

Cours 8

Compléments probabilistes

Plan :

1. Tribu cylindrique et fonctions aléatoires
2. Tension, théorème de Prokhorov
3. Théorème de représentation de Skorokhod

1) Tribu cylindrique et fonctions aléatoires

Pour considérer des fonctions aléatoires i.e. variables aléatoires à valeurs dans des espaces fonctionnels, souvent appelés processus stochastiques, il faut définir une tribu sur ces espaces. C'est l'objet de la **tribu cylindrique**.

Pour simplifier, on se concentrera sur $\mathbb{R}^{[0,1]} = \{f : [0, 1] \rightarrow \mathbb{R}\}$.

Rappel : Si $(\Omega, \mathcal{A}, \mathbb{P})$ est un espace de probabilité et $(E, \mathcal{E}), (F, \mathcal{F})$ sont deux espaces mesurés :

- Si $f : E \rightarrow F$, alors $\sigma(f) = \{f^{-1}(B) : B \in \mathcal{F}\}$ est la plus petite tribu sur E qui rend $f : E \rightarrow (F, \mathcal{F})$ mesurable.
- Si $(f_i)_{i \in I}$ est une famille de fonctions $E \rightarrow F$, alors $\sigma(f_i, i \in I) = \sigma(\{f_i^{-1}(B_i) : i \in I, B_i \in \mathcal{F}\})$ est la plus petite tribu sur E qui rend toutes les fonctions $f_i : E \rightarrow (F, \mathcal{F})$ mesurables.

Définition (Tribu cylindrique) On considère $\mathbb{R}^{[0,1]} = \{f : [0, 1] \rightarrow \mathbb{R}\}$. Pour tout $t \in [0, 1]$, on pose :

$$\Pi_t : \mathbb{R}^{[0,1]} \rightarrow \mathbb{R}, \quad f \mapsto f(t),$$

appelée projection. On définit sur $\mathbb{R}^{[0,1]}$ la tribu

$$\mathcal{B}(\mathbb{R})^{\otimes [0,1]} = \sigma(\Pi_t : 0 \leq t \leq 1) = \sigma(\Pi_t^{-1}(B) : 0 \leq t \leq 1, B \in \mathcal{B}(\mathbb{R}))$$

qui est la plus petite tribu sur $\mathbb{R}^{[0,1]}$ qui rend toutes les projections mesurables, appelée **tribu produit** ou **tribu cylindrique**.

Définition (Cylindres) Les ensembles de la forme :

$$\Pi_{t_1}^{-1}(A_1) \cap \Pi_{t_2}^{-1}(A_2) \cap \dots \cap \Pi_{t_k}^{-1}(A_k) = \{f : [0, 1] \rightarrow \mathbb{R}, f(t_1) \in A_1, \dots, f(t_k) \in A_k\}$$

pour $0 \leq t_1 < t_2 < \dots < t_k \leq 1$ et $A_1, \dots, A_k \in \mathcal{B}(\mathbb{R})$, sont appelés **cylindres**. Ils forment un π -système générateur de la tribu cylindrique.

Exemple $\left\{f : [0, 1] \rightarrow \mathbb{R} : -1 < f\left(\frac{1}{2}\right) < 2 \text{ et } f(1) > \sqrt{2}\right\} = \Pi_{\frac{1}{2}}^{-1}(] - 1, 2[) \cap \Pi_1^{-1}(] \sqrt{2}, \infty[)$ est un cylindre.

Proposition Une fonction $X : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}^{[0,1]}, \mathcal{B}(\mathbb{R})^{\otimes [0,1]})$, $\omega \mapsto (X_t(\omega))_{0 \leq t \leq 1}$ est mesurable si et seulement si pour tout $t \in [0, 1]$, $X(t) : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, $\omega \mapsto X_t(\omega)$ est mesurable.

Preuve :

\Rightarrow Si X est mesurable, $\Pi_t \circ X$ est mesurable comme composée de fonctions mesurables.

\Leftarrow Il suffit de montrer que pour tout cylindre $C = \Pi_{t_1}^{-1}(A_1) \cap \dots \cap \Pi_{t_k}^{-1}(A_k)$, on a $X^{-1}(C) \in \mathcal{A}$, car les cylindres engendrent la tribu cylindrique. Mais

$$X^{-1}(C) = \{\omega \in \Omega : X_{t_1}(\omega) \in A_1, \dots, X_{t_k}(\omega) \in A_k\}$$

est mesurable par hypothèse.

Conséquences probabilistes

- $(X_t)_{0 \leq t \leq 1}$ est une variable aléatoire à valeurs dans $\mathbb{R}^{[0,1]}$ si et seulement si $\forall t \in [0, 1]$, X_t est une variable aléatoire réelle.
- Si $(X_t)_{0 \leq t \leq 1}$ est une variable aléatoire à valeurs dans $\mathbb{R}^{[0,1]}$, sa loi est caractérisée par les probabilités des cylindres : $\mathbb{P}(X_{t_1} \in A_1, \dots, X_{t_k} \in A_k)$, appelées **marginales fini-dimensionnelles**. En particulier, deux v.a. $(X_t)_{0 \leq t \leq 1}$ et $(Y_t)_{0 \leq t \leq 1}$ à valeurs dans $\mathbb{R}^{[0,1]}$ ont même loi ssi $\forall 0 \leq t_1 < \dots < t_k \leq 1$, on a

$$(X_{t_1}, \dots, X_{t_k}) \stackrel{\text{loi}}{=} (Y_{t_1}, \dots, Y_{t_k})$$

Proposition

Soient $(X_t)_{0 \leq t \leq 1}$ et $(Y_t)_{0 \leq t \leq 1}$ deux v.a. à valeurs dans $\mathbb{R}^{[0,1]}$ définis sur le même espace de probabilité. Alors : $(X_t)_{0 \leq t \leq 1} \perp\!\!\!\perp (Y_t)_{0 \leq t \leq 1}$ ssi $\forall 0 \leq s_1 < s_2 < \dots < s_k \leq 1$ on a $(X_{s_1}, \dots, X_{s_k}) \perp\!\!\!\perp (Y_{s_1}, \dots, Y_{s_k})$.

Preuve :

\Rightarrow Ceci provient du principe de composition.

\Leftarrow On a alors $\mathbb{P}(X \in C, Y \in C') = \mathbb{P}(X \in C) \mathbb{P}(Y \in C')$ pour tous cylindres C, C' . Les cylindres forment un π -système générateur de la tribu cylindrique, le résultat en découle.

2) Tension, théorème de Prokhorov

La notion de **tension** est en quelque sorte l'équivalent probabiliste de la **compacité**. On note ici $\mathcal{M}_1(\mathbb{R})$ l'ensemble des mesures de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ et \Rightarrow la convergence étroite.

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

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

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 une famille bornée dans L^1 (i.e. $\exists C > 0, \sup_{i \in I} \mathbb{E}[|X_i|] \leq C$), alors $(X_i)_{i \in I}$ est tendue. En effet, par l'inégalité de Markov :

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

Remarque Une **suite** $(\mu_n)_{n \geq 1}$ de $\mathcal{M}_1(\mathbb{R})$ est tendue ssi

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

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

Théorème (Prokhorov) Soit $(\mu_n)_{n \geq 1}$ une suite de $\mathcal{M}_1(\mathbb{R})$. Alors, $(\mu_n)_{n \geq 1}$ est tendu $\iff \forall$ extraction ϕ, \exists extraction ψ t.q. $(\mu_{\phi \circ \psi(n)})_{n \geq 1}$ converge étroitement.

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

Preuve :

⊞ 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, pour n fixé, l'inégalité $\limsup_{i \rightarrow +\infty} \mu_i(\mathbb{R} \setminus [-n, n]) > \varepsilon$ implique qu'il existe une infinité d'indices i tels que $\mu_i(\mathbb{R} \setminus [-n, n]) > \varepsilon$. On peut alors choisir un de ces indices strictement supérieur à $\phi(n-1)$ et le noter $\phi(n)$). Par hypothèse, soit ψ une extraction t.q. $(\mu_{\phi \circ \psi(n)})$ convergeant étroitement vers $\mu \in \mathcal{M}_1(\mathbb{R})$. Alors pour $x > 0$ t.q. $\mu(\{x\}) = \mu(\{-x\}) = 0$, pour n assez grand, $[-x, x] \subset [-\psi(n), \psi(n)]$. Donc

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

En faisant $x \rightarrow \infty$, on obtient $\mu(\mathbb{R}) \leq 1 - \varepsilon$, ce qui est absurde.

⊞ Quitte à travailler avec $\mu_{\phi(n)}$, on suppose $\phi = \text{identité}$.

Étape 1 : Identifier une limite potentielle. On pose $F_n = F_{\mu_n}$ la fonction de répartition de μ_n . Pour tout $q \in \mathbb{Q}$, la suite $(F_n(q))_{n \geq 1}$ est bornée dans $[0, 1]$. Par un argument diagonal (procédé d'extraction diagonale), il existe une extraction ψ telle que pour tout $q \in \mathbb{Q}$, $(F_{\psi(n)}(q))_{n \geq 1}$ converge vers une limite notée $F_\infty(q)$. On pose alors

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

On vérifie que F est croissante et continue à droite.

Soit $\varepsilon > 0$ et $A > 0$ t.q. $\sup_{n \geq 1} \mu_n(\mathbb{R} \setminus [-A, A]) \leq \varepsilon$.

Soit $q \in \mathbb{Q}$ avec $q > A$. Alors $F_n(q) \geq 1 - \varepsilon$. En passant à la limite, $F_\infty(q) \geq 1 - \varepsilon$, d'où $F(A) \geq 1 - \varepsilon$. On montre de même que $F(-A) \leq \varepsilon$. Donc $\lim_{-\infty} F = 0$ et $\lim_{+\infty} F = 1$. Ainsi, il existe une unique mesure $\mu \in \mathcal{M}_1(\mathbb{R})$ telle que $F = F_\mu$ (c'est la mesure de Lebesgue-Stieltjes associée à F , image de la mesure de Lebesgue sur $]0, 1[$ par $x \mapsto \inf\{y \in \mathbb{R} : F(y) \geq x\}$).

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

Soit $x \in \mathbb{R}$ t.q. $\mu(\{x\}) = 0$. Alors

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

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

$$F_\infty(q) = \lim_{n \rightarrow \infty} F_{\psi(n)}(q) \leq \liminf_{n \rightarrow \infty} F_{\psi(n)}(x) \leq \limsup_{n \rightarrow \infty} F_{\psi(n)}(x) \leq \limsup_{n \rightarrow \infty} F_{\psi(n)}(q') = F_\infty(q')$$

En faisant $q \uparrow x$ et $q' \downarrow x$, on conclut que $F_{\psi(n)}(x) \xrightarrow[n \rightarrow \infty]{} F(x)$, ce qui prouve la convergence étroite.

Lemme des sous-sous-suites pour les mesures

Soient $\mu_n, \mu \in \mathcal{M}_1(\mathbb{R})$. Alors $\mu_n \Rightarrow \mu$ ssi $\forall \phi$ extraction, $\exists \psi$ extraction t.q. $\mu_{\phi \circ \psi} \Rightarrow \mu$.

Preuve :

⊞ OK.

⊞ Par l'absurde, soit $\varepsilon > 0$, $f \in \mathcal{C}_b(\mathbb{R})$ et ϕ extraction t.q. $\forall n, |\mu_{\phi(n)}(f) - \mu(f)| \geq \varepsilon$. Par hypothèse, $\exists \psi$ extraction t.q. $\mu_{\phi \circ \psi(n)} \Rightarrow \mu$. Alors $\mu_{\phi \circ \psi(n)}(f) \rightarrow \mu(f)$, contradiction.

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

1. $(\mu_n)_{n \geq 1}$ est tendue.
2. Pour toute sous-suite convergente vers une mesure $\nu \in \mathcal{M}_1(\mathbb{R})$, on a $\nu = \mu$ (Unicité de la limite).

Preuve :

⊞ Clair.

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

. Or (μ_n) est tendue, donc $\exists \psi$ extraction et $\nu \in \mathcal{M}_1(\mathbb{R})$ t.q. $\mu_{\phi \circ \psi(n)} \Rightarrow \nu$. Donc $\nu = \mu$ par hypothèse, absurde.

Remarque Le théorème de Prokhorov reste vrai pour des variables aléatoires à valeurs dans des espaces métriques généraux (complets ayant une suite dénombrable dense), comme par exemple $\mathcal{C}([0, 1], \mathbb{R})$ muni de la norme uniforme. En pratique, pour montrer une convergence en loi dans ces espaces, on procède comme suit.

Pour montrer que $X_n \xrightarrow{(d)} X$:

1. **TENSION** : On montre que la suite (X_n) est tendue.
2. **UNICITÉ DE LA LIMITE** : On montre que si $X_{\phi(n)} \xrightarrow{(d)} Y$, alors $Y \stackrel{(d)}{=} X$ en utilisant une propriété qui caractérise les mesures de probabilité.

3) Théorème de représentation de Skorokhod

On présente ici un résultat qui permet en quelque sorte de "transformer" une convergence en loi en convergence presque sûre.

Théorème Soient $(X_n)_{n \geq 1}$ et X des variables aléatoires réelles telles que $X_n \xrightarrow{\text{loi}} X$. Alors il existe un espace de probabilité $(\Omega', \mathcal{A}', \mathbb{P}')$ et des v.a. $(Y_n)_{n \geq 1}, Y : \Omega' \rightarrow \mathbb{R}$ t.q.

1. $\forall n \geq 1$ fixé, $Y_n \stackrel{\text{loi}}{=} X_n, Y \stackrel{\text{loi}}{=} X$.
2. $Y_n \xrightarrow[n \rightarrow \infty]{\text{p.s.}} Y$.

Preuve :

Pour simplifier, nous allons supposer que les fonctions de répartition F_n de X_n et F de X sont continues et strictement croissantes (et donc des bijections de \mathbb{R} dans $]0, 1[$). Dans le cas général, on utilise l'inverse généralisé. (cf. cours 3)

Prenons $(\Omega', \mathcal{A}, \mathbb{P}') = (]0, 1[, \mathcal{B}(]0, 1[), \text{Lebesgue})$ et $U : x \mapsto x$ de sorte que U suit la loi uniforme. On définit $Y_n = F_n^{-1}(U), Y = F^{-1}(U)$.

Comme $X_n \xrightarrow{\text{loi}} X$, on a $F_n \rightarrow F$ simplement, et donc $F_n^{-1} \rightarrow F^{-1}$ simplement. On en déduit que $Y_n \rightarrow Y$ partout (donc presque sûrement). On a vu au cours 3 que Y_n a pour fonction de répartition F_n , donc $Y_n \stackrel{\text{loi}}{=} X_n$. De même $Y \stackrel{\text{loi}}{=} X$, ce qui conclut.

⚠ Toute information sur les lois jointes des X_n est perdue lors de cette construction (par exemple, si les (X_n) étaient indépendantes, les (Y_n) n'ont aucune raison de l'être).

Remarque Ce théorème est également vrai pour des variables aléatoires à valeurs dans des espaces métriques généraux (complets ayant une suite dénombrable dense), mais la preuve est plus délicate.

Application Soient X_n, X des v.a. réelles telles que $X_n \xrightarrow{\text{loi}} X$. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction t.q. f est presque sûrement continue en X . Alors :

$$f(X_n) \xrightarrow{\text{loi}} f(X)$$

Preuve :

D'après le théorème de représentation de Skorokhod, il existe des variables $(Y_n)_{n \geq 1}$ et Y définies sur un même espace de probabilité t.q. $Y_n \stackrel{\text{loi}}{=} X_n, Y \stackrel{\text{loi}}{=} X$ et $Y_n \xrightarrow{\text{p.s.}} Y$.

Comme $Y \stackrel{\text{loi}}{=} X$, f est presque sûrement continue en Y . Donc $f(Y_n) \xrightarrow{\text{p.s.}} f(Y)$.

Or, la convergence presque sûre implique la convergence en loi. Donc $f(Y_n) \xrightarrow{\text{loi}} f(Y)$. Mais $f(Y_n) \stackrel{\text{loi}}{=} f(X_n)$ et $f(Y) \stackrel{\text{loi}}{=} f(X)$, d'où le résultat.

Exemple Si $X_n \xrightarrow{\text{loi}} X$ avec $X > 0$ p.s., alors $\frac{1}{X_n} \xrightarrow{\text{loi}} \frac{1}{X}$.