leq n}$ i.i.d. uniformes sur $[0, 1]$ et en posant

$$X' = \sum_{i=1}^n \mathbb{1}_{\{U_i \leq p\}}, \quad Y' = \sum_{i=1}^n \mathbb{1}_{\{U_i \leq q\}}.$$

□

Plan. Théorème de représentation de Skorokhod ; distance en variation totale ; distance de Lévy–Prokhorov ; monotonie et ordre stochastique.

Références pour un bestiaire de distances sur $\mathcal{M}_1(E)$.

- Alison L. Gibbs, Francis Edward Su, *On choosing and bounding probability metrics*.
- Svante Janson, *Probability distances*.

C.1 Théorème de représentation de Skorokhod

Théorème C.2 (Représentation de Skorokhod). Soit (E, d) séparable, et soient $(X_i)_{i \geq 1}$ et X des variables aléatoires à valeurs dans E telles que $X_n \xrightarrow{\text{loi}} X$. On peut définir des variables aléatoires $(X'_i)_{i \geq 1}$ et X' sur un même espace de probabilité telles que, pour tout $n \geq 1$,

$$X_n \stackrel{\text{loi}}{=} X'_n, \quad X \stackrel{\text{loi}}{=} X', \quad X'_n \xrightarrow{\text{p.s.}} X'.$$

Voir la feuille d'exercices pour la preuve dans le cas $E = \mathbb{R}$. En pratique, ce théorème est utilisé pour se ramener à un « raisonnement déterministe » à partir d'une convergence en loi (on fait « comme si » $X_n \rightarrow X$ de manière déterministe).

Attention. Toute information sur les lois jointes des X_i est perdue chez les X'_i .

Proposition C.3 (Application). Soient $X, X_1, X_2, \dots$ des variables aléatoires réelles telles que $X_n \xrightarrow{\text{loi}} X$. On suppose que

$$\lim_{a \rightarrow \infty} \sup_{n \geq 1} \mathbb{E}[|X_n| \mathbb{1}_{\{|X_n| \geq a\}}] = 0. \quad (*)$$

Alors $\mathbb{E}[X_n] \rightarrow \mathbb{E}[X]$.

Démonstration. D'après le théorème de représentation de Skorokhod, on peut supposer que $X_n \xrightarrow{\text{p.s.}} X$ et que $(*)$ est satisfaite (car c'est une propriété des lois individuelles des X_n). Alors $X_n \xrightarrow{\mathbb{P}} X$ et la conclusion découle du fait que

$$X_n \xrightarrow{\mathbb{P}} X, \quad (X_n)_{n \geq 1} \text{ uniformément intégrable} \implies X_n \xrightarrow{L^1} X.$$

□

Pour la preuve, on admet l'existence de variables aléatoires indépendantes ayant des lois fixées et on utilise le lemme suivant.

Lemme C.4. Soit (E, d) séparable et $\mu \in \mathcal{M}_1(E)$. Pour tout $\varepsilon > 0$, il existe une partition $E = \bigsqcup_{i \geq 1} A_i$ en ensembles mesurables telle que

$$\forall i \geq 1, \quad \text{diam}(A_i) \leq \varepsilon, \quad \mu(\partial A_i) = 0.$$

Démonstration. Soit $(x_i)_{i \geq 1}$ une suite dénombrable dense. Pour tout $i \geq 1$,

$$\{r \in [\varepsilon/4, \varepsilon/2] : \mu(\partial B(x_i, r)) > 0\}$$

est dénombrable. On choisit alors $\varepsilon/4 \leq r_i \leq \varepsilon/2$ tel que $\mu(\partial B(x_i, r_i)) = 0$, et on note $B_i = B(x_i, r_i)$. On définit par récurrence

$$A_1 = B_1, \quad A_i = B_1^c \cap \dots \cap B_{i-1}^c \cap B_i,$$

et on vérifie que $\mu(\partial A_i) = 0$. En effet, $\partial A = \partial(A^c)$ et $\partial(A \cup B) \subset \partial A \cup \partial B$, donc

$$\mu(\partial A) = 0 \implies \mu(\partial A^c) = 0, \quad \mu(\partial A) = \mu(\partial B) = 0 \implies \mu(\partial(A \cup B)) = 0.$$

□

Preuve du théorème de représentation de Skorokhod. Notons μ_n la loi de X_n et μ la loi de X .

Étape 1. Soit $p \geq 1$ et $(A_i^p)_{i \geq 1}$ une partition donnée par le lemme pour $\varepsilon_p = 2^{-p}$ avec μ . Soit k_p tel que

$$\mu \left(\bigcup_{i=k_p+1}^{\infty} A_i^p \right) < \varepsilon_p.$$

On pose $A_0^p = \bigcup_{i=k_p+1}^{\infty} A_i^p$, de sorte que $\mu(A_0^p) < \varepsilon_p$ et $\mu(\partial A_0^p) = 0$. En particulier, $A_0^p, A_1^p, \dots, A_{k_p}^p$ est une partition de E . Par convergence en loi, on peut trouver a_p tel que

$$n \geq a_p \implies \mu_n(A_i^p) \geq (1 - \varepsilon_p)\mu(A_i^p) \quad \text{pour } 0 \leq i \leq k_p. \quad (*)$$

En particulier, $\mu_n(A_i^p) = 0 \implies \mu(A_i^p) = 0$. Sans perte de généralité, on suppose $a_1 < a_2 < \dots$ et $n \geq a_1$. Soit $p_n \geq 1$ l'unique entier tel que $a_{p_n} \leq n < a_{p_n+1}$. Lorsque $n \rightarrow \infty$, on a $p_n \rightarrow \infty$.

Étape 2. L'idée est essentiellement de définir $X'_n = X_n \mid X_n \in A_i^{p_n}$ lorsque $X \in A_i^{p_n}$. Il faut cependant ajuster à la marge pour que $X'_n \stackrel{\text{loi}}{=} X_n$. On va définir X'_n comme suit :

$$X'_n = \begin{cases} Y_{ni}, & U \leq 1 - \varepsilon_{p_n}, X' \in A_i^{p_n}, \mu_n(A_i^{p_n}) > 0, \\ Z_n, & U > 1 - \varepsilon_{p_n}, \end{cases}$$

avec $(Y_{ni}), (Z_n)_{n \geq 1}$, U et X' des variables aléatoires indépendantes définies sur le même espace de probabilité telles que

- U est uniforme sur $[0, 1]$;
- $X \stackrel{\text{loi}}{=} X'$;
- $\mathcal{L}(Y_{ni}) = \mu_n(\cdot \mid A_i^{p_n})$;
- la loi de Z_n est définie par

$$\mathbb{P}(Z_n \in A) = \frac{1}{\varepsilon_{p_n}} \sum_{\substack{0 \leq i \leq k_{p_n} \\ \mu_n(A_i^{p_n}) > 0}} \mu_n(A \mid A_i^{p_n}) (\mu_n(A_i^{p_n}) - (1 - \varepsilon_{p_n}) \mu(A_i^{p_n})).$$

Cette quantité définit bien une mesure positive par $(\star)$, et de masse totale 1, car

$$\mu_n(A_i^{p_n}) = 0 \implies \mu(A_i^{p_n}) = 0.$$

On vérifie que $X'_n \stackrel{\text{loi}}{=} X_n$:

$$\begin{aligned} \mathbb{P}(X'_n \in A) &= \sum_{\substack{0 \leq i \leq k_{p_n} \\ \mu_n(A_i^{p_n}) > 0}} \mathbb{P}(U \leq 1 - \varepsilon_{p_n}, X' \in A_i^{p_n}, Y_{ni} \in A) \\ &\quad + \mathbb{P}(U > 1 - \varepsilon_{p_n}, Z_n \in A) \\ &= (1 - \varepsilon_{p_n}) \sum_{\substack{0 \leq i \leq k_{p_n} \\ \mu_n(A_i^{p_n}) > 0}} \underbrace{\mathbb{P}(X' \in A_i^{p_n})}_{\mu(A_i^{p_n})} \mu_n(A \mid A_i^{p_n}) + \varepsilon_{p_n} \mathbb{P}(Z_n \in A) \\ &= \mu_n(A), \end{aligned}$$

car

$$\mathbb{P}(Z_n \in A) = \frac{1}{\varepsilon_{p_n}} \left(\mu_n(A) - (1 - \varepsilon_{p_n}) \sum_{\substack{0 \leq i \leq k_{p_n} \\ \mu_n(A_i^{p_n}) > 0}} \mu(A_i^{p_n}) \mu_n(A \mid A_i^{p_n}) \right).$$

Étape 3. On vérifie que $X'_n \xrightarrow{\text{p.s.}} X'$. Soit

$$E_p = \{X' \notin A_0^p, U \leq 1 - \varepsilon_p\}.$$

On a

$$\mathbb{P}(E_p^c) \leq \mathbb{P}(X' \in A_0^p) + \mathbb{P}(U > 1 - \varepsilon_p) \leq \frac{1}{2^{p-1}}.$$

Donc $\sum_{p \geq 1} \mathbb{P}(E_p^c) < \infty$. Par Borel–Cantelli, presque sûrement, à partir d’un certain rang, E_p est réalisé. Donc, presque sûrement, pour n assez grand, E_{p_n} est réalisé. Par définition de X'_n , cela implique alors qu’il existe $1 \leq i \leq k_{p_n}$ tel que $X', X'_n \in A_i^{p_n}$, et donc $d(X', X'_n) \leq \varepsilon_{p_n}$. Ainsi $X'_n \xrightarrow{\text{p.s.}} X'$. $\square$

Remarque C.5. Lorsque E est polonais, il est possible de montrer que l’espace de probabilité commun peut être pris comme $[0, 1]$ muni de la mesure de Lebesgue.

C.2 Distance en variation totale

Soit $(S, \mathcal{B})$ un espace mesurable. Pour les énoncés faisant intervenir $\mathbb{P}(X = Y)$ ou $\mathbb{P}(X \neq Y)$, on suppose de plus que la diagonale

$$\Delta_S = \{(x, x) : x \in S\}$$

appartient à $\mathcal{B} \otimes \mathcal{B}$. C’est notamment le cas lorsque S est un espace métrique séparable muni de sa tribu borélienne.

Définition C.6. Pour $\mu, \nu \in \mathcal{M}_1(S)$, on pose

$$d_{VT}(\mu, \nu) = \sup_{A \in \mathcal{B}} |\mu(A) - \nu(A)|.$$

Pour des variables aléatoires X, Y à valeurs dans S , on pose

$$d_{VT}(X, Y) = \sup_{A \in \mathcal{B}} |\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)|.$$

On montre les résultats suivants :

- d_{VT} est une distance sur $\mathcal{M}_1(S)$ qui le rend complet.
- Si μ et ν sont absolument continues par rapport à une mesure σ -finie ρ sur $(S, \mathcal{B})$, alors

$$d_{VT}(\mu, \nu) = \frac{1}{2} \int_S \left| \frac{d\mu}{d\rho} - \frac{d\nu}{d\rho} \right| d\rho.$$

Ceci permet de montrer que

$$d_{VT}(\mu, \nu) = \frac{1}{2} \sup \left\{ \left| \int f d\mu - \int f d\nu \right| : f : S \rightarrow \mathbb{R} \text{ mesurable, } \|f\|_\infty \leq 1 \right\},$$

et que, si X et Y ont respectivement pour lois μ et ν ,

$$d_{VT}(\mu, \nu) = \inf \{ \mathbb{P}(X' \neq Y') : (X', Y') \text{ couplage de } X, Y \},$$

avec l’infimum qui est atteint.

- Lorsque S est un espace métrique muni de sa tribu borélienne, $d_{LP} \leq d_{VT}$.
- Lorsque S est dénombrable, muni de la distance discrète, et $\mathcal{B} = \mathcal{P}(S)$, la convergence pour d_{VT} est équivalente à la convergence en loi (**attention** : faux en général), et

$$d_{VT}(\mu, \nu) = \frac{1}{2} \sum_{x \in S} |\mu(\{x\}) - \nu(\{x\})|.$$

Remarque C.7. Une p -métrique sur S est une fonction δ à valeurs dans $[0, \infty]$, définie sur l'ensemble des lois jointes de variables aléatoires (X, Y) à valeurs dans S , telle que, pour des variables aléatoires X, Y, Z définies sur le même espace :

- (i) si $X = Y$ presque sûrement, alors $\delta(X, Y) = 0$;
- (ii) $\delta(X, Y) = \delta(Y, X)$;
- (iii) $\delta(X, Y) \leq \delta(X, Z) + \delta(Z, Y)$.

La quantité

$$\underline{\delta}(X, Y) = \inf\{\delta(X', Y') : X' \stackrel{\text{loi}}{=} X, Y' \stackrel{\text{loi}}{=} Y\}$$

est appelée *métrique minimale* associée à δ . Ainsi, d_{VT} est la métrique minimale associée à $\delta(X, Y) = \mathbb{P}(X \neq Y)$.

Remarque C.8. Pour un ensemble non vide $\mathcal{F}$ de fonctions mesurables bornées $S \rightarrow \mathbb{R}$, on pose, pour X, Y variables aléatoires à valeurs dans S ,

$$d_{\mathcal{F}}(X, Y) = \sup\{|\mathbb{E}[f(X)] - \mathbb{E}[f(Y)]| : f \in \mathcal{F}\}.$$

Cette quantité est une *pseudo-distance duale*, éventuellement infinie, qui ne dépend que des lois de X et Y . Elle définit une distance entre les lois considérées lorsqu'elle est finie et que $\mathcal{F}$ les sépare. Ainsi, d_{VT} est la distance duale définie par l'une des trois familles

$$\begin{aligned}\mathcal{F} &= \{f : S \rightarrow \mathbb{R} \text{ mesurable, } \|f\|_{\infty} \leq \tfrac{1}{2}\}, \\ \mathcal{F} &= \{f : S \rightarrow \mathbb{R} \text{ mesurable, } 0 \leq f \leq 1\}, \\ \mathcal{F} &= \{\mathbb{1}_A : A \in \mathcal{B}\}.\end{aligned}$$

Référence. Notes de cours *Probability Theory : The Coupling Method* de Frank den Hollander (section 2) :

<http://websites.math.leidenuniv.nl/probability/lecturenotes/CouplingLectures.pdf>.

C.3 Distance de Lévy–Prokhorov

Soit (E, d) métrique séparable. On rappelle que, pour $\mu, \nu \in \mathcal{M}_1(E)$,

$$d_{LP}(\mu, \nu) = \inf\{\varepsilon > 0 : \forall A \in \mathcal{B}(E), \mu(A) \leq \nu(A^{\varepsilon}) + \varepsilon\},$$

où $A^{\varepsilon} = \{x \in E : d(x, A) < \varepsilon\}$.

Définition C.9. Pour des variables aléatoires X, Y définies sur le même espace de probabilité et à valeurs dans E , on pose

$$\alpha(X, Y) = \inf\{\varepsilon > 0 : \mathbb{P}(d(X, Y) \geq \varepsilon) < \varepsilon\}.$$

C'est la métrique de Ky–Fan, et on peut vérifier qu'elle métrise la convergence en probabilité.

Théorème C.10 (Strassen). Soit (E, d) séparable, et soient $\mu, \nu \in \mathcal{M}_1(E)$. Alors

$$d_{LP}(\mu, \nu) = \inf\{\alpha(X, Y) : (X, Y) \text{ couplage avec } \mathcal{L}(X) = \mu, \mathcal{L}(Y) = \nu\}.$$

Si μ et ν sont tendues (toujours le cas si E est complet), alors l'infimum est atteint.

Preuve d'une inégalité. Soit (X, Y) un tel couplage. On montre que $d_{LP}(\mu, \nu) \leq \alpha(X, Y)$. Soit $\varepsilon > \alpha(X, Y)$, de sorte que $\mathbb{P}(d(X, Y) \geq \varepsilon) < \varepsilon$. Alors

$$\begin{aligned} \mu(A) = \mathbb{P}(X \in A) &\leq \mathbb{P}(X \in A, d(X, Y) < \varepsilon) + \mathbb{P}(X \in A, d(X, Y) \geq \varepsilon) \\ &\leq \mathbb{P}(Y \in A^\varepsilon) + \varepsilon = \nu(A^\varepsilon) + \varepsilon, \end{aligned}$$

d'où $d_{LP}(\mu, \nu) \leq \varepsilon$. □

L'autre inégalité est délicate et peut se montrer en utilisant le lemme des mariages ; voir Billingsley, *Convergence of Probability Measures*, deuxième édition, théorème 6.9, ou Dudley, *Real Analysis and Probability*, corollaire 11.6.4.

Remarque C.11. Ainsi, d_{LP} est la métrique minimale associée à α .

On finit cette partie par un résultat utile de couplage « à la Skorokhod ».

Théorème C.12 (Dudley). Soit (E, d) séparable, et soient X_n, Y_n des variables aléatoires à valeurs dans E telles que $d_{LP}(X_n, Y_n) \rightarrow 0$. Alors il existe un espace de probabilité et des variables aléatoires $(X'_n, Y'_n)_{n \geq 1}$ définies dessus telles que

$$\forall n \geq 1, \quad X'_n \stackrel{\text{loi}}{=} X_n, \quad Y'_n \stackrel{\text{loi}}{=} Y_n, \quad d(X'_n, Y'_n) \rightarrow 0 \quad \text{p.s.}$$

Voir Dudley, *Real Analysis and Probability*, théorème 11.7.1.

C.4 Monotonie, ordre stochastique

Ici, (E, d) est un espace polonais.

Définition C.13. Une relation $\preceq$ est un ordre partiel si, pour tous $x, y, z \in E$:

- (i) $x \preceq x$;
- (ii) $x \preceq y, y \preceq z \implies x \preceq z$;
- (iii) $x \preceq y, y \preceq x \implies x = y$.

Exemple C.14. — $\mathbb{R}$ muni de l'ordre usuel.

— $\{0, 1\}^{\mathbb{Z}^2}$, avec $(a_i)_{i \in \mathbb{Z}^2} \preceq (b_i)_{i \in \mathbb{Z}^2}$ si, pour tout i , $a_i = 1 \implies b_i = 1$. Si $a_i = 1$, on dit que le site i est occupé, vide sinon.

Définition C.15. Un événement A est dit *croissant* si $x \in A \implies \{y \in E : x \preceq y\} \subset A$.

Exemple C.16. Pour $E = \{0, 1\}^{\mathbb{Z}^2}$, l'événement

$$A = \{\text{il existe un chemin infini constitué de sites occupés adjacents}\}$$

est croissant.

Définition C.17. Soient $\mu, \nu \in \mathcal{M}_1(E)$. On dit que ν domine stochastiquement μ , et on note $\mu \preceq \nu$, si

$$\forall A \in \mathcal{B}(E) \text{ croissant, } \mu(A) \leq \nu(A),$$

ou, de manière équivalente,

$$\int_E f \, d\mu \leq \int_E f \, d\nu \quad \forall f : E \rightarrow \mathbb{R} \text{ mesurable, bornée, croissante.}$$

Par croissante, on veut dire $x \preceq y \implies f(x) \leq f(y)$.

Exemple C.18. Pour $E = \mathbb{R}$, $\mu \preceq \nu$ si et seulement si, pour tout $x \in \mathbb{R}$, $\mu([-\infty, x]) \geq \nu([-\infty, x])$.

Théorème C.19 (Strassen). On suppose que $\{(x, y) \in E^2 : x \preceq y\}$ est fermé. Soient $\mu, \nu \in \mathcal{M}_1(E)$. Alors les deux assertions suivantes sont équivalentes :

- (i) $\mu \preceq \nu$;
- (ii) il existe un couplage (X, Y) tel que $\mathcal{L}(X) = \mu$, $\mathcal{L}(Y) = \nu$ et $\mathbb{P}(X \preceq Y) = 1$.

La preuve repose sur de l'analyse fonctionnelle; voir Lindvall, *Lectures on the Coupling Method*, chapitre IV, ou Lindvall, *On Strassen's Theorem on Stochastic Domination*, ECP 4 (1999), 51–59.

5 Théorème de Donsker et applications

Exercices associés : feuille 5, p. 177.

Dans ce chapitre

5.1	Le théorème de Donsker	55
5.2	Loi de l'arcsinus	58
5.3	Théorème de Lévy	60

5.1 Le théorème de Donsker

L'inégalité suivante sera utile pour démontrer le théorème d'invariance de Donsker.

Proposition 5.1 (Inégalité de Skorokhod). Soient $(X_i)_{i \geq 1}$ des variables aléatoires i.i.d. à valeurs dans $\mathbb{R}^d$, centrées, telles que $c = \mathbb{E}[|X_1|^2] < \infty$. On pose $S_0 = 0$ et $S_n = X_1 + \dots + X_n$ pour $n \geq 1$. Alors, pour tous $\lambda > 0$ et $n \geq 1$,

$$\mathbb{P} \left(\max_{1 \leq i \leq n} |S_i| > \lambda \sqrt{n} \right) \leq 2\mathbb{P} \left(|S_n| > (\lambda - \sqrt{2c})\sqrt{n} \right),$$

où $|\cdot|$ désigne la norme euclidienne.

Démonstration. Si $c = 0$, alors $X_i = 0$ presque sûrement pour tout i , et l'inégalité est immédiate. Supposons $c > 0$. Si $0 < \lambda < \sqrt{2c}$, le membre de droite vaut 2. Il reste donc à traiter $\lambda \geq \sqrt{2c}$. Soit $T = \inf\{k \geq 1 : |S_k| > \lambda \sqrt{n}\}$; on écrit

$$\begin{aligned} \mathbb{P} \left(\max_{1 \leq i \leq n} |S_i| > \lambda \sqrt{n} \right) &= \mathbb{P}(T \leq n) \\ &\leq \mathbb{P}(|S_n| > (\lambda - \sqrt{2c})\sqrt{n}) + \mathbb{P}(T \leq n, |S_n| \leq (\lambda - \sqrt{2c})\sqrt{n}). \end{aligned}$$

Or, pour $1 \leq j \leq n$, sur l'événement $\{T = j\}$, on a $|S_j| > \lambda \sqrt{n}$; et si de plus $|S_n| \leq (\lambda - \sqrt{2c})\sqrt{n}$, alors $|S_n - S_j| > \sqrt{2cn}$ par inégalité triangulaire. Par indépendance et l'inégalité de Markov,

$$\begin{aligned} \mathbb{P}(T = j, |S_n - S_j| > \sqrt{2cn}) &= \mathbb{P}(T = j)\mathbb{P}(|S_{n-j}| > \sqrt{2cn}) \\ &\leq \mathbb{P}(T = j) \frac{(n-j)\mathbb{E}[|X_1|^2]}{2cn} \leq \frac{\mathbb{P}(T = j)}{2}. \end{aligned}$$

On conclut que

$$\mathbb{P} \left(\max_{1 \leq i \leq n} |S_i| > \lambda \sqrt{n} \right) \leq \mathbb{P}(|S_n| > (\lambda - \sqrt{2c})\sqrt{n}) + \frac{1}{2}\mathbb{P}(T \leq n),$$

d'où le résultat. □

On commence par quelques notations. Soient $(X_i)_{i \geq 1}$ des variables aléatoires i.i.d. dans $\mathbb{R}^d$ telles que $\mathbb{E}[X_1] = 0$ et $\text{Cov}(X_1) = I_d$, la matrice identité dans $\mathbb{R}^d$. On pose

$S_0 = 0$, $S_n = X_1 + \cdots + X_n$ pour $n \geq 1$, et, par interpolation linéaire,

$$S_t = (1 - \{t\})S_{[t]} + \{t\}S_{[t]+1} \quad (t \geq 0).$$

Enfin, pour $n \geq 1$, on pose

$$S_t^{(n)} = \frac{S_{nt}}{\sqrt{n}}, \quad t \geq 0,$$

de sorte que $S^{(n)}$ est une variable aléatoire à valeurs dans $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$.

Théorème 5.2 (Donsker). La suite $S^{(n)}$ converge en loi dans $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ vers un processus $(B_t)_{t \geq 0}$ appelé *mouvement brownien standard de dimension d*. Sa loi est caractérisée par les propriétés suivantes : $B_0 = 0$ presque sûrement et, pour tous $0 = t_0 \leq t_1 \leq \cdots \leq t_k$, les variables aléatoires $(B_{t_i} - B_{t_{i-1}})_{1 \leq i \leq k}$ sont indépendantes, respectivement de loi gaussienne $\mathcal{N}(0, (t_i - t_{i-1})I_d)$.

Démonstration. Étape 1 : convergence des lois fini-dimensionnelles. On fixe $0 = t_0 < t_1 < \cdots < t_k$. D'après le TCL,

$$\left(\frac{S_{[nt_i]} - S_{[nt_{i-1}]}}{\sqrt{n}} : 1 \leq i \leq k \right) \xrightarrow{\text{loi}} (N_1, N_2, \dots, N_k),$$

où $N_1, \dots, N_k$ sont indépendantes et N_i est de loi $\mathcal{N}(0, (t_i - t_{i-1})I_d)$. Par image continue,

$$\left(\frac{S_{[nt_i]}}{\sqrt{n}} : 1 \leq i \leq k \right) \xrightarrow{\text{loi}} (N_1, N_1 + N_2, \dots, N_1 + \cdots + N_k).$$

Pour enlever les parties entières, on remarque que

$$\left| S_t^{(n)} - \frac{S_{[nt]}}{\sqrt{n}} \right| \leq \frac{|X_{[nt]+1}|}{\sqrt{n}},$$

qui converge en probabilité vers 0 à t fixé. Donc

$$\left(S_{t_i}^{(n)} - \frac{S_{[nt_i]}}{\sqrt{n}} : 1 \leq i \leq k \right) \xrightarrow{\mathbb{P}} 0.$$

Par le lemme de Slutsky, on en déduit que

$$(S_{t_i}^{(n)} : 1 \leq i \leq k) \xrightarrow{\text{loi}} (N_1, N_1 + N_2, \dots, N_1 + \cdots + N_k),$$

ce qui montre la convergence des lois fini-dimensionnelles.

Étape 2 : tension. Il s'agit de vérifier que, pour tout $N \geq 1$, les restrictions $(S_{[0, N]}^{(n)})_{n \geq 1}$ sont tendues. Pour simplifier les notations, on le fait pour $N = 1$ (le cas général étant identique). Avec des hypothèses de moments supplémentaires, on peut utiliser le critère de Kolmogorov. Dans le cas général, on va montrer que

$$\forall \varepsilon, \eta > 0, \exists \delta \in]0, 1[, \forall t \in [0, 1], \limsup_{n \rightarrow \infty} \frac{1}{\delta} \mathbb{P} \left(\sup_{t \leq s \leq (t+\delta) \wedge 1} |S_s^{(n)} - S_t^{(n)}| > \eta \right) \leq \varepsilon,$$

en utilisant l'inégalité maximale de Skorokhod.

a) Pour cela, on remarque d'abord que, pour $0 < \delta < 1$ et $0 \leq t \leq 1$,

$$\sup_{t \leq s \leq t+\delta} |S_{ns} - S_{nt}| \leq 2 \max_{0 \leq i \leq \lfloor n\delta \rfloor + 2} |S_{\lfloor nt \rfloor + i} - S_{\lfloor nt \rfloor}|.$$

En effet, on écrit

$$|S_{ns} - S_{nt}| \leq |S_{ns} - S_{\lfloor nt \rfloor}| + |S_{nt} - S_{\lfloor nt \rfloor}|.$$

D'une part,

$$|S_{nt} - S_{\lfloor nt \rfloor}| \leq |S_{\lfloor nt \rfloor + 1} - S_{\lfloor nt \rfloor}|,$$

car $(S_{\lfloor nt \rfloor + u} : 0 \leq u \leq 1)$ décrit un segment. D'autre part,

$$|S_{ns} - S_{\lfloor nt \rfloor}| \leq \max(|S_{\lfloor ns \rfloor} - S_{\lfloor nt \rfloor}|, |S_{\lfloor ns \rfloor + 1} - S_{\lfloor nt \rfloor}|),$$

car $(S_{\lfloor ns \rfloor + u} : 0 \leq u \leq 1)$ décrit un segment. Comme $\lfloor ns \rfloor + 1 \leq \lfloor nt \rfloor + \lfloor n\delta \rfloor + 2$, on en déduit l'inégalité désirée.

b) Ainsi, pour tous $\eta > 0$, $\delta \in]0, 1[$ et $t \in [0, 1]$,

$$\begin{aligned} \mathbb{P} \left(\sup_{t \leq s \leq t+\delta} |S_s^{(n)} - S_t^{(n)}| > \eta \right) &\leq \mathbb{P} \left(\max_{1 \leq i \leq \lfloor n\delta \rfloor + 2} |S_{\lfloor nt \rfloor + i} - S_{\lfloor nt \rfloor}| > \frac{\eta\sqrt{n}}{2} \right) \\ &= \mathbb{P} \left(\max_{1 \leq i \leq \lfloor n\delta \rfloor + 2} |S_i| > \frac{\eta\sqrt{n}}{2\sqrt{\lfloor n\delta \rfloor + 2}} \sqrt{\lfloor n\delta \rfloor + 2} \right) \\ &\leq 2\mathbb{P} \left(\frac{|S_{\lfloor n\delta \rfloor + 2}|}{\sqrt{\lfloor n\delta \rfloor + 2}} > \frac{\eta}{2} \frac{1}{\sqrt{\delta + 2/n}} - \sqrt{2d} \right), \end{aligned}$$

d'après l'inégalité de Skorokhod, car $\mathbb{E}[|X_1|^2] = d$. D'après le TCL,

$$\limsup_{n \rightarrow \infty} \mathbb{P} \left(\sup_{t \leq s \leq t+\delta} |S_s^{(n)} - S_t^{(n)}| > \eta \right) \leq 2\mathbb{P} \left(|\mathcal{N}(0, I_d)| > \frac{\eta}{2\sqrt{\delta}} - \sqrt{2d} \right).$$

Or $\mathbb{P}(|\mathcal{N}(0, I_d)| \geq x)$ décroît exponentiellement vite quand $x \rightarrow \infty$. Donc on peut trouver $0 < \delta < 1$ tel que

$$\frac{1}{\delta} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\sup_{t \leq s \leq t+\delta} |S_s^{(n)} - S_t^{(n)}| > \eta \right) \leq \varepsilon.$$

□

Corollaire 5.3. Soient $(X_i)_{i \geq 1}$ des variables aléatoires réelles i.i.d. centrées. Notons $\sigma^2 = \mathbb{E}[X_1^2] \in]0, \infty[$. Alors, pour tout $a > 0$,

$$\mathbb{P} \left(\max_{0 \leq i \leq n} S_i > a\sigma\sqrt{n} \right) \rightarrow 2 \int_a^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

Démonstration. En notant $S_t^{(n)} = S_{nt}/(\sigma\sqrt{n})$ pour $0 \leq t \leq 1$, on a

$$(S_t^{(n)} : 0 \leq t \leq 1) \xrightarrow{\text{loi}} (B_t)_{0 \leq t \leq 1}$$

d'après le théorème de Donsker, où $(B_t)_{0 \leq t \leq 1}$ est un mouvement brownien standard. Puisque

$$\sup_{0 \leq t \leq 1} S_t^{(n)} = \frac{1}{\sigma\sqrt{n}} \max_{0 \leq i \leq n} S_i$$

et que l'application $\sup : \mathcal{C}([0, 1], \mathbb{R}) \rightarrow \mathbb{R}$, $f \mapsto \sup f$, est continue, le résultat s'ensuit en utilisant le fait que

$$\sup_{0 \leq t \leq 1} B_t \stackrel{\text{loi}}{=} |B_1|$$

(ce qu'on peut voir avec le principe de réflexion). □

Une propriété du « continu » (la loi explicite de $\sup_{0 \leq t \leq 1} B_t$) a été utilisée pour obtenir un résultat asymptotique dans le discret (la fonction de répartition de $\max_{0 \leq k \leq n} S_k$). Dans les deux exemples suivants, c'est un résultat du discret, passant à la limite, qui permet d'obtenir une information intéressante sur le mouvement brownien.

5.2 Loi de l'arcsinus

Théorème 5.4. Soit $(B_t)_{0 \leq t \leq 1}$ un mouvement brownien standard. On pose

$$T = \sup\{t \in [0, 1] : B_t = 0\}.$$

Alors T a pour densité $\frac{1}{\pi\sqrt{x(1-x)}} \mathbb{1}_{\{0 < x < 1\}}$, ou, de façon équivalente,

$$\mathbb{P}(T \leq t) = \frac{2}{\pi} \arcsin(\sqrt{t}) \quad (0 \leq t \leq 1).$$

Démonstration. Étape 1. On pose $S_n = X_1 + \dots + X_n$. Les variables $(X_i)_{i \geq 1}$ sont i.i.d., de loi $\mathbb{P}(X_i = \pm 1) = 1/2$, et on note

$$T_{2n} = \max\{0 \leq k \leq 2n : S_k = 0\}.$$

Montrons que $T_{2n}/(2n)$ converge en loi vers une variable aléatoire de densité

$$\frac{1}{\pi\sqrt{x(1-x)}} \mathbb{1}_{\{0 < x < 1\}}.$$

Pour $0 \leq j \leq n$, on a

$$\mathbb{P}(T_{2n} = 2j) = \mathbb{P}(S_{2j} = 0) \mathbb{P}(S_1 \neq 0, \dots, S_{2(n-j)} \neq 0).$$

Or

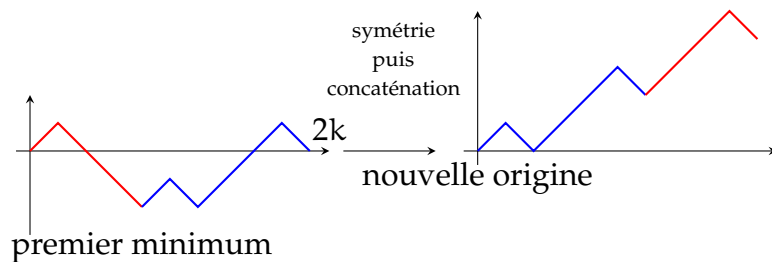
$$\forall k \geq 1, \quad \mathbb{P}(S_1 \neq 0, \dots, S_{2k} \neq 0) = \mathbb{P}(S_{2k} = 0). \quad (*)$$

Démontrons cette identité surprenante par un joli argument combinatoire.

Tout d'abord,

$$\begin{aligned}\mathbb{P}(S_1 \neq 0, \dots, S_{2k} \neq 0) &= 2\mathbb{P}(S_1 > 0, \dots, S_{2k} > 0) \\ &= 2 \times \frac{1}{2} \mathbb{P}(S_1 \geq 0, \dots, S_{2k-1} \geq 0) \\ &= \mathbb{P}(S_1 \geq 0, \dots, S_{2k} \geq 0).\end{aligned}$$

La deuxième égalité s'obtient en changeant l'origine du repère en $(1, 1)$; la dernière par parité, puisque $S_{2k-1} \geq 0 \implies S_{2k} \geq 0$. Construisons maintenant une bijection entre les chemins $(S_1, \dots, S_{2k})$ tels que $S_{2k} = 0$ et ceux tels que $S_1 \geq 0, \dots, S_{2k} \geq 0$.



On prend la partie avant le premier instant où l'infimum est atteint. On l'enlève et on fait une symétrie d'axe (Oy) . On la concatène à droite de ce qui reste (partie bleue) et on place l'origine du repère au premier point de la partie bleue (voir la figure). Après transformation, on obtient un chemin tel que $S_1 \geq 0, \dots, S_{2k} \geq 0$. On vérifie que c'est une bijection (la transformation inverse se décrit aisément). Ceci montre (*). Ainsi,

$$\mathbb{P}(T_{2n} = 2j) = \mathbb{P}(S_{2j} = 0) \mathbb{P}(S_{2(n-j)} = 0).$$

Or

$$\mathbb{P}(S_{2k} = 0) = \binom{2k}{k} \frac{1}{2^{2k}} \underset{k \rightarrow \infty}{\sim} \frac{1}{\sqrt{\pi k}}.$$

Donc, pour $0 < c < 1$,

$$\mathbb{P}(T_{2n} = 2\lfloor cn \rfloor) \underset{n \rightarrow \infty}{\sim} \frac{1}{n} \frac{1}{\pi \sqrt{c(1-c)}}.$$

Appliquons le lemme 6.3 avec $k = 1$, $X_n = T_{2n}/2$ et $\alpha_n = n$. La limite locale ci-dessus vaut pour $c \in]0, 1[$; pour $c \notin [0, 1]$, la probabilité considérée est nulle pour n assez grand. On obtient donc, pour presque tout $c \in \mathbb{R}$, la densité $\phi(c) = (\pi \sqrt{c(1-c)})^{-1}$ sur $]0, 1[$, prolongée par 0 ailleurs. Comme son intégrale vaut 1, le lemme donne la convergence en loi de $T_{2n}/(2n)$ annoncée.

Étape 2. On travaille sur le sous-espace fermé

$$\mathcal{C}_0 = \{f \in \mathcal{C}([0, 1], \mathbb{R}) : f(0) = 0\}.$$

Pour $f \in \mathcal{C}_0$, on pose

$$T(f) = \max\{t \in [0, 1] : f(t) = 0\}.$$

Cette application est bien définie et mesurable : pour $\alpha \in [0, 1]$, l'ensemble $\{f : T(f) \geq \alpha\}$ est fermé. Elle n'est pas continue partout : pour $f_n(x) = x/n$, on a $T(f_n) = 0$, alors que $T(0) = 1$.

Elle est toutefois continue en toute fonction f telle que $0 < T(f) < 1$ et qui prend les deux signes dans tout voisinage de $T(f)$. En effet, une paire de signes opposés près de $T(f)$ reste de signes opposés après une perturbation uniforme assez petite, ce qui fournit un zéro voisin par le théorème des valeurs intermédiaires. D'autre part, pour tout $\varepsilon > 0$ tel que $T(f) + \varepsilon \leq 1$, le minimum de $|f|$ sur $[T(f) + \varepsilon, 1]$ est strictement positif : aucune petite perturbation uniforme ne peut avoir de zéro dans cet intervalle.

Presque sûrement, $B_1 \neq 0$ et B prend les deux signes arbitrairement près de 0, donc $0 < T(B) < 1$. De plus, 0 n'est la valeur d'aucun extremum local intérieur. En effet, pour des rationnels $0 < p < q < 1$, le minimum de B sur $[p, q]$ s'écrit

$$B_p + \inf_{0 \leq u \leq q-p} (B_{p+u} - B_p).$$

Les deux termes sont indépendants et B_p a une densité : ce minimum n'a donc pas d'atome en 0. Il en va de même du maximum. Une réunion dénombrable sur les intervalles rationnels exclut alors tout extremum local intérieur de valeur 0. Ainsi B prend les deux signes dans tout voisinage de $T(B)$, et vérifie presque sûrement la condition de continuité précédente. D'après Donsker et le théorème de l'application continue,

$$T\left(\left(\frac{S_{2nt}}{\sqrt{2n}} : 0 \leq t \leq 1\right)\right) \xrightarrow{\text{loi}} T(B).$$

Pour la marche simple considérée à l'étape 1, les zéros de l'interpolation linéaire se trouvent aux instants entiers. Le membre de gauche vaut donc $T_{2n}/(2n)$, ce qui identifie la loi de $T(B)$. $\square$

Corollaire 5.5. Soient $(X_i)_{i \geq 1}$ des variables aléatoires réelles i.i.d. telles que $\mathbb{E}[X_1] = 0$ et $\mathbb{E}[X_1^2] \in]0, \infty[$. On pose $S_0 = 0$, $S_k = X_1 + \dots + X_k$, et on prolonge S par interpolation linéaire. Alors

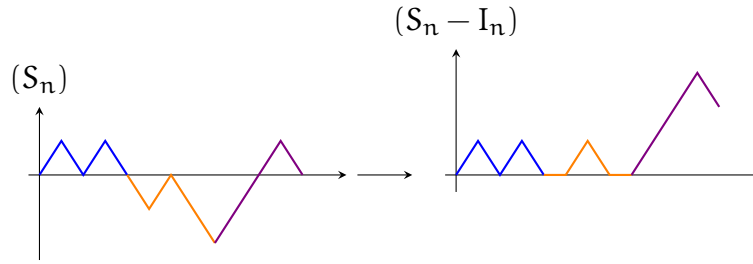
$$\mathbb{P}\left(\frac{1}{2n} \max\{u \in [0, 2n] : S_u = 0\} \leq t\right) \longrightarrow \frac{2}{\pi} \arcsin(\sqrt{t}) \quad (0 < t < 1).$$

5.3 Théorème de Lévy

Soit $(B_t)_{t \geq 0}$ un mouvement brownien et $I_t = \inf_{[0, t]} B$ l'infimum courant.

Théorème 5.6 (Lévy). On a $(B_t - I_t)_{t \geq 0} \stackrel{\text{loi}}{=} (|B_t|)_{t \geq 0}$.

Démonstration. Soient $(X_i)_{i \geq 1}$ des variables aléatoires i.i.d. telles que $\mathbb{P}(X_i = \pm 1) = 1/2$. On pose $S_n = X_1 + \dots + X_n$ et $I_n = \min_{0 \leq k \leq n} S_k$.



Si on regarde les probabilités de transition de $S_n - I_n$ et celles de $|S_n|$, on voit que ce sont les mêmes en dehors de l'état 0, où $S_n - I_n$ peut rester coincé avec probabilité $1/2$. Posons $R_n = S_n - I_n$ et

$$\theta_n = \#\{1 \leq k \leq n : R_k \neq R_{k-1}\}, \quad \theta_0 = 0.$$

Un pas est supprimé exactement lorsque la marche atteint un nouveau minimum strict, de sorte que

$$n - \theta_n = -I_n.$$

La loi forte des grands nombres entraîne $I_n/n \rightarrow 0$ presque sûrement. En prolongeant θ par interpolation linéaire, on obtient donc, presque sûrement, pour tout $T > 0$,

$$\sup_{0 \leq t \leq T} \left| \frac{\theta_{nt}}{n} - t \right| \rightarrow 0.$$

En effet, ce supremum est majoré par $n^{-1} \max_{0 \leq k \leq \lceil nT \rceil} |S_k|$, qui tend vers 0.

Pour $m \in \mathbb{N}$, posons

$$\tau_m = \inf\{k \in \mathbb{N} : \theta_k = m\}, \quad Y_m = R_{\tau_m}.$$

Ces temps sont presque sûrement finis puisque $\theta_n/n \rightarrow 1$. La chaîne $(Y_m)_{m \geq 0}$ a les mêmes transitions que $(|S_m|)_{m \geq 0}$, y compris en 0; leurs lois sont donc égales. Prolongeons R et Y par interpolation linéaire. On a alors, pour tout $t \geq 0$,

$$R_t = Y_{\theta_t}.$$

Par Donsker,

$$(Y_{nt}/\sqrt{n})_{t \geq 0} \xrightarrow{\text{loi}} (|B_t|)_{t \geq 0}.$$

La convergence du changement de temps vers l'identité, le lemme de Slutsky et la continuité de la composition donnent

$$(R_{nt}/\sqrt{n})_{t \geq 0} \xrightarrow{\text{loi}} (|B_t|)_{t \geq 0}.$$

D'autre part, $R_t = S_t - \inf_{0 \leq u \leq t} S_u$ pour la marche simple interpolée. L'application qui retranche l'infimum courant est continue pour la convergence uniforme sur les compacts; Donsker donne donc aussi

$$(R_{nt}/\sqrt{n})_{t \geq 0} \xrightarrow{\text{loi}} (B_t - I_t)_{t \geq 0}.$$

L'unicité de la limite en loi conclut. □

6 Le pont brownien

Exercices associés : feuille 5, p. 177.

Dans ce chapitre

6.1	Définition du pont brownien	62
6.2	Le théorème local limite	63
6.3	Convergence vers le pont brownien	66

6.1 Définition du pont brownien

Soit $(B_t)_{0 \leq t \leq 1}$ un mouvement brownien standard.

Théorème 6.1 (Définition du pont brownien). Lorsque $\varepsilon \downarrow 0$, la loi de $(B_s)_{0 \leq s \leq 1}$ sachant que $\{|B_1| \leq \varepsilon\}$ converge en loi vers une mesure de probabilité $\mathbb{P}_0$ sur $\mathcal{C}([0, 1], \mathbb{R})$, appelée *loi du pont brownien*. Par ailleurs, la loi de $(B_s - sB_1)_{0 \leq s \leq 1}$ est $\mathbb{P}_0$.

Informellement, le pont brownien est « $(B_t)_{0 \leq t \leq 1}$ sachant $B_1 = 0$ ».

Démonstration. Posons $b_s = B_s - sB_1$ pour $0 \leq s \leq 1$. La famille $((b_s)_{0 \leq s \leq 1}, B_1)$ est gaussienne et, par la covariance brownienne,

$$\text{Cov}(b_s, B_1) = \text{Cov}(B_s, B_1) - s \text{Var}(B_1) = s - s = 0.$$

Ainsi b est indépendant de B_1 . Soit $F : \mathcal{C}([0, 1], \mathbb{R}) \rightarrow \mathbb{R}$ bornée et lipschitzienne, de constante L . Comme $\|B - b\|_\infty = |B_1|$, l'indépendance donne, pour tout $\varepsilon > 0$,

$$\begin{aligned} |\mathbb{E}[F(B) \mid |B_1| \leq \varepsilon] - \mathbb{E}[F(b)]| &= |\mathbb{E}[F(B) - F(b) \mid |B_1| \leq \varepsilon]| \\ &\leq L \mathbb{E}[\|B - b\|_\infty \mid |B_1| \leq \varepsilon] \leq L\varepsilon. \end{aligned}$$

La caractérisation de la convergence en loi par les fonctions lipschitziennes bornées conclut : la limite est la loi de b . □

Dans la suite, on note $p_t(x) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/(2t)}$ la densité de B_t , pour $t > 0$.

Proposition 6.2. La loi du pont brownien est caractérisée par la relation suivante d'absolue continuité : pour $0 < t < 1$ et $F : \mathcal{C}([0, t], \mathbb{R}) \rightarrow \mathbb{R}$ continue bornée ou mesurable positive,

$$\mathbb{E}[F(b_s : 0 \leq s \leq t)] = \mathbb{E} \left[F(B_s : 0 \leq s \leq t) \frac{p_{1-t}(-B_t)}{p_1(0)} \right].$$

Démonstration. Il suffit de montrer le résultat pour $F : \mathcal{C}([0, t], \mathbb{R}) \rightarrow \mathbb{R}$ continue bornée. On en déduit le résultat pour $F = \mathbb{1}_A$ avec A fermé par approximation, puis $F = \mathbb{1}_A$ avec A borélien par classe monotone, puis F étagée par linéarité, et enfin $F \geq 0$ mesurable par approximation et convergence monotone. On pose $W_u = B_{t+u} - B_t$, de

sorte que $(W_u)_{0 \leq u \leq 1-t}$ est un mouvement brownien indépendant de $(B_u)_{0 \leq u \leq t}$. Alors

$$\begin{aligned} & \mathbb{E}[F(B_s : 0 \leq s \leq t) \mid |B_1| \leq \varepsilon] \\ &= \frac{1}{\mathbb{P}(|B_1| \leq \varepsilon)} \mathbb{E}[F(B_s : 0 \leq s \leq t) \mathbb{1}_{\{|B_t + W_{1-t}| \leq \varepsilon\}}] \\ &= \frac{1}{\mathbb{P}(|B_1| \leq \varepsilon)} \mathbb{E}[F(B_s : 0 \leq s \leq t) \mathbb{1}_{\{W_{1-t} \in [-B_t - \varepsilon, -B_t + \varepsilon]\}}] \\ &= \frac{1}{\int_{-\varepsilon}^{\varepsilon} p_1(x) dx} \mathbb{E} \left[F(B_s : 0 \leq s \leq t) \int_{-B_t - \varepsilon}^{-B_t + \varepsilon} p_{1-t}(x) dx \right]. \end{aligned}$$

Or

$$\frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} p_1(x) dx \longrightarrow p_1(0), \quad \frac{1}{2\varepsilon} \int_{-B_t - \varepsilon}^{-B_t + \varepsilon} p_{1-t}(x) dx \longrightarrow p_{1-t}(-B_t).$$

De plus, p_{1-t} étant bornée,

$$\frac{1}{2\varepsilon} \int_{-B_t - \varepsilon}^{-B_t + \varepsilon} p_{1-t}(x) dx \leq \|p_{1-t}\|_{\infty}.$$

Par convergence dominée,

$$\mathbb{E}[F(B_s : 0 \leq s \leq t) \mid |B_1| \leq \varepsilon] \longrightarrow \mathbb{E} \left[F(B_s : 0 \leq s \leq t) \frac{p_{1-t}(-B_t)}{p_1(0)} \right].$$

Comme aussi $\mathbb{E}[F(B_s : 0 \leq s \leq t) \mid |B_1| \leq \varepsilon] \rightarrow \mathbb{E}[F(b_s : 0 \leq s \leq t)]$, on en déduit le résultat. $\square$

6.2 Le théorème local limite

Lemme 6.3. Soit $k \geq 1$ entier ; pour $1 \leq i \leq k$, soit (α_n^i) une suite de réels strictement positifs telle que $\alpha_n^i \rightarrow \infty$. On considère une variable aléatoire $(X_n^1, \dots, X_n^k)$ à valeurs dans $\mathbb{Z}^k$ telle que, pour presque tout $(\lambda_1, \dots, \lambda_k) \in \mathbb{R}^k$ au sens de la mesure de Lebesgue,

$$\alpha_n^1 \alpha_n^2 \cdots \alpha_n^k \mathbb{P}((X_n^1, \dots, X_n^k) = (\lfloor \lambda_1 \alpha_n^1 \rfloor, \dots, \lfloor \lambda_k \alpha_n^k \rfloor)) \longrightarrow \phi(\lambda_1, \dots, \lambda_k),$$

où ϕ est une densité de probabilité sur $\mathbb{R}^k$. Alors

$$\left(\frac{X_n^1}{\alpha_n^1}, \dots, \frac{X_n^k}{\alpha_n^k} \right) \xrightarrow{\text{loi}} X,$$

où X a pour densité ϕ .

Démonstration. On va utiliser le lemme de Scheffé (voir l'exercice 1.4) : si X_n a pour densité f_n sur $\mathbb{R}^k$ et si $f_n \rightarrow f$ presque partout, avec f une densité de probabilité, alors

$X_n \xrightarrow{\text{loi}} X$, où X a pour densité f . Soit $(Y_n^1, \dots, Y_n^k)$ une variable aléatoire de densité

$$\phi_n(\lambda_1, \dots, \lambda_k) = \alpha_n^1 \cdots \alpha_n^k \mathbb{P}((X_n^1, \dots, X_n^k) = (\lfloor \lambda_1 \alpha_n^1 \rfloor, \dots, \lfloor \lambda_k \alpha_n^k \rfloor))$$

sur $\mathbb{R}^k$, et soit

$$(Z_n^1, \dots, Z_n^k) = (\lfloor \alpha_n^1 Y_n^1 \rfloor, \dots, \lfloor \alpha_n^k Y_n^k \rfloor),$$

qui a la même loi que $(X_n^1, \dots, X_n^k)$. Puisque $\phi_n \rightarrow \phi$ presque partout, le lemme de Scheffé donne $(Y_n^1, \dots, Y_n^k) \xrightarrow{\text{loi}} X$; et comme

$$\left| Y_n^i - \frac{Z_n^i}{\alpha_n^i} \right| \leq \frac{1}{\alpha_n^i} \rightarrow 0 \quad (1 \leq i \leq k),$$

on en déduit le résultat. $\square$

Définition 6.4. Une variable aléatoire réelle X est dite *lattice* s'il existe $b \in \mathbb{R}$ et $h \in \mathbb{R}_+^*$ tels que $\mathbb{P}(X \in b + h\mathbb{Z}) = 1$. Si X n'est pas presque sûrement constante, il existe un plus grand tel h , appelé *span* de X . S'il vaut 1, on dit que X est *apériodique*. Dans le cas dégénéré $X = b$ presque sûrement, tout $h > 0$ convient et il n'existe pas de span fini maximal.

Exemple 6.5. Si $\mathbb{P}(X = 1) = \mathbb{P}(X = -1) = 1/2$, X a pour span 2.

Lemme 6.6. Soit X une variable aléatoire réelle et ϕ sa fonction caractéristique.

- (1) X est lattice si et seulement s'il existe $t \in \mathbb{R}^*$ tel que $|\phi(t)| = 1$.
- (2) X a pour span $h > 0$ si et seulement si $|\phi(2\pi/h)| = 1$ et $|\phi(t)| < 1$ pour $t \in]0, 2\pi/h[$.

Démonstration. (1) Si $\mathbb{P}(X \in b + h\mathbb{Z}) = 1$, alors $\mathbb{P}((X - b)/h \in \mathbb{Z}) = 1$, de sorte que

$$\mathbb{E} \left[e^{2i\pi(X-b)/h} \right] = 1, \quad \left| \mathbb{E} \left[e^{2i\pi X/h} \right] \right| = |e^{2i\pi b/h}| = 1.$$

Réciproquement, si $|\phi(t_0)| = 1$ pour $t_0 \in \mathbb{R}^*$, soit $a \in \mathbb{R}$ tel que $\phi(t_0) = e^{ia}$. Alors $\mathbb{E}[e^{i(t_0 X - a)}] = 1$. Donc $\mathbb{E}[\cos(t_0 X - a)] = 1$. Comme $\cos \leq 1$, on en déduit que, presque sûrement, $\cos(t_0 X - a) = 1$. Donc, presque sûrement, $t_0 X - a \in 2\pi\mathbb{Z}$, et

$$\mathbb{P} \left(X \in \frac{a}{t_0} + \frac{2\pi}{t_0} \mathbb{Z} \right) = 1.$$

(2) La preuve de (1) montre qu'il existe $b \in \mathbb{R}$ et $h > 0$ tels que $\mathbb{P}(X \in b + h\mathbb{Z}) = 1$ si et seulement si $|\phi(2\pi/h)| = 1$, d'où le résultat. $\square$

Théorème 6.7 (Théorème local limite). Soient $(X_i)_{i \geq 1}$ des variables aléatoires réelles i.i.d. apériodiques à valeurs dans $\mathbb{Z}$, telles que $\alpha = \mathbb{E}[X_1]$ et $\sigma^2 = \mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2 \in]0, \infty[$. On note $S_n = X_1 + \dots + X_n$. Alors

$$\sup_{k \in \mathbb{Z}} \left| \sqrt{n} \mathbb{P}(S_n = k) - \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left(-\frac{1}{2} \left(\frac{k - \alpha n}{\sigma\sqrt{n}} \right)^2 \right) \right| \rightarrow 0.$$

Démonstration. Posons $X'_i = X_i - \alpha$, $S'_n = S_n - \alpha n$. On note $\phi(t) = \mathbb{E}[e^{itX'_1}]$. Puisque $\mathbb{E}[e^{itS'_n}] = \sum_{j \in \mathbb{Z} - \alpha n} e^{itj} \mathbb{P}(S'_n = j)$, il vient, pour $k \in \mathbb{Z} - \alpha n$,

$$\mathbb{P}(S'_n = k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-itk} \mathbb{E}[e^{itS'_n}] dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-itk} \phi(t)^n dt.$$

Ainsi, $u \in \mathbb{R}$ étant choisi de sorte que $u\sigma\sqrt{n} \in \mathbb{Z} - \alpha n$,

$$\sigma\sqrt{n} \mathbb{P}(S'_n = u\sigma\sqrt{n}) = \frac{1}{2\pi} \int_{-\pi\sigma\sqrt{n}}^{\pi\sigma\sqrt{n}} e^{-itu} \phi \left(\frac{t}{\sigma\sqrt{n}} \right)^n dt.$$

Mais, pour tout $u \in \mathbb{R}$,

$$\frac{1}{\sqrt{2\pi}} e^{-u^2/2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itu - t^2/2} dt.$$

Posons $q(t) = e^{-t^2/2}$ et

$$g_n(t) = \mathbb{1}_{\{|t| \leq \pi\sigma\sqrt{n}\}} \phi \left(\frac{t}{\sigma\sqrt{n}} \right)^n.$$

Les deux formules d'inversion précédentes donnent

$$\sup_{k \in \mathbb{Z}} \left| \sigma\sqrt{n} \mathbb{P}(S_n = k) - \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{(k - \alpha n)^2}{2n\sigma^2} \right) \right| \leq \frac{1}{2\pi} \|g_n - q\|_{L^1(\mathbb{R})}.$$

Il suffit donc de montrer que $g_n \rightarrow q$ dans $L^1(\mathbb{R})$. Le centrage et le moment d'ordre deux donnent

$$\phi(t) = 1 - \frac{t^2\sigma^2}{2} + t^2g(t), \quad g(0) = 0, \quad g \text{ continue.} \quad (*)$$

En effet, le reste de Taylor satisfait

$$|e^{ix} - 1 - ix + x^2/2| \leq \min(|x|^3/6, x^2).$$

Ainsi $|g(t)| \leq \mathbb{E}[\min(|t|X'_1|^3/6, (X'_1)^2)] \rightarrow 0$ par convergence dominée. De plus, $(*)$ entraîne $|\phi(t)|^2 = 1 - \sigma^2 t^2 + o(t^2)$. On peut donc choisir $0 < \delta < \pi$ tel que

$$|\phi(t)| \leq e^{-\sigma^2 t^2/4} \quad (|t| \leq \delta).$$

Posons

$$h_n(t) = \mathbb{1}_{\{|t| \leq \delta \sigma \sqrt{n}\}} \phi \left(\frac{t}{\sigma \sqrt{n}} \right)^n.$$

Pour tout $t \in \mathbb{R}$, le développement (*) donne $h_n(t) \rightarrow q(t)$, tandis que $|h_n(t)| \leq e^{-t^2/4}$. Par convergence dominée, $\|h_n - q\|_{L^1(\mathbb{R})} \rightarrow 0$. Enfin, l'apériodicité et la continuité de ϕ donnent

$$\rho = \sup_{\delta \leq |t| \leq \pi} |\phi(t)| < 1, \quad \|g_n - h_n\|_{L^1(\mathbb{R})} \leq 2\pi \sigma \sqrt{n} \rho^n \rightarrow 0.$$

L'inégalité triangulaire prouve $\|g_n - q\|_{L^1(\mathbb{R})} \rightarrow 0$. La borne issue de l'inversion, divisée par σ , conclut. $\square$

Corollaire 6.8. Sous les hypothèses du théorème local limite, en supposant $\mathbb{E}[X_1] = 0$, on a :

- (1) $\mathbb{P}(S_n = 0) \sim \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{\sqrt{n}}$;
- (2) il existe $C > 0$ tel que, pour tous $n \geq 1$ et $k \in \mathbb{Z}$, $\mathbb{P}(S_n = k) \leq C/\sqrt{n}$;
- (3) si (x_n) est une suite d'entiers telle que $x_n/(\sigma\sqrt{n}) \rightarrow x \in \mathbb{R}$, alors

$$\sqrt{\sigma^2 n} \mathbb{P}(S_n = x_n) \rightarrow \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Remarque 6.9. Il est possible de montrer l'amélioration suivante, utile pour $|k|$ grand :

$$\sup_{k \in \mathbb{Z}} \max \left(1, \left(\frac{k - an}{\sqrt{n}} \right)^2 \right) \left| \sigma \sqrt{n} \mathbb{P}(S_n = k) - \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{k - an}{\sigma \sqrt{n}} \right)^2 \right) \right| \rightarrow 0.$$

Voir Spitzer, *Principles of Random Walk*, chapitre II, section 7, p. 90.

6.3 Convergence vers le pont brownien

On rappelle la notation $p_t(x) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/(2t)}$.

Théorème 6.10 (Donsker conditionné). Soient $(X_i)_{i \geq 1}$ des variables aléatoires i.i.d. à valeurs dans $\mathbb{Z}$, apériodiques, telles que $\mathbb{E}[X_1] = 0$ et $\sigma^2 = \mathbb{E}[X_1^2] \in]0, \infty[$. On pose $S_n = X_1 + \dots + X_n$ et on définit S_t par interpolation linéaire. Alors

$$\left(\frac{S_{nt}}{\sigma \sqrt{n}} : 0 \leq t \leq 1 \right) \text{ sous } \mathbb{P}(\cdot \mid S_n = 0) \xrightarrow{\text{loi}} (b_t)_{0 \leq t \leq 1}.$$

Remarque 6.11. L'apériodicité implique que $\mathbb{P}(S_n = 0) > 0$ pour tout n assez grand (conséquence du théorème local limite, par exemple).

Démonstration. Étape 1 : convergence sur $[0, s]$, avec $0 < s < 1$ fixé.

Soit $F : \mathcal{C}([0, s], \mathbb{R}) \rightarrow \mathbb{R}$ continue bornée. Alors

$$\begin{aligned} \mathbb{E} \left[F \left(\frac{S_{nt}}{\sigma\sqrt{n}} : 0 \leq t \leq s \right) \middle| S_n = 0 \right] \\ = \mathbb{E} \left[F \left(\frac{S_{nt}}{\sigma\sqrt{n}} : 0 \leq t \leq s \right) \frac{Q_{n-[ns]}(-S_{[ns]})}{Q_n(0)} \right], \end{aligned}$$

avec $Q_n(k) = \mathbb{P}(S_n = k)$. Fixons $s' \in]s, 1[$. Par Donsker et le théorème de représentation de Skorokhod, on peut construire des copies des trajectoires sur $[0, s']$ telles que

$$\sup_{0 \leq t \leq s'} \left| \frac{S_{nt}}{\sigma\sqrt{n}} - B_t \right| \longrightarrow 0 \quad \text{presque sûrement.}$$

On conserve les mêmes notations pour ces copies. Comme $[ns]/n \rightarrow s$ et que B est continu, on a en particulier

$$\frac{S_{[ns]}}{\sigma\sqrt{n}} \longrightarrow B_s \quad \text{presque sûrement.}$$

En utilisant le théorème local limite, on obtient :

(i) il existe $C > 0$ tel que, pour tout n assez grand,

$$\frac{Q_{n-[ns]}(-S_{[ns]})}{Q_n(0)} \leq C$$

(avec les assertions (1) et (2) du corollaire 6.8);

(ii) presque sûrement,

$$\frac{Q_{n-[ns]}(-S_{[ns]})}{Q_n(0)} \longrightarrow \frac{p_1(-B_s/\sqrt{1-s})}{\sqrt{1-s} p_1(0)} = \frac{p_{1-s}(-B_s)}{p_1(0)}$$

(avec l'assertion (3) du corollaire 6.8).

Par convergence dominée, on conclut que

$$\begin{aligned} \mathbb{E} \left[F \left(\frac{S_{nt}}{\sigma\sqrt{n}} : 0 \leq t \leq s \right) \middle| S_n = 0 \right] \\ \longrightarrow \mathbb{E} \left[F(B_t : 0 \leq t \leq s) \frac{p_{1-s}(-B_s)}{p_1(0)} \right] = \mathbb{E}[F(b_t : 0 \leq t \leq s)]. \end{aligned}$$

Étape 2 : convergence des lois fini-dimensionnelles sur $[0, 1]$. Ceci provient de l'étape 1, avec le fait qu'en $t = 1$ les deux processus valent 0.

Étape 3 : tension sur $[0, 1]$. On utilise un argument de renversement du temps : en posant $\hat{S}_i^n = S_n - S_{n-i}$, on a clairement

$$(S_i)_{0 \leq i \leq n} \stackrel{\text{loi}}{=} (\hat{S}_i^n)_{0 \leq i \leq n},$$

et cette égalité en loi reste vraie sachant $S_n = 0$. On en déduit que, sachant $S_n = 0$,

$$S^{1,n} := \left(\frac{S_{nt}}{\sigma\sqrt{n}} : 0 \leq t \leq \frac{3}{4} \right) \xrightarrow{\text{loi}} \left(b_t : 0 \leq t \leq \frac{3}{4} \right),$$

$$S^{2,n} := \left(\frac{S_{n(1-t)}}{\sigma\sqrt{n}} : 0 \leq t \leq \frac{3}{4} \right) \xrightarrow{\text{loi}} \left(\tilde{b}_t : 0 \leq t \leq \frac{3}{4} \right),$$

avec $\tilde{b}$ un pont brownien. Il s'agit ici de deux convergences marginales séparées, qui suffisent à établir la tension des deux suites. On en déduit que $S^{1,n}$ et $S^{2,n}$ sont tendus. Du fait que, pour $\delta < 1/4$,

$$\omega \left(\left(\frac{S_{nt}}{\sigma\sqrt{n}} : 0 \leq t \leq 1 \right), \delta \right) = \max(\omega(S^{1,n}, \delta), \omega(S^{2,n}, \delta)),$$

on conclut que $(S_{nt}/(\sigma\sqrt{n}) : 0 \leq t \leq 1)$ sous $\mathbb{P}(\cdot \mid S_n = 0)$ est tendu. $\square$

7 Processus discontinus : l'espace de Skorokhod

Exercices associés : feuille 6, p. 196.

Dans ce chapitre

7.1	Propriétés de base	69
7.2	Topologie de Skorokhod (J_1) sur $\mathbb{D}$	71
7.3	Convergence étroite dans $\mathbb{D}$	74
7.4	Critères de tension	75
7.5	Compléments facultatifs	78

On note $\mathbb{D}$ (ou $\mathbb{D}([0, 1], \mathbb{R}^d)$) l'ensemble des fonctions $x : [0, 1] \rightarrow \mathbb{R}^d$ qui sont

- continues à droite : pour tout $t < 1$, $x(t+) = \lim_{s \downarrow t} x(s) = x(t)$;
- ayant des limites à gauche : pour tout $0 < t \leq 1$, $x(t-) = \lim_{s \uparrow t} x(s)$ existe.

Pour $x \in \mathbb{D}$, on dit que x est *càdlàg* et on note $\Delta x(t) = x(t) - x(t-)$ pour $0 < t \leq 1$, et $\Delta x(0) = 0$. Les fonctions et processus càdlàg apparaissent naturellement en probabilités :

- fonctions de répartition;
- processus de Poisson;
- processus de Lévy (càdlàg, à accroissements stationnaires et indépendants), qui apparaissent notamment dans l'extension du théorème de Donsker à certaines variables aléatoires de variance infinie;
- toute martingale par rapport à une filtration complète et continue à droite admet une modification càdlàg.

Attention. $(\mathbb{D}, \|\cdot\|_\infty)$ est un espace de Banach, mais il n'est pas séparable.

Prendre $x^{(u)} = \mathbb{1}_{[u, 1]}$: $\|x^{(u)} - x^{(v)}\|_\infty = 1$ pour $u \neq v$.

Il faut chercher une autre topologie sur $\mathbb{D}$.

Références.

- Billingsley, *Convergence of Probability Measures*, deuxième édition.
- Jacod–Shiryaev, *Limit Theorems for Stochastic Processes*, chapitre VI.
- Kallenberg, *Foundations of Modern Probability*, deuxième édition.

Remarque 7.1. On se place dans $\mathbb{D}([0, 1], \mathbb{R}^d)$ pour simplifier, mais ce qui suit s'adapte sans souci au cas de $\mathbb{D}([0, 1], S)$, avec (S, d) polonais.

7.1 Propriétés de base

Exemple 7.2. Soit $(r_i)_{i \geq 1}$ une énumération des rationnels de $]0, 1[$. On a

$$x(t) = \sum_{i=1}^{\infty} \frac{1}{2^i} \mathbb{1}_{\{r_i \leq t\}} \in \mathbb{D}.$$

Proposition 7.3. Soit $x \in \mathbb{D}$.

- (1) Pour tout $\varepsilon > 0$, $\#\{t \in [0, 1] : |\Delta x(t)| > \varepsilon\} < \infty$.
- (2) $\{t \in [0, 1] : \Delta x(t) \neq 0\}$ est dénombrable.
- (3) x est bornée.
- (4) x est limite uniforme de fonctions en escalier.

Idée de preuve. (1) Si l'ensemble est infini, on prend une suite (t_n) dedans. Quitte à extraire, on suppose $t_n \rightarrow t$. On distingue les deux cas $t_n \downarrow t$ et $t_n \uparrow t$ pour une contradiction.

(2) Cela découle de (1).

(3) Par l'absurde : si $|x(t_n)| > n$, on extrait $t_n \rightarrow t$ avec (t_n) monotone, et on obtient $\max(|x(t)|, |x(t-)|) = \infty$, absurde.

(4) On introduit $\omega(x, A) = \sup_{s, t \in A} |x(t) - x(s)|$ pour $A \subset [0, 1]$. On utilise alors le lemme suivant. □

Lemme 7.4. Pour tout $\varepsilon > 0$, il existe $0 = t_0 < t_1 < \dots < t_r = 1$ tel que

$$\max_{0 \leq i \leq r-1} \omega(x, [t_i, t_{i+1}[) \leq \varepsilon.$$

Idée de preuve. Soit $\varepsilon > 0$. On note

$$T = \sup \left\{ t \in [0, 1] : \exists 0 = t_0 < t_1 < \dots < t_r = t, \max_{0 \leq i \leq r-1} \omega(x, [t_i, t_{i+1}[) \leq \varepsilon \right\}.$$

Alors $x(0) = x(0+) \implies T > 0$; l'existence de $x(T-)$ implique que le supremum est atteint; et $x(T) = x(T+)$ implique que $T < 1$ est impossible. □

On conclut la preuve de (4) en prenant

$$x_n = \sum_{i=0}^{r-1} x(t_i) \mathbb{1}_{[t_i, t_{i+1}[} \quad \text{sur } [0, 1[, \quad x_n(1) = x(1),$$

avec une partition associée à $\varepsilon \downarrow 0$. On définit alors le « module de continuité modifié ».

Notation. Pour $x : [0, 1] \rightarrow \mathbb{R}^d$ et $0 < \delta < 1$, on pose

$$\omega'(x, \delta) = \inf_{\substack{r \geq 1, 0=t_0 < t_1 < \dots < t_r=1 \\ t_{i+1}-t_i > \delta}} \max_{0 \leq i \leq r-1} \omega(x, [t_i, t_{i+1}[).$$

Exemple 7.5. Soient $0 < \varepsilon < 1$ et $x_n = n \mathbb{1}_{[1-\varepsilon, 1]}$. Pour $0 < \delta < 1$, on a

$$\omega'(x_n, \delta) = \begin{cases} 0, & \delta < \min(\varepsilon, 1 - \varepsilon), \\ n, & \delta \geq \min(\varepsilon, 1 - \varepsilon). \end{cases}$$

Proposition 7.6. Soient $x, y : [0, 1] \rightarrow \mathbb{R}^d$.

- (1) $x \in \mathbb{D}$ si et seulement si $\omega'(x, \delta) \rightarrow 0$ lorsque $\delta \rightarrow 0$ (à rapprocher de x continue si et seulement si $\omega(x, \delta) \rightarrow 0$).
- (2) $\omega'(x, \delta) \leq 2\|x - y\|_\infty + \omega'(y, \delta)$.
- (3) Si $x \in \mathbb{D}$ et $0 < \delta < 1$, alors

$$(a) \quad \omega'(x, \delta) \leq \omega(x, 2\delta),$$

$$(b) \quad \omega(x, \delta) \leq 2\omega'(x, \delta) + \sup_{0 \leq t \leq 1} |\Delta x(t)|.$$

Preuve laissée en exercice (on a déjà vu le sens direct pour (1)); voir Billingsley, p. 123.

7.2 Topologie de Skorokhod (J_1) sur $\mathbb{D}$

Deux fonctions x et y sont proches pour la norme uniforme si le graphe de $x(t)$ peut être modifié en celui de $y(t)$ par une petite perturbation des ordonnées, en gardant les abscisses fixées. Dans $\mathbb{D}$, on va aussi permettre une petite déformation temporelle.

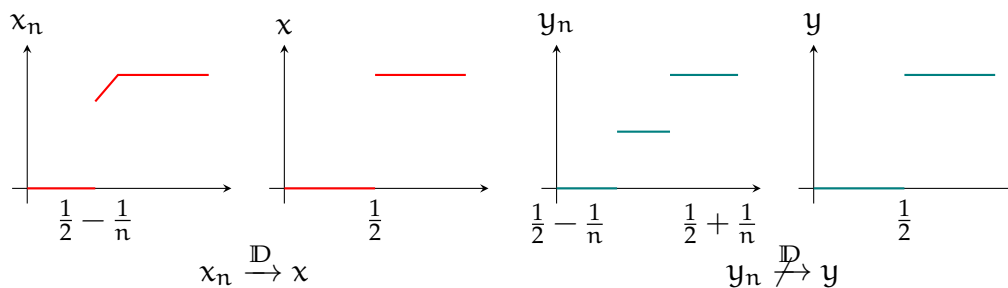
Définition 7.7. Un *changement de temps* est une fonction $\lambda : [0, 1] \rightarrow [0, 1]$ telle que $\lambda(0) = 0$, $\lambda(1) = 1$, et λ est strictement croissante et continue. On note Λ l'ensemble des changements de temps.

Définition 7.8. Soient $x_n \in \mathbb{D}$ et $x \in \mathbb{D}$. On dit que $x_n \rightarrow x$ *au sens de Skorokhod* si, pour tout $n \geq 1$, il existe $\lambda_n \in \Lambda$ tel que

$$\|\lambda_n - \text{Id}\|_\infty \rightarrow 0, \quad \|x_n \circ \lambda_n - x\|_\infty \rightarrow 0.$$

De manière équivalente, on peut demander $\|x_n - x \circ \lambda_n\|_\infty \rightarrow 0$. On note $x_n \xrightarrow{\mathbb{D}} x$.

Exemple 7.9. Les premiers graphes illustrent une convergence dans $\mathbb{D}$: un saut légèrement décalé, de hauteur tendant vers 1, est suivi d'une petite variation continue. Les seconds illustrent l'absence de convergence dans $\mathbb{D}$ lorsque deux sauts de hauteur $1/2$ se rapprochent pour former un saut de hauteur 1.



Remarque 7.10. Pour $\lambda_n, \lambda \in \Lambda$, $\|\lambda_n - \lambda\|_\infty \rightarrow 0$ si et seulement si λ_n converge simplement vers λ (théorème de Dini).

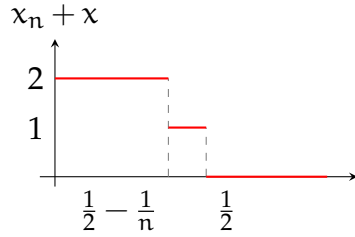
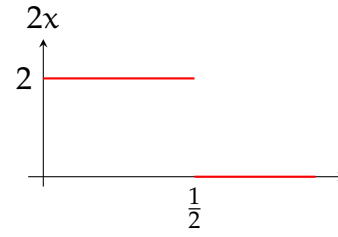
Remarque 7.11. Si on pose, pour $x, y \in \mathbb{D}$,

$$d_S(x, y) = \inf_{\lambda \in \Lambda} \max(\|\lambda - \text{Id}\|_\infty, \|x \circ \lambda - y\|_\infty),$$

on vérifie que d_S est une distance sur $\mathbb{D}$ qui métrise la notion de convergence précédente, et que $(\mathbb{D}, d_S)$ est un espace métrique séparable.

Attention. $(\mathbb{D}, d_S)$ n'est pas une topologie d'espace vectoriel normé : si

$x_n = \mathbb{1}_{[0, 1/2 - 1/n[}$ et $x = \mathbb{1}_{[0, 1/2[}$, on a $x_n \xrightarrow{\mathbb{D}} x$, mais $x_n + x \not\xrightarrow{\mathbb{D}} 2x$.



On vérifie cependant que, si $x_n \xrightarrow{\mathbb{D}} x$ et $y_n \xrightarrow{\mathbb{D}} y$ avec y continue, alors $x_n + y_n \xrightarrow{\mathbb{D}} x + y$.

Pour $t \in [0, 1]$, on note $\Pi_t : \mathbb{D} \rightarrow \mathbb{R}^d$, $x \mapsto x(t)$.

Proposition 7.12 (Exemples de fonctionnelles continues). (1) $x \mapsto \|x\|_\infty$ est continue pour d_S .

(2) $x \mapsto \sup_{0 \leq t \leq 1} |\Delta x(t)|$ est continue pour d_S .

(3) Π_0 et Π_1 sont continues.

(4) Soit $t \in]0, 1[$. Π_t est continue en toute fonction x continue en t .

(5) Soit $f : \mathbb{R}^d \rightarrow \mathbb{R}$ continue. $x \mapsto \int_0^1 f(x(s)) ds$ est continue pour d_S .

Démonstration. Soit $x_n \xrightarrow{\mathbb{D}} x$ et $\lambda_n \in \Lambda$ tels que $x_n \circ \lambda_n \rightarrow x$ uniformément et $\lambda_n \rightarrow \text{Id}$ uniformément.

(1) On a

$$|\|x_n\|_\infty - \|x\|_\infty| = |\|x_n \circ \lambda_n\|_\infty - \|x\|_\infty| \leq \|x_n \circ \lambda_n - x\|_\infty \rightarrow 0.$$

(2) On a $|\Delta x(t)| = |x(t) - x(t-)| \leq 2\|x\|_\infty$, et puisque

$$\sup_{0 \leq t \leq 1} |\Delta x_n(t)| = \sup_{0 \leq t \leq 1} |\Delta(x_n \circ \lambda_n)(t)|,$$

on obtient

$$\left| \sup_{0 \leq t \leq 1} |\Delta x_n(t)| - \sup_{t \in [0, 1]} |\Delta x(t)| \right| \leq \sup_{t \in [0, 1]} |\Delta(x_n \circ \lambda_n - x)(t)| \leq 2\|x_n \circ \lambda_n - x\|_\infty \rightarrow 0.$$

(3) Ceci provient du fait que, pour tout $\lambda \in \Lambda$, on a $\lambda(0) = 0$ et $\lambda(1) = 1$.

(4) On écrit

$$|x_n(t) - x(t)| \leq \|x_n - x \circ \lambda_n^{-1}\|_\infty + |x \circ \lambda_n^{-1}(t) - x(t)| \rightarrow 0$$

par continuité de x en t .

(5) Soit $x_n \xrightarrow{\mathbb{D}} x$. Comme $\sup_n \|x_n\|_\infty \vee \|x\|_\infty < \infty$, on peut supposer f bornée. Posons $D_x = \{t \in [0, 1] : \Delta x(t) \neq 0\}$, dénombrable. Alors

$$\int_a^b f(x_n(t)) dt = \int_a^b \underbrace{f(x_n(t)) \mathbb{1}_{\{t \notin D_x\}}}_{\rightarrow f(x(t)) \mathbb{1}_{\{t \notin D_x\}}} dt.$$

On conclut par convergence dominée. $\square$

Remarque 7.13. En général, les fonctions $f_1 : x \mapsto x(t)$, $f_2 : x \mapsto \omega'(x, \delta)$ et

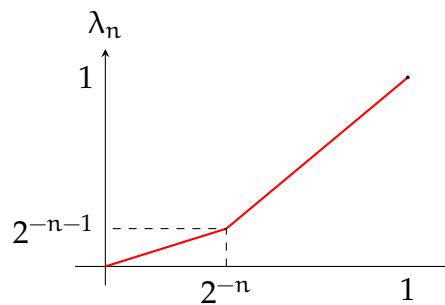
$$f_3 : x \mapsto \sup_{a \leq t \leq b} |\Delta x(t)| \quad (0 \leq a \leq b \leq 1)$$

ne sont pas continues (bien que $x \mapsto \omega(x, \delta)$ soit continue pour $x \in \mathcal{C}([0, 1])$). Par exemple, fixons $0 < \delta < 1/2$ et considérons, pour n assez grand pour que $\delta + 2^{-n} < 1 - \delta$, $x_n = \mathbb{1}_{[0, \delta + 2^{-n}[}$ et $x = \mathbb{1}_{[0, \delta[}$. Alors $x_n \xrightarrow{\mathbb{D}} x$, mais

$$\begin{aligned} x_n(\delta) &= 1 \not\rightarrow x(\delta) = 0, \\ \omega'(x_n, \delta) &= 0 \not\rightarrow \omega'(x, \delta) = 1, \\ x_n(\delta) - x_n(\delta-) &= 0 \not\rightarrow x(\delta) - x(\delta-) = -1. \end{aligned}$$

On peut cependant montrer que f_2 et f_3 sont semi-continues supérieurement.

Exemple 7.14. Soit $x_n = \mathbb{1}_{[0, 2^{-n}[}$. Si on considère le changement de temps λ_n représenté ci-dessous, on voit que $d_S(x_n, x_{n+1}) \leq 2^{-n}$.



En particulier, (x_n) est de Cauchy mais ne converge pas (la seule limite possible est la fonction nulle). Ainsi, $(\mathbb{D}, d_S)$ n'est pas complet.

Cependant :

Théorème 7.15. Il existe une distance d'_S sur $\mathbb{D}$ telle que les topologies engendrées par d_S et d'_S sont les mêmes, et $(\mathbb{D}, d'_S)$ est complet.

Voir Billingsley, théorème 12.2.

7.3 Convergence étroite dans $\mathbb{D}$

Proposition 7.16. La tribu produit sur $\mathbb{D}$ (la plus petite tribu rendant mesurables les applications $\Pi_t : x \mapsto x(t)$) est la tribu borélienne de $(\mathbb{D}, d_S)$.

Pour $t < 1$ et pour les entiers p assez grands pour que $t + 1/p \leq 1$, l'application

$$\Pi_t^p : x \mapsto p \int_t^{t+1/p} x(s) \, ds$$

est continue, donc mesurable. Puisque $\Pi_t = \lim_{p \rightarrow \infty} \Pi_t^p$ par continuité à droite, Π_t est mesurable. Enfin, Π_1 est continue. Pour l'inclusion réciproque, notons $\mathcal{T}$ la tribu produit et posons

$$A_n x(t) = x(k2^{-n}) \quad \text{si } k2^{-n} \leq t < (k+1)2^{-n}, \quad A_n x(1) = x(1).$$

L'application $A_n : (\mathbb{D}, \mathcal{T}) \rightarrow (\mathbb{D}, \mathcal{B}(\mathbb{D}))$ est mesurable : elle ne dépend que d'un nombre fini de coordonnées et l'application qui associe à ces valeurs la fonction en escalier correspondante est continue pour la norme uniforme, donc pour la topologie J_1 .

Vérifions que $A_n x \rightarrow x$ dans $\mathbb{D}$ pour tout $x \in \mathbb{D}$. Fixons $\varepsilon > 0$ et choisissons une partition $0 = t_0 < \dots < t_r = 1$ telle que $\omega(x, [t_i, t_{i+1}[) \leq \varepsilon$ pour tout i . Pour n assez grand, les points $s_i^{(n)} = 2^{-n} \lceil 2^n t_i \rceil$ sont distincts. Soit λ_n le changement de temps affine par morceaux qui envoie chaque t_i sur $s_i^{(n)}$. Alors $\|\lambda_n - \text{Id}\|_\infty \leq 2^{-n}$. Si $t \in [t_i, t_{i+1}[$, la coordonnée de grille utilisée par $A_n x(\lambda_n(t))$ appartient encore à $[t_i, t_{i+1}[$; ainsi

$$\|A_n x \circ \lambda_n - x\|_\infty \leq \varepsilon.$$

Au temps 1, on a aussi $A_n x(\lambda_n(1)) = A_n x(1) = x(1)$. On en déduit $d_S(A_n x, x) \rightarrow 0$. L'identité de $(\mathbb{D}, \mathcal{T})$ dans $(\mathbb{D}, \mathcal{B}(\mathbb{D}))$ est donc une limite ponctuelle d'applications mesurables à valeurs dans un espace métrique, d'où $\mathcal{B}(\mathbb{D}) \subset \mathcal{T}$.

Remarque 7.17. La tribu produit est aussi la plus petite tribu rendant mesurables les applications $(\Pi_t)_{t \in D}$, où $D \subset [0, 1]$ est dense et contient 1. En effet,

$$\Pi_t = \lim_{\substack{s \in D \\ s \downarrow t}} \Pi_s \quad \text{ponctuellement.}$$

Ainsi, les lois fini-dimensionnelles caractérisent les mesures de $\mathcal{M}_1(\mathbb{D})$. Cependant :

Attention. La convergence des lois fini-dimensionnelles n'est pas impliquée par la convergence en loi (les Π_t ne sont pas continues).

Définition 7.18. Soit X une variable aléatoire dans $\mathbb{D}$. On dit que $t \in [0, 1]$ est un *point de discontinuité fixe* de X si $\mathbb{P}(|\Delta X_t| > 0) > 0$. On note D_X l'ensemble de ces points.

Proposition 7.19. Si X est une variable aléatoire dans $\mathbb{D}$, D_X est au plus dénombrable.

Démonstration. Pour $\varepsilon, \delta > 0$, posons

$$D_X^{\varepsilon, \delta} = \{t \in [0, 1] : \mathbb{P}(|\Delta X_t| > \varepsilon) \geq \delta\}.$$

Si cet ensemble était infini, il contiendrait une suite de temps déterministes distincts (t_n) . Chaque trajectoire càdlàg n'a qu'un nombre fini de sauts de taille supérieure à ε ; ainsi

$$\mathbb{1}_{\{|\Delta X_{t_n}| > \varepsilon\}} \longrightarrow 0 \quad \text{presque sûrement.}$$

Par convergence dominée, $\mathbb{P}(|\Delta X_{t_n}| > \varepsilon) \rightarrow 0$, ce qui contredit la minoration par δ . Tous les $D_X^{\varepsilon, \delta}$ sont donc finis, et

$$D_X = \bigcup_{p, q \geq 1} D_X^{1/p, 1/q}$$

est au plus dénombrable. □

Théorème 7.20. Soient X^n, X des variables aléatoires à valeurs dans $\mathbb{D}$. Alors $X^n \xrightarrow{\text{loi}} X$ si et seulement si :

- (i) $X^n \xrightarrow{\text{fidi}} X$ en dehors de $D_X \setminus \{1\}$, c'est-à-dire, pour tous $t_1, \dots, t_k \in ([0, 1] \setminus D_X) \cup \{1\}$,

$$(X_{t_1}^n, \dots, X_{t_k}^n) \xrightarrow{\text{loi}} (X_{t_1}, \dots, X_{t_k});$$

- (ii) (X^n) est tendue dans $\mathcal{M}_1(\mathbb{D})$.

Démonstration. Le sens direct pour (ii) provient du théorème de Prokhorov. Pour (i), si $t \notin D_X$, alors Π_t est continue $\mathbb{P}_X$ -presque partout. L'application Π_1 est quant à elle continue sur tout $\mathbb{D}$, même lorsque $1 \in D_X$. Le théorème de continuité donne donc les convergences fini-dimensionnelles annoncées.

Réciproquement, par le théorème de Prokhorov, il suffit de voir que si $X^{\varphi(n)} \xrightarrow{\text{loi}} Y$, alors $X \stackrel{\text{loi}}{=} Y$. Pour cela, par (i), X et Y ont les mêmes lois fini-dimensionnelles sur $([0, 1] \setminus (D_X \cup D_Y)) \cup \{1\}$, et ont donc la même loi. □

7.4 Critères de tension

L'analogue du théorème d'Arzelà–Ascoli dans $\mathbb{D}$ est le suivant.

Théorème 7.21. Une famille $\mathcal{F} \subset \mathbb{D}$ est relativement compacte si et seulement si

- (1) $\sup_{x \in \mathcal{F}} \|x\|_\infty < \infty$;
- (2) $\forall \varepsilon > 0, \exists \delta > 0, \sup_{x \in \mathcal{F}} \omega'(x, \delta) \leq \varepsilon$.

Comme pour les fonctions continues, on en déduit :

Théorème 7.22. Une suite $(X^n)_{n \geq 1}$ de variables aléatoires dans $\mathbb{D}$ est tendue si et seulement si

- (1) $(\|X^n\|_\infty)_{n \geq 1}$ est tendue dans $\mathbb{R}$;
- (2) $\forall \varepsilon, \eta > 0, \exists \delta > 0, \limsup_{n \rightarrow \infty} \mathbb{P}(\omega'(X^n, \delta) > \eta) \leq \varepsilon.$

Remarque 7.23. On ne peut pas remplacer (1) par « $(X^n(0))_{n \geq 1}$ est tendue », comme dans $\mathcal{C}([0, 1], \mathbb{R})$: prendre par exemple $X^n = n\mathbb{1}_{[U, 1]}$, avec U uniforme sur $[0, 1]$.

On peut cependant démontrer que, lorsque (2) est vérifiée, $(1) \iff (1') \iff (1'')$, avec

(1') Il existe $T \subset [0, 1]$ dense, contenant 1, tel que pour tout $t \in T$, $(X^n(t))_{n \geq 1}$ est tendue.

(1'') Les suites $(X^n(0))_{n \geq 1}$ et $(\sup_{0 \leq t \leq 1} |\Delta X^n(t)|)_{n \geq 1}$ sont tendues.

Voir Billingsley, deuxième édition, corollaire au théorème 13.2, p. 140.

Application. Soient $(U_i)_{i \geq 1}$ des variables aléatoires i.i.d. uniformes sur $[0, 1]$ et $\lambda > 0$. On pose, pour $t \leq 1$,

$$X_t^{(n)} = \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{1}_{\{U_k \leq \lambda/n\}}.$$

Alors $(X_t^{(n)})_{0 \leq t \leq 1} \xrightarrow{\text{loi}} (N_t^\lambda)_{0 \leq t \leq 1}$, où N^λ est un processus de Poisson d'intensité λ .

Idée de preuve. Il suffit de considérer les entiers $n \geq \lambda$. *Lois fini-dimensionnelles.* N^λ n'a pas de discontinuités fixes. Soient $0 = t_0 < t_1 < \dots < t_k \leq 1$. Les accroissements $X_{t_{i+1}}^{(n)} - X_{t_i}^{(n)}$, $0 \leq i < k$, sont indépendants à n fixé et de lois

$$\text{Bin}\left(\lfloor nt_{i+1} \rfloor - \lfloor nt_i \rfloor, \frac{\lambda}{n}\right) \xrightarrow{\text{loi}} \text{Poi}(\lambda(t_{i+1} - t_i)).$$

Tension. D'abord, $\|X^{(n)}\|_\infty \stackrel{\text{loi}}{=} \text{Bin}(n, \lambda/n)$ converge en loi, donc est tendue.

Pour $0 < \delta < 1/2$, notons $A_n(\delta)$ l'événement où deux sauts distincts de $X^{(n)}$ sont espacés d'au plus δ , et $B_n(\delta)$ l'événement où un saut appartient à $]0, \delta] \cup [1 - \delta, 1]$. Sur le complément de $A_n(\delta) \cup B_n(\delta)$, la subdivision formée par 0, les temps de saut et 1 a toutes ses mailles strictement supérieures à δ . Par conséquent, $\omega'(X^{(n)}, \delta) = 0$.

Pour $n \geq \lambda$, l'indépendance des indicatrices donne

$$\mathbb{P}(A_n(\delta)) \leq \sum_{\substack{1 \leq i < j \leq n \\ j-i \leq n\delta}} \mathbb{P}(U_i \leq \lambda/n, U_j \leq \lambda/n) \leq n \lfloor n\delta \rfloor \frac{\lambda^2}{n^2} \leq \lambda^2 \delta.$$

Il y a au plus $2n\delta + 2$ points de la grille dans $]0, \delta] \cup [1 - \delta, 1]$, donc

$$\mathbb{P}(B_n(\delta)) \leq 2\lambda\delta + \frac{2\lambda}{n}.$$

Ainsi, pour tout $\eta > 0$,

$$\limsup_{n \rightarrow \infty} \mathbb{P}(\omega'(X^{(n)}, \delta) > \eta) \leq (\lambda^2 + 2\lambda)\delta \xrightarrow[\delta \downarrow 0]{} 0,$$

ce qui achève la preuve de tension. □

On conclut par plusieurs autres critères de tension. Le premier est un analogue du critère de Kolmogorov adapté aux trajectoires càdlàg (voir Billingsley, deuxième édition, section 13, notamment la preuve du théorème 13.5).

Théorème 7.24 (Critère de Billingsley). Soit $(X^n)_{n \geq 1}$ une suite de variables aléatoires dans $\mathbb{D}$. On suppose que :

- (1) $(\|X^n\|_\infty)_{n \geq 1}$ est tendue ;
- (2) il existe $a, b > 0$ et une fonction continue et croissante $F : [0, 1] \rightarrow \mathbb{R}_+$ telle que, pour tout $n \geq 1$ et tous $0 \leq r \leq s \leq t \leq 1$,

$$\mathbb{E}[|X^n(t) - X^n(s)|^a |X^n(s) - X^n(r)|^a] \leq (F(t) - F(r))^{1+b};$$

- (3) pour tout $\varepsilon > 0$,

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} (\mathbb{P}(|X^n(\delta) - X^n(0)| \geq \varepsilon) + \mathbb{P}(|X^n(1) - X^n(1 - \delta)| \geq \varepsilon)) = 0.$$

Alors $(X^n)_{n \geq 1}$ est tendue.

Remarque 7.25. Les conditions (1) et (2) seules ne suffisent pas : la suite $X^n = \mathbb{1}_{[0,1/n[}$ les vérifie, mais n'est pas tendue dans $\mathbb{D}$. La condition (3) n'est cependant pas nécessaire à toute suite tendue : elle exclut notamment la suite constante $X^n = \mathbb{1}_{\{1\}}$.

Sous l'hypothèse (2), avec F continue, la condition (3) est satisfaite si, pour une variable aléatoire X à valeurs dans $\mathbb{D}$, on a $X^n \xrightarrow{\text{fidi}} X$ sur $[0, 1] \setminus D_X$ et si X est presque sûrement continu en 1. La borne de (2) permet de comparer les incréments aux bords à ceux calculés en des temps voisins hors de D_X ; la continuité de F permet ensuite de faire tendre ces temps vers les extrémités.

On conclut par le critère d'Aldous, particulièrement utile pour les marches aléatoires ou, plus généralement, les martingales.

Pour $X^n \in \mathbb{D}$, on note $(\mathcal{F}_t^n)_{0 \leq t \leq 1}$ sa filtration naturelle et

$$\mathcal{T}^n = \{\text{temps d'arrêt pour } \mathcal{F}^n \text{ qui sont } \leq 1\}.$$

Théorème 7.26 (Aldous). On suppose que

- (1) $(\|X^n\|_\infty)_{n \geq 1}$ est tendue ;
- (2) pour toute suite $h_n \downarrow 0$ et toute suite $\tau_n \in \mathcal{T}^n$,

$$\left| X^n_{(\tau_n + h_n) \wedge 1} - X^n_{\tau_n} \right| \xrightarrow{\mathbb{P}} 0.$$

Alors $(X^n)_{n \geq 1}$ est tendue.

Voir par exemple Kallenberg, théorème 16.11, pour une preuve.

Ce théorème permet par exemple de démontrer le résultat spectaculaire suivant.

Théorème 7.27. Soient $(X_i^n)_{i,n \geq 1}$ des variables aléatoires telles que, pour tout $n \geq 1$, $(X_i^n)_{i \geq 1}$ sont i.i.d. On pose $S_k^n = X_1^n + \dots + X_k^n$ et on suppose qu'il existe une suite d'entiers $m_n \rightarrow \infty$ telle que $S_{m_n}^n$ converge en loi vers une limite notée X .

Alors il existe $(X_t)_{0 \leq t \leq 1} \in \mathbb{D}$ à accroissements stationnaires et indépendants (c'est-à-dire un processus de Lévy) tel que $X_1 \stackrel{\text{loi}}{=} X$ et

$$(S_{[m_n t]}^n : 0 \leq t \leq 1) \xrightarrow{\text{loi}} (X_t)_{0 \leq t \leq 1}.$$

Voir Kallenberg, théorème 16.14.

Si $(X_i)_{i \geq 1}$ sont i.i.d., centrées, de variance finie, en prenant $X_i^n = X_i/\sqrt{n}$ et $m_n = n$, c'est essentiellement le théorème de Donsker.

7.5 Compléments facultatifs

7.5.1 L'espace $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$

On note $\mathbb{D}_m = \mathbb{D}([0, m], \mathbb{R}^d)$, $m > 0$, et d_m la distance de Skorokhod naturellement définie dessus. Soit $\mathbb{D}_\infty = \mathbb{D}([0, \infty[, \mathbb{R}^d)$.

Définition 7.28. On dit que $x_n \xrightarrow{\mathbb{D}_\infty} x$ si, pour tout $n \geq 1$, il existe $\lambda_n \in \Lambda_\infty$, où

$$\Lambda_\infty = \{\text{homéomorphismes croissants de } \mathbb{R}_+ \text{ dans } \mathbb{R}_+\},$$

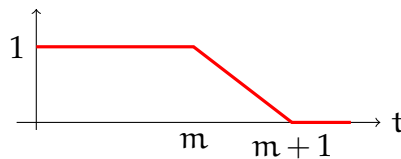
tels que

- 1) $\|\lambda_n - \text{Id}\|_\infty \rightarrow 0$;
- 2) pour tout $m \geq 1$, $\|(x_n \circ \lambda_n - x) \mathbb{1}_{[0, m]}\|_\infty \rightarrow 0$.

Remarque 7.29. L'application de restriction $\mathbb{D}_\infty \rightarrow \mathbb{D}_m$, $x \mapsto x \mathbb{1}_{[0, m]}$, n'est pas continue dans $\mathbb{D}_\infty$ si x a un saut en m .

Pour définir une distance d_∞ telle que $d_\infty(x_n, x) \rightarrow 0$ si et seulement si $x_n \xrightarrow{\mathbb{D}_\infty} x$, on introduit, pour tout $m \geq 1$, la fonction

$$g_m(t) = \begin{cases} 1, & 0 \leq t \leq m, \\ m+1-t, & m \leq t \leq m+1, \\ 0, & t \geq m+1. \end{cases}$$



On pose, pour $x, y \in \mathbb{D}_\infty$,

$$d_\infty(x, y) = \sum_{m \geq 1} 2^{-m} (d_{m+1}((g_m x)|_{[0, m+1]}, (g_m y)|_{[0, m+1]}) \wedge 1).$$

Cette distance métrise la convergence définie ci-dessus. La topologie obtenue est polonaise : elle admet une distance compatible complète et séparable.

Critères de tension. Pour $m \geq 1$ et $0 < \delta < m$, on définit

$$\omega'_m(x, \delta) = \inf_{\substack{0=t_0 < t_1 < \dots < t_r=m \\ t_{i+1}-t_i > \delta, 0 \leq i \leq r-1}} \max_{0 \leq i \leq r-1} \omega(x, [t_i, t_{i+1}[).$$

Une suite (X^n) est tendue dans $\mathbb{D}_\infty$ si et seulement si, pour chaque entier $m \geq 1$, la suite $((g_m X^n)|_{[0, m+1]})$ est tendue dans $\mathbb{D}_{m+1}$. On applique donc les critères précédents avec $\|g_m X^n\|_\infty$ et $\omega'_{m+1}(g_m X^n, \delta)$. Voir Billingsley, section 16.

7.5.2 $\mathcal{C}$ -tension

On remarque que $\mathcal{C} = \mathcal{C}([0, 1], \mathbb{R}^d)$ est un fermé de $\mathbb{D}$, car c'est

$$\left\{ x \in \mathbb{D} : \sup_{0 \leq t \leq 1} |\Delta x(t)| = 0 \right\}.$$

Définition 7.30. Une suite (X^n) de variables aléatoires de $\mathbb{D}$ est dite $\mathcal{C}$ -tendue si elle est tendue dans $\mathbb{D}$ ($\mathbb{D}$ -tendue) et si toute valeur d'adhérence est portée par $\mathcal{C}$: si $X^{\varphi(n)} \xrightarrow{\text{loi}} X$, alors $\mathbb{P}(X \in \mathcal{C}) = 1$.

Proposition 7.31. Soient X^n, X des variables aléatoires dans $\mathbb{D}$ telles que $X^n \xrightarrow{\text{loi}} X$. On suppose que

$$\sup_{0 \leq t \leq 1} |\Delta X^n(t)| \xrightarrow{\mathbb{P}} 0.$$

Alors $\mathbb{P}(X \in \mathcal{C}) = 1$.

Démonstration. L'application

$$J : \mathbb{D} \longrightarrow \mathbb{R}_+, \quad J(x) = \sup_{0 \leq t \leq 1} |\Delta x(t)|$$

est continue, comme démontré dans la proposition sur les fonctionnelles continues. Ainsi $J(X^n) \xrightarrow{\text{loi}} J(X)$. Or $J(X^n) \xrightarrow{\mathbb{P}} 0$ par hypothèse ; par unicité de la limite en loi, $J(X) = 0$ presque sûrement. Une fonction càdlàg sans saut étant continue, on obtient $\mathbb{P}(X \in \mathcal{C}) = 1$. $\square$

Voir aussi Billingsley, théorème 13.4, p. 142.

Théorème 7.32. Soit (X^n) une suite de variables aléatoires de $\mathbb{D}$. Elle est $\mathcal{C}$ -tendue si et seulement si

- (1) $(X^n)_{n \geq 1}$ est tendue dans $\mathbb{R}^d$;

(2) pour tous $\varepsilon, \eta > 0$, il existe $\delta > 0$ tel que

$$\limsup_{n \rightarrow \infty} \mathbb{P}(\omega(X^n, \delta) > \eta) \leq \varepsilon.$$

Voir Billingsley, corollaire p. 142.

On en déduit aisément une version càdlàg du théorème de Donsker : si $(X_i)_{i \geq 1}$ sont i.i.d., centrées, de variance $\sigma^2 \in]0, \infty[$, en notant $S_n = X_1 + \dots + X_n$, on a

$$\left(\frac{S_{[nt]}}{\sigma\sqrt{n}} : 0 \leq t \leq 1 \right) \xrightarrow{\text{loi}} (B_t : 0 \leq t \leq 1),$$

où la convergence a lieu en loi dans $\mathbb{D}$.

7.5.3 Processus ponctuels

Dans le cas des processus ponctuels, la convergence des lois fini-dimensionnelles implique la convergence en loi.

Théorème 7.33. On note $\mathbb{D}_c$ l'ensemble des fonctions de comptage sur $[0, 1]$, c'est-à-dire les fonctions càdlàg, à valeurs dans $\mathbb{Z}$, croissantes, dont les sauts valent 1 aux points de discontinuité. Soient X^n, X des variables aléatoires à valeurs dans $\mathbb{D}_c$. On suppose que $\mathbb{P}(X(1-) = X(1)) = 1$.

Alors $X^n \xrightarrow{\text{loi}} X$ si et seulement s'il existe $T \subset [0, 1]$ dénombrable dense, contenant 0 et 1, tel que, pour tous $t_1, \dots, t_k \in T$,

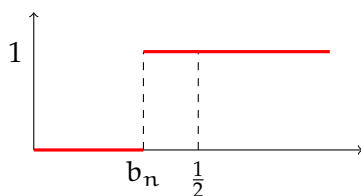
$$(X_{t_1}^n, \dots, X_{t_k}^n) \xrightarrow{\text{loi}} (X_{t_1}, \dots, X_{t_k}).$$

Voir Billingsley, théorème 12.6.

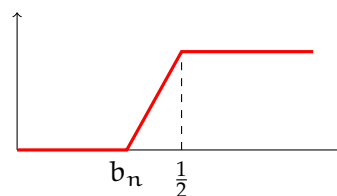
7.5.4 Autres topologies

Nous avons considéré la topologie de Skorokhod J_1 sur $\mathcal{D}([0, 1], \mathbb{R})$, mais il en existe trois autres (J_2 , M_1 , M_2) pour lesquelles les convergences suivantes ont lieu, avec

$$x = \mathbb{1}_{[1/2, 1]}, \quad a_n = \frac{1}{2} - \frac{1}{2n}, \quad b_n = \frac{1}{2} - \frac{1}{n}, \quad c_n = \frac{1}{2} - \frac{3}{2n} \quad (n \geq 4).$$

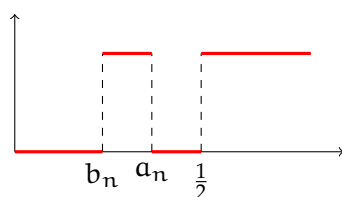


J_1



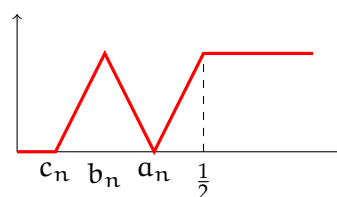
M_1

pas J_1 , pas J_2



J_2

pas J_1 , pas M_1



M_2

pas J_2 , pas M_1 , pas J_1

Voir W. Whitt, *Stochastic Process Limits*, chapitre 11.5.

8 Mesures aléatoires de Poisson

Exercices associés : feuille 7, p. 203.

Dans ce chapitre

8.1	Rappels sur les lois de Poisson	82
8.2	Mesures aléatoires de Poisson	83
8.3	Points multiples	87
8.4	Formules de Campbell (additive et exponentielle)	89
8.5	Transformations de nuages poissonniens	92
8.6	Nuages de Poisson avec paramètre temporel	93
8.7	Formule de Palm	95

Un des objectifs est de définir une notion de « pluie » dans un espace assez général, où les gouttes arrivent indépendamment et rarement selon une répartition décrite par une mesure μ . Lorsque μ est infinie, l'expression « un point choisi au hasard selon μ » n'a pas vraiment de sens. Mais on va pouvoir donner un sens à « une infinité de points choisis selon μ » à travers les mesures aléatoires de Poisson (on verra que les lois de Poisson apparaissent naturellement).

Références. Kingman, *Poisson Processes*; Last et Penrose, *Lectures on the Poisson Process*.

8.1 Rappels sur les lois de Poisson

Pour $\lambda \in [0, \infty]$, on dit qu'une variable aléatoire X à valeurs dans $\mathbb{N} \cup \{+\infty\}$ suit une loi de Poisson de paramètre λ , et on note $X \sim \mathcal{P}(\lambda)$, si

- pour $\lambda = 0$, $X = 0$ presque sûrement ;
- pour $\lambda = \infty$, $X = \infty$ presque sûrement ;
- pour $0 < \lambda < \infty$, $\mathbb{P}(X = k) = e^{-\lambda} \lambda^k / k!$, pour $k \geq 0$.

On a $\mathbb{E}[e^{-uX}] = \exp(-\lambda(1 - e^{-u}))$ pour $u > 0$ (avec la convention $0 \times \infty = 0$). Les deux propriétés suivantes seront bien utiles.

Proposition 8.1. — *Superposition.* Si $(\lambda_i)_{i \in I}$ est une famille au plus dénombrable de $[0, \infty]$ et $(X_i)_{i \in I}$ des variables aléatoires indépendantes avec $X_i \sim \mathcal{P}(\lambda_i)$, alors

$$\sum_{i \in I} X_i \sim \mathcal{P}\left(\sum_{i \in I} \lambda_i\right).$$

- *Raréfaction.* Soit $X \sim \mathcal{P}(\lambda)$ avec $\lambda \in [0, \infty]$. Soient $(U_k)_{k \geq 1}$ des variables aléatoires i.i.d. à valeurs dans $\mathbb{N}$, indépendantes de X , avec $\mathbb{P}(U_k = i) = p_i$ pour $i \geq 1$. Alors les variables

$$X_i = \sum_{k=1}^X \mathbb{1}_{\{U_k=i\}}, \quad i \geq 1,$$

sont indépendantes et $X_i \sim \mathcal{P}(\lambda p_i)$ (avec $0 \times \infty = 0$).

La preuve, qui peut être faite en utilisant les fonctions caractéristiques, est laissée en exercice.

8.2 Mesures aléatoires de Poisson

On se place dans le cadre assez général d'un espace mesuré $(E, \mathcal{E})$.

Définition 8.2. On note $\mathcal{M}_s(E)$ l'ensemble des mesures μ sur E pouvant être écrites sous la forme $\mu = \sum_{i \in I} \mu_i$, avec I dénombrable et μ_i mesure finie. Les éléments de $\mathcal{M}_s(E)$ sont appelés *mesures s-finies*.

Exemple 8.3. Une mesure σ -finie est s-finie, mais la réciproque est fausse (prendre $\mu = \sum_{k=1}^{\infty} \delta_1$). L'avantage de travailler avec $\mathcal{M}_s(E)$ est la stabilité par somme dénombrable (ce n'est pas le cas des mesures σ -finies).

Remarque 8.4. On peut vérifier que les théorèmes de Fubini s'appliquent aux mesures s-finies en utilisant le théorème de convergence monotone (voir le théorème A.13 dans Last et Penrose).

Définition 8.5. On note $\mathcal{M}_{\text{atom}}(E)$ l'ensemble des sommes dénombrables d'atomes de la forme $\sum_{i \in I} \delta_{x_i}$, avec I dénombrable et $(x_i)_{i \in I} \in E^I$. On munit $\mathcal{M}_{\text{atom}}(E)$ de la plus petite tribu rendant mesurables les applications

$$\mathcal{M}_{\text{atom}}(E) \longrightarrow \mathbb{N} \cup \{+\infty\}, \quad \mu \longmapsto \mu(A), \quad A \in \mathcal{E}.$$

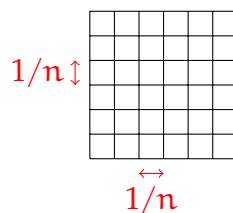
Exemple 8.6. 1) $N : \Omega \rightarrow \mathcal{M}_{\text{atom}}(E)$ est une variable aléatoire si et seulement si, pour tout $A \in \mathcal{E}$ et tout $k \geq 0$, $\{N(A) = k\}$ est mesurable, car

$$\{N(A) = \infty\} = \left(\bigcup_{k \geq 0} \{N(A) = k\} \right)^c.$$

2) L'application $E \rightarrow \mathcal{M}_{\text{atom}}(E)$, $x \mapsto \delta_x$, est mesurable.

3) Pour $A \in \mathcal{E}$, l'application $\mathcal{M}_{\text{atom}}(E) \rightarrow \mathcal{M}_{\text{atom}}(E)$, $\mu \mapsto \mu|_A$, est mesurable, où $\mu|_A(B) = \mu(A \cap B)$ pour $B \in \mathcal{E}$.

Avant de donner la définition de mesure aléatoire de Poisson, une motivation : si on veut définir une pluie « uniforme » sur $\mathbb{R}^2$, il est naturel de discrétiser $\mathbb{R}^2$ par un quadrillage de maille $1/n$ et de choisir un point dans chaque cellule, indépendamment, avec probabilité $1/n^2$.



Alors le nombre de points qui tombent dans une région A est approximativement de loi $\text{Bin}(n^2 \text{Aire}(A), 1/n^2)$, qui converge en loi vers $\mathcal{P}(\text{Aire}(A))$, et les nombres de points qui tombent dans des régions disjointes sont indépendants.

Définition 8.7. Soient $(E, \mathcal{E})$ un espace mesuré et μ une mesure s -finie sur E . Une *mesure aléatoire de Poisson* (ou *nuage poissonien*) d'intensité μ est une variable aléatoire N à valeurs dans $\mathcal{M}_{\text{atom}}(E)$ telle que

- ① pour tout $A \in \mathcal{E}$, $N(A) \sim \mathcal{P}(\mu(A))$;
- ② si $(A_i)_{1 \leq i \leq k}$ est une famille de parties mesurables disjointes de E , les variables $(N(A_i))_{1 \leq i \leq k}$ sont indépendantes.

Remarque 8.8. — Pour chaque A fixé, presque sûrement, $N(A) = \infty \iff \mu(A) = \infty$.

— Si les singletons sont mesurables dans E et si $\mu(\{x\}) > 0$, alors $\mathbb{P}(N(\{x\}) \geq 2) > 0$.

— $\mathbb{E}[e^{-\lambda N(A)}] = e^{-\mu(A)(1-e^{-\lambda})}$.

Théorème 8.9. Pour toute mesure $\mu \in \mathcal{M}_s(E)$, il existe un nuage poissonien d'intensité μ , unique en loi dans $\mathcal{M}_{\text{atom}}(E)$.

Pour démontrer ce théorème, on a recours au lemme utile suivant.

Lemme 8.10 (Stabilité par somme). Soit $(N_i)_{i \geq 1}$ une famille dénombrable de nuages poissoniens indépendants d'intensités μ_i . Alors $N = \sum_{i \geq 1} N_i$ est un nuage poissonien d'intensité $\mu = \sum_{i \geq 1} \mu_i$.

Démonstration. Soit $A \in \mathcal{E}$. Par le principe de superposition des lois de Poisson,

$$N(A) \sim \mathcal{P}\left(\sum_{i \geq 1} \mu_i(A)\right) = \mathcal{P}(\mu(A)).$$

Si $(A_j)_{1 \leq j \leq k}$ sont des parties mesurables disjointes, les variables $(N_i(A_j))_{i \geq 1, 1 \leq j \leq k}$ sont indépendantes. Par regroupement par paquets, les variables $(\sum_{i=1}^{\infty} N_i(A_j))_{1 \leq j \leq k}$ sont indépendantes. $\square$

Preuve de l'existence des nuages poissoniens. D'après le lemme précédent, il suffit de traiter le cas où μ est finie. Si $\mu(E) = 0$, le nuage vide $M = 0$ convient. Supposons donc $0 < \mu(E) < \infty$. Soient $L \sim \mathcal{P}(\mu(E))$ et $(X_i)_{i \geq 1}$ des variables aléatoires i.i.d. de loi $\mu(\cdot)/\mu(E)$, indépendantes de L . On pose

$$M = \sum_{i=1}^L \delta_{X_i}.$$

Vérifions que M convient. C'est bien une variable aléatoire (voir l'exemple de mesurabilité de $x \mapsto \delta_x$).

Soient $(A_j)_{1 \leq j \leq k}$ des parties mesurables disjointes de E et posons $B = (\bigcup_{j=1}^k A_j)^c$. On définit, pour $i \geq 1$,

$$U_i = \begin{cases} j & \text{si } X_i \in A_j, \\ 0 & \text{si } X_i \in B. \end{cases}$$

Les variables $(U_i)_{i \geq 1}$ sont i.i.d. D'après le principe de raréfaction pour les lois de Poisson, les variables

$$M(A_j) = \sum_{i=1}^L \mathbb{1}_{\{X_i \in A_j\}} = \sum_{i=1}^L \mathbb{1}_{\{U_i=j\}}$$

sont indépendantes et $M(A_j) \sim \mathcal{P}(\mu(E)\mu(A_j)/\mu(E)) = \mathcal{P}(\mu(A_j))$. $\square$

Preuve de l'unicité en loi des nuages poissonniens. Les parties de la forme

$$\{\nu \in \mathcal{M}_{\text{atom}}(E) : \nu(A_1) = k_1, \dots, \nu(A_\ell) = k_\ell\},$$

avec A_i mesurables et $k_i \geq 0$, forment un π -système générateur de la tribu de $\mathcal{M}_{\text{atom}}(E)$. Or une telle partie s'écrit comme réunion finie disjointe de parties de la forme

$$\{\nu \in \mathcal{M}_{\text{atom}}(E) : \nu(B_1) = j_1, \dots, \nu(B_n) = j_n\}, \quad (7)$$

avec (B_i) mesurables disjoints et $j_i \geq 0$. Ainsi une mesure de probabilité sur $\mathcal{M}_{\text{atom}}(E)$ est caractérisée par ses valeurs sur les ensembles de la forme (7).

Si deux nuages poissonniens vérifient les deux conditions de la définition pour la même intensité, leurs lois sont égales sur les ensembles (7), donc sont les mêmes. $\square$

Remarque 8.11. La preuve précédente montre que, si μ est s-finie, on peut construire un nuage poissonnien sous la forme

$$M = \sum_{i=1}^K \delta_{X_i}, \quad (8)$$

où K prend ses valeurs dans $\mathbb{N} \cup \{+\infty\}$ et les X_i sont des variables aléatoires à valeurs dans E . Lorsque $\mu(E) = \infty$, on a $K = \infty$ presque sûrement. Les positions X_i ne sont en général ni indépendantes ni de même loi.

Si M est un nuage poissonnien, en général il ne s'écrit pas sous la forme (8), car on ne peut en général pas énumérer les atomes de manière mesurable.

En effet, soit $E = [0, 1]$ et supposons qu'il existe des applications mesurables $\pi_i : \mathcal{M}_{\text{atom}}(E) \rightarrow E$ et $L : \mathcal{M}_{\text{atom}}(E) \rightarrow \mathbb{N} \cup \{+\infty\}$ telles que

$$\forall \nu \in \mathcal{M}_{\text{atom}}(E), \quad \nu = \sum_{i=1}^{L(\nu)} \delta_{\pi_i(\nu)}.$$

Soient λ la mesure de Lebesgue sur $[0, 1]$, $\mu = \sum_{n \geq 1} \lambda$ et M un nuage poissonnien d'intensité μ . Pour chaque borélien A fixé, $M(A)$ est presque sûrement égal à 0 si

$\lambda(A) = 0$ et à ∞ sinon. La tribu $\sigma(M)$ est donc constituée d'événements de probabilité 0 ou 1 : ses générateurs ont cette propriété, et les événements de probabilité 0 ou 1 forment une tribu.

Comme $M(E) = \infty$ presque sûrement, on a $L(M) = \infty$ presque sûrement. La variable $Z = \pi_1(M) \in [0, 1]$ est indépendante d'elle-même, donc

$$\text{Var}(Z) = \mathbb{E}[Z^2] - \mathbb{E}[Z]^2 = 0.$$

Ainsi $Z = c$ presque sûrement pour un $c \in [0, 1]$ déterministe. L'énumération impose alors $M(\{c\}) \geq 1$ presque sûrement, alors que $\mu(\{c\}) = 0$ impose $M(\{c\}) = 0$ presque sûrement. C'est une contradiction : des hypothèses additionnelles de « finitude » sont donc nécessaires.

Si E est un sous-ensemble mesurable d'un espace métrique polonais, en notant $\mathcal{M}_{\text{atom}}^{<\infty}(E)$ les mesures finies de $\mathcal{M}_{\text{atom}}(E)$, il est possible d'énumérer de manière mesurable les atomes des mesures de $\mathcal{M}_{\text{atom}}^{<\infty}(E)$ (proposition 6.2 dans Last et Penrose). On en déduit que, si M est un nuage poissonien d'intensité μ σ -finie sur E polonais, alors on peut écrire

$$M = \sum_{i=1}^{L(M)} \delta_{\pi_i(M)},$$

avec L et les π_i mesurables (corollaire 6.5 dans Last et Penrose).

Conséquence importante. Si on souhaite démontrer une propriété qui ne dépend que de la loi de M , alors on peut supposer sans perte de généralité que M est de la forme (8), construite comme dans la preuve de l'existence des nuages poissoniens.

Exemple 8.12 (Un processus de Poisson est un nuage poissonien). Soient $\lambda > 0$ et $(X_i)_{i \geq 1}$ des variables aléatoires i.i.d. de loi $\text{Exp}(\lambda)$. On pose

$$M = \sum_{k=1}^{\infty} \delta_{X_1 + \dots + X_k}.$$

Alors M est un nuage poissonien d'intensité $\lambda \mathbb{1}_{\{x \geq 0\}} dx$.

Démonstration. Posons $S_n = X_1 + \dots + X_n$. La loi forte donne $S_n/n \rightarrow 1/\lambda$ presque sûrement, donc $S_n \rightarrow \infty$. Fixons $T > 0$ et posons $N_T = M([0, T])$. Pour $n \geq 1$, le changement de variables $s_i = x_1 + \dots + x_i$, puis l'indépendance de X_{n+1} , donnent

$$\begin{aligned} \mathbb{P}(N_T = n, S_1 \in ds_1, \dots, S_n \in ds_n) \\ &= \lambda^n e^{-\lambda s_n} e^{-\lambda(T-s_n)} \mathbb{1}_{\{0 < s_1 < \dots < s_n < T\}} ds_1 \cdots ds_n \\ &= \left(e^{-\lambda T} \frac{(\lambda T)^n}{n!} \right) \left(\frac{n!}{T^n} \mathbb{1}_{\{0 < s_1 < \dots < s_n < T\}} \right) ds_1 \cdots ds_n. \end{aligned}$$

Le second facteur est la densité des statistiques d'ordre de n variables uniformes indépendantes sur $[0, T]$. Le cas $n = 0$ donne $\mathbb{P}(N_T = 0) = \mathbb{P}(X_1 > T) = e^{-\lambda T}$. Ainsi $N_T \sim \mathcal{P}(\lambda T)$ et, conditionnellement à $N_T = n$, les positions ordonnées ont la loi de ces

statistiques d'ordre. La construction à intensité finie montre donc que $M|_{[0,T]}$ est un nuage poissonien d'intensité $\lambda \mathbb{1}_{[0,T]}(t) dt$.

Pour conclure, soient $A_1, \dots, A_r$ des boréliens disjoints de $\mathbb{R}_+$ et $k_1, \dots, k_r \in \mathbb{N}$. Comme $M(A_j \cap [0, T]) \uparrow M(A_j)$, la continuité décroissante des probabilités donne

$$\mathbb{P}(M(A_j) \leq k_j \text{ pour tout } j) = \lim_{T \rightarrow \infty} \prod_{j=1}^r e^{-\theta_{j,T}} \sum_{\ell=0}^{k_j} \frac{\theta_{j,T}^\ell}{\ell!}, \quad \theta_{j,T} = \lambda \text{Leb}(A_j \cap [0, T]).$$

Chaque facteur tend vers la fonction de répartition en k_j d'une loi $\mathcal{P}(\lambda \text{Leb}(A_j))$; il tend vers zéro si ce paramètre est infini. Les mêmes égalités s'appliquent à toute sous-famille des A_j . Elles donnent les lois de Poisson des comptes et leur indépendance, y compris lorsque certains comptes sont presque sûrement infinis. Ceci conclut. $\square$

Proposition 8.13 (Stabilité par restriction). Soit M un nuage poissonien d'intensité μ . Soient $(A_i)_{i \in I}$ des parties mesurables disjointes de E . Alors les $(M|_{A_i})_{i \in I}$ sont des nuages poissoniens indépendants, et $M|_{A_i}$ est d'intensité $\mu|_{A_i}$.

C'est une conséquence immédiate de la définition.

On en déduit le résultat suivant, important en pratique pour se ramener au cas d'une mesure d'intensité finie.

Corollaire 8.14. Soit M une mesure aléatoire de Poisson d'intensité μ supposée σ -finie. Alors on peut écrire $M = \sum_{i=1}^{\infty} M_i$, avec $(M_i)_{i \geq 1}$ des mesures aléatoires de Poisson indépendantes d'intensités finies.

Démonstration. On écrit $E = \bigcup_{i=1}^{\infty} A_i$ avec $(A_i)_{i \geq 1}$ des éléments disjoints de $\mathcal{E}$ de μ -mesure finie. Alors $M = \sum_{i=1}^{\infty} M|_{A_i}$, avec $(M|_{A_i})_{i \geq 1}$ des mesures aléatoires de Poisson indépendantes d'intensités $\mu|_{A_i}$ finies. $\square$

Remarque 8.15. Si μ est s -finie, il n'est pas clair que l'on peut décomposer M de manière « mesurable » en $M = \sum_{i=1}^{\infty} M_i$, avec $(M_i)_{i \geq 1}$ des mesures aléatoires de Poisson indépendantes d'intensités finies. Mais lorsque μ est σ -finie, on vient de voir que c'est possible.

8.3 Points multiples

On suppose que $\{x\} \in \mathcal{E}$ pour tout $x \in E$. Les atomes d'une mesure $\sum_{i \in I} \delta_{x_i} \in \mathcal{M}_{\text{atom}}(E)$ ne sont pas forcément distincts.

Si M est un nuage poissonien d'intensité μ , $M(\{x\}) \sim \mathcal{P}(\mu(\{x\}))$. Donc

$$\forall x \in E, \mu(\{x\}) = 0 \iff \forall x \in E, \mathbb{P}(M(\{x\}) \in \{0, 1\}) = 1.$$

On va en fait voir que

$$\forall x \in E, \mu(\{x\}) = 0 \iff \mathbb{P}(\forall x \in E, M(\{x\}) \in \{0, 1\}) = 1,$$

sous de bonnes hypothèses garantissant la mesurabilité de l'événement.

Définition 8.16. On note

$$\mathcal{M}_{\text{atom}}^{\text{simple}}(E) = \left\{ \mu = \sum_{i \in I} \delta_{x_i} \in \mathcal{M}_{\text{atom}}(E) : x_i \neq x_j \text{ pour } i \neq j \right\}.$$

Rappelons que $\mathcal{M}_{\text{atom}}(E)$ est muni de la plus petite tribu pour laquelle les applications $\mu \mapsto \mu(A)$, à valeurs dans $\mathbb{N} \cup \{+\infty\}$, sont mesurables pour tout $A \in \mathcal{E}$.

Remarque 8.17. En général, $\mathcal{M}_{\text{atom}}^{\text{simple}}(E)$ n'est pas une partie mesurable de $\mathcal{M}_{\text{atom}}(E)$. En effet, soient $(X_i)_{i \geq 1}$ des variables aléatoires i.i.d. uniformes sur $[0, 1] = E$. Posons $M_1 = \sum_{n \geq 1} \delta_{X_n}$ et $M_2 = \sum_{n \geq 1} 2\delta_{X_n}$. Alors M_1 et M_2 sont deux mesures aléatoires de Poisson d'intensité s-finie μ donnée par

$$\mu(A) = \begin{cases} 0 & \text{si } \lambda(A) = 0, \\ \infty & \text{si } \lambda(A) > 0, \end{cases}$$

avec λ la mesure de Lebesgue sur $[0, 1]$. La mesure μ est s-finie car $\mu = \sum_{n=1}^{\infty} \lambda$. Les variables M_1 et M_2 ont même loi, $M_1 \in \mathcal{M}_{\text{atom}}^{\text{simple}}(E)$ presque sûrement et $M_2 \notin \mathcal{M}_{\text{atom}}^{\text{simple}}(E)$; donc $\mathcal{M}_{\text{atom}}^{\text{simple}}(E)$ n'est pas mesurable.

Des hypothèses de finitude et de mesurabilité sont nécessaires.

Lemme 8.18. On suppose que $D = \{(x, x) : x \in E\} \in \mathcal{E} \otimes \mathcal{E}$. Alors

$$\mathcal{M}_{\text{atom}}^{\text{simple},f}(E) = \{\mu \in \mathcal{M}_{\text{atom}}^{\text{simple}}(E) : \mu(E) < \infty\}$$

est mesurable dans $\mathcal{M}_{\text{atom}}(E)$.

Démonstration. Posons $\mathcal{M}_{\text{atom}}^f(E) = \{\mu \in \mathcal{M}_{\text{atom}}(E) : \mu(E) < \infty\}$, partie mesurable de $\mathcal{M}_{\text{atom}}(E)$, et

$$\Phi : \mathcal{M}_{\text{atom}}^f(E) \longrightarrow \mathcal{M}_{\text{atom}}^f(E \times E), \quad \sum_{i=1}^m \delta_{x_i} \longmapsto \sum_{\substack{1 \leq i, j \leq m \\ i \neq j}} \delta_{(x_i, x_j)}.$$

Puisque $\mathcal{M}_{\text{atom}}^{\text{simple},f}(E) = \{\mu \in \mathcal{M}_{\text{atom}}^f(E) : \Phi(\mu)(D) = 0\}$, il suffit de vérifier que $\mu \mapsto \Phi(\mu)(D)$ est mesurable. Pour cela, considérons

$$\mathcal{C} = \{C \in \mathcal{E} \otimes \mathcal{E} : \mu \mapsto \Phi(\mu)(C) \text{ est mesurable sur } \mathcal{M}_{\text{atom}}^f(E)\}.$$

Pour μ finie et $A, B \in \mathcal{E}$, $\Phi(\mu)(A \times B) = \mu(A)\mu(B) - \mu(A \cap B)$. On vérifie que $\mathcal{C}$ est une classe monotone contenant les pavés mesurables. Ceci conclut. On peut de même montrer que Φ est mesurable. $\square$

Proposition 8.19. On suppose que μ est σ -finie et que $D = \{(x, x) : x \in E\} \in \mathcal{E} \otimes \mathcal{E}$. Soit M un nuage poissonien d'intensité μ . Alors

$$\forall x \in E, \mu(\{x\}) = 0 \iff M \in \mathcal{M}_{\text{atom}}^{\text{simple}}(E) \text{ presque sûrement.}$$

Cela signifie qu'il existe un événement A tel que $\mathbb{P}(A) = 1$ et $\omega \in A \Rightarrow M(\omega) \in \mathcal{M}_{\text{atom}}^{\text{simple}}(E)$. Si $\mathbb{P}$ est complète (tout sous-ensemble d'un événement de probabilité nulle est mesurable), c'est donc équivalent à $\mathbb{P}(M \in \mathcal{M}_{\text{atom}}^{\text{simple}}(E)) = 1$. Le cas échéant, on dit que M est *sans point multiple*.

On rappelle qu'en général $\mathcal{M}_{\text{atom}}^{\text{simple}}(E)$ n'est pas mesurable dans $\mathcal{M}_{\text{atom}}(E)$.

Démonstration. Pour le sens réciproque, supposons que $\mu(\{x\}) = c > 0$ pour un $x \in E$. Alors

$$\mathbb{P}(M(\{x\}) \geq 2) = 1 - e^{-c} - ce^{-c} > 0,$$

ce qui contredit l'absence presque sûre de points multiples.

Pour le sens direct, traitons d'abord le cas μ finie. Si $\mu(E) = 0$, alors $M(E) = 0$ presque sûrement, donc M est simple presque sûrement. Supposons désormais $\mu(E) > 0$. D'après le lemme, $\mathbb{P}(M \in \mathcal{M}_{\text{atom}}^{\text{simple},f}(E))$ ne dépend que de la loi de M . Ainsi, on peut supposer que M est construit à partir de variables aléatoires comme dans la preuve de l'existence des nuages poissoniens :

$$M = \sum_{i=1}^L \delta_{X_i},$$

avec $L \sim \mathcal{P}(\mu(E))$ et $(X_i)_{i \geq 1}$ i.i.d. de loi $\mu(\cdot)/\mu(E)$, indépendantes de L . Alors, pour $1 \leq i \neq j$,

$$\mathbb{P}(X_i = X_j) = \frac{1}{\mu(E)^2} \iint_{E^2} \mathbb{1}_{\{x=y\}} \mu(dx) \mu(dy) = 0.$$

Donc $\mathbb{P}(M \in \mathcal{M}_{\text{atom}}^{\text{simple},f}(E)) = 1$.

Dans le cas général σ -fini, on écrit $E = \bigcup_{i=1}^{\infty} A_i$ avec les A_i disjoints et $\mu(A_i) < \infty$ pour tout $i \geq 1$. On a vu que $M = \sum_{i=1}^{\infty} M|_{A_i}$, où les $(M|_{A_i})_{i \geq 1}$ sont des nuages poissoniens indépendants d'intensités respectives $\mu|_{A_i}$, donc finies. Soit alors

$$A = \bigcap_{i=1}^{\infty} \{M|_{A_i} \in \mathcal{M}_{\text{atom}}^{\text{simple},f}(E)\}.$$

D'après le cas fini, $\mathbb{P}(A) = 1$. De plus $\omega \in A \Rightarrow M(\omega) \in \mathcal{M}_{\text{atom}}^{\text{simple}}(E)$. □

8.4 Formules de Campbell (additive et exponentielle)

Dans toute cette partie, $\mu \in \mathcal{M}_s(E)$ est fixée et M est un nuage poissonien d'intensité μ , avec un espace mesuré $(E, \mathcal{E})$ quelconque.

Théorème 8.20 (Formule additive). 1) *Cas positif.* Soit $f : E \rightarrow \mathbb{R}_+$ mesurable. Alors la variable aléatoire $M(f) = \int_E f(x)M(dx)$ a pour espérance

$$\mathbb{E}[M(f)] = \int_E f(x)\mu(dx).$$

2) *Cas intégrable.* Soit $f : E \rightarrow \mathbb{R}$, μ -intégrable. Alors $M(f) = \int_E f(x)M(dx)$ est intégrable et

$$\mathbb{E}[M(f)] = \int_E f(x)\mu(dx).$$

Démonstration. 1) On écrit $f \geq 0$ comme limite croissante de fonctions étagées $f_n = \sum_{i=1}^{N_n} c_i \mathbb{1}_{A_i}$. En particulier,

$$\mathbb{E}[M(f_n)] = \sum_{i=1}^{N_n} c_i \mu(A_i) = \mu(f_n).$$

Puis $M(f_n) \uparrow M(f)$. Donc $M(f)$ est mesurable et, par convergence monotone, $\mathbb{E}[M(f)] = \mu(f)$.

2) On écrit $f = f^+ - f^-$ et on applique 1) à f^+ et f^- . □

Théorème 8.21 (Formule exponentielle). 1) *Cas positif.* Soit $f : E \rightarrow \mathbb{R}_+$ mesurable. On a

$$\mathbb{E}[e^{-M(f)}] = \exp\left(-\int_E (1 - e^{-f(x)})\mu(dx)\right).$$

2) *Cas intégrable.* Soit $f : E \rightarrow \mathbb{R}$, μ -intégrable. On a

$$\mathbb{E}[e^{iM(f)}] = \exp\left(\int_E (e^{if(x)} - 1)\mu(dx)\right).$$

Démonstration. 1) On écrit $f \geq 0$ comme limite croissante de fonctions étagées $f_n = \sum_{i=1}^{N_n} c_i \mathbb{1}_{A_i}$, avec les A_i disjoints. Alors $M(f_n) = \sum_{i=1}^{N_n} c_i M(A_i)$ est une combinaison linéaire de variables de Poisson indépendantes, et on conclut en utilisant la forme de la transformée de Laplace d'une variable de Poisson et la convergence monotone.

2) On écrit $f = \lim_{n \rightarrow \infty} f_n$, avec $f = f^+ - f^-$, $f_n = f_n^+ - f_n^-$, $0 \leq f_n^+ \uparrow f^+$, $0 \leq f_n^- \uparrow f^-$ et f_n^+ , f_n^- étagées. Comme en 1), un calcul explicite donne

$$\mathbb{E}[e^{iM(f_n)}] = \exp\left(\int_E (e^{if_n(x)} - 1)\mu(dx)\right).$$

D'une part, le membre de droite converge vers $\exp(\int_E (e^{if(x)} - 1)\mu(dx))$ par convergence dominée, en utilisant

$$|e^{if_n(x)} - 1| \leq |f_n(x)| = f_n^+(x) + f_n^-(x) \leq f^+(x) + f^-(x) = |f(x)|.$$

D'autre part, $M(f_n) = M(f_n^+) - M(f_n^-)$, avec $M(f_n^+) \uparrow M(f^+)$ et $M(f_n^-) \uparrow M(f^-)$ par convergence monotone. Donc $M(f_n) \rightarrow M(f)$. Comme $|e^{iM(f_n)}| = 1$, on a aussi $\mathbb{E}[e^{iM(f_n)}] \rightarrow \mathbb{E}[e^{iM(f)}]$ par convergence dominée, ce qui conclut. $\square$

Remarque 8.22 (Caractérisation par la transformée de Laplace). Soit μ une mesure s -finie. Une variable aléatoire N à valeurs dans $\mathcal{M}_{\text{atom}}(E)$ est un nuage poissonien d'intensité μ si et seulement si, pour toute fonction mesurable $f : E \rightarrow \mathbb{R}_+$,

$$\mathbb{E}[e^{-N(f)}] = \exp \left(- \int_E (1 - e^{-f(x)}) \mu(dx) \right).$$

Il suffit d'utiliser les fonctions $f = \sum_{i=1}^n c_i \mathbb{1}_{A_i}$, avec $c_i \geq 0$ et les A_i mesurables disjoints, puis l'unicité de la transformée de Laplace jointe.

Si μ est σ -finie, on peut aussi utiliser la caractérisation suivante : pour toute fonction réelle μ -intégrable f , on a $N(|f|) < \infty$ presque sûrement et

$$\mathbb{E}[e^{iN(f)}] = \exp \left(\int_E (e^{if(x)} - 1) \mu(dx) \right).$$

On applique d'abord cette identité aux combinaisons linéaires d'indicatrices d'ensembles de μ -mesure finie, puis on épuise E par de tels ensembles. Cette seconde caractérisation ne s'étend pas aux mesures seulement s -finies.

Pour $f : E \rightarrow \mathbb{R}_+$ mesurable, d'après la formule additive,

$$\mathbb{E} \left[\int f(x) M(dx) \right] < \infty \iff \int f d\mu < \infty.$$

Le résultat suivant donne une condition nécessaire et suffisante pour que $\int f(x) M(dx) < \infty$ presque sûrement.

Proposition 8.23. Soit $f : E \rightarrow \mathbb{R}_+$ mesurable. On a l'alternative :

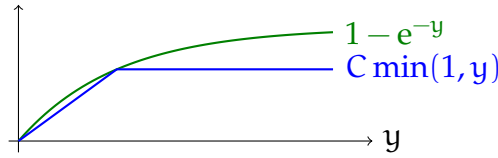
- (a) si $\int_E \min(1, f) d\mu < \infty$, alors presque sûrement $\int_E f(x) M(dx) < \infty$;
- (b) si $\int_E \min(1, f) d\mu = \infty$, alors presque sûrement $\int_E f(x) M(dx) = \infty$.

Démonstration. On remarque que, pour $x \in [0, \infty]$, $x < \infty \iff \lim_{\lambda \downarrow 0} e^{-\lambda x} = 1$. Ainsi, par la formule exponentielle,

$$\begin{aligned} \mathbb{P} \left(\int f(x) M(dx) < \infty \right) &= \lim_{\lambda \downarrow 0} \mathbb{E}[e^{-\lambda M(f)}] \\ &= \lim_{\lambda \downarrow 0} \exp \left(- \int (1 - e^{-\lambda f(x)}) \mu(dx) \right). \end{aligned}$$

Si $\int_E \min(1, f) d\mu < \infty$, pour $\lambda \leq 1$, on a $1 - e^{-\lambda f(x)} \leq \min(1, f(x))$, et le résultat s'ensuit par convergence dominée.

Si $\int_E \min(1, f) d\mu = \infty$, on remarque que $C \min(1, y) \leq 1 - e^{-y}$ pour une certaine constante $C = 1 - e^{-1}$.



Ainsi, pour $0 < \lambda < 1$,

$$\int (1 - e^{-\lambda f(x)}) \mu(dx) \geq C \int \min(1, \lambda f(x)) \mu(dx) \geq \lambda C \int \min(1, f(x)) \mu(dx) = \infty.$$

Donc $\mathbb{P}(\int f(x) M(dx) < \infty) = 0$. □

Exemple 8.24. Si $\mu(dx) = x^{-3/2} \mathbb{1}_{\{x>0\}} dx$ et $f(x) = x$, alors $M(f) < \infty$ presque sûrement, mais $\mathbb{E}[M(f)] = \infty$.

8.5 Transformations de nuages poissonniens

Proposition 8.25 (Stabilité par image). Soient M un nuage poissonnien d'intensité μ s -finie sur $(E, \mathcal{E})$ et $f : (E, \mathcal{E}) \rightarrow (F, \mathcal{G})$ mesurable. Alors $f_* M$ est un nuage poissonnien d'intensité $f_* \mu$ sur F .

Démonstration. Tout d'abord, vérifions que f_* est l'application

$$\mathcal{M}_{\text{atom}}(E) \longrightarrow \mathcal{M}_{\text{atom}}(F), \quad \sum_{i \in I} \delta_{x_i} \longmapsto \sum_{i \in I} \delta_{f(x_i)},$$

et qu'elle est mesurable. Pour $B \in \mathcal{G}$, on a bien

$$\left(\sum_{i \in I} \delta_{f(x_i)} \right) (B) = \#\{i \in I : x_i \in f^{-1}(B)\} = \left(\sum_{i \in I} \delta_{x_i} \right) (f^{-1}(B)).$$

Pour $B \in \mathcal{G}$ et $k \geq 0$,

$$\begin{aligned} f_*^{-1}(\{\nu \in \mathcal{M}_{\text{atom}}(F) : \nu(B) = k\}) &= \{\nu \in \mathcal{M}_{\text{atom}}(E) : (f_* \nu)(B) = k\} \\ &= \{\nu \in \mathcal{M}_{\text{atom}}(E) : \nu(f^{-1}(B)) = k\}, \end{aligned}$$

qui est mesurable. Ainsi $f_* M \in \mathcal{M}_{\text{atom}}(F)$ et, si $B \in \mathcal{G}$,

$$(f_* M)(B) = M(f^{-1}(B)) \sim \mathcal{P}(\mu(f^{-1}(B))) = \mathcal{P}((f_* \mu)(B)).$$

Par ailleurs, si $B_1, \dots, B_k \in \mathcal{G}$ sont disjoints, leurs préimages $f^{-1}(B_1), \dots, f^{-1}(B_k)$ sont disjoints, donc les $(f_* M)(B_i) = M(f^{-1}(B_i))$ sont indépendantes. □

Remarque 8.26. Un autre avantage de travailler avec des mesures s -finies ici plutôt que σ -finies est que l'image d'une mesure σ -finie n'est en général pas σ -finie.

Proposition 8.27 (Stabilité par marquage). On suppose que $M = \sum_{k=1}^L \delta_{X_k}$, avec $L \in \mathbb{N} \cup \{+\infty\}$ et $X_k \in E$ des variables aléatoires, est un nuage poissonien d'intensité $\mu \in \mathcal{M}_s(E)$. Soient $(Y_i)_{i \geq 1}$ des variables aléatoires i.i.d. dans $(F, \mathcal{G})$, de loi μ' , indépendantes de $(L, X_1, X_2, \dots)$. Alors

$$\widetilde{M} = \sum_{k=1}^L \delta_{(X_k, Y_k)}$$

est un nuage poissonien dans $E \times F$ d'intensité $\mu \otimes \mu'$.

Démonstration. Soit $f : E \times F \rightarrow \mathbb{R}_+$ mesurable. On calcule

$$\begin{aligned} \mathbb{E}[\exp(-\widetilde{M}(f))] &= \mathbb{E} \left[\mathbb{E} \left[\exp \left(- \sum_{k=1}^L f(X_k, Y_k) \right) \middle| L, (X_k)_{k \geq 1} \right] \right] \\ &= \mathbb{E} \left[\prod_{k=1}^L \int_F \exp(-f(X_k, y)) \mu'(dy) \right] = \mathbb{E}[\exp(-M(F))], \end{aligned}$$

avec $F(x) = -\ln \int_F e^{-f(x, y)} \mu'(dy) \geq 0$. Donc, d'après la formule exponentielle,

$$\begin{aligned} \mathbb{E}[\exp(-\widetilde{M}(f))] &= \exp \left(- \int_E (1 - e^{-F(x)}) \mu(dx) \right) \\ &= \exp \left(- \int_{E \times F} (1 - e^{-f(x, y)}) (\mu \otimes \mu')(dx dy) \right). \end{aligned}$$

D'où le résultat. □

Exemple 8.28. Soit $M = \sum_{k=1}^L \delta_{X_k}$ un nuage poissonien d'intensité μ sur E , donné avec cette énumération aléatoire, et soit $(Y_i)_{i \geq 1}$ une suite i.i.d. de loi Bernoulli(p), indépendante de $(L, X_1, X_2, \dots)$. Alors

$$M_p = \sum_{k=1}^L Y_k \delta_{X_k}$$

est un nuage poissonien d'intensité $p\mu$, car M_p s'identifie à la restriction à $E \times \{1\}$ de $\widetilde{M} = \sum_{k=1}^L \delta_{(X_k, Y_k)}$, d'intensité $\mu \otimes \text{Bernoulli}(p)$.

8.6 Nuages de Poisson avec paramètre temporel

On se place dans le cadre d'un espace mesuré $(E, \mathcal{E})$, avec $\{x\} \in \mathcal{E}$ pour tout $x \in E$. On suppose que $\mathbb{P}$ est complète et que μ est σ -finie sur E .

Définition 8.29. Un *processus ponctuel de Poisson* sur E d'intensité μ est un nuage poissonien M sur $\mathbb{R}_+ \times E$ d'intensité $dt \otimes \mu$.

Proposition 8.30. Presque sûrement, pour tout $t \geq 0$, $M(\{t\} \times E) \leq 1$.

Démonstration. Choisissons des ensembles mesurables $E_n \uparrow E$ tels que $\mu(E_n) < \infty$, et posons $\pi(t, x) = t$. Pour $m, n \geq 1$, la stabilité par restriction et par image montre que

$$N_{m,n} = \pi_*(M|_{[0,m] \times E_n})$$

est un nuage poissonien sur $\mathbb{R}_+$ d'intensité $\mu(E_n) \mathbb{1}_{[0,m]}(t) dt$. Cette intensité est finie et sans atome, donc $N_{m,n}$ est presque sûrement simple.

Sur l'intersection dénombrable de ces événements de probabilité 1, on a, pour tous $m, n \geq 1$ et $t \in [0, m]$,

$$M(\{t\} \times E_n) = N_{m,n}(\{t\}) \leq 1.$$

En faisant $n \rightarrow \infty$ puis en choisissant $m \geq t$, on obtient $M(\{t\} \times E) \leq 1$ pour tout $t \geq 0$. $\square$

Définition 8.31. Pour $t \geq 0$, on pose

$$e(t) = \begin{cases} x & \text{si } M(\{(t, x)\}) = 1, \\ \partial & \text{sinon, où } \partial \text{ est un point cimetière.} \end{cases}$$

Proposition 8.32. 1) Pour tout $t \geq 0$, $e(t)$ est une variable aléatoire.

2) Si $\mu(E) = \infty$, $\{t : e(t) \neq \partial\}$ est dense presque sûrement.

3) Si $0 < \mu(E) < \infty$, presque sûrement, $\{t : e(t) \neq \partial\}$ est infini discret et peut être ordonné avec des variables aléatoires $0 < T_1 < T_2 < \dots$ qui vérifient :

- (a) $T_1, T_2 - T_1, T_3 - T_2, \dots$ sont i.i.d. de loi $\text{Exp}(\mu(E))$;
- (b) les variables $(e(T_i))_{i \geq 1}$ sont i.i.d. de loi $\mu(\cdot)/\mu(E)$;
- (c) $(e(T_i))_{i \geq 1}$ est indépendante de $(T_i)_{i \geq 1}$.

Démonstration. 1) Pour $A \in \mathcal{E}$,

$$\{e(t) \in A\} = \{M(\{t\} \times A) = 1\}, \quad \{e(t) = \partial\} = \{M(\{t\} \times E) = 0\}.$$

2) Si $\mu(E) = \infty$, pour tous $a < b$ dans $\mathbb{Q}_+$, $M([a, b] \times E) \sim \mathcal{P}(\infty)$, donc vaut ∞ presque sûrement.

3) La loi de $(T_i, e(T_i))_{i \geq 1}$ ne dépend que de la loi de M . Par exemple,

$$T_1 = \inf\{t \geq 0 : M([0, t] \times E) \geq 1\}, \quad \{e(T_i) \in A\} = \{M(\{T_i\} \times A) = 1\}.$$

La propriété à démontrer ne dépend donc que de la loi de M . Construisons alors M comme suit. Soit $P = \sum_{i=1}^{\infty} \delta_{T_i}$ un nuage poissonien d'intensité $\mu(E) dt$ sur $\mathbb{R}_+$ et $(Y_i)_{i \geq 1}$ des variables i.i.d. de loi $\mu(\cdot)/\mu(E)$, indépendantes de P . Alors, par marquage,

$$M = \sum_{i=1}^{\infty} \delta_{(T_i, Y_i)}$$

est un nuage poissonien d'intensité $dt \otimes \mu$. Par ailleurs, on a vu que $(P([0, t]))_{t \geq 0}$ est un processus de Poisson d'intensité $\mu(E)$, et on peut bien ordonner $T_1 < T_2 < \dots$ qui vérifient (a). Les propriétés (b) et (c) découlent du fait que $e(T_i) = Y_i$. $\square$

Exemple 8.33. Prenons $\mu(dx) = x^{-3/2} \mathbb{1}_{\{x>0\}} dx$ sur $E = \mathbb{R}_+^*$. On a $\mu(E) = \infty$. Pour $t \geq 0$, on pose

$$X_t = \sum_{s \leq t} e(s) \mathbb{1}_{\{e(s) \neq \emptyset\}} = M(f_t), \quad f_t(s, x) = x \mathbb{1}_{\{s \leq t\}}.$$

Comme $\int_{\mathbb{R}_+ \times \mathbb{R}_+^*} (1 \wedge f_t(s, x)) ds \mu(dx) < \infty$, pour tout $t \geq 0$, presque sûrement $X_t < \infty$; par monotonie, presque sûrement, $X_t < \infty$ pour tout $t \geq 0$. Néanmoins, pour tout $t > 0$, $\int_{\mathbb{R}_+ \times \mathbb{R}_+^*} f_t(s, x) ds \mu(dx) = \infty$, donc $\mathbb{E}[X_t] = \infty$.

En utilisant les propriétés des mesures aléatoires de Poisson, on peut vérifier que $(X_t)_{t \geq 0}$ est un processus à accroissements stationnaires et indépendants, et il est càdlàg par construction : c'est un processus de Lévy croissant (un *subordonneur*).

8.7 Formule de Palm

Théorème 8.34 (Formule de Palm / équation de Mecke). Soient μ σ -finie et M un nuage poissonien d'intensité μ . Soit $f : E \times \mathcal{M}_{\text{atom}}(E) \rightarrow \mathbb{R}_+$ mesurable. Alors

$$\mathbb{E} \left[\int_E f(x, M) M(dx) \right] = \int_E \mathbb{E}[f(x, M + \delta_x)] \mu(dx). \quad (9)$$

La formule de Palm décrit le nuage vu depuis un point distingué : le membre de gauche somme sur tous les points du nuage, avec leur multiplicité, tandis que le membre de droite ajoute un point en x à un nuage de Poisson de loi initiale et intègre contre $\mu(dx)$. Cette identité ne correspond pas, en général, au conditionnement ordinaire par $M(\{x\}) \geq 1$; la distinction est essentielle lorsque μ possède des atomes. On peut voir ce résultat comme une extension de la formule additive prenant en compte les atomes.

Démonstration. L'application $(x, \nu) \mapsto \nu + \delta_x$ est conjointement mesurable, puisque, pour tout $A \in \mathcal{E}$,

$$(\nu + \delta_x)(A) = \nu(A) + \mathbb{1}_A(x).$$

Choisissons $E_n \uparrow E$ tels que $\mu(E_n) < \infty$, et posons

$$H = \{\nu \in \mathcal{M}_{\text{atom}}(E) : \nu(E_n) < \infty \text{ pour tout } n\}.$$

Cet ensemble est mesurable et $\mathbb{P}(M \in H) = 1$. En remplaçant si nécessaire M par la mesure nulle sur l'événement nul $\{M \notin H\}$, nous pouvons travailler avec une version à valeurs dans H . Sur H , l'application $\nu \mapsto \nu|_{E_n}$ est un noyau fini. Un argument de classe monotone donne donc la mesurabilité de

$$\nu \mapsto \int_{E_n} f(x, \nu) \nu(dx).$$

En faisant croître n , on obtient celle de l'intégrale sur E . Ces observations, avec la mesurabilité conjointe de l'ajout d'un atome et le théorème de Tonelli, justifient les deux membres de (9) et montrent qu'ils ne dépendent que de la loi de M .

Étape 1 : μ finie. Si $\mu(E) = 0$, alors $M = 0$ presque sûrement et les deux membres de la formule sont nuls. Supposons donc $\mu(E) > 0$. On écrit $M = \sum_{i=1}^L \delta_{X_i}$, avec $L \sim \mathcal{P}(\mu(E))$ et $(X_i)_{i \geq 1}$ i.i.d. de loi $\mu(\cdot)/\mu(E)$, indépendantes de L . Alors

$$\begin{aligned} \mathbb{E} \left[\int_E f(x, M) M(dx) \right] &= \sum_{k=1}^{\infty} \frac{\mu(E)^k}{k!} e^{-\mu(E)} \mathbb{E} \left[\sum_{i=1}^k f \left(X_i, \sum_{j=1}^k \delta_{X_j} \right) \right] \\ &= \sum_{k=1}^{\infty} \frac{\mu(E)^k}{k!} e^{-\mu(E)} k \mathbb{E} \left[f \left(X_k, \sum_{j=1}^k \delta_{X_j} \right) \right] \\ &= \sum_{k=1}^{\infty} \frac{\mu(E)^k}{k!} e^{-\mu(E)} k \int_E \mathbb{E} \left[f \left(x, \sum_{j=1}^{k-1} \delta_{X_j} + \delta_x \right) \right] \frac{\mu(dx)}{\mu(E)} \\ &= \int_E \left(\sum_{k=1}^{\infty} \frac{\mu(E)^{k-1}}{(k-1)!} e^{-\mu(E)} \mathbb{E} \left[f \left(x, \sum_{j=1}^{k-1} \delta_{X_j} + \delta_x \right) \right] \right) \mu(dx) \\ &= \int_E \mathbb{E}[f(x, M + \delta_x)] \mu(dx). \end{aligned}$$

Étape 2 : μ σ -finie. Écrivons $E = \bigsqcup_{i \geq 1} A_i$, avec $\mu(A_i) < \infty$, et posons $M_i = M|_{A_i}$ et $R_i = M|_{A_i^c}$. Ces deux restrictions sont indépendantes. Pour chaque mesure atomique r fixée, appliquons le cas fini à M_i et à la fonction $(x, v) \mapsto f(x, v + r)$, puis intégrons en r suivant la loi de R_i . L'indépendance et le théorème de Tonelli donnent

$$\mathbb{E} \left[\int_{A_i} f(x, M) M(dx) \right] = \int_{A_i} \mathbb{E}[f(x, M + \delta_x)] \mu(dx).$$

En sommant sur $i \geq 1$ et en appliquant à nouveau Tonelli, on obtient (9). $\square$

Remarque 8.35. Si les atomes de M peuvent être énumérés de manière mesurable (c'est le cas si μ est σ -finie et E polonais), la formule de Palm peut s'écrire

$$\mathbb{E} \left[\int_E f(x, M - \delta_x) M(dx) \right] = \int_E \mathbb{E}[f(x, M)] \mu(dx).$$

Partie II

Exercices

Sommaire

1	Convergence étroite dans $\mathbb{R}^k$	98
2	Différents modes de convergence (presque sûre, en probabilité, L^1)	111
3	Mesures de probabilité sur un espace métrique	125
4	L'espace des fonctions continues sur un compact	155
5	Théorème de Donsker, pont brownien	177
6	L'espace des fonctions càdlàg	196
7	Mesures aléatoires de Poisson	203

1 Convergence étroite dans $\mathbb{R}^k$

Cours associé : [chapitre 1](#), p. 6.

On note $\mathcal{M}_1(\mathbb{R}^k)$ l'ensemble des mesures de probabilités sur $\mathbb{R}^k$, avec $k \geq 1$ un entier fixé, et $\Rightarrow$ désigne la convergence étroite dans cet espace.

1.1 Énoncés

Exercice 1.1. [\[corrigé p. 101\]](#) Soit $(\mu_n)_{n \geq 1}$ une suite de mesures de $\mathcal{M}_1(\mathbb{R})$ et $\mu \in \mathcal{M}_1(\mathbb{R})$. On suppose que $\mu_n \Rightarrow \mu$. Y a-t-il des implications entre les assertions suivantes ?

- (a) μ_n est à densité pour n assez grand
- (b) μ est à densité
- (c) μ_n est atomique pour n assez grand
- (d) μ est atomique.

Rappelons qu'une mesure atomique est une mesure qui s'écrit $\sum a_i \delta_{b_i}$ pour des suites $a_i \in \mathbb{R}_+$ et $b_i \in \mathbb{R}$ et que dans $\mathbb{R}^n$ par mesure à densité on entend par rapport à la mesure de Lebesgue (l'usage est juste de dire à densité).

* * *

Exercice 1.2. [\[corrigé p. 101\]](#) Montrer qu'une famille $(\mu_i)_{i \in I}$ de mesures de $\mathcal{M}_1(\mathbb{R})$ est tendue si et seulement si il existe une fonction mesurable $f : \mathbb{R} \rightarrow [0, \infty)$ telle que $f(x) \rightarrow \infty$ pour $|x| \rightarrow \infty$ et $\sup_{i \in I} \int_{\mathbb{R}} f d\mu_i < \infty$.

* * *

Exercice 1.3. [\[corrigé p. 102\]](#) Soit $(X_n)_{n \geq 1}$ des variables aléatoires réelles. Montrer que la suite $(X_n)_{n \geq 1}$ est tendue si et seulement si pour tout $\varepsilon > 0$ et pour toute suite $(c_n)_{n \geq 1}$ de réels strictement positifs telle que $c_n \rightarrow 0$ on a $\mathbb{P}(c_n |X_n| \geq \varepsilon) \rightarrow 0$.

* * *

Exercice 1.4. [\[corrigé p. 102\]](#) (**Théorème de Riesz et lemme de Scheffé**) Soit $(E, \mathcal{A}, \mu)$ un espace mesuré, avec μ une mesure positive (pas forcément finie). Soit $(f_n)_{n \geq 1}$ une suite de fonctions mesurables de E dans $\mathbb{R}$ telles que :

$$f_n \rightarrow f \quad \mu - \text{presque partout.} \quad (1)$$

- (1) Soit $p \in [1, \infty[$. On suppose que la convergence (1) a lieu, que pour tout $n \geq 1$, $f_n \in L^p(\mu)$, $f \in L^p(\mu)$ et $\|f_n\|_p \rightarrow \|f\|_p$ quand $n \rightarrow \infty$. Démontrer que $f_n \rightarrow f$ dans $L^p(\mu)$ (théorème de Riesz).

On pourra introduire la fonction $g_n = 2^p(|f_n|^p + |f|^p) - |f_n - f|^p$.

- (2) Montrer le lemme de Scheffé :

$$\text{si (1), } f_n \text{ et } f \text{ sont } \mu \text{ intégrables, } f_n \geq 0 \text{ et } \int_E f_n d\mu \xrightarrow{n \rightarrow \infty} \int_E f d\mu$$

$$\text{alors } \int_E |f_n - f| d\mu \xrightarrow{n \rightarrow \infty} 0.$$

- (3) Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires positives intégrables qui converge presque sûrement vers X . On suppose que $\mathbb{E}[X_n] \rightarrow \mathbb{E}[X] < \infty$. Montrer que X_n converge vers X dans L^1 .
- (4) Dans cette question, on suppose $E = \mathbb{R}$ et $\mathcal{A} = \mathcal{B}(\mathbb{R})$. Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires réelles telles que X_n ait pour loi $f_n d\mu$. Montrer que, si f_n converge μ -presque partout vers une densité de probabilité f , alors $X_n \Rightarrow X$, où X a pour loi $f d\mu$.

* * *

Exercice 1.5. [corrigé p. 103] (**Transformée de Laplace**) Soit $\mu \in \mathcal{M}_1(\mathbb{R}_+)$. Sa transformée de Laplace est définie par l'intégrale $L_\mu(t) = \int_{\mathbb{R}_+} e^{-tx} \mu(dx)$ pour $t \geq 0$.

- (1) Vérifier que L_μ est de classe C^∞ sur $]0, \infty[$.
- (2) Montrer que pour tout $x > 0$ on a

$$\mu([0, x[) + \frac{1}{2}\mu(\{x\}) = \lim_{t \rightarrow \infty} \sum_{k=0}^{\lfloor tx \rfloor} \frac{(-t)^k}{k!} L_\mu^{(k)}(t),$$

où $L_\mu^{(k)}$ désigne la dérivée k -ième de L_μ . En déduire que si $\mu, \nu \in \mathcal{M}_1(\mathbb{R}_+)$ ont même transformée de Laplace, alors $\mu = \nu$.

- (3) Soit $(\mu_n)_{n \geq 1}$ une suite de mesures de $\mathcal{M}_1(\mathbb{R}_+)$. On suppose que L_{μ_n} converge simplement sur $\mathbb{R}_+$ vers une fonction L continue à droite en 0. Montrer qu'il existe $\mu \in \mathcal{M}_1(\mathbb{R}_+)$ telle que $L = L_\mu$ et $\mu_n \Rightarrow \mu$.

On pourra s'inspirer de la preuve du théorème de Lévy.

* * *

Exercice 1.6. [corrigé p. 105] Soit $(\mu_n)_{n \geq 1}$ une suite de mesures de $\mathcal{M}_1(\mathbb{R})$ et $\mu \in \mathcal{M}_1(\mathbb{R})$.

- (1) On suppose que $(\mu_n)_{n \geq 1}$ est tendue. Montrer que

$$\forall \varepsilon > 0, \quad \exists \delta > 0, \quad \forall n \geq 1, \quad |x - y| \leq \delta \implies |\phi_{\mu_n}(x) - \phi_{\mu_n}(y)| \leq \varepsilon.$$

(en d'autres termes, la suite $(\phi_{\mu_n})_{n \geq 1}$ est uniformément équicontinue).

- (2) On suppose que $\mu_n \Rightarrow \mu$. Montrer que ϕ_{μ_n} converge vers ϕ_μ uniformément sur tout compact. Donner un exemple où la convergence n'est pas uniforme.
- (3) La réciproque de l'énoncé de la première question est-elle vraie, autrement dit est-ce que si la suite $(\phi_{\mu_n})_{n \geq 1}$ est uniformément équicontinue alors la suite $(\mu_n)_{n \geq 1}$ est tendue?

* * *

Exercice 1.7. [corrigé p. 106] Soit $(\mu_n)_{n \geq 1}$ une suite de mesures de $\mathcal{M}_1(\mathbb{R})$ et $\mu \in \mathcal{M}_1(\mathbb{R})$ sans atomes. Montrer que $\mu_n \Rightarrow \mu$ si et seulement si $\sup_{s \in \mathbb{R}} |F_{\mu_n}(s) - F_\mu(s)| \rightarrow 0$.

Exercice 1.8. [corrigé p. 107] (**Théorème de Glivenko-Cantelli**) Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires réelles indépendantes et de même loi μ . On considère $\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$ la mesure empirique de ces variables aléatoires. On note F_n la fonction de répartition de μ_n et F celle de μ . Le but de cet exercice est de démontrer le théorème de Glivenko-Cantelli :

$$\text{presque sûrement, } \lim_{n \rightarrow \infty} \sup_{s \in \mathbb{R}} |F_n(s) - F(s)| = 0. \quad (*)$$

(1) Montrer (*) lorsque μ est la loi uniforme sur $[0, 1]$.

Pour $0 < x < 1$, on pose $G(x) = \inf\{s \in \mathbb{R} : F(s) \geq x\}$ (appelé inverse généralisé de F). Il est possible de vérifier que G est croissante et continue à gauche sur $]0, 1[$. Soit $(Y_i)_{i \geq 1}$ une suite de variables aléatoires indépendantes de loi uniforme sur $]0, 1[$.

(2) (a) Montrer que pour tout $t \in \mathbb{R}$ et $x \in]0, 1[$ on a

$$F(t) \geq x \iff t \geq G(x). \quad (2)$$

(b) Montrer que $(G(Y_i))_{i \geq 1}$ sont des variables indépendantes de même loi μ .

(c) Montrer que les suites de fonctions $(F_n)_{n \geq 1}$ et $(A_n \circ F)_{n \geq 1}$ ont même loi, où A_n est la fonction de répartition empirique de $Y_1, \dots, Y_n$.

(3) En déduire (*).

* * *

Exercice 1.9. [corrigé p. 108] (**Caractérisation des fonctions de répartition dans $\mathbb{R}^k$**)

Pour $x = (x_1, \dots, x_k) \in \mathbb{R}^k$ et $y = (y_1, \dots, y_k) \in \mathbb{R}^k$ on note $x \leq y$ si $x_i \leq y_i$ pour tout $1 \leq i \leq k$. Soit $F : \mathbb{R}^k \rightarrow [0, 1]$ une fonction. On dit qu'elle est *continue à droite* si $F(x^n) \rightarrow F(x)$ lorsque $x^n \downarrow x$. On dit qu'elle est *propre* si $F(x) \rightarrow 1$ lorsque $\min_i x_i \rightarrow \infty$ et $F(x) \rightarrow 0$ lorsque $\min_i x_i \rightarrow -\infty$. On dit qu'elle est une *fonction de répartition* s'il existe une mesure $\mu \in \mathcal{M}_1(\mathbb{R}^k)$ telle que $F(x) = \mu(\{y \in \mathbb{R}^k : y \leq x\})$.

(1) Donner un exemple de fonction $F : \mathbb{R}^2 \rightarrow [0, 1]$ continue à droite, propre, croissante en chacune de ses variables et qui n'est pas une fonction de répartition.

On dit que F est à *accroissements positifs* si pour tout pavé $]x, y] =]x_1, y_1] \times \dots \times]x_k, y_k]$ on a $F([x, y]) := \sum_u s(u)F(u) \geq 0$, où la somme est prise sur tous les coins u de $]x, y]$ et $s(u) = (-1)^p$ avec $p = \sum_{i=1}^k \mathbb{1}_{u_i = x_i}$.

(2) Montrer que $F : \mathbb{R}^k \rightarrow [0, 1]$ est une fonction de répartition si et seulement si elle est continue à droite, propre, et est à accroissements positifs.

Pour une preuve probabiliste de la réciproque, on pourra poser $\mathbf{1} = (1, \dots, 1)$ et justifier l'existence, pour tout $n \geq 1$, d'une probabilité μ_n à support dans $(2^{-n}\mathbb{Z})^k$ telle que

$$\mu_n(\{2^{-n}x\}) = F([2^{-n}(x - \mathbf{1}), 2^{-n}x]), \quad x \in \mathbb{Z}^k.$$

On construira ensuite des variables X^n de loi μ_n telles que $X^m - 2^{-m}\mathbf{1} < X^n \leq X^m$ coordonnée par coordonnée pour $m < n$, puis on considérera $X = \lim_n X^n$.

Exercice 1.10. [corrigé p. 109] Soit (E, d) un espace métrique muni de sa tribu borélienne et $f : E \rightarrow \mathbb{R}$ mesurable. Montrer que l'ensemble des points de discontinuité de f est mesurable.

Indication. Pour $\varepsilon, \delta > 0$, on pourra vérifier que $U_{\varepsilon, \delta} := \{x \in E : \exists y, z \in B(x, \varepsilon), |f(y) - f(z)| > \delta\}$ est ouvert.

* * *

Exercice 1.11. [corrigé p. 109] Soit μ_n, μ des mesures de probabilité sur $\mathbb{R}$. On note λ la mesure de Lebesgue sur $\mathbb{R}$ et $\phi_\nu(t) = \int e^{itx} \nu(dx)$ pour une mesure de probabilité ν sur $\mathbb{R}$.

- (1) Démontrer que $\mu_n \Rightarrow \mu$ si et seulement si la fonction caractéristique de μ_n converge simplement vers la fonction caractéristique de μ pour presque tout point par rapport à la mesure de Lebesgue.
- (2) Est-il vrai que $\mu_n \Rightarrow \mu$ si et seulement si la fonction caractéristique de μ_n converge simplement sur un ensemble dénombrable dense vers la fonction caractéristique de μ ?

* * *

Exercice 1.12. [corrigé p. 110] Soit $(X_n)_{n \geq 1}$, X des variables aléatoires réelles. Soit D un ensemble dénombrable dense de $\mathbb{R}$.

- (1) Alix dit : « si X_n converge en loi vers X lorsque $n \rightarrow \infty$, alors pour tout $u \in D$ on a $\mathbb{P}(X_n \leq u) \rightarrow \mathbb{P}(X \leq u)$ lorsque $n \rightarrow \infty$. » A-t-il raison ? Justifiez votre réponse.
- (2) Billie dit : « si pour tout $u \in D$ on a $\mathbb{P}(X_n \leq u) \rightarrow \mathbb{P}(X \leq u)$ lorsque $n \rightarrow \infty$, alors X_n converge en loi vers X lorsque $n \rightarrow \infty$. » A-t-elle raison ? Justifiez votre réponse.

1.2 Corrigés

Corrigé de l'exercice 1.1. [énoncé p. 98] Non :

- si μ_n est à densité, μ peut aussi bien être à densité (prendre $\mu = \mu_n$) qu'atomique (prendre $\mu_n(dx) = \frac{n}{2} \mathbb{1}_{[-1/n, 1/n]}(x) dx$, qui converge étroitement vers δ_0), voire singulière par rapport à la mesure de Lebesgue (penser à l'escalier du diable).
- si μ_n est atomique, μ peut aussi bien être atomique (prendre $\mu = \mu_n$) ou bien à densité (prendre $\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{i/n}$, qui converge étroitement vers la mesure de Lebesgue sur $[0, 1]$).

* * *

Corrigé de l'exercice 1.2. [énoncé p. 98] $\Leftarrow$ Posons $C = 1 + \sup_{i \in I} \int_{\mathbb{R}} f d\mu_i$. Soit $\varepsilon > 0$ et $A > 0$ tel que $f(x) \geq C/\varepsilon$ pour $|x| \geq A$. On écrit, pour $i \in I$,

$$C \geq \int_{\mathbb{R}} f d\mu_i \geq \int_{\mathbb{R} \setminus [-A, A]} f d\mu_i \geq \frac{C}{\varepsilon} \int_{\mathbb{R} \setminus [-A, A]} d\mu_i,$$

de sorte que $\mu_i(\mathbb{R} \setminus [-A, A]) \leq \varepsilon$.

$\Rightarrow$ Soit $(A_n)_{n \geq 1}$ une suite de réels strictement positifs tels que pour tout $n \geq 1$ et $i \in I$ on ait

$$\mu_i(\mathbb{R} \setminus [-A_n, A_n]) \leq \frac{1}{n^3}.$$

Sans perte de généralité, on peut supposer que la suite (A_n) est strictement croissante et $A_n \rightarrow \infty$. On définit alors f par $f(x) = 0$ si $|x| \leq A_1$ et $f(x) = n$ si $A_n < |x| \leq A_{n+1}$, de sorte que f est mesurable, $f(x) \rightarrow \infty$, et pour $i \in I$,

$$\begin{aligned} \int_{\mathbb{R}} f d\mu_i &= \sum_{n=1}^{\infty} \int_{[-A_{n+1}, A_{n+1}] \setminus [-A_n, A_n]} f d\mu_i \\ &\leq \sum_{n=1}^{\infty} n \mu_i([-A_{n+1}, A_{n+1}] \setminus [-A_n, A_n]) \\ &\leq \sum_{n=1}^{\infty} n \mu_i(\mathbb{R} \setminus [-A_n, A_n]) \\ &\leq \sum_{n=1}^{\infty} \frac{1}{n^2}, \end{aligned}$$

d'où le résultat.

* * *

Corrigé de l'exercice 1.3. [énoncé p. 98] Supposons que la suite $(X_n)_{n \geq 1}$ est tendue. Soit $\varepsilon > 0$ et $(c_n)_{n \geq 1}$ une suite de réels strictement positifs telle que $c_n \rightarrow 0$. Soit $\eta > 0$. Par tension, il existe $M > 0$ tel que $\mathbb{P}(|X_n| \geq M) \leq \eta$. Pour n assez grand, $\varepsilon/c_n \geq M$, et alors

$$\mathbb{P}(c_n |X_n| \geq \varepsilon) \leq \mathbb{P}(|X_n| \geq M) \leq \eta,$$

d'où le résultat.

Réciproquement, raisonnons par l'absurde en supposant que la suite $(X_n)_{n \geq 1}$ n'est pas tendue. On peut alors trouver $\varepsilon > 0$ et une extraction $\phi(n)$ telle que $\mathbb{P}(|X_{\phi(n)}| \geq n) \geq \varepsilon$ pour tout n . On définit alors $c_i = 1$ pour $1 \leq i < \phi(1)$ et $c_i = 1/k$ pour $\phi(k) \leq i < \phi(k+1)$. Cette suite est strictement positive, tend vers 0 et vérifie $c_{\phi(k)} = 1/k$. Alors, pour tout $n \geq 1$,

$$\mathbb{P}(c_{\phi(n)} |X_{\phi(n)}| \geq 1) \geq \varepsilon,$$

contradiction.

* * *

Corrigé de l'exercice 1.4. [énoncé p. 98]

(1) Par inégalité triangulaire, on remarque que

$$\begin{aligned} |f_n - f|^p &\leq (|f_n| + |f|)^p \leq 2^p \max(|f_n|^p, |f|^p) \\ &\leq 2^p (|f_n|^p + |f|^p), \end{aligned}$$

de sorte que $g_n \geq 0$. Appliquons le lemme de Fatou à g_n :

$$\begin{aligned} 2^{p+1} \int_E |f|^p d\mu &= \int_E \left(\liminf_{n \rightarrow \infty} g_n \right) d\mu \\ &\leq \liminf_{n \rightarrow \infty} \int_E 2^p (|f_n|^p + |f|^p) d\mu \\ &\quad - \limsup_{n \rightarrow \infty} \int_E |f_n - f|^p d\mu. \end{aligned}$$

On en déduit que

$$\limsup_{n \rightarrow \infty} \int_E |f_n - f|^p d\mu \leq 0,$$

d'où le résultat.

- (2) C'est simplement le théorème de Riesz pour $p = 1$.
 (3) C'est simplement le lemme de Scheffé.
 (4) Soit $s \in \mathbb{R}$. On montre que $\mathbb{P}(X_n \leq s) \rightarrow \mathbb{P}(X \leq s)$. Pour cela, en notant $A =]-\infty, s]$, on remarque que

$$\left| \int_A f_n d\mu - \int_A f d\mu \right| \leq \int_{\mathbb{R}} |f_n - f| d\mu \xrightarrow{n \rightarrow \infty} 0,$$

ce qui conclut.

Remarque. La même preuve montre qu'en fait $\mathbb{P}(X_n \in A) \rightarrow \mathbb{P}(X \in A)$ pour tout borélien A . Ceci implique la convergence en loi, mais la réciproque n'est pas vraie (prendre par exemple $\mu_n = \delta_{1/n} \Rightarrow \delta_0$ avec $A = \{0\}$).

* * *

Corrigé de l'exercice 1.5. [énoncé p. 99]

- (1) Sur tout intervalle compact $[a, b] \subset]0, \infty[$ et pour tout entier $k \geq 0$,

$$\left| \frac{\partial^k}{\partial t^k} e^{-tu} \right| = u^k e^{-tu} \leq u^k e^{-au}.$$

Le majorant est borné sur $\mathbb{R}_+$, donc μ -intégrable. La dérivation sous l'intégrale, à tout ordre, donne $L_\mu \in C^\infty([0, \infty[)$ et

$$L_\mu^{(k)}(t) = (-1)^k \int_{\mathbb{R}_+} u^k e^{-tu} \mu(du).$$

- (2) Soit $x > 0$. D'après la question précédente et le théorème de Fubini,

$$\sum_{k=0}^{\lfloor tx \rfloor} \frac{(-t)^k}{k!} L_\mu^{(k)}(t) = \int_{\mathbb{R}_+} \sum_{k=0}^{\lfloor tx \rfloor} \frac{(tu)^k}{k!} e^{-tu} \mu(du).$$

Or

$$\sum_{k=0}^{\lfloor tx \rfloor} \frac{(tu)^k}{k!} e^{-tu} = \mathbb{P}(\text{Poisson}(tu) \leq \lfloor tx \rfloor).$$

Montrons que

$$\mathbb{P}(\text{Poisson}(tu) \leq \lfloor tx \rfloor) \xrightarrow{t \rightarrow \infty} \begin{cases} 0 & \text{si } x < u \\ \frac{1}{2} & \text{si } x = u \\ 1 & \text{si } x > u \end{cases}$$

en distinguant les trois cas :

- $x = u$: En notant $(Y_i)_{i \geq 1}$ des variables aléatoires i.i.d. de Poisson de paramètre x , posons $m = \lfloor t \rfloor$ et prenons une variable Z_t de loi de Poisson de paramètre $(t - m)x$, indépendante des Y_i . Alors

$$\begin{aligned} \mathbb{P}(\text{Poisson}(tx) \leq \lfloor tx \rfloor) \\ = \mathbb{P}\left(\frac{Y_1 + \dots + Y_m - mx + Z_t - (t - m)x}{\sqrt{tx}} \leq \frac{\lfloor tx \rfloor - tx}{\sqrt{tx}}\right). \end{aligned}$$

Le TCL donne $(Y_1 + \dots + Y_m - mx)/\sqrt{tx} \Rightarrow \mathcal{N}(0, 1)$. De plus, $\text{Var}(Z_t) = (t - m)x \leq x$, donc $(Z_t - (t - m)x)/\sqrt{tx} \rightarrow 0$ en probabilité, tandis que $(\lfloor tx \rfloor - tx)/\sqrt{tx} \rightarrow 0$. Le lemme de Slutsky et la continuité de la fonction de répartition de la loi normale en 0 donnent la limite 1/2.

- $x < u$: alors

$$\begin{aligned} \mathbb{P}(\text{Poisson}(tu) \leq \lfloor tx \rfloor) &= \mathbb{P}(tu - \lfloor tx \rfloor \leq tu - \text{Poisson}(tu)) \\ &\leq \mathbb{P}(tu - \lfloor tx \rfloor \leq |tu - \text{Poisson}(tu)|) \\ &\leq \frac{\text{Var}(\text{Poisson}(tu))}{(tu - \lfloor tx \rfloor)^2} \\ &\sim \frac{u}{t(u - x)^2} \end{aligned}$$

lorsque $t \rightarrow \infty$, qui tend donc vers 0.

- $x > u$: alors, de même,

$$\begin{aligned} \mathbb{P}(\text{Poisson}(tu) \leq \lfloor tx \rfloor) &= 1 - \mathbb{P}(\text{Poisson}(tu) > \lfloor tx \rfloor) \\ &= 1 - \mathbb{P}(\text{Poisson}(tu) - tu > \lfloor tx \rfloor - tu) \\ &\geq 1 - \mathbb{P}(|\text{Poisson}(tu) - tu| > \lfloor tx \rfloor - tu) \end{aligned}$$

et Si $u = 0$, la probabilité considérée vaut 0. Si $u > 0$, pour t assez grand,

$$\mathbb{P}(|\text{Poisson}(tu) - tu| > \lfloor tx \rfloor - tu) \leq \frac{\text{Var}(\text{Poisson}(tu))}{(\lfloor tx \rfloor - tu)^2} \sim \frac{u}{t(x - u)^2} \rightarrow 0.$$

Ainsi,

$$\sum_{k=0}^{\lfloor tx \rfloor} \frac{(tu)^k}{k!} e^{-tu} \xrightarrow{t \rightarrow \infty} \mathbb{1}_{[0, x[}(u) + \frac{1}{2} \mathbb{1}_{\{x\}}(u).$$

En utilisant le théorème de convergence dominée, on conclut que

$$\begin{aligned} & \int_{\mathbb{R}_+} \sum_{k=0}^{\lfloor tx \rfloor} \frac{(tu)^k}{k!} e^{-tu} \mu(du) \\ & \xrightarrow{t \rightarrow \infty} \int_{\mathbb{R}_+} \left(\mathbb{1}_{[0,x[}(u) + \frac{1}{2} \mathbb{1}_{\{x\}}(u) \right) \mu(du) \\ & = \mu([0, x]) + \frac{1}{2} \mu(\{x\}). \end{aligned}$$

Supposons maintenant que $\mu, \nu \in \mathcal{M}_1(\mathbb{R}_+)$ ont même transformée de Laplace. Notons D l'ensemble des atomes de μ et de ν , qui est au plus dénombrable. Alors d'après le résultat précédent, $\mu([x, y]) = \nu([x, y])$ pour tous $x, y \in \mathbb{R} \setminus D$. Puisque $\mathbb{R} \setminus D$ est dense, on en déduit que μ et ν coïncident sur tout intervalle ouvert et donc $\mu = \nu$ par application du lemme des classes monotones.

- (3) On remarque tout d'abord que si $\mu_n \Rightarrow \mu$, alors L_{μ_n} converge simplement vers L_μ (la fonction $x \mapsto e^{-tx}$ étant continue bornée sur $\mathbb{R}_+$).

Il suffit de montrer que $(\mu_n)_{n \geq 1}$ est tendue. Toute limite étroite d'une sous-suite est portée par $\mathbb{R}_+$ et sa transformée de Laplace est L , puisque $x \mapsto e^{-tx}$ est continue bornée sur $\mathbb{R}_+$. La question (2) montre que toutes ces limites sont égales. La tension et l'unicité des valeurs d'adhérence donnent donc l'existence de $\mu \in \mathcal{M}_1(\mathbb{R}_+)$ telle que $\mu_n \Rightarrow \mu$ et $L = L_\mu$.

Pour montrer la tension, on utilise, pour tout $K > 0$,

$$(1 - e^{-1})\mu_n([K, \infty[) \leq \int_{\mathbb{R}_+} (1 - e^{-x/K}) \mu_n(dx) = 1 - L_{\mu_n}(1/K).$$

Par convergence simple des transformées de Laplace,

$$\limsup_{n \rightarrow \infty} \mu_n([K, \infty[) \leq \frac{1 - L(1/K)}{1 - e^{-1}} \rightarrow 0 \quad (K \rightarrow \infty),$$

car $L(0) = 1$ et L est continue à droite en zéro. On absorbe les premiers termes, en nombre fini, en augmentant K : la suite $(\mu_n)_{n \geq 1}$ est donc tendue.

* * *

Corrigé de l'exercice 1.6. [énoncé p. 99]

- (1) Soit $\varepsilon > 0$ et $M > 0$ tel que pour tout $n \geq 1$, $\mu_n(\mathbb{R} \setminus [-M, M]) \leq \varepsilon$. Par uniforme continuité, il existe $\delta > 0$ tel que $|x - y| \leq \delta$ et $t \in [-M, M]$ impliquent $|e^{itx} - e^{ity}| \leq \varepsilon$. On écrit alors :

$$\begin{aligned} |\phi_{\mu_n}(x) - \phi_{\mu_n}(y)| & \leq \int_{[-M, M]} |e^{itx} - e^{ity}| \mu_n(dt) + \int_{\mathbb{R} \setminus [-M, M]} |e^{itx} - e^{ity}| \mu_n(dt) \\ & \leq \varepsilon + 2\mu_n(\mathbb{R} \setminus [-M, M]) \\ & \leq 3\varepsilon \end{aligned}$$

- (2) On sait que ϕ_{μ_n} converge simplement vers ϕ_μ . C'est alors un résultat général : si une suite de fonctions réelles sur un compact est uniformément équicontinue et converge simplement, alors elle converge uniformément. Pour le démontrer dans notre cas précis, soit $\varepsilon > 0$ et K un compact. Soit $\delta > 0$ tel que l'implication de la première question est vraie. En passant à la limite, remarquons tout d'abord que $|x - y| \leq \delta$ implique $|\phi_\mu(x) - \phi_\mu(y)| \leq \varepsilon$. Par compacité, on peut recouvrir K par un nombre fini de boules $(B(x_i, \delta))_{1 \leq i \leq k}$. Soit N tel que $n \geq N$ implique $\max_{1 \leq i \leq k} |\phi_{\mu_n}(x_i) - \phi_\mu(x_i)| \leq \varepsilon$. Alors, pour $n \geq N$, soit $x \in K$. En notant $1 \leq i \leq k$ l'entier tel que $x \in B(x_i, \delta)$:

$$|\phi_{\mu_n}(x) - \phi_\mu(x)| \leq |\phi_{\mu_n}(x) - \phi_{\mu_n}(x_i)| + |\phi_{\mu_n}(x_i) - \phi_\mu(x_i)| + |\phi_\mu(x_i) - \phi_\mu(x)| \leq 3\varepsilon,$$

ce qui conclut.

En prenant par exemple $X_n = 1/n$ et $X = 0$, on a $\phi_{X_n}(t) = e^{it/n} \rightarrow \phi_X(t)$ pour tout $t \in \mathbb{R}$, mais $\|\phi_{X_n} - \phi_X\|_\infty = 2$ pour tout $n \geq 1$.

- (3) Oui. La borne de queue utilisée dans la preuve du théorème de Lévy donne, pour une constante universelle $c > 0$,

$$\begin{aligned} \sup_n \mu_n(\mathbb{R} \setminus [-K, K]) &\leq cK \sup_n \int_{-1/K}^{1/K} (1 - \operatorname{Re} \phi_{\mu_n}(u)) du \\ &\leq 2c \sup_n \sup_{|u| \leq 1/K} |1 - \phi_{\mu_n}(u)| \rightarrow 0 \quad (K \rightarrow \infty). \end{aligned}$$

En effet, $\phi_{\mu_n}(0) = 1$ pour tout n et la famille est équicontinue en zéro. Cela démontre directement la tension.

* * *

Corrigé de l'exercice 1.7. [énoncé p. 99] La réciproque est claire. Le sens direct est une conséquence du deuxième théorème de Dini. Donnons ici une approche dans le contexte de l'exercice. Fixons $k \geq 2$. L'application F_μ étant continue, croissante, de limite nulle en $-\infty$ et de limite 1 en ∞ , il existe des points $s_1 < \dots < s_{k-1}$ tels que $F_\mu(s_i) = \frac{i}{k}$ pour tout $1 \leq i \leq k-1$. Par convergence simple, pour n assez grand, pour tout $1 \leq i \leq k-1$,

$$\frac{i-1}{k} \leq F_{\mu_n}(s_i) \leq \frac{i+1}{k}.$$

Par convention, posons $s_0 = -\infty$ et $s_k = \infty$. Il vient que pour tout n assez grand, pour tout $s \in \mathbb{R}$, en choisissant s_i tel que $s_i \leq s < s_{i+1}$:

$$F_{\mu_n}(s) \leq F_{\mu_n}(s_{i+1}) \leq \frac{i+2}{k} = F_\mu(s_i) + \frac{2}{k} \leq F_\mu(s) + \frac{2}{k}$$

et de même

$$F_{\mu_n}(s) \geq F_{\mu_n}(s_i) \geq \frac{i-1}{k} = F_\mu(s_{i+1}) - \frac{2}{k} \geq F_\mu(s) - \frac{2}{k}.$$

Ainsi, pour n assez grand,

$$\sup_{s \in \mathbb{R}} |F_{\mu_n}(s) - F_\mu(s)| \leq \frac{2}{k},$$

ce qui conclut.

Corrigé de l'exercice 1.8. [énoncé p. 100]

- (1) On suit la même approche que pour l'exercice 7. Fixons $k \geq 2$. Il existe des points $s_1 < \dots < s_{k-1}$ tels que $F(s_i) = \frac{i}{k}$ pour tout $1 \leq i \leq k-1$. D'après la loi des grands nombres, presque sûrement, pour n assez grand, pour tout $1 \leq i \leq k-1$,

$$\frac{i-1}{k} \leq F_{\mu_n}(s_i) \leq \frac{i+1}{k}.$$

Par convention, posons $s_0 = -\infty$ et $s_k = \infty$. Il vient que pour tout n assez grand, pour tout $s \in \mathbb{R}$, en choisissant s_i tel que $s_i \leq s < s_{i+1}$:

$$F_{\mu_n}(s) \leq F_{\mu_n}(s_{i+1}) \leq \frac{i+2}{k} = F_{\mu}(s_i) + \frac{2}{k} \leq F_{\mu}(s) + \frac{2}{k}$$

et de même

$$F_{\mu_n}(s) \geq F_{\mu_n}(s_i) \geq \frac{i-1}{k} = F_{\mu}(s_{i+1}) - \frac{2}{k} \geq F_{\mu}(s) - \frac{2}{k}.$$

Ainsi, presque sûrement, pour n assez grand,

$$\sup_{s \in \mathbb{R}} |F_{\mu_n}(s) - F(s)| \leq \frac{2}{k}.$$

Ceci étant vrai pour tout $k \geq 2$, ceci conclut.

- (2) (a) Si $F(t) \geq x$, alors clairement $\inf\{s \in \mathbb{R} : F(s) \geq x\} \leq t$ et donc $G(x) \leq t$. Réciproquement, montrons que $F(t) < x$ implique $t < G(x)$. Puisque F est continue à droite, il existe $\varepsilon > 0$ tel que $F(t + \varepsilon) < x$. On a alors $G(x) \geq t + \varepsilon > t$.
- (b) Elles sont indépendantes de même loi et, pour tout $t \in \mathbb{R}$, la question précédente donne

$$\mathbb{P}(G(Y_1) \leq t) = \mathbb{P}(Y_1 \leq F(t)) = F(t),$$

d'où le résultat.

- (c) Les suites $(X_i)_{i \geq 1}$ et $(G(Y_i))_{i \geq 1}$ ont même loi. On peut donc, pour étudier une propriété de la suite entière, travailler avec $X_i = G(Y_i)$ pour tout i . Alors, simultanément pour tout n et tout t ,

$$F_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{G(Y_i) \leq t\}} = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{Y_i \leq F(t)\}} = A_n(F(t)).$$

- (3) D'après (1), presque sûrement,

$$\lim_{n \rightarrow \infty} \sup_{u \in [0,1]} |A_n(u) - u| = 0.$$

Comme $F(s) \in [0,1]$, on a alors presque sûrement

$$\sup_{s \in \mathbb{R}} |A_n(F(s)) - F(s)| \leq \sup_{u \in [0,1]} |A_n(u) - u| \rightarrow 0.$$

L'égalité en loi des suites entières établie en (c) transfère cette propriété presque sûre à $(F_n)_{n \geq 1}$. L'événement est mesurable, car les fonctions de répartition sont continues à droite et chaque supremum sur $\mathbb{R}$ peut être pris sur $\mathbb{Q}$.

Corrigé de l'exercice 1.9. [énoncé p. 100]

- (1) Prenons $F(x_1, x_2) = \mathbb{1}_{\{x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \geq 1\}}$. Cette fonction est continue à droite, propre et croissante en chacune de ses variables. Si elle était la fonction de répartition d'une mesure $\mu \in \mathcal{M}_1(\mathbb{R}^2)$, on aurait

$$\mu([0, 1] \times]0, 1]) = F(1, 1) - F(1, 0) - F(0, 1) + F(0, 0) = -1,$$

ce qui est impossible.

- (2) L'implication est claire (si F est la fonction de répartition de $\mu \in \mathcal{M}_1(\mathbb{R}^k)$, la propriété d'être à accroissements positifs provient du fait que $\sum_u s(u)F(u) = \mu([x, y])$ – en reprenant les notations de l'énoncé).

Posons $\mathbf{1} = (1, \dots, 1)$. Puisque F est propre et à accroissements positifs, il existe bien des mesures de probabilité μ_n à support dans $(2^{-n}\mathbb{Z})^k$ telles que $\mu_n(\{2^{-n}x\}) = F([2^{-n}(x - \mathbf{1}), 2^{-n}x])$ pour $x \in \mathbb{Z}^k$.

Ensuite, par construction, F est finiment additive sur les pavés, de sorte que pour $x \in \mathbb{Z}^k$ et $1 \leq m < n$:

$$\mu_m(2^{-m}x - \mathbf{1}, x] = \mu_n(2^{-m}x - \mathbf{1}, x]).$$

On construit X^1 de loi μ_1 , puis, conditionnellement à $X^n = z$, on choisit

$$X^{n+1} = z - 2^{-n-1}i, \quad i \in \{0, 1\}^k,$$

avec probabilité

$$\frac{\mu_{n+1}(\{z - 2^{-n-1}i\})}{\mu_n(\{z\})}.$$

Le dénominateur est strictement positif pour μ_n -presque tout z . La finie additivité des accroissements donne

$$\sum_{i \in \{0, 1\}^k} \mu_{n+1}(\{z - 2^{-n-1}i\}) = \mu_n(\{z\}),$$

de sorte que ces probabilités définissent bien une transition. On obtient par récurrence X^n de loi μ_n et, coordonnée par coordonnée, $X^m - 2^{-m}\mathbf{1} < X^n \leq X^m$ pour $m < n$. La suite est décroissante et minorée par $X^1 - 2^{-1}\mathbf{1}$; elle converge donc vers une variable $X \in \mathbb{R}^k$.

Vérifions que la fonction de répartition de X est F . Pour $x \in \mathbb{R}^k$ dyadique et n assez grand, $\mathbb{P}(X^n \leq x) = F(x)$. Comme $X^n \downarrow X$ coordonnée par coordonnée, on obtient

$$\mathbb{P}(X < x) \leq F(x) \leq \mathbb{P}(X \leq x).$$

En appliquant également cette inégalité au point dyadique $x + 2^{-m}\mathbf{1}$, on a, pour tout $m \geq 1$,

$$F(x) \leq \mathbb{P}(X \leq x) \leq \mathbb{P}(X < x + 2^{-m}\mathbf{1}) \leq F(x + 2^{-m}\mathbf{1}).$$

La continuité à droite donne $\mathbb{P}(X \leq x) = F(x)$ aux points dyadiques. En approchant un point quelconque par une suite décroissante de points dyadiques, on conclut par continuité à droite de F et continuité décroissante des probabilités.

Corrigé de l'exercice 1.10. [énoncé p. 101] Vérifions tout d'abord que $U_{\varepsilon, \delta}$ est bien ouvert. Si $x \in U_{\varepsilon, \delta}$ et $y, z \in B(x, \varepsilon)$ sont tels que $|f(y) - f(z)| > \delta$, alors pour tout x' tel que $d(x, x') < \varepsilon - \max(d(x, y), d(x, z))$ on a $x' \in U_{\varepsilon, \delta}$. En particulier, $U_{\varepsilon, \delta}$ est mesurable. On conclut en remarquant que l'ensemble des points de discontinuité de f s'écrit

$$\bigcup_{\substack{\delta > 0 \\ \delta \in \mathbb{Q}}} \bigcap_{\substack{\varepsilon > 0 \\ \varepsilon \in \mathbb{Q}}} U_{\varepsilon, \delta}.$$

* * *

Corrigé de l'exercice 1.11. [énoncé p. 101]

- (1) L'implication est claire. Pour la réciproque ; on adapte la preuve du théorème de Lévy 1.18. Comme dans cette preuve, on a l'existence d'une constante $c > 0$ telle que pour tout $A > 0$

$$\begin{aligned} \mu_n(\mathbb{R} \setminus [-A, A]) &\leq cA \int_{-1/A}^{1/A} (1 - \operatorname{Re} \phi_{\mu_n}(u)) du \\ &\xrightarrow{n \rightarrow \infty} cA \int_{-1/A}^{1/A} (1 - \operatorname{Re} \phi_{\mu}(u)) du \\ &\xrightarrow{A \rightarrow \infty} 0, \end{aligned}$$

où la première convergence a lieu en vertu de la convergence presque partout de ϕ_{μ_n} vers ϕ_{μ} et du théorème de convergence dominée, et la seconde en vertu de la continuité de ϕ_{μ} en 0.

Ainsi $(\mu_n)_{n \geq 1}$ est tendue. Supposons que $\mu_{\psi(n)} \Rightarrow \nu$. D'après le théorème de Prokhorov, il suffit de vérifier que $\mu = \nu$. On a alors $\phi_{\mu_{\psi(n)}}(t) \rightarrow \phi_{\nu}(t)$ pour tout $t \in \mathbb{R}$. Or $\phi_{\mu_{\psi(n)}}(t) \rightarrow \phi_{\mu}(t)$ pour presque tout $t \in \mathbb{R}$. Donc $\phi_{\mu} = \phi_{\nu}$ presque partout. Étant des fonctions continues, ces deux fonctions sont alors égales partout, ce qui conclut.

- (2) C'est faux, on peut par exemple prendre $\mu_n = \delta_{2\pi n!}$.

Corrigé de l'exercice 1.12. [énoncé p. 101]

- (1) Alix a tort. Par exemple si $D = \mathbb{Q}$, si $X_n = 1/n$ et $X = 0$, X_n converge en loi vers X , mais $\mathbb{P}(X_n \leq 0) \not\rightarrow \mathbb{P}(X \leq 0)$ lorsque $n \rightarrow \infty$.
- (2) Billie a raison.

Première solution. En notant E l'ensemble des fonctions de la forme $\sum_{i=1}^n \alpha_i \mathbb{1}_{[a_i, b_i]}$ avec $\alpha_i \in \mathbb{R}$ et $a_i, b_i \in D$ avec $a_i < b_i$, on a $\mathbb{E}[f(X_n)] \rightarrow \mathbb{E}[f(X)]$ pour tout $f \in E$. Or par uniforme continuité, toute fonction continue à support compact est limite uniforme de fonctions dans E , ce qui montre que X_n converge en loi vers X d'après la proposition 1.5.

Deuxième solution. Notons F et F_n les fonctions de répartition respectives de X et de X_n . Soit x un point de continuité de F . Par croissance, on a pour $u < x < v$ avec $u, v \in D$:

$$F(x) - F_n(v) \leq F(x) - F_n(x) \leq F(x) - F_n(u).$$

Ainsi, en faisant $n \rightarrow \infty$:

$$F(x) - F(v) \leq \liminf_{n \rightarrow \infty} (F(x) - F_n(x)) \leq \limsup_{n \rightarrow \infty} (F(x) - F_n(x)) \leq F(x) - F(u).$$

On conclut par continuité de F en x en faisant tendre u, v vers x .

2 Différents modes de convergence (presque sûre, en probabilité, L^1)

Cours associé : [chapitre 2](#), p. 14.

2.1 Énoncés

Exercice 2.1. [\[corrigé p. 114\]](#) (**Méthode des moments**) Soit X une variable aléatoire réelle admettant des moments de tout ordre. On suppose que la loi de X est *caractérisée par ses moments*, c'est-à-dire que si Y est une variable aléatoire admettant des moments de tout ordre vérifie $\mathbb{E}[X^k] = \mathbb{E}[Y^k]$ pour tout entier $k \geq 1$, alors X et Y ont même loi.

- (1) Soit $(Z_n)_{n \geq 1}$ une suite de variables aléatoires réelles uniformément intégrables qui converge en loi vers Z . Montrer que $\mathbb{E}[Z_n] \rightarrow \mathbb{E}[Z]$.
- (2) Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires réelles admettant des moments de tout ordre. On suppose que pour tout entier $k \geq 1$, $\mathbb{E}[X_n^k] \rightarrow \mathbb{E}[X^k]$. Montrer que $X_n \Rightarrow X$.

Remarque. Voir l'exercice 2.5 pour des conditions suffisantes pour qu'une variable aléatoire soit caractérisée par ses moments.

* * *

Exercice 2.2. [\[corrigé p. 115\]](#) Soit (E, d) un espace métrique séparable et $X, X_1, X_2, \dots$ des variables aléatoires définies sur un même espace de probabilité et à valeurs dans E . Montrer que si $X_n \rightarrow X$ presque sûrement ou en probabilité, alors $X_n \Rightarrow X$.

* * *

Exercice 2.3. [\[corrigé p. 116\]](#) Soit $(X_i)_{i \geq 1}$ des variables aléatoires réelles indépendantes de même loi.

- (1) Montrer que si $\mathbb{E}[|X_1|] < \infty$, la suite $(\max(X_1, \dots, X_n)/n)_{n \geq 2}$ est uniformément intégrable.
- (2) La réciproque est-elle vraie ?

* * *

Exercice 2.4. [\[corrigé p. 116\]](#) (**Théorème de représentation de Skorokhod pour $\mathbb{R}$**) Soit X une variable aléatoire réelle. On note F sa fonction de répartition et on pose, pour $0 < u < 1$, $F^{-1}(u) = \inf\{x \in \mathbb{R} : F(x) \geq u\}$. On rappelle (cf feuille d'exercices 1) que $F^{-1}(u) \leq x \iff u \leq F(x)$. Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires réelles qui converge en loi vers X .

- (1) Vérifier que F^{-1} est une fonction croissante continue à gauche de $]0, 1[$ dans $\mathbb{R}$.
- (2) Montrer que pour tout $u \in]0, 1[$, on a

$$F^{-1}(u) \leq \liminf_{n \rightarrow \infty} F_n^{-1}(u) \leq \limsup_{n \rightarrow \infty} F_n^{-1}(u) \leq F^{-1}(u+),$$

où $F^{-1}(u+)$ est la limite à droite de F^{-1} en u .

- (3) En déduire que si $X, X_1, X_2, \dots$ sont des variables aléatoires réelles telles que $X_n \Rightarrow X$, il existe un espace de probabilité et des variables aléatoires $X', X'_1, X'_2, \dots$ définies dessus telles que $X \stackrel{\text{loi}}{=} X', X_i \stackrel{\text{loi}}{=} X'_i$ pour tout $i \geq 1$ et X'_n converge presque sûrement vers X' lorsque $n \rightarrow \infty$.

* * *

Exercice 2.5. [corrigé p. 117] Soit X une variable aléatoire réelle admettant des moments de tout ordre. On pose $m_k = \mathbb{E}[X^k]$, $M_k = \mathbb{E}[|X|^k]$ pour $k \geq 1$ et $\phi(t) = \mathbb{E}[e^{itX}]$ pour $t \in \mathbb{R}$.

- (1) Montrer que le rayon de convergence de la série entière $\sum_{k=0}^{\infty} \frac{m_k z^k}{k!}$ est non nul si et seulement si le rayon de convergence de $\sum_{k=0}^{\infty} \frac{M_k z^k}{k!}$ est non nul.

On suppose dans la suite que ces rayons de convergence sont non nuls et on note R le rayon de convergence de $\sum_{k=0}^{\infty} \frac{M_k z^k}{k!}$.

- (2) Montrer que pour tout $h \in \mathbb{R}$,

$$\left| e^{ih} - \sum_{k=0}^K \frac{(ih)^k}{k!} \right| \leq \frac{|h|^{K+1}}{(K+1)!}.$$

- (3) Montrer que pour tout $t \in \mathbb{R}$ et $|h| < R$ on a le développement en série entière $\phi(t+h) = \sum_{k=0}^{\infty} \frac{\phi^{(k)}(t)}{k!} h^k$.
- (4) En déduire que la loi de X est caractérisée par ses moments.

* * *

Exercice 2.6. [corrigé p. 118] Soit $(X_i)_{i \in \mathbb{I}}$ une famille de variables aléatoires uniformément intégrables. En reprenant la construction de la fonction croissante $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ telle que $\lim_{x \rightarrow \infty} \frac{\phi(x)}{x} = \infty$ et $\sup_{i \in \mathbb{I}} \mathbb{E}[\phi(|X_i|)] < \infty$ (preuve du critère de de la Vallée Poussin), vérifier que ϕ peut être choisie de sorte que pour tout $x \geq 1$, $\phi'(2x) \leq 2\phi'(x)$, puis que ϕ peut être choisie de sorte que $\phi(x) \leq x^2$ pour $x \geq 1$.

* * *

Exercice 2.7. [corrigé p. 118] Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires réelles uniformément intégrable.

- (1) Pour $n \geq 1$ on pose $S_n = (X_1 + \dots + X_n)/n$. Montrer que la suite de variables aléatoires $(S_n)_{n \geq 1}$ est uniformément intégrable.
- (2) Montrer que l'adhérence dans L^1 de $\{X_n : n \geq 1\}$ forme une famille uniformément intégrable.

* * *

Exercice 2.8. [corrigé p. 119] Soit (E, d) un espace métrique et $(x_n)_{n \geq 1}$ une suite à valeurs dans E . On suppose que $(\delta_{x_n})_{n \geq 1}$ converge en loi vers une variable aléatoire X . Que dire de X ?

Exercice 2.9. [corrigé p. 119] Soit $(X_n)_{n \geq 1}$ et X des variables aléatoires à valeurs dans un espace métrique séparable. Pour chaque mode de convergence que vous connaissez, étudier si le fait que toute sous-suite de $(X_n)_{n \geq 1}$ a une sous-suite qui converge vers X implique que X_n converge vers X .

* * *

Exercice 2.10. [corrigé p. 119] Soit (X_n) et (Y_n) des variables aléatoires réelles telles que $0 \leq X_n \leq Y_n$. On suppose que X_n converge en probabilité et que Y_n converge dans L^1 . Montrer que X_n converge dans L^1 .

* * *

Exercice 2.11. [corrigé p. 119] Soient $p, q > 1$ tels que $1/p + 1/q = 1$. Soit $X \in L^q(\Omega, \mathcal{A}, \mathbb{P})$ et soit $(Y_i)_{i \in I}$ une famille de variables aléatoires bornée dans L^p (c'est-à-dire $\sup_{i \in I} \mathbb{E}[|Y_i|^p]^{1/p} < \infty$). Démontrer que la famille $(XY_i)_{i \in I}$ est uniformément intégrable.

* * *

Exercice 2.12. [corrigé p. 120] Soit $\lambda > 0$ et soit X une variable aléatoire dont la loi est donnée par $\mathbb{P}(X \geq a) = a^{-\lambda}$ pour tout $a \geq 1$. Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires i.i.d. de même loi que X et on pose

$$T_n = \left(\prod_{i=1}^n X_i \right)^{1/n}.$$

Déterminer les entiers n tels que $T_n \in L^1$. A-t-on

$$\mathbb{E}[|T_n - e^{1/\lambda}|] \rightarrow 0 \quad \text{lorsque } n \rightarrow \infty ?$$

Justifier votre réponse.

* * *

Exercice 2.13. [corrigé p. 121] Soit $(X_n)_{n \geq 1}$ et $(Y_n)_{n \geq 1}$ deux suites de variables aléatoires réelles définies sur le même espace de probabilité telles que $Y_n \geq X_n \geq 0$ et $\mathbb{E}[Y_n] = 1$ pour tout $n \geq 1$. On suppose que la suite de variables aléatoires $(X_n)_{n \geq 1}$ converge en loi vers une variable aléatoire positive X telle que $\mathbb{E}[X] = 1$.

- (1) Montrer que la suite de variables aléatoires $((X_n, Y_n))_{n \geq 1}$ converge en loi vers (X, X) .
- (2) Montrer que la suite de variables aléatoires $(Y_n)_{n \geq 1}$ est uniformément intégrable.

* * *

Exercice 2.14. [corrigé p. 122] Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires à valeurs dans $[0, 1]$ et $p > 0$.

- (1) On suppose que $\mathbb{E}[X_n] \rightarrow 1$ lorsque $n \rightarrow \infty$. Montrer que pour tout $p > 0$, $\mathbb{E}[|X_n - 1|^p] \rightarrow 0$ lorsque $n \rightarrow \infty$.
- (2) On suppose que $\mathbb{E}[X_n] \rightarrow \alpha$ et $\mathbb{E}[X_n^2] \rightarrow \alpha^2$ avec $\alpha \in [0, 1[$ lorsque $n \rightarrow \infty$. Montrer que

$$\mathbb{E}[|X_n - \alpha|^p] \rightarrow 0 \quad \text{lorsque } n \rightarrow \infty.$$

Indication. Dans les deux cas, on pourra commencer par montrer que X_n converge en probabilité.

* * *

Exercice 2.15. [corrigé p. 122] Soit $(X_i)_{i \geq 1}$ et X des variables aléatoires réelles. On pose $S_n = (X_1 + \dots + X_n)/n$ pour $n \geq 1$. On considère l'assertion suivante :

$$(A) : \quad \text{si } X_n \xrightarrow[n \rightarrow \infty]{} X, \quad \text{alors } S_n \xrightarrow[n \rightarrow \infty]{} X.$$

Est-ce que l'assertion (A) est vraie pour les modes de convergence suivants ?

(1) Convergence p.s.

(2) Convergence L^1

(3) Convergence en probabilité

(4) Convergence en loi.

Justifiez vos réponses.

* * *

Exercice 2.16. [corrigé p. 123] Soient $(X_n)_{n \geq 1}$ et $(Y_n)_{n \geq 1}$ des variables aléatoires réelles. On suppose que pour tout $n \geq 1$, la loi de X_n est absolument continue par rapport à la loi de Y_n de densité f_n . Cela signifie que $f_n : \mathbb{R} \rightarrow \mathbb{R}_+$ est une fonction mesurable telle que pour toute fonction mesurable positive $F : \mathbb{R} \rightarrow \mathbb{R}_+$ ou pour toute fonction mesurable bornée $F : \mathbb{R} \rightarrow \mathbb{R}$ on a :

$$\mathbb{E}[F(X_n)] = \mathbb{E}[f_n(Y_n)F(Y_n)].$$

(1) On suppose que $f_n(Y_n)$ converge en probabilité vers 1 lorsque $n \rightarrow \infty$ et que Y_n converge en loi vers une variable aléatoire Y . Montrer que X_n converge en loi vers Y .

(2) On suppose que f_n est continue pour tout $n \geq 1$, que la suite $(f_n)_{n \geq 1}$ converge uniformément sur $\mathbb{R}$ vers une fonction f , que Y_n converge presque sûrement vers Y et enfin que $\mathbb{E}[f(Y)] = 1$. Démontrer que X_n converge en loi.

* * *

Exercice 2.17. [corrigé p. 124] Soient $(X_k^n)_{k, n \geq 1}$ et $(X_k)_{k \geq 1}$ des variables aléatoires définies sur un même espace de probabilité, à valeurs dans un espace métrique séparable E , telles que, pour tout $k \geq 1$, presque sûrement, $X_k^n = X_k$ pour tout n assez grand. Démontrer qu'il existe une suite déterministe d'entiers $A_n \rightarrow \infty$ telle que

$$\mathbb{P}(\text{pour tout } k \in \{1, 2, \dots, A_n\} \text{ on a } X_k^n = X_k) \xrightarrow[n \rightarrow \infty]{} 1.$$

2.2 Corrigés

Corrigé de l'exercice 2.1. [énoncé p. 111]

- (1) *Première façon.* D'après le théorème de représentation de Skorokhod ([exercice 2.4](#)), puisque le caractère uniformément intégrable ne dépend que des lois individuelles, on peut supposer que la convergence $Z_n \rightarrow Z$ a lieu presque sûrement (et a fortiori

en probabilité). D'après la proposition 2.8, cette convergence a aussi lieu dans L^1 , ce qui démontre le lemme

Deuxième façon. Pour $K \geq 0$, on définit $h_K(x) = \max(-K, \min(x, K))$. Cette fonction est continue bornée. L'uniforme intégrabilité donne $C = \sup_n \mathbb{E}[|Z_n|] < \infty$. Par convergence en loi,

$$\mathbb{E}[h_K(|Z|)] = \lim_{n \rightarrow \infty} \mathbb{E}[h_K(|Z_n|)] \leq C.$$

La convergence monotone, lorsque $K \rightarrow \infty$, entraîne $\mathbb{E}[|Z|] \leq C$. De plus, $|x - h_K(x)| \leq |x| \mathbb{1}_{\{|x| > K\}}$, donc

$$|\mathbb{E}[Z_n] - \mathbb{E}[Z]| \leq \mathbb{E}[|Z_n| \mathbb{1}_{\{|Z_n| > K\}}] + |\mathbb{E}[h_K(Z_n)] - \mathbb{E}[h_K(Z)]| + \mathbb{E}[|Z| \mathbb{1}_{\{|Z| > K\}}].$$

Pour K fixé, le terme central tend vers 0. Ainsi,

$$\limsup_{n \rightarrow \infty} |\mathbb{E}[Z_n] - \mathbb{E}[Z]| \leq \sup_n \mathbb{E}[|Z_n| \mathbb{1}_{\{|Z_n| > K\}}] + \mathbb{E}[|Z| \mathbb{1}_{\{|Z| > K\}}].$$

En faisant $K \rightarrow \infty$, on conclut par uniforme intégrabilité et intégrabilité de Z .

(2) On montre la tension puis l'unicité de la limite.

Première étape : tension. Puisque $(\mathbb{E}[X_n^2])_{n \geq 2}$ converge, la suite $(X_n)_{n \geq 1}$ est tendue.

Seconde étape : unicité de la limite. Soit ϕ une extraction et Y une variable aléatoire telle que $X_{\phi(n)} \Rightarrow Y$. Soit $k \geq 1$ un entier. Puisque $(\mathbb{E}[(X_{\phi(n)})^{2k}])_{n \geq 1}$ converge, $(X_{\phi(n)}^k)_{n \geq 1}$ est uniformément intégrable. Vérifions que

$$\mathbb{E}[|Y|^k] < \infty \quad \text{et} \quad \mathbb{E}[X_{\phi(n)}^k] \rightarrow \mathbb{E}[Y^k]. \quad (3)$$

Le lemme implique (3), et on obtient $\mathbb{E}[X^k] = \mathbb{E}[Y^k]$. Puisque X est caractérisée par ses moments, on en déduit que X et Y ont la même loi, ce qui conclut.

* * *

Corrigé de l'exercice 2.2. [énoncé p. 111] Si $X_n \rightarrow X$ presque sûrement, et si $f : E \rightarrow \mathbb{R}$ est une fonction continue bornée, alors $f(X_n) \rightarrow f(X)$ presque sûrement et puisque $|f(X_n)| \leq \|f\|_\infty < \infty$, le théorème de convergence dominée implique $\mathbb{E}[f(X_n)] \rightarrow \mathbb{E}[f(X)]$.

Si $X_n \rightarrow X$ en probabilité, soit $f : E \rightarrow \mathbb{R}$ une fonction continue bornée. On raisonne par l'absurde et on suppose que $\mathbb{E}[f(X_n)] \not\rightarrow \mathbb{E}[f(X)]$. Il existe alors $\epsilon > 0$ et une extraction ϕ telle que $|\mathbb{E}[f(X_{\phi(n)})] - \mathbb{E}[f(X)]| \geq \epsilon$ pour tout $n \geq 1$. Or $X_{\phi(n)}$ converge en probabilité vers X , il existe donc une extraction ψ telle que $X_{\phi(\psi(n))} \rightarrow X$ presque sûrement. Alors $f(X_{\phi(\psi(n))}) \rightarrow f(X)$ presque sûrement. Comme $|f(X_{\phi(\psi(n))})| \leq \|f\|_\infty$, la convergence dominée donne $\mathbb{E}[f(X_{\phi(\psi(n))})] \rightarrow \mathbb{E}[f(X)]$, contradiction.

Corrigé de l'exercice 2.3. [énoncé p. 111]

- (1) Si
- $\mathbb{E} [|X_1|] < \infty$
- , on écrit

$$\left| \frac{\max(X_1, \dots, X_n)}{n} \right| \leq \frac{\sum_{k=1}^n |X_k|}{n}.$$

D'après la loi des grands nombres, la suite $\frac{\sum_{k=1}^n |X_k|}{n}$ converge dans L^1 vers $\mathbb{E} [|X_1|]$. Elle est donc uniformément intégrable, ce qui implique que la suite $(\frac{\max(X_1, \dots, X_n)}{n})_{n \geq 1}$ est uniformément intégrable.

- (2) Oui, si les variables aléatoires sont positives : le cas échéant,
- $0 \leq X_1/2 \leq \max(X_1, X_2)/2$
- et donc
- $\mathbb{E} [X_1] < \infty$
- .

Non en général : par exemple, en prenant X_1 une variable aléatoire qui a pour densité $\frac{1}{|x|^2} \mathbb{1}_{x \leq -1}$, on voit que $\max(X_1, \dots, X_n)$ a pour densité $\frac{n}{|x|^{n+1}} \mathbb{1}_{x \leq -1}$, de sorte que

$$\mathbb{E} \left[\left| \frac{\max(X_1, \dots, X_n)}{n} \right|^{3/2} \right] = \frac{1}{n^{3/2}} \int_1^\infty \frac{nx^{3/2}}{x^{n+1}} dx = \frac{1}{n^{3/2}} \frac{2n}{2n-3}.$$

La suite $(\frac{\max(X_1, \dots, X_n)}{n})_{n \geq 2}$ étant bornée dans $L^{3/2}$, elle est uniformément intégrable, bien que X_1 ne soit pas intégrable.

* * *

Corrigé de l'exercice 2.4. [énoncé p. 111]

- (1) Pour la croissance de
- F^{-1}
- , si
- $u < v$
- et si
- $F(x) \geq v$
- , on a alors
- $F(x) \geq u$
- , de sorte que
- $F^{-1}(u) \leq x$
- , et en passant à l'inf sur
- x
- on obtient
- $F^{-1}(u) \leq F^{-1}(v)$
- .

Pour le caractère continu à gauche, supposons que $u_n \uparrow u$ mais que $F^{-1}(u_n) \not\rightarrow F^{-1}(u)$. Il existe alors $\varepsilon > 0$ tel que $F^{-1}(u_n) \leq F^{-1}(u) - \varepsilon$ pour tout $n \geq 1$. Donc $u_n \leq F(F^{-1}(u) - \varepsilon)$. En passant à la limite, $u \leq F(F^{-1}(u) - \varepsilon)$ et donc $F^{-1}(u) \leq F^{-1}(u) - \varepsilon$, absurde.

- (2) Soit
- $0 < u < 1$
- .

Pour montrer que $\limsup_{n \rightarrow \infty} F_n^{-1}(u) \leq F^{-1}(u+)$, il suffit de vérifier que

$$F(x) \geq u + \varepsilon \implies \limsup_{n \rightarrow \infty} F_n^{-1}(u) \leq x.$$

On montre la contraposée : supposons qu'il existe une extraction ϕ et $\eta > 0$ tels que $x + \eta \leq F_{\phi(n)}^{-1}(u)$ pour n assez grand. On choisit un point $x_0 \in [x, x + \eta[$ en lequel F est continu. Alors $F_{\phi(n)}(x_0) < u$ pour n assez grand, et donc par convergence en loi de X_n vers X on en déduit que $F(x) \leq F(x_0) \leq u < u + \varepsilon$.

Pour montrer que

$$F^{-1}(u) \leq \liminf_{n \rightarrow \infty} F_n^{-1}(u),$$

considérons un point de continuité x de F tel que $x > \liminf_{n \rightarrow \infty} F_n^{-1}(u)$. Il existe alors une extraction ϕ telle que $x \geq F_{\phi(n)}^{-1}(u)$ pour tout $n \geq 1$. Ainsi, $F_{\phi(n)}(x) \geq u$ et en passant à la limite on obtient $F(x) \geq u$, et donc $x \geq F^{-1}(u)$. Ceci montre que $F^{-1}(u) \leq \liminf_{n \rightarrow \infty} F_n^{-1}(u)$.

- (3) Soit U une variable aléatoire uniforme sur $]0, 1[$. On pose $X' = F^{-1}(U)$ et $X'_n = F_n^{-1}(U)$. D'après la feuille d'exercices 1, $X \stackrel{\text{loi}}{=} X'$ et $X_i \stackrel{\text{loi}}{=} X'_i$ pour tout $i \geq 1$. D'après la question précédente, $X'_n(\omega) \rightarrow X'(\omega)$ si F^{-1} est continue en $U(\omega)$. Puisque F^{-1} a un nombre au plus dénombrable de discontinuités, il s'ensuit que F^{-1} est presque sûrement continue en U et le résultat désiré en découle.

* * *

Corrigé de l'exercice 2.5. [énoncé p. 112]

- (1) L'implication $\boxed{\Leftarrow}$ est claire car $|m_k| \leq M_k$. Pour l'implication $\boxed{\Rightarrow}$, l'idée est de montrer que les rayons de convergence de

$$\sum_{\substack{k=0 \\ k \text{ pair}}}^{\infty} \frac{M_k z^k}{k!} \quad \text{et} \quad \sum_{\substack{k=0 \\ k \text{ impair}}}^{\infty} \frac{M_k z^k}{k!}$$

sont non nuls. Pour le premier, ceci provient du fait que $m_k = M_k$ pour k pair, et pour le second de l'inégalité $|x|^{2j-1} \leq 1 + x^{2j}$ pour tout $x \in \mathbb{R}$ et $j \geq 1$.

- (2) La formule de Taylor avec reste intégral donne pour $h \in \mathbb{R}$ et $K \geq 1$:

$$e^{ih} = \sum_{k=0}^K \frac{(ih)^k}{k!} + \frac{(ih)^{K+1}}{K!} \int_0^1 (1-u)^K e^{iuh} du,$$

de sorte que

$$\left| e^{ih} - \sum_{k=0}^K \frac{(ih)^k}{k!} \right| \leq \frac{|h|^{K+1}}{K!} \int_0^1 (1-u)^K du = \frac{|h|^{K+1}}{(K+1)!}.$$

- (3) D'après la question précédente, on peut écrire

$$e^{i(t+h)x} = \sum_{k=0}^K \frac{(ihx)^k}{k!} e^{itx} + R_K(x, h)$$

avec $|R_K(x, h)| \leq \frac{|hx|^{K+1}}{(K+1)!}$. En particulier,

$$\phi(t+h) = \sum_{k=0}^K \frac{\mathbb{E}[(iX)^k e^{itX}]}{k!} h^k + \mathbb{E}[R_K(X, h)]$$

Or $\phi^{(k)}(t) = \mathbb{E}[(iX)^k e^{itX}]$, et par définition de R , si $|h| < R$, $\frac{|h|^{K+1} M_{K+1}}{(K+1)!} \rightarrow 0$, et le résultat s'ensuit.

- (4) Puisque ϕ caractérise la loi de X , il suffit de montrer que ϕ s'exprime à partir des moments de X . D'après la question précédente, en prenant $t = 0$, ϕ s'exprime en fonction des moments de X sur $] -R, R[$. Puis, en redéveloppant en série entière autour de points proches de R et de $-R$, on en déduit que ϕ s'exprime en fonction des moments de X sur $] -2R, 2R[$, et ensuite sur $] -nR, nR[$ pour tout $n \geq 1$ par récurrence, d'où le résultat.

Remarque. Il est possible de trouver des variables aléatoires X et Y , admettant des moments de tout ordre, telles que $\mathbb{E}[X^k] = \mathbb{E}[Y^k]$ pour tout $k \geq 1$ mais telles que leurs lois ne sont pas les mêmes. Un exemple explicite est donné par les densités

$$\frac{1}{x\sqrt{2\pi}} \exp\left(\frac{-(\ln x)^2}{2}\right) \mathbb{1}_{x>0} \quad \text{et} \quad (1 + \sin(2\pi \ln x)) \frac{1}{x\sqrt{2\pi}} \exp\left(\frac{-(\ln x)^2}{2}\right) \mathbb{1}_{x>0}.$$

* * *

Corrigé de l'exercice 2.6. [énoncé p. 112] Par uniforme intégrabilité, choisissons des seuils $(x_m)_{m \geq 1}$ tels que $x_m \geq 2m$, $x_{m+1} > 4x_m$ et

$$\sup_{i \in I} \mathbb{E}[|X_i| \mathbb{1}_{\{|X_i| \geq x_m\}}] \leq 2^{-m}.$$

Posons $h(t) = 0$ pour $t \leq 1$, $h(t) = t - 1$ pour $1 \leq t \leq 2$, et $h(t) = 1$ pour $t \geq 2$, puis

$$g(x) = 1 + \sum_{m \geq 1} h(x/x_m), \quad \phi(x) = \int_0^x g(u) du.$$

Les sommes sont localement finies. Ainsi g est continue croissante et tend vers l'infini, et ϕ est de classe C^1 , croissante, convexe, avec $\phi(x)/x \rightarrow \infty$. De plus,

$$\phi(x) \leq x + \sum_{m \geq 1} x \mathbb{1}_{\{x \geq x_m\}},$$

donc $\sup_i \mathbb{E}[\phi(|X_i|)] \leq \sup_i \mathbb{E}[|X_i|] + \sum_m 2^{-m} < \infty$. Un terme $h(2x/x_m) - h(x/x_m)$ ne peut être non nul que si $x/2 < x_m < 2x$. L'espacement des seuils garantit qu'il y a au plus un tel terme. Par conséquent,

$$\phi'(2x) = g(2x) \leq g(x) + 1 \leq 2g(x) = 2\phi'(x).$$

Enfin, $g(x) = 1$ sur $[0, 2]$, et, pour $x \geq 2$, $g(x) \leq 1 + \#\{m : 2m < x\} \leq 1 + x/2 \leq x$. Ainsi, pour $x \geq 1$,

$$\phi(x) = 1 + \int_1^x g(u) du \leq 1 + \int_1^x u du = \frac{x^2 + 1}{2} \leq x^2.$$

* * *

Corrigé de l'exercice 2.7. [énoncé p. 112] Pour les deux questions on utilise la caractérisation " ε - δ " de l'uniforme intégrabilité. Tout d'abord, puisque $(X_n)_{n \geq 1}$ est uniformément intégrable, elle est bornée dans L^1 : il existe $C > 0$ tel que $\mathbb{E}[|X_n|] \leq C$ pour tout $n \geq 1$. Aussi, pour $\varepsilon > 0$ fixé, il existe $\delta > 0$ tel que pour tout événement A avec $\mathbb{P}(A) \leq \delta$ on a $\mathbb{E}[|X_n| \mathbb{1}_A] \leq \varepsilon$ pour tout $n \geq 1$.

- (1) Par inégalité triangulaire et linéarité de l'espérance, pour tout $n \geq 1$ on a $\mathbb{E}[|S_n|] \leq C$ et pour tout événement A avec $\mathbb{P}(A) \leq \delta$ on a

$$\mathbb{E}[|S_n| \mathbb{1}_A] \leq \frac{1}{n} \sum_{k=1}^n \mathbb{E}[|X_k| \mathbb{1}_A] \leq \varepsilon,$$

ce qui montre que la suite de variables aléatoires $(S_n)_{n \geq 1}$ est uniformément intégrable.

- (2) Soit X dans l'adhérence de $\{X_n : n \geq 1\}$ pour la norme de L^1 . Il existe une suite $(Z_j)_{j \geq 1}$ d'éléments de cet ensemble telle que $\mathbb{E}[|Z_j - X|] \rightarrow 0$. Pour tout événement A avec $\mathbb{P}(A) \leq \delta$,

$$\mathbb{E}[|X|] \leq \mathbb{E}[|X - Z_j|] + C, \quad \mathbb{E}[|X| \mathbb{1}_A] \leq \mathbb{E}[|X - Z_j|] + \varepsilon.$$

En faisant $j \rightarrow \infty$, on obtient $\mathbb{E}[|X|] \leq C$ et $\mathbb{E}[|X| \mathbb{1}_A] \leq \varepsilon$. Les mêmes constantes C et δ conviennent à toute l'adhérence, qui est donc uniformément intégrable.

* * *

Corrigé de l'exercice 2.8. [énoncé p. 112] Notons $\mu = \mathcal{L}(X)$ et, pour $N \geq 1$,

$$F_N = \overline{\{x_n : n \geq N\}}.$$

Le théorème de porte-manteau donne $\mu(F_N) = 1$, puisque $\delta_{x_n}(F_N) = 1$ pour $n \geq N$. Ainsi $\mu(\bigcap_N F_N) = 1$: cette intersection est non vide. Choisissons x dans cette intersection. Par définition des F_N , il existe une sous-suite $x_{n_j} \rightarrow x$. On a alors $\delta_{x_{n_j}} \Rightarrow \delta_x$, d'où $\mu = \delta_x$ par unicité de la limite étroite. Ainsi $X = x$ presque sûrement. Enfin, pour tout $\varepsilon > 0$, porte-manteau donne $\liminf_n \mathbb{1}_{\{x_n \in B(x, \varepsilon)\}} \geq 1$, donc $x_n \rightarrow x$.

* * *

Corrigé de l'exercice 2.9. [énoncé p. 113] C'est vrai pour des modes de convergence dont la convergence est définie à travers une distance : ainsi c'est vrai pour la convergence L^p , la convergence en loi et la convergence en probabilité $X_n \rightarrow X$ en probabilité ssi $\mathbb{E}[\min(d(X_n, X), 1)] \rightarrow 0$.

En revanche, c'est faux pour la convergence p.s. Par exemple si $(X_n)_{n \geq 1}$ est une suite de variables aléatoires de Bernoulli indépendantes telle que $\mathbb{P}(X_n = 1) = 1 - \mathbb{P}(X_n = 0) = 1/n$, la suite $(X_n)_{n \geq 1}$ prend p.s. une infinité de fois les valeurs 0 et 1 par Borel-Cantelli, mais il est clair que de toute sous-suite ϕ en ré-extrayant ψ de sorte que $\phi \circ \psi(n) \geq n^2$, la suite $(X_{\phi \circ \psi(n)})_{n \geq 1}$ converge presque sûrement vers 0.

* * *

Corrigé de l'exercice 2.10. [énoncé p. 113] Puisque Y_n converge dans L^1 , la famille (Y_n) est uniformément intégrable. Pour tout $A > 0$, la domination $0 \leq X_n \leq Y_n$ donne

$$\sup_n \mathbb{E}[X_n \mathbb{1}_{\{X_n \geq A\}}] \leq \sup_n \mathbb{E}[Y_n \mathbb{1}_{\{Y_n \geq A\}}] \rightarrow 0 \quad (A \rightarrow \infty).$$

La famille (X_n) est donc uniformément intégrable. Sa convergence en probabilité entraîne sa convergence dans L^1 par la proposition 2.8.

* * *

Corrigé de l'exercice 2.11. [énoncé p. 113] *Première solution.* On utilise la caractérisation $\varepsilon - \delta$ de l'uniforme intégrabilité. Tout d'abord la famille $(XY_i)_{i \in I}$ est bornée dans L^1 , puisqu'en vertu de l'inégalité de Holder :

$$\mathbb{E}[|XY_i|] \leq \mathbb{E}[|X|^q]^{1/q} \mathbb{E}[|Y_i|^p]^{1/p}$$

et que $(Y_i)_{i \in I}$ est bornée dans L^p .

Ensuite, soit $\varepsilon > 0$ et $C = \sup_{i \in I} \mathbb{E} [|Y_i|^p]^{1/p}$. Si $C = 0$, tous les produits XY_i sont nuls presque sûrement. Sinon, puisque $|X|^q \in L^1$, il existe $\delta > 0$ tel que $\mathbb{P}(A) \leq \delta$ implique $\mathbb{E} [|X|^q \mathbb{1}_A] \leq \varepsilon^q / C^q$. Par l'inégalité de Hölder, on a alors pour tout $i \in I$:

$$\mathbb{E} [|XY_i| \mathbb{1}_A] \leq \mathbb{E} [|X|^q \mathbb{1}_A]^{1/q} \mathbb{E} [|Y_i|^p]^{1/p} \leq \varepsilon,$$

d'où le résultat.

Deuxième solution. Soit $i \in I$. On commence par utiliser l'inégalité $ab \leq a^p/p + b^q/q$ pour $a, b \geq 0$:

$$\begin{aligned} \mathbb{E} [|XY_i| \mathbb{1}_{|XY_i| > A}] &\leq \mathbb{E} [|XY_i| \mathbb{1}_{|X|^q/q + |Y_i|^p/p > A}] \\ &\leq \mathbb{E} [|XY_i| (\mathbb{1}_{|X|^q/q > A/2} + \mathbb{1}_{|Y_i|^p/p > A/2})] \\ &= \mathbb{E} [|Y_i| |X| \mathbb{1}_{|X|^q/q > A/2}] + \mathbb{E} [|X| |Y_i| \mathbb{1}_{|Y_i|^p/p > A/2}] \end{aligned}$$

Pour contrôler le premier terme, on utilise l'inégalité de Hölder :

$$\mathbb{E} [|Y_i| |X| \mathbb{1}_{|X|^q/q > A/2}] \leq \mathbb{E} [|Y_i|^p]^{1/p} \mathbb{E} [|X|^q \mathbb{1}_{|X|^q/q > A/2}]^{1/q}.$$

Le premier terme est borné uniformément en i et le deuxième terme tend vers 0 par convergence dominée car $X \in L^q(\Omega, \mathcal{A}, \mathbb{P})$.

Le contrôle de $\mathbb{E} [|X| |Y_i| \mathbb{1}_{|Y_i|^p/p > A/2}]$ est un peu plus délicat. Pour $R > 0$ on écrit en utilisant l'inégalité d'Hölder

$$\begin{aligned} \mathbb{E} (|XY_i| \mathbb{1}_{|Y_i|^p/p > A/2}) &\leq R \mathbb{E} (|Y_i| \mathbb{1}_{|Y_i|^p/p > A/2}) + \mathbb{E} (|XY_i| \mathbb{1}_{|X| > R}) \\ &\leq R \mathbb{E} (|Y_i| \mathbb{1}_{|Y_i|^p/p > A/2}) + \|X \mathbb{1}_{|X| > R}\|_q \|Y_i\|_p. \end{aligned}$$

Puisque la famille $(Y_i)_{i \in I}$ est uniformément intégrable et bornée dans L^p , on a

$$\sup_{i \in I} \mathbb{E} (|XY_i| \mathbb{1}_{|Y_i|^p/p > A/2}) \leq \underbrace{R \sup_{i \in I} \mathbb{E} (|Y_i| \mathbb{1}_{|Y_i|^p/p > A/2})}_{\rightarrow 0 \text{ quand } A \rightarrow \infty} + \|X \mathbb{1}_{|X| > R}\|_q \underbrace{\sup_{i \in I} \|Y_i\|_p}_{< \infty}.$$

Enfin, comme précédemment,

$$\|X \mathbb{1}_{|X| > R}\|_q \xrightarrow{R \rightarrow \infty} 0$$

par convergence dominée. Ceci conclut.

* * *

Corrigé de l'exercice 2.12. [énoncé p. 113] Par indépendance, $\mathbb{E} [T_n] = \mathbb{E} [X^{1/n}]^n$, et le calcul du moment donne

$$\mathbb{E} [T_n] = \begin{cases} \left(\frac{\lambda n}{\lambda n - 1} \right)^n, & \lambda n > 1, \\ +\infty, & \lambda n \leq 1. \end{cases}$$

Ainsi, $T_n \in L^1$ exactement lorsque $n > 1/\lambda$. La convergence demandée est vraie pour tout $\lambda > 0$; elle porte donc sur une queue de la suite.

En calculant par exemple sa fonction de répartition, on vérifie que $\ln(X)$ suit une loi exponentielle de paramètre λ , de sorte que T_n converge presque sûrement vers $e^{1/\lambda}$ par la loi forte des grands nombres.

La formule déjà obtenue donne

$$\mathbb{E}[T_n] = \left(1 - \frac{1}{\lambda n}\right)^{-n} \rightarrow e^{1/\lambda}.$$

Sur la queue $n > 1/\lambda$, les variables T_n sont positives et intégrables. Comme $T_n \rightarrow e^{1/\lambda}$ presque sûrement, le lemme de Scheffé donne directement

$$\mathbb{E}[|T_n - e^{1/\lambda}|] \rightarrow 0.$$

* * *

Corrigé de l'exercice 2.13. [énoncé p. 113]

- (1) La suite de variables aléatoires $(Y_n)_{n \geq 1}$ étant bornée dans L^1 , elle est tendue. La suite de variables aléatoires $(X_n)_{n \geq 1}$ est également tendue (d'après le théorème de Prokhorov, puisqu'elle converge en loi). La suite de variables aléatoires $((X_n, Y_n))_{n \geq 1}$ est donc tendue. Il existe ainsi une extraction φ et une variable aléatoire (U, V) telle que $(X_{\varphi(n)}, Y_{\varphi(n)}) \rightarrow (U, V)$ en loi, avec $U = X$ en loi. D'après le théorème de représentation de Skorokhod, on peut supposer que cette convergence a lieu p.s. On a également $X_{\varphi(n)} \leq Y_{\varphi(n)}$ p.s. (c'est une propriété qui ne dépend que de la loi de $(X_{\varphi(n)}, Y_{\varphi(n)})$). En particulier, $U \leq V$ p.s. D'après le lemme de Fatou, on a $\mathbb{E}[V] \leq 1$, et comme $\mathbb{E}[U] = 1$, on en déduit que $\mathbb{E}[V] = 1$. Donc $\mathbb{E}[U] = \mathbb{E}[V]$ avec $U \leq V$ p.s., ce qui entraîne $U = V$ p.s. Le même raisonnement s'applique à toute sous-suite convergente : chaque valeur d'adhérence a donc la loi de (X, X) . La tension implique alors la convergence de la suite entière vers cette loi.

Autre solution. La convergence en loi et la positivité donnent $1 = \mathbb{E}[X] \leq \liminf_n \mathbb{E}[X_n]$, tandis que $\mathbb{E}[X_n] \leq \mathbb{E}[Y_n] = 1$. Donc $\mathbb{E}[X_n] \rightarrow 1$. Par l'inégalité de Markov, pour tout $\varepsilon > 0$,

$$\mathbb{P}(|Y_n - X_n| > \varepsilon) \leq \frac{1}{\varepsilon} \mathbb{E}[Y_n - X_n] \rightarrow 0.$$

D'après le lemme de Slutsky, $(Y_n - X_n, X_n) \Rightarrow (0, X)$. La continuité de $(a, b) \mapsto (b, a + b)$ donne alors $(X_n, Y_n) \Rightarrow (X, X)$.

- (2) D'après le théorème de représentation de Skorokhod, on peut supposer que Y_n converge p.s. vers X . Comme $\mathbb{E}[Y_n] \rightarrow 1 = \mathbb{E}[X]$ (d'après la preuve de (1)), le lemme de Scheffé entraîne que $Y_n \rightarrow X$ dans L^1 , ce qui implique par résultat du cours que la suite $(Y_n)_{n \geq 1}$ est uniformément intégrable.

Corrigé de l'exercice 2.14. [énoncé p. 113] Dans les deux cas, montrons d'abord que X_n converge en probabilité vers α (en posant $\alpha = 1$ dans la première question).

Cas (1) : on a

$$\mathbb{P}(|X_n - 1| \geq \varepsilon) = \mathbb{P}(1 - X_n \geq \varepsilon) \leq \frac{1}{\varepsilon}(1 - \mathbb{E}[X_n]) \xrightarrow{n \rightarrow \infty} 0.$$

Cas (2) : on a

$$\begin{aligned} \mathbb{P}(|X_n - \alpha| > \varepsilon) &= \mathbb{P}((X_n - \alpha)^2 > \varepsilon^2) \leq \frac{\mathbb{E}[(X_n - \alpha)^2]}{\varepsilon^2} \\ &= \frac{\mathbb{E}[X_n^2] - 2\alpha\mathbb{E}[X_n] + \alpha^2}{\varepsilon^2} \rightarrow \frac{\alpha^2 - 2\alpha^2 + \alpha^2}{\varepsilon^2} = 0. \end{aligned}$$

Pour conclure, on remarque que la suite $(\mathbb{E}[|X_n - \alpha|^p])_{n \geq 1}$ est bornée; il suffit de montrer que 0 est sa seule valeur d'adhérence. Soit ϕ extraction telle que $(\mathbb{E}[|X_{\phi(n)} - \alpha|^p])_{n \geq 1}$ converge. Puisque X_n converge en probabilité vers α , il existe une extraction ψ telle que $X_{\phi \circ \psi(n)}$ converge presque sûrement vers α . Par convergence dominée, $(\mathbb{E}[|X_{\phi \circ \psi(n)} - \alpha|^p])_{n \geq 1}$ converge vers 0, ce qui conclut. Alternativement, on peut utiliser le théorème de super convergence dominée en remarquant que $|X_n - \alpha|^p$ est uniformément intégrable car bornée.

* * *

Corrigé de l'exercice 2.15. [énoncé p. 114]

(1) Vrai d'après le lemme de Césaro.

(2) Vrai, en écrivant par exemple

$$\mathbb{E}[|S_n - X|] \leq \frac{\sum_{k=1}^n \mathbb{E}[|X_k - X|]}{n}$$

et en utilisant le lemme de Césaro.

(3) Faux, par exemple si $(X_n)_{n \geq 1}$ sont des variables aléatoires indépendantes avec $\mathbb{P}(X_n = n) = \frac{1}{n}$ et $\mathbb{P}(X_n = 0) = 1 - 1/n$. Alors $X_n \rightarrow 0$ en probabilité. Vérifions que $S_n \not\rightarrow 0$ en probabilité. Pour $\varepsilon \in]0, 1/2[$, on remarque que

$$\mathbb{P}(S_{2n} \geq \varepsilon) \geq \mathbb{P}(\exists k \in \{n, n+1, \dots, 2n\} : X_k = k).$$

Pour $n \geq 2$, par indépendance,

$$\mathbb{P}(\exists k \in \{n, \dots, 2n\} : X_k = k) = 1 - \prod_{k=n}^{2n} \left(1 - \frac{1}{k}\right) = 1 - \frac{n-1}{2n} = \frac{n+1}{2n} \rightarrow \frac{1}{2}.$$

Ainsi $\mathbb{P}(S_{2n} \geq \varepsilon) \not\rightarrow 0$.

(4) Faux, en prenant par exemple $X_{2n} = -X_{2n+1} = N$ avec N une loi $\mathcal{N}(0, 1)$.

Corrigé de l'exercice 2.16. [énoncé p. 114]

- (1) Soit $F : \mathbb{R} \rightarrow \mathbb{R}$ continue bornée, et posons $g_n = f_n(Y_n)$. On a $g_n \geq 0$ et $\mathbb{E}[g_n] = 1$. Pour tout $\varepsilon > 0$,

$$\mathbb{E}[|g_n - 1|] = 2\mathbb{E}[(1 - g_n)_+] \leq 2\varepsilon + 2\mathbb{P}(|g_n - 1| \geq \varepsilon).$$

La convergence en probabilité donne donc $\mathbb{E}[|g_n - 1|] \rightarrow 0$. Par l'identité de changement de loi,

$$|\mathbb{E}[F(X_n)] - \mathbb{E}[F(Y_n)]| = |\mathbb{E}[(g_n - 1)F(Y_n)]| \leq \|F\|_\infty \mathbb{E}[|g_n - 1|] \rightarrow 0.$$

Comme $Y_n \Rightarrow Y$, on conclut que $X_n \Rightarrow Y$.

Autre solution. Soit F fermé, $0 < \varepsilon < 1$, $g_n = f_n(Y_n)$ et $B_n = \{|g_n - 1| > \varepsilon\}$. L'identité de changement de loi donne

$$\mathbb{P}(X_n \in F) = \mathbb{E}[g_n \mathbb{1}_{\{Y_n \in F\}}] \leq (1 + \varepsilon)\mathbb{P}(Y_n \in F) + \mathbb{E}[g_n \mathbb{1}_{B_n}].$$

Puisque $\mathbb{E}[g_n] = 1$ et $g_n \geq 1 - \varepsilon$ sur B_n^c ,

$$\mathbb{E}[g_n \mathbb{1}_{B_n}] = 1 - \mathbb{E}[g_n \mathbb{1}_{B_n^c}] \leq \varepsilon + \mathbb{P}(B_n).$$

Ainsi, par convergence en probabilité et par le théorème de porte-manteau,

$$\limsup_n \mathbb{P}(X_n \in F) \leq (1 + \varepsilon)\mathbb{P}(Y \in F) + \varepsilon.$$

On fait tendre ε vers 0 pour conclure.

Autre solution. Soit $F : \mathbb{R} \rightarrow \mathbb{R}$ continue bornée. Soit $\varepsilon > 0$. Comme $\mathbb{E}[F(Y_n)] \rightarrow \mathbb{E}[F(Y)]$, il suffit de montrer que $|\mathbb{E}[F(X_n)] - \mathbb{E}[F(Y_n)]| \rightarrow 0$. Pour cela, on écrit

$$|\mathbb{E}[F(X_n)] - \mathbb{E}[F(Y_n)]| = |\mathbb{E}[(f_n(Y_n) - 1)F(Y_n)]| \leq \|F\|_\infty \mathbb{E}[|f_n(Y_n) - 1|].$$

Il suffit donc de démontrer que $f_n(Y_n) \rightarrow 1$ dans L^1 . Ceci provient du théorème de représentation de Skorokhod, combiné avec le lemme de Scheffé car $\mathbb{E}[f_n(Y_n)] = 1$.

- (2) La fonction f est continue comme limite uniforme de fonctions continues. Posons $g_n = f_n(Y_n)$ et $g = f(Y)$. Puisque $Y_n \rightarrow Y$ presque sûrement,

$$|g_n - g| \leq \|f_n - f\|_\infty + |f(Y_n) - f(Y)| \rightarrow 0 \quad \text{presque sûrement.}$$

De plus, $g_n, g \geq 0$ et $\mathbb{E}[g_n] = \mathbb{E}[g] = 1$. Le lemme de Scheffé donne donc $\mathbb{E}[|g_n - g|] \rightarrow 0$. Pour toute fonction $F : \mathbb{R} \rightarrow \mathbb{R}$ continue bornée, l'identité de changement de loi entraîne

$$\begin{aligned} |\mathbb{E}[F(X_n)] - \mathbb{E}[gF(Y)]| &= |\mathbb{E}[g_n F(Y_n)] - \mathbb{E}[gF(Y)]| \\ &\leq \|F\|_\infty \mathbb{E}[|g_n - g|] + \mathbb{E}[|g(F(Y_n) - F(Y))|] \rightarrow 0. \end{aligned}$$

Le dernier terme tend vers zéro par convergence dominée, avec majorant $2\|F\|_\infty g$, intégrable. Ainsi, X_n converge en loi vers la mesure de probabilité de densité f par rapport à la loi de Y .

Corrigé de l'exercice 2.17. [énoncé p. 114] Par hypothèse, pour tout $k \geq 1$

$$\mathbb{P} \left(\bigcup_{n_0 \geq 1} \bigcap_{n \geq n_0} \{X_k^n = X_k\} \right) = 1.$$

L'union en n_0 étant croissante, on a donc $\mathbb{P}(X_k^n = X_k) \rightarrow 1$ lorsque $n \rightarrow \infty$, et ce pour tout $k \geq 1$.

Il existe donc une suite $(i_M)_{M \geq 1}$ d'entiers qu'on peut supposer strictement croissante telle que pour tout $M \geq 1$ et $n \geq i_M$ on a

$$\sum_{k=1}^M \mathbb{P}(X_k^n \neq X_k) \leq \frac{1}{M}.$$

Pour $1 \leq n < i_1$, on pose $A_n = 1$. Pour $n \geq i_1$, on pose $A_n = M$ si $i_M \leq n < i_{M+1}$. Vérifions que cette suite convient. Tout d'abord, on a clairement $A_n \rightarrow \infty$. On écrit

$$\begin{aligned} \mathbb{P}(\{\text{pour tout } k \in \{1, 2, \dots, A_n\} \text{ on a } X_k^n = X_k\}^c) &\leq \sum_{k=1}^{A_n} \mathbb{P}(X_k^n \neq X_k) \\ &\leq \frac{1}{A_n}, \end{aligned}$$

qui tend vers 0.

3 Mesures de probabilité sur un espace métrique

Cours associé : [chapitre 3](#), p. 18.

3.1 Énoncés

Exercice 3.1. [corrigé p. 133] Soit $(x_n)_{n \geq 1}$ une suite de nombres réels et $\mu \in \mathcal{M}_1(\mathbb{R})$ tels que $\delta_{x_n} \Rightarrow \mu$. Montrer qu'il existe $x \in \mathbb{R}$ tel que $\mu = \delta_x$.

* * *

Exercice 3.2. [corrigé p. 133] Soit (E, d) un espace métrique, $x \in E$ un élément fixe et $(X_n)_{n \geq 1}$ des variables aléatoires à valeurs dans E . Montrer que $X_n \Rightarrow x$ si et seulement si $X_n \rightarrow x$ en probabilité.

* * *

Exercice 3.3. [corrigé p. 134] (**Lemme de Slutsky**) Soient E et F deux espaces métriques séparables, $y \in F$ un élément fixé, $X, X_1, X_2, \dots$ et $(Y_i)_{i \geq 1}$ des variables aléatoires à valeurs dans respectivement E et F telles que $X_n \Rightarrow X$ et $Y_n \Rightarrow y$. Montrer que $(X_n, Y_n) \Rightarrow (X, y)$.

* * *

Exercice 3.4. [corrigé p. 134] (**Produits dénombrables d'espaces métriques**) Soit $((E_i, d_i))_{i \geq 1}$ une suite d'espaces métriques séparables. On pose $E = E_1 \times E_2 \times \dots$ et

$$\forall x = (x_i)_{i \geq 1}, y = (y_i)_{i \geq 1} \in E, \quad d(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)}.$$

- (1) Montrer que (E, d) est séparable.
- (2) Montrer qu'une suite $(x^n)_{n \geq 1} \in E$ converge vers $x \in E$ si et seulement si pour tout $k \geq 1$, $x_k^n \rightarrow x_k$ lorsque $n \rightarrow \infty$.
- (3) Montrer que (E, d) est complet si tous les E_i le sont.
- (4) Montrer que (E, d) est compact si tous les E_i le sont.

* * *

Exercice 3.5. [corrigé p. 135] (**Un exemple d'espace métrique non séparable**) Montrer que $\mathcal{C}_b(\mathbb{R})$, l'espace des fonctions continues bornées de $\mathbb{R}$ dans $\mathbb{R}$, muni de la norme uniforme n'est pas séparable.

* * *

Exercice 3.6. [corrigé p. 135] (**Principe des lois accompagnantes**) Sur un même espace de probabilité, on considère des variables aléatoires $(Y_{n,k})_{n,k \geq 1}$ et $(X_n)_{n \geq 1}$ à valeurs dans un espace métrique séparable (E, d) .

On suppose que pour tout $k \geq 1$, la suite $(Y_{n,k}, n \geq 1)$ converge en loi vers une limite Y_k . Enfin, on suppose que pour tout $\varepsilon > 0$,

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(d(X_n, Y_{n,k}) \geq \varepsilon) = 0.$$

- (1) On suppose que $Y_k \Rightarrow Y$ lorsque $k \rightarrow \infty$. Montrer que $X_n \Rightarrow Y$.
- (2) On suppose de plus que (E, d) est complet. Montrer que $(Y_k)_{k \geq 1}$ converge en loi.
On pourra utiliser avec profit la distance de Lévy–Prokhorov.

* * *

Exercice 3.7. [corrigé p. 136] Soit (E, d) un espace métrique complet et $A \subset E$. Montrer que A est précompact si, et seulement si, A est relativement compact.

* * *

Exercice 3.8. [corrigé p. 136] Soit $X, X_1, X_2, \dots$ des variables aléatoires réelles telles que $X_n \Rightarrow X$ et $X_n \geq 0$. Montrer que $\mathbb{E}[X] \leq \liminf_{n \rightarrow \infty} \mathbb{E}[X_n]$.

* * *

Exercice 3.9. [corrigé p. 136] Soit (E, d) un espace métrique séparable.

- (1) Montrer que l'application $F : x \mapsto \delta_x$ est un homéomorphisme de E sur $E_0 := \{\delta_y : y \in E\}$, où ce dernier est vu comme sous-espace de $\mathcal{M}_1(E)$ muni de la topologie de la convergence étroite.

Soit $\mu_n = \delta_{x_n}$ une suite de E_0 qui converge étroitement vers μ . On suppose par l'absurde que $(x_n)_{n \geq 1}$ n'admet aucune sous-suite convergente.

- (2) Montrer que $\{x_n, n \geq 1\}$ est un fermé, ainsi que toutes ses parties (finies ou infinies). Montrer que cela est contradictoire avec la convergence étroite de μ_n . En déduire que E_0 est un fermé de $\mathcal{M}_1(E)$.
- (3) En déduire que si $\mathcal{M}_1(E)$ est compact, alors E est compact.

* * *

Exercice 3.10. [corrigé p. 137] Soit $(E_i)_{i \geq 1}$ une suite d'espaces polonais. On pose $E = E_1 \times E_2 \times \dots$, muni de la même distance d que dans l'exercice 3.4 :

$$\forall x, y \in E, \quad d(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)}.$$

Montrer qu'une famille $\Gamma \subset \mathcal{M}_1(E)$ est tendue si et seulement si pour tout $i \geq 1$, $\{\pi_i \mu : \mu \in \Gamma\}$ est tendu, où $\pi_i : E \rightarrow E_i$ est la projection canonique et $\pi_i \mu$ la mesure image de μ par π_i .

* * *

Exercice 3.11. [corrigé p. 137] Soit (E, d) un espace métrique et $A \subset E$. On voit A comme un espace métrique muni de la distance d .

- (1) Montrer que les ouverts de (A, d) sont de la forme $A \cap O$ avec O ouvert de (E, d) .
- (2) Montrer que les éléments de $\mathcal{B}(A)$ sont de la forme $A \cap B$ avec $B \in \mathcal{B}(E)$.
- (3) Soit $K \subset A$ un compact de (A, d) . Montrer que K est fermé dans E .

Exercice 3.12. [corrigé p. 138] Montrer qu'il n'existe pas de distance sur $\mathbb{R}^{[0,1]}$ qui métrise la convergence presque partout (on pourra montrer l'existence d'une suite de fonctions qui ne converge pas presque sûrement, mais telle que de toute sous-suite on peut ré-extraire une sous-sous-suite qui converge presque sûrement vers 0).

* * *

Exercice 3.13. [corrigé p. 138] Soit (E, d) un espace métrique.

- (1) On suppose que E est séparable. Montrer que tout ouvert de E est union dénombrable de boules ouvertes.
- (2→*) La réciproque est-elle vraie ?

* * *

Exercice 3.14. [corrigé p. 139] Soit (E, d) un espace métrique séparable et complet. Soit $I : E \rightarrow [0, \infty]$ une fonction telle que pour tout $c \in [0, \infty[$, $\{x \in E : I(x) \leq c\}$ est compact. Soit $(\mathbb{P}_n)_{n \geq 1}$ une suite de $\mathcal{M}_1(E)$. On considère les propriétés (A), (B) et (C) suivantes :

(A) Pour tout fermé F de E , on a

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln \mathbb{P}_n(F) \leq -\inf_{x \in F} I(x).$$

(B) Pour tout compact K de E , on a

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln \mathbb{P}_n(K) \leq -\inf_{x \in K} I(x).$$

(C) Pour tout $M \geq 0$, il existe un compact K_M de E tel que

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln \mathbb{P}_n(E \setminus K_M) \leq -M.$$

- (1) Démontrer que (C) et (B) impliquent (A).
- (2) Démontrer que (A) implique (C).

* * *

Exercice 3.15. [corrigé p. 140] Soit $d \geq 1$ un entier. Soient $\mu, \nu \in \mathcal{M}_1(\mathbb{R}^d)$.

- (1) Démontrer que la famille $\mathcal{F}$ des mesures de probabilité de $\mathcal{M}_1(\mathbb{R}^d \times \mathbb{R}^d)$ de la forme $P_{(X,Y)}$ où X et Y sont des variables aléatoires définies sur le même espace de probabilité avec X ayant loi μ et Y ayant loi ν est compact dans $\mathcal{M}_1(\mathbb{R}^d \times \mathbb{R}^d)$ muni de la distance de Lévy-Prokhorov d_{LP} (ici on note P_U pour la loi d'une variable aléatoire U).
- (2) On suppose que $\int_{\mathbb{R}^d} |x| \mu(dx) < \infty$ et $\int_{\mathbb{R}^d} |x| \nu(dx) < \infty$. Démontrer que l'infimum

$$I = \inf\{\mathbb{E}[|X - Y|] : X \text{ a pour loi } \mu \text{ et } Y \text{ a pour loi } \nu\}$$

est atteint.

Exercice 3.16. [corrigé p. 141] Soit (E, d) un espace métrique séparable. Démontrer qu'il existe un ensemble dénombrable H de fonctions continues bornées de E dans $\mathbb{R}$ tel que pour tout $\mu \in \mathcal{M}_1(E)$ et pour toute suite $(\mu_n)_{n \geq 1}$ d'éléments de $\mathcal{M}_1(E)$, μ_n converge étroitement vers μ si et seulement si pour tout $f \in H$ on a $\mu_n(f) \rightarrow \mu(f)$.

* * *

Exercice 3.17. [corrigé p. 141] Pour $x = (x_1, x_2, \dots) \in \mathbb{R}^{\mathbb{N}^*}$, on note $\sup(x) = \sup_{i \geq 1} x_i$. Ce supremum prend ses valeurs dans $\mathbb{R} \cup \{+\infty\}$, muni de sa topologie usuelle.

- (1) Soit $X, X^1, X^2, \dots$ des variables aléatoires à valeurs dans $\mathbb{R}^{\mathbb{N}}$ (muni de la tribu et de la distance définies à la section 3.3) telles que $X^n \Rightarrow X$. Alexandra dit : on a alors $\sup(X^n) \Rightarrow \sup(X)$. A-t-elle raison ? Justifiez votre réponse.
- (2) On considère $\ell^1 = \{x \in \mathbb{R}^{\mathbb{N}} : \sum_{i=1}^{\infty} |x_i| < \infty\}$, muni de la distance $d(x, y) = \sum_{i=1}^{\infty} |x_i - y_i|$ et de la tribu borélienne associée. Soit $X, X^1, X^2, \dots$ des variables aléatoires à valeurs dans ℓ^1 telles que $X^n \Rightarrow X$. Béatrice dit : on a alors $\sup(X^n) \Rightarrow \sup(X)$. A-t-elle raison ? Justifiez votre réponse.

* * *

Exercice 3.18. [corrigé p. 142] Soit (E, d) un espace métrique et notons d_{LP} la distance de Lévy–Prokhorov sur $\mathcal{M}_1(E)$.

- (1) Montrer que

$$d_{LP}(\mu, \nu) = \inf\{r > 0 : \forall F \text{ fermé de } E, \mu(F) \leq \nu(F^r) + r\}$$

où $F^r = \{x \in E : d(x, F) < r\}$.

- (2) Que se passe-t-il si on remplace “fermé” par “ouvert” ? Autrement dit, a-t-on

$$d_{LP}(\mu, \nu) = \inf\{r > 0 : \forall O \text{ ouvert de } E, \mu(O) \leq \nu(O^r) + r\}?$$

* * *

Exercice 3.19. [corrigé p. 142] Soit μ une mesure de probabilité tendue sur un espace métrique (E, d) . Montrer qu'il existe un borélien $A \subset E$ avec (A, d) séparable et $\mu(A) = 1$.

* * *

Exercice 3.20. [corrigé p. 143] Considérons une suite $(E_n, d_n)_{n \geq 1}$ d'espaces métriques munis de leurs tribus boréliennes. Pour tout $n \geq 1$, soit μ_n une mesure de probabilité sur E_n , et $(X_n(k))_{k \geq 1}$ une suite de variables aléatoires i.i.d. à valeurs dans E_n de loi μ_n . On suppose que, pour tout $n \geq 1$, l'application $(x, y) \mapsto d_n(x, y)$ est mesurable pour la tribu produit $\mathcal{B}(E_n) \otimes \mathcal{B}(E_n)$ (c'est notamment le cas si E_n est séparable).

On note $\mathcal{M}_{\infty}(\mathbb{R}) = \mathbb{R}^{\mathbb{N}^* \times \mathbb{N}^*} = \{(a_{i,j})_{i,j \geq 1} : a_{i,j} \in \mathbb{R} \text{ pour } i, j \geq 1\}$ l'ensemble des matrices réelles “infinies”, muni de la distance

$$d((a_{i,j})_{i,j \geq 1}, (b_{i,j})_{i,j \geq 1}) = \sum_{i,j=1}^{\infty} \frac{|a_{i,j} - b_{i,j}|}{|a_{i,j} - b_{i,j}| + 1} \times \frac{1}{2^{i+j}}$$

et de sa tribu borélienne associée.

Montrer que les trois assertions suivantes sont équivalentes :

- (i) la suite $((d_n(X_n(i), X_n(j)))_{i,j \geq 1})_{n \geq 1}$ est tendue dans $\mathcal{M}_\infty(\mathbb{R})$;
- (ii) la suite $(d_n(X_n(1), X_n(2)))_{n \geq 1}$ est tendue dans $\mathbb{R}$;
- (iii) pour tout $n \geq 1$, il existe $x_n \in E_n$ tel que la suite $(d_n(x_n, X_n(1)))_{n \geq 1}$ est tendue dans $\mathbb{R}$.

* * *

Exercice 3.21. [corrigé p. 144] Soit (E, d) un espace métrique et (μ_n) une suite de mesures de probabilité sur E . On suppose que pour toute fonction continue bornée $f : E \rightarrow \mathbb{R}$ la suite $(\int f d\mu_n)_{n \geq 1}$ converge.

- (1) On suppose que E est compact. Montrer que $(\mu_n)_{n \geq 1}$ converge étroitement.
- (2) (*) On suppose que E est polonais. Montrer que $(\mu_n)_{n \geq 1}$ converge étroitement.

* * *

Exercice 3.22. [corrigé p. 145] Soit E un espace vectoriel normé séparable, $n \geq 2$ un entier et $(X_i)_{1 \leq i \leq n}$ des variables i.i.d.

- (1) On suppose que la v.a. X_1 est tendue. Montrer que la v.a. $X_1 + \dots + X_n$ est tendue.
- (2) La réciproque est-elle vraie ? Justifiez votre réponse.

* * *

Exercice 3.23. [corrigé p. 146] Soit (E, d) un espace métrique polonais. On munit l'ensemble $\mathcal{M}_1(E)$ des mesures de probabilité sur E de la distance de Lévy-Prokhorov d_{LP} . Soit $(M_n)_{n \geq 1}$ une suite de variables aléatoires à valeurs dans $\mathcal{M}_1(E)$.

- (1) Soit M une variable aléatoire à valeurs dans $\mathcal{M}_1(E)$. Montrer que les propriétés suivantes sont équivalentes :
 - (a) $d_{LP}(M_n, M) \rightarrow 0$ en probabilité.
 - (b) pour toute fonction continue bornée $f : E \rightarrow \mathbb{R}$ on a $\int f dM_n \rightarrow \int f dM$ en probabilité.
 - (c) pour toute extraction φ il existe une extraction ψ telle que presque sûrement $M_{\varphi \circ \psi(n)}$ converge vers M dans $(\mathcal{M}_1(E), d_{LP})$.

On pourra utiliser le fait suivant (conséquence de la construction du lemme 3.30, en prenant les rayons $1/N$) : il existe un ensemble dénombrable $\mathcal{F} \subset \mathcal{C}_b(E, \mathbb{R})$ tel que, pour tout $\mu \in \mathcal{M}_1(E)$ et tout $\varepsilon > 0$, il existe $f_1, \dots, f_K \in \mathcal{F}$ et $\delta > 0$ tels que

$$\left\{ \nu \in \mathcal{M}_1(E) : \left| \int f_i d\nu - \int f_i d\mu \right| < \delta \text{ pour tout } 1 \leq i \leq K \right\} \subset B(\mu, \varepsilon),$$

où $B(\mu, \varepsilon)$ désigne la boule ouverte dans $\mathcal{M}_1(E)$ centrée en μ et de rayon ε pour d_{LP} .

- (2) Soit $\mu \in \mathcal{M}_1(E)$ une mesure de probabilité (déterministe). Conditionnellement à M_n , soient Π_n^1 et Π_n^2 deux variables aléatoires indépendantes de loi M_n (c'est-à-dire que $\mathbb{E}[f(\Pi_n^1)|M_n] = \int_E f dM_n$ pour toute fonction mesurable positive f). On suppose que (Π_n^1, Π_n^2) converge en loi lorsque $n \rightarrow \infty$ vers (Π^1, Π^2) , où Π^1 et Π^2 sont deux variables aléatoires indépendantes de loi μ . Démontrer que $d_{LP}(M_n, \mu) \rightarrow 0$ en probabilité.

* * *

Exercice 3.24. [corrigé p. 147] Soit (E, d) un espace métrique séparable et X_n, Y_n des variables aléatoires à valeurs dans E . On suppose que X_n converge presque sûrement et que $\forall \varepsilon > 0, \mathbb{P}(d(X_n, Y_n) \geq \varepsilon) \rightarrow 0$.

- (1) Montrer que Y_n converge en probabilité.
- (2) Est-il vrai que Y_n converge presque sûrement ? Justifiez votre réponse (si la réponse est oui, donnez une preuve ; si la réponse est non, donnez un contre-exemple avec l'espace E de votre choix).

* * *

Problème 3.25. [corrigé p. 147]

- (1) Démontrer qu'il existe un ensemble dénombrable D de fonctions lipschitziennes à support compact de $\mathbb{R}$ dans $\mathbb{R}$ tel que pour toute suite $(\mu_n)_{n \geq 1}$ d'éléments de $\mathcal{M}_1(\mathbb{R})$ et $\mu \in \mathcal{M}_1(\mathbb{R})$, μ_n converge étroitement vers μ si et seulement si pour tout $f \in D$ on a $\mu_n(f) \rightarrow \mu(f)$.
- (2) Démontrer que toute fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ lipschitzienne à support compact peut s'écrire comme la différence de deux fonctions lipschitziennes croissantes bornées.
Indication. Choisir $a < b$ tels que $\text{supp}(f) \subset]a, b[$ et, si f est L -lipschitzienne, considérer la rampe tronquée $H(x) = L \min(\max(x - a, 0), b - a)$.
- (3) Soit Z une variable aléatoire réelle à densité. Soit $(Z_i)_{i \geq 1}$ des variables aléatoires réelles i.i.d. de même loi que Z et P_n le polynôme $P_n(X) = \prod_{i=1}^n (X - Z_i) \in \mathbb{R}[X]$. On note $(Y_i^n)_{1 \leq i \leq n-1}$ les racines de P'_n . Démontrer que presque sûrement, $\frac{1}{n-1} \sum_{i=1}^{n-1} \delta_{Y_i^n}$ converge étroitement lorsque n tend vers l'infini.
- (4) Soit $\mathcal{E} = \{a_1, \dots, a_K\} \subset \mathbb{C}$ un ensemble fini. Soit $(Z_i)_{i \geq 1}$ des variables aléatoires complexes i.i.d. à valeurs dans $\mathcal{E}$. Soit P_n le polynôme $P_n(X) = \prod_{i=1}^n (X - Z_i) \in \mathbb{C}[X]$. On note $(Y_i^n)_{1 \leq i \leq n-1}$ les racines complexes de P'_n . Démontrer que presque sûrement, $\frac{1}{n-1} \sum_{i=1}^{n-1} \delta_{Y_i^n}$ converge étroitement lorsque n tend vers l'infini.

* * *

Problème 3.26. [corrigé p. 148] On considère $\mathbb{U} = \bigcup_{n=0}^{\infty} (\mathbb{N}^*)^n$, où $(\mathbb{N}^*)^0 = \{\emptyset\}$ par convention. Un élément $u \in \mathbb{U}$ est ainsi une suite finie d'entiers strictement positifs, appelé aussi mot. Si $u \in (\mathbb{N}^*)^n$, on note $|u| = n$ sa longueur. On définit $\partial \mathbb{U} = (\mathbb{N}^*)^{\mathbb{N}^*}$ et on pose $\overline{\mathbb{U}} = \mathbb{U} \cup \partial \mathbb{U}$. Pour $u \in \mathbb{U}$ et $v \in \overline{\mathbb{U}}$, on note uv la concaténation de u et de v (en particulier $\emptyset u = u \emptyset$ pour tout mot u), et on note $u \preceq uv$.

Pour $u, v \in \overline{\mathbb{U}}$ avec $u \neq v$, on note $u \wedge v \in \mathbb{U}$ le plus long mot pour lequel il existe $x, y \in \overline{\mathbb{U}}$ tels que $u = (u \wedge v)x$ et $v = (u \wedge v)y$. On pose enfin $u \wedge u = u$.

(1) On définit pour $u, v \in \overline{\mathbb{U}}$:

$$d(u, v) = \begin{cases} 0 & \text{si } u = v, \\ e^{-|u \wedge v|} & \text{sinon.} \end{cases}$$

Montrer que d est une distance sur $\overline{\mathbb{U}}$ qui vérifie $d(u, w) \leq \max(d(u, v), d(v, w))$ pour tous $u, v, w \in \overline{\mathbb{U}}$ et qui rend cet espace polonais.

(2) Montrer que dans $\overline{\mathbb{U}}$ les boules ouvertes sont fermées.

(3) Pour $u \in \mathbb{U}$, on note $T(u) = \{uv : v \in \overline{\mathbb{U}}\}$. Soit $(\mu_n)_{n \geq 1}$ une suite de mesures de probabilité sur $\overline{\mathbb{U}}$ muni de sa tribu borélienne. Montrer que $(\mu_n)_{n \geq 1}$ converge étroitement vers une mesure de probabilité μ sur $\overline{\mathbb{U}}$ si et seulement si pour tout $u \in \mathbb{U}$,

$$\mu_n(\{u\}) \rightarrow \mu(\{u\}) \quad \text{et} \quad \mu_n(T(u)) \rightarrow \mu(T(u)).$$

* * *

Problème 3.27. [corrigé p. 149] Soit (E, d) un espace métrique séparable. On munit $\mathcal{M}_1(E)$ de la distance de Lévy-Prokhorov d_{LP} .

On pourra utiliser le fait suivant (conséquence du lemme 3.30) : pour tout $\mu \in \mathcal{M}_1(E)$ et tout $\varepsilon > 0$, il existe $f_1, \dots, f_K \in \mathcal{C}_b(E, \mathbb{R})$ et $\delta > 0$ tels que

$$\left\{ \nu \in \mathcal{M}_1(E) : \left| \int f_i d\nu - \int f_i d\mu \right| < \delta \text{ pour tout } 1 \leq i \leq K \right\} \subset B(\mu, \varepsilon),$$

où $B(\mu, \varepsilon)$ désigne la boule ouverte dans $\mathcal{M}_1(E)$ centrée en μ et de rayon ε pour d_{LP} .

Soit $\mu \in \mathcal{M}_1(E)$ une mesure de probabilité (déterministe) fixée, et $(M_n)_{n \geq 1}$ une suite de variables aléatoires à valeurs dans $\mathcal{M}_1(E)$.

Première partie.

(1) Démontrer que les assertions suivantes sont équivalentes :

- (a) pour tout $f \in \mathcal{C}_b(E, \mathbb{R})$, $\int f dM_n \rightarrow \int f d\mu$ dans L^1 .
- (b) M_n converge en loi vers μ (vue comme variable aléatoire à valeurs dans $\mathcal{M}_1(E)$);
- (c) M_n converge en probabilité vers μ (vue comme variable aléatoire à valeurs dans $\mathcal{M}_1(E)$);
- (d) pour tout $f \in \mathcal{C}_b(E, \mathbb{R})$, $\int f dM_n$ converge en probabilité vers $\int f d\mu$;
- (e) lorsque $E = \mathbb{R}$, en notant $\Phi_n(t) = \int e^{itx} M_n(dx)$ et $\Phi(t) = \int e^{itx} \mu(dx)$, on a $\Phi_n(t) \rightarrow \Phi(t)$ en probabilité pour tout $t \in \mathbb{R}$.

(2) On suppose dans cette question que $E = \mathbb{R}$ et que μ est caractérisée par ses moments.

- (a) On suppose que pour tout $k \geq 1$, $\int x^k dM_n$ converge en probabilité vers $\int x^k d\mu$. Montrer que M_n converge en loi vers μ .
- (b) On suppose que pour tout $k \geq 1$, $\mathbb{E} [\int x^k dM_n]$ converge vers $\int x^k d\mu$ et que $\text{Var}(\int x^k dM_n) \rightarrow 0$. Montrer que M_n converge en loi vers μ .

Deuxième partie. On suppose dans la suite que les variables aléatoires $(M_n)_{n \geq 1}$ sont définies sur un même espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$.

(3) Démontrer que les quatre assertions suivantes sont équivalentes :

- (a) pour presque tout $\omega \in \Omega$, M_n converge étroitement vers μ ;
- (b) pour presque tout $\omega \in \Omega$, pour tout $f \in \mathcal{C}_b(E, \mathbb{R})$, $\int f dM_n \rightarrow \int f d\mu$;
- (c) pour tout $f \in \mathcal{C}_b(E, \mathbb{R})$, pour presque tout $\omega \in \Omega$, $\int f dM_n \rightarrow \int f d\mu$;
- (d) lorsque $E = \mathbb{R}$, pour tout $t \in \mathbb{R}$, on a presque sûrement $\int e^{itx} dM_n(dx) \rightarrow \int e^{itx} \mu(dx)$.

- (4) On suppose dans cette question que $E = \mathbb{R}$ et que μ est caractérisée par ses moments. On suppose que pour tout $k \geq 1$, $\int x^k dM_n$ converge pour presque tout $\omega \in \Omega$ vers $\int x^k d\mu$. Montrer que pour presque tout $\omega \in \Omega$, M_n converge étroitement vers μ .
- (5) Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires i.i.d. de loi μ . Que dire du comportement, lorsque $n \rightarrow \infty$, de $M_n = \frac{1}{n} \sum_{k=1}^n \delta_{X_k}$?

* * *

Problème 3.28. [corrigé p. 152] Soit (E, d) un espace métrique **compact** muni de sa tribu borélienne. On note $\mathcal{M}_1(E)$ l'ensemble des mesures de probabilité sur E . Pour $\mu \in \mathcal{M}_1(E)$ et $\varepsilon > 0$ on pose

$$I_\mu(\varepsilon) = \inf_{x \in E} \mu(\overline{B}(x, \varepsilon)),$$

où $\overline{B}(x, \varepsilon) = \{y \in E : d(x, y) \leq \varepsilon\}$ désigne la boule fermée de rayon ε centrée en x .

Première partie.

- (1) Calculer $\lim_{M \rightarrow \infty} I_\mu(M)$.
- (2) Montrer que I_μ est càdlàg.
- (3) Montrer que si pour tout $x \in E$ on a $\mu(\{x\}) = 0$, alors $\lim_{\varepsilon \downarrow 0} I_\mu(\varepsilon) = 0$.
- (4) Camille dit : « La réciproque de l'énoncé de la question (3) est aussi vraie. ». A-t-il raison ? Justifiez votre réponse.

Deuxième partie. Soit $(\mu_n)_{n \geq 1}$ une suite de $\mathcal{M}_1(E)$ et $\mu \in \mathcal{M}_1(E)$.

- (5) On suppose que μ_n converge étroitement vers μ . Montrer que pour tout $\varepsilon > 0$ on a

$$\limsup_{n \rightarrow \infty} I_{\mu_n}(\varepsilon) \leq I_\mu(\varepsilon).$$

On dit qu'une mesure de probabilité sur E est *de support plein* si la masse qu'elle donne à tout ouvert non vide est strictement positive.

- (6) On suppose que μ_n converge étroitement vers μ . Montrer que les trois assertions suivantes sont équivalentes :
- (a) μ est de support plein ;
 - (b) $I_\mu(\varepsilon) > 0$ pour tout $\varepsilon > 0$;
 - (c) $\liminf_{n \rightarrow \infty} I_{\mu_n}(\varepsilon) > 0$ pour tout $\varepsilon > 0$.

Troisième partie. On suppose dans la suite que $(M_n)_{n \geq 1}$ est une suite de variables aléatoires à valeurs dans $(\mathcal{M}_1(E), d_{LP})$ et que M est une variable aléatoire à valeurs dans $(\mathcal{M}_1(E), d_{LP})$.

- (7) On suppose que M_n converge en loi vers M . Démontrer que les deux assertions suivantes sont équivalentes :
- (a) M est presque sûrement de support plein ;
 - (b) pour tous $\varepsilon, \varepsilon' > 0$ il existe $\delta, N > 0$ tels que

$$\forall n \geq N, \quad \mathbb{P}(I_{M_n}(\varepsilon) < \delta) < \varepsilon'.$$

- (8) Dominique dit : « M_n converge en loi vers M si et seulement si sur tout intervalle compact de $\mathbb{R}_+^*$, I_{M_n} converge en loi vers I_M pour la topologie de Skorokhod. » A-t-elle raison ? Justifiez votre réponse.

3.2 Corrigés

Corrigé de l'exercice 3.1. [énoncé p. 125] Tout d'abord, la suite (x_n) est bornée. En effet, sinon, il existe une extraction ϕ telle que $|x_{\phi(n)}| \rightarrow \infty$. Par portemanteau, pour tout $A > 0$, on a alors $0 = \liminf_{n \rightarrow \infty} \delta_{x_{\phi(n)}}([-A, A]) \geq \mu([-A, A]) = 0$ et donc $\mu([-A, A]) = 0$ pour tout $A > 0$, absurde.

Vérifions que (x_n) a une seule valeur d'adhérence. Si $x_{\phi(n)} \rightarrow x$, alors par portemanteau pour tout $\varepsilon > 0$, $1 = \limsup_{n \rightarrow \infty} \delta_{x_{\phi(n)}}([x - \varepsilon, x + \varepsilon]) \leq \mu([x - \varepsilon, x + \varepsilon])$. Ceci étant vrai pour tout $\varepsilon > 0$, on en déduit que $\mu(\{x\}) = 1$. Puisqu'il ne peut pas y avoir $x \neq y$ tels que $\mu(\{x\}) = \mu(\{y\}) = 1$, ceci montre que $\mu = \delta_x$ où x est l'unique valeur d'adhérence de $(x_n)_{n \geq 1}$.

* * *

Corrigé de l'exercice 3.2. [énoncé p. 125]

$\Leftarrow$ Supposons que $X_n \rightarrow x$ en probabilité. Soient $f \in \mathcal{C}_b(E)$ et $\varepsilon > 0$. Par continuité de f en x , il existe $\delta > 0$ tel que $d(y, x) < \delta$ implique $|f(y) - f(x)| < \varepsilon$. Alors

$$|\mathbb{E}[f(X_n)] - f(x)| \leq \varepsilon + 2\|f\|_\infty \mathbb{P}(d(X_n, x) \geq \delta).$$

En faisant $n \rightarrow \infty$, puis $\varepsilon \downarrow 0$, on obtient $\mathbb{E}[f(X_n)] \rightarrow f(x)$, donc $X_n \Rightarrow x$. La distance à un point fixé étant continue, donc mesurable, cet argument est valable sur tout espace métrique.

$\Rightarrow$ Pour $\varepsilon > 0$, en notant $B(x, \varepsilon)$ la boule ouverte centrée en x de rayon ε , de sorte que son complémentaire est fermé, par porte-manteau :

$$\limsup_{n \rightarrow \infty} \mathbb{P}(X_n \notin B(x, \varepsilon)) \leq \mathbb{P}(x \notin B(x, \varepsilon)) = 0.$$

Donc pour tout $\varepsilon > 0$, $\mathbb{P}(d(X_n, x) \geq \varepsilon) \rightarrow 0$, donc $X_n \rightarrow x$ en probabilité.

* * *

Corrigé de l'exercice 3.3. [énoncé p. 125] Par hypothèse, $(X_n, y) \Rightarrow (X, y)$. Munissons le produit de la distance

$$d((x_1, x_2), (y_1, y_2)) = d_E(x_1, y_1) + d_F(x_2, y_2).$$

Par hypothèse,

$$d((X_n, y), (X_n, Y_n)) = d_F(Y_n, y) \rightarrow 0 \quad \text{en probabilité.}$$

Il s'ensuit que $(X_n, Y_n) \Rightarrow (X, y)$ d'après un résultat du cours.

* * *

Corrigé de l'exercice 3.4. [énoncé p. 125] Tout d'abord, on vérifie sans difficultés que d est une distance et que pour tous $N \geq 1, x, y \in E$,

$$\sum_{n=N}^{\infty} \frac{1}{2^n} \frac{d_n(x_n, y_n)}{1 + d_n(x_n, y_n)} \leq \frac{1}{2^{N-1}}.$$

- (1) Soit D_i un ensemble dénombrable dense de E_i et $a_i \in E_i$ un élément fixé de E_i . On note D l'ensemble des suites qui s'écrivent sous la forme $(x_1, \dots, x_K, a_{K+1}, \dots)$ avec $K \geq 1$ un entier, $x_1 \in D_1, \dots, x_K \in D_K$. Soit $\varepsilon > 0$ et $y = (y_1, y_2, \dots) \in E$. On choisit N tel que $1/2^{N-1} < \varepsilon$ et un élément $x = (x_1, \dots, x_N, a_{N+1}, \dots) \in D$ tel que $d_i(x_i, y_i) \leq \varepsilon/N$ pour tout $1 \leq i \leq N$. On a alors $d(x, y) \leq 2\varepsilon$.
- (2) $\Rightarrow$ Pour tout $i \geq 1$, on a

$$\frac{1}{2^i} \frac{d_i(x_i, y_i)}{1 + d_i(x_i, y_i)} \leq d(x, y).$$

Ainsi $d(x^n, x) \rightarrow 0$ implique $d_i(x_i^n, x_i) \rightarrow 0$ pour chaque $i \geq 1$.

$\Leftarrow$ Supposons que $x_k^n \rightarrow x_k$ pour tout $k \geq 1$ lorsque $n \rightarrow \infty$. On choisit N tel que $1/2^{N-1} < \varepsilon$ et un entier n_0 tel que $n \geq n_0$ implique $d_i(x_i^n, x_i) \leq \varepsilon/N$ pour tout $1 \leq i \leq N$. Alors $d(x^n, x) \leq 2\varepsilon$ pour $n \geq n_0$.

- (3) Soit $(x^n)_{n \geq 1}$ une suite de Cauchy de (E, d) . Alors pour tout $i \geq 1$, $(x_i^n)_{n \geq 1}$ est une suite de Cauchy dans (E_i, d_i) , donc converge vers un élément noté x_i . On en déduit que $x^n \rightarrow (x_1, x_2, \dots)$ d'après la question précédente.
- (4) Soit $(x^n)_{n \geq 1}$ une suite de (E, d) . Pour $i \geq 1$, chaque suite $(x_i^n)_{n \geq 1}$ étant à valeurs dans E_i , elle admet une valeur d'adhérence. Par procédé diagonal, il existe une extraction ϕ telle que $(x_i^{\phi(n)})_{n \geq 1}$ converge pour tout $i \geq 1$ vers une valeur notée x_i . D'après la question (2), on en déduit que $x^{\phi(n)} \rightarrow (x_1, x_2, \dots)$.

Corrigé de l'exercice 3.5. [énoncé p. 125] Considérons le sous-ensemble K de $\mathcal{C}_b(\mathbb{R})$ constitué des fonctions affines sur chaque intervalle $[m, m+1]$, $m \in \mathbb{Z}$, et qui valent 0 ou 1 sur $\mathbb{Z}$. Chaque suite de $\{0, 1\}^{\mathbb{Z}}$ définit exactement une telle fonction, donc K est non dénombrable. Si $f, g \in K$ sont distinctes, leurs valeurs diffèrent en un entier et $\|f - g\|_{\infty} = 1$. Les boules ouvertes $(B(f, 1/2))_{f \in K}$ sont donc deux à deux disjointes. Toute partie dénombrable de $\mathcal{C}_b(\mathbb{R})$ manque l'une de ces boules : cet espace n'est pas séparable.

* * *

Corrigé de l'exercice 3.6. [énoncé p. 125]

- (1) On va appliquer porte-manteau. Soit F fermé, $\varepsilon > 0$. Notons $F_{\varepsilon} = \{x \in E : d(x, F) \leq \varepsilon\}$, qui est fermé. Alors

$$\mathbb{P}(X_n \in F) \leq \mathbb{P}(Y_{n,k} \in F_{\varepsilon}) + \mathbb{P}(d(X_n, Y_{n,k}) \geq \varepsilon).$$

Puisque $Y_{n,k} \Rightarrow Y_k$, par porte-manteau,

$$\limsup_{n \rightarrow \infty} \mathbb{P}(X_n \in F) \leq \mathbb{P}(Y_k \in F_{\varepsilon}) + \limsup_{n \rightarrow \infty} \mathbb{P}(d(X_n, Y_{n,k}) \geq \varepsilon).$$

Or $Y_k \Rightarrow Y$, donc en passant à la $\limsup$ quand $k \rightarrow \infty$, par porte-manteau :

$$\limsup_{n \rightarrow \infty} \mathbb{P}(X_n \in F) \leq \mathbb{P}(Y \in F_{\varepsilon}).$$

En faisant $\varepsilon \downarrow 0$, on conclut que

$$\limsup_{n \rightarrow \infty} \mathbb{P}(X_n \in F) \leq \mathbb{P}(Y \in F),$$

et donc $X_n \Rightarrow Y$ par porte-manteau.

- (2) Puisque (E, d) est séparable complet, $(\mathcal{M}_1(E), d_{LP})$ est séparable complet. Il suffit donc de montrer que $(Y_k)_{k \geq 1}$ est de Cauchy dans $(\mathcal{M}_1(E), d_{LP})$. À cet effet, on écrit

$$d_{LP}(Y_p, Y_q) \leq d_{LP}(Y_p, Y_{n,p}) + d_{LP}(Y_{n,p}, Y_{n,q}) + d_{LP}(Y_{n,q}, Y_q),$$

de sorte que

$$d_{LP}(Y_p, Y_q) \leq \limsup_{n \rightarrow \infty} d_{LP}(Y_{n,p}, Y_{n,q}).$$

Ensuite, on écrit

$$d_{LP}(Y_{n,p}, Y_{n,q}) \leq d_{LP}(Y_{n,p}, X_n) + d_{LP}(Y_{n,q}, X_n).$$

Or pour tout borélien A ,

$$\mathbb{P}(X_n \in A) \leq \mathbb{P}(Y_{n,k} \in A^{\varepsilon}) + \mathbb{P}(d(X_n, Y_{n,k}) \geq \varepsilon).$$

En choisissant K tel que

$$k \geq K \implies \limsup_{n \rightarrow \infty} \mathbb{P}(d(X_n, Y_{n,k}) \geq \varepsilon) \leq \varepsilon/2,$$

on en déduit que pour $k \geq K$, $\limsup_{n \rightarrow \infty} d_{LP}(X_n, Y_{n,k}) \leq \varepsilon$. Ainsi, pour tout $p, q \geq K$,

$$\limsup_{n \rightarrow \infty} d_{LP}(Y_{n,p}, Y_{n,q}) \leq 2\varepsilon,$$

ce qui conclut.

Corrigé de l'exercice 3.7. [énoncé p. 126]

$\boxed{\Leftarrow}$ Si $\bar{A}$ est compact, pour tout $\varepsilon > 0$ il peut être recouvert par un nombre fini de boules de rayon ε , et c'est a fortiori le cas pour A .

$\boxed{\Rightarrow}$ Tout d'abord, remarquons que $\bar{A}$ est également précompact (si A est inclus dans une union de boules rayon ε , $\bar{A}$ est inclus dans une union de boules de rayon 2ε). Par caractérisation séquentielle de la compacité dans un espace métrique, il suffit de montrer que toute suite $(x_n)_{n \geq 1}$ de $\bar{A}$ a une valeur d'adhérence. Par précompacité, on peut construire par récurrence des extractions $\psi_1, \psi_2, \dots$ et des éléments $z_1, z_2, \dots$ tels que pour tout $n \geq 1$,

$$x_{\psi_1(n)} \in B(z_1, \varepsilon), \dots, x_{\psi_1 \circ \dots \circ \psi_k(n)} \in B(z_k, \varepsilon/2^k), \dots$$

et $B(z_{k+1}, \varepsilon/2^{k+1}) \subset B(z_k, \varepsilon/2^{k-1})$. Par procédé diagonal, on pose $\phi(n) = \psi_1 \circ \dots \circ \psi_n(n)$ et on vérifie que la suite $(x_{\phi(n)})_{n \geq 1}$ est de Cauchy, donc convergente car E est complet.

* * *

Corrigé de l'exercice 3.8. [énoncé p. 126] D'après le théorème de représentation de Skorokhod, quitte à changer d'espace de probabilité, on peut supposer que la convergence $X_n \rightarrow X$ a lieu presque sûrement. En particulier, $X \geq 0$ presque sûrement et le résultat désiré découle du lemme de Fatou.

* * *

Corrigé de l'exercice 3.9. [énoncé p. 126]

- (1) Il est clair que si $x_n \rightarrow x$, alors $f(x_n) \rightarrow f(x)$ pour toute fonction $f : E \rightarrow \mathbb{R}$ continue bornée, donc F est continue. Par ailleurs, F est clairement injective : si $\delta_x = \delta_y$ avec $x \neq y$, on a alors $1 = \delta_x(\{x\}) = \delta_y(\{x\}) = 0$. Enfin, si $\delta_{x_n} \rightarrow \delta_x$, en considérant les fonctions $f_K(u) = \max(1 - Kd(x, u), 0)$ et en remarquant que $\delta_z(f_K) = 0$ si $d(z, x) > 1/K$, on voit que $x_n \rightarrow x$. Ceci conclut.
- (2) Toute partie de $\{x_n : n \geq 1\}$ est fermée : un point de son adhérence qui ne lui appartiendrait pas fournirait une sous-suite convergente de (x_n) . De plus, chaque valeur est prise un nombre fini de fois, sans quoi il existerait une sous-suite constante. On peut donc extraire une suite $(x_{\phi(n)})$ de valeurs deux à deux distinctes. Les ensembles fermés

$$F_0 = \{x_{\phi(2n)} : n \geq 1\}, \quad F_1 = \{x_{\phi(2n-1)} : n \geq 1\}$$

sont disjoints. Par porte-manteau, pour $j = 0, 1$, $1 = \limsup_n \delta_{x_{\phi(n)}}(F_j) \leq \mu(F_j)$, ce qui est impossible.

Pour montrer que E_0 est un fermé de $\mathcal{M}_1(E)$, considérons une suite $\mu_n = \delta_{x_n}$ une suite de E_0 qui converge étroitement vers μ . D'après ce qui précède, il existe une extraction telle que $(x_{\phi(n)})_{n \geq 1}$ converge vers une limite notée x . Alors $\delta_{x_{\phi(n)}} \Rightarrow \delta_x$ et donc $\mu = \delta_x$. Ainsi, E_0 est fermé.

- (3) Soit $(x_n)_{n \geq 1}$ une suite de E . Alors $(\delta_{x_n})_{n \geq 1}$ est une suite de E_0 , compact (car fermé dans $\mathcal{M}_1(E)$ compact). Il existe donc une extraction ϕ telle que $(\delta_{x_{\phi(n)}})$ converge étroitement, et d'après ce qui précède il existe $x \in E$ tel que $x_{\phi(n)} \rightarrow x$, d'où le résultat.

* * *

Corrigé de l'exercice 3.10. [énoncé p. 126]

$\Leftarrow$ Supposons que pour tout $i \geq 1$, $\{\pi_i \mu : \mu \in \Gamma\}$ est tendu. Soit $\varepsilon > 0$ et pour tout $i \geq 1$ choisissons un compact $K_i \subset E_i$ tel que pour tout $\mu \in \Gamma$ on ait $\pi_i \mu({}^c K_i) \leq \varepsilon/2^i$. On pose alors

$$K = \prod_{i=1}^{\infty} K_i = \bigcap_{i=1}^{\infty} \pi_i^{-1}(K_i).$$

Par procédé diagonal on vérifie que K est un compact de E . Par ailleurs, pour tout $\mu \in \Gamma$

$$\mu({}^c K) \leq \sum_{i=1}^{\infty} (\mu(\pi_i^{-1}({}^c K_i))) = \sum_{i=1}^{\infty} (1 - \pi_i \mu(K_i)) \leq \varepsilon.$$

$\Rightarrow$ Soit $\varepsilon > 0$ et K un compact de E tel que $\mu(K) \geq 1 - \varepsilon$ pour tout $\mu \in \Gamma$. Par continuité des projections canoniques, les $K_i = \pi_i(K)$ sont compacts. Par ailleurs, pour $i \geq 1$ et $\mu \in \Gamma$, puisque $K \subset \pi_i^{-1}(K_i)$, on a

$$\pi_i \mu(K_i) \geq \mu(K) \geq 1 - \varepsilon.$$

* * *

Corrigé de l'exercice 3.11. [énoncé p. 126]

- (1) L'inclusion $I : (A, d) \rightarrow (E, d)$ étant continue, l'image réciproque de tout ouvert est un ouvert, donc les éléments de la forme $A \cap O$ avec O ouvert de (E, d) sont des ouverts de (A, d) .

Réciproquement, soit U un ouvert de (A, d) . Pour tout $x \in U$, il existe alors $r_x > 0$ tel que $B_A(x, r_x) \subset U$. Alors

$$U = \bigcup_{x \in U} B_A(x, r_x).$$

Or $B_A(x, r_x) = B_E(x, r_x) \cap A$. Donc

$$U = \bigcup_{x \in U} (B_E(x, r_x) \cap A) = \left(\bigcup_{x \in U} B_E(x, r_x) \right) \cap A,$$

d'où le résultat en prenant $O = \bigcup_{x \in U} B_E(x, r_x)$, ouvert dans E .

- (2) L'inclusion $I : (A, d) \rightarrow (E, d)$ étant continue donc mesurable, on en déduit que les éléments de la forme $A \cap B$ avec $B \in \mathcal{B}(E)$ sont des éléments de $\mathcal{B}(A)$.

Réciproquement, les éléments de la forme $A \cap B$ avec $B \in \mathcal{B}(E)$ contiennent les ouverts de A (d'après la question précédente) et forment une tribu, donc contiennent les éléments de $\mathcal{B}(A)$.

- (3) Soit $(x_n)_{n \geq 1}$ une suite d'éléments de K qui converge vers $x \in E$. Par compacité dans K , il existe une extraction ϕ et $y \in K$ tels que $x_{\phi(n)} \rightarrow y$. Ainsi $y = x \in K$, donc K est fermé.

* * *

Corrigé de l'exercice 3.12. [énoncé p. 127] Pour $n \geq 1$, écrivons $n = 2^N + k$ avec $0 \leq k < 2^N$, et posons

$$f_n(x) = \mathbb{1}_{[k2^{-N}, (k+1)2^{-N}]}(x).$$

À chaque niveau N , chaque point de $[0, 1]$ appartient à au moins un intervalle du balayage et, pour $N \geq 2$, n'appartient pas à tous. La suite $(f_n(x))$ prend donc indéfiniment les valeurs 0 et 1 : elle ne converge en aucun point. En revanche, $\text{Leb}(\{f_n = 1\}) = 2^{-N} \rightarrow 0$. De toute sous-suite, on peut ainsi extraire (f_{n_j}) telle que $\sum_j \text{Leb}(\{f_{n_j} = 1\}) < \infty$. Par Borel–Cantelli, $f_{n_j} \rightarrow 0$ presque partout. Le critère de convergence par sous-suites dans un espace métrique exclut donc toute distance métrisant la convergence presque partout. Ce raisonnement reste valable sur le quotient des fonctions mesurables modulo l'égalité presque partout.

* * *

Corrigé de l'exercice 3.13. [énoncé p. 127]

- (1) Soit $(x_i)_{i \geq 1}$ une suite dénombrable dense de E . Si U est un ouvert, on écrit

$$U = \bigcup_{i, m \geq 1: B(x_i, 1/m) \subset U} B(x_i, 1/m).$$

En effet, l'inclusion $\supset$ est claire. Pour l'inclusion $\subset$, soit $x \in U$ et $\varepsilon > 0$ tel que $B(x, \varepsilon) \subset U$. Choisissons un entier $m > 2/\varepsilon$, puis x_i tel que $d(x, x_i) < 1/(2m)$. Alors $x \in B(x_i, 1/m)$ et, pour $z \in B(x_i, 1/m)$, $d(z, x) < 3/(2m) < \varepsilon$, donc $B(x_i, 1/m) \subset U$.

- (2) Oui! Supposons que tout ouvert de E est union dénombrable de boules ouvertes. On raisonne par l'absurde en supposant que (E, d) est non séparable.

Vérifions d'abord que qu'il existe $r > 0$ et un sous-ensemble $E_1 \subset E$ non dénombrable tel que pour tous $x \neq y$ dans E_1 on a $d(x, y) > r$. Pour chaque $r = 1/n$, considérons un sous-ensemble maximal (pour l'inclusion) de E ayant cette propriété (qui existe d'après le lemme de Zorn). S'ils sont tous dénombrables, alors clairement E est séparable.

Considérons maintenant l'ensemble $E_2 \subset E_1$ des points $x \in E_1$ pour lesquels il existe un point $y_x \in E$ tel que $0 < d(x, y_x) < r/10$.

Cas 1. E_2 est non dénombrable. On pose alors

$$U = \bigcup_{x \in E_2} B(x, d(x, y_x)),$$

qui est une union disjointe. Considérons une boule ouverte $B(z, a) \subset U$. Puisque $z \in U$, il existe $x \in E$ tel que $d(z, x) < d(x, y_x)$. Mais

$$a \leq d(z, y_x) \leq d(z, x) + d(x, y_x) < 2d(x, y_x) < r/5$$

car $y_x \notin U$ et a fortiori $y_x \notin B(z, a)$. Par conséquent $B(z, a)$ est inclus dans une unique boule $B(x, d(x, y_x))$, et il faut donc un nombre non dénombrable de boules ouvertes pour couvrir U , absurde.

Cas 2. $E_3 = E_1 \setminus E_2$ est non dénombrable. Chaque point de E_3 est isolé dans E , donc toute partie de E_3 est ouverte dans E . Pour $x \in E_3$, posons

$$R(x) = \sup\{a > 0 : B(x, a) \cap E_3 \text{ est dénombrable}\}.$$

Alors $0 < R(x) < \infty$ et $B(x, R(x)) \cap E_3$ est dénombrable. Pour toute partie dénombrable $S \subset E_3$ et tout $x \in E_3 \setminus S$, on peut choisir des points distincts $z_j \in E_3 \setminus (S \cup \{x\})$ tels que $R(x) \leq d(x, z_j) < R(x) + 1/j$: chaque boule de rayon strictement supérieur à $R(x)$ est non dénombrable, tandis que la boule de rayon $R(x)$ est dénombrable. Appelons étoile centrée en x un ensemble $\{x\} \cup \{z_j : j \geq 1\}$ ainsi construit. Par Zorn, il existe une famille maximale d'étoiles deux à deux disjointes. Elle est non dénombrable : sinon leur réunion serait dénombrable et la construction précédente permettrait d'ajouter une étoile disjointe. Soit U l'ensemble de leurs centres. Si $U = \bigcup_{i \geq 1} B(x_i, r_i)$, les centres x_i appartiennent à U . Pour au moins un i , $r_i > R(x_i)$, sinon chaque boule, étant incluse dans $U \subset E_3$, serait dénombrable. Mais cette boule contient alors des points z_j de l'étoile centrée en x_i , alors que cette étoile rencontre U seulement en x_i . Contradiction.

* * *

Corrigé de l'exercice 3.14. [énoncé p. 127]

- (1) Supposons (B) et (C). Soit F fermé et $\alpha = \inf_{x \in F} I(x)$. Si $\alpha = 0$, la majoration demandée découle de $\mathbb{P}_n(F) \leq 1$. Sinon, fixons $0 \leq c < \alpha$ et un compact K_c donné par (C). Comme $F \cap K_c$ est compact, (B) implique

$$\limsup_n \frac{1}{n} \ln \mathbb{P}_n(F \cap K_c) \leq -\alpha \leq -c.$$

De plus, $\limsup_n n^{-1} \ln \mathbb{P}_n(E \setminus K_c) \leq -c$. La majoration

$$\mathbb{P}_n(F) \leq \mathbb{P}_n(F \cap K_c) + \mathbb{P}_n(E \setminus K_c)$$

donne donc $\limsup_n n^{-1} \ln \mathbb{P}_n(F) \leq -c$. Faire tendre c vers α si $\alpha < \infty$, ou vers $+\infty$ si $\alpha = +\infty$, conclut.

- (2) Soit $(x_i)_{i \geq 1}$ une suite dense dans E et fixons $M > 0$. Pour $k, \ell \geq 1$, posons

$$O_{k,\ell} = \bigcup_{i=1}^{\ell} B(x_i, 1/k).$$

Fixons $M' > M$. Par compacité de $\{I \leq kM'\}$, il existe j_k tel que $\{I \leq kM'\} \subset O_{k,j_k}$. Par (A),

$$\limsup_n \frac{1}{n} \ln \mathbb{P}_n(E \setminus O_{k,j_k}) \leq -kM'.$$

Il existe donc n_k tel que $\mathbb{P}_n(E \setminus O_{k,j_k}) \leq e^{-n_k M}$ pour $n \geq n_k$. Puisque $O_{k,\ell} \uparrow E$ lorsque $\ell \rightarrow \infty$, on peut agrandir j_k en ℓ_k pour obtenir

$$\mathbb{P}_n(E \setminus O_{k,\ell_k}) \leq e^{-n_k M} \quad \text{pour tout } n \geq 1.$$

Posons $K_M = \bigcap_{k \geq 1} \overline{O_{k,\ell_k}}$. Cet ensemble est fermé et précompact, car pour chaque k il est inclus dans une union finie de boules fermées de rayon $1/k$. Il est donc compact dans E complet. Enfin,

$$\mathbb{P}_n(E \setminus K_M) \leq \sum_{k \geq 1} e^{-n_k M} = \frac{e^{-nM}}{1 - e^{-nM}},$$

d'où $\limsup_n n^{-1} \ln \mathbb{P}_n(E \setminus K_M) \leq -M$. Pour le niveau $M = 0$, le compact obtenu au niveau 1 convient.

Remarque. La propriété (C) s'appelle tension exponentielle. Cet exercice est inspiré des Lemma 2.5 et Lemma 2.6 dans James Lynch. Jayaram Sethuraman. "Large Deviations for Processes with Independent Increments." Ann. Probab. 15 (2) 610 - 627, April, 1987. <https://doi.org/10.1214/aop/1176992161> et constitue à présent une brique de base de la théorie des grandes déviations.

* * *

Corrigé de l'exercice 3.15. [énoncé p. 127]

- (1) Soit $(\mathbb{P}_{(X_n, Y_n)})_{n \geq 1}$ une suite d'éléments de $\mathcal{F}$. On montre que cette suite admet une sous-suite qui converge étroitement vers un élément de $\mathcal{F}$. Montrons d'abord que cette suite est tendue.

Soit $\varepsilon > 0$. Soient K_1 et K_2 deux compacts de $\mathbb{R}^d$ tels que $\mu(K_1) \geq 1 - \varepsilon$ et $\nu(K_2) \geq 1 - \varepsilon$. Alors pour tout $n \geq 1$:

$$\mathbb{P}((X_n, Y_n) \in K_1 \times K_2) \geq 1 - 2\varepsilon.$$

Ainsi, d'après le théorème de Prokhorov, il existe une sous-suite φ et une variable aléatoire (U, V) telle que $(X_{\varphi(n)}, Y_{\varphi(n)})$ converge en loi vers (U, V) . Il suffit de vérifier que U et X ont la même loi, ainsi que V et Y . Pour cela, la continuité de $(x, y) \mapsto x$ et $(x, y) \mapsto y$ implique que $X_{\varphi(n)}$ converge en loi vers U et $Y_{\varphi(n)}$ converge en loi vers V . Puisque $X_{\varphi(n)}$ a la même loi que X et $Y_{\varphi(n)}$ a la même loi que Y , le résultat s'ensuit.

- (2) Soit $P_{(X_n, Y_n)}$ une suite d'éléments de $\mathcal{F}$ telle que

$$\mathbb{E} [|X_n - Y_n|] \rightarrow I.$$

D'après la question (1), il existe une sous-suite φ et un élément $P_{(X, Y)}$ de $\mathcal{F}$ tels que $(X_{\varphi(n)}, Y_{\varphi(n)})$ converge en loi vers (X, Y) . Quitte à renuméroter cette sous-suite, supposons que (X_n, Y_n) converge en loi vers (X, Y) . Montrons que

$$\mathbb{E} [|X - Y|] = I.$$

Pour tout $R \geq 0$, la fonction $(x, y) \mapsto |x - y| \wedge R$ étant continue bornée,

$$\mathbb{E}[|X_n - Y_n| \wedge R] \xrightarrow{n \rightarrow \infty} \mathbb{E}[|X - Y| \wedge R].$$

Puisque $\mathbb{E}[|X_n - Y_n| \wedge R] \leq \mathbb{E}[|X_n - Y_n|]$, on en déduit que

$$\mathbb{E}[|X - Y| \wedge R] \leq I.$$

En faisant $R \rightarrow \infty$ on obtient $\mathbb{E}[|X - Y|] \leq I$. On a clairement $I \leq \mathbb{E}[|X - Y|]$ par définition de I , d'où le résultat.

Remarque. La famille $\mathcal{F}$ est appelée ensemble des couplages de X et Y , et le couplage de la question (2) est appelé couplage (ou transport) optimal pour le coût L^1 . Cet exercice est inspiré de la Section 5.4.3 de [ce document](#).

* * *

Corrigé de l'exercice 3.16. [énoncé p. 128] Soit $(x_i)_{i \geq 1}$ une suite dense dans E . Pour $K, N \geq 1$ et $\emptyset \neq I \subset \{1, \dots, K\}$, posons

$$f_I^{K,N}(x) = \max \left(1 - Nd \left(x, \bigcup_{i \in I} B(x_i, 1/N) \right), 0 \right),$$

et posons $f_{\emptyset}^{K,N} = 0$. La famille H de ces fonctions est dénombrable et incluse dans $\mathcal{C}_b(E)$. La convergence étroite entraîne la convergence des intégrales de ses éléments. Réciproquement, supposons $\mu_n(f) \rightarrow \mu(f)$ pour tout $f \in H$. Fixons $\varepsilon > 0$, choisissons un entier $N > 3/\varepsilon$, puis K tel que

$$\mu \left(\bigcup_{i=1}^K B(x_i, 1/N) \right) > 1 - 1/N.$$

Les fonctions $f_I^{K,N}$ sont exactement celles construites dans le lemme 3.30, avec $\eta = 1/N$. Pour n assez grand, les 2^K différences $|\mu_n(f_I^{K,N}) - \mu(f_I^{K,N})|$ sont toutes inférieures à $1/N$. Le lemme donne donc

$$d_{LP}(\mu_n, \mu) \leq 3/N < \varepsilon.$$

Ainsi $d_{LP}(\mu_n, \mu) \rightarrow 0$, donc $\mu_n \xrightarrow{\text{étroite}} \mu$.

* * *

Corrigé de l'exercice 3.17. [énoncé p. 128]

- (1) Non. Prenons $X^n = e_n$ déterministe, où $(e_n)_k = \mathbb{1}_{\{k=n\}}$. Chaque coordonnée de e_n tend vers 0, donc $e_n \rightarrow 0$ dans $\mathbb{R}^{\mathbb{N}}$, et a fortiori $X^n \Rightarrow 0$. Pourtant $\sup(X^n) = 1$ pour tout n , alors que $\sup(0) = 0$.
- (2) Oui! On vérifie que la fonction

$$\begin{aligned} \sup : \ell^1 &\rightarrow \mathbb{R} \\ (x_i)_{i \geq 1} &\mapsto \sup_{i \geq 1} x_i \end{aligned}$$

est continue (c'est un argument déterministe), et le résultat s'ensuit par stabilité de la convergence en loi par composition par une fonction continue.

Corrigé de l'exercice 3.18. [énoncé p. 128]

(1) Notons

$$d_F(\mu, \nu) = \inf\{r > 0 : \forall F \text{ fermé de } E, \mu(F) \leq \nu(F^r) + r\}.$$

Tout d'abord, on a $d_{LP}(\mu, \nu) \geq d_F(\mu, \nu)$. En effet, si on choisit $r > d_{LP}(\mu, \nu)$, on a alors

$$\mu(A) \leq \nu(A^r) + r$$

pour tout borélien A et a fortiori pour tout A fermé, ce qui implique que $d_F(\mu, \nu) \leq r$. Ainsi $d_{LP}(\mu, \nu) \geq d_F(\mu, \nu)$.

Choisissons maintenant $r > d_F(\mu, \nu)$, de sorte que

$$\mu(F) \leq \nu(F^r) + r$$

pour tout fermé F . Pour un borélien A , on a alors

$$\mu(A) \leq \mu(\overline{A}) \leq \nu(\overline{A}^r) + r = \nu(A^r) + r$$

car $\overline{A}^r = A^r$ (en effet, l'inclusion $A^r \subset \overline{A}^r$ est claire, et pour l'autre inclusion, si $x \in \overline{A}^r$, on trouve $y \in \overline{A}$ tel que $d(x, y) < r$, puis $z \in A$ tel que $d(y, z) < r - d(x, y)$, de sorte que $d(x, z) < r$). Cela entraîne que $d_{LP}(\mu, \nu) \leq r$, ce qui conclut.

(2) Oui! Comme à la question précédente, on a $d_{LP}(\mu, \nu) \geq d_O(\mu, \nu)$. Pour l'autre inégalité, soit $r > d_O(\mu, \nu)$ de sorte que

$$\mu(O) \leq \nu(O^r) + r$$

pour tout ouvert O . Soit A un borélien; nous allons vérifier que $\mu(A) \leq \nu(A^{r+\eta}) + r + \eta$ pour tout $\eta > 0$, ce qui impliquera que $d_{LP}(\mu, \nu) \leq r + \eta$ et conclura. Pour cela, on écrit

$$\mu(A) \leq \mu(A^\eta) \leq \nu((A^\eta)^r) + r \leq \nu(A^{\eta+r}) + r + \eta.$$

* * *

Corrigé de l'exercice 3.19. [énoncé p. 128] Par tension, pour tout $n \geq 1$, il existe un compact $K_n \subset E$ tel que $\mu(K_n) \geq 1 - 1/n$. Posons

$$A = \bigcup_{n=1}^{\infty} K_n.$$

Alors pour tout $n \geq 1$, $\mu(A) \geq 1 - 1/n$, de sorte qu'en faisant $n \rightarrow \infty$ on obtient $\mu(A) = 1$. Par ailleurs, A est séparable, étant union dénombrable d'ensembles séparables (puisqu'ils sont compacts).

Corrigé de l'exercice 3.20. [énoncé p. 128]

(i) $\implies$ (ii) Ceci découle que l'application

$$\begin{aligned} \mathcal{M}_\infty(\mathbb{R}) &\rightarrow \mathbb{R} \\ (a_{i,j})_{i,j \geq 1} &\mapsto a_{1,2} \end{aligned}$$

est continue.

(ii) $\implies$ (i) Soit $\varepsilon > 0$. Pour $i, j \geq 1$, soit $C_{i,j} > 0$ tel que pour tout $n \geq 1$ on a

$$\mathbb{P}(d_n(X_n(1), X_n(2)) \geq C_{i,j}) \leq \frac{\varepsilon}{2^{i+j}}.$$

On vérifie (en utilisant le procédé d'extraction diagonal) que

$$K = \{(a_{i,j})_{i,j \geq 1} \in \mathcal{M}_\infty(\mathbb{R}) : |a_{i,j}| \leq C_{i,j} \text{ pour tous } i, j \geq 1\}$$

est un compact de $\mathcal{M}_\infty(\mathbb{R})$, et que

$$\begin{aligned} \mathbb{P}((d_n(X_n(i), X_n(j)))_{i,j \geq 1} \notin K) &\leq \sum_{i,j=1}^{\infty} \mathbb{P}(d_n(X_n(i), X_n(j)) \geq C_{i,j}) \\ &\leq \sum_{i,j=1}^{\infty} \mathbb{P}(d_n(X_n(1), X_n(2)) \geq C_{i,j}) \\ &\leq \sum_{i,j=1}^{\infty} \frac{\varepsilon}{2^{i+j}} \\ &= \varepsilon. \end{aligned}$$

(iii) $\implies$ (ii) Soit $x_n \in E_n$ tel que la suite $(d_n(x_n, X_n(1)))_{n \geq 1}$ est tendue dans $\mathbb{R}$. Fixons $\varepsilon > 0$ et $C > 0$ tel que $\mathbb{P}(d_n(x_n, X_n(1)) \geq C/2) \leq \varepsilon/2$. Alors par inégalité triangulaire

$$\begin{aligned} \mathbb{P}(d_n(X_n(1), X_n(2)) > C) &\leq \mathbb{P}(d_n(x_n, X_n(1)) > C/2) + \mathbb{P}(d_n(x_n, X_n(2)) > C/2) \\ &= 2\mathbb{P}(d_n(x_n, X_n(1)) > C/2) \\ &\leq \varepsilon \end{aligned}$$

(ii) $\implies$ (iii) Pour tout $\varepsilon > 0$, soit $C_\varepsilon > 0$ tel que $\mathbb{P}(d_n(X_n(1), X_n(2)) > C_\varepsilon) \leq \varepsilon$ pour tout $n \geq 1$

Puisque d'après le théorème de Fubini

$$\mathbb{P}(d_n(X_n(1), X_n(2)) > C) = \int_{E_n} \mu_n(dx) \mathbb{P}(d_n(x, X_n(1)) > C),$$

il existe $x_n^\varepsilon \in E_n$ tel que $\mathbb{P}(d_n(x_n^\varepsilon, X_n(1)) > C_\varepsilon) \leq \varepsilon$ pour tout $n \geq 1$.

Vérifions que $x_n = x_n^{1/4}$ convient. Pour $\varepsilon < 1/2$, l'événement

$$\{d_n(x_n^{1/4}, X_n(1)) \leq C_{1/4}\} \cap \{d_n(x_n^\varepsilon, X_n(1)) \leq C_\varepsilon\}$$

est non vide car de probabilité au moins $1 - \varepsilon - 1/4$. On en déduit que $d_n(x_n^{1/4}, x_n^\varepsilon) \leq C_\varepsilon + C_{1/4}$. Ainsi,

$$\mathbb{P}(d_n(x_n^{1/4}, X_n(1)) > 2C_\varepsilon + C_{1/4}) \leq \mathbb{P}(d_n(x_n^\varepsilon, X_n(1)) > C_\varepsilon) \leq \varepsilon,$$

ce qui conclut.

Corrigé de l'exercice 3.21. [énoncé p. 129]

- (1) Pour f continue bornée, notons $\ell(f)$ la limite de $(\int f d\mu_n)_{n \geq 1}$. Puisque E est compact, $(\mu_n)_{n \geq 1}$ est tendue. Soit ψ extraction et μ mesure de probabilité sur E tels que $\mu_{\psi(n)} \Rightarrow \mu$. En particulier, $\ell(f) = \int f d\mu$.

Vérifions que $\mu_n \Rightarrow \mu$. D'après le théorème de Prokhorov, il suffit de vérifier que si $\mu_{\phi(n)} \Rightarrow \nu$ alors $\nu = \mu$. Par hypothèse $\mu_{\phi(n)}(f) \rightarrow \ell(f)$, et donc $\int f d\nu = \ell(f) = \int f d\mu$. On conclut que $\nu = \mu$.

- (2) Posons $\ell(f) = \lim_n \mu_n(f)$ pour $f \in \mathcal{C}_b(E)$. Nous allons montrer que (μ_n) est tendue. La conclusion suivra ensuite exactement comme en (1), par Prokhorov et unicité de la limite.

Supposons par l'absurde que (μ_n) ne soit pas tendue. Il existe alors $r > 0$ et $\eta > 0$ tels que, pour toute partie finie $F \subset E$, il existe n vérifiant

$$\mu_n \left(E \setminus \bigcup_{x \in F} B(x, r) \right) > \eta.$$

En effet, si cette propriété était fausse, pour tout $\varepsilon > 0$ et chaque $k \geq 1$, on pourrait choisir une réunion finie de boules de rayon 2^{-k} dont la masse est au moins $1 - \varepsilon 2^{-k}$ pour toutes les μ_n . L'intersection des unions correspondantes de boules fermées serait fermée et précompacte, donc compacte puisque E est complet, et de masse au moins $1 - \varepsilon$ pour toutes les μ_n . De plus, le témoin n peut être choisi arbitrairement grand : les mesures en nombre fini correspondant aux indices initiaux sont tendues, et peuvent être absorbées en ajoutant un nombre fini de boules à la réunion.

Posons $n_0 = 0$ et construisons par récurrence une extraction (n_j) , des compacts K_j, H_j et des fonctions $h_j : E \rightarrow [0, 1]$. Supposons les objets d'indices strictement inférieurs à j construits. Choisissons une réunion finie A_j de boules ouvertes de rayon r dont les centres forment un $r/4$ -réseau de $K_1 \cup \dots \cup K_{j-1}$; pour $j = 1$, posons $A_1 = \emptyset$. Choisissons ensuite $n_j > n_{j-1}$ assez grand pour que

$$\mu_{n_j}(A_j^c) > \eta, \quad \sum_{i < j} |\mu_{n_j}(h_i) - \ell(h_i)| < \eta/8.$$

C'est possible par la propriété précédente et par convergence des intégrales des fonctions h_i déjà construites. Prenons un compact K_j tel que $\mu_{n_j}(K_j^c) < \eta/8$, et posons

$$H_j = K_j \setminus A_j, \quad h_j(x) = \left(1 - \frac{4}{r} d(x, H_j) \right)_+.$$

Alors H_j est compact et $\mu_{n_j}(h_j) > 7\eta/8$. Par construction, H_j est à distance au moins $3r/4$ des K_i , $i < j$. Les ensembles $\{h_j > 0\}$ sont donc deux à deux à distance au moins $r/4$, et, pour $j > i$, $\{h_j > 0\} \cap K_i = \emptyset$.

Les fonctions h_j sont positives et leurs ensembles de positivité sont disjoints. Ainsi $\sum_{j=1}^m h_j \leq 1$ pour tout m , puis $\sum_{j \geq 1} \ell(h_j) \leq 1$ par positivité de ℓ . Choisissons J tel que $\sum_{j \geq J} \ell(h_j) < \eta/8$ et posons

$$f = \sum_{\substack{j \geq J \\ j \text{ pair}}} h_j.$$

Comme les ensembles de positivité sont disjoints, cette somme est aussi le supremum des fonctions considérées. Elle est donc $4/r$ -lipschitzienne et comprise entre 0 et 1. Pour $j \geq J$ pair, $\mu_{n_j}(f) > 7\eta/8$. Pour $j \geq J$ impair, on a au contraire

$$\begin{aligned} \mu_{n_j}(f) &\leq \sum_{\substack{J \leq i < j \\ i \text{ pair}}} \mu_{n_j}(h_i) + \mu_{n_j}(K_j^c) \\ &\leq \sum_{i \geq J} \ell(h_i) + \sum_{i < j} |\mu_{n_j}(h_i) - \ell(h_i)| + \mu_{n_j}(K_j^c) < 3\eta/8. \end{aligned}$$

La suite $(\mu_n(f))$ ne converge donc pas, contradiction. La famille (μ_n) est tendue, ce qui conclut la preuve.

Ce résultat est dû à Alexandrov. Voir Ramamoorthi, R.V., Rao, B.V. & Sethuraman, J., *A note on weak convergence*, Sankhya A 74, 269–276 (2012), <https://doi.org/10.1007/s13171-012-0016-6>.

* * *

Corrigé de l'exercice 3.22. [énoncé p. 129]

- (1) Soit $\varepsilon > 0$ et K compact de E tel que $\mathbb{P}(X_1 \in K) \geq 1 - \varepsilon/n$. Soit

$$\begin{aligned} \Phi : E^n &\rightarrow E \\ (x_1, \dots, x_n) &\mapsto x_1 + \dots + x_n \end{aligned}$$

continue. Alors en notant $K_n = \Phi(K \times K \times \dots \times K)$, compact comme image continue d'un compact on a

$$\mathbb{P}(X_1 + \dots + X_n \notin K_n) \leq \mathbb{P}(\exists 1 \leq i \leq n : X_i \notin K) \leq \sum_{i=1}^n \mathbb{P}(X_i \notin K) \leq \sum_{i=1}^n \frac{\varepsilon}{n} = \varepsilon,$$

de sorte que $\mathbb{P}(X_1 + \dots + X_n \in K_n) \geq 1 - \varepsilon$, d'où le résultat.

- (2) La réponse est oui. Pour cela, il suffit de démontrer que si X et Y sont deux variables aléatoires indépendantes telles que $X + Y$ est tendue, alors X et Y sont tendues. Le résultat voulu découle alors par récurrence.

À cet effet, soit $\varepsilon > 0$ et K compact tel que $\mathbb{P}(X + Y \in K) > 1 - \varepsilon$. Notons $\mathbb{P}_X$ et $\mathbb{P}_Y$ les lois respectives de X et Y de sorte que par Fubini

$$1 - \varepsilon < \int_{E \times E} \mathbb{1}_K(x + y) \mathbb{P}_X(dx) \mathbb{P}_Y(dy) = \int_E \mathbb{P}_Y(K - x) \mathbb{P}_X(dx).$$

Il s'ensuit qu'il existe $x \in E$ tel que $\mathbb{P}_Y(K - x) > 1 - \varepsilon$. Puisque $K - x$ est compact comme image d'un compact par l'application $z \mapsto z - x$, ceci montre que Y est tendue, et par symétrie X l'est également. Ceci conclut.

Corrigé de l'exercice 3.23. [énoncé p. 129]

- (1) L'équivalence entre (a) et (c) provient du fait général que $X_n \rightarrow X$ en probabilité si et seulement si de toute sous-suite on peut extraire une sous-sous-suite qui converge p.s. vers X .

Montrons que (c) implique (b). Soit $f : E \rightarrow \mathbb{R}$ continue bornée. Il suffit de montrer que pour tout extraction ϕ il existe une extraction ψ telle que $\int f dM_{\phi \circ \psi(n)} \rightarrow \int f dM$ presque sûrement. Soit ϕ extraction. Par (c), soit ψ extraction telle que presque sûrement $M_{\phi \circ \psi(n)}$ converge vers M dans $(\mathcal{M}_1(E), d_{LP})$. L'espace métrique E étant polonais, la convergence au sens de d_{LP} implique la convergence étroite, et donc presque sûrement $M_{\phi \circ \psi(n)}$ converge étroitement vers M et donc $\int f dM_{\phi \circ \psi(n)} \rightarrow \int f dM$ presque sûrement.

Montrons finalement (b) implique (a). Soit ϕ extraction. On montre qu'il existe une extraction ψ telle que presque sûrement $d_{LP}(M_{\phi \circ \psi(n)}, M) \rightarrow 0$. Tout d'abord, pour toute fonction $f \in \mathcal{F}$, on a $\int f dM_{\phi(n)} \rightarrow \int f dM$ en probabilité, de sorte qu'il existe une extraction ψ telle que $\int f dM_{\phi \circ \psi(n)} \rightarrow \int f dM$ presque sûrement. Comme $\mathcal{F}$ est dénombrable, par extraction diagonale, il existe une extraction ψ telle que

$$\text{presque sûrement, } \forall f \in \mathcal{F}, \int f dM_{\phi \circ \psi(n)} \xrightarrow{n \rightarrow \infty} \int f dM. \quad (4)$$

Vérifions que presque sûrement $d_{LP}(M_{\phi \circ \psi(n)}, M) \rightarrow 0$. Soit $\varepsilon > 0$. D'après le rappel, il existe $f_1, \dots, f_K \in \mathcal{F}$ et $\delta > 0$ tels que

$$\left\{ \nu \in \mathcal{M}_1(E) : \left| \int f_i d\nu - \int f_i dM \right| < \delta \text{ pour tout } 1 \leq i \leq K \right\} \subset B(M, \varepsilon).$$

⚡ Attention. Les fonctions $f_1, \dots, f_K$, ainsi que K et δ , peuvent dépendre de la réalisation de M . On raisonne trajectoire par trajectoire sur l'événement de probabilité 1 de (4); aucune sélection mesurable n'est nécessaire.

Par (4), presque sûrement, pour n assez grand on a

$$\left| \int f_i dM_{\phi \circ \psi(n)} - \int f_i dM \right| < \delta$$

pour tout $1 \leq i \leq K$. Il s'ensuit que presque sûrement, pour n assez grand on a $d_{LP}(M_{\phi \circ \psi(n)}, M) \leq \varepsilon$, et le résultat s'ensuit.

- (2) Soit $f \in \mathcal{C}_b(E, \mathbb{R})$ et posons $m = \int_E f d\mu$ ainsi que $M_n(f) = \int_E f dM_n$. L'indépendance conditionnelle des deux échantillons donne directement

$$\mathbb{E} \left[(M_n(f) - m)^2 \right] = \mathbb{E} \left[(f(\Pi_n^1) - m)(f(\Pi_n^2) - m) \right]. \quad (5)$$

La fonction $(x, y) \mapsto (f(x) - m)(f(y) - m)$ est continue bornée. Par convergence en loi du couple, le membre de droite converge donc vers

$$\mathbb{E} \left[(f(\Pi^1) - m)(f(\Pi^2) - m) \right] = \mathbb{E} \left[f(\Pi^1) - m \right] \mathbb{E} \left[f(\Pi^2) - m \right] = 0,$$

car les variables limites sont indépendantes et de loi μ . Ainsi $M_n(f) \rightarrow \mu(f)$ dans L^2 , donc en probabilité. La question (1) donne $d_{LP}(M_n, \mu) \rightarrow 0$ en probabilité.

Remarque. La première question est inspirée de la Proposition 2.2 de [cet article](#) et la deuxième question provient du Lemma 3.1 dans [cet article](#).

* * *

Corrigé de l'exercice 3.24. [énoncé p. 130]

- (1) Notons X la limite presque sûre de X_n . On écrit pour $\varepsilon > 0$:

$$\mathbb{P}(d(Y_n, X) \geq 2\varepsilon) \leq \mathbb{P}(d(Y_n, X_n) \geq \varepsilon) + \mathbb{P}(d(X_n, X) \geq \varepsilon).$$

Le premier terme de la somme converge vers 0 par hypothèse, et le second tend vers 0 également car la convergence presque sûre entraîne la convergence en probabilité.

- (2) Non : en prenant $X_n = X$ il suffit de prendre un cas où une convergence en probabilité a lieu sans convergence presque sûre. On peut par exemple prendre $E = \mathbb{R}$, $X_n = X = 0$, $(Y_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes telles que $\mathbb{P}(Y_n = 1) = 1/n$ et $\mathbb{P}(Y_n = 0) = 1 - 1/n$.

* * *

Corrigé du problème 3.25. [énoncé p. 130]

- (1) On peut utiliser des polynômes à coefficients rationnels multipliés par des fonctions plateaux régularisées, dans l'esprit des familles de fonctions test de la section 1.2.
- (2) Soit L une constante de Lipschitz de f , et choisissons $a < b$ tels que $\text{supp}(f) \subset]a, b[$. Posons

$$c(x) = \max(a, \min(x, b)), \quad H(x) = L(c(x) - a),$$

$$F = \frac{H + f}{2}, \quad G = \frac{H - f}{2}.$$

Puisque $f(x) = f(c(x))$, pour tous $x < y$,

$$|f(y) - f(x)| \leq L(c(y) - c(x)) = H(y) - H(x).$$

Ainsi F et G sont croissantes. Elles sont bornées et L -lipschitziennes, puisque H et f le sont, et $f = F - G$.

- (3) Puisque Z est à densité, les racines de P_n sont presque sûrement différentes. D'après le théorème de Rolle, les racines de P'_n sont réelles et situées entre les zéros de P_n . Notons ainsi $Z_1^n < Z_2^n < \dots < Z_n^n$ les racines de P_n et quitte à renuméroter les Y_i^n on peut supposer que

$$Z_1^n < Y_1^n < Z_2^n < \dots < Y_{n-1}^n < Z_n^n.$$

D'après les questions (1) et (2), il suffit de traiter chaque fonction croissante lipschitzienne bornée f intervenant dans les décompositions des éléments de l'ensemble dénombrable D . Par entrelacement,

$$\frac{\sum_{k=1}^n f(Z_k) - f(Z_n^n)}{n-1} \leq \frac{1}{n-1} \sum_{k=1}^{n-1} f(Y_k^n) \leq \frac{\sum_{k=1}^n f(Z_k) - f(Z_1^n)}{n-1}.$$

Les deux bornes convergent presque sûrement vers $\mathbb{E}[f(Z)]$ par la loi forte des grands nombres et la bornitude de f . L'intersection des événements de probabilité un pour cette famille dénombrable conclut que les mesures empiriques des racines de P'_n convergent presque sûrement vers la loi de Z .

- (4) Posons $N_i^n = \sum_{j=1}^n \mathbb{1}_{\{Z_j = a_i\}}$, $J_n = \{i : N_i^n > 0\}$ et $K_n = |J_n|$. Le polynôme

$$P_n(X) = \prod_{i \in J_n} (X - a_i)^{N_i^n}$$

a une dérivée qui possède, en chaque a_i , $i \in J_n$, une racine de multiplicité $N_i^n - 1$, ainsi que $K_n - 1$ autres racines, comptées avec multiplicité, notées W_j^n . Pour $f : \mathbb{C} \rightarrow \mathbb{R}$ continue bornée,

$$\frac{1}{n-1} \sum_{j=1}^{n-1} f(Y_j^n) = \frac{1}{n-1} \left(\sum_{i \in J_n} (N_i^n - 1) f(a_i) + \sum_{j=1}^{K_n-1} f(W_j^n) \right).$$

Par conséquent, pour $n \geq 2$,

$$\left| \frac{1}{n-1} \sum_{j=1}^{n-1} f(Y_j^n) - \frac{1}{n} \sum_{j=1}^n f(Z_j) \right| \leq \frac{2K \|f\|_\infty}{n-1}.$$

Par la loi forte des grands nombres, presque sûrement, simultanément pour $1 \leq i \leq K$, $N_i^n/n \rightarrow \mathbb{P}(Z_1 = a_i)$. Sur cet événement, la mesure empirique des Z_j converge étroitement vers la loi de Z_1 , et l'estimation précédente montre la même convergence pour la mesure empirique des racines de P'_n .

* * *

Corrigé du problème 3.26. [énoncé p. 130]

- (1) La séparation et la symétrie de d sont claires. Pour l'inégalité triangulaire, on remarque que pour $u, v, w \in \overline{\mathcal{U}}$ on a $d(u, w) \leq \max(d(u, v), d(v, w))$. En effet, si $d(u, w) > d(u, v)$, alors $u \wedge w \preceq u \wedge v$, de sorte que $v \wedge w = u \wedge w$ et $d(u, w) = d(v, w)$.

Pour la séparabilité, $\mathcal{U}$ est dense dans $\overline{\mathcal{U}}$. En effet, si $u \in \partial \mathcal{U}$ et $[u]_k \in \mathcal{U}$ désigne les k premiers entiers de u , on a clairement $d([u]_k, u) \rightarrow 0$ quand $k \rightarrow \infty$.

Pour tout $u \in \overline{\mathcal{U}}$ et $k \geq 1$, notons $[u]_k$ son préfixe de longueur k si u a au moins k lettres, et u lui-même sinon. Soit (u_n) une suite de Cauchy. La troncature est 1-lipschitzienne; deux mots distincts de longueur au plus k sont à distance au moins e^{-k} . Ainsi, pour chaque k , $[u_n]_k$ est constant à partir d'un certain rang, de valeur $v(k)$. Ces valeurs sont compatibles : $[v(\ell)]_k = v(k)$ pour $k \leq \ell$. Elles définissent un unique mot fini ou infini v dont elles sont les troncatures, et $u_n \rightarrow v$. L'espace est donc complet, et, étant séparable, polonais.

- (2) Soit $u \in \overline{\mathcal{U}}$ et $r > 0$. Vérifions que $B(u, r) = \{v \in \overline{\mathcal{U}} : d(u, v) < r\}$ est fermé. Pour cela, soit $v \in \overline{\mathcal{U}} \setminus B(u, r)$ de sorte que $d(u, v) \geq r$, et vérifions que $B(v, r) \cap B(u, r) = \emptyset$. En effet, par l'absurde, si $w \in B(v, r) \cap B(u, r)$, on a $d(w, v) < r$ et $d(w, u) < r$, et alors $d(u, v) \leq \max(d(u, w), d(w, v)) < r$, absurde.

- (3) On vérifie tout d'abord que les ensembles de la forme $\{u\}$ et $T(u)$ pour $u \in \mathbb{U}$ sont à la fois ouverts et fermés. Ceci provient de la question précédente, en remarquant que pour $u \in (\mathbb{N}^*)^n$, on a $\{u\} = B(u, e^{-n-1})$ et $T(u) = B(u, e^{-n+1/2})$.

L'implication provient alors immédiatement du théorème de porte-manteau.

Pour la réciproque, par séparabilité, tout ouvert peut s'écrire comme union dénombrable de boules ouvertes. On remarque que les boules ouvertes de $\overline{\mathbb{U}}$ sont précisément de la forme $\{u\}$ et $T(u)$ pour $u \in \mathbb{U}$, et que deux boules ouvertes sont soit disjointes, soit incluses l'une dans l'autre. Par suite, tout ouvert O de $\overline{\mathbb{U}}$ peut s'écrire comme union dénombrable disjointe d'ensembles de la forme précédente (ouverts et fermés). Ainsi, en écrivant $O = \coprod_{k \geq 1} O_k$, d'après le lemme de Fatou :

$$\mu(O) = \sum_{k=1}^{\infty} \mu(O_k) = \sum_{k=1}^{\infty} \liminf_{n \rightarrow \infty} \mu_n(O_k) \leq \liminf_{n \rightarrow \infty} \sum_{k=1}^{\infty} \mu_n(O_k) = \liminf_{n \rightarrow \infty} \mu_n(O),$$

ce qui conclut.

* * *

Corrigé du problème 3.27. [énoncé p. 131]

- (1) L'équivalence entre (b) et (c) provient du fait que la convergence en loi et la convergence en probabilité vers une constante sont équivalentes.

Montrons que (c) implique (d). Soit $f \in \mathcal{C}_b(E, \mathbb{R})$. Pour montrer que $\int f dM_n$ converge en probabilité vers $\int f d\mu$ on va montrer que pour toute extraction ϕ , il existe une extraction ψ telle que $\int f dM_{\phi \circ \psi(n)}$ converge presque sûrement vers $\int f d\mu$. Soit donc ϕ extraction. Comme M_n converge en probabilité vers μ , il existe une extraction ψ telle que presque sûrement $M_{\phi \circ \psi(n)}$ converge (pour d_{LP}) vers μ . Alors presque sûrement $M_{\phi \circ \psi(n)}$ converge étroitement vers μ , donc presque sûrement $\int f dM_{\phi \circ \psi(n)}$ converge vers $\int f d\mu$.

Alternativement, pour montrer que (c) implique (d) on peut remarquer que pour $f \in \mathcal{C}_b(E, \mathbb{R})$, l'application $\nu \mapsto \int f d\nu$ est continue, et que la convergence en probabilité est stable par composition par une fonction continue.

Montrons que (d) implique (c). Soit $\varepsilon > 0$. Il s'agit de montrer que

$$\mathbb{P}(d_{LP}(M_n, \mu) \geq \varepsilon) \rightarrow 0,$$

c'est-à-dire que $\mathbb{P}(M_n \notin B(\mu, \varepsilon)) \rightarrow 0$. D'après le fait donné en début d'exercice, il existe $f_1, \dots, f_K \in \mathcal{C}_b(E, \mathbb{R})$ et $\delta > 0$ tels que

$$\left\{ \nu \in \mathcal{M}_1(E) : \left| \int f_i d\nu - \int f_i d\mu \right| < \delta \text{ pour tout } 1 \leq i \leq K \right\} \subset B(\mu, \varepsilon).$$

Alors

$$\mathbb{P}(M_n \notin B(\mu, \varepsilon)) \leq \sum_{i=1}^K \mathbb{P}\left(\left| \int f_i dM_n - \int f_i d\mu \right| \geq \delta\right)$$

qui tend vers 0 par hypothèse.

Le fait que (a) implique (d) provient immédiatement de l'inégalité de Markov.

Le fait que (d) implique (a) provient du fait que la suite $(\int f dM_n)_{n \geq 1}$ est uniformément intégrable (car bornée par $\|f\|_\infty$).

L'implication (c) $\Rightarrow$ (e) découle de la continuité de $\nu \mapsto \int e^{itx} \nu(dx)$, en considérant séparément les parties réelle et imaginaire.

Montrons (e) $\Rightarrow$ (c). Comme expliqué dans la remarque suivante, on peut travailler sur un espace de probabilité commun. Pour tout t , la convergence en probabilité et la borne $|\Phi_n(t) - \Phi(t)| \leq 2$ entraînent $\mathbb{E}[|\Phi_n(t) - \Phi(t)|] \rightarrow 0$. Par Fubini et convergence dominée,

$$\mathbb{E} \left[\int_{\mathbb{R}} e^{-|t|} |\Phi_n(t) - \Phi(t)| dt \right] \rightarrow 0.$$

Les intégrandes sont conjointement mesurables, car elles sont mesurables en ω et continues en t . De toute sous-suite, on peut donc extraire (n_j) telle que

$$\sum_j \mathbb{E} \left[\int_{\mathbb{R}} e^{-|t|} |\Phi_{n_j}(t) - \Phi(t)| dt \right] < \infty.$$

Par Tonelli, presque sûrement, $\Phi_{n_j}(t) \rightarrow \Phi(t)$ pour Lebesgue-presque tout t . L'exercice 1.11(1) donne alors $M_{n_j} \xrightarrow{\text{étroite}} \mu$ presque sûrement, donc $d_{LP}(M_{n_j}, \mu) \rightarrow 0$ presque sûrement. Le critère par sous-suites entraîne (c).

En particulier, pour tout compact $K \subset \mathbb{R}$, $\sup_{t \in K} |\Phi_n(t) - \Phi(t)|$ converge en probabilité vers zéro : on applique l'exercice 1.6(2) le long des sous-suites presque sûrement convergentes.

Remarque. Stricte sensu, il faut faire attention car les variables aléatoires $(M_n)_{n \geq 1}$ ne sont pas forcément définies sur le même espace de probabilité, de sorte qu'on ne peut pas parler de convergence p.s. (ce point m'a été signalé par Dimitri Faure). Cependant, sans perte de généralité, on peut supposer que les variables aléatoires $(M_n)_{n \geq 1}$ sont définies sur le même espace de probabilité, par exemple en considérant une suite $(M'_n)_{n \geq 1}$ de variables aléatoires indépendantes définies sur le même espace de probabilité telles que pour tout $n \geq 1$ les lois de M_n et M'_n sont les mêmes (on utilise l'existence de mesures produit), et en travaillant avec M'_n à la place de M_n (les propriétés considérées ne dépendent que des lois individuelles).

- (2) (a) Démontrons que M_n converge en probabilité vers μ en montrant que pour toute extraction ϕ il existe une extraction ψ telle que presque sûrement $M_{\phi \circ \psi(n)}$ converge vers μ (i.e. presque sûrement $d_{LP}(M_{\phi \circ \psi(n)}, \mu) \rightarrow 0$). Soit donc ϕ extraction. Par procédé diagonal, il existe une extraction ψ telle que p.s. pour tout $k \geq 1$, $\int x^k dM_{\phi \circ \psi(n)}$ converge vers $\int x^k d\mu$. D'après un exercice de la feuille d'exercices, p.s. $M_{\phi \circ \psi(n)}$ converge étroitement vers μ , donc p.s. $d_{LP}(M_{\phi \circ \psi(n)}, \mu) \rightarrow 0$.

Alternativement, sans redéfinir les variables aléatoires sur le même espace de probabilité, il est possible de vérifier que (M_n) est tendue et pour l'identification de la limite utiliser le principe des lois accompagnantes en considérant la fonctionnelle $\nu \mapsto \int \max(\min(x^k, K), -K) d\nu$ avec $K \rightarrow \infty$.

(b) D'après l'inégalité de Bienaymé-Tchebychev, pour $\varepsilon > 0$ on a

$$\begin{aligned} \mathbb{P} \left(\left| \int x^k dM_n - \int x^k d\mu \right| \geq 2\varepsilon \right) \\ \leq \mathbb{P} \left(\left| \int x^k dM_n - \mathbb{E} \left[\int x^k dM_n \right] \right| \geq \varepsilon \right) \\ + \mathbb{P} \left(\left| \mathbb{E} \left[\int x^k dM_n \right] - \int x^k d\mu \right| \geq \varepsilon \right) \\ \leq \frac{\text{Var}(\int x^k dM_n)}{\varepsilon^2} + \mathbb{1}_{|\mathbb{E}[\int x^k dM_n] - \int x^k d\mu| \geq \varepsilon} \end{aligned}$$

qui tend vers 0 par hypothèse, et on peut appliquer la question (a).

(3) L'équivalence entre (a) et (b) est simplement la définition de la convergence étroite.

Le fait que (b) implique (c) est clair.

Pour montrer que (c) implique (b), fixons $\varepsilon > 0$ et montrons que p.s., pour n assez grand $d_{LP}(M_n, \mu) \leq \varepsilon$. D'après le fait donné en début d'exercice, il existe $f_1, \dots, f_K \in \mathcal{C}_b(E, \mathbb{R})$ et $\delta > 0$ tels que

$$\left\{ \nu \in \mathcal{M}_1(E) : \left| \int f_i d\nu - \int f_i d\mu \right| < \delta \text{ pour tout } 1 \leq i \leq K \right\} \subset B(\mu, \varepsilon).$$

Par hypothèse, p.s. pour n assez grand

$$\left| \int f_i dM_n - \int f_i d\mu \right| < \delta \text{ pour tout } 1 \leq i \leq K$$

Ainsi, p.s. pour n assez grand $d_{LP}(M_n, \mu) \leq \varepsilon$.

Le fait que (b) implique (d) est clair : sur l'événement de probabilité 1 où pour tout $f \in \mathcal{C}_b(E, \mathbb{R})$, $\int f dM_n \rightarrow \int f d\mu$ on a $\int e^{itx} dM_n(dx) \rightarrow \int e^{itx} \mu(dx)$ pour tout $t \in \mathbb{R}$.

Pour montrer que (d) implique (b), l'idée est d'utiliser le théorème de Fubini pour l'interversion. Notons

$$A = \left\{ (\omega, t) \in \Omega \times \mathbb{R} : \int e^{itx} dM_n(\omega)(dx) \not\rightarrow \int e^{itx} \mu(dx) \right\}$$

ainsi que ses sections

$$A_\omega = \{t \in \mathbb{R} : (\omega, t) \in A\}, \quad A^t = \{\omega \in \Omega : (\omega, t) \in A\}.$$

Par hypothèse, pour tout $t \in \mathbb{R}$ on a $\mathbb{P}(A^t) = 0$. Appliquons alors le théorème de Fubini, en notant λ la mesure de Lebesgue :

$$0 = \int_{\mathbb{R}} \mathbb{P}(A^t) \lambda(dt) = \int_{\Omega \times \mathbb{R}} \mathbb{1}_A d\mathbb{P} \otimes d\lambda = \int_{\Omega} \lambda(A_\omega) d\mathbb{P}.$$

On en déduit que $\lambda(A_\omega) = 0$ pour presque tout $\omega \in \Omega$. Ainsi, p.s. pour λ presque tout $t \in \mathbb{R}$ on a $\int e^{itx} dM_n(\omega)(dx) \rightarrow \int e^{itx} \mu(dx)$. On conclut avec la question (1) de l'exercice 1.11 (théorème de Lévy sous convergence presque partout en fréquence).

- (4) C'est une application directe de l'exercice de la feuille d'exercices précédemment mentionné.
- (5) D'après la loi forte des grands nombres, pour tout $f \in \mathcal{C}_b(E, \mathbb{R})$, pour presque tout $\omega \in \Omega$,

$$\int f dM_n = \frac{1}{n} \sum_{k=1}^n f(X_k) \rightarrow \mathbb{E}[f(X_1)] = \int f d\mu.$$

D'après la question (3), on en déduit que p.s. M_n converge étroitement vers μ (a fortiori M_n converge en probabilité et en loi vers μ).

* * *

Corrigé du problème 3.28. [énoncé p. 132]

- (1) Par compacité, E est borné. Il existe donc $M > 0$ tel que pour tout $x \in E$ on a $\overline{B}(x, M) = E$. Ainsi $\lim_{M \rightarrow \infty} I_\mu(M) = 1$.
- (2) Il est clair que I_μ est croissante, ce qui implique l'existence des limites à gauche. Pour démontrer la continuité à droite, raisonnons par l'absurde en supposant que I_μ n'est pas continue à droite en $\varepsilon_0 > 0$. Alors par croissance

$$I_\mu(\varepsilon_0) < \lim_{\varepsilon \downarrow \varepsilon_0} I_\mu(\varepsilon).$$

On peut donc trouver $x_0 \in E$ et $\delta > 0$ tels que pour tout $n \geq 1$,

$$\mu(\overline{B}(x_0, \varepsilon_0)) + \delta < I_\mu(\varepsilon_0 + 1/n).$$

En particulier, pour tout $n \geq 1$ on a

$$\mu(\overline{B}(x_0, \varepsilon_0)) + \delta < \mu(\overline{B}(x_0, \varepsilon_0 + 1/n)).$$

Comme $\mu(\overline{B}(x_0, \varepsilon_0 + 1/n)) \rightarrow \mu(\overline{B}(x_0, \varepsilon_0))$ lorsque $n \rightarrow \infty$, on obtient une contradiction.

- (3) Pour tout $x \in E$ et tout $r > 0$,

$$\inf_{y \in E} \mu(\{y\}) \leq I_\mu(r) \leq \mu(\overline{B}(x, r)).$$

En faisant $r \downarrow 0$, puis en prenant l'infimum sur x , on obtient

$$\lim_{r \downarrow 0} I_\mu(r) = \inf_{x \in E} \mu(\{x\}).$$

La limite est donc nulle si μ est sans atome. Elle est même toujours nulle lorsque E est infini : sinon une infinité de points distincts auraient tous une masse minorée par la même constante strictement positive.

- (4) Camille se trompe : par exemple, si $E = [0, 1]$ et si λ désigne la mesure de Lebesgue, $\mu = \frac{1}{2}\delta_0 + \frac{1}{2}\lambda$ vérifie $\lim_{\varepsilon \downarrow 0} I_\mu(\varepsilon) = 0$ mais $\mu(\{0\}) > 0$.
- (5) Fixons $\varepsilon > 0$. Pour tout $x \in E$, le théorème de porte-manteau appliqué au fermé $\overline{B}(x, \varepsilon)$ donne

$$\limsup_{n \rightarrow \infty} I_{\mu_n}(\varepsilon) \leq \limsup_{n \rightarrow \infty} \mu_n(\overline{B}(x, \varepsilon)) \leq \mu(\overline{B}(x, \varepsilon)).$$

En prenant l'infimum sur x , on obtient $\limsup_n I_{\mu_n}(\varepsilon) \leq I_\mu(\varepsilon)$.

- (6) Supposons (a). Si $I_\mu(\varepsilon) = 0$, choisissons x_j tel que $\mu(\overline{B}(x_j, \varepsilon)) \rightarrow 0$. Par compacité, quitte à extraire, $x_j \rightarrow x$. Pour j assez grand,

$$\overline{B}(x, \varepsilon/2) \subset \overline{B}(x_j, \varepsilon).$$

Ainsi $\mu(B(x, \varepsilon/2)) = 0$, contrairement au support plein. Donc (a) implique (b).

Supposons (b) et fixons $\varepsilon > 0$. Choisissons $x_1, \dots, x_k$ tels que $E = \bigcup_{i=1}^k B(x_i, \varepsilon/2)$. Par porte-manteau,

$$\liminf_n \mu_n(B(x_i, \varepsilon/2)) \geq \mu(B(x_i, \varepsilon/2)) \geq \mu(\overline{B}(x_i, \varepsilon/4)) \geq I_\mu(\varepsilon/4) > 0.$$

Pour tout $x \in E$, un indice i vérifie $d(x, x_i) < \varepsilon/2$, donc $B(x_i, \varepsilon/2) \subset \overline{B}(x, \varepsilon)$. Ainsi

$$I_{\mu_n}(\varepsilon) \geq \min_{1 \leq i \leq k} \mu_n(B(x_i, \varepsilon/2)),$$

et sa limite inférieure est strictement positive. Cela prouve (c).

Enfin, supposons (c). Soit O un ouvert non vide et choisissons $x \in E$, $r > 0$ tels que $\overline{B}(x, r) \subset O$. D'après (5),

$$\mu(O) \geq I_\mu(r) \geq \limsup_n I_{\mu_n}(r) \geq \liminf_n I_{\mu_n}(r) > 0.$$

Donc (c) implique (a), ce qui conclut.

- (7) Puisque E est séparable, $(\mathcal{M}_1(E), d_{LP})$ l'est également. On peut donc utiliser le théorème de représentation de Skorokhod et supposer que la convergence de M_n vers M a lieu presque sûrement dans $(\mathcal{M}_1(E), d_{LP})$.
- Supposons (a) et fixons $\varepsilon, \varepsilon' > 0$. D'après (6), sur la représentation presque sûrement convergente,

$$\liminf_n I_{M_n}(\varepsilon) > 0 \quad \text{presque sûrement.}$$

Les événements $\{I_{M_n}(\varepsilon) \geq 1/j \text{ pour tout } n \geq N\}$, en faisant croître j et N , recouvrent donc un événement de probabilité un. Il existe ainsi j, N tels que cet événement ait probabilité strictement supérieure à $1 - \varepsilon'$. Avec $\delta = 1/j$, on obtient $\mathbb{P}(I_{M_n}(\varepsilon) < \delta) < \varepsilon'$ pour tout $n \geq N$.

Réciproquement, supposons (b). Fixons $k \geq 1$ et $\varepsilon' > 0$, puis choisissons δ, N comme dans (b) pour $\varepsilon = 1/k$. D'après (5), l'application $v \mapsto I_v(1/k)$ est semi-continue supérieurement, donc $\{v : I_v(1/k) < \delta\}$ est ouvert. Par porte-manteau,

$$\mathbb{P}(I_M(1/k) < \delta) \leq \liminf_n \mathbb{P}(I_{M_n}(1/k) < \delta) \leq \varepsilon'.$$

Ainsi $\mathbb{P}(I_M(1/k) = 0) \leq \varepsilon'$. Comme $\varepsilon' > 0$ est arbitraire, cette probabilité vaut zéro. Presque sûrement, $I_M(1/k) > 0$ pour tout $k \geq 1$; par monotonie et (6), M est de support plein.

- (8) Non, les deux implications sont fausses. Donnons des contre exemples dans le cas déterministe. Si $E = [0, 1]$, $\mu_{2n} = \delta_0$ et $\mu_{2n+1} = \delta_1$, la suite $(I_{\mu_n})_{n \geq 1}$ est constante mais $(\mu_n)_{n \geq 1}$ ne converge pas en loi.

Si $E = [0, 2]$, prenons, pour $n \geq 2$,

$$\mu_n = \frac{1}{3}\delta_0 + \frac{1}{3}\delta_{1/2} + \frac{1}{3}\delta_{1/2+1/n}.$$

Alors

$$I_{\mu_n}(\varepsilon) = \begin{cases} 0 & \text{si } 0 < \varepsilon < 3/2 - 1/n, \\ 1/3 & \text{si } 3/2 - 1/n \leq \varepsilon < 3/2, \\ 2/3 & \text{si } 3/2 \leq \varepsilon < 2, \\ 1 & \text{si } 2 \leq \varepsilon. \end{cases}$$

Ces valeurs minimales sont réalisées par la boule de centre 2; pour les rayons des trois dernières lignes, toute boule contient au moins respectivement le dernier atome, les deux derniers atomes, puis tous les atomes. Or μ_n converge étroitement vers $\mu = \frac{1}{3}\delta_0 + \frac{2}{3}\delta_{1/2}$, pour laquelle

$$I_\mu(\varepsilon) = \begin{cases} 0 & \text{si } 0 < \varepsilon < 3/2, \\ 2/3 & \text{si } 3/2 \leq \varepsilon < 2, \\ 1 & \text{si } 2 \leq \varepsilon. \end{cases}$$

Sur le compact $[1, 7/4]$, I_{μ_n} prend la valeur $1/3$ pour tout n assez grand, tandis que I_μ ne prend que 0 et $2/3$. Pour tout changement de temps croissant bijectif λ_n de ce compact, $\|I_{\mu_n} \circ \lambda_n - I_\mu\|_\infty \geq 1/3$. La convergence J_1 est donc impossible.

Remarque. Ce problème est inspiré de l'article suivant :

Benedikt Stufler, Mass and radius of balls in Gromov-Hausdorff-Prokhorov convergent sequences, preprint disponible sur arxiv, arxiv :2201.12251

<https://arxiv.org/abs/2201.12251>

Ce problème montre comment vérifier qu'une mesure limite est de support plein à partir d'estimées sur les mesures qui l'approchent. En revanche, la condition $I_\mu(r) \rightarrow 0$ lorsque $r \downarrow 0$ ne caractérise pas l'absence d'atomes, comme l'illustre la question (4).

4 L'espace des fonctions continues sur un compact

Cours associé : [chapitre 4](#), p. 35.

4.1 Énoncés

Exercice 4.1. [corrigé p. 161] Pour tout $n \geq 0$, on se donne deux processus X^n, Y^n dans l'espace $\mathcal{C}([0, 1], \mathbb{R})$, définis sur un même espace de probabilités (qui peut éventuellement dépendre de n). On suppose que $X^n \rightarrow X$ et $Y^n \rightarrow Y$ en loi.

- (1) Alix dit : la suite $((X^n, Y^n))_{n \geq 0}$ est une suite tendue dans $\mathcal{C}([0, 1], \mathbb{R}^2)$. A-t-elle raison ?
- (2) On voit maintenant le couple (X^n, Y^n) comme un unique processus

$$t \mapsto (X_t^n, Y_t^n), \quad t \in [0, 1]$$

dans $\mathcal{C}([0, 1], \mathbb{R}^2)$. Billie dit : la suite $(X^n, Y^n)_{n \geq 0}$ est tendue dans $\mathcal{C}([0, 1], \mathbb{R}^2)$. A-t-il raison ?

- (3) On voit maintenant le couple (X^n, Y^n) comme une fonction $(X_s^n, Y_t^n)_{s, t \in [0, 1]}$ de $\mathcal{C}([0, 1]^2, \mathbb{R}^2)$. Camille dit : la suite $(X^n, Y^n)_{n \geq 0}$ est tendue dans $\mathcal{C}([0, 1]^2, \mathbb{R}^2)$. A-t-elle raison ?

* * *

Exercice 4.2. [corrigé p. 161] Soit $(X^n)_{n \geq 1}$ une suite de processus dans $\mathcal{C}([0, 1], \mathbb{R}^k)$. On suppose que $X^n \Rightarrow X$ dans $\mathcal{C}([0, 1], \mathbb{R}^k)$. Soit $f : \mathbb{R}^k \rightarrow \mathbb{R}$ une fonction continue. On pose

$$Y_n = \int_0^1 f(X_s^n) ds, \quad Y = \int_0^1 f(X_s) ds.$$

Montrer que $(X^n, Y_n) \Rightarrow (X, Y)$.

* * *

Exercice 4.3. [corrigé p. 162] Soient X^n, X des processus croissants, continus de $[0, 1]$ dans $\mathbb{R}$. Montrer que $X^n \Rightarrow X$ si et seulement si X^n converge vers X au sens des marginales fini-dimensionnelles.

* * *

Exercice 4.4. [corrigé p. 162]

- (1) Pour $a \geq 0$, l'application $f \mapsto t_a(f) = \inf\{t \geq 0 : f(t) \geq a\}$ définie sur $\mathcal{C} = \mathcal{C}(\mathbb{R}_+, \mathbb{R})$ à valeurs dans $\mathbb{R}_+ \cup \{\infty\}$ est-elle continue ? Sinon a-t-elle des points de continuité, et si oui, lesquels ?
- (2) Soit $(X^n, n \geq 0)$ une suite de processus dans $\mathcal{C}$. On suppose que X^n converge en loi vers X pour la topologie uniforme sur les compacts. On suppose également que $X_0^n = 0$ pour tout $n \geq 1$. On note, pour $a \geq 0$,

$$T_a^n = t_a(X^n), \quad T_a = t_a(X).$$

Discuter la convergence (ou non) de T_a^n vers T_a . Discuter le cas particulier où X est le mouvement brownien standard (cette dernière question se traite mieux avec la notion de propriété de Markov forte pour le mouvement brownien, voir le cours de calcul stochastique).

* * *

Exercice 4.5. [corrigé p. 164] Montrer que $\{f \in \mathbb{R}^{[0,1]} : f \text{ est continue}\}$ n'est pas un élément de la tribu produit $\mathcal{B}(\mathbb{R})^{\otimes [0,1]}$.

Indication. On pourra considérer une variable aléatoire U uniforme sur $[0, 1]$ et considérer la fonction $X(t) = \mathbb{1}_{t \neq U}$ pour $0 \leq t \leq 1$.

* * *

Exercice 4.6. [corrigé p. 164] On considère un espace de probabilités $(\Omega, \mathcal{F}, \mathbb{P})$ muni d'une filtration, c'est-à-dire d'une famille $(\mathcal{F}_t, t \geq 0)$ de tribus telle que $\mathcal{F}_s \subset \mathcal{F}_t$ pour tout $s \leq t$. On rappelle qu'une martingale est une famille de variables aléatoires $(X_t, t \geq 0)$ telle que $X_t \in L^1$ pour tout $t \geq 0$ et telle que pour tout $s \leq t$ on a $\mathbb{E}[X_t | \mathcal{F}_s] = X_s$.

Soit $(X^n, n \geq 0)$ une suite de processus de $\mathcal{C} = \mathcal{C}([0, 1], \mathbb{R})$, intégrables et adaptés par rapport à des filtrations $(\mathcal{F}_t^n)$, c'est-à-dire que X_t^n est mesurable par rapport à $\mathcal{F}_t^n$ pour tout $t \in [0, 1]$. On suppose que X^n est une $\mathcal{F}^n$ -martingale, et que X^n converge en loi vers X dans $\mathcal{C}$. On suppose également que pour tout $t \in [0, 1]$, $(X_t^n)_{n \geq 1}$ est uniformément intégrable.

Montrer que $(X_t, 0 \leq t \leq 1)$ est une $(\mathcal{F}_t)$ -martingale où $\mathcal{F}_t = \sigma(X_s, 0 \leq s \leq t)$.

* * *

Exercice 4.7. [corrigé p. 164] On note $\mathcal{C}$ l'ensemble des fonctions continues de $[0, 1]$ dans $\mathbb{R}$. Soit $(X_n(t) : 0 \leq t \leq 1)_{n \geq 1}$ une suite de variables aléatoires à valeurs dans $\mathcal{C}$. On note $X_n = (X_n(t) : 0 \leq t \leq 1)$.

(1) Démontrer que la suite $(X_n)_{n \geq 1}$ est tendue dans $\mathcal{C}$ si et seulement si les deux conditions suivantes sont vérifiées :

- (a) la suite $(X_n(0))_{n \geq 1}$ est tendue dans $\mathbb{R}$;
- (b) pour toute suite $\delta_k \rightarrow 0$ et pour toute suite $(n_k)_{k \geq 1}$, on a

$$\omega(X_{n_k}, \delta_k) \xrightarrow[k \rightarrow \infty]{(\mathbb{P})} 0,$$

où $\omega(f, \delta) = \sup_{|s-t| \leq \delta, 0 \leq s, t \leq 1} |f(s) - f(t)|$ désigne le module de continuité.

(2) Soit $(X_n^i)_{n \geq 1, 1 \leq i \leq n}$ une suite de variables aléatoires à valeurs dans $\mathcal{C}$ telle que pour tout $n \geq 1$, les variables aléatoires $(X_n^i)_{1 \leq i \leq n}$ ont la même loi que X_n (mais on ne les suppose pas indépendantes). On suppose que :

- (I) la suite de variables aléatoires $(X_n)_{n \geq 1}$ est tendue dans $\mathcal{C}$;
 - (II) la suite de variables aléatoires $\sup_{t \in [0, 1]} |X_n(t)|$ est uniformément intégrable.
- Démontrer que la suite de variables aléatoires $(Y_n)_{n \geq 1}$ est tendue dans $\mathcal{C}$, où Y_n est définie par

$$Y_n(t) = \frac{1}{n} \sum_{i=1}^n X_n^i(t), \quad t \in [0, 1].$$

- (3) Le résultat de la question (2) reste-t-il vrai si on remplace (II) par (II') défini par (II') la suite de variables aléatoires $\sup_{t \in [0,1]} |X_n(t)|$ est bornée dans L^1 ? Justifiez votre réponse.

* * *

Exercice 4.8. [corrigé p. 166] Soient $(X_i)_{i \geq 1}$ des variables aléatoires i.i.d. à valeurs dans $\mathcal{C}([0, 1])$, l'espace des fonctions continues de $[0, 1]$ dans $\mathbb{R}$ muni de la norme uniforme. On suppose qu'il existe $C, \varepsilon > 0$ tels que

$$\begin{aligned} \forall 0 \leq s, t \leq 1, \quad \mathbb{E} \left[(X_1(s) - X_1(t))^2 \right] &\leq C|s - t|^{1+\varepsilon}, \\ \forall 0 \leq s \leq 1, \quad \mathbb{E} [X_1(s)] &= 0 \quad \text{et} \quad \mathbb{E} [X_1(s)^2] < \infty. \end{aligned}$$

On pose

$$Z_n = \frac{X_1 + X_2 + \cdots + X_n}{\sqrt{n}}.$$

- (1) Pour $0 \leq s, t \leq 1$, calculer $\mathbb{E} [(Z_n(s) - Z_n(t))^2]$.
- (2) Montrer que $(Z_n)_{n \geq 1}$ converge en loi dans $\mathcal{C}([0, 1])$.

* * *

Exercice 4.9. [corrigé p. 166] Pour $x \in \mathbb{R}$, on note $\mathbb{P}_x$ la loi d'un mouvement brownien réel standard partant de x . Soit $(x_n)_{n \geq 1}$ une suite de nombres réels.

- (1) Lorsque $(x_n)_{n \geq 1}$ est bornée, montrer que la suite $(\mathbb{P}_{x_n})_{n \geq 1}$ est tendue dans $\mathcal{C}([0, 1], \mathbb{R})$.
- (2) Le résultat reste-t-il vrai si la suite $(x_n)_{n \geq 1}$ n'est plus supposée bornée ? Justifiez votre réponse.

* * *

Exercice 4.10. [corrigé p. 167] Soient X et Y deux variables aléatoires à valeurs dans $\mathcal{C}([0, 1], \mathbb{R})$, l'ensemble des fonctions continues de $[0, 1]$ dans $\mathbb{R}$ muni de la topologie de la convergence uniforme. On suppose que pour tout $\varepsilon > 0$, il existe un couple de variables aléatoires $(X_\varepsilon, Y_\varepsilon)$ à valeurs dans $\mathcal{C}([0, 1], \mathbb{R})^2$ tel que X_ε a la même loi que X , Y_ε a la même loi que Y et $\mathbb{P} (\|X_\varepsilon - Y_\varepsilon\|_\infty > \varepsilon) < \varepsilon$. Montrer que X et Y ont même loi.

* * *

Problème 4.11. [corrigé p. 167] Dans cet exercice, on note $\mathbb{R}^{[0,1]}$ l'ensemble des fonctions de $[0, 1]$ dans $\mathbb{R}$ muni de la tribu produit $\mathcal{B}(\mathbb{R})^{\otimes [0,1]}$ et $\mathcal{C} = \mathcal{C}([0, 1], \mathbb{R})$ l'espace vectoriel des fonctions continues de $[0, 1]$ à valeurs réelles muni de la norme uniforme et la tribu borélienne associée. On considère un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$.

- (1) Montrer que la tribu borélienne de $\mathcal{C}$ est égale à l'ensemble des A de la forme $A = B \cap \mathcal{C}$ pour $B \in \mathcal{B}(\mathbb{R})^{\otimes [0,1]}$.

Si $X, Y : \Omega \rightarrow \mathbb{R}^{[0,1]}$ sont deux variables aléatoires, on dit que Y est une *modification* (ou *version*) de X si pour tout $t \in [0, 1]$, on a $\mathbb{P}(X(t) = Y(t)) = 1$.

(2) Soient $X, Y : \Omega \rightarrow \mathbb{R}^{[0,1]}$ deux variables aléatoires.

(a) Vérifier que si Y est une modification de X , alors X et Y ont la même loi.

(b) Billie dit : la réciproque de (a) est vraie. A-t-il raison ? Justifiez votre réponse.

Dans toute la suite de cet exercice, on fixe une variable aléatoire $Y : \Omega \rightarrow \mathbb{C}$. (Ainsi Y_ω est une fonction continue pour tout $\omega \in \Omega$.)

(3) Soit $X : \Omega \rightarrow \mathbb{R}^{[0,1]}$ une variable aléatoire. On suppose que X et Y ont la même loi.

(a) Notons Y^* la variable aléatoire Y vue comme fonction à valeurs dans $\mathbb{R}^{[0,1]}$. Démontrer qu'il existe une application mesurable $F : \mathbb{R}^{[0,1]} \rightarrow \mathbb{C}$ telle que $Y = F \circ Y^*$.

Indication. On pourra utiliser le lemme de Doob-Dynkin, qui est le résultat suivant. Lorsque f, g sont deux fonctions mesurables définies sur $(\Omega, \mathcal{F})$ et à valeurs dans respectivement $(S, \mathcal{S})$ et $(T, \mathcal{T})$, si $f : (\Omega, \sigma(g)) \rightarrow (S, \mathcal{S})$ est mesurable et si $(S, \mathcal{S})$ est un espace polonais muni de sa tribu borélienne, alors il existe une application mesurable $h : (T, \mathcal{T}) \rightarrow (S, \mathcal{S})$ telle que $f = h \circ g$.

(b) Démontrer qu'il existe une variable aléatoire $X' : \Omega \rightarrow \mathbb{C}$ telle que X' soit une modification de X .

(4) (a) Soient $(U_n)_{n \geq 1}$ et $(V_n)_{n \geq 1}$ deux suites de variables aléatoires à valeurs dans un espace métrique (E, d) complet séparable. On suppose que les deux suites $(U_n)_{n \geq 1}$ et $(V_n)_{n \geq 1}$ ont même loi. Montrer que $\{(U_n)_{n \geq 1} \text{ converge}\}$ (c'est-à-dire l'ensemble $\{\omega \in \Omega : U_n(\omega) \text{ converge quand } n \rightarrow \infty\}$) est un événement (i.e. est mesurable) et que

$$\mathbb{P}((U_n)_{n \geq 1} \text{ converge}) = \mathbb{P}((V_n)_{n \geq 1} \text{ converge}).$$

(b) On considère une suite de variables aléatoires $(X_n)_{n \geq 1}$ à valeurs dans $\mathbb{R}^{[0,1]}$ telle que X_n converge vers Y au sens des marginales fini-dimensionnelles. On suppose aussi que pour tout $q \in [0, 1] \cap \mathbb{Q}$, la suite $(X_n(q))_{n \geq 1}$ converge presque sûrement. Montrer qu'il existe une variable aléatoire $X : \Omega \rightarrow \mathbb{C}$ telle que pour tout $q \in [0, 1] \cap \mathbb{Q}$, la suite $(X_n(q))_{n \geq 1}$ converge presque sûrement vers $X(q)$.

* * *

Problème 4.12. [corrigé p. 169]

Première partie. Soit (E, d) un espace métrique séparable et $(x_n)_{n \geq 1}$ une suite dense. On note $\mathcal{M}_1(E)$ l'ensemble des mesures de probabilité sur E muni de sa tribu borélienne, et $\text{Lip}_1(E)$ l'ensemble des fonctions de E dans $\mathbb{R}$ 1-lipschitziennes (c'est-à-dire telles que $|f(x) - f(y)| \leq d(x, y)$ pour tous $x, y \in E$).

Soit $\mu_n, \mu \in \mathcal{M}_1(E)$. L'objectif de cette partie est de montrer que $\mu_n \Rightarrow \mu$ si et seulement si $\int f d\mu_n \rightarrow \int f d\mu$ pour toute fonction f de la forme

$$f(x) = \max(c_0, c_1 - d(x, x_1), \dots, c_n - d(x, x_n))$$

avec $n \geq 1$ et $c_0, \dots, c_n \in \mathbb{R}$.

- (1) Justifier l'implication.
- (2) Vérifier que pour montrer que $\mu_n \Rightarrow \mu$ il suffit de vérifier que $\int f d\mu_n \rightarrow \int f d\mu$ pour toute fonction $f \in \text{Lip}_1(E)$ bornée.
- (3) Démontrer la réciproque.
Indication. Pour $f \in \text{Lip}_1(E)$ bornée telle que $f \geq c_0$, on pourra vérifier que pour tout $x \in E$ on a

$$f(x) = \sup_{n \geq 1} \max(c_0, f(x_n) - d(x, x_n)).$$
- (4) Si E est borné (c'est-à-dire si $\sup_{x,y \in E} d(x,y) < \infty$), montrer que $\mu_n \Rightarrow \mu$ si et seulement si $\int f d\mu_n \rightarrow \int f d\mu$ pour toute fonction $f(x)$ polynomiale en les variables $d(x, x_1), \dots, d(x, x_n)$ pour tout $n \geq 1$.

Deuxième partie. Considérons l'ensemble E des fonctions mesurables de $[0, 1] \rightarrow \mathbb{R}$, considérées à égalité presque partout près (c'est-à-dire qu'on identifie deux fonctions égales presque partout) muni de la distance

$$d(f, g) = \int_0^1 \min(1, |f(t) - g(t)|) dt. \quad (6)$$

- (5) En notant λ la mesure de Lebesgue, montrer que $f_n \rightarrow f$ dans E si et seulement si, pour tout $\varepsilon > 0$ on a $\lambda(\{x \in [0, 1] : |f_n(x) - f(x)| \geq \varepsilon\}) \rightarrow 0$ quand $n \rightarrow \infty$.
- (6) On suppose que $f_n \rightarrow f$ dans E . Montrer qu'on peut trouver une extraction ϕ telle que $f_{\phi(n)}$ converge presque partout vers f .
- (7) Démontrer que E est polonais.

Troisième partie. Considérons l'espace métrique $\mathcal{C}$ des fonctions continues de $[0, 1] \rightarrow \mathbb{R}$ muni de la distance d définie par (6).

- (8) Vérifier que $(\mathcal{C}, d)$ est séparable et borné. Est-il complet ? Justifier votre réponse.

Soient $(X_n)_{n \geq 1}$ des variables aléatoires à valeurs dans $\mathcal{C}$ et X une variable aléatoire à valeurs dans $\mathcal{C}$.

- (9) Démontrer que si X_n converge en loi vers X , alors pour tout $k \geq 1$, si $(U_i)_{1 \leq i \leq k}$ sont i.i.d. uniformes sur $[0, 1]$ (aussi indépendantes de X_n, X) on a la convergence en loi de $(X_n(U_1), \dots, X_n(U_k))$ vers $(X(U_1), \dots, X(U_k))$.
- (10) On suppose que les marginales fini-dimensionnelles de X_n convergent en loi vers celles de X . Démontrer que X_n converge en loi vers X .
- (11) [Question bonus hors barème] Démontrer que X_n converge en loi vers X si et seulement si, pour tout $k \geq 1$, en notant $V_1^k, \dots, V_k^k$ le réarrangement croissant de k variables aléatoires i.i.d. uniformes sur $[0, 1]$ (aussi indépendantes de X_n et X) on a la convergence en loi de $(X_n(V_1^k), \dots, X_n(V_k^k))$ vers $(X(V_1^k), \dots, X(V_k^k))$.

Problème 4.13. [corrigé p. 174] Soient $(E_i)_{i \geq 1}$ des variables aléatoires exponentielles indépendantes de paramètre 1. On pose $S_0 = 0$ et $S_i = E_1 + \dots + E_i$ pour $i \geq 1$. Pour $t \geq 0$ on pose $P(t) = \text{Card}\{i \geq 1 : S_i \leq t\}$, de sorte que P est un processus de Poisson de paramètre 1.

Première partie.

- (1) Montrer que pour tout $T \geq 0$ on a $\sup_{0 \leq u \leq T} \left| \frac{P(un)}{n} - u \right| \xrightarrow[n \rightarrow \infty]{\text{p.s.}} 0$.
- (2) Est-il vrai que $\sup_{u \geq 0} \left| \frac{P(un)}{n} - u \right| \xrightarrow[n \rightarrow \infty]{\text{p.s.}} 0$? Justifiez votre réponse.

Deuxième partie. Soit $(Y_n)_{n \geq 1}$ une suite de variables aléatoires à valeurs dans $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$, muni de la topologie de la convergence uniforme sur tout compact. On suppose qu'il existe une variable aléatoire $Y \in \mathcal{C}(\mathbb{R}_+, \mathbb{R})$ telle que pour tout $T \geq 0$, $(Y_n(t) : 0 \leq t \leq T)$ converge en loi vers $(Y(t) : 0 \leq t \leq T)$ dans $\mathcal{C}([0, T], \mathbb{R})$ lorsque $n \rightarrow \infty$.

- (3) Démontrer que pour tout $T \geq 0$ la convergence suivante a lieu en probabilité

$$\sup_{0 \leq t \leq T} \left| Y_n \left(\frac{P(tn)}{n} \right) - Y_n(t) \right| \xrightarrow[n \rightarrow \infty]{(\mathbb{P})} 0.$$

Troisième partie. Soient $T > 0$, $\beta : \mathbb{R} \rightarrow \mathbb{R}_+$ une fonction lipschitzienne et $z_0 \in \mathbb{R}_+$.

- (4) Justifier que l'équation différentielle $z'(t) = \beta(z(t))$ avec $z(0) = z_0$ admet une unique solution sur $[0, T]$, qui sera notée z dans la suite.

Pour $n \geq 1$, on fixe $Z_n(0) \in \mathbb{R}$ et on suppose que $(Z_n(t))_{0 \leq t \leq T}$ vérifie pour tout $0 \leq t \leq T$:

$$Z_n(t) = Z_n(0) + P \left(n \int_0^t \beta \left(\frac{Z_n(s)}{n} \right) ds \right),$$

On pose $\bar{Z}_n(t) = Z_n(t)/n$ et on suppose que $\bar{Z}_n(0) \rightarrow z_0$ lorsque $n \rightarrow \infty$.

- (5) Montrer que

$$\sup_{0 \leq t \leq T} |\bar{Z}_n(t) - z(t)| \xrightarrow[n \rightarrow \infty]{\text{p.s.}} 0.$$

Indication. On pourra utiliser le lemme de Gronwall.

* * *

Exercice 4.14. [corrigé p. 176] Soit K un espace métrique compact et E un espace métrique. On note $C(K, E)$ l'espace des fonctions continues de K dans E muni de la distance $d(f, g) = \sup_{x \in K} d_E(f(x), g(x))$. Montrer que $\{B \cap C(K, E) : B \in \mathcal{B}(E)^{\otimes K}\}$ est la plus petite tribu sur $C(K, E)$ qui rend toutes les projections

$$\begin{aligned} \Pi_x : C(K, E) &\rightarrow E \\ f &\mapsto f(x) \end{aligned}$$

pour $x \in K$ mesurables (autrement dit, la trace de la tribu produit $\mathcal{B}(E)^{\otimes K}$ sur $C(K, E)$ est la tribu produit sur $C(K, E)$).

4.2 Corrigés

Corrigé de l'exercice 4.1. [énoncé p. 155]

- (1) Oui! Soit $\varepsilon > 0$. D'après le théorème de Prokhorov, il existe deux compacts K_1, K_2 de $\mathcal{C}([0, 1], \mathbb{R})$ tels que $\mathbb{P}(X^n \in K_1) \geq 1 - \varepsilon$ et $\mathbb{P}(Y^n \in K_2) \geq 1 - \varepsilon$ pour tout entier $n \geq 0$. Alors $K_1 \times K_2$ est un compact de $\mathcal{C}([0, 1], \mathbb{R})^2$, et $\mathbb{P}((X^n, Y^n) \in K_1 \times K_2) \geq 1 - 2\varepsilon$, ce qui montre que $((X^n, Y^n))_{n \geq 0}$ est une suite tendue dans $\mathcal{C}([0, 1], \mathbb{R})^2$.
- (2) Oui! Tout d'abord, il est clair que pour tout $t \in [0, 1]$, $(X_t^n, Y_t^n)_{n \geq 1}$ est tendue. Pour contrôler le module de continuité, munissons $\mathbb{R}^2$ de la norme $\|\cdot\|_1$. Pour une fonction $f = (f_1, f_2) \in \mathcal{C}([0, 1], \mathbb{R}^2)$,

$$\omega(f, \delta) = \sup\{|f_1(s) - f_1(t)| + |f_2(s) - f_2(t)| : |s - t| \leq \delta\} \leq \omega(f_1, \delta) + \omega(f_2, \delta).$$

Ainsi,

$$\mathbb{P}(\omega((X^n, Y^n), \delta) \geq \eta) \leq \mathbb{P}(\omega(X^n, \delta) \geq \eta/2) + \mathbb{P}(\omega(Y^n, \delta) \geq \eta/2),$$

ce qui implique la tension.

Remarque. Alternativement, on peut utiliser le fait que la fonction

$$\begin{aligned} \mathcal{C}([0, 1], \mathbb{R})^2 &\rightarrow \mathcal{C}([0, 1], \mathbb{R}^2) \\ (f, g) &\mapsto (f(t), g(t))_{0 \leq t \leq 1} \end{aligned}$$

est continue et la stabilité de la tension par composition par une fonction continue.

- (3) Oui! Il suffit de vérifier que la fonction

$$\begin{aligned} \Phi : \mathcal{C}([0, 1], \mathbb{R})^2 &\rightarrow \mathcal{C}([0, 1]^2, \mathbb{R}^2) \\ (f, g) &\mapsto (f(s), g(t))_{0 \leq s \leq 1, 0 \leq t \leq 1} \end{aligned}$$

est continue, et le résultat en découlera par stabilité de la tension par composition par une fonction continue. À cet effet, en munissant $\mathbb{R}^2$ de la norme $\|\cdot\|_1$, et en notant $d(F, G) = \sup_{(s,t) \in [0,1]^2} \|F(s, t) - G(s, t)\|_1$ pour $F, G \in \mathcal{C}([0, 1]^2, \mathbb{R}^2)$, on remarque que

$$d(\Phi(f, g), \Phi(f', g')) = \|f - f'\|_\infty + \|g - g'\|_\infty.$$

Il en découle que si $f_n \rightarrow f$ et $g_n \rightarrow g$, alors $\Phi(f_n, g_n) \rightarrow \Phi(f, g)$, d'où le résultat.

* * *

Corrigé de l'exercice 4.2. [énoncé p. 155] D'après le théorème de représentation de Skorokhod, on peut supposer que la convergence $X^n \rightarrow X$ a lieu presque sûrement. Par continuité de l'application

$$\begin{aligned} \mathcal{C}([0, 1], \mathbb{R}^k) &\rightarrow \mathbb{R} \\ g &\mapsto \int_0^1 f(g(s)) ds, \end{aligned}$$

on en déduit que $Y_n \rightarrow Y$ presque sûrement. Ainsi, $(X^n, Y_n) \rightarrow (X, Y)$ presque sûrement, et donc en loi.

Remarque. On peut aussi directement dire que l'application

$$\begin{aligned} \Phi : \mathcal{C}([0, 1], \mathbb{R}^k) &\rightarrow \mathcal{C}([0, 1], \mathbb{R}^k) \times \mathbb{R} \\ g &\mapsto (g, \int_0^1 f(g(s)) ds) \end{aligned}$$

est continue, et donc que $\Phi(X^n)$ converge en loi vers $\Phi(X)$.

* * *

Corrigé de l'exercice 4.3. [énoncé p. 155] $\Rightarrow$ Cette implication est toujours vraie.

$\Leftarrow$ Puisque $(X^n(0))_{n \geq 1}$ converge en loi, il suffit de vérifier que la suite $(X^n)_{n \geq 1}$ est tendue. Le découpage utilisé dans la preuve du corollaire 4.19, avec $\delta = 1/k$, donne, pour toute fonction $f \in \mathcal{C}([0, 1], \mathbb{R})$,

$$\omega(f, 1/k) \leq 3 \max_{0 \leq i \leq k-1} \sup_{r \in [0, 1/k]} |f(r + i/k) - f(i/k)|.$$

En particulier, si f est croissante, on a

$$\omega(f, 1/k) \leq 3 \max_{0 \leq i \leq k-1} |f((i+1)/k) - f(i/k)|.$$

Par ailleurs, par convergence des marginales fini-dimensionnelles de X^n vers celles de X , on a

$$\max_{0 \leq i \leq k-1} |X^n((i+1)/k) - X^n(i/k)| \xrightarrow[n \rightarrow \infty]{(\text{loi})} \max_{0 \leq i \leq k-1} |X((i+1)/k) - X(i/k)|.$$

Par continuité de X , on a, pour tout $\eta > 0$,

$$\mathbb{P} \left(\max_{0 \leq i \leq k-1} |X((i+1)/k) - X(i/k)| \geq \eta \right) \xrightarrow[k \rightarrow \infty]{} 0.$$

On en déduit que pour tout $\varepsilon, \eta > 0$, on peut trouver $k \geq 1$ tel que

$$\limsup_{n \rightarrow \infty} \mathbb{P}(\omega(X^n, 1/k) \geq \eta) \leq \varepsilon.$$

* * *

Corrigé de l'exercice 4.4. [énoncé p. 155]

- (1) Non, elle n'est pas continue. Dans l'exemple ci-dessous, $f_n \rightarrow f$ uniformément sur tous les compacts, mais $3 = t_a(f_n) \not\rightarrow t_a(f) = 1$. En munissant $[0, \infty]$ de sa topologie usuelle, les points de continuité sont exactement les fonctions f telles que $t_a(f) = \infty$, ou $f(0) > a$, ou encore $t_a(f) < \infty$ et, pour tout $\varepsilon > 0$, il existe $t \in [t_a(f), t_a(f) + \varepsilon]$ tel que $f(t) > a$. Lorsque $t_a(f) > 0$, cette dernière condition équivaut à l'absence de maximum local en $t_a(f)$.

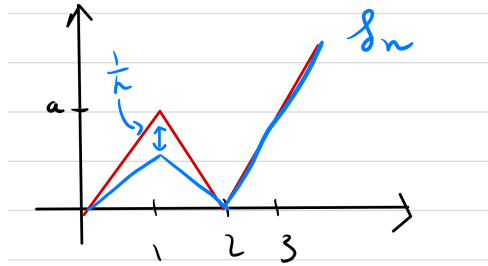


FIGURE 1 – Discontinuité du premier temps d'atteinte : f_n converge vers f , mais $t_a(f_n) = 3 \neq t_a(f) = 1$.

- (2) Tout d'abord, il est clair que $0 = T_0^n$ converge vers $T_0 = 0$. On suppose donc $a > 0$. D'après la question précédente, si presque sûrement, X n'admet pas de maximum local en T_a lorsque $T_a < \infty$, alors T_a^n converge en loi vers T_a (utiliser par exemple le théorème de représentation de Skorokhod). Lorsque X est le mouvement brownien standard, on sait que $T_a < \infty$ presque sûrement, et puisque pour tout $\epsilon > 0$, $\mathbb{P}\left(\sup_{[0,\epsilon]} X > 0\right) = 1$, d'après la propriété de Markov forte appliquée au temps d'arrêt T_a , presque sûrement X n'admet pas de maximum local en T_a . Ainsi, dans ce cas, T_a^n converge en loi vers T_a .

Maintenant, si avec probabilité strictement positive $T_a < \infty$ et X admet un maximum local en T_a , alors on peut trouver X^n tel que T_a^n ne converge pas en loi vers T_a . Puisque $a > 0$ et $X_0 = 0$, on a $T_a > 0$ et

$$\begin{aligned} \{T_a < \infty \text{ et } X \text{ admet un maximum local en } T_a\} \\ = \bigcup_{k \geq 1} \{T_a < k, X_t \leq a \text{ pour tout } t \in [T_a, T_a + 1/k]\}. \end{aligned}$$

Choisissons k tel que l'événement A figurant dans cette union ait une probabilité strictement positive, et posons

$$X_t^n = X_t - \frac{\min(t, 1)}{n} \mathbb{1}_A.$$

Alors $X_0^n = 0$ et $X^n \rightarrow X$ uniformément sur tout compact. Sur A , $T_a^n \geq T_a + 1/k$, et hors de A , $T_a^n = T_a$. Pour $M > 2k$, on obtient donc

$$\mathbb{E}[T_a^n \wedge M] - \mathbb{E}[T_a \wedge M] \geq \mathbb{P}(A)/k > 0.$$

La fonction $u \mapsto u \wedge M$ étant continue bornée sur $[0, \infty]$, ceci exclut la convergence en loi de T_a^n vers T_a .

Corrigé de l'exercice 4.5. [énoncé p. 156] On raisonne par l'absurde en supposant que $A = \{f \in \mathbb{R}^{[0,1]} : f \text{ est continue}\} \in \mathcal{B}(\mathbb{R})^{\otimes [0,1]}$. La variable aléatoire $X = (X(t) : 0 \leq t \leq 1)$ a les mêmes marginales fini-dimensionnelles que la fonction constante égale à 1. Elle a donc la même loi que la fonction constante égale à 1. Ainsi

$$\mathbb{P}(X \in A) = \mathbb{P}(1 \in A) = 1.$$

C'est absurde car X n'est pas continu.

* * *

Corrigé de l'exercice 4.6. [énoncé p. 156] Tout d'abord, on remarque que par uniforme intégrabilité, pour tout $t \in [0, 1]$, X_t est intégrable et $\mathbb{E}[X_t] = \mathbb{E}[X_0]$.

Fixons $0 \leq s \leq t \leq 1$. Pour tout entier $m \geq 1$, tous $0 \leq t_1 \leq \dots \leq t_m \leq s$ et toute fonction continue bornée $h : \mathbb{R}^m \rightarrow \mathbb{R}$, la propriété de martingale donne

$$\mathbb{E}[X_s^n h(X_{t_1}^n, \dots, X_{t_m}^n)] = \mathbb{E}[X_t^n h(X_{t_1}^n, \dots, X_{t_m}^n)],$$

car $h(X_{t_1}^n, \dots, X_{t_m}^n)$ est $\mathcal{F}_s^n$ -mesurable. Par représentation de Skorokhod, on peut supposer $X^n \rightarrow X$ presque sûrement dans $\mathcal{C}$. Les deux produits convergent alors presque sûrement vers les produits correspondants pour X . Ils sont uniformément intégrables, car h est bornée et les suites (X_s^n) et (X_t^n) le sont. Ainsi,

$$\mathbb{E}[X_s h(X_{t_1}, \dots, X_{t_m})] = \mathbb{E}[X_t h(X_{t_1}, \dots, X_{t_m})].$$

Le théorème de la classe monotone fonctionnelle étend cette identité à toute variable bornée $\mathcal{F}_s$ -mesurable, puisque les fonctions cylindriques continues bornées engendrent $\mathcal{F}_s$. En particulier, pour tout $A \in \mathcal{F}_s$,

$$\mathbb{E}[X_s \mathbb{1}_A] = \mathbb{E}[X_t \mathbb{1}_A].$$

Comme X est adapté et intégrable, c'est une martingale pour sa filtration naturelle.

* * *

Corrigé de l'exercice 4.7. [énoncé p. 156]

- (1) D'après le cours, la suite $(X_n)_{n \geq 1}$ est tendue dans $\mathcal{C}$ si et seulement si la suite $(X_n(0))_{n \geq 1}$ est tendue dans $\mathbb{R}$ et

$$\forall \varepsilon > 0, \forall \eta > 0, \exists \delta > 0, \quad \limsup_{n \rightarrow \infty} \mathbb{P}(\omega(X_n, \delta) \geq \eta) \leq \varepsilon. \quad (7)$$

Il s'agit donc de démontrer que (7) est équivalent à (b). Démontrons que les deux négations sont équivalentes :

- La négation de (7) est : il existe $\varepsilon, \eta > 0$ tels que pour tout $\delta > 0$,

$$\limsup_{n \rightarrow \infty} \mathbb{P}(\omega(X_n, \delta) \geq \eta) > \varepsilon.$$

- La négation de (b) est : il existe une suite $\delta_k \rightarrow 0$ et une suite $n_k \rightarrow \infty$ tels que $\omega(X_{n_k}, \delta_k)$ ne converge pas en probabilité vers 0.

En supposant la négation de (7), il existe $\varepsilon, \eta > 0$ et une suite (n_k) telle que pour tout $k \geq 1$:

$$\mathbb{P}(\omega(X_{n_k}, 1/k) \geq \eta) > \varepsilon.$$

Ainsi en prenant $\delta_k = 1/k$ on a bien que $\omega(X_{n_k}, \delta_k)$ ne converge pas en probabilité vers 0.

En supposant la négation de (b), il existe une suite $\delta_k \rightarrow 0$, $\eta, \varepsilon > 0$ et une suite (n'_k) tendant vers l'infini (après extraction ; pour un ensemble fini d'indices, les modules convergent en probabilité vers zéro) telle que pour tout $k \geq 1$

$$\mathbb{P}(\omega(X_{n'_k}, \delta_k) \geq \eta) > \varepsilon.$$

Ainsi, pour tout $\delta > 0$, pour tout k assez grand tel que $\delta_k < \delta$, on a

$$\mathbb{P}(\omega(X_{n'_k}, \delta) \geq \eta) > \varepsilon$$

On obtient donc une limite supérieure au moins égale à ε , donc strictement supérieure à $\varepsilon/2$, ce qui démontre la négation de (7).

- (2) On utilise le résultat de la question (1). On démontre tout d'abord que $(Y_n(0))_{n \geq 1}$ est tendue en montrant qu'elle est bornée dans L^1 . Ceci provient du fait que par linéarité

$$\mathbb{E}[|Y_n(0)|] \leq \mathbb{E}[|X_n(0)|] \leq \mathbb{E}\left[\sup_{t \in [0,1]} |X_n(t)|\right],$$

et puisque la suite de variables aléatoires $\sup_{t \in [0,1]} |X_n(t)|$ est uniformément intégrable elle est bornée dans L^1 .

Pour la tension, soit une suite $\delta_k \rightarrow 0$ et une suite $(n_k)_{k \geq 1}$. Vérifions que $\omega(Y_{n_k}, \delta_k)$ converge en probabilité vers 0. Tout d'abord, puisque la suite $(X_n)_{n \geq 1}$ est tendue dans $\mathcal{C}$, $\omega(X_{n_k}, \delta_k)$ converge en probabilité vers 0. D'après l'hypothèse (II), cette suite est également uniformément intégrable et donc $\mathbb{E}[\omega(X_{n_k}, \delta_k)] \rightarrow 0$. Ainsi :

$$\omega(Y_{n_k}, \delta_k) \leq \frac{1}{n_k} \sum_{i=1}^{n_k} \omega(X_{n_k}^i, \delta_k),$$

de sorte que

$$\mathbb{E}[\omega(Y_{n_k}, \delta_k)] \leq \mathbb{E}[\omega(X_{n_k}, \delta_k)] \xrightarrow[k \rightarrow \infty]{} 0.$$

Il s'ensuit que $\omega(Y_{n_k}, \delta_k)$ converge dans L^1 vers 0 et donc en probabilité.

- (3) Non. Prenons par exemple une suite $(B_n)_{n \geq 1}$ de variables aléatoires de Bernoulli de paramètres respectifs $1/n$ et prenons $(X_n^i)_{1 \leq i \leq n}$ i.i.d. de même loi que X_n défini par

$$X_n(t) = \begin{cases} n^2 B_n t & \text{si } t \leq 1/n \\ n B_n & \text{si } t \geq 1/n. \end{cases}$$

On a $\mathbb{P}(\|X_n\|_\infty > \varepsilon) \leq 1/n$ pour tout $\varepsilon > 0$. Ainsi $X_n \rightarrow 0$ en probabilité dans $\mathcal{C}$, et la suite (X_n) est tendue : la condition (I) est vérifiée. De plus, $\mathbb{E}[\|X_n\|_\infty] = 1$, ce

qui donne (II'). En revanche, la probabilité qu'un seul des X_n^i soit non nul vaut $(1 - 1/n)^{n-1} \rightarrow e^{-1}$; sur cet événement, $\omega(Y_n, 1/n) = 1$. La condition (1b) échoue donc, et (Y_n) n'est pas tendue dans $\mathcal{C}$.

Remarque. Cet exercice est inspiré de l'appendice A de l'article Berzunza Ojeda G, Janson S. The distance profile of rooted and unrooted simply generated trees. *Combinatorics, Probability and Computing*. 2022;31(3):368-410. doi:10.1017/S0963548321000304.

* * *

Corrigé de l'exercice 4.8. [énoncé p. 157]

- (1) On trouve que $\mathbb{E}[(Z_n(s) - Z_n(t))^2] = \mathbb{E}[(X_1(s) - X_1(t))^2]$.
- (2) On vérifie d'abord la tension avec le critère de tension de Kolmogorov. D'une part, pour tout $s \in [0, 1]$,

$$\mathbb{E}[Z_n(s)^2] = \mathbb{E}[X_1(s)^2] < \infty,$$

de sorte que la suite $(Z_n(s))_{n \geq 1}$ est bornée dans L^2 , donc tendue. D'autre part, par la question précédente, on a bien, pour tous $s, t \in [0, 1]$,

$$\mathbb{E}[(Z_n(s) - Z_n(t))^2] \leq C|s - t|^{1+\varepsilon}.$$

Pour identifier la limite, vérifions la convergence des marginales de dimension finie de Z_n . Pour $0 \leq t_1 < \dots < t_k \leq 1$, on peut écrire

$$(Z_n(t_1), \dots, Z_n(t_k)) = \frac{Y_1 + \dots + Y_n}{\sqrt{n}},$$

où $Y_1 = (X_1(t_1), \dots, X_1(t_k))$ avec les $(Y_i)_{1 \leq i \leq n}$ i.i.d. à valeurs dans $\mathbb{R}^k$. D'après le théorème central limite multidimensionnel, $(Z_n(t_1), \dots, Z_n(t_k))$ converge en loi lorsque $n \rightarrow \infty$ vers un vecteur gaussien centré sur $\mathbb{R}^k$ de matrice de covariance $(\mathbb{E}[X_1(t_i)X_1(t_j)])_{1 \leq i, j \leq k}$.

Ceci conclut.

* * *

Corrigé de l'exercice 4.9. [énoncé p. 157]

- (1) Posons $M = \sup_n |x_n| < \infty$ et choisissons, pour $\varepsilon > 0$, un compact $K \subset \mathcal{C}([0, 1], \mathbb{R})$ tel que $\mathbb{P}_0(K) \geq 1 - \varepsilon$. L'application

$$[-M, M] \times K \longrightarrow \mathcal{C}([0, 1], \mathbb{R}), \quad (x, f) \longmapsto x + f$$

est continue. Son image K' est donc compacte, et $\mathbb{P}_{x_n}(K') \geq \mathbb{P}_0(K) \geq 1 - \varepsilon$ pour tout n . La famille $(\mathbb{P}_{x_n})$ est tendue.

- (2) Non. En raisonnant par l'absurde, si $(\mathbb{P}_{x_n})_{n \geq 1}$ est tendue, par image continue en évaluant à l'instant 0, la suite de mesures $(\delta_{x_n})_{n \geq 1}$ serait tendue dans $\mathbb{R}$, ce qui n'est pas le cas en prenant une sous-suite non bornée.

Corrigé de l'exercice 4.10. [énoncé p. 157] Soit $F : \mathcal{C}([0, 1], \mathbb{R}) \rightarrow \mathbb{R}$ une fonction lipschitzienne bornée, de constante de Lipschitz L pour la norme uniforme. Pour $\varepsilon > 0$, on utilise le couple $(X_\varepsilon, Y_\varepsilon)$ donné dans l'énoncé. Puisque ses marginales ont les lois de X et Y , on a

$$\begin{aligned} |\mathbb{E}[F(X)] - \mathbb{E}[F(Y)]| &\leq \mathbb{E}[|F(X_\varepsilon) - F(Y_\varepsilon)|] \\ &\leq L\varepsilon + 2\|F\|_\infty \mathbb{P}(\|X_\varepsilon - Y_\varepsilon\|_\infty > \varepsilon) \\ &\leq (L + 2\|F\|_\infty)\varepsilon. \end{aligned}$$

La deuxième inégalité s'obtient en séparant l'événement $\{\|X_\varepsilon - Y_\varepsilon\|_\infty \leq \varepsilon\}$ et son complémentaire. En faisant tendre ε vers zéro, on obtient $\mathbb{E}[F(X)] = \mathbb{E}[F(Y)]$ pour toute fonction lipschitzienne bornée F . Ces fonctions caractérisent les mesures de probabilité d'après le théorème de portemanteau, donc X et Y ont même loi.

* * *

Corrigé du problème 4.11. [énoncé p. 157]

- (1) On a vu dans le cours que la tribu borélienne de $\mathcal{C}$ est égale à sa tribu produit, i.e. la plus petite tribu rendant les projections $\pi_t : \mathcal{C} \rightarrow \mathbb{R}$ mesurables pour $t \in [0, 1]$. Vérifions que celle-ci est égale à $\mathcal{F} = \{B \cap \mathcal{C} : B \in \mathcal{B}(\mathbb{R})^{\otimes [0,1]}\}$.

Tout d'abord, $\mathcal{F}$ est bien une tribu. Ensuite, pour tout $t \in [0, 1]$ et $A \in \mathcal{B}(\mathbb{R})$, $\{f \in \mathcal{C} : f(t) \in A\} = \{f \in \mathbb{R}^{[0,1]} : f(t) \in A\} \cap \mathcal{C}$, la tribu $\mathcal{F}$ rend bien toutes les projections mesurables.

Enfin, soit $\mathcal{G}$ une tribu sur $\mathcal{C}$ qui rend toutes les projections mesurables et considérons

$$\{B \in \mathcal{B}(\mathbb{R})^{\otimes [0,1]} : B \cap \mathcal{C} \in \mathcal{G}\}.$$

On vérifie que c'est une tribu de $\mathbb{R}^{[0,1]}$ qui contient les cylindres de $\mathbb{R}^{[0,1]}$, elle est donc égale à $\mathcal{B}(\mathbb{R})^{\otimes [0,1]}$, ce qui conclut.

- (2) (a) Si Y est une modification de X , pour $0 \leq t_1 < t_2 < \dots < t_n \leq 1$ on a

$$(X(t_1), \dots, X(t_n)) = (Y(t_1), \dots, Y(t_n)) \quad \text{presque sûrement,}$$

de sorte que ces deux vecteurs aléatoires ont bien même loi.

- (b) Non, il a tort : on prend par exemple ε une variable aléatoire de Bernoulli de paramètre $1/2$ et on pose $X(t) = \varepsilon$ et $Y(t) = 1 - \varepsilon$ pour $0 \leq t \leq 1$.

- (3) (a) Remarquons qu'en notant $I : \mathcal{C} \rightarrow \mathbb{R}^{[0,1]}$ l'inclusion (mesurable d'après (1)), on a $Y^* = I \circ Y$ (en particulier, Y^* est mesurable comme composée d'applications mesurables). Vérifions que $Y : (\Omega, \sigma(Y^*)) \rightarrow \mathcal{C}$ est mesurable. Soit A un ensemble mesurable de $\mathcal{C}$. D'après (1), il s'écrit $A = B \cap \mathcal{C} = I^{-1}(B)$ pour $B \in \mathcal{B}(\mathbb{R})^{\otimes [0,1]}$, de sorte que

$$Y^{-1}(A) = (Y^*)^{-1}(B)$$

est bien dans $\sigma(Y^*)$.

D'après le lemme de Doob-Dynkin, puisque $\mathcal{C}$ est polonais, il existe une application mesurable $F : \mathbb{R}^{[0,1]} \rightarrow \mathcal{C}$ telle que $Y = F \circ Y^*$.

- (b) On pose $X' = F \circ X$, qui est bien une variable aléatoire à valeurs dans $\mathcal{C}$. Vérifions que pour tout $t \in [0, 1]$, $\mathbb{P}(X(t) = X'(t)) = 1$. Remarquons que puisque X et Y ont la même loi, X et Y^* ont la même loi. Donc $(X, X') = (X, F(X))$ et $(Y^*, Y) = (Y^*, F(Y^*))$ ont même loi. Ainsi, pour tout $t \in [0, 1]$,

$$\mathbb{P}(X(t) = X'(t)) = \mathbb{P}(Y^*(t) = Y(t)) = 1.$$

- (4) (a) Puisque E est complet, une suite d'éléments de E converge si et seulement si elle est de Cauchy. Ainsi,

$$\{(u_n)_{n \geq 1} \text{ converge}\} = \bigcap_{n \geq 1} \bigcup_{N \geq 1} \bigcap_{p, q \geq N} \left\{ d(u_p, u_q) \leq \frac{1}{n} \right\},$$

et les résultats désirés s'ensuivent.

- (b) Pour $q \in [0, 1] \cap \mathbb{Q}$, notons $Z(q)$ la limite presque sûre de $(X_n(q))$, définie arbitrairement sur l'événement négligeable où elle n'existe pas. Les marginales fini-dimensionnelles de Z aux temps rationnels coïncident avec celles de Y . Pour chaque $t \in [0, 1] \setminus \mathbb{Q}$, fixons une suite rationnelle $(q_j(t))$ convergeant vers t . Par (4a) et par continuité de Y , $(Z(q_j(t)))_j$ converge presque sûrement; définissons $Z(t)$ par cette limite, et par zéro si elle n'existe pas. Chaque $Z(t)$ est mesurable. Pour toute famille finie de temps, la convergence presque sûre simultanée de ces approximations rationnelles montre que les lois conjointes correspondantes de Z et de Y coïncident. Ainsi Z est une variable aléatoire de $\mathbb{R}^{[0,1]}$ de même loi que Y .

D'après (3b), il existe une variable aléatoire $X : \Omega \rightarrow \mathcal{C}$ qui soit une modification de Z . Pour tout rationnel q , $X_n(q) \rightarrow Z(q) = X(q)$ presque sûrement, ce qui conclut.

Remarque. Le résultat reste vrai si Y n'est pas défini sur le même espace de probabilité que la suite $(X_n)_{n \geq 1}$: on peut encore prolonger Z à $[0, 1]$ en une variable aléatoire de $\mathbb{R}^{[0,1]}$ ayant les mêmes marginales fini-dimensionnelles que Y en utilisant la question (4a).

En effet, pour tout $t \in [0, 1] \setminus \mathbb{Q}$, considérons une suite de rationnels $q_n \rightarrow t$. Par continuité de Y , $Y(q_n)$ converge p.s. vers $Y(t)$. Puisque Z et Y ont les mêmes marginales fini-dimensionnelles, les deux suites $(Y(q_n))_{n \geq 1}$ et $(Z(q_n))_{n \geq 1}$ ont même loi. Par (4a), il existe donc une variable aléatoire notée $Z(t)$, limite presque sûre de $(Z(q_n))_{n \geq 1}$. Il s'ensuit que Z a les mêmes marginales fini-dimensionnelles que Y .

Remarque. Cet exercice est inspiré par le Theorem 3.24 dans O. Kallenberg, Foundations of Modern Probability (2nd Edition), Springer.

Corrigé du problème 4.12. [énoncé p. 158]**Première partie.**

- (1) Ceci découle simplement du fait qu'une fonction de la forme

$$f(x) = \max(c_0, c_1 - d(x, x_1), \dots, c_n - d(x, x_n))$$

est continue bornée.

- (2) Par linéarité de l'intégrale, on a alors
- $\int f d\mu_n \rightarrow \int f d\mu$
- pour toute fonction
- f
- lipschitzienne bornée, ce qui implique que
- $\mu_n \Rightarrow \mu$
- par le théorème de portemanteau.

- (3) Commençons par vérifier l'indication. Soit
- $f \in \text{Lip}_1(E)$
- bornée telle que
- $f \geq c_0$
- .

Si $n \geq 1$, on a $f(x) \geq c_0$ et $f(x) \geq f(x_n) - d(x, x_n)$ car f est 1-lipschitzienne. Donc $f(x) \geq \sup_{n \geq 1} \max(c_0, f(x_n) - d(x, x_n))$.

Ensuite fixons $\varepsilon > 0$ et vérifions qu'on peut trouver $n \geq 1$ tel que $f(x) \leq \max(c_0, f(x_n) - d(x, x_n)) + \varepsilon$. Par densité, soit x_n tel que $d(x, x_n) \leq \varepsilon/2$. Alors

$$\begin{aligned} f(x) - \varepsilon &= f(x_n) + f(x) - f(x_n) - \varepsilon \\ &\leq f(x_n) + d(x, x_n) - 2d(x, x_n) \\ &= f(x_n) - d(x, x_n), \end{aligned}$$

d'où le résultat.

Pour démontrer la réciproque, soit $f \in \text{Lip}_1(E)$ bornée et soit c_0 tel que $f \geq c_0$. Pour $i \geq 1$, posons

$$f_i(x) = \max(c_0, f(x_1) - d(x, x_1), \dots, f(x_i) - d(x, x_i)).$$

La suite (f_i) croît simplement vers f . Chaque f_i est exactement de la forme des fonctions test figurant dans l'hypothèse; celle-ci donne donc, pour i fixé,

$$\int f_i d\mu_n \rightarrow \int f_i d\mu.$$

Comme $\int f_i d\mu_n \leq \int f d\mu_n$, on en déduit que

$$\int f_i d\mu \leq \liminf_{n \rightarrow \infty} \int f d\mu_n.$$

Par convergence monotone, en faisant $i \rightarrow \infty$ on conclut que

$$\int f d\mu \leq \liminf_{n \rightarrow \infty} \int f d\mu_n.$$

En remplaçant f par $-f$, le résultat désiré en découle.

- (4) Comme pour (1), l'implication est immédiate. Pour la réciproque, soit
- $A > 0$
- tel que
- $d(x, y) \leq A$
- pour tout
- $x, y \in E$
- . On remarque que l'application

$$\begin{aligned} \Phi : [0, A]^n &\longrightarrow \mathbb{R} \\ (t_1, \dots, t_n) &\mapsto \max(c_0, c_1 - t_1, \dots, c_n - t_n) \end{aligned}$$

peut être approchée uniformément par des polynômes en $t_1, \dots, t_n$ d'après le théorème de Stone-Weierstrass, ce qui permet de conclure.

Deuxième partie.

(5) Pour l'implication, on écrit, pour $\varepsilon \in]0, 1[$, en utilisant l'inégalité de Markov :

$$\begin{aligned} & \lambda\{x \in [0, 1] : |f_n(x) - f(x)| \geq \varepsilon\} \\ &= \lambda\{x \in [0, 1] : \min(1, |f_n(x) - f(x)|) \geq \varepsilon\} \\ &\leq \frac{1}{\varepsilon} \int_0^1 \min(1, |f_n(t) - f(t)|) dt, \end{aligned}$$

qui tend vers 0.

Pour la réciproque, supposons que $\lambda(\{x \in [0, 1] : |f_n(x) - f(x)| \geq \varepsilon\}) \rightarrow 0$ pour tout $\varepsilon > 0$. Soit $\varepsilon \in]0, 1[$ et notons $A = \{x \in [0, 1] : |f_n(x) - f(x)| \geq \varepsilon\}$. Alors

$$\int_0^1 \min(1, |f_n(t) - f(t)|) dt \leq \int_A 1 dt + \int_{A^c} \varepsilon dt \leq \lambda(A) + \varepsilon,$$

de sorte que $d(f_n, f) \leq 2\varepsilon$ pour n assez grand.

(6) Supposons que $f_n \rightarrow f$ dans E . On choisit une extraction ϕ telle que pour tout $n \geq 1$ on ait

$$\lambda\left(\left\{x \in [0, 1] : |f_{\phi(n)}(x) - f(x)| \geq \frac{1}{2^n}\right\}\right) \leq \frac{1}{2^n}.$$

D'après le lemme de Borel-Cantelli, $f_{\phi(n)}$ converge presque partout vers f .

(7) L'espace $L^1([0, 1], dt)$ est séparable, et $d(f, g) \leq \|f - g\|_1$. Il est dense dans E : les troncatures $f^{(m)} = (-m) \vee (f \wedge m)$ appartiennent à L^1 et convergent vers f pour d , par convergence dominée. Une partie dénombrable dense de L^1 est donc dense dans E .

Pour la complétude, soit (f_n) une suite de Cauchy dans E . Choisissons une extraction ϕ telle que $d(f_{\phi(j+1)}, f_{\phi(j)}) \leq 2^{-j}$. Alors

$$\lambda\{|f_{\phi(j+1)} - f_{\phi(j)}| > 2^{-j}\} \leq 2^{-j}.$$

Par Borel-Cantelli, les différences successives sont presque partout finalement majorées par 2^{-j} . La suite extraite est donc presque partout de Cauchy et converge vers une fonction réelle mesurable f , définie arbitrairement sur l'ensemble exceptionnel. Par convergence dominée, $d(f_{\phi(j)}, f) \rightarrow 0$. La propriété de Cauchy entraîne alors $d(f_n, f) \rightarrow 0$.

Troisième partie. Pour $t \in [0, 1]$ et $\varepsilon > 0$, posons $I_{t,\varepsilon} = [0, 1] \cap [t - \varepsilon, t + \varepsilon]$ et

$$A_{t,\varepsilon}(f) = \frac{1}{\lambda(I_{t,\varepsilon})} \int_{I_{t,\varepsilon}} \arctan(f(u)) du.$$

Comme $|\arctan x - \arctan y| \leq \pi \min(1, |x - y|)$, l'application $A_{t,\varepsilon}$ est continue pour d . Pour toute $f \in \mathcal{C}$, la continuité de f donne $A_{t,1/m}(f) \rightarrow \arctan(f(t))$. Ainsi l'évaluation $\Pi_t(f) = f(t) = \tan(\lim_m A_{t,1/m}(f))$ est mesurable. L'application $(f, t) \mapsto f(t)$ est alors conjointement mesurable, puisque les trajectoires sont continues. Ceci justifie également les évaluations en des temps aléatoires et les applications de Fubini ci-dessous.

- (8) La séparabilité provient par exemple du fait que l'espace métrique des fonctions continues de $[0, 1]$ dans $\mathbb{R}$ munies de la norme uniforme est séparable, combiné avec le fait $d(f, g) \leq \|f - g\|_\infty$ pour tout $f, g \in \mathcal{C}$.

Ensuite, il est clair que $d(f, g) \leq 1$ pour tous $f, g \in \mathcal{C}$.

Enfin, $(\mathcal{C}, d)$ n'est pas complet. En effet, si on pose

$$f_n(x) = \begin{cases} 0 & \text{si } x \leq \frac{1}{2}, \\ n(x - \frac{1}{2}) & \text{si } \frac{1}{2} < x < \frac{1}{2} + \frac{1}{n}, \\ 1 & \text{si } x \geq \frac{1}{2} + \frac{1}{n}. \end{cases}$$

la suite f_n est de Cauchy pour $\|\cdot\|_1$ donc pour d , mais elle ne converge pas. En effet, si elle converge vers une fonction continue $f \in \mathcal{C}$, elle converge aussi vers f dans E . Or $f_n \rightarrow \mathbb{1}_{[1/2, 1]}$ pour $\|\cdot\|_1$ donc dans E . Ainsi $f = \mathbb{1}_{[1/2, 1]}$ presque partout, ce qui est absurde.

- (9) Soit $F : \mathbb{R}^k \rightarrow \mathbb{R}$ continue bornée. Il suffit de montrer que

$$\mathbb{E}[F(X_n(U_1), \dots, X_n(U_k))] \xrightarrow{n \rightarrow \infty} \mathbb{E}[F(X(U_1), \dots, X(U_k))].$$

À cet effet écrivons

$$\begin{aligned} & \mathbb{E}[F(X_n(U_1), \dots, X_n(U_k))] \\ &= \mathbb{E} \left[\int_{[0, 1]^k} F(X_n(t_1), \dots, X_n(t_k)) dt_1 \cdots dt_k \right] \\ &= \mathbb{E}[\Phi(X_n)], \end{aligned}$$

où

$$\begin{aligned} \Phi : \mathcal{C} &\longrightarrow \mathbb{R} \\ f &\longmapsto \int_{[0, 1]^k} F(f(t_1), \dots, f(t_k)) dt_1 \cdots dt_k. \end{aligned}$$

Il suffit alors de remarquer que Φ est continue bornée. Elle est bornée car F l'est. Si $f_n \rightarrow f$ pour d , toute sous-suite possède, d'après (6), une sous-suite extraite convergeant presque partout vers f . Par convergence dominée sur $[0, 1]^k$, les images par Φ de cette sous-suite extraite convergent vers $\Phi(f)$. Le critère de convergence par sous-suites donne donc $\Phi(f_n) \rightarrow \Phi(f)$.

- (10) Soit $(f_n)_{n \geq 1}$ une suite dense dans $\mathcal{C}$. D'après la question (4), il suffit de vérifier que pour tout $k \geq 1$ et tout polynôme P en k variables on a

$$\mathbb{E}[P(d(X_n, f_1), \dots, d(X_n, f_k))] \xrightarrow{n \rightarrow \infty} \mathbb{E}[P(d(X, f_1), \dots, d(X, f_k))]. \quad (8)$$

Or

$$\begin{aligned} & \mathbb{E}[P(d(X_n, f_1), \dots, d(X_n, f_k))] \\ &= \mathbb{E} \left[P \left(\int_0^1 \min(1, |X_n(t) - f_1(t)|) dt, \dots, \int_0^1 \min(1, |X_n(t) - f_k(t)|) dt \right) \right] \end{aligned}$$

est une combinaison linéaire d'éléments de la forme

$$\int_{[0,1]^N} \mathbb{E} \left[\prod_{i=1}^N \min(1, |X_n(t_i) - g_i(t_i)|) \right] dt_1 \cdots dt_N$$

avec $N \geq 1$ et $g_i \in \{f_1, \dots, f_k\}$ pour tout $1 \leq i \leq N$. Or, à $t_1, \dots, t_N \in [0, 1]$ fixés, on a

$$\mathbb{E} \left[\prod_{i=1}^N \min(1, |X_n(t_i) - g_i(t_i)|) \right] = \mathbb{E} [\Psi(X_n(t_1), \dots, X_n(t_N))]$$

avec

$$\begin{aligned} \Psi : \quad \mathbb{R}^N &\longrightarrow \mathbb{R} \\ (x_1, \dots, x_N) &\mapsto \prod_{i=1}^N \min(1, |x_i - g_i(t_i)|) \end{aligned}$$

continue bornée. Ainsi, par convergence dominée

$$\begin{aligned} \int_{[0,1]^N} \mathbb{E} \left[\prod_{i=1}^N \min(1, |X_n(t_i) - g_i(t_i)|) \right] dt_1 \cdots dt_N \\ \xrightarrow{n \rightarrow \infty} \int_{[0,1]^N} \mathbb{E} \left[\prod_{i=1}^N \min(1, |X(t_i) - g_i(t_i)|) \right] dt_1 \cdots dt_N. \end{aligned}$$

Ceci démontre (8) et conclut.

- (11) Supposons d'abord que X_n converge en loi vers X . Pour $k \geq 1$ et $F : \mathbb{R}^k \rightarrow \mathbb{R}$ continue bornée, posons

$$\Phi(f) = k! \int_{0 < t_1 < \dots < t_k < 1} F(f(t_1), \dots, f(t_k)) dt_1 \cdots dt_k.$$

Comme en (9), la question (6) et la convergence dominée montrent que Φ est continue bornée pour d . Puisque $(V_1^k, \dots, V_k^k)$ a pour densité $k!$ sur ce simplexe, l'indépendance donne

$$\begin{aligned} \mathbb{E} [F(X_n(V_1^k), \dots, X_n(V_k^k))] &= \mathbb{E} [\Phi(X_n)] \\ &\xrightarrow{n \rightarrow \infty} \mathbb{E} [\Phi(X)] = \mathbb{E} [F(X(V_1^k), \dots, X(V_k^k))]. \end{aligned}$$

Ceci démontre le sens direct.

Pour la réciproque, on peut prendre les $V_1^k, \dots, V_k^k$ comme le réarrangement des k premiers termes d'une même suite $(U_i)_{i \geq 1}$ de variables i.i.d. uniformes sur $[0, 1]$, indépendante de $(X_n)_{n \geq 1}$ et de X . Notons $\rho(x, y) = \min(1, |x - y|)$. Pour $g \in \mathcal{C}$ fixée, on va approcher $d(f, g)$ par

$$D_{k,g}(f) = \frac{1}{k} \sum_{i=1}^k \rho(f(V_i^k), g(i/k)).$$

Comparons cette quantité avec

$$A_{k,g}(f) = \frac{1}{k} \sum_{i=1}^k \rho(f(V_i^k), g(V_i^k)) = \frac{1}{k} \sum_{i=1}^k \rho(f(U_i), g(U_i)).$$

Pour f déterministe, les k termes de cette moyenne sont i.i.d., à valeurs dans $[0, 1]$, d'espérance $d(f, g)$ et de variance au plus $1/4$. Ainsi

$$\mathbb{E} [|A_{k,g}(f) - d(f, g)|] \leq \text{Var}(A_{k,g}(f))^{1/2} \leq \frac{1}{2\sqrt{k}}.$$

Posons ensuite

$$\Delta_k = \max_{1 \leq i \leq k} |V_i^k - i/k|, \quad \omega_g(\delta) = \sup_{\substack{s, t \in [0, 1] \\ |s-t| \leq \delta}} |g(s) - g(t)|.$$

Si F_k désigne la fonction de répartition empirique de $U_1, \dots, U_k$, l'absence presque sûre d'ex aequo donne $F_k(V_i^k) = i/k$, donc

$$\Delta_k \leq \sup_{0 \leq t \leq 1} |F_k(t) - t| \xrightarrow[k \rightarrow \infty]{p.s.} 0$$

d'après le théorème de Glivenko–Cantelli ([exercice 1.8](#)). Or

$$|D_{k,g}(f) - A_{k,g}(f)| \leq \min(1, \omega_g(\Delta_k)).$$

La continuité uniforme de g et la convergence dominée donnent ainsi

$$\sup_{f \in \mathcal{C}} \mathbb{E} [|D_{k,g}(f) - d(f, g)|] \leq \frac{1}{2\sqrt{k}} + \mathbb{E} [\min(1, \omega_g(\Delta_k))] =: \varepsilon_{k,g} \rightarrow 0.$$

Cette borne ne dépend pas de f . Par indépendance, elle reste donc valable pour $f = X_n$, uniformément en n , ainsi que pour $f = X$.

Soient maintenant $g_1, \dots, g_r \in \mathcal{C}$ et P un polynôme en r variables. Pour k fixé, l'hypothèse donne

$$\mathbb{E} [P(D_{k,g_1}(X_n), \dots, D_{k,g_r}(X_n))] \xrightarrow[n \rightarrow \infty]{} \mathbb{E} [P(D_{k,g_1}(X), \dots, D_{k,g_r}(X))],$$

car l'expression testée est une fonction continue bornée de $(X_n(V_1^k), \dots, X_n(V_k^k))$. Le polynôme P est lipschitzien sur $[0, 1]^r$: il existe $L \geq 0$ tel que $|P(a) - P(b)| \leq L \sum_{j=1}^r |a_j - b_j|$ pour $a, b \in [0, 1]^r$. Par inégalité triangulaire, la borne précédente entraîne

$$\limsup_{n \rightarrow \infty} |\mathbb{E} [P(d(X_n, g_1), \dots, d(X_n, g_r))] - \mathbb{E} [P(d(X, g_1), \dots, d(X, g_r))]| \leq 2L \sum_{j=1}^r \varepsilon_{k,g_j}.$$

En faisant tendre k vers l'infini, on obtient la convergence des espérances de tous les polynômes en ces distances. On choisit alors les g_j dans une suite dense de $(\mathcal{C}, d)$, qui est séparable et borné d'après (8), et la question (4) conclut que X_n converge en loi vers X .

Remarque. Une partie de ce problème est inspirée d'Aldous, D., & Pitman, J. (2002). Invariance principles for non-uniform random mappings and trees. In *Asymptotic Combinatorics with Application to Mathematical Physics* (pp. 113-147). Springer, Dordrecht.

Corrigé du problème 4.13. [énoncé p. 160]

(1) On commence par montrer que

$$\frac{P(t)}{t} \xrightarrow[t \rightarrow \infty]{p.s.} 1. \quad (9)$$

Il est clair que $P(t) \rightarrow \infty$ lorsque $t \rightarrow \infty$ car P est croissante et $P(S_n) = n$. On remarque que $S_{P(t)} \leq t < S_{P(t)+1}$. Soit $0 < \varepsilon < 1$. D'après la loi des grands nombres, il existe $N > 0$ tel que pour $n \geq N$ on a $|S_n/n - 1| \leq \varepsilon$. Alors pour t assez grand pour que $P(t) \geq N$ on a

$$(1 - \varepsilon)P(t) \leq S_{P(t)} \leq t < S_{P(t)+1} \leq (1 + \varepsilon)(P(t) + 1).$$

Il s'ensuit que

$$\frac{1}{1 + \varepsilon} - \frac{1}{t} \leq \frac{P(t)}{t} \leq \frac{1}{1 - \varepsilon}$$

et (9) en découle.

Revenons à la convergence uniforme sur $[0, T]$. Sur l'événement de probabilité 1 où (9) a lieu, fixons $\varepsilon > 0$ et choisissons $N > 0$ tel que $|P(t)/t - 1| \leq \varepsilon$ pour $t \geq N$. En séparant les temps $nu \leq N$ et $nu \geq N$, on obtient

$$\sup_{0 \leq u \leq T} \left| \frac{P(nu)}{n} - u \right| \leq \frac{P(N) + N}{n} + T\varepsilon.$$

Le premier terme tend vers zéro. On conclut en faisant $\varepsilon \downarrow 0$.

(2) Par définition, on remarque que pour $u = S_K/n$ on a

$$\left| \frac{P(un)}{n} - u \right| = \left| \frac{K}{n} - \frac{S_K}{n} \right| = \frac{|S_K - K|}{n}.$$

Pour tout $M > 0$, le théorème central limite donne

$$\mathbb{P}(|S_K - K| > M) \rightarrow 1 \quad (K \rightarrow \infty).$$

En effet, $(S_K - K)/\sqrt{K}$ converge en loi vers une loi normale centrée réduite, qui n'a pas d'atome en zéro, et $M/\sqrt{K} \rightarrow 0$. Par conséquent

$$\mathbb{P} \left(\sup_{K \geq 1} |S_K - K| > M \right) \geq \sup_{K \geq 1} \mathbb{P}(|S_K - K| > M) = 1.$$

En intersectant ces événements pour les entiers $M \geq 1$, on obtient $\sup_K |S_K - K| = \infty$ presque sûrement. L'identité précédente implique donc, simultanément pour tout entier $n \geq 1$,

$$\sup_{u \geq 0} \left| \frac{P(un)}{n} - u \right| = \infty \quad \text{presque sûrement.}$$

(3) Fixons $T > 0$ et posons

$$\Delta_n = \sup_{0 \leq t \leq T} \left| \frac{P(nt)}{n} - t \right|, \quad \omega_{T+1}(f, \delta) = \sup_{\substack{0 \leq s, t \leq T+1 \\ |s-t| \leq \delta}} |f(s) - f(t)|.$$

D'après (1), $\Delta_n \rightarrow 0$ presque sûrement. Pour $0 < \delta \leq 1$, sur $\{\Delta_n \leq \delta\}$, les temps t et $P(nt)/n$ appartiennent à $[0, T+1]$ et leur distance est au plus δ . Ainsi, pour tout $\eta > 0$,

$$\begin{aligned} \mathbb{P} \left(\sup_{0 \leq t \leq T} \left| Y_n \left(\frac{P(nt)}{n} \right) - Y_n(t) \right| > \eta \right) \\ \leq \mathbb{P}(\Delta_n > \delta) + \mathbb{P}(\omega_{T+1}(Y_n, \delta) > \eta). \end{aligned}$$

La convergence en loi des restrictions de Y_n à $[0, T+1]$ implique leur tension. Le critère de tension par les modules de continuité donne donc

$$\lim_{\delta \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(\omega_{T+1}(Y_n, \delta) > \eta) = 0.$$

En faisant $n \rightarrow \infty$, puis $\delta \downarrow 0$ dans la majoration précédente, on obtient la convergence en probabilité demandée. Pour $T = 0$, l'écart est identiquement nul puisque $P(0) = 0$ presque sûrement. Aucune indépendance entre P et Y_n n'est nécessaire.

(4) Ceci provient du théorème de Cauchy-Lipschitz global (β est lipschitzienne).

(5) *Première étape.* Vérifions d'abord que, presque sûrement, il existe $K > 0$ tel que $|\bar{Z}_n(t)| \leq K$ pour tout $n \geq 1$ et tout $0 \leq t \leq T$.

D'après la question (1), presque sûrement il existe deux constantes $a, b > 0$ telles que $P(t) \leq a + bt$ pour tout $t \geq 0$. Par ailleurs, comme β est lipschitzienne, il existe deux constantes $A, B > 0$ telles que $|\beta(z)| \leq A + B|z|$ pour $z \in \mathbb{R}$. Alors pour $0 \leq t \leq T$, par définition de Z_n :

$$|\bar{Z}_n(t)| \leq |\bar{Z}_n(0)| + \frac{a}{n} + b \int_0^t \beta(\bar{Z}_n(s)) ds \leq |\bar{Z}_n(0)| + \frac{a}{n} + b \int_0^t (A + B|\bar{Z}_n(s)|) ds.$$

On en déduit que

$$|\bar{Z}_n(t)| \leq |\bar{Z}_n(0)| + \frac{a}{n} + bAT + bB \int_0^t |\bar{Z}_n(s)| ds.$$

D'après le lemme de Gronwall, il en découle que pour tout $0 \leq t \leq T$:

$$|\bar{Z}_n(t)| \leq \left(|\bar{Z}_n(0)| + \frac{a}{n} + bAT \right) e^{bBT}.$$

Comme $|\bar{Z}_n(0)|$ converge lorsque $n \rightarrow \infty$, la quantité de droite est bornée, ce qui conclut la première étape.

Deuxième étape. On montre que presque sûrement il existe $M > 0$ et $L > 0$ tels que pour tout $t \in [0, T]$ on a

$$|\bar{Z}_n(t) - z(t)| \leq |\bar{Z}_n(0) - z_0| + \sup_{0 \leq s \leq t} \frac{1}{n} |\hat{P}(nMs)| + L \int_0^t |\bar{Z}_n(s) - z(s)| ds$$

où on a posé $\hat{P}(t) = P(t) - t$.

En utilisant le fait que $z(t) = z_0 + \int_0^t \beta(z(s))ds$, on écrit

$$\begin{aligned} \bar{Z}_n(t) - z(t) &= \bar{Z}_n(0) - z_0 + \frac{1}{n} \hat{P} \left(n \int_0^t \beta(\bar{Z}_n(s)) ds \right) \\ &\quad - \left(\int_0^t \beta(z(s)) ds - \int_0^t \beta(\bar{Z}_n(s)) ds \right). \end{aligned}$$

Soit $L > 0$ une constante de Lipschitz pour β et $M > 0$ tel que $|\beta(x)| \leq M$ pour tout $-K \leq x \leq K$. Alors pour $0 \leq t \leq T$:

$$|\bar{Z}_n(t) - z(t)| \leq |\bar{Z}_n(0) - z_0| + \sup_{0 \leq s \leq t} \frac{1}{n} |\hat{P}(nMs)| + L \int_0^t |\bar{Z}_n(s) - z(s)| ds.$$

Ceci conclut la deuxième étape.

Troisième étape. D'après le lemme de Gronwall, pour tout $t \in [0, T]$ on a

$$|\bar{Z}_n(t) - z(t)| \leq \left(|\bar{Z}_n(0) - z_0| + \sup_{0 \leq s \leq t} \frac{1}{n} |\hat{P}(nMs)| \right) e^{Lt}.$$

Donc

$$\sup_{0 \leq t \leq T} |\bar{Z}_n(t) - z(t)| \leq \left(|\bar{Z}_n(0) - z_0| + \sup_{0 \leq s \leq T} \frac{1}{n} |\hat{P}(nMs)| \right) e^{LT},$$

qui tend presque sûrement vers 0 par hypothèse et grâce à la question (1).

Remarque. Ce problème est inspiré du Theorem 5.2 de l'ouvrage [Stochastic Epidemic Models and Their Statistical Analysis](#) (Andersson & Britton).

* * *

Corrigé de l'exercice 4.14. [énoncé p. 160] Notons $\mathcal{F} = \{B \cap C(K, E) : B \in \mathcal{B}(E)^{\otimes K}\}$, qui est bien une tribu sur $C(K, E)$.

Puisque pour tout $x \in K$ et $A \in \mathcal{B}(E)$, $\{f \in C(K, E) : f(x) \in A\} = \{f \in E^K : f(x) \in A\} \cap C(K, E) \in \mathcal{F}$, la tribu $\mathcal{F}$ rend bien les projections Π_x mesurables.

Soit maintenant $\mathcal{G}$ une tribu sur $C(K, E)$ qui rend toutes les projections Π_x mesurables et considérons

$$\{B \in \mathcal{B}(E)^{\otimes K} : B \cap C(K, E) \in \mathcal{G}\}.$$

On vérifie que c'est une tribu qui contient les cylindres de $\mathcal{B}(E)^{\otimes K}$, elle est donc égale à $\mathcal{B}(E)^{\otimes K}$, ce qui conclut.

5 Théorème de Donsker, pont brownien

Cours associé : [chapitre 5](#), p. 55 et [chapitre 6](#), p. 62.

5.1 Énoncés

Exercice 5.1. [corrigé p. 183] Soit $(B_t)_{t \geq 0}$ un mouvement brownien standard à valeurs dans $\mathbb{R}$. On pose, pour tout $c > 0$:

$$B_t^{(c)} = \frac{1}{\sqrt{c}} B_{ct}.$$

- (1) Montrer que la famille de variables aléatoires à valeurs dans $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$ $(B^{(c)}, c > 0)$ est tendue.
- (2) On suppose que $0 < c < 1$. Montrer l'égalité en loi

$$((B_{c+t}, t \geq 0), (B_s^{(c)}, 0 \leq s \leq 1)) \stackrel{(\text{loi})}{=} ((\sqrt{c}B'_1 + B_t, t \geq 0), (B'_s, 0 \leq s \leq 1)),$$

où B' est un processus indépendant de B , et de même loi.

- (3) On considère des réels strictement positifs $0 < t_1 < t_2 < \dots < t_k \leq 1$ et $0 < t'_1 < \dots < t'_{k'} \leq 1$ et on garde les notations de la question précédente. Montrer la convergence en loi

$$(B_{t_1}, \dots, B_{t_k}, B_{t'_1}^{(c)}, \dots, B_{t'_{k'}}^{(c)}) \xrightarrow[c \rightarrow 0]{(\text{loi})} (B_{t_1}, \dots, B_{t_k}, B'_{t'_1}, \dots, B'_{t'_{k'}}).$$

- (4) En conclure que $(B, B^{(c)})$ converge en loi dans $\mathcal{C}(\mathbb{R}_+, \mathbb{R})^2$ vers (B, B') lorsque $c \rightarrow 0$. Expliquer pourquoi ce résultat peut paraître surprenant.
- (5) Que devient la question précédente si on fait cette fois tendre $c \rightarrow \infty$?

* * *

Exercice 5.2. [corrigé p. 184] Soient $(X_i)_{i \geq 1}$ des variables aléatoires réelles i.i.d. centrées de variance 1. On pose $S_0 = 0$, $S_n = X_1 + \dots + X_n$, ainsi que $A_0 = 0$, $A_n = S_1 + \dots + S_n$. On définit $(S_t)_{t \geq 0}$ et $(A_t)_{t \geq 0}$ par interpolation linéaire. On note enfin, pour $0 \leq t \leq 1$,

$$S_t^{(n)} = \frac{S_{nt}}{\sqrt{n}}, \quad A_t^{(n)} = \frac{A_{nt}}{n^{3/2}}.$$

Montrer que $(S^{(n)}, A^{(n)})$ converge en loi dans $\mathcal{C}([0, 1], \mathbb{R})^2$ vers une limite dont on précisera la loi.

* * *

Exercice 5.3. [corrigé p. 184] Soit $(b_t)_{0 \leq t \leq 1}$ un pont brownien standard.

Montrer que, presque sûrement, pour tous $s, t \in]0, 1[$, si $s \neq t$ et si s et t sont des instants de maximum local de b , alors $b_s \neq b_t$.

On pourra admettre que, presque sûrement, les valeurs des maxima locaux d'un mouvement brownien sont deux à deux distinctes.

Exercice 5.4. [corrigé p. 185] (Loi forte des grands nombres fonctionnelle)

- (1) Soient $(X_i)_{i \geq 1}$ des variables aléatoires i.i.d. réelles telles que $\mathbb{E}[|X_1|] < \infty$. On pose $S_0 = 0$, $S_n = X_1 + \dots + X_n$ et on définit S_t par interpolation linéaire pour $t \geq 0$. On note finalement

$$S_t^{(n)} = \frac{S_{nt}}{n}, \quad t \geq 0, \quad n \geq 1.$$

Montrer que presque sûrement, la convergence suivante a lieu pour la topologie de la convergence uniforme sur tout compact

$$(S_t^{(n)})_{t \geq 0} \xrightarrow{n \rightarrow \infty} (\mathbb{E}[X_1] t)_{t \geq 0}.$$

- (2) Soit $(N_t)_{t \geq 0}$ un processus de Poisson standard avec $N_0 = 0$. Montrer que presque sûrement, la convergence suivante a lieu pour la topologie de la convergence uniforme sur tout compact

$$\left(\frac{1}{n} N_{nt} \right)_{t \geq 0} \xrightarrow{n \rightarrow \infty} (t)_{t \geq 0}.$$

* * *

Exercice 5.5. [corrigé p. 185] (Détails de la preuve du théorème de Lévy) On considère une marche aléatoire symétrique simple $(S_n)_{n \geq 0}$ (c'est-à-dire $S_0 = 0$ et $S_n = X_1 + \dots + X_n$ avec $(X_i)_{i \geq 1}$ i.i.d. et $\mathbb{P}(X_1 = \pm 1) = \frac{1}{2}$). On pose

$$I_n = \min_{0 \leq k \leq n} S_k, \quad \theta_0 = 0,$$

et, pour $n \geq 1$,

$$\theta_n = \text{Card}\{1 \leq k \leq n : S_k - I_k \neq S_{k-1} - I_{k-1}\}.$$

On définit aussi $\theta_0^{-1} = 0$ et $\theta_k^{-1} = \inf\{n \geq 1 : \theta_n = k\}$ pour $k \geq 1$. Finalement, on pose $X_n = S_{\theta_n^{-1}} - I_{\theta_n^{-1}}$ et on définit $(X_t)_{t \geq 0}$ par interpolation linéaire.

- (1) Justifier que $(X_t)_{t \geq 0}$ et $(|S_t|)_{t \geq 0}$ ont même loi.
 (2) Vérifier que $(S_n - I_n)_{n \geq 0}$ est une chaîne de Markov récurrente nulle et en déduire que

$$\frac{1}{n} \sum_{k=1}^n \mathbb{1}_{S_k - I_k = 0} \xrightarrow[n \rightarrow \infty]{\text{p.s.}} 0.$$

- (3) Montrer que, pour tout $T > 0$,

$$\max_{0 \leq k \leq \lfloor nT \rfloor} \left| \frac{\theta_k^{-1}}{n} - \frac{k}{n} \right| \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0,$$

et en déduire que

$$\left(\frac{X_{nt}}{\sqrt{n}} : t \geq 0 \right) \xrightarrow[n \rightarrow \infty]{(\text{loi})} (B_t - I_t : t \geq 0).$$

- (4) Conclure que $(B_t - I_t : t \geq 0)$ et $(|B_t| : t \geq 0)$ ont même loi.

Exercice 5.6. [corrigé p. 187] Soit X une variable aléatoire réelle non constante, de fonction caractéristique notée ϕ . On rappelle que X est dite *lattice* s'il existe $b \in \mathbb{R}$ et $h > 0$ tel que $\mathbb{P}(X \in b + h\mathbb{Z}) = 1$ et que son *span* est le plus grand tel h .

- (1) Montrer que X est *lattice* si et seulement s'il existe $t_0 \in \mathbb{R}_*$ tel que $|\phi(t_0)| = 1$.
- (2) Montrer que X a pour *span* h si et seulement si $|\phi(2\pi/h)| = 1$ et $|\phi(t)| < 1$ pour $0 < |t| < 2\pi/h$.

On suppose maintenant que X est à valeurs entières. On rappelle que X est dite *apériodique* si son *span* vaut 1.

- (3) On suppose que X est *lattice* de *span* h . A-t-on forcément $h \in \mathbb{Z}$?
- (4) Est-il vrai que si le PGCD du support $\text{supp}(X)$ de X vaut 1, alors X est *apériodique*?
- (5) On suppose qu'il existe $k \in \text{supp}(X)$ tel que $\text{PGCD}(\{i - k : i \in \text{supp}(X)\}) = 1$. Montrer que X est *apériodique*.

* * *

Exercice 5.7. [corrigé p. 188] Donner un exemple d'une famille de variables aléatoires $(X_i^{(n)})_{i,n \geq 1}$ à valeurs entières telles que pour tout $n \geq 1$, les variables aléatoires $(X_i^{(n)})_{i \geq 1}$ sont i.i.d. *apériodiques*, et telles que $S_n^{(n)} = X_1^{(n)} + \dots + X_n^{(n)}$ satisfait à un théorème central limite mais pas à un théorème local limite.

* * *

Exercice 5.8. [corrigé p. 188] Soit $n \geq 1$ un entier. On considère une urne avec $2n$ boules. Parmi celles-ci, n sont étiquetées $+1$ et n sont étiquetées -1 . On tire successivement $2n$ boules, uniformément parmi les boules présentes dans l'urne à chaque tirage, on pose $S_0 = 0$ et pour $1 \leq k \leq 2n$ on note S_k la somme des k premières boules tirées. On définit S_t pour $0 \leq t \leq 2n$ par interpolation linéaire. Étudier la convergence en loi de la suite

$$\left(\frac{S_{2nt}}{\sqrt{n}} : 0 \leq t \leq 1 \right)$$

dans l'espace $\mathcal{C}([0, 1], \mathbb{R})$ dans les deux cas suivants :

- (1) les tirages se font **avec** remise,
- (2) les tirages se font **sans** remise.

* * *

Exercice 5.9. [corrigé p. 189] Soit $(X_i)_{i \geq 1}$ une suite de variables aléatoires à valeurs dans $\mathbb{Z}$, indépendantes, de même loi, de variance finie σ^2 . On suppose que $\mathbb{E}[X_1] = 0$ et que $\mathbb{P}(X_1 = k) > 0$ pour tout $k \in \mathbb{Z}$. Pour $n \geq 1$, on pose $S_n = X_1 + \dots + X_n$ et on note μ_n la loi de X_1 sachant que $S_n = 0$.

- (1) Démontrer que la suite de mesures de probabilité $(\mu_n)_{n \geq 1}$ converge étroitement.
- (2) Alix dit : en notant Y_n une variable aléatoire de loi μ_n , la suite de variables aléatoires $(Y_n)_{n \geq 1}$ est uniformément intégrable. A-t-elle raison?

Problème 5.10. [corrigé p. 189]

Première partie. On considère des variables aléatoires réelles $I_n, I_n(\delta), I, I(\delta)$ pour $\delta > 0$ et $n \geq 1$ telles que :

- (i) presque sûrement $I(\delta) \rightarrow I$ lorsque $\delta \rightarrow 0$,
- (ii) pour tout $\delta > 0$, $I_n(\delta)$ converge en loi vers $I(\delta)$ lorsque $n \rightarrow \infty$,
- (iii) pour tout $n \geq 1$ et $\delta > 0$, $\mathbb{E} [|I_n(\delta) - I_n|] \leq \delta$.

(1) Montrer que $I_n \rightarrow I$ en loi lorsque $n \rightarrow \infty$.

On considère dans la suite des variables aléatoires $(X_i)_{i \geq 1}$ i.i.d. à valeurs dans $\mathbb{Z}$ apériodiques avec $\mathbb{E}[X_1] = 0$, $\mathbb{E}[X_1^2] = 1$. On pose $S_0 = 0$ et $S_n = X_1 + \dots + X_n$ pour $n \geq 1$. Les fonctions tronquées $|x|^{-\gamma} \mathbb{1}_{|x| > \delta}$ et $|x|^{-\gamma} \mathbb{1}_{0 < |x| \leq \delta}$ sont, par convention, égales à 0 en $x = 0$. On note $(B_t)_{t \geq 0}$ un mouvement brownien réel standard issu de 0. On fixe également $0 < \gamma < 1$.

Deuxième partie.

- (2) Montrer que pour tout $\delta > 0$, presque sûrement la mesure de Lebesgue de $\{t \in [0, 1] : B_t = \delta\}$ est nulle.
- (3) Montrer que pour tout $\delta > 0$ la convergence suivante a lieu en loi

$$\frac{1}{n^{1-\gamma/2}} \sum_{i=1}^n \frac{1}{|S_i|^\gamma} \mathbb{1}_{|S_i| > \delta \sqrt{n}} \xrightarrow[n \rightarrow \infty]{(\text{loi})} \int_0^1 \frac{1}{|B_t|^\gamma} \mathbb{1}_{|B_t| > \delta} dt.$$

Troisième partie.

- (4) Montrer que presque sûrement $t \mapsto \frac{1}{|B_t|^\gamma}$ est intégrable sur $[0, 1]$ par rapport à la mesure de Lebesgue.
- (5) Montrer que la convergence suivante a lieu en loi

$$\frac{1}{n^{1-\gamma/2}} \sum_{i=1, S_i \neq 0}^n \frac{1}{|S_i|^\gamma} \xrightarrow[n \rightarrow \infty]{(\text{loi})} \int_0^1 \frac{1}{|B_t|^\gamma} dt.$$

(*) Que se passe-t-il pour $\gamma = 1$?

* * *

Problème 5.11. [corrigé p. 192] L'objet de ce problème est d'étudier le comportement asymptotique d'une marche aléatoire qui peut parfois faire plusieurs sauts identiques à la suite.

On note $\mathcal{C}$ l'ensemble des fonctions continues de $[0, 1]$ dans $\mathbb{R}$ muni de la norme uniforme.

Les trois premières parties peuvent être traitées indépendamment les unes des autres.

Première partie. Soit $H \in [1/2, 1]$. On considère une variable aléatoire $W = (W_t)_{0 \leq t \leq 1}$ à valeurs dans $\mathcal{C}$ vérifiant les propriétés suivantes :

- a) $(W_t)_{0 \leq t \leq 1}$ est un processus gaussien centré, c'est-à-dire que pour tout $k \geq 1$, pour tous $0 \leq t_1 < \dots < t_k \leq 1$, pour tous $a_1, \dots, a_k \in \mathbb{R}$, la variable aléatoire $a_1 W_{t_1} + \dots + a_k W_{t_k}$ est une variable aléatoire gaussienne centrée.
- b) Pour tous $s, t \in [0, 1]$ on a $\mathbb{E} [W_s W_t] = \frac{1}{2} (s^{2H} + t^{2H} - |s - t|^{2H})$.

La famille $(\mathbb{E} [W_s W_t])_{0 \leq s, t \leq 1}$ est appelée matrice de covariance de W . On rappelle que la loi d'un processus gaussien centré est caractérisée par sa matrice de covariance. En effet, si $(X_1, \dots, X_n)$ est un vecteur gaussien (c'est-à-dire que toute combinaison linéaire des coordonnées suit une loi normale) centré sa fonction caractéristique est, en notant $u = (u_1, \dots, u_n) \in \mathbb{R}^n$,

$$\phi_{(X_1, \dots, X_n)}(u) = \mathbb{E} \left[e^{i(u_1 X_1 + \dots + u_n X_n)} \right] = e^{-\frac{1}{2} u^T K u},$$

où ^{T} désigne la transposition et $K = (\mathbb{E} [X_i X_j])_{1 \leq i, j \leq n}$ est la matrice de covariance.

Enfin, on note $\mathbb{P}_H$ la loi de W (l'existence de $\mathbb{P}_H$ sera démontrée plus tard dans ce problème).

- (1) Si Z est une variable aléatoire de loi normale $\mathcal{N}(0, \sigma^2)$ centrée de variance σ^2 , démontrer que $\mathbb{E} [Z^4] = 3\sigma^4$.
- (2) On suppose l'existence de $\mathbb{P}_H$ pour tout $1/2 \leq H \leq 1$. Démontrer que l'application

$$\begin{array}{ccc} \Phi & : & [1/2, 1] \rightarrow (\mathcal{M}_1(\mathcal{C}), d_{LP}) \\ H & \mapsto & \mathbb{P}_H \end{array}$$

est continue.

- (3) Démontrer l'existence de $\mathbb{P}_H$ dans le cas où $H = 1$.

Deuxième partie. Soit $H \in]1/2, 1[$. Dans cette partie, $(Y_i)_{i \geq 1}$ sont des variables aléatoires réelles centrées de variance finie. On pose $S_0 = 0$ et $S_n = Y_1 + \dots + Y_n$ pour $n \geq 1$ entier. Enfin, on définit S_t pour $t \geq 0$ par interpolation linéaire et on pose pour $n \geq 1$ et $t \in [0, 1]$:

$$Z_n(t) = \frac{S_{nt}}{n^H}.$$

- (4) On suppose dans cette question que les propriétés suivantes sont satisfaites :
 - a) pour tout $n \geq 1$ et $i \geq 1$, $S_{n+i} - S_n$ a la même loi que S_i .
 - b) $\mathbb{E} [S_n^2] = O(n^{2H})$ lorsque $n \rightarrow \infty$.
 - c) (Z_n) converge en loi au sens des fini-dimensionnelles dans $\mathcal{C}$.

Démontrer que Z_n converge en loi dans $\mathcal{C}$ lorsque $n \rightarrow \infty$.

- (5) On suppose dans cette question que les propriétés suivantes sont satisfaites :
 - a) $(Y_i)_{i \geq 1}$ est une famille gaussienne.

- b) pour tout $n \geq 1$ et $i \geq 1$, $S_{n+i} - S_n$ a la même loi que S_i .
 c) $\mathbb{E}[S_n^2] \sim c \cdot n^{2H}$ lorsque $n \rightarrow \infty$, avec $c > 0$.

Démontrer que Z_n converge en loi lorsque $n \rightarrow \infty$ vers $(\sqrt{c} \cdot W_t : 0 \leq t \leq 1)$ où la loi de W a été définie dans la première partie.

Troisième partie. Soit $p \in [0, 1]$. Dans cette partie, $(X_n)_{n \geq 1}$ sont des variables aléatoires indépendantes telles que X_1 suit une loi de Bernoulli de paramètre $1/2$ et pour $n \geq 2$ la variable aléatoire X_n suit une loi de Bernoulli de paramètre $1 - p$. Pour $n \geq 1$, on pose $Y_n = (-1)^{X_1 + \dots + X_n}$, puis $S_0 = 0$ et $S_n = Y_1 + \dots + Y_n$. Ainsi, les sauts de la marche (S_n) sont ± 1 , le premier saut est ± 1 avec probabilités $1/2$ et pour $n \geq 2$ le n -ième saut est le $(n-1)$ -ième saut avec probabilité p , et l'opposé du $(n-1)$ -ième saut avec probabilité $1 - p$. Enfin, on définit S_t pour $t \geq 0$ par interpolation linéaire.

- (6) Démontrer que pour tous $i, n \geq 1$ on a $\mathbb{E}[Y_i Y_{i+n}] = (2p - 1)^n$.
 (7) Dans le cas où $p = 1/2$, démontrer que $(S_{nt}/\sqrt{n} : 0 \leq t \leq 1)$ converge en loi dans $\mathcal{C}$ vers la loi de W définie dans la première partie pour $H = 1/2$.

Quatrième partie. Soit $H \in]1/2, 1[$. Soit P une variable aléatoire à valeurs dans $[1/2, 1]$ de loi

$$(1 - H)2^{3-2H}(1 - x)^{1-2H} \mathbb{1}_{[1/2, 1]}(x) dx.$$

Dans cette partie, on note $\mathbf{X} = (X_n)_{n \geq 1}$ une suite de variables aléatoires qui, conditionnellement à P , sont indépendantes et telles que X_1 suit une loi de Bernoulli de paramètre $1/2$ et pour $n \geq 2$ la variable aléatoire X_n suit une loi de Bernoulli de paramètre $1 - P$. Formellement, si pour $p \in [0, 1]$ μ_p est la loi d'une suite de variables aléatoires indépendantes de Bernoulli dont la première a paramètre $1/2$ et les autres ont paramètre $1 - p$, la loi $\mathbb{P}_{\mathbf{X}}$ de $\mathbf{X}$ est caractérisée par le fait que pour tout événement A de $\{0, 1\}^{\mathbb{N}^*}$ on a

$$\mathbb{P}_{\mathbf{X}}(A) = \int_0^1 \mu_x(A) \cdot (1 - H)2^{3-2H}(1 - x)^{1-2H} \mathbb{1}_{[1/2, 1]}(x) dx.$$

Soit maintenant $(\mathbf{X}^k)_{k \geq 1}$ une suite de variables aléatoires i.i.d. de même loi que $\mathbf{X}$. Pour $k \geq 1$, on pose $\mathbf{X}^k = (X_n^k)_{n \geq 1}$. Pour $k, n \geq 1$, on pose $Y_n^k = (-1)^{X_1^k + \dots + X_n^k}$.

- (8) Démontrer que lorsque $k \rightarrow \infty$, $\left(\frac{Y_{n+1}^1 + \dots + Y_{n+1}^k}{\sqrt{k}} : n \geq 1 \right)$ converge en loi au sens des marginales fini-dimensionnelles vers un processus gaussien $(G_n)_{n \geq 1}$ centré qui vérifie :

- a) pour tout $n \geq 1$, $\mathbb{E}[G_n^2] = 1$.
 b) pour tout $n \geq 1$ et $i \geq 1$, la quantité $r(n) = \mathbb{E}[G_i G_{i+n}]$ ne dépend pas de i et on a

$$r(n) \underset{n \rightarrow \infty}{\sim} \frac{1}{c_H^2} \frac{H(2H - 1)}{n^{2-2H}} \quad \text{avec } c_H = \sqrt{\frac{H(2H - 1)}{\Gamma(3 - 2H)}}.$$

Remarque. On admettra que $\int_{1/2}^1 (2u-1)^n (1-u)^{1-2H} du \sim \frac{1}{c_H^2} \frac{H(2H-1)}{(1-H)^{2^{3-2H}}} \cdot \frac{1}{n^{2-2H}}$ lorsque $n \rightarrow \infty$.

- (9) On pose $V_0 = 0$ et $V_n = G_1 + \dots + G_n$ pour $n \geq 1$ et on définit V_t pour $t \geq 0$ par interpolation linéaire. Démontrer que

$$\left(c_H \frac{V_{nt}}{n^H} : 0 \leq t \leq 1 \right) \xrightarrow[n \rightarrow \infty]{(loi)} (W_t : 0 \leq t \leq 1),$$

où la loi de W a été définie dans la première partie.

5.2 Corrigés

Corrigé de l'exercice 5.1. [énoncé p. 177]

- (1) Pour tout $c > 0$, le processus $B^{(c)}$ est continu, gaussien centré, et

$$\mathbb{E} \left[B_s^{(c)} B_t^{(c)} \right] = \frac{1}{c} \min(cs, ct) = \min(s, t).$$

Il a donc la loi du mouvement brownien. La famille de ses lois est réduite à une seule mesure de probabilité sur l'espace polonais $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$: elle est tendue.

- (2) Le processus $(B_{c+t} - B_c)_{t \geq 0}$ est un mouvement brownien indépendant de $(B_s)_{0 \leq s \leq c}$, et donc de $(B_s^{(c)})_{0 \leq s \leq 1}$. Comme $B_c = \sqrt{c} B_1^{(c)}$, l'égalité en loi annoncée en découle.
- (3) D'après (2),

$$\begin{aligned} (B_{c+t_i} : 1 \leq i \leq k; B_{t'_j}^{(c)} : 1 \leq j \leq k') \\ \stackrel{(loi)}{=} (\sqrt{c} B_1' + B_{t_i} : 1 \leq i \leq k; B_{t'_j}' : 1 \leq j \leq k'). \end{aligned}$$

Le membre de droite converge presque sûrement vers le vecteur voulu lorsque $c \rightarrow 0$. Par continuité de B , $\max_{1 \leq i \leq k} |B_{c+t_i} - B_{t_i}| \rightarrow 0$ presque sûrement. Le lemme de Slutsky permet donc de remplacer B_{c+t_i} par B_{t_i} .

- (4) La tension des deux marginales entraîne celle des couples : si un compact K de $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$ vérifie $\mathbb{P}(B \notin K) \leq \varepsilon$, alors

$$\mathbb{P} \left((B, B^{(c)}) \notin K \times K \right) \leq 2\varepsilon.$$

Le raisonnement de (2)–(3) s'applique sur tout $[0, T]$: il suffit de remplacer c par cT dans le décalage du premier processus et $\sqrt{c} B_1'$ par $\sqrt{c} B_T'$. Les lois finidimensionnelles des couples convergent donc vers celles de deux mouvements browniens indépendants. Avec la tension, ceci donne la convergence en loi dans $\mathcal{C}(\mathbb{R}_+, \mathbb{R})^2$.

Ce résultat peut paraître surprenant car, pour chaque c , le processus $B^{(c)}$ est entièrement déterminé par B , alors que les deux processus limites sont indépendants.

- (5) La limite est encore un couple de mouvements browniens indépendants. En effet, si $W = B^{(c)}$, alors W est un mouvement brownien et $W^{(1/c)} = B$. Ainsi

$$(B, B^{(c)}) \stackrel{(\text{loi})}{=} (B^{(1/c)}, B).$$

On applique (4) avec $1/c \rightarrow 0$, puis on échange les deux coordonnées.

* * *

Corrigé de l'exercice 5.2. [énoncé p. 177] La limite est le couple

$$\left((B_t)_{0 \leq t \leq 1}, \left(\int_0^t B_s ds \right)_{0 \leq t \leq 1} \right),$$

où B est un mouvement brownien standard. Pour le voir, posons $J_t^{(n)} = \int_0^t S_s^{(n)} ds$. L'application

$$f \mapsto \left(f, \left(t \mapsto \int_0^t f(s) ds \right) \right)$$

est continue de $\mathcal{C}([0, 1], \mathbb{R})$ dans son carré, pour les normes uniformes. Le théorème de Donsker et le théorème de l'application continue donnent donc la convergence annoncée pour $(S^{(n)}, J^{(n)})$.

Il reste à comparer $A^{(n)}$ et $J^{(n)}$. L'interpolation linéaire de A donne l'identité

$$A_t^{(n)} = \int_0^t \frac{S_{[ns]}}{\sqrt{n}} ds.$$

Sur l'intervalle $[(i-1)/n, i/n]$, $|S_{[ns]} - S_{ns}| \leq |X_i|$ pour presque tout s . Ainsi

$$\sup_{0 \leq t \leq 1} |A_t^{(n)} - J_t^{(n)}| \leq \frac{1}{n\sqrt{n}} \sum_{i=1}^n |X_i|,$$

dont l'espérance vaut $\mathbb{E}[|X_1|] / \sqrt{n}$ et tend vers 0. Le lemme de Slutsky conclut.

* * *

Corrigé de l'exercice 5.3. [énoncé p. 177] Fixons $r \in]0, 1[$. D'après la Proposition 6.2, la loi de $(b_t)_{0 \leq t \leq r}$ est absolument continue par rapport à celle d'un mouvement brownien sur $[0, r]$. La propriété admise pour le mouvement brownien implique donc que, presque sûrement, les valeurs de b en ses instants de maximum local dans $]0, r[$ sont deux à deux distinctes.

Pour $m \geq 2$, posons $r_m = 1 - 1/m$. En intersectant les événements précédents pour tous les m , on obtient un événement de probabilité 1. Sur cet événement, si $s, t \in]0, 1[$ sont deux instants distincts de maximum local de b , on choisit m assez grand pour que $\max(s, t) < r_m$. Ces deux instants sont aussi des instants de maximum local de la restriction de b à $[0, r_m]$, donc $b_s \neq b_t$. Ceci conclut.

Corrigé de l'exercice 5.4. [énoncé p. 178]

- (1) Posons $m = \mathbb{E}[X_1]$ et $R_k = S_k - km$. La loi forte des grands nombres donne $R_k/k \rightarrow 0$ presque sûrement. Sur cet événement, pour tout $\varepsilon > 0$, il existe K tel que $|R_k| \leq \varepsilon k$ pour $k \geq K$. Pour tout $T > 0$,

$$\frac{1}{n} \max_{0 \leq k \leq \lceil nT \rceil} |R_k| \leq \frac{1}{n} \max_{0 \leq k < K} |R_k| + \varepsilon \frac{\lceil nT \rceil}{n}.$$

En faisant tendre n vers l'infini, puis ε vers 0, on voit que le membre de gauche tend vers 0. L'interpolation linéaire donne alors

$$\sup_{0 \leq t \leq T} \left| \frac{S_{nt}}{n} - mt \right| \leq \frac{1}{n} \max_{0 \leq k \leq \lceil nT \rceil} |R_k| \rightarrow 0.$$

Ceci vaut sur le même événement de probabilité 1 pour tout $T > 0$.

- (2) Les accroissements $(N_k - N_{k-1})_{k \geq 1}$ sont i.i.d. de loi de Poisson de paramètre 1. D'après (1), presque sûrement, pour tout $T > 0$,

$$\frac{1}{n} \max_{0 \leq k \leq \lceil nT \rceil} |N_k - k| \rightarrow 0.$$

Comme $N_{\lfloor nt \rfloor} \leq N_{nt} \leq N_{\lceil nt \rceil}$, on obtient

$$\sup_{0 \leq t \leq T} \left| \frac{N_{nt}}{n} - t \right| \leq \frac{1}{n} \max_{0 \leq k \leq \lceil nT \rceil} |N_k - k| + \frac{1}{n} \rightarrow 0.$$

* * *

Corrigé de l'exercice 5.5. [énoncé p. 178] Posons $R_n = S_n - I_n$ et $\tau_k = \theta_k^{-1}$. Un pas de R est nul exactement lorsque S atteint un nouveau minimum strict. Par conséquent,

$$n - \theta_n = -I_n.$$

La loi forte des grands nombres, sous sa forme fonctionnelle démontrée dans l'exercice précédent, entraîne $I_n/n \rightarrow 0$ presque sûrement. Ainsi $\theta_n/n \rightarrow 1$, et tous les temps τ_k sont presque sûrement finis.

- (1) La chaîne R a pour probabilités de transition

$$p(0,0) = p(0,1) = \frac{1}{2}, \quad p(j,j-1) = p(j,j+1) = \frac{1}{2} \quad (j \geq 1).$$

La chaîne $(X_k)_{k \geq 0} = (R_{\tau_k})_{k \geq 0}$ s'obtient en supprimant les pas nuls en 0. Par indépendance des accroissements de la marche aux temps d'arrêt, elle part de 0, passe de 0 à 1 avec probabilité 1, et, de tout $j \geq 1$, passe à $j-1$ ou $j+1$ avec probabilités 1/2. Ce sont exactement les transitions de $(|S_k|)_{k \geq 0}$, donc ces chaînes ont même loi. Enfin, pour une marche à pas ± 1 , $|S_t|$ est l'interpolation linéaire des $|S_k|$: aucun segment ne traverse 0 en son intérieur. L'égalité en loi vaut donc aussi pour les processus continus.

- (2) La chaîne R est irréductible sur $\mathbb{N}$. Jusqu'à son premier passage en 0, elle évolue comme une marche symétrique simple. Pour cette marche issue de $j \in \{1, \dots, M-1\}$, notons H_0 et H_M les temps d'atteinte de 0 et de M . Les formules de ruine donnent

$$\mathbb{P}_j(H_0 < H_M) = 1 - \frac{j}{M}, \quad \mathbb{E}_1[H_0 \wedge H_M] = M - 1.$$

On les retrouve en résolvant, avec les conditions aux bords, les relations $q_j = (q_{j-1} + q_{j+1})/2$ et $u_j = 1 + (u_{j-1} + u_{j+1})/2$: $q_j = 1 - j/M$ avec $q_0 = 1, q_M = 0$, et $u_j = j(M-j)$. En faisant $M \rightarrow \infty$, on obtient $\mathbb{P}_j(H_0 < \infty) = 1$ et $\mathbb{E}_1[H_0] = \infty$. Ainsi R est récurrente. Issue de 0, elle fait son premier pas vers 1 avec probabilité $1/2$; son temps de premier retour en 0 a donc une espérance infinie. Elle est récurrente nulle.

Pour en déduire l'assertion demandée, notons $0 = T_0 < T_1 < T_2 < \dots$ ses retours successifs en 0. Les variables $T_i - T_{i-1}$ sont i.i.d. positives, d'espérance infinie. Pour chaque entier $M \geq 1$, la loi forte donne

$$\liminf_{j \rightarrow \infty} \frac{T_j}{j} \geq \mathbb{E}[\min(T_1, M)] \quad \text{presque sûrement.}$$

En faisant tendre M vers l'infini, on voit que $T_j/j \rightarrow \infty$ presque sûrement. Si $L_n = \max\{j : T_j \leq n\}$, alors

$$\frac{1}{n} \sum_{k=1}^n \mathbb{1}_{R_k=0} = \frac{L_n}{n} \leq \frac{L_n}{T_{L_n}} \rightarrow 0$$

presque sûrement, l'inégalité étant utilisée pour $L_n \geq 1$.

- (3) La convergence $\theta_n/n \rightarrow 1$ entraîne $\tau_k/k \rightarrow 1$ presque sûrement : en effet, $\theta_{\tau_k} = k$ et $\tau_k \rightarrow \infty$. De plus, $\tau_{k+1} \geq \tau_k + 1$, donc $\tau_k - k$ est croissante et positive. Pour tout $T > 0$,

$$\max_{0 \leq k \leq \lfloor nT \rfloor} \left| \frac{\tau_k}{n} - \frac{k}{n} \right| = \frac{\tau_{\lfloor nT \rfloor} - \lfloor nT \rfloor}{n} \rightarrow 0 \quad \text{presque sûrement.}$$

On a donc même une convergence presque sûre dans l'assertion demandée.

Prolongeons R par interpolation linéaire et posons $R_t^{(n)} = R_{nt}/\sqrt{n}$. Pour la marche simple interpolée, $R_t = S_t - \inf_{0 \leq s \leq t} S_s$. L'application qui retranche l'infimum courant est continue pour la convergence uniforme sur les compacts, puisque

$$\sup_{t \leq T} \left| \left(f(t) - \inf_{s \leq t} f(s) \right) - \left(g(t) - \inf_{s \leq t} g(s) \right) \right| \leq 2 \sup_{t \leq T} |f(t) - g(t)|.$$

Le théorème de Donsker donne donc $R^{(n)} \Rightarrow (B_t - I_t)_{t \geq 0}$.

Soit λ_n la fonction affine par morceaux prenant la valeur τ_k/n au temps k/n . D'après ce qui précède, λ_n converge presque sûrement vers l'identité, uniformément sur les compacts. Le lemme de Slutsky et la continuité de la composition donnent

$$R^{(n)} \circ \lambda_n \Rightarrow (B_t - I_t)_{t \geq 0}.$$

Il faut tenir compte de l'interpolation : si $\tau_{k+1} = \tau_k + 1$, les trajectoires $X_{nt}/\sqrt{n}$ et $R_{\lambda_n(t)}^{(n)}$ coïncident sur $[k/n, (k+1)/n]$. Sinon, R reste en 0 entre τ_k et $\tau_{k+1} - 1$, puis monte à 1, tandis que X interpole directement de 0 à 1. Dans tous les cas,

$$\sup_{t \geq 0} \left| \frac{X_{nt}}{\sqrt{n}} - R_{\lambda_n(t)}^{(n)} \right| \leq \frac{1}{\sqrt{n}}.$$

Ceci donne la convergence en loi demandée.

- (4) D'après (1) et le théorème de Donsker, $(X_{nt}/\sqrt{n})_{t \geq 0}$ converge aussi en loi vers $(|B_t|)_{t \geq 0}$, l'application $f \mapsto |f|$ étant continue. L'unicité de la limite en loi donne

$$(B_t - I_t)_{t \geq 0} \stackrel{(\text{loi})}{=} (|B_t|)_{t \geq 0}.$$

* * *

Corrigé de l'exercice 5.6. [énoncé p. 179]

- (1) Si $\mathbb{P}(X \in b + h\mathbb{Z}) = 1$, alors $|\phi(2\pi/h)| = 1$. Réciproquement, si $|\phi(t_0)| = 1$, la variable de module 1 $e^{it_0 X}$ est presque sûrement constante. Il existe donc $b \in \mathbb{R}$ tel que

$$\mathbb{P}\left(X \in b + \frac{2\pi}{|t_0|}\mathbb{Z}\right) = 1.$$

- (2) La question (1) montre que tout $t \neq 0$ tel que $|\phi(t)| = 1$ fournit un pas lattice $2\pi/|t|$. Ainsi, h est le span si et seulement si h est un pas lattice et aucun pas strictement supérieur à h n'est possible, ce qui équivaut exactement aux deux conditions annoncées.
- (3) Oui (pour des variables aléatoires non constantes) : si $\mathbb{P}(X \in b + h\mathbb{Z}) = 1$, en prenant $k_1 \neq k_2$ dans le support de X , on a $k_1 = b + ha_1$ et $k_2 = b + ha_2$ pour des entiers a_1 et a_2 , de sorte que h est rationnel. Supposons maintenant par l'absurde que le span de X est rationnel non entier. En l'écrivant sous la forme $h = p/q$ avec $p, q \geq 1$ premiers entre eux, et en considérant $k_1 - k_2$, on voit que p divise $k_1 - k_2$. Puisque le span d'une variable aléatoire est invariant par l'addition d'une même quantité à tous les éléments de son support, on peut supposer que tous les éléments du support de X sont divisibles par p . La variable aléatoire $Y = X/p$ a alors pour span $1/q$, ce qui est absurde car le span d'une variable aléatoire entière vaut clairement au moins 1.
- (4) Non, prendre par exemple X de support $\{1, 3\}$; le span vaut 2.
- (5) Soit $k \in \text{supp}(X)$ tel que $\text{PGCD}(\{i - k : i \in \text{supp}(X)\}) = 1$. D'après (3), le span h est un entier strictement positif. Comme le support est inclus dans un translaté de $h\mathbb{Z}$, h divise chaque différence $i - k$, donc leur PGCD. Ainsi $h = 1$.

Corrigé de l'exercice 5.7. [énoncé p. 179] Pour $n \geq 2$, prenons dans chaque ligne des variables i.i.d. de loi

$$\mathbb{P}(X_1^{(n)} = 0) = 1 - e^{-n} - n^{-1/2}, \quad \mathbb{P}(X_1^{(n)} = 1) = e^{-n}, \quad \mathbb{P}(X_1^{(n)} = 2) = n^{-1/2}.$$

Ces trois probabilités sont positives : la loi est apériodique. Pour $n = 1$, prenons une loi de Bernoulli de paramètre $1/2$. Posons $m_n = \mathbb{E}[S_n^{(n)}]$ et $s_n^2 = \text{Var}(S_n^{(n)})$. Pour $n \geq 2$,

$$m_n = n(e^{-n} + 2n^{-1/2}), \quad s_n^2 = n(e^{-n} + 4n^{-1/2} - (e^{-n} + 2n^{-1/2})^2) \sim 4\sqrt{n}.$$

Les variables centrées de chaque ligne sont bornées en valeur absolue par 2, tandis que $s_n \rightarrow \infty$. La condition de Lindeberg est donc satisfaite, d'où

$$\frac{S_n^{(n)} - m_n}{s_n} \Rightarrow \mathcal{N}(0, 1), \quad \frac{S_n^{(n)} - m_n}{2n^{1/4}} \Rightarrow \mathcal{N}(0, 1).$$

En revanche, la convergence locale

$$\sup_{k \in \mathbb{Z}} \left| s_n \mathbb{P}(S_n^{(n)} = k) - \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(k - m_n)^2}{2s_n^2}\right) \right| \rightarrow 0$$

est fautive. En effet, choisissons un entier impair k_n avec $|k_n - m_n| \leq 1$. Une somme impaire contient au moins un terme égal à 1, donc

$$s_n \mathbb{P}(S_n^{(n)} = k_n) \leq s_n \mathbb{P}(S_n^{(n)} \text{ est impair}) \leq s_n n e^{-n} \rightarrow 0,$$

alors que le terme gaussien en k_n tend vers $1/\sqrt{2\pi}$.

* * *

Corrigé de l'exercice 5.8. [énoncé p. 179] Soit $(X_i)_{i \geq 1}$ une suite de variables aléatoires i.i.d. de loi $\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = 1/2$ et notons $S'_n = X_1 + \dots + X_n$ pour $n \geq 0$. On a $\mathbb{E}[X_1] = 0$ et $\mathbb{E}[X_1^2] = 1$.

- (1) On a convergence en loi vers $\sqrt{2}$ fois le mouvement brownien sur $[0, 1]$. On remarque que $(S_i)_{0 \leq i \leq 2n}$ a la même loi que $(S'_i)_{0 \leq i \leq 2n}$, et le résultat découle du théorème de Donsker.
- (2) On montre que la suite converge en loi vers $\sqrt{2}$ fois le pont brownien. Les trajectoires $(S_i)_{0 \leq i \leq 2n}$ ont la même loi que $(S'_i)_{0 \leq i \leq 2n}$ conditionnées par $S'_{2n} = 0$. Pour utiliser exactement la version apériodique du théorème, posons

$$Y_i = \frac{X_{2i-1} + X_{2i}}{2}, \quad T_k = Y_1 + \dots + Y_k = \frac{S'_{2k}}{2}.$$

Les Y_i sont i.i.d., centrées, de variance $1/2$, et leur loi sur $\{-1, 0, 1\}$, de probabilités $1/4, 1/2, 1/4$, est apériodique. En interpolant T linéairement, Donsker conditionné donne

$$\left(\frac{T_{nt}}{\sqrt{n/2}} : 0 \leq t \leq 1 \right) \text{ sous } \mathbb{P}(\cdot | T_n = 0) \Rightarrow b.$$

Sur chaque intervalle de deux pas, les deux interpolations vérifient $|S'_{2nt} - 2T_{nt}| \leq 1$. Après division par $\sqrt{n}$, cet écart tend uniformément vers zéro. Comme $T_n = 0$ équivaut à $S'_{2n} = 0$, le lemme de Slutsky donne la limite $\sqrt{2}b$ annoncée.

Corrigé de l'exercice 5.9. [énoncé p. 179]

(1) On écrit, pour $k \in \mathbb{Z}$:

$$\mu_n(k) = \mathbb{P}(X_1 = k | S_n = 0) = \mathbb{P}(X_1 = k) \cdot \frac{\mathbb{P}(S_{n-1} = -k)}{\mathbb{P}(S_n = 0)}.$$

D'après le théorème local limite le numérateur et le dénominateur sont tous les deux équivalents à $\frac{1}{\sqrt{2\pi\sigma^2 n}}$, de sorte que

$$\mu_n(k) \xrightarrow{n \rightarrow \infty} \mathbb{P}(X_1 = k).$$

Ainsi, μ_n converge étroitement vers la loi de X_1 .

(2) Oui, elle a raison. Vérifions que $(Y_n)_{n \geq 1}$ est bornée dans L^2 , ce qui entraînera le résultat désiré. On a

$$\mathbb{E}[Y_n^2] = \sum_{k \in \mathbb{Z}} k^2 \mathbb{P}(X_1 = k) \cdot \frac{\mathbb{P}(S_{n-1} = -k)}{\mathbb{P}(S_n = 0)}.$$

D'après le théorème local limite uniforme, il existe $C > 0$ tel que, pour tout $n \geq 2$ et $k \in \mathbb{Z}$,

$$\frac{\mathbb{P}(S_{n-1} = -k)}{\mathbb{P}(S_n = 0)} \leq C.$$

En effet, le numérateur est uniformément $O(n^{-1/2})$ et le dénominateur est équivalent à $(2\pi\sigma^2 n)^{-1/2}$. On absorbe les valeurs initiales de n dans C puisque $\mathbb{P}(S_n = 0) > 0$. Ainsi $\mathbb{E}[Y_n^2] \leq C \mathbb{E}[X_1^2]$ pour $n \geq 2$. Enfin, $Y_1 = 0$ presque sûrement. La suite est donc bornée dans L^2 , et par conséquent uniformément intégrable.

* * *

Corrigé du problème 5.10. [énoncé p. 180]

(1) C'est une variante du principe des lois accompagnantes. On montre que pour tout fermé F

$$\limsup_{n \rightarrow \infty} \mathbb{P}(I_n \in F) \leq \mathbb{P}(I \in F). \quad (10)$$

D'après l'inégalité de Markov, pour $\eta > 0$ on a $\mathbb{P}(|I_n(\delta) - I_n| \geq \eta) \leq \frac{\delta}{\eta}$. Donc en notant F^η le η -voisinage fermé de F on a

$$\limsup_{n \rightarrow \infty} \mathbb{P}(I_n \in F) \leq \limsup_{n \rightarrow \infty} \mathbb{P}(I_n(\delta) \in F^\eta) + \frac{\delta}{\eta} \leq \mathbb{P}(I(\delta) \in F^\eta) + \frac{\delta}{\eta}$$

en vertu de la convergence en loi de $I_n(\delta)$ vers $I(\delta)$. En faisant $\delta \rightarrow 0$, puis $\eta \rightarrow 0$ on obtient le résultat voulu (10).

Autre solution. Alternativement, on peut utiliser la caractérisation de la convergence en loi via les fonctions lipschitziennes bornées (théorème de portemanteau).

(2) On applique le théorème de Fubini à $(\omega, t) \mapsto \mathbb{1}_{B_t=\delta}$:

$$\mathbb{E} \left[\int_0^1 \mathbb{1}_{B_t=\delta} dt \right] = \int_0^1 \mathbb{P}(B_t = \delta) dt = 0,$$

donc p.s. $\int_0^1 \mathbb{1}_{B_t=\delta} dt = 0$, ce qui donne le résultat.

(3) Pour $\delta > 0$, posons

$$g_\delta(x) = \begin{cases} |x|^{-\gamma}, & |x| > \delta, \\ 0, & |x| \leq \delta. \end{cases}$$

On a l'identité exacte

$$\frac{1}{n^{1-\gamma/2}} \sum_{i=1}^n |S_i|^{-\gamma} \mathbb{1}_{|S_i| > \delta\sqrt{n}} = \int_0^1 g_\delta \left(\frac{S_{\lceil nt \rceil}}{\sqrt{n}} \right) dt. \quad (11)$$

Par Donsker et la représentation de Skorokhod, on peut réaliser les marches interpolées $S_{nt}/\sqrt{n}$ de sorte qu'elles convergent uniformément presque sûrement sur $[0, 1]$ vers B . Leurs versions en escalier $S_{\lceil nt \rceil}/\sqrt{n}$ convergent alors aussi uniformément vers B , par continuité uniforme de la trajectoire limite. Les seules discontinuités de g_δ sont δ et $-\delta$. D'après (2), appliqué aussi à $-B$, ces niveaux sont visités pendant un ensemble de temps de mesure de Lebesgue nulle, presque sûrement. Enfin $|g_\delta| \leq \delta^{-\gamma}$. La convergence dominée donne donc la convergence annoncée.

(4) On montre que

$$\mathbb{E} \left[\int_0^1 \frac{1}{|B_t|^\gamma} dt \right] < \infty \quad (12)$$

ce qui impliquera que p.s. $\int_0^1 \frac{1}{|B_t|^\gamma} dt < \infty$. D'après le théorème de Fubini et d'après la propriété de scaling du mouvement brownien,

$$\mathbb{E} \left[\int_0^1 \frac{1}{|B_t|^\gamma} dt \right] = \int_0^1 \mathbb{E} \left[\frac{1}{|B_t|^\gamma} \right] dt = \int_0^1 \frac{1}{t^{\gamma/2}} dt \cdot \mathbb{E} \left[\frac{1}{|B_1|^\gamma} \right].$$

Puisque $\gamma/2 < 1$, il suffit de vérifier que $\mathbb{E}[1/|B_1|^\gamma] < \infty$. Pour cela, on écrit simplement

$$\mathbb{E} \left[\frac{1}{|B_1|^\gamma} \right] = \int_{\mathbb{R}} \frac{1}{|x|^\gamma} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx < \infty$$

car $\gamma < 1$, et (12) s'ensuit.

(5) On pose

$$I_n = \frac{1}{n^{1-\gamma/2}} \sum_{i=1, S_i \neq 0}^n \frac{1}{|S_i|^\gamma}, \quad I_n(\delta) = \frac{1}{n^{1-\gamma/2}} \sum_{i=1}^n \frac{1}{|S_i|^\gamma} \mathbb{1}_{|S_i| > \delta\sqrt{n}}$$

et

$$I = \int_0^1 \frac{1}{|B_t|^\gamma} dt, \quad I(\delta) = \int_0^1 \frac{1}{|B_t|^\gamma} \mathbb{1}_{|B_t| > \delta} dt$$

et on vérifie qu'on peut appliquer (1).

Le point (i) découle par convergence monotone, le point (ii) provient de la question (3). Le point délicat est (iii). Pour cela, écrivons :

$$\mathbb{E} [|I_n(\delta) - I_n|] = \frac{1}{n^{1-\gamma/2}} \sum_{i=1}^n \mathbb{E} \left[\frac{1}{|S_i|^\gamma} \mathbb{1}_{0 < |S_i| \leq \delta\sqrt{n}} \right].$$

Par corollaire du théorème local limite, il existe une constante $C > 0$ telle que $\mathbb{P}(S_i = k) \leq \frac{C}{\sqrt{i}}$ pour tout $k \in \mathbb{Z}$ et $i \geq 1$ (dans la suite, C est une constante dépendant de la loi des incréments et de γ qui peut changer de ligne en ligne). Ainsi

$$\begin{aligned} \mathbb{E} \left[\frac{1}{|S_i|^\gamma} \mathbb{1}_{0 < |S_i| \leq \delta\sqrt{n}} \right] &= \sum_{k=1}^{[\delta\sqrt{n}]} \frac{1}{k^\gamma} \mathbb{P}(|S_i| = k) \\ &\leq C \sum_{k=1}^{[\delta\sqrt{n}]} \frac{1}{k^\gamma} \cdot \frac{1}{\sqrt{i}} \\ &\leq C \cdot \delta^{1-\gamma} \cdot n^{1/2-\gamma/2} \cdot \frac{1}{\sqrt{i}}. \end{aligned}$$

Donc

$$\begin{aligned} \frac{1}{n^{1-\gamma/2}} \sum_{i=1}^n \mathbb{E} \left[\frac{1}{|S_i|^\gamma} \mathbb{1}_{0 < |S_i| \leq \delta\sqrt{n}} \right] &\leq C \frac{\delta^{1-\gamma} \cdot n^{1/2-\gamma/2}}{n^{1-\gamma/2}} \sum_{i=1}^n \frac{1}{\sqrt{i}} \\ &\leq C \frac{\delta^{1-\gamma} \cdot n^{1/2-\gamma/2}}{n^{1-\gamma/2}} \cdot n^{1/2} \leq C \delta^{1-\gamma}. \end{aligned}$$

Pour appliquer exactement (iii), posons

$$r(\delta) = (\delta/C)^{1/(1-\gamma)}.$$

Remplaçons $I_n(\delta)$, $I(\delta)$ par $I_n(r(\delta))$, $I(r(\delta))$. La borne obtenue est alors

$$\mathbb{E} [|I_n(r(\delta)) - I_n|] \leq \delta,$$

et les propriétés (i) et (ii) sont conservées.

(*) Pour $\gamma = 1$, on a

$$\int_0^1 \frac{dt}{|B_t|} = +\infty \quad \text{presque sûrement.}$$

On peut le voir avec le temps local brownien, normalisé par $\int_0^1 h(B_t) dt = \int_{\mathbb{R}} h(x) L_1^x dx$. On admet ici les propriétés usuelles : $x \mapsto L_1^x$ est continue et $L_1^0 > 0$ presque sûrement. La formule d'occupation et la divergence de $\int_0^\varepsilon dx/x$ donnent alors l'assertion.

La quantité discrète diverge également :

$$J_n := \frac{1}{\sqrt{n}} \sum_{\substack{1 \leq i \leq n \\ S_i \neq 0}} \frac{1}{|S_i|} \rightarrow +\infty \quad \text{en probabilité.}$$

En effet, pour $\delta > 0$ fixé, la preuve de (3) reste valable avec $\gamma = 1$ et donne

$$J_n(\delta) := \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\mathbb{1}_{\{|S_i| > \delta\sqrt{n}\}}}{|S_i|} \Rightarrow J(\delta) := \int_0^1 \frac{\mathbb{1}_{\{|B_t| > \delta\}}}{|B_t|} dt,$$

avec la convention de l'énoncé pour les termes nuls. Comme $J_n \geq J_n(\delta)$, pour tout $M > 0$,

$$\liminf_{n \rightarrow \infty} \mathbb{P}(J_n > M) \geq \mathbb{P}(J(\delta) > M).$$

Or $J(\delta) \uparrow +\infty$ presque sûrement lorsque $\delta \downarrow 0$. Le membre de droite tend donc vers 1.

* * *

Corrigé du problème 5.11. [énoncé p. 180]

- (1) Si $\sigma = 0$, alors $Z = 0$ presque sûrement et l'identité est immédiate. Supposons donc $\sigma > 0$. On a

$$\mathbb{E}[Z^4] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x^4 e^{-x^2/(2\sigma^2)} dx.$$

En intégrant par parties, il vient

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x^3 \cdot x e^{-\frac{x^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} 3x^2 \cdot \sigma^2 e^{-\frac{x^2}{2\sigma^2}} dx = 3\sigma^4.$$

- (2) Soit $H_n \rightarrow H$ avec $H \in [1/2, 1]$. On montre que $\mathbb{P}_{H_n}$ converge étroitement vers $\mathbb{P}_H$. On suppose que W^n a pour loi $\mathbb{P}_{H_n}$.

Convergence des marginales fini-dimensionnelles. Soient $0 \leq t_1 < t_2 < \dots < t_k \leq 1$. Pour démontrer la convergence en loi de $((W_{t_1}^n, \dots, W_{t_k}^n))_{n \geq 1}$, vu la forme des fonctions caractéristiques des vecteurs gaussiens, il suffit de démontrer que les matrices de covariance convergent d'après le théorème de Lévy. On a

$$\mathbb{E}[W_{t_i}^n W_{t_j}^n] = \frac{1}{2} (t_i^{2H_n} + t_j^{2H_n} - |t_i - t_j|^{2H_n}) \xrightarrow{n \rightarrow \infty} \frac{1}{2} (t_i^{2H} + t_j^{2H} - |t_i - t_j|^{2H}),$$

de sorte que les marginales fini-dimensionnelles de $\mathbb{P}_{H_n}$ convergent vers celles de $\mathbb{P}_H$.

Tension. Les accroissements sont gaussiens centrés et

$$\mathbb{E}[(W_s^n - W_t^n)^2] = |s - t|^{2H_n}.$$

D'après (1), pour tous n et $s, t \in [0, 1]$,

$$\mathbb{E}[(W_s^n - W_t^n)^4] = 3|s - t|^{4H_n} \leq 3|s - t|^2,$$

car $H_n \geq 1/2$. Comme $W_0^n = 0$ presque sûrement, le critère de Kolmogorov donne la tension, avec les mêmes constantes pour tous $H_n \in [1/2, 1]$. La convergence des marginales fini-dimensionnelles conclut.

- (3) Soit N une variable aléatoire gaussienne centrée réduite. On pose $W_s = sN$ pour $0 \leq s \leq 1$. Ainsi $(W_s)_{0 \leq s \leq 1}$ est un vecteur gaussien de covariance $\mathbb{E}[W_s W_t] = st$, dont la loi est donc $\mathbb{P}_1$.
- (4) Il suffit de vérifier le critère de tension de Kolmogorov. On montre qu'il existe $C > 0$ et $\eta > 0$ tels que pour n assez grand, pour tout $s, t \in [0, 1]$:

$$\mathbb{E} \left[|Z_n(s) - Z_n(t)|^2 \right] \leq C|s - t|^{1+\eta}.$$

Choisissons $C' > 0$ tel que $\mathbb{E} [S_n^2]^{1/2} \leq C'n^H$ pour tout $n \geq 1$. Soient $s \leq t \in [0, 1]$.

Premier cas : $\lfloor ns \rfloor = \lfloor nt \rfloor = k$. Alors $S_{ns} - S_{nt} = n(s - t)(S_{k+1} - S_k)$. Comme $0 \leq n(t - s) \leq 1$, on obtient

$$\mathbb{E} \left[|Z_n(s) - Z_n(t)|^2 \right] = n^{2-2H}(t-s)^2 \mathbb{E} [S_1^2] \leq \mathbb{E} [S_1^2] (t-s)^{2H}.$$

La constante est indépendante de n, s, t .

Deuxième cas : $\lfloor ns \rfloor < \lfloor nt \rfloor$. Posons $k = \lfloor ns \rfloor$ et $\ell = \lfloor nt \rfloor$. Alors en utilisant le fait

$$|S_{nt} - S_{ns}| \leq |S_{nt} - S_\ell| + |S_\ell - S_{k+1}| + |S_{k+1} - S_{ns}|,$$

le premier cas et l'inégalité $x \leq x^H$ pour $x \in [0, 1]$, on a

$$\begin{aligned} & \mathbb{E} \left[(S_{ns} - S_{nt})^2 \right]^{1/2} \\ & \leq \mathbb{E} \left[(S_{nt} - S_\ell)^2 \right]^{1/2} + \mathbb{E} \left[(S_\ell - S_{k+1})^2 \right]^{1/2} + \mathbb{E} \left[(S_{k+1} - S_{ns})^2 \right]^{1/2} \\ & \leq C(nt - \ell) + C'(\ell - (k+1))^H + C(k+1 - ns) \\ & \leq C(nt - \ell)^H + C'(\ell - (k+1))^H + C(k+1 - ns)^H \\ & \leq C(nt - ns)^H + C'(nt - ns)^H + C(nt - ns)^H \\ & \leq (2C + C')n^H(t-s)^H, \end{aligned}$$

ce qui implique

$$\mathbb{E} \left[|Z_n(s) - Z_n(t)|^2 \right] \leq C''(t-s)^{2H}.$$

Comme $2H > 1$, le critère de tension de Kolmogorov est vérifié.

- (5) Les hypothèses (a) et (b) de la question (4) sont vérifiées; sa preuve donne donc la tension de (Z_n) . Nous allons établir la convergence des lois fini-dimensionnelles sans supposer préalablement l'existence de W . Pour cela, il suffit de faire converger les matrices de covariance des vecteurs gaussiens correspondants. Fixons $0 \leq s < t \leq 1$ et utilisons $xy = (x^2 + y^2 - (x - y)^2)/2$:

$$\begin{aligned} & \frac{1}{n^{2H}} \mathbb{E} [S_{\lfloor ns \rfloor} S_{\lfloor nt \rfloor}] \\ & = \frac{1}{2n^{2H}} \mathbb{E} [S_{\lfloor ns \rfloor}^2] + \frac{1}{2n^{2H}} \mathbb{E} [S_{\lfloor nt \rfloor}^2] - \frac{1}{2n^{2H}} \mathbb{E} [(S_{\lfloor nt \rfloor} - S_{\lfloor ns \rfloor})^2] \\ & = \frac{1}{2n^{2H}} \mathbb{E} [S_{\lfloor ns \rfloor}^2] + \frac{1}{2n^{2H}} \mathbb{E} [S_{\lfloor nt \rfloor}^2] - \frac{1}{2n^{2H}} \mathbb{E} [S_{\lfloor nt \rfloor - \lfloor ns \rfloor}^2] \\ & \rightarrow \frac{c}{2} (s^{2H} + t^{2H} - |s - t|^{2H}). \end{aligned}$$

Pour conclure, pour tout $t \in [0, 1]$,

$$\frac{|S_{nt} - S_{[nt]}|}{n^H} \leq \frac{|S_{[nt]+1} - S_{[nt]}|}{n^H} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0.$$

En effet, l'accroissement du numérateur a la même loi que S_1 (également lorsque $[nt] = 0$). Le lemme de Slutsky transfère donc ces limites gaussiennes aux processus interpolés. Les matrices de covariance limites sont positives au sens large, comme limites de matrices de covariance. La tension et la convergence de toutes les marginales fini-dimensionnelles donnent alors une limite en loi Z à valeurs dans $\mathcal{C}$. En posant $W = Z/\sqrt{c}$, on construit un processus continu, gaussien centré, de covariance

$$\mathbb{E}[W_s W_t] = \frac{1}{2}(s^{2H} + t^{2H} - |s - t|^{2H}).$$

Sa loi est donc $\mathbb{P}_H$. L'application de ce résultat en (9) établira son existence pour tout $H \in]1/2, 1[$.

(6) On a

$$\mathbb{E}[Y_i Y_{i+n}] = \mathbb{E}[(-1)^{X_{i+1} + \dots + X_{i+n}}] = \mathbb{E}[(-1)^{X_2}]^n = (p - (1-p))^n = (2p-1)^n.$$

(7) C'est une conséquence du théorème de Donsker : dans le cas $p = 1/2$, $(Y_n)_{n \geq 1}$ sont i.i.d. avec $\mathbb{P}(Y_1 = 1) = \mathbb{P}(Y_1 = -1) = 1/2$.

(8) Pour $m \geq 1$ fixé, les vecteurs

$$(Y_2^j, \dots, Y_{m+1}^j), \quad j \geq 1,$$

sont i.i.d., centrés et bornés. Le théorème central limite multidimensionnel donne donc

$$\left(\frac{1}{\sqrt{k}} \sum_{j=1}^k Y_{i+1}^j : 1 \leq i \leq m \right) \Rightarrow \mathcal{N}(0, K_m),$$

où

$$(K_m)_{i\ell} = \mathbb{E}[Y_{i+1}^1 Y_{\ell+1}^1] = \mathbb{E}[(2p-1)^{|i-\ell|}].$$

Ces lois gaussiennes sont cohérentes lorsque m varie ; le théorème d'extension de Kolmogorov fournit un processus gaussien centré $(G_i)_{i \geq 1}$ ayant ces marginales. On a $\mathbb{E}[G_i^2] = 1$ et, pour $n \geq 1$,

$$r(n) = \mathbb{E}[G_i G_{i+n}] = \mathbb{E}[(2p-1)^n] = (1-H)2^{3-2H} \int_{1/2}^1 (2u-1)^n (1-u)^{1-2H} du.$$

L'équivalent admis dans l'énoncé donne le résultat.

- (9) Posons $r(0) = 1$. Le processus (G_i) est gaussien centré et sa covariance $r(|i - j|)$ ne dépend que du décalage. Ainsi $V_{n+i} - V_n$ a la même loi que V_i . De plus,

$$\mathbb{E} [V_n^2] = \sum_{i,j=1}^n r(|i - j|) = n + 2 \sum_{k=1}^{n-1} (n - k)r(k).$$

D'après (8), $r(k) \sim c_H^{-2} H(2H - 1)k^{2H-2}$. Comme $2H - 2 > -1$, une comparaison série-intégrale donne

$$\sum_{k=1}^{n-1} (n - k)k^{2H-2} \sim \frac{n^{2H}}{(2H - 1)2H}.$$

Par conséquent,

$$\mathbb{E} [V_n^2] \sim c_H^{-2} n^{2H},$$

le terme n étant négligeable puisque $2H > 1$. Les hypothèses de (5) sont donc satisfaites avec $c = c_H^{-2}$, ce qui donne la convergence de $c_H V_{nt}/n^H$ vers W_t .

Remarque. Le processus W s'appelle le *mouvement brownien fractionnaire*; ce problème est inspiré de l'article de Nathanaël Enriquez "A simple construction of the fractional Brownian motion." *Stochastic Processes and their Applications* 109.2 (2004) : 203-223.

6 L'espace des fonctions càdlàg

Cours associé : [chapitre 7](#), p. 69.

6.1 Énoncés

Exercice 6.1. [corrigé p. 198] On note $\mathbb{D} = \mathbb{D}([0, 1], \mathbb{R})$ l'ensemble des fonctions càdlàg sur $[0, 1]$ à valeurs réelles muni de la topologie J_1 de Skorokhod définie en cours. Pour une variable aléatoire Z , on note F_Z sa fonction de répartition définie par $F_Z(t) = \mathbb{P}(Z \leq t)$ pour $t \in \mathbb{R}$, vue comme élément de $\mathbb{D}$ par restriction à $[0, 1]$.

Soient $(X_n)_{n \geq 1}$, X des variables aléatoires réelles à valeurs dans $[\frac{1}{2}, 1]$.

- (1) Alix dit : "si X_n converge en loi, alors F_{X_n} converge dans $\mathbb{D}$." A-t-il raison ? Justifiez votre réponse.
- (2) Billie dit : "si F_{X_n} converge dans $\mathbb{D}$, alors X_n converge en loi." A-t-elle raison ? Justifiez votre réponse.

* * *

Problème 6.2. [corrigé p. 199]

Les trois parties de ce problème sont indépendantes, sauf la question (6) qui utilise des résultats de parties précédentes.

Notations et rappels. Pour un espace vectoriel E et $A \subset E$, on note $\text{co}(A)$ l'enveloppe convexe de A , définie par

$$\text{co}(A) = \left\{ \sum_{i=1}^n \lambda_i a_i : n \geq 1, a_i \in A \text{ et } \lambda_i \in [0, 1] \text{ pour } 1 \leq i \leq n, \sum_{i=1}^n \lambda_i = 1 \right\}.$$

On note $\mathbb{D} = \mathbb{D}([0, 1], \mathbb{R})$ l'espace des fonctions càdlàg de $[0, 1]$ dans $\mathbb{R}$ muni de la topologie J_1 de Skorokhod. Pour $x \in \mathbb{D}$ et $0 < t \leq 1$, on note $\Delta x(t) = x(t) - x(t-)$ et $\Delta x(0) = 0$. Pour $0 < \delta < 1$ et $x \in \mathbb{D}$ on note

$$\omega'(x, \delta) = \inf \max_{1 \leq i \leq n} \sup_{s, t \in [t_{i-1}, t_i[} |x(s) - x(t)|,$$

où l'infimum est pris sur toutes les partitions $0 = t_0 < t_1 < \dots < t_{n-1} < t_n = 1$ de $[0, 1]$ telles que $\min_{1 \leq i \leq n} (t_i - t_{i-1}) > \delta$.

Pour un espace métrique E et $A \subset E$, on rappelle que A est dit précompact si pour tout $\varepsilon > 0$, A peut être recouvert par un nombre fini de boules de rayon ε , et que A est dit relativement compact si $\overline{A}$ est compact. Si E est complet, on rappelle qu'une partie est précompacte si et seulement si elle est relativement compacte.

On rappelle qu'une famille $\mathcal{F} \subset \mathbb{D}$ est relativement compacte si et seulement si les deux conditions suivantes sont vérifiées :

- $\sup_{x \in \mathcal{F}} \|x\|_\infty < \infty$
- $\forall \varepsilon > 0, \exists \delta > 0, \sup_{x \in \mathcal{F}} \omega'(x, \delta) \leq \varepsilon$.

Première partie. Soit $(E, \|\cdot\|)$ un espace vectoriel normé complet séparable.

- (1) Si $K \subset E$ est compact, montrer que l'enveloppe convexe de K , notée $\text{co}(K)$ est précompacte.
- (2) Montrer que toute mesure de probabilité sur E est convexe-tendue, au sens où pour tout $\varepsilon > 0$, il existe un compact convexe K tel que $\mu(K) \geq 1 - \varepsilon$.

Deuxième partie.

- (3) On considère

$$\mathcal{F}_0 = \left\{ \mathbb{1}_{[s,1]} : \frac{1}{4} \leq s \leq \frac{3}{4} \right\}.$$

Justifier que $\mathcal{F}_0$ est relativement compact dans $\mathbb{D}$ et que $\text{co}(\mathcal{F}_0)$ n'est pas relativement compact dans $\mathbb{D}$.

Pour $\mathcal{F} \subset \mathbb{D}$ et $\varepsilon > 0$, on note

$$S_\varepsilon(\mathcal{F}) = \left\{ t \in [0, 1] : \sup_{x \in \mathcal{F}} |\Delta x(t)| > \varepsilon \right\}.$$

- (4) On suppose que $K \subset \mathbb{D}$ est une partie relativement compacte. Montrer que si $\text{co}(K)$ est relativement compact, alors pour tout $\varepsilon > 0$, $\text{Card}(S_\varepsilon(K)) < \infty$.

Troisième partie. Soit X une variable aléatoire à valeurs dans $\mathbb{D}$. On dit que X est continu en probabilité si pour toute suite $t_n \rightarrow t$, la suite $(X(t_n))_{n \geq 1}$ converge en probabilité vers $X(t)$ lorsque $n \rightarrow \infty$.

- (5) Donner un exemple où $\forall \omega \in \Omega$, X_ω n'est pas continu mais X est continu en probabilité.
- (6) On suppose que X est continu en probabilité et que sa loi est convexe-tendue. Démontrer que X est continu p.s.
- (7) Toute mesure de probabilité sur $\mathbb{D}$ est-elle convexe-tendue? Est-ce contradictoire avec (2)?

Question bonus : On suppose que $K \subset \mathbb{D}$ est une partie relativement compacte. Montrer que si $\text{Card}(S_\varepsilon(K)) < \infty$ pour tout $\varepsilon > 0$ alors $\text{co}(K)$ est relativement compact.

* * *

Problème 6.3. [corrigé p. 201] On note $\mathcal{C}$ l'ensemble des fonctions continues de $[0, 1]$ dans $\mathbb{R}$ muni de la topologie de la convergence uniforme et $\mathbb{D}$ l'ensemble des fonctions càdlàg de $[0, 1]$ dans $\mathbb{R}$ muni de la topologie de Skorokhod introduite en cours. Soit $(U_i)_{i \geq 1}$ des variables aléatoires i.i.d. uniformes sur $[0, 1]$. On pose pour $n \geq 1$ et $t \in [0, 1]$:

$$F_n(t) = \frac{1}{n} \text{Card}(\{i \in \{1, 2, \dots, n\} : U_i \leq t\}), \quad B_n(t) = \sqrt{n}(F_n(t) - t).$$

L'objectif de ce problème est de montrer que B_n converge en loi dans $\mathbb{D}$ vers le pont brownien.

On dit que $(M_1, \dots, M_n)$ suit la loi multinomiale $M(k, p_1, \dots, p_n)$ si $k \geq 1$, $p_i \geq 0$, $\sum_{i=1}^n p_i = 1$, et si, pour tous entiers $m_i \geq 0$ tels que $\sum_{i=1}^n m_i = k$,

$$\mathbb{P}((M_1, \dots, M_n) = (m_1, \dots, m_n)) = \frac{k!}{\prod_{i=1}^n m_i!} \prod_{i=1}^n p_i^{m_i}.$$

Cette probabilité vaut 0 ailleurs; on convient que $0^0 = 1$.

On pose, pour $1 \leq j \leq n$:

$$N_j = \text{Card}(\{i \in \{1, 2, \dots, n\} : U_i \in [(j-1)/n, j/n]\}).$$

(1) Démontrer que $(N_1, \dots, N_n)$ suit une loi multinomiale et identifier ses paramètres.

Soit $(P_i)_{i \geq 1}$ des variables aléatoires i.i.d. de Poisson de paramètre 1. On pose pour $0 \leq k \leq n$:

$$S_k = \sum_{i=1}^k (P_i - 1)$$

et on définit S_t pour $t \geq 0$ par interpolation linéaire. Soient enfin $\hat{B}_n \in \mathcal{C}$ la fonction obtenue en interpolant linéairement B_n aux points $\{k/n, 0 \leq k \leq n\}$ et $\hat{S}_n \in \mathcal{C}$ une variable aléatoire dont la loi est la loi conditionnelle de $(\frac{1}{\sqrt{n}} S_{nt})_{0 \leq t \leq 1}$ sachant que $S_n = 0$.

(2) Démontrer que $\hat{B}_n$ et $\hat{S}_n$ ont même loi.

(3) Justifier que $\hat{B}_n$ converge en loi dans $\mathbb{D}$ vers le pont brownien.

(4) Démontrer que $\sup_{0 \leq t \leq 1} |\hat{B}_n(t) - B_n(t)| \rightarrow 0$ en probabilité.

(5) Conclure.

6.2 Corrigés

Corrigé de l'exercice 6.1. [énoncé p. 196]

(1) Non. Pour $n \geq 4$, prenons $\mathcal{L}(X_n) = \frac{1}{2}\delta_{1/2+1/n} + \frac{1}{2}\delta_{1/2+2/n}$, et posons $X_n = 1/2$ pour $n < 4$. Ces variables prennent bien leurs valeurs dans $[1/2, 1]$ et convergent en loi vers $1/2$. Si F_{X_n} convergeait dans $\mathbb{D}$ vers F , la convergence aux points de continuité, la continuité à droite et la convergence au point terminal 1 imposeraient $F = \mathbb{1}_{[1/2, 1]}$. Pourtant, pour tout changement de temps λ_n , il existe t_n tel que $\lambda_n(t_n) = 1/2 + 1/n$, et

$$|F_{X_n}(\lambda_n(t_n)) - F(t_n)| = \frac{1}{2}.$$

La convergence J_1 est donc impossible.

(2) Oui. Soit F la limite dans $\mathbb{D}$ et choisissons des changements de temps λ_n tels que $F_{X_n} \circ \lambda_n \rightarrow F$ uniformément. Les fonctions $F_{X_n} \circ \lambda_n$ étant croissantes, F est croissante. Elle est càdlàg, et la convergence aux extrémités donne $F(0) = 0$ et $F(1) = 1$. En prolongeant F par 0 sur $] -\infty, 0[$ et par 1 sur $]1, \infty[$, on obtient donc

une fonction de répartition, notée encore F , d'une loi de probabilité ν sur $\mathbb{R}$. La convergence J_1 donne $F_{X_n}(t) \rightarrow F(t)$ en tout point de continuité de F dans $[0, 1]$; hors de cet intervalle, la convergence est immédiate. Le critère de convergence des fonctions de répartition entraîne donc $\mathcal{L}(X_n) \Rightarrow \nu$.

* * *

Corrigé du problème 6.2. [énoncé p. 196]

- (1) Soit $\varepsilon > 0$. Choisissons $x_1, \dots, x_N \in K$ tels que $K \subset \bigcup_{i=1}^N B(x_i, \varepsilon)$. Toute combinaison convexe d'éléments de K est à distance strictement inférieure à ε de $P = \text{co}(\{x_1, \dots, x_N\})$. Le simplexe des coefficients étant compact, P est compact. Recouvrons donc P par un nombre fini de boules de rayon ε : les mêmes centres et le rayon 2ε recouvrent $\text{co}(K)$. Comme ε est arbitraire, $\text{co}(K)$ est précompact.
- (2) Soit $\varepsilon > 0$. D'après le théorème de Prokhorov, il existe un compact C tel que $\mu(C) \geq 1 - \varepsilon$. On obtient le résultat désiré en prenant $K = \overline{\text{co}(C)}$.
- (3) On a bien $\sup_{x \in \mathcal{F}_0} \|x\|_\infty < \infty$. D'autre part, pour $\varepsilon > 0$, en prenant $\delta < 1/4$, on a $\sup_{x \in \mathcal{F}_0} \omega'(x, \delta) = 0$. Ainsi, $\mathcal{F}_0$ est relativement compact. En revanche, en prenant

$$x_n = \frac{1}{2} \mathbb{1}_{[1/2-1/n, 1]} + \frac{1}{2} \mathbb{1}_{[1/2+1/n, 1]} \in \text{co}(\mathcal{F}_0), \quad n \geq 4,$$

la suite (x_n) n'a aucune sous-suite convergente, de sorte que $\text{co}(\mathcal{F}_0)$ n'est pas relativement compact dans $\mathbb{D}$.

- (4) Raisonnons par l'absurde. Si $S_\varepsilon(K)$ est infini pour un $\varepsilon > 0$, il existe des temps distincts $u_n \rightarrow u_0$ et des fonctions $x_n \in K$ tels que $|\Delta x_n(u_n)| > \varepsilon$. On peut écarter l'éventuel terme $u_n = 1$; par convention, $u_n \neq 0$. Par relative compacité de K , choisissons $\delta_0 > 0$ tel que

$$\sup_{x \in K} \omega'(x, \delta_0) < \frac{\varepsilon}{2}.$$

Pour chaque n , choisissons une partition à mailles strictement supérieures à δ_0 dont l'oscillation maximale pour x_n est strictement inférieure à $\varepsilon/2$. Le saut en u_n force ce point à être une coupure intérieure de la partition. Les oscillations de x_n sur $[u_n - \delta_0, u_n[$ et $[u_n, u_n + \delta_0[$ sont donc inférieures à $\varepsilon/2$.

Fixons $0 < \delta < \delta_0$ et choisissons $m \neq n$ tels que $0 < |u_m - u_n| < \delta$. Chaque point est alors à l'intérieur d'une maille de la partition choisie pour l'autre fonction, de sorte que

$$|\Delta x_m(u_n)| \leq \frac{\varepsilon}{2}, \quad |\Delta x_n(u_m)| \leq \frac{\varepsilon}{2}.$$

Pour $z = (x_m + x_n)/2 \in \text{co}(K)$, l'inégalité triangulaire inverse donne donc

$$|\Delta z(u_m)| > \frac{\varepsilon}{4}, \quad |\Delta z(u_n)| > \frac{\varepsilon}{4}.$$

Une partition à mailles strictement supérieures à δ ne peut contenir ces deux points parmi ses coupures. Au moins un de ces sauts est à l'intérieur d'une maille, dont l'oscillation est supérieure à $\varepsilon/4$. Ainsi $\omega'(z, \delta) \geq \varepsilon/4$. Pour tout $0 < \delta < \delta_0$, on a construit un tel $z \in \text{co}(K)$, ce qui contredit la relative compacité de $\text{co}(K)$.

- (5) Prenons $\Omega =]0, 1[$ muni de la mesure de Lebesgue et $U(\omega) = \omega$. Posons $X(t) = \mathbb{1}_{\{U \leq t\}}$. Chaque trajectoire a un saut, donc est discontinue. Pour $0 < \varepsilon \leq 1$,

$$\mathbb{P}(|X(t_n) - X(t)| \geq \varepsilon) \leq \mathbb{P}(U \in [\min(t_n, t), \max(t_n, t)]) = |t_n - t| \rightarrow 0.$$

Pour $\varepsilon > 1$, la probabilité est nulle.

- (6) Soit $\eta > 0$ et K un compact convexe tel que $\mathbb{P}(X \in K) \geq 1 - \eta$. Pour $\varepsilon > 0$, l'ensemble $D_\varepsilon = S_\varepsilon(K)$ est fini d'après (4). Sur $\{X \in K\}$, on a $|\Delta X(t)| \leq \varepsilon$ pour tout $t \notin D_\varepsilon$. Pour chaque $t > 0$ fixé, en faisant tendre s vers t par valeurs inférieures, $X(s) \rightarrow X(t-)$ presque sûrement et $X(s) \rightarrow X(t)$ en probabilité. Ainsi $\Delta X(t) = 0$ presque sûrement. Cette propriété vaut simultanément sur l'ensemble fini D_ε . Par conséquent,

$$\mathbb{P}\left(\sup_{0 \leq t \leq 1} |\Delta X(t)| \leq \varepsilon\right) \geq 1 - \eta.$$

Faisons $\varepsilon \downarrow 0$, puis $\eta \downarrow 0$. Une fonction càdlàg sans saut est continue, ce qui conclut.

- (7) Soit X l'exemple de la question (5). D'après les questions (5) et (6), sa loi n'est pas convexe-tendue. Ce n'est pas contradictoire avec (2) car $\mathbb{D}$ muni de la topologie J_1 de Skorokhod n'est pas une topologie d'espace vectoriel normé.

Remarque. Ce problème est inspiré par la Remark 2.5 dans

Basse-O'Connor, Andreas, and Jan Rosiński. "On the uniform convergence of random series in Skorokhod space and representations of cadlag infinitely divisible processes." The Annals of Probability 41.6 (2013) : 4317-4341.

<https://arxiv.org/pdf/1111.1682.pdf>

Question bonus. Fixons $\varepsilon > 0$. La relative compacité de K donne $\delta > 0$ tel que $\sup_{x \in K} \omega'(x, \delta) < \varepsilon$. Choisissons une partition commune $0 = s_0 < \dots < s_q = 1$ contenant tous les points de $S_\varepsilon(K)$ et dont toutes les mailles ont une longueur strictement inférieure à δ . C'est possible puisque $S_\varepsilon(K)$ est fini.

Pour chaque $x \in K$, choisissons aussi une partition à mailles strictement supérieures à δ sur lesquelles les oscillations de x sont inférieures à ε . L'intervalle ouvert $]s_j, s_{j+1}[$ contient au plus une coupure de cette seconde partition. S'il en contient une, le saut correspondant est de taille au plus ε , puisque cette coupure n'appartient pas à $S_\varepsilon(K)$. Les oscillations de part et d'autre sont inférieures à ε , donc

$$\sup_{x \in K} \omega(x, [s_j, s_{j+1}[) \leq 3\varepsilon.$$

La même borne vaut pour toute combinaison convexe d'éléments de K . Pour $z \in \text{co}(K)$, posons

$$Tz = \sum_{j=0}^{q-1} z(s_j) \mathbb{1}_{[s_j, s_{j+1}[} + z(1) \mathbb{1}_{\{1\}}.$$

Alors $\|z - Tz\|_\infty \leq 3\varepsilon$. Comme $\sup_{x \in K} \|x\|_\infty < \infty$, les fonctions Tz , $z \in \text{co}(K)$, forment une partie bornée d'un espace de dimension finie. Elles sont donc précompactes pour la norme uniforme. L'arbitraire de ε montre que $\text{co}(K)$ est lui-même précompact pour cette norme. L'espace $(\mathbb{D}, \|\cdot\|_\infty)$ étant complet, son adhérence uniforme est compacte. La convergence uniforme impliquant la convergence J_1 , on obtient la relative compacité souhaitée.

Voir aussi Peter Z. Daffer et Robert L. Taylor, « Laws of Large Numbers for $D[0, 1]$ », *Annals of Probability* 7(1), 85–95 (1979), théorème 6.

* * *

Corrigé du problème 6.3. [énoncé p. 197]

- (1) Pour des entiers positifs $m_1, \dots, m_n$ tels que $m_1 + \dots + m_n = n$ on a

$$\mathbb{P}((N_1, \dots, N_n) = (m_1, \dots, m_n)) = \binom{n}{m_1, \dots, m_n} \frac{1}{n^n} = \frac{n!}{\prod_{i=1}^n m_i!} \frac{1}{n^n}$$

de sorte que $(N_1, \dots, N_n)$ suit la loi multinomiale $M(n, 1/n, \dots, 1/n)$.

- (2) Il suffit de vérifier que les deux vecteurs suivants ont même loi :

$$(\widehat{B}_n(0), \widehat{B}_n(1/n), \dots, \widehat{B}_n(n/n))$$

et $(\widehat{S}_n(0), \widehat{S}_n(1/n), \dots, \widehat{S}_n(n/n))$

Remarquons que pour $0 \leq k \leq n$ on a

$$\widehat{B}_n(k/n) = \frac{1}{\sqrt{n}} \sum_{i=1}^k (N_i - 1), \quad \widehat{S}_n(k/n) \stackrel{\text{loi}}{=} \frac{1}{\sqrt{n}} \sum_{i=1}^k (P_i - 1) \quad \text{sous } \mathbb{P}(\cdot | S_n = 0).$$

Il suffit donc de vérifier que, sous la probabilité conditionnelle $\mathbb{P}(\cdot | S_n = 0)$, le vecteur $(P_1, \dots, P_n)$ suit la loi multinomiale

$$M(n, 1/n, \dots, 1/n).$$

En vertu du conditionnement, le support de sa loi est bien l'ensemble des entiers positifs $m_1, \dots, m_n$ tels que $m_1 + \dots + m_n = n$. Pour de tels entiers, on a

$$\mathbb{P}((P_1, \dots, P_n) = (m_1, \dots, m_n) | S_n = 0) = \frac{1}{\mathbb{P}(S_n = 0)} \prod_{i=1}^n e^{-1} \frac{1}{m_i!}.$$

Or $\widehat{S}_n + n$ suit une loi de Poisson de paramètre n , de sorte que $\mathbb{P}(S_n = 0) = e^{-n} \frac{n^n}{n!}$. Ainsi

$$\mathbb{P}((P_1, \dots, P_n) = (m_1, \dots, m_n) | S_n = 0) = \frac{n!}{\prod_{i=1}^n m_i!} \frac{1}{n^n},$$

d'où le résultat.

- (3) Puisque $\mathbb{E}[P_1 - 1] = 0$ et $\text{Var}(P_1 - 1) = 1$, d'après le théorème de Donsker conditionné vu en cours $\widehat{S}_n$ converge en loi dans $\mathcal{C}$ vers le pont brownien. Comme $\widehat{B}_n$ et $\widehat{S}_n$ ont même loi, il en est de même de $\widehat{B}_n$. Ensuite, si $f_n, f \in \mathcal{C}$ sont telles que $f_n \rightarrow f$ dans $\mathcal{C}$, alors $f_n \rightarrow f$ dans $\mathbb{D}$ (prendre le changement de temps égal à l'identité). Autrement dit, l'inclusion de $\mathcal{C}$ dans $\mathbb{D}$ est continue. On conclut par stabilité de la convergence en loi par composition par une fonction continue.

(4) *Première solution.* On commence par écrire

$$\sup_{0 \leq t \leq 1} |\widehat{B}_n(t) - B_n(t)| \leq \frac{1}{\sqrt{n}} \max_{1 \leq k \leq n} N_k.$$

En remarquant que les N_i suivent des lois $\text{Bin}(n, 1/n)$, on a alors pour tout $\varepsilon > 0$:

$$\begin{aligned} \mathbb{P} \left(\max_{1 \leq k \leq n} N_k \geq \varepsilon \sqrt{n} \right) &\leq n \mathbb{P} (N_1 \geq \varepsilon \sqrt{n}) \\ &\leq n \mathbb{E} \left[e^{N_1} \right] e^{-\varepsilon \sqrt{n}} \\ &= n \left(1 + \frac{e-1}{n} \right)^n e^{-\varepsilon \sqrt{n}} \sim n e^{e-1} e^{-\varepsilon \sqrt{n}} \end{aligned}$$

qui converge vers 0 lorsque $n \rightarrow \infty$.

Deuxième solution. Si $\frac{i}{n} \leq t < \frac{i+1}{n}$ pour $0 \leq i \leq n-1$, on remarque que

$$B_n \left(\frac{i}{n} \right) - \frac{1}{\sqrt{n}} \leq B_n(t) \leq B_n \left(\frac{i+1}{n} \right) + \frac{1}{\sqrt{n}},$$

$$\min\{B_n(i/n), B_n((i+1)/n)\} \leq \widehat{B}_n(t) \leq \max\{B_n(i/n), B_n((i+1)/n)\}.$$

Comme $\widehat{B}_n(i/n) = B_n(i/n)$ pour $0 \leq i \leq n$ on en déduit que

$$\sup_{0 \leq t \leq 1} |\widehat{B}_n(t) - B_n(t)| \leq \max_{0 \leq i \leq n-1} \left| \widehat{B}_n \left(\frac{i+1}{n} \right) - \widehat{B}_n \left(\frac{i}{n} \right) \right| + \frac{1}{\sqrt{n}}.$$

Or $\widehat{B}_n$ converge en loi dans $\mathcal{C}$ vers le pont brownien. Ainsi pour tout $\varepsilon, \eta > 0$ il existe $\delta > 0$ tel que pour n assez grand

$$\mathbb{P} \left(\omega(\widehat{B}_n, \delta) > \eta \right) \leq \varepsilon.$$

On en déduit que pour $n > 1/\delta$ assez grand on a

$$\max_{0 \leq i \leq n-1} \left| \widehat{B}_n \left(\frac{i+1}{n} \right) - \widehat{B}_n \left(\frac{i}{n} \right) \right| \leq \eta$$

avec probabilité au moins $1 - \varepsilon$. Ceci conclut.

(5) Tout d'abord, en notant d_S la distance de Skorokhod

$$d_S(f, g) = \inf_{\lambda \in \Lambda} \max(\|\lambda - \text{Id}\|_\infty, \|f \circ \lambda - g\|_\infty)$$

avec Λ l'ensemble des changements de temps. En prenant $\lambda = \text{Id}$, on obtient directement

$$d_S(\widehat{B}_n, B_n) \leq \|\widehat{B}_n - B_n\|_\infty \xrightarrow{\mathbb{P}} 0$$

d'après (4).

Ainsi, $\widehat{B}_n$ converge en loi vers le pont brownien dans $\mathbb{D}$ et $d_S(\widehat{B}_n, B_n)$ converge en probabilité vers 0. On en déduit que B_n converge en loi vers le pont brownien.

Remarque. Ce problème est inspiré de l'article suivant :

Jean-François Marckert. "One more approach to the convergence of the empirical process to the Brownian bridge." Electronic Journal of Statistics, 2 118-126 2008

<https://arxiv.org/pdf/0710.3296.pdf>

7 Mesures aléatoires de Poisson

Cours associé : [chapitre 8](#), p. 82.

7.1 Énoncés

Exercice 7.1. [corrigé p. 205] Soit M un nuage poissonien d'intensité $dx dy$ sur $\mathbb{R}^2$. On note M_θ et M_R les nuages de Poisson sur $\mathbb{R}/2\pi\mathbb{Z}$ et sur $\mathbb{R}_+$ obtenus à partir de M en considérant les angles et les distances à l'origine des points de M . Donner les intensités de ces deux nuages. Sont-ils indépendants?

* * *

Exercice 7.2. [corrigé p. 206] Soit M un nuage poissonien sur $\mathbb{R}_+$ d'intensité $\lambda(t)dt$, avec $\lambda : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ une fonction continue telle que $\int_0^\infty \lambda(t)dt = \infty$.

- (1) Justifier qu'on peut écrire M sous la forme $M = \sum_{k=1}^\infty \delta_{T_k}$ où $0 < T_1 < T_2 < \dots$ sont des variables aléatoires.
- (2) On pose $a(t) = \int_0^t \lambda(s)ds$. Montrer que $M' = \sum_{k \geq 1} \delta_{a(T_k)}$ est un nuage poissonien sur $\mathbb{R}_+$ d'intensité dx .

On suppose à partir de maintenant que $\lambda(t) = \frac{t}{1+t}$ pour $t \geq 0$.

- (3) Calculer la loi de T_1 et celle de (T_1, T_2) .
- (4) Montrer que presque sûrement

$$\sum_{n \geq 1} \frac{1}{T_n^2} < \infty \quad \text{et} \quad \sum_{n \geq 1} \frac{1}{T_n} = \infty.$$

- (5) Montrer que $T_{n+1} - T_n$ converge en loi quand $n \rightarrow \infty$ et préciser la loi limite.

* * *

Exercice 7.3. [corrigé p. 207] Soit μ une mesure de probabilité sur $\mathbb{R}$ et N un nuage poissonien d'intensité $dx \otimes \mu$ sur $\mathbb{R} \times \mathbb{R}$. On note $N = \sum_{n \in \mathbb{Z}} \delta_{(X_n, V_n)}$ avec

$$\dots < X_{-k} < \dots < X_0 < 0 < X_1 < \dots < X_\ell < \dots,$$

et on voit N comme une distribution de particules sur la droite réelle, chaque particule n étant initialement en position X_n avec vitesse V_n .

- (a) On pose $N_t = \sum_{n \in \mathbb{Z}} \delta_{(X_n + tV_n, V_n)}$. Montrer que N_t est encore un nuage poissonien d'intensité $dx \otimes \mu$.
- (b) On suppose $\int_{\mathbb{R}} |v| \mu(dv) < \infty$. Pour chaque particule de vitesse $V_n \neq 0$, soit

$$T_n = -\frac{X_n}{V_n}$$

son temps d'atteinte de 0, éventuellement négatif. Les particules de vitesse nulle ne contribuent pas aux arrivées. Montrer que

$$\sum_{n \in \mathbb{Z}: V_n \neq 0} \delta_{(T_n, V_n)}$$

est une mesure aléatoire de Poisson sur $\mathbb{R} \times \mathbb{R}$ et déterminer son intensité.

- (c) Calculer $\mathbb{P}(\forall s \in [0, t], \forall n \in \mathbb{Z}, X_n + sV_n \notin [a, b])$ où $[a, b]$ est un intervalle fixé et déterminer la loi de $\inf\{t \geq 0 : \exists s \in [0, t], \exists n \in \mathbb{Z}, X_n + sV_n \in [a, b]\}$.

* * *

Exercice 7.4. [corrigé p. 208] Soit M un nuage poissonien sur $\mathbb{R}^2$ d'intensité $\lambda dx dy$ pour $\lambda > 0$, écrit sous la forme $M = \sum_{i \geq 1} \delta_{X_i}$, où l'énumération $(X_i)_{i \geq 1}$ est une fonction mesurable de M fixée à l'avance.

- (1) Calculer la loi de $\inf_{i \geq 1} |X_i|$.
- (2) Soit $(Y_n)_{n \geq 1}$ des variables aléatoires i.i.d. à valeurs dans $\mathbb{R}^2$ indépendantes de M . Montrer que $M' = \sum_{i \geq 1} \delta_{X_i + Y_i}$ a la même loi que M .
- (3) Soit maintenant $(R_n)_{n \geq 1}$ des variables aléatoires i.i.d. réelles positives indépendantes de M . On considère l'ensemble aléatoire

$$A = \bigcup_{i \geq 1} B(X_i, R_i),$$

où $B(x, r)$ est la boule de centre x et de rayon r dans $\mathbb{R}^2$.

- (a) Calculer la loi du nombre de boules de A qui recouvrent un point donné $x \in \mathbb{R}^2$.
- (b) Montrer que si $\mathbb{E}[R_1^2] = \infty$ alors pour tout $x \in \mathbb{R}^2$ on a presque sûrement $x \in A$.
- (c) En déduire que si $\mathbb{E}[R_1^2] = \infty$ alors $A = \mathbb{R}^2$ presque sûrement.
- (d) Montrer que si $\mathbb{E}[R_1^2] < \infty$ alors $A \neq \mathbb{R}^2$ presque sûrement.
- (4) On suppose que $R_i = r > 0$. On note $\mathcal{C}(0)$ l'amas de l'origine, c'est-à-dire les points accessibles depuis l'origine par un chemin qui reste dans A .

- (a) On note $\mathbb{P}_{\lambda, r}$ la loi de A avec intensité $\lambda > 0$ et rayon $r > 0$. Montrer que

$$\mathbb{P}_{\lambda, r}(\mathcal{C}(0) \text{ non borné}) = \mathbb{P}_{1, \sqrt{\lambda}r}(\mathcal{C}(0) \text{ non borné}) = \mathbb{P}_{\lambda r^2, 1}(\mathcal{C}(0) \text{ non borné}).$$

- (b) Montrer que $\mathbb{P}_{\lambda, r}(\mathcal{C}(0) \text{ non borné})$ est croissante en λ et en r .

* * *

Exercice 7.5. [corrigé p. 210] (**Théorème de Rényi**) Soit M une variable aléatoire à valeurs dans $M_{\text{atom}}^s(\mathbb{R}_+)$ (sous-ensemble des mesures $\mu \in M_{\text{atom}}(\mathbb{R}_+)$ dont les atomes sont deux à deux distincts), presque sûrement finie sur les compacts. On suppose que pour tout borélien A de $\mathbb{R}_+$ on a

$$\mathbb{P}(M(A) = 0) = e^{-\lambda(A)},$$

où λ est la mesure de Lebesgue sur $\mathbb{R}_+$.

- (1) Soient $I_1, \dots, I_m$ des intervalles deux à deux disjoints. Montrer que les événements $\{M(I_j) = 0 : 1 \leq j \leq m\}$ sont indépendants.

- (2) Pour tout $n, k \geq 0$ on pose $D_{n,k} = [k2^{-n}, (k+1)2^{-n}[$. Soit J un sous-intervalle borné de $\mathbb{R}_+$, d'intérieur non vide. Pour $n \geq 0$, on note $K_n = \text{Card}\{k \geq 0 : D_{n,k} \subseteq J\}$ et

$$M_n(J) = \sum_{k \geq 0 : D_{n,k} \subseteq J} \mathbb{1}_{\{M(D_{n,k}) \neq 0\}}.$$

Déterminer la loi de $M_n(J)$ en termes de n et de K_n .

- (3) Montrer que l'on a $M_n(J) \uparrow M(J)$ presque-sûrement lorsque $n \rightarrow \infty$ (on considérera les atomes de M dans J et on vérifiera dans un premier temps qu'ils sont tous distincts de $\inf J$ et $\sup J$) et en déduire la loi de $M(J)$.
- (4) Soient $J_1, \dots, J_k$ des sous-intervalles bornés deux à deux disjoints de $\mathbb{R}_+$. Montrer que les variables aléatoires $M(J_1), \dots, M(J_k)$ sont indépendantes.
- (5) On note $N_t = M([0, t])$, $t \geq 0$, le processus de comptage associé à M . Montrer que N est un processus de Poisson d'intensité 1. En déduire que M est une mesure de Poisson d'intensité $\lambda(dx)$.

* * *

Exercice 7.6. [corrigé p. 211] Trouver une tribu $\mathcal{E}$ sur $[0, 1]$ et une mesure de probabilité μ sur $([0, 1], \mathcal{E})$ telle que $\mu(B) \in \{0, 1\}$ pour tout $B \in \mathcal{E}$ mais telle que μ ne soit pas une masse de Dirac.

7.2 Corrigés

Corrigé de l'exercice 7.1. [énoncé p. 203] Le point 0 n'appartient presque sûrement pas au nuage. Par stabilité par image et passage en coordonnées polaires, les intensités cherchées sont

$$\mu_R(dr) = 2\pi r dr, \quad \mu_\theta(B) = \int_B \int_0^\infty r dr d\theta = \begin{cases} 0 & \text{si } \lambda(B) = 0, \\ \infty & \text{si } \lambda(B) > 0, \end{cases}$$

où λ désigne la mesure de longueur sur $\mathbb{R}/2\pi\mathbb{Z}$. La mesure μ_θ est s-finie, puisqu'elle est égale à $\sum_{j \geq 1} \lambda$, mais elle n'est pas σ -finie.

Les deux nuages sont indépendants. En effet, pour chaque borélien B fixé, $M_\theta(B)$ est presque sûrement égal à 0 ou à ∞ , suivant que $\lambda(B) = 0$ ou $\lambda(B) > 0$. Tous les événements qui engendrent $\sigma(M_\theta)$ ont donc une probabilité égale à 0 ou à 1. Comme les événements de probabilité 0 ou 1 forment une tribu, il en est de même de tout événement de $\sigma(M_\theta)$. Une telle tribu est indépendante de toute autre tribu, en particulier de $\sigma(M_R)$. On utilise ici la tribu des mesures aléatoires définie dans le cours, engendrée par les applications d'évaluation.

Corrigé de l'exercice 7.2. [énoncé p. 203]

- (1) L'intensité est sans atome et finie sur les compacts. Le nuage est donc presque sûrement simple, fini sur tout compact et sans point en 0. De plus, $M(\mathbb{R}_+) = \infty$ presque sûrement, puisque son intensité totale est infinie. Ses atomes peuvent ainsi être rangés dans l'ordre croissant. Les variables

$$T_k = \inf\{t \geq 0 : M([0, t]) \geq k\}$$

sont mesurables, car $\{T_k \leq t\} = \{M([0, t]) \geq k\}$, et vérifient presque sûrement $0 < T_1 < T_2 < \dots$ et $T_k \rightarrow \infty$.

- (2) Par stabilité par image, il suffit de montrer que l'image de $\lambda(t)dt$ par a est la mesure de Lebesgue. L'application a est continue, croissante, $a(0) = 0$ et $a(t) \rightarrow \infty$. Pour $x \geq 0$, posons

$$b(x) = \sup\{t \geq 0 : a(t) \leq x\}.$$

Alors $b(x) < \infty$, $a(b(x)) = x$ et $a^{-1}([0, x]) = [0, b(x)]$. Par conséquent,

$$\int_{a^{-1}([0, x])} \lambda(t) dt = a(b(x)) = x.$$

Ces égalités caractérisent la mesure de Lebesgue. Cet argument reste valable lorsque a n'est pas strictement croissante.

- (3) On a $a(t) = t - \log(1 + t)$. Pour $t \geq 0$,

$$\mathbb{P}(T_1 > t) = \mathbb{P}(M([0, t]) = 0) = e^{-a(t)} = (1 + t)e^{-t}.$$

Ainsi T_1 a pour densité $te^{-t}\mathbb{1}_{\{t>0\}}$. D'après (2), les variables $S_k = a(T_k)$ sont les temps successifs d'un processus de Poisson d'intensité 1. En particulier, (S_1, S_2) a pour densité $e^{-v}\mathbb{1}_{\{0 < u < v\}}$ en (u, v) . Le changement de variables $(u, v) = (a(s), a(t))$ donne la densité de (T_1, T_2) :

$$\lambda(s)\lambda(t)e^{-a(t)}\mathbb{1}_{\{0 < s < t\}} = \frac{st}{1+s}e^{-t}\mathbb{1}_{\{0 < s < t\}}.$$

- (4) Posons $S_0 = 0$ et $E_k = S_k - S_{k-1}$ pour $k \geq 1$. Les variables (E_k) sont i.i.d. de loi exponentielle de paramètre 1. La loi forte des grands nombres donne $S_n/n \rightarrow 1$ presque sûrement. Comme $T_n \rightarrow \infty$ et $a(t)/t \rightarrow 1$, on en déduit

$$\frac{T_n}{n} = \frac{T_n}{a(T_n)} \frac{S_n}{n} \rightarrow 1 \quad \text{presque sûrement.}$$

Les deux assertions résultent alors de la comparaison avec $\sum_n n^{-2}$ et $\sum_n n^{-1}$.

- (5) Posons $G_n = T_{n+1} - T_n$ et $E_{n+1} = a(T_{n+1}) - a(T_n)$, de loi $\text{Exp}(1)$. Comme $a'(t) = t/(1+t)$ est croissante et majorée par 1,

$$\frac{T_n}{1+T_n} G_n \leq E_{n+1} \leq G_n.$$

Ainsi

$$0 \leq G_n - E_{n+1} \leq \frac{E_{n+1}}{T_n} \xrightarrow{\mathbb{P}} 0,$$

puisque (E_{n+1}) est tendue et $T_n \rightarrow \infty$ presque sûrement. Le lemme de Slutsky donne donc $T_{n+1} - T_n \xrightarrow{\text{loi}} \text{Exp}(1)$.

Corrigé de l'exercice 7.3. [énoncé p. 203]

- (a) Le nuage N_t est l'image de N par $(x, v) \mapsto (x + tv, v)$. Pour toute fonction mesurable positive f , le changement de variables $y = x + tv$ donne

$$\int_{\mathbb{R}} \int_{\mathbb{R}} f(x + tv, v) dx \mu(dv) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(y, v) dy \mu(dv).$$

L'intensité est donc inchangée, et la stabilité par image conclut.

- (b) On restreint d'abord N à $\mathbb{R} \times (\mathbb{R} \setminus \{0\})$, puis on applique la transformation $(x, v) \mapsto (-x/v, v)$. La stabilité par restriction et par image montre que la mesure obtenue est poissonnienne. Le changement de variables $s = -x/v$ donne, pour $f \geq 0$ mesurable,

$$\int_{\mathbb{R} \setminus \{0\}} \int_{\mathbb{R}} f(-x/v, v) dx \mu(dv) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(s, v) |v| ds \mu(dv).$$

L'intensité cherchée est donc $ds \otimes (|v| \mu(dv))$. En particulier, les temps d'arrivée forment un nuage de Poisson d'intensité $m_1 ds$, avec $m_1 = \int_{\mathbb{R}} |v| \mu(dv)$.

- (c) Posons $\ell = b - a$ et

$$A_t = \{(x, v) : \text{il existe } s \in [0, t] \text{ tel que } x + sv \in [a, b]\}.$$

Pour une vitesse $v \geq 0$, les positions initiales correspondantes forment l'intervalle $[a - tv, b]$; pour $v < 0$, elles forment $[a, b - tv]$. Dans les deux cas, sa longueur est $\ell + t|v|$. Ainsi, pour $t > 0$,

$$\mathbb{P}(\forall s \in [0, t], \forall n \in \mathbb{Z}, X_n + sV_n \notin [a, b]) = \mathbb{P}(N(A_t) = 0) = e^{-\ell - tm_1}.$$

Pour $t = 0$, cette probabilité vaut $e^{-\ell}$.

Notons H le temps défini dans l'énoncé, avec la convention $\inf \emptyset = \infty$. Si $m_1 < \infty$, il n'y a qu'un nombre fini de particules visitant $[a, b]$ pendant un intervalle de temps borné. Le premier temps de visite est donc atteint dès qu'il est fini, et

$$\mathbb{P}(H > t) = e^{-\ell - tm_1}, \quad t \geq 0.$$

Pour $0 < m_1 < \infty$, on obtient

$$\mathcal{L}(H) = (1 - e^{-\ell})\delta_0 + e^{-\ell} \text{Exp}(m_1).$$

Si $m_1 = 0$, toutes les particules sont immobiles, et la loi sur $[0, \infty]$ est $(1 - e^{-\ell})\delta_0 + e^{-\ell}\delta_{\infty}$.

Si l'hypothèse d'intégrabilité de (b) n'est pas conservée et $m_1 = \infty$, la probabilité d'absence de visite vaut 0 pour tout $t > 0$. En appliquant ceci aux temps $1/k$, on obtient $H = 0$ presque sûrement, même si l'intervalle peut être vide de particules à l'instant 0.

Corrigé de l'exercice 7.4. [énoncé p. 204]

- (1) Posons $D = \inf_{i \geq 1} |X_i|$. Le nuage est presque sûrement fini sur les compacts. Ainsi, pour $t \geq 0$,

$$\mathbb{P}(D > t) = \mathbb{P}(M(\overline{B}(0, t)) = 0) = e^{-\lambda \pi t^2}.$$

En particulier, D^2 suit la loi exponentielle de paramètre $\lambda\pi$, et D a pour densité $2\lambda\pi t e^{-\lambda\pi t^2} \mathbb{1}_{\{t>0\}}$.

- (2) Notons ν la loi de Y_1 . Puisque l'énumération $(X_i)_{i \geq 1}$ est une fonction mesurable de M , elle est indépendante de $(Y_i)_{i \geq 1}$. Par stabilité par marquage, $\sum_{i \geq 1} \delta_{(X_i, Y_i)}$ est donc un nuage poissonien d'intensité $\lambda dz \otimes \nu(dy)$. Son image par $(z, y) \mapsto z + y$ est M' . Pour tout borélien $C \subset \mathbb{R}^2$, l'intensité de cette image vaut, par Fubini et invariance par translation de la mesure de Lebesgue,

$$\lambda \int_{\mathbb{R}^2} \nu(dy) \int_{\mathbb{R}^2} \mathbb{1}_{\{z+y \in C\}} dz = \lambda \text{Aire}(C).$$

La stabilité par image et l'unicité en loi des nuages poissoniens donnent le résultat.

- (3) Notons maintenant ρ la loi de R_1 . De la même façon, $\widetilde{M} = \sum_{i \geq 1} \delta_{(X_i, R_i)}$ est un nuage poissonien d'intensité $\mu(dz, dr) = \lambda dz \rho(dr)$.

- (a) Pour $x \in \mathbb{R}^2$, posons $H_x = \{(z, r) : x \in B(z, r)\}$. Le nombre de boules contenant x est $\widetilde{M}(H_x)$, donc suit une loi de Poisson de paramètre

$$\mu(H_x) = \lambda \int_0^\infty \pi r^2 \rho(dr) = \lambda \pi \mathbb{E}[R_1^2].$$

Cette formule vaut aussi pour un paramètre infini, auquel cas ce nombre est presque sûrement infini. Elle ne dépend pas du choix de boules ouvertes ou fermées, leurs frontières étant de mesure de Lebesgue nulle.

- (b) Si $\mathbb{E}[R_1^2] = \infty$, pour tout x fixé, $\widetilde{M}(H_x) = \infty$ presque sûrement. En particulier, $\mathbb{P}(x \in A) = 1$.
- (c) Il faut ici obtenir un événement de probabilité 1 sur lequel tous les points sont recouverts simultanément. Fixons un entier $m \geq 1$ et considérons les boules telles que $R_i > |X_i| + m$. Chacune contient $\overline{B}(0, m)$. Leur nombre suit une loi de Poisson de paramètre

$$\lambda \int_0^\infty \pi(r-m)_+^2 \rho(dr) = \infty.$$

En effet, $(r-m)_+^2 \geq r^2/4$ pour $r \geq 2m$, et $\mathbb{E}[R_1^2 \mathbb{1}_{\{R_1 \geq 2m\}}] = \infty$. Ainsi $\overline{B}(0, m) \subset A$ presque sûrement. En intersectant ces événements pour $m \geq 1$, on obtient $A = \mathbb{R}^2$ presque sûrement.

- (d) Supposons $a = \lambda \pi \mathbb{E}[R_1^2] < \infty$ et posons $p = e^{-a} > 0$. D'après (a), un point donné n'est pas recouvert avec probabilité p . Pour en déduire l'existence presque sûre d'un point non recouvert, on contrôle la dépendance entre deux points éloignés.

Pour $h \in \mathbb{R}^2$, posons

$$c(h) = \mu(H_0 \cap H_h) = \lambda \int_0^\infty \text{Aire}(B(0, r) \cap B(h, r)) \rho(dr).$$

Deux boules de rayon r sont disjointes lorsque $2r < |h|$, donc

$$0 \leq c(h) \leq \lambda \pi \mathbb{E} \left[R_1^2 \mathbb{1}_{\{2R_1 \geq |h|\}} \right] \longrightarrow 0 \quad (|h| \rightarrow \infty).$$

La probabilité que deux points x, y ne soient pas recouverts est la probabilité que $\widetilde{M}(H_x \cup H_y) = 0$, soit

$$\mathbb{P}(x \notin A, y \notin A) = e^{-2\alpha + c(y-x)}.$$

Ainsi, en posant $Z_j = \mathbb{1}_{\{(j,0) \notin A\}}$, on a $\mathbb{E}[Z_j] = p$ et

$$\text{Cov}(Z_i, Z_j) = p^2 (e^{c((j-i,0))} - 1).$$

Ces covariances sont bornées et tendent vers 0 lorsque $|i - j| \rightarrow \infty$. Par conséquent,

$$\text{Var} \left(\frac{1}{n} \sum_{j=1}^n Z_j \right) = \frac{p(1-p)}{n} + \frac{2p^2}{n^2} \sum_{h=1}^{n-1} (n-h) (e^{c((h,0))} - 1) \longrightarrow 0.$$

Pour la dernière convergence, il suffit de séparer la somme en $h \leq H$ et $h > H$, puis de faire tendre d'abord n , puis H , vers l'infini. L'inégalité de Bienaymé–Tchebychev donne alors

$$\mathbb{P}(Z_1 = \dots = Z_n = 0) \leq \frac{1}{p^2} \text{Var} \left(\frac{1}{n} \sum_{j=1}^n Z_j \right) \longrightarrow 0.$$

Presque sûrement, au moins un des points $(j, 0)$ n'appartient donc pas à A , ce qui prouve $A \neq \mathbb{R}^2$.

- (4) (a) Pour $s > 0$, l'image de M par l'homothétie $z \mapsto sz$ est un nuage poissonnien d'intensité $(\lambda/s^2) dz$. L'image de A est l'union des boules $B(sX_i, sr)$, et l'image de l'amas de l'origine est l'amas de l'origine dans cet ensemble. Une homothétie préserve le caractère non borné. Par conséquent,

$$\mathbb{P}_{\lambda, r}(\mathcal{C}(0) \text{ non borné}) = \mathbb{P}_{\lambda/s^2, sr}(\mathcal{C}(0) \text{ non borné}).$$

Le choix $s = \sqrt{\lambda}$ donne le premier changement d'échelle; le choix $s = 1/r$ donne le second.

- (b) À nuage fixé, si $r \leq r'$, alors $\bigcup_i B(X_i, r) \subset \bigcup_i B(X_i, r')$. Tout chemin du premier ensemble est un chemin du second, donc la probabilité considérée est croissante en r .

Si $0 < \lambda \leq \lambda'$, prenons deux nuages indépendants M_1, M_2 d'intensités respectives λdz et $(\lambda' - \lambda) dz$. Par superposition, $M_1 + M_2$ a l'intensité $\lambda' dz$. L'union des boules de rayon r centrées aux points de M_1 est incluse dans celle associée à $M_1 + M_2$. Le même argument sur les chemins prouve la croissance en λ .

Corrigé de l'exercice 7.5. [énoncé p. 204]

- (1) Pour toute partie
- A
- de
- $\{1, \dots, m\}$
- ,

$$\mathbb{P} \left(\bigcap_{j \in A} \{M(I_j) = 0\} \right) = \mathbb{P} \left(M \left(\bigcup_{j \in A} I_j \right) = 0 \right) = e^{-\sum_{j \in A} \lambda(I_j)} = \prod_{j \in A} \mathbb{P} (M(I_j) = 0).$$

C'est exactement l'indépendance des événements considérés.

- (2) D'après (1), les indicatrices sommées sont indépendantes, et chacune suit la loi de Bernoulli de paramètre
- $1 - e^{-2^{-n}}$
- . Par conséquent,

$$\mathcal{L}(M_n(J)) = \text{Bin}(K_n, 1 - e^{-2^{-n}}).$$

- (3) Pour chaque
- $x \geq 0$
- fixé,
- $\mathbb{P}(M(\{x\}) = 0) = 1$
- . Il n'y a donc presque sûrement aucun atome en
- $\inf J$
- ni en
- $\sup J$
- . De plus, les atomes dans
- J
- sont en nombre fini et sont distincts. Pour chaque réalisation vérifiant ces propriétés, dès que
- n
- est assez grand, les cellules dyadiques contenant ces atomes sont contenues dans
- J
- et contiennent chacune un seul atome. On a donc finalement
- $M_n(J) = M(J)$
- .

D'autre part, une cellule dyadique contenue dans J se partage en deux cellules du niveau suivant, elles aussi contenues dans J . Si elle contient un atome, au moins l'une de ses deux cellules en contient un. Ainsi $M_n(J) \leq M_{n+1}(J)$ et $M_n(J) \uparrow M(J)$ presque sûrement.

On a $K_n 2^{-n} \rightarrow \lambda(J)$ et $1 - e^{-2^{-n}} \sim 2^{-n}$. Pour $z \in [0, 1]$, la convergence dominée donne alors

$$\mathbb{E} \left[z^{M(J)} \right] = \lim_{n \rightarrow \infty} (1 - (1 - z)(1 - e^{-2^{-n}}))^{K_n} = e^{-\lambda(J)(1-z)}.$$

C'est la fonction génératrice d'une loi de Poisson de paramètre $\lambda(J)$, donc $M(J) \sim \mathcal{P}(\lambda(J))$.

- (4) Pour n fixé, les cellules dyadiques contenues dans les différents J_i sont deux à deux disjointes. D'après (1), les variables $M_n(J_1), \dots, M_n(J_k)$ sont donc indépendantes. En faisant tendre n vers l'infini dans leur fonction génératrice jointe, et en utilisant (3), on obtient l'indépendance de $M(J_1), \dots, M(J_k)$. Si l'un des intervalles est réduit à un point ou est vide, sa masse est nulle presque sûrement, et la conclusion reste valable.
- (5) Presque sûrement, $N_0 = 0$ et les trajectoires de N sont croissantes, càdlàg, à sauts de taille 1, par simplicité et finitude locale de M . Pour $0 = t_0 < t_1 < \dots < t_k$,

$$N_{t_i} - N_{t_{i-1}} = M([t_{i-1}, t_i]), \quad 1 \leq i \leq k,$$

et, d'après (3) et (4), ces variables sont indépendantes et de lois $\mathcal{P}(t_i - t_{i-1})$. Ainsi N est un processus de Poisson d'intensité 1.

Enfin, la loi d'une mesure aléatoire localement finie sur $\mathbb{R}_+$ est déterminée par les lois jointes de ses masses sur les intervalles $[0, t]$, $t \in \mathbb{Q}_+$. En effet, sur chaque

$[0, m]$, les évaluations sur les intervalles à extrémités rationnelles engendrent toutes les évaluations boréliennes par un argument de classe monotone ; on fait ensuite tendre m vers l'infini. Ces lois jointes sont ici celles d'un nuage poissonien d'intensité λ . Donc M est une mesure de Poisson de cette intensité.

* * *

Corrigé de l'exercice 7.6. [énoncé p. 205] On prend la tribu dénombrable-codénombrable

$$\mathcal{E} = \{B \subseteq [0, 1] : B \text{ est dénombrable ou } [0, 1] \setminus B \text{ est dénombrable}\},$$

et on pose

$$\mu(B) = \begin{cases} 0 & \text{si } B \text{ est dénombrable,} \\ 1 & \text{si } [0, 1] \setminus B \text{ est dénombrable.} \end{cases}$$

Les deux cas sont disjoints puisque $[0, 1]$ n'est pas dénombrable. La famille $\mathcal{E}$ est stable par complémentaire. Une réunion dénombrable de ses éléments est dénombrable si tous le sont, et codénombrable dès que l'un d'eux l'est : c'est donc bien une tribu.

Vérifions l'additivité dénombrable de μ . Dans une famille dénombrable d'éléments deux à deux disjoints de $\mathcal{E}$, il y a au plus un ensemble codénombrable. S'il n'y en a aucun, la réunion est dénombrable et les deux membres de l'égalité d'additivité sont nuls. S'il y en a un, tous les autres sont dénombrables et les deux membres valent 1. Ainsi μ est une mesure de probabilité. Enfin, pour tout $x \in [0, 1]$, $\{x\} \in \mathcal{E}$ et $\mu(\{x\}) = 0 \neq 1 = \delta_x(\{x\})$: μ n'est donc aucune masse de Dirac.