



Centre de Télé-Enseignement Sciences
Université de Provence

MASTER SCIENCES

Mention Mathématiques et Applications -- (2M1MAP)

Expédition dans la semaine n°	Etape	Code UE	N° d'envoi de l'UE
	2M1MAP	2TMAP14	5

Nom de l'UE : Théorie de la mesure et Probabilités (UE 1-4)

– **Contenu de l'envoi : Polycopié, chapitres 10 et 11 et corrigés des exercices des chapitres 10 et 11**

- Guide du travail à effectuer

Semaine 1 :

Etudier le chapitre 10, sections 1, 2, 3 (Transformation de Fourier)

Exercices proposés (avec corrigés) : 10.1, 10.5 (1ere question)

Semaine 2 :

Etudier le chapitre 10, sections 4 et 6 (Transf. de Fourier dans L2, Fonctions caractéristiques)

Exercices proposés (avec corrigés) : 10.10, 10.11

Semaine 3 :

Etudier le chapitre 11, section 1 (Espérance conditionnelle)

Exercices proposés (avec corrigés) : 11.1, 11.3, 11.7

Semaine 4 :

Etudier le chapitre 11, section 2 (Martingales)

Exercices proposés (avec corrigés) : 11.10, 11.23

Le deuxième devoir est à rendre à la réception de ce cinquième envoi (mi-mars)

- Coordonnées de l'enseignant responsable de l'envoi

Thierry Gallouet, LATP-CMI, 39 rue Joliot Curie, 13453 marseille cedex 13.

email : gallouet@cmi.univ-mrs.fr

fax : 04 91 11 35 52

Vous pouvez aussi consulter la page web: <http://www.cmi.univ-mrs.fr/~gallouet/tele.d>

et me poser des questions par email.



Secrétariat : Centre de Télé-Enseignement Sciences Université de Provence
case 35 3, place Victor Hugo 13331 Marseille Cedex 03 Tél : +33 (0)4 91 10 63 97 Fax : +33 (0)4 91 10 63 16

ctes@up.univ-mrs.fr

<http://www.ctes.univ-mrs.fr>

<http://www.telesup.univ-mrs.fr>

Pour rapprocher la connaissance

Chapitre 10

Transformation de Fourier

10.1 Introduction et notations

La notion de série de Fourier permet d'analyser les fonctions définies d'un compact $[a, b]$ de \mathbb{R} dans \mathbb{R} (ou \mathbb{C}). La notion de transformée de Fourier permet d'analyser les fonctions définies de \mathbb{R} (ou \mathbb{R}^N) dans \mathbb{R} (ou \mathbb{C}). La transformée de Fourier est une notion employée par exemple en théorie du signal, en théorie des probabilités et pour l'analyse des équations aux dérivées partielles.

Dans toute la suite, on considèrera l'espace mesuré $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ et on notera $d\lambda_N(x) = dx$. Soit $N \geq 1$, les espaces $\mathcal{L}_\mathbb{C}^1(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ et $L_\mathbb{C}^1(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ ont été définis dans la section 4.10 et les espaces $\mathcal{L}_\mathbb{C}^2(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ et $L_\mathbb{C}^2(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ ont été définis dans la section 6.2. On rappelle aussi que si f est une fonction définie de \mathbb{R}^N dans \mathbb{C} , la fonction f est mesurable si et seulement si ses parties réelle et imaginaire sont mesurables (chaque ensemble, \mathbb{R}^N , \mathbb{R} ou \mathbb{C} , étant muni de sa tribu borélienne). On peut, bien sûr, aussi définir les espaces $\mathcal{L}_\mathbb{C}^p(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ et $L_\mathbb{C}^p(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ pour tout $p \in [1, \infty]$.

Définition 10.1 (Espaces $L_\mathbb{C}^p(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$) Soit $N \geq 1$, $p \in [1, \infty]$ et f une fonction mesurable de \mathbb{R}^N dans \mathbb{C} (c'est-à-dire $f^{-1}(A) \in \mathcal{B}(\mathbb{R}^N)$ pour tout $A \in \mathcal{B}(\mathbb{C})$). On dit que $f \in L_\mathbb{C}^p(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ si $|f| \in L_\mathbb{R}^p(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ et on définit $\|f\|_p$ par $\|f\|_p = \||f|\|_p$ où $\||f|\|_p$ est la norme de $|f|$ dans $L^p(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ (vue au chapitre 6). L'espace $L_\mathbb{C}^p(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ est l'espace $\mathcal{L}_\mathbb{C}^p(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ quotienté par la relation d'équivalence " $= p.p.$ ". C'est un espace de Banach (complexe), c'est-à-dire un e.v.n. (sur \mathbb{C}) complet.

Remarque 10.1 Soit $N \geq 1$, $p \in [1, \infty]$ et f une fonction définie de \mathbb{R}^N dans \mathbb{C} . Il est facile de voir que $f \in L_\mathbb{C}^p(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ si et seulement si ses parties réelle et imaginaire sont dans $L_\mathbb{R}^p(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$.

10.2 Transformation de Fourier dans L^1

10.2.1 Définitions et premières propriétés

Soit $N \geq 1$. Pour $x = (x_1, \dots, x_N)^t \in \mathbb{R}^N$ et $t = (t_1, \dots, t_N)^t \in \mathbb{R}^N$, on note $x \cdot t$ le produit scalaire euclidien de x et t , c'est-à-dire $x \cdot t = \sum_{i=1}^N x_i t_i$. Dans ce chapitre, On note aussi $L_\mathbb{C}^p(\mathbb{R}^N)$ l'espace $L_\mathbb{C}^p(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$, pour $p \in [1, \infty]$.

Définition 10.2 (Transformée de Fourier dans L^1)

Soit $N \geq 1$ et $f \in L^1_{\mathbb{C}}(\mathbb{R}^N)$. Pour $t \in \mathbb{R}^N$, l'application $x \mapsto e^{-ix \cdot t} f(x)$ (définie de \mathbb{R}^N dans \mathbb{C}) appartient à $L^1_{\mathbb{C}}(\mathbb{R}^N)$. On définit alors $\hat{f}(t)$ par :

$$\hat{f}(t) = (2\pi)^{-\frac{N}{2}} \int f(x) e^{-ix \cdot t} dx. \quad (10.1)$$

La fonction \hat{f} (définie de \mathbb{R}^N dans \mathbb{C}) s'appelle transformée de Fourier de f .

On note $C_0(\mathbb{R}^N, \mathbb{C}) = \{g \in C(\mathbb{R}^N, \mathbb{C}) \text{ t.q. } g(t) \rightarrow 0 \text{ quand } |t| \rightarrow +\infty\}$. On rappelle que $C_0(\mathbb{R}^N, \mathbb{C})$ est un espace de Banach quand il est muni de la norme de la convergence uniforme, c'est-à-dire :

$$\|\varphi\|_u = \max_{x \in \mathbb{R}^N} |\varphi(x)|.$$

Proposition 10.1 Soit $N \geq 1$. Soit F l'application qui à f (appartenant à $L^1_{\mathbb{C}}(\mathbb{R}^N)$) associe sa transformée de Fourier. F est une application linéaire continue de $L^1_{\mathbb{C}}(\mathbb{R}^N)$ dans $C_0(\mathbb{R}^N, \mathbb{C})$.

DÉMONSTRATION :

- Le théorème de continuité sous le signe \int (théorème 4.9) appliqué à la fonction $(x, t) \mapsto e^{-ix \cdot t} f(x)$ entraîne immédiatement que \hat{f} est continue.
- On montre maintenant que $\hat{f} \in C_0(\mathbb{R}, \mathbb{R})$.
- Cas $N = 1$. On remarque que pour $t \neq 0$, on a, comme $e^{i\pi} = -1$,

$$\hat{f}(t) = -(2\pi)^{-\frac{1}{2}} \int e^{-i(x - \frac{\pi}{t})t} f(x) dx,$$

et donc avec un changement de variable simple,

$$\hat{f}(t) = -(2\pi)^{-\frac{1}{2}} \int e^{-iyt} f(y + \frac{\pi}{t}) dy.$$

On en déduit que

$$2\hat{f}(t) = (2\pi)^{-\frac{1}{2}} \int e^{-ixt} (f(x) - f(x + \frac{\pi}{t})) dx$$

et donc que $|\hat{f}(t)| \leq \frac{1}{2} (2\pi)^{-\frac{1}{2}} \|f(\cdot) - f(\cdot + \frac{\pi}{t})\|_1$. Le théorème de continuité en moyenne dans L^1 (théorème 5.6) donne alors le fait que $\hat{f}(t) \rightarrow 0$ quand $|t| \rightarrow \infty$.

- Cas $N > 1$. On reprend la même méthode. Pour $t \neq 0$, $t = (t_1, \dots, t_N)^t$, il existe $j \in \{1, \dots, N\}$ t.q. $|t_j| = \max_{k=1, \dots, N} |t_k|$. On a alors, comme $e^{i\pi} = -1$, en notant e_j le j -ième vecteur de base de \mathbb{R}^N ,

$$\hat{f}(t) = -(2\pi)^{-\frac{N}{2}} \int e^{-i(x - \frac{\pi}{t_j} e_j) \cdot t} f(x) dx,$$

et donc avec un changement de variable simple,

$$\hat{f}(t) = -(2\pi)^{-\frac{N}{2}} \int e^{-iy \cdot t} f(y + \frac{\pi}{t_j} e_j) dy.$$

On en déduit que $|\hat{f}(t)| \leq \frac{1}{2} (2\pi)^{-\frac{N}{2}} \|f(\cdot) - f(\cdot + \frac{\pi}{t_j} e_j)\|_1$. Le théorème de continuité en moyenne dans L^1 (théorème 8.2) donne alors le fait que $\hat{f}(t) \rightarrow 0$ quand $|t| \rightarrow \infty$.

La transformée a la propriété intéressante de transformer la convolution en produit. Ceci est montré dans la proposition suivante.

Proposition 10.2 Soient f et $g \in L^1_{\mathbb{C}}(\mathbb{R}^N)$, alors $\widehat{f \star g} = (2\pi)^{\frac{N}{2}} \hat{f} \hat{g}$.

DÉMONSTRATION : Par la proposition 7.9, on $f \star g \in L^1_{\mathbb{C}}(\mathbb{R}^N)$ et pour p.p. $x \in \mathbb{R}^N$, $f \star g(x) = \int f(x-y)g(y)dy$. On a donc, pour tout $t \in \mathbb{R}^N$,

$$\begin{aligned} \widehat{f \star g}(t) &= (2\pi)^{-\frac{N}{2}} \int \left(\int f(x-y)g(y)dy \right) e^{-ix \cdot t} dx = \\ &= (2\pi)^{-\frac{N}{2}} \int \left(\int f(x-y)g(y)e^{-i(x-y) \cdot t} e^{-iy \cdot t} dy \right) dx. \end{aligned}$$

En appliquant le théorème de Fubini (théorème 7.3) à la fonction $(x, y) \mapsto f(x-y)g(y)e^{-i(x-y) \cdot t} e^{-iy \cdot t}$ (qui est bien intégrable sur \mathbb{R}^{2N} car son module est la fonction $(x, y) \mapsto |f(x-y)g(y)|$ dont l'intégrale sur \mathbb{R}^{2N} est égale à $\|f\|_1 \|g\|_1$), on obtient :

$$\widehat{f \star g}(t) = (2\pi)^{-\frac{N}{2}} \int \left(\int f(x-y)e^{-i(x-y) \cdot t} dx \right) g(y)e^{-iy \cdot t} dy.$$

Comme, pour tout $y \in \mathbb{R}^N$, $\int f(x-y)e^{-i(x-y) \cdot t} dx = \int f(z)e^{-iz \cdot t} dz = (2\pi)^{\frac{N}{2}} \hat{f}(t)$, on en déduit :

$$\widehat{f \star g}(t) = \hat{f}(t) \int g(y)e^{-iy \cdot t} dy = (2\pi)^{\frac{N}{2}} \hat{f}(t) \hat{g}(t),$$

ce qui est le résultat annoncé. ■

10.2.2 Théorème d'inversion

Il est naturel de se poser les deux questions suivantes :

- (i) La transformée de Fourier d'une fonction f caractérise-t-elle la fonction f (c'est-à-dire si $\hat{f} = \hat{g}$, a-t-on $f = g$ p.p.) ?
- (ii) peut-on retrouver la fonction à partir de sa transformée de Fourier ?

Les réponses à ces questions sont fournies par le théorème d'inversion de Fourier :

Théorème 10.1 (Inversion partielle de la transformée de Fourier) Soit $N \geq 1$ et $f \in L^1_{\mathbb{C}}$ t.q. $\hat{f} \in L^1_{\mathbb{C}}$. On a alors $f = \widehat{\hat{f}(-\cdot)}$ p.p., c'est-à-dire :

$$f(t) = (2\pi)^{-\frac{N}{2}} \int \hat{f}(x)e^{ix \cdot t} dx, \text{ pour presque tout } t \in \mathbb{R}^N.$$

DÉMONSTRATION : La démonstration fait l'objet de l'exercice 10.1. ■

Une conséquence de ce théorème est l'injectivité de l'application F , qui fournit donc une réponse positive à la question (i). En effet, soient f et $g \in L^1_{\mathbb{C}}$ t.q. $\hat{f} = \hat{g}$; alors par linéarité, $\widehat{f-g} = 0$ et donc $\widehat{f-g} \in L^1_{\mathbb{C}}$. En appliquant le théorème d'inversion, on a donc $f = g$ p.p..

Ce théorème apporte aussi une réponse partielle à la question (ii) : on peut calculer f à partir de \hat{f} dès que $\hat{f} \in L^1$. Il faut remarquer à ce propos que L^1 n'est pas stable par transformation de Fourier (voir exercice 10.2).

10.2.3 Régularité et décroissance à l'infini

Proposition 10.3 (Différentiabilité, dimension 1)

1. Soit $f \in L^1_{\mathbb{C}}(\mathbb{R}) \cap C^1(\mathbb{R}, \mathbb{C})$ t.q. $f' \in L^1_{\mathbb{C}}(\mathbb{R})$ (où f' est la dérivée de f). Alors, pour tout $t \in \mathbb{R}$, $\widehat{f'}(t) = (it)\widehat{f}(t)$.
2. Soit $f \in L^1_{\mathbb{C}}(\mathbb{R})$ t.q. $(\cdot)f \in L^1_{\mathbb{C}}(\mathbb{R})$ (où $(\cdot)f$ est l'application qui à $x \in \mathbb{R}$ associe $xf(x)$). Alors, $\widehat{f} \in C^1(\mathbb{R}, \mathbb{C})$ et, pour tout $t \in \mathbb{R}$, $\widehat{f'}(t) = \widehat{(-i\cdot)f}(t)$.

DÉMONSTRATION : La démonstration du premier item consiste à faire une intégration par parties sur l'intervalle $[-n, n]$ puis à faire tendre n vers l'infini (en remarquant que $f(\pm n) \rightarrow 0$, quand $n \rightarrow \infty$, voir l'exercice 5.8).

Le deuxième item est une conséquence immédiate du théorème 4.10 (théorème de dérivation sous le signe \int). ■

La transformation de Fourier "transforme" donc la dérivation en multiplication par la fonction $(i\cdot)$, et la multiplication par $(-i\cdot)$ en dérivation. Cette propriété est utilisée, par exemple, pour la résolution d'équations différentielles (qui sont ainsi transformées en équations algébriques).

Cette propriété se généralise au cas de la dimension N et pour un ordre k de dérivation quelconque. On introduit pour ce faire les notations suivantes : soient $\alpha = (\alpha_1, \dots, \alpha_N)^t \in \mathbb{N}^N$ un "multi-indice" et f une fonction de \mathbb{R}^N dans \mathbb{C} . On définit $|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_N$ et :

$$D^\alpha f = \left(\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_N}}{\partial x_N^{\alpha_N}} \right) f.$$

Proposition 10.4 (Différentiabilité, dimension N)

1. Soit $N \geq 1$ et $k \geq 1$. Soit $f \in C^k(\mathbb{R}^N, \mathbb{C})$ t.q. $D^\alpha f \in L^1_{\mathbb{C}}(\mathbb{R}^N)$ pour tout $\alpha \in \mathbb{N}^N$ t.q. $|\alpha| \leq k$. Alors, $\widehat{D^\alpha f}(t) = (it)^\alpha \widehat{f}(t)$ pour tout $t \in \mathbb{R}$ et pour tout $\alpha \in \mathbb{N}^N$ t.q. $|\alpha| \leq k$, avec $(it)^\alpha = (it_1)^{\alpha_1} (it_2)^{\alpha_2} \cdots (it_N)^{\alpha_N}$.
2. Soit f t.q. $(\cdot)^\alpha f \in L^1_{\mathbb{C}}(\mathbb{R}^N)$ pour tout $\alpha \in \mathbb{N}^N$ tel que $|\alpha| \leq k$. Alors, $\widehat{f} \in C^k(\mathbb{R}^N, \mathbb{C})$ et $D^\alpha \widehat{f} = \widehat{(-i\cdot)^\alpha f}$ pour tout $\alpha \in \mathbb{N}^N$ tel que $|\alpha| \leq k$.

La proposition 10.4 montre que la dérivabilité de f entraîne la décroissance de \widehat{f} à l'infini ("plus f est dérivable, plus \widehat{f} décroît vite à l'infini"), et réciproquement. Cette remarque incite à définir l'espace des fonctions à décroissance rapide (souvent appelé "espace de Schwartz"), noté $\mathcal{S}(\mathbb{R}^N, \mathbb{C})$ ou \mathcal{S}_N en abrégé.

$$\mathcal{S}_N = \{f \in C^\infty(\mathbb{R}^N, \mathbb{C}) \text{ t.q. pour tout } \alpha \text{ et } \beta \in \mathbb{N}^N, \sup_{x \in \mathbb{R}^N} |(x)^\alpha D^\beta f(x)| < +\infty\}.$$

On va montrer l'invariance par transformation de Fourier de cet espace. On commence par remarquer que $\mathcal{S}_N \subset L^1_{\mathbb{C}}(\mathbb{R}^N)$. En effet, si $f \in \mathcal{S}_N$, en prenant des choix convenables de α et β dans la définition de \mathcal{S}_N , on remarque qu'il existe $C \in \mathbb{R}$ t.q. $(1 + |x|^{N+1})|f(x)| \leq C$. On en déduit que $f \in L^1_{\mathbb{C}}(\mathbb{R}^N)$. Plus généralement, si $f \in \mathcal{S}_N$, on remarque que $(\cdot)^\beta D^\alpha f \in L^1_{\mathbb{C}}(\mathbb{R}^N)$ pour tout $\alpha, \beta \in \mathbb{N}^N$. On obtient alors la proposition 10.5.

Proposition 10.5 Soit $N \geq 1$ et $f \in \mathcal{S}(\mathbb{R}^N, \mathbb{C})$. Pour tout $\alpha, \beta \in \mathbb{N}^N$ on a :

$$D^\alpha \widehat{((-i\cdot)^\beta f)} = (i\cdot)^\alpha \widehat{(-i\cdot)^\beta f} = (i\cdot)^\alpha D^\beta \widehat{f}, \quad (10.2)$$

$$\widehat{(-i\cdot)^\alpha D^\beta f} = D^\alpha \widehat{D^\beta f} = D^\alpha [(i\cdot)^\beta \widehat{f}]. \quad (10.3)$$

DÉMONSTRATION : La démonstration est une adaptation simple de celle de la proposition 10.4. ■

La proposition 10.5 et le théorème 10.1 nous permettent alors de remarquer que la transformée de Fourier est une bijection de \mathcal{S}_N dans \mathcal{S}_N .

Proposition 10.6 Soit $N \geq 1$. L'application F qui à f associe sa transformée de Fourier est une bijection de \mathcal{S}_N dans \mathcal{S}_N . De plus, pour tout $f \in \mathcal{S}_N$, on a $f = \widehat{\widehat{f}}(-\cdot)$.

DÉMONSTRATION : En utilisant la proposition 10.5, on montre facilement que F envoie \mathcal{S}_N dans \mathcal{S}_N . Le théorème d'inversion (théorème 10.1) donne alors que f est injective et que $f = \widehat{\widehat{f}}(-\cdot)$ pour tout $f \in \mathcal{S}_N$. De cette dernière formule, on déduit que F est surjective (et donc bijective) de \mathcal{S}_N dans \mathcal{S}_N . ■

10.3 Transformée de Fourier d'une mesure signée

Il est facile d'étendre la définition de la transformation de Fourier à une mesure signée sur les boréliens de \mathbb{R}^N . Plus précisément, si m est une mesure signée sur les boréliens de \mathbb{R}^N ($N \geq 1$), la définition 10.3 définit la fonction \widehat{m} (continue et bornée de \mathbb{R}^N dans \mathbb{C}) et, si $m = f\lambda_N$, avec $f \in L^1_{\mathbb{R}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$ (c'est-à-dire que m est la mesure signée de densité f par rapport à la mesure de Lebesgue sur les boréliens de \mathbb{R}^N), on a $\widehat{m}(t) = \widehat{f}(t)$ pour tout $t \in \mathbb{R}^N$.

Définition 10.3 (Transformée de Fourier d'une mesure signée)

Soit $N \geq 1$ et m une mesure signée sur les boréliens de \mathbb{R}^N . Soit $t \in \mathbb{R}^N$, l'application $x \mapsto e^{-ix \cdot t}$ (définie de \mathbb{R}^N dans \mathbb{C}) appartient à $L^1_{\mathbb{C}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), m)$ (car elle est continue, donc borélienne, et bornée). On définit alors $\widehat{m}(t)$ par :

$$\widehat{m}(t) = (2\pi)^{-\frac{N}{2}} \int e^{-ix \cdot t} dm(x). \quad (10.4)$$

La fonction \widehat{m} (définie de \mathbb{R}^N dans \mathbb{C}) s'appelle transformée de Fourier de m .

On rappelle que $C_b(\mathbb{R}^N, \mathbb{C}) = \{g \in C(\mathbb{R}^N, \mathbb{C}) \text{ t.q. } \sup_{t \in \mathbb{R}^N} |g(t)| < \infty\}$ et que $C_b(\mathbb{R}^N, \mathbb{C})$ est un espace de Banach quand il est muni de la norme de la convergence uniforme.

Proposition 10.7 Soit $N \geq 1$. Soit m une mesure signée sur les boréliens de \mathbb{R}^N . La fonction \widehat{m} appartient à $C_b(\mathbb{R}^N, \mathbb{C})$.

DÉMONSTRATION : Le fait que \widehat{m} est bornée est simple. On utilise la décomposition de Hahn (proposition 2.6), c'est-à-dire le fait que $m = m^+ - m^-$, où m^{\pm} sont deux mesures finies étrangères. On remarque alors que, pour tout $t \in \mathbb{R}^N$, on a $|\widehat{m}(t)| \leq |m|(\mathbb{R}^N) < +\infty$, où $|m| = m^+ + m^-$.

Le fait que \widehat{m} est continue est une conséquence immédiate du théorème de continuité sous le signe \int (théorème 4.9) appliqué à la fonction $(x, t) \mapsto e^{-ix \cdot t}$ (et avec les mesures finies m^{\pm}). ■

Comme pour la transformation de Fourier dans L^1 , on peut se demander si \widehat{m} caractérise la mesure signée m et si on peut retrouver m à partir de \widehat{m} . La proposition suivante s'intéresse à ces deux questions.

Proposition 10.8 Soit $N \geq 1$.

1. Soit m et μ deux mesures signées sur les boréliens de \mathbb{R}^N . On suppose que $\widehat{m} = \widehat{\mu}$. alors, $m = \mu$.

2. Soit m une mesure signée sur les boréliens de \mathbb{R}^N . On suppose que $\hat{m} \in \mathcal{L}_{\mathbb{C}}^1(\mathbb{R}^N)$. Alors, $m = f\lambda_N$ avec $f = \hat{m}(\cdot)$ p.p. (c'est-à-dire p.p. pour la mesure λ_N).

DÉMONSTRATION : La démonstration de cette proposition fait l'objet de l'exercice 10.5. ■

Nous avons vu dans la section précédente qu'il y avait un lien entre les propriétés de décroissance de f à l'infini et les propriétés de régularité de \hat{f} (propositions 10.3, 10.4 et 10.5). Dans le même esprit, on peut remarquer que, si f est une fonction de \mathbb{R} dans \mathbb{R}_+ , borélienne et intégrable, la continuité de \hat{f} en 0 donne une précision sur la convergence vers 0, quand $a \rightarrow +\infty$, de $\int_{]-a, a[} |f(x)| dx$. Nous donnons dans le lemme 10.1 cette précision dans le cas plus général d'une mesure (positive) finie. Ce lemme permet de démontrer le théorème 10.3 (lui-même utile pour démontrer le théorème central limite, théorème 10.4).

Lemme 10.1 Soit m une mesure (positive) finie (et non nulle) sur $\mathcal{B}(\mathbb{R})$. On pose $M = m(\mathbb{R})$ et, pour $t \in \mathbb{R}$,

$$\psi(t) = \sqrt{2\pi} \frac{\hat{m}(t) + \hat{m}(-t)}{2}.$$

(La fonction ψ prend donc ses valeurs dans \mathbb{R} , et même dans $[-M, M]$, est continue en 0 et $\psi(0) = M$.) On a alors, pour tout $a > 0$,

$$m(\{x, |x| \geq a\}) \leq a \int_0^{2/a} (M - \psi(t)) dt.$$

DÉMONSTRATION : On peut supposer $M = 1$ (il suffit, si $M \neq 1$, de remplacer la mesure m par la mesure m/M). On remarque tout d'abord que pour tout $t \in \mathbb{R}$ on a

$$\hat{m}(t) + \hat{m}(-t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (e^{-ixt} + e^{ixt}) dm(x) = \frac{2}{\sqrt{2\pi}} \int_{\mathbb{R}} \cos(xt) dm(x).$$

et donc

$$\psi(t) = \int_{\mathbb{R}} \cos(xt) dm(x).$$

On en déduit bien que $\psi(t) \in \mathbb{R}$, $|\psi(t)| \leq 1$ et que $\psi(0) = 1$.

Soit $a > 0$, on pose $\varepsilon = 2/a$ et $T = \int_0^\varepsilon (1 - \psi(t)) dt$. En utilisant le théorème de Fubini-Tonelli (théorème 7.2) on obtient

$$T = \int_0^\varepsilon \left(\int_{\mathbb{R}} (1 - \cos(xt)) dm(x) \right) dt = \int_{\mathbb{R}} \left(\int_0^\varepsilon (1 - \cos(xt)) dt \right) dm(x) = \int_{\mathbb{R}} \left(\varepsilon - \frac{\sin(\varepsilon x)}{x} \right) dm(x).$$

Mais, $(\varepsilon - \frac{\sin(\varepsilon x)}{x}) = \varepsilon(1 - \frac{\sin(\varepsilon x)}{\varepsilon x}) \geq \varepsilon/2$ si $|\varepsilon x| \geq 2$, ce qui est vrai si $|x| \geq a$ car $\varepsilon a = 2$. On a donc

$$T = \int_{\mathbb{R}} \left(\varepsilon - \frac{\sin(\varepsilon x)}{x} \right) dm(x) \geq \frac{\varepsilon}{2} m(\{x, |x| \geq a\}).$$

Obtient donc, finalement,

$$m(\{x, |x| \geq a\}) \leq \frac{2T}{\varepsilon} \leq \frac{2}{\varepsilon} \int_0^\varepsilon (1 - \psi(t)) dt = a \int_0^{2/a} (1 - \psi(t)) dt.$$

Ce qui est le résultat annoncé. ■

10.4 Transformation de Fourier dans L^2

On aimerait ici définir la transformée de Fourier d'un élément de $L^2_{\mathbb{C}}(\mathbb{R}^N) = L^2_{\mathbb{C}}(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda_N)$. On rappelle que l'espace $L^2_{\mathbb{C}}(\mathbb{R}^N)$ est un espace de Hilbert et que le produit scalaire sur $L^2_{\mathbb{C}}(\mathbb{R}^N)$ est défini par (en notant $dt = d\lambda_N(t)$) :

$$(f/g)_2 = \int f(t)\overline{g(t)}dt.$$

Il est clair que la définition de \hat{f} qu'on a donnée pour $f \in L^1_{\mathbb{C}}(\mathbb{R}^N)$ ne s'applique pas pour un élément de $L^2_{\mathbb{C}}(\mathbb{R}^N)$. Pour définir la transformée de Fourier d'un élément de $L^2_{\mathbb{C}}(\mathbb{R}^N)$, on va utiliser la densité de \mathcal{S}_N dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ (on peut montrer que $\mathcal{S}_N \subset L^2_{\mathbb{C}}(\mathbb{R}^N)$, comme on a montré que $\mathcal{S}_N \subset L^1_{\mathbb{C}}(\mathbb{R}^N)$). On va d'abord remarquer que la transformée de Fourier envoie \mathcal{S}_N dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ et que c'est une isométrie pour la norme de $L^2_{\mathbb{C}}(\mathbb{R}^N)$. On utilisera ensuite la densité de \mathcal{S}_N dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ pour définir la transformée de Fourier des éléments de $L^2_{\mathbb{C}}(\mathbb{R}^N)$.

Proposition 10.9 Soit $N \geq 1$ et $f, g \in \mathcal{S}_N$ (on a donc, en particulier, $f, g \in L^2_{\mathbb{C}}(\mathbb{R}^N)$). Alors $\hat{f}, \hat{g} \in L^2_{\mathbb{C}}(\mathbb{R}^N)$ et $(f/g)_2 = (\hat{f}/\hat{g})_2$. En particulier, $\|f\|_2 = \|\hat{f}\|_2$.

DÉMONSTRATION : Soit $f, g \in \mathcal{S}_N$. Comme $f, \hat{f} \in \mathcal{S}_N \subset L^1_{\mathbb{C}}(\mathbb{R}^N)$, on peut appliquer le théorème d'inversion (théorème 10.1). Il donne $f = \hat{\hat{f}}(-\cdot)$ et donc :

$$(f/g)_2 = \int \hat{\hat{f}}(-t)\overline{\hat{g}(t)}dt = (2\pi)^{-\frac{N}{2}} \int \left(\int e^{ix \cdot t} \hat{f}(x)dx \right) \overline{\hat{g}(t)}dt.$$

On utilise maintenant le théorème de Fubini (théorème 7.3). Il s'applique car $\hat{f}, \hat{g} \in L^1_{\mathbb{C}}(\mathbb{R}^N)$. On obtient :

$$(f/g)_2 = (2\pi)^{-\frac{N}{2}} \int \left(\int e^{ix \cdot t} \overline{\hat{g}(t)}dt \right) \hat{f}(x)dx = \int \overline{\hat{g}(t)}\hat{f}(t)dt = (\hat{f}/\hat{g})_2,$$

ce qui termine la démonstration. ■

La proposition 10.9 permet de définir, par un argument de densité, la transformée de Fourier dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$.

Théorème 10.2 Soit $N \geq 1$. Il existe une application linéaire continue \bar{F} de $L^2_{\mathbb{C}}(\mathbb{R}^N)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ t.q. :

1. Si $f \in L^1_{\mathbb{C}}(\mathbb{R}^N) \cap L^2_{\mathbb{C}}(\mathbb{R}^N)$, on a alors $\bar{F}(f) = \hat{f}$ p.p.
2. Pour tout $f, g \in L^2_{\mathbb{C}}(\mathbb{R}^N)$, on a $(f/g)_2 = (\bar{F}(f)/\bar{F}(g))_2$.
3. Pour tout $f \in L^2_{\mathbb{C}}(\mathbb{R}^N)$, on a $f = \bar{F}(\bar{F}(f))(-\cdot)$.
4. \bar{F} est une bijection de $L^2_{\mathbb{C}}(\mathbb{R}^N)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$.

Pour $f \in L^2_{\mathbb{C}}(\mathbb{R}^N)$, $\bar{F}(f)$ s'appelle la transformée de Fourier de f . Compte tenu du premier item, on notera en général, \hat{f} la transformée de Fourier de f si $f \in L^1_{\mathbb{C}}(\mathbb{R}^N)$ (et alors $\hat{f} \in C_0(\mathbb{R}^N, \mathbb{C})$) ou si $f \in L^2_{\mathbb{C}}(\mathbb{R}^N)$ (et alors $\hat{f} \in L^2_{\mathbb{C}}(\mathbb{R}^N)$).

DÉMONSTRATION : L'application $f \mapsto \hat{f}$ est définie sur \mathcal{S}_N , qui est un sous espace de $L^2_{\mathbb{C}}(\mathbb{R}^N)$, et prend ses valeurs dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ (car $\mathcal{S}_N \subset L^2_{\mathbb{C}}(\mathbb{R}^N)$, en confondant, comme d'habitude, un élément de $L^2_{\mathbb{C}}(\mathbb{R}^N)$ avec l'un de ses représentants). Comme cette application est linéaire, continue pour la norme de $L^2_{\mathbb{C}}(\mathbb{R}^N)$, et que \mathcal{S}_N est dense dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ (car $\mathcal{S}_N \supset C_c^\infty(\mathbb{R}^N, \mathbb{C})$ et $C_c^\infty(\mathbb{R}^N, \mathbb{C})$ est dense dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$, voir le théorème 8.4), on en déduit que cette application se prolonge en une application, notée \bar{F} , de $L^2_{\mathbb{C}}(\mathbb{R}^N)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$.

Plus précisément, soit $f \in L^2_{\mathbb{C}}(\mathbb{R}^N)$. Il existe $(f_n)_{n \in \mathbb{N}} \subset \mathcal{S}_N$ t.q. $f_n \rightarrow f$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ quand $n \rightarrow \infty$. La suite $(f_n)_{n \in \mathbb{N}}$ est donc de Cauchy dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$. La proposition 10.9 donne alors que la suite $(\hat{f}_n)_{n \in \mathbb{N}}$ est donc aussi de Cauchy dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$. Elle converge donc dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$. On aimerait définir $\bar{F}(f)$ comme étant la limite (dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$) de la suite $(\hat{f}_n)_{n \in \mathbb{N}}$. Ceci est possible à condition que cette limite ne dépende que de f et pas du choix de la suite $(f_n)_{n \in \mathbb{N}}$ convergeant dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ vers f . Or, ce dernier point est facile car si $(g_n)_{n \in \mathbb{N}}$ est une autre suite convergeant dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ vers f , on a $\|\hat{f}_n - \hat{g}_n\|_2 = \|f_n - g_n\|_2 \rightarrow 0$ quand $n \rightarrow \infty$. On a ainsi défini \bar{F} de $L^2_{\mathbb{C}}(\mathbb{R}^N)$ dans lui même.

La linéarité de \bar{F} découle immédiatement du fait que l'application $f \mapsto \hat{f}$ est linéaire de \mathcal{S}_N dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$. Enfin, soit $f \in L^2_{\mathbb{C}}(\mathbb{R}^N)$ et $(f_n)_{n \in \mathbb{N}} \subset \mathcal{S}_N$ t.q. $f_n \rightarrow f$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$. La proposition 10.9 donne que $\|\hat{f}_n\|_2 = \|f_n\|_2$ pour tout $n \in \mathbb{N}$. En passant à la limite quand $n \rightarrow \infty$, on en déduit $\|\bar{F}(f)\|_2 = \|f\|_2$. Ce qui prouve la continuité de \bar{F} de $L^2_{\mathbb{C}}(\mathbb{R}^N)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$. On montre maintenant les 4 items du théorème.

1. Soit $f \in L^1_{\mathbb{C}}(\mathbb{R}^N) \cap L^2_{\mathbb{C}}(\mathbb{R}^N)$. En reprenant la démonstration du théorème 8.4, il est facile de voir qu'il existe une suite $(f_n)_{n \in \mathbb{N}} \subset C^{\infty}_c(\mathbb{R}^N, \mathbb{C})$ t.q. $f_n \rightarrow f$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ et $L^1_{\mathbb{C}}(\mathbb{R}^N)$ lorsque $n \rightarrow +\infty$ (car dans la démonstration du théorème 8.4, la suite construite pour converger vers f dans L^p ne dépend pas de p). On en déduit que $\hat{f}_n \rightarrow \hat{f}$ uniformément sur \mathbb{R}^N lorsque $n \rightarrow +\infty$ (car $f_n \rightarrow f$ dans $L^1_{\mathbb{C}}(\mathbb{R}^N)$) et que $\hat{f}_n \rightarrow \bar{F}(f)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ lorsque $n \rightarrow +\infty$ (car $f_n \rightarrow f$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$) et donc que, après extraction éventuelle d'une sous suite, on peut supposer que $\hat{f}_n \rightarrow \bar{F}(f)$ p.p. quand $n \rightarrow \infty$. On en déduit bien que $\hat{f} = \bar{F}(f)$ p.p..
2. Soit $f, g \in L^2_{\mathbb{C}}(\mathbb{R}^N)$. Il existe deux suites $(f_n)_{n \in \mathbb{N}} \subset \mathcal{S}_N$ et $(g_n)_{n \in \mathbb{N}} \subset \mathcal{S}_N$ t.q. $f_n \rightarrow f, g_n \rightarrow g$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$. La proposition 10.9 donne $(\hat{f}_n/\hat{g}_n)_2 = (f_n/g_n)_2$ pour tout $n \in \mathbb{N}$. En passant à la limite quand $n \rightarrow \infty$ on obtient bien $(\bar{F}(f)/\bar{F}(g))_2 = (f/g)_2$.
3. Soit $f \in L^2$. Soit $(f_n)_{n \in \mathbb{N}} \subset \mathcal{S}_N$ t.q. $f_n \rightarrow f$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ quand $n \rightarrow \infty$. On a donc $\hat{f}_n \rightarrow \bar{F}(f)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$, ce qui donne aussi $\hat{f}_n(\cdot) \rightarrow \bar{F}(f)(\cdot)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ et donc $\hat{f}_n(\cdot) \rightarrow \bar{F}(\bar{F}(f))(\cdot)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$ quand $n \rightarrow +\infty$. La proposition 10.6 donne $f_n = \hat{f}_n(\cdot)$ pour tout $n \in \mathbb{N}$. On en déduit donc (par unicité de la limite dans L^2) que $f = \bar{F}(\bar{F}(f))(\cdot)$ p.p..
4. L'injectivité de \bar{F} (de $L^2_{\mathbb{C}}(\mathbb{R}^N)$ dans $L^2_{\mathbb{C}}(\mathbb{R}^N)$) découle du fait que $\|\bar{F}(f)\|_2 = \|f\|_2$ et que \bar{F} est linéaire. La surjectivité est une conséquence immédiate du troisième item.

■

10.5 Résolution d'une EDO ou d'une EDP

On donne ici deux exemples simples d'utilisation de la transformation de Fourier pour la résolution d'une équation différentielle (souvent notée EDO pour Equation Différentielle Ordinaire) ou d'une équation aux dérivées partielles (souvent notée EDP pour Equation aux Dérivées Partielles).

Soit $N \geq 0, (a_0, \dots, a_N) \in \mathbb{R}^{N+1}$ et $g \in \mathcal{S}_1$ (donnés). On cherche $f \in \mathcal{S}_1$ qui vérifie :

$$a_N f^{(N)}(x) + \dots + a_1 f'(x) + a_0 f(x) = g(x), \text{ pour tout } x \in \mathbb{R}. \quad (10.5)$$

Si $f \in \mathcal{S}_1$ vérifie (10.5), elle vérifie nécessairement, par transformation de Fourier :

$$a_N \widehat{f^{(N)}}(t) + \dots + a_1 \widehat{f'}(t) + a_0 \widehat{f}(t) = \widehat{g}(t), \text{ pour tout } t \in \mathbb{R}. \quad (10.6)$$

c'est-à-dire :

$$a_N (it)^N \widehat{f}(t) + \dots + a_1 it \widehat{f}(t) + a_0 \widehat{f}(t) = \widehat{g}(t), \text{ pour tout } t \in \mathbb{R}. \quad (10.7)$$

En posant $p(t) = a_N(it)^N + \dots + a_1 it + a_0$ et en supposant que p ne s'annule pas sur \mathbb{R} , on a alors :

$$\hat{f}(t) = \frac{\hat{g}(t)}{p(t)}. \quad (10.8)$$

Comme $g \in \mathcal{S}_1$ et que p ne s'annule pas sur \mathbb{R} , on peut montrer que $\frac{\hat{g}}{p} \in \mathcal{S}_1$. En utilisant maintenant le théorème d'inversion, ou la proposition 10.6, on obtient :

$$f(t) = \hat{f}(-t) = \widehat{\left(\frac{\hat{g}}{p}\right)}(-t), \text{ pour tout } t \in \mathbb{R}. \quad (10.9)$$

On a donc montré que f est nécessairement donnée par (10.9). Réciproquement, il est facile de voir que la fonction donnée par (10.9) est solution de (10.5), c'est donc l'unique solution dans \mathcal{S}_1 de (10.5) (en supposant que p ne s'annule pas sur \mathbb{R}).

Soit $N \geq 1$, on cherche $u : \mathbb{R}^N \rightarrow \mathbb{R}$, de classe C^2 (c'est-à-dire deux fois continûment dérivable) t.q.

$$-\Delta u(x) = 0, \text{ pour tout } x \in \mathbb{R}^N. \quad (10.10)$$

Cherchons $u \in \mathcal{S}_N$ (et donc $\Delta u \in \mathcal{S}_N$) solution de (10.10). On a donc $\widehat{\Delta u} = 0$ (partout sur \mathbb{R}^N), c'est-à-dire $|t|^2 \hat{u}(t) = 0$ et donc $\hat{u}(t) = 0$, pour tout $t \neq 0$. Comme \hat{u} est continue, ceci entraîne que $\hat{u} = 0$ (partout sur \mathbb{R}^N), et donc $u = 0$. La fonction identiquement égale à 0 est donc la seule solution de (10.10) dans \mathcal{S}_N .

On peut effectuer un raisonnement analogue si on cherche u de classe C^2 et dans $L^2_{\mathbb{R}}(\mathbb{R}^N)$ (on peut aussi le faire si u est seulement dans $L^2_{\mathbb{R}}(\mathbb{R}^N)$, il faut alors convenablement définir Δu). On obtient encore que la seule solution de (10.10) est $u = 0$.

Par contre, ce résultat d'unicité n'est plus vrai si on ne demande pas à u d'être de carré intégrable. En effet, les fonctions constantes sont toutes solutions de (10.10) (et on peut montrer que ce sont les seules fonctions, de classe C^2 et bornées, solutions de (10.10)).

10.6 Fonction caractéristique d'un vecteur aléatoire

Définition 10.4 Soit (Ω, \mathcal{A}, P) un espace probabilisé, $d \geq 1$ et X un v.a. de dimension d . On appelle fonction caractéristique de X la fonction de \mathbb{R}^d dans \mathbb{C} définie par :

$$\varphi_X(u) = \int_{\Omega} e^{iX \cdot u} dP = E(e^{iX \cdot u}), \text{ pour } u \in \mathbb{R}^d.$$

Dans la définition précédente, la fonction $e^{iX \cdot u}$ est bien intégrable sur Ω (pour tout $u \in \mathbb{R}^d$) car la fonction $x \mapsto e^{ix \cdot u}$ est continue bornée de \mathbb{R}^d dans \mathbb{C} . En notant p_X la loi de X , on remarque également que $\varphi_X = (2\pi)^{d/2} \hat{p}_X(-\cdot)$. La section 10.3 donne alors quelques propriétés élémentaires de la fonction caractéristique d'un v.a..

Proposition 10.10 Soit (Ω, \mathcal{A}, P) un espace probabilisé, $d \geq 1$ et X un v.a. de dimension d . La fonction caractéristique de X vérifie alors les propriétés suivantes :

1. $\varphi_X \in C_b(\mathbb{R}^d, \mathbb{C})$.
2. La loi de X est entièrement déterminée par φ_X (c'est-à-dire que si Y est un autre v.a. de dimension d et que $\varphi_X = \varphi_Y$, on a nécessairement $p_X = p_Y$).

3. Si $\varphi_X \in L^1_{\mathbb{C}}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda_d)$, la loi de X a une densité par rapport à la mesure de Lebesgue sur $\mathcal{B}(\mathbb{R}^d)$, c'est-à-dire $p_X = f\lambda_d$, et :

$$f(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{-ix \cdot u} \varphi_X(u) du, \quad \lambda_d\text{-p.s. en } x \in \mathbb{R}^d.$$

DÉMONSTRATION : Comme $\varphi_X = (2\pi)^{d/2} \hat{p}_X(\cdot)$, le premier item est donnée par la proposition 10.7 (qui donne $\hat{p}_X \in C_b(\mathbb{R}^d, \mathbb{C})$).

Le deuxième item est donnée par le premier item de la proposition 10.8.

Pour le troisième item, on remarque que le deuxième item de la proposition 10.8 donne $p_X = f\lambda_d$ avec $f = \hat{p}_X(\cdot)$, ce qui donne λ_d -p.s. en $x \in \mathbb{R}^d$:

$$f(x) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{ix \cdot t} \hat{p}_X(t) dt = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{-ix \cdot t} \varphi_X(t) dt.$$

■

Les fonctions caractéristiques peuvent être utilisées pour montrer l'indépendance de v.a.r. (ou de v.a.). on donne un exemple dans la proposition suivante.

Proposition 10.11 Soit (Ω, \mathcal{A}, P) un espace probabilisé, $d > 1$ et X_1, \dots, X_d d v.a.r. Alors, les v.a.r X_1, \dots, X_d sont indépendantes si et seulement si on a $\varphi_X(u) = \prod_{j=1}^d \varphi_{X_j}(u_j)$ pour tout $u = (u_1, \dots, u_d)^t \in \mathbb{R}^d$, où X est la v.a. de composantes X_1, \dots, X_d .

DÉMONSTRATION : D'après le théorème 9.2 les v.a.r. X_1, \dots, X_d sont indépendantes si et seulement $p_X = p_{X_1} \otimes \dots \otimes p_{X_d}$. Par la proposition 10.8, ces deux mesures sont égales si et seulement si leurs transformées de Fourier sont égales, c'est-à-dire si et seulement si :

$$\varphi_X(u) = \int_{\mathbb{R}^d} e^{ix \cdot u} d(p_{X_1} \otimes \dots \otimes p_{X_d}) \text{ pour tout } u \in \mathbb{R}^d.$$

Comme $e^{ix \cdot u} = \prod_{j=1}^d e^{ix_j u_j}$ pour tout $x = (x_1, \dots, x_d)^t$ et tout $u = (u_1, \dots, u_d)^t$, la définition de la mesure produit donne

$$\int_{\mathbb{R}^d} e^{ix \cdot u} d(p_{X_1} \otimes \dots \otimes p_{X_d}) = \prod_{j=1}^d \varphi_{X_j}(u_j),$$

ce qui termine la démonstration de cette proposition. ■

Il est intéressant aussi de connaître le lien entre convergence en loi et convergence simple des fonctions caractéristiques. Nous donnons ce lien dans le deuxième item du théorème 10.3 (dû à Paul Lévy).

Théorème 10.3 (Convergence en loi et fonctions caractéristiques) Soit (Ω, \mathcal{A}, P) un espace probabilisé, $d \geq 1$ et $(X_n)_{n \in \mathbb{N}^*}$ une suite de v.a. de dimension d .

1. On suppose que la fonction caractéristique de X_n converge simplement (quand $n \rightarrow \infty$) vers une fonction φ continue en 0. Il existe alors un v.a. X t.q. $X_n \rightarrow X$ en loi, quand $n \rightarrow \infty$.
2. Soit X un v.a. de dimension d . Alors, $X_n \rightarrow X$ en loi, quand $n \rightarrow \infty$, si et seulement si $\varphi_{X_n}(u) \rightarrow \varphi_X(u)$, quand $n \rightarrow \infty$, pour tout $u \in \mathbb{R}^d$.

DÉMONSTRATION : Le deuxième item du théorème est une conséquence facile du premier item. En effet, si $X_n \rightarrow X$ en loi, on a, bien sûr, $\varphi_{X_n}(u) \rightarrow \varphi_X(u)$ pour tout $u \in \mathbb{R}^d$ (car la fonction $x \mapsto e^{ix \cdot u}$ est continue et bornée de \mathbb{R} dans \mathbb{C} , pour tout $u \in \mathbb{R}^d$). Réciproquement, on suppose que $\varphi_{X_n}(u) \rightarrow \varphi_X(u)$ pour tout $u \in \mathbb{R}^d$. Comme la fonction φ_X est continue (et donc, en particulier continue en 0), le premier item donne que X_n converge en loi vers un v.a. Y . Mais ceci donne que X et Y ont même fonction caractéristique et donc même loi. On en déduit bien que X_n tend vers X en loi.

Nous démontrons maintenant le premier item du théorème 10.3. Cette démonstration se fait en deux étapes. Dans la première étape on montre que la suite des lois des X_n est tendue (on utilise ici le lemme 10.1 et la continuité de φ). Puis dans la seconde étape on conclut grâce à la proposition 9.6.

Étape 1. On montre dans cette étape que la suite des lois des X_n est tendue. Pour cela, il suffit de montrer que la suite des lois de chaque composante des X_n est tendue. Soit donc $i \in \{1, \dots, d\}$ et $(X_n^{(i)})_{n \in \mathbb{N}}$ la suite des i -ème composantes des X_n . On note m_n la loi de $X_n^{(i)}$ et, pour tout $t \in \mathbb{R}$,

$$\psi_n(t) = \frac{1}{2} \varphi_{X_n^{(i)}}(t) + \varphi_{X_n^{(i)}}(-t) = \sqrt{2\pi} \frac{\hat{m}_n(t) + \hat{m}_n(-t)}{2}$$

Le lemme 10.1 donne alors, pour tout $a > 0$ et $n \in \mathbb{N}$,

$$m_n(\{x, |x| \geq a\}) \leq a \int_0^{2/a} (1 - \psi_n(t)) dt. \quad (10.11)$$

On a $\varphi_{X_n^{(i)}}(t) = \varphi_{X_n}(te_i)$ (ou e_i est le i -ème vecteur de la base canonique de \mathbb{R}^d), la fonction $\psi_n(t)$ converge donc pour tout $t \in \mathbb{R}$ et sa limite est une fonction ψ (de \mathbb{R} dans \mathbb{R}) continue en 0. Comme $\psi_n(0) = 1$ pour tout $n \in \mathbb{N}$, on a aussi $\psi(0) = 1$.

Soit $\varepsilon > 0$, comme ψ est continue en 0 et $\psi(0) = 1$, il existe $a_0 > 0$ t.q. $\max_{t \in [0, 2/a_0]} (1 - \psi(t)) \leq \varepsilon$, et donc

$$a_0 \int_0^{2/a_0} (1 - \psi(t)) dt \leq 2\varepsilon.$$

La fonction ψ_n converge simplement vers ψ et $0 \leq (1 - \psi_n) \leq 1$ le théorème de convergence dominée donne donc

$$\lim_{n \rightarrow \infty} a_0 \int_0^{2/a_0} (1 - \psi_n(t)) dt = a_0 \int_0^{2/a_0} (1 - \psi(t)) dt.$$

Il existe donc n_0 t.q.

$$n \geq n_0 \Rightarrow \lim_{n \rightarrow \infty} a_0 \int_0^{2/a_0} (1 - \psi_n(t)) dt \leq 3\varepsilon.$$

Avec (10.11) on a donc

$$n \geq n_0 \Rightarrow m_n(\{x, |x| \geq a_0\}) \leq 3\varepsilon.$$

D'autre part, en utilisant la continuité décroissante d'une mesure, pour tout $n \in \{1, \dots, n_0 - 1\}$, il existe a_n t.q. $m_n(\{x, |x| \geq a_n\}) \leq 3\varepsilon$. En prenant $a = \max\{a_0, a_1, \dots, a_{n_0-1}\}$ on a donc

$$m_n(\{x, |x| \geq a\}) \leq 3\varepsilon \text{ pour tout } n \in \mathbb{N}^*.$$

Ceci montre bien que la suite $(m_n)_{n \in \mathbb{N}^*}$ est tendue et termine la première étape.

Étape 2. Dans cette seconde étape il suffit d'utiliser la proposition 9.6. Cette proposition nous donne l'existence d'une sous suite de la suite $(X_n)_{n \in \mathbb{N}^*}$ et d'un v.a. X t.q. cette sous suite tend vers X en loi. On en déduit, en

particulier que $\varphi_X = \varphi$. Il reste à montrer que toute la suite $(X_n)_{n \in \mathbb{N}^*}$ tend vers X en loi (et pas seulement une sous suite). Pour le montrer, on raisonne par l'absurde. Si la suite $(X_n)_{n \in \mathbb{N}^*}$ ne converge pas en loi vers X , il existe $\phi \in C_b(\mathbb{R}^d, \mathbb{R})$ t.q. $E(\phi(X_n)) \not\rightarrow E(\phi(X))$. Il existe donc $\varepsilon > 0$ et une sous suite de la suite $(X_n)_{n \in \mathbb{N}^*}$, que nous noterons aussi $(X_n)_{n \in \mathbb{N}^*}$ (pour ne pas alourdir les notations), t.q.

$$|E(\phi(X_n)) - E(\phi(X))| \geq \varepsilon. \quad (10.12)$$

Mais, la proposition 9.6 nous donne l'existence d'une sous suite de cette sous suite et d'un v.a. Y t.q. la sous suite de la sous suite tend vers Y en loi. Comme la convergence en loi donne la convergence simple des fonctions caractéristiques (et que $\varphi_{X_n} \rightarrow \varphi$ simplement), on en déduit que $\varphi_Y = \varphi = \varphi_X$. On a donc $P_X = P_Y$. La sous suite de la sous suite tend donc vers X en loi. En contradiction avec (10.12). On a bien ainsi montré finalement que $X_n \rightarrow X$ en loi. ■

On donne maintenant la fonction caractéristique d'un vecteur gaussien.

Proposition 10.12 Soit (Ω, \mathcal{A}, P) un espace probabilisé.

1. Soit X une v.a.r. gaussienne. On note m son espérance (donc, $m = E(X) \in \mathbb{R}$) et σ^2 sa variance (donc, $\sigma^2 = E((X - m)^2) \geq 0$) de sorte que $X \sim \mathcal{N}(m, \sigma^2)$. On a alors :

$$\varphi_X(u) = e^{imu} e^{-\frac{\sigma^2}{2} u^2}, \text{ pour tout } u \in \mathbb{R}.$$

2. Soit $d \geq 1$ et X une v.a. gaussien de dimension d . On note m son espérance (donc, $m = E(X) \in \mathbb{R}^d$) et D sa matrice de covariance (donc, D est une matrice symétrique, semi-définie positive, son terme à la ligne j et la colonne k est donné par la covariance des composantes d'indices j et k de X) de sorte que $X \sim \mathcal{N}(m, D)$ (proposition 9.11). On a alors :

$$\varphi_X(u) = e^{im \cdot u} e^{-\frac{1}{2} D u \cdot u}, \text{ pour tout } u \in \mathbb{R}^d.$$

DÉMONSTRATION : Soit X une v.a.r. gaussienne, $X \sim \mathcal{N}(m, \sigma^2)$. On suppose tout d'abord que $\sigma^2 = 0$, on a alors $X = m$ p.s. et donc, pour tout $u \in \mathbb{R}$, $\varphi_X(u) = e^{imu}$, ce qui est bien la formule annoncée. On suppose maintenant que $\sigma > 0$, on a alors, pour tout $u \in \mathbb{R}$:

$$\varphi_X(u) = \int_{\mathbb{R}} e^{ixu} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) dx.$$

Avec le changement de variable $y = \frac{x-m}{\sigma}$, on obtient :

$$\varphi_X(u) = e^{imu} \int_{\mathbb{R}} e^{iy\sigma u} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy,$$

ce qui donne $\varphi_X(u) = e^{imu} \frac{1}{\sqrt{2\pi}} \psi(\sigma u)$ avec $\psi(t) = \int_{\mathbb{R}} e^{iyt} e^{-\frac{y^2}{2}} dy$ pour tout $t \in \mathbb{R}$. Comme la fonction $y \mapsto e^{y^2}$ est paire, on a aussi :

$$\psi(t) = \int_{\mathbb{R}} \cos(yt) e^{-\frac{y^2}{2}} dy, \text{ pour tout } t \in \mathbb{R}.$$

Pour calculer $\psi(t)$, on remarque que le théorème de dérivabilité sous le signe intégrale s'applique (théorème 4.10) et donne que ψ est de classe C^1 et

$$\psi'(t) = \int_{\mathbb{R}} (-y) \sin(yt) e^{-\frac{y^2}{2}} dy.$$

En intégrant par parties cette dernière intégrale (en fait, on intègre par parties sur $[-n, n]$ puis on fait tendre n vers l'infini), on obtient :

$$\psi'(t) = - \int_{\mathbb{R}} t \cos(yt) e^{-\frac{y^2}{2}} dy = -t\psi(t), \text{ pour tout } t \in \mathbb{R},$$

ce qui donne $\psi(t) = \psi(0)e^{-\frac{t^2}{2}}$. Comme $\psi(0) = \int_{\mathbb{R}} e^{-\frac{y^2}{2}} dy = \sqrt{2\pi}$, on en déduit que $\psi(t) = \sqrt{2\pi}e^{-\frac{t^2}{2}}$ (pour tout $t \in \mathbb{R}$) et donc que $\varphi_X(u) = e^{imu}e^{-\frac{\sigma^2 u^2}{2}}$, ce qui est bien la formule annoncée.

On suppose maintenant que $d > 1$ et que X est un v.a. gaussien. Soit $u \in \mathbb{R}^d$. On a $\varphi_X(u) = E(e^{iX \cdot u}) = \varphi_{X \cdot u}(1)$. Pour connaître $\varphi_X(u)$, il suffit donc de connaître la fonction caractéristique de la v.a.r. $X \cdot u$. Comme X est un v.a. gaussien, la v.a.r. $X \cdot u$ est une v.a.r. gaussienne. Sa moyenne est $E(X \cdot u) = E(X) \cdot u = m \cdot u$ et sa variance est (voir l'exercice 173) :

$$\sigma^2 = E((X \cdot u - m \cdot u)^2) = E(u^t(X - m)(X - m)^t u) = u^t \text{Cov}(X)u = u^t D u.$$

On a donc, d'après la première partie de cette démonstration (c'est-à-dire le cas $d = 1$),

$$\varphi_X(u) = \varphi_{X \cdot u}(1) = e^{im \cdot u} e^{-\frac{u^t D u}{2}},$$

ce qui est bien la formule annoncée (car $u^t D u = D u \cdot u$). ■

On montre enfin le théorème central limite.

Théorème 10.4 (Théorème central limite) Soit (Ω, \mathcal{A}, P) un espace probabilisé, $d \geq 1$ et $(X_n)_{n \in \mathbb{N}^*}$ une suite de v.a. de dimension d i.i.d.. On suppose que $E(|X_1|^2) < \infty$. On pose $E(X_1) = m$, $D = \text{Cov}(X_1)$ et, pour $n \in \mathbb{N}^*$,

$$Y_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i - m).$$

La suite $(Y_n)_{n \in \mathbb{N}^*}$ converge alors en loi vers tout v.a. Y dont la loi est la loi normale multidimensionnelle $\mathcal{N}(0, D)$. (Voir la définition 9.3 et la proposition 9.11 pour la définition de cette loi normale multidimensionnelle.)

DÉMONSTRATION : D'après le théorème 10.3 et la proposition 10.12, il suffit de montrer que, pour tout $u \in \mathbb{R}^d$,

$$\lim_{n \rightarrow \infty} \varphi_{Y_n}(u) = e^{-\frac{u^t D u}{2}},$$

où φ_{Y_n} est la fonction caractéristique du v.a. Y_n .

Soit $u \in \mathbb{R}^d$. On a

$$\varphi_{Y_n}(u) = E(e^{iY_n \cdot u}) = E(e^{\frac{i}{\sqrt{n}} \sum_{p=1}^n (X_p - m) \cdot u}) = E\left(\prod_{p=1}^n e^{\frac{i}{\sqrt{n}} (X_p - m) \cdot u}\right).$$

Les v.a. X_p ($p = 1, \dots, n$) étant indépendants et identiquement distribués, on a donc, par la proposition 9.9 (cette proposition est écrite pour des fonctions à valeurs réelles mais elle s'étend à des fonctions à valeurs complexes en décomposant les fonctions en parties réelles et imaginaires),

$$\varphi_{Y_n}(u) = E(e^{\frac{i}{\sqrt{n}} (X_1 - m) \cdot u})^n. \tag{10.13}$$

On pose $z_n = E(e^{\frac{i}{\sqrt{n}}(X_1-m) \cdot u})$ et on calcule maintenant z_n en faisant un développement limité des fonctions cos et sin. Il existe deux fonctions ε_1 et ε_2 appartenant à $C_b(\mathbb{R}, \mathbb{R})$ t.q. $\lim_{x \rightarrow 0} \varepsilon_i(x) = 0$, $i = 1, 2$ et

$$\cos(x) = 1 - \frac{x^2}{2} + x^2 \varepsilon_1(x) \text{ et } \sin(x) = x + x^2 \varepsilon_2(x).$$

On en déduit que

$$\int_{\Omega} \cos\left(\frac{1}{\sqrt{n}}(X_1 - m) \cdot u\right) dP = 1 - \frac{1}{2n} \int_{\Omega} ((X_1 - m) \cdot u)^2 dP + \frac{1}{n} \int_{\Omega} ((X_1 - m) \cdot u)^2 \varepsilon_1\left(\frac{1}{\sqrt{n}}(X_1 - m) \cdot u\right) dP,$$

et

$$\int_{\Omega} \sin\left(\frac{1}{\sqrt{n}}(X_1 - m) \cdot u\right) dP = \frac{1}{2n} \int_{\Omega} (X_1 - m) \cdot u dP + \frac{1}{n} \int_{\Omega} ((X_1 - m) \cdot u)^2 \varepsilon_2\left(\frac{1}{\sqrt{n}}(X_1 - m) \cdot u\right) dP.$$

Comme $E(|X_1|^2) < +\infty$, le théorème de convergence dominée donne

$$\lim_{n \rightarrow \infty} \int_{\Omega} ((X_1 - m) \cdot u)^2 \varepsilon_i\left(\frac{1}{\sqrt{n}}(X_1 - m) \cdot u\right) dP = 0 \text{ pour } i = 1, 2.$$

D'autre part, on a

$$\int_{\Omega} ((X_1 - m) \cdot u)^2 dP = \int_{\Omega} u^t (X_1 - m) (X_1 - m)^t u dP = u^t \left(\int_{\Omega} (X_1 - m) (X_1 - m)^t dP \right) u = u^t D u.$$

et

$$\int_{\Omega} (X_1 - m) \cdot u dP = \left(\int_{\Omega} (X_1 - m) dP \right) \cdot u = 0.$$

En posant $a = \frac{u^t D u}{2}$, on a donc

$$\int_{\Omega} \cos\left(\frac{1}{\sqrt{n}}(X_1 - m) \cdot u\right) dP = 1 - \frac{b_n}{n}, \quad \int_{\Omega} \sin\left(\frac{1}{\sqrt{n}}(X_1 - m) \cdot u\right) dP = \frac{c_n}{n},$$

avec $\lim_{n \rightarrow \infty} b_n = a$ et $\lim_{n \rightarrow \infty} c_n = 0$. En posant $a_n = b_n + i c_n$, on a donc $z_n = (1 - \frac{a_n}{n})$ avec $\lim_{n \rightarrow \infty} a_n = a$. Avec (10.13), ceci donne

$$\varphi_{Y_n}(u) = z_n^n = \left(1 - \frac{a_n}{n}\right)^n.$$

Comme $a \in \mathbb{R}$, le lemme 10.2 donne le résultat, c'est-à-dire $\lim_{n \rightarrow \infty} \varphi_{Y_n}(u) = e^{-a} = e^{-\frac{u^t D u}{2}}$. ■

Lemme 10.2 Soit $a \in \mathbb{R}$ et $(a_n)_{n \in \mathbb{N}}$ une suite de nombres complexes t.q. $\lim_{n \rightarrow \infty} a_n = a$. On a alors

$$\lim_{n \rightarrow \infty} \left(1 - \frac{a_n}{n}\right)^n = e^{-a}.$$

DÉMONSTRATION : Ce lemme est bien connu si $a_n \in \mathbb{R}$ pour tout $n \in \mathbb{N}$. Il suffit de remarquer que $a_n < n$ pour n assez grand et que $\lim_{n \rightarrow \infty} n \ln\left(1 - \frac{a_n}{n}\right) = -\lim_{n \rightarrow \infty} a_n = -a$. La difficulté est ici que a_n peut ne pas être un nombre réel.

On pose $z_n = 1 - \frac{a_n}{n}$ et $y_n = 1 - \frac{a}{n}$. Comme la suite $(a_n)_{n \in \mathbb{N}}$ est bornée (car elle est convergente), il existe $M \in \mathbb{R}$ t.q. $|a_n| \leq M$ pour tout $n \in \mathbb{N}$. On a donc $|z_n| \leq 1 + \frac{M}{n}$ (et aussi $|y_n| \leq 1 + \frac{M}{n}$). On a alors (en écrivant $\varphi(1) - \varphi(0) = \int_0^1 \varphi'(t) dt$ avec $\varphi \in C^\infty(\mathbb{R}, \mathbb{C})$ définie par $\varphi(t) = (tz_n + (1-t)y_n)^n$)

$$z_n^n - y_n^n = n(z_n - y_n) \int_0^1 (tz_n + (1-t)y_n)^{n-1} dt.$$

Comme $|tz_n + (1-t)y_n| \leq t|z_n| + (1-t)|y_n| \leq 1 + \frac{M}{n} \leq e^{\frac{M}{n}}$ (pour tout $n \in \mathbb{N}$), on en déduit que

$$|z_n^n - y_n^n| \leq n|z_n - y_n|e^M = |a - a_n|e^M.$$

On a donc $\lim_{n \rightarrow \infty} z_n^n = \lim_{n \rightarrow \infty} y_n^n = \lim_{n \rightarrow \infty} (1 - \frac{a}{n})^n = e^{-a}$. ■

10.7 Exercices

10.7.1 Transformation de Fourier dans L^1

Exercice 10.1 (Résultat partiel d'inversion de Fourier dans L^1) Corrigé 182 page 503

Soit $H(t) = e^{-|t|}$, $t \in \mathbb{R}$. On pose, pour $\lambda > 0$:

$$h_\lambda(x) = (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} H(\lambda t) e^{itx} dt, x \in \mathbb{R}. \quad (10.14)$$

1. Montrer que $h_\lambda(x) = (\frac{2}{\pi})^{\frac{1}{2}} \frac{\lambda}{\lambda^2 + x^2}$, et $\int_{\mathbb{R}} h_\lambda(x) dx = (2\pi)^{\frac{1}{2}}$.

2. Soit $f \in L^1_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, montrer que pour tout $x \in \mathbb{R}$, on a :

$$f \star h_\lambda(x) = \int_{\mathbb{R}} H(\lambda t) \hat{f}(t) e^{ixt} dt. \quad (10.15)$$

3. Soit g une fonction mesurable bornée de \mathbb{R} dans \mathbb{C} , continue en 0. Montrer que $g \star h_\lambda(0) \rightarrow \sqrt{2\pi}g(0)$ quand $\lambda \rightarrow 0$. [Utiliser 1. et le théorème de convergence dominée.]

4. Soit $f \in L^1_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, montrer que :

$$\|f \star h_\lambda - \sqrt{2\pi}f\|_1 \rightarrow 0 \text{ lorsque } \lambda \rightarrow 0. \quad (10.16)$$

[On pourra utiliser la continuité en moyenne et la question précédente avec $g(y) = \int |f(x-y) - f(x)| dx$.]

5. Dédurre de ce qui précède le théorème d'inversion de Fourier, théorème 10.1.

Exercice 10.2 1. Calculer la transformée de Fourier de $1_{[-a,a]}$, $a \in \mathbb{R}_+$. En déduire que L^1 n'est pas stable par transformation de Fourier.

2. On pose $g_n = 1_{[-n,n]}$. Calculer $f \star g_n$, et montrer qu'il existe $h_n \in L^1$ t.q. $\hat{h}_n = f \star g_n$. Montrer que la suite $f \star g_n$ est bornée dans $C_0(\mathbb{R}, \mathbb{R})$ alors que la suite h_n n'est pas bornée dans L^1 . En déduire que la transformée de Fourier n'est pas surjective de L^1 dans C_0 .

Exercice 10.3 Soit $f \in \mathcal{S}$, on pose $L(f)(x) = f''(x) + xf(x)$.

1. Montrer que $L(f) = 0$ entraîne $f = 0$.
2. Soit $k \in \mathcal{S}$, montrer que l'équation différentielle $h'(t) = k(t)$ a une solution dans \mathcal{S} si et seulement si $\int k(t)dt = 0$.
3. Soit $g \in \mathcal{S}$, étudier l'existence et l'unicité des solutions dans \mathcal{S} de l'équation $L(f) = g$. On pourra remarquer que pour $h \in \mathcal{S}$, on a :

$$i \frac{d}{dt}(he^{-i\frac{t^3}{3}}) - t^2(he^{-i\frac{t^3}{3}}) = ih'(t)e^{-i\frac{t^3}{3}}.$$

Exercice 10.4 On note L^p l'espace $L^p_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$.

1. Soient $f, g \in L^1$, montrer que $f\hat{g} \in L^1$, $g\hat{f} \in L^1$ et $\int f\hat{g}d\lambda = \int g\hat{f}d\lambda$.
2. Soit $B = [-\frac{1}{2}, \frac{1}{2}]$. Montrer que $1_B \star 1_B(t) = (1 - |t|)^+$.
3. On pose $\theta_n = (1 - \frac{|t|}{n})^+$, $n \in \mathbb{N}^*$. Dédurre de la question précédente que :

$$\hat{\theta}_n(y) = \frac{4}{\sqrt{2\pi}} \frac{\sin^2(\frac{ny}{2})}{ny^2}, \forall y \in \mathbb{R}.$$

[Se ramener à θ_1 ...]

4. Soit $f \in L^1 \cap L^\infty$ t.q. $\hat{f}(t) \in \mathbb{R}_+$ pour tout $t \in \mathbb{R}$. On se propose de montrer que $\hat{f} \in L^1$ (et donc que le théorème d'inversion s'applique)
 - (a) On note $\varphi_n = \theta_n \hat{f}$; montrer que $\varphi_n \uparrow \hat{f}$ et $\int \varphi_n d\lambda \uparrow \int \hat{f} d\lambda$ lorsque $n \rightarrow +\infty$.
 - (b) Montrer qu'il existe $\alpha \geq 0$ indépendant de n tel que $\int \hat{\theta}_n(y) dy = \alpha, \forall n \in \mathbb{N}^*$. En déduire que $\hat{f} \in L^1$.

10.7.2 Transformée de Fourier d'une mesure signée

Exercice 10.5 (Caractérisation de m par \hat{m}) *Corrigé 183 page 505*

Soit $d \geq 1$.

1. Soit m et μ deux mesures signées sur les boréliens de \mathbb{R}^d . On suppose que $\hat{m} = \hat{\mu}$.
 - (a) Soit $\varphi \in L^1_{\mathbb{C}}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda_d)$. Montrer que $\int \varphi dm = \int \varphi d\mu$.
 - (b) Montrer que $\int \varphi dm = \int \varphi d\mu$ pour tout $\varphi \in \mathcal{S}(\mathbb{R}^d, \mathbb{C})$ (et donc pour tout $\varphi \in C_c^\infty(\mathbb{R}^d, \mathbb{R})$).
 - (c) Montrer que $m = \mu$ (On rappelle qu'une fonction de $\varphi \in C_c(\mathbb{R}^d, \mathbb{R})$ est limite uniforme de fonctions de $\varphi \in C_c^\infty(\mathbb{R}^d, \mathbb{R})$).
2. Soit m une mesure signée sur les boréliens de \mathbb{R}^d . On suppose que $\hat{m} \in L^1_{\mathbb{C}}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda_d)$. Montrer que m est la mesure de densité f par rapport à la mesure de Lebesgue avec $f = \hat{\hat{m}}(\cdot)$.

10.7.3 Transformation de Fourier dans L^2

Exercice 10.6 On note ici \hat{f} la transformée de Fourier, pour $f \in L^1$ ou L^2 .

1. Soient $f, g \in \mathcal{S}$, Montrer que $\widehat{fg} = \frac{1}{\sqrt{2\pi}} \hat{f} \star \hat{g}$.
2. Soient $f, g \in L^2$, Montrer que $\widehat{fg} = \frac{1}{\sqrt{2\pi}} \hat{f} \star \hat{g}$.

10.7.4 Fonction Caractéristique d'une v.a.r.

Exercice 10.7 (Calcul de fonctions caractéristiques)

Soit (Ω, \mathcal{A}, P) un espace de probabilités et X une v.a. réelle. Calculer la fonction caractéristique φ_X de X dans les cas suivants :

1. $X = a$ p.s. ($a \in \mathbb{R}$).
2. $X \sim \mathcal{B}(p)$ (loi de Bernoulli de paramètre p : $\mathcal{P}[X = 1] = p = 1 - \mathcal{P}[X = 0]$).
3. X suit une loi exponentielle de paramètre λ (avec $\lambda > 0$).

Exercice 10.8 (Loi normale et vecteur gaussien)

Soit (Ω, \mathcal{A}, P) un espace probabilisé, $d \geq 1$ et X un v.a. de dimension d . Soit $m \in \mathbb{R}^d$ et D une matrice s.d.p. (de taille $d \times d$). Si $X \sim \mathcal{N}(m, D)$, où $\mathcal{N}(m, D)$ est définie par la définition 9.3, montrer que X est un vecteur gaussien.

Exercice 10.9 (Vecteurs gaussiens et densité)

Soit (Ω, \mathcal{A}, P) un espace de probabilités, $d \geq 1$ et $X = (X_1, \dots, X_d)$ un v.a. de dimension d .

On suppose que X est un vecteur gaussien (c'est-à-dire que $\sum_{i=1}^d a_i X_i$ suit une loi gaussienne pour tout $a_1, \dots, a_d \in \mathbb{R}$). On note m la moyenne de X et D la matrice de covariance de X . Montrer que la loi de X est de densité par rapport à la mesure de Lebesgue (sur \mathbb{R}^d) si et seulement si D est inversible. Si D est inversible, montrer que la loi de X est la loi $\mathcal{N}(m, D)$ donnée dans la définition 9.3.

Exercice 10.10 (V.a. gaussiens, indépendance, covariance) *Corrigé 184 page 506*

Soit (Ω, \mathcal{A}, P) un espace de probabilités, $d \geq 1$ et $X = (X_1, \dots, X_d)$ un v.a. de dimension d .

1. On suppose ici que $d = 2$.
 - (a) On suppose que X_1 et X_2 suivent des lois gaussiennes et sont indépendantes, montrer que X est un vecteur gaussien et que $\text{Cov}(X_1, X_2) = 0$.
 - (b) On suppose que X est un vecteur gaussien et que $\text{Cov}(X_1, X_2) = 0$. Montrer que X_1 et X_2 sont indépendantes.
2. On suppose toujours $d = 2$. Donner un exemple pour lequel X_1 et X_2 sont gaussiennes mais X n'est pas un vecteur gaussien. [On pourra, par exemple, choisir (Ω, \mathcal{A}, P) et X_1, X_2 de manière à avoir $\text{Cov}(X_1, X_2) = 0$ sans que X_1, X_2 soient indépendantes, voir l'exercice 4.44.]
3. On suppose que X est un vecteur gaussien et que les composantes de X sont indépendantes deux à deux. Montrer que X_1, \dots, X_d sont indépendantes.

Exercice 10.11 (Suite de v.a.r.i.i.d. de Poisson) *Corrigé 185 page 508*

Soit (Ω, \mathcal{A}, P) un espace de probabilité et X une v.a. de Poisson de paramètre λ ($\lambda > 0$). On rappelle que, $P[X = k] = e^{-\lambda} \frac{\lambda^k}{k!}$, pour tout $k \in \mathbb{N}$ et que $E[X] = \lambda$, $\text{Var}[X] = \lambda$.

1. Calculer la fonction caractéristique φ_X de X .
2. Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de v.a. indépendantes et de Poisson de paramètre λ .
 - (a) Soit $n > 1$. Dédurre de la question 1 la loi de la v.a. $Y_n = \sum_{p=1}^n X_p$.
 - (b) Utiliser le théorème central limite pour démontrer que

$$e^{-n} \sum_{k=0}^n \frac{n^k}{k!} \rightarrow \frac{1}{2} \text{ quand } n \rightarrow \infty.$$

Exercice 10.12 (Sur les lois des grands nombres) Corrigé 186 page 509

Soit (Ω, \mathcal{A}, P) un espace probabilisé et X une v.a.r. dont la loi est la loi de Cauchy c'est-à-dire que $P_X = f\lambda$, avec $f(x) = \frac{1}{\pi(1+x^2)}$ pour $x \in \mathbb{R}$ (on rappelle que λ désigne la mesure de Lebesgue sur les boréliens de \mathbb{R}).

1. Montrer que X n'est pas une v.a.r. intégrable.
2. Soit $n \in \mathbb{N}^*$, montrer que $P(\{|X| > n\}) \geq \frac{1}{n\pi}$, où $\{|X| > n\} = \{\omega \in \Omega, |X(\omega)| > n\}$. [On pourra remarquer que $\frac{2}{1+x^2} \geq \frac{1}{x^2}$ pour tout $x \geq 1$.]
3. Pour $x \in \mathbb{R}$, on pose $g(x) = e^{-|x|}$. Montrer que

$$\frac{2}{1+u^2} = \int_{\mathbb{R}} e^{iux} g(x) dx = \sqrt{2\pi} \hat{g}(u) \text{ pour tout } u \in \mathbb{R}.$$

En déduire que la fonction caractéristique de X , notée φ_X vérifie $\varphi_X(u) = e^{-|u|}$ pour tout $u \in \mathbb{R}$. [On pourra utiliser de théorème d'inversion de Fourier qui donne $\hat{g} = g(\cdot)$ p.p. si $g, \hat{g} \in L^1_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$.]

On se donne maintenant une suite $X_1, X_2, \dots, X_n, \dots$ de v.a.r.i.i.d. et on suppose que la loi de X_1 (et donc de tous les X_n) est la loi de Cauchy. Pour $n \in \mathbb{N}^*$, on pose :

$$Z_n = \frac{1}{n} \left(\sum_{p=1}^n X_p \right).$$

4. Soit $n \in \mathbb{N}^*$.

(a) Montrer que

$$\varphi_{Z_n}(u) = \varphi_{X_1}^n \left(\frac{u}{n} \right) \text{ pour tout } u \in \mathbb{R}.$$

(b) Donner la loi de Z_n .

5. Pour $n \in \mathbb{N}^*$, on pose $A_n = \{|X_n| > n\}$ et $B_n = \cup_{p \geq n} A_p$. on pose aussi $B = \cap_{n \in \mathbb{N}^*} B_n$.

(a) Soit $n \in \mathbb{N}^*$. En utilisant la question 2, montrer que

$$P(B_n^c) = \prod_{p \geq n} P(A_p^c) \leq \prod_{p \geq n} \left(1 - \frac{1}{p\pi} \right).$$

En déduire que $P(B_n^c) = 0$ et donc que $P(B_n) = 1$.

(b) Montrer que $P(B) = 1$. [Cette question est une conséquence de la question (a) mais elle peut aussi être faite directement en appliquant le lemme de Borel-Cantelli et la question 2.]

(c) Soit $\omega \in B$. Montrer que $\frac{X_n(\omega)}{n} \not\rightarrow 0$ quand $n \rightarrow \infty$.

(d) Soit $\omega \in \Omega$. Montrer que si la suite $(Z_n(\omega))_{n \in \mathbb{N}^*}$ converge vers une limite finie on a alors

$$\lim_{n \rightarrow \infty} \frac{X_n(\omega)}{n} = 0.$$

[Ecrire X_n en fonction de Z_n et Z_{n-1} .]

(e) Déduire des trois questions précédentes que la suite $(Z_n)_{n \in \mathbb{N}^*}$ ne converge pas p.s. vers une limite finie quand $n \rightarrow \infty$.

6. Pour tout $n \in \mathbb{N}^*$, montrer que

$$Z_{2n} - Z_n = \frac{U_n + V_n}{2},$$

où U_n et V_n sont deux v.a.r. indépendantes, de loi de Cauchy.

En déduire que la suite $(Z_n)_{n \in \mathbb{N}^*}$ ne converge pas en probabilité.

7. Pour quelles raisons ne peut-on appliquer, à la suite $(Z_n)_{n \in \mathbb{N}^*}$, les lois forte et faible des grands nombres ?

Exercice 10.13 (Sur la loi d'un vecteur aléatoire)

Soit (Ω, \mathcal{A}, P) un espace de probabilité et X un v.a. de dimension d . Montrer que la loi de X est uniquement déterminée par la donnée des lois de toutes les v.a.r. $a \cdot X$, $a \in \mathbb{R}^d$, $|a| = 1$. [On pourra utiliser la fonction caractéristique de X .]

Exercice 10.14 (Limite en loi de Gaussiennes)

Soit (Ω, \mathcal{A}, P) un espace de probabilité, X une v.a.r. et $(X_n)_{n \in \mathbb{N}}$ une suite de v.a.r. gaussiennes. On suppose que X_n tend en loi vers X , quand $n \rightarrow \infty$. On note m_n l'espérance de X_n et σ_n^2 la variance de X_n (avec $\sigma_n \geq 0$). Si $\sigma_n > 0$, la v.a.r. X_n a donc pour loi une loi dont la densité par rapport à la mesure de Lebesgue est

$$f_n(x) = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{(x-m_n)^2}{2\sigma_n^2}} \text{ pour } x \in \mathbb{R}.$$

Si $\sigma_n = 0$, la loi de X_n est δ_{m_n} .

Si Z est une v.a.r., on note φ_Z la fonction caractéristique de Z (on rappelle que φ_Z est une fonction continue de \mathbb{R} dans \mathbb{C} et que $\varphi_Z(0) = 1$). On rappelle que la fonction caractéristique de X_n est

$$\varphi_{X_n}(t) = e^{-\frac{\sigma_n^2}{2}t^2} e^{im_n t} \text{ pour } t \in \mathbb{R}.$$

1. Soit Z une v.a.r. Montrer qu'il existe $t \in \mathbb{R}^*$ t.q. $|\varphi_Z(t)| > 0$.
2. Montrer que la suite $(\sigma_n^2)_{n \in \mathbb{N}}$ est convergente dans \mathbb{R} . [On pourra utiliser la convergence de $|\varphi_{X_n}(t)|$ vers $|\varphi_X(t)|$ pour t bien choisi.]
3. Calculer $P(\{X_n \geq m_n\})$ et $P(\{X_n \leq m_n\})$ et en déduire que la suite $(m_n)_{n \in \mathbb{N}}$ est bornée. [On pourra utiliser le fait que la suite $(P_{X_n})_{n \in \mathbb{N}}$ est tendue, d'après la proposition 9.5.]
4. Montrer que la suite $(m_n)_{n \in \mathbb{N}}$ est convergente. En déduire que X est une v.a.r. gaussienne.

Exercice 10.15 (Caractérisation de $X = 0$ p.s.)

Soit (Ω, \mathcal{A}, p) un espace probabilisé. Si Z est une v.a.r., on note φ_Z la fonction caractéristique de Z , c'est-à-dire $\varphi_Z(t) = E(e^{iZt})$ pour tout $t \in \mathbb{R}$.

1. Soit X une variable aléatoire réelle et $a > 0$. On suppose que $\varphi_X(t) = 1$ pour tout $t \in [-a, a]$.

(a) Montrer que, pour tout $t \in [-a, a]$, $\int_{\Omega} [1 - \cos(tX(\omega))] dp(\omega) = 0$.

(b) Montrer que $X = 0$ p.s..

2. Soit X et Y deux variables aléatoires réelles indépendantes.

(a) Montrer que $\varphi_{X+Y} = \varphi_X \varphi_Y$.

(b) On suppose maintenant que $X + Y$ et Y ont même loi. Montrer que $X = 0$ p.s..

Exercice 10.16 (Indépendance et variables gaussiennes)

Soit (Ω, \mathcal{A}, p) un espace probabilisé et X, Y deux v.a.r.i.i.d. t.q. $E(X) = 0$ et $\text{Var}(X) = 1$.

On rappelle une caractérisation de l'indépendance avec les fonctions caractéristiques, donnée dans la proposition 10.11 : Soit Z_1, Z_2 deux v.a.r., on note Z le v.a. (de dimension 2) dont les composantes sont Z_1 et Z_2 . Les v.a.r. Z_1 et Z_2 sont indépendantes si et seulement si les fonctions caractéristiques de Z, Z_1 et Z_2 (notées φ_Z, φ_{Z_1} et φ_{Z_2}) vérifient $\varphi_Z(u_1, u_2) = \varphi_{Z_1}(u_1)\varphi_{Z_2}(u_2)$ pour tout $u_1, u_2 \in \mathbb{R}$.

1. On suppose dans cette question que X et Y sont des v.a.r. gaussiennes. Montrer que $X + Y$ et $X - Y$ sont des v.a.r. indépendantes. [On pourra utiliser le rappel.]

On suppose maintenant, pour toute la suite de l'exercice, que $X + Y$ et $X - Y$ sont indépendantes. Le but des questions suivantes est de démontrer que X et Y sont des variables gaussiennes. On note toujours φ_X la fonction caractéristique de X .

2. Montrer que $\lim_{t \rightarrow 0} \frac{1 - \varphi_X(t)}{t^2} = \frac{1}{2}$. [Utiliser $E(X) = 0, \text{Var}(X) = 1$ et un développement limité des fonctions cos et sin.]
3. Montrer que $\varphi_X(2t) = \varphi_X(t)^3 \varphi_X(-t)$ pour tout $t \in \mathbb{R}$ [Utiliser l'indépendance de $X + Y$ et $X - Y$]. Montrer que $\varphi_X(t) \neq 0$ pour tout $t \in \mathbb{R}$.
4. Soit $\rho(t) = \frac{\varphi_X(t)}{\varphi_X(-t)}$. Montrer que $\rho(t) = \rho(t/2^k)^{2^k}$ pour tout $t \in \mathbb{R}$ et tout $k \in \mathbb{N}$. Montrer que $\rho(t) = 1$ pour tout $t \in \mathbb{R}$ [utiliser la question 2]. Montrer que φ_X prend ses valeurs dans \mathbb{R}_+^* .
5. Montrer que $\varphi_X(t+s)\varphi_X(t-s) = \varphi_X(t)^2\varphi_X(s)^2$ pour tout $t, s \in \mathbb{R}$. En déduire que $\varphi_X(nt) = \varphi_X(t)^{n^2}$ pour tout $t \in \mathbb{R}$ et tout $n \in \mathbb{N}$.
6. Montrer que X (et donc aussi Y) est une v.a.r. gaussienne.

Exercice 10.17 (Un exemple...)

Soit (Ω, \mathcal{A}, P) un espace de probabilité et (X, Y) un couple de variables aléatoires réelles, à valeurs dans $[-1, 1]^2$. On suppose que la loi de ce couple a une densité f par rapport à la mesure de Lebesgue (sur $\mathcal{B}(\mathbb{R}^2)$) avec :

$$f(x, y) = \frac{1 + xy(x^2 - y^2)}{4} 1_{[-1, 1]^2}(x, y), \quad (x, y)^t \in \mathbb{R}^2.$$

1. Montrer que les lois des v.a. X et Y ont des densités par rapport à la mesure de Lebesgue (sur $\mathcal{B}(\mathbb{R})$). Calculer ces densités. X et Y sont-elles indépendantes ?
2. Calculer les espérances de X, Y et XY . Que vaut $\text{cov}(X, Y)$?
3. Calculer les fonctions caractéristiques de X, Y et $X + Y$.

Chapitre 11

Espérance conditionnelle et martingales

11.1 Espérance conditionnelle

Nous commençons par définir l'espérance conditionnée par une tribu.

Définition 11.1 Soit (Ω, \mathcal{A}, P) un espace probabilisé, X une v.a.r. intégrable et \mathcal{B} une tribu incluse dans \mathcal{A} . On appelle "Espérance, conditionnée par \mathcal{B} , de X " ou "Espérance conditionnelle de X par rapport à \mathcal{B} " l'ensemble des applications Z de Ω dans \mathbb{R} , \mathcal{B} -mesurable, intégrable et t.q. :

$$E(ZU) = E(XU) \text{ pour toute application } U \text{ de } \Omega \text{ dans } \mathbb{R}, \mathcal{B}\text{-mesurable, bornée.} \quad (11.1)$$

On note $E(X|\mathcal{B})$ cette espérance conditionnelle (c'est donc un ensemble de fonctions). (Noter que dans (11.1), les applications ZU et XU sont bien des v.a.r. intégrables).

Cette définition peut sembler un peu abrupte. On montrera dans la proposition 11.1 que, sous les hypothèses de la définition 11.1, l'espérance conditionnelle existe, c'est-à-dire que l'ensemble $E(X|\mathcal{B})$ est non vide, et que $E(X|\mathcal{B})$ est "unique", ceci signifiant que si $Z_1, Z_2 \in E(X|\mathcal{B})$, on a nécessairement $Z_1 = Z_2$ p.s..

L'ensemble $E(X|\mathcal{B})$, défini dans la définition 11.1, est une ensemble de v.a.r. (car Z \mathcal{B} -mesurable implique Z \mathcal{A} -mesurable) mais, en pratique, on confond cet ensemble avec l'un de ces éléments (comme on confond un élément de L^P avec l'un de ses représentants). Si Z est une v.a.r. \mathcal{B} -mesurable intégrable et t.q. $E(ZU) = E(XU)$ pour toute v.a.r. U \mathcal{B} -mesurable bornée, on écrira donc $Z = E(X|\mathcal{B})$ p.s. au lieu d'écrire $Z \in E(X|\mathcal{B})$.

Avant de démontrer l'existence et l'unicité de l'espérance conditionnelle, donnons quelques exemples simples. Soit (Ω, \mathcal{A}, P) un espace probabilisé et X une v.a.r. intégrable. Prenons tout d'abord $\mathcal{B} = \{\emptyset, \Omega\}$. Il est alors facile de voir (exercice 11.1) que $E(X|\mathcal{B})$ est réduit à un seul élément et que cet élément est la fonction constante et égale à $E(X)$.

Soit maintenant $A \in \mathcal{A}$ t.q. $0 < P(A) < 1$ et $\mathcal{B} = \{\emptyset, A, A^c, \Omega\}$ (qui est bien une tribu incluse dans \mathcal{A}). On peut ici montrer (exercice 11.1) que $E(X|\mathcal{B})$ est aussi réduit à un seul élément et cet élément est la fonction Z définie par :

$$Z = \frac{E(X1_A)}{P(A)} 1_A + \frac{E(X1_{A^c})}{P(A^c)} 1_{A^c}.$$

La quantité $\frac{E(X1_A)}{P(A)}$ s'appelle "espérance de X sachant A ". On a ainsi fait le lien entre "espérance de X sachant un événement" et "espérance de X par rapport à une tribu" (ou "selon une tribu").

Dans les deux exemples précédents, l'ensemble $E(X|\mathcal{B})$ était réduit à un seul élément. Voici maintenant un exemple où $E(X|\mathcal{B})$ n'est pas réduit à un seul élément. On prend $B \in \mathcal{A}$ t.q. $P(B) = 1$ et $B^c \neq \emptyset$ (c'est le cas, par exemple, si P est une mesure diffuse, que \mathcal{A} contient les singletons et que B^c est formé d'un nombre fini ou dénombrable de points de Ω). On prend encore $\mathcal{B} = \{\emptyset, B, B^c, \Omega\}$. Pour $a \in \mathbb{R}$, on pose $Z_a = E(X)1_B + a1_{B^c}$. On peut alors montrer (exercice 11.2) que $E(X|\mathcal{B}) = \{Z_a, a \in \mathbb{R}\}$. L'ensemble $E(X|\mathcal{B})$ n'est donc pas réduit à un élément.

On montre maintenant l'existence et l'unicité de l'espérance, conditionnée par une tribu, d'une v.a.r. intégrable.

Proposition 11.1 *Soit (Ω, \mathcal{A}, P) un espace probabilisé et \mathcal{B} une tribu incluse dans \mathcal{A} . Soit X une v.a.r. intégrable. Alors :*

1. (Existence) $E(X|\mathcal{B}) \neq \emptyset$.
2. (Unicité) $Z_1, Z_2 \in E(X|\mathcal{B}) \Rightarrow Z_1 = Z_2$ p.s..

DÉMONSTRATION : On démontre d'abord l'unicité de $E(X|\mathcal{B})$. Puis, on démontre l'existence de $E(X|\mathcal{B})$ si X est de carré intégrable, puis l'existence si X est positive (et intégrable) et enfin l'existence si X est seulement intégrable. En fait, la partie "existence si X est de carré intégrable" est inutile. Elle n'est pas utilisée pour la suite de la démonstration mais elle est éventuellement intéressante pour la compréhension de l'espérance conditionnelle.

Unicité. Soit $Z_1, Z_2 \in E(X|\mathcal{B})$. On pose $U = \text{sign}(Z_1 - Z_2)$ (on rappelle que la fonction sign est définie par $\text{sign}(s) = -1$ si $s < 0$, $\text{sign}(s) = 1$ si $s > 0$ et (par exemple) $\text{sign}(0) = 0$). Comme la fonction sign est borélienne de \mathbb{R} dans \mathbb{R} et que $(Z_1 - Z_2)$ est \mathcal{B} -mesurable, la fonction U est bien \mathcal{B} -mesurable. Elle est aussi bornée, on a donc en utilisant 11.1 avec $Z = Z_1$ et $Z = Z_2$, $E(XU) = E(Z_1U)$ et $E(XU) = E(Z_2U)$. Ceci donne $E((Z_1 - Z_2)U) = 0$ et donc $E(|Z_1 - Z_2|) = 0$. On en déduit $Z_1 = Z_2$ p.s..

Existence si X est de carré intégrable. On note $P_{\mathcal{B}}$ la restriction de P (qui est une mesure sur \mathcal{A}) à \mathcal{B} (tribu incluse dans \mathcal{A}). La mesure $P_{\mathcal{B}}$ est donc une probabilité sur \mathcal{B} . On note H l'espace de Hilbert $L^2_{\mathbb{R}}(\Omega, \mathcal{B}, P_{\mathcal{B}})$ et, pour $V \in H$, on pose :

$$T(V) = \int_{\Omega} XV dP.$$

Il est clair que $T(V)$ est bien définie. En étant précis, on remarque que $T(V) = \int_{\Omega} Xv dP$, où $v \in L^2_{\mathbb{R}}(\Omega, \mathcal{B}, P_{\mathcal{B}})$ est un représentant de V (et cette quantité ne dépend pas du représentant choisi). C'est pour définir T que nous avons besoin que X soit de carré intégrable.

L'application T est linéaire continue de H dans \mathbb{R} (et on a $\|T\| \leq \|X\|_2$). On peut donc appliquer le théorème de Riesz dans les espaces de Hilbert (théorème 6.9), il donne l'existence de $Z \in L^2(\Omega, \mathcal{B}, P_{\mathcal{B}})$ t.q. :

$$T(V) = \int_{\Omega} ZV dP \text{ pour tout } V \in H. \tag{11.2}$$

Comme $Z \in L^2(\Omega, \mathcal{B}, P_{\mathcal{B}})$, la fonction Z est bien \mathcal{B} -mesurable et intégrable (elle est même de carré intégrable). On montre maintenant que Z vérifie (11.1) (et donc que $Z \in E(X|\mathcal{B})$). Soit U une application \mathcal{B} -mesurable bornée de Ω dans \mathbb{R} . On a $U \in L^2(\Omega, \mathcal{B}, P_{\mathcal{B}})$, on peut donc utiliser (11.2) avec pour V la classe de U et on obtient

$$E(XU) = T(U) = \int_{\Omega} ZU dP = E(ZU).$$

L'application Z vérifie donc (11.1). ce qui prouve que $Z \in E(X|\mathcal{B})$.

Plus précisément, un développement du raisonnement ci avant (que les courageux peuvent faire) permet d'interpréter l'application $X \mapsto E(X|\mathcal{B})$ comme l'opérateur de projection orthogonale de $L^2(\Omega, \mathcal{A}, P)$ dans le sous espace vectoriel fermé formé à partir de $L^2(\Omega, \mathcal{B}, P_{\mathcal{B}})$.

Existence si X est positive et intégrable. On utilise ici le théorème de Radon-Nikodym (théorème 6.11, qui se démontre d'ailleurs avec le théorème de Riesz dans les espaces de Hilbert, théorème 6.9). On note toujours $p_{\mathcal{B}}$ la restriction de P à \mathcal{B} (de sorte que $P_{\mathcal{B}}$ est une probabilité sur \mathcal{B}).

Pour $B \in \mathcal{A}$, on pose $m(A) = \int_{\Omega} X 1_A dP$. On définit ainsi une mesure finie, m , sur \mathcal{A} , c'est la mesure de densité X par rapport à P . On note maintenant $m_{\mathcal{B}}$ la restriction de cette mesure à la tribu \mathcal{B} . La mesure $m_{\mathcal{B}}$ est absolument continue par rapport à la mesure $P_{\mathcal{B}}$ (car $B \in \mathcal{B}$, $P_{\mathcal{B}}(B) = 0$ implique que $P(B) = 0$ et donc $X 1_B = 0$ p.s. et donc $m(B) = 0$ et donc $m_{\mathcal{B}}(B) = 0$). Le théorème de Radon-Nikodym (théorème 6.11) donne alors l'existence de Z , \mathcal{B} -mesurable positive, t.q. $m_{\mathcal{B}} = Z P_{\mathcal{B}}$ (c'est-à-dire que $m_{\mathcal{B}}$ est la mesure sur \mathcal{B} de densité Z par rapport à $P_{\mathcal{B}}$).

La fonction Z est intégrable car $\int_{\Omega} Z dP = \int_{\Omega} Z dP_{\mathcal{B}} = m_{\mathcal{B}}(\Omega) = m(\Omega) = E(X) < +\infty$ (et $E(Z) = E(X)$). Il reste à montrer que Z vérifie (11.1) (ce qui donnera que $Z \in E(X|\mathcal{B})$). On remarque tout d'abord que $E(Z 1_B) = \int_{\Omega} Z 1_B dP = m_{\mathcal{B}}(B) = \int_{\Omega} X 1_B dP = E(X 1_B)$ pour tout $B \in \mathcal{B}$. Par linéarité positive, on a donc, pour toute fonction \mathcal{B} -étagée positive,

$$E(ZU) = \int_{\Omega} ZU dP = \int_{\Omega} ZU dP_{\mathcal{B}} = \int_{\Omega} U dm = \int_{\Omega} UX dP = E(XU). \quad (11.3)$$

Par convergence monotone, on en déduit alors que (11.3) est encore vraie pour toute fonctions \mathcal{B} -mesurable positive. Enfin, en utilisant $U = U^+ - U^-$ (et en remarquant que ZU et XU sont intégrables), on conclut que (11.3) est vraie pour toute fonction U \mathcal{B} -mesurable bornée de Ω dans \mathbb{R} . (On a ici repris un argument vu dans la remarque 6.22.) L'application Z vérifie donc (11.1). Ce qui prouve que $Z \in E(X|\mathcal{B})$.

Existence si X est seulement intégrable. Comme les fonctions X^+ et X^- sont positives et intégrables, il existe $Z_1 \in E(X^+|\mathcal{B})$ et $Z_2 \in E(X^-|\mathcal{B})$. On pose $Z = Z_1 - Z_2$. L'application Z est \mathcal{B} -mesurable et intégrable (car Z_1 et Z_2 le sont) et, pour tout fonction U \mathcal{B} -mesurable bornée, on a :

$$E(ZU) = E(Z_1U) - E(Z_2U) = E(X^+U) - E(X^-U) = E(XU).$$

L'application Z vérifie donc (11.1). Ce qui prouve que $Z \in E(X|\mathcal{B})$. ■

Soit (Ω, \mathcal{A}, P) un espace probabilisé et \mathcal{B} une tribu incluse dans \mathcal{A} . On a défini l'espérance, conditionnée par \mathcal{B} , d'une v.a.r. intégrable. On va maintenant montrer qu'on peut étendre la définition à des v.a.r. qui ne sont pas intégrables mais qui sont positives (la démonstration est déjà essentiellement dans la démonstration de la proposition 11.1). Pour cela, on va commencer par donner une "p.s.-caractérisation" de $E(X|\mathcal{B})$ lorsque X est une v.a.r. positive et intégrable. Cette caractérisation n'utilisant pas l'intégrabilité de X on aura ainsi une définition de $E(X|\mathcal{B})$ lorsque X est une v.a.r. positive. Ceci est fait dans la proposition 11.2 et la définition 11.2.

Proposition 11.2 Soit (Ω, \mathcal{A}, P) un espace probabilisé et \mathcal{B} une tribu incluse dans \mathcal{A} .

1. Soit X une v.a.r. intégrable positive. Alors, $Z \in E(X|\mathcal{B})$ si et seulement si Z est \mathcal{B} -mesurable, intégrable, ≥ 0 p.s. et t.q. :

$$E(ZU) = E(XU), \quad (11.4)$$

pour toute application U de Ω dans \mathbb{R} , \mathcal{B} -mesurable et positive.

2. Soit X une v.a.r. positive. On note $\bar{E}(X|\mathcal{B})$ l'ensemble des applications \mathcal{B} -mesurables positives vérifiant (11.4). On a alors :

(a) (Existence) $\bar{E}(X|\mathcal{B}) \neq \emptyset$.

(b) (Unicité) $Z_1, Z_2 \in \bar{E}(X|\mathcal{B}) \Rightarrow Z_1 = Z_2$ p.s..

DÉMONSTRATION : On commence par montrer le premier item. Si $Z \in E(X|\mathcal{B})$, la fonction Z est bien \mathcal{B} -mesurable intégrable et vérifie (11.1). Elle vérifie donc (11.4) en ajoutant “ U bornée”. Pour montrer que $Z \geq 0$ p.s., on prend $U = 1_B$ avec $B = \{Z < 0\}$ (U est bien \mathcal{B} -mesurable bornée). On obtient $E(ZU) = E(XU) \geq 0$. Comme $ZU \leq 0$, on a donc $ZU = 0$ p.s. et donc $Z \geq 0$ p.s.. Enfin, pour montrer que Z vérifie (11.4) (c’est-à-dire avec U \mathcal{B} -mesurable positive mais non nécessairement bornée), il suffit d’utiliser le théorème de convergence monotone (théorème 4.2) en introduisant $U_n = U1_{B_n}$ avec, pour $n \in \mathbb{N}$, $B_n = \{U \leq n\}$.

Réciproquement, si Z est \mathcal{B} -mesurable, intégrable, ≥ 0 p.s. et vérifie (11.4), il est facile de voir que Z vérifie (11.1). En effet, si U est \mathcal{B} -mesurable bornée, on utilise (11.4) avec les parties positive et négative de U pour obtenir (11.1). Donc, $Z \in E(X|\mathcal{B})$.

On montre maintenant le deuxième item de la proposition.

Existence. On reprend la démonstration de la proposition 11.1. On rappelle que $p_{\mathcal{B}}$ est la restriction de P à \mathcal{B} , $m = XP$ (c’est-à-dire la mesure de densité X par rapport à P) et $m_{\mathcal{B}}$ est la restriction de m à \mathcal{B} . La mesure $m_{\mathcal{B}}$ est absolument continue par rapport à la mesure $P_{\mathcal{B}}$. La mesure m n’est pas finie si X n’est pas intégrable. On ne peut donc pas appliquer directement le théorème 6.11 (qui demande que $m_{\mathcal{B}}$ soit finie). Pour résoudre cette petite difficulté, on pose, pour $n \in \mathbb{N}$ $m_n = X1_{\{n \leq X < n+1\}}$ (qui est une mesure sur \mathcal{A}) et $m_{n,\mathcal{B}}$ sa restriction \mathcal{B} . La mesure $m_{n,\mathcal{B}}$ est absolument continue par rapport à la mesure $P_{\mathcal{B}}$ et est finie. Le théorème de Radon-Nikodym donne alors l’existence de Z_n , \mathcal{B} -mesurable positive, t.q. $m_{n,\mathcal{B}} = Z_n P_{\mathcal{B}}$. On pose alors $Z = \sum_{n \in \mathbb{N}} Z_n$. La fonction Z est \mathcal{B} -mesurable positive, il reste à montrer que Z vérifie (11.4) (ce qui donnera que $Z \in \bar{E}(X|\mathcal{B})$). Soit U une application \mathcal{B} -mesurable positive de Ω dans \mathbb{R} . Pour tout $n \in \mathbb{N}$, on a

$$E(Z_n U) = \int_{\Omega} Z_n U dP = \int_{\Omega} Z_n U dP_{\mathcal{B}} = \int_{\Omega} U d m_{n,\mathcal{B}} = \int_{\Omega} U X 1_{\{n \leq X < n+1\}} dP.$$

En sommant cette dernière égalité pour $n \in \mathbb{N}$, on obtient (par le corollaire 4.1 sur les séries à termes positifs)

$$E(ZU) = E(XU).$$

L’application Z vérifie donc (11.4). Ce qui prouve que $Z \in \bar{E}(X|\mathcal{B})$.

Unicité. Soit $Z_1, Z_2 \in \bar{E}(X|\mathcal{B})$. prenons $U = (\text{sign}(Z_1 - Z_2))^+$ (qui est bien \mathcal{B} -mesurable et positive). On a donc, par (11.4), $E(Z_1 U) = E(Z_2 U) = E(XU)$, mais on ne peut rien en déduire car il est possible que $E(XU) = +\infty$. On va donc modifier légèrement U . Soit $n \in \mathbb{N}^*$. On pose $B_n = \{Z_1 \leq n\} \cap \{Z_2 \leq n\}$ et $U_n = U1_{B_n}$. La fonction U_n est encore \mathcal{B} -mesurable et positive et (11.4) donne $E(Z_1 U_n) = E(Z_2 U_n)$. Comme $0 \leq E(Z_1 U_n) = E(Z_2 U_n) \leq n$, on en déduit $E((Z_1 - Z_2)U_n) = 0$. Mais, $(Z_1 - Z_2)U_n \geq 0$. En faisant tendre n vers l’infini, le théorème de convergence monotone (théorème 4.1) donne $E((Z_1 - Z_2)U) = 0$, c’est-à-dire $E((Z_1 - Z_2)^+) = 0$ et donc $Z_1 \leq Z_2$ p.s.. En changeant les rôles de Z_1 et Z_2 on a aussi $Z_2 \leq Z_1$ p.s.. D’où $Z_1 = Z_2$ p.s.. ■

Définition 11.2 Soit (Ω, \mathcal{A}, P) un espace probabilisé, X une v.a.r. positive et \mathcal{B} une tribu incluse dans \mathcal{A} . On appelle “Espérance, conditionnée par \mathcal{B} , de X ” ou “Espérance conditionnelle de X par rapport à \mathcal{B} ” l’ensemble des applications Z de Ω dans \mathbb{R} , \mathcal{B} -mesurable, positive et t.q. :

$$E(ZU) = E(XU), \tag{11.5}$$

pour toute application U de Ω dans \mathbb{R} , \mathcal{B} -mesurable et positive.

On note $\bar{E}(X|\mathcal{B})$ cette espérance conditionnelle (c’est donc un ensemble de fonctions). (Noter que dans (11.1), les applications ZU et XU sont bien des v.a.r. positives, leur intégrale sur Ω est donc bien définie et appartient à $\bar{\mathbb{R}}_+$).

La proposition 11.2 nous donne l'existence et l'unicité (p.s.) de l'espérance conditionnelle lorsque X est une v.a.r. positive. Sous les hypothèses de la définition 11.2, si X est de plus intégrable, on a donc deux définitions de l'espérance conditionnelle de X par rapport à \mathcal{B} , notée $E(X|\mathcal{B})$ et $\bar{E}(X|\mathcal{B})$. La proposition 11.2 montre que $Z_1 \in E(X|\mathcal{B})$ et $Z_2 \in \bar{E}(X|\mathcal{B})$ implique $Z_1 = Z_2$ p.s.. En pratique, comme on confond $E(X|\mathcal{B})$ avec l'un de ces éléments et $\bar{E}(X|\mathcal{B})$ avec l'un de ces éléments, on a donc $E(X|\mathcal{B}) = \bar{E}(X|\mathcal{B})$ p.s.. Il est donc inutile de conserver la notation $\bar{E}(X|\mathcal{B})$ et on conservera la notation $E(X|\mathcal{B})$ dans les deux cas, c'est-à-dire " X v.a.r. intégrable" et " X v.a.r. positive".

Nous donnons maintenant quelques propriétés de l'espérance conditionnelle.

Proposition 11.3 Soit (Ω, \mathcal{A}, P) un espace probabilisé, \mathcal{B} une tribu incluse dans \mathcal{A} et X une v.a.r. Soit $p \in]1, \infty[$ et q le nombre conjugué de p (i.e. $q = p/(p-1)$ si $p < +\infty$ et $q = 1$ si $p = \infty$). On suppose que $X \in \mathcal{L}_{\mathbb{R}}^p(\Omega, \mathcal{A}, P)$. Soit $Z \in E(X|\mathcal{B})$. Alors, $Z \in \mathcal{L}_{\mathbb{R}}^p(\Omega, \mathcal{A}, P)$ et $E(ZU) = E(XU)$ pour toute application U (de Ω dans \mathbb{R}) \mathcal{B} -mesurable t.q. $|U|^q$ soit intégrable.

DÉMONSTRATION : La démonstration fait partie de l'exercice 11.6. En fait, le cas $p = 2$ a déjà été vu dans la démonstration de la proposition 11.1. ■

Proposition 11.4 (Inégalité de Jensen généralisée)

Soit (Ω, \mathcal{A}, P) un espace probabilisé, \mathcal{B} une tribu incluse dans \mathcal{A} et X une v.a.r. de carré intégrable. Soit φ une fonction convexe de \mathbb{R} dans \mathbb{R} . On suppose que $\varphi(X)$ est intégrable. On a alors $E(\varphi(X)|\mathcal{B}) \geq \varphi(E(X|\mathcal{B}))$ p.s..

DÉMONSTRATION : D'après le lemme 11.1, comme φ est convexe, il existe c , fonction croissante de \mathbb{R} dans \mathbb{R} (et donc fonction borélienne de \mathbb{R} dans \mathbb{R}) t.q., pour tout $x, a \in \mathbb{R}$, $\varphi(x) - \varphi(a) \geq c(a)(x - a)$.

Soit $Z \in E(X|\mathcal{B})$. On a donc pour tout $\omega \in \Omega$:

$$\varphi(X(\omega)) - \varphi(Z(\omega)) \geq c(Z(\omega))(X(\omega) - Z(\omega)). \quad (11.6)$$

On aimerait intégrer cette inégalité sur un élément (bien choisi) de \mathcal{B} mais cela n'est pas possible car les v.a.r. $\varphi(Z)$ et $c(Z)(X - Z)$ peuvent ne pas être intégrables (bien que Z et X soient intégrables). Pour $p \in \mathbb{N}^*$, on introduit donc $A_p = \{|Z| \leq p\}$ de sorte que les v.a.r. $1_{A_p}c(Z)(X - Z)$ et $1_{A_p}\varphi(Z)$ sont intégrables (noter que $c(Z)$ est bornée sur A_p car c est croissante). On pose aussi $A = \{E(\varphi(X)|\mathcal{B}) - \varphi(Z) < 0\}$ et $B_p = A_p \cap A$. Soit $p \in \mathbb{N}^*$, l'inégalité (11.6) donne $1_{B_p}(\varphi(X) - \varphi(Z)) \geq 1_{B_p}c(Z)(X - Z)$ et donc, en intégrant sur Ω :

$$\int_{B_p} (\varphi(X) - \varphi(Z))dP \geq \int_{B_p} c(Z)(X - Z)dP. \quad (11.7)$$

Comme Z et $E(\varphi(X)|\mathcal{B})$ sont \mathcal{B} -mesurables, on a $B_p \in \mathcal{B}$ (et donc 1_{B_p} est \mathcal{B} -mesurable). On a aussi $c(Z)$ \mathcal{B} -mesurable (car c est borélienne) et donc $1_{B_p}c(Z)$ \mathcal{B} -mesurable. On en déduit :

$$\int_{B_p} c(Z)(X - Z)dP = E(1_{B_p}c(Z)(X - Z)) = 0 \text{ (car } Z \in E(X|\mathcal{B})),$$

et

$$\int_{B_p} (\varphi(X) - \varphi(Z))dP = E(1_{B_p}(\varphi(X) - \varphi(Z))) = E(1_{B_p}(E(\varphi(X)|\mathcal{B}) - \varphi(Z))).$$

Avec (11.7), on en déduit :

$$\int_{B_p} (E(\varphi(X)|\mathcal{B}) - \varphi(Z))dP \geq 0.$$

Comme $E(\varphi(X)|\mathcal{B}) - \varphi(Z) < 0$ sur B_p (car $B_p \subset A$), on a donc $P(B_p) = 0$ et donc $P(A) = P(\cup_{p \in \mathbb{N}^*} B_p) = 0$. Ce qui donne bien $E(\varphi(X)|\mathcal{B}) \geq \varphi(Z)$ p.s.. ■

Lemme 11.1 Soit φ une fonction convexe de \mathbb{R} dans \mathbb{R} , il existe alors c , fonction croissante de \mathbb{R} dans \mathbb{R} (et donc fonction borélienne de \mathbb{R} dans \mathbb{R}) t.q., pour tout $x, a \in \mathbb{R}$, $\varphi(x) - \varphi(a) \geq c(a)(x - a)$.

DÉMONSTRATION : Si φ est dérivable sur \mathbb{R} , la fonction c existe et est unique, elle est donnée par $c = \varphi'$. L'existence de c est légèrement plus difficile si φ n'est pas dérivable sur tout \mathbb{R} (et on perd l'unicité de c).

Soit $a \in \mathbb{R}$, on considère le fonction $h_a : x \mapsto \frac{\varphi(x) - \varphi(a)}{x - a}$ qui est définie sur $\mathbb{R} \setminus \{a\}$. La convexité de φ permet de montrer que h_a est croissante (c'est-à-dire que $x, y \in \mathbb{R} \setminus \{a\}$, $x > y \Rightarrow h_a(x) \geq h_a(y)$). La fonction h_a a donc une limite à gauche (et à droite) en tout point, y compris au point a . On pose (par exemple) :

$$c(a) = \lim_{x \rightarrow a, x < a} h_a(x).$$

Il est facile de vérifier que la fonction c ainsi définie est croissante de \mathbb{R} dans \mathbb{R} et vérifie, pour tout $x, a \in \mathbb{R}$, $\varphi(x) - \varphi(a) \geq c(a)(x - a)$. ■

On définit maintenant l'espérance conditionnelle par rapport à une v.a.r. ou un v.a.

Définition 11.3 Soit (Ω, \mathcal{A}, P) un espace probabilisé et X une v.a.r. (ou un v.a. de dimension d , $d \geq 1$). Soit Y une v.a.r. intégrable ou une v.a.r. positive. On appelle "Espérance, conditionnée par X , de Y " ou "Espérance conditionnelle de Y par rapport à X " (ou "Espérance conditionnelle de Y sachant X ") l'ensemble $E(Y|\sigma(X))$, où $\sigma(X)$ est la tribu engendrée par X . On note $E(Y|X)$ cette espérance conditionnelle, de sorte que $E(Y|X) = E(Y|\sigma(X))$. (L'ensemble $E(Y|X)$ est donc un ensemble de v.a.r. et, comme d'habitude, on confond $E(Y|X)$ avec l'un de ces éléments.)

Pour caractériser $E(Y|X)$ (sous les hypothèses de la définition 11.3) et pour calculer cette espérance conditionnelle, on utilise, en général, le théorème 3.1 que nous rappelons sous une forme légèrement plus précise (donnée dans la démonstration du théorème 3.1).

Théorème 11.1 (Y mesurable par rapport à X) Soit (Ω, \mathcal{A}, P) un espace probabilisé et X, Y deux v.a.r.. On note $\sigma(X)$ la tribu engendrée par X . Alors :

- La v.a.r. Y est $\sigma(X)$ -mesurable si et seulement si il existe f , fonction borélienne de \mathbb{R} dans \mathbb{R} , t.q. $Y = f(X)$.
- La v.a.r. Y est $\sigma(X)$ -mesurable bornée si et seulement si il existe f , fonction borélienne bornée de \mathbb{R} dans \mathbb{R} , t.q. $Y = f(X)$.
- La v.a.r. Y est $\sigma(X)$ -mesurable positive si et seulement si il existe f , fonction borélienne positive de \mathbb{R} dans \mathbb{R} , t.q. $Y = f(X)$.

DÉMONSTRATION : La démonstration de ce théorème est donnée dans la démonstration du théorème 3.1. ■

Voici une conséquence immédiate de ce théorème, utilisée pour calculer $E(Y|X)$

Proposition 11.5 (Calcul de $E(Y|X)$) Soit (Ω, \mathcal{A}, P) un espace probabilisé et X, Y deux v.a.r.. Soit Z une application de Ω dans \mathbb{R} .

1. On suppose que Y est intégrable. Alors, $Z \in E(Y|X)$ si et seulement si il existe ψ application borélienne de \mathbb{R} dans \mathbb{R} t.q. $Z = \psi(X)$, $\psi(X)$ est intégrable et

$$E(\psi(X)\varphi(X)) = E(Y\varphi(X)), \tag{11.8}$$

pour toute application φ de \mathbb{R} dans \mathbb{R} , borélienne bornée.

2. On suppose que Y est positive. Alors, $Z \in E(Y|X)$ si et seulement si il existe ψ application borélienne positive de \mathbb{R} dans \mathbb{R} t.q. $Z = \psi(X)$ et

$$E(\psi(X)\varphi(X)) = E(Y\varphi(X)), \quad (11.9)$$

pour toute application φ de \mathbb{R} dans \mathbb{R} , borélienne positive.

DÉMONSTRATION : La démonstration est une conséquence facile du théorème 11.1. ■

La conséquence de la proposition 11.5 est que (sous les hypothèses de la proposition) l'on cherche $E(Y|X)$ sous la forme d'une fonction $\psi(X)$ (avec ψ de \mathbb{R} dans \mathbb{R}) vérifiant (11.8) (ou (11.9)). On raisonne, en général, par "condition nécessaire sur ψ " et, comme on sait que $E(X|Y)$ existe, il est même inutile de vérifier que la fonction $\psi(X)$ que l'on trouve (qui est, en général, définie p.s.) est bien intégrable (ou positive).

La proposition 11.5 montre également que (sous les hypothèses de la proposition 11.5) la fonction Y est une fonction de X (c'est-à-dire $Y = \psi(X)$ pour un certain ψ de \mathbb{R} dans \mathbb{R}) si et seulement si $E(Y|X) = Y$. Pour montrer que la v.a.r. Y est une fonction d'une autre v.a.r. X , il suffit donc de montrer que $E(Y|X) = Y$. En comparaison, le calcul de la covariance entre X et Y (après "normalisation") s'intéresse seulement à l'existence ou non d'une dépendance affine de Y en fonction de X . Voir, à ce propos, l'exercice 11.14.

Remarque 11.1 (Deux propriétés de l'espérance conditionnelle) Voici deux propriétés qui nous seront utiles dans la section suivante. Soit (Ω, \mathcal{A}, P) un espace probabilisé, \mathcal{B} une sous tribu de \mathcal{A} et X une v.a.r. intégrable.

1. Soit V une v.a.r. \mathcal{B} -mesurable bornée. On a alors $E(XV|\mathcal{B}) = VE(X|\mathcal{B})$ p.s.. (Voir l'exercice 11.4.)
2. On suppose que $\sigma(X)$ et \mathcal{B} sont des tribus indépendantes. On a alors $E[X|\mathcal{B}] = E[X]$ p.s.. En particulier, si Y est une v.a.r. indépendante de X (c'est-à-dire que $\sigma(X)$ et $\sigma(Y)$ sont des tribus indépendantes), on a alors $E[X|Y] = E[X]$ p.s.. (Voir l'exercice 11.5.)

11.2 Martingales

Définition 11.4 (Filtration et processus) Soit (Ω, \mathcal{A}, P) un espace probabilisé

1. On appelle "filtration" une suite de tribus $(\mathcal{B}_n)_{n \in \mathbb{N}}$ t.q. $\mathcal{B}_n \subset \mathcal{B}_{n+1} \subset \mathcal{A}$, pour tout $n \in \mathbb{N}$.
2. On appelle "processus réel" une suite de v.a.r..
3. Soit $(\mathcal{B}_n)_{n \in \mathbb{N}}$ une filtration et $(X_n)_{n \in \mathbb{N}}$ un processus réel. On dit que $(X_n)_{n \in \mathbb{N}}$ est adapté à la filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$ si, pour tout $n \in \mathbb{N}$, X_n est \mathcal{B}_n -mesurable.

Définition 11.5 (Martingale) Soit (Ω, \mathcal{A}, P) un espace probabilisé, $(\mathcal{B}_n)_{n \in \mathbb{N}}$ une filtration et $(X_n)_{n \in \mathbb{N}}$ un processus réel (c'est-à-dire une suite de v.a.r.).

1. (Martingale) La suite $(X_n)_{n \in \mathbb{N}}$ est une martingale par rapport à la filtration $(\mathcal{B}_n)_n$ si on a, pour tout $n \in \mathbb{N}$,
 - (a) X_n est \mathcal{B}_n -mesurable et intégrable,
 - (b) $E(X_{n+1}|\mathcal{B}_n) = X_n$ p.s..
2. (Sous et sur martingale) la suite $(X_n)_{n \in \mathbb{N}}$ est une sous-martingale [resp. sur-martingale] par rapport à la filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$ si on a, pour tout $n \in \mathbb{N}$,
 - (a) X_n est \mathcal{B}_n -mesurable et intégrable,
 - (b) $E(X_{n+1}|\mathcal{B}_n) \geq X_n$ p.s. [resp. $E(X_{n+1}|\mathcal{B}_n) \leq X_n$ p.s.].

Remarque 11.2 Soit (Ω, \mathcal{A}, P) un espace probabilisé, $(\mathcal{B}_n)_{n \in \mathbb{N}}$ une filtration et $(X_n)_{n \in \mathbb{N}}$ une martingale par rapport à la filtration $(\mathcal{B}_n)_n$. Pour tout $n \in \mathbb{N}$, on a, comme 1_Ω est \mathcal{B}_n -intégrable bornée, et $E(X_{n+1}|\mathcal{B}_n) = X_n$ p.s.,

$$E(X_{n+1}) = \int_{\Omega} X_{n+1} 1_{\Omega} dP = \int_{\Omega} E(X_{n+1}|\mathcal{B}_n) 1_{\Omega} dP = \int_{\Omega} X_n 1_{\Omega} dP = E(X_n).$$

On a donc $E(X_n) = E(X_0)$ pour tout $n \in \mathbb{N}$.

Exemple 11.1 (Exemples de Martingales) Soit (Ω, \mathcal{A}, P) un espace probabilisé.

1. Soit $(\mathcal{B}_n)_{n \in \mathbb{N}}$ une filtration et X une v.a.r. intégrable. On pose $X_n = E(X|\mathcal{B}_n)$. La suite $(X_n)_{n \in \mathbb{N}}$ est une martingale (par rapport à la filtration $(\mathcal{B}_n)_n$). (Voir l'exercice 11.21.)
2. Soit X_0 une v.a.r. intégrable et $(J_n)_{n \in \mathbb{N}^*}$ une suite de v.a.r. intégrable et de moyenne nulle. On suppose que la suite formée de X_0 et $(J_n)_{n \in \mathbb{N}^*}$ est une suite de v.a.r. indépendantes. On pose alors, pour $n \in \mathbb{N}$, $X_{n+1} = X_n + J_{n+1}$ et \mathcal{B}_n la tribu engendrée par X_0, \dots, X_n . La suite $(X_n)_{n \in \mathbb{N}}$ est une martingale (par rapport à la filtration $(\mathcal{B}_n)_n$). (Voir l'exercice 11.22.)

Définition 11.6 (Temps d'arrêt) Soit (Ω, \mathcal{A}, P) un espace probabilisé, $(\mathcal{B}_n)_{n \in \mathbb{N}}$ une filtration et T une v.a. à valeurs dans $\mathbb{N} \cup \{+\infty\}$ (c'est-à-dire une application mesurable de Ω , muni de la tribu \mathcal{A} , dans $\mathbb{N} \cup \{+\infty\}$, muni de la tribu formée de l'ensemble des parties de $\mathbb{N} \cup \{+\infty\}$). L'application T s'appelle un temps d'arrêt si, pour tout $n \in \mathbb{N}$, $\{T = n\} \in \mathcal{B}_n$. Si T est un temps d'arrêt, on note \mathcal{B}_T la tribu définie par $\mathcal{B}_T = \{A \in \mathcal{A} \text{ t.q., pour tout } n \in \mathbb{N}, A \cup \{T = n\} \in \mathcal{B}_n\}$.

Le théorème suivant montre qu'une martingale "arrêtée" est encore une martingale.

Théorème 11.2 (Martingale arrêtée à un temps d'arrêt) Soit (Ω, \mathcal{A}, P) un espace de probabilité muni d'une filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$ et $(X_n)_{n \in \mathbb{N}}$ une martingale (par rapport à la filtration $(\mathcal{B}_n)_n$). Soit ν un temps d'arrêt. Pour $n \in \mathbb{N}$, on pose $Y_n = X_{\nu \wedge n}$ (On rappelle que $\nu \wedge n(\omega) = \min\{\nu(\omega), n\}$ et donc que $X_{\nu \wedge n}(\omega) = X_{\min\{\nu(\omega), n\}}(\omega)$, pour tout $\omega \in \Omega$.) Alors, la suite $(Y_n)_{n \in \mathbb{N}}$ est encore une martingale (par rapport à la filtration $(\mathcal{B}_n)_n$).

DÉMONSTRATION : Soit $n \in \mathbb{N}$, on a $Y_n = X_n 1_{\{T > n\}} + \sum_{k=0}^n X_k 1_{\{T=k\}}$. On en déduit tout d'abord que Y_n est intégrable (comme somme finie de fonctions intégrables). Puis, on montre que Y_n est \mathcal{B}_n -mesurable. Pour cela, on remarque que $\{T = k\} \in \mathcal{B}_k \subset \mathcal{B}_n$, pour $k \leq n$, et que

$$\{T > n\} = \left(\cup_{k=0}^n \{T = k\} \right)^c \in \mathcal{B}_n.$$

Enfin, on remarque que X_k est \mathcal{B}_n -mesurable pour tout $k \leq n$. Grâce à la stabilité des fonctions mesurables par somme et produit, on obtient bien, finalement, que Y_n est \mathcal{B}_n -mesurable.

Il reste maintenant à montrer que $E(Y_{n+1}|\mathcal{B}_n) = Y_n$ p.s., pour tout $n \in \mathbb{N}$. Soit $n \in \mathbb{N}$, on a

$$Y_{n+1} = X_{n+1} 1_{\{T > n+1\}} + \sum_{k=0}^{n+1} X_k 1_{\{T=k\}} = X_{n+1} 1_{\{T \geq n+1\}} + \sum_{k=0}^n X_k 1_{\{T=k\}}.$$

Par linéarité de l'espérance conditionnelle, on a donc

$$E(Y_{n+1}|\mathcal{B}_n) = E(X_{n+1} 1_{\{T \geq n+1\}}|\mathcal{B}_n) + \sum_{k=0}^n E(X_k 1_{\{T=k\}}|\mathcal{B}_n).$$

Comme $\{T \geq n+1\} = (\cup_{k=0}^n \{T = k\})^c \in \mathcal{B}_n$, la remarque 11.1 (et le fait que $E(X_{n+1}|\mathcal{B}_n) = X_n$ p.s.) donne $E(X_{n+1}1_{\{T \geq n+1\}}|\mathcal{B}_n) = 1_{\{T \geq n+1\}}E(X_{n+1}|\mathcal{B}_n) = X_n 1_{\{T \geq n+1\}}$. Puis comme, pour $k \in \{0, \dots, n\}$, $X_k 1_{\{T=k\}}$ est \mathcal{B}_n -mesurable, on a $E(X_k 1_{\{T=k\}}|\mathcal{B}_n) = X_k 1_{\{T=k\}}$. On obtient ainsi

$$E(Y_{n+1}|\mathcal{B}_n) = X_n 1_{\{T \geq n+1\}} + \sum_{k=0}^n X_k 1_{\{T=k\}} = Y_n \text{ p.s.}$$

Ce qui termine la démonstration. ■

On conclut cette section par un théorème, sans démonstration, sur la convergence des martingales.

Théorème 11.3 (Convergence p.s. d'une martingale)

Soit (Ω, \mathcal{A}, P) un espace probabilisé, $(\mathcal{B}_n)_{n \in \mathbb{N}}$ une filtration et $(X_n)_{n \in \mathbb{N}}$ une suite de v.a.r. On suppose que $(X_n)_{n \in \mathbb{N}}$ est une martingale par rapport à $(\mathcal{B}_n)_{n \in \mathbb{N}}$.

1. On suppose que la suite $(X_n)_{n \in \mathbb{N}}$ est bornée dans $L^1_{\mathbb{R}}(\Omega, \mathcal{A}, P)$. Alors il existe une v.a.r. intégrable, X , t.q. $X_n \rightarrow X$ p.s., quand $n \rightarrow \infty$.
2. On suppose $X_n \geq 0$ pour tout $n \in \mathbb{N}$. Alors, il existe une v.a.r. intégrable, X , t.q. $X_n \rightarrow X$ p.s., quand $n \rightarrow \infty$.

On peut noter que le deuxième item du théorème 11.3 est une conséquence du premier car, pour une martingale, on a toujours $E(X_n) = E(X_0)$ pour tout $n \in \mathbb{N}$ (et les X_n sont toujours intégrales). Si $X_n \geq 0$, on a donc $\|X_n\|_1 = E(X_0) < +\infty$.

11.3 Exercices

11.3.1 Espérance conditionnelle

Exercice 11.1 (Espérance conditionnelle selon une tribu) Corrigé 187 page 511

Soit (Ω, \mathcal{A}, P) un espace probabilisé et Y une variable aléatoire réelle intégrable. Dans les trois cas suivants, montrer que $E[Y|\mathcal{B}]$ est réduit à un élément et déterminer $E[Y|\mathcal{B}]$ (en fonction de Y et \mathcal{B}).

1. La tribu \mathcal{B} la tribu grossière, c'est-à-dire $\mathcal{B} = \{\emptyset, \Omega\}$.
2. Soit $B \in \mathcal{A}$ t.q. $0 < P[B] < 1$. On prend pour \mathcal{B} la tribu engendrée par B .
3. Soit $(B_n)_{n \in \mathbb{N}^*} \subset \mathcal{A}$ t.q. $B_n \cap B_m = \emptyset$ si $n \neq m$, $\Omega = \cup_{n \in \mathbb{N}^*} B_n$ et $0 < P(B_n) < 1$ pour tout $n \in \mathbb{N}^*$. On prend pour \mathcal{B} la tribu engendrée par $(B_n)_{n \in \mathbb{N}^*}$ (c'est-à-dire $\mathcal{B} = \{\cup_{n \in J} B_n, J \subset \mathbb{N}^*\}$).

Exercice 11.2 (Espérance conditionnelle selon une tribu (2))

Soit (Ω, \mathcal{A}, P) un espace probabilisé et X une v.a.r. intégrable. Soit $B \in \mathcal{A}$ t.q. $P(B) = 1$ et $B^c \neq \emptyset$ (c'est le cas, par exemple, si P est une mesure diffuse, que \mathcal{A} contient les singletons et que B^c est formé d'un nombre fini ou dénombrable de points de Ω). On pose $\mathcal{B} = \{\emptyset, B, B^c, \Omega\}$. Pour $a \in \mathbb{R}$, on pose $Z_a = E(X)1_B + a1_{B^c}$. Montrer que $E(X|\mathcal{B}) = \{Z_a, a \in \mathbb{R}\}$.

Exercice 11.3 (Espérance conditionnelle selon une v.a.r.) Corrigé 188 page 512

Soit (Ω, \mathcal{A}, P) un espace probabilisé, X une variable aléatoire réelle et Y une variable aléatoire réelle intégrable.

1. On suppose qu'il existe $a \in \mathbb{R}$ t.q. $X = a$ p.s.. Donner un élément de $E[Y|X]$.
2. On suppose que X prend p.s. deux valeurs x_1 ou x_2 avec $x_1 \neq x_2$. Donner un élément de $E[Y|X]$.

3. On suppose que X est une v.a. prenant p.s. ses valeurs dans un ensemble dénombrable $\{x_n, n \in \mathbb{N}^*\}$ avec $P(X = x_n) \neq 0$ pour tout $n \in \mathbb{N}^*$. Donner un élément de $E[Y|X]$.

Exercice 11.4 (Egalité d'espérances conditionnelles)

Soit (Ω, \mathcal{A}, P) un espace probabilisé, \mathcal{B} une sous-tribu de \mathcal{A} , X une v.a.r. intégrable et V une v.a.r. \mathcal{B} -mesurable bornée. Montrer que $E[XV|\mathcal{B}] = E[X|\mathcal{B}]V$ p.s..

Exercice 11.5 (Espérance conditionnelle et indépendance)

Soit (Ω, \mathcal{A}, P) un espace probabilisé, et X une v.a.r. intégrable.

1. Soit Y une v.a.r. indépendante de X . Montrer que $E[X|Y] = E[X]$ p.s..
2. Soit \mathcal{B} une sous tribu de \mathcal{A} . On suppose que $\sigma(X)$ et \mathcal{B} sont des tribus indépendantes. Montrer que $E[X|\mathcal{B}] = E[X]$ p.s..

Exercice 11.6 (Espérance conditionnelle d'une v.a.r. appartenant à \mathcal{L}^p)

Soit (Ω, \mathcal{A}, P) un espace de probabilités, \mathcal{G} une sous-tribu de \mathcal{A} et Y une v.a.r. intégrable. On pose $Z = E[Y|\mathcal{G}]$ (plus précisément, on confond ici, comme d'habitude, la classe $E[Y|\mathcal{G}]$ avec l'un de ses éléments).

1. On suppose, dans cette question, qu'il existe $M \in \mathbb{R}$ t.q. $|Y| \leq M$ p.s.. Montrer que $|Z| \leq M$ p.s.. et que $E[Z\Phi] = E[Y\Phi]$ pour tout $\Phi \mathcal{G}$ -mesurable et intégrable.
2. Soit $p \in]1, \infty[$ et $q = \frac{p}{p-1}$. On suppose que $|Y|^p$ est intégrable, montrer que $|Z|^p$ est intégrable et que $E[Z\Phi] = E[Y\Phi]$ pour tout $\Phi \mathcal{G}$ -mesurable et t.q. $|\Phi|^q$ soit intégrable.

Exercice 11.7 (Calcul de $E(\exp(XY)|X)$ si $Y \sim \mathcal{N}(0, 1)$) Corrigé 189 page 514

Soit (Ω, \mathcal{A}, P) un espace de probabilités et X, Y deux v.a. réelles indépendantes. On suppose que Y suit une loi gaussienne centrée réduite et que $E[\exp(X^2/2)] < \infty$. Montrer que $\exp(XY)$ est intégrable et déterminer $E[\exp(XY)|X]$.

Exercice 11.8 (Espérance selon une somme de v.a.r.i.i.d)

Soit (Ω, \mathcal{A}, P) un espace de probabilités.

1. Soit X et Y deux v.a.r. intégrables. Montrer que $E(X|X+Y) + E(Y|X+Y) = X+Y$. On suppose maintenant que X et Y sont indépendantes et de même loi. Montrer que

$$E(X|X+Y) = E(Y|X+Y) = \frac{X+Y}{2}.$$

2. Soit $n \in \mathbb{N}^*$ et X_1, \dots, X_n des v.a.r. indépendantes, de même loi et intégrables. On note $S_n = \sum_{k=1}^n X_k$. Montrer que $E(X_1|S_n) = S_n/n$.

Exercice 11.9 (Une condition nécessaire et suffisante pour avoir $X = Y$ p.s.)

Soit (Ω, \mathcal{A}, P) un espace de probabilités et X, Y deux v.a.r. intégrables. On suppose que $E[X|Y] = Y$ p.s. et $E[Y|X] = X$ p.s..

1. Soit $c \in \mathbb{R}$. Montrer que

$$E[(X - Y)1_{X>c, Y>c}] = E[(Y - X)1_{X>c \geq Y}] \leq 0.$$

En déduire que $E[(X - Y)1_{X>c, Y>c}] = 0$, puis que $P[X > c \geq Y] = 0$.

2. Montrer que $X = Y$ p.s..

3. on suppose maintenant que X et Y sont de carré intégrables. Montrer qu'une démonstration (beaucoup) plus directe de la question 2 est possible en calculant $E((X - Y)^2)$.

N.B. Le cas où X et Y sont seulement intégrables (traité dans les questions 1 et 2) peut aussi se faire avec la question 3 en "tronquant" les v.a.r. X et Y .

Exercice 11.10 (Espérance du produit et produit des espérances) *Corrigé 190 page 514*

Soit (Ω, \mathcal{A}, P) un espace de probabilités et X, Y deux v.a. intégrables t.q. XY est intégrable et $E[X|Y] = E(X)$ p.s.. Montrer que $E(XY) = E(X)E(Y)$.

Exercice 11.11 (Égalité de lois donne égalité d'espérances conditionnelles)

Soit (Ω, \mathcal{A}, P) un espace de probabilités et X, Y, Z trois v.a.r.. On suppose que X et Y sont intégrables et que $(X, Z) \sim (Y, Z)$.

1. Montrer que $E[X|Z] = E[Y|Z]$ p.s..
2. Soit f une fonction borélienne de \mathbb{R} dans \mathbb{R} . On suppose que $f(X)$ et $f(Y)$ sont intégrables (ou que f est à valeurs dans \mathbb{R}_+). Montrer que $E[f(X)|Z] = E[f(Y)|Z]$ p.s..

Exercice 11.12 (Convergence faible et espérances conditionnelles)

Soit (Ω, \mathcal{A}, P) un espace de probabilités, \mathcal{B} une sous-tribu de \mathcal{A} , X une v.a. intégrable et $(X_n)_{n \in \mathbb{N}}$ une suite de v.a. intégrables. On suppose que $X_n \rightarrow X$ faiblement dans $L^1(\Omega, \mathcal{A}, P)$ quand $n \rightarrow \infty$ (ce qui est équivalent à dire que, pour tout v.a. U bornée, on a $\lim_{n \rightarrow \infty} E[X_n U] = E[XU]$).

1. Montrer que pour toute v.a. U , \mathcal{B} -mesurable et bornée, on a

$$\lim_{n \rightarrow \infty} E[E[X_n|\mathcal{B}]U] = E[E[X|\mathcal{B}]U].$$

2. Montrer que $E[X_n|\mathcal{B}] \rightarrow E[X|\mathcal{B}]$ faiblement dans $L^1(\Omega, \mathcal{A}, P)$ quand $n \rightarrow \infty$.

Exercice 11.13 (Minoration d'une espérance conditionnelle)

Soit (Ω, \mathcal{A}, P) un espace de probabilité, \mathcal{B} une sous-tribu de \mathcal{A} et Y est une v.a.r. positive. Soit U une v.a.r. positive, \mathcal{B} -mesurable et t.q. pour toute v.a.r. positive \mathcal{B} -mesurable Z on ait $E(YZ) \geq E(UZ)$. Montrer que $E(Y|\mathcal{B}) \geq U$ p.s..

Exercice 11.14 (Dépendance linéaire et dépendance non linéaire)

Soit (Ω, \mathcal{A}, P) un espace de probabilité et X, Y deux v.a.r. non constantes.

1. (Dépendance linéaire.) On pose $\bar{X} = \frac{X - E(X)}{\sqrt{\text{Var}(X)}}$ et $\bar{Y} = \frac{Y - E(Y)}{\sqrt{\text{Var}(Y)}}$.

- (a) Montrer que $|\text{Cov}(\bar{X}, \bar{Y})| \leq 1$.
- (b) Montrer que $|\text{Cov}(\bar{X}, \bar{Y})| = 1$ si et seulement si il existe $\alpha, \beta \in \mathbb{R}$, $\alpha \neq 0$, t.q. $Y = \alpha X + \beta$.
- (c) Donner un exemple pour lequel $Y = f(X)$ (avec f fonction borélienne de \mathbb{R} dans \mathbb{R}) et $\text{Cov}(\bar{X}, \bar{Y}) = 0$.

2. (Dépendance.)

- (a) On suppose que X et Y sont indépendantes. Montrer que $E(Y|X) = E(Y)$.
- (b) Montrer qu'il existe f (fonction borélienne de \mathbb{R} dans \mathbb{R}) t.q. $Y = f(X)$ si et seulement $E(Y|X) = Y$.

Exercice 11.15 (Lorsque $E(Y|X) = X$ p.s....) *Corrigé 191 page 515*

Soit (Ω, \mathcal{A}, P) un espace probabilisé et X, Y deux v.a.r. de carré intégrable.

On suppose que $E(Y|X) = X$ p.s..

- 1.(a) Soit φ une fonction borélienne de \mathbb{R} dans \mathbb{R} . On suppose que la v.a.r. $\varphi(X)$ est de carré intégrable. Montrer que $\int_{\Omega} Y\varphi(X)dP = \int_{\Omega} X\varphi(X)dP$.
- (b) Montrer que $E(Y) = E(X)$ et $E(XY) = E(X^2)$.
- 2.(a) Montrer que $E(X^2) \leq E(Y^2)$.
- (b) Montrer que $Y = X$ p.s. si et seulement si $E(Y^2) = E(X^2)$.

Exercice 11.16 (V.a. gaussien et espérance conditionnelle) *Corrigé 192 page 515*

Soit (Ω, \mathcal{A}, P) un espace probabilisé et $(X, Y)^t$ un v.a. gaussien de dimension 2, On note a l'espérance de X , b l'espérance de Y et D la matrice de covariance du v.a. $(X, Y)^t$ (on a donc $D_{1,1} = \text{Var}(X)$, $D_{2,2} = \text{Var}(Y)$ et $D_{1,2} = D_{2,1} = \text{Cov}(X, Y)$). On suppose que $\text{Var}(Y) > 0$.

1. Calculer $\alpha, \beta \in \mathbb{R}$ (en fonction de a, b et D) de manière à avoir $E[X] = E[\alpha + \beta Y]$ et $E[XY] = E[(\alpha + \beta Y)Y]$.
- Avec α et β ainsi déterminés, on définit la fonction affine l de \mathbb{R} dans \mathbb{R} par $l(s) = \alpha + \beta s$ pour $s \in \mathbb{R}$ et on définit la v.a.r. Z par $Z = X - l(Y)$.
2. Montrer que $(Z, Y)^t$ est un v.a. gaussien. Montrer que $\text{Cov}(Z, Y) = 0$. En déduire que Z et Y sont des v.a.r. indépendantes [On pourra utiliser la question 1(b) de l'exercice 10.10].
3. Montrer que $E[X|Y] = l(Y)$ p.s.
4. Calculer (en fonction de D) $\text{Var}(Z)$.

Dans la suite, on note $\sigma = \sqrt{\text{Var}(Z)}$ et, pour $a \in \mathbb{R}$ on note μ_a la loi normale $\mathcal{N}(a, \sigma^2)$.

5. Soit f une fonction borélienne bornée de \mathbb{R} dans \mathbb{R} . Pour $a \in \mathbb{R}$, on pose $\psi(a) = \int_{\mathbb{R}} f d\mu_a$.

- (a) Montrer que ψ est une fonction continue (de \mathbb{R} dans \mathbb{R}) si $\sigma > 0$.
- (b) Montrer que $E[f(X)|Y] = \psi(l(Y))$ p.s.

Exercice 11.17 (Indépendance et espérance)

Soit (Ω, \mathcal{A}, P) un espace probabilisé et $\mathcal{B}_1, \mathcal{B}_2$ deux sous-tribus de \mathcal{A} . On pose $\mathcal{B} = \sigma(\mathcal{B}_1 \cup \mathcal{B}_2)$ (c'est-à-dire que \mathcal{B} est la tribu engendrée par \mathcal{B}_1 et \mathcal{B}_2). On pose aussi $\mathcal{C} = \{B_1 \cap B_2; B_1 \in \mathcal{B}_1, B_2 \in \mathcal{B}_2\}$.

1. Montrer que $\mathcal{B} = \sigma(\mathcal{C})$ (c'est-à-dire que \mathcal{B} est la tribu engendrée par \mathcal{C}).
- Soit Y une v.a.r. intégrable. On note $\sigma(Y)$ la tribu engendrée par Y . On suppose que $\sigma(Y)$ et \mathcal{B}_1 sont indépendantes de \mathcal{B}_2 .
2. Soit U_1 une v.a.r. \mathcal{B}_1 -mesurable et bornée et U_2 une v.a.r. \mathcal{B}_2 -mesurable et intégrable. Montrer que YU_1 et U_2 sont des v.a.r. indépendantes. En déduire que YU_1U_2 est intégrable et que $E(YU_1U_2) = E(YU_1)E(U_2)$.
3. On pose $Z_1 = E[Y|\mathcal{B}_1]$ (la v.a.r. Z_1 est \mathcal{B}_1 -mesurable intégrable et $E(YU) = E(Z_1U)$ pour toute v.a.r. \mathcal{B}_1 -mesurable bornée).
- (a) Montrer que $E(Y1_C) = E(Z_11_C)$ pour tout $C \in \mathcal{C}$.
- (b) Montrer que $E(Y1_B) = E(Z_11_B)$ pour tout $B \in \mathcal{B}$.
4. Montrer que $E[Y|\mathcal{B}] = E[Y|\mathcal{B}_1]$ p.s..

Exercice 11.18 (Espérance conditionnelle d'une somme)

Soit (Ω, \mathcal{A}, P) un espace probabilisé et $(X_i)_{i \in \mathbb{N}^*}$ une suite de v.a.r.i.i.d.. On suppose que X_1 est intégrable (on a donc aussi X_i intégrable pour tout $i \in \mathbb{N}^*$). Pour $n \in \mathbb{N}^*$, on pose $S_n = X_1 + \dots + X_n$.

1. (Cette question reprend l'exercice 11.8.) Soit $n > 1$. Montrer que pour tout $k \in \{1, \dots, n\}$, $E[X_k|S_n] = E[X_1|S_n]$ p.s.. [On pourra remarquer que les v.a. $(X_1, X_2, \dots, X_k, \dots, X_n)$ et $(X_k, X_2, \dots, X_1, \dots, X_n)$ ont même loi.]
- En déduire que $E[X_1|S_n] = \frac{S_n}{n}$ p.s..

2. Soit $n \in \mathbb{N}^*$. On note \mathcal{D}_n la tribu engendrée par l'ensemble des v.a.r. $S_k, k \geq n$. Montrer que $E[X_1 | \mathcal{D}_n] = E[X_1 | \mathcal{S}_n] = \frac{S_n}{n}$ p.s.. [On pourra utiliser l'exercice 11.17 en choisissant convenablement \mathcal{B}_1 et \mathcal{B}_2 .]

Exercice 11.19 (Convergence d'une suite d'espérance conditionnelles)

Soient (E, T, p) un espace probabilisé et $(T_n)_{n \in \mathbb{N}}$ une famille de tribus sur E t.q. $T_n \subset T_{n+1}$, pour tout $n \in \mathbb{N}$, et $\cup_{n \in \mathbb{N}} T_n = T$. Soit $X \in L^2_{\mathbb{R}}(E, T, p)$ et $E(X | T_n)$ l'espérance conditionnelle de X par rapport à la tribu T_n . Nous allons montrer que $E(X | T_n)$ converge vers X dans $L^2_{\mathbb{R}}(E, T, p)$ lorsque $n \rightarrow +\infty$.

1. Montrer qu'il existe $e \in L^2_{\mathbb{R}}(E, T, p)$ et une sous-suite de la suite $(E(X | T_n))_{n \in \mathbb{N}}$ qui converge faiblement vers e dans $L^2_{\mathbb{R}}(E, T, p)$.
2. Montrer que $\int XY dp = \int eY dp$, pour tout $Y \in F = \overline{\cup_{n \in \mathbb{N}} L^2_{\mathbb{R}}(E, T_n, p)}$.
3. Montrer que $F = L^2_{\mathbb{R}}(E, T, p)$ et en déduire que $e = X$ p.s..
4. Montrer que $\|E(X | T_n)\|_2 \leq \|X\|_2$ et en déduire que la suite $(e(X, T_n))_{n \in \mathbb{N}}$ converge vers X dans $L^2_{\mathbb{R}}(E, T, p)$.

Exercice 11.20 (Projection sur $L^2(\Omega, \tau(X), P)$ versus projection sur $ev(X)$)

Soit (Ω, \mathcal{A}, P) un espace de probabilités. Si \mathcal{B} est une sous-tribu de \mathcal{A} , on note encore P la restriction de P à \mathcal{B} , de sorte que (Ω, \mathcal{B}, P) est encore un espace de probabilités. On rappelle que $\mathcal{L}^2(\Omega, \mathcal{B}, P) \subset \mathcal{L}^2(\Omega, \mathcal{A}, P)$ et que l'on peut considérer $L^2(\Omega, \mathcal{B}, P)$ comme un s.e.v. de $L^2(\Omega, \mathcal{A}, P)$. L'espace $L^2(\Omega, \mathcal{B}, P)$ est donc un s.e.v. fermé de $L^2(\Omega, \mathcal{A}, P)$.

Soit $X \in L^2(\Omega, \mathcal{A}, P)$ (en choisissant un représentant de X , X est donc une v.a.). On note $ev(X)$ le s.e.v. engendré par X (noter que $ev(X)$ est un s.e.v. fermé de $L^2(\Omega, \mathcal{A}, P)$).

Si V est un s.e.v. fermé de $L^2(\Omega, \mathcal{A}, P)$, on note P_V l'opérateur de projection orthogonale sur V .

1. On suppose que X est constante et non nulle (c'est-à-dire qu'il existe $a \in \mathbb{R}^*$ t.q. $X = a$ p.s.). Montrer que $P_{ev(X)} = P_{L^2(\Omega, \tau(X), P)}$ (i.e. $ev(X) = L^2(\Omega, \tau(X), P)$).
2. On suppose que X n'est pas constante. Montrer que pour toute sous-tribu \mathcal{B} de \mathcal{A} , $P_{ev(X)} \neq P_{L^2(\Omega, \mathcal{B}, P)}$ (i.e. $ev(X) \neq L^2(\Omega, \mathcal{B}, P)$).

Remarque : Si Y est une v.a. de carré intégrable, on a $E[Y | X] = E[Y | \tau(X)] = P_{L^2(\Omega, \tau(X), P)} Y$.

11.3.2 Martingales

Exercice 11.21 (Martingale construite avec les espérances d'une v.a.)

Soit (Ω, \mathcal{A}, P) un espace de probabilité muni d'une filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$. Soit X une v.a.r. intégrable. Montrer que la suite $(X_n)_{n \in \mathbb{N}}$ définie par $X_n = E[X | \mathcal{B}_n]$ (pour tout $n \in \mathbb{N}$) est une martingale par rapport à la filtration (\mathcal{B}_n) .

Montrer que $(X_n)_{n \in \mathbb{N}}$ est aussi une martingale pour la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$ définie par $\mathcal{F}_n = \tau(X_1, \dots, X_n)$ (pour tout $n \in \mathbb{N}$).

Exercice 11.22 (Martingale d'un jeu équilibré)

Soit (Ω, \mathcal{A}, P) un espace probabilisé, X_0 une v.a.r. intégrable et $(J_n)_{n \in \mathbb{N}^*}$ une suite de v.a.r. intégrable et de moyenne nulle. On suppose que la suite formée de X_0 et $(J_n)_{n \in \mathbb{N}^*}$ est une suite de v.a.r. indépendantes. On pose alors, pour $n \in \mathbb{N}$, $X_{n+1} = X_n + J_{n+1}$ et \mathcal{B}_n la tribu engendrée par X_0, \dots, X_n . Montrer que la suite $(X_n)_{n \in \mathbb{N}}$ est une martingale (par rapport à la filtration $(\mathcal{B}_n)_n$).

Exercice 11.23 (Quelques propriétés des martingales) Corrigé 193 page 518

Soit (Ω, \mathcal{A}, P) un espace de probabilité muni d'une filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$ (c'est-à-dire d'une suite croissante de sous tribus de \mathcal{A}) et $(X_n)_{n \in \mathbb{N}}$ une suite de de v.a.r. (c'est-à-dire un processus réel). On suppose que X_n est intégrable pour tout $n \in \mathbb{N}$.

1. On suppose que $(X_n)_{n \in \mathbb{N}}$ est une sous-martingale (par rapport à la filtration $(\mathcal{B}_n)_n$) Montrer que la suite $(E(X_n))_{n \in \mathbb{N}}$ est croissante.
2. On suppose que $(X_n)_{n \in \mathbb{N}}$ est une martingale (par rapport à la filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$). Montrer que $E(X_{n+m} | \mathcal{B}_n) = X_n$ p.s. pour tout $m \geq 0$.
3. Soit φ une fonction convexe de \mathbb{R} dans \mathbb{R} . On suppose que $(X_n)_{n \in \mathbb{N}}$ est une martingale (par rapport à la filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$) et que $\varphi(X_n)$ est intégrable pour tout $n \in \mathbb{N}$ (on rappelle que $\varphi(X_n)$ est une notation pour désigner $\varphi \circ X_n$). Montrer que $(\varphi(X_n))_{n \in \mathbb{N}}$ est une sous-martingale (par rapport à la filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$).

Exercice 11.24 (Séries de Fourier et martingales)

Soit (Ω, \mathcal{A}, P) un espace de probabilité et \mathcal{B} une sous tribu de \mathcal{A} . Montrer que

$$X \in L^2(\Omega, \mathcal{B}, P) \Rightarrow X^+ \in L^2(\Omega, \mathcal{B}, P).$$

$$(L^2(\Omega, \mathcal{B}, P) = L^2_{\mathbb{R}}(\Omega, \mathcal{B}, P).)$$

On prend maintenant $\Omega =]0, 1[$, $\mathcal{A} = \mathcal{B}(]0, 1[)$ et $P = \lambda$ (la mesure de Lebesgue sur $\mathcal{B}(]0, 1[)$). On pose $H = L^2_{\mathbb{C}}(\Omega, \mathcal{A}, P)$. Pour $p \in \mathbb{Z}$, on définit $e_p \in H$ par $e_p(x) = \exp(2i\pi px)$ pour $x \in]0, 1[$. On rappelle que $\{e_p, p \in \mathbb{Z}\}$ est une base hilbertienne de H . pour $n \in \mathbb{N}$, on pose $V_n = \text{ev}\{e_p, -n \leq p \leq n\}$ (c'est un s.e.v. fermé de H).

Soit $X \in H$. On sait que $X_n \rightarrow X$ dans H , quand $n \rightarrow \infty$, avec $X_n = P_{V_n} X$ où P_{V_n} désigne l'opérateur de projection orthogonale sur V_n (X_n est donc une somme partielle de la série de Fourier de X).

1. Montrer que $P_{V_n}(X_{n+1}) = X_n$ pour tout $n \in \mathbb{N}$.
2. Soit $n \in \mathbb{N}^*$, montrer qu'il n'existe pas de sous-tribu \mathcal{B} de \mathcal{A} t.q. $V_n = L^2_{\mathbb{C}}(\Omega, \mathcal{B}, P)$. [On pourra, par exemple, commencer par remarquer que V_n est formé de fonctions analytiques.]
3. Soit $n \in \mathbb{N}^*$. On pose $\mathcal{B}_n = \tau(e_p, -n \leq p \leq n)$. A-t-on $V_n \subset L^2_{\mathbb{C}}(\Omega, \mathcal{B}_n, P)$ ou $L^2_{\mathbb{C}}(\Omega, \mathcal{B}_n, P) \subset V_n$?

Exercice 11.25 (Quelques questions sur les temps d'arrêt)

Soit (Ω, \mathcal{A}, P) un espace de probabilité muni d'une filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$.

1. Soit $(X_n)_{n \in \mathbb{N}}$ une suite de v.a.r., adaptée à la filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$. Soit $E \in \mathcal{B}(\mathbb{R})$ et τ défini de Ω dans \mathbb{R} par

$$\tau(\omega) = 1 \text{ si } X_1(\omega) \in E \text{ et } \tau = 5 \text{ si } X_1(\omega) \notin E.$$

Montrer que τ est un temps d'arrêt.

2. Soit ν et τ deux temps d'arrêt par rapport à la filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$. Montrer que $\nu + \tau$ est encore un temps d'arrêt (par rapport à la filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$).
3. Soit ν et τ deux temps d'arrêt, par rapport à la filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$, et t.q. $\nu \leq \tau$ p.s.. Soit \mathcal{B}_{ν} et \mathcal{B}_{τ} les deux tribus associées. Montrer que $\mathcal{B}_{\nu} \subset \mathcal{B}_{\tau}$. [Si T est un temps d'arrêt, on note $\mathcal{B}_T = \{A \in \mathcal{A} \text{ t.q., pour tout } n \in \mathbb{N}, A \cap \{T = n\} \in \mathcal{B}_n\}$.]

Exercice 11.26 (Martingale arrêtée à un temps d'arrêt)

Soit (Ω, \mathcal{A}, P) un espace de probabilité muni d'une filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$ et $(X_n)_{n \in \mathbb{N}}$, $(\phi_n)_{n \in \mathbb{N}^*}$ deux suites de v.a.r.. On suppose que, pour tout $n \in \mathbb{N}$, X_n est intégrable et que, pour tout $n \in \mathbb{N}^*$, ϕ_n est bornée. Pour $n \in \mathbb{N}^*$, on pose $(\Delta X)_n = X_n - X_{n-1}$ et $(\phi \cdot X)_n = \sum_{k=1}^n \phi_k(\Delta X)_k$ (ceci est une intégrale stochastique discrète).

1. On suppose que $(X_n)_{n \in \mathbb{N}}$ est une martingale et que $(\phi_n)_{n \in \mathbb{N}^*}$ est prévisible (c'est-à-dire que ϕ_{n+1} est \mathcal{B}_n -mesurable pour tout $n \in \mathbb{N}$). Montrer que $((\phi \cdot X)_n)_{n \in \mathbb{N}^*}$ est une martingale.
2. (Exemple) Soit ν un temps d'arrêt. On prend ici $\phi_n = 1_{\{\nu \geq n\}}$ pour tout $n \in \mathbb{N}^*$. Montrer que $(\phi_n)_{n \in \mathbb{N}^*}$ est prévisible et $(\phi \cdot X)_n = X_{\nu \wedge n}$ pour tout $n \in \mathbb{N}^*$. (On rappelle que $\nu \wedge n(\omega) = \min\{\nu(\omega), n\}$ et donc que $X_{\nu \wedge n}(\omega) = X_{\min\{\nu(\omega), n\}}(\omega)$, pour tout $\omega \in \Omega$.)
Remarque : La question précédente permet de montrer qu'une martingale arrêtée à un temps d'arrêt est encore une martingale et donne donc une démonstration du théorème 11.2.

Exercice 11.27 (Caractérisation des martingales régulières)

Soit (Ω, \mathcal{A}, P) un espace de probabilité muni d'une filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$ et $(X_n)_{n \in \mathbb{N}}$ une martingale bornée dans L^1 (c'est-à-dire t.q. $\sup\{E(|X_n|), n \in \mathbb{N}\} < \infty$).

1. Montrer que la suite $(X_n)_{n \in \mathbb{N}}$ est équi-intégrable si et seulement si $(X_n)_{n \in \mathbb{N}}$ est une martingale régulière (c'est-à-dire que la suite $(X_n)_{n \in \mathbb{N}}$ converge dans $L^1_{\mathbb{R}}(\Omega, \mathcal{A}, P)$). [On pourra utiliser le théorème de Vitali.]
2. On suppose qu'il existe $q > 0$ t.q. $\sup\{E(|X_n|^q), n \in \mathbb{N}\} < \infty$. Montrer que $(X_n)_{n \in \mathbb{N}}$ est une martingale régulière.

Exercice 11.28 (Une condition pour qu'un temps d'arrêt soit intégrable)

Soit (Ω, \mathcal{A}, P) un espace de probabilité muni d'une filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$ et T un temps d'arrêt par rapport à la filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$.

1. On suppose qu'il existe $\varepsilon_0 > 0$ et $n_0 \in \mathbb{N}$ t.q.

$$P[T > kn_0] \leq (1 - \varepsilon_0)^{k-1} \text{ pour tout } k \in \mathbb{N}^*. \quad (11.10)$$

Montrer que $E[T] < \infty$.

2. On suppose qu'il existe $\varepsilon_0 > 0$ et $n_0 \in \mathbb{N}$ t.q. $\mathcal{P}[n + n_0 \geq T > n] \geq \varepsilon_0 P[T > n]$ pour tout $n \in \mathbb{N}$. Montrer que l'inégalité (11.10) est vraie.
3. On suppose qu'il existe $\varepsilon_0 > 0$ et $n_0 \in \mathbb{N}$ t.q. $E[1_{\{T \leq n + n_0\}} | \mathcal{B}_n] > \varepsilon_0$ p.s., pour tout $n \in \mathbb{N}$. Montrer que $P[T \in]n, n + n_0] \geq \varepsilon_0 P[T > n]$. En déduire que $E[T] < \infty$.

[On rappelle que pour une v.a.r. X , on a $E[|X|] < \infty$ si et seulement si $\sum_{n \in \mathbb{N}} \mathcal{P}[|X| \geq n] < \infty$.]

Exercice 11.29 (On joue à pile ou face...)

Soit (Ω, \mathcal{A}, P) un espace de probabilité et $(J_n)_{n \in \mathbb{N}^*}$ une suite de v.a.r.i.i.d. de loi $m = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$ (c'est-à-dire que, pour tout $n \in \mathbb{N}^*$, $P[J_n = 1] = P[J_n = -1] = 1/2$). On pose $X_0 = 0$ et, pour $n \in \mathbb{N}$, $X_{n+1} = X_n + J_{n+1}$. Soit a et b deux entiers strictement positifs et, pour tout $\omega \in \Omega$,

$$T(\omega) = \inf\{n \geq 0 \text{ t.q. } X_n(\omega) = -a \text{ ou } X_n(\omega) = b\} \text{ si il existe } n \in \mathbb{N} \text{ t.q. } X_n(\omega) \in \{-a, b\},$$

$$T(\omega) = \infty \text{ sinon.}$$

On note $p_a = P(\{\omega \in \Omega, \text{ t.q. } X_{T(\omega)}(\omega) = -a\})$ (p_a est donc la probabilité que X_n atteigne $-a$ avant b). Pour $n \in \mathbb{N}$, on désigne par \mathcal{B}_n la tribu engendrée par X_0, \dots, X_n et on pose, pour $\omega \in \Omega$, $Y_n(\omega) = X_{\min(n, T(\omega))}(\omega)$.

1. Montrer que $(\mathcal{B}_n)_{n \in \mathbb{N}}$ est une filtration et que T est un temps d'arrêt pour cette filtration.
2. Montrer que $E[T] < \infty$.
3. Montrer que $(X_n)_{n \in \mathbb{N}}$ est une martingale (par rapport à la filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$).
4. Montrer que $(Y_n)_{n \in \mathbb{N}}$ est une martingale (par rapport à la filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$) et que $p_a = b/(a + b)$.

N.B. : On peut aussi montrer que $E[T] = ab$.

Exercice 11.30 (Il est temps de s'arrêter)

Soit (Ω, \mathcal{A}, P) un espace de probabilité muni d'une filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$ et $(X_n)_{n \in \mathbb{N}}$ une suite de v.a.r., adaptée à la filtration $(\mathcal{F}_n)_{n \in \mathbb{N}}$. On suppose aussi que $E[|X_n|] < \infty$ pour tout $n \in \mathbb{N}$.

1. Montrer que pour tout temps d'arrêt (par rapport à $(\mathcal{F}_n)_{n \in \mathbb{N}}$) τ borné,

$$E[|X_\tau|] < \infty.$$

On suppose dans la suite que pour tout temps d'arrêt (par rapport à $(\mathcal{F}_n)_{n \in \mathbb{N}}$) τ borné,

$$E[X_\tau] = E[X_0].$$

2. Soit $n \in \mathbb{N}$ et $\Lambda_n \in \mathcal{F}_n$. On définit σ_n par

$$\sigma_n(\omega) = n \text{ si } \omega \in \Lambda_n \text{ et } \sigma_n(\omega) = n + 1 \text{ si } \omega \notin \Lambda_n.$$

Montrer que σ_n est un temps d'arrêt (par rapport à $(\mathcal{F}_n)_{n \in \mathbb{N}}$).

3. Soit $n \in \mathbb{N}$ et ν_n défini par $\nu_n(\omega) = n + 1$ pour tout $\omega \in \Omega$. Montrer que ν_n est aussi un temps d'arrêt (par rapport à $(\mathcal{F}_n)_{n \in \mathbb{N}}$).

4. En remarquant que $X_\tau = X_\tau 1_A + X_\tau 1_{A^c}$ (pour tout temps d'arrêt τ et tout événement A), montrer que $(X_n)_{n \in \mathbb{N}}$ est une martingale.

Chapitre 21

Transformation de Fourier

21.1 Transformée de Fourier dans L^1

Corrigé 182 (Résultat partiel d'inversion de Fourier dans L^1)

Soit $H(t) = e^{-|t|}$, $t \in \mathbb{R}$. On pose, pour $\lambda > 0$:

$$h_\lambda(x) = (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} H(\lambda t) e^{itx} dt, x \in \mathbb{R}. \quad (21.1)$$

1. Montrer que $h_\lambda(x) = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \frac{\lambda}{\lambda^2 + x^2}$, et $\int_{\mathbb{R}} h_\lambda(x) dx = (2\pi)^{\frac{1}{2}}$.

corrigé

Soit $x \in \mathbb{R}$. Comme H est paire, on a $(2\pi)^{\frac{1}{2}} h_\lambda(x) = \int_{\mathbb{R}} e^{-|\lambda t|} \cos(tx) dt$ et donc

$$(2\pi)^{\frac{1}{2}} h_\lambda(x) = 2 \lim_{n \rightarrow \infty} \int_0^n e^{-\lambda t} \cos(tx) dt.$$

En intégrant deux fois par parties, on remarque que

$$\int_0^n e^{-\lambda t} \cos(tx) dt = \frac{1}{\lambda} - \left(\frac{x}{\lambda}\right)^2 \int_0^n e^{-\lambda t} \cos(tx) dt + a_n,$$

avec $\lim_{n \rightarrow \infty} a_n = 0$. Ceci donne

$$\int_0^n e^{-\lambda t} \cos(tx) dt = \frac{\lambda}{\lambda^2 + x^2} (1 + \lambda a_n),$$

et donc $h_\lambda(x) = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \frac{\lambda}{\lambda^2 + x^2}$.

Pour calculer $\int_{\mathbb{R}} h_\lambda(x) dx$, on utilise le changement de variable $x = \lambda y$, on obtient

$$\int_{\mathbb{R}} h_\lambda(x) dx = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \int_{\mathbb{R}} \frac{1}{1 + y^2} dy = (2\pi)^{\frac{1}{2}}.$$

2. Soit $f \in L^1_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, montrer que pour tout $x \in \mathbb{R}$, on a :

$$f \star h_{\lambda}(x) = \int_{\mathbb{R}} H(\lambda t) \hat{f}(t) e^{ixt} dt. \quad (21.2)$$

corrigé

Noter que λ désigne ici à la fois le paramètre λ introduit au début de l'énoncé et la mesure de Lebesgue. Cette maladresse de notation semble toutefois sans gravité.

Comme $f \in L^1_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ et $h_{\lambda} \in L^{\infty}_{\mathbb{R}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, $f \star h_{\lambda}(x)$ est défini pour tout $x \in \mathbb{R}$ et on a :

$$f \star h_{\lambda}(x) = \int_{\mathbb{R}} f(t) h_{\lambda}(x-t) dt = (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} f(t) \left(\int_{\mathbb{R}} H(\lambda y) e^{i(x-t)y} dy \right) dt.$$

Comme $f \in L^1_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ et $H(\lambda \cdot) \in L^{\infty}_{\mathbb{R}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, on peut utiliser le théorème de Fubini (théorème 7.3) pour inverser l'ordre d'intégration et obtenir :

$$\begin{aligned} f \star h_{\lambda}(x) &= (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} H(\lambda y) e^{ixy} \left(\int_{\mathbb{R}} f(t) e^{-ity} dt \right) dy \\ &= \int_{\mathbb{R}} H(\lambda y) e^{ixy} \hat{f}(y) dy. \end{aligned}$$

3. Soit g une fonction mesurable bornée de \mathbb{R} dans \mathbb{C} , continue en 0. Montrer que $g \star h_{\lambda}(0) \rightarrow \sqrt{2\pi}g(0)$ quand $\lambda \rightarrow 0$. [Utiliser 1. et le théorème de convergence dominée.]

corrigé

On utilise maintenant le fait que $h_{\lambda} \in L^1_{\mathbb{R}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ et $g \in L^{\infty}_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ pour remarquer que $g \star h_{\lambda}(x)$ est défini pour tout $x \in \mathbb{R}$. Pour $x = 0$, on a :

$$g \star h_{\lambda}(0) = \int_{\mathbb{R}} g(x) h_{\lambda}(x) dx = \left(\frac{2}{\pi} \right)^{\frac{1}{2}} \int_{\mathbb{R}} \frac{\lambda}{\lambda^2 + x^2} g(x) dx.$$

Avec le changement de variable $y = \frac{x}{\lambda}$, on obtient :

$$g \star h_{\lambda}(0) = \left(\frac{2}{\pi} \right)^{\frac{1}{2}} \int_{\mathbb{R}} g(\lambda y) \frac{1}{1+y^2} dy.$$

Comme $|g(\lambda y) \frac{1}{1+y^2}| \leq \|g\|_{\infty} \frac{1}{1+y^2}$ (avec $\|g\|_{\infty} = \sup_{x \in \mathbb{R}} |g(x)|$) et que la fonction $y \mapsto \frac{1}{1+y^2}$ est intégrable sur \mathbb{R} , on peut utiliser le théorème de convergence dominée pour en déduire (grâce à la continuité de g en 0) :

$$\lim_{\lambda \rightarrow 0} g \star h_{\lambda}(0) = \left(\frac{2}{\pi} \right)^{\frac{1}{2}} g(0) \int_{\mathbb{R}} \frac{1}{1+y^2} dy = (2\pi)^{\frac{1}{2}} g(0).$$

4. Soit $f \in L^1_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, montrer que :

$$\|f \star h_{\lambda} - \sqrt{2\pi}f\|_1 \rightarrow 0 \text{ lorsque } \lambda \rightarrow 0. \quad (21.3)$$

[On pourra utiliser la continuité en moyenne et la question précédente avec $g(y) = \int |f(x-y) - f(x)| dx$.]

corrigé

Comme $\int_{\mathbb{R}} h_{\lambda}(y) dy = \sqrt{2\pi}$, on a :

$$\begin{aligned} \|f \star h_{\lambda} - \sqrt{2\pi}f\|_1 &= \int_{\mathbb{R}} \left| \int_{\mathbb{R}} (f(x-y) - f(x)) h_{\lambda}(y) dy \right| dx \\ &\leq \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x-y) - f(x)| h_{\lambda}(y) dy \right) dx. \end{aligned}$$

En utilisant le théorème de Fubini-Tonelli (théorème 7.2), on en déduit :

$$\|f \star h_{\lambda} - \sqrt{2\pi}f\|_1 \leq \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x-y) - f(x)| dx \right) h_{\lambda}(y) dy = g \star h_{\lambda}(0),$$

avec $g(y) = \int |f(x-y) - f(x)| dx$.

Le théorème de continuité en moyenne dans L^1 (théorème 5.6, écrit pour de fonctions à valeurs réelles mais la généralisation est immédiate pour des fonctions à valeurs complexes) donne que g est continue en 0 et donc aussi continue de \mathbb{R} dans \mathbb{R} (en remarquant que $|g(y) - g(z)| \leq g(y-z)$) et donc aussi mesurable de \mathbb{R} dans \mathbb{C} . Elle est également bornée (car $|g(y)| \leq 2\|f\|_1$, pour tout $y \in \mathbb{R}$). On peut donc utiliser la question précédente, elle donne que $\lim_{\lambda \rightarrow 0} g \star h_{\lambda}(0) = 0$ et donc que $\lim_{\lambda \rightarrow 0} \|f \star h_{\lambda} - \sqrt{2\pi}f\|_1 = 0$.

5. Dédurre de ce qui précède le théorème d'inversion de Fourier, théorème 10.1.

corrigé

On note $L^1 = L^1_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$. Soit $f \in L^1$ (on a donc $\hat{f} \in C_0(\mathbb{R}, \mathbb{C})$). On suppose que $\hat{f} \in L^1$. Soit $(\lambda_n)_{n \in \mathbb{N}} \subset \mathbb{R}_+^*$ une suite t.q. $\lim_{n \rightarrow \infty} \lambda_n = 0$. Comme $f \in L^1$, la question précédente nous donne que $f \star h_{\lambda_n} \rightarrow \sqrt{2\pi}f$ dans L^1 et la question 2 nous donne pour tout $n \in \mathbb{N}$ et tout $x \in \mathbb{R}$:

$$f \star h_{\lambda_n}(x) = \int_{\mathbb{R}} H(\lambda_n t) \hat{f}(t) e^{ixt} dt.$$

On utilise maintenant le théorème de convergence dominée qui s'applique parce que $\hat{f} \in L^1$ et que l'on a, pour tout x et tout n , $|H(\lambda_n t) \hat{f}(t) e^{ixt}| \leq |\hat{f}(t)|$. Comme $\lim_{n \rightarrow \infty} H(\lambda_n x) = 1$, pour tout $x \in \mathbb{R}$, on a donc, pour tout $x \in \mathbb{R}$:

$$\lim_{n \rightarrow \infty} f \star h_{\lambda_n}(x) = \int_{\mathbb{R}} \hat{f}(t) e^{ixt} dt = \sqrt{2\pi} \hat{\hat{f}}(-x).$$

Enfin, comme $f \star h_{\lambda_n} \rightarrow \sqrt{2\pi}f$ dans L^1 , on peut supposer, après extraction d'une sous suite, que $f \star h_{\lambda_n} \rightarrow \sqrt{2\pi}f$ p.p.. On a donc, finalement (par unicité de la limite dans \mathbb{R}), $\sqrt{2\pi}f = \sqrt{2\pi} \hat{\hat{f}}(-\cdot)$ p.p., c'est-à-dire $f = \hat{\hat{f}}(-\cdot)$ p.p..

21.2 Transformée de Fourier d'une mesure signée

Corrigé 183 (Caractérisation de m par \hat{m})

Soit $d \geq 1$.

1. Soit m et μ deux mesures signées sur les boréliens de \mathbb{R}^d . On suppose que $\hat{m} = \hat{\mu}$.

(a) Soit $\varphi \in L^1_{\mathbb{C}}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda_d)$. Montrer que $\int \hat{\varphi} dm = \int \hat{\varphi} d\mu$.

corrigé

On remarque que $\int \hat{\varphi} dm = (2\pi)^{-\frac{d}{2}} \int \left(\int e^{-ix \cdot t} \varphi(x) dx \right) dm(t)$. (Les intégrales sont toutes sur \mathbb{R}^d). La mesure signée m peut se décomposer en différence de deux mesures positives étrangères $m = m^+ - m^-$ (décomposition de Hahn, proposition 2.6). Comme

$$\int \int |e^{-ix \cdot t} \varphi(x)| dx dm^{\pm}(t) = \|\varphi\|_1 m^{\pm}(\mathbb{R}^d) < +\infty,$$

on peut utiliser le théorème de Fubini-Tonelli (théorème 7.2) avec les mesures λ_d et m^+ et les mesures λ_d et m^- . On obtient ainsi :

$$\int \hat{\varphi} dm = (2\pi)^{-\frac{d}{2}} \int \left(\int e^{-ix \cdot t} dm(t) \right) \varphi(x) dx = \int \hat{m}(x) \varphi(x) dx.$$

Le même raisonnement donne $\int \hat{\varphi} d\mu = \int \hat{\mu}(x) \varphi(x) dx$. Comme $\hat{m}(x) = \hat{\mu}(x)$ pour tout $x \in \mathbb{R}^d$, on en déduit bien $\int \hat{\varphi} dm = \int \hat{\varphi} d\mu$.

(b) Montrer que $\int \varphi dm = \int \varphi d\mu$ pour tout $\varphi \in \mathcal{S}(\mathbb{R}^d, \mathbb{C})$ (et donc pour tout $\varphi \in C_c^{\infty}(\mathbb{R}^d, \mathbb{R})$).

corrigé

Comme $\mathcal{S}(\mathbb{R}^d, \mathbb{C}) \subset L^1_{\mathbb{C}}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda_d)$, la question précédente donne $\int \hat{\varphi} dm = \int \hat{\varphi} d\mu$ pour tout $\varphi \in \mathcal{S}(\mathbb{R}^d, \mathbb{C})$. Or, l'application $\varphi \mapsto \hat{\varphi}$ est une bijection dans $\mathcal{S}(\mathbb{R}^d, \mathbb{C})$ (proposition 10.6). On a donc $\int \varphi dm = \int \varphi d\mu$ pour tout $\varphi \in \mathcal{S}(\mathbb{R}^d, \mathbb{C})$.

(c) Montrer que $m = \mu$ (On rappelle qu'une fonction de $\varphi \in C_c(\mathbb{R}^d, \mathbb{R})$ est limite uniforme de fonctions de $\varphi \in C_c^{\infty}(\mathbb{R}^d, \mathbb{R})$).

corrigé

Soit $\varphi \in C_c(\mathbb{R}^d, \mathbb{R})$ et $(\varphi_n)_{n \in \mathbb{N}} \subset C_c^{\infty}(\mathbb{R}^d, \mathbb{R})$ t.q. $\varphi_n \rightarrow \varphi$, uniformément sur \mathbb{R}^d , quand $n \rightarrow \infty$. La question précédente donne $\int \varphi_n dm = \int \varphi_n d\mu$ pour tout $n \in \mathbb{N}$. On utilise alors le théorème de convergence dominée (ce qui est possible car les mesures m^{\pm} et μ^{\pm} sont des mesures finies), il donne $\int \varphi dm = \int \varphi d\mu$.

La proposition 5.4 donne alors $m = \mu$.

2. Soit m une mesure signée sur les boréliens de \mathbb{R}^d . On suppose que $\hat{m} \in L^1_{\mathbb{C}}(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda_d)$. Montrer que m est la mesure de densité f par rapport à la mesure de Lebesgue avec $f = \hat{m}(\cdot)$.

21.3 Fonction caractéristique d'un v.a.

Corrigé 184 (V.a. gaussiens, indépendance, covariance)

Soit (Ω, \mathcal{A}, P) un espace de probabilités, $d \geq 1$ et $X = (X_1, \dots, X_d)$ un v.a. de dimension d .

1. On suppose ici que $d = 2$.

(a) On suppose que X_1 et X_2 suivent des lois gaussiennes et sont indépendantes, montrer que X est un vecteur gaussien et que $\text{Cov}(X_1, X_2) = 0$.

corrigé

Comme X_1 et X_2 suivent des lois gaussiennes, il existe $m_1, m_2 \in \mathbb{R}$ et $\sigma_1, \sigma_2 \in \mathbb{R}_+$ t.q. $X_k \sim \mathcal{N}(m_k, \sigma_k^2)$ pour $i = 1, 2$. On a donc $\varphi_{X_k}(u) = e^{i u m_k} e^{-\frac{\sigma_k^2 u^2}{2}}$ pour $k = 1, 2$ et pour tout $u \in \mathbb{R}$.

Soit $a_1, a_2 \in \mathbb{R}$. On calcule alors la fonction caractéristique de la v.a.r. $a_1X_1 + a_2X_2$. Soit $u \in \mathbb{R}$, on a :

$$\varphi_{a_1X_1+a_2X_2}(u) = \int_{\Omega} e^{i(a_1X_1+a_2X_2)u} dP = \int_{\Omega} e^{ia_1X_1u} e^{ia_2X_2u} dP$$

En utilisant l'indépendance de X_1 et X_2 , on en déduit :

$$\varphi_{a_1X_1+a_2X_2}(u) = \int_{\Omega} e^{ia_1X_1u} dP \int_{\Omega} e^{ia_2X_2u} dP = \varphi_{X_1}(a_1u) \varphi_{X_2}(a_2u).$$

Ce qui donne $\varphi_{a_1X_1+a_2X_2}(u) = e^{iu(a_1m_1+a_2m_2)} e^{-\frac{u^2(\sigma_1^2 a_1^2 + \sigma_2^2 a_2^2)}{2}}$. Comme la loi d'une v.a.r. est entièrement déterminée par sa fonction caractéristique (voir la proposition 10.10), on en déduit que $a_1X_1 + a_2X_2 \sim \mathcal{N}(m, \sigma^2)$ avec $m = a_1m_1 + a_2m_2$ et $\sigma = \sqrt{\sigma_1^2 a_1^2 + \sigma_2^2 a_2^2}$. Ceci prouve bien que X est un vecteur gaussien.

L'indépendance de X_1 et X_2 permet aussi de calculer la fonction caractéristique de X . Soit $u = (u_1, u_2)^t \in \mathbb{R}^2$:

$$\varphi_X(u) = \int_{\Omega} e^{i(X_1u_1+X_2u_2)} dP = e^{i(u_1m_1+u_2m_2)} e^{-\frac{u_1^2\sigma_1^2+u_2^2\sigma_2^2}{2}}.$$

Ceci prouve que la matrice de covariance de X est diagonale (proposition 10.12) et donc que

$$\text{Cov}(X_1, X_2) = 0.$$

- (b) On suppose que X est un vecteur gaussien et que $\text{Cov}(X_1, X_2) = 0$. Montrer que X_1 et X_2 sont indépendantes.

corrigé

D'après la proposition 10.11, pour montrer que X_1 et X_2 sont indépendantes, il suffit de montrer que $\varphi_X(u) = \prod_{j=1}^2 \varphi_{X_j}(u_j)$ pour tout $u = (u_1, u_2)^t \in \mathbb{R}^2$. Ceci est une conséquence facile du fait que $\text{Cov}(X_1, X_2) = 0$. En effet, la matrice de covariance de X est alors diagonale et on a bien (grâce à la proposition 10.12), en reprenant les notations de la question précédente (c'est-à-dire $X_k \sim \mathcal{N}(m_k, \sigma_k^2)$ pour $i = 1, 2$),

$$\varphi_X(u) = e^{i(u_1m_1+u_2m_2)} e^{-\frac{u_1^2\sigma_1^2+u_2^2\sigma_2^2}{2}} = \varphi_{X_1}(u_1) \varphi_{X_2}(u_2),$$

c'est-à-dire $\varphi_X(u) = \prod_{j=1}^2 \varphi_{X_j}(u_j)$ pour tout $u = (u_1, u_2)^t \in \mathbb{R}^2$.

2. On suppose toujours $d = 2$. Donner un exemple pour lequel X_1 et X_2 sont gaussiennes mais X n'est pas un vecteur gaussien. [On pourra, par exemple, choisir (Ω, \mathcal{A}, P) et X_1, X_2 de manière à avoir $\text{Cov}(X_1, X_2) = 0$ sans que X_1, X_2 soient indépendantes, voir l'exercice 4.44.]

corrigé

On considère les deux v.a.r. S et X de l'exercice 4.44 et on prend $X_1 = SX$ et $X_2 = X$. Les v.a.r. X_1 et X_2 sont gaussiennes (on a $X_1 \sim \mathcal{N}(0, 1)$ et $X_2 \sim \mathcal{N}(0, 1)$), elles sont dépendantes et on a $\text{Cov}(X_1, X_2) = 0$ (voir l'exercice 4.44). On en déduit que le v.a. $X = (X_1, X_2)$ n'est pas gaussien (sinon les v.a.r. seraient indépendantes, d'après la question précédente).

3. On suppose que X est un vecteur gaussien et que les composantes de X sont indépendantes deux à deux. Montrer que X_1, \dots, X_d sont indépendantes.

corrigé

La démonstration est ici similaire à celle de la question 1(b). D'après la proposition 10.11, pour montrer que les v.a.r. X_1, \dots, X_d sont indépendantes, il suffit de montrer que $\varphi_X(u) = \prod_{j=1}^d \varphi_{X_j}(u_j)$ pour tout $u = (u_1, \dots, u_d)^t \in \mathbb{R}^d$. Ceci est une conséquence facile du fait que les v.a.r. X_1, \dots, X_d sont indépendantes deux à deux. En effet, cette hypothèse d'indépendance deux à deux donne que la matrice de covariance de X est diagonale et on a alors (grâce à la proposition 10.12), avec $X_k \sim \mathcal{N}(m_k, \sigma_k^2)$ pour $i = 1, \dots, d$,

$$\varphi_X(u) = e^{i \sum_{k=1}^d u_k m_k} e^{-\sum_{k=1}^d \frac{u_k^2 \sigma_k^2}{2}} = \varphi_{X_1}(u_1) \dots \varphi_{X_d}(u_d),$$

c'est-à-dire $\varphi_X(u) = \prod_{j=1}^d \varphi_{X_j}(u_j)$ pour tout $u = (u_1, \dots, u_d)^t \in \mathbb{R}^d$.

Corrigé 185 (Suite de v.a.r.i.i.d. de Poisson)

Soit (Ω, \mathcal{A}, P) un espace de probabilité et X une v.a. de Poisson de paramètre λ ($\lambda > 0$). On rappelle que, $P[X = k] = e^{-\lambda} \frac{\lambda^k}{k!}$, pour tout $k \in \mathbb{N}$ et que $E[X] = \lambda$, $Var[X] = \lambda$.

1. Calculer la fonction caractéristique φ_X de X .

corrigé

Soit $u \in \mathbb{R}$, on a

$$\varphi_X(u) = \int_{\Omega} e^{iXu} dP = \sum_{k \in \mathbb{N}} e^{iku} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k \in \mathbb{N}} \frac{(\lambda e^{iu})^k}{k!} = e^{\lambda(e^{iu} - 1)}.$$

2. Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de v.a. indépendantes et de Poisson de paramètre λ .

(a) Soit $n > 1$. Dédurre de la question 1 la loi de la v.a. $Y_n = \sum_{p=1}^n X_p$.

corrigé

Soit $u \in \mathbb{R}$. On a $\varphi_{Y_n}(u) = \int_{\Omega} e^{iY_n u} dP = \int_{\Omega} \prod_{p=1}^n e^{iX_p u} dP$. Comme les v.a.r. X_1, \dots, X_n sont indépendantes, on en déduit que

$$\varphi_{Y_n}(u) = \prod_{p=1}^n \int_{\Omega} e^{iX_p u} dP = e^{n\lambda(e^{iu} - 1)}.$$

Comme la loi d'une v.a.r. est entièrement déterminée par sa fonction caractéristique (proposition 10.10), on en déduit que Y_n est une v.a. de Poisson de paramètre $n\lambda$.

(b) Utiliser le théorème central limite pour démontrer que

$$e^{-n} \sum_{k=0}^n \frac{n^k}{k!} \rightarrow \frac{1}{2} \text{ quand } n \rightarrow \infty.$$

corrigé

On suppose maintenant que $\lambda = 1$. la suite $(X_n)_{n \in \mathbb{N}^*}$ est une suite de v.a.r.i.i.d. de carrés intégrables et on a $E(X_1) = 1$ et $Var(X_1) = 1$. Pour $n \in \mathbb{N}^*$, on pose

$$Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i - 1) = \frac{1}{\sqrt{n}} (Y_n - n).$$

Le théorème central limite (théorème 10.4) donne que la suite $(P_{Z_n})_{n \in \mathbb{N}^*}$ converge étroitement vers la loi normale $\mathcal{N}(0, 1)$. Comme une loi normale est diffuse (c'est-à-dire qu'elle ne charge pas les points), on en déduit que $P(Z_n \leq 0)$ tend vers $1/2$ quand $n \rightarrow \infty$ (voir l'exercice 6.60 et noter que $P(Z_n \leq 0) = P_{Z_n}(\cdot - \infty, 0]$). Or, $P(Z_n \leq 0) = P(Y_n \leq n)$ et, comme Y_n est une v.a. de Poisson de paramètre n , on a :

$$P(Y_n \leq n) = \sum_{k=0}^n P(Y_n = k) = \sum_{k=0}^n e^{-n} \frac{n^k}{k!}.$$

On a donc $e^{-n} \sum_{k=0}^n \frac{n^k}{k!} \rightarrow 1/2$, quand $n \rightarrow \infty$.

Corrigé 186 (Sur les lois des grands nombres)

Soit (Ω, \mathcal{A}, P) un espace probabilisé et X une v.a.r. dont la loi est la loi de Cauchy c'est-à-dire que $P_X = f\lambda$, avec $f(x) = \frac{1}{\pi(1+x^2)}$ pour $x \in \mathbb{R}$ (on rappelle que λ désigne la mesure de Lebesgue sur les boréliens de \mathbb{R}).

1. Montrer que X n'est pas une v.a.r. intégrable.
2. Soit $n \in \mathbb{N}^*$, montrer que $P(\{|X| > n\}) \geq \frac{1}{n\pi}$, où $\{|X| > n\} = \{\omega \in \Omega, |X(\omega)| > n\}$. [On pourra remarquer que $\frac{2}{1+x^2} \geq \frac{1}{x^2}$ pour tout $x \geq 1$.]
3. Pour $x \in \mathbb{R}$, on pose $g(x) = e^{-|x|}$. Montrer que

$$\frac{2}{1+u^2} = \int_{\mathbb{R}} e^{iux} g(x) dx = \sqrt{2\pi} \hat{g}(u) \text{ pour tout } u \in \mathbb{R}.$$

En déduire que la fonction caractéristique de X , notée φ_X vérifie $\varphi_X(u) = e^{-|u|}$ pour tout $u \in \mathbb{R}$. [On pourra utiliser de théorème d'inversion de Fourier qui donne $\hat{\hat{g}} = g(-\cdot)$ p.p. si $g, \hat{g} \in L^1_{\mathbb{C}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$.]

On se donne maintenant une suite $X_1, X_2, \dots, X_n, \dots$ de v.a.r.i.i.d. et on suppose que la loi de X_1 (et donc de tous les X_n) est la loi de Cauchy. Pour $n \in \mathbb{N}^*$, on pose :

$$Z_n = \frac{1}{n} \left(\sum_{p=1}^n X_p \right).$$

4. Soit $n \in \mathbb{N}^*$.

(a) Montrer que

$$\varphi_{Z_n}(u) = \varphi_{X_1}^n \left(\frac{u}{n} \right) \text{ pour tout } u \in \mathbb{R}.$$

(b) Donner la loi de Z_n .

5. Pour $n \in \mathbb{N}^*$, on pose $A_n = \{|X_n| > n\}$ et $B_n = \cup_{p \geq n} A_p$. on pose aussi $B = \cap_{n \in \mathbb{N}^*} B_n$.

(a) Soit $n \in \mathbb{N}^*$. En utilisant la question 2, montrer que

$$P(B_n^c) = \prod_{p \geq n} P(A_p^c) \leq \prod_{p \geq n} \left(1 - \frac{1}{p\pi} \right).$$

En déduire que $P(B_n^c) = 0$ et donc que $P(B_n) = 1$.

- (b) Montrer que $P(B) = 1$. [Cette question est une conséquence de la question (a) mais elle peut aussi être faite directement en appliquant le lemme de Borel-Cantelli et la question 2.]

(c) Soit $\omega \in B$. Montrer que $\frac{X_n(\omega)}{n} \not\rightarrow 0$ quand $n \rightarrow \infty$.

(d) Soit $\omega \in \Omega$. Montrer que si la suite $(Z_n(\omega))_{n \in \mathbb{N}^*}$ converge vers une limite finie on a alors

$$\lim_{n \rightarrow \infty} \frac{X_n(\omega)}{n} = 0.$$

[Ecrire X_n en fonction de Z_n et Z_{n-1} .]

(e) Dédire des trois questions précédentes que la suite $(Z_n)_{n \in \mathbb{N}^*}$ ne converge pas p.s. vers une limite finie quand $n \rightarrow \infty$.

6. Pour tout $n \in \mathbb{N}^*$, montrer que

$$Z_{2n} - Z_n = \frac{U_n + V_n}{2},$$

où U_n et V_n sont deux v.a.r. indépendantes, de loi de Cauchy.

En déduire que la suite $(Z_n)_{n \in \mathbb{N}^*}$ ne converge pas en probabilité.

7. Pour quelles raisons ne peut-on appliquer, à la suite $(Z_n)_{n \in \mathbb{N}^*}$, les lois forte et faible des grands nombres ?

~~corrigé~~

En attente

Chapitre 22

Espérance conditionnelle et martingales

22.1 Espérance conditionnelle

Corrigé 187 (Espérance conditionnelle selon une tribu)

Soit (Ω, \mathcal{A}, P) un espace probabilisé et Y une variable aléatoire réelle intégrable. Dans les trois cas suivants, montrer que $E[Y|\mathcal{B}]$ est réduit à un élément et déterminer $E[Y|\mathcal{B}]$ (en fonction de Y et \mathcal{B}).

1. La tribu \mathcal{B} la tribu grossière, c'est-à-dire $\mathcal{B} = \{\emptyset, \Omega\}$.

~~corrigé~~

Soit Z une application de Ω dans \mathbb{R} , \mathcal{B} -mesurable. Soit $a \in \text{Im}(Z)$ (on suppose, bien sûr, $\Omega \neq \emptyset$). On a alors $\{Z = a\} = \{\omega \in \Omega; Z(\omega) = a\} \neq \emptyset$. Comme Z est \mathcal{B} -mesurable, on a donc $\{Z = a\} = \Omega$. Une application \mathcal{B} -mesurable est donc une fonction constante (réciproquement, une fonction constante est bien \mathcal{B} -mesurable). Si $Z \in E(Y|\mathcal{B})$, il existe donc $a \in \mathbb{R}$ t.q. $Z(\omega) = a$ pour tout $\omega \in \Omega$. Le réel a doit alors vérifier $E(aU) = E(UY)$ pour toute application U , \mathcal{B} -mesurable de Ω dans \mathbb{R} . On a donc $ab = E(ab) = E(bY) = bE(Y)$ pour tout $b \in \mathbb{R}$. La seule solution est donc $a = E(Y)$. L'ensemble $E[Y|\mathcal{B}]$ est donc réduit à un seul élément, la fonction constante et égale à $E(Y)$.

2. Soit $B \in \mathcal{A}$ t.q. $0 < P[B] < 1$. On prend pour \mathcal{B} la tribu engendrée par B .

~~corrigé~~

Soit Z une application de Ω dans \mathbb{R} , \mathcal{B} -mesurable. Les parties B et B^c sont non vides (car de probabilité strictement positive). Soit $\omega_1 \in B$ et $a = Z(\omega_1)$. On a alors $\{Z = a\} = \{\omega \in \Omega; Z(\omega) = a\} \neq \emptyset$. Comme Z est \mathcal{B} -mesurable, on a donc $\{Z = a\} = B$ ou Ω et donc $\{Z = a\} \supset B$. De même, soit $\omega_2 \in B^c$ et $b = Z(\omega_2)$, on a $\{Z = b\} \supset B^c$. Une application \mathcal{B} -mesurable est donc une fonction constante sur B et B^c . Réciproquement, une fonction constante sur B et B^c est bien \mathcal{B} -mesurable.

Si $Z \in E(Y|\mathcal{B})$, il existe donc $a, b \in \mathbb{R}$ t.q. $Z = a1_B + b1_{B^c}$. Les réels a, b doivent alors vérifier $E(ZU) = E(UY)$ pour toute application U \mathcal{B} -mesurable de Ω dans \mathbb{R} , c'est-à-dire :

$$a\alpha P(B) + b\beta P(B^c) = \alpha \int_B Y dP + \beta \int_{B^c} Y dP \text{ pour tout } \alpha, \beta \in \mathbb{R}.$$

Comme $P(B) > 0$ et $P(B^c) = 1 - P(B) > 0$, la seule solution est donc :

$$a = \frac{\int_B Y dP}{P(B)} \text{ et } b = \frac{\int_{B^c} Y dP}{P(B^c)}.$$

L'ensemble $E[Y|\mathcal{B}]$ est donc réduit à un seul élément, la fonction Z définie par

$$Z = \frac{\int_B Y dP}{P(B)} 1_B + \frac{\int_{B^c} Y dP}{P(B^c)} 1_{B^c}.$$

3. Soit $(B_n)_{n \in \mathbb{N}^*} \subset \mathcal{A}$ t.q. $B_n \cap B_m = \emptyset$ si $n \neq m$, $\Omega = \cup_{n \in \mathbb{N}^*} B_n$ et $0 < P(B_n) < 1$ pour tout $n \in \mathbb{N}^*$. On prend pour \mathcal{B} la tribu engendrée par $(B_n)_{n \in \mathbb{N}^*}$ (c'est-à-dire $\mathcal{B} = \{\cup_{n \in J} B_n, J \subset \mathbb{N}^*\}$).

corrigé

On reprend le même raisonnement que dans les deux questions précédentes. On remarque d'abord qu'une application Z de Ω dans \mathbb{R} est \mathcal{B} -mesurable si et seulement si il existe une suite $(\alpha_n)_{n \in \mathbb{N}^*} \subset \mathbb{R}$ t.q. $Z = \sum_{n \in \mathbb{N}^*} \alpha_n 1_{B_n}$. (Cette Série est bien convergente en tout point de Ω car les B_n sont disjoints deux à deux.

Si $Z \in E(Y|\mathcal{B})$, il existe donc $(a_n)_{n \in \mathbb{N}^*} \subset \mathbb{R}$ t.q. $Z = \sum_{n \in \mathbb{N}^*} a_n 1_{B_n}$. La suite $(a_n)_{n \in \mathbb{N}^*}$ doit alors être telle que Z soit intégrable et que $E(ZU) = E(UY)$ pour toute application U \mathcal{B} -mesurable bornée de Ω dans \mathbb{R} , c'est-à-dire t.q.

$$\sum_{n \in \mathbb{N}^*} |a_n| P(B_n) < +\infty \text{ et } \sum_{n \in \mathbb{N}^*} \alpha_n a_n P(B_n) = \sum_{n \in \mathbb{N}^*} \alpha_n \int_{B_n} Y dP,$$

pour toute suite bornée $(\alpha_n)_{n \in \mathbb{N}^*} \subset \mathbb{R}$. Comme $P(B_n) > 0$, la seule solution est donc :

$$a_n = \frac{\int_{B_n} Y dP}{P(B_n)} \text{ pour tout } n \in \mathbb{N}^*.$$

Comme on sait que l'ensemble $E[Y|\mathcal{B}]$ est non vide, il est inutile de vérifier que la fonction Z trouvée est intégrable (puisque cette fonction est la seule fonction pouvant appartenir à $E[Y|\mathcal{B}]$). L'ensemble $E[Y|\mathcal{B}]$ est donc réduit à un seul élément. Cet élément est la fonction Z définie par $Z = \sum_{n \in \mathbb{N}^*} \frac{\int_{B_n} Y dP}{P(B_n)} 1_{B_n}$.

Corrigé 188 (Espérance conditionnelle selon une v.a.r.)

Soit (Ω, \mathcal{A}, P) un espace probabilisé, X une variable aléatoire réelle et Y une variable aléatoire réelle intégrable.

1. On suppose qu'il existe $a \in \mathbb{R}$ t.q. $X = a$ p.s.. Donner un élément de $E[Y|X]$.

corrigé

On utilise la proposition 11.5. Soit $Z \in E[Y|X]$, il existe alors ψ , fonction borélienne de \mathbb{R} dans \mathbb{R} , t.q. $Z = \psi(X)$, $\psi(X)$ est intégrable et, pour toute application φ de \mathbb{R} dans \mathbb{R} , borélienne bornée,

$$E(\psi(X)\varphi(X)) = E(Y\varphi(X)).$$

On a donc $Z = \psi(a)$ p.s. et en prenant pour φ une fonction t.q. $\varphi(a) = 1$ dans l'égalité précédente, on obtient $\psi(a) = E(Y)$. On a donc finalement $Z = E(Y)$ p.s.. La fonction constante et égale à $E(Y)$ est un élément de $E[Y|X]$. Plus précisément, la fonction $\psi(X)$ est un élément de $E[Y|X]$, dès que ψ est une fonction borélienne de \mathbb{R} dans \mathbb{R} et t.q. $\psi(a) = E(Y)$.

2. On suppose que X prend p.s. deux valeurs x_1 ou x_2 avec $x_1 \neq x_2$. Donner un élément de $E[Y|X]$.

corrigé

On pose $A_1 = \{X = x_1\}$ et $A_2 = \{X = x_2\}$. On suppose que $P(A_1) > 0$ et $P(A_2) > 0$ (sinon, on est ramené à la question précédente). Noter que $A_1 \cap A_2 = \emptyset$ et $P(A_1) + P(A_2) = 1$. On utilise encore la proposition 11.5.

Soit $Z \in E[Y|X]$, il existe alors ψ , fonction borélienne de \mathbb{R} dans \mathbb{R} , t.q. $Z = \psi(X)$, $\psi(X)$ est intégrable et, pour toute application φ de \mathbb{R} dans \mathbb{R} , borélienne bornée,

$$E(\psi(X)\varphi(X)) = E(Y\varphi(X)). \quad (22.1)$$

On a donc $Z = \psi(x_1)1_{A_1} + \psi(x_2)1_{A_2}$ p.s.. Soit $\alpha_1, \alpha_2 \in \mathbb{R}$, en prenant pour φ une fonction borélienne bornée de \mathbb{R} dans \mathbb{R} t.q. $\varphi(x_1) = \alpha_1$ et $\varphi(x_2) = \alpha_2$ dans l'égalité (22.1), on obtient :

$$\psi(x_1)\alpha_1 P(A_1) + \psi(x_2)\alpha_2 P(A_2) = \alpha_1 \int_{A_1} Y dP + \alpha_2 \int_{A_2} Y dP,$$

pour tout $\alpha_1, \alpha_2 \in \mathbb{R}$. Comme $P(A_i) > 0$ pour $i = 1, 2$, on en déduit que $\psi(x_i) = \frac{\int_{A_i} Y dP}{P(A_i)}$, pour $i = 1, 2$, et donc que

$$Z = \frac{\int_{A_1} Y dP}{P(A_1)} 1_{A_1} + \frac{\int_{A_2} Y dP}{P(A_2)} 1_{A_2} \text{ p.s..}$$

Ici encore, la fonction $\psi(X)$ est un élément de $E[Y|X]$ dès que ψ est une fonction borélienne de \mathbb{R} dans \mathbb{R} et t.q. $\psi(x_i) = \frac{\int_{A_i} Y dP}{P(A_i)}$ pour $i = 1, 2$ (un exemple possible est donc $\psi(x_i) = \frac{\int_{A_i} Y dP}{P(A_i)}$ pour $i = 1, 2$ et $\psi(x) = 0$ pour $x \notin \{x_1, x_2\}$).

3. On suppose que X est une v.a. prenant p.s. ses valeurs dans un ensemble dénombrable $\{x_n, n \in \mathbb{N}^*\}$ avec $P(X = x_n) \neq 0$ pour tout $n \in \mathbb{N}^*$. Donner un élément de $E[Y|X]$.

corrigé

On peut supposer que les x_n sont différents deux à deux. Pour $n \in \mathbb{N}^*$, on pose $A_n = \{X = x_n\}$. Les ensembles A_n sont disjoints deux à deux, $P(A_n) > 0$ pour tout $n \in \mathbb{N}^*$ et $\sum_{n=1}^{\infty} P(A_n) = 1$.

On utilise encore la proposition 11.5. Soit $Z \in E[Y|X]$ (Noter que, comme on sait que $E[Y|X]$ est non vide, il existe $Z \in E[Y|X]$). Il existe alors ψ , fonction borélienne de \mathbb{R} dans \mathbb{R} , t.q. $Z = \psi(X)$, $\psi(X)$ est intégrable et, pour toute application φ de \mathbb{R} dans \mathbb{R} , borélienne bornée,

$$E(\psi(X)\varphi(X)) = E(Y\varphi(X)). \quad (22.2)$$

On a donc $Z = \sum_{n \in \mathbb{N}^*} \psi(x_n)1_{A_n}$ p.s.. Soit $(\alpha_n)_{n \in \mathbb{N}^*} \subset \mathbb{R}$ une suite bornée de \mathbb{R} . Dans l'égalité (22.2), on prend pour φ une fonction borélienne bornée de \mathbb{R} dans \mathbb{R} t.q. $\varphi(x_n) = \alpha_n$ pour tout $n \in \mathbb{N}^*$ (un tel φ existe, on peut prendre, par exemple, $\varphi(x) = 0$ si x est différent de tous les x_n), on obtient :

$$\sum_{n \in \mathbb{N}^*} \psi(x_n)\alpha_n P(A_n) = \sum_{n \in \mathbb{N}^*} \alpha_n \int_{A_n} Y dP,$$

pour toute suite bornée $(\alpha_n)_{n \in \mathbb{N}^*} \subset \mathbb{R}$. Comme $P(A_n) > 0$ pour tout $n \in \mathbb{N}^*$, on en déduit que $\psi(x_n) = \frac{\int_{A_n} Y dP}{P(A_n)}$, pour tout $n \in \mathbb{N}^*$, et donc que

$$Z = \sum_{n \in \mathbb{N}^*} \frac{\int_{A_n} Y dP}{P(A_n)} 1_{A_n} \text{ p.s..}$$

Enfin, ici encore, la fonction $\psi(X)$ est un élément de $E[Y|X]$ dès que ψ est une fonction borélienne de \mathbb{R} dans \mathbb{R} et t.q. $\psi(x_n) = \frac{\int_{A_n} Y dP}{P(A_n)}$ pour tout $n \in \mathbb{N}^*$. Un exemple possible est donc $\psi(x_n) = \frac{\int_{A_n} Y dP}{P(A_n)}$ pour tout $n \in \mathbb{N}^*$ et $\psi(x) = 0$ pour $x \notin \{x_n, n \in \mathbb{N}^*\}$. Cet exemple donne la fonction $\sum_{n \in \mathbb{N}^*} \frac{\int_{A_n} Y dP}{P(A_n)} 1_{A_n}$.

Corrigé 189 (Calcul de $E(\exp(XY)|X)$ si $Y \sim \mathcal{N}(0, 1)$)

Soit (Ω, \mathcal{A}, P) un espace de probabilités et X, Y deux v.a. réelles indépendantes. On suppose que Y suit une loi gaussienne centrée réduite et que $E[\exp(X^2/2)] < \infty$. Montrer que $\exp(XY)$ est intégrable et déterminer $E[\exp(XY)|X]$.

corrigé

La v.a.r. $\exp(XY)$ est positive. On calcule $\int_{\Omega} e^{XY} dP$ en utilisant l'indépendance de X et Y (et le théorème 9.2, qui donne que $P_{(X,Y)} = P_X \otimes P_Y$) et le fait que $Y \sim \mathcal{N}(0, 1)$:

$$\int_{\Omega} e^{XY} dP = \int_{\mathbb{R}} \int_{\mathbb{R}} e^{xy} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy dP_X(x).$$

En remarquant que $xy - \frac{y^2}{2} = -\frac{1}{2}(x-y)^2 + \frac{1}{2}x^2$ on obtient :

$$\begin{aligned} \int_{\Omega} e^{XY} dP &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-\frac{(x-y)^2}{2}} dy \right) e^{\frac{x^2}{2}} dP_X(x) = \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-\frac{z^2}{2}} dz \right) e^{\frac{x^2}{2}} dP_X(x), \end{aligned}$$

et donc $\int_{\Omega} e^{XY} dP = \int_{\mathbb{R}} e^{\frac{x^2}{2}} dP_X(x) = E(e^{\frac{X^2}{2}}) < +\infty$. Ce qui donne que e^{XY} est une v.a.r. intégrable.

Selon la proposition 11.5 on cherche un élément de $E[e^{XY}|X]$ sous la forme $\psi(X)$ où ψ est une fonction borélienne de \mathbb{R} dans \mathbb{R} , t.q. $Z = \psi(X)$, $\psi(X)$ est intégrable et, pour toute application φ de \mathbb{R} dans \mathbb{R} , borélienne bornée,

$$E(\psi(X)\varphi(X)) = E(e^{XY}\varphi(X)).$$

Soit φ une application borélienne bornée de \mathbb{R} dans \mathbb{R} , on calcule $E(e^{XY}\varphi(X))$ en utilisant, comme précédemment, l'indépendance de X et Y et le fait que $Y \sim \mathcal{N}(0, 1)$:

$$\begin{aligned} E(e^{XY}\varphi(X)) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{xy}\varphi(x) e^{-\frac{y^2}{2}} dy dP_X(x) = \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-\frac{(x-y)^2}{2}} dy \right) e^{\frac{x^2}{2}} \varphi(x) dP_X(x). \end{aligned}$$

On a donc $E(e^{XY}\varphi(X)) = \int_{\mathbb{R}} e^{\frac{x^2}{2}} \varphi(x) dP_X(x) = E(e^{-\frac{X^2}{2}} \varphi(X))$. Ceci nous montre que $e^{-\frac{X^2}{2}}$ est un élément de $E[e^{XY}|X]$ et donc (comme on confond $E[e^{XY}|X]$ avec l'un de des éléments) $E[e^{XY}|X] = e^{-\frac{X^2}{2}}$ p.s..

Corrigé 190 (Espérance du produit et produit des espérances)

Soit (Ω, \mathcal{A}, P) un espace de probabilités et X, Y deux v.a. intégrables t.q. XY est intégrable et $E[X|Y] = E(X)$ p.s.. Montrer que $E(XY) = E(X)E(Y)$.

corrigé

Grâce à la proposition 11.5 on a, pour toute application φ de \mathbb{R} dans \mathbb{R} , borélienne bornée,

$$E(E[X|Y]\varphi(Y)) = E(X\varphi(Y))$$

Comme $E[X|Y] = E(X)$ p.s., on en déduit, pour toute application φ de \mathbb{R} dans \mathbb{R} , borélienne bornée,

$$E(X)E(\varphi(Y)) = E(X\varphi(Y)). \tag{22.3}$$

Soit $n \in \mathbb{N}^*$. Pour $s \in \mathbb{R}$, on pose $T_n(s) = \max\{-n, \min\{s, n\}\}$. La fonction T_n est borélienne (car continue) bornée (par n) de \mathbb{R} dans \mathbb{R} . On peut donc utiliser (22.3) avec $\varphi = T_n$. On obtient $E(X)E(T_n(Y)) = E(XT_n(Y))$.

Comme Y est intégrable, on a, par convergence dominée, $\lim_{n \rightarrow \infty} E(T_n(Y)) = E(Y)$ (noter que $|T_n(Y)| \leq |Y|$). Comme XY est intégrable (et c'est uniquement ici que cette hypothèse est utilisée), on a, par convergence dominée,

$$\lim_{n \rightarrow \infty} E(XT_n(Y)) = E(XY)$$

(noter que $|XT_n(Y)| \leq |XY|$).

En passant à limite quand $n \rightarrow \infty$ sur l'égalité $E(X)E(T_n(Y)) = E(XT_n(Y))$, on a donc $E(X)E(Y) = E(XY)$.

Corrigé 191 (Lorsque $E(Y|X) = X$ p.s....)

Soit (Ω, \mathcal{A}, P) un espace probabilisé et X, Y deux v.a.r. de carré intégrable.

On suppose que $E(Y|X) = X$ p.s..

1.(a) Soit φ une fonction borélienne de \mathbb{R} dans \mathbb{R} . On suppose que la v.a.r. $\varphi(X)$ est de carré intégrable. Montrer que $\int_{\Omega} Y\varphi(X)dP = \int_{\Omega} X\varphi(X)dP$.

(b) Montrer que $E(Y) = E(X)$ et $E(XY) = E(X^2)$.

2.(a) Montrer que $E(X^2) \leq E(Y^2)$.

(b) Montrer que $Y = X$ p.s. si et seulement si $E(Y^2) = E(X^2)$.

corrigé

En attente

Corrigé 192 (V.a. gaussien et espérance conditionnelle)

Soit (Ω, \mathcal{A}, P) un espace probabilisé et $(X, Y)^t$ un v.a. gaussien de dimension 2, On note a l'espérance de X , b l'espérance de Y et D la matrice de covariance du v.a. $(X, Y)^t$ (on a donc $D_{1,1} = \text{Var}(X)$, $D_{2,2} = \text{Var}(Y)$ et $D_{1,2} = D_{2,1} = \text{Cov}(X, Y)$). On suppose que $\text{Var}(Y) > 0$.

1. Calculer $\alpha, \beta \in \mathbb{R}$ (en fonction de a, b et D) de manière à avoir $E[X] = E[\alpha + \beta Y]$ et $E[XY] = E[(\alpha + \beta Y)Y]$.

corrigé

On a $E[X] = a$, $E[Y] = b$, $E[Y^2] = \text{Var}(Y) + E[Y]^2 = D_{2,2} + b^2$ et

$$E[XY] = E[(X - E(X))(Y - E(Y))] + E[X]E[Y] = D_{1,2} + ab.$$

On cherche donc α et β t.q.

$$\begin{aligned} \alpha + b\beta &= a, \\ b\alpha + (D_{2,2} + b^2)\beta &= D_{1,2} + ab. \end{aligned}$$

Comme $D_{2,2} \neq 0$, ce système de 2 équations à 2 inconnues (qui sont α et β) a bien une unique solution. Cette solution est

$$\beta = \frac{D_{1,2}}{D_{2,2}}, \quad \alpha = a - b \frac{D_{1,2}}{D_{2,2}}.$$

Avec α et β ainsi déterminés, on définit la fonction affine l de \mathbb{R} dans \mathbb{R} par $l(s) = \alpha + \beta s$ pour $s \in \mathbb{R}$ et on définit la v.a.r. Z par $Z = X - l(Y)$.

2. Montrer que $(Z, Y)^t$ est un v.a. gaussien. Montrer que $\text{Cov}(Z, Y) = 0$. En déduire que Z et Y sont des v.a.r. indépendantes [On pourra utiliser la question 1(b) de l'exercice 10.10].

corrigé

Soit $a_1, a_2 \in \mathbb{R}$. On va montrer que la v.a.r. $a_1Z + a_2Y$ suit une loi gaussienne (ceci montre bien que le vecteur aléatoire $(Z, Y)^t$ est un vecteur gaussien). Comme $Z = X - l(Y) = X - \alpha - \beta Y$, on a

$$a_1Z + a_2Y = a_1X + (a_2 - a_1\beta)Y - a_1\alpha.$$

Comme $(X, Y)^t$ est un v.a. gaussien, la v.a.r. $a_1X + (a_2 - a_1\beta)Y$ suit une loi gaussienne. Il existe donc $m \in \mathbb{R}$ et $\sigma \in \mathbb{R}_+$ t.q. $a_1X + (a_2 - a_1\beta)Y \sim \mathcal{N}(m, \sigma^2)$. On a alors

$$a_1X + (a_2 - a_1\beta)Y - a_1\alpha \sim \mathcal{N}(m - a_1\alpha, \sigma^2).$$

ceci prouve que $a_1X + (a_2 - a_1\beta)Y - a_1\alpha$ suit une loi gaussienne. La v.a.r. $a_1Z + a_2Y$ suit donc une loi gaussienne pour tout $a_1, a_2 \in \mathbb{R}$. Ceci prouve bien que le vecteur aléatoire $(Z, Y)^t$ est un vecteur gaussien.

On calcule maintenant $\text{Cov}(Z, Y)$. On rappelle que $Z = X - (\alpha + \beta Y)$. On remarque d'abord que $E[Z] = E[X] - E[\alpha + \beta Y] = 0$ (grâce à la première relation satisfaite par α et β). Puis,

$$\text{Cov}(Z, Y) = E[ZY] = E[XY] - E[(\alpha + \beta Y)Y].$$

La seconde relation satisfaite par α et β donne alors $\text{Cov}(Z, Y) = 0$.

Comme $(Z, Y)^t$ est un vecteur gaussien et que $\text{Cov}(Z, Y) = 0$, la question 1(b) de l'exercice 10.10 donne que Z et Y sont indépendantes.

3. Montrer que $E[X|Y] = l(Y)$ p.s.

corrigé

La v.a.r. $l(Y)$ est intégrable (car sa loi est gaussienne). Pour montrer que $E[X|Y] = l(Y)$ p.s., il suffit, d'après la proposition 11.5, de montrer que $E(l(Y)\varphi(Y)) = E(X\varphi(Y))$ pour toute application φ de \mathbb{R} dans \mathbb{R} , borélienne bornée.

Soit donc φ de \mathbb{R} dans \mathbb{R} , borélienne bornée. Comme Z et Y sont indépendantes et que $E[Z] = 0$, on a $E[Z\varphi(Y)] = E[Z]E[\varphi(Y)] = 0$. Comme $Z = X - l(Y)$ on a donc

$$E[X\varphi(Y)] = E[l(Y)\varphi(Y)].$$

Ce qui prouve bien que $E[X|Y] = l(Y)$ p.s.

4. Calculer (en fonction de D) $\text{Var}(Z)$.

corrigé

Comme $E[Z] = 0$ et $E[Zl(Y)] = E[Z]E[l(Y)] = 0$, on a

$$\begin{aligned} \text{Var}(Z) &= E[Z^2] = E[XZ] = E[X^2] - E[Xl(Y)] \\ &= E[X^2] - \alpha E[X] - \beta E[XY]. \end{aligned}$$

Il suffit maintenant d'utiliser les valeurs de α et β et le fait que

$$E[X^2] = D_{1,1} + a^2, \quad E[X] = a, \quad E[X, Y] = D_{1,2} + ab.$$

On obtient

$$\text{Var}(Z) = D_{1,1} - \frac{D_{1,2}^2}{D_{2,2}}$$

Dans la suite, on note $\sigma = \sqrt{\text{Var}(Z)}$ et, pour $a \in \mathbb{R}$ on note μ_a la loi normale $\mathcal{N}(a, \sigma^2)$.

5. Soit f une fonction borélienne bornée de \mathbb{R} dans \mathbb{R} . Pour $a \in \mathbb{R}$, on pose $\psi(a) = \int_{\mathbb{R}} f d\mu_a$.

(a) Montrer que ψ est une fonction continue (de \mathbb{R} dans \mathbb{R}) si $\sigma > 0$.

corrigé

On utilise ici le théorème 4.9. On remarque que

$$\psi(a) = \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} f(x) e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}\sigma} \int_{\mathbb{R}} F(x, a) dx,$$

avec $F(x, a) = f(x) e^{-\frac{(x-a)^2}{2\sigma^2}}$. La fonction $a \mapsto F(x, a)$ est continue, pour tout $x \in \mathbb{R}$. Pour montrer que ψ est continue, il suffit (d'après le théorème 4.9) de montrer que la fonction $x \mapsto F(x, a)$ est dominée, localement uniformément par rapport à a , par une fonction intégrable. Nous montrons maintenant cette domination.

Soit $M > 0$. Pour tout $a \in]-M, M[$ et tout $x \in \mathbb{R}$, on a $|F(x, a)| \leq g_M(x)$ avec

$$g_M(x) = |f(x)| e^{-\frac{((|x|-M)^+)^2}{2\sigma^2}}.$$

(La vérification de $|F(x, a)| \leq g_M(x)$ peut se faire en distinguant les cas $x \in [-M, M]$, $x > M$ et $x < -M$.) La fonction g_M est bien intégrable (pour la mesure de Lebesgue). Le théorème 4.9 donne alors que ψ est continue sur $] - M, M[$. Comme M est arbitraire, on en déduit que ψ est continue sur \mathbb{R} .

(b) Montrer que $E[f(X)|Y] = \psi(l(Y))$ p.s.

corrigé

Si $\sigma > 0$, la fonction ψ est continue, elle est donc borélienne. On remarque aussi que ψ est bornée (en effet, soit $M \in \mathbb{R}_+$ t.q. $|f(x)| \leq M$ pour tout $x \in \mathbb{R}$. On a alors aussi $|\psi(a)| \leq M$ pour tout $a \in \mathbb{R}$).

Si $\sigma = 0$, on a $\psi(a) = f(a)$ pour tout $a \in \mathbb{R}$. La fonction ψ est donc aussi borélienne bornée.

Comme ψ est borélienne, $\psi(l(Y))$ est donc une v.a.r.. Comme ψ est bornée, la v.a.r. $\psi(l(Y))$ est bornée et donc intégrable. Pour montrer que $E[f(X)|Y] = \psi(l(Y))$ p.s., il suffit, comme à la question précédente (cf proposition 11.5), de montrer que $E[\psi(l(Y))\varphi(Y)] = E[f(X)\varphi(Y)]$ pour toute application φ de \mathbb{R} dans \mathbb{R} , borélienne bornée.

Soit donc φ borélienne bornée de \mathbb{R} dans \mathbb{R} . On a, comme Z et Y sont indépendantes,

$$\begin{aligned} E[f(X)\varphi(Y)] &= E[f(Z + l(Y))\varphi(Y)] \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(z + l(y)) dP_Z(z) \right) \varphi(y) dP_Y(y). \end{aligned}$$

On utilise maintenant le fait que $Z \sim \mathcal{N}(0, \sigma^2)$, on en déduit

$$\int_{\mathbb{R}} f(z + l(y)) dP_Z(z) = \int_{\mathbb{R}} f(z + l(y)) d\mu_0(z).$$

le changement de variable $\bar{z} = z + l(y)$ donne alors

$$\int_{\mathbb{R}} f(z + l(y)) d\mu_0(z) = \int_{\mathbb{R}} f(\bar{z}) d\mu_{l(y)}(\bar{z}) = \psi(l(y)).$$

On a donc, finalement,

$$E[f(X)\varphi(Y)] = \int_{\mathbb{R}} \psi(l(y))\varphi(y) dP_Y(y) = E[\psi(l(Y))\varphi(Y)].$$

Ce qui prouve bien $E[f(X)|Y] = \psi(l(Y))$ p.s..

22.2 Martingales

Corrigé 193 (Quelques propriétés des martingales)

Soit (Ω, \mathcal{A}, P) un espace de probabilité muni d'une filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$ (c'est-à-dire d'une suite croissante de sous tribus de \mathcal{A}) et $(X_n)_{n \in \mathbb{N}}$ une suite de de v.a.r. (c'est-à-dire un processus réel). On suppose que X_n est intégrable pour tout $n \in \mathbb{N}$.

1. On suppose que $(X_n)_{n \in \mathbb{N}}$ est une sous-martingale (par rapport à la filtration $(\mathcal{B}_n)_n$). Montrer que la suite $(E(X_n))_{n \in \mathbb{N}}$ est croissante.

—————**corrigé**—————

Soit $n \in \mathbb{N}$. Comme $(X_n)_{n \in \mathbb{N}}$ est une sous-martingale, on a $E(X_{n+1}|\mathcal{B}_n) \geq X_n$ p.s. et donc

$$E(E(X_{n+1}|\mathcal{B}_n)) \geq E(X_n).$$

Or (comme les fonctions constantes sont \mathcal{B}_n -mesurables bornées),

$$E(E(X_{n+1}|\mathcal{B}_n)) = E(X_{n+1}).$$

On a donc $E(X_{n+1}) \geq E(X_n)$.

2. On suppose que $(X_n)_{n \in \mathbb{N}}$ est une martingale (par rapport à la filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$). Montrer que $E(X_{n+m}|\mathcal{B}_n) = X_n$ p.s. pour tout $m \geq 0$.

—————**corrigé**—————

Pour $m = 0$, le fait que $E(X_n|\mathcal{B}_n) = X_n$ p.s., pour tout $n \in \mathbb{N}$, découle du fait que X_n est \mathcal{B}_n -mesurable. On montre maintenant la propriété demandée par récurrence sur $m \in \mathbb{N}^*$.

Pour $m = 1$ le fait que $E(X_{n+1}|\mathcal{B}_n) = X_n$ p.s., pour tout $n \in \mathbb{N}$, est donné dans la définition de martingale.

Soit $m \in \mathbb{N}^*$. On suppose que $E(X_{n+m}|\mathcal{B}_n) = X_n$ p.s., pour tout $n \in \mathbb{N}$. On veut montrer que $E(X_{n+m+1}|\mathcal{B}_n) = X_n$ p.s., pour tout $n \in \mathbb{N}$. Soit $n \in \mathbb{N}$. Comme $(X_n)_{n \in \mathbb{N}}$ est une martingale, on a $E(X_{n+m+1}|\mathcal{B}_{n+m}) = X_{n+m}$ p.s. et donc :

$$E(X_{n+m+1}U) = E(X_{n+m}U) \text{ pour tout } U \text{ } \mathcal{B}_{n+m}\text{-mesurable bornée.}$$

Comme $\mathcal{B}_n \subset \mathcal{B}_{n+m}$, on a donc aussi

$$E(X_{n+m+1}U) = E(X_{n+m}U) \text{ pour tout } U \text{ } \mathcal{B}_n\text{-mesurable bornée.}$$

L'hypothèse de récurrence donne $E(X_{n+m}|\mathcal{B}_n) = X_n$ p.s.. On a donc :

$$E(X_{n+m}U) = E(X_nU) \text{ pour tout } U \text{ } \mathcal{B}_n\text{-mesurable bornée.}$$

On a en déduit :

$$E(X_{n+m+1}U) = E(X_nU) \text{ pour tout } U \text{ } \mathcal{B}_n\text{-mesurable bornée.}$$

Ce qui montre que $E(X_{n+m+1}|\mathcal{B}_n) = X_n$ p.s. et termine la récurrence.

3. Soit φ une fonction convexe de \mathbb{R} dans \mathbb{R} . On suppose que $(X_n)_{n \in \mathbb{N}}$ est une martingale (par rapport à la filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$) et que $\varphi(X_n)$ est intégrable pour tout $n \in \mathbb{N}$ (on rappelle que $\varphi(X_n)$ est une notation pour désigner $\varphi \circ X_n$). Montrer que $(\varphi(X_n))_{n \in \mathbb{N}}$ est une sous-martingale (par rapport à la filtration $(\mathcal{B}_n)_{n \in \mathbb{N}}$).

—————**corrigé**—————

On remarque tout d'abord que $\varphi(X_n)$ est bien \mathcal{B}_n -mesurable (car X_n est \mathcal{B}_n -mesurable et φ est borélienne), pour tout $n \in \mathbb{N}$. Pour montrer que $(\varphi(X_n))_{n \in \mathbb{N}}$ est une sous-martingale, il suffit alors d'utiliser la proposition 11.4 sur l'inégalité de Jensen. Soit $n \in \mathbb{N}$. La proposition 11.4 donne

$$E(\varphi(X_{n+1})|\mathcal{B}_n) \geq \varphi(E(X_{n+1}|\mathcal{B}_n)) \text{ p.s..}$$

Comme $E(X_{n+1}|\mathcal{B}_n) = X_n$ p.s., on en déduit $E(\varphi(X_{n+1})|\mathcal{B}_n) \geq \varphi(X_n)$ p.s., ce qui montre bien que $(\varphi(X_n))_{n \in \mathbb{N}}$ est une sous-martingale.

That's all folks

Corrigé 122 (Sur les suites convergentes en mesure et bornées dans \mathcal{L}^p)

Soit (E, T, m) un espace mesuré fini. Pour $r \in [1, +\infty]$, on note \mathcal{L}^r l'espace $\mathcal{L}^r_{\mathbb{R}}(E, T, m)$. Soit $p \in]1, +\infty]$, $(f_n)_{n \in \mathbb{N}}$ une suite bornée de \mathcal{L}^p (c'est-à-dire $\sup_{n \in \mathbb{N}} \|f_n\|_p < +\infty$) et f une fonction mesurable X dans \mathbb{R} . On suppose que $f_n \rightarrow f$ en mesure quand $n \rightarrow \infty$.

1. Montrer que $f \in \mathcal{L}^p$. [On rappelle que, comme $f_n \rightarrow f$ en mesure, il existe une sous suite de la suite $(f_n)_{n \in \mathbb{N}}$ qui converge p.p. vers f .]

————— corrigé —————

On note $M = \sup_{n \in \mathbb{N}} \|f_n\|_p$. Il existe une sous suite de la suite $(f_n)_{n \in \mathbb{N}}$ qui converge p.p. vers f . On note $(f_{\varphi(n)})_{n \in \mathbb{N}}$ (avec φ strictement croissante de \mathbb{N} dans \mathbb{N}) une telle sous suite.

Si $p < +\infty$, on applique alors le lemme de Fatou à la suite $(g_n)_{n \in \mathbb{N}}$ avec $g_n = |f_{\varphi(n)}|^p$. Comme $g_n \rightarrow |f|^p$ p.p., on en déduit que

$$\int |f|^p dm \leq \liminf_{n \rightarrow \infty} \int g_n dm \leq M^p < +\infty.$$

Ce qui prouve que $f \in \mathcal{L}^p$. (Ce résultat est aussi vrai si $p = 1$.)

Si $p = \infty$, comme $|f_{\varphi(n)}| \leq M$ p.p., on en déduit que $|f| \leq M$ p.p. et donc que $f \in \mathcal{L}^\infty$.

2. Soit $q \in [1, p[$. Montrer que $\|f_n - f\|_q \rightarrow 0$ quand $n \rightarrow \infty$. [Pour $\varepsilon > 0$, on pourra utiliser, avec $A_{n,\varepsilon} = \{|f_n - f| \leq \varepsilon\}$, l'égalité suivante :

$$\int |f_n - f|^q dm = \int_{A_{n,\varepsilon}} |f_n - f|^q dm + \int_{A_{n,\varepsilon}^c} |f_n - f|^q dm.]$$

————— corrigé —————

Pour tout $\varepsilon > 0$ et $n \in \mathbb{N}$, on a

$$\int |f_n - f|^q dm = \int_{A_{n,\varepsilon}} |f_n - f|^q dm + \int_{A_{n,\varepsilon}^c} |f_n - f|^q dm \leq \varepsilon^q m(E) + \int |f_n - f|^q 1_{A_{n,\varepsilon}^c} dm. \quad (17.8)$$

On suppose tout d'abord $p < +\infty$. En appliquant l'inégalité de Hölder à la dernière intégrale (avec l'exposant p/q et son conjugué), on obtient

$$\int |f_n - f|^q dm \leq \varepsilon^q m(E) + \left(\int |f_n - f|^p dm \right)^{\frac{q}{p}} m(A_{n,\varepsilon}^c)^{1-\frac{q}{p}}.$$

Comme $\|f_n\|_p \leq M$ et $\|f\|_p \leq M$, on en déduit

$$\int |f_n - f|^q dm \leq \varepsilon^q m(E) + (2^{p+1} M^p)^{\frac{q}{p}} m(A_{n,\varepsilon}^c)^{1-\frac{q}{p}}.$$

Soit maintenant $\eta > 0$. On choisit tout d'abord $\varepsilon > 0$ t.q. $\varepsilon^q m(E) \leq \eta$. Puis, comme $f_n \rightarrow f$ en mesure, on a $\lim_{n \rightarrow \infty} m(A_{n,\varepsilon}^c) = 0$. Il existe donc $n_0 \in \mathbb{N}$ t.q.

$$n \geq n_0 \Rightarrow (2^{p+1} M^p)^{\frac{q}{p}} m(A_{n,\varepsilon}^c)^{1-\frac{q}{p}} \leq \eta$$

et donc

$$n \geq n_0 \Rightarrow \int |f_n - f|^q dm \leq 2\eta.$$

Ce qui montre bien que $\|f_n - f\|_q \rightarrow 0$ quand $n \rightarrow \infty$.

On suppose maintenant que $p = \infty$. Comme $|f_n| \leq M$ p.p. et $|f| \leq M$ p.p., l'inégalité 17.8 donne

$$\int |f_n - f|^q dm \leq \varepsilon^q m(E) + 2^q M^q m(A_{n,\varepsilon}^c).$$

On conclut alors, comme dans le cas $p < +\infty$, que $\|f_n - f\|_q \rightarrow 0$ quand $n \rightarrow \infty$.

3. (Question plus difficile.) Donner un exemple pour lequel $\|f_n - f\|_p \not\rightarrow 0$ quand $n \rightarrow \infty$.

corrigé

Pour cet exemple, on prend $(E, T, m) = (]0, 1[, \mathcal{B}(]0, 1[), \lambda)$ et $p = 2$. Pour $n \in \mathbb{N}^*$, on définit f_n par

$$f_n(x) = \sqrt{n} \text{ si } 0 < x < \frac{1}{n} \text{ et } f_n(x) = 0 \text{ si } \frac{1}{n} < x < 1.$$

Avec ce choix, la suite $(f_n)_{n \in \mathbb{N}^*}$ est bornée dans \mathcal{L}^2 , elle converge en mesure vers la fonction nulle mais $\|f_n\|_2 \not\rightarrow 0$ quand $n \rightarrow \infty$.

Corrigé 138 (Orthogonalité et v.a.r. indépendantes)

Soit (Ω, \mathcal{A}, P) un espace probabilisé et $(X_n)_{n \in \mathbb{N}}$ une suite de v.a.r. indépendantes. On suppose que X_n est de carré intégrable pour tout $n \in \mathbb{N}$ et que $(X_n/X_m)_2 = 0$ si $n \neq m$ (On note L^2 l'espace $L^2_{\mathbb{R}}(\Omega, \mathcal{A}, p)$ et $(\cdot/\cdot)_2$ le produit scalaire dans L^2). On suppose enfin que $\sum_{n=0}^{\infty} \text{Var}(X_n) < +\infty$.

1. Montrer qu'il existe au plus un entier $n \geq 0$ t.q. $E(X_n) \neq 0$.

corrigé

Soit $n \neq m$. Comme X_n et X_m sont deux v.a.r. indépendantes et intégrables, la proposition 4.10 (deuxième item) donne que $X_n X_m$ est intégrable et

$$E(X_n X_m) = E(X_n)E(X_m).$$

Comme $E(X_n X_m) = \int_{\Omega} X_n X_m dP = (X_n/X_m)_2 = 0$, on en déduit qu'il est impossible que $E(X_n)$ et $E(X_m)$ soient tous les deux non nuls. On en déduit bien qu'il existe au plus un entier $n \geq 0$ t.q. $E(X_n) \neq 0$.

2. Montrer que la série $\sum_{n=0}^{\infty} X_n$ est convergente dans L^2 .

corrigé

Pour tout $n \in \mathbb{N}$, on a $\text{Var}(X_n) = E(X_n^2) - E(X_n)^2$. Pour tout $n \in \mathbb{N}$ sauf éventuellement un seul, on a donc

$$\text{Var}(X_n) = E(X_n^2) = \int_{\Omega} X_n^2 dP = \|X_n\|_2^2.$$

La série $\sum_{n=0}^{+\infty} \|X_n\|_2^2$ est donc convergente. On en déduit que la série $\sum_{n=0}^{\infty} X_n$ est convergente dans L^2 , comme cela est démontré dans le premier exercice sur les espaces de Hilbert (corrigé 125).

Devoir 2, corrigé de l'exercice 7.4

Corrigé 159 (Fonction dont les traces sont boréliennes)

Pour $B \subset \mathbb{R}^2$, on note $t(B)$ l'ensemble des $x_1 \in \mathbb{R}$ t.q. $(x_1, 0) \in B$. On pose $T = \{B \subset \mathbb{R}^2; t(B) \in \mathcal{B}(\mathbb{R})\}$.

Soit $\theta \in]0, \frac{\pi}{2}[$. Pour $x = (x_1, x_2)^t \in \mathbb{R}^2$, on pose $g(x) = (x_1 \cos \theta - x_2 \sin \theta, x_1 \sin \theta + x_2 \cos \theta)^t$.

1. Montrer que T est une tribu contenant $\mathcal{B}(\mathbb{R}^2)$.

corrigé

On montre assez facilement que T est une tribu sur \mathbb{R}^2 . En effet, on peut remarquer, par exemple, que $\mathbb{R}^2 \in T$ (car $t(\mathbb{R}^2) = \mathbb{R} \in \mathcal{B}(\mathbb{R})$), T est stable par union dénombrable (si $B_n \in T$ pour tout $n \in \mathbb{N}$, on a $t(\cup_{n \in \mathbb{N}} B_n) = \cup_{n \in \mathbb{N}} t(B_n) \in \mathcal{B}(\mathbb{R})$ par stabilité de $\mathcal{B}(\mathbb{R})$ par union dénombrable) et T est stable par passage au complémentaire (si $B \in T$, on a $t(B^c) = t(B)^c \in \mathcal{B}(\mathbb{R})$ par stabilité de $\mathcal{B}(\mathbb{R})$ par passage au complémentaire).

Pour montrer que $T \supset \mathcal{B}(\mathbb{R}^2)$, on pose

$$\mathcal{C} = \{A_1 \times A_2, A_1, A_2 \in \mathcal{B}(\mathbb{R})\}.$$

On sait que \mathcal{C} engendre $\mathcal{B}(\mathbb{R}^2)$. On remarque maintenant que $\mathcal{C} \subset T$ (car si $A_1, A_2 \in \mathcal{B}(\mathbb{R})$, on a $t(A_1 \times A_2) = A_1$ ou \emptyset selon que $0 \in A_2$ ou $0 \notin A_2$). La tribu T contenant \mathcal{C} , elle contient nécessairement la tribu engendrée par \mathcal{C} , c'est-à-dire $\mathcal{B}(\mathbb{R}^2)$.

2. Soit $A \subset \mathbb{R}$ t.q. $A \notin \mathcal{B}(\mathbb{R})$. On pose $B = A \times \{0\}$.

(a) Montrer que $B \notin \mathcal{B}(\mathbb{R}^2)$.

corrigé

Comme $t(B) = A \notin \mathcal{B}(\mathbb{R})$, on a $B \notin T$ et donc $B \notin \mathcal{B}(\mathbb{R}^2)$ car $T \supset \mathcal{B}(\mathbb{R}^2)$.

(b) On pose $f = 1_B \circ g$. Montrer que la fonction f n'est pas une fonction borélienne de \mathbb{R}^2 dans \mathbb{R} mais que les fonctions $f(x_1, \cdot)$ et $f(\cdot, x_2)$ sont boréliennes de \mathbb{R} dans \mathbb{R} , pour tout $x_1, x_2 \in \mathbb{R}$.

corrigé

La fonction g est linéaire bijective de \mathbb{R}^2 dans \mathbb{R}^2 . On note h sa fonction réciproque. La fonction h est donc aussi linéaire. La fonction h est donc borélienne (car continue) et on a $1_B = f \circ h$. Comme $B \notin \mathcal{B}(\mathbb{R}^2)$, la fonction 1_B n'est pas borélienne. On en déduit que f n'est pas borélienne (si f était borélienne, on aurait 1_B borélienne car composée de fonctions boréliennes).

Soit $x_1 \in \mathbb{R}$. On va montrer que la fonction $f(x_1, \cdot)$ est borélienne. On pose $x_2 = -x_1 \frac{\sin \theta}{\cos \theta}$. On distingue alors deux cas possibles :

Premier cas. On suppose que $x_1 \cos \theta - x_2 \sin \theta \notin A$. La fonction $f(x_1, \cdot)$ est alors la fonction nulle (c'est-à-dire que $f(x_1, x) = 0$ pour tout $x \in \mathbb{R}$). La fonction $f(x_1, \cdot)$ est donc borélienne.

Second cas. On suppose que $x_1 \cos \theta - x_2 \sin \theta \in A$. La fonction $f(x_1, \cdot)$ est alors la fonction nulle partout sauf au point x_2 où elle vaut 1 (c'est-à-dire que $f(x_1, x) = 1_{\{x_2\}}(x)$ pour tout $x \in \mathbb{R}$). La fonction $f(x_1, \cdot)$ est donc borélienne.

De manière analogue on peut montrer que $f(\cdot, x_2)$ est borélienne pour tout $x_2 \in \mathbb{R}$.

Devoir 2, corrigé de l'exercice 9.5

Corrigé 182 (Exemple de loi pour le max de deux v.a.r.)

On considère deux variables aléatoires réelles indépendantes X et Y qui ont pour loi de probabilité conjointe la loi dans le plan \mathbb{R}^2 de densité $f(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$ par rapport à la mesure de Lebesgue de \mathbb{R}^2 , σ étant un nombre réel donné non nul. Quelle est la loi de probabilité de la variable aléatoire $Z = \max(|X|, |Y|)$?

corrigé

On peut supposer $\sigma > 0$. Soit φ une fonction de \mathbb{R} dans \mathbb{R} , borélienne bornée. On va chercher la loi de Z en exprimant convenablement $E(\varphi(Z))$ grâce à la connaissance de la loi du couple (X, Y) .

$$E(\varphi(Z)) = \int_{\mathbb{R}} \int_{\mathbb{R}} \varphi(\max(|x|, |y|)) \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) dx dy.$$

On en déduit facilement que

$$E(\varphi(Z)) = \frac{8}{2\pi\sigma^2} \int_0^\infty \varphi(x) \exp\left(-\frac{x^2}{2\sigma^2}\right) \left(\int_0^x \exp\left(-\frac{y^2}{2\sigma^2}\right) dy\right) dx.$$

Pour $u \geq 0$, on pose $e(u) = \int_0^u \exp\left(-\frac{\xi^2}{2}\right) d\xi$. On a alors

$$E(\varphi(Z)) = \frac{4}{\pi\sigma} \int_0^\infty \varphi(x) \exp\left(-\frac{x^2}{2\sigma^2}\right) e\left(\frac{x}{\sigma}\right) dx.$$

La loi de Z a donc une densité par rapport à la mesure de Lebesgue de \mathbb{R} et cette densité g est donnée par

$$g(x) = \frac{4}{\pi\sigma} 1_{\mathbb{R}_+}(x) \exp\left(-\frac{x^2}{2\sigma^2}\right) e\left(\frac{x}{\sigma}\right).$$

L'hypothèse d'indépendance de X et Y est inutile pour cet exercice.
