Théorèmes limites et Applications

Igor Kortchemski

Notes du cours de M2 enseigné au M2 de l'aléatoire (Université Paris-Saclay) entre 2020 et 2025

Exercices et problèmes corrigés : https://igor-kortchemski.perso.math.cnrs.fr/M2/TheoremesLimites-exercices.pdf

La première partie du cours est consacrée à l'étude de la notion de convergence en loi, d'abord dans $\mathbb{R}^n$, puis dans un cadre assez général où les objets aléatoires considérés prennent leurs valeurs dans un espace métrique complet et séparable. On démontre en particulier le théorème de Prokhorov qui caractérise les parties de mesures de probabilité relativement compactes pour la topologie de la convergence en loi.

Dans un second temps, cette théorie est appliquée à l'étude de la convergence en loi dans l'espace des fonctions continues à valeurs réelles, puis dans l'espace des fonctions càdlàg à valeurs réelles. On établit en particulier le théorème de Donsker, selon lequel une marche aléatoire à pas indépendants et de même loi converge après renormalisation vers un mouvement brownien.

La dernière partie du cours est consacrée à l'étude des mesures aléatoires de Poisson.

Sommaire

I] Convergence étroite dans $\mathbb{R}^k$

1) Définition p6
2) Restriction des fonctions test p7
3) Fonctions de répartition, tension ($k=1$) p8
4) Fonctions caractéristiques, théorème de Lévy p14

Intermède : séparabilité p18

II] Autres modes de convergence

1) Convergence presque sûre, en probabilité p19
2) Uniforme intégrabilité et convergence L^1 p20

III] Mesures de probabilité sur un espace métrique

1) Régularité de mesures p23
2) Théorème de porte-manteau p24
3) Convergence étroite dans $\mathbb{R}^{\mathbb{N}}$ p27
4) Tension et théorème de Prokhorov p30
5) Propriétés topologiques de $M_1(E)$: distance de Lévy-Prokhorov p37
6) Compléments de topologie : topologie faible-* sur $M_1(E)$ p41

IV] L'espace des fonctions continues sur un compact

1) Caractère polonais p46
2) Marginales fini-dimensionnelles p47
3) Convergence en loi p51
4) Compacité (Arzelà-Ascoli) p52
5) Tension p55
6) Critère de tension de Kolmogorov p56
7) Fonctions continues sur $\mathbb{R}_+$ p61

Interlude: Couplages

1) Théorème de représentation de Skorokhod p64
2) Distance en variation totale p69
3) Distance de Lévy-Prokhorov p70
4) Monotonie, ordre stochastique p72

V] Théorème de Donsker et applications

1) Le théorème de Donsker p73
2) Loi de l'arcsinus p77
3) Théorème de Lévy sur le mouvement brownien p79

VI] Le pont brownien

1) Définition p80
2) Le théorème local limite p81
3) Convergence vers le pont brownien (Donsker conditionné) p85

VII) Quelques mots sur les processus discontinus : l'espace $\mathbb{D}$ de Skorokhod

1) Propriétés de base p88
2) Topologie J_1 de Skorokhod p90
3) Convergence étroite dans $\mathbb{D}$ p93
4) Critères de tension p95
5) Compléments facultatifs
   a) $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$ p99
   b) C-tension p100
   c) Processus ponctuels p101
   d) Autres topologies sur $\mathbb{D}$. p101

VIII) Mesures aléatoires de Poisson

1) Rappels sur les lois de Poisson p102
2) Mesures aléatoires de Poisson p103
3) Points multiples p111
4) Formule de Campbell p114
5) Transformations de nuages poissoniens p117
6) Nuages de Poisson avec paramètre temporel. p119
7) Formule de Palm p121

Notations :

- (E, d) espace métrique
- $\mathcal{B}_E$ ou $\mathcal{B}$ tribu borélienne (+ petite qui contient les ouverts)
- $\mathcal{M}_1(E)$: mesures de proba sur E
- Si $f : E \to \mathbb{R}$ est mesurable ≥ 0 ou μ-intégrable, on note

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

- $\mathcal{C}_b(E)$: fonctions continues bornées $E \to \mathbb{R}$
- $\mathcal{C}_c(E)$: fonction continues $E \to \mathbb{R}$ à support compact

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

I] Convergence étroite dans $\mathbb{R}^k$

Plan : 1) Définition

2) Restriction des fonctions tests

3) Fonctions de répartition, réunion ($k=1$)

4) Fonctions caractéristiques, théorème de Lévy

1) Définition

Définition Soit (E,d) un espace métrique, $\mu_n, \mu \in M_1(E)$. On dit que (μ_n) converge étroitement vers μ si

$$\forall f \in C_b(E), \quad \mu_n(f) \longrightarrow \mu(f).$$

On note alors $\mu_n \Rightarrow \mu$

Exemples

- Si $x_n \longrightarrow x \in \mathbb{R}^k$, alors $\delta_{x_n} \Rightarrow \delta_x$
- $\mu_n = \frac{1}{n+1} \sum_{i=0}^{n} \delta_{i/n} \Rightarrow$ Lebesgue$([0,1])$

Formulation probabiliste Si X_n, X sont des v.a de lois respectives μ_n, μ, on dit que

$$X_n \xrightarrow{\text{loi}} X \text{ si } \mu_n \Rightarrow \mu, \text{ i.e } \forall f \in C_b(E), \ \mathbb{E}[f(X_n)] \longrightarrow \mathbb{E}[f(X)].$$

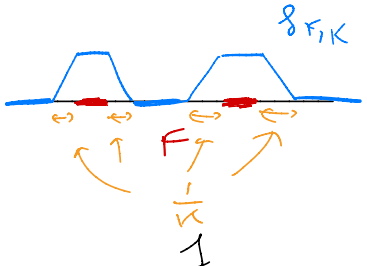
Proposition Soit (E,d) un espace métrique, $\mu_n, \mu, \nu \in M_1(E)$.

Si $\mu_n \Rightarrow \mu$ et $\nu_n \Rightarrow \mu$, alors $\mu = \nu$

Preuve : On a alors $\forall f \in C_b(E), \ \mu(f) = \nu(f)$

Soit F fermé et $K > 0$. On pose

$$f_{F,K}(x) = \max(1 - K\,d(x,F), 0)$$



Comme $g_{F,\varepsilon}$ converge simplement vers 1_F, par convergence dominée, on en déduit que $\mu(F) = \nu(F)$, et donc $\mu = \nu$ par le théorème des classes monotones

Dans la suite de ce cours 1, on se restreint à $\mathbb{R}^k$, muni de la distance euclidienne.

2) Restriction des fonctions test

Théorème ("petit porte-manteau") Soit $\mu_n, \mu \in M_1(\mathbb{R}^k)$. On a :

$\mu_n \Rightarrow \mu$ ssi $\forall f : \mathbb{R}^k \to \mathbb{R}_+$, mesurable, bornée, continue μ presque partout (i.e $\mu(\{x \in \mathbb{R}^k : f \text{ continue en } x\}) = 1$) on a $\mu_n(f) \to \mu(f)$

Voir p24 pour une preuve d'un énoncé plus général dans un cadre plus général.

Proposition Soit $\mu_n, \mu \in M_1(\mathbb{R}^k)$. Soit $H \subset \{\text{fonctions } \mathbb{R}^k \to \mathbb{R} \text{ mesurables bornées}\}$ tq $C_c(\mathbb{R}^k) \subset \bar{H}$, l'adhérence étant prise pour la norme uniforme. Alors

$\forall f \in H$, $\mu_n(f) \xrightarrow[n\to\infty]{} \mu(f)$ implique $\mu_n \Rightarrow \mu$

Preuve : **Etape 1 :** On montre $\forall f \in C_c(\mathbb{R}^k)$ $\mu_n(f) \to \mu(f)$ implique $\mu_n \Rightarrow \mu$.

Idée : argument de troncature. Soit $\gamma_r : \mathbb{R}^k \to [0,1]$ continue tq $\gamma_r(x) = 1$ pour $\|x\| \le r$, $\gamma_r(x) = 0$ pour $\|x\| \ge r+1$

Soit $f \in C_b(\mathbb{R})$. Alors

$$|\mu_n(f) - \mu(f)| \le |\mu_n(f) - \mu_n(f\gamma_r)| + |\mu_n(f\gamma_r) - \mu(f\gamma_r)| + |\mu(f\gamma_r) - \mu(f)|$$

Quand $n \to \infty$: $\limsup$ 1er terme $\le \|f\|_\infty (1 - \mu(\gamma_r))$ $\le \|f\|_\infty (1 - \mu(\gamma_r))$

$\limsup$ 2e terme $= 0$

Ainsi $\limsup_{n\to\infty} |\mu_n(f) - \mu(f)| \le 2\|f\|_\infty (1 - \mu(\gamma_R))$

Comme $\mu(\gamma_R) \xrightarrow[R\to\infty]{} 1$ (par convergence dominée), cela conclut l'étape 1.

Etape 2 Soit $f \in \mathcal{C}_c(\mathbb{R}^k)$. Soit $g \in H$ tq $\|f - g\|_\infty \le \varepsilon$. Alors

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

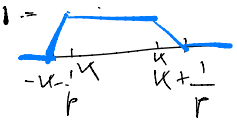
Donc $\mu_n(f) \to \mu(f)$. On conclut par l'étape 1.

Application : Soit $\mu \in M_1(\mathbb{R})$ et $(X_n)_{n\ge 1}$ des v.a. i.i.d. de loi μ.

Soit $\mu_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$. Alors p.s. $\mu_n \Rightarrow \mu$. En particulier, toute probabilité de $M_1(\mathbb{R})$ est limite étroite de mesures atomiques.

Preuve Soit $(f_k)_{k\ge 1}$ une suite de $\mathcal{C}_c(\mathbb{R})$ dense dans $\mathcal{C}_c(\mathbb{R})$,

prendre par exemple les polynômes à coefficients rationnels $\times$ fonctions plateau $T_{k,p}$ avec : $k, p \in \mathbb{N}$



On pose $H = \{f_k : k \ge 1\}$

Comme $\mu_n(f_k) = \frac{1}{n}\sum_{i=1}^n f_k(X_i) \xrightarrow[n\to\infty]{p.s.} \mathbb{E}[f_k(X_1)] = \mu(f_k)$ d'après la loi des grands nombres

Donc p.s. $\forall f \in H$, $\mu_n(f) \to \mu(f)$, d'où le résultat par la prop

3) Fonctions de répartition ($k=1$) et fonction

Définition Si $\mu \in M_1(\mathbb{R})$, on note $F_\mu(x) = \mu(]-\infty, x])$ pour $x \in \mathbb{R}$ sa fonction de répartition

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

• Réciproquement, si $F : \mathbb{R} \to [0,1]$ est croissante, continue à droite et vérifie $\lim_{-\infty} F = 0$, $\lim_{+\infty} F = 1$, alors $\exists \mu \in M_1(\mathbb{R})$ tq $F = F_\mu$. En effet, on peut prendre pour μ image de la mesure de Lebesgue sur $]0,1[$ par

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

(c'est la mesure de Lebesgue-Stieltjes associée à F, cf exercice)

• μ a un atome en $x \iff \Delta F_\mu(x) = F_\mu(x) - F_\mu(x-) = \mu(\{x\}) > 0$

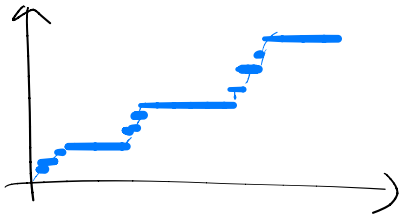
• μ a un nombre dénombrable d'atomes (car $\forall k \geq 1$, $\#\{x \in \mathbb{R},\ \mu(\{x\}) \geq \frac{1}{k}\} \leq k$)

• μ n'a pas d'atome $\iff F_\mu$ est continue

• μ a une densité g (% Lebesgue) $\Rightarrow F_\mu$ est dérivable presque partout de dérivée g et $F_\mu(x) = \int_{-\infty}^{x} g(t)\,dt$.

⚠ F_μ est toujours presque partout dérivable, mais n'est pas toujours l'intégrale de sa dérivée (c'est en fait le cas si $\mu \ll$ mesure de Lebesgue),

prendre l'escalier du diable / Cantor :
(dérivée = 0 presque partout)



Proposition Soit $\mu_n, \mu \in M_1(\mathbb{R})$.

$\mu_n \Rightarrow \mu$ ssi $F_{\mu_n}(x) \xrightarrow[n\to\infty]{} F_\mu(x)$ $\forall x \in \mathbb{R}$ tq $\Delta F_\mu(x) = 0$

Preuve : • Si $\mu_n \Rightarrow \mu$ et si $\Delta F_\mu(a) = 0$, alors $g(x) = \mathbb{1}_{x \le a}$ est continue μ-presque partout, et donc par petit porte-manteau,

$$\mu_n(g) \longrightarrow \mu(g)$$

• Réciproquement, on a alors $\mu_n(g) \longrightarrow \mu(g)$ 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 $C_c(\mathbb{R})$ est limite uniforme de telles fonctions. On conclut avec la proposition de 2).

$\sim$

Définition On dit qu'une famille $(\mu_i)_{i \in I}$ de mesures de $M_1(\mathbb{R})$ est tendue si $\forall \varepsilon > 0, \exists A > 0, \sup_{i \in I} \mu_i(\mathbb{R} \setminus [-A, A]) \le \varepsilon$

Remarque Une suite $(\mu_n)_{n \ge 1}$ de $M_1(\mathbb{R})$ est tendue si

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

Ceci provient du fait qu'une mesure $\mu \in M_1(\mathbb{R})$ est tendue puisque $\lim_{A \to \infty} \mu(\mathbb{R} \setminus [-A, A]) = 0$

Le résultat suivant est un résultat de type "compacité" et est un cas particulier du théorème de Prokhorov, qui sera vu au Cours 3.

Théorème Soit $(\mu_n)_{n\geq 1}$ une suite de $M_1(\mathbb{R})$. Alors

$(\mu_n)_{n\geq 1}$ est tendue $\Leftrightarrow$ $\forall$ extraction φ, $\exists$ extraction ψ t.q. $(\mu_{\varphi\circ\psi(n)})_{n\geq 1}$ converge étroitement

(par extraction, on veut dire sous-suite, ou, plus formellement, une injection croissante $\mathbb{N}^* \to \mathbb{N}^*$)

Preuve $\boxed{\Leftarrow}$ Par l'absurde, on suppose $\exists \varepsilon > 0$ t.q. $\forall n \geq 1$,

$\mu_{\phi(n)}(\mathbb{R}\setminus[-n,n]) > \varepsilon$ (on peut prendre ϕ croissante car à $i \geq 1$ fixé $\lim_{n\to\infty} \mu_i(\mathbb{R}\setminus[-n,n]) = 0$)

Par hypothèse, soit ψ extraction t.q. $(\mu_{\phi\circ\psi(n)})_{n\geq 1}$ converge étroitement vers $\mu \in M_1(\mathbb{R})$. Alors pour tout $x > 0$ t.q. $\mu(\{x\}) = \mu(\{-x\}) = 0$, pour n assez grand, $[-x, x] \subset [-\psi(n), \psi(n)]$. Donc

$$\mu([-x,x]) = \lim_{n\to\infty} \mu_{\phi\circ\psi(n)}([-x,x]) \leq \liminf_{n\to\infty} \mu_{\phi\circ\psi(n)}([-\psi(n),\psi(n)]) \leq 1-\varepsilon.$$

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

$\boxed{\Rightarrow}$ Quitte à travailler avec $\mu_{\phi(n)}$, on suppose $\phi = \mathrm{id}$

Étape 1 : identifier une limite potentielle. On pose $F_n = F_{\mu_n}$

$\forall q \in \mathbb{Q}$, $(F_n(q))_{n\geq 1}$ est borné. Par argument diagonal, $\exists \psi$ extraction t.q. $\forall q \in \mathbb{Q}$, $(F_{\psi(n)}(q))_{n\geq 1}$ 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 $A > 0$ tq $\sup_{n \geq 1} \mu_n(\mathbb{R} \setminus [-A, A]) \leq \varepsilon$.

Soit $q \in \mathbb{Q}$, $q > A$. Alors $F_n(q) \geq 1 - \varepsilon$

En faisant $n \to \infty$, on obtient $F_\infty(q) \geq 1 - \varepsilon$.

D'où $F(A) \geq 1 - \varepsilon$.

On montre de même que $F(-A) \leq \varepsilon$.

Donc $\lim_{-\infty} F = 0$, $\lim_{+\infty} F = 1$.

Ainsi $\exists\, \mu \in M_1(\mathbb{R})$ tq $F = F_\mu$ (c'est la mesure de Lebesgue-Stieltjes associée à F, cf remarque précédente)

Étape 2 : On vérifie que $\mu_{\varphi(n)} \Rightarrow \mu$.

Soit $x \in \mathbb{R}$ tq $\mu(\{x\}) = 0$. 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 $F(x) = \lim_{u \nearrow x} F(u)$)

Ainsi, si $q < x < q'$ avec $q, q' \in \mathbb{Q}$:

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

En faisant $q \uparrow x$, $q' \downarrow x$, on conclut que $F_{\varphi(n)}(x) \xrightarrow[n \to \infty]{} F(x)$.

$\rightsquigarrow$

Remarques: • Il est tentant de reformuler en disant que $\{\mu_n, n\geq 1\}$ est d'adhérence compacte dans $M_1(\mathbb{R})$ muni d'une distance qui métrise la convergence étroite. Ceci sera effectivement fait dans le cours mais c'est prématuré à ce stade.

• L'extension à $\mathbb{R}^n$ est similaire (cf Proposition 5.21 dans Kallenberg

Définition 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 Si $(X_i)_{i\in I}$ est borné dans L_1 (i.e $\exists C>0$ t.q $\sup_{i\in I}\mathbb{E}[|X_i|]<C$) 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 des sous-sous-suites pour les mesures

Soit $\mu_n, \mu \in M_1(\mathbb{R})$. Alors $\mu_n \Rightarrow \mu$ ssi $\forall \varphi$ extraction, $\exists \psi$ extraction t.q $\mu_{\varphi\circ\psi(n)} \Rightarrow \mu$.

Preuve: $\Rightarrow$ ok

$\Leftarrow$ Par l'absurde, soit $\varepsilon>0$, $f\in C_b(\mathbb{R})$ et φ extraction t.q $\forall n$, $|\mu_{\varphi(n)}(f)-\mu(f)|\geq\varepsilon$. Par hypothèse, $\exists \psi$ extraction t.q $\mu_{\varphi\circ\psi(n)} \Rightarrow \mu$. Alors $\mu_{\varphi\circ\psi(n)}(f) \to \mu(f)$, contradiction.

$\rightsquigarrow$

Corollaire : Soit $(\mu_n)_{n\geq 1}$ une suite de $M_1(\mathbb{R})$, $\mu \in M_1(\mathbb{R})$.

Alors $\mu_n \Rightarrow \mu$ ssi $\begin{cases} (\mu_n)_{n\geq 1} \text{ est tendue} \\ \text{si } \mu_{\phi(n)} \Rightarrow \nu \text{ pour une extraction } \phi \text{ et } \nu \in M_1(\mathbb{R}), \text{ alors } \mu = \nu \end{cases}$

Preuve : $\Rightarrow$ Clair

$\Leftarrow$ On raisonne par l'absurde. D'après le lemme précédent, $\exists \varphi$ extraction t.q. $\forall \psi$ extraction, $\mu_{\varphi\circ\psi(n)} \not\Rightarrow \mu$.

Or $(\mu_{\varphi(n)})$ est tendue, donc $\exists \psi$-extraction et $\nu \in M_1(\mathbb{R})$ t.q. $\mu_{\varphi\circ\psi(n)} \Rightarrow \nu$. Donc $\nu = \mu$ par hypothèse, absurde.

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

1. On montre que (μ_n) est tendue [Tension]
2. On prend φ extraction et ν t.q. $\mu_{\phi(n)} \Rightarrow \nu$.

On montre que $\nu = \mu$ en utilisant une propriété qui caractérise une mesure. [Unicité de la limite]

4) Fonctions caractéristiques (Thm de Lévy)

Définition Pour $\mu \in M_1(\mathbb{R}^k)$, on note $\phi_\mu(u) = \int_{\mathbb{R}^k} e^{i u\cdot x} \mu(dx)$, $u \in \mathbb{R}^k$, la fonction caractéristique de μ (produit scalaire)

Il est clair que $\mu_n \Rightarrow \mu$ implique $\phi_{\mu_n} \xrightarrow{CVS} \phi_\mu$ (convergence simple)

Pour la réciproque, on a le résultat plus précis suivant :

Théorème (Lévy) Soit $\mu_n \in M_1(\mathbb{R}^k)$ tq

① $\forall u \in \mathbb{R}^k$, $\phi_{\mu_n}(u)$ converge vers une limite notée $\phi(u)$

② $u \mapsto \phi(u)$ est continue en 0.

Alors $\exists \mu \in M_1(\mathbb{R}^k)$ tq $\phi = \phi_\mu$ et $\mu_n \Rightarrow \mu$

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

Lemme (injectivité de la fonction caractéristique)

Soit $\mu, \nu \in M_1(\mathbb{R})$ tq $\phi_\mu = \phi_\nu$. Alors $\mu = \nu$.

Idée de preuve : on convole avec une gaussienne : soit $g_\varepsilon(x) = \frac{1}{\sqrt{2\pi\varepsilon}} e^{-\frac{x^2}{2\varepsilon}}$

et on considère $\mu * g_\varepsilon(y) = \int \mu(dx)\, g_\varepsilon(x-y)$.

C'est une densité de proba $\mathcal{C}^\infty$, on vérifie que $\mu * g_\varepsilon(y)\,dy \underset{\varepsilon \to 0}{\Rightarrow} \mu$

Or $\phi_{g_\varepsilon}(\xi) = e^{-\frac{\varepsilon \xi^2}{2}} = g_{1/\varepsilon}(\xi)\sqrt{2\pi/\varepsilon}$

Donc $g_\varepsilon(\xi) = \frac{1}{\sqrt{2\pi\varepsilon}}\, \phi_{g_{1/\varepsilon}}(\xi)$.

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

$= \frac{1}{\sqrt{2\pi\varepsilon}} \int d\xi\, e^{i\xi\cdot(x-y)}\, g_{\frac{1}{\varepsilon}}(\xi)$

$\overset{\text{Fubini}}{=} \frac{1}{\sqrt{2\pi}} \int d\xi\, g_{\frac{1}{\varepsilon}}(\xi)\, e^{-i\xi y}\, \phi_\mu(\xi)$.

Ainsi, $\mu * g_\varepsilon$ s'exprime en fonction de ϕ_μ et ε ! On conclut en faisant $\varepsilon \to 0$

$\sim$

Lemme (lien queue / fonction caractéristique) $\exists c > 0$ tq

$\forall \mu \in M_1(\mathbb{R})$, $\forall K > 0$, on a $\mu(\mathbb{R} \setminus [-K, K]) \le c K \int_{-\frac{1}{K}}^{\frac{1}{K}} (1 - \mathrm{Re}\, \phi_\mu(u))\, du$

Preuve : On a

$$K\int_{-\frac{1}{K}}^{\frac{1}{K}}(1-\mathrm{Re}\,\phi_\mu(u))\,du = K\int_{-\frac{1}{K}}^{K}\left(1-\int_{\mathbb{R}}\cos(ux)\,\mu(dx)\right)du \overset{\text{(Fubini)}}{=} \int_{\mathbb{R}}\mu(dx)\,K\int_{-\frac{1}{K}}^{\frac{1}{K}}(1-\cos(ux))\,du$$

$$= 2\int_{\mathbb{R}}\mu(dx)\left(1-\frac{\sin(x/K)}{x/K}\right)$$

$$\geq 2\int_{\mathbb{R}\setminus[-K,K]}\mu(dx)\left(1-\frac{\sin(x/K)}{x/K}\right) \quad \text{Or} \quad \frac{\sin(y)}{y} \leq \sin(1) \text{ pour } |y|\geq 1$$

$$\geq 2\,(1-\sin(1))\,\mu(\mathbb{R}\setminus[-K,K]) \sim$$

Preuve du théorème de Lévy

Étape 1 : Tension

On a $\mu_n(\mathbb{R}\setminus[-A,A]) \leq cA\int_{-\frac{1}{A}}^{\frac{1}{A}}(1-\mathrm{Re}\,\phi_n(u))\,du \xrightarrow[n\to\infty]{} cA\int_{-\frac{1}{A}}^{\frac{1}{A}}(1-\mathrm{Re}\,\phi(u))\,du \xrightarrow[A\to\infty]{} 0$

On en déduit que $\limsup_{A\to\infty}\limsup_{n\to\infty}\mu_n(\mathbb{R}\setminus[-A,A]) = 0$ (ϕ continue en 0)

Donc $(\mu_n)_{n\geq 1}$ est tendue. En particulier, $\exists\,\psi_0$ extraction et $\mu\in\mathcal{M}_1(\mathbb{R})$ tq $\mu_{\psi_0(n)}\Rightarrow\mu$. Alors $\phi_{\mu_{\psi_0(n)}} \xrightarrow{CVS} \phi_\mu$. Donc $\phi=\phi_\mu$, ce qui montre déjà que ϕ est la fct caractéristique d'une mesure de proba.

Étape 2 : unicité de la limite

Soit ψ extraction et $\nu\in\mathcal{M}_1(\mathbb{R})$ tq $\mu_{\psi(n)}\Rightarrow\nu$.

Alors $\phi_{\mu_{\psi(n)}} \xrightarrow{CVS} \phi_\nu$. Or $\phi_{\mu_n} \xrightarrow{CVS} \phi=\phi_\mu$.

Donc $\phi_\mu=\phi_\nu$ et $\mu=\nu$ par injectivité de la fonction caractéristique $\sim$

On retrouve une structure du raisonnement en deux étapes, qui est le crédo de ce cours. Il y a essentiellement 3 ingrédients :

① Tension

② Une certaine propriété ? parce à la limite étroite

(pour le thm de Lévy $\mu_n \Rightarrow \mu$ implique $\phi_{\mu_n} \xrightarrow{CVS} \phi_\mu$)

③ Cette propriété ? caractérise les mesures de proba.

Souvent ② + ③ est appelé "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 proba (en vue de l'unicité de la limite).

Interlude : Séparabilité

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

Par exemple, sont séparables $(\ell_\infty(\mathbb{R}), \|\cdot\|_\infty)$ et tout espace métrique compact (utiliser Borel-Lebesgue), et $(\ell_b(\mathbb{R}), \|\cdot\|_\infty)$ ne l'est pas (exercice)

- La séparabilité est une propriété très utile lorsqu'on considère des produits d'espaces (Si E, F sont 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 v.a à valeurs dans (E, δ) séparable, alors $d(X,Y)$ est mesurable comme composée de

$$(\Omega, \mathcal{F}) \longrightarrow (E\times E, \mathcal{B}(E)\otimes\mathcal{B}(E)) \quad (=)$$
$$\omega \longmapsto (X(\omega), Y(\omega))$$

et $(E\times E, \mathcal{B}(E\times E)) \longrightarrow \mathbb{R}$ (mesurable car continue)
$$(x, y) \longmapsto d(x,y)$$

- La séparabilité est une notion topologique, au sens où si E et F sont homéomorphes, alors E séparable $\Leftrightarrow$ F séparable (NB: la complétude n'est pas une notion topologique)

II) Autres modes de convergence

Plan : 1) Convergence presque sûre, en probabilité
2) Uniforme intégrabilité et convergence L^1

1) Convergence presque sûre, en probabilité

Soit $(X_n)_{n\geq 1}$ et X des v.a définies sur le même espace de proba, à valeurs dans (E,d) séparable

Définition

- On dit que X_n converge presque surement (p-s) vers X et on note $X_n \xrightarrow{p\text{-}s} X$ si
$$\mathbb{P}\left(\{\omega\in\Omega : X_n(\omega)\xrightarrow[n\to\infty]{} X(\omega)\}\right)=1$$
- On dit que X_n converge en probabilité vers X et on note $X_n \xrightarrow{\mathbb{P}} X$ si
$$\forall \varepsilon>0,\ \mathbb{P}(d(X_n,X)>\varepsilon)\xrightarrow[n\to\infty]{} 0$$

Remarques · $\mathbb{P}(X_n\to X)$ 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 ps 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 On a $X_n \xrightarrow{\mathbb{P}} X$ ssi de toute sous-suite $(X_{\varphi(n)})$ on peut reextraire une sous-sous-suite $(X_{\varphi\circ\psi(n)})$ qui converge ps vers X

En particulier, $X_n \xrightarrow{p\text{-}s} X$ implique $X_n \xrightarrow{\mathbb{P}} X$

Preuve Vérifions tout d'abord que $\boxed{X_n \xrightarrow{\mathbb{P}} X \iff \mathbb{E}[d(X_n,X)\wedge 1]\longrightarrow 0}$ (minimum)

$\Leftarrow$ On a pour $0<\varepsilon<1$ $\mathbb{P}(d(X_n,X)>\varepsilon)\leq \frac{1}{\varepsilon}\mathbb{E}[d(X_n,X)\wedge 1]\longrightarrow 0$

$\Rightarrow$ On écrit pour $\varepsilon>0$ $\mathbb{E}[d(X_n,X)\wedge 1]\leq \mathbb{E}[d(X_n,X)\mathbb{1}_{d(X_n,X)\leq\varepsilon}]+\mathbb{E}[d(X_n,X)\wedge 1\,\mathbb{1}_{d(X_n,X)\geq\varepsilon}]$
$$\leq \varepsilon + \mathbb{P}(d(X_n,X)\geq\varepsilon)$$
Donc $\limsup_{n\to\infty}\mathbb{E}[d(X_n,X)\wedge 1]=0$.

Revenons à la preuve de la proposition :

$\Rightarrow$ Soit φ extraction. Il existe alors une extraction ψ telle que $\forall n \geq 1$, $\mathbb{E}[d(X_{\varphi\circ\psi(n)}, X) \wedge 1] \leq \frac{1}{2^n}$

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

Donc p.s $\sum_{n=1}^{\infty} d(X_{\varphi\circ\psi(n)}, X) \wedge 1 < \infty$. Donc $X_{\varphi\circ\psi(n)} \xrightarrow[n\to\infty]{p.s} X$.

$\Leftarrow$ Par l'absurde, si $X_n \xrightarrow{\mathbb{P}} X$, on trouve $\varepsilon > 0$ et une extraction φ t.q $\forall n$ $\mathbb{E}[d(X_{\varphi(n)}, X) \wedge 1] \geq \varepsilon$

Soit alors ψ extraction t.q $X_{\varphi\circ\psi(n)} \xrightarrow{p.s} X$.

Par convergence dominée, $\mathbb{E}[d(X_{\varphi\circ\psi(n)}, X) \wedge 1] \longrightarrow 0$.

Absurde

2) Uniforme intégrabilité et cv dans L^1

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

Définition On dit que (X_n) converge vers X dans L^1 et on note $X_n \xrightarrow{L^1} X$ si

$$\mathbb{E}[|X_n - X|] \xrightarrow[n\to\infty]{} 0$$

Le thm de convergence dominée permet de passer de la cv ps à la cv L^1 avec une hypothèse de domination. Aussi, le lemme de Scheffé (cf feuille d'exercices 1) dit que $X_n \geq 0$, $X \geq 0$, $X_n \xrightarrow{p.s} X$, $\mathbb{E}[X_n] \to \mathbb{E}[X] \Rightarrow X_n \xrightarrow{L^1} X$.

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

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

de variables aléatoires.

Théorème Soit $(X_i)_{i\in I}$ des v.a réelles intégrables. Les 3 conditions suivantes sont équivalentes :

(1) $\sup_{i\in I} \mathbb{E}[|X_i| \mathbb{1}_{|X_i|\geq x}] \xrightarrow[x\to\infty]{} 0$

(2) $\sup_{i\in I} \mathbb{E}[|X_i|] < \infty$ et $\forall \varepsilon > 0, \exists \delta > 0, \forall A$ mesurable, $\mathbb{P}(A) \leq \delta \Rightarrow \sup_{i\in I} \mathbb{E}[|X_i| \mathbb{1}_A] \leq \varepsilon$

(3) il existe une fonction croissante $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ t.q. $\frac{\varphi(x)}{x} \xrightarrow[x\to\infty]{} +\infty$ et $\sup_{i\in I} \mathbb{E}[\varphi(|X_i|)] < \infty$ [critère de de la Vallée Poussin]

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

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

Corollaire : Une somme de deux familles UI est UI

Preuve du Thm : **(1) ⇒ (2)** Pour la bornitude dans L^1, on écrit pour x assez grand :

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

Ensuite, on fixe $\varepsilon > 0$. Par hypothèse, il existe x_ε t.q. $\sup_{i\in I} \mathbb{E}[|X_i| \mathbb{1}_{|X_i|\geq x_0}] \leq \varepsilon$

Alors, si $\mathbb{P}(A) \leq \varepsilon / x_\varepsilon$, pour $i \in I$ on a

$$\mathbb{E}[|X_i| \mathbb{1}_A] = \mathbb{E}[|X_i| \mathbb{1}_A \mathbb{1}_{|X_i|\leq x_\varepsilon}] + \mathbb{E}[|X_i| \mathbb{1}_A \mathbb{1}_{|X_i|> x_\varepsilon}]$$

$$\leq x_\varepsilon \mathbb{P}(A) + \varepsilon \leq^? \varepsilon$$

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

(3) ⇒ (1) Soit $\varepsilon > 0$ et $x_\varepsilon > 0$ t.q. $x_\varepsilon \leq \varepsilon \varphi(x_\varepsilon)$. 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) ⇒ (3) Soit (x_n) une suite strictement croissante telle que

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

On pose alors $\varphi(x) = \sum_{m=1}^{\infty} (x - x_m)_+$, $x \geq 0$, qui est croissante.

De plus $\frac{\varphi(x)}{x} = \sum_{m=1}^{\infty} \left(1 - \frac{x_m}{x}\right)_+ \xrightarrow[x\to\infty]{} +\infty$ par convergence monotone

et $\mathbb{E}[\varphi(|X_i|)] \leq \sum_{m=1}^{\infty} \mathbb{E}\left[|X_i| \mathbb{1}_{|X_i|\geq x_m}\right] \leq 1.$

$\rightsquigarrow$

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

Remarque On a vu (cf feuille exos 1) que $(X_i)_{i\in I}$ est tendue ssi $\exists \varphi : \mathbb{R}_+ \to \mathbb{R}_+$ t.q. $\lim_{+\infty} \varphi = +\infty$

et $\sup_{i\in I} \mathbb{E}[\varphi(|X_i|)] < \infty$ (c'est donc une notion plus faible que UI et **UI ⇒ tension**.

Proposition Soit $(X_n)_{n\geq 1}$ des v.a. réelles intégrables et X une v.a. réelle.

On a $X_n \xrightarrow[\text{et } X\in L^1]{L^1} X$ ssi $X_n \xrightarrow{\mathbb{P}} X$ et $(X_n)_{n\geq 1}$ est UI.

Preuve : $\Rightarrow$ On suppose $X_n \xrightarrow{L^1} X$. Alors $X_n \xrightarrow{\mathbb{P}} X$ (Markov). Pour mq $(X_n)_{n\geq 1}$ est UI

il suffit de mq $(X_n - X)_{n\geq 1}$ est UI, puisque la somme de deux familles UI est UI par

caractérisation (2). Pour cela, si $\varepsilon > 0$, on choisit n_0 t.q. $n \geq n_0 \Rightarrow \mathbb{E}[|X_n - X|] \leq \varepsilon$ pour $n \geq n_0$

et x_0 t.q. $x \geq x_0 \Rightarrow \max_{2\leq i\leq n_0} \mathbb{E}\left[|X_i - X| \mathbb{1}_{|X_i - X|\geq x}\right] \leq \varepsilon$. Alors $\sup_{i\geq 1} \mathbb{E}\left[|X_i - X| \mathbb{1}_{|X_i - X|\geq x}\right] \leq \varepsilon$.

$\Leftarrow$ Vérifions d'abord que $X \in L^1$. Soit φ extraction t.q. $X_{\varphi(n)} \xrightarrow{ps} X$. Alors, d'après

le lemme de Fatou, $\mathbb{E}[|X|] \leq \liminf_{n\to\infty} \mathbb{E}[|X_{\varphi(n)}|] < \infty$ 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

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

$\rightsquigarrow$

III) Mesures de probabilité sur un espace métrique

Plan : 1) Régularité de mesures
2) Théorème de porte-manteau
3) Convergence étroite dans $\mathbb{R}^d$
4) Tension et théorème de Prokhorov
5) Propriétés topologiques de $M_1(E)$: distance de Lévy-Prokhorov
6) Compléments de topologie : topologie faible-$*$ sur $M_1(E)$

1) Régularité de mesures

Théorème Soit (E,d) un espace métrique et $\mu \in M_1(E)$. Pour tout $A \in \mathcal{B}(E)$,

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

Preuve : c'est une preuve à la "théorie de la mesure". On note

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

- Si O ouvert, $O \in \mathcal{A}$: pour l'inf, ok. Pour le sup, soit $F_n = \{x \in E : d(x, {}^cO) \geq \frac{1}{n}\}$, fermé.
Comme $O = \bigcup_{n \geq 1}^{\uparrow} F_n$, on a $\mu(O) = \lim_{n\to\infty}^{\uparrow} \mu(F_n)$.

- $\mathcal{A}$ est une tribu : $\emptyset \in \mathcal{A}$: ok. Passage au complémentaire : ok.
Soit $(A_n)_{n\geq 1}$ une suite de $\mathcal{A}$ et F_n fermé, O_n ouvert tq $F_n \subset A_n \subset O_n$ et $\mu(O_n \setminus F_n) \leq \frac{\varepsilon}{2^n}$.
Alors $O = \bigcup_{n\geq 1} O_n$ est un ouvert contenant $\bigcup_{n\geq 1} A_n$ tq $\mu(O \setminus \bigcup_{n\geq 1} A_n) \leq \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} \leq \varepsilon$.
Cependant, $F' = \bigcup_{n\geq 1} F_n$ n'est pas forcément fermé. Pour contourner, on remarque que
$\bigcup_{n=1}^{k} F_n \uparrow F'$ et donc $\exists N \geq 1$ tq $\mu(F' \setminus \bigcup_{n=1}^{N} F_n) \leq \varepsilon$. On pose $F = \bigcup_{n=1}^{N} F_n$
On a alors $\mu(O \setminus F) \leq \mu(O \setminus F') + \mu(F' \setminus F) \leq 2\varepsilon$.

$\sim$

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

2) Théorème de Porte-manteau

Ici (E,d) est un espace métrique

Théorème (Porte-manteau/Alexandrov) Soit $(\mu_n)_{n\geq 1}$ une suite de $M_1(E)$ et $\mu \in M_1(E)$.

Ses propriétés suivantes sont équivalentes:

(1) $\mu_n \Rightarrow \mu$

(2) $\forall f: E \to \mathbb{R}$ lipschitzienne bornée, $\mu_n(f) \longrightarrow \mu(f)$

(3) $\forall F$ fermé, $\limsup_{n\to\infty} \mu_n(F) \leq \mu(F)$ [intuition: la masse ne peut pas sortir de fermés]

(4) $\forall O$ ouvert, $\liminf_{n\to\infty} \mu_n(O) \geq \mu(O)$

(5) $\forall A \in \mathcal{B}(E)$ tq $\mu(\partial A)=0$, on a $\lim_{n\to\infty} \mu_n(A) = \mu(A)$.

(6) $\forall f: E \to \mathbb{R}$ mesurable bornée et continue en μ-presque tout point, on a $\mu_n(f) \longrightarrow \mu(f)$.

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

$X_n \xrightarrow{loi} X \iff \forall f: E\to\mathbb{R}$ lip bornée $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)] \iff \limsup_{n\to\infty} \mathbb{P}(X_n \in F) \leq \mathbb{P}(X\in F)$ etc.

Avant la preuve, voyons quelques applications.

Proposition Soient X_n, X v.a à valeurs dans (E,d) telles que $X_n \xrightarrow{loi} X$. Soit $f: E \to \mathbb{R}$ mesurable presque sûrement continue en X. Alors $f(X_n) \xrightarrow{loi} f(X)$

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

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

Preuve: On applique portemanteau. Soit F fermé, et notons $F_\varepsilon = \{x\in S : d(x,F)\leq\varepsilon\}$, 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 $\limsup_{n\to\infty} \mathbb{P}(Y_n\in F) \leq \limsup_{n\to\infty} \mathbb{P}(X_n\in F_\varepsilon) = \mathbb{P}(X\in F_\varepsilon)$.

Or $\mathbb{P}(X\in F_\varepsilon) \xrightarrow[\varepsilon\to 0]{} \mathbb{P}(X\in F)$ (car F fermé) ce qui conclut.

Corollaire (Lemme de Slutsky) Soient (E, d_E), $(E', d_{E'})$ séparables.

$X_n \Rightarrow X$ dans E, $X'_n \xrightarrow{\mathbb{P}} x'$ dans E' avec $x' \in E'$ fixé

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

Preuve : • On munit $E \times E'$ de la distance $d((e,e'),(b,b')) = d_E(a,b) + d_{E'}(a',b')$.

• Ainsi, $(X_n, x') \Rightarrow (X, x')$ (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.

$\rightsquigarrow$

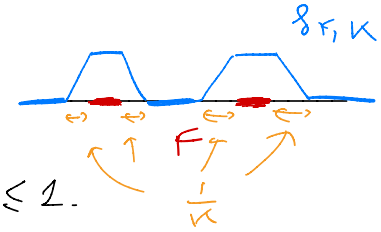
Application : dans $\mathbb{R}$, si $X_n \xrightarrow{\text{loi}} X$, $Y_n \xrightarrow{\mathbb{P}} 1$, alors $\frac{X_n}{Y_n} \xrightarrow{\text{loi}} X$. En effet, $(X_n, Y_n) \xrightarrow{\text{loi}} (X, 1)$ par Slutsky, puis $g(X_n, Y_n) \xrightarrow{\text{loi}} g(X, 1)$ avec $g(x,y) = \begin{cases} \frac{x}{y} & \text{si } y \neq 0 \\ 0 & \text{sinon} \end{cases}$, p.s. continue en $(X, 1)$

Preuve du théorème de Porte-manteau

$\boxed{(1) \Rightarrow (2)}$ OK

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

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



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

Par convergence dominée, $\mu(g_{F,K}) \xrightarrow[K \to \infty]{} \mu(F)$, d'où le résultat

$\boxed{(3) \Rightarrow (4)}$ OK par passage au complémentaire

$\boxed{(3)+(4) \Rightarrow (5)}$ Soit $A \in \mathcal{B}(E)$ tq $\mu(\partial A) = 0$, de sorte que $\mu(\mathring{A}) = \mu(A) = \mu(\bar{A})$.

Alors $\mu(A) = \mu(\mathring{A}) \leq \liminf_{n \to \infty} \mu_n(\mathring{A}) \leq \limsup_{n \to \infty} \mu_n(A) \leq \limsup_{n \to \infty} \mu_n(\bar{A})$

$\leq \mu(\bar{A}) = \mu(A)$.

Donc tous les $\leq$ sont des $=$.

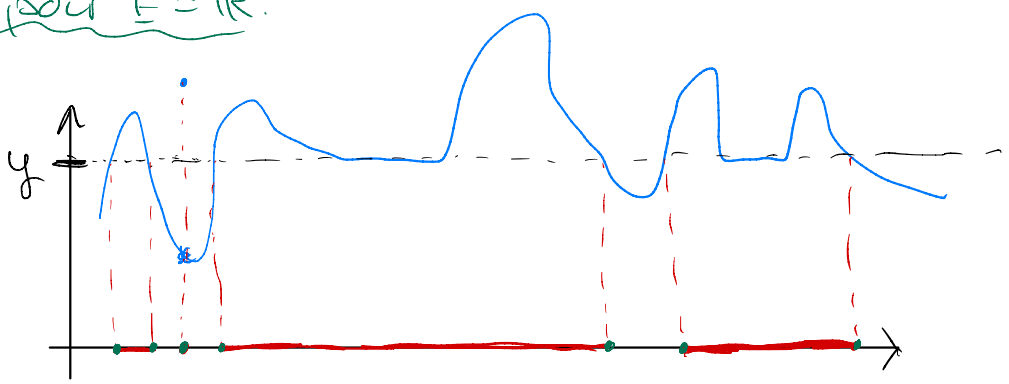
(5)⇒(6) Soit f continue μ presque partout, bornée par M. Quitte à faire $f = f_+ - f_-$ on peut supposer $f \geq 0$. Alors, par Fubini,

$$\mu_n(f) = \int_E f(x)\,\mu_n(dx) = \int_E \int_0^{+\infty} \mathbb{1}_{y \leq f(x)}\,dy\,\mu_n(dx) = \int_0^{+\infty} dy \int_0^{+\infty} \mathbb{1}_{y \leq f(x)}\,\mu_n(dx)$$

$$= \int_0^{+\infty} dy\ \mu_n(\underbrace{\{f \geq y\}}_{A_y}) = \int_0^M dy\ \mu_n(A_y)$$

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 $E = \mathbb{R}$:



En rouge $A_y = \{x : f(x) \geq y\}$, en vert ∂A_y.

En effet, soit $x_n \in A_y$ t.q. $x_n \to x \in \partial A_y$. Si $x \in D$, il n'y a rien à faire. Si $x \notin D$: comme $f(x_n) \geq y$ $\forall n$, on a alors $f(x) \geq y$. Alors soit $f(x) = y$, soit $f(x) > y$, auquel cas $x \in \overset{\circ}{A}_y$, pas possible.

Or il y a au plus un nombre dénombrable de y t.q. $\mu(\{f = y\}) > 0$ (pour un entier $n \geq 1$, il y a au plus n y t.q. $\mu(\{f = y\}) \geq \frac{1}{n}$ car μ est une probe)

Donc pour Lebesgue presque tout y, $\mu(\partial A_y) = 0$ et $\mu_n(A_y) \xrightarrow[n \to \infty]{} \mu(A_y)$

Par convergence dominée, on conclut que

$$\mu_n(f) \xrightarrow[n \to \infty]{} \int_0^M dy\ \mu(A_y) = \mu(f)$$

(6)⇒(1) OK

~

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.

(a) Tribu sur $\mathbb{R}^{\mathbb{N}}$

Définition : 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 $\pi_k : \mathbb{R}^{\mathbb{N}} \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mesurables
$(x_i)_{i\geq 1} \mapsto x_k$
Elle est parfois appelée tribu cylindrique.

Remarques 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 $(\Omega, \mathcal{F}) \to (\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R})^{\otimes \mathbb{N}})$
$\omega \mapsto (x_i(\omega))_{i\geq 1}$
est mesurable · En particulier $(X_i)_{i\geq 1}$ est une v.a (à valeurs dans $\mathbb{R}^{\mathbb{N}}$) ssi $\forall i \geq 1$, X_i est une v.a.

Définition Soit $\mu \in M_1(\mathbb{R}^{\mathbb{N}})$. Pour $1 \leq i_1 < \dots < i_k$, on note $\mu_{i_1, \dots, i_k}$ la mesure image de μ par l'application $\mathbb{R}^{\mathbb{N}} \to \mathbb{R}^k$
$(x_i)_{i\geq 1} \mapsto (x_{i_1}, \dots, x_{i_k})$. 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 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 (fdi en abrégé)

Proposition Une mesure $\mu \in M_1(\mathbb{R}^{\mathbb{N}})$ est caractérisée par ses marginales fidi

En particulier, la loi d'une v.a. $(X_i)_{i\geq 1} \in \mathbb{R}^{\mathbb{N}}$ est caractérisée par ses lois fidi

Preuve : ceci provient du fait que les cylindres forment un π-système générateur de $\mathcal{B}(\mathbb{R})^{\otimes \mathbb{N}}$, et donc deux mesures de $M_1(\mathbb{R}^{\mathbb{N}})$ qui coïncident dessus sont égales d'après le lemme des classes monotones.

En termes probabilistes, $(X_i)_{i\geq 1} \overset{\text{loi}}{=} (Y_i)_{i\geq 1} \Leftrightarrow \forall 1 \leq i_1 < \dots < i_k$,
$(X_{i_1}, \dots, X_{i_k}) \overset{\text{loi}}{=} (Y_{i_1}, \dots, Y_{i_k})$

⑤ 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 Pour $(x_i)_{i\geq 1}, (y_i)_{i\geq 1} \in \mathbb{R}^{\mathbb{N}}$, on pose

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

Si $(x_i)_{i\geq 1}$ et $(y_i)_{i\geq 1}$ coïncident sur les K premiers termes (i.e $x_i = y_i$ 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 $(x_i)_{i\geq 1} \in \mathbb{R}^{\mathbb{N}}$, $d(x^n, x) \to 0 \Leftrightarrow \forall i \geq 1, x_i^n \underset{n\to\infty}{\longrightarrow} x_i$

Remarque 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 (i.e collection d'ouverts) pour laquelle toutes les projections π_k sont continues.

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

Preuve : • $\mathcal{B}(\mathbb{R})^{\otimes\mathbb{N}} \subset \mathcal{B}(\mathbb{R}^{\mathbb{N}})$ car les projections π_k sont continues.

• Pour l'autre inclusion, soit $(z^j)_{j\geq 1}$ une suite dense dans $\mathbb{R}^{\mathbb{N}}$. On définit $N^j_{k,M} = \{(y_i) \in \mathbb{R}^{\mathbb{N}} : |y_i - z^j_i| < \frac{1}{M} \text{ pour } 1 \leq i \leq k\}$, ouvert de $\mathbb{R}^{\mathbb{N}}$, appelé ouvert élémentaire. Comme il y a un nombre dénombrable de tels ouverts, il suffit de vérifier que tout ouvert de $\mathbb{R}^{\mathbb{N}}$ est union d'ouverts élémentaires (car ils sont dans $\mathcal{B}(\mathbb{R})^{\otimes\mathbb{N}}$)
Pour cela, il suffit de montrer que $\forall O$ ouvert, $\forall x \in O$, il existe un ouvert élémentaire contenant x et inclus dans O. Pour cela, soit $\varepsilon > 0$ tq $B(x, 3\varepsilon) \subset O$.
On choisit alors M et k tq $\frac{1}{2^k} < \varepsilon$ et $\frac{1}{M} < \varepsilon$, et z^j tq $d(z^j, x) < \frac{1/M}{1+1/M} 2^{-k}$.
On vérifie alors $x \in N^j_{k,M}$ et que $N^j_{k,M} \subset B(x, 3\varepsilon)$

c) 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 Soit $(X^n)_{n\geq 1}$, X des v.a de $\mathbb{R}^{\mathbb{N}}$. Alors

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

Preuve : $\Rightarrow$ Clair par continuité de $\pi_{i_1,\dots,i_k} : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^k$, $(x_i)_{i\geq 1} \mapsto (x_{i_1},\dots,x_{i_k})$

$\Leftarrow$ Soit $f:\mathbb{R}^{\mathbb{N}}\to\mathbb{R}$ lipschitzienne bornée. Soit $\delta>0, \varepsilon>0$ tq $d(x,y)\leq\delta \Rightarrow |f(x)-f(y)|\leq\varepsilon$. pour $x,y\in\mathbb{R}^{\mathbb{N}}$

On choisit alors K tq $\frac{1}{2^{K+1}}<\delta$ et on écrit

$$|\mathbb{E}[f(X^n_1,\dots)]-\mathbb{E}[f(X_1,\dots)]| \leq \underbrace{\mathbb{E}[|f(X^n_1,\dots)-f(X^n_1,\dots,X^n_K,0,0,\dots)|]}_{\leq\varepsilon}$$

$$+\underbrace{\mathbb{E}[|f(X^n_1,\dots X^n_K,0,0\dots)-f(X_1,\dots X_K,0,0\dots)|]}_{\xrightarrow[n\to\infty]{}0 \text{ par hypothèse}} + \underbrace{\mathbb{E}[|f(X_1,\dots,X_K,0,0\dots)-f(X_1,\dots)|]}_{\leq\varepsilon}$$

Ceci conclut par porte-manteau

4) Tension et théorème de Prokhorov

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

Définition Soit (E,d) un espace métrique et $A\subset E$.

On dit que A est relativement compact si $\bar{A}$ est compact

Remarque A est relativement compact $\iff$ de toute suite de A on peut extraire une sous-suite qui converge

En particulier $\{x_n : n\geq 1\}$ est relativement compact $\iff \forall$ extraction ϕ, $\exists\ \psi$ extraction tq $(x_{\phi\circ\psi(n)})_{n\geq 1}$ converge

Lemme A précompact $\iff \bar{A}$ précompact

Preuve : $\Leftarrow$ Clair

⇒ Si $A \subset \bigcup_{n\geq 1} B(x_n, \varepsilon)$, vérifions que $\bar{A} \subset \bigcup_{n\geq 1} B(x_n, 2\varepsilon)$.

Soit $y \in \bar{A}$ et $y_n \to y$ avec $y_n \in A$. Soit n_0 tq $d(y, y_{n_0}) < \varepsilon$

Soit n_1 tq $y_{n_0} \in B(x_{n_1}, \varepsilon)$. Alors $y \in B(x_{n_1}, 2\varepsilon)$ par inégalité triangulaire

$\Leftarrow$

Avec la séparabilité, la complétude joue un rôle important dans le thm 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) pas complet (par ex $E = \mathbb{R}$, $d_1(x,y) = |x-y|$, $d_2(x,y) = |\arctan(x) - \arctan(y)|$)

Un des intérêts de la complétude est d'avoir une caractérisation des compacts grâce à la notion de précompacité

Définition : Soit (E,d) un espace métrique et $A \subset E$.

- On dit que A est précompact si $\forall \varepsilon > 0$, A peut être recouvert par un nombre fini de boules de rayon ε.

Exemple : Un espace métrique compact est précompact et complet

Le résultat suivant sera très utile [exercice] pour la preuve.

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

A précompact $\Leftrightarrow$ A relativement compact

Définition : Soit (E,d) un espace métrique

- On dit qu'une mesure $\mu \in M_1(E)$ est tendue si $\forall \varepsilon > 0, \exists K$ compact tq $\mu(K) \geq 1-\varepsilon$.
- On dit qu'une famille $(\mu_i)_{i\in I}$ de mesures est tendue si

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

Une famille de variables aléatoires est tendue si la famille de leurs lois est

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

Preuve : Soit $\varepsilon>0$ et K compact tq $\inf_{i\in I}\mu_i(K)>1-\varepsilon$. 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 (un peu simplifiée) On dit que (E,d) est polonais s'il est séparable et complet.

NB : 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 = métrique, complet, séparable.

Proposition Soit (E,d) polonais et $\mu\in M_1(E)$. Alors μ est tendue.

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

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

On pose alors $K=\bigcap_{n\geq 1}\bigcup_{k=1}^{N_n}\overline{B}\left(x_k,\frac{1}{n}\right)$.

K est fermé et précompact. Il est donc compact.

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

Théorème (Prokhorov) Soit (E,d) métrique et $(\mu_n)_{n\geq 1}$ une suite de $M_1(E)$

1. Si $(\mu_n)_{n\geq 1}$ est tendue, alors $\forall$ extraction φ, $\exists$ extraction ψ tq $(\mu_{\varphi\circ\psi(n)})_{n\geq 1}$ converge étroitement
2. Si (E,d) est polonais, la réciproque est vraie

On commence par un lemme :

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

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

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

- 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 tq $d_E(x, z^i) < \frac{d_E(x,y)}{2}$.

On ne peut avoir $d_E(x, z^i) = d_E(y, z^i)$ par inégalité triangulaire

- Supposons $\Phi(x^n) \to \Phi(x)$ Par l'absurde, supposons $x^n \not\to x$.

Il existe alors $\varepsilon > 0$ et une extraction ϕ tq $\forall n$, $d_E(x^{\phi(n)}, x) \geq \varepsilon$

On choisit z^i tq $d_E(z^i, x) < \frac{\varepsilon}{2}$. Alors $d_E(x^{\phi(n)}, z^i) \geq \frac{\varepsilon}{2}$.

Or $d_E(x^{\phi(n)}, z^i) \to d_E(x, z^i)$, absurde.

Donc Φ est un homéomorphisme sur son image $\sim$

Revenons à la preuve du thm de Prokhorov.

Preuve de (1) (c'est le sens le plus utile, mais aussi le plus difficile.

L'idée est de traiter successivement les cas $\mathbb{R}^1$, $\mathbb{R}^{\mathbb{N}}$, $E \subset \mathbb{R}^{\mathbb{N}}$, E

- Le cas $E = \mathbb{R}$ a été vu au cours 1, le cas $\mathbb{R}^d$ se traite de même. (cf feuille d'exos ou Kallenberg, prop 5.2.1)
- Cas $E = \mathbb{R}^{\mathbb{N}}$: Notons $\hat{\Pi}_k : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^k$, $(x_i)_{i\geq 1} \mapsto (x_1, \dots, x_k)$ continue. L'image d'un compact par une fonction continue étant compacte, $\forall k \geq 1$, $(\hat{\Pi}_k \mu_n)_{n\geq 1}$ est tendue dans $\mathbb{R}^k$.

Par procédé diagonal, $\forall$ extraction φ, $\exists$ extraction ψ tq si $\sigma = \varphi \circ \psi$, $\forall k \geq 1$, $(\hat{\Pi}_k \mu_{\sigma(n)})_{n\geq 1}$ converge étroitement vers une même $\nu_k \in M_1(\mathbb{R}^k)$

Les $(\nu_k)_{k\geq 1}$ étant cohérentes (i.e $\hat{\pi}_k \nu_{k+1} = \nu_k$), d'après le théorème d'extension de Kolmogorov, il existe $\nu \in M_1(\mathbb{R}^{\mathbb{N}})$ tq $\hat{\pi}_k \nu = \nu_k \ \forall k \geq 1$.

Il s'ensuit que $\mu_{\sigma(n)} \Rightarrow \nu$, car dans $\mathbb{R}^{\mathbb{N}}$ la convergence étroite est équivalente à la convergence étroite des marginales fini-dimensionnelles.

- Cas E σ-compact (E = union dénombrables de compacts)

E étant séperable, 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 $\mathcal{B}(E) = \{A \in \mathcal{B}(\mathbb{R}^{\mathbb{N}}) : A \subset E\}$ (exo)

On étend μ_n sur $\mathbb{R}^{\mathbb{N}}$ en posant $\hat{\mu}_n(A) = \mu_n(A \cap E)$ pour $A \in \mathcal{B}(\mathbb{R}^{\mathbb{N}})$, i.e $\hat{\mu}_n$ est la même image de μ_n par l'inclusion $I : E \to \mathbb{R}^{\mathbb{N}}$, continue ($x \mapsto x$).

(a) Alors $(\hat{\mu}_n)_{n\geq 1}$ est tendue comme image de $(\mu_n)_{n\geq 1}$ par I.

Soit φ extraction. D'après l'étape précédente, $\exists \psi$ extraction tq si $\sigma = \varphi \circ \psi$, $(\hat{\mu}_{\sigma(n)})_{n\geq 1}$ converge étroitement vers $\hat{\mu} \in M_1(\mathbb{R}^{\mathbb{N}})$

(b) Vérifions que $\hat{\mu}(E) = 1$. Soit K_i compact de E tq $\mu_{\sigma(n)}(K_i) \geq 1 - \frac{1}{i} \ \forall n \geq 1$.

Puisque K_i est fermé dans $\mathbb{R}^{\mathbb{N}}$ (exo), par Porte-manteau,

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

En faisant $i \to \infty$, on obtient $\hat{\mu}(E) = 1$.

On définit alors μ sur E en posant $\mu(A) = \hat{\mu}(A)$ pour $A \in \mathcal{B}(E)$.

(c) Vérifions que $\mu_{\sigma(n)} \Rightarrow \mu$. On utilise porte-manteau.

Soit O ouvert de E. Alors $O = E \cap \hat{O}$ avec $\hat{O}$ ouvert de $\mathbb{R}^N$ (exo),

et $\liminf_{n\to\infty} \mu_{\sigma(n)}(O) = \liminf_{n\to\infty} \hat{\mu}_{\sigma(n)}(\hat{O}) \geq \hat{\mu}(\hat{O}) = \hat{\mu}(\hat{O} \cap E)$ (car $\hat{\mu}(E)=1$)
$= \mu(O)$

• Cas général Soit K_i compact tq $\forall n \geq 1$, $\mu_n(K_i) \geq 1 - \frac{1}{i}$.

On pose $E_0 = \bigcup_{i=1}^{\infty} K_i$, de sorte que $\mu_n(E_0) = 1 \ \forall n \geq 1$.

On applique le cas précédent à $\tilde{\mu}_n$ défini sur E_0 par $\tilde{\mu}_n(A) = \mu_n(A)$:

(En particulier, comme $\mu_n(E_0)=1$, $\mu_n(A) = \tilde{\mu}_n(A \cap E_0)$ pour $A \in \mathcal{B}(E)$)

($(\tilde{\mu}_n)_{n\geq 1}$ est bien tendue car $\forall i \geq 1$, $\forall n \geq 1$, $\tilde{\mu}_n(K_i) \geq 1 - 1/i$ avec K_i compact)

Soit ψ extraction, $\exists$ φ-extraction tq si $\sigma = \varphi \circ \psi$, alors $\tilde{\mu}_{\sigma(n)} \Rightarrow \tilde{\mu} \in M_1(E_0)$

On définit alors μ par $\mu(A) = \tilde{\mu}(A \cap E_0)$ pour $A \in \mathcal{B}(E)$.

On conclut que $\mu_{\sigma(n)} \Rightarrow \mu$, car $\mu_{\sigma(n)}$ et μ sont les mesures image de $\tilde{\mu}_{\sigma(n)}$ et $\tilde{\mu}$ par l'inclusion $I : E_0 \to E$, continue.
$x \mapsto x$

~

Preuve de (2) On suppose (E,d) polonais et $\forall \psi, \exists \varphi$ tq $\mu_{\phi\circ\varphi(n)}$ cv étroitement

Soit $(x_i)_{i\geq 1}$ une suite dense.

Vérifions d'abord que $\forall k \geq 1, \exists N_k \geq 1, \forall n \geq 1, \ \mu_n\left(\bigcup_{i=1}^{N_k} B(x_i, \frac{1}{k})\right) \geq 1 - \frac{\varepsilon}{2^k}$

Sinon, $\exists k \geq 1$ et une suite $\phi(n)$ tq $\forall n$, $\mu_{\phi(n)}\left(\bigcup_{i=1}^{n} B(x_i, \frac{1}{k})\right) < 1 - \frac{\varepsilon}{2^k}$.

Soit alors φ tq $\mu_{\phi\circ\varphi(n)} \Rightarrow \mu$.

Par Porte-Manteau, pour tout $m \geq 1$

$$1 - \frac{\varepsilon}{2^k} > \liminf_{n\to\infty} \mu_{\phi(n)}\left(\bigcup_{i=1}^{n} B(x_i, \frac{1}{k})\right) \geq \liminf_{n\to\infty} \mu_{\phi(n)}\left(\bigcup_{i=1}^{m} B(x_i, \frac{1}{k})\right)$$
$$\geq \mu\left(\bigcup_{i=1}^{m} B(x_i, \frac{1}{k})\right)$$

Donc en faisant $m \to \infty$, $\mu(E) \leq 1 - \frac{\varepsilon}{2^k}$, absurde.

On pose alors $K = \bigcap_{k \geq 1} \bigcup_{i=1}^{N_k} \overline{B}(a_i, \frac{1}{k})$.

Comme précédent pour montrer qu'une mesure de proba sur un espace polonais est tendue, on voit que K est compact et que $\forall n \geq 1$, $\mu_n(K) \geq 1 - \varepsilon$

$\rightsquigarrow$

Corollaire Soit (E,d) polonais, $(\mu_n)_{n\geq 1}$ suite de $M_1(\mathbb{R})$, $\mu \in M_1(\mathbb{R})$. Alors

$\mu_n \Rightarrow \mu$ ssi $\begin{cases} (\mu_n)_{n\geq 1} \text{ est tendue} \\ \text{Si } \phi \text{ extraction et } \nu \in M_1(\mathbb{R}) \text{ tq } \mu_{\phi(n)} \Rightarrow \nu, \text{ alors } \mu = \nu \end{cases}$

La preuve est la même que celle vue pour $E = \mathbb{R}$ au cours 1.

Reformulation probabiliste : Soit $(X_n)_{n\geq 1}$, X des v.a à valeurs dans un espace polonais. Alors

$X_n \xrightarrow{loi} X$ ssi $\begin{cases} (X_n) \text{ est tendue, i.e } \forall \varepsilon > 0, \exists K_\varepsilon \text{ compact, } \limsup_{n\to\infty} \mathbb{P}(X_n \notin K_\varepsilon) \leq \varepsilon \\ \text{si } X_{\phi(n)} \xrightarrow{loi} Y, \text{ alors } X \overset{loi}{=} Y. \end{cases}$

Application Soit $(X_n)_{n\geq 1}$ et $(Y_n)_{n\geq 1}$ des va à valeurs dans (E,d) et (E',d') polonais. On suppose $X_n \xrightarrow{loi} X$, $Y_n \xrightarrow{loi} Y$ et $X_n \perp\!\!\!\perp Y_n$ $\forall n$.

Alors $(X_n, Y_n) \xrightarrow{loi} (\tilde{X}, \tilde{Y})$ avec $\tilde{X} \overset{loi}{=} X$, $\tilde{Y} \overset{loi}{=} Y$ et $\tilde{X} \perp\!\!\!\perp \tilde{Y}$

Preuve : Étape 1 : $((X_n, Y_n))_{n\geq 1}$ est tendue.

En effet, puisque $X_n \xrightarrow{loi} X$, $Y_n \xrightarrow{loi} Y$, pour $\varepsilon > 0$, $\exists$ compacts K, K' de E, E' tq $\mathbb{P}(X_n \in K) \geq 1 - \varepsilon$ 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 $\forall f \in \mathcal{C}_b(E)$, $\forall g \in \mathcal{C}_b(E')$,

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

$$\downarrow n \to \infty \qquad\qquad \downarrow n \to \infty$$

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

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

$\sim$

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

Soit (E, d) un espace métrique.

Définition 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 d_{LP} est une distance

Preuve : • symétrie : ok. • inégalité triangulaire : provient de $(A^\varepsilon)^\eta \subset A^{\varepsilon+\eta}$

• Séparation : si $d_{LP}(\mu, \nu) = 0$, alors pour tout fermé F,

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

Comme $F = \bigcap_{n \geq 1} F^{\frac{1}{n}}$, en faisant $n \to \infty$ on en déduit $\mu(F) \leq \nu(F)$ et $\nu(F) = \mu(F)$

Donc $\mu = \nu$ sur les fermés Donc $\mu = \nu$

$\sim$

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

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

Preuve : Soit $\varepsilon>0$ tq $\forall A\in\mathcal{B}_E,\ \mu(A)\le\nu(A^\varepsilon)+\varepsilon$. Soit $A\in\mathcal{B}_E$.

On vérifie que $\left({}^c(A^\varepsilon)\right)^\varepsilon \subset {}^cA$, de sorte que

$$\nu(A) = 1-\nu({}^cA) \le 1-\nu\left(\left({}^c(A^\varepsilon)\right)^\varepsilon\right) \le 1-\mu\left({}^c(A^\varepsilon)\right)+\varepsilon = \mu(A^\varepsilon)+\varepsilon$$

La distance de Lévy-Prokhorov est liée à la convergence étroite :

Proposition Soit $(\mu_n)_{n\ge1}$ une suite de $M_1(E)$ et $\mu\in M_1(E)$.

1. $d_{LP}(\mu_n,\mu)\longrightarrow 0$ implique $\mu_n\Rightarrow\mu$
2. Si E est séparable, $\mu_n\Rightarrow\mu$ implique $d_{LP}(\mu_n,\mu)\longrightarrow 0$

Ainsi, lorsque E est séparable, d_{LP} métrise la convergence étroite sur $M_1(E)$

Preuve : 1. On suppose $d_{LP}(\mu_n,\mu)\longrightarrow 0$. Soit F fermé, $\varepsilon>0$. Soit n_0 tq $n\ge n_0$ implique $\mu_n(F)\le\mu(F^\varepsilon)+\varepsilon$. Donc $\limsup_{n\to\infty}\mu_n(F)\le\mu(F^\varepsilon)+\varepsilon$.

En faisant $\varepsilon\to 0$, on en déduit que $\limsup_{n\to\infty}\mu_n(F)\le\mu(F)$

Donc par porte-manteau, $\mu_n\Rightarrow\mu$

2. Soit E séparable et $\mu_n\Rightarrow\mu$.

Soit $\varepsilon>0$. On montre pour n assez grand, $\forall A\in\mathcal{B}_E,\ \mu(A)\le\mu_n(A^{2\varepsilon})+2\varepsilon$

On remarque que pour $A\in\mathcal{B}_E$, comme A^ε est ouvert, par porte-manteau,

$$\mu(A)\le\mu(A^\varepsilon)\le\mu_n(A^\varepsilon)+\varepsilon \quad \text{pour } n \text{ assez grand.}$$

La difficulté est de le faire uniformément en A. Pour ça, on utilise la séparabilité pour se restreindre à un nombre fini de A.

Soit $(x_i)_{i\geq 1}$ une suite dense dans E.

Puisque $E = \bigcup_{i=1}^{\infty} B(x_i, \varepsilon)$, il existe N tq $\mu\left(\bigcup_{i=1}^{N} B(x_i,\varepsilon)\right) \geq 1-\varepsilon$.

Par porte-manteau, il existe n_0 tq $n \geq n_0$ implique

$$\mu_n(O) \geq \mu(O) - \varepsilon \quad \text{pour tout } O \text{ de la forme } B(x_{i_1},\varepsilon) \cup \dots \cup B(x_{i_k},\varepsilon)$$

avec $k \in \mathbb{N}$ et $1 \leq i_1 < i_2 < \dots < i_k \in \mathbb{N}$

Soit $A \in \mathcal{B}_E$ quelconque et posons $O_A = \bigcup_{\substack{1\leq i\leq N \\ B(x_i,\varepsilon)\cap A \neq \emptyset}} B(x_i,\varepsilon)$

Si $x \in A \setminus O_A$, $x \in {}^c\left(\bigcup_{i=1}^{N} B(x_i,\varepsilon)\right)$, donc est dans un ensemble de mesure $\leq \varepsilon$,

donc $\mu(A) \leq \mu(O_A) + \varepsilon \leq \mu_n(O_A) + 2\varepsilon$

Or $O_A \subset A^{2\varepsilon}$. Donc $\mu(A) \leq \mu_n(A^{2\varepsilon}) + 2\varepsilon$.

Donc $d_{LP}(\mu, \mu_n) \leq 2\varepsilon$ d'après le lemme précédent

~

Corollaire du thm de Prokhorov Si (E,d) est compact, alors $(\mathcal{M}_1(E), d_{LP})$ est compact

Preuve : Soit $(\mu_n)_{n\geq 1}$ une suite de $\mathcal{M}_1(E)$. Il suffit de mq $(\mu_n)_{n\geq 1}$ admet une valeur d'adhérence. Comme E est compact, $(\mu_n)_{n\geq 1}$ est tendue

Il existe donc φ extraction et $\mu \in \mathcal{M}_1(E)$ tq $\mu_{\varphi(n)} \Rightarrow \mu$.

Comme E est séparable, $\mu_{\varphi(n)} \to \mu$ dans $(\mathcal{M}_1(E), d_{LP})$.

~

Remarque Lorsque (E,d) est séparable, le théorème de Prokhorov se reformule en : si $A \subset \mathcal{M}_1(E)$, la famille $\{\mu : \mu \in A\}$ est tendue ssi A est relativement compact dans $(\mathcal{M}_1(E), d_{LP})$

Théorème Soit (E,d) un espace métrique

① Si E est séparable, alors $(M_1(E), d_{LP})$ est séparable

② Si E est complet séparable, alors $(M_1(E), d_{LP})$ est complet séparable

Preuve : ① Soit $(x_i)_{i\geq 1}$ une suite dense de E. On note l'ensemble des mesures de la forme $\nu = \sum_{m\in M} r_m \delta_{x_m}$ avec $M \subset \mathbb{N}$, $\#M < \infty$ et (r_m) des rationnels de somme 1 est dense. Soit $\mu \in M_1(E)$

Soit $\varepsilon > 0$, $M \subset \mathbb{N}$, $\#M < \infty$ tq $\mu\left(\bigcup_{m\in M} B(x_m, \varepsilon)\right) \geq 1-\varepsilon$.

En remplaçant les $B(x_m,\varepsilon)$ par des intersections de celles-ci, on écrit

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

On pose alors $\nu = \sum_{m\in M} r_m \delta_{x_m}$, avec (r_m) choisis tq $\sum_{m\in M} |r_m - \mu(B_m)| \leq \varepsilon$

Si $A \in \mathcal{B}(E)$, on pose $\tilde{A} = \bigsqcup_{m\in M,\, B_m \cap A \neq \emptyset} B_m$.

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 $\tilde{A} \subset A^{\varepsilon}$, 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 propriétés des suites de Cauchy). Soit $\varepsilon > 0$, et φ extraction tq

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

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

Alors pour $i \geq \varphi(n)$, $\mu_i\left(K_n^{\varepsilon/2^{n+1}}\right) \geq \mu_{\varphi(n)}(K_n) - \frac{\varepsilon}{2^{n+1}} \geq 1 - \frac{\varepsilon}{2^n}$

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

En posant $K = \bigcap_{n\geq 1} \overline{K_n^{\frac{\varepsilon}{2^{n+1}}}}$, on vérifie que K est compact (car précompact et fermé) et de mesure $\geq 1-\varepsilon$. Ceci conclut.

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, appelés ouverts tel que :

  ① Toute intersection finie d'éléments de $\mathcal{O}$ appartient à $\mathcal{O}$

  ② 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 $\mathcal{B}$, i.e $\forall U\in\mathcal{O}, \forall x\in U, \exists V\in\mathcal{B}, x\in V\subset U$.

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 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 $\mathcal{P}(E)$ tq $\cup\mathcal{B}=E$ et tq

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

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 $\forall U \in \mathcal{O}$ tq $x \in U$, $\exists N$ tq $n \geq N \Rightarrow x_n \in U$

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 $\tau_2 = \{A \subset \mathbb{N} : 1 \notin A$ ou $\sum_{n \in \mathbb{N} \setminus A} \frac{1}{n} < +\infty\}$)

- 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, appelé topologie induite par d. Un espace topologique $(E, \mathcal{O})$ est dit métrisable s'il existe une di

- Les notions de "être à base dénombrable d'ouverts", de "séparabilité", de "convergence", de "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, pas 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 $M_1(E)$ pour laquelle les fonctions $\mu \mapsto \mu(f)$ sont continues pour $f \in C_b(E)$ (i.e la topologie initiale sur $M_1(E)$ définie par les applications $\mu \mapsto \mu(f)$ pour $f \in C_b(E)$), appelée topologie faible-étoile sur $M_1(E)$ (aussi abrégé en faible-$*$)

Lemme La famille de parties

$$B_{(\mu, \varepsilon, f_1, \dots, f_n)} = \{ \nu \in M_1(E) : |\mu(f_i) - \nu(f_i)| < \varepsilon \quad \forall 1 \le i \le n \}$$

pour $\varepsilon > 0$, $f_1, \dots, f_n \in C_b(E)$ et $\mu \in M_1(E)$, constitue une base d'ouverts de la topologie faible-$*$ sur $M_1(E)$

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

~

Proposition 1) Soit $\mu_n, \mu \in M_1(E)$. Alors $\mu_n \Rightarrow \mu$ ssi $\mu_n \to \mu$ au sens de la topologie faible-$*$

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

Preuve : 1) On suppose $\mu_n \Rightarrow \mu$. Soit U ouvert faible-$*$ contenant μ. Il existe alors un ouvert de la forme $B_{(\mu, \varepsilon, f_1, \dots, f_k)}$ inclus dans U.

Puisque $\mu_n \Rightarrow \mu$, $\max_{1 \le i \le k} |\mu_n(f_i) - \mu(f_i)| \to 0$ et donc $\mu_n \in B_{(\mu, \varepsilon, f_1, \dots, f_k)}$ pour n assez grand.

• On suppose $\mu_n \to \mu$ au sens de la topologie faible-$*$. Alors $\forall f \in C_b(E)$, $\forall \varepsilon > 0$, $\mu_n \in B_{(\mu, \varepsilon, f)}$ pour n assez grand, i.e. $|\mu_n(f) - \mu(f)| \le \varepsilon$.

2) • Soit U un ouvert faible-$*$ et $\mu \in U$. Raisonnons par l'absurde en supposant que $\forall \varepsilon > 0$, $B_{d_{LP}}(\mu, \varepsilon) \not\subset U$. Il existe alors une suite $\mu_n \in M_1(E)$ t.q. $\mu_n \xrightarrow{d_{LP}} \mu$ et $\mu_n \notin U$ $\forall n \ge 0$. Mais on a vu que $\mu_n \xrightarrow{d_{LP}} \mu$ implique $\mu_n \Rightarrow \mu$, et par 1) cela implique que $\mu_n \to \mu$ pour la topologie faible-$*$, absurde

• On montre que $\forall \mu \in M_1(E)$, $\forall \varepsilon > 0$, $B_{d_{LP}}(\mu, \varepsilon)$ est un voisinage de μ pour la topologie faible-$*$. La preuve est très proche de celle utilisée pour montrer que $\mu_n \Rightarrow \mu$ implique $d_{LP}(\mu_n, \mu) \to 0$

Soit $(x_i)_{i \ge 1}$ dense dans E et N t.q. $\mu\left(\bigcup_{i \ge N} B(x_i, \varepsilon)\right) < \varepsilon$.

Pour $I \subset \{1, 2, \dots, N\}$ on note $B_I = \bigcup_{i \in I} B(x_i, \varepsilon)$ et on pose

$$f_I(x) = \max\left(1 - \frac{1}{\varepsilon} d(x, B_I), 0\right)$$

de sorte que $\mathbb{1}_{B_I} \le f_I \le \mathbb{1}_{B_I^\varepsilon}$.

Soit alors $V = \left\{ \nu \in M_1(E) : |\nu(f_I) - \mu(f_I)| < \varepsilon \ \forall I \subset \{1, 2, \dots, N\} \right\}$, voisinage de μ pour la topologie faible-$*$.

Montrons que $V \subset B_{d_{LP}}(\mu, 2\varepsilon)$. Soit $A \in \mathcal{B}(E)$ et posons

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

Comme $(B_{I_A})^\varepsilon \subseteq A^{2\varepsilon}$, on a

$$
\begin{aligned}
\mu(A) &\leq \mu(B_{I_A}) + \mu(A \setminus B_{I_A}) \\
&\leq \mu(B_{I_A}) + \mu\Big(\bigcup_{i>N} B(x_i, \varepsilon)\Big) \\
&\leq \mu(\delta_{I_A}) + \varepsilon \\
&\leq \nu(\delta_{I_A}) + 2\varepsilon \\
&\leq \nu((B_{I_A})^{\varepsilon}) + 2\varepsilon \\
&\leq \nu(A^{2\varepsilon}) + 2\varepsilon,
\end{aligned}
$$

d'où le résultat

$\sim$

Lorsque E est un espace topologique métrisable et que $M_1(E)$ est muni de la topologie faible-$*$, il est possible de montrer que

E compact $\Leftrightarrow$ $M_1(E)$ compact

E séparable $\Leftrightarrow$ $M_1(E)$ séparable

E polonais $\Leftrightarrow$ $M_1(E)$ polonais

IV/ L'espace des fonctions continues sur un compact

Plan : 1) Caractère polonais
2) Marginales fini-dimensionnelles
3) Convergence en loi dans cet espace
4) Compacité dans cet espace (Arzelà-Ascoli) et tension
5) Tension
6) Critère de tension de Kolmogorov
7) Fonctions continues sur $\mathbb{R}_+$

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

1) Caractère polonais

Proposition Soit (K,ρ) un espace métrique compact et (E,d) un espace polonais
Alors $(C(K,E), d)$ est polonais.

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

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

$$C_{m,n} = \left\{ f \in C(K,E) : \forall x,y \in E,\ \rho(x,y) < \frac{1}{m} \Rightarrow d(f(x),f(y)) < \frac{1}{n} \right\}$$

Puisque K est compact, pour tout $m \geq 1$, il existe un ensemble fini K_m tq $K = \bigcup_{x \in K_m} B(x, \frac{1}{2m})$. 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 $< \frac{1}{n}$

Pour un k_m-uplet $(i_1, \dots, i_{k_m})$ d'entiers ≥ 1 on note

$$F^{m,n}_{(i_1,\dots,i_{k_m})} = \{ g \in C_{m,n} : g(x_j^m) \in L^n_{i_j} \text{ pour tout } 1 \leq j \leq k_m \}$$

Lorsque $F^{m,n}_{(i_1,\dots,i_{k_m})} \neq \emptyset$, on choisit une fonction $g^{m,n}_{(i_1,\dots,i_{k_m})} \in F^{m,n}_{(i_1,\dots,i_{k_m})}$

Soit $D_{m,n}$ l'ensemble de telles fonctions.

Vérifions que $\forall g \in C_{m,n}$, $\exists\, g \in D_{m,n}$ tq $\forall 1 \leq j \leq k_m$, $d(g(x_j^m), g(x_j^n)) \leq \frac{1}{n}$

Soit $g \in C_{m,n}$. On observe que $C_{m,n} = \bigcup_{i_1,\dots,i_{k_m} \geq 1} F^{m,n}_{(i_1,\dots,i_{k_m})}$

Choisissons $i_1, \dots, i_{k_m}$ tels que $g(x_j^m) \in L^n_{i_j}$ pour $1 \leq j \leq k_m$. En particulier

En particulier $F^{m,n}_{(i_1,\dots,i_{k_m})} \neq \emptyset$. On peut donc prendre

$g = g^{m,n}_{(i_1,\dots,i_{k_m})}$ qui convient (car $\operatorname{diam}(L^n_i) < \frac{1}{n}$, $\forall i$).

Soit maintenant $D = \bigcup_{m,n} D_{m,n}$, dénombrable. Vérifions que D est dense dans $C(K,E)$. Soit $g \in C(K,E)$ et $\varepsilon > 0$. Fixons $n > 3/\varepsilon$. Puisque g est uniformément continue, $\exists m \geq 1$ tq $g \in C_{m,n}$. Soit alors $g \in D_{m,n}$ tq $d(g(x_j^m), g(x_j^m)) < \frac{1}{n}$ $\forall 1 \leq j \leq k$.

Par inégalité triangulaire, on déduit $\forall x \in K$, avec x_j^m tq $x \in B(x_j^m, \frac{1}{2m})$,

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

2) Marginales fini-dimensionnelles

a) 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 $\mathbb{R}^{\mathbb{N}}$ dans le cours 2)

Définition : 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 $\pi_t : E^I \longrightarrow (E, \mathcal{B}(E))$ mesurables
$(x_i)_{i\in I} \longmapsto x_t$
Elle est parfois appelée tribu cylindrique

Remarques 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 $(\Omega, \mathcal{F}) \to (E^I, \mathcal{B}(E)^{\otimes I})$ est mesurable ssi $\forall i \in I$ $\omega \mapsto x_i(\omega)$
$\omega \longmapsto (x_i(\omega))_{i\in I}$
est mesurable. En particulier $(X_i)_{i\in I}$ est une v.a (à valeurs dans E^I) ssi $\forall i \in I$, X_i est une v.a.

Définition Soit $\mu \in M_1(E^I)$. 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 $E^I \longrightarrow E^k$. Si $(X_i)_{i\in I}$ a pour loi μ,
$(x_i)_{i\in I} \mapsto (x_{i_1}, \dots, x_{i_k})$
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 (fdfdi en abrégé)

Proposition Une mesure $\mu \in M_1(E^I)$ est caractérisée par ses marginales fdfdi
En particulier, la loi d'une v.a. $(X_i)_{i\in I} \in E^I$ est caractérisée par ses lois fdfdi

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

$\sim$

L'avantage de ce point de vue est l'existence de v.a :

Théorème d'extension de Kolmogorov On se donne sur E^I une famille de marginales fini-dimensionnelles $\{\mu_{i_1,\dots,i_k}\}$ cohérentes, au sens où si $\mathcal{L}((X_{i_1},\dots,X_{i_k})) = \mu_{i_1,\dots,i_k}$, alors $\forall 1 \le j \le k, \mathcal{L}((X_{i_1},\dots,\widehat{X_{i_j}},\dots X_{i_k})) = \mu_{i_1,\dots,\widehat{i_j},\dots,i_k}$

Alors il existe $\mu \in \mathcal{M}_1(E^I)$ dont les marginales fidi sont celles ci-dessus

le $\widehat{\ }$ signifie qu'on enlève le terme.

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

Pb en général E^I n'est pas polonais. Le cadre n'est pas bien adapté à la convergence étroite !

b) L'espace des fonctions continues

Notons ici $C = C(K,E)$ 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 $\mathcal{B}(E)^{\otimes K}$, i.e $\mathcal{P}_C = \{A \cap C : A \in \mathcal{B}(E)^{\otimes K}\}$. De manière équivalente, c'est la plus petite tribu sur C qui rend toutes les projections $\pi_t : \begin{array}{ccc} C & \longrightarrow & E \\ g & \longmapsto & g(t) \end{array}$ mesurables [exo]

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

Preuve : ⊛ Puisque $\forall x \in K$, $\pi_x : C \to E$ 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 $F : \begin{array}{ccc} C & \to & \mathbb{R} \\ f & \mapsto & d(f, f_0) \end{array}$. Puisque K est compact, il existe $D \subset K$ dénombrable dense.

Par continuité, $d(f, f_0) = \sup_{x \in D} d(f(x), f_0(x))$, et puisque $\forall x \in D$, $\begin{array}{ccc} (C, \mathcal{P}_C) & \to & \mathbb{R} \\ f & \mapsto & d(f(x), f_0(x)) \end{array}$ est mesurable (c'est la composée $\phi \circ \pi_x$ avec $\phi : \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ y & \mapsto & d(y, f_0(x)) \end{array}$ continue donc mesurable), $F : (C, \mathcal{P}_C) \to \mathbb{R}$ est mesurable comme le sup d'une famille dénombrable d'applications mesurables. Ainsi, $\forall r > 0$, $F^{-1}([0, r[) \in \mathcal{P}_C$, i.e $\{f \in C : d(f, f_0) < r\} \in \mathcal{P}_C$

Étape 2 : Par séparabilité de C, tout ouvert de C est union dénombrable de boules ouvertes $[\varepsilon, r[$. Cela conclut.

$\sim$

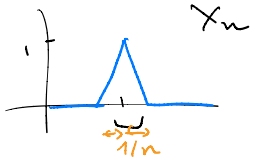
D'après ce qui précède, les marginales fidi $\{\mu_{t_1, \dots, t_k} ; t_1, \dots, t_k \in K\}$ d'une mesure $\mu \in \mathcal{M}_1(C)$ la caractérisent.

⚠ L'ensemble $\{f \in \mathbb{R}^{[0,1]} : f \text{ est continue}\}$ n'est pas mesurable par rapport à $\mathcal{B}(\mathbb{R})^{\otimes [0,1]}$ $[\varepsilon, r[$

3) Convergence en loi dans cet espace

Définition Soit $\mu_n, \mu \in M_1(C)$. On dit que μ_n converge vers μ au sens des fidi si $\forall t_1,\dots,t_k \in K$, $(\mu_n)_{t_1,\dots,t_k} \Rightarrow \mu_{t_1,\dots,t_k}$. On note $\mu_n \xrightarrow{f.d.} \mu$. De même, une suite $(X_n)_{n\geq 1}$ de v.a. à valeurs dans C converge vers X au sens des fidi si $\forall t_1,\dots,t_k \in K$ $(X_n(t_1),\dots,X_n(t_k)) \xrightarrow{\text{loi}} (X(t_1),\dots,X(t_k))$.

! 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 fausse : Soit U uniforme sur $[0,1]$, $X_n(t) = \max(0, 1-n|U-t|)$ pour $0 \leq t \leq 1$



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

Cependant, on a le résultat suivant :

Théorème Soit $\mu_n, \mu \in M_1(C)$.

1. $\mu_n \Rightarrow \mu$ ssi $\mu_n \xrightarrow{f.d.} \mu$ 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.

Preuve : 1. $\Rightarrow$ Ok par continuité des projections et le thm de Prokhorov.

$\Leftarrow$ D'après le thm de Prokhorov, il suffit de montrer que si $\mu_{\phi(n)} \Rightarrow \nu$ alors $\mu = \nu$. Mais alors $\mu_{\phi(n)} \xrightarrow{f.d.} \nu$. Donc μ et ν ont les mêmes fidi, donc sont égales

(2) Par tension, $\exists$ extraction et $\mu \in M_1(C)$ tq $\mu_{\phi(n)} \Rightarrow \mu$.

Alors $\mu_{\phi(n)} \xrightarrow{fd} \mu$ et donc $\mu_n \xrightarrow{fd} \mu$. Donc $\mu_n \Rightarrow \mu$ par (1)

Ainsi, dans $C([0,1], E)$, $(X_n(t))_{0\le t\le 1} \xrightarrow{loi} (X(t))_{0\le t\le 1}$ ssi $\forall 0\le t_1 < \dots < t_k \le 1$ on a $(X_n(t_1), \dots, X_n(t_k)) \xrightarrow{loi} (X(t_1), \dots, X(t_k))$ et $(X_n)_{n\ge 1}$ est tendue.

Remarque Étant donnée une famille $\tilde{\mathcal{K}}$ de fidi cohérentes, d'après le thm d'extension de Kolmogorov, il existe $\mu \in M_1(\mathbb{R}^{[0,1]})$ dont les fidi sont $\tilde{\mathcal{K}}$. Mais rien ne garantit qu'il existe une telle $\mu \in M_1(C)$.

D'après le théorème précédent, c'est le cas si les fidis $\tilde{\mathcal{K}}$ sont limites fidi d'une famille tendue de $M_1(C)$.

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

4) Compacité dans C

La notion d'équicontinuité est fondamentale.

Définition Une famille $\mathcal{K} \subset C(X,Y)$ est dite équicontinue si

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x,y \in X,\ \forall f \in \mathcal{K},\ d_X(x,y) \le \eta \Rightarrow d_Y(f(x), f(y)) \le \varepsilon$$

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

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

le module de continuité de f.

c'est équivalent à $\forall \varepsilon > 0,\ \exists \eta > 0,\ \sup_{f\in\mathcal{K}} \omega(f,\eta) \le \varepsilon$

Remarques Pour K compact :

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

Théorème (Arzela-Ascoli) Soit (K,d) compact et (E,d) un espace métrique complet. Soit $\mathcal{F} \subset \mathcal{C}(K,E)$. Alors $\mathcal{F}$ est relativement compacte dans $\mathcal{C}(K,E)$ ssi

- $\mathcal{F}$ est équicontinue
- $\exists D \subset K$ dénombrable dense tq $\forall x \in D$, $\mathcal{F}(x) = \{f(x), f \in \mathcal{F}\}$ est relativement compact dans E

Preuve $\boxed{\Leftarrow}$ Soit $(f_n)_{n \geq 1}$ une suite de $\mathcal{F}$. On trouve une sous-suite qui converge uniformément. Ecrivons $D = \{a_n : n \geq 1\}$. Puisque $(f_n(a_k))_{n \geq 1} \in \overline{\mathcal{F}(a_k)}$, compact dans E, il existe par procédé diagonal une extraction φ telle que $(f_{\varphi(n)}(a_k))_{n \geq 1}$ converge $\forall k \geq 1$.

Posons $g_n(x) = f_{\varphi(n)}(x)$ pour simplifier les notations.

- Vérifions que $\forall x \in K$, $(g_n(x))_{n \geq 1}$ converge.

Comme E est complet, il suffit de montrer qu'elle est de Cauchy.

Pour cela, on fixe $\varepsilon > 0$ et $\eta > 0$ tq $\sup_{f \in \mathcal{F}} \omega(f,\eta) \leq \varepsilon$.

Par compacité, il existe $L \geq 1$ tq $K = \bigcup_{i=1}^{L} B(a_i, \eta)$.

Pour $1 \leq i \leq L$, puisque $(g_n(a_i))_{n \geq 1}$ converge, elle est de Cauchy.

Il existe donc N tq $p,q \geq N$, $1 \leq k \leq L \Rightarrow d_E(g_p(a_k), g_q(a_k)) \leq \varepsilon$

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

$$d_E(g_p(x), g_q(x)) \leq d_E(g_p(x), g_p(a_k)) + d_E(g_p(a_k), g_q(a_k)) + d_E(g_q(a_k), g_q(x))$$

$$\leq 3\varepsilon.$$

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

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

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

$(\Rightarrow)$ • Soit $x \in K$. Montrons que $\overline{\mathcal{F}(x)}$ est compact

Si l'application $\Pi_x : \begin{array}{ccc} C(K,E) & \to & E \\ f & \mapsto & f(x) \end{array}$ 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 compacité de $\overline{\mathcal{F}}$, il existe $f_1, \dots, f_N \in \mathcal{F}$ tq $\mathcal{F} \subset \bigcup_{k=1}^{N} B(f_k, \varepsilon)$.

Par uniforme continuité, on peut trouver $\eta > 0$ tq

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

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

$\sim$

Remarque : Si dans E les boules fermées sont compactes (comme dans $\mathbb{R}^k$) et si K est convexe, on a l'amélioration :

$\mathcal{F}$ relativement compacte $(\Leftrightarrow)$ • $\mathcal{F}$ équicontinue
• $\mathcal{F}(x_0)$ relativement compact pour un $x_0 \in K$.

En effet, on montre alors que $\mathcal{F}(x)$ est dans une boule fermée pour x dans une boule autour de x_0, puis de proche en proche pour tout $x \in K$.

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

Voyons maintenant une application probabiliste

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

① $\forall \varepsilon>0, \forall \eta>0, \exists \delta>0$ tq $\limsup_{n\to\infty} \mathbb{P}(w(X_n,\delta)\geq\eta)\leq\varepsilon$

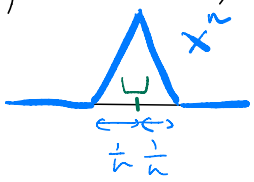
② $\forall t\in[0,1]\cap\mathbb{Q}$, $(X^n(t))_{n\geq 1}$ est tendue

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

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

Exemple : 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 : $\Rightarrow$ Soit $\varepsilon>0$ et K un compact de $\mathcal{C}$ tel que $\forall n\geq 1$, $\mathbb{P}(X^n\notin K)\leq\varepsilon$. Soit $\eta>0$.

① Par Arzela-Ascoli, on peut trouver δ tel que $\sup_{g\in K} w(g,\delta)<\eta$.

On a alors $\mathbb{P}(w(X_n,\delta)\geq\eta)\leq\mathbb{P}(X^n\notin K)\leq\varepsilon$

② Soit $t\in[0,1]$. Il existe A compact de E tel que $g(t)\in A$ $\forall g\in K$. 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 A_i des compacts de E tq $\forall n\geq 1, \mathbb{P}(X^n(t_i)\notin A_i)\leq\frac{\varepsilon}{2^{i+1}}$, et pour tout $k\geq 1$ on peut trouver $\delta_k>0$ tel que $\forall n\geq 1$, $\mathbb{P}\left(w(X^n,\delta_k)>\frac{1}{2^k}\right)\leq\frac{\varepsilon}{2^k}$.

En posant $K_\varepsilon=\{g\in\mathcal{C} : \forall i\geq 1, g(t_i)\in A_i, \forall k\geq 1, w(g,\delta_k)\leq\frac{1}{2^k}\}$, on voit que $\mathbb{P}(X^n\notin K_\varepsilon)\leq\varepsilon$, que K_ε est fermé ($g\mapsto w(g,\delta_k)$ est continue) et que K_ε est compact par Arzela-Ascoli

Remarque ① est équivalent à $\forall$ suite $\delta_k \to 0$, $\forall$ suite (n_k), $w(X_{n_k}, \delta_k) \xrightarrow{(\mathbb{P})} 0$
[exo]

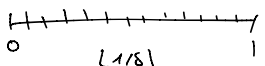
Corollaire: Soit $(X^n)_{n\geq 1}$ des variables aléatoires dans $\mathcal{C}$ telles que

① $\forall \varepsilon > 0$, $\forall \eta > 0$, $\exists \delta < 1$, $\forall t \in [0,1]$, $\limsup_{n\to\infty} \frac{1}{\delta} \mathbb{P}\left(\sup_{t \leq s \leq (t+\delta)\wedge 1} d(X_s^n, X_t^n) > \eta \right) \leq \varepsilon$

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

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

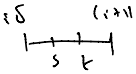
Preuve : On vérifie la condition ① du théorème. Pour cela, on découpe $[0,1]$ en intervalles de longueur δ.



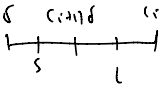
Soit $f \in \mathcal{C}$. Vérifions que $\{ w(f,\delta) > 3\eta \} \subset \bigcup_{i=0}^{[1/\delta]} \left\{ \sup_{i\delta \leq s \leq (i+1)\delta\wedge 1} d(f(s), f(i\delta)) \geq \eta \right\}$.

On suppose que $\forall\, 0 \leq i \leq [1/\delta]$ on a $\sup_{i\delta \leq s \leq (i+1)\delta\wedge 1} d(f(s), f(i\delta)) < \eta$.

Si $|s-t| < \delta$:

Cas 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 

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

Ainsi, $\mathbb{P}(w(X^n, \delta) > 3\eta) \leq ([1/\delta]+1) \max_{t \in \{0, \delta, 2\delta, \dots, [1/\delta]\delta\}} \mathbb{P}\left(\sup_{t \leq s \leq (t+\delta)\wedge 1} d(X_s^n, X_t^n) > \eta \right)$

$\leq 2\varepsilon$ pour n assez grand, car $[1/\delta] + 1 \leq \frac{1}{\delta} + 1 \leq \frac{2}{\delta}$.

$\sim$

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

Notations • Soit $D_n = \{\frac{i}{2^n} : 0 \le i \le 2^n\}$ les dyadiques de niveau n dans $[0,1]$ et $D = \bigcup_{n \ge 1} D_n$

• Pour $x = (x_t)_{0 \le t \le 1} \in E^{[0,1]}$ 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-Hölder, $M_\gamma(x) < \infty$.

Théorème (critère de Kolmogorov)

(1) Soit $X = (X_t)_{0 \le t \le 1}$ une v.a à valeurs dans $E^{[0,1]}$. On suppose que

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

Alors pour tout $0 < \gamma < \frac{\beta}{\alpha}$, $\mathbb{E}[M_\gamma(x)^\alpha] \le C(\alpha, \beta, \gamma, c)$

(2) Soit $(X^n)_{n \ge 1}$ des v.a à valeurs dans $\mathcal{C}$. On suppose que

(a) $\forall t \in [0,1]$, $(X^n_t)_{n \ge 1}$ est tendue

(b) $\exists \alpha, \beta, c > 0$ tq $\forall n \ge 1$, $\forall s,t \in [0,1]$, $\mathbb{E}[d(X^n_s, X^n_t)^\alpha] \le c\,|s-t|^{1+\beta}$

Alors $(X^n)_{n \ge 1}$ est tendue, et toute limite de sous-suite est p.s γ Hölder pour tout $0 < \gamma < \frac{\beta}{\alpha}$.

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 inégalité triangulaire le long d'un chemin de dyadiques "consécutifs".

Notons $\Delta_k = \max\{d(X_s, X_t) : s,t \in D_k,\ |s-t| = \frac{1}{2^k}\}$.

Lemme $M_\gamma(x) \le 2^{\gamma+1} \cdot \left(\sum_{k=1}^{\infty} 2^{k\gamma} \Delta_k\right)$

Preuve du lemme : On commence par écrire

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

Le point clé est de remarquer que pour tout $m\geq 1$,

$$\sup_{\substack{|s-t|\leq\frac{1}{2^m} \\ s,t\in D}} d(X_s,X_t) \leq 2\sum_{k=m}^{\infty}\Delta_k \quad (\star\star\star)$$

En effet, si $|s-t|\leq\frac{1}{2^m}$ et $s\leq t$, notons s_m le plus grand élément de D_m $\leq$ à s et t_m le plus grand élément de D_m $\leq$ à t. On a soit $s_m=t_m$, soit $t_m=s_m+\frac{1}{2^m}$

On peut alors écrire pour $\ell\geq 1$ et $a_1,\dots,a_r\in\{0,1\}$ (dépendants de s)

$$s_m=\frac{\ell}{2^m} \quad \text{et} \quad s=\frac{\ell}{2^m}+\frac{a_1}{2^{m+1}}+\dots+\frac{a_r}{2^{m+r}}.$$

De sorte qu'en posant $a_0=0$,

$$\begin{aligned} d(X_{s_m},X_s) &\leq \sum_{i=1}^{r} d\left(X_{\frac{\ell}{2^m}+\frac{a_1}{2^{m+1}}+\dots+\frac{a_i}{2^{m+i}}}, X_{\frac{\ell}{2^m}+\frac{a_1}{2^{m+1}}+\dots+\frac{a_{i-1}}{2^{m+i-1}}}\right) \\ &\leq \sum_{i=1}^{r}\Delta_{m+i} \\ &\leq \sum_{k=m+1}^{\infty}\Delta_k. \end{aligned}$$

De même $d(X_{t_m},X_t)\leq\sum_{k=m+1}^{\infty}\Delta_k$.

Ainsi, $d(X_s,X_t)\leq \underbrace{d(X_{s_m},X_{t_m})}_{\leq\Delta_m}+2\sum_{k=m+1}^{\infty}\Delta_k \leq 2\sum_{k=m}^{+\infty}\Delta_k$,

ce qui montre $(\star\star\star)$

En injectant $(\star\star\star)$ dans $(\star\star)$ on obtient

$$\begin{aligned} M_\gamma(X) &\leq \sup_{m\geq 1} 2^{(m+1)\gamma+1}\left(\sum_{k=m}^{+\infty}\Delta_k\right) \\ &\leq 2^{\gamma+1}\sup_{m\geq 1}\left(\sum_{k=m}^{\infty}\Delta_k 2^{k\gamma}\right) = 2^{\gamma+1}\left(\sum_{k=1}^{\infty}2^{k\gamma}\Delta_k\right), \end{aligned}$$

ce qui démontre le lemme. ∎

Preuve du Théorème ① D'après le lemme,

$$\mathbb{E}[M_\gamma(X)^\alpha] \le 2^{\gamma\alpha+\alpha}\,\mathbb{E}\Big[\Big(\sum_{k=1}^{\infty} 2^{k\gamma}\,\Delta_k\Big)^\alpha\Big]$$

D'après l'inégalité de Minkowski, on en déduit que

$$\mathbb{E}\Big[\Big(\sum_{k=1}^{+\infty} 2^{k\gamma}\Delta_k\Big)^\alpha\Big]^{\frac{1}{\alpha\vee 1}} \le \sum_{k=1}^{+\infty} 2^{k\gamma\frac{\alpha}{\alpha\vee 1}}\,\mathbb{E}[\Delta_k^\alpha]^{\frac{1}{\alpha\vee 1}}$$

Puisque $\#\{(s,t) : t-s=\frac{1}{2^k},\ s,t\in D_K\} = 2^k$, on a par hypothèse (*)

$$\mathbb{E}[\Delta_k^\alpha] \le 2^k\cdot c\cdot\Big(\frac{1}{2^k}\Big)^{1+\beta} = \frac{c}{2^{k\beta}}$$

Ainsi $\mathbb{E}[M_\gamma(X)^\alpha]^{\frac{1}{\alpha\vee 1}} \le 2^{\frac{\gamma\alpha+\alpha}{\alpha\vee 1}} \sum_{k=1}^{+\infty} c^{\frac{1}{\alpha\vee 1}}\, 2^{\frac{k}{\alpha\vee 1}(\alpha\gamma-\beta)} < +\infty$ car $\gamma < \beta/\alpha$

Ainsi, on a bien $\mathbb{E}[M_\gamma(X)^\alpha] \le C(\alpha,\beta,\gamma,c)$.

② Pour la réunion, il suffit de vérifier que $\forall \varepsilon>0, \forall \eta>0, \exists \delta>0,\ \mathbb{P}(w(X^n,\delta)\ge\eta)\le\varepsilon$.

Par définition de $M_\gamma(X^n)$, $\forall s,t\in D$ on a $d(X^n_s, X^n_t)\le M_\gamma(X^n)|s-t|^\gamma$. Par continuité, c'est aussi vrai $\forall s,t\in[0,1]$

Ainsi, si $|s-t|\le\delta$, on a $d(X^n_s, X^n_t)\le M_\gamma(X^n)\delta^\gamma$ de sorte que $w(X^n,\delta)\le M_\gamma(X^n)\cdot\delta^\gamma$.

Par inégalité de Markov, $\mathbb{P}(w(X^n,\delta)\ge\eta) \le \frac{1}{\eta^\alpha}\delta^{\gamma\alpha}\,\mathbb{E}[M_\gamma(X^n)^\alpha]$

On en déduit le résultat par ①.

- Enfin, supposons $X^{\varphi(n)} \xrightarrow{\text{loi}} X$ dans $\mathcal{C}$ et montrons que X est p.s γ-Hölder.

Il suffit de vérifier que X vérifie (*). En effet, par ① on aura alors $M_\gamma(X)<\infty$ p.s et X étant continu, cela impliquera que X est p.s γ-Hölder.

Pour cela, on utilise le fait que si $Y_n \xrightarrow{\text{loi}} Y$ avec $Y_n, Y\ge 0$, alors $\mathbb{E}[Y] \le \liminf_{n\to\infty} \mathbb{E}[Y_n]$.

En effet, on a $\mathbb{P}(Y_n\ge u)\to\mathbb{P}(Y\ge u)$ pour presque tout u par rapport à la mesure de Lebesgue, donc d'après le lemme de Fatou, $\mathbb{E}[Y] = \int_0^\infty \mathbb{P}(Y\ge u)\,du = \int_0^{+\infty} \lim_{n\to\infty} \mathbb{P}(Y_n\ge u)\,du \le \liminf_{n\to\infty} \int_0^\infty \mathbb{P}(Y_n\ge u)\,du = \liminf_{n\to\infty} \mathbb{E}[Y_n]$

On applique ceci à $d(X^n_s, X^n_t)^\alpha \xrightarrow{\text{loi}} d(X_s, X_t)^\alpha$:

$$\mathbb{E}[d(X_s, X_t)^\alpha] \leq \liminf_{n\to\infty} \mathbb{E}[d(X^n_s, X^n_t)^\alpha] \leq c|s-t|^{1+\beta}$$

$\sim$

Remarque le 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 $\forall 0<\gamma<\frac{\beta}{\alpha}$ il existe un processus $(\tilde{X}_t)_{0\leq t\leq 1}$ γ-Hölder tq $\forall t\in[0,1]$, $\mathbb{P}(X_t = \tilde{X}_t) = 1$.

En effet, puisque $\mathbb{E}[M_\gamma(X)^\alpha] < \infty$, il existe A de proba 1 tq $\forall \omega\in A$, X est γ-Hölder sur D. Pour $\omega\in A$, on peut donc prolonger X en une fonction γ-Hölder notée $\tilde{X}$. Pour $\omega\notin A$, on pose $\tilde{X}=0$.

Vérifions que pour $t\in[0,1]$, $\mathbb{P}(X_t = \tilde{X}_t) = 1$. Par (*) et l'inégalité de Markov si $s_n \to t$ et $s_n \in D$, $X_{s_n} \xrightarrow{\mathbb{P}} X_t$. D'autre part, par construction, $X_{s_n} \xrightarrow{p.s.} \tilde{X}_t$. Donc $X_{s_n} \xrightarrow{\mathbb{P}} \tilde{X}_t$. Donc $\mathbb{P}(X_t = \tilde{X}_t) = 1$.

Application : Soit $(X_i)_{i\geq 1}$ des variables aléatoires réelles iid tq $\mathbb{E}[X_1]=0$ et $\mathbb{E}[X_1^4]<\infty$

On pose $S_0=0$, $S_n = X_1+\cdots+X_n$ pour $n\geq 1$ et $S_t = (1-\{t\})S_{\lfloor t\rfloor} + \{t\}S_{\lfloor t\rfloor+1}$ pour $t\geq 0$. Enfin on définit $S^{(n)}_t = \frac{S_{nt}}{\sqrt{n}}$ pour $0\leq t\leq 1$ de sorte que $S^{(n)} \in \mathcal{C} = \mathcal{C}([0,1],\mathbb{R})$.

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

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

Preuve : On va appliquer le critère de Kolmogorov. Tout d'abord, d'après le TCL, $\forall t\in[0,1]$, $(S^{(n)}_t)_{n\geq 1}$ converge en loi donc est tendue.

Vérifions tout d'abord que $\exists C>0$, $\forall n \geq 1$, $\mathbb{E}[S_n^4] \leq C \cdot n^2$. (★)

Pour cela, on écrit $\mathbb{E}[S_n^4] = \sum_{1 \leq i_1, \dots, i_4 \leq n} \mathbb{E}[X_{i_1} X_{i_2} X_{i_3} X_{i_4}]$

Or $\mathbb{E}[X_{i_1} \cdots X_{i_4}] = 0$ si l'un des indices apparaît une seule fois

Ainsi $\mathbb{E}[S_n^4] = \sum_{i=1}^{n} \mathbb{E}[X_1^4] + \sum_{1 \leq i < j \leq n} \sum_{\substack{i_1,i_2,i_3,i_4 : \\ \#\{k : i_k = i\} = 2 \\ \#\{k : i_k = j\} = 2}} \mathbb{E}[X_i^2]^2$ quantité qui ne dépend pas de n

Donc $\mathbb{E}[S_n^4] \leq Cn + C'n^2 \leq C'' \cdot n^2$.

Ensuite, fixons $s, t \in [0,1]$, $s<t$.

- Si $ns, nt \in \mathbb{N}$, on écrit $\mathbb{E}[|S_{nt} - S_{ns}|^4] = \mathbb{E}[S_{n(t-s)}^4] \leq C(nt-ns)^2$ (★★)
- Si $\lfloor ns \rfloor = \lfloor nt \rfloor = k$, $S_{nt} - S_{ns} = n(t-s)(S_{k+1} - S_k)$, donc

$$\mathbb{E}[(S_{nt} - S_{ns})^4] = n^4 \underbrace{(t-s)^2}_{\leq \frac{1}{n}} (t-s)^2 \mathbb{E}[X_1^4] \leq C(nt-ns)^2$$

- Si $\lfloor ns \rfloor \neq \lfloor nt \rfloor$, écrivons $\lfloor ns \rfloor = k$, $\lfloor nt \rfloor = \ell$ avec $k < \ell$.

Alors en écrivant $|S_{nt} - S_{ns}| \leq |S_{nt} - S_\ell| + |S_\ell - S_{k+1}| + |S_{k+1} - S_{ns}|$, avec

l'inégalité de Minkowski on obtient avec (★) et (★★) :

$$\mathbb{E}[(S_{nt} - S_{ns})^4]^{\frac{1}{4}} \leq C^{\frac{1}{4}} \left(\underbrace{(nt-\ell)}_{\leq nt-ns}^{\frac{1}{2}} + \underbrace{(\ell-(k+1))}_{\leq nt-ns}^{\frac{1}{2}} + \underbrace{(k+1-ns)}_{\leq nt-ns}^{\frac{1}{2}} \right) \leq 3C^{\frac{1}{4}} \sqrt{n} \sqrt{t-s}$$

Ainsi, $\mathbb{E}[(S_s^{(n)} - S_t^{(n)})^4] = \mathbb{E}\left[\left(\frac{S_{nt} - S_{ns}}{\sqrt{n}}\right)^4\right] \leq C(t-s)^2$ pour $0 \leq s,t \leq 1$.

Le critère de Kolmogorov s'applique pour tout $\gamma \in]0, \frac{1}{4}[$.

~

7) Fonctions continues sur $\mathbb{R}_+$

En pratique, les processus aléatoires sont souvent indicés par $\mathbb{R}_+$ plutôt que par $[0,1]$.

Comme l'espace $\mathcal{C}(\mathbb{R}_+, E)$ la distance uniforme $D(f,g) = \sup_{t \geq 0} d(f(t), g(t))$ peut valoir $+\infty$, il faut faire attention.

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

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

Remarque facultative En termes topologiques, c'est la topologie compacte-ouverte sur $\mathcal{C}(\mathbb{R}_+, E)$, dont une base de voisinages 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 :

- $f_n \to f$ dans $\mathcal{C}(\mathbb{R}_+, E)$ ssi $f_n \to f$ uniformément sur les compacts
- $\mathcal{C}(\mathbb{R}_+, E)$ est polonais
- Les marginales fini-dimensionnelles caractérisent les mesures de $\mathcal{M}_1(\mathcal{C}(\mathbb{R}_+, E))$ car la tribu borélienne sur $\mathcal{C}(\mathbb{R}_+, E)$ coincide 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 exo)

Proposition • Une partie $\mathcal{F}$ de $\mathcal{C}(\mathbb{R}_+, E)$ est relativement compacte ssi $\forall N \ge 1$ 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 $(X^n)_{n \ge 1}$ de v.a. de $\mathcal{C}(\mathbb{R}_+, E)$ est tendue ssi $\forall N \ge 1$ la suite des restrictions $(X^n_{|[0,N]})_{n \ge 1}$ 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.

Intermède : couplages

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

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

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

Preuve : On construit un couplage dans lequel $\mathrm{Bin}(n,p) \leq \mathrm{Bin}(n,q)$ en considérant $(U_i)_{1 \leq i \leq n}$ iid sur $[0,1]$ et en posant $X' = \sum_{i=1}^{n} \mathbb{1}_{U_i \leq p}$, $Y' = \sum_{i=1}^{n} \mathbb{1}_{U_i \leq q}$.

$\rightsquigarrow$

Plan : 1) Théorème de représentation de Skorokhod

2) Distance en variation totale

3) Distance de Lévy-Prokhorov

4) Monotonie, ordre stochastique

Références pour un bestiaire de distances sur $M_1(E)$:

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

1) Théorème de représentation de Skorokhod

Théorème Soit (E,d) séparable, $(X_i)_{i \geq 1}$, X des v.a à valeurs dans E telles que $X_n \xrightarrow{\text{loi}} X$. On peut définir des v.a $(X'_i)_{i \geq 1}$, X' sur un m espace de proba tq $\forall n \geq 1$, $X_n \overset{\text{loi}}{=} X'_n$, $X \overset{\text{loi}}{=} X'$ et $X'_n \xrightarrow{p.s} X'$.

Voir feuille d'exos 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" $X_n \to X$ de manière déterministe)

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

Application Soit $X_1, X_2, X_3, \dots$ des v.a réelles t.q $X_n \xrightarrow{\text{loi}} X$. On suppose que $\lim_{a \to \infty} \sup_{n \geq 1} \mathbb{E}[|X_n| \mathbb{1}_{|X_n| \geq a}] = 0$ (*)

Alors $\mathbb{E}[X_n] \to \mathbb{E}[X]$

Preuve : D'après le thm de représentation de Skorokhod, on peut supposer que $X_n \xrightarrow{p.s.} X$ et que (*) (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

$$X_n \xrightarrow{\mathbb{P}} X \text{ , } (X_n)_{n \geq 1} \cup I \Rightarrow X_n \xrightarrow{L^1} X$$

$\sim$

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

Lemme Soit (E,d) séparable et $\mu \in M_1(E)$. Pour tout $\varepsilon > 0$, il existe une partition $E = \bigsqcup_{i \geq 1} A_i$ en ensembles mesurables tel que

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

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

$$\{r \in [\tfrac{\varepsilon}{4}, \tfrac{\varepsilon}{2}] : \mu(\partial B(x_i, r)) > 0\} \text{ est dénombrable.}$$

On choisit alors $\frac{\varepsilon}{2} \leq r_i \leq \varepsilon$ $\forall q$ $\mu(\underbrace{\partial B(x_i, r_i)}_{B_i}) = 0$

On définit par récurrence $A_1 = B_1, \dots, A_i = B_1^c \cap \dots \cap B_{i-1}^c \cap B_i$

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

$\mu(\partial A) = 0 \Rightarrow \mu(\partial A^c) = 0$ et $\mu(\partial A) = \mu(\partial B) = 0 \Rightarrow \mu(\partial(A \cup B)) = 0$)

$\sim$

Preuve du théorème de représentation de Skorokhod.

Notons μ_n la loi de X_n, 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 $\varepsilon_p = \frac{1}{2^p}$ 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

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

En particulier, $\mu_n(A_i^p) = 0 \Rightarrow \mu(A_i^p) = 0$

Sans perte de généralité, on suppose $a_1 < a_2 < \dots$ et $n > a_1$. Soit $p_n \geq 1$ l'unique entier tel que $a_{p_n} \leq n < a_{p_n+1}$. Lorsque $n \to \infty$, on a $p_n \to \infty$.

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 $X'_n \overset{\text{loi}}{=} X_n$.

On va définir X'_n comme suit :

$$X'_n = \sum_{\substack{i=0,\\ \mu_n(A_i^{p_n})>0}}^{k_{p_n}} \mathbb{1}_{\{U \le 1-\varepsilon_{p_n},\, X' \in A_i^{p_n}\}} Y_{n,i} + \mathbb{1}_{U>1-\varepsilon_{p_n}} Z_n$$

avec $(Y_{n,i})_{n,i\ge 1}$, $(Z_n)_{n\ge 1}$, U, X' v.a. indépendantes définies sur le même espace de probabilité :

- U uniforme sur $[0,1]$
- $X \overset{\text{loi}}{=} X'$
- $\text{loi}(Y_{n,i}) = \mu_n(\cdot \mid A_i^{p_n})$
- $\text{loi}(Z_n)$ définie par

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

quantité qui définit bien une mesure positive par (*) et de masse totale 1 (car $\mu_n(A_i^{p_n})=0 \Rightarrow \mu(A_i^{p_n})=0$)

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

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

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

Étape 3 On vérifie que $X'_n \xrightarrow{ps} X'$.

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

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

Donc $\sum_{p\ge 1} \mathbb{P}((E_p)^c) < \infty$.

Donc par Borel-Cantelli, ps, à partir d'un certain rang E_p est réalisé

Donc ps pour n assez grand E_{P_n} est réalisé.

Par définition de X'_n, cela implique alors $\exists\, 1 \le i \le k_{P_n}$ tq $X', X'_n \in A_i^{P_n}$

et donc $d(X, X'_n) \le \varepsilon_{P_n}$

Donc $X'_n \xrightarrow{p\text{-}s} X'$.

2

Remarque Lorsque E est polonais, il est possible de montrer que l'espace de proba commun peut être pris comme $[0,1]$ muni de la mesure de Lebesgue

2) Distance en variation totale

Soit $(S, \mathcal{B})$ un espace mesuré.

Définition Pour $\mu, \nu \in M_1(S)$, 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 $M_1(S)$ qui le rend complet
- Si μ et ν sont absolument continues par rapport à une même σ-finie sur $(S, \mathcal{B})$, alors $d_{VT}(\mu,\nu) = \frac{1}{2} \int_E \left| \frac{d\mu}{d\rho} - \frac{d\nu}{d\rho} \right| d\rho$

Ceci permet de montrer que :

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

et que $d_{VT}(\mu,\nu) = \inf \{ \mathbb{P}(X' \neq Y') : (X', Y') \text{ couplage de } X, Y \}$, avec l'inf 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, $\mathcal{B} = \mathcal{P}(S)$ la convergence pour d_{VT} est équivalente à la convergence en loi (⚠ faux en général) et $d_{VT}(\mu,\nu) = \frac{1}{2} \sum_{x \in S} |\mu(x) - \nu(x)|$

Remarque : 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 v.a. X,Y,Z définies sur le même espace :

① si $X=Y$ p.s. alors $\delta(X,Y)=0$ ② $\delta(X,Y)=\delta(Y,X)$ ③ $\delta(X,Y)\leq\delta(X,Z)+\delta(Z,Y)$.

La quantité $\hat{\delta}(X,Y)=\inf\{\delta(X',Y') : X'\overset{loi}{=}X,\ Y'\overset{loi}{=}Y\}$ est appelée <u>métrique minimale</u> associée à δ.

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

Remarque Pour un ensemble $\mathcal{F}$ de fonctions mesurables $S\to\mathbb{R}$, on pose pour X,Y v.a. à valeurs dans S, $d_{\mathcal{F}}(X,Y)=\sup\{|\mathbb{E}[f(X)]-\mathbb{E}[f(Y)]| : f\in\mathcal{F}\}$, appelée <u>métrique duale définie par $\mathcal{F}$</u>. Ainsi, d_{VT} est la métrique duale définie par l'une des 3 familles $\mathcal{F}=\{f:S\to\mathbb{R} \text{ mesurable}, \|f\|_\infty\leq\frac{1}{2}\}$, $\mathcal{F}=\{f:S\to\mathbb{R} \text{ mesurable}, 0\leq f\leq 1\}$ ou $\mathcal{F}=\{\mathbb{1}_A : A\in\mathcal{B}(S)\}$

Référence : .Notes de cours "Probability theory: the Coupling Method" de Frank den Hollander : http://websites.math.leidenuniv.nl/probability/lecturenotes/CouplingLectures.pdf (Section 2)

3) Distance de Lévy-Prokhorov

Soit (E,d) métrique séparable. On rappelle que $d_{LP}(\mu,\nu)=\inf\{\varepsilon>0 : \forall A\in\mathcal{B}(E),\ \mu(A)\leq\nu(A^\varepsilon)+\varepsilon\}$ pour $\mu,\nu\in\mathcal{M}_1(E)$, où $A^\varepsilon=\{x\in E : d(x,A)<\varepsilon\}$.

Définition Pour des variables aléatoires X,Y définies sur le même espace de proba 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 (Strassen) Soit (E,d) séparable, $\mu,\nu \in \mathcal{M}_1(E)$. Alors

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

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

Preuve d'une inégalité: Soit (X,Y) un tel couplage. On a $d_{LP}(\mu,\nu) \le \alpha(X,Y)$.

Soit $\varepsilon > \alpha(X,Y)$, de sorte que $\mathbb{P}(d(X,Y)\ge\varepsilon)<\varepsilon$.

Alors $\mu(A) = \mathbb{P}(X\in A) \le \mathbb{P}(X\in A, d(X,Y)<\varepsilon) + \mathbb{P}(X\in A, d(X,Y)\ge\varepsilon)$

$$\le \mathbb{P}(Y\in A^\varepsilon) + \varepsilon = \nu(A^\varepsilon) + \varepsilon,$$

d'où $d_{LP}(\mu,\nu) \le \varepsilon$

$\rightsquigarrow$

L'autre inégalité est délicate et peut se montrer en utilisant le Lemme des mariages, cf

Billingsley, Convergence of probability measures (2nd Ed), Thm 6.9

ou Dudley, Real Analysis and probability, Corollary 11.6.4.

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

On finit cette partie par un résultat utile de couplage "à la Skorokhod":

Théorème (Dudley) Soit (E,d) séparable, X_n, Y_n des v.a. à valeurs dans E tq $d_{LP}(X_n,Y_n)\to 0$

Alors il existe un espace de proba avec des v.a. $(X'_n, Y'_n)_{n\ge1}$ définies dessus tq

- $\forall n\ge1$, $X'_n \overset{\text{loi}}{=} X_n$, $Y'_n \overset{\text{loi}}{=} Y_n$
- ps, $d(X'_n, Y'_n)\to 0$

cf Dudley, Real Analysis and probability, Theorem 11.7.1

4) Monotonie, ordre stochastique

Ici (E,d) est un espace polonais.

Définition Une relation $\leq$ est un ordre partiel si $\forall x,y,z \in E$:

1) $x \leq x$ 2) $x \leq y, y \leq z \Rightarrow x \leq z$ 3) $x \leq y, y \leq x \Rightarrow x = y$

Exemples • $\mathbb{R}$ muni de l'ordre usuel

• $\{0,1\}^{\mathbb{Z}^2}$ avec $(a_i)_{i \in \mathbb{Z}^2} \leq (b_i)_{i \in \mathbb{Z}^2}$ si $a_i = 1 \Rightarrow b_i = 1$ (si $a_i = 1$, on dit que le site i est occupé, vide sinon)

Définition Un événement A est dit croissant si $x \in A \Rightarrow \{y \in E : x \leq y\} \in A$.

Exemple Pour $E = \{0,1\}^{\mathbb{Z}^2}$, $A = \{\exists$ chemin infini constitué de sites occupés adjacents$\}$ est croissant

Définition Soit $\mu, \nu \in M_1(E)$. On dit que ν domine stochastiquement μ et on note $\mu \leq \nu$ si $\forall A \in \mathcal{B}(E)$ croissant, $\mu(A) \leq \nu(A)$, ou, de manière équivalente

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

(Par croissante on veut dire $x \leq y \Rightarrow f(x) \leq f(y)$)

Exemple Pour $E = \mathbb{R}$, $\mu \leq \nu \Leftrightarrow \forall x \in \mathbb{R}$, $\mu(]-\infty, x]) \geq \nu(]-\infty, x])$.

Théorème (Strassen) On suppose que $\{(x,y) \in E^2 : x \leq y\}$ est fermé. Soit $\mu, \nu \in M_1(E)$. Alors (1) (E) (?)

(1) $\mu \leq \nu$

(2) $\exists$ couplage (X,Y) tq loi$(X) = \mu$, loi$(X) = \nu$ et $\mathbb{P}(X \leq X') = 1$

La preuve repose sur de l'analyse fonctionnelle, cf

Lindvall, *Lectures on the Coupling Method*, chapter IV

ou Lindvall, *On Strassen's theorem on stochastic domination*, ECP 4 (1999), 51-59

V Théorème de Donsker et applications

Plan : 1) Le théorème de Donsker

2) Loi de l'arcsinus

3) Théorème de Lévy sur le mouvement brownien

1) Le théorème de Donsker

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

Proposition (inégalité de Skorokhod) Soit $(X_i)_{i\geq 1}$ des v.a. iid à valeurs dans $\mathbb{R}^d$, centrées avec $\mathbb{E}[|X_1|^2] = c$. Alors $\forall \lambda > 0$, $\forall n \geq 1$, $\mathbb{P}\left(\max_{1\leq i\leq n} |S_i| > \lambda\sqrt{n}\right) \leq 2\,\mathbb{P}\left(|S_n| > (\lambda - \sqrt{2c})\sqrt{n}\right)$,
où $|\cdot|$ désigne la norme euclidienne.

Preuve : Soit T le temps d'arrêt $T = \inf\{k\geq 1 : |S_k| > \lambda\sqrt{n}\}$. On suppose $\lambda > \sqrt{2c}$ et on écrit

$$\mathbb{P}\left(\max_{1\leq i\leq n} |S_i| > \lambda\sqrt{n}\right) = \mathbb{P}(T \leq n)$$

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

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 $|S_n - S_j| \geq \sqrt{2cn}$ par inégalité triangulaire.

Par indépendance et l'inégalité de Markov,

$$\mathbb{P}\left(T = j, |S_n - S_j| > \sqrt{2cn}\right) = \mathbb{P}(T = j)\,\mathbb{P}\left(|S_{n-j}| > \sqrt{2cn}\right)$$

$$\leq \mathbb{P}(T = j)\,\frac{(n-j)\,\mathbb{E}[|X_1|^2]}{2cn} \leq \frac{\mathbb{P}(T = j)}{2}$$

On conclut que $\mathbb{P}\left(\max_{1\leq i\leq n} |S_i| > \lambda\sqrt{n}\right) \leq \mathbb{P}\left(|S_n| > (\lambda - \sqrt{2c})\sqrt{n}\right) + \frac{1}{2}\,\mathbb{P}(T \leq n)$, d'où le résultat

$\sim$

On commence par quelques notations. Soit $(X_i)_{i\geq 1}$ des v.a. iid dans $\mathbb{R}^d$ telles que $\mathbb{E}[X_1]=0$ et $\mathrm{Cov}(X_1)=I_d$, la matrice identité dans $\mathbb{R}^d$. On pose $S_0=0$, $S_n=X_1+\dots+X_n$ pour $n\geq 1$ et par interpolation linéaire $S_t=(1-\{t\})S_{\lfloor t\rfloor}+\{t\}S_{\lfloor t\rfloor+1}$ pour $t\geq 0$. Enfin, pour $n\geq 1$, on pose $S^{(n)}_t=\frac{S_{nt}}{\sqrt{n}}$, $t\geq 0$, de sorte que $S^{(n)}$ est une variable aléatoire à valeurs dans $\mathcal{C}(\mathbb{R}_+,\mathbb{R}^d)$.

Théorème (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 le fait que pour tout $0=t_0\leq t_1\leq\dots\leq t_k$, les variables aléatoires $(B_{t_i}-B_{t_{i-1}})_{1\leq i\leq k}$ sont indépendantes, respectivement de loi gaussienne $\mathcal{N}(0,(t_i-t_{i-1})I_d)$.*

Preuve : Etape 1 : convergence des fidi

Or fixe $0=t_0<t_1<\dots<t_k$. D'après le TCL,

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

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

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

Pour enlever les parties entières, on remarque que $\left|S^{(n)}_t-\frac{S_{\lfloor nt\rfloor}}{\sqrt{n}}\right|\leq\frac{|X_{\lfloor nt\rfloor+1}|}{\sqrt{n}}$, qui converge en probabilité vers 0 à t fixé.

Donc $\left(S^{(n)}_{t_i}-\frac{S_{\lfloor nt_i\rfloor}}{\sqrt{n}} : 1\leq i\leq k\right)\xrightarrow{\mathbb{P}}0$.

Par le lemme de Slutsky, on en déduit que

$$\left(S^{(n)}_{t_i} : 1\leq i\leq k\right)\xrightarrow{\text{loi}}(N_1,N_1+N_2,\dots,N_1+\dots+N_k),$$

ce qui montre la convergence des fidi.

Étape 2 : Tension.

Il s'agit de vérifier que $\forall N \geq 1$, les restrictions $(S^{(n)}_{|[0,N]})_{n\geq 1}$ sont tendues. Pour simplifier les notations, on le fait pour $N=1$ (le cas général étant identique).

Avec des hypothèses de moment supplémentaires, on peut utiliser le critère de Kolmogorov.

Dans le cas général, on va montrer que

$$\forall \varepsilon, \eta > 0 \text{ fixés}, \ \forall t \in [0,1], \ \limsup_{n\to\infty} \frac{1}{\delta} \mathbb{P}\Big(\sup_{t\leq s\leq (t+\delta)\wedge 1} |S^{(n)}_s - S^{(n)}_t| > \eta \Big) \leq \varepsilon$$

en utilisant l'inégalité maximale de Skorokhod

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

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

En effet, on écrit $|S_{ns} - S_{nt}| \leq |S_{ns} - S_{\lfloor nt\rfloor}| + |S_{nt} - S_{\lfloor nt\rfloor}|$.

D'une part, $|S_{nt} - S_{\lfloor nt\rfloor}| \leq |S_{\lfloor nt\rfloor+1} - S_{\lfloor nt\rfloor}|$ ($(S_{\lfloor nt\rfloor+u} : 0\leq u\leq 1)$ décrit un segment)

D'autre part, $|S_{ns} - S_{\lfloor nt\rfloor}| \leq \max\big(|S_{\lfloor ns\rfloor} - S_{\lfloor nt\rfloor}|, |S_{\lfloor ns\rfloor+1} - S_{\lfloor nt\rfloor}|\big)$

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, $\forall \eta>0$, $\forall \delta\in]0,1[$, $\forall t\in[0,1]$,

$$\mathbb{P}\Big(\sup_{t\leq s\leq t+\delta} |S^{(n)}_s - S^{(n)}_t| > \eta \Big) \leq \mathbb{P}\Big(\max_{1\leq i\leq \lfloor n\delta\rfloor+2} |S_{\lfloor nt\rfloor+i} - S_{\lfloor nt\rfloor}| > \eta\sqrt{n}/2 \Big)$$

$$= \mathbb{P}\Big(\max_{1\leq i\leq \lfloor n\delta\rfloor+2} |S_i| > \Big(\frac{\eta}{2} \frac{\sqrt{n}}{\sqrt{\lfloor n\delta\rfloor+2}} \Big) \cdot \sqrt{\lfloor n\delta\rfloor+2} \Big)$$

$$\leq 2\,\mathbb{P}\Big(\frac{|S_{\lfloor n\delta\rfloor+2}|}{\sqrt{\lfloor n\delta\rfloor+2}} > \frac{\eta}{2} \frac{1}{\sqrt{\delta+2/n}} - \sqrt{2k} \Big)$$ d'après l'inégalité de Skorokhod, car $\mathbb{E}[|X_1|^2] = d$.

D'après le TCL,

$$\limsup_{n\to\infty} \mathbb{P}\left(\sup_{t\le s\le t+\delta} |S_s^{(n)} - S_t^{(n)}| > \eta \right) \le 2\,\mathbb{P}\left(|N(0,\Sigma_d)| > \frac{\eta}{2} \times \frac{1}{\sqrt{\delta}} - \sqrt{2d} \right)$$

Or $\mathbb{P}(|N(0,\Sigma_d)| \ge x)$ décroit exponentiellement vite quand $x \to \infty$. Donc on peut trouver $\alpha \le \delta < 1$ tel que $\frac{1}{\delta} \limsup_{n\to\infty} \mathbb{P}\left(\sup_{t\le s\le t+\delta} |S_s^{(n)} - S_t^{(n)}| > \eta \right) \le \varepsilon$.

$\sim$

Corollaire Soit $(X_i)_{i\ge 1}$ des variables aléatoires réelles centrées. Notons $\sigma^2 = \mathbb{E}[X_1^2]$. Alors

$$\forall a > 0,\quad \mathbb{P}\left(\max_{0\le i\le n} S_i > a\sigma\sqrt{n} \right) \xrightarrow[n\to\infty]{} 2\int_a^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx$$

Preuve : En notant $S_t^{(n)} = \frac{S_{nt}}{\sigma\sqrt{n}}$ pour $0\le t\le 1$, on a $(S_t^{(n)} : 0\le t\le 1) \xrightarrow[n\to\infty]{\text{loi}} (B_t)_{0\le t\le 1}$ d'après le théorème de Donsker, où $(B_t)_{0\le t\le 1}$ est un mouvement brownien standard.

Puisque $\sup_{0\le t\le 1} S_t^{(n)} = \max_{0\le i\le n} S_i$ et que $\sup : \mathcal{B}([0,1],\mathbb{R}) \to \mathbb{R}$ est $f \mapsto \sup f$ continue, le résultat s'ensuit en utilisant le fait que

$$\sup_{0\le t\le 1} B_t \overset{\text{loi}}{=} |B_1| \quad \text{(ce qu'on peut voir avec le principe de reflexion)}$$

Une propriété du "continu" (loi explicite de $\sup_{0\le t\le 1} B_t$) a été utilisée pour obtenir un résultat asymptotique dans le discret (fonction de répartition de $\max_{0\le k\le n} S_k$).

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.

2) loi de l'arcsinus

Théorème Soit $(B_t)_{0\le t\le 1}$ 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 $\mathbb{P}(T\le t) = \frac{2}{\pi}\arcsin(\sqrt{t})$ pour $0\le t\le 1$.

Preuve : Etape 1 On pose $S_n = X_1+\dots+X_n$ avec $(X_i)_{i\ge1}$ iid tq $\mathbb{P}(X_1 = \pm 1) = \frac{1}{2}$ et on note $T_{2n} = \max\{0\le k\le 2n : S_k = 0\}$. Montrons que $\frac{T_{2n}}{2n} \xrightarrow{\text{loi}}$ v.a. de densité $\frac{1}{\pi\sqrt{x(1-x)}}\mathbb{1}_{0<x<1}$.

Pour $0\le j\le n$, 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 $\boxed{\forall k\ge 1,\ \mathbb{P}(S_1\neq 0,\dots,S_{2k}\neq 0) = \mathbb{P}(S_{2k}=0)}$ (*)

Démontrons cette identité surprenante par un joli argument combinatoire

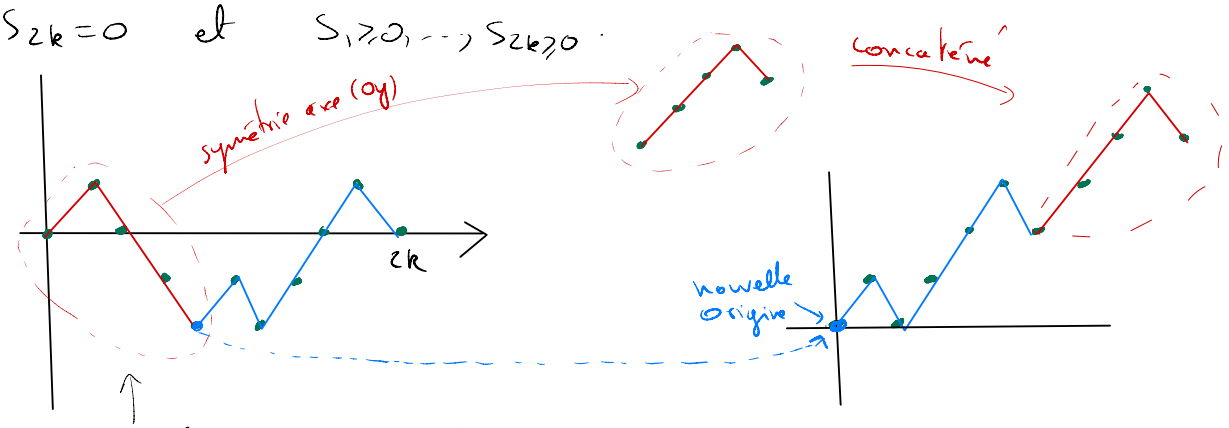
Tout d'abord, $\mathbb{P}(S_1\neq 0,\dots,S_{2k}\neq 0) = 2\,\mathbb{P}(S_1>0,\dots,S_{2k}>0)$

$= 2\times\frac{1}{2}\mathbb{P}(S_1\ge 0,\dots,S_{2k-1}\ge 0)$ en changeant l'origine du repère en $(1,1)$

$= \mathbb{P}(S_1\ge 0,\dots,S_{2k}\ge 0)$ (par parité, $S_{2k-1}\ge 0 \iff S_{2k}\ge 0$)

Construisons maintenant une bijection entre chemins $(S_1,\dots,S_{2k})$ tels que

$S_{2k}=0$ et $S_1\ge 0,\dots,S_{2k}\ge 0$.



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 concaténe à droite de ce qui reste (partie bleue) et on place l'origine du repère au 1er point de la partie bleue (cf 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 $\mathbb{P}(S_{2k} = 0) = \binom{2k}{k} \frac{1}{2^{2k}} \underset{k\to\infty}{\sim} \frac{1}{\sqrt{\pi k}}$.

Donc $\mathbb{P}(T_{2n} = L2c\,ns) \underset{n\to\infty}{\sim} \frac{1}{n} \frac{1}{\pi\sqrt{c(1-c)}}$ pour $0 < c < 1$.

Ceci implique le résultat (exo, ou cf cours 7 la semaine prochaine : un théorème local limite implique une convergence en loi)

Etape 2 Pour $f \in \mathcal{B} = \mathcal{B}([0,1], \mathbb{R})$, soit $T(f) = \sup\{t \in [0,1] : f(t) = 0\}$. On ne peut pas directement appliquer le théorème de Donsker car T n'est pas continue.

(prendre par exemple $f_n(x) = \frac{x}{n}$ pour $0 \leq x \leq 1$: $0 = T(f_n) \not\to T(0) = 1$)

Mais on vérifie que si 0 n'est pas une valeur de min ou max local, alors T est continue en f.

Par ailleurs, ps 0 n'est pas une valeur de min ou max local d'un mouvement brownien

D'après le thm de Donsker, $\left(\frac{S_{nt}}{\sqrt{n}} : 0 \leq t \leq 1\right) \xrightarrow{\text{loi}} B$, donc $T\left(\left(\frac{S_{nt}}{\sqrt{n}} : 0 \leq t \leq 1\right)\right) \xrightarrow{\text{loi}} T(B)$

(on utilise le fait que si $X_n \xrightarrow{\text{loi}} X$ et si f est ps continue en X, alors $f(X_n) \xrightarrow{\text{loi}} f(X)$)

$\sim$

Corollaire Si $(X_i)_{i\geq 1}$ sont iid avec $\mathbb{E}[X_1] = 0$ et $\mathbb{E}[X_1^2] \in]0, \infty[$, alors

$$\mathbb{P}\left(\frac{1}{2n} \max\{0 \leq k \leq 2n : S_k = 0\} \leq t\right) \xrightarrow[n\to\infty]{} \frac{2}{\pi} \arcsin(\sqrt{t}) \quad \text{pour } 0 < t < 1.$$

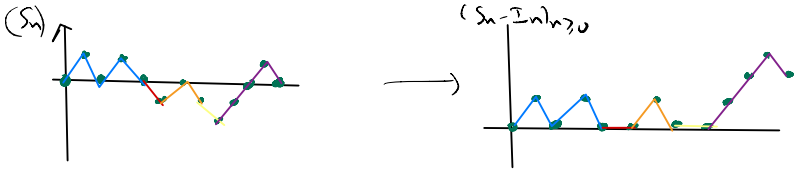
où $S_k = X_1 + \dots + X_k$.

3) Théorème de Lévy

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

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

Preuve : Soit $(X_i)_{i\geq 1}$ des v.a. iid t.q. $\mathbb{P}(X_1 = \pm 1) = 1/2$, on pose $S_n = X_1 + \dots + X_n$ et $I_n = \min_{0\leq k\leq n} S_k$



Si on regarde les probas 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 proba 1/2.

On note $\vartheta_n = \#\{k \leq n-1 : S_k - I_k \neq S_{k-1} - I_{k-1}\}$ le nombre de pas "effectifs" de $(S_k - I_k)_{k\geq 0}$ jusqu'au temps n, et $\vartheta_t^{-1} = \inf\{s : \vartheta_s = t\}$ son "inverse". Alors

$$(S_{\vartheta_n^{-1}} - I_{\vartheta_n^{-1}})_{n\geq 0} \overset{loi}{=} (|S_n|)_{n\geq 0}$$

Mais d'après le théorème de Donsker $\left(\frac{1}{\sqrt{n}} |S_{nt}|\right)_{t\geq 0} \xrightarrow{loi} (|B_t|)_{t\geq 0}$.

Pour contrôler ϑ_n, on vérifie que $(S_n - I_n)_{n\geq 0}$ est une chaîne de Markov récurrente nulle (ou même invariante ∞), de sorte que $\frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{S_k - I_k = 0} \xrightarrow[n\to\infty]{p.s.} 0$.

On en déduit aisément que $\left(\frac{\vartheta^{-1}_{nt}}{n} : t\geq 0\right) \xrightarrow{loi} (t)_{t\geq 0}$, puis que

$$\left(\frac{S_{\vartheta^{-1}_{nt}} - I_{\vartheta^{-1}_{nt}}}{\sqrt{n}}\right)_{t\geq 0} \xrightarrow{loi} (B_t - I_t)_{t\geq 0}$$ (détails laissés en exo)

On en conclut que $(B_t - I_t)_{t\geq 0} \overset{loi}{=} (|B_t|)_{t\geq 0}$.

~

VI : le pont brownien

Plan : 1) Définition du pont brownien

2) Le théorème local limite

3) Convergence vers le pont brownien (Donsker conditionné)

1) Définition du pont brownien

Soit $(B_t)_{0\le t\le 1}$ un mouvement brownien standard

Définition/Théorème Lorsque $\varepsilon \downarrow 0$, la loi de $(B_s)_{0\le s\le 1}$ sachant que $\{|B_1|\le\varepsilon\}$ 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 $(B_s - sB_1)_{0\le s\le 1}$ est $\mathbb{P}_0$.

Informellement, le pont brownien est $(B_t)_{0\le t\le 1}$ "sachant $B_1 = 0$"

Preuve : *Étape 1 :* On pose $b_s = B_s - sB_1$ pour $0\le s\le 1$. On vérifie d'abord que $(b_s)_{0\le s\le 1} \perp\!\!\!\perp B_1$.

Puisque $((b_s)_{0\le s\le 1}, B_1)$ est un vecteur gaussien, il suffit de vérifier que $\forall s\in[0,1]$, $\mathrm{Cov}(b_s, B_1) = 0$.

Pour cela, on calcule $\mathrm{Cov}(b_s, B_1) = \mathbb{E}[(B_s - sB_1)B_1] = \mathbb{E}[B_sB_1] - s = \mathbb{E}[B_s(B_1 - B_s) + B_s^2] - s$

$= \mathbb{E}[B_s]\mathbb{E}[B_1 - B_s] + s - s = 0$.

Étape 2 : Soit $F : \mathcal{C}([0,1],\mathbb{R}) \to \mathbb{R}$ lipschitzienne bornée. On note $p_t(x) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}$ la densité de B_t.

$$\mathbb{E}[F(B_s : 0\le s\le 1) \,|\, |B_1|\le\varepsilon] = \frac{1}{\mathbb{P}(|B_1|\le\varepsilon)} \mathbb{E}[F(b_s + sB_1 : 0\le s\le 1)\, \mathbb{1}_{|B_1|\le\varepsilon}]$$

$$= \frac{1}{\mathbb{P}(|B_1|\le\varepsilon)} \int_{-\varepsilon}^{\varepsilon} p_1(x)\,dx\, \mathbb{E}[F(b_s + sx : 0\le s\le 1)]$$

Ainsi,

$$|\mathbb{E}[F(B_s : 0\le s\le 1) \,|\, |B_1|\le\varepsilon] - \mathbb{E}[F(b_s : 0\le s\le 1)]| \le \frac{1}{\int_{-\varepsilon}^{\varepsilon} p_1(x)\,dx} \int_{-\varepsilon}^{\varepsilon} p_1(x)\,dx \Big| \mathbb{E}[F(b_s + sx : 0\le s\le 1)] - \mathbb{E}[F(b_s : 0\le s\le 1)] \Big|$$

$$\xrightarrow[\varepsilon\to 0]{} 0 \quad \text{car } F \text{ lipschitzienne}$$

$\sim$

Proposition La loi du pont brownien est caractérisée par la relation suivante d'absolue continuité : Pour $0<t<1$ et $F : \mathcal{C}([0,1],\mathbb{R}) \to \mathbb{R}$ continue bornée ou mesurable ≥ 0

$$\mathbb{E}[F(b_s : 0\leq s\leq t)] = \mathbb{E}\left[F(B_s : 0\leq s\leq t)\ \frac{p_{1-t}(-B_t)}{p_1(0)}\right].$$

Preuve : Il suffit de montrer le résultat pour $F : \mathcal{C}([0,1],\mathbb{R}) \to \mathbb{R}$ 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 MB indépendant de $(B_u)_{0\leq u\leq t}$

Alors $\mathbb{E}[F(B_s : 0\leq s\leq t) \mid |B_1|\leq\varepsilon] = \frac{1}{\mathbb{P}(|B_1|\leq\varepsilon)}\mathbb{E}[F(B_s : 0\leq s\leq t)\ \mathbb{1}_{\{|B_t+W_{1-t}|\leq\varepsilon\}}]$

$$= \frac{1}{\mathbb{P}(|B_1|\leq\varepsilon)}\mathbb{E}\left[F(B_s : 0\leq s\leq t)\ \mathbb{1}_{W_{1-t}\in[-B_t-\varepsilon,\,-B_t+\varepsilon]}\right]$$

$$= \frac{1}{\int_{-\varepsilon}^{\varepsilon} p_1(x)\,dx}\mathbb{E}\left[F(B_s : 0\leq s\leq t)\int_{-B_t-\varepsilon}^{-B_t+\varepsilon} p_{1-t}(x)\,dx\right]$$

Or $\frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon} p_1(x)\,dx \xrightarrow[\varepsilon\to 0]{} p_1(0)$ et $\frac{1}{2\varepsilon}\int_{-B_t-\varepsilon}^{-B_t+\varepsilon} p_{1-t}(x)\,dx \xrightarrow[\varepsilon\to 0]{} p_{1-t}(-B_t)$.

De plus, p_{1-s} étant borné, $\frac{1}{2\varepsilon}\int_{-B_t-\varepsilon}^{-B_t+\varepsilon} p_{1-t}(x)\,dx \leq \|p_{1-t}\|_\infty$.

Par convergence dominée, $\mathbb{E}[F(B_s : 0\leq s\leq t) \mid |B_1|\leq\varepsilon] \xrightarrow[\varepsilon\to 0]{} \mathbb{E}\left[F(B_s : 0\leq s\leq t)\frac{p_{1-t}(-B_t)}{p_1(0)}\right]$

Comme aussi $\mathbb{E}[F(B_s : 0\leq s\leq t) \mid |B_1|\leq\varepsilon] \xrightarrow[\varepsilon\to 0]{} \mathbb{E}[F(b_s : 0\leq s\leq t)]$, on en déduit le résultat.

2) Le théorème local limite

Lemme Soit $k\geq 1$ entier, pour $1\leq i\leq k$ soit (α_n^i) une suite de réels >0 tq $\alpha_n^i \xrightarrow[n\to\infty]{} +\infty$. On considère une variable aléatoire $(X_n^1,\dots,X_n^k)$ à valeurs dans $\mathbb{Z}^k$ tq $\forall(\lambda_1,\dots,\lambda_k)\in\mathbb{R}^k$,

$$\alpha_n^1\alpha_n^2\cdots\alpha_n^k\ \mathbb{P}\left((X_n^1,\dots,X_n^k) = (\lfloor\lambda_1\alpha_n^1\rfloor,\dots,\lfloor\lambda_k\alpha_n^k\rfloor)\right) \xrightarrow[n\to\infty]{} \phi(\lambda_1,\dots,\lambda_k)$$

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 densité ϕ.

Preuve : On va utiliser le Lemme de Scheffé (cf feuille 1) : si X_n a pour densité f_n sur $\mathbb{R}^k$ et si $f_n \xrightarrow{pp} f$ avec f densité de proba, alors $X_n \xrightarrow{loi} X$ avec X à densité f.

Soit $(Y_n^1, \dots, Y_n^k)$ une v.a. de densité $\phi_n(\lambda_1, \dots, \lambda_k) = \alpha_n^1 \cdots \alpha_n^k \mathbb{P}\left((X_n^1, \dots, X_n^k) = (\lfloor \lambda_1 \alpha_n^1 \rfloor, \dots, \lfloor \lambda_k \alpha_n^k \rfloor)\right)$ 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 $\phi_n \to \phi$, $(Y_n^1, \dots, Y_n^k) \xrightarrow{loi} X$ et comme $\left| Y_n^i - \frac{Z_n^i}{\alpha_n^i} \right| \le \frac{1}{\alpha_n^i} \xrightarrow[n \to \infty]{} 0$ $\forall 1 \le i \le k$,

on en déduit le résultat.

$\sim$

Définition Une variable aléatoire X réelle est dite lattice s'il existe $b \in \mathbb{R}$ et $h \in \mathbb{R}_+^*$ tq $\mathbb{P}(X \in b + h\mathbb{Z}) = 1$

Le plus grand tel h est appelé le span de X (le fait qu'il soit bien défini provient du lemme ci-dessous)

S'il vaut 1, on dit que X est apériodique

Exemple Si $\mathbb{P}(X=1) = \mathbb{P}(X=-1)$, X a pour span 2

Lemme Soit X v.a. réelle

(1) X est lattice ssi $\exists t \in \mathbb{R}^*$ tq $|\phi(t)| = 1$

(2) X a pour span $h > 0$ ssi $|\phi(\frac{2\pi}{h})| = 1$ et $|\phi(t)| < 1$ pour $t \in]0, \frac{2\pi}{h}[$.

Preuve : (1) Si $\mathbb{P}(X \in b + h\mathbb{Z}) = 1$, alors $\mathbb{P}\left(\frac{X-b}{h} \in \mathbb{Z}\right) = 1$, de sorte que $\mathbb{E}\left[e^{2i\pi \frac{X-b}{h}}\right] = 1$ et

$$\left|\mathbb{E}\left[e^{\frac{2i\pi X}{h}}\right]\right| = \left|e^{2i\pi b}\right| = 1$$

Réciproquement, si $|\phi(t_0)| = 1$ pour $t_0 \in \mathbb{R}^*$, soit $a \in \mathbb{R}$ tq $\phi(t_0) = e^{ia}$. Alors $\mathbb{E}\left[e^{i(t_0 X - a)}\right] = 1$

Donc $\mathbb{E}[\cos(t_0 X - a)] = 1$. Comme $\cos \le 1$, on en déduit que p.s $\cos(t_0 X - a) = 1$

Donc p.s $t_0 X - a \in 2\pi\mathbb{Z}$, donc $\mathbb{P}\left(X \in \frac{a}{t_0} + \frac{2\pi}{t_0}\mathbb{Z}\right) = 1$.

(2) La preuve de (1) montre que $\exists b \in \mathbb{R}$ et $h > 0$ tq $\mathbb{P}(X \in b + h\mathbb{Z}) = 1$ ssi $|\phi(\frac{2\pi}{h})| = 1$,

d'où le résultat

$\sim$

Théorème (local limite) Soit $(X_i)_{i\geq 1}$ des v.a réelles i.i.d apériodiques à valeurs dans $\mathbb{Z}$ tq $a=\mathbb{E}[X_1]$ et $\sigma^2=\mathbb{E}[X_1^2]-\mathbb{E}[X_1]^2\in]0,\infty[$. On note $S_n=X_1+\cdots+X_n$. Alors

$$\sup_{k\in\mathbb{Z}}\left|\sqrt{n}\,\mathbb{P}(S_n=k)-\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2}\left(\frac{k-an}{\sigma\sqrt{n}}\right)^2\right)\right|\longrightarrow_{n\to\infty}0.$$

Preuve : Posons $X_i'=X_i-a$, $S_n'=S_n-an$. On note $\varphi(t)=\mathbb{E}[e^{itX_1'}]$

Puisque $\mathbb{E}[e^{itS_n'}]=\sum_{j\in\mathbb{Z}-an}e^{itj}\,\mathbb{P}(S_n'=j)$, il vient $\mathbb{P}(S_n'=k)=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{-itk}\,\mathbb{E}[e^{itS_n'}]\,dt=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{-itk}\,\varphi(t)^n\,dt$ pour $k\in\mathbb{Z}-an$

Ainsi, $u\in\mathbb{R}$ étant choisi de sorte que $u\sigma\sqrt{n}-an\in\mathbb{Z}$

$$\sigma\sqrt{n}\,\mathbb{P}(S_n'=u\sigma\sqrt{n})=\frac{1}{2\pi}\int_{-\pi\sigma\sqrt{n}}^{\pi\sigma\sqrt{n}}e^{-itu}\,\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)^n dt.$$

Mais $\forall u\in\mathbb{R}$,

$$\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u^2}=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{-itu-t^2/2}\,dt.$$

Ainsi, en fixant $A>0$ et $0<\varepsilon\leq 1$, pour n assez grand

$$\left|\sigma\sqrt{n}\,\mathbb{P}(S_n'=u\sigma\sqrt{n})-\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u^2}\right|\leq\frac{1}{2\pi}\left(|I_1^{(A)}|+|I_2^{(A,\varepsilon)}|+|I_3^{(\varepsilon)}|+|I_4^{(A)}|\right)$$

avec $I_1=\int_{-A}^{A}e^{-itu}\left(\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)^n-e^{-t^2/2}\right)dt$, $I_2=\int_{A<|t|<\varepsilon\sigma\sqrt{n}}e^{-itu}\,\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)^n dt$,

$I_3=\int_{\varepsilon\sigma\sqrt{n}<|t|<\pi\sigma\sqrt{n}}e^{-itu}\,\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)^n dt$, $I_4=\int_{|t|>A}e^{-itu-t^2/2}\,dt$

On montre $\forall\varepsilon'>0$ fixé, on peut trouver $A>0$ et $0<\varepsilon\leq 1$ tq pour n assez grand $|I_i|\leq\varepsilon'$ pour $1\leq i\leq 4$.

On écrit $\varphi(t)=1-\frac{t^2}{2}\sigma^2+t^2\gamma(t^2)$ avec $\gamma(0)=0$ et γ continue. (*)

Pour le voir, on peut utiliser l'inégalité $\left|e^{ix}-\sum_{k=0}^{n}\frac{(ix)^k}{k!}\right|\leq\min\left(\frac{|x|^{n+1}}{(n+1)!},\frac{2|x|^n}{n!}\right)$

(cf Lemma 3.3.19 dans Durrett, Probability theory and examples, 5th Ed),

qui implique $|\varphi(t)-1-t^2\sigma^2/2|\leq t^2\,\underbrace{\mathbb{E}[\min(|t||X|^3,2X^2)]}_{\to 0\text{ par convergence dominée}}$

Pour I_1 On peut écrire $I_1=\int_{-A}^{A}g_n(u,t)\,dt$ avec $|g_n(u,t)|\leq 1+e^{-t^2/2}$ pour tout $u\in\mathbb{R}$, $t\in[-A,A]$, $n\geq 1$

et $\sup_{u\in\mathbb{R}}|g_n(u,t)|\to 0$ par (*).

Par convergence dominée, $\forall A>0$, $|I_1|\to 0$ uniformément en u.

Pour I_4 On a $|I_4| \leq 2\int_A^\infty e^{-t^2/2}\,dt$.

Pour I_3 Par apériodicité, $|\phi(t)|<1$ pour tout $0<t<2\pi$. Il existe donc $c>0$ tq $\left|\phi\left(\frac{t}{\sigma\sqrt{n}}\right)\right| \leq e^{-c}$ pour tout $\varepsilon\sigma\sqrt{n} < |t| < \pi\sigma\sqrt{n}$, de sorte que

$$|I_3| \leq 2\int_{\varepsilon\sigma\sqrt{n}}^{+\pi\sigma\sqrt{n}} e^{-cn}\,dt \leq 2\pi\sigma\sqrt{n}\,e^{-cn}$$

Pour I_2 On montre que $|\varphi(t)| \leq \exp\left(-\frac{t^2\sigma^2}{4}\right)$ pour t assez petit.

Par (*), $|\phi(t)|^2 = \phi(t)\overline{\phi(t)} = \left(1-\frac{t^2\sigma^2}{2}+t^2\eta(t^2)\right)\left(1-\frac{t^2\sigma^2}{2}+t^2\overline{\eta(t^2)}\right) = 1-\sigma^2t^2+o(t^2)$,

et donc $|\phi(t)| = 1-\frac{\sigma^2t^2}{2}+o(t^2)$, d'où le résultat par le DL de exp en 0.

Ainsi, si $\varepsilon>0$ est choisi suffisamment petit,

$$|I_2| \leq 2\int_A^{\varepsilon\sigma\sqrt{n}} \left(e^{-\left(\frac{t}{\sigma\sqrt{n}}\right)^2\frac{\sigma^2}{4}}\right)dt \leq 2\int_A^\infty e^{-t^2/4}\,dt.$$

Ceci conclut la preuve ~

Corollaire Sous les hypothèses du théorème local limite, en supposant $\mathbb{P}[X_1]=0$, on a

(1) $\mathbb{P}(S_n=0) \sim \frac{1}{\sqrt{2\pi\sigma^2}} \times \frac{1}{\sqrt{n}}$

(2) $\exists C>0$ tq $\forall n\geq 1, \forall k\geq 1, \mathbb{P}(S_n=k) \leq \frac{C}{\sqrt{n}}$

(3) si (x_n) est une suite d'entiers tq $\frac{x_n}{\sigma\sqrt{n}} \to x \in \mathbb{R}$, alors

$$\sqrt{\sigma^2 n}\,\mathbb{P}(S_n=x_n) \underset{n\to\infty}{\longrightarrow} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Remarque Il est possible de montrer l'amélioration suivante (utile pour $|k|$ grand) :

$$\sup_{k\in\mathbb{Z}} \max\left(1, \left(\frac{k-an}{\sqrt{n}}\right)^2\right)\left|\sigma\sqrt{n}\,\mathbb{P}(Z_n=k) - \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{k-an}{\sigma\sqrt{n}}\right)^2\right)\right| \longrightarrow 0$$

(cf Principles of Random Walk, Spitzer (Chap II, Sect 7, P10))

3) Convergence vers le pont brownien

On rappelle la notation $p_t(x) = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t}$

Théorème (Donsker conditionné) Soit $(X_i)_{i\geq 1}$ 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 $S_n = X_1 + \cdots + X_n$ et on définit S_t par interpolation linéaire. Alors

$$\left(\frac{S_{nt}}{\sigma\sqrt{n}} : 0 \leq t \leq 1 \right) \text{ sous } \mathbb{P}(\cdot \mid S_n = 0) \xrightarrow[n\to\infty]{\text{loi}} (b_t)_{0\leq t\leq 1}$$

Remarque L'apériodicité implique que $\mathbb{P}(S_n = 0) > 0$ pour tout n assez grand (conséquence du théorème local limite par ex)

Preuve : **Étape 1** : convergence sur $[0,s]$ avec $0 < s < 1$ fixé.

Soit $F : \mathcal{C}([0,s], \mathbb{R}) \to \mathbb{R}$ continue bornée. Alors

$$\mathbb{E}\left[F\left(\frac{S_{nt}}{\sigma\sqrt{n}} : 0 \leq t \leq s \right) \middle| S_n = 0 \right] = \mathbb{E}\left[F\left(\frac{S_{nt}}{\sigma\sqrt{n}} : 0 \leq t \leq s \right) \frac{q_{n-r_{ns}}(-S_{r_{ns}})}{q_n(0)} \right], \text{ avec } q_r(k) = \mathbb{P}(S_r = k).$$

D'après le théorème de représentation de Skorokhod, on peut supposer que

$$\left(\frac{S_{nt}}{\sigma\sqrt{n}} : 0 \leq t \leq s \right) \xrightarrow{p.s} (B_t : 0 \leq t \leq s).$$

En utilisant le théorème local limite, on obtient

1) $\exists C > 0$ t.q. $\forall n \geq 1$, $\frac{q_{n-r_{ns}}(-S_{r_{ns}})}{q_n(0)} \leq C$ (avec le corollaire (1) et (2) de la page précédente)

2) p.s. $\frac{q_{n-r_{ns}}(-S_{r_{ns}})}{q_n(0)} \xrightarrow[n\to\infty]{} \frac{p_1\left(-\frac{B_s}{\sqrt{1-s}}\right)}{\sqrt{1-s}\, p_1(0)} = \frac{p_{1-s}(-B_s)}{p_1(0)}$. (avec le corollaire (3))

Par convergence dominée, on conclut que

$$\mathbb{E}\left[F\left(\frac{S_{nt}}{\sigma\sqrt{n}} : 0 \leq t \leq s \right) \middle| S_n = 0 \right] \xrightarrow[n\to\infty]{} \mathbb{E}\left[F(B_t : 0 \leq t \leq s) \frac{p_{1-s}(-B_s)}{p_1(0)} \right] = \mathbb{E}\left[F(b_t : 0 \leq t \leq 1) \right].$$

Étape 2 : Convergence des fdd sur $[0,1]$

Ceci provient de l'étape 1 (avec le fait que en $t=1$ les 2 processus valent 0)

Étape 3 : Tension sur $[0,1]$

On utilise un argument de renversement des temps :

en posant $\hat{S}_i^n = S_n - S_{n-i}$, on a clairement $(S_i)_{0\le i\le n} \overset{\text{loi}}{=} (\hat{S}_i^n)_{0\le i\le n}$

et cette égalité en loi reste vraie sachant $S_n = 0$.

On en déduit que sachant $S_n = 0$,

$$S^{1,n} := \left(\frac{S_{nt}}{\sigma\sqrt{n}} : 0 \le t \le \frac{3}{4}\right) \xrightarrow[n\to\infty]{\text{loi}} \left(b_t : 0 \le t \le \frac{3}{4}\right)$$

$$S^{2,n} := \left(\frac{S_{n(1-t)}}{\sigma\sqrt{n}} : 0 \le t \le \frac{3}{4}\right) \xrightarrow[n\to\infty]{\text{loi}} \left(\tilde{b}_t : 0 \le t \le \frac{3}{4}\right)$$

avec $\tilde{b}$ pont brownien indépendant de b.

On en déduit que $S^{1,n}$ et $S^{2,n}$ sont tendus, et du fait que pour $\delta < \frac{1}{4}$,

$$w\left(\left(\frac{S_{nt}}{\sigma\sqrt{n}} : 0 \le t \le 1\right), \delta\right) = \max\left(w(S^{1,n}, \delta), w(S^{2,n}, \delta)\right),$$

on conclut que $\left(\frac{S_{nt}}{\sigma\sqrt{n}} : 0 \le t \le 1\right)$ sous $\mathbb{P}(\cdot \mid S_n = 0)$ est tendu.

$\rightsquigarrow$

VII : Quelques mots sur les processus discontinus : l'espace de Skorokhod

On note $\mathbb{D}$ (ou $\mathbb{D}([0,1],\mathbb{R}^d)$) l'ensemble des fonctions $x:[0,1]\to\mathbb{R}^d$ qui sont :

- continues à droite $(\forall t<1,\ x(t+)=\lim_{s\downarrow t} x(s)=x(t))$
- ayant des limites à gauche $(\forall t\leq 1,\ x(t-)=\lim_{s\uparrow t} x(s))$

Pour $x\in\mathbb{D}$, on dit que x est càdlàg et on note $\Delta x(t)=x(t)-x(t-)$ $(0\leq t\leq 1)$ et $\Delta x(0)=0$.

Les fonctions/processus càdlàg apparaissent naturellement en probabilités :

- fonctions de répartition
- Processus de Poisson
- processus de Lévy (càdlàg à accroissements stationnaires et indépendants), qui apparaissent naturellement dans l'extension du thm de Donsker à certaines v.a de variance ∞
- Toute martingale % à une filtration complète et càd admet une modif càdlàg

⚠ $(\mathbb{D},\|\cdot\|_\infty)$ 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 $\mathbb{D}$.

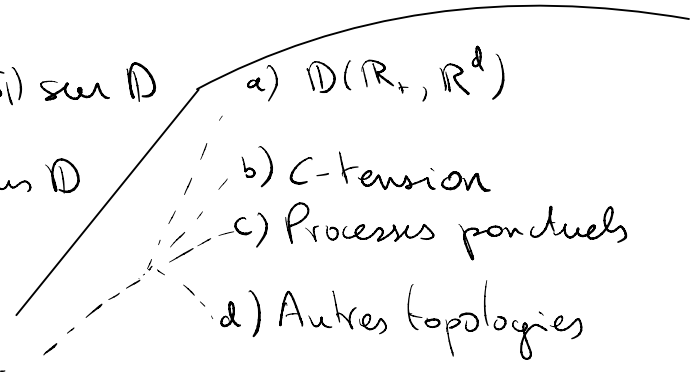
Plan : 1) Propriétés de base

2) Topologie de Skorokhod (J_1) sur $\mathbb{D}$

3) Convergence étroite dans $\mathbb{D}$

4) Critères de tension

5) Compléments facultatifs



Références : • Billingsley, Convergence of probability measures, 2nd Ed.

• Jacod-Shiraev, Limit theorems for stochastic processes. (Chap VI)

• Kallenberg, Foundations of Modern Probability, 2nd Ed.

Remarque On se place dans $\mathbb{D}([0,1], \mathbb{R}^d)$ pour simplifier, mais ce qui suit s'adapte sans soucis au cas de $\mathbb{D}([0,1], S)$ avec (S,d) polonais.

1) Propriétés de base

Ex : Soit $(r_i)_{i\geq 1}$ une énumération des rationnels de $]0,1[$. On a $x(t) = \sum_{i=1}^{\infty} \frac{1}{2^i} \mathbb{1}_{\{r_i \geq t\}} \in \mathbb{D}$.

Prop Soit $x \in \mathbb{D}$

1) $\forall \varepsilon > 0, \#\{t \in [0,1], |\Delta x(t)| > \varepsilon\} < \infty$

2) $\{t \in [0,1] : \Delta x(t) \neq 0\}$ est dénombrable

3) x est bornée

4) x est lim uniforme de fonctions en escalier

Idée de preuve : 1) Si le # est ∞, on prend une suite $\infty(t_n)$ dedans. Quitte à extraire, on suppose $t_n \to t$. On distingue les deux cas $t_n \downarrow t$, $t_n \uparrow t$ pour une contradiction.

2) Ok par 1)

3) Par l'absurde : si $|x(t_n)| > n$, on extrait $t_n \to t$ avec (t_n) monotone et on obtient $\max(|x(t)|, |x(t-)|) = \infty$, absurde

4) On introduit $w(x, A) = \sup_{s,t \in A} |x(t) - x(s)|$ pour $A \subset [0,1]$. On utilise alors :

Lemme $\forall \varepsilon > 0, \exists\, 0 = t_0 < t_1 < \cdots < t_r = 1$ tq $\max_{0 \leq i \leq r-1} w(x, [t_i, t_{i+1}[) \leq \varepsilon$.

Idée de preuve : Soit $\varepsilon > 0$. On note $T = \sup\{ t \in [0,1] : \exists\, 0 = t_0 < t_1 < \cdots < t_r = t,\ \max_{0 \le i \le r-1} (x, [t_i, t_{i+1}[) \le \varepsilon \}$

Alors $x(0) = x(0+) \Rightarrow T > 0$, $x(T-)$ existe $\Rightarrow$ le sup est atteint, $x(T) = x(T+) \Rightarrow T < 1$

est impossible

$\sim$

On conclut en prenant $x_n = \sum_{i=1}^{n-1} x(t_i)\, \mathbb{1}_{[t_i, t_{i+1}[}$ sur $[0,1[$ et $x_n(1) = x(1)$

$\sim$

On définit alors le "module de continuité modifié" :

Notation Pour $x : [0,1] \to \mathbb{R}^d$ et $\delta > 0$, on pose $w'(x, \delta) = \inf_{r \ge 1} \max_{0 \le i \le r-1} w(x, [t_i, t_{i+1}[)$

$0 = t_0 < t_1 < \cdots < t_r = 1$
tq $|t_{i+1} - t_i| > \delta$

Exemple : Soit $x_n = n\, \mathbb{1}_{[1-\varepsilon, 1]}$. Alors $w'(x_n, \delta) = 0$ si $\delta < \varepsilon$
$= n$ si $\delta \ge \varepsilon$

Proposition Soient $x, y : [0,1] \to \mathbb{R}^d$.

1) $x \in D \iff w'(x, \delta) \underset{\delta \to 0}{\longrightarrow} 0$ (à rapprocher de x continu $\iff w(x, \delta) \underset{\delta \to 0}{\longrightarrow} 0$)

2) $w'(x, \delta) \le 2 \|x - y\|_\infty + w'(y, \delta)$

3) Si $x \in D$, $\delta > 0$, (a) $w'(x, \delta) \le w(x, 2\delta)$

(b) $w(x, \delta) \le 2\, w'(x, \delta) + \sup_{0 \le t \le 1} |\Delta x(t)|$

Preuve laissée en exo (on a déjà vu $\Rightarrow$ pour 1)), cf Billingsley p 123

2) Topologie de Skorokhod $(T,\|\cdot\|)$ 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 Un changement de temps est une fonction $\lambda : [0,1] \to [0,1]$ t.q. $\lambda(0)=0$, $\lambda(1)=1$, λ est strictement croissante et continue. On note $\Lambda = \{\text{changements de temps}\}$.

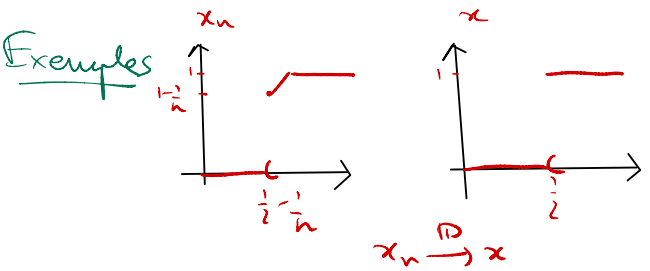
Définition Soit $x_n \in \mathbb{D}, x \in \mathbb{D}$. On dit que $x_n \to x$ au sens de Skorokhod si

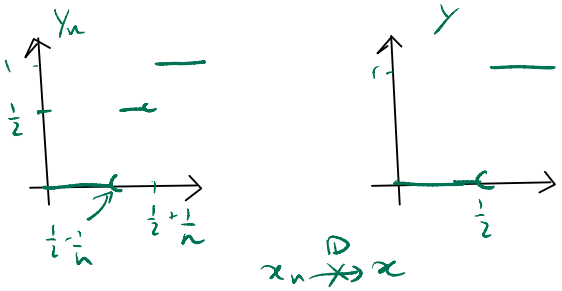
$\forall n \geq 1, \exists \lambda_n \in \Lambda$ t.q. $\|\lambda_n - Id\|_\infty \xrightarrow[n\to\infty]{} 0$ et $\|x_n \circ \lambda_n - x\|_\infty \xrightarrow[n\to\infty]{} 0$

(ou $\|x_n - x\circ\lambda_n\|_\infty \to 0$ de manière équivalente)

On note $x_n \xrightarrow{\mathbb{D}} x$

Exemples





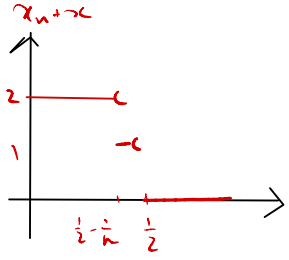
Remarque Pour $\lambda_n, \lambda \in \Lambda$, $\|\lambda_n - \lambda\|_\infty \to 0 \iff \lambda_n$ converge simplement (thm de Dini)

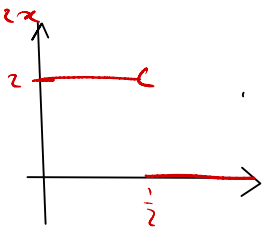
Remarque : Si on pose, pour $x, y \in \mathbb{D}$, $d_S(x,y) = \inf_{\lambda\in\Lambda} \max\left(\|\lambda - Id\|_\infty, \|x\circ\lambda - y\|\right)$,

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

⚠ $(\mathbb{D}, d_S)$ n'est pas une topologie d'espace vectoriel normé : si $x_n = \mathbb{1}_{[0,\frac{1}{2}-\frac{1}{n}[}$ et $x = \mathbb{1}_{[0,\frac{1}{2}[}$,

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 $\Pi_t : \begin{array}{ccc} \mathbb{D} & \to & \mathbb{R} \\ x & \mapsto & x(t) \end{array}$

Prop (exemples de fonctionnelles continues)

1) $x \mapsto \|x\|_\infty$ est continue pour d_S

2) $x \mapsto \sup_{0\le t\le 1} |\Delta x(t)|$ 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 $f:\mathbb{R}\to\mathbb{R}$ continue. $x \mapsto \int_0^1 f(x(s))\,ds$ est continue pour d_S

Preuve : Soit $x_n \xrightarrow{\mathbb{D}} x$, et $\lambda_n\in\Lambda$ tq $x_n\circ\lambda_n \xrightarrow{CVU} x$ et $\lambda_n \xrightarrow{CVU} Id$.

1) On a

$$\big|\|x_n\|_\infty - \|x\|_\infty\big| = \big|\|x_n\circ\lambda_n\|_\infty - \|x\|_\infty\big| \le \|x_n\circ\lambda_n - x\|_\infty \longrightarrow 0$$

2) On a $|\Delta x(t)| = |x(t)-x(t-)| \le 2\|x\|_\infty$, et puisque $\sup_{0\le t\le 1}|\Delta x_n(t)| = \sup_{0\le t\le 1}|\Delta x_n\circ\lambda_n(t)|$

$$\Big|\sup_{0\le t\le 1}|\Delta x_n(t)| - \sup_{t\in[0,1]}|\Delta x(t)|\Big| \le \sup_{t\in[0,1]}|\Delta(x_n\circ\lambda_n - x)| \le 2\|x_n\circ\lambda_n - x\|_\infty \longrightarrow 0.$$

3) Ceci provient du fait que pour tout $\lambda\in\Lambda$ on a $\lambda(0)=0$ et $\lambda(1)=1$

4) On écrit

$$|x_n(t)-x(t)| \le \|x_n - x\circ\lambda_n^{-1}\|_\infty + |x\circ\lambda_n^{-1}(t) - x(t)| \xrightarrow[n\to\infty]{} 0$$

par continuité de x en t.

5) Soit $x_n \xrightarrow{\mathbb{D}} x$. Comme $\sup\|x_n\|_\infty \vee \|x\|_\infty < \infty$, on peut supposer f bornée.

Posons $D_x = \{t\in[0,1],\ \Delta x(t)\neq 0\}$, dénombrable. Alors

$$\int_a^b f(x_n(t))\,dt = \int_a^b \underbrace{f(x_n(t))\,\mathbb{1}_{t\notin D_x}}_{\xrightarrow[n\to\infty]{CVS} f(x(t))\mathbb{1}_{t\notin D_x}}\,dt.$$

On conclut par convergence dominée

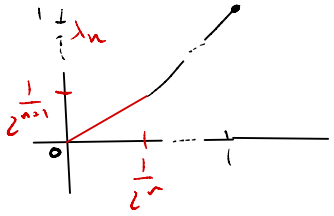
Remarques En général les fonctions $f_1 : x \mapsto x(t)$, $f_2 : x \mapsto w'(x,\delta)$ et $f_3 : x \mapsto \sup_{a \le t \le b} |\Delta x(t)|$ pour $0 \le a \le b \le 1$ ne sont pas continues (bien que $x \mapsto w(x,\delta)$ est continue pour $x \in \mathcal{C}([0,1])$).

Par exemple, si $x_n = \mathbb{1}_{[0, \delta + \frac{1}{2^n}[}$ et $x = \mathbb{1}_{[0,\delta[}$, $x_n \xrightarrow{D} x$ mais $x_n(\delta) = 1 \not\to x(\delta) = 0$,

$w'(x_n, \delta) = 0 \not\to w'(x,\delta) = 1$ et $x_n(\delta) - x_n(\delta -) = 0 \not\to x(\delta) - x(\delta -) = 1$

On peut cependant montrer que f_2 est semi-continue inférieurement et que f_3 est semi-continue supérieurement.

Exemple Soit $x_n = \mathbb{1}_{[0, \frac{1}{2^n}]}$. Si on considère λ_n :



On voit que $d_S(x_n, x_{n+1}) \le \frac{1}{2^n}$.

En particulier, (x_n) est de Cauchy mais ne converge pas (la seule limite possible est la fonction nulle). Ainsi, (D, d_S) n'est pas complet.

Cependant :

Théorème Il existe une distance d'_S sur D telle que les topologies engendrées par d_S et d'_S sont les mêmes, et (D, d'_S) est complet.

Cf Billingsley, Theorem 12.2.

3) Convergence étroite dans $\mathbb{D}$

Proposition La tribu produit sur $\mathbb{D}$ (plus petite tribu rendant les applications $\Pi_t : x \mapsto x(t)$ mesurables) est la tribu borélienne de $(\mathbb{D}, d_S)$

Le fait que Π_t soit mesurable provient du fait que $\Pi_t^p : x \mapsto p\int_t^{t+1/p} x(s)\,ds$ est continue, donc mesurable, puis que $\Pi_t = \lim_{p\to\infty} \Pi_t^p$ par continuité à droite

Pour montrer que la tribu borélienne de $(\mathbb{D}, d_S)$ est incluse dans la tribu produit, on procède un peu comme dans le cas continu

Remarque La tribu produit est aussi la plus petite tribu rendant mesurables les applications $(\Pi_t)_{t\in D}$ où $D \subset [0,1]$ dense contenant 1.

En effet, $\Pi_t = \lim_{\substack{s\in D \\ s\searrow t}} \Pi_s$ (ponctuellement)

Ainsi, les fidi caractérisent les mesures de $M_1(\mathbb{D})$. Cependant :

⚠ La cv des fidi n'est pas impliquée par la cv en loi (les Π_t ne sont pas continus)

Définition Soit X v.a 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.

Prop Si X v.a dans $\mathbb{D}$, D_X est au plus dénombrable

Idée de preuve : montrer que $\forall \varepsilon, \delta > 0$, $D_X^{\varepsilon,\delta} = \{t \in [0,1] : \mathbb{P}(|\Delta X_t| > \varepsilon) \geq \delta\}$ est fini.

Théorème Soit $X^n, X \in \mathbb{D}$ des v.a. Alors $X^n \xrightarrow{\text{loi}} X$ ssi

(i) $X^n \xrightarrow{\text{fidi}} X$ en dehors de $D_X \cup \{0\}$, i.e $\forall t_1, \dots, t_k \in ([0,1] \setminus D_X) \cup \{0\}$

$(X^n_{t_1}, \dots, X^n_{t_k}) \xrightarrow{\text{loi}} (X_{t_1}, \dots, X_{t_k})$

(ii) (X^n) est tendue dans $M_1(\mathbb{D})$

Preuve : $\Rightarrow$ (ii) provient du thm de Prokhorov

Pour (i), on remarque que si $t \notin D_x$, alors $\mathbb{P}_X(dx)$ ps, Π_t est continue en x.

$\Leftarrow$ Par le thm de Prokhorov, il suffit de mq si $X^{\varphi(n)} \xrightarrow{\text{loi}} Y$,

alors $X \overset{\text{loi}}{=} Y$.

Pour cela, par (i), on a que X et Y ont les fidi sur $([0,1] / D_X \cup D_Y) \cup \{0\}$

et ont donc même loi

4) Critères de tension

L'analogue du théorème d'Arzela-Ascoli dans $\mathbb{D}$ est :

Théorème Une famille $\mathcal{F} \subset \mathbb{D}$ est relativement compacte ssi

(1) $\sup_{x \in \mathcal{F}} \|x\|_\infty < \infty$

(2) $\forall \varepsilon > 0, \exists \delta > 0, \sup_{x \in \mathcal{F}} w'(x, \delta) \leq \varepsilon$

Comme pour les fonctions continues, on en déduit

Théorème Soit $(X^n)_{n \geq 1}$ de v.a de $\mathbb{D}$ est tendue ssi

(1) $(\|X^n\|_\infty)_{n \geq 1}$ est tendue dans $\mathbb{R}$

(2) $\forall \varepsilon, \eta > 0, \exists \delta > 0, \limsup_{n \to \infty} \mathbb{P}(w'(X^n, \delta) > \eta) \leq \varepsilon$

Remarque : On ne peut pas remplacer (1) par "$(x_n(0))_{n \geq 1}$ est tendu" comme dans $\mathcal{C}([0,1], \mathbb{R}^d)$, prendre par exemple $X^n = n \mathbb{1}_{[U,1]}$ avec U uniforme sur $[0,1]$

On peut cependant démontrer que lorsque (2) est vérifié, (1) $\Leftrightarrow$ (1') $\Leftrightarrow$ (1'') avec (1') et (1'') définis par :

(1') $\exists T \subset [0,1]$ dense contenant 1 t.q $\forall t \in T, (X^n(t))_{n \geq 1}$ est tendu

(1'') $(X^n(0))_{n \geq 1}$ et $\left(\sup_{0 \leq t \leq 1} |\Delta X^n(t)| \right)_{n \geq 1}$ sont tendues.

(cf Billingsley p140)

Application : Soit $(U_i)_{i \geq 1}$ des va iid uniformes sur $[0,1]$ et $\lambda > 0$. On pose pour $t \in [0,1]$ $X_t^{(n)} = \sum_{k=1}^{\lfloor nt \rfloor} \mathbb{1}_{U_k \leq \frac{\lambda}{n}}$. 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 : • $\mathrm{Poi}(N^\lambda$ n'a pas de discontinuités fixes). Soit $0<t_1<\dots<t_k\le 1$.

Alors les $(X^{(n)}_{\lfloor nt_{i+1}\rfloor} - X^{(n)}_{\lfloor nt_i\rfloor})$ sont $\perp\!\!\!\perp$ à n fixé, de loi $\mathrm{Bin}(\lfloor nt_{i+1}\rfloor - \lfloor nt_i\rfloor, \frac{\lambda}{n}) \xrightarrow{\text{loi}} \mathrm{Poi}(\lambda(t_{i+1}-t_i))$.

• tension * $\|X^{(n)}\|_\infty \overset{\text{loi}}{=} \mathrm{Bin}(n, \frac{\lambda}{n})$ converge en loi, donc est tendue

* contrôle de $w'(X^n, \delta)$.

On a $\limsup_{n\to\infty} \mathbb{P}(\underbrace{\exists \text{ deux sauts de } X^n \text{ espacés d'au plus } \delta}_{A_n(\delta)}) \xrightarrow[\delta\to 0]{} 0$, ce qui suffit car

$\{w'(X_n,\delta)>\eta\} \subset A_n(\delta)$.

Pour cela on regarde $\mathbb{P}(A_n(\delta) \mid X^n_1 = k)$.

Sachant $X^n_1 = k$, les sauts de X^n sont un ss-ensemble uniforme à k éléments de $\{\frac{1}{n}, \frac{2}{n}, \dots, \frac{n}{n}\}$.

Si on a deux sauts espacés de moins de δ, on peut retirer un tel saut qui est le plus à droite, et marquer le saut à distance $\le\delta$ le plus proche.

Cela permet de voir que (choix de $k-1$ sauts ; choix du saut marqué ; ajout d'un saut à distance $\le\delta$ à sa droite)

$$\mathbb{P}(A_n(\delta) \mid X^n_1 = k) \le \frac{\binom{n}{k-1}(k-1)(\delta n)}{\binom{n}{k}}$$

$$\le \delta k^3 \quad \text{(après calculs)}.$$

Ainsi, $\mathbb{P}(A_n(\delta)) \le \delta \sum_{k\ge 0} k^3\, \mathbb{P}(\mathrm{Bin}(n,\frac{\lambda}{n}) = k) \le C\delta$,

d'où le résultat.

$\rightsquigarrow$

On conclut par plusieurs autres critères de tension. Le premier est un équivalent du critère de Kolmogorov (cf Billingsley, Theorem 13.5)

Théorème (critère de Billingsley) Soit $(X^n)_{n\ge 1}$ une suite de v.a. dans $\mathbb{D}$. On suppose que :

(1) $(\|X^n\|_\infty)_{n\ge 1}$ est tendue

(2) Il existe $a, b > 0$ et une fonction croissante $F : [0,1] \to \mathbb{R}_+$ telle que

$$\forall n \geq 1,\ \forall\, r \leq s \leq t,\ \mathbb{E}\left[|X^n(t) - X^n(s)|^a \, |X^n(s) - X^n(r)|^a \right] \leq (F(t) - F(r))^{1+b}$$

(3) $\forall \varepsilon > 0$,

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \left(\mathbb{P}(|X^n(\delta) - X^n(0)| \geq \varepsilon) + \mathbb{P}(|X^n(1) - X^n(1-\delta)| \geq \varepsilon) \right) = 0$$

Remarques : • la condition (3) est nécessaire : prendre par exemple $X^n = \mathbb{1}_{[0, \frac{1}{n}[}$

• La condition (3) est satisfaite si pour un certain $X \in \mathbb{D}$, $X_n \xrightarrow{\text{fidi}} X$ au sens des fidi sur $[0,1] \setminus D_X$ et si X est p.s. continu en 1.

On conclut par le critère d'Aldous, particulièrement utile pour des marches aléatoires ou plus généralement des martingales.

Pour $X^n \in \mathbb{D}$, on note $(\mathcal{F}^n_t)_{0 \leq t \leq 1}$ sa filtration naturelle. On note

$\mathcal{T}^n = \{$ temps d'arrêt pour $\mathcal{F}^n$ qui sont $\leq 1 \}$

Théorème (Aldous) On suppose que

(1) $(\|X^n\|_\infty)_{n \geq 1}$ est tendue

(2) $\forall$ suite $h_n \to 0$, $\forall$ 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

Cf par exemple Kallenberg, Theorem 16.11 pour une preuve.

Ce théorème permet par exemple de démontrer le résultat spectaculaire suivant :

Théorème Soit $(X_i^n)_{i,n\geq 1}$ des v.a. t.q. $\forall n \geq 1$, $(X_i^n)_{i\geq 1}$ sont i.i.d. On pose $S_k^n = X_1^n + \cdots + X_k^n$ et on suppose qu'il existe une suite $m_n \to \infty$ d'entiers t.q. $S_{m_n}^n$ 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 (i.e un processus de Lévy) tel que $X_1 \overset{\text{loi}}{=} X$ et

$$\left(S^n_{\lfloor m_n t \rfloor} : 0 \leq t \leq 1\right) \xrightarrow{\text{loi}} (X_t)_{0\leq t\leq 1}$$

(cf Kallenberg, Theorem 16.14)

Si $(X_i)_{i\geq 1}$ sont i.i.d centrés de variance finie, en prenant $X_i^n = \frac{X_i}{\sqrt{n}}$ et $m_n = n$, c'est essentiellement le théorème de Donsker.

5) Compléments facultatifs

a) $\mathbb{D}(\mathbb{R}_+, \mathbb{R})$

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 $\mathbb{D}_\infty = \mathbb{D}([0,\infty[, \mathbb{R}^d)$.

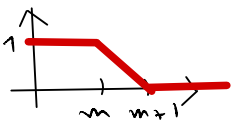
Déf On dit que $x_n \xrightarrow{\mathbb{D}_\infty} x$ si $\forall n \geq 1$, $\exists \lambda_n \in \Lambda_\infty = \{$ homéomorphismes croissants de $\mathbb{R}_+$ dans $\mathbb{R}_+\}$ t.q.

1) $\|\lambda_n - Id\|_\infty \xrightarrow[n\to\infty]{} 0$

2) $\forall m \geq 1$, $\|(x_n \circ \lambda_n - x)\mathbb{1}_{[0,m]}\|_\infty \xrightarrow[n\to\infty]{} 0$

⚠ $\mathbb{D}_\infty \to \mathbb{D}_m$, $x \mapsto x\mathbb{1}_{[0,m]}$ n'est pas continue dans $\mathbb{D}_\infty$ (si x a un saut en m)

Pour définir une distance d_∞ t.q. $d_\infty(x_n, x) \to 0$ ssi $x_n \xrightarrow{\mathbb{D}_\infty} x$, on introduit $\forall m \geq 1$ la fonction $g_m(t)$:



et on pose pour $x, y \in \mathbb{D}_\infty$:

$$d_\infty(x,y) = \sum_{m\geq 0} \frac{d_m(g_m x, g_m y) \wedge 1}{2^m}$$

et on vérifie que $(\mathbb{D}_\infty, d_\infty)$ est polonais.

Critères de tension : on remplace w' par des w'_m, $m \geq 1$

où $w'_m(x,\delta) = \inf_{\substack{0=t_0<t_1<\dots<t_r=m \\ |t_{i+1}-t_i|>\delta \\ \text{pour } 0\leq i\leq r-1}} w(x, [t_i, t_{i+1}[)$

[Cf Billingsley, Section 16]

b) C-tension

On remarque que $\mathcal{C} = \mathcal{C}([0,1], \mathbb{R}^d)$ est un fermé de $\mathbb{D}$, car c'est

$\{x \in \mathbb{D} : \sup_{0 \le t \le 1} |\Delta x(t)| = 0\}$.

Définition Une suite (X^n) de v.a de $\mathbb{D}$ est dite C-tendue si elle est tendue dans $\mathbb{D}$ ($\mathbb{D}$-tendue) et si toute valeur d'adhérence est portée par $\mathcal{C}$. (i.e si $X^{\varphi(n)} \xrightarrow{\text{loi}} X$, alors $\mathbb{P}(X \in \mathcal{C}) = 1$)

Proposition Soit X^n, X des v.a dans $\mathbb{D}$ tq $X^n \xrightarrow{\text{loi}} X$. On suppose que $\sup_{0 \le t \le 1} |\Delta X^n(t)| \xrightarrow{\mathbb{P}} 0$. Alors $\mathbb{P}(X \in \mathcal{C}) = 1$

Cf Billingsley, theorem 13.4 p 142

Théorème Soit (X^n) une suite de v.a de $\mathbb{D}$. Elle est $\mathcal{C}$-tendue ssi (1) $(X_0^n)_{n \ge 1}$ est tendue dans $\mathbb{R}$

(2) $\forall \varepsilon, \eta > 0$, $\exists \delta > 0$, $\limsup_{n \to \infty} \mathbb{P}(w(X^n, \delta) > \eta) \le \varepsilon$

Cf Billingsley, Corollary p 142

On en déduit aisément une version càdlàg du théorème de Donsker :

Si $(X_i)_{i \ge 1}$ sont iid, centrés, de variance $\sigma^2 \in]0, \infty[$, en notant $S_n = X_1 + \dots + X_n$, on a

$$\left(\frac{S_{\lfloor nt \rfloor}}{\sigma \sqrt{n}} : 0 \le t \le 1 \right) \xrightarrow{\text{loi}} (B_t : 0 \le t \le 1),$$

où la convergence a lieu en loi dans $\mathbb{D}$.

c) Processus ponctuels

Dans le cas des processus ponctuels, la convergence des fidi implique la convergence en loi

Théorème On note D_c l'ensemble des fonctions de comptage sur $[0,1]$, i.e. des fonctions càdlàg, à valeurs dans $\mathbb{Z}$, croissantes, dont les sauts valent 1 aux points de discontinuités. Soit X^n, X v.a à valeurs dans D_c.

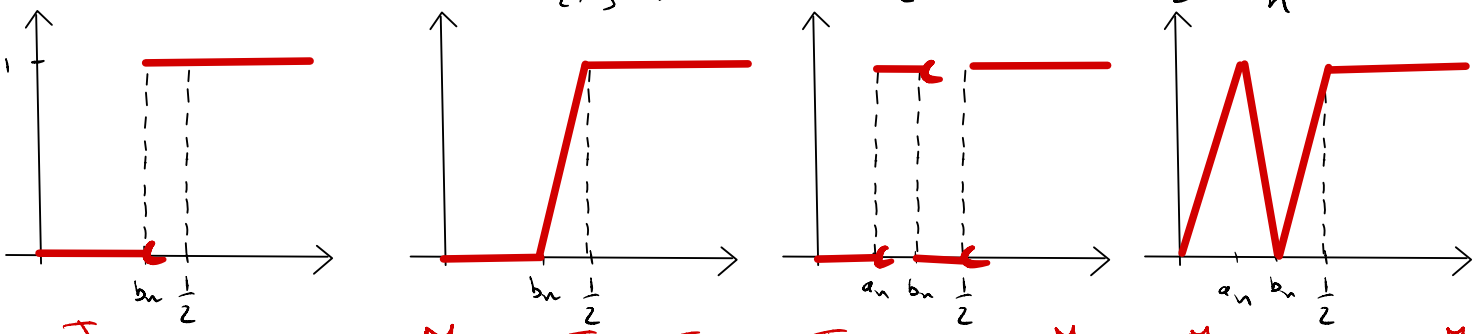
Alors $X^n \xrightarrow{\text{loi}} X$ ssi $\exists T \subset [0,1]$ dénombrable dense tq

$$\forall t_1, \dots, t_k \in T, \quad (X^n_{t_1}, \dots, X^n_{t_k}) \xrightarrow{\text{loi}} (X_{t_1}, \dots, X_{t_k})$$

(cf Billingsley, thm 12.6)

d) Autres topologies

Nous avons considéré la topologie de Skorokhod J_1 sur $D([0,1],\mathbb{R})$, mais il en existe trois autres (J_2, M_1, M_2) pour lesquels les convergences suivantes ont lieu, avec $x = \mathbb{1}_{[\frac{1}{2},1]}$, $a_n = \frac{1}{2} - \frac{1}{2n}$, $b_n = \frac{1}{2} - \frac{1}{n}$:



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

cf W. Whitt, Stochastic process limits, Chap 11.5

VIII Mesures aléatoires de Poisson

Un des objectifs est de définir une notion de "pluie" dans un espace assez général, où les gouttes arrivent indépendamment et aléatoirement 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).

Plan :
1) Rappels sur les lois de Poisson
2) Mesures aléatoires de Poisson
3) Points multiples
4) Formules de Campbell
5) Transformations de nuages poissoniens
6) Nuages de Poisson avec paramètre temporel.
7) Formule de Palm

Références : • Kingman, Poisson processes

• Last, Penrose, Lectures on the Poisson process

1) Rappels sur les lois de Poisson

Pour $\lambda \in [0, \infty]$, on dit qu'une v.a. 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 $\lambda = 0$, $X = 0$ p.s.
- Pour $\lambda = \infty$, $X = \infty$ p.s.
- Pour $0 < \lambda < \infty$, $\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$ 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 très utiles

Proposition

- [Superposition] Si $(\lambda_i)_{i\in I}$ est une famille au plus dénombrable de $[0,\infty]$ et $(X_i)_{i\in I}$ des v.a indépendantes avec $X_i \sim \mathcal{P}(\lambda_i)$, alors $\sum_{i\in I} X_i \sim \mathcal{P}(\sum_{i\in I}\lambda_i)$
- [Marquage] Soit $X\sim\mathcal{P}(\lambda)$ avec $\lambda\in[0,\infty]$. Soit $(U_k)_{k\geq 1}$ des v.a 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 v.a. $X_i = \sum_{k=1}^{X} \mathbb{1}_{U_k=i}$ pour $i\geq 1$ sont indépendantes et $X_i\sim\mathcal{P}(\lambda p_i)$ (avec $0\times\infty=0$)

La preuve, pouvant être faite en utilisant les fonctions caractéristiques est laissée en exercice.

2) Mesures aléatoires de Poisson

On se place dans le cadre assez général d'un espace mesuré $(E,\mathcal{E})$.

Déf 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

Ex 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 (pas le cas des mesures σ-finies)

Remarque On peut vérifier que les théorèmes de Fubini s'appliquent pour des mesures s-finies en utilisant le théorème de convergence monotone (cf theorem A.B dans Last & Penrose)

Déf On note $\mathcal{M}_{atom}(E)$ 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 $\mathcal{M}_{atom}(E)$ de la plus petite tribu rendant mesurables les applications $\mu \mapsto \mu(A)$ pour tout $A \in \mathcal{E}$

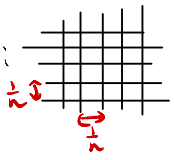
$$\mathcal{M}_{atom}(E) \to \mathbb{N} \cup \{+\infty\}$$

Exemples 1) $N : \Omega \to \mathcal{M}_{atom}(E)$ est une v.a. ssi $\forall A \in \mathcal{E}, \forall k \geq 0, \{N(A) = k\}$ est mesurable

(car $\{N(A) = \infty\} = \left(\bigcup_{k=1}^{\infty} \{N(A) = k\}\right)^c$)

2) Pour $x \in E$, $E \to \mathcal{M}_{atom}(E)$, $x \mapsto \delta_x$ est mesurable

3) Pour $A \in \mathcal{E}$, $\mathcal{M}_{atom}(E) \to \mathcal{M}_{atom}(E)$, $\mu \mapsto \mu_{\mid A}$ est mesurable, où $\mu_{\mid A}$ est défini par $\mu_{\mid A}(B) = \mu(A \cap B)$ 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 $\mathbb{R}^2$:  et de choisir un point dans chaque cellule indépendamment avec probabilité $\frac{1}{n^2}$.

Alors le nombre de points qui tombent dans une région A est $\sim \mathrm{Bin}(n^2 \mathrm{Aire}(A), \frac{1}{n^2})$ qui converge en loi vers $\mathcal{P}(\mathrm{Aire}(A))$, et le nombre de points qui tombent dans des régions disjointes sont indépendants

Définition Soit $(E, \mathcal{E})$ espace mesuré et μ s-finie sur E. Une mesure aléatoire de Poisson (ou nuage poissonien) d'intensité μ est une variable aléatoire N à valeurs dans $\mathcal{M}_{atom}(E)$ telle que

① $\forall A \in \mathcal{B}(E)$, $N(A) \sim \mathcal{P}(\mu(A))$

② Si $(A_i)_{1 \leq i \leq k}$ est une famille de parties mesurables disjointes de E, alors les v.a. $(N(A_i))_{1 \leq i \leq k}$ sont indépendantes.

Remarques : • Presque sûrement, $N(A) = \infty \Leftrightarrow \mu(A) = \infty$

• si les singletons sont mesurables dans E, si $\mu(\{x\}) > 0$, $\mathbb{P}(N(\{x\}) \geq 2) > 0$.

• $\mathbb{E}[e^{-\lambda N(A)}] = e^{-\mu(A)(1-e^{-\lambda})}$

Théorème Pour toute mesure $\mu \in \mathcal{M}_s(E)$, il existe un nuage poissonien d'intensité μ, unique en loi dans $\mathcal{M}_{atom}(E)$.

Pour démontrer ce théorème, on a recours au lemme utile suivant

Lemme (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 $\mu = \sum_{i \geq 1} \mu_i$

Preuve : • Soit $A \in \mathcal{E}$. Par principe de superposition des lois de Poisson,

$$N(A) \sim \mathcal{P}\left(\sum_{i \in I} \mu_i(A)\right) = \mathcal{P}(\mu(A))$$

• Si $(A_j)_{1 \leq j \leq k}$ sont des parties mesurables disjointes, les v.a. $(N_i(A_j))_{i \geq 1, 1 \leq j \leq k}$ sont indépendantes. Par regroupement par paquets on en déduit que $\left(\sum_{i=1}^{\infty} N_i(A_j)\right)_{1 \leq j \leq k}$ sont indépendants

$\sim$

Preuve de l'existence des nuages poissoniens : d'après le lemme précédent, il suffit de traiter le cas où μ est une mesure finie.

Soit $L \sim \mathcal{P}(\mu(E))$ et $(X_i)_{i \geq 1}$ des v.a. iid de loi $\frac{\mu(\cdot)}{\mu(E)}$ indépendantes de L

On pose $$M = \sum_{i=1}^{L} \delta_{X_i}$$

Vérifions que M convient. C'est bien une v.a. (cf exemple 2 page 3)

Soit $(A_j)_{1\le j\le k}$ des parties mesurables disjointes de E et posons $B = \left(\bigcup_{j=1}^{k} A_j\right)^c$. On définit pour $i \ge 1$

$$U_i = \begin{cases} j & \text{si } X_i \in A_j \\ 0 & \text{si } X_i \in B \end{cases}$$

Les variables aléatoires $(U_i)_{i\ge 1}$ sont i.i.d, donc d'après le principe de raréfaction pour les lois de Poisson, les v.a

$$N(A_j) = \sum_{i=1}^{L} \mathbb{1}_{X_i \in A_j} = \sum_{i=1}^{L} \mathbb{1}_{U_i = j}$$

sont indépendantes et $N(A_j) \sim \mathcal{P}\left(\mu(E)\cdot \frac{\mu(A_j)}{\mu(E)}\right) = \mathcal{P}(\mu(A_j))$.

$\sim$

Preuve de l'unicité en loi des nuages poissoniens

Les parties de la forme $\{\mu \in M_{atom}(E) : \mu(A_1) = k_1, \dots, \mu(A_\ell) = k_\ell\}$ avec A_i mesurable et $k_i \ge 0$ forment un π-système générateur de $M_{atom}(E)$. Or une telle partie s'écrit comme réunion finie disjointe de parties de la forme $\{\mu \in M_{atom}(E) : \mu(B_1) = j_1, \dots, \mu(B_n) = j_n\}$ (*)

avec (B_i) mesurables disjoints et $j_i \ge 0$. Ainsi une mesure de probabilité sur E est caractérisée par ses valeurs sur les ensembles de la forme (*). Dans le cas d'un nuage poissonien, si deux nuages poissoniens vérifient ① et ②, leurs lois sont égales sur les ensembles de la forme (*), donc sont les mêmes

$\sim$

Remarque La preuve précédente montre que si μ est s-finie, il existe des variables aléatoires $(X_i)_{i\geq 1}$ à valeurs dans E et K à valeurs dans $\mathbb{N}\cup\{+\infty\}$ toutes indépendantes telles que

$$M = \sum_{i=1}^{K} \delta_{X_i} \quad (*)$$

est un nuage poissonien d'intensité μ. (avec $K=\infty$ ps si $\mu(E)=\infty$)

⚠ les X_i n'ont en général pas la même loi

⚠ Si M est un nuage poissonien, en général il ne s'écrit pas sous la forme (*) car en général, on ne peut pas énumérer les atomes de manière mesurable.

En effet, soit $E=[0,1]$ et supposons par l'absurde qu'il existe des applications mesurables $\pi_i : \mathcal{M}_{atom}(E) \to E$ et $L : \mathcal{M}_{atom}(E) \to \mathbb{N}\cup\{+\infty\}$

$$\forall \mu \in \mathcal{M}_{atom}(E),\quad \mu = \sum_{i=1}^{L(\mu)} \delta_{\pi_i(\mu)}$$

Soit λ la mesure de Lebesgue sur $[0,1]$ et $\mu = \sum_{i=1}^{\infty} \lambda \in \mathcal{M}_S([0,1])$.

Ainsi $\mu(A)=0$ ssi $\lambda(A)=0$ et $\mu(A)=\infty$ ssi $\lambda(A)>0$.

Soit M un nuage poissonien d'intensité μ. Ainsi, $\forall A\in\mathcal{B}([0,1])$ ps $M(A)=0$ ou ∞. Ainsi $2M \overset{\text{loi}}{=} M$ et on vérifie que $M \perp\!\!\!\perp M$. Alors

$$(\pi_i(2M))_{i\geq 1} \perp\!\!\!\perp (\pi_i(2M))_{i\geq 1}$$

Puisque μ est sans atome, $\forall i\geq 1$, $\forall x\in[0,1]$, $\mathbb{P}(\pi_i(2M)=x)=0$.

Par Fubini, $\forall i\neq j$, $\mathbb{P}(\pi_i(2M)=\pi_j(2M))=0$.

Donc $\mathbb{P}(\forall i\neq j, \pi_i(2M)\neq\pi_j(2M))=1$, absurde.

⇝ il faut des hypothèses additionnelles de "finitude"

- Si E est un sous-ensemble mesurable d'un espace métrique polonais, on notera $\mathcal{M}_{atom}^{<\infty}(E)$ les mesures finies de $\mathcal{M}_{atom}(E)$,

il est possible de numéroter de manière mesurable les atomes des mesures de $\mathcal{M}^{<\infty}_{\text{atom}}(E)$ (Proposition 6.2 dans Last & Penrose)

On en déduit que si M est un nuage poissonien d'intensité μ σ-finie sur E polonais, alors on peut écrire

$$M = \sum_{i=1}^{L(M)} \delta_{\pi_i(M)} \quad \text{avec } L, \pi_i \text{ mesurables.}$$

(cf Corollary 6.5 dans Last & Penrose)

Toutefois:

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 (*) ci-dessus, construite comme dans la preuve de l'existence de nuages poissoniens

Exemple (un processus de Poisson est un nuage poissonien)

Soit $\lambda > 0$ et $(X_i)_{i\geq 1}$ des v.a iid $\mathcal{E}xp(\lambda)$. On pose $M = \sum_{k=1}^{\infty} \delta_{X_1 + \dots + X_k}$. Alors M est un nuage poissonien d'intensité $\lambda \mathbb{1}_{x\geq 0}\, dx$.

Preuve: D'après la propriété de sans mémoire de la loi exponentielle, les v.a $(M_{[i,i+1[})_{i\geq 0}$ sont iid. Il suffit donc de vérifier que $M_{[0,1[} \overset{\mathcal{L}}{=} M_{[0,1]}$ est un nuage poissonien d'intensité $\lambda \mathbb{1}_{[0,1]}(x)\, dx$.

On sait qu'un tel nuage a la loi de $\sum_{i=1}^{L} \delta_{U_i}$ avec $L \sim \mathcal{P}oi(\lambda)$ et $(U_i)_{i\geq 1}$ iid sur $[0,1]$. Vérifions que $M_{[0,1]}$ est de cette forme.

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

$$\mathbb{P}(M([0,1])=n) = \mathbb{P}(S_n \le 1, S_{n+1} > 1)$$

$$= \int_{\{x_i>0\}} \prod_{i=1}^{n+1} (\lambda e^{-\lambda x_i} dx_i) \, \mathbb{1}_{\{x_1+\cdots+x_n \le 1 < x_1+\cdots+x_{n+1}\}}$$

$$\overset{[Y_i = x_1+\cdots+x_i]}{=} \int_{\{0\le y_1 \le \cdots \le y_{n+1}\}} \lambda^{n+1} e^{-\lambda y_{n+1}} \mathbb{1}_{\{y_n \le 1 < y_{n+1}\}} \, dy_1 \, dy_2 \cdots dy_{n+1}$$

$$= \lambda^n \int_{\{0\le y_1 \le \cdots \le y_n \le 1\}} dy_1 \cdots dy_n \; e^{-\lambda}$$

$$= \frac{\lambda^n}{n!} \int_{0 \le y_1, \ldots, y_n \le 1} dy_1 \cdots dy_n \; e^{-\lambda}$$

$$= \frac{\lambda^n}{n!} e^{-\lambda}.$$

Donc $M([0,1]) \sim \mathcal{P}(\lambda)$.

Déterminons la loi de $(S_1, \ldots, S_n)$ sachant $M([0,1]) = n$:

$$\mathbb{E}[g(S_1,\ldots,S_n) \mid M([0,1])=n] = \frac{\displaystyle\int_{\{x_i>0\}} \prod_{i=1}^{n+1} (\theta e^{-\theta x_i} dx_i) \, g(x_1, x_1+x_2, \ldots, x_1+\cdots+x_n) \, \mathbb{1}_{x_1+\cdots+x_n \le 1 < x_1+\cdots+x_{n+1}}}{\frac{\lambda^n}{n!} e^{-\lambda}}$$

même type de calculs que juste avant $= [\cdots] = n! \displaystyle\int_{0 \le y_1 < \cdots < y_n \le 1} dy_1 \cdots dy_n \, g(y_1, \ldots, y_n)$

Ainsi, $(S_1, \ldots, S_n)$ sachant $M([0,1])=n$ a la loi des statistiques d'ordre de n v.a. uniformes iid sur $[0,1]$, ce qui conclut. $\sim$

Proposition (Stabilité par restriction) Soit M nuage poissonien d'intensité μ. Soit $(A_i)_{i\in I}$ des parties disjointes de E. Alors $(M_{|A_i})_{i\in I}$ sont des nuages poissoniens $\perp\!\!\!\perp$ 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 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

Preuve : On écrit $E = \bigcup_{i=1}^{\infty} A_i$ avec $(A_i)_{i\geq 1}$ 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 avec $M_{|A_i}$ d'intensité $\mu_{|A_i}$, finie.

∞

Remarque Si μ est s-finie, ce n'est pas clair qu'on peut décomposer M de manière "mesurable" $M = \sum_{i=1}^{\infty} M_i$ avec $(M_i)_{i\geq 1}$ mesures aléatoires de Poisson indépendantes d'intensités finies. Mais lorsque μ est σ-finie, on vient de voir que c'est possible.

3) Points multiples

On suppose que $\{x\} \in \mathcal{E}$ $\forall x \in E$.

Les atomes d'une mesure $\sum_{i \in I} \delta_{x_i} \in \mathcal{M}_{atom}(E)$ 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 voudrait 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 On note

$$\mathcal{M}_{atom}^{simple}(E) = \left\{ \mu = \sum_{i \in I} \delta_{x_i} \in \mathcal{M}_{atom}(E) : x_i \neq x_j \text{ pour } i \neq j \right\}$$

Rappelons que $\mathcal{M}_{atom}(E)$ est muni de la plus petite tribu pour laquelle les $\begin{array}{ccc} \mathcal{M}_{atom}(E) & \to & \mathbb{N} \cup \{+\infty\} \\ \mu & \mapsto & \mu(A) \end{array}$ sont mesurables pour tout $A \in \mathcal{E}$.

⚠ En général, $\mathcal{M}_{atom}^{simple}(E)$ n'est pas une partie mesurable de $\mathcal{M}_{atom}(E)$.

En effet, soit $(X_i)_{i \geq 1}$ des v.a. iid 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 $\mu(A) = \begin{cases} 0 \text{ si } \lambda(A) = 0 \\ \infty \text{ si } \lambda(A) = \infty \end{cases}$ avec λ la mesure de Lebesgue sur $[0,1]$. (μ est s-finie car $\mu = \sum_{n=1}^{\infty} \lambda$).

M_1 et M_2 ont même loi, $M_1 \in \mathcal{M}_{atom}^{simple}(E)$ et $M_2 \notin \mathcal{M}_{atom}^{simple}(E)$, donc $\mathcal{M}_{atom}^{simple}(E)$ n'est pas mesurable

Une hypothèse de finitude et de mesurabilité sont nécessaires :

Lemme On suppose que $D = \{(x,x) : x \in E\} \in \mathcal{E} \otimes \mathcal{E}$

Alors $\mathcal{M}_{atom}^{simple,f}(E) = \{\mu \in \mathcal{M}_{atom}^{simple}(E) : \mu(E) < \infty\}$ est mesurable dans $\mathcal{M}_{atom}(E)$.

Preuve : Posons $\mathcal{M}_{atom}^{f}(E) = \{\mu \in \mathcal{M}_{atom}(E) : \mu(E) < \infty\}$, partie mesurable de $\mathcal{M}_{atom}(E)$

et $$\Phi : \mathcal{M}_{atom}^{f}(E) \longrightarrow \mathcal{M}_{atom}^{f}(E \times E)$$
$$\sum_{i=1}^{m} \delta_{x_i} \longmapsto \sum_{\substack{i,j=1 \\ i \neq j}}^{m} \delta_{(x_i, x_j)}$$

Puisque $\mathcal{M}_{atom}^{simple,f}(E) = \{\mu \in \mathcal{M}_{atom}^{f}(E) : \Phi(\mu)(D) = 0\}$, il suffit de vérifier que $\mu \longmapsto \Phi(\mu)(D)$ est mesurable. Pour cela, considérons

$$\mathcal{C} = \{C \in \mathcal{E} \otimes \mathcal{E} : \begin{cases} \mathcal{M}_{atom}^{f}(E) \to \mathbb{R} \\ \mu \mapsto \Phi(\mu)(C) \text{ est mesurable} \end{cases}\}.$$

Puisque pour μ finie et $A, B \in \mathcal{E}$, on a $\Phi(\mu)(A \times B) = \mu(A)\,\mu(B) - \mu(A \cap B)$.

On vérifie que $\mathcal{C}$ est une classe monotone, contenant les pavés mesurables.

Ceci conclut. (N.B. On peut de même montrer que Φ est mesurable)

$\sim$

Proposition 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

$\forall x \in E \quad \mu(\{x\}) = 0 \iff$ p.s. $M \in \mathcal{M}_{atom}^{simple}(E)$ (c'est-à-dire qu'il existe un événement A tq $\mathbb{P}(A) = 1$ et $\omega \in A \Rightarrow M(\omega) \in \mathcal{M}_{atom}^{simple}(E)$)

Si $\mathbb{P}$ est complète (càd que tout sous-ensemble d'un événement de proba nulle est mesurable), c'est donc $\iff \mathbb{P}(M \in \mathcal{M}_{atom}^{simple}(E)) = 1$

Le cas échéant, on dit que M est sans point multiple

⚠ On rappelle qu'en général $\mathcal{M}_{atom}^{simple}(E)$ n'est pas mesurable dans E.

Preuve : ⇐ Supposons que $\mu(\{x\}) = c > 0$ pour un $x \in E$. Alors

$\mathbb{P}(M(\{x\}) \geq 2) = 1 - e^{-c} - c e^{-c} > 0$, donc $\mathbb{P}(M \notin \mathcal{M}_{atom}^{simple}(E)) > 0$,

absurde

⇒ • Traitons d'abord le cas μ fini. Alors d'après le lemme $\mathbb{P}(M \in \mathcal{M}_{atom}^{simple, f}(E))$ ne dépend que de la loi de M. Ainsi, on peut supposer que M est construite 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 $(x_i)_{i \geq 1}$ iid de loi $\frac{\mu(\cdot)}{\mu(E)}$, $\perp$ de L.

Alors pour $1 \leq i \neq j$, $\mathbb{P}(X_i = X_j) = \frac{1}{\mu(E)^2} \iint_{E^2} 1_{x=y} \, \mu(dx) \mu(dy) = 0$.

Donc $\mathbb{P}(M \in \mathcal{M}_{atom}^{simple, f}(E)) = 1$.

• Dans le cas général σ-fini $\bigcup_{i=1}^{\infty} A_i = E$ avec $(A_i)_{i \geq 1}$ disjoints et $\mu(A_i) < \infty \ \forall i \geq 1$, on a vu que $M = \sum_{i=1}^{\infty} M_{1_{A_i}}$, où $(M_{1_{A_i}})_{i \geq 1}$ sont des nuages poissoniens indépendants d'intensités respectives $\mu_{1_{A_i}}$ (et donc finies).

Soit alors $\Omega = \bigcap_{i=1}^{\infty} \{ M_{1_{A_i}} \in \mathcal{M}_{atom}^{simple, f}(E) \}$

ainsi d'après le cas μ fini on a $\mathbb{P}(\Omega) = 1$.

De plus $\omega \in \Omega \Rightarrow M(\omega) \in \mathcal{M}_{atom}^{simple}(E)$.

$\sim$

4) Formules de Campbell (additive, 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 (formule additive)

1) (cas positif) Soit $f : E \to \mathbb{R}_+$ mesurable. Alors la v.a $M(f) = \int_E f(x) M(dx)$ a pour espérance $\mathbb{E}[M(f)] = \int_E f(x) \mu(dx)$

2) (cas intégrable) Soit $f : E \to \mathbb{R}$, μ-intégrable. Alors la v.a $M(f) = \int_E f(x) M(dx)$ est intégrable, et $\mathbb{E}[M(f)] = \int f(x) \mu(dx)$

Preuve 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) avec f^+ et f^-

$\sim$

Théorème (formule exponentielle)

1) (cas positif) Soit $f : E \to \mathbb{R}_+$ mesurable. On a

$$\mathbb{E}\left[e^{-M(f)} \right] = \exp\left(- \int_E \mu(dx) \left(1 - e^{-f(x)} \right) \right).$$

2) (cas intégrable) Soit $f : E \to \mathbb{R}$ μ-intégrable. On a

$$\mathbb{E}\left[e^{iM(f)} \right] = \exp\left(\int_E \mu(dx) \left(e^{if(x)} - 1 \right) \right)$$

Preuve : 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 $(A_i)_{i\geq 1}$ disjoints. Alors $M(f_n) = \sum_{i=1}^{N_n} c_i N(A_i)$ est une combinaison linéaire de v.a de Poisson $\perp\!\!\!\perp$, et on conclut en utilisant la forme de la transfo de Laplace d'une v.a de Poisson et convergence monotone

2) On écrit $f = \lim_{n\to\infty} f_n$ 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), on a calcul explicite donc

$$\mathbb{E}[e^{iM(f_n)}] = \exp\left(\int_E \mu(dx)\left(e^{if_n(x)} - 1\right)\right).$$

Mais d'une part, $\exp\left(\int_E \mu(dx)\left(e^{if_n(x)}-1\right)\right) \xrightarrow[n\to\infty]{} \exp\left(\int_E \mu(dx)\left(e^{if(x)}-1\right)\right)$ par convergence dominée, en utilisant $|e^{if_n(x)} - 1| \leq |f_n(x)| = f_n^+ + f_n^- \leq f^+ + f^- = |f|$.

D'autre part, $M(f_n) = M(f_n^+) - M(f_n^-)$, et $M(f_n^+) \nearrow M(f^+)$ et $M(f_n^-) \nearrow M(f^-)$ par convergence monotone. Donc $M(f_n) \xrightarrow[n\to\infty]{} M(f)$.

Comme $|e^{iM(f_n)}| = 1$, on a aussi $\mathbb{E}[e^{iM(f_n)}] \xrightarrow[n\to\infty]{} \mathbb{E}[e^{iM(f)}]$ par convergence dominée, ce qui conclut.

$\sim$

Remarque importante La formule exponentielle permet une caractérisation "à la Laplace" des mesures de Poisson : si N est une v.a à valeurs dans $\mathcal{M}_{aton}(E)$ telle que $\mathbb{E}[e^{-N(f)}] = \exp\left(-\int_E \mu(dx)(1-e^{-f(x)})\right)$ pour $f \geq 0$ ou $\mathbb{E}[e^{iN(f)}] = \exp\left(\int_E \mu(dx)(e^{if(x)}-1)\right)$ pour $f : E \to \mathbb{R}$ μ intégrable, alors N est un nuage poissonien d'intensité μ

(prendre simplement $f = \sum_{i=1}^{n} c_i \mathbb{1}_{A_i}$ avec $(A_i)_{1\leq i\leq n}$ disjoints)

Pour $f: E \to \mathbb{R}_+$ mesurable, d'après la formule additive $\mathbb{E}\left[\int f(x) M(dx)\right] < \infty \Leftrightarrow \int f d\mu < \infty$

le résultat suivant donne une CNS pour $\int f(x) M(dx) < \infty$ p.s

Proposition Soit $f: E \to \mathbb{R}_+$ mesurable. On a l'alternative :

(a) Si $\int_E \min(1, f)\, d\mu < \infty$, alors p.s. $\int_E f(x) M(dx) < \infty$

(b) Si $\int_E \min(1, f)\, d\mu = \infty$, alors p.s. $\int_E f(x) M(dx) = \infty$

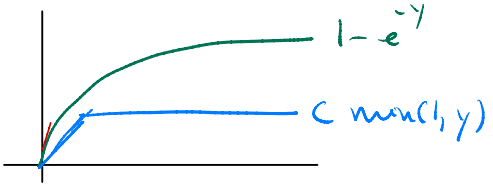
Preuve : On remarque que pour $x \in [0, \infty]$, $x < \infty \Leftrightarrow \lim_{\lambda \to 0} e^{-\lambda x} = 1$

Ainsi, $\mathbb{P}\left(\int f(x) M(dx) < \infty\right) = \lim_{\lambda \to 0} \mathbb{E}\left[e^{-\lambda M(f)}\right]$

$= \lim_{\lambda \to 0} \exp\left(-\int \mu(dx)\left(1 - e^{-\lambda f(x)}\right)\right)$ (formule expo)

(1) Si $\int_E \min(1, f)\, d\mu < \infty$, pour $\lambda \leq 1$ on a $1 - e^{-\lambda f(x)} \leq \min(1, f)$ et le résultat s'ensuit par convergence dominée

(2) Si $\int_E \min(1, f)\, d\mu = \infty$, on remarque que $\exists\, C \quad \min(1, y) \leq 1 - e^{-y}$



pour une certaine constante C $(= 1 - e^{-1})$

Ainsi, pour $\lambda < 1$, $\int \mu(dx)\left(1 - e^{-\lambda f(x)}\right) \geq C \int \mu(dx) \min(1, \lambda x) \geq \lambda C \int \mu(dx) \min(1, x) = \infty$

Donc $\mathbb{P}\left(\int f(x) M(dx) < \infty\right) = 0$

$\sim$

Exemple Si $\mu = \frac{1}{x^{3/2}} \mathbb{1}_{x>0}\, dx$ et $f(x) = x$, alors $M(f) < \infty$ p.s

mais $\mathbb{E}[M(f)] = \infty$

5) Transformations de nuages poissoniens

Proposition (stabilité par image)

Soit M un nuage poissonien d'intensité μ s-finie sur $(E, \mathcal{E})$ et $f : (E, \mathcal{E}) \to (F, \mathcal{F})$ mesurable. Alors $f_*(M)$ est un nuage poissonien d'intensité $f_*\mu$ sur F

Preuve : Tout d'abord vérifions que f_* est $\mathcal{M}_{atom}(E) \to \mathcal{M}_{atom}(F)$ et $\sum_{i\in I} \delta_{x_i} \mapsto \sum_{i\in I} \delta_{f(x_i)}$

qu'elle est mesurable.

- Pour $B \in \mathcal{F}$, on a bien $\left(\sum_{i\in I} \delta_{f(x_i)}\right)(B) = \#\{i \in I : x_i \in f^{-1}(B)\} = \left(\sum_{i\in I} \delta_{x_i}\right)(f^{-1}(B))$
- On a pour $B \in \mathcal{F}$, $k \geq 0$:

$$f_*^{-1}\left(\{\nu \in \mathcal{M}_{atom}(F) : \nu(B) = k\}\right) = \{\mu \in \mathcal{M}_{atom}(E) : f_*\mu(B) = k\}$$
$$= \{\mu \in \mathcal{M}_{atom}(E) : \mu(f^{-1}(B)) = k\},$$

mesurable

Ainsi, $f_*(M) \in \mathcal{M}_{atom}(F)$, et si $B \in \mathcal{F}$,

$$f_*(M)(B) = M(f^{-1}(B)) \sim \mathcal{P}(\mu(f^{-1}(B))) \sim \mathcal{P}(f_*\mu)$$

Par ailleurs, si $B_1, \dots, B_k \in \mathcal{F}$ sont disjoints, $f^{-1}(B_1), \dots, f^{-1}(B_k)$ sont disjoints donc les $f_*(M)(B_i) = M(f^{-1}(B_i))$ sont $\perp\!\!\!\perp$.

$\sim$

Remarque Un autre avantage de travailler avec des mesures σ-finies ici plutôt que σ-finies est que l'image d'une mesure σ-finie n'est en général pas σ-finie

Proposition (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 $\mu \in \mathcal{M}_s(E)$.

Soit $(Y_i)_{i \geq 1}$ des v.a. i.i.d dans $(F, \tilde{\mathcal{F}})$ de loi μ', indépendantes de $(L, X_1, X_2, \dots)$

Alors $\tilde{M} = \sum_{k=1}^{L} \delta_{(X_k, Y_k)}$ est un nuage poissonien dans $E \times F$ d'intensité $\mu \otimes \mu'$

Preuve : Soit $f : E \times F \to \mathbb{R}_+$ mesurable. On calcule :

$$\mathbb{E}[\exp(-\tilde{M}(f))] = \mathbb{E}\left[\mathbb{E}\left[\exp\left(-\sum_{k=1}^{L} f(X_k, Y_k) \;\middle|\; L, (X_k)_{k \geq 1}\right]\right]\right.$$

$$= \mathbb{E}\left[\prod_{k=1}^{L} \int_F \mu'(dy) \exp(-f(X_k, y))\right]$$

$$= \mathbb{E}[\exp(-M(F))]$$

avec $F(x) = -\ln \int_F \mu'(dy)\, e^{-f(x,y)} \geq 0$.

Donc d'après la formule exponentielle

$$\mathbb{E}[\exp(-\tilde{M}(f))] = \exp\left(-\int_E (1 - e^{-F(x)})\, \mu(dx)\right)$$

$$= \exp\left(-\int_{E \times F} (1 - e^{-f(x,y)})\, \mu \otimes \mu'(dx\, dy)\right).$$

D'où le résultat

$\sim$

Exemple

Soit $M = \sum_{k=1}^{L} \delta_{X_k}$ nuage poissonien d'intensité μ sur E et $(Y_i)_{i \geq 1}$ iid $\perp\!\!\!\perp$ Bernoulli(p). Alors $M_p = \sum_{k=1}^{L} Y_k \delta_{X_k}$ est un nuage poissonien d'intensité $p\mu$, car M_p est la restriction à $E \times \{1\}$ de $\tilde{M} = \sum_{k=1}^{L} \delta_{(X_k, Y_k)}$ d'intensité $\mu \otimes$ Bernoulli(p)

6) Nuages de Poisson avec paramètre temporel

On se place dans le cadre d'un espace mesuré $(E,\mathcal{E})$ avec $\{x\}\in\mathcal{E}$ $\forall x\in E$, on suppose que $\mathcal{E}$ est complète et que μ est σ-finie sur E.

Déf 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 Presque sûrement, $\forall t\geq 0$ $M(\{t\}\times E)\leq 1$

Preuve : On montre $\forall T>0$, $A_T=\{\exists t\in[0,T],\ M(\{t\}\times E)\geq 2\}$ est inclus dans un événement négligeable. Par σ-finitude, il suffit de le faire pour μ finie.

En effet, en écrivant $E=\bigcup_{i=1}^{\infty} A_i$ avec $A_i\in\mathcal{E}$ disjoints et $\mu(A_i)<\infty$, on a

$$A_T=\bigcup_{\substack{i,j=1\\ i\neq j}}^{\infty}\{\exists t\in[0,T],\ M(\{t\}\times(A_i\cup A_j))\geq 2\}$$

Comme $\mu(A_i\cup A_j)<\infty$, il suffit de traiter le cas μ finie.

Supposons donc μ finie

Pour $T>0$: $A_T\subset B_{T,n}$ avec $B_{T,n}=\{\exists\, 1\leq k\leq n,\ M(\left[\frac{k-1}{n}T,\frac{k}{n}T\right]\times E)\geq 2\}$.

Or $\mathbb{P}(B_{T,n})\leq \sum_{k=1}^{n}(1-e^{-\lambda_n}(1+\lambda_n))$ avec $\lambda_n=\frac{\mu(E)T}{n}$

$$\leq \sum_{k=1}^{n}\lambda_n^2=\frac{(\mu(E)T)^2}{n}\xrightarrow[n\to\infty]{}0$$

Ainsi $A_T\subset\bigcap_{n\geq 1}B_{T,n}$ avec $\mathbb{P}(\bigcap_{n\geq 1}B_{T,n})=0$.

Déf Pour $t\geq 0$, on pose $e(t)=\begin{cases} x & \text{si } M(\{(t,x)\})=1\\ \partial & \text{point cimetière sinon}\end{cases}$

Proposition 1) $\forall t \geq 0$, $e(t)$ est une v.a.

2) Si $\mu(E) = \infty$, $\{t : e(t) \neq \partial\}$ est dense p.s

3) Si $0 < \mu(E) < \infty$, p.s $\{t : e(t) \neq \partial\}$ est infini discret et peut être ordonné avec des v.a $0 < T_1 < T_2 < \cdots$ qui vérifient

(a) $T_1, T_2 - T_1, T_3 - T_2, \ldots$ sont ind $\exp(\mu(E))$

(b) les v.a $(e(T_i))_{i \geq 1}$ sont ind de loi $\mu(\cdot)/\mu(S)$

(c) $(e(T_i))_{i \geq 1} \perp\!\!\!\perp (T_i)_{i \geq 1}$

Preuve : 1) Pour $A \subset E$, $\{e(t) \in A\} = \{M(\{t\} \times A) = 1\}$
$\{e(t) = \partial\} = \{M(\{t\} \times E) = 0\}$

2) Si $\mu(E) = \infty$, $\forall a < b \in \mathbb{Q}_+$ on a $M([a,b] \times E) \sim \text{Poi}(\infty) = \infty$ p.s.

3) La loi de $(T_i, e(T_j))_{i,j \geq 1}$ ne dépend que de la loi de M, par exemple

$T_1 = \inf\{t \geq 0 : M([0,t] \times E) \geq 1\}$ et $\{e(T_1) \in A\} = \{M(\{T_1\} \times A) = 1\}$.

La prop 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 $(Y_i)_{i \geq 1}$ ind $\sim \mu(\cdot)/\mu(E)$. 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 au cours dernier que $(P([0,t]))_{t \geq 0}$ est un processus de Poisson d'intensité $\mu(E)$ et on peut bien ordonner $T_1 < T_2 < \cdots$ qui vérifient (a).

Les propriétés (b) et (c) découlent du fait que $e(T_i) = Y_i$.

Exemple Prenons $\mu = \frac{1}{x^{3/2}} \mathbb{1}_{x>0}\, dx$ sur $E = \mathbb{R}_+^*$. On a $\mu(E) = \infty$

Pour $t \geq 0$ on pose $X_t = \sum_{s \leq t} e(s)\, \mathbb{1}_{e(s) \neq \partial} = M(g_t)$

avec $g_t(s,x) = x\, \mathbb{1}_{s \leq t}$

Comme $\int_{\mathbb{R}_+ \times \mathbb{R}_+^*} ds \otimes \mu(dx)\ 1 \wedge g_t(s,x) < \infty$, $\forall t \geq 0$, p.s $X_t < \infty$, et par monotonie p.s $\forall t \geq 0$ $X_t < \infty$.

Néanmoins, comme $\forall t \geq 0$ $\int_{\mathbb{R}_+ \times \mathbb{R}_+^*} ds \otimes \mu(dx)\, g_t(s,x) = \infty$ on a $\mathbb{E}[X_t] = \infty$

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 subordinateur)

7) Formule de Palm

Théorème (Formule de Palm/équation de Mecke) Soit μ σ-finie et M un nuage poissonien d'intensité μ. Soit $f : E \times \mathcal{M}_{atom}(E) \to \mathbb{R}_+$ mesurable. Alors

$$\mathbb{E}\left[\int f(x, M)\, M(dx)\right] = \int_E \mathbb{E}[f(x, M + \delta_x)]\, \mu(dx) \quad (\mathcal{P})$$

On peut intuitivement interpréter la formule de Palm comme suit : sachant que M a un atome en x, ce qui advient avec intensité $\mu(dx)$, la mesure M a même loi que $M + \delta_x$, i.e la mesure de Poisson à laquelle on a rajouté un atome en x

On peut voir ce résultat comme une extension de la formule additive prenant en compte les atomes.

Preuve : Tout d'abord, comme $\mu \mapsto \mu + \delta_x$ est mesurable, par approximation par des fonctions étagées, on vérifie que $\mu \mapsto \int f(x,\mu)\,\mu(dx)$ et $\mu \mapsto f(x, \mu+\delta_x)$ sont mesurables. Ainsi (♀) ne dépend que de la loi de M.

Étape 1 : μ finie.

On écrit M sous la forme $M = \sum_{i=1}^{L} \delta_{X_i}$ avec $L \sim \mathrm{Poi}(\mu(E))$ et $(X_i)_{i\geq 1}$ iid de loi $\frac{\mu(\cdot)}{\mu(E)}$ indépendants de L. Alors

$$\mathbb{E}\left[\int f(x,M)\,M(dx)\right] = \sum_{k=1}^{+\infty} \frac{\mu(E)^k}{k!} e^{-\mu(E)} \mathbb{E}\left[\sum_{i=1}^{k} f\Big(X_i, \sum_{j=1}^{k} \delta_{X_j}\Big)\right]$$

$$= \sum_{k=1}^{+\infty} \frac{\mu(E)^k}{k!} e^{-\mu(E)} \cdot k\, \mathbb{E}\left[f\Big(X_k, \sum_{j=1}^{k} \delta_{X_j}\Big)\right]$$

$$= \sum_{k=1}^{+\infty} \frac{\mu(E)^k}{k!} e^{-\mu(E)} \cdot k \int_E \frac{\mu(dx)}{\mu(E)}\, \mathbb{E}\left[f\Big(x, \sum_{j=1}^{k-1} \delta_{X_j} + \delta_x\Big)\right]$$

$$= \int_E \mu(dx) \left(\sum_{k=1}^{\infty} \frac{\mu(E)^{k-1}}{(k-1)!} e^{-\mu(E)}\, \mathbb{E}\left[f\Big(x, \sum_{j=1}^{k-1} \delta_{X_j} + \delta_x\Big)\right]\right)$$

$$= \int_E \mu(dx)\, \mathbb{E}\left[f(x, M+\delta_x)\right]$$

Étape 2 : Écrivons $M = \sum_{i\geq 1} M_i$ avec M_i mesures poissoniennes indépendantes d'intensités μ_i finies. Pour $n \geq 1$, posons $S_n = \sum_{i=1}^{n} M_i$, $R_n = \sum_{i=n+1}^{\infty} M_i$ et $s_n = \sum_{i=1}^{n} \mu_i$.

Par convergence monotone,

$$\mathbb{E}\left[\int f(x,M)\,M(dx)\right] = \lim_{n\to\infty} \mathbb{E}\left[\int f(x, S_n + R_n)\, S_n(dx)\right]$$

$$= \lim_{n\to\infty} \mathbb{E}\left[\int F(x, S_n)\, S_n(dx)\right]$$

avec $F(x,\mu) = \mathbb{E}[f(x, \mu + R_n)]$ car $S_n \perp\!\!\!\perp R_n$

Puisque S_n est d'intensité finie,

$$\mathbb{E}\left[\int F(x, S_n) S_n(dx)\right] = \int_E \mathbb{E}\left[F(x, S_n + \delta_x)\right] S_n(dx)$$

$$= \int_E \mathbb{E}\left[g(x, M + \delta_x)\right] S_n(dx)$$

$$\xrightarrow[n\to\infty]{} \int_E \mathbb{E}\left[g(x, M + \delta_x)\right] \mu(dx)$$

$$\sim$$

Remarque 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

$$\mathbb{E}\left[\int g(x, M - \delta_x) M(dx)\right] = \int_E \mathbb{E}\left[g(x, M)\right] \mu(dx)$$