cours 21, le lundi 12 avril 2010

Fonction caractéristique d'une variable aléatoire

La fonction caractéristique φ_{X} d'une variable aléatoire X sur $(\Omega, \mathcal{F}, \mathrm{P})$, définie par

$$(*) \qquad \varphi_{\mathrm{X}}(t) = \mathrm{E}\, \mathrm{e}^{\mathrm{i}t\mathrm{X}} = \int_\Omega \mathrm{e}^{\mathrm{i}t\mathrm{X}(\omega)}\, \mathrm{dP}(\omega) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}tx}\, \mathrm{dP}_{\mathrm{X}}(x),$$

est une fonction continue de $t \in \mathbb{R}$, par les théorèmes de continuité sous l'intégrale : en effet, la fonction sous (disons) la deuxième intégrale de $(*)$, à savoir

$$f(t, \omega) = \mathrm{e}^{\mathrm{i}t\mathrm{X}(\omega)}$$

est continue en t et admet la majoration-égalité suivante pour son module,

$$|f(t, \omega)| = |\, \mathrm{e}^{\mathrm{i}t\mathrm{X}(\omega)}\, | = 1,$$

majorant indépendant du paramètre t ; ce majorant est une fonction constante, intégrable par rapport à la mesure finie P, donc le théorème de continuité s'applique à l'intégrale fonction de t.

La fonction φ_{X} est bornée par 1,

$$|\varphi_{\mathrm{X}}(t)| = \Bigl| \int_\Omega \mathrm{e}^{\mathrm{i}t\mathrm{X}(\omega)}\, \mathrm{dP}(\omega) \Bigr| \le \int_\Omega |\, \mathrm{e}^{\mathrm{i}t\mathrm{X}(\omega)}\, |\, \mathrm{dP}(\omega) = \int_\Omega \mathrm{dP}(\omega) = 1,$$

et le maximum du module de φ_{X} est atteint au point $t = 0$,

$$\varphi_{\mathrm{X}}(0) = \mathrm{E}\, \mathrm{e}^0 = 1.$$

Pour voir si on peut dériver φ_{X}, on regarde la dérivée partielle de la fonction $f(t, \omega)$ mentionnée plus haut, à savoir

$$\frac{\partial f}{\partial t}(t, \omega) = \mathrm{i}\mathrm{X}(\omega)\, \mathrm{e}^{\mathrm{i}t\mathrm{X}(\omega)}\,;$$

la majoration du module de cette dérivée partielle va de soi,

$$\bigl| \mathrm{i}\mathrm{X}(\omega)\, \mathrm{e}^{\mathrm{i}t\mathrm{X}(\omega)} \bigr| = |\mathrm{X}(\omega)|.$$

Le théorème de la théorie de Lebesgue donne : pour que φ_{X} soit dérivable, il suffit que X soit intégrable ; dans ce cas,

$$\varphi'_{\mathrm{X}}(t) = \mathrm{i}\, \mathrm{E}\bigl(\mathrm{X}\, \mathrm{e}^{\mathrm{i}t\mathrm{X}}\bigr) = \mathrm{i} \int_\Omega \mathrm{X}(\omega)\, \mathrm{e}^{\mathrm{i}t\mathrm{X}(\omega)}\, \mathrm{dP}(\omega) = \mathrm{i} \int_{\mathbb{R}} x\, \mathrm{e}^{\mathrm{i}tx}\, \mathrm{dP}_{\mathrm{X}}(x).$$

En fait, la fonction caractéristique est de classe C^1 dans ce cas, car la dérivée partielle $\frac{\partial f}{\partial t}$ est continue par rapport à t, donc φ'_{X} est continue par le théorème de continuité.

On peut continuer à dériver : si $\mathrm{E}\, |\mathrm{X}|^n < +\infty$ pour un entier $n \ge 1$, alors φ_{X} est de classe C^n ; en particulier, quand $\mathrm{X} \in \mathrm{L}^2$,

$$\varphi''_{\mathrm{X}}(t) = -\, \mathrm{E}\, \mathrm{X}^2\, \mathrm{e}^{\mathrm{i}t\mathrm{X}}\,.$$

Quand X a des moments de tous les ordres n, la fonction φ_{X} est de classe C^∞ et

$$\forall n \ge 0, \quad \varphi_{\mathrm{X}}^{(n)}(0) = \mathrm{i}^n\, \mathrm{E}\, \mathrm{X}^n.$$

Rappel-exemple : v.a. U uniforme sur $[-1,1]$. Dans ce cas tous les moments existent, et de plus la fonction caractéristique φ_{U} est non seulement C^∞, mais analytique, somme d'une série entière de rayon de convergence $+\infty$; en effet, on a calculé

$$\varphi_{\mathrm{U}}(t) = \frac{\sin t}{t} = \sum_{n=0}^{+\infty} (-1)^n \frac{t^{2n}}{(2n+1)!}.$$

On vérifie que le développement en série de Taylor

$$\varphi_{\mathrm{U}}(0) + \frac{\varphi_{\mathrm{U}}''(0)}{2} t^2 + \cdots = 1 - \frac{t^2}{6} + \cdots$$

redonne $\mathrm{E}\,\mathrm{U} = 0$, $\mathrm{E}\,\mathrm{U}^2 = -\varphi_{\mathrm{U}}''(0) = 1/3$.

Dans le cas de la loi gaussienne centrée réduite, ou de sa densité

$$g(x) = \frac{\mathrm{e}^{-x^2/2}}{\sqrt{2\pi}},$$

on a vu que

$$\widehat{g}(t) = \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}\,tx}\, \mathrm{e}^{-x^2/2} \,\frac{\mathrm{d}x}{\sqrt{2\pi}} = \mathrm{e}^{-t^2/2}\,.$$

Comme $\widehat{g}$ est proportionnelle à g, on retrouve évidemment la fonction g à partir de sa transformée de Fourier, en récrivant (compte tenu de la parité des fonctions) l'égalité précédente sous la forme

$$\mathrm{e}^{-x^2/2} = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\,tx}\, \widehat{g}(t) \,\frac{\mathrm{d}t}{\sqrt{2\pi}} \ ;$$

on a donc

$$g(x) = \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\,tx}\, \widehat{g}(t) \,\frac{\mathrm{d}t}{2\pi} \ ;$$

on verra que ce n'est pas un phénomène isolé : c'est un cas particulier de la *formule d'inversion de Fourier*.

On va modifier légèrement l'énoncé du théorème annoncé la dernière fois. On a eu l'occasion de dire qu'il y a eu un peu de flottement entre deux définitions possibles de la fonction de répartition,

$$\mathrm{P}(\mathrm{X} \le t) \quad \text{ou bien} \quad \mathrm{P}(\mathrm{X} < t)\,;$$

le choix qui a été fait, $\mathrm{F}_{\mathrm{X}}(t) = \mathrm{P}(\mathrm{X} \le t)$, donne une fonction de répartition continue à droite, comme on l'a démontré dans le cours. Dans un souci de conciliation, on aurait pu prendre (on ne l'a pas fait, et pour de bonnes raisons)

$$\mathrm{F}^*_{\mathrm{X}}(t) = \frac{\mathrm{P}(\mathrm{X} < t) + \mathrm{P}(\mathrm{X} \le t)}{2}.$$

Ce choix pourrait sembler raisonnable, mais il a le défaut de n'être continu, ni à droite, ni à gauche. Néanmoins, on a

$$(**) \qquad \mathrm{F}_{\mathrm{X}}(t) = \lim_{h \searrow 0} \mathrm{F}^*_{\mathrm{X}}(t+h),$$

donc cette fonction symétrisée F_X^* détermine aussi la loi de X ; on a en effet pour tout réel $h > 0$,

$$\mathrm{F}_\mathrm{X}(t) = \mathrm{P}(\mathrm{X} \le t) \le \mathrm{P}(\mathrm{X} < t+h) \le \mathrm{F}_\mathrm{X}^*(t+h) \le \mathrm{P}(\mathrm{X} \le t+h) = \mathrm{F}_\mathrm{X}(t+h)\,;$$

cet encadrement et la continuité à droite de F_X prouvent l'égalité $(**)$ ci-dessus.

On va montrer un résultat d'inversion de Fourier pour les mesures. On y a fait le choix contestable de tout exprimer dans le langage des probabilités : une mesure μ de masse 1 sur $\mathbb{R}$ est vue comme la loi $\mu = \mathrm{P}_\mathrm{X}$ d'une variable aléatoire X, et la considération de la transformée de Fourier de μ est équivalente à la considération de la fonction caractéristique φ_X. On va chercher à calculer la mesure des intervalles $]a, b]$ pour μ, exprimée à partir de la fonction de répartition F_X comme $\mathrm{F}_\mathrm{X}(b) - \mathrm{F}_\mathrm{X}(a)$. En réalité, c'est la fonction symétrisée F_X^* qui va apparaître.

Théorème. *Pour tous réels $a < b$, on a*

$$\mathrm{F}_\mathrm{X}^*(b) - \mathrm{F}_\mathrm{X}^*(a) = \lim_{\varepsilon \searrow 0} \int_a^b \Bigl(\frac{1}{2\pi} \int_\mathbb{R} \varphi_\mathrm{X}(v)\, \mathrm{e}^{-\varepsilon^2 v^2/2}\, \mathrm{e}^{-\mathrm{i}vx}\, \mathrm{d}v \Bigr)\, \mathrm{d}x.$$

L'apparition du noyau gaussien dans la formule n'est pas une fatalité : on aurait pu utiliser d'autres fonctions dont on aurait vérifié d'avance que la transformée de Fourier donne une densité de probabilité sur $\mathbb{R}$, par exemple

$$k(v) = \frac{1}{2\pi}\, \mathrm{e}^{-|v|}$$

dont la transformée de Fourier $\widehat{k}$ est la *densité de Cauchy*

$$x \in \mathbb{R} \ \rightarrow\ \frac{1}{\pi(1+x^2)}.$$

On obtiendrait alors par la même preuve la formule

$$\mathrm{F}_\mathrm{X}^*(b) - \mathrm{F}_\mathrm{X}^*(a) = \lim_{\varepsilon \searrow 0} \int_a^b \Bigl(\frac{1}{2\pi} \int_\mathbb{R} \varphi_\mathrm{X}(v)\, \mathrm{e}^{-\varepsilon|v|}\, \mathrm{e}^{-\mathrm{i}vx}\, \mathrm{d}v \Bigr)\, \mathrm{d}x.$$

Ce qui est nécessaire, c'est d'introduire un facteur qui fait converger les intégrales, qui ne seraient pas en général convergentes.

Remarque. La loi de Cauchy, comme la loi gaussienne, est symétrique par rapport à 0 (leurs densités sont des fonctions paires). Il en résulte, pour une v.a. de Cauchy ou une v.a. gaussienne Y, que $\mathrm{P}(\mathrm{Y} > 0) = \mathrm{P}(\mathrm{Y} \le 0) = 1/2$; en effet

$$\mathrm{P}(\mathrm{Y} > 0) = \int_0^{+\infty} \mathrm{e}^{-x^2/2}\, \frac{\mathrm{d}x}{\sqrt{2\pi}} = \frac{1}{2} \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2/2}\, \frac{\mathrm{d}x}{\sqrt{2\pi}} = \frac{1}{2}.$$

C'est ce 1/2 qui va faire apparaître la demi-somme qui est dans la définition de F_X^*.

Lemme. *Si* Y *est une variable aléatoire gaussienne centrée indépendante de* X, *on a*

$$\mathrm{F}^*_\mathrm{X}(b) - \mathrm{F}^*_\mathrm{X}(a) = \lim_{\varepsilon \searrow 0} \mathrm{P}(\{a < \mathrm{X} + \varepsilon \mathrm{Y} \le b\}).$$

Preuve du lemme. — Si x, y sont des réels quelconques, on vérifie que

$$\mathbf{1}_{\{a<x+\varepsilon y \le b\}} \underset{\varepsilon \searrow 0}{\longrightarrow} \mathbf{1}_{(x=a)\,\&\,(y>0)} + \mathbf{1}_{a<x<b} + \mathbf{1}_{(x=b)\,\&\,(y\le 0)}.$$

Quand $\varepsilon \searrow 0$, on a donc en tout point ω de Ω

$$\mathbf{1}_{\{a<\mathrm{X}(\omega)+\varepsilon \mathrm{Y}(\omega) \le b\}} \to \mathbf{1}_{(\mathrm{X}(\omega)=a)\,\&\,(\mathrm{Y}(\omega)>0)} + \mathbf{1}_{\{a<\mathrm{X}(\omega)<b\}} + \mathbf{1}_{(\mathrm{X}(\omega)=b)\,\&\,(\mathrm{Y}(\omega)\le 0)},$$

qui tend en étant dominé par la constante 1, intégrable pour P. On a donc

$$\mathrm{P}(\{a < \mathrm{X} + \varepsilon\mathrm{Y} \le b\}) \to \mathrm{P}((\mathrm{X} = a)\,\&\,(\mathrm{Y} > 0)) + \mathrm{P}(\{a < \mathrm{X} < b\}) + \mathrm{P}((\mathrm{X} = b)\,\&\,(\mathrm{Y} \le 0))$$

On a dit que $\mathrm{P}(\mathrm{Y} > 0) = \mathrm{P}(\mathrm{Y} \le 0) = 1/2$, et on continue le calcul de la limite, égale à

$$\frac{1}{2}\mathrm{P}(\mathrm{X} = a) + \mathrm{P}(\{a < \mathrm{X} < b\}) + \frac{1}{2}\mathrm{P}(\mathrm{X} = b) =$$

$$= \frac{\mathrm{P}(a \le \mathrm{X} < b) + \mathrm{P}(a < \mathrm{X} \le b)}{2} = \mathrm{F}^*_\mathrm{X}(b) - \mathrm{F}^*_\mathrm{X}(a).$$

Preuve du théorème. — On suppose que l'espace Ω est assez riche pour qu'on puisse définir sur $(\Omega, \mathcal{F}, \mathrm{P})$ une variable gaussienne Y (centrée réduite) indépendante de X (voir remarque plus bas). D'après le lemme précédent, il suffit de montrer que pour tout $\varepsilon > 0$, on a

$$\mathrm{P}(a < \mathrm{X} + \varepsilon\mathrm{Y} \le b) = \int_a^b \Bigl(\frac{1}{2\pi} \int_\mathbb{R} \varphi_\mathrm{X}(v)\, \mathrm{e}^{-\varepsilon^2 v^2/2}\, \mathrm{e}^{-\mathrm{i} vz}\, \mathrm{d}v \Bigr)\, \mathrm{d}z.$$

La loi du couple (X, Y) est le produit des deux lois, donc

$$p(a, b) := \mathrm{P}(\{a < \mathrm{X} + \varepsilon\mathrm{Y} \le b\}) = \int_{\mathbb{R}^2} \mathbf{1}_{\{a<x+\varepsilon y\le b\}}\, \mathrm{e}^{-y^2/2}\, \frac{\mathrm{d}y}{\sqrt{2\pi}}\, \mathrm{dP}_\mathrm{X}(x).$$

En injectant la représentation par Fourier de la densité gaussienne, on obtient

$$p(a, b) = \int_{\mathbb{R}^2} \mathbf{1}_{\{a<\varepsilon y+x\le b\}} \Bigl(\int_\mathbb{R} \frac{1}{2\pi}\, \mathrm{e}^{-u^2/2}\, \mathrm{e}^{-\mathrm{i} uy}\, \mathrm{d}u \Bigr)\, \mathrm{d}y\, \mathrm{dP}_\mathrm{X}(x).$$

Cette intégrale en trois variables peut être traitée par Fubini, car la fonction de trois variables

$$h(x, y, u) = \mathbf{1}_{\{a<\varepsilon y+x\le b\}}\, \mathrm{e}^{-u^2/2}\, \mathrm{e}^{-\mathrm{i} uy}$$

est borélienne et intégrable sur $\mathbb{R}^3$ par rapport à la mesure considérée,

$$\int_{\mathbb{R}^3} \mathbf{1}_{\{a<\varepsilon y+x\le b\}}\, \mathrm{e}^{-u^2/2}\, |\, \mathrm{e}^{-\mathrm{i} uy}\, |\, \mathrm{d}y\, \mathrm{d}u\, \mathrm{dP}_\mathrm{X}(x) =$$

$$= \int_\mathbb{R} \Bigl(\int_{\mathbb{R}^2} \mathbf{1}_{\{a<\varepsilon y+x\le b\}}\, \mathrm{e}^{-u^2/2}\, \mathrm{d}y\, \mathrm{d}u \Bigr)\, \mathrm{dP}_\mathrm{X}(x) = \frac{b-a}{\varepsilon}\, \sqrt{2\pi} < +\infty.$$

On écrit par conséquent avec Fubini

$$p(a,b) = \frac{1}{2\pi} \int_{\mathbb{R}} \Bigl(\int_{\mathbb{R}^2} \mathbf{1}_{\{a<\varepsilon y + x \le b\}} \, \mathrm{e}^{-u^2/2} \, \mathrm{e}^{-\mathrm{i} uy} \, \mathrm{d}y \, \mathrm{d}u \Bigr) \mathrm{dP_X}(x),$$

et pour x fixé, par le changement de variables affine $(y,u) \in \mathbb{R}^2 \to (z,v) = (\varepsilon y + x, u/\varepsilon)$ de déterminant égal à 1, on a $(z-x)v = yu$, $\mathrm{d}z\,\mathrm{d}v = \mathrm{d}y\,\mathrm{d}u$ et on obtient, après ce changement dans l'intégrale double intérieure,

$$p(a,b) = \frac{1}{2\pi} \int_{\mathbb{R}} \Bigl(\int_{\mathbb{R}^2} \mathbf{1}_{\{a<z \le b\}} \, \mathrm{e}^{-\varepsilon^2 v^2/2} \, \mathrm{e}^{-\mathrm{i} v(z-x)} \, \mathrm{d}z \, \mathrm{d}v \Bigr) \mathrm{dP_X}(x)\,;$$

puis, à nouveau par un Fubini que le lecteur justifiera,

$$\begin{aligned} \mathrm{P}(\{a < \mathrm{X} + \varepsilon \mathrm{Y} \le b\}) &= \frac{1}{2\pi} \int_{\mathbb{R}^2} \mathbf{1}_{\{a<z \le b\}} \, \mathrm{e}^{-\varepsilon^2 v^2/2} \, \mathrm{e}^{-\mathrm{i} vz} \Bigl(\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i} vx} \, \mathrm{dP_X}(x) \Bigr) \mathrm{d}z \, \mathrm{d}v \\ &= \int_a^b \Bigl(\frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{e}^{-\varepsilon^2 v^2/2} \, \mathrm{e}^{-\mathrm{i} vz} \, \varphi_{\mathrm{X}}(v) \, \mathrm{d}v \Bigr) \mathrm{d}z, \end{aligned}$$

ce qui termine la preuve du théorème.

Remarque. On peut toujours « enrichir » l'espace Ω pour introduire la variable gaussienne de la preuve précédente. Il suffit de remplacer Ω par $\Omega_1 = \Omega \times \mathbb{R}$, muni de la tribu produit $\mathcal{F} \otimes \mathcal{B}$ et de la probabilité produit tensoriel $\mathrm{P} \otimes \gamma$, où γ est la probabilité gaussienne centrée réduite sur $\mathbb{R}$. Pour $\omega_1 = (\omega, x) \in \Omega_1$, on pose

$$\mathrm{X}_1(\omega_1) = \mathrm{X}(\omega), \quad \mathrm{Y}(\omega_1) = x.$$

Alors Y est gaussienne, indépendante de X_1, et X_1 a la même loi que X.

Proposition. *Si μ et ν, mesures positives finies sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, ont la même transformée de Fourier, elles sont égales. Plus précisément, pour tous les réels $a < b$ on a*

$$\mu\bigl(]a,b]\bigr) + \mu\bigl([a,b[\bigr) = 2 \lim_{\varepsilon \searrow 0} \int_a^b \Bigl(\frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i} tx - \varepsilon^2 t^2/2} \, \widehat{\mu}(t) \, \mathrm{d}t \Bigr) \mathrm{d}x.$$

Preuve. — Si $\mu(\mathbb{R}) = \widehat{\mu}(0) = 0$, c'est trivial, sinon on multiplie par un coefficient pour avoir une probabilité. On pose $\Omega = \mathbb{R}$, $\mathcal{F}$ la tribu borélienne et $\mathrm{P} = \mu$; la loi de $\mathrm{X} : x \in \Omega \to x \in \mathbb{R}$ est égale à μ, et on applique le résultat précédent avec les adaptations nécessaires (ou bien on refait tous les calculs de la preuve du théorème en changeant de langage).

Proposition. *La transformation de Fourier est injective sur $\mathrm{L}^1(\mathbb{R})$.*

Preuve. — La transformation de Fourier est linéaire sur $\mathrm{L}^1(\mathbb{R})$, il suffit donc de vérifier que son noyau ne contient que le vecteur nul de $\mathrm{L}^1(\mathbb{R})$. Si f est à valeurs complexes, on note que

$$\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i} tx} \, \overline{f(x)} \, \mathrm{d}x = \overline{\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i} tx} f(x) \, \mathrm{d}x} = \overline{\widehat{f}(-t)} = 0,$$

donc la fonction conjuguée $\overline{f}$ a une transformée de Fourier nulle ; il en résulte que les transformées des fonctions réelles $\operatorname{Re} f = (f + \overline{f})/2$ et $\operatorname{Im} f$ sont nulles aussi. Il suffit maintenant de traiter le cas de fonctions à valeurs réelles : l'injectivité dans le cas réel donnera $\operatorname{Re} f = \operatorname{Im} f = 0$, donc $f = 0$.

Si f est réelle, intégrable et $\widehat{f} = 0$, on a pour tout t

$$0 = \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}tx} f(x) \, \mathrm{d}x = \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}tx} f^{+}(x) \, \mathrm{d}x - \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}tx} f^{-}(x) \, \mathrm{d}x,$$

ce qui montre que les mesures positives finies $\mathrm{d}\mu^{+}(x) = f^{+}(x) \, \mathrm{d}x$ et $\mathrm{d}\mu^{-}(x) = f^{-}(x) \, \mathrm{d}x$ ont la même transformée de Fourier, donc ces deux mesures sont égales sur la tribu borélienne de $\mathbb{R}$. Posons

$$\mathrm{A} = \{f \geq 0\} \in \mathcal{B}_{\mathbb{R}} \, ;$$

alors,

$$\int_{\mathbb{R}} f^{+}(x) \, \mathrm{d}x = \int_{\mathbb{R}} \mathbf{1}_{\mathrm{A}}(x) f^{+}(x) \, \mathrm{d}x = \mu^{+}(\mathrm{A}) = \mu^{-}(\mathrm{A}) = \int_{\mathbb{R}} \mathbf{1}_{\mathrm{A}}(x) f^{-}(x) \, \mathrm{d}x = 0,$$

car la fonction $\mathbf{1}_{\mathrm{A}} f^{-}$ est nulle ; il en résulte que f^{+} est nulle presque partout, et de même pour f^{-} en travaillant avec $\{f \leq 0\}$. On a bien montré que $f = 0_{\mathrm{L}^1}$.