Igor Kortchemski – igor.kortchemski@polytechnique.edu

Devoir maison

À rendre individuellement le mardi 8 novembre (rendu papier ou rendu électronique).

Les questions ne sont pas classées par ordre croissant de difficulté.

La précision et la clarté de la rédaction sera prise en compte dans l'évaluation.

Toute question peut être admise et utilisée ultérieurement, à condition de l'indiquer clairement.

Les résultats du cours ou des feuilles d'exercices peuvent être utilisés sans démonstration, à condition de les citer clairement.

Le DM est long et comporte des questions délicates : il est normal de ne pas réussir certaines questions.

* * *

Exercice 1. Soit $(X_n)_{n\geq 1}$ et X des variables aléatoires à valeurs dans un espace métrique séparable. Pour chaque mode de convergence que vous connaissez, étudier si le fait que toute sous-suite de $(X_n)_{n\geq 1}$ a une sous-suite qui converge vers X implique que X_n converge vers X.

Éléments de correction :
C'est vrai pour des modes de convergence dont la convergence est définie à travers une distance : ainsi c'est vrai pour la convergence L^p, la convergence en loi et la convergence en probabilité $X_n \to X$ en probabilité ssi $\mathbb{E}[\min(d(X_n,X),1)] \to 0$.

En revanche, c'est faux pour la convergence p.s. Par exemple si $(X_n)_{n\geq 1}$ est une suite de variables aléatoires de Bernoulli indépendantes telle que $\mathbb{P}(X_n = 1) = 1-\mathbb{P}(X_n = 0) = 1/n$, la suite $(X_n)_{n\geq 1}$ prend p.s. une infinité de fois les valeurs 0 et 1 par Borel-Cantelli, mais il est clair que de toute sous-suite ϕ en ré-extrayant ψ de sorte que $\phi \circ \psi(n) \geq n^2$, la suite $(X_{\phi\circ\psi(n)})_{n\geq 1}$ converge presque sûrement vers 0. □

* * *

Exercice 2. Soit μ_n, μ des mesures de probabilité sur $\mathbb{R}$. On note λ la mesure de Lebesgue sur $\mathbb{R}$ et $\phi_\nu(t) = \int e^{itx}\nu(\mathrm{d}x)$ pour une mesure de probabilité ν sur $\mathbb{R}$.

(1) Démontrer que $\mu_n \implies \mu$ si et seulement si la fonction caractéristique de μ_n converge simplement vers la fonction caractéristique de μ pour presque tout point par rapport à la mesure de Lebesgue.

(2) *[Question ouverte : bonus hors barème]* Est-il vrai que $\mu_n \implies \mu$ si et seulement si la fonction caractéristique de μ_n converge simplement sur un ensemble dénombrable dense vers la fonction caractéristique de μ ?

Éléments de correction :

(1) L'implication est claire. Pour la réciproque; on adapte la preuve du théorème de Lévy 1.

Comme dans le cours 1, on a l'existence d'une constante $c > 0$ telle que pour tout $A > 0$

$$\mu_n(\mathbb{R}\backslash[-A,A]) \leq cA \int_{-1/A}^{1/A} (1-\mathrm{Re}\phi_{\mu_n}(u))\mathrm{d}u \quad \underset{n\to\infty}{\longrightarrow} \quad cA \int_{-1/A}^{1/A} (1-\mathrm{Re}\phi_{\mu}(u))\mathrm{d}u \quad \underset{A\to\infty}{\longrightarrow} \quad 0,$$

où la première convergence a lieu en vertu de la converge presque partout de ϕ_{μ_n} vers ϕ_μ et du théorème de converge dominée, et la seconde en vertu de la continuité de ϕ_μ en 0.

Ainsi $(\mu_n)_{n\geq 1}$ est tendue. Supposons que $\mu_{\psi(n)} \implies \nu$. D'après le théorème de Prokhorov, il suffit de vérifier que $\mu = \nu$. On a alors $\phi_{\mu_{\psi(n)}}(t) \to \phi_\nu(t)$ pour tout $t \in \mathbb{R}$. Or $\phi_{\mu_{\psi(n)}}(t) \to \phi_\mu(t)$ pour presque tout $t \in \mathbb{R}$. Donc $\phi_\mu = \phi_\nu$ presque partout. Étant des fonctions continues, ces deux fonctions sont alors égales partout, ce qui conclut.

(2) C'est faux, on peut par exemple prendre $\mu_n = \delta_{2\pi n!}$.

□

* * *

Exercice 3. Soit (E,d) une espace métrique et (μ_n) une suite de mesures de probabilité sur E. On suppose que pour toute fonction continue bornée $f : E \to \mathbb{R}$ la suite $(\int f \mathrm{d}\mu_n)_{n\geq 1}$ converge.

(1) On suppose que E est compact. Montrer que $(\mu_n)_{n\geq 1}$ converge étroitement.

(2) *[question bonus hors barème]* On suppose que E est polonais. Montrer que $(\mu_n)_{n\geq 1}$ converge étroitement.

Éléments de correction :

(1) Pour f continue bornée, notons $\ell(f)$ la limite de $(\int f \mathrm{d}\mu_n)_{n\geq 1}$. Puisque E est compact, $(\mu_n)_{n\geq 1}$ est tendue. Soit ψ extraction et μ mesure de probabilité sur E tels que $\mu_{\psi(n)} \implies \mu$. En particulier, $\ell(f) = \int f \mathrm{d}\mu$.
Vérifions que $\mu_n \implies \mu$. D'après le théorème de Prokhorov, il suffit de vérifier que si $\mu_{\phi(n)} \implies \nu$ alors $\nu = \mu$. Par hypothèse $\mu_{\phi(n)}(f) \to \ell(f)$, et donc $\int f \mathrm{d}\nu = \ell(f) = \int f \mathrm{d}\mu$. On conclut que $\nu = \mu$.

(2) Il s'agit d'un résultat dû à Alexandrov, voir par exemple Ramamoorthi, R.V., Rao, B.V. & Sethuraman, J. A note on weak convergence. Sankhya A 74, 269–276 (2012). https://doi.org/10.1007/s13171-012-0016-6 pour une preuve.

□

* * *

Petit problème 4. Soit (E,d) un espace métrique séparable. On munit $\mathcal{M}_1(E)$ de la distance de Lévy-Prokhorov d_{LP}.

On pourra utiliser le fait suivant (démontré dans les compléments du cours 3) : pour tout $\mu \in \mathcal{M}_1(E)$ et pour tout $\varepsilon > 0$, il existe $f_1, \ldots, f_K \in \mathcal{C}_b(E,\mathbb{R})$ et $\delta > 0$ tels que

$$\left\{ \nu \in \mathcal{M}_1(E) : \left| \int f_i d\nu - \int f_i d\mu \right| < \delta \text{ pour tout } 1 \leq i \leq K \right\} \subset B(\mu, \varepsilon),$$

 Igor Kortchemski – igor.kortchemski@polytechnique.edu

où $B(\mu,\varepsilon)$ désigne la boule ouverte dans $\mathcal{M}_1(E)$ centrée en μ et de rayon ε pour d_{LP}.

Soit $\mu \in \mathcal{M}_1(E)$ une mesure de probabilité (déterministe) fixée, et $(M_n)_{n\geq 1}$ une suite de variables aléatoires à valeurs dans $\mathcal{M}_1(E)$.

Première partie.

(1) Démontrer que les assertions suivantes sont équivalentes :
   (a) pour tout $f \in \mathcal{C}_b(E,\mathbb{R})$, $\int f \mathrm{d}M_n \to \int f \mathrm{d}\mu$ dans $L^1(\mathbb{R})$.
   (b) M_n converge en loi vers μ (vue comme variable aléatoire à valeurs dans $\mathcal{M}_1(E)$) ;
   (c) M_n converge en probabilité vers μ (vue comme variable aléatoire à valeurs dans $\mathcal{M}_1(E)$) ;
   (d) pour tout $f \in \mathcal{C}_b(E,\mathbb{R})$, $\int f \mathrm{d}M_n$ converge en probabilité vers $\int f \mathrm{d}\mu$;
   (e) lorsque $E=\mathbb{R}$, en notant $\Phi_n(t)=\int e^{itx}M_n(\mathrm{d}x)$ et $\Phi(t)=\int e^{itx}\mu(\mathrm{d}x)$ pour tout $t\in\mathbb{R}$, il existe $\delta>0$ tel que $\sup_{|t|\leq\delta}|\Phi_n(t)-\Phi(t)|$ converge en probabilité vers 0 et $\Phi_n(t)$ converge en probabilité vers $\Phi(t)$ pour tout $t\in\mathbb{R}$.

(2) On suppose dans cette question que $E=\mathbb{R}$ et que μ est caractérisée par ses moments.
   (a) On suppose que pour tout $k\geq 1$, $\int x^k \mathrm{d}M_n$ converge en probabilité vers $\int x^k \mathrm{d}\mu$. Montrer que M_n converge en loi vers μ.
   (b) On suppose que pour tout $k\geq 1$, $\mathbb{E}[\int x^k \mathrm{d}M_n]$ converge vers $\int x^k \mathrm{d}\mu$ et que $\mathrm{Var}(\int x^k \mathrm{d}M_n)\to 0$. Montrer que M_n converge en loi vers μ.

Deuxième partie. On suppose dans la suite que les variables aléatoires $(M_n)_{n\geq 1}$ sont définies sur un même espace de probabilité $(\Omega,\mathcal{F},\mathbb{P})$.

(3) Démontrer que les quatre assertions suivantes sont équivalentes :
   (a) pour presque tout $\omega\in\Omega$, M_n converge étoitement vers μ ;
   (b) pour presque tout $\omega\in\Omega$, pour tout $f \in \mathcal{C}_b(E,\mathbb{R})$, $\int f \mathrm{d}M_n \to \int f \mathrm{d}\mu$;
   (c) pour tout $f \in \mathcal{C}_b(E,\mathbb{R})$, pour presque tout $\omega\in\Omega$, $\int f \mathrm{d}M_n \to \int f \mathrm{d}\mu$;
   (d) lorsque $E=\mathbb{R}$, pour tout $t\in\mathbb{R}$, on a presque sûrement $\int e^{itx}\mathrm{d}M_n(\mathrm{d}x) \to \int e^{itx}\mu(\mathrm{d}x)$.

(4) On suppose dans cette question que $E=\mathbb{R}$ et que μ est caractérisée par ses moments. On suppose que pour tout $k\geq 1$, $\int x^k \mathrm{d}M_n$ converge pour presque tout $\omega\in\Omega$ vers $\int x^k \mathrm{d}\mu$. Montrer que pour presque tout $\omega\in\Omega$, M_n converge étroitement vers μ.

(5) Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires i.i.d. de loi μ. Que dire du comportement, lorsque $n\to\infty$, de $M_n=\frac{1}{n}\sum_{k=1}^n \delta_{X_k}$?

Éléments de correction :

(1) L'équivalence entre (b) et (c) provient du fait que la convergence en loi et la convergence en probabilité vers une constante sont équivalentes.

Montrons que (c) implique (d). Soit $f \in \mathcal{C}_b(E,\mathbb{R})$. Pour montrer que $\int f \mathrm{d}M_n$ converge en probabilité vers $\int f \mathrm{d}\mu$ on va montrer que pour toute extraction ϕ, il existe une extraction ψ telle que que $\int f \mathrm{d}M_{\phi\circ\psi(n)}$ converge presque sûrement vers $\int f \mathrm{d}\mu$. Soit donc ϕ extraction. Comme

M_n converge en probabilité vers μ, il existe une extraction ψ telle que presque sûrement $M_{\phi\circ\phi(n)}$ converge (pour d_{LP}) vers μ. Alors presque sûrement $M_{\phi\circ\phi(n)}$ converge étroitement vers μ, donc presque sûrement $\int f \mathrm{d}M_{\phi\circ\psi(n)}$ converge vers $\int f \mathrm{d}\mu$.

Alternativement, pour montrer que (c) implique (d) on peut remarquer que pour $f \in \mathcal{C}_b(E,\mathbb{R})$, l'application $\nu \mapsto |\int f \mathrm{d}\nu$ est continue, et que la convergence en probabilité est stable par composition par une fonction continue.

Montrons que (d) implique (c). Soit $\varepsilon > 0$ et montrons que $\mathbb{P}(\mathrm{d}_{\mathrm{LP}}(M_n,\mu) > \varepsilon) \to 0$, c'est-à-dire que $\mathbb{P}(M_n \notin B(\mu,\varepsilon)) \to 0$. D'après le fait donné en début d'exercice, il existe $f_1,\ldots,f_K \in \mathcal{C}_b(E,\mathbb{R})$ et $\delta > 0$ tels que

$$\left\{\nu \in \mathcal{M}_1(E) : \left|\int f_i d\nu - \int f_i d\mu\right| < \delta \text{ pour tout } 1 \le i \le K\right\} \subset B(\mu,\varepsilon).$$

Alors

$$\mathbb{P}(M_n \notin B(\mu,\varepsilon)) \le \sum_{i=1}^{K} \mathbb{P}\left(\left|\int f_i \mathrm{d}M_n - \int f_i \mathrm{d}\mu\right| > \delta\right)$$

qui tend vers 0 par hypothèse.

Le fait que (a) implique (d) provient immmédiatement de l'inégalité de Markov.

Le fait que (d) implique (a) provient du fait que la suite $(\int f \mathrm{d}M_n)_{n\ge1}$ est uniformément intégrable (car bornée par $\|f\|_\infty$).

Montrons que (c) implique (e). Soit $t \in \mathbb{R}$. Il suffit de montrer que pour toute extraction ϕ, il existe une extraction ψ telle que presque sûrement $\Phi_{\phi\circ\psi(n)}(t)$ converge vers $\Phi(t)$ et $\sup_{|t|\le 1} |\Phi_{\phi\circ\psi(n)}(t) - \Phi(t)|$ converge vers 0. À cet effet, par (c), pour toute extraction ϕ, il existe une extraction ψ telle que p.s. $M_{\phi\circ\psi(n)}$ converge étroitement vers μ. D'après l'Exercice 6 (2) de la Feuille 1 d'exercices, on a bien que p.s. $\Phi_{\phi\circ\psi(n)}$ converge vers $\Phi(t)$ sur tout compact.

Montrons que (e) implique (d). Soit $f \in \mathcal{C}_b(E,\mathbb{R})$. Il suffit de montrer que pour toute extraction ϕ, il existe une extraction ψ telle que presque sûrement $\int f \mathrm{d}M_{\phi\circ\psi(n)}$ converge en probabilité vers $\int f \mathrm{d}\mu$.
Soit ϕ extraction. Vérifions d'abord qu'il existe une extraction ψ telle que p.s. $(M_{\phi\circ\psi(n)})_{n\ge1}$ est tendue. Soit $\delta > 0$ et ψ extraction telle que $\sup_{|t|\le\delta} |\Phi_{\phi\circ\psi(n)}(t) - \Phi(t)|$ converge presque sûrement vers 0. Comme dans le cours 1, en utilisant l'inégalité

$$M_{\phi\circ\psi(n)}(\mathbb{R}\backslash[-A,A]) \le cA \int_{-1/A}^{1/A} (1 - \mathrm{Re}\Phi_{\phi\circ\psi(n)}(t))\mathrm{d}t$$

en passant à la limite lorsque $n \to \infty$ et en utilisant la continuité de Φ en 0, on conclut que $(M_{\phi\circ\psi(n)})_{n\ge1}$ est tendue.
Soit maintenant $f \in \mathcal{C}_b(\mathbb{R},\mathbb{R})$ et vérifions que $\int f \mathrm{d}M_n$ converge en probabilité vers $\int f \mathrm{d}\mu$ en trouvant une extraction Ψ telle que $\int f \mathrm{d}M_{\phi\circ\psi(n)}$ converge en probabilité vers $\int f \mathrm{d}\mu$ (cela suffit, cf Exercice 1). Soit $\varepsilon > 0$. D'après le paragraphe précédent, par tension, p.s. on peut trouver $A > 0$ tel que $\mu(\mathbb{R}\backslash[-A,A]) \le \varepsilon$ et $M_{\phi\circ\psi(n)}(\mathbb{R}\backslash[-A,A]) \le \varepsilon$ pour tout $n \ge 1$. A fortiori, si g est une

fonction continue à support compact inclus dans $[-A,A]$, on a

$$\left|\int f \mathrm{d}M_{\phi\circ\psi(n)} - \int f \mathrm{d}\mu\right| \leq 2\varepsilon\|f\|_\infty + \left|\int g \mathrm{d}M_{\phi\circ\psi(n)} - \int g \mathrm{d}\mu\right|.$$

Sans perte de géneralité on peut donc supposer que f est à support compact. D'après le théorème de Stone-Weierstrass, on peut trouver des complexes $(\alpha_j)_{1\leq j\leq k}$ et des réels $(t_j)_{1\leq j\leq k}$ tels que

$$\sup_{|x|\leq A}\left|f(x) - \sum_{j=1}^{k}\alpha_j e^{it_j x}\right| \leq \varepsilon.$$

Alors

$$\left|\int_{-A}^{A} f(x) M_{\phi\circ\psi(n)}(\mathrm{d}x) - \int_{-A}^{A} f(x)\mu(\mathrm{d}x)\right| \leq 2\varepsilon + \sum_{j=1}^{k}|\alpha_j||\Phi_{\phi\circ\psi(n)}(t_j) - \phi(t_j)|,$$

qui tend en probabilité vers 0 par hypothèse. Le résultat désiré en découle.

Remarque. Strico-sensu, il faut faire attention car les variables aléatoires $(M_n)_{n\geq 1}$ ne sont pas forcément définies sur le même espace de probabilité, de sorte qu'on ne peut pas parler de convergence p.s. (ce point m'a été signalé par Dimitri Faure). Cependant, sans perte de généralité, on peut supposer que les variables aléatoires $(M_n)_{n\geq 1}$ sont définies sur le même espace de probabilité, par exemple en considérant une suite $(M_n')_{n\geq 1}$ de variables aléatoires indépendantes définies sur le même espace de probabilité telles que pour tout $n\geq 1$ les lois de M_n et M_n' sont les mêmes (on utilise l'existence de mesures produit), et en travaillant avec M_n' à la place de M_n (les propriétés considérées ne dépendent que des lois individuelles).

(2) (a) Démontrons que M_n converge en probabilité vers μ en montrant que pour toute extraction ϕ il existe une extraction ψ telle que presque sûrement $M_{\phi\circ\psi(n)}$ converge vers M (i.e. presque sûrement $\mathrm{d}_{\mathrm{LP}}(M_{\phi\circ\psi(n)}, M) \to 0$). Soit donc ϕ extraction. Par procédé diagonal, il existe une extraction ψ telle que p.s. pour tout $k\geq 1$, $\int x^k \mathrm{d}M_{\phi\circ\psi(n)}$ converge vers $\int x^k \mathrm{d}\mu$. D'après un exercice de la feuille d'exercices, p.s. $M_{\phi\circ\psi(n)}$ converge étroitement vers M, donc p.s. $\mathrm{d}_{\mathrm{LP}}(M_{\phi\circ\psi(n)}, M) \to 0$.

Alternativement, sans redéfinir les variables aléatoires sur le même espace de probabilité, il est possible de vérifier que (M_n) est tendue et pour l'identification de la limite utiliser le principe des lois accompagnantes en considérant la fonctionnelle $\nu \mapsto \int \max(\min(x^k, K), -K)\mathrm{d}\nu$ avec $K\to\infty$.

(b) D'après l'inégalité de Bienaymé-Tchebychev, pour $\varepsilon > 0$ on a

$$\begin{aligned}
&\mathbb{P}\left(\left|\int x^k \mathrm{d}M_n - \int x^k \mathrm{d}\mu\right| \geq 2\varepsilon\right)\\
&\qquad \leq \mathbb{P}\left(\left|\int x^k \mathrm{d}M_n - \mathbb{E}\left[\int x^k \mathrm{d}M_n\right]\right| \geq \varepsilon\right) + \mathbb{P}\left(\left|\mathbb{E}\left[\int x^k \mathrm{d}M_n\right] - \int x^k \mathrm{d}\mu\right| \geq \varepsilon\right)\\
&\qquad \leq \frac{\mathrm{Var}(\int x^k \mathrm{d}M_n)}{\varepsilon^2} + \mathbb{1}_{|\mathbb{E}[\int x^k \mathrm{d}M_n] - \int x^k \mathrm{d}\mu|\geq\varepsilon}
\end{aligned}$$

qui tend vers 0 par hypothèse, et on peut appliquer la question (a).

(3) L'équivalence entre (a) et (b) est simplement la définition de la convergence étroite.

Le fait que (b) implique (c) est clair.

Pour montrer que (c) implique (b), fixons $\varepsilon > 0$ et montrons que p.s., pour n assez grand $\mathrm{d}_{\mathrm{LP}}(M_n, \mu) \leq \varepsilon$. D'après le fait donné en début d'exercice, il existe $f_1, \ldots, f_K \in \mathcal{C}_b(E, \mathbb{R})$ et $\delta > 0$ tels que

$$\left\{\nu \in \mathcal{M}_1(E) : \left| \int f_i d\nu - \int f_i d\mu \right| < \delta \text{ pour tout } 1 \leq i \leq K \right\} \subset B(\mu, \varepsilon).$$

Par hypothèse, p.s. pour n assez grand

$$\left| \int f_i dM_n - \int f_i d\mu \right| < \delta \text{ pour tout } 1 \leq i \leq K$$

Ainsi, p.s. pour n assez grand $\mathrm{d}_{\mathrm{LP}}(M_n, \mu) \leq \varepsilon$.

Alternativement, sans redéfinir les variables aléatoires sur le même espace de probabilité, il est possible d'utiliser la question (3) du problème 5.

Le fait que (b) implique (d) est clair : sur l'événement de probabilité 1 où pour tout $f \in \mathcal{C}_b(E, \mathbb{R})$, $\int f \mathrm{d}M_n \to \int f \mathrm{d}\mu$ on a $\int e^{itx} \mathrm{d}M_n(\mathrm{d}x) \to \int e^{itx} \mu(\mathrm{d}x)$ pour tout $t \in \mathbb{R}$.

Pour montrer que (d) implique (b), l'idée est d'utiliser le théorème de Fubini pour l'interversion. Notons

$$A = \left\{(\omega, t) \in \Omega \times \mathbb{R} : \int e^{itx} \mathrm{d}M_n(\omega)(\mathrm{d}x) \nrightarrow \int e^{itx} \mu(\mathrm{d}x)\right\}$$

ainsi que ses sections

$$A_\omega = \{t \in \mathbb{R} : (\omega, t) \in A\}, \qquad A^t = \{\omega \in \Omega : (\omega, t) \in A\}.$$

Par hypothèse, pour tout $t \in \mathbb{R}$ on a $\mathbb{P}(A^t) = 0$. Appliquons alors le théorème de Fubini, en notant λ la mesure de Lebesgue :

$$0 = \int_{\mathbb{R}} \mathbb{P}(A^t) \lambda(\mathrm{d}t) = \int_{\Omega \times \mathbb{R}} \mathbb{1}_A \mathrm{d}\mathbb{P} \otimes \mathrm{d}\lambda = \int_\Omega \lambda(A_\omega) \mathrm{d}\mathbb{P}.$$

On en déduit que $\lambda(A_\omega) = 0$ pour presque tout $\omega \in \Omega$. Ainsi, p.s. pour λ presque tout $t \in \mathbb{R}$ on a $\int e^{itx} \mathrm{d}M_n(\omega)(\mathrm{d}x) \to \int e^{itx} \mu(\mathrm{d}x)$. On conclut avec la première question de l'exercice 2.

(4) C'est une application directe de l'exercice de la feuille d'exercices précédemment mentionné.

(5) D'après la loi forte des grands nombres, pour tout $f \in \mathcal{C}_b(E, \mathbb{R})$, pour presque tout $\omega \in \Omega$,

$$\int f \mathrm{d}M_n = \frac{1}{n} \sum_{k=1}^{n} f(X_k) \to \mathbb{E}[f(X_1)] = \int f \mathrm{d}\mu.$$

D'après la question (3), on en déduit que p.s. M_n converge étroitement vers μ (a fortiori M_n converge en probabilité et en loi vers μ).

□

* * *

Petit problème 5.

Première partie. Soit (E,d) un espace métrique séparable et $(x_n)_{n\geq 1}$ une suite dense. On note $\mathcal{M}_1(E)$ l'ensemble des mesures de probabilité sur E muni de sa tribu borélienne, et $\mathrm{Lip}_1(E)$ l'ensemble des fonctions de E dans $\mathbb{R}$ 1-lipschitziennes (c'est-à-dire telles que $d(f(x),f(y)) \leq |x-y|$ pour tous $x,y \in E$).

Soit $\mu_n, \mu \in \mathcal{M}_1(E)$. L'objectif de cette partie est de montrer que $\mu_n \implies \mu$ si et seulement si $\int f \mathrm{d}\mu_n \to \int f \mathrm{d}\mu$ pour toute fonction f de la forme

$$f(x) = \max(c_0, c_1 - d(x,x_1), \ldots, c_n - d(x,x_n))$$

avec $n \geq 1$ et $c_0, \ldots, c_n \in \mathbb{R}$.

(1) Justifier l'implication.

(2) Vérifier que pour montrer que $\mu_n \implies \mu$ il suffit de vérifier que $\int f \mathrm{d}\mu_n \to \int f \mathrm{d}\mu$ pour toute fonction $f \in \mathrm{Lip}_1(E)$ bornée.

(3) Démontrer la réciproque.
Indication. Pour $f \in \mathrm{Lip}_1(E)$ bornée telle que $f \geq c_0$, on pourra vérifier que pour tout $x \in E$ on a

$$f(x) = \sup_{n\geq 1} \max(c_0, f(x_n) - d(x,x_n)).$$

(4) Si E est borné (c'est-à-dire si $\sup_{x,y\in E} d(x,y) < \infty$), montrer que $\mu_n \implies \mu$ si et seulement si $\int f \mathrm{d}\mu_n \to \int f \mathrm{d}\mu$ pour toute fonction $f(x)$ polynomiale en les variables $d(x,x_1), \ldots, d(x,x_n)$ pour tout $n \geq 1$.

Deuxième partie. Considérons l'ensemble E des fonctions mesurables de $[0,1] \to \mathbb{R}$, considérées à égalité presque partout près (c'est-à-dire qu'on identifie deux fonctions égales presque partout) muni de la distance

$$\mathrm{d}(f,g) = \int_0^1 \min(1, |f(t) - g(t)|) \mathrm{d}t. \tag{1}$$

(5) En notant λ la mesure de Lebesgue, montrer que $f_n \to f$ dans E si et seulement pour tout $\varepsilon > 0$ on a $\lambda(\{x \in [0,1] : |f_n(x) - f(x)| \geq \varepsilon\}) \to 0$ quand $n \to \infty$.

(6) On suppose que $f_n \to f$ dans E. Montrer qu'on peut trouver une extraction ϕ telle que $f_{\phi(n)}$ converge presque partout vers f.

(7) Démontrer que E est polonais.

Troisième partie. Considérons l'espace métrique $\mathcal{C}$ des fonctions continues de $[0,1] \to \mathbb{R}$ muni de la distance d définie par (1).

(8) Vérifier que $(\mathcal{C}, d)$ est séparable et borné. Est-il complet ? Justifier votre réponse.

Soit $(X_n)_{n\geq 1}$ des variables aléatoires à valeurs dans $\mathcal{C}$ et X une variable aléatoire à valeurs dans $\mathcal{C}$.

(9) Démontrer que si X_n converge en loi vers X, alors pour tout $k \geq 1$, si $(U_i)_{1\leq i\leq k}$ sont i.i.d. uniformes sur $[0,1]$ (aussi indépendantes de X_n, X) on a la convergence en loi de $(X_n(U_1), \ldots, X_n(U_k))$ vers $(X(U_1), \ldots, X(U_k))$.

(10) On suppose que les marginales fini-dimensionnelles de X_n convergent en loi vers celles de X. Démontrer que X_n converge en loi vers X.

(11) *[Question bonus hors barème]* Démontrer que X_n converge en loi vers X si et seulement si, pour tout $k \geq 1$, en notant $V_1^k, \ldots, V_k^k$ le réarrangement croissant de k variables aléatoires i.i.d. uniformes sur $[0,1]$ (aussi indépendantes de X_n et X) on a la convergence en loi de $(X_n(V_1^k), \ldots, X_n(V_k^k))$ vers $(X(V_1^k), \ldots, X(V_k^k))$.

Éléments de correction :

Première partie.

(1) Ceci simplement du fait qu'une fonction de type $f(x) = \max(c_0, c_1 - d(x, x_1), \ldots, c_n - d(x, x_n))$ est continue bornée.

(2) Par linéarité de l'intégrale, on a alors $\int f \, \mathrm{d}\mu_n \to \int f \, \mathrm{d}\mu$ pour toute fonction f lipschitzienne bornée, ce qui implique que $\mu_n \implies \mu$ par le théorème de porte-manteau.

(3) Commençons par vérifier l'indication. Soit $f \in \mathrm{Lip}_1(E)$ bornée telle que $f \geq c_0$.

Si $n \geq 1$, on a $f(x) \geq c_0$ et $f(x) \geq f(x_n) - d(x, x_n)$ car f est 1-lipschtzienne. Donc $f(x) \geq \sup_{n\geq 1} \max(c_0, f(x_n) - d(x, x_n))$.

Ensuite fixons $\varepsilon > 0$ et vérifions qu'on peut trouver $n \geq 1$ tel que $f(x) \leq \max(c_0, f(x_n) - d(x, x_n)) + \varepsilon$. Par densité, soit x_n tel que $d(x, x_n) \leq \varepsilon/2$. Alors

$$f(x) - \varepsilon = f(x_n) + f(x) - f(x_n) - \varepsilon \leq f(x_n) + d(x, x_n) - 2d(x, x_n) = f(x_n) - d(x, x_n),$$

d'où le résultat.

Pour démontrer la réciproque, soit $f \in \mathrm{Lip}_1(E)$ bornée et soit c_0 tel que $f \geq c_0$. Posons

$$f_k(x) = \max_{1\leq i\leq k} \max(c_0, f(x_i) - d(x, x_i)),$$

de sorte que f_n converge simplement vers $f(x)$ en étant croissante. Par ailleurs, f_n est 1-lipschitzienne bornée, donc

$$\int f_i \mathrm{d}\mu_n \quad \underset{n\to\infty}{\longrightarrow} \quad \int f_i \mathrm{d}\mu.$$

Comme $\int f_i d\mu_n \leq \int f \,\mathrm{d}\mu_n$, on en déduit que

$$\int f_i \mathrm{d}\mu \leq \liminf_{n\to\infty} \int f \,\mathrm{d}\mu_n.$$

Par convergence monotone, en faisant $i \to \infty$ on conclut que

$$\int f \,\mathrm{d}\mu \leq \liminf_{n\to\infty} \int f \,\mathrm{d}\mu_n.$$

En remplaçant f par $-f$, le résultat désiré en découle.

(4) Comme pour (1), l'implication est immédiate. Pour la réciproque, soit $A > 0$ tel que $d(x,y) \leq A$ pour tout $x, y \in E$. On remarque que l'application

$$\Phi: \begin{array}{ccc} [0,A]^n & \longrightarrow & \mathbb{R} \\ (t_1,\ldots,t_n) & \mapsto & \max(c_0, c_1 - t_1, \ldots, c_n - t_n) \end{array}$$

peut être approchée uniformément par des polynômes en $t_1,\ldots,t_n$ d'après le théorème de Stone-Weierstrass, ce qui permet de conclure.

Deuxième partie.

(5) Pour l'implication, on écrit, pour $\varepsilon \in]0,1[$, en utilisant l'inégalité de Markov :

$$\lambda\{x\in[0,1] : |f_n(x)-f(x)| \geq \varepsilon\} = \lambda\{x \in [0,1] : \min(1, |f_n(x)-f(x)|) \geq \varepsilon\} \leq \frac{1}{\varepsilon}\int_0^1 \min(1,|f_n(t)-f(t)|)\mathrm{d}t,$$

qui tend vers 0.

Pour la réciproque, supposons que $\lambda(\{x \in [0,1] : |f_n(x) - f(x)| \geq \varepsilon\}) \to 0$ pour tout $\varepsilon > 0$. Soit $\varepsilon \in]0,1[$ et notons $A = \{x \in [0,1] : |f_n(x) - f(x)| \geq \varepsilon\}$. Alors

$$\int_0^1 \min(1, |f_n(t) - f(t)|)\mathrm{d}t \leq \int_A 1\,\mathrm{d}t + \int_{A^c} \varepsilon \mathrm{d}t \leq \lambda(A) + \varepsilon,$$

de sorte que $\mathrm{d}(f_n, f) \leq \mathrm{d}t \leq 2\varepsilon$ pour n assez grand.

(6) Supposons que $f_n \to f$ dans E. On choisit une extraction ϕ telle que pour tout $n \geq 1$ on ait

$$\lambda\left(\left\{x \in [0,1] : |f_{\phi(n)}(x) - f(x)| \geq \frac{1}{2^n}\right\}\right) \leq \frac{1}{2^n}.$$

D'après le lemme de Borel-Cantelli, $f_{\phi(n)}$ converge presque partout vers f.

(7) La séparabilité provient du fait que $L^1([0,1], \mathrm{d}\lambda)$ et que la convergence dans $L^1([0,1], \mathrm{d}\lambda)$ implique la convergence pour d car $d(f,g) \leq \|f-g\|_1$.

Pour la complétude, soit (f_n) une suite de Cauchy dans E. Il suffit de monter qu'elle a une valeur d'adhérence. En utilisant la question précédente, soit ϕ une extraction telle que $f_{\phi(n)}$ converge presque partout vers une fonction notée f. Le théorème de convergence dominée entraîne immédiatement que $\mathrm{d}(f_{\phi(n)}, f) \to 0$, ce qui conclut.

Troisième partie. Tout d'abord, remarquons que pour tout $t \in [0,1]$,

$$\begin{array}{rccc} \Pi_t : & \mathcal{C} & \longrightarrow & \mathbb{R} \\ & f & \mapsto & f(t) \end{array}$$

est mesurable. En effet, en posant pour tout $\varepsilon > 0$

$$\begin{array}{rccc} \Pi_t^\varepsilon : & \mathcal{C} & \longrightarrow & \mathbb{R} \\ & f & \mapsto & \frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon} f(t)\mathrm{d}t \end{array},$$

on voit que Π_t^ε est continu (c'est par exemple une conséquence de (6) et du théorème de convergence dominée) donc mesurable, et $\Pi_t = \lim_{\varepsilon\to 0}\Pi_t^\varepsilon$ est mesurable comme limite d'applications mesurables.

(8) La séparabilité provient par exemple du fait que l'espace métrique des fonctions continues de $[0,1]$ dans $\mathbb{R}$ munies de la norme uniforme est séparable, combiné avec le fait $d(f,g) \leq \|f-g\|_\infty$ pour tout $f,g \in \mathcal{C}$.

Ensuite, il est clair que $d(f,g) \leq 1$ pour tous $f,g \in \mathcal{C}$.

Enfin, $(\mathcal{C}, d)$ n'est pas complet. En effet, si on pose

$$f_n(x) = \begin{cases} 0 & \text{si } x \leq \frac{1}{2} \\ \frac{1}{2} + \frac{n}{2}\left(x - \frac{1}{2}\right) & \text{si } x \leq \frac{1}{2} + \frac{1}{n} \\ 1 & \text{si } x \geq \frac{1}{2} + \frac{1}{n}, \end{cases}$$

la suite f_n est de Cauchy pour $\|\cdot\|_1$ donc pour d, mais elle ne converge pas. En effet, si elle converge vers une fonction continue $f \in \mathcal{C}$, elle converge aussi vers f dans E. Or $f_n \to \mathbb{1}_{[1/2,1]}$ pour $\|\cdot\|_1$ donc dans E. Ainsi $f = \mathbb{1}_{[1/2,1]}$ presque partout, ce qui est absurde.

(9) Soit $F : \mathbb{R}^k \to \mathbb{R}$ continue bornée. Il suffit de montrer que

$$\mathbb{E}[F(X_n(U_1), \ldots, X_n(U_k))] \quad \underset{n\to\infty}{\longrightarrow} \quad \mathbb{E}[F(X(U_1), \ldots, X(U_k))].$$

À cet effet écrivons

$$\mathbb{E}[F(X_n(U_1), \ldots, X_n(U_k))] = \mathbb{E}\left[\int_{[0,1]^k} F(X_n(t_1), \ldots, X_n(t_k))\mathrm{d}t_1 \cdots \mathrm{d}t_k\right] = \mathbb{E}\left[\Phi(X_n)\right],$$

où

$$\begin{array}{rccc} \Phi : & \mathcal{C} & \longrightarrow & \mathbb{R} \\ & f & \mapsto & \int_{[0,1]^k} F(f(t_1), \ldots, f(t_k))\mathrm{d}t_1 \cdots \mathrm{d}t_k \end{array}.$$

Il suffit alors de remarquer que Φ est continue bornée. En effet, le caractère bornée provient du fait que F est bornée. Pour la continuité, soit $f_n \to f$ dans $\mathcal{C}$. Puisque $(\Phi(f_n))_{n\geq 1}$ est bornée, il suffit de montrer qu'elle a unique valeur d'adhérence. D'après la question (6), il existe une extraction ϕ telle que $f_{\phi(n)} \to f$ presque partout. Alors par convergence dominée $\Phi(f_{\phi(n)}) \to \Phi(f)$, ce qui conclut.

(10) Soit $(f_n)_{n\geq 1}$ une suite dense dans $\mathcal{C}$. D'après la question (4), il suffit de vérifier que pour tout $k \geq 1$ et tout polynôme P en k variables on a

$$\mathbb{E}[P(\mathrm{d}(X_n, f_1), \dots, \mathrm{d}(X_n, f_k))] \underset{n\to\infty}{\longrightarrow} \mathbb{E}[P(\mathrm{d}(X, f_1), \dots, \mathrm{d}(X, f_k))]. \tag{2}$$

Or

$$\mathbb{E}[P(\mathrm{d}(X_n, f_1), \dots, \mathrm{d}(X_n, f_k))] = \mathbb{E}\left[P\left(\int_0^1 \min(1, |X_n(t) - f_1(t)|)\mathrm{d}t, \dots, \int_0^1 \min(1, |X_n(t) - f_k(t)|)\mathrm{d}t\right)\right]$$

est une combinaison linéaire d'éléments de la forme

$$\int_{[0,1]^N} \mathbb{E}\left[\prod_{i=1}^N \min(1, |X_n(t_i) - g_i(t_i)|)\right] \mathrm{d}t_1 \cdots \mathrm{d}t_N$$

avec $N \geq 1$ et $g_i \in \{f_1, \dots, f_k\}$ pour tout $1 \leq i \leq N$. Or, à $t_1, \dots, t_N \in [0,1]$ fixés, on a

$$\mathbb{E}\left[\prod_{i=1}^N \min(1, |X_n(t_i) - g_i(t_i)|)\right] = \mathbb{E}[\Psi(X_n(t_1), \dots, X_n(t_N))]$$

avec

$$\begin{array}{cccc} \Psi : & \mathbb{R}^N & \longrightarrow & \mathbb{R} \\ & (x_1, \dots, x_N) & \mapsto & \prod_{i=1}^N \min(1, x_i - g_i(t_i)) \end{array}$$

continue bornée. Ainsi, par convergence dominée

$$\int_{[0,1]^N} \mathbb{E}\left[\prod_{i=1}^N \min(1, |X_n(t_i) - g_i(t_i)|)\right] \mathrm{d}t_1 \cdots \mathrm{d}t_N \underset{n\to\infty}{\longrightarrow} \int_{[0,1]^N} \mathbb{E}\left[\prod_{i=1}^N \min(1, |X(t_i) - g_i(t_i)|)\right] \mathrm{d}t_1 \cdots \mathrm{d}t_N.$$

Ceci démontre (2) et conclut.

(11) Dans le cadre général de l'espace E, voir la Proposition 29 (en annexe) de Aldous, D., & Pitman, J. (2002). Invariance principles for non-uniform random mappings and trees. In Asymptotic Combinatorics with Application to Mathematical Physics (pp. 113-147). Springer, Dordrecht.

$\square$