Université Paris-Dauphine **2021-2022**
Cours de "Intégrale de Lebesgue et probabilités" Décembre 2020

Chapitre 9 - Constructions de mesures (hors programme)

Table des matières

1 Théorèmes de Carathéodory

On commence par une défintion et un résultat très général.

Définition 1.1. *On dit que $\mu^* : \mathscr{P}(E) \to [0,\infty]$ est une mesure extérieure si*
(i) $\mu^*(\emptyset) = 0$ *;*
(ii) μ^* *est croissante :* $\mu^*(A) \le \mu^*(B)$ *si* $A \subset B$ *;*
(iii) μ^* *est σ-sous-additive : pour toute suite* (A_n) *de* $\mathscr{P}(E)$*, on a*

$$\mu^*\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) \le \sum_{n\in\mathbb{N}} \mu^*(A_n).$$

Théorème 1.2 (Carathéodory)**.** *Etant donnée une mesure extérieure sur un ensemble E, on définit l'ensemble de parties*

$$\mathcal{A} := \big\{A \subset E;\ \mu^*(C) = \mu^*(C\cap A) + \mu^*(C\cap A^c)\ \forall C \subset E\big\}.$$

Alors
(i) $\mathcal{A}$ *est une tribu (qui contient toute les parties* $A \subset E$ *telles que* $\mu^*(A) = 0$*) ;*
(ii) $\widetilde{\mu} := \mu^*|_{\mathcal{A}} : \mathcal{A} \to [0,\infty]$ *est une mesure positive.*

Idée de la preuve. • Pour tous $A, C \subset E$, on a toujours

$$\mu^*(C) \leq \mu^*(C \cap A) + \mu^*(C \cap A^c),$$

par σ-sous-additivité, et c'est donc juste l'inégalité inverse qu'il faut démontrer. En particulier, si $\mu^*(A) = 0$, on a

$$\mu^*(C) \geq \mu^*(C \cap A^c) = \mu^*(C \cap A) + \mu^*(C \cap A^c),$$

par croissance, ce qui démontre que $A \in \mathcal{A}$.

• Pour $A, B \in \mathcal{A}$ et $C \subset E$, on a

$$\begin{aligned} \mu^*(C \cap (A \cup B)) &= \mu^*(C \cap (A \cup B) \cap A) + \mu^*(C \cap (A \cup B) \cap A^c) \\ &= \mu^*(C \cap A) + \mu^*(C \cap B \cap A^c). \end{aligned}$$

On en déduit

$$\begin{aligned} &\mu^*(C \cap (A \cup B)) + \mu^*(C \cap (A \cup B)^c) = \\ &= \mu^*(C \cap A) + \mu^*(C \cap B \cap A^c) + \mu^*(C \cap A^c \cap B^c) \\ &= \mu^*(C \cap A) + \mu^*(C \cap A^c) + \mu^*(C \cap A^c) = \mu^*(C), \end{aligned}$$

ce qui prouve que $\mathcal{A}$ est stable par union finie. On en déduit immédiatement que $\mathcal{A}$ est une algèbre.

• On se donne maintenant une suite $(B_k)_{k \geq 1}$ d'éléments de $\mathcal{A}$ deux à deux disjoints. On montre par récurrence que pour tout $C \subset E$ et tout $n \geq 1$

$$\mu^*(C) = \sum_{k=1}^{n} \mu^*(C \cap B_k) + \mu^*\Big(C \cap \Big(\bigcup_{k=1}^{n} B_k\Big)^c\Big).$$

Par croissance de μ^*, on a alors

$$\mu^*(C) \geq \sum_{k=1}^{n} \mu^*(C \cap B_k) + \mu^*\big(C \cap B^c\big), \quad B := \bigcup_{k=1}^{\infty} B_k,$$

pour tout $n \geq 1$, et donc

$$\mu^*(C) \geq \sum_{k=1}^{\infty} \mu^*(C \cap B_k) + \mu^*\big(C \cap B^c\big). \tag{1}$$

Par σ-sous-additivité de μ^*, cela prouve

$$\mu^*(C) \geq \mu^*\big(C \cap B\big) + \mu^*\big(C \cap B^c\big).$$

Par σ-sous-additivité de μ^*, l'inégalité inverse est immédiate, ce qui prouve que $B \in \mathcal{A}$. Si $(A_k)_{k \geq 1}$ est une suite quelconque de $\mathcal{A}$, on définit la suite $(B_k)_{k \geq 1}$ d'éléments de $\mathcal{A}$ deux à deux disjoints, en posant $B_1 := A_1, \ldots, B_k := A_k \backslash B_{k-1}$ et on observe que $\cup A_k = \cup B_k \in \mathcal{A}$, ce qui termine la preuve du fait que $\mathcal{A}$ est une tribu.

• Soit à nouveau $(B_k)_{k \geq 1}$ une suite d'éléments de $\mathcal{A}$ deux à deux disjoints. En prenant $C := \cup B_k$ dans (1) et en utilisant la notation $B := \cup B_k$, on trouve

$$\mu^*\Big(\bigcup_{k=1}^{\infty} B_k\Big) \geq \sum_{k=1}^{\infty} \mu^*\Big(\Big(\bigcup_{k=1}^{\infty} B_k\Big) \cap B_k\Big) + \mu^*\big(B \cap B^c\big) = \sum_{k=1}^{\infty} \mu^*\big(B_k\big).$$

Par hypothèse de σ-sous-additivité de μ^*, cette inégalité est en fait une égalité, c'est-à-dire que μ^* est σ-additive sur les éléments de $\mathcal{A}$. □

Nous donnons maintenant un théorème d'extension et une remarque qui permettent de construire une mesure dans la plupart des situations à partir d'une définition naïve de la mesure (ou prémesure).

Théorème 1.3 (d'extension de Carathéodory)**.** *Soient*

(i) $\mathcal{B}$ une semi-algèbre de E ;

(ii) $\mu^\sharp : \mathcal{B} \to [0,\infty]$ une application σ-finie et σ-additive au sens où

— *il existe une suite (B_n) de $\mathcal{B}$ telle que $\cup_n B_n = E$ et $\mu(B_n) < \infty$;*

— *pour toute suite (B_n) d'éléments disjoints de $\mathcal{B}$ telle que $\cup_n B_n \in \mathcal{B}$ on a*

$$\mu^\sharp\big(\bigcup_n B_n\big) = \sum_n \mu^\sharp(B_n).$$

Alors il existe une unique mesure σ-finie μ sur $\sigma(\mathcal{B})$ telle que $\mu_{|\mathcal{B}} = \mu^\sharp$.

Idée de la preuve. Pour tout $A \in \mathscr{P}(E)$, on définit

$$\mu^*(A) := \inf\Big\{\sum_{k=1}^{\infty} \mu^\sharp(B_k);\ B_k \in \mathcal{B},\ A \subset \bigcup_k B_k\Big\}. \tag{2}$$

• On vérifie que μ^* est une mesure extérieure. Les hypothèses impliquent que $\mu^\sharp$ est croissante et que $\mu^\sharp(\emptyset) = 0$, d'où on déduit les mêmes propriétés pour μ^*. Pour montrer la σ-sous additivité de μ^*, on considère une suite (A_n) de $\mathscr{P}(E)$ telle que $\mu^*(A_n) < \infty$ (sinon il n'y a rien à démontrer) et pour $\varepsilon > 0$, il existe une suite $(B_{k,n})$ de $\mathcal{B}$ telle que

$$A_n \subset \bigcup_k B_{k,n} \quad \text{et} \quad \sum_{k=1}^{\infty} \mu^\sharp(B_{k,n}) \le \mu^*(A_n) + \varepsilon 2^{-n}.$$

On en déduit

$$\mu^*\Big(\bigcup_n A_n\Big) \le \sum_{k,n=1}^{\infty} \mu^\sharp(B_{k,n}) \le \sum_{n=1}^{\infty} \mu^*(A_n) + \varepsilon,$$

et la σ-sous additivité de μ^*, du fait que $\varepsilon > 0$ peut être pris arbitrairement petit. Le Théorème 1.2 nous dit que l'on peut construire une tribu $\mathcal{A}$ ainsi qu'une mesure $\widetilde{\mu}$ sur $\mathcal{A}$ à partir de la mesure extérieure μ^*.

• La suite de la preuve est un peu fastidieuse ! En voici les principales étapes. On commence par démontrer que μ^* coïncide avec $\mu^\sharp$ sur $\mathcal{B}$. On montre ensuite que

$$\forall\, A, C \in \mathcal{B}, \quad \mu^\sharp(C \cap A) + \mu^*(C \cap A^c) = \mu^\sharp(C), \tag{3}$$

et on en déduit que

$$\forall\, A \in \mathcal{B},\ \forall\, C \subset E, \quad \mu^*(C \cap A) + \mu^*(C \cap A^c) = \mu^*(C),$$

ce qui montrer que $\mathcal{B} \subset \mathcal{A}$, et donc $\sigma(\mathcal{B}) \subset \mathcal{A}$. On conclut en prenant $\mu := \widetilde{\mu}_{|\sigma(\mathcal{B})}$. □

En résumé, la construction de μ se fait de la manière suivante :

$$\begin{tikzcd}
\mu^\sharp \arrow[r] \arrow[d, leftrightarrow] & \mu^* \arrow[r, "\text{Théorème 1.2}"] \arrow[d, leftrightarrow] & \widetilde{\mu} \arrow[r] \arrow[d, leftrightarrow] & \mu := \widetilde{\mu}_{|\sigma(\mathcal{B})} \arrow[d, leftrightarrow] \\
\mathcal{B} & \mathcal{P}(E) & \mathcal{A} \text{ tribu } \supset \mathcal{B} & \sigma(\mathcal{B})
\end{tikzcd}$$

Remarque 1.4. *On peut montrer le même type de résultat dans un cadre un peu plus général. On suppose que $\mathcal{B}$ est un π-système sur un ensemble E et $\mu^\sharp : \mathcal{B} \to [0,\infty]$ est une application (disons une prémesure au sens où au minimum $\mu^\sharp(\emptyset) = 0$ et $\mu^\sharp(A \cup B) \le \mu^\sharp(A) + \mu^\sharp(B)$). On définit la mesure extérieure μ^* sur $\mathscr{P}(E)$ à l'aide de* (2) *et on fait l'hypothèse que* (3) *est vérifiée. Alors, la conslusion du Théorème 1.3 est encore vraie. On peut montrer que la propriété* (3) *est équivalente à la proporiété (ii) de l'énoncé du Théorème 1.3 lorsque $\mathcal{B}$ est une semi-algèbre.*

2 Construction de la mesure de Lebesgue

Théorème 2.1 (Lebesgue)**.** *Il existe une mesure λ borélienne sur $\mathbb{R}^d$ telle que*

$$\lambda(]a_1, b_1[\times \cdots \times]a_d, b_d[) = \prod_{i=1}^{d} (b_i - a_i),$$

pour tout pavé $]a_1, b_1[\times \dots]a_d, b_d[\subset \mathbb{R}^d$. De plus, pour toute fonction $f \in C_c(\mathbb{R}^d)$, on a

$$\int_{\mathbb{R}} f \, d\lambda = \int_{\mathbb{R}} f(x) \, dx, \tag{4}$$

où le terme de droite désigne l'intégrale au sens de Riemann.

Ebauche de la preuve (Cas $d = 1$ pour simplifier, on peut alors passer au cas de la dimension supérieure $d \geq 2$ à l'aide de la théorie du Fubini-Tonelli). Pour un intervalle (a, b) de $\bar{\mathbb{R}}$, $a < b$, on pose

$$\lambda^\sharp((a, b)) = b - a.$$

Ici, les parenthèses (et) désignent indistinctement les crochets [et]. On sait (depuis le Chapitre 1) que l'ensemble $\mathcal{B}$ des unions finies d'intervalles forme une semi-algèbre. On peut appliquer le Théorème 1.3 de Carathéodory qui nous dit qu'il existe une tribu $\mathscr{L}(\mathbb{R})$ contenant $\mathcal{B}$ et une mesure $\tilde{\lambda}$ sur $\mathscr{L}(\mathbb{R})$ qui coïncide avec $\lambda^\sharp$ sur $\mathcal{B}$. En particulier $\mathscr{L}(\mathbb{R})$ contient également la tribu borélienne $\mathscr{B}(\mathbb{R})$ qui a été définie comme étant la plus petite tribu contenant $\mathcal{B}$. On note $\lambda := \tilde{\lambda}_{|\mathscr{B}(\mathbb{R})}$. Enfin, pour $f \in C_c(\mathbb{R})$, on a

$$\int_{\mathbb{R}} f \, d\lambda = \lim_{n\to\infty} \int_{\mathbb{R}} \sum_{i=-\infty}^{+\infty} f(\frac{i}{n}) \mathbf{1}_{(\frac{i}{n}, \frac{i+1}{n})} \, d\lambda = \lim_{n\to\infty} \sum_{i=-\infty}^{+\infty} \frac{1}{n} \, f(\frac{i}{n}) = \int_{\mathbb{R}} f(x) \, dx,$$

où on utilise le théorème de convergence dominée pour justifier la première limite et la définition de l'intégrale de Riemann pour justifier la seconde limite. On notera que les deux sommes sont en fait finies pour tout $n \geq 1$ fixé. □

Remarque 2.2. *(1) Il est possible ici de définir simplement la longueur (ou prémesure) des ouverts en posant pour un ouvert $\mathcal{O}$ de $\mathbb{R}$*

$$\mu^\sharp(\mathcal{O}) = \sum_{j\in J} (b_i - a_i) = \sup_{\mathcal{U}\subset\mathcal{O}} \mu^\sharp(\mathcal{U}),$$

où $\mathcal{O} = \bigcup_{j\in J}]a_j b_j[$, $J \subset \mathbb{N}$, et où le supremum est pris sur l'ensemble des ouverts $\mathcal{U}$ qui sont union **finie** *d'intervalles. Avec cette deuxième définition, on ne peut pas appliquer directement le Théorème 1.3 d'extension de Carathéodory car l'ensemble des ouverts ne forme pas une algèbre. On peut néanmoins utiliser la généralisation de la Remarque 1.4 comme nous le verrons dans la preuve du Théorème 4.1.*

(2) Le Théorème 2.1 peut être généralisé afin de construire les mesures de Stieltjes. Considérons $F : \mathbb{R} \to \mathbb{R}$ une application croissante et continue à droite. On définit

$$\mu^\sharp([a, b[) = F(b) - F(a),$$

pour tout $-\infty < a \leq b \leq \infty$, avec la convention $F(\infty) := \lim F(n)$. Puisque l'ensemble des intervalles semi-ouverts forme une semi-algèbre $\mathcal{A}$, le Théorème 1.3 de Carathéodory nous dit encore qu'il existe une (unique) mesure borélienne μ_F telle que $\mu_{F|\mathcal{A}} = \mu^\sharp$. Le cas de la mesure de Lebesgue correspond au choix $F(x) = x$.

3 Complétion d'une tribu

Soit $(E, \mathscr{B}, \mu)$ un espace mesuré.

Définition 3.1. *- On dit que $N \subset E$ est μ-négligeable (ou simplement négligeable) si $N \subset B \in \mathscr{B}$ et $\mu(B) = 0$. On note $N \in \mathcal{N}_\mu$.*

- La tribu $\mathscr{B}$ est dite complète pour la mesure μ si elle contient tous les négligeables.

Proposition 3.2. *L'ensemble $\mathscr{A} := \{B \cup N,\ B \in \mathscr{B},\ N\ \mu\text{-négligeable}\}$ est la plus petite tribu μ-complète contenant $\mathscr{B}$. La mesure μ peut être prolongée d'une unique façon en une mesure $\widetilde{\mu}$ sur $\mathscr{A}$. La mesure $\widetilde{\mu}$ est appelée la mesure complétée de μ.*

Remarque 3.3. *Il convient de remarquer que les notations sont cohérentes ! Si $(E, \mathscr{B}, \mu)$ est un espace mesuré avec E polonais et μ borelienne et σ-finie, alors la mesure $\widetilde{\mu}$ ainsi construite coïncide avec celle construite à l'aide du Théorème 1.3 :*

$$\begin{tikzcd}
\mu^\sharp := \mu \arrow[r, "\textit{Théorème 1.3}"] \arrow[d, leftrightarrow] & \widetilde{\mu} \arrow[d, leftrightarrow] & \mu \arrow[l, "\textit{Proposition 3.2}"'] \arrow[d, leftrightarrow] \\
\mathscr{B} & \mathscr{A} = \mathscr{B} + \mathcal{N}_\mu & \mathscr{B}
\end{tikzcd}$$

Exemples 3.4. *Sur $\mathbb{R}$, on définit la mesure extérieure*

$$\lambda^*(A) := \inf\{\lambda^\sharp(\mathcal{O});\ \mathcal{O} \text{ ouvert},\ A \subset \mathcal{O}\}, \quad \lambda^\sharp(]a, b[) = b - a.$$

On appelle tribu de Lebesgue, on note $\mathscr{L}(\mathbb{R})$, la tribu que l'on obtient à partir de λ^* grâce au Théorème 1.2. La tribu de Lebesgue $\mathscr{L}(\mathbb{R})$ est la tribu complétée de la tribu de Borel $\mathscr{B}(\mathbb{R})$, et $\mathscr{L}(\mathbb{R}) \neq \mathscr{B}(\mathbb{R})$. Pour voir ce dernier point, on définit l'ensemble de Cantor C de la manière suivante : on note $C_0 = [0, 1]$, $C_1 = [0, 1/3] \cup [2/3, 1]$, $C_2 := [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1]$, ..., puis $C := \lim_n C_n$. Il est clair que C est négligeable et on peut montrer qu'il n'est pas dénombrable. Donc l'ensemble des parties de C est de cardinal strictement supérieur à la puissance du continu et le cardinal de $\mathscr{L}(\mathbb{R})$ n'est donc pas plus petit. D'autre part, on peut montrer que le cardinal de $\mathscr{B}(\mathbb{R})$ est celui de $\mathbb{R}$. On a donc card$\mathscr{L}(\mathbb{R}) >$ card$\mathscr{B}(\mathbb{R})$. On peut montrer à l'aide de l'axiome du choix qu'il existe des sous-ensembles de $\mathbb{R}$ qui ne sont pas Lebesgue mesurable.

4 Théorèmes de représentation de Riesz-Markov

Théorème 4.1 (Riesz-Markov)**.** *Soit T une forme linéaire positive sur $C_c(\mathbb{R}^d)$, c'est-à-dire, une application $T : C_c(\mathbb{R}^d) \to \mathbb{R}$ telle que*

(i) T est linéaire : $T(\phi + \alpha\psi) = T(\phi) + \alpha T(\psi)$ pour tout $\alpha \in \mathbb{R}$, $\phi, \psi \in C_c(\mathbb{R}^d)$;

(ii) T est positive : $T(\phi) \geq 0$ si $\phi \in C_c(\mathbb{R}^d)$ et $\phi \geq 0$.

Alors, il existe une unique mesure borélienne σ-finie μ sur $\mathbb{R}^d$ telle que

$$T(\phi) = \int_{\mathbb{R}^d} \phi(x)\, \mu(dx).$$

Idée de la preuve. • Pour un ouvert non vide $\mathcal{O} \subset \mathbb{R}^d$, on définit

$$\mu^\sharp(\mathcal{O}) := \sup\{T(\phi);\ \phi \in C_c(\mathbb{R}^d),\ 0 \leq \phi \leq 1,\ \operatorname{supp}\phi \subset \mathcal{O}\}$$

et $\mu^\sharp(\emptyset) = 0$. On vérifie que

(a) $\mu^\sharp(\mathcal{O}_1) \leq \mu^\sharp(\mathcal{O}_2)$ si $\mathcal{O}_1 \subset \mathcal{O}_2$;

(b) $\mu^\sharp(\mathcal{O}_1 \cup \mathcal{O}_2) \leq \mu^\sharp(\mathcal{O}_1) + \mu^\sharp(\mathcal{O}_2$;

(c) $\mu^\sharp(\cup\mathcal{O}_n) \leq \sum_n \mu^\sharp(\mathcal{O}_n)$, pour toute suite d'ouverts $(\mathcal{O}_n)$.

• Pour un ensemble quelconque $A \subset \mathbb{R}^d$, on définit

$$\mu^*(A) := \inf\{\mu^\sharp(\mathcal{O});\ \mathcal{O} \text{ ouvert},\ A \subset \mathcal{O}\}.$$

On vérifie que μ^* est une mesure extérieure. On note $\mathscr{A}$ la tribu associée et définie dans le Théorème 1.2 de Carathéodory.

• On montre ensuite que la famille $\mathscr{T}$ des ouverts est incluse dans $\mathscr{A}$ et que μ^* est régulière au sens où pour tout $A \in \mathscr{A}$ et tout $\varepsilon > 0$, il existe $\mathcal{O}$ ouvert et K compact tels que

$$K \subset A \subset \mathcal{O} \ \text{ et } \ \mu^*(\mathcal{O}) - \varepsilon \le \mu^*(A) \le \mu^*(K) + \varepsilon.$$

On en déduit que la propriété (3) est vérifiée.

• On peut alors conclure que $\mathscr{A}$ contient la tribu borélienne $\mathscr{B}$ et qu'il existe une mesure μ définie sur $\mathscr{A}$, donc sur $\mathscr{B}$, par restriction de μ^*. Pour $0 \le \phi \in C_c(\mathbb{R}^d)$ fixé, on observe que

$$\frac{1}{n}\sum_{j\ge 1} \mathbf{1}_{\phi \ge j/n} \le \phi \le \frac{1}{n}\sum_{j\ge 0} \mathbf{1}_{\phi > j/n}.$$

Par régularité de μ, on a

$$\frac{1}{n}\sum_{j\ge 1} \mu(\{\phi > j/n\} \le T(\phi) \le \frac{1}{n}\sum_{j\ge 0} \mu(\{\phi \ge j/n\}).$$

On conclut que

$$T(\phi) = \lim_{n\to\infty} \frac{1}{n}\sum_{j\ge 1} \mu(\{\phi > j/n\} = \lim_{n\to\infty} \frac{1}{n}\sum_{j\ge 1} \mu(\{\phi > j/n\} = \int_{\mathbb{R}^d} \phi\, d\mu.$$

L'unicité a été démontrée au chapitre 2. □

Remarque. (1) Dans le début de la preuve, on doit pouvoir également définir une application $\mu^\sharp$ sur la semi-algèbre des pavés et appliquer directement le Théorème 1.3 ….

(2) On peut également déduire la construction de la mesure de Lebesgue à partir du Théorème 4.1 de représentation de Riesz-Markov comme alternative au Théorème 2.1. On procède de la manière suivante. On définit l'application $T : C_c(\mathbb{R}) \to \mathbb{R}$ par

$$T(\phi) := \int_{\mathbb{R}} \phi(x)\, dx,$$

où le terme de droite désigne l'intégrale de Riemann (obtenue comme limite de séries de Riemann). Cette application est évidemment une forme linéaire positive. Le Théorème 4.1 permet de conclure à l'existence d'une mesure positive (de Lebesgue) telle que (4). Pour $[a,b] \subset \mathbb{R}$ on définit $\phi_n \in C_c(\mathbb{R})$ affine par morceaux telle que $\phi_n(x) = 1$ pour tout $x \in [a,b]$ et $\phi_n \searrow \mathbf{1}_{[a,b]}$. On obtient

$$\lambda([a,b]) = \lim_n \int_{\mathbb{R}} \phi_n d\lambda = \lim_n \int_{\mathbb{R}} \phi_n(x)\, dx = b - a,$$

en utilisant le théorème de convergence dominée de Lebesgue dans la première limite et un calcul élémentaire (aires de triangles) pour la deuxième limite. □

Théorème 4.2 (variante 1)**.** *Soit T une forme linéaire contine sur $C_0(\mathbb{R}^d)$, i.e. $\exists\, C \ge 0$ telle que $\forall\, \phi \in C_0(\mathbb{R}^d)$, $|T(\phi)| \le C\, \|\phi\|_\infty$. Alors, il existe deux mesures boréliennes finies $\mu_\pm$ sur $\mathbb{R}^d$ telle que*

$$T(\phi) = \int_{\mathbb{R}^d} \phi(x)\, \mu_+(dx) - \int_{\mathbb{R}^d} \phi(x)\, \mu_-(dx).$$

Il n'y a évidemment pas unicité du couple (μ_+, μ_-) puisque $\mu_+ - \mu_- = (\mu_+ + \nu) - (\mu_- + \nu)$ pour toute mesure positive et σ-finie ν, de sorte que le couple $(\mu_+ + \nu, \mu_- + \nu)$ conviendra aussi. Toutefois, sous la condition supplémentaire que les mesures sont étrangères, i.e. $\mu_+ \perp \mu_-$ au sens du Théorème de Radon-Nikodym, alors cette décomposition est unique. On dit que $\sigma := \mu_+ - \mu_-$ est une mesure signée.

Théorème 4.3 (variante 2)**.** *Soit T une forme linéaire continue sur $L^p(\mathbb{R}^d)$, $1 \le p < \infty$, i.e. $\exists\, C \ge 0$ telle que $\forall\, \phi \in L^p(\mathbb{R}^d)$, $|T(\phi)| \le C\, \|\phi\|_p$. Alors il existe une unique fonction $f \in L^{p'}(\mathbb{R}^d)$ telle que*

$$T(\phi) = \int_{\mathbb{R}^d} \phi(x)\, f(x)\, dx.$$

5 Théorèmes de Prokhorov et de Lévy

Théorème 5.1 (Prokhorov)**.** *Soit (μ_n) une suite de mesures de probabilité qui est tendue, c'est-à-dire qui vérifie*

$$\forall\, \varepsilon > 0,\ \exists\, R > 0,\ \forall\, n \in \mathbb{N}, \quad \mu_n(B_R^c) \le \varepsilon.$$

Alors, il existe une sous-suite (μ_{n_k}) et une de mesure de probabilité μ telle que $\mu_{n_k} \rightharpoonup \mu$ faiblement.

Ebauche de la preuve. Pour tout $\varphi \in C_c(\mathbb{R}^d)$, on peut extraire une sous-suite $(\mu_{n'})$ et un réel ℓ_φ tels que

$$\int_{\mathbb{R}^d} \varphi d\mu_{n'} \to \ell_\varphi.$$

Par un argument de séparabilité de $C_c(\mathbb{R}^d)$ (existence d'un sous-ensemble dénombrable et dense dans $C_c(\mathbb{R}^d)$) et d'extraction diagonale de Cantor, on peut trouver une sous-suite (μ_{n_k}) telle que cette convergence a lieu pour tout $\varphi \in C_c(\mathbb{R}^d)$ (on inverse l'ordre des quantificateurs). Il est alors facile de voir que $T : C_c(\mathbb{R}^d) \to \mathbb{R}$, $\varphi \mapsto T(\varphi) := \ell_\varphi$ est une application linéaire et positive. D'après le Théorème 4.1 de Riesz-Markov, il existe une mesure positive telle que

$$T(\varphi) = \int_{\mathbb{R}^d} \varphi d\mu.$$

Comme pour une suite (χ_R) de $C_c(\mathbb{R}^d)$, $\mathbf{1}_{B_R} \le \chi \le \mathbf{1}_{B_{2R}}$, on a

$$\mu(B_R) \le \int_{\mathbb{R}^d} \chi_R d\mu = \lim_n \int_{\mathbb{R}^d} \chi_R d\mu_n \le 1,$$

et en fixant $\varepsilon > 0$ et choisissant R convenablement grâce au critère de tension, on a

$$\mu(B_{2R}) \ge \int_{\mathbb{R}^d} \chi_R d\mu = \lim_n \int_{\mathbb{R}^d} \chi_R d\mu_n \ge 1 - \lim_n \int_{B_R^c} d\mu_n \ge 1 - \varepsilon,$$

on en déduit $\mu(\mathbb{R}^d) = 1$. Cela prouve que μ est une mesure de probabilité. □

Théorème 5.2 (de Lévy - version forte)**.** *Soit (μ_n) une suite de mesures de probabilité telle que $\hat{\mu}_n \to \phi$ ponctuellement avec ϕ continue en 0. Alors ϕ est la transformation de Fourier d'une mesure μ et $\mu_n \rightharpoonup \mu$ faiblement.*

Ebauche de la preuve. Le point essentiel est d'établit la tension de la suite (μ_n). On note $\phi_n := \hat{\mu}_n$. Pour tout $u > 0$ (petit), on écrit

$$\begin{aligned}
\frac{1}{u}\int_{-u}^{u} (1 - \phi_n(\xi))\, d\xi &= \frac{1}{u}\int_{-u}^{u}\int_{\mathbb{R}} (1 - e^{ix\xi})\, d\mu_n(x)\, d\xi \\
&= \frac{1}{u}\int_{\mathbb{R}}\int_{-u}^{u} (1 - e^{ix\xi})\, d\xi\, d\mu_n(x) \\
&= \frac{2}{u}\int_{\mathbb{R}} \Big(u - \frac{e^{ixu} - e^{-ixu}}{2i\xi}\Big)\, d\mu_n(x) \\
&= 2\int_{\mathbb{R}} \Big(1 - \frac{\sin(xu)}{xu}\Big)\, d\mu_n(x) \\
&\ge \int_{\mathbb{R}} 2\Big(1 - \frac{\sin(xu)}{xu}\Big)\, \mathbf{1}_{xu \ge 1}\, d\mu_n(x),
\end{aligned}$$

où on a utiliser le théorème de Fubini à la deuxième ligne. En définissant

$$K := \inf_{|y|\geq 1} 2\Big(1 - \frac{\sin y}{y}\Big) > 0,$$

on obtient

$$\frac{1}{u}\int_{-u}^{u}(1-\phi_n(\xi))\,d\xi \geq K\int_{\mathbb{R}} \mathbf{1}_{xu\geq 1}\,d\mu_n(x) = K\mu_n\big([-1/u,1/u]^c\big).$$

Par hypothèse de convergence de ϕ_n vers ϕ et de continuité de ϕ en 0, on a $\phi(0)=1$ et $\phi(s)\geq 1-\varepsilon/2$ pour tout $\varepsilon>0$ et tout $|s|\leq u$, u assez petit, de sorte que

$$\mu_n\big([-1/u,1/u]^c\big) \leq K^{-1}\,\frac{1}{u}\int_{-u}^{u}(1-\phi_n(\xi))\,d\xi \leq K^{-1}\,\varepsilon,$$

pour tout n assez grand. Cela prouve la tension de (μ_n). Grâce au Théorème 5.1 de Prokhorov, on en déduit l'existence d'une sous-suite (μ_{n_k}) et d'une mesure de probabilité μ telles que $\mu_{n_k} \rightharpoonup \mu$ faiblement. On obtient alors $\phi_{n_k}\to\hat{\mu}=\phi$, et par injectivité de la transformation de Fourier, et donc unicité de la limite, on obtient que $\mu_n \rightharpoonup \mu$. □

6 Loi conditionnelle

Définition 6.1. *Soient $(E,\mathcal{E})$ et $(F,\mathcal{F})$ deux espace mesurables. On appelle probabilité (ou noyau) de transition de E dans F une application*

$$K : E\times\mathcal{F} \;\to\; [0,1]$$

telle que
- *pour tout $x\in E$, $K(x,\cdot)$ est une mesure de probabilité sur $(F,\mathcal{F})$;*
- *pour tout $A\in\mathcal{F}$, $K(\cdot,A)$ est une application mesurable sur $(E,\mathcal{E})$.*

Si f est mesurable et bornée sur $(F,\mathcal{F})$ et λ est une mesure de probabilité sur $(E,\mathcal{E})$, on définit la fonction mesurable Kf sur E et la mesure de probabilité λK sur F par

$$(Kf)(x) := \int_E f(y)K(x,dy), \quad (\lambda K)(A) := \int_E K(x,A)\lambda(dx).$$

Définition 6.2. *Soient X une variable aléatoire à valeurs dans $(E,\mathcal{E})$ et Y une variable aléatoire à valeurs dans $(F,\mathcal{F})$. On appelle loi conditionnelle de Y sachant X tout noyau de transition K telle que pour toute fonction mesurable et bornée sur $(F,\mathcal{F})$, on a*

$$\mathbf{E}(\varphi(Y)|X) = \int_F \varphi(y)K(X,dy).$$

Théorème 6.3 (de Jirina)**.** *Soient X et Y deux variables aléatoires réelles. Il existe alors une loi conditionnelle de Y sachant X.*

7 Espaces produits et tribus produits

Nous complétons l'étude commencée dans le premier chapitre des tribus produits (Chapitres 1.1 & 4.1) qui peuvent être générées par des semi-algèbres (Chapitre 1.2) convenablements définies.

Exercice 7.1 (Tribus produits)**.** *1) Soient E un ensemble et $\mathscr{A}$ une algèbre de parties de E. Pour toute famille finie d'entiers $i_1, \dots, i_n$ distincts et toute famille $A_{i_k} \in \mathscr{A}$, $1 \le k \le n$, on définit le cylindre élémentaire*

$$C = \prod_j C_j, \quad C_j = A_{i_k} \text{ si } j = i_k, \quad C_j = E \text{ si } j \notin \{i_1, \dots, i_n\}.$$

L'ensemble des cylindres élémentaires forme une semi-algèbre de $E^{\mathbb{N}}$, et donc l'ensemble des réunions finies de cylindres élémentaires forme une algèbre de parties de $E^{\mathbb{N}}$. On appelle tribu produit de $E^{\mathbb{N}}$ la tribu engendrée par les cylindres, on la note $\mathscr{A}^{\otimes \mathbb{N}}$.

2) Soit $(E_t)_{t \in T}$ une famille d'ensembles indexée par un ensemble T d'indices (quelconque), chacun étant muni d'une algèbre $\mathscr{A}_t$ de parties de E_t. Pour tout $J \subset T$ sous-ensemble fini et $A_t \in \mathscr{A}_t$, $t \in J$, on définit le cylindre élémentaire

$$C = \prod_{t \in T} C_t, \quad C_t = A_t \text{ si } t \in J, \quad C_t = E_t \text{ si } t \notin J.$$

L'ensemble des cylindres élémentaires forme une semi-algèbre de $\prod_{t \in T} E_t$, et donc l'ensemble des réunions finies de cylindres élémentaires forme une algèbre de parties de $\prod_{t \in T} E_t$. On appelle tribu produit de $\prod_{t \in T} E_t$ la tribu engendrée par les cylindres, on la note $\otimes_{t \in T} \mathscr{A}_t$. On note donc $(\prod_{t \in T} E_t, \otimes_{t \in T} \mathscr{A}_t)$ l'espace mesurable produit.

Proposition 7.2. *Si (E_n, d_n) est une suite d'espaces métriques séparable, alors l'espace produit $E = \prod_n E_n$ peut être muni de la distance $d((x_n), (y_n)) = \sup_n (d_n(x, y_n) \wedge 2^{-n})$ et $\mathscr{B}(E) = \otimes_{n \in \mathbb{N}} \mathscr{B}(E_n)$.*

Proposition 7.3. *Soient $(E, \mathscr{A})$, $(F_t, \mathscr{B}_t)_{t \in T}$ des espaces mesurables et $f : E \to \prod F_t$. On note $\pi_\tau : \prod F_t \to F_\tau$ la projection canonique, i.e. $\pi_\tau((x_t)_{t \in T}) = x_\tau$. Alors f est mesurable si, et seulement si, les $\pi_\tau \circ f$ sont mesurables pour tout $\tau \in T$. En particulier, toute fonction f à valeurs dans $\mathbb{C}$ est mesurable si, et seulement si, $\Re e f$ et $\Im m f$ sont mesurables.*

Preuve de la Proposition 7.3. Le sens direct est une conséquence immédiate de la Proposition 4.2 du Chapitre 1 et du fait que les projections canoniques $\pi_\tau : (\prod F_t, \otimes_{t \in T} \mathscr{B}_t) \to (F_\tau, \mathscr{B}_\tau)$, $x = (x_t)_{t \in T} \mapsto \pi_\tau(x) = x_\tau$ sont mesurables. Afin de démontrer l'implication réciproque, on définit l'ensemble de parties

$$\mathcal{E} := \Big\{ \prod B_t ; \exists \tau \in T,\, B_t = F_t \ \forall t \ne \tau,\, B_\tau \in \mathscr{B}_\tau \Big\},$$

qui satisfait $\sigma(\mathcal{E}) = \otimes \mathscr{B}_t$ d'après (une généralisation de) l'Exercice 1.11 du Chapitre 1. Pour tout $B \in \mathcal{E}$, en utilisant les notations introduites dans la définition de $\mathcal{E}$, on a $f^{-1}(B) = (\pi_\tau \circ f)^{-1}(B_\tau)$ est mesurable, soit donc $f^{-1}(\mathcal{E}) \subset \mathscr{A}$. On conclut grâce à la Proposition 4.3 du Chapitre 1. $\square$

8 Suite de variables aléatoires indépendantes

La dernière question que l'on se pose est celle de savoir exhiber une suite de var iid, ou même plus généralement, une suite de var indépendantes de lois prescrites. On se donne donc une suite de mesures (μ_n) sur $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$.

Commençons par un cas plus simple. Pour tout $N \in \mathbb{N}^*$, on peut définir la "mesure produit" ν_N sur $\mathbb{R}^N$, par

$$\nu_N := \prod_{i=1}^N \mu_i,$$

donc l'existence est assurée par la théorie de Fubini-Tonelli. On construit alors une famille (X_n), $1 \le n \le N$, de var indépendantes telle que $X_n \sim \mu_n$ de la manière suivante. On pose $\Omega := \mathbb{R}^N$,

$\mathscr{A} := \mathscr{B}(\mathbb{R}^N)$, $\mathbf{P} := \nu_N$, et $X_n : \Omega \to \mathbb{R}$, $X_n(\omega) = \omega_n$, la n-ième coordonnée de $\omega = (\omega_1, \dots, \omega_N)$, puisqu'alors

$$\mathbf{P}(X_n \in A) = \int_{\mathbb{R}^N} \mathbf{1}_{x_n \in A}\, d\nu_N(x) = \int_{\mathbb{R}} \mathbf{1}_{y \in A}\, d\mu_n(y) = \mu_n(A).$$

Venons-en au cas d'une suite (dénombrable). La plus grosse difficulté est de construire la mesure de probabilité sur $\mathbb{R}^{\mathbb{N}}$. On munit $\mathbb{R}^{\mathbb{N}}$ de la tribu produit définie comme la plus petite tribu engendrée par les tribus boréliennes sur chaque coordonnée. Sur l'algèbre des "cylindres", c'est-à-dire, des ensembles de la forme

$$A := A_1 \times \cdots \times A_N \times \mathbb{R} \times \cdots \times \mathbb{R} \times \dots, \quad A_i \in \mathscr{B}(\mathbb{R}),$$

on définit

$$\mu^\sharp(A) := \nu_N(A \cap \mathbb{R}^N) = \prod_{i=1}^N \mu_i(A_i).$$

On applique le Théorème 1.3 de Carathéodory afin d'obtenir une mesure μ sur $\mathbb{R}^{\mathbb{N}}$ qui coïncide avec $\mu^\sharp$ sur l'algèbre des cylindres. On conclut de manière similaire au cas précédent. □

Remarques. (1) La tribu sur $\mathbb{R}^N$ est également la tribu borélienne associée à la distance

$$d(x, y) := \sum_{n=1} 2^{-n}(|x_n - y_n| \wedge 1),$$

pour toutes suites $x = (x_n)$, $y = (y_n) \in \mathbb{R}^N$.
(2) On introduit la suite (ν_N) de mesures de probabilités sur $\mathbb{R}^{\mathbb{N}}$ définies par

$$\nu_N(A) := \prod_{i=1}^N \mu_i(A \cap \mathbb{R}^N).$$

Si on sait montrer que (ν_N) est tendue et que l'on étend le Théorème 5.1 de Prokhorov en une version valable sur $\mathbb{R}^{\mathbb{N}}$, on en déduit l'existence d'une mesure μ obtenue comme limite d'une suite extraite des (ν_N) ... .

9 Compléments sur la mesure de Lebesgue et la régularité

Revenons sur la façon de définir le volume dans $\mathbb{R}^d$. On procède en plusieurs étapes.

(1) On définit le volume $\mathrm{vol}(Q)$ des cubes (pas nécessairement ouverts ou fermés) relativement compacts Q de $\mathbb{R}^d$ par la formule habituelle.

(2) On définit la classe $\mathcal{Q}$ des ensembles

$$Q := \bigcup_{j \in J} Q_j,\ J \subset \mathbb{N},\ Q_j \text{ cube relativement compact pour tout } j \in J,$$

deux à deux disjoints, puis on pose

$$\mathrm{vol}(Q) := \sum_{j \in J} \mathrm{vol}(Q_j).$$

L'ensemble de parties $\mathcal{Q}$ est un π-système mais pas une tribu.

(3) On définit la mesure extérieure de tout sous-ensemble $A \subset \mathbb{R}^d$ à l'aide de la formule

$$\lambda^*(A) := \inf_{Q \in \mathcal{Q},\, Q \supset A} \mathrm{vol}(Q).$$

La mesure extérieure λ^* n'est pas une mesure sur $\mathscr{P}(\mathbb{R}^d)$.
(3′) Alternativement, on peut définir

$$\lambda^*(A) := \inf\{ \sum_{j\in J} \mathrm{vol}(Q_j),\ A \subset \bigcup_{j\in J} Q_j \}$$

où $J \subset \mathbb{N}$ et les Q_j sont des cubes compacts, non nécessairement disjoints.

(3″) Altenativement encore, on dit qu'un ensemble $\mathcal{U}$ est une union presque disjointe d'ensembles (U_j) si les ouverts $\mathrm{Int}(U_j)$ sont disjoints deux à deux. On observera qu'un ouvert quelconque est une union au plus dénombre et presque disjointe de cubes fermés. Pour tout ensemble $\mathcal{U}$ qui est union au plus dénombre et presque disjointe de cubes Q_j, note $\mathcal{U} \in \mathcal{Q}'$ et on pose

$$\mathrm{vol}(\mathcal{U}) := \sum_{j\in J} \mathrm{vol}(Q_j).$$

On définit la mesure extérieure de tout sous-ensemble $A \subset \mathbb{R}^d$ à l'aide de la formule

$$\lambda^*(A) := \inf_{\mathcal{U}\in\mathcal{Q}',\mathcal{U}\supset A} \mathrm{vol}(\mathcal{U}).$$

(4) On observe que pour tout ouvert $O \subset \mathbb{R}^d$, on a

$$\lambda^*(O) = \mathrm{vol}(O).$$

(5) On définit la classe $\mathscr{L}$ des ensembles mesurables $A \subset \mathbb{R}^d$ au sens de Borel-Lebesgue (on dira BL-mesurables) comme ceux que l'on peut arbitrairement approximer à l'aide des ouverts :

$$\forall\, \varepsilon > 0, \quad \exists\, O \in \mathscr{T}, \quad A \subset O, \quad \lambda^*(O \backslash A) < \varepsilon. \tag{5}$$

On pose alors

$$\lambda(A) := \inf_{O\in\mathscr{T},O\supset A} \mathrm{vol}(O), \quad \text{si } A \in \mathscr{L}.$$

Il convient alors de montrer que $\mathscr{L}$ est une tribu (qui contient par définition les ouverts) et que λ est une mesure.

Nous avions défini la tribu des ensembles mesurables d'une autre façon que nous rappelons. On dit qu'un ensemble A est mesurable au sens de Carathéodory (on dit C-mesurable) si

$$\lambda^*(B) = \lambda^*(A \cap B) + \lambda^*(A^c \cap B), \quad \forall\, B \subset E. \tag{6}$$

Théorème 9.1. *Au sens de la mesure extérieure de Lebesgue, un ensemble est BL-mesurable si, et seulement si, il est C-mesurable.*

Plus généralement :

Théorème 9.2. *Soit μ une mesure borélienne sur (E, d). Elle est régulière au sens où*

$$\forall\, A \in \mathscr{B}(E), \quad \mu(A) = \sup_{K \text{ compact} \subset A} \mu(K) = \inf_{O \text{ ouvert} \supset A} \mu(O)$$

dans les deux cas suivants :

(i) E est localement compact séparable et μ est finie sur les compacts ;

(ii) E est polonais (complet séparable) et μ est finie.