Comme $\tilde{f} + 1/n$ est une fonction positive de $\mathcal{P}$ et $\omega \in \mathcal{N}^c$, on a

$$\Lambda_\omega(\tilde{f} + 1/n) \geq 0, \text{ d'où } \Lambda_\omega(f) \geq -2/n.$$

En faisant tendre n vers l'infini, on obtient $\Lambda_\omega(f) \geq 0$.

On pose alors, pour $f \in C(\Omega)$:

$$\tilde{\Lambda}_\omega(f) = \mathbb{E}_\mathbb{P}(f)\mathbb{1}_\mathcal{N}(\omega) + \Lambda_\omega(f)\mathbb{1}_{\mathcal{N}^c}(\omega).$$

Comme $\mathcal{N} \in \mathcal{S}$, $\tilde{\Lambda}_\omega(f)$ est mesurable pour $f \in \mathcal{P}$. Mais pour $f \in C(\Omega)$, il existe une suite f_n d'éléments de $\mathcal{P}$ telle que f_n converge uniformément vers f et $\tilde{\Lambda}_\omega(f_n)$ converge partout (pour tout ω) vers $\tilde{\Lambda}_\omega(f)$: $\omega \mapsto \tilde{\Lambda}_\omega(f)$ est donc $\mathcal{S}$-mesurable comme limite ponctuelle d'applications $\mathcal{S}$-mesurables.

Comme $\mathcal{N}$ est de mesure nulle, on a encore pour tout $f \in \mathcal{P}$:

$$\forall A \in \mathcal{S} \quad \int_A f \, d\mathbb{P} = \int_A \tilde{\Lambda}_\omega(f) \, d\mathbb{P}(\omega). \tag{3.3}$$

Pour tout A, les formes linéaires $f \mapsto \int_A f \, d\mathbb{P}$ et $f \mapsto \int_A \tilde{\Lambda}_\omega(f) \, d\mathbb{P}(\omega)$ sont continues sur $C(\Omega)$. Comme elles coïncident sur $\mathcal{P}$, elles coïncident sur $C(\Omega)$.

D'après le théorème de représentation de Riesz, pour tout $\omega \in \Omega$, il existe $\mathbb{P}_\omega^\mathcal{S}$ telle que pour tout $f \in C(\Omega)$, $\tilde{\Lambda}_\omega(f) = \int_\Omega f \, d\mathbb{P}_\omega^\mathcal{S}$.

Comme $\omega \mapsto \int_\Omega f \, d\mathbb{P}_\omega^\mathcal{S}$ est $\mathcal{S}$-mesurable pour tout $f \in C(\Omega)$, la définition de $\mathcal{T}$ entraîne, avec le théorème fondamental de la mesurabilité, que $\omega \mapsto \mathbb{P}_\omega^\mathcal{S}$ est $\mathcal{S}$-mesurable [4].

Soit maintenant K un compact de Ω. On peut construire une suite bornée $(f_n)_{n\geq 1}$ de fonctions continues qui convergent simplement vers $\mathbb{1}_K$ – on peut prendre par exemple $f_n(x) = (1 + nd(K,x))^{-1}$.

Par convergence dominée, pour tout ω, $\int_\Omega f_n \, d\mathbb{P}_\omega^\mathcal{S}$ converge vers $\mathbb{P}_\omega^\mathcal{S}(K)$, ce qui entraîne en particulier que $\omega \mapsto \mathbb{P}_\omega^\mathcal{S}(K)$ est $\mathcal{S}$-mesurable. Comme

$$\forall A \in \mathcal{S} \quad \int_A f_n \, d\mathbb{P} = \int_A \int_\Omega f_n \, d\mathbb{P}_\omega^\mathcal{S} \, d\mathbb{P}(\omega),$$

4. voir l'exercice 42

En appliquant encore une fois le théorème de convergence dominée, on a

$$\forall A \in \mathcal{S} \quad \int_A \mathbb{1}_K \, d\mathbb{P} = \int_A \int_\Omega \mathbb{1}_K \, d\mathbb{P}^{\mathcal{S}}_\omega \, d\mathbb{P}(\omega).$$

Passons maintenant au cas d'un borélien quelconque : comme P est une mesure de probabilités sur $[0,1]^d \subset \mathbb{R}^d$, P est régulière, c'est à dire que pour B borélien de Ω, on peut trouver une suite croissante (K_n) de compacts et une suite (O_n) décroissante d'ouverts, avec $K_n \subset B \subset O_n$ et $\lim \mathbb{P}(O_n \backslash K_n) = 0$. Pour une preuve de cette propriété, on pourra par exemple se reporter à l'annexe B. de Garet–Kurtzmann [17].

Avec le lemme de Fatou

$$\begin{aligned}\int \lim_{n\to+\infty} (\mathbb{P}^{\mathcal{S}}_\omega(O_n) - \mathbb{P}^{\mathcal{S}}_\omega(K_n)) \, d\mathbb{P} &\le \lim_{n\to+\infty} \int (\mathbb{P}^{\mathcal{S}}_\omega(O_n) - \mathbb{P}^{\mathcal{S}}_\omega(K_n)) \, d\mathbb{P} \\ &= \lim_{n\to+\infty} \mathbb{P}(O_n \backslash K_n) = 0.\end{aligned}$$

Comme la suite $(\mathbb{P}^{\mathcal{S}}_\omega(O_n) - \mathbb{P}^{\mathcal{S}}_\omega(K_n))_{n\ge 0}$ est décroissante, il s'ensuit que pour $\mathbb{P}$-presque tout ω, $\mathbb{P}^{\mathcal{S}}_\omega(O_n) - \mathbb{P}^{\mathcal{S}}_\omega(K_n)$ tend vers 0. Ainsi, pour $\mathbb{P}$-presque tout ω, $\mathbb{P}^{\mathcal{S}}_\omega(K_n)$ tend vers $\mathbb{P}^{\mathcal{S}}_\omega(B)$.

On a montré que, à un $\mathbb{P}$-négligeable près, l'application $\omega \mapsto \mathbb{P}^{\mathcal{S}}_\omega(B)$ est $\mathcal{S}$-mesurable. L'identité $\int_A \mathbb{1}_{K_n} \, d\mathbb{P} = \int_A \mathbb{P}^{\mathcal{S}}_\omega(K_n) \, d\mathbb{P}$ entraîne à la limite $\int_A \mathbb{1}_B \, d\mathbb{P} = \int_A \mathbb{P}^{\mathcal{S}}_\omega(B) \, d\mathbb{P}$. Ainsi, $\mathbb{P}^{\mathcal{S}}_\omega(B)$ est bien une version de la probabilité conditionnelle de B sachant $\mathcal{S}$.

Le passage au fonctions étagées, puis aux fonctions mesurables positives, est classique et laissé au lecteur.

✓

3.2.2 Loi d'un vecteur sachant un autre

Vu ce qu'on vient de faire pour définir la probabilité conditionnelle sachant une tribu, on a une définition naturelle pour la loi d'un vecteur Y sachant un vecteur X : ce sera tout simplement la loi de Y sous la probabilité conditionnée par la tribu $\sigma(X)$.

Essai de définition : Si X et Y sont des variables aléatoires à valeurs dans des espaces polonais E et F, définies sur un espace de probabilité polonais $(\Omega, \mathcal{B}(\Omega), \mathbb{P})$, on appelle loi de Y sachant X l'application

$$\omega \mapsto (\mathbb{P}_\omega^{\sigma(X)})_Y,$$

où $\omega \mapsto \mathbb{P}_\omega^{\sigma(X)}$ est une version de la loi conditionnelle de $\mathbb{P}$ sachant $\sigma(X)$.[5]

On rappelle que la notation m_Y désigne la mesure image, de m par Y, en particulier, si f est mesurable bornée, avec le théorème de transfert

$$\int_F f\ d\left((\mathbb{P}_\omega^{\sigma(X)})_Y\right) = \int_\Omega f(Y)\ d\mathbb{P}_\omega^{\sigma(X)},$$

ainsi $\omega \mapsto \int f\ d\left((\mathbb{P}_\omega^{\sigma(X)})_Y\right)$ est une version de l'espérance conditionnelle de $f(Y)$ sachant X.

Malheureusement, cette définition semble un peu trop abstraite pour pouvoir être exploitée directement. On va essayer tout de même essayer d'utiliser le théorème 32, mais en s'appuyant sur l'étude d'un cas particulier plus précis.

Un cas particulier

Un cas particulier important est le cas où l'espace Ω est le produit $\Omega = E \times F$ de deux espaces polonais[6].On prend pour $\mathbb{P}$ une mesure μ quelconque sur $(E \times F, \mathcal{B}(E \times F))$[7].

Théorème 33. *Soient E et F des espaces polonais, μ une mesure de probabilité sur $E \times F$. On note X la projection sur la première coordonnée, Y la projection sur la deuxième coordonnée, et μ_X la mesure image de μ par X. Alors, il existe une famille $(\tilde{\mu}_x)_{x\in E}$ de mesures de probabilités telle que $x \mapsto \tilde{\mu}_x$ soit $(E, \mathcal{B}(E))$-$(M_1(F), \mathcal{T})$-mesurable, et telle que pour toute application f mesurable positive, on a*

$$\int_{E\times F} f\ d\mu = \int_E \left(\int f(x,y)\ d\tilde{\mu}_x(y)\right)\ d\mu_X(x).$$

5. A priori, cette application dépend de la version de la loi conditionnelle choisie. On verra plus tard, comme d'habitude, qu'il y a identité à un négligeable près.

6. Typiquement $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$

7. On préfère temporairement la lettre μ à la la lettre $\mathbb{P}$ car nous sommes sur un espace particulier.

En particulier, pour A borélien de E et B borélien de $\mathcal{F}$, on a

$$\mu(A \times B) = \int_E \mathbb{1}_A(x)\tilde{\mu}_x(B)\, d\mu_X(x) \tag{3.4}$$

La mesure $B \mapsto \tilde{\mu}_x(B)$ est couramment appelée "loi de Y sachant $X = x$."

Démonstration. Ici, La tribu $\mathcal{S} = \sigma(X)$ est formée par les ensembles de la forme $A \times F$, où A décrit l'ensemble des boréliens de E.

Prenons, comme précédemment une famille dénombrable $(O_n)_{n \geq 1}$ d'ouverts de E telle que deux mesures qui coïncident sur les O_n sont égales.

Soit $n \geq 1$. Comme $O_n \times F \in \mathcal{S}$, on a pour μ-presque tout ω :

$$\mu_\omega^{\mathcal{S}}(O_n \times F) = \mathbb{1}_{O_n \times F}(\omega) = \delta_\omega(O_n \times F),$$

autrement dit il existe Ω_n avec $\mu(\Omega_n) = 1$ et $\mu_\omega^{\mathcal{S}}(O_n \times F) = \delta_\omega(O_n \times F)$ pour $\omega \in \Omega_n$.

Maintenant, pour $\omega \in \tilde{\Omega} = \cap_{n \geq 1} \Omega_n$, les mesures $A \mapsto \mu_\omega^{\mathcal{S}}(A \times F)$ et $A \mapsto \delta_\omega(A \times F)$ coïncident. En particulier, si $\omega \in \tilde{\Omega}$ s'écrit $\omega = (x, y)$, avec x dans E et y dans F, on a

$$\mu_\omega^{\mathcal{S}}(\{x\} \times F) = \delta_\omega(\{x\} \times F) = 1. \tag{3.5}$$

Comme $(x, y) \mapsto \mu_{(x,y)}^{\mathcal{S}}$ est $\mathcal{S}$-mesurable, $\mu_{(x,y)}^{\mathcal{S}}$ ne dépend que de x [8]. On se fixe $y_0 \in F$ quelconque et on peut poser pour $x \in E$:

$$\tilde{\mu}_x(B) = \mu_{(x,y_0)}^{\mathcal{S}}(E \times B). \tag{3.6}$$

~~Notons que $\tilde{\Omega}_X$ est mesurable. Pour $x \notin \tilde{\Omega}_X$, on pose $\tilde{\mu}_x = \delta_{\omega_0}$, où ω_0 est un point quelconque.~~

8. On laisse ce point à préciser en exercice. On notera qu'on n'a pas besoin ici du lemme de Doob.