Igor Kortchemski – igor.kortchemski@polytechnique.edu

Feuille d'exercices (cours 1) : convergence étroite dans $\mathbb{R}^k$

On note $\mathcal{M}_1(\mathbb{R}^k)$ l'ensemble des mesures de probabilités sur $\mathbb{R}^k$, avec $k \geq 1$ un entier fixé, et $\Rightarrow$ désigne la convergence étroite dans cet espace.

1 Exercices à chercher pour le mardi 26 septembre

Ces exercices seront corrigés au début de la séance du mardi 26 septembre

Exercice 1. Soit $(\mu_n)_{n\geq 1}$ une suite de mesures de $\mathcal{M}_1(\mathbb{R})$ et $\mu \in \mathcal{M}_1(\mathbb{R})$. On suppose que $\mu_n \Rightarrow \mu$. Y a-t-il des implications entre les assertions suivantes ?

(a) μ_n est à densité pour n assez grand

(b) μ est à densité

(c) μ_n est atomique pour n assez grand

(d) μ est atomique.

Rappelons qu'une mesure atomique est une mesure qui s'écrit $\sum a_i \delta_{b_i}$ pour des suites $a_i \in \mathbb{R}_+$ et $b_i \in \mathbb{R}$ et que dans $\mathbb{R}^n$ par mesure à densité on entend par rapport à la mesure de Lebesgue (l'usage est juste de dire à densité).

Corrigé :

Non :

- si μ_n est à densité, μ peut aussi bien être à densité (prendre $\mu = \mu_n$) qu'atomique (prendre $\mu_n(dx) = 2n\mathbb{1}_{[-1/n,1/n]}(x)dx$, qui converge étroitement vers δ_0), voire singulière par rapport à la mesure de Lebesgue (penser à l'escalier du diable).
- si μ_n est atomique, μ peut aussi bien être atomique (prendre $\mu = \mu_n$) ou bien à densité (prendre $\mu_n = \frac{1}{n}\sum_{i=1}^n \delta_{i/n}$, qui converge étroitement vers la mesure de Lebesgue sur $[0, 1]$).

□

Exercice 2. Montrer qu'une famille $(\mu_i)_{i\in I}$ de mesures de $\mathcal{M}_1(\mathbb{R})$ est tendue si et seulement si il existe une fonction mesurable $f : \mathbb{R} \to [0, \infty)$ telle que $f(x) \to \infty$ pour $|x| \to \infty$ et $\sup_{i\in I} \int_{\mathbb{R}} f \, d\mu_i < \infty$.

Corrigé :

$\boxed{\Longleftarrow}$ Posons $C = \sup_{i\in I} \int_{\mathbb{R}} f \, d\mu_i$. Soit $\varepsilon > 0$ et $A > 0$ tel que $|f(x)| \geq C/(\varepsilon)$ pour $|x| \geq A$. On écrit, pour $i \in I$,

$$C \geq \int_{\mathbb{R}} f \, d\mu \geq \int_{\mathbb{R}\backslash[-A,A]} f \, d\mu \geq \frac{C}{\varepsilon} \int_{\mathbb{R}\backslash[-A,A]} d\mu_i,$$

de sorte que $\mu_i(\mathbb{R}\backslash[-A, A]) \leq \varepsilon$.

$\boxed{\Longrightarrow}$ Soit $(A_n)_{n\geq 1}$ une suite de réels strictement positifs tels que pour tout $n \geq 1$ et $i \in I$ on ait

$$\mu_i(\mathbb{R}\backslash[-A_n, A_n]) \leq \frac{1}{n^3}.$$

Sans perte de généralité, on peut supposer que la suite (A_n) est strictement croissante et $A_n \to \infty$. On définit alors f par $f(x) = 0$ si $|x| < A_1$ et $f(x) = n$ si $A_n \leq |x| < A_{n+1}$, de sorte que f est mesurable, $f(x) \to \infty$, et pour $i \in I$,

$$\begin{aligned} \int_{\mathbb{R}} f \, \mathrm{d}\mu_i &= \sum_{n=1}^{\infty} \int_{[-A_{n+1}, A_{n+1}] \backslash [-A_n, A_n]} f \, d\mu_i \\ &\leq \sum_{n=1}^{\infty} n \mu_i([-A_{n+1}, A_{n+1}] \backslash [-A_n, A_n]) \\ &\leq \sum_{n=1}^{\infty} n \mu_i(\mathbb{R} \backslash [-A_n, A_n]) \\ &\leq \sum_{n=1}^{\infty} \frac{1}{n^2}, \end{aligned}$$

d'où le résultat. □

2 Exercices additionnels (facultatif)

Exercice 3. Soit $(X_n)_{n\geq 1}$ des variables aléatoires réelles. Montrer que la suite $(X_n)_{n\geq 1}$ est tendue si et seulement si pour tout $\varepsilon > 0$ et pour toute suite $(c_n)_{n\geq 1}$ de réels strictement positifs telle que $c_n \to 0$ on a $\mathbb{P}(c_n|X_n| \geq \varepsilon) \to 0$.

Corrigé :
Supposons que la suite $(X_n)_{n\geq 1}$ est tendue. Soit $\varepsilon > 0$ et $(c_n)_{n\geq 1}$ une suitede réels positifs telle que $c_n \to 0$. Soit $\eta > 0$. Par tension, il existe $M > 0$ tel que $\mathbb{P}(|X_n| \geq M) \leq \eta$. Pour n assez grand, $\varepsilon / c_n \geq M$, et alors

$$\mathbb{P}(c_n|X_n| \geq \varepsilon) \leq \mathbb{P}(|X_n| \geq M) \leq \eta,$$

d'où le résultat.

Réciproquement, raisonnons par l'absurde en supposant que la suite $(X_n)_{n\geq 1}$ n'est pas tendue. On peut alors trouver $\varepsilon > 0$ et une extraction $\phi(n)$ telle que $\mathbb{P}\left(|X_{\phi(n)}| \geq n\right) \geq \varepsilon$ pour tout n. On définit alors la suite $(c_n)_{n\geq 1}$ en posant $c_{\phi(k)} = k$ pour tout $k \geq 1$ et $c_i = k - 1$ pour $\phi(k-1) \leq i < \phi(k)$. Alors, pour tout $n \geq 1$,

$$\mathbb{P}\left(c_{\phi(n)}|X_{\phi(n)}| \geq 1\right) \geq \varepsilon,$$

contradiction. □

Exercice 4. – *(Théorème de Riesz et lemme de Scheffé)* – Soit $(E, \mathcal{A}, \mu)$ un espace mesuré, avec μ une mesure positive (pas forcément finie). Soit $(f_n)_{n\geq 1}$ une suite de fonctions mesurables de E dans $\mathbb{R}$ telles que :

$$f_n \to f \quad \mu - \text{presque partout}. \tag{1}$$

(1) On suppose que la convergence (1) a lieu, que pour tout $n \geq 1$, $f_n \in L^p(\mu)$, $f \in L^p(\mu)$ et $\|f_n\|_p \to \|f\|_p$ quand $n \to \infty$. Démontrer que $f_n \to f$ dans $L^p(\mu)$ (théorème de Riesz).

On pourra introduire la fonction $g_n = 2^p(|f_n|^p + |f|^p) - |f_n - f|^p$.

Igor Kortchemski – igor.kortchemski@polytechnique.edu

(2) Montrer le lemme de Scheffé :

$$\text{si } (1), f_n \text{ et } f \text{ sont } \mu \text{ intégrables,} \quad f_n \geq 0 \quad \text{et} \quad \int_E f_n d\mu \underset{n\to\infty}{\longrightarrow} \int_E f\, d\mu$$

alors $\int_E |f_n - f| d\mu \underset{n\to\infty}{\longrightarrow} 0$.

(3) Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires positives qui converge presque sûrement vers X. On suppose que $\mathbb{E}[X_n] \to \mathbb{E}[X]$. Montrer que X_n converge vers X dans L^1.

(4) Soit $(X_n)_{n\geq 1}$ des variables aléatoires réelles telles que pour tout $n \geq 1$, X_n a pour loi $f_n d\mu$ (i.e. sa loi a une densité f_n par rapport à μ). Montrer que si f_n converge μ-presque partout vers une densité de probabilité f, alors $X_n \Rightarrow X$, où X est une variable aléatoire de loi $f\, d\mu$.

Corrigé :

(1) Par inégalité triangulaire, on remarque que $|f_n - f|^p \leq (|f_n| + |f|)^p \leq 2^p \max(|f_n|^p, |f|^p) \leq 2^p(|f_n|^p + |f|^p)$ de sorte que $g_n \geq 0$. Appliquons le lemme de Fatou à g_n :

$$2^{p+1} \int_E |f|^p \mathrm{d}\mu = \int_E (\liminf_{n\to\infty} g_n) \mathrm{d}\mu \leq \liminf_{n\to\infty} \int_E 2^p(|f_n|^p + |f|^p)\mathrm{d}\mu - \limsup_{n\to\infty} \int_E |f_n - f|^p \mathrm{d}\mu.$$

On en déduit que

$$\limsup_{n\to\infty} \int_E |f_n - f|^p \mathrm{d}\mu \leq 0,$$

d'où le résultat.

(2) C'est simplement le théorème de Riesz pour $p = 1$.

(3) C'est simplement le lemme de Scheffé.

(4) Soit $s \in \mathbb{R}$. On montre que $\mathbb{P}(X_n \leq s) \to \mathbb{P}(X \leq s)$. Pour cela, en notant $A =]-\infty, s]$, on remarque que

$$\left| \int_A f_n d\mu - \int_A f\, d\mu \right| \leq \int_{\mathbb{R}} |f_n - f| d\mu \underset{n\to\infty}{\longrightarrow} 0,$$

ce qui conclut.

Remarque. La même preuve montre qu'en fait $\mathbb{P}(X_n \in A) \to \mathbb{P}(X \in A)$ pour tout borélien A. Ceci implique la convergence en loi, mais la réciproque n'est pas vraie (prendre par exemple $\mu_n = \delta_{1/n} \Rightarrow \delta_0$ avec $A = \{0\}$).

□

Exercice 5. – *(Transformée de Laplace)* – Soit $\mu \in \mathcal{M}_1(\mathbb{R}_+)$. Sa transformée de Laplace est définie par l'intégrale $L_\mu(t) = \int_{\mathbb{R}_+} e^{-tx} \mu(\mathrm{d}x)$ pour $t \geq 0$.

(1) Vérifier que L_μ est de classe C^∞ sur $]0, \infty[$.

(2) Montrer que pour tout $x > 0$ on a

$$\mu([0, x[) + \frac{1}{2}\mu(\{x\}) = \lim_{t\to\infty} \sum_{k=0}^{\lfloor tx \rfloor} \frac{(-t)^k}{k!} L_\mu^{(k)}(t),$$

Igor Kortchemski – igor.kortchemski@polytechnique.edu

où $L_\mu^{(k)}$ désigne la dérivée k-ième de L_μ. En déduire que si $\mu, \nu \in \mathcal{M}_1(\mathbb{R}_+)$ ont même transformée de Laplace, alors $\mu = \nu$.

(3) Soit $(\mu_n)_{n\geq 1}$ une suite de mesures de $\mathcal{M}_1(\mathbb{R}_+)$. On suppose que L_{μ_n} converge simplement sur $\mathbb{R}_+$ vers une fonction L continue à droite en 0. Montrer qu'il existe $\mu \in \mathcal{M}_1(\mathbb{R}_+)$ telle que $L = L_\mu$ et $\mu_n \Rightarrow \mu$.

On pourra s'inspirer de la preuve du théorème de Lévy.

Corrigé :

(1) Ceci provient du fait que pour tout entier $k \geq 1$, la dérivée k-ième de $t \mapsto e^{-tx}$ est μ-intégrable, puisque la fonction $x \mapsto x^k e^{-tx}$ est bornée sur $\mathbb{R}_+$ pour tout $t > 0$. En particulier,

$$L_\mu^{(k)}(t) = (-1)^k \int_{\mathbb{R}_+} u^k e^{-tu} \mu(\mathrm{d}u).$$

(2) Soit $x > 0$. D'après la question précédente et le théorème de Fubini,

$$\sum_{k=0}^{\lfloor tx \rfloor} \frac{(-t)^k}{k!} L_\mu^{(k)}(t) = \int_{\mathbb{R}_+} \sum_{k=0}^{\lfloor tx \rfloor} \frac{(tu)^k}{k!} e^{-tu} \mu(\mathrm{d}u).$$

Or

$$\sum_{k=0}^{\lfloor tx \rfloor} \frac{(tu)^k}{k!} e^{-tu} = \mathbb{P}\left(\mathsf{Poisson}(tu) \leq \lfloor tx \rfloor\right).$$

Montrons que

$$\mathbb{P}\left(\mathsf{Poisson}(tu) \leq \lfloor tx \rfloor\right) \quad \underset{t\to\infty}{\longrightarrow} \quad \begin{cases} 0 & \text{si } x < u \\ \frac{1}{2} & \text{si } x = u \\ 1 & \text{si } x > u \end{cases}$$

en distinguant les trois cas :

– $x = u$: En notant $(Y_i)_{i\geq 1}$ des variables aléatoires i.i.d. de Poisson de paramètre x, on a

$$\mathbb{P}\left(\mathsf{Poisson}(tx) \leq \lfloor tx \rfloor\right) = \mathbb{P}\left(\frac{Y_1 + \cdots + Y_{\lfloor t \rfloor} - \lfloor t \rfloor x + \mathsf{Poisson}(\lfloor tx \rfloor - \lfloor t \rfloor x) - (\lfloor tx \rfloor - \lfloor t \rfloor x)}{\sqrt{tx}} \leq 0\right)$$

avec $\mathsf{Poisson}(\lfloor tx \rfloor - \lfloor t \rfloor x)$ une variable aléatoire de Poisson de paramètre $\lfloor tx \rfloor - \lfloor t \rfloor x$ indépendante de $(Y_i)_{i\geq 1}$. D'après le TCL,

$$\mathbb{P}\left(\frac{Y_1 + \cdots + Y_{\lfloor t \rfloor} - \lfloor t \rfloor x}{\sqrt{tx}} \leq 0\right) \quad \underset{t\to\infty}{\longrightarrow} \quad \frac{1}{2}.$$

Igor Kortchemski – igor.kortchemski@polytechnique.edu

En remarquant que $|\lfloor tx\rfloor - \lfloor t\rfloor x| \leq 1+x$, on voit que $\frac{\mathsf{Poisson}(\lfloor tx\rfloor-\lfloor t\rfloor x)-(\lfloor tx\rfloor-\lfloor t\rfloor x)}{\sqrt{tx}} \to 0$ en probabilité lorsque $t\to\infty$ (ceci provient par exemple de l'inégalité de Markov)) donc d'après le lemme de Slutsky on conclut que $\mathbb{P}(\mathsf{Poisson}(tx) \leq \lfloor tx\rfloor) \to \frac{1}{2}$ lorsque $t\to\infty$.

Remarque. Plus formellement, on a utilisé le fait que si $X_n+Y_n \Rightarrow Z$, avec $X_n \perp\!\!\!\perp Y_n$ et Y_n qui converge en probabilité vers 0, alors $X_n \Rightarrow Z$.

– $x<u$: alors

$$\begin{aligned}\mathbb{P}(\mathsf{Poisson}(tu) \leq \lfloor tx\rfloor) &= \mathbb{P}(tu-\lfloor tx\rfloor \leq tu-\mathsf{Poisson}(tu))\\ &\leq \mathbb{P}(tu-\lfloor tx\rfloor \leq |tu-\mathsf{Poisson}(tu)|)\\ &\leq \frac{\mathsf{Var}(\mathsf{Poisson}(tu))}{(tu-\lfloor tx\rfloor)^2}\\ &\sim \frac{u}{t(u-x)}\end{aligned}$$

lorsque $t\to\infty$, qui tend donc vers 0.

– $x>u$: alors, de même,

$$\begin{aligned}\mathbb{P}(\mathsf{Poisson}(tu) \leq \lfloor tx\rfloor) &= 1-\mathbb{P}(\mathsf{Poisson}(tu) > \lfloor tx\rfloor)\\ &= 1-\mathbb{P}(\mathsf{Poisson}(tu)-tu > \lfloor tx\rfloor - tu)\\ &\geq 1-\mathbb{P}(|\mathsf{Poisson}(tu)-tu| > \lfloor tx\rfloor - tu)\end{aligned}$$

et

$$\mathbb{P}(|\mathsf{Poisson}(tu)-tu| > \lfloor tx\rfloor - tu) \leq \frac{\mathsf{Var}(\mathsf{Poisson}(tu))}{(\lfloor tx\rfloor - tu)^2} \sim \frac{u}{t(x-u)} \to 0.$$

Ainsi,

$$\sum_{k=0}^{\lfloor tx\rfloor} \frac{(tu)^k}{k!}e^{-tu} \quad \underset{t\to\infty}{\longrightarrow} \quad \mathbb{1}_{[0,x[}(u)+\frac{1}{2}\mathbb{1}_{\{x\}}(u).$$

En utilisant le théorème de convergence dominée, on conclut que

$$\int_{\mathbb{R}_+}\sum_{k=0}^{\lfloor tx\rfloor} \frac{(tu)^k}{k!}e^{-tu}\mu(\mathrm{d}u) \quad \underset{t\to\infty}{\longrightarrow} \quad \int_{\mathbb{R}_+}(\mathbb{1}_{[0,x[}(u)+\frac{1}{2}\mathbb{1}_{\{x\}}(u))\mu(d\mathrm{u}) = \mu([0,x[)+\frac{1}{2}\mu(\{x\}).$$

Supposons maintenant que $\mu,\nu \in \mathcal{M}_1(\mathbb{R}_+)$ ont même transformée de Laplace. Notons D l'ensemble des atomes de μ et de ν, qui est au plus dénombrable. Alors d'après le résultat précédent, $\mu([x,y[)=\nu([x,y[)$ pour tous $x,y\in\mathbb{R}\backslash D$. Puisque $\mathbb{R}\backslash D$ est dense, on en déduit que μ et ν coincident sur tout intervalle ouvert et donc $\mu=\nu$ par application du lemme des classes monotones.

(3) On remarque tout d'abord que si $\mu_n \Rightarrow \mu$, alors L_{μ_n} converge simplement vers μ (la fonction $x\mapsto e^{-tx}$ étant continue bornée sur $\mathbb{R}_+$).

Igor Kortchemski – igor.kortchemski@polytechnique.edu

Il suffit de montrer que $(\mu_n)_{n\geq 1}$ est tendue. En effet, on conclut alors la preuve comme pour le théorème de Lévy : soit μ la limite en loi le long d'une sous-suite ϕ_0 (qui existe par tension). On raisonne par l'absurde et on suppose qu'il existe $\varepsilon > 0$, une extraction ϕ et une fonction $f : \mathbb{R} \to \mathbb{R}$ continue bornée telle que $|\mu_{\phi(n)}(f) - \mu(f)| \geq \varepsilon$. Par tension, il existe une extraction ψ telle que $\mu_{\phi\circ\psi(n)} \Rightarrow \nu$ pour une certaine mesure $\nu \in M1(\mathbb{R})$. Comme $L_{\mu_{\phi_0(n)}}$ converge simplement vers L_μ, comme $L_{\mu_{\phi_0(n)}}$ converge simplement vers L_μ et comme L_{μ_n} converge simplement vers L, on en déduit que $L = L_\mu = L_\nu$ et donc $\mu = \nu$.

Pour montrer que $(\mu_n)_{n\geq 1}$ est tendue, on écrit

$$\begin{aligned}
K\int_0^{1/K}(1-L_{\mu_n}(t))\mathrm{d}t &= K\int_0^{1/K}\left(1-\int_0^\infty e^{-tx}\mu(\mathrm{d}x)\right)\mathrm{d}t\\
&= K\int_0^\infty\int_0^{1/K}(1-e^{-tx})\mathrm{d}t\mu_n(\mathrm{d}x)\\
&\geq K\int_K^\infty\int_0^{1/K}(1-e^{-tx})\mathrm{d}t\mu_n(\mathrm{d}x)\\
&\geq K\int_K^\infty\int_0^{1/K}(1-e^{-Kt})\mathrm{d}t\mu_n(\mathrm{d}x)\\
&= \frac{1}{e}\mu_n([K,\infty[).
\end{aligned}$$

Or par convergence dominée $K\int_0^{1/K}(1-L_{\mu_n}(t))\mathrm{d}t \to K\int_0^{1/K}(1-L_\mu(t))\mathrm{d}t$ et par continuité à droite $K\int_0^{1/K}(1-L_\mu(t))\mathrm{d}t \to 0$ lorsque $K\to\infty$. On en déduit aisément la tension de $(\mu_n)_{n\geq 1}$.

□

Exercice 6. Soit $(\mu_n)_{n\geq 1}$ une suite de mesures de $\mathcal{M}_1(\mathbb{R})$ et $\mu \in \mathcal{M}_1(\mathbb{R})$.

(1) On suppose que $(\mu_n)_{n\geq 1}$ est tendue. Montrer que

$$\forall \varepsilon > 0, \quad \exists \delta > 0, \quad \forall n \geq 1, \quad |x-y| \leq \delta \implies |\phi_{\mu_n}(x) - \phi_{\mu_n}(y)| \leq \varepsilon.$$

(en d'autres termes, la suite $(\phi_{\mu_n})_{n\geq 1}$ est uniformément équicontinue).

(2) On suppose que $\mu_n \Rightarrow \mu$. Montrer que ϕ_{μ_n} converge vers ϕ_μ uniformément sur tout compact. Donner un exemple où la convergence n'est pas uniforme.

(3) La réciproque de l'énoncé de la première question est-elle vraie, autrement dit est-ce que si la suite $(\phi_{\mu_n})_{n\geq 1}$ est uniformément équicontinue alors la suite $(\mu_n)_{n\geq 1}$ est tendue ?

Corrigé :

(1) Soit $\varepsilon > 0$ et $M > 0$ tel que pour tout $n \geq 1$, $\mu_n(\mathbb{R}\backslash[-M,M]) \leq \varepsilon$. Par uniforme continuité, il

Igor Kortchemski – igor.kortchemski@polytechnique.edu

existe $\delta > 0$ tel que $|x-y| \leq \delta$ et $t \in [-M,M]$ impliquent $|e^{itx} - e^{ity}| \leq \varepsilon$. On écrit alors :

$$\begin{aligned} |\phi_{\mu_n}(x) - \phi_{\mu_n}(y)| &= \int_{[-M,M]} |e^{itx} - e^{ity}| \mu_n(dt) + \int_{\mathbb{R}\setminus[-M,M]} |e^{itx} - e^{ity}| \mu_n(dt) \\ &\leq \varepsilon + 2\mu_n(\mathbb{R}\setminus[-M,M]) \\ &\leq 3\varepsilon \end{aligned}$$

(2) On sait que ϕ_{μ_n} converge simplement vers ϕ_μ. C'est alors un résultat général : si une suite de fonctions réelles sur un compact est uniformément équicontinue et converge simplement, alors elle converge uniformément. Pour le démontrer dans notre cas précis, soit $\varepsilon > 0$ et K un compact. Soit $\delta > 0$ tel que l'implication de la première question est vraie. En passant à la limite, remarquons tout d'abord que $|x-y| \leq \delta$ implique $|\phi_\mu(x) - \phi_\mu(y)| \leq \varepsilon$. Par compacité, on peut recouvrir K par un nombre fini de boules $(B(x_i,\delta))_{1 \leq i \leq k}$. Soit N tel que $n \geq N$ implique $\max_{1 \leq i \leq k} |\phi_{\mu_n}(x_i) - \phi_\mu(x)| \leq \varepsilon$. Alors, pour $n \geq N$, soit $x \in K$. En notant $1 \leq i \leq k$ l'entier tel que $x \in B(x_i, \delta)$:

$$|\phi_{\mu_n}(x) - \phi_\mu(x)| \leq |\phi_{\mu_n}(x) - \phi_{\mu_n}(x_i)| + |\phi_{\mu_n}(x_i) - \phi_\mu(x_i)| + |\phi_\mu(x_i) - \phi_\mu(x)| \leq 3\varepsilon,$$

ce qui conclut.

En prenant par exemple $X_n = \mathcal{N}(0, 1-1/n)$ et $X = \mathcal{N}(0,1)$, on a $\phi_{X_n}(t) = e^{-t^2(1-1/n)^2/2} \to \phi_X(t)$ pour tout $t \in \mathbb{R}$, mais $\|\phi_{X_n} - \phi_X\|_\infty = \infty$ pour tout $n \geq 1$.

(3) Oui : si $(\phi_{\mu_n})_{n \geq 1}$ est uniformément équicontinue, comme $|\phi_{\mu_n}(0)| = 1$, d'après le théorème d'Arzela-Ascoli, la suite $(\phi_{\mu_n})_{n \geq 1}$ a des sous-suites qui convergent uniformément sur tout compact. Par procédé diagonal (en se restreignant par exemple à des compacts $[-N,N]$ avec N entier), on trouve une extraction γ telle que $(\phi_{\mu_{\gamma(n)}})_{n \geq 1}$ converge uniformément sur tout compact vers une fonction limite continue ϕ. D'après le théorème de Lévy, ϕ est la fonction caractéristique d'une mesure de probabilité vers laquelle $(\phi_{\mu_{\gamma(n)}})_{n \geq 1}$ converge étroitement.

□

Exercice 7. Soit $(\mu_n)_{n \geq 1}$ une suite de mesures de $\mathcal{M}_1(\mathbb{R})$ et $\mu \in \mathcal{M}_1(\mathbb{R})$ sans atomes. Montrer que $\mu_n \Rightarrow \mu$ si et seulement si $\sup_{s \in \mathbb{R}} |F_{\mu_n}(s) - F_\mu(s)| \to 0$.

Corrigé :

La réciproque est claire. Le sens direct est une conséquence du deuxième théorème de Dini. Donnons ici une approche dans le contexte de l'exercice. Fixons $k \geq 2$. L'application F_μ étant continue, croissante, de limite nulle en $-\infty$ et de limite 1 en ∞, il existe des points $s_1 < \cdots < s_{k-1}$ tels que $F_\mu(s_i) = \frac{i}{k}$ pour tout $1 \leq i \leq k$. Par convergence simple, pour n assez grand, pour tout $1 \leq i \leq k$,

$$\frac{i-1}{k} \leq F_{\mu_n}(s_i) \leq \frac{i+1}{k}.$$

Par convention, posons $s_0 = -\infty$ et $s_k = \infty$. Il vient que pour tout n assez grand, pour tout $s \in \mathbb{R}$, en

Igor Kortchemski – igor.kortchemski@polytechnique.edu

choisissant s_i tel que $s_i \leq s < s_{i+1}$:

$$F_{\mu_n}(s) \leq F_{\mu_n}(s_{i+1}) \leq \frac{i+2}{k} = F_\mu(s_i) + \frac{2}{k} \leq F_\mu(s) + \frac{2}{k}$$

et de même

$$F_{\mu_n}(s) \geq F_{\mu_n}(s_i) \geq \frac{i-1}{k} = F_\mu(s_{i+1}) - \frac{2}{k} \geq F_\mu(s) - \frac{2}{k}.$$

Ainsi, pour n assez grand,

$$\sup_{s \in \mathbb{R}} |F_{\mu_n}(s) - F_\mu(s)| \leq \frac{2}{k},$$

ce qui conclut. □

Exercice 8. – *(Théorème de Glivenko-Cantelli)* – Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles indépendantes et de même loi μ. On considère $\mu_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}$ la mesure empirique de ces variables aléatoires. On note F_n la fonction de répartition de μ_n et F celle de μ. Le but de cet exercice est de le théorème de Glivenko-Cantelli :

$$\text{presque sûrement,} \qquad \lim_{n\to\infty} \sup_{s\in\mathbb{R}} |F_n(s) - F(s)| = 0. \qquad (\star)$$

(1) Montrer $(\star)$ lorsque μ est la loi uniforme sur $[0,1]$.

Pour $0 \leq x \leq 1$, on pose $G(x) = \inf\{s \in \mathbb{R} : F(s) \geq x\}$ (appelé inverse généralisé de F). Il est possible de vérifier que G est croissante, continue à gauche en tout point. Soit $(Y_i)_{1\leq i\leq n}$ des variables aléatoires indépendantes de loi uniforme sur $[0,1]$.

(2) (a) Montrer que pour tout $t \in \mathbb{R}$ et $x \in [0,1]$ on a

$$F(t) \geq x \iff t \geq G(x). \qquad (2)$$

(b) Montrer que $(G(Y_i))_{1\leq i\leq n}$ sont des variables indépendantes de même loi μ.

(c) Montrer que F_n et $A_n \circ F$ ont même loi, où A_n est la fonction de répartition empirique des $Y_1, \ldots, Y_n$.

(3) En déduire $(\star)$.

Corrigé :

(1) On suit la même approche que pour l'exercice 7. Fixons $k \geq 2$. Il existe des points $s_1 < \cdots < s_{k-1}$ tels que $F(s_i) = \frac{i}{k}$ pour tout $1 \leq i \leq k$. D'après la loi des grands nombres, presque sûrement, pour n assez grand, pour tout $1 \leq i \leq k$,

$$\frac{i-1}{k} \leq F_{\mu_n}(s_i) \leq \frac{i+1}{k}.$$

Par convention, posons $s_0 = -\infty$ et $s_k = \infty$. Il vient que pour tout n assez grand, pour tout $s \in \mathbb{R}$,

Igor Kortchemski – igor.kortchemski@polytechnique.edu

en choisissant s_i tel que $s_i \leq s < s_{i+1}$:

$$F_{\mu_n}(s) \leq F_{\mu_n}(s_{i+1}) \leq \frac{i+2}{k} = F_\mu(s_i) + \frac{2}{k} \leq F_\mu(s) + \frac{2}{k}$$

et de même

$$F_{\mu_n}(s) \geq F_{\mu_n}(s_i) \geq \frac{i-1}{k} = F_\mu(s_{i+1}) - \frac{2}{k} \geq F_\mu(s) - \frac{2}{k}.$$

Ainsi, presque sûrement, pour n assez grand,

$$\sup_{s \in \mathbb{R}} |F_{\mu_n}(s) - F(s)| \leq \frac{2}{k}.$$

Ceci étant vrai pour tout $k \geq 2$, ceci conclut.

(2) (a) Si $F(t) \geq x$, alors clairement $\inf\{s \in \mathbb{R} : F(s) \geq x\} \leq t$ et donc $G(x) \leq t$. Réciproquement, montrons que $F(t) < x$ implique $t < G(x)$. Puisque F est continue à droite, il existe $\varepsilon > 0$ tel que $F(t+\varepsilon) < x$. On a alors $G(x) \geq t + \varepsilon > t$.

(b) Elles sont clairement indépendantes de même loi et d'après la question précédente, pour $0 \leq t \leq 1$,

$$\mathbb{P}(G(Y_1) \leq t) = \mathbb{P}(F(t) \geq Y_1) = F(t),$$

d'où le résultat.

(c) Sans perte de généralité, on peut supposer que $X_i = G(Y_1)$. On a alors

$$F_n(t) = \frac{1}{n}\mathrm{Card}(\{1 \leq i \leq n : G(Y_i) \leq t\} = \frac{1}{n}\mathrm{Card}(\{1 \leq i \leq n : Y_i \leq F(t)\} = A_n(F(t)).$$

(3) D'après (1), on a

$$\text{presque sûrement,} \qquad \lim_{n \to \infty} \sup_{s \in \mathbb{R}} |A_n(s) - s| = 0,$$

ce qui implique

$$\text{presque sûrement,} \qquad \lim_{n \to \infty} \sup_{s \in \mathbb{R}} |A_n(F(s)) - F(s)| = 0,$$

et le résultat désiré en découle par (c).

□

Exercice 9. – *(Caractérisation des fonctions de répartition dans $\mathbb{R}^k$.)* – Pour $x = (x_1, \ldots, x_k) \in \mathbb{R}^k$ et $y = (y_1, \ldots, y_k) \in \mathbb{R}^k$ on note $x \leq y$ si $x_i \leq y_i$ pour tout $1 \leq i \leq k$. Soit $F : \mathbb{R}^k \to [0,1]$ une fonction. On dit qu'elle est *continue à droite* si $F(x^n) \to F(x)$ lorsque $x^n \downarrow x$. On dit qu'elle est *propre* si $F(x) \to 1$ lorsque $\min_i x_i \to \infty$ et $F(x) \to 0$ lorsque $\min_i x_i \to 0$. On dit qu'elle est une *fonction de répartition* s'il existe une mesure $\mu \in \mathcal{M}_1(\mathbb{R}^k)$ telle que $F(x) = \mu(\{y \in \mathbb{R}^k : y \leq x\})$.

(1) Donner un exemple de fonction $F : \mathbb{R}^2 \to [0,1]$ continue à droite, propre, croissante en chacune de ses variables et qui n'est pas une fonction de répartition.

On dit que F est à *accroissements positifs* si pour tout pavé $]x,y] =]x_1,y_1] \times \cdots \times]x_k,y_k]$ on a $F(]x,y]) := \sum_u s(u)F(u) \geq 0$, où la somme est prise sur tous les coins u de $]x,y]$ et $s(u) = (-1)^p$ avec $p = \sum_{i=1}^k \mathbb{1}_{u_i = y_i}$.

(2) Montrer que $F : \mathbb{R}^k \to [0,1]$ est une fonction de répartition si et seulement si elle est continue à

droite, propre, et est à accroissements positifs.

Pour une preuve probabiliste de la réciproque, on pourra justifier l'existence pour tout $n \geq 1$ *de mesures de probabilité* μ_n *à support dans* $(2^{-n}\mathbb{Z})^k)$ *telles que* $\mu_n(x/2^n) = F(]2^{-n}(x-1), 2^{-n}x])$ *pour* $x \in \mathbb{Z}^k$*, et de variables aléatoires* $(X^n)_{n\geq 1}$ *telles que* X^n *soit de loi* μ_n *et* $X^m - 2^{-m} < X^n \leq X^m$ *pour tout* $m < n$*, et enfin considérer* $X = \lim_{n\to\infty} X^n$.

Corrigé :

(1) Prenons $F(x_1, x_2) = \mathbb{1}_{x_1+x_2\geq 0}$. S'il existe $\mu \in \mathcal{M}_1(\mathbb{R}^k)$ telle que F soit sa fonction de répartition, on a alors $\mu([-1,2]\times[-1,2]) = F(2,2) - F(2,-1) - F(-1,2) + F(-1,-1) = -1$, absurde.

(2) L'implication est claire (si F est la fonction de répartition de $\mu \in \mathcal{M}_1(\mathbb{R}^k)$, la propriété d'être à accroissements positifs provient du fait que $\sum_u s(u)F(u) = \mu(]x,y])$ – en reprenant les notations de l'énoncé).

Puisque F est propre et à accroissements positifs, il existe bien des mesures de probabilité μ_n à support dans $(2^{-n}\mathbb{Z})^k)$ telles que $\mu_n(x/2^n) = F(]2^{-n}(x-1), 2^{-n}x])$ pour $x \in \mathbb{Z}^k$.

Ensuite, par construction, F est finiment additive sur les pavés, de sorte que pour $x \in \mathbb{Z}^k$ et $1 \leq m < n$:

$$\mu_m(2^{-m}]x-1,x]) = \mu_n(2^{-m}(x-1,x]).$$

Ceci permet de construire par récurrence une suite de variables aléatoires $(X^n)_{n\geq 1}$ telles que X^n soit de loi μ_n et $X^m - 2^{-m} < X^n \leq X^m$ pour tout $m < n$ à partir d'une suite de variables aléatoires i.i.d. uniformes sur $[0,1]$. En effet, supposons $X^1, \ldots, X^n$ construites. Puisque

$$\sum_{i_1,\ldots,i_k = 0 \text{ ou } 1} \mu_{n+1}\left(\left]X_1^n - \frac{i_1}{2^{n+1}}, X_1^n - \frac{i_1}{2^{n+1}} - 2^{-n-1}\right], \ldots, \left]X_k^n - \frac{i_k}{2^{n+1}}, X_1^n - \frac{i_k}{2^{n+1}} - 2^{-n-1}\right]\right)$$
$$= \mu_n\left(\left]X^n - \frac{1}{2^n}, X^n\right]\right),$$

on construit X^{n+1} sachant X^n de sorte que pour $i_1, \ldots, i_k = 0$ ou 1,

$$X^{n+1} = \left(X_1^n - \frac{i_1}{2^{n+1}}, \ldots, X_k^n - \frac{i_k}{2^{n+1}}\right)$$

avec probabilité

$$\frac{1}{\mu_n\left(\left]X^n - \frac{1}{2^n}, X^n\right]\right)} \mu_{n+1}\left(\left]X_1^n - \frac{i_1}{2^{n+1}}, X_1^n - \frac{i_1}{2^{n+1}} - 2^{-n-1}\right], \ldots, \left]X_k^n - \frac{i_k}{2^{n+1}}, X_1^n - \frac{i_k}{2^{n+1}} - 2^{-n-1}\right]\right).$$

On a bien $X^m - 2^{-m} < X^n \leq X^m$ pour tout $m < n$, ce qui permet par monotonie de définir $X = \lim_{n\to\infty} X^n$.

Igor Kortchemski – igor.kortchemski@polytechnique.edu

Vérifions que la fonction de répartition de X est F. Soit $x \in \mathbb{R}^k$ dyadique. En particulier, $\mathbb{P}(X^n \leq x) = F(x)$. Par ailleurs, d'après le lemme de Fatou,

$$\mathbb{P}(X < x) = \mathbb{E}\left[\lim_{n\to\infty} \mathbb{1}_{X^n < x}\right] \leq \liminf_{n\to\infty} \mathbb{P}(X^n < x) \leq \liminf_{n\to\infty} \mathbb{P}(X^n \leq x) \leq F(x) = \mathbb{P}(X^n \leq x) \leq \mathbb{P}(X \leq x).$$

Ainsi,

$$F(x) \leq \mathbb{P}(X \leq x) < \mathbb{P}(X < x + 2^{-n}) \leq F(x + 2^{-n}).$$

En faisant tendre n vers l'infini, on conclut en utilisant la continuité à droite de F.

□

Exercice 10. Soit (E,d) un espace métrique muni de sa tribu borélienne et $f : E \to \mathbb{R}$ mesurable. Montrer que l'ensemble des points de discontinuité de f est mesurable.

Indication. Pour $\varepsilon, \delta > 0$, on pourra vérifier que $U_{\varepsilon,\delta} := \{x \in E : \exists y, z \in B(x,\varepsilon), |f(y) - f(z)| > \delta\}$ est ouvert.

Corrigé :

Vérifions tout d'abord que $U_{\varepsilon,\delta}$ est bien ouvert. Si $x \in U_{\varepsilon,\delta}$ et $y, z \in B(x,\varepsilon)$ sont tels que $|f(y) - f(z)| > \delta$, alors pour tout x' tel que $d(x,x') < \varepsilon - \max(d(x,y), d(x,z))$ on a $x' \in U_{\varepsilon,\delta}$. En particulier, $U_{\varepsilon,\delta}$ est mesurable. On conclut en remarquant que l'ensemble des points de discontinuité de f s'écrit

$$\bigcup_{\substack{\delta>0\\ \delta\in\mathbb{Q}}} \bigcap_{\substack{\varepsilon>0\\ \varepsilon\in\mathbb{Q}}} U_{\varepsilon,\delta}.$$

□