cours 16, le lundi 22 mars 2010

Preuve de l'inégalité de Jensen : compléments sur la convexité

Lemme des trois pentes. *On suppose que φ est une fonction convexe définie sur un intervalle* $\mathrm{I} \subset \mathbb{R}$. *Si u, v, w sont trois points de* I *tels que $u < v < w$, on a*

$$\frac{\varphi(v) - \varphi(u)}{v - u} \le \frac{\varphi(w) - \varphi(u)}{w - u} \le \frac{\varphi(w) - \varphi(v)}{w - v}.$$

Preuve. — Posons $t = (v - u)/(w - u)$; alors $0 < t < 1$ et

$$v = u + (v - u) = u + t(w - u) = (1 - t)u + tw,$$

donc par la convexité de φ on a

$$\varphi(v) = \varphi((1 - t)u + tw) \le (1 - t)\varphi(u) + t\varphi(w) = \varphi(u) + t\big(\varphi(w) - \varphi(u)\big), \tag{1}$$

donc

$$\varphi(v) - \varphi(u) \le t(\varphi(w) - \varphi(u)) \quad \text{et} \quad \frac{\varphi(v) - \varphi(u)}{v - u} \le \frac{\varphi(w) - \varphi(u)}{w - u},$$

et pour l'autre inégalité, on transforme (1) en

$$(1 - t)(\varphi(w) - \varphi(u)) \le \varphi(w) - \varphi(v) \quad \text{donc} \quad \frac{\varphi(w) - \varphi(u)}{w - u} \le \frac{\varphi(w) - \varphi(v)}{w - v}.$$

Conséquence. *La fonction φ admet une dérivée à droite à valeurs dans $[-\infty, +\infty[$ en tout point $t \in$* I *qui n'est pas un majorant de* I, *et une dérivée à gauche à valeurs dans $]-\infty, +\infty]$ en tout point qui n'est pas un minorant de* I.

Preuve. — Considérons un point $t \in \mathrm{I}$ qui n'est pas un majorant de I ; d'après le lemme précédent, la fonction pente $y \to (\varphi(y) - \varphi(t))/(y - t)$, définie pour $y > t$ élément de I, est croissante. La pente a donc une limite à droite $\varphi'_d(t)$ au point t, à condition d'admettre la valeur $-\infty$ pour limite possible. La vérification de l'affirmation sur la dérivée à gauche est identique.

Une fonction convexe définie sur un intervalle fermé $\mathrm{I} = [a, b]$ peut ne pas être continue aux extrémités de l'intervalle. Les propriétés de la fonction convexe sont plus agréables si on considère l'intérieur de l'intervalle, ou bien si, comme dans l'énoncé qui suit, on suppose tout de suite que l'intervalle I est ouvert.

Corollaire. *Soit $\varphi : \mathrm{I} \to \mathbb{R}$ une fonction convexe sur un intervalle* ouvert I *; en tout point de* I, *φ admet une dérivée à droite et une dérivée à gauche finies, donc φ est continue sur l'ouvert* I. *Si $t_0 < t_1$ sont deux points de* I, *on a $\varphi'_g(t_0) \le \varphi'_d(t_0) \le \varphi'_g(t_1) \le \varphi'_d(t_1)$: les deux fonctions φ'_g et φ'_d sont finies et croissantes sur* I *et $\varphi'_g \le \varphi'_d$; la fonction φ'_d est continue à droite et φ'_g continue à gauche sur* I. *L'ensemble* D *des points t de* I *où $\varphi'_g(t) < \varphi'_d(t)$ est au plus dénombrable, la fonction φ est dérivable en tout point $t \in \mathrm{I} \setminus \mathrm{D}$.*

Preuve. — Si t est un point de l'ouvert I, ce point n'est ni minorant ni majorant de I ; il y a d'après ce qui précède une dérivée à gauche et une dérivée à droite pour φ au point t, et la dérivée à gauche est plus petite que la dérivée à droite : supposons que y, z soient deux points de I tels que $y < t < z$, et supposons que $h > 0$ soit assez petit pour que $y < t-h$ et $t+h < z$; appliquons le lemme aux points $y < t-h < t$, $t-h < t < t+h$ et $t < t+h < z$; on obtient

$$\frac{\varphi(t)-\varphi(y)}{t-y} \le \frac{\varphi(t)-\varphi(t-h)}{h} \le \frac{\varphi(t+h)-\varphi(t)}{h} \le \frac{\varphi(z)-\varphi(t)}{z-t} \ ;$$

en passant à la limite quand $h \to 0$ on obtient

$$\frac{\varphi(t)-\varphi(y)}{t-y} \le \varphi'_g(t) \le \varphi'_d(t) \le \frac{\varphi(z)-\varphi(t)}{z-t},$$

ce qui montre que les deux demi-dérivées sont finies, et entraîne la continuité de φ au point t. Si $t_0 < t_1$ sont deux points de I, on obtient en appliquant ce qui précède à $t = t_0$ et à $t = t_1$,

$$\varphi'_g(t_0) \le \varphi'_d(t_0) \le \frac{\varphi(t_1)-\varphi(t_0)}{t_1-t_0} \le \varphi'_g(t_1) \le \varphi'_d(t_1). \tag{2}$$

Montrons la continuité à droite de φ'_d en tout point t_0 de I ; par (2) et par la définition de la dérivée à droite, il existe $t_1 > t_0$ tel que

$$\varphi'_d(t_0) \le \frac{\varphi(t_1)-\varphi(t_0)}{t_1-t_0} \le \varphi'_d(t_0) + \varepsilon\,;$$

par la continuité de φ au point t_0, on peut trouver $h_0 > 0$ assez petit pour que $t_0+h_0 < t_1$ et pour que pour tout $h \in]0, h_0[$, on ait

$$\frac{\varphi(t_1)-\varphi(t_0+h)}{t_1-t_0-h} \le \frac{\varphi(t_1)-\varphi(t_0)}{t_1-t_0} + \varepsilon \le \varphi'_d(t_0) + 2\varepsilon\,;$$

d'après (2), appliqué à $t_0 + h < t_1$,

$$\varphi'_d(t_0) \le \varphi'_d(t_0+h) \le \frac{\varphi(t_1)-\varphi(t_0+h)}{t_1-t_0-h} \le \varphi'_d(t_0) + 2\varepsilon,$$

ce qui termine la preuve de la continuité à droite de φ'_d, la preuve de la continuité à gauche de φ'_g est analogue.

L'ensemble D des points de discontinuité (qui sont des sauts) de la fonction croissante φ'_d est au plus dénombrable, et de plus, les propriétés d'encadrement des dérivées gauche et droite,

$$t-h < t < t+h \ \Rightarrow\ \varphi'_g(t-h) \le \varphi'_d(t-h) \le \varphi'_g(t) \le \varphi'_d(t) \le \varphi'_g(t+h) \le \varphi'_d(t+h)$$

entraînent que φ'_g admet exactement le même ensemble D de discontinuités. En chaque point t de $\mathrm{I} \setminus \mathrm{D}$, la fonction φ est dérivable et $\varphi'(t) = \varphi'_g(t) = \varphi'_d(t)$.

Minorantes affines
Si t_0 est intérieur à I et si α est un réel qui vérifie $\varphi'_g(t_0) \le \alpha \le \varphi'_d(t_0)$, il résulte de (2) que

$$t \in \mathrm{I} \ \Rightarrow \ \varphi(t) - \varphi(t_0) \ge \alpha(t - t_0). \tag{3}$$

On peut récrire cette inégalité sous la forme

$$\forall t \in \mathrm{I}, \quad \varphi(t) \ge \varphi(t_0) + \varphi'_d(t_0)(t - t_0).$$

Dans le cas où φ est dérivable au point t_0, la fonction $t \to \varphi(t_0) + \varphi'(t_0)(t - t_0)$ est la fonction affine tangente à φ au point t_0. L'inégalité dit que le graphe de la fonction convexe est *au-dessus* de ses tangentes (pour toutes les tangentes qui existent).

Si μ est une mesure finie sur $(\mathrm{X}, \mathcal{F})$, si f est μ-intégrable à valeurs dans I, l'existence d'une minorante affine de φ implique que $x \in \mathrm{X} \to \varphi(f(x))$ est *quasi-intégrable*, c'est-à-dire que la partie négative de $\varphi(f)$ est intégrable. En effet, si

$$\forall t \in \mathrm{I}, \quad \alpha t + \beta \le \varphi(t),$$

alors $\varphi(f(x)) \ge \alpha f(x) + \beta =: h(x)$ qui est intégrable (la constante β est μ-intégrable puisque μ est finie). On a

$$-\varphi(f(x)) \le -h(x) \le |h(x)|,$$

donc $\varphi(f))^- = \max(0, -\varphi(f)) \le |h|$ est intégrable. On peut poser

$$\int_{\mathrm{X}} \varphi(f(x)) \,\mathrm{d}\mu(x) = \int_{\mathrm{X}} \varphi(f(x))^+ \,\mathrm{d}\mu(x) - \int_{\mathrm{X}} \varphi(f(x))^- \,\mathrm{d}\mu(x)$$

qui peut prendre la valeur $+\infty$, mais ne prend pas la forme indéterminée embarrassante $+\infty - (+\infty)$.

La preuve de Jensen résulte de (3) appliquée au point

$$t_0 = m = \int_{\mathrm{X}} f(x) \,\mathrm{d}\mu(x) \in \mathrm{I}.$$

V. Fubini

V.1. Construction de la mesure produit

L'idée pour construire la mesure de Lebesgue λ_2 sur le carré $[0,1]^2$ est de découper un sous-ensemble C du carré en tranches verticales C_x,

$$\mathrm{C}_x = \{y \in [0,1] : (x,y) \in \mathrm{C}\}, \qquad x \in [0,1]$$

et de poser

$$\lambda_2(\mathrm{C}) = \int_0^1 \lambda_1(\mathrm{C}_x) \,\mathrm{d}x = \int_0^1 \Bigl(\int_0^1 \mathbf{1}_{\mathrm{C}}(x,y) \,\mathrm{d}y \Bigr) \,\mathrm{d}x.$$

Il va falloir travailler pour justifier l'existence des intégrales précédentes.

On considère deux espaces mesurables $(\mathrm{X},\mathcal{F})$ et $(\mathrm{Y},\mathcal{G})$ et l'espace mesurable produit $(\mathrm{X}\times\mathrm{Y},\mathcal{F}\otimes\mathcal{G})$.

Lemme 1. *Si ν est une mesure σ-finie sur $(\mathrm{Y},\mathcal{G})$, alors pour tout ensemble $\mathrm{C}\in\mathcal{F}\otimes\mathcal{G}$ on a*

a.— *pour tout $x\in\mathrm{X}$, la fonction $y\to\mathbf{1}_\mathrm{C}(x,y)$ est $\mathcal{G}$-mesurable sur Y,*

b.— *la fonction $x\to\int_\mathrm{Y}\mathbf{1}_\mathrm{C}(x,y)\,\mathrm{d}\nu(y)$ est $\mathcal{F}$-mesurable sur X, à valeurs dans* $[0,+\infty]$.

Preuve. — Commençons par expliquer qu'il suffit de traiter le cas des mesures finies ; si le résultat est vrai pour toute mesure finie, alors il est encore vrai pour ν mesure σ-finie : en effet, soit $(\mathrm{Y}_n)\subset\mathcal{G}$ une suite croissante d'ensembles de mesure finie pour ν, qui tend en croissant vers Y ; pour tout entier n, la mesure à densité $\mathrm{d}\nu_n(y)=\mathbf{1}_{\mathrm{Y}_n}(y)\,\mathrm{d}\nu(y)$ est finie, donc la fonction

$$x\to\int_\mathrm{Y}\mathbf{1}_\mathrm{C}(x,y)\,\mathrm{d}\nu_n(y)=\int_\mathrm{Y}\mathbf{1}_{\mathrm{Y}_n}(y)\mathbf{1}_\mathrm{C}(x,y)\,\mathrm{d}\nu(y)$$

est $\mathcal{F}$-mesurable par hypothèse ; mais par le TCM pour ν, cette suite de fonctions tend simplement sur X, en croissant, vers la fonction $x\to\int_\mathrm{Y}\mathbf{1}_\mathrm{C}(x,y)\,\mathrm{d}\nu(y)$, qui sera donc $\mathcal{F}$-mesurable : la propriété **b.** est vraie pour ν ; d'autre part, la propriété **a.** ne dépend pas du fait que ν soit finie ou non. On va donc supposer ν mesure finie dans la suite de la preuve.

On veut montrer que la classe $\mathcal{D}$ des parties C du produit $\mathrm{X}\times\mathrm{Y}$ telles que

$$y\to\mathbf{1}_\mathrm{C}(x,y)\ \ \mathcal{G}\text{-mesurable pour tout } x,\quad x\to\int_\mathrm{Y}\mathbf{1}_\mathrm{C}(x,y)\,\mathrm{d}\nu(y)\ \ \mathcal{F}\text{-mesurable},$$

contient la tribu produit $\mathcal{F}\otimes\mathcal{G}$. On utilisera le lemme de classe monotone. On montre d'abord que $\mathcal{D}$ contient le π-système $\mathcal{P}$ de parties de $\mathrm{X}\times\mathrm{Y}$ formé des pavés mesurables $\mathrm{C}=\mathrm{A}\times\mathrm{B}$, $\mathrm{A}\in\mathcal{F}$ et $\mathrm{B}\in\mathcal{G}$. Ces ensembles produits forment un π-système ; en effet $\mathrm{X}\times\mathrm{Y}$ est un pavé mesurable, et

$$(\mathrm{A}_1\times\mathrm{B}_1)\cap(\mathrm{A}_2\times\mathrm{B}_2)=(\mathrm{A}_1\cap\mathrm{A}_2)\times(\mathrm{B}_1\cap\mathrm{B}_2)\in\mathcal{P}.$$

Si $\mathrm{C}=\mathrm{A}\times\mathrm{B}$ est un élément de $\mathcal{P}$, on a $\mathbf{1}_\mathrm{C}(x,y)=\mathbf{1}_\mathrm{A}(x)\mathbf{1}_\mathrm{B}(y)$, donc $y\to\mathbf{1}_\mathrm{C}(x,y)$ est la fonction $\mathbf{1}_\mathrm{A}(x)\mathbf{1}_\mathrm{B}$, multiple de $\mathbf{1}_\mathrm{B}$, donc $\mathcal{G}$-mesurable, et la deuxième est $\nu(\mathrm{B})\mathbf{1}_\mathrm{A}$, multiple de $\mathbf{1}_\mathrm{A}$, donc $\mathcal{F}$-mesurable. On a donc $\mathcal{P}\subset\mathcal{D}$.

La classe $\mathcal{D}$ est un λ-système : on montre la stabilité par différence propre grâce à la linéarité : si $\mathrm{C}\subset\mathrm{D}$ sont deux ensembles de $\mathcal{D}$, alors pour tout $x\in\mathrm{X}$

$$y\ \to\ \mathbf{1}_{\mathrm{D}\setminus\mathrm{C}}(x,y)=\mathbf{1}_\mathrm{D}(x,y)-\mathbf{1}_\mathrm{C}(x,y)$$

est $\mathcal{G}$-mesurable, et comme ν est supposée finie, on peut écrire

$$\int_\mathrm{Y}1_{\mathrm{D}\setminus\mathrm{C}}(x,y)\,\mathrm{d}\nu(y)=\int_\mathrm{Y}\mathbf{1}_\mathrm{D}(x,y)\,\mathrm{d}\nu(y)-\int_\mathrm{Y}\mathbf{1}_\mathrm{C}(x,y)\,\mathrm{d}\nu(y)$$

qui est donc $\mathcal{F}$-mesurable. Le passage à la limite croissante d'une suite $(\mathrm{D}_n) \subset \mathcal{D}$, de réunion D, résulte par limite simple de mesurables et TCM : pour tout $x \in \mathrm{X}$, on a que

$$\{y \to \mathbf{1}_{\mathrm{D}_n}(x,y)\} \nearrow \{y \to \mathbf{1}_{\mathrm{D}}(x,y)\},$$

donc la limite est $\mathcal{G}$-mesurable, et

$$\{x \to \int_{\mathrm{Y}} \mathbf{1}_{\mathrm{D}_n}(x,y)\,\mathrm{d}\nu(y)\} \nearrow \{x \to \int_{\mathrm{Y}} \mathbf{1}_{\mathrm{D}}(x,y)\,\mathrm{d}\nu(y)\}$$

qui est donc $\mathcal{F}$-mesurable.

Puisque $\mathcal{D}$ est un λ-système qui contient le π-système $\mathcal{P}$, la classe $\mathcal{D}$ contient la tribu $\mathcal{F} \otimes \mathcal{G}$ engendrée par la classe $\mathcal{P}$ des pavés mesurables. On en déduit que pour tout ensemble C de la tribu produit, on a les deux propriétés **a.** et **b.** de la définition de $\mathcal{D}$. Le lemme est démontré.

Théorème. *On suppose que μ, ν sont deux mesures σ-finies. Il existe une mesure ξ sur $(\mathrm{X} \times \mathrm{Y}, \mathcal{F} \otimes \mathcal{G})$ telle que*

$$\forall \mathrm{A} \in \mathcal{F},\ \forall \mathrm{B} \in \mathcal{G}, \quad \xi(\mathrm{A} \times \mathrm{B}) = \mu(\mathrm{A})\nu(\mathrm{B}).$$

De plus, cette mesure ξ est unique.

Cette mesure s'appelle la *mesure produit*, on la note $\xi = \mu \otimes \nu$.

Preuve. — Grâce aux propriétés **a**. et **b**. du lemme, on peut considérer la fonction ξ sur $\mathcal{F} \otimes \mathcal{G}$ définie par

$$\forall \mathrm{C} \in \mathcal{F} \otimes \mathcal{G}, \quad \xi(\mathrm{C}) = \int_{\mathrm{X}} \Bigl(\int_{\mathrm{Y}} \mathbf{1}_{\mathrm{C}}(x,y)\,\mathrm{d}\nu(y) \Bigr)\,\mathrm{d}\mu(x),$$

et montrer que c'est une mesure sur $(\mathrm{X} \times \mathrm{Y}, \mathcal{F} \otimes \mathcal{G})$ par la version séries du TCM, appliquée d'abord à ν, puis à μ : si $(\mathrm{C}_n)_{n \geqslant 0} \subset \mathcal{F} \otimes \mathcal{G}$ est une suite d'ensembles deux à deux disjoints et $\mathrm{C} = \bigcup_{n \geqslant 0} \mathrm{C}_n$, on a

$$\mathbf{1}_{\mathrm{C}}(x,y) = \sum_{n \geqslant 0} \mathbf{1}_{\mathrm{C}_n}(x,y),$$

donc on aura successivement

$$\xi(\mathrm{C}) = \int_{\mathrm{X}} \Bigl(\int_{\mathrm{Y}} \sum_{n \geqslant 0} \mathbf{1}_{\mathrm{C}_n}(x,y)\,\mathrm{d}\nu(y) \Bigr)\,\mathrm{d}\mu(x)$$

$$= \int_{\mathrm{X}} \sum_{n \geqslant 0} \Bigl(\int_{\mathrm{Y}} \mathbf{1}_{\mathrm{C}_n}(x,y)\,\mathrm{d}\nu(y) \Bigr)\,\mathrm{d}\mu(x) = \sum_{n \geqslant 0} \int_{\mathrm{X}} \Bigl(\int_{\mathrm{Y}} \mathbf{1}_{\mathrm{C}_n}(x,y)\,\mathrm{d}\nu(y) \Bigr)\,\mathrm{d}\mu(x) = \sum_{n \geqslant 0} \xi(\mathrm{C}_n).$$

Comme les deux mesures μ et ν sont supposées σ-finies, il existe deux suites de parties $(\mathrm{X}_n) \subset \mathcal{F}$ et $(\mathrm{Y}_n) \subset \mathcal{G}$, croissantes, telles que $\mu(\mathrm{X}_n) < +\infty$, $\nu(\mathrm{Y}_n) < +\infty$, et qui croissent respectivement vers X et vers Y.

Si ξ_1 et ξ_2 sont deux mesures vérifiant $\xi(\mathrm{A} \times \mathrm{B}) = \mu(\mathrm{A})\nu(\mathrm{B})$ pour tous A, B, elles sont égales sur la classe $\mathcal{C}$ des produits de mesure finie,

$$\mathcal{C} = \{\mathrm{A} \times \mathrm{B} : \mathrm{A} \in \mathcal{F}, \mu(\mathrm{A}) < +\infty, \mathrm{B} \in \mathcal{G}, \nu(\mathrm{B}) < +\infty\}.$$

Cette classe $\mathcal{C}$ est stable par intersection finie et contient la suite $(\mathrm{X}_n \times \mathrm{Y}_n)$. De plus la classe $\mathcal{C}$ engendre la tribu $\mathcal{F} \otimes \mathcal{G}$: pour tout pavé mesurable $\mathrm{A} \times \mathrm{B}$, on voit que $\mathrm{A} \times \mathrm{B}$ est la réunion dénombrable des pavés de mesure finie

$$(\mathrm{A} \times \mathrm{B}) \cap (\mathrm{X}_n \times \mathrm{Y}_n) = (\mathrm{A} \cap \mathrm{X}_n) \times (\mathrm{B} \cap \mathrm{Y}_n) \in \mathcal{C},$$

donc tous les générateurs $\mathrm{A} \times \mathrm{B}$ de la tribu $\mathcal{F} \otimes \mathcal{G}$ sont dans $\sigma(\mathcal{C})$. On a donc $\mathcal{F} \otimes \mathcal{G} \subset \sigma(\mathcal{C})$, et l'inclusion réciproque est claire car $\mathcal{C}$ est contenue dans la famille génératrice pour $\mathcal{F} \otimes \mathcal{G}$ formée de tous les pavés mesurables.

L'unicité permet de montrer que

$$\forall \mathrm{C} \in \mathcal{F} \otimes \mathcal{G}, \quad \xi(\mathrm{C}) = \int_{\mathrm{Y}} \Bigl(\int_{\mathrm{X}} \mathbf{1}_{\mathrm{C}}(x, y) \, \mathrm{d}\mu(x) \Bigr) \, \mathrm{d}\nu(y).$$

Il suffit de se convaincre que la preuve qui a été faite en commençant (c'est-à-dire, dans l'intégrale « intérieure ») par intégrer en y peut être faite aussi bien en commençant par intégrer en x.

V.2. Intégration par rapport à une mesure produit

V.2.1. *Fubini positif*

Théorème. *On suppose que μ, ν sont deux mesures σ-finies sur $(\mathrm{X}, \mathcal{F})$ et sur $(\mathrm{Y}, \mathcal{G})$ respectivement. Pour toute fonction f sur $\mathrm{X} \times \mathrm{Y}$, qui est $\mathcal{F} \otimes \mathcal{G}$-mesurable à valeurs dans $[0, +\infty]$, on a :*

— *pour tout $x \in \mathrm{X}$, la fonction $y \to f(x, y)$ est $\mathcal{G}$-mesurable,*
— *la fonction $x \to \int_{\mathrm{Y}} f(x, y) \, \mathrm{d}\nu(y)$ est $\mathcal{F}$-mesurable, et*

$$\int_{\mathrm{X} \times \mathrm{Y}} f \, \mathrm{d}(\mu \otimes \nu) = \int_{\mathrm{X}} \Bigl(\int_{\mathrm{Y}} f(x, y) \, \mathrm{d}\nu(y) \Bigr) \, \mathrm{d}\mu(x) = \int_{\mathrm{Y}} \Bigl(\int_{\mathrm{X}} f(x, y) \, \mathrm{d}\mu(x) \Bigr) \, \mathrm{d}\nu(y).$$

Preuve. — La stratégie est toujours la même : on commence par les indicatrices d'ensembles de la tribu $\mathcal{F} \otimes \mathcal{G}$, puis les $\mathcal{F} \otimes \mathcal{G}$-étagées ≥ 0, puis les $\mathcal{F} \otimes \mathcal{G}$-mesurables ≥ 0.

Si $f = \mathbf{1}_{\mathrm{C}}$, avec $\mathrm{C} \in \mathcal{F} \otimes \mathcal{G}$, on obtient le résultat par le lemme 1 et par la définition de la mesure $\xi = \mu \otimes \nu$; si $\varphi = \sum_{j=1}^n c_j \mathbf{1}_{\mathrm{C}_j}$ est $\mathcal{F} \otimes \mathcal{G}$-étagée, on a $x \to \varphi(x, y)$ mesurable par combinaison linéaire de fonctions mesurables, puis par positive-linéarité de l'intégrale

$$\Bigl\{ x \to \int_{\mathrm{Y}} \Bigl(\sum_{j=1}^n c_j \mathbf{1}_{\mathrm{C}_j}(x, y) \Bigr) \, \mathrm{d}\nu(y) \Bigr\} = \Bigl\{ x \to \sum_{j=1}^n c_j \int_{\mathrm{Y}} \mathbf{1}_{\mathrm{C}_j}(x, y) \, \mathrm{d}\nu(y) \} \Bigr\}$$

est $\mathcal{F}$-mesurable et

$$\int_{\mathrm{X}} \Bigl(\int_{\mathrm{Y}} \Bigl(\sum_{j=1}^n c_j \mathbf{1}_{\mathrm{C}_j}(x, y) \Bigr) \, \mathrm{d}\nu(y) \Bigr) \, \mathrm{d}\mu(x) = \sum_{j=1}^n c_j \int_{\mathrm{X}} \Bigl(\int_{\mathrm{Y}} \mathbf{1}_{\mathrm{C}_j}(x, y) \, \mathrm{d}\nu(y) \Bigr) \, \mathrm{d}\mu(x)$$

$$= \sum_{j=1}^n c_j \int_{\mathrm{X} \times \mathrm{Y}} \mathbf{1}_{\mathrm{C}_j} \, \mathrm{d}\xi = \int_{\mathrm{X} \times \mathrm{Y}} \Bigl(\sum_{j=1}^n c_j \mathbf{1}_{\mathrm{C}_j} \Bigr) \, \mathrm{d}\xi = \int_{\mathrm{X} \times \mathrm{Y}} \varphi \, \mathrm{d}\xi$$

d'où le résultat dans le cas étagé.

Pour finir il faut appliquer le TCM trois fois. Si une suite (φ_n) de fonctions $\mathcal{F}\otimes\mathcal{G}$-étagées tend simplement en croissant vers f, on a d'abord pour tout x que $y \to f(x,y)$ est $\mathcal{G}$-mesurable comme limite croissante des $y \to \varphi_n(x,y)$; ensuite, on a par le TCM pour ν

$$\int_{\mathrm{Y}} \varphi_n(x,y)\,\mathrm{d}\nu(y) \nearrow \int_{\mathrm{Y}} f(x,y)\,\mathrm{d}\nu(y)$$

ce qui montre que $x \to \int_{\mathrm{Y}} f(x,y)\,\mathrm{d}\nu(y)$ est $\mathcal{F}$-mesurable ; le TCM pour ξ montre que

$$\int_{\mathrm{X}\times\mathrm{Y}} f\,\mathrm{d}\xi = \lim_n \int_{\mathrm{X}\times\mathrm{Y}} \varphi_n\,\mathrm{d}\xi$$

qui est égale, d'après le cas étagé déjà traité, à

$$\lim_n \int_{\mathrm{X}} \Big(\int_{\mathrm{Y}} \varphi_n(x,y)\,\mathrm{d}\nu(y)\Big)\,\mathrm{d}\mu(x)\,;$$

puis le TCM pour μ donne

$$\lim_n \int_{\mathrm{X}} \Big(\int_{\mathrm{Y}} \varphi_n(x,y)\,\mathrm{d}\nu(y)\Big)\,\mathrm{d}\mu(x) = \int_{\mathrm{X}} \Big(\int_{\mathrm{Y}} f(x,y)\,\mathrm{d}\nu(y)\Big)\,\mathrm{d}\mu(x),$$

et l'égalité vraie pour les étagées φ_n est passée à la limite.

Comme la mesure peut-être définie en intervertissant les intégrales, l'autre égalité découle de la même preuve.