15 - UNE FORMULE INTÉGRALE POUR LES MESURES PRODUITS (MISE À JOUR)

www.daniel-saada.eu

Rappelons que si m_1 est une mesure de probabilité sur un univers Ω_1 muni d'une tribu $\mathcal{T}_1$, si m_2 est une mesure de probabilité sur $(\Omega_2, \mathcal{T}_2)$, il existe sur $(\Omega = \Omega_1 \times \Omega_2, \mathcal{T} = \mathcal{T}_1 \otimes \mathcal{T}_2)$ une mesure de probabilité m et une seule, notée $m_1 \otimes m_2$, telle que $m(T_1 \times T_2) = m_1(T_1)m_2(T_2)$ pour tous T_1 dans $\mathcal{T}_1$ et T_2 dans $\mathcal{T}_2$. Comme m est σ-additive, si A dans la tribu produit $\mathcal{T}$ est une réunion dénombrable *disjointe* de pavés $T_n \times T'_n$, $A = \sum_n T_n \times T'_n$, alors $m(A) = \sum_n m(T_n).m(T'_n)$. On pourra consulter mon livre, paragraphe **13.2**, page **212**.

Qu'en est-il pour les événements de la tribu qui ne sont pas des réunions dénombrables de pavés, comme la diagonale de Ω quand Ω_1 et Ω_2 sont égaux et non dénombrables, et à condition que la diagonale soit dans la tribu produit ? La réponse est donnée dans l'encadré ci-après, mais rappelons d'abord (p. **210**) que si A est dans la tribu $\mathcal{T}_1 \otimes \mathcal{T}_2$ toutes les sections A_x de A sont dans la tribu $\mathcal{T}_2$ et toutes les sections A_y sont dans la tribu $\mathcal{T}_1$, où pour x dans Ω_1, $A_x = \{y \in \Omega_2 : (x,y) \in A\}$, pour y dans Ω_2, $A_y = \{x \in \Omega_1 : (x,y) \in A\}$:

> Pour tout A de la tribu produit, $x \to m_2(A_x)$ est m_1 - intégrable sur Ω_1,
> $y \to m_1(A_y)$ est m_2 - intégrable sur Ω_2 et $m(A) = \int_{\Omega_1} m_2(A_x) = \int_{\Omega_2} m_1(A_y)$.

Cette formule a un double intérêt : elle montre la nécessité du calcul intégral, elle explique pourquoi l'aire sous la courbe d'équation $y = f(x)$ s'obtient en intégrant $f(x)$ qui est la longueur de la section en x de la surface sous la courbe (*la surface est l'intégrale d'une longueur et le volume l'intégrale d'une surface*).

Pour établir la formule intégrale qui donne $m(A)$, on introduit la famille des événements de $\mathcal{T}$ pour lesquels elle est vraie :

$$\mathcal{F} = \left\{ A \in \mathcal{T} : x \to m_2(A_x) \text{ est m}_2\text{- intégrable et } m(A) = \int_{\Omega_1} m_2(A_x) \right\} ;$$

comme $x \to m_2(A_x)$ est positive et bornée, il revient au même de dire que

$$\mathcal{F} = \left\{ A \in \mathcal{T} : x \to m_2(A_x) \text{ est } \mathcal{T}_2\text{- mesurable et } m(A) = \int_{\Omega_1} m_2(A_x) \right\}$$

mais nous n'aurons pas besoin de ce raffinement.

Nous allons prouver que $\mathcal{F}$ est une famille monotone (p. **103**) *qui contient tous les pavés de $\mathcal{T}$; on en déduira que $\mathcal{F}$ est $\mathcal{T}$ tout entier en vertu du Théorème de la classe monotone* (p. **104**).

1) $\mathcal{F}$ est une famille monotone de $\Omega_1 \times \Omega_2$

a) <u>$\mathcal{F}$ est stable par réunion dénombrable croissante</u>

Si $A = \bigcup_n A_n$, avec $A_n \subset A_{n+1}$, alors d'une part $A \in \mathcal{T}$, d'autre part $m(A) = \lim_n m(A_n) = \lim_n \int_{\Omega_1} m_2(A_n)_x$.

Justifions ensuite $\lim_n \int_{\Omega_1} m_2(A_n)_x = \int_{\Omega_1} \lim_n m_2(A_n)_x = \int_{\Omega_1} m_2(A)_x$. D'abord, $A_x = \bigcup_n (A_n)_x$ et $(A_n)_x \subset (A_{n+1})_x$, donc $A_x = \lim_n (A_n)_x$ et $m(A_x) = \lim_n m(A_n)_x$; ensuite on applique, au choix, le Théorème de convergence monotone de **Beppo-Levi** ou le Théorème de convergence dominée de **Lebesgue**, à la suite des fonctions $x \to m(A_n)_x$:

cette suite de fonctions est croissante, de limite $x \to m(A)_x$, positive et majorée par $m_1(\Omega_1).m_2(\Omega_2)$.
Chacun des deux théorèmes assure que $x \to m_2(A_x)$ est m_1-intégrable .

b) $\mathcal{F}$ <u>est stable par différence croissante</u> : A et B dans $\mathcal{F}$ et $A \subset B$ impliquent $B - A \in \mathcal{F}$

En effet, $B - A \in \mathcal{T}$ et comme $m(B-A) = m(B) - m(A)$:

$m(B-A) = \int_{\Omega_1} m_2(B_x) - \int_{\Omega_1} m_2(A_x) = \int_{\Omega_1} \left(m_2(B_x - A_x)\right) = \int_{\Omega_1} m_2(B-A)_x$, donc $B - A \in \mathcal{F}$.

c) $\mathcal{F}$ <u>contient</u> $\Omega_1 \times \Omega_2$
Pour tout x de Ω_1, $(\Omega_1 \times \Omega_2)_x = \Omega_2$, et donc $\int_{\Omega_1} m(\Omega_1 \times \Omega_2)_x = \int_{\Omega_1} m(\Omega_2) = m(\Omega_1).m(\Omega_2) = m(\Omega_1 \times \Omega_2)$.

2) $\mathcal{F}$ **contient les pavés de** $\mathcal{T}_1 \otimes \mathcal{T}_2$

Lorsque $A = A_1 \times A_2$ est dans $\mathcal{T}$, on sait que $A_1 \in \mathcal{T}_1$ et $A_2 \in \mathcal{T}_2$ (p. **210**), et alors $A_x = A_2$ si $x \in A_1$ et $A_x = \varnothing$ sinon ; évidemment, $x \to m_2(A_x)$ est m_1-intégrable donc $\int_{\Omega_1} m_2(A_x) = \int_{A_1} m_2(A_2) = m_2(A_2).m_1(A_1) = m(A)$ et A est bien dans $\mathcal{F}$.

3) $\mathcal{F}$ **coïncide avec la tribu** $\mathcal{T}$
Par définition, $\mathcal{T}_1 \otimes \mathcal{T}_2$ est engendrée par les pavés et l'ensemble des pavés, $\mathcal{P}(\mathcal{T}_1) \times \mathcal{P}(\mathcal{T}_2)$ et est stable par intersection dénombrable puisque $\bigcap_n A_n \times B_n = \bigcap_n A_n \times \bigcap_n B_n$.
Le Théorème de la classe monotone affirme que la famille monotone engendrée par une famille stable par intersection *finie* est la tribu engendrée par cette famille : il en résulte que $\mathcal{T} \subset \mathcal{F}$. Comme $\mathcal{F}$ est contenue dans $\mathcal{T}$, $\mathcal{F} = \mathcal{T}$.

Deuxième démonstration. On prouve d'abord que, pour A dans $\mathcal{T}$, $x \to m_2(A_x)$ est m_2-intégrable , puis que la fonction $A \to \mu(A) = \int_{\Omega_1} m_2(A_x)$ est une mesure de probabilité sur la tribu produit, enfin que pour tout (T_1, T_2) on a $\mu(T_1 \times T_2) = m_1(T_1).m_2(T_2)$; grâce à l'unicité de la mesure produit, $\mu = m_1 \otimes m_2 = m$.

<u>APPLICATIONS</u>

1) *Une diagonale est négligeable pour la mesure produit si l'une des deux mesures est diffuse*
Soit m_1 et m_2 des mesures de probabilité sur $(\Omega, \mathcal{T})$ dont l'une au moins est diffuse ($\mathcal{T}$ contient alors tous les singletons, ou événements élémentaires, de Ω). La mesure $m = m_1 \otimes m_2$ est diffuse sur $(\Omega^2, \mathcal{T} \times \mathcal{T})$ puisque par construction $(m_1 \otimes m_2)(x, y) = m_1\{x\}.m_2\{y\}$. Quand la diagonale Δ de Ω^2 est dans la tribu produit, $m(\Delta) = 0$ parce que, si m_1 est diffuse par exemple, $m_1(\Delta_x) = m_1\{x\} = 0$ pour tout x de Ω .
En particulier, Δ est négligeable pour $m \otimes m$ si m est diffuse.

2) *Une courbe paramétrée du plan est-elle de surface nulle* ?

On se donne une fonction continue F de $[0,1]$ dans $\mathbb{R}^2$, représentée par $t \to F(t) = \left(f(t), g(t)\right)$, f et g étant des fonctions numériques continues sur $[0,1]$. Soit C la « courbe » image de $[0,1]$ par F : C est fermée et bornée dans le plan donc est mesurable pour m , m désignant la mesure de Borel sur les compacts de $\mathbb{R}^2$ (la *surface*). Aucune confusion n'étant à craindre, on désignera aussi par m la mesure de Borel sur $[0,1]$ (la *longueur*).

On sait qu'il existe une fonction *continue* F telle que $C = [0,1]^2$ (courbe de **Peano**), d'où $m(C) = 1$, résultat dérangeant. On sait aussi que $m(C) = 0$ lorsque les fonctions numériques f et g sont *toutes les deux* C^1 sur $[0,1]$ (page **215**, *Un résultat sur les surfaces*).

Nous allons prouver que C est encore de mesure nulle *si l'une seulement* des fonctions f ou g *est* C^1 (l'autre demeurant continue). Supposons par exemple g à dérivée bornée et f continue sans plus. La section de C en le réel x est l'ensemble des y pour lesquels il existe t tel que $x = f(t)$ et $y = g(t)$, donc $C_x = g\left(f^{-1}(x)\right)$, t décrivant $f^{-1}(x)$. Puisque $[0,1] = \sum_{x\in\mathbb{R}} f^{-1}(x)$, la mesure du fermé $f^{-1}(x)$ est nulle pour tous les réels x sauf sur un ensemble au plus dénombrable D de $\mathbb{R}$ *parce que la famille des réels positifs* $m\left(f^{-1}(x)\right)$ *est sommable* (il suffit que m soit additive). La fonction g étant à dérivée bornée, $C_x = g\left(f^{-1}(x)\right)$ est de mesure nulle, comme $f^{-1}(x)$, sur $\mathbb{R} - D$ (p. **156**). Il en résulte que l'intégrale de $x \to m(C_x)$ sur $\mathbb{R} - D$ vaut zéro ; comme D est négligeable pour m :

$$\int_0^1 m(C_x)dx = m(C) = 0 .$$

3) *Loi de* $Z(x, y) = X(x) + Y(y)$ *quand* X *et* Y *sont à densité*

X et Y sont deux variables aléatoires réelles de densité f et g sur $(\Omega, \mathcal{T}, p)$; Z est (p. **217**) une variable aléatoire sur Ω^2 muni de la tribu produit $\mathcal{T} \otimes \mathcal{T}$ (et de la probabilité produit $q = p \otimes p$). On sait déjà que Z et $X+Y$ ont même espérance (p. **218**). Nous allons montrer de plus que Z a, résultat facile à retenir, *la même fonction de répartition que* $X+Y$ *quand* X *et* Y *sont indépendantes*. Nous en donnons trois démonstrations.

Démonstration 1, avec les outils du texte :
Pour a réel, posons $E = \{\omega\in \Omega : X(\omega) + Y(\omega) \le a\}$ et $\mathcal{E} = \{(\omega_1, \omega_2) \in \Omega^2 : X(\omega_1) + Y(\omega_2) \le a\}$
et prouvons que $p(E) = q(\mathcal{E})$. Le calcul de $p(E)$ est classique quand X et Y sont indépendantes:

$p(E) = \int_{-\infty}^{+\infty} F(a-t)g(t)dt = \int_{-\infty}^{+\infty} F(t)g(a-t)dt$, F étant la fonction de répartition de X

ou

$p(E) = \int_{-\infty}^{+\infty} G(a-t)f(t)dt = \int_{-\infty}^{+\infty} G(t)f(a-t)dt$, G étant la fonction de répartition de Y .

Pour évaluer $q(\mathcal{E})$, on passe par les sections : $\mathcal{E}_{\omega_1} = \{\omega_2 \in \Omega : X(\omega_1) + Y(\omega_2) \le a\}$; il vient
$p(\mathcal{E}_{\omega_1}) = G\left(a - X(\omega_1)\right)$ et donc $q(\mathcal{E}) = \int_\Omega G\left(a - X(\omega_1)\right)d\omega_1$.
Or
$\int_\Omega G\left(a - X(\omega_1)\right)d\omega_1 = \int_{\mathbb{R}} G(x)\varphi(x)dx$, où φ est une densité de $\omega \to a - X(\omega)$.
Un petit calcul donne $\varphi(x) = f(a-x)$, ce qui termine la démonstration.

Deuxième démonstration, avec le délicat Théorème de la fonction caractéristique :

si $E(e^{itU}) = E(e^{itV})$ pour tout réel t , les variables aléatoires U et V ont même loi.

(Consulter par exemple D. Foata et A. Fuchs, *Calcul des probabilités*, Masson, édition 1996, p. 171, Th. 7.1.)

$$E(e^{itZ}) = \int_{\Omega^2} e^{itZ(x,y)} dp(x)dp(y) = \int_{\Omega} e^{itX(x)} dp(x) \int_{\Omega^2} e^{itY(y)} dp(y) = E(e^{itX}).E(e^{itY})$$

Comme X et Y sont indépendantes, il en de même de e^{itX} et e^{itY} et alors $E(e^{itX}).E(e^{itY}) = E(e^{it(X+Y)})$,
donc Z et $X+Y$ ont même loi.

Troisième démonstration, avec la décomposition de Z en somme de v.a. indépendantes

En effet, $Z = U + V$, avec $U(x,y) = X(x)$ et $V(x,y) = Y(y)$. Les fonctions U et V dont des variables aléatoires sur Ω^2 : par exemple, $(U \le a) = (X \le a) \times \Omega$, qui est dans la tribu produit. La loi de U est celle de X, la loi de V est celle de Y : par exemple, $q(U \le a) = q\big((X \le a) \times \Omega\big) = p(X \le a)$. On en déduit

$$E(U) = E(X),\ E(V) = E(Y),\ E(Z) = E(U) + E(V) = E(X) + E(Y).$$

Les variables U et V sont indépendantes, par exemple

$$q(U \le a \text{ et } V \le b) = q\big((X \le a) \times (Y \le b)\big) = p(X \le a).p(Y \le b) = q(U \le a).q(V \le b).$$

Voilà pourquoi la loi de Z est celle de $X+Y$ quand on suppose X et Y indépendantes.
De plus, $V(Z) = V(U) + V(V) = V(X) + V(Y)$.

****