Leçon 4

Rappels de la théorie de l'intégration : théorème de Fubini-Tonelli, inégalité de Jensen

1 Théorème de Fubini-Tonelli

Le théorème (ou les théorèmes) de Fubini [1]-Tonelli [2] décrivent l'intégration par rapport aux mesures produit.

Comme il a été rappelé en Leçon 2, si μ_1 et μ_2 sont deux mesures σ-finies sur des espaces mesurables respectifs $(X_1, \mathcal{A}_1)$ et $(X_2, \mathcal{A}_2)$, la mesure produit $\mu = \mu_1 \otimes \mu_2$ est construite sur la σ-algèbre $\mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2$ produit sur l'espace produit $X = X_1 \times X_2$ par la règle

$$\mu(A) \;=\; \mu_1 \otimes \mu_2(A_1 \times A_2) \;=\; \mu_1(A_1)\, \mu_2(A_2)$$

pour tout pavé (ou rectangle) $A = A_1 \times A_2$, $A_1 \in \mathcal{A}_1$, $A_2 \in \mathcal{A}_2$.

Dans ce qui suit, $(X, \mathcal{A}, \mu)$ est l'espace mesuré produit $X = X_1 \times X_2$ muni de la σ-algèbre produit $\mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2$ et de la mesure produit $\mu = \mu_1 \otimes \mu_2$.

Théorème 1 (Théorème de Tonelli). *Soit $f : (X, \mathcal{A}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ une fonction mesurable positive sur l'espace produit X muni de la σ-algèbre produit $\mathcal{A}$; alors*

$$\int_X f \, d\mu \;=\; \int_{X_1} \bigg(\int_{X_2} f(x_1, x_2) d\mu_2(x_2) \bigg) d\mu_1(x_1)$$

ainsi que la formule analogue en renversant le rôle des indices 1 et 2.

Rappeler que s'agissant d'une fonction positive, les expressions de part et d'autre de l'égalité du théorème peuvent être éventuellement infinies.

Il convient de préciser le sens du membre de droite dans cet énoncé. Pour chaque $x_1 \in X_1$ fixé, la fonction $x_2 \in X_2 \to f(x_1, x_2)$ est mesurable par rapport à la σ-algèbre $\mathcal{A}_2$ (ce qui demande vérification), et $\int_{X_2} f(x_1, x_2) d\mu_2(x_2)$ est son intégrale par rapport à μ_2. Cette intégrale est alors elle-même une fonction mesurable de la variable $x_1 \in X_1$ par rapport à la σ-algèbre $\mathcal{A}_1$, dont il est pris

1. Guido Fubini, mathématicien italien (1879–1943).
2. Leonida Tonelli, mathématicien italien (1885–1946).

l'intégrale par rapport à μ_1. Une partie significative de l'énoncé de Tonelli est l'égalité des intégrations partielles dans un ordre et dans l'autre.

Comme dans tous les raisonnements construisant une intégrale de Lebesgue, toutes les assertions précédentes se vérifient pour commencer sur une fonction indicatrice $f = \mathbb{1}_A$ de pavé $A = A_1 \times A_2$. On parle d'intégration partielle, d'abord en la seconde variable puis en la première, ou réciproquement.

Le théorème de Fubini concerne les fonctions de signe quelconque, et requiert, ainsi qu'il est attendu, une hypothèse d'intégrabilité. Il précède historiquement le théorème de Tonelli.

Théorème 2 (Théorème de Fubini)**.** *Soit $f : (X, \mathcal{A}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ une fonction sur l'espace produit X muni de la σ-algèbre produit $\mathcal{A}$ intégrable par rapport à la mesure produit μ, autrement dit $\int_X |f| d\mu < \infty$; alors*

$$\int_X f\, d\mu = \int_{X_1} \Big(\int_{X_2} f(x_1, x_2) d\mu_2(x_2) \Big) d\mu_1(x_1)$$

ainsi que la formule analogue en renversant le rôle des indices 1 *et* 2*.*

Dans la pratique, le théorème de Fubini s'appuie sur celui de Tonelli en vérifiant au préalable la condition d'intégrabilité $\int_X |f| d\mu < \infty$ par intégration partielle, suivant l'une ou l'autre des coordonnées, puisque que $|f|$ est mesurable positive.

Il est illustratif d'appliquer les théorèmes de Fubini et Tonelli à la situation « idéale » où la fonction à intégrer est un produit $f(x) = f_1(x_1) f_2(x_2)$, $x = (x_1, x_2) \in X$, donnant lieu à une séparation des variables. Dans ce cas, sous les conditions de mesurabilité et d'intégrabilité requises,

$$\int_X f\, d\mu = \Big(\int_{X_1} f_1 d\mu_1 \Big) \Big(\int_{X_2} f_2 d\mu_2 \Big).$$

Si μ_1 et μ_2 sont la mesure de comptage sur $\{1, \ldots, n\}$, l'identité précédente recouvre des propriétés classiques du type

$$\sum_{k,\ell=1}^{n} a_k b_\ell = \sum_{k=1}^{n} a_k \sum_{\ell=1}^{n} b_\ell$$

pour des réels (ou complexes) a_k, b_ℓ, $k, \ell = 1, \ldots, n$.

Un autre exemple instructif d'application des théorèmes de Fubini et Tonelli permet de retrouver le théorème de convergence monotone pour les séries (cf. Leçon 3). En effet, si f_n, $n \in \mathbb{N}$, est une famille de fonctions mesurables positives sur un espace mesuré $(X_1, \mathcal{A}_1, \mu_1)$ (avec toutefois μ_1 σ-finie), alors

$$\int_{X_1} \sum_{n\in\mathbb{N}} f_n \, d\mu_1 = \sum_{n\in\mathbb{N}} \int_{X_1} f_n \, d\mu_1.$$

Pour cela, considérer $X_2 = \mathbb{N}$ muni de la σ-algèbre des parties et de la mesure $\mu_2 = \sum_{n\in\mathbb{N}} \delta_n$ de comptage. La conclusion découle alors de l'application du théorème de Tonelli à la fonction

$$f(x_1, n) = f_n(x_1), \quad x_1 \in X_1, \ n \in \mathbb{N} = X_2.$$

Le cas des fonctions de signe quelconque s'obtient du théorème de Fubini.

Les théorèmes de Fubini et Tonelli s'étendent sans difficulté à des mesures produits de plus de 2 facteurs.

Pour plus de simplicité, les deux théorèmes seront souvent confondus ou utilisés simultanément sous le nom de théorème de Fubini-Tonelli.

2 Décomposition sphérique de la mesure de Lebesgue

La théorie de la mesure et de l'intégration de Lebesgue s'applique à des mesures construites et définies sur d'autres espaces que les espaces euclidiens traditionnels. L'exemple de la mesure de Lebesgue, ou mesure uniforme, sur la sphère est instructif. Celle-ci peut être construite par une décomposition de la mesure de Lebesgue λ^d dans $\mathbb{R}^d$; en dimension 2, ce sont les fameuses coordonnées polaires.

Dans ce qui suit, $\mathbb{S}^{d-1} = \{x \in \mathbb{R}^d ; |x| = 1\}$ désigne la sphère euclidienne unité dans $\mathbb{R}^d$, munie de la tribu borélienne (trace) $\mathcal{B}(\mathbb{S}^{d-1}) = \mathcal{B}(\mathbb{R}^d) \cap \mathbb{S}^{d-1}$.

Théorème 3. *Sur $(\mathbb{S}^{d-1}, \mathcal{B}(\mathbb{S}^{d-1}))$, il existe une unique mesure σ^{d-1}, invariante par rotation, telle que pour toute fonction $f : \mathbb{R}^d \to \mathbb{R}$ borélienne, positive ou intégrable par rapport à λ^d,*

$$\int_{\mathbb{R}^d} f \, d\lambda^d = \int_{]0,\infty[} \int_{\mathbb{S}^{d-1}} r^{d-1} f(rs) \, d\lambda^1(r) d\sigma^{d-1}(s).$$

La mesure σ^{d-1} est construite en posant

$$\sigma^{d-1}(B) = d \, \lambda^d \big(\Psi^{-1}(B) \cap U \big)$$

pour tout $B \in \mathcal{B}(\mathbb{S}^{d-1})$, où $U = \{x \in \mathbb{R}^d ; |x| \leq 1\}$ est la boule euclidienne unité de $\mathbb{R}^d$ et $\Psi(x) = \frac{x}{|x|} \in \mathbb{S}^{d-1}$ si $x \in \mathbb{R}^d \setminus \{0\}$ (correspondant géométriquement à la mesure de Lebesgue du cône issu de l'origine interceptant B sur la sphère) ; σ^{d-1} est ainsi la mesure image de λ^d sur $U \setminus \{0\}$ par l'application Ψ. (À ce titre, elle est invariante par rotation.) Pour se convaincre alors de l'identité du théorème, considérer la partie

$$A = \big\{ x \in \mathbb{R}^d ; \, |x| \in [a, b] \, , \Psi(x) \in B \big\}$$

où $0 < a < b$ et $B \in \mathcal{B}(\mathbb{S}^{d-1})$, pour laquelle il est aisé de vérifier que

$$\int_{\mathbb{R}^d} \mathbb{1}_A \, d\lambda^d = \lambda^d(A) = (b^d - a^d) \, \lambda^d \big(\Psi^{-1}(B) \cap U \big).$$

Par ailleurs, par définition de A et le théorème de Fubini-Tonelli,

$$\begin{aligned}\int_{]0,\infty[}\int_{\mathbb{S}^{d-1}} & r^{d-1}\mathbb{1}_A(rs)\,d\lambda^1(r)d\sigma^{d-1}(s)\\ &= \int_{[a,b]} r^{d-1}d\lambda^1(r)\int_{\mathbb{S}^{d-1}}\mathbb{1}_B(s)d\sigma^{d-1}(s)\\ &= \frac{1}{d}\,(b^d-a^d)\,\sigma^{d-1}(B)\\ &= (b^d-a^d)\,\lambda^d\big(\Psi^{-1}(B)\cap U\big).\end{aligned}$$

La formule du théorème est donc satisfaite pour les fonctions indicatrices, correspondant à la brique élémentaire de la théorie de l'intégration. Elle s'étend aux fonctions étagées positives, puis, par convergence monotone, à toutes les fonctions positives, et intégrables.

Lorsque $d = 2$, le théorème correspond exactement au changement de variable en coordonnées polaires, avec pour σ^1 la mesure de Lebesgue sur $[0, 2\pi[$, r étant le jacobien. En plus grande dimension, les points de $\mathbb{S}^{d-1}$ peuvent être paramêtrés par des angles. Le choix classique des coordonnées sphériques dans $\mathbb{R}^3$ est illustré dans l'Exercice 8, Leçon 5.

3 Inégalité de Jensen

L'inégalité de Jensen [3] est une inégalité importante de convexité qui ne s'applique qu'aux mesures de masse totale égale à 1 (de probabilité), ou finie mais après une normalisation appropriée.

Une *fonction convexe* ϕ sur un intervalle ouvert (pour simplifier) I de $\mathbb{R}$ vérifie par définition

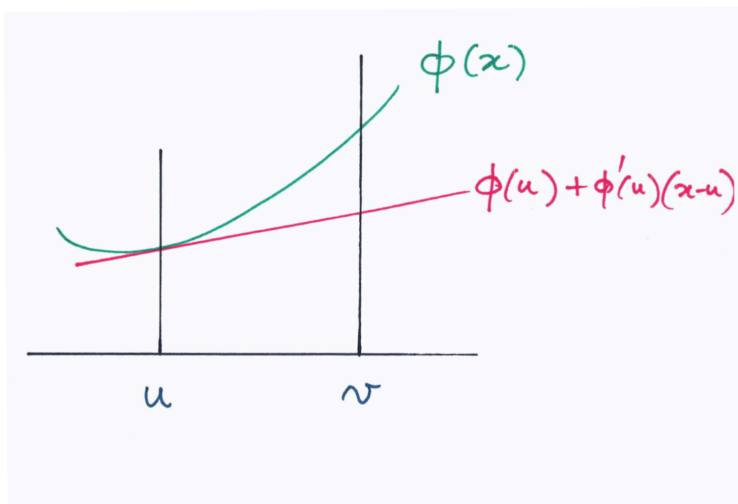
$$\phi\big(\theta u + (1-\theta)v\big) \;\leq\; \theta\phi(u) + (1-\theta)\phi(v)$$

3. Johan Jensen, mathématicien et ingénieur danois (1859–1925).

pour tous $u, v \in I$ et tout $\theta \in [0, 1]$. Les propriétés des fonctions convexes sont classiques ; pour les besoins de cette leçon, il suffira de rappeler qu'une fonction convexe $\phi : I \to \mathbb{R}$ est continue et admet, en tout point de I, une dérivée à gauche et à droite, et que la propriété de convexité exprime que la fonction est au-dessus de ses tangentes au sens où

$$\phi(u) - \phi(v) \le \phi'(u)(u - v)$$

pour $u, v \in I$, $\phi'(u)$ désignant la dérivée, à gauche ou à droite, au point u. Un dessin pourra confirmer intuitivement cette propriété :



Si ϕ est de classe $\mathcal{C}^2$ sur I, la formule de Taylor permet de montrer qu'elle est convexe si et seulement si $\phi'' \ge 0$ sur I. Une fonction ϕ *concave* si $-\phi$ est convexe. Les fonctions $\phi(u) = u^2$, $\phi(u) = e^{\alpha u}$, $\alpha \in \mathbb{R}$, sont convexes sur $I = \mathbb{R}$, la fonction $\phi(u) = \ln(u)$ est concave sur $I =]0, \infty[$.

Théorème 4 (Inégalité de Jensen). *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré de masse totale égale à 1, $\mu(X) = 1$; soit $\phi : I \to \mathbb{R}$ une fonction convexe, et soit f une fonction mesurable sur $(X, \mathcal{A})$ à valeurs dans I, intégrable et telle que $\phi(f)$ est intégrable également ; alors*

$$\phi\Big(\int_X f \, d\mu\Big) \le \int_X \phi(f) \, d\mu.$$

Démonstration. Il suffit d'utiliser la propriété des tangentes pour $u = \int_X f d\mu$ et $v = f(x)$ pour tout $x \in X$; ainsi

$$\phi(u) - \phi\big(f(x)\big) \;\le\; \phi'(u)\big(u - f(x)\big).$$

L'intégrale conservant le sens des inégalités, l'intégration de cette dernière par rapport à μ fournit

$$\phi(u) - \int_X \phi(f) d\mu \;\le\; \phi'(u)\left(u - \int_X f\, d\mu\right)$$

où il a été utilisé que $\int_X c d\mu = c$ pour toute fonction constante (égale à c) car $\mu(X) = 1$. Comme $u = \int_X f d\mu$, le membre de droite est nul, ce qui conduit à l'inégalité annoncée. $\square$

Une récurrence sur la définition de la convexité d'une fonction ϕ sur un intervalle I de $\mathbb{R}$ indique que si $u_1, \ldots, u_n \in I$, et $p_1, \ldots, p_n \in [0, 1]$ tels que $\sum_{k=1}^n p_k = 1$, alors

$$\phi\left(\sum_{k=1}^n p_k u_k\right) \;\le\; \sum_{k=1}^n p_k \phi(u_k).$$

C'est une autre façon de lire l'inégalité de Jensen par rapport à la mesure (de probabilité) $\sum_{k=1}^n p_k \delta_{u_k}$, qui peut aussi conduire à une démonstration de celle-ci. En effet, si $f = \sum_{k=1}^n c_k \mathbb{1}_{A_k}$ est une fonction étagée (de signe quelconque),

$$\int_X f d\mu \;=\; \sum_{k=1}^n c_k \mu(A_k)$$

et

$$\int_X \phi(f) d\mu \;=\; \sum_{k=1}^n \phi(c_k)\mu(A_k)$$

où $\sum_{k=1}^n \mu(A_k) = 1$.

Beaucoup d'autres démonstrations de l'inégalité de Jensen sont disponibles, l'une d'elle utilise la loi des grands nombres (Exercice 1, Leçon 19).

La fonction $\phi(u) = |u|$ est certes convexe sur $\mathbb{R}$, mais l'inégalité

$$\left| \int_X f \, d\mu \right| \leq \int_X |f| \, d\mu$$

est une conséquence des propriétés de comparaison de l'intégrale, ou inégalité triangulaire, et est satisfaite même si μ n'est pas de probabilité.

La mesure de Lebesgue λ sur $\mathbb{R}$ n'est certes pas de probabilité, sa restriction à un intervalle borné $[a, b]$ est toutefois finie de sorte que l'inégalité de Jensen s'applique une fois cette dernière normalisée, donc

$$\phi\left(\frac{1}{b-a} \int_{[a,b]} f \, d\lambda \right) \leq \frac{1}{b-a} \int_{[a,b]} \phi(f) \, d\lambda.$$

Par exemple

$$\left(\int_{[a,b]} f \, d\lambda \right)^2 \leq (b-a) \int_{[a,b]} f^2 \, d\lambda$$

(sous les conditions d'intégrabilité requises).

Exercices

(*Une étoile * désignera une question de difficulté supérieure.*)

Exercice 1. Le théorème de Fubini s'applique-t-il à la fonction

$$f(x,y) = \frac{x^2 - y^2}{(x^2+y^2)^2}\,, \quad (x,y) \in [0,1]^2$$

(par rapport à la mesure de Lebesgue sur $[0,1]^2$) ?

Exercice 2. Les fonctions $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ suivantes sont-elles intégrables sur $[0,1]^2$ par rapport à la mesure de Lebesgue ?

$$f(x,y) \;=\; \frac{x-y}{(x^2+y^2)^{\frac{3}{2}}}\,; \qquad f(x,y) \;=\; \frac{1}{(1-xy)^p}\,, \quad p > 0.$$

Exercice 3. Calculer $\int_E (x^3+y^3)\, d\lambda(x,y)$ où $E = \{x, y \geq 0\,; \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1\}$ $(a, b > 0)$.

Exercice 4 (*Intégrale gaussienne*). En utilisant le théorème de Fubini-Tonelli, puis un changement de variables en coordonnées polaires, démontrer que

$$\left(\int_{\mathbb{R}} e^{-\frac{1}{2}x^2} d\lambda\right)^2 \;=\; \int_{]0,2\pi[} \int_{]0,\infty[} \rho\, e^{-\frac{1}{2}\rho^2} d\lambda(\theta) d\lambda(\rho).$$

En déduire que

$$\int_{\mathbb{R}} e^{-\frac{1}{2}x^2} d\lambda \;=\; \sqrt{2\pi}\,.$$

Exercice 5. Soit μ une mesure de probabilité sur un espace mesurable $(X, \mathcal{A})$, et soit $f : X \to \mathbb{R}$ de carré intégrable pour μ. Démontrer que

$$\int_X f^2 d\mu - \left(\int_X f \, d\mu \right)^2 = \frac{1}{2} \int_X \int_X \big[f(x) - f(y) \big]^2 d\mu(x) d\mu(y).$$

(Le membre de gauche est une variance au sens de la Leçon 9.)

Exercice 6* (*Intégrale de Dirichlet*[4]). On se propose de déterminer la valeur (si elle existe) de l'intégrale

$$I = \int_0^\infty \frac{\sin(x)}{x} \, dx.$$

a) Donner un sens à cette intégrale. La fonction $x \in]0, \infty[\mapsto \frac{\sin(x)}{x}$ est-elle intégrable au sens de Lebesgue ?

b) Soit $a > 0$; démontrer que la fonction $f : (x, y) \mapsto e^{-xy} \sin(x)$ est intégrable sur $]0, a[\times]0, \infty[$ par rapport à la mesure de Lebesgue λ sur $\mathbb{R}^2$.

c) Poser $I_a = \int_{]0,a[\times]0,\infty[} f d\lambda$. Exprimer I_a de deux façons différentes, et calculer la limite $\lim_{a \to \infty} I_a$. En déduire la valeur de l'intégrale I.

Exercice 7. Cet exercice a pour but de calculer le volume (c'est-à-dire la mesure de Lebesgue) ω_d de la boule euclidienne unité (fermée) B de $\mathbb{R}^d$ à partir de l'intégrale gaussienne de l'Exercice 4.

a) Démontrer que

$$\int_{\mathbb{R}^d} e^{-\frac{1}{2}|x|^2} d\lambda^d = (2\pi)^{\frac{d}{2}}.$$

b) Utiliser la décomposition de la mesure de Lebesgue λ^d dans $\mathbb{R}^d$ en partie radiale et sphérique (Théorème 3) pour obtenir que

$$\int_{\mathbb{R}^d} e^{-\frac{1}{2}|x|^2} d\lambda^d = \int_{]0,\infty[} \int_{\mathbb{S}^{d-1}} r^{d-1} e^{-\frac{1}{2}r^2} \, d\lambda^1(r) d\sigma^{d-1}(s).$$

4. Johann Dirichlet, mathématicien prussien (1805–1859).

c) En déduire que

$$(2\pi)^{\frac{d}{2}} \;=\; 2^{\frac{d}{2}-1}\,\Gamma\Big(\frac{d}{2}\Big)\,\sigma^{d-1}(\mathbb{S}^{d-1})$$

où $\Gamma(p) = \int_{]0,\infty[} x^{p-1}e^{-x}d\lambda$ est la fonction Gamma. (Rappeler que $\Gamma(p+1) = p\,\Gamma(p)$.)

d) De la définition de σ^{d-1}, déduire que $\omega_d = \frac{1}{d}\,\sigma^{d-1}(\mathbb{S}^{d-1})$, et conclure que

$$\omega_d \;=\; \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}\,.$$

Exercice 8 (*Théorème du retour de Poincaré*). Soient $(X,\mathcal{A},\mu)$ un espace mesuré avec μ finie, et $T : X \to X$ une application mesurable de X, sous laquelle μ est invariante, i.e. $\mu \circ T^{-1} = \mu$. Les itérées de T sont désignées par T^k, $k \geq 0$ (avec la convention $T^0 = \mathrm{Id}$). Soit également $A \in \mathcal{A}$ tel que $\mu(A) > 0$.

a) Pour tout $x \in X$, poser $N(x) = \sum_{k\geq 0} \mathbb{1}_A(T^k(x))$, et $G(x) = e^{-N(x)}$. Vérifier que G est une application mesurable (dans les boréliens de $\mathbb{R}$) et intégrable par rapport à μ.

b) Poser $H = G \circ T$. Démontrer que $G = e^{-\mathbb{1}_A}H$ et que $\int_X H d\mu = \int_X G d\mu$. En déduire que $G = H$ μ-presque partout.

c*) Conclure que pour μ-presque tout $x \in A$, la suite $T^k(x)$, $k \geq 0$, passe une infinité de fois dans A.