Intégration et Probabilités – TD 1
Corrigé partiel

Exercice 1. Pour la deuxième partie de l'exercice, on raisonne par l'absurde. Supposons que $\mathcal{F} = \cup_{n\in\mathbb{N}}\mathcal{F}_n$ soit une tribu. Puisque $\{2n\} \in \mathcal{F}_{2n+1}$, $\{2n\} \in \mathcal{F}$, et comme $\mathcal{F}$ est supposée être une tribu, par réunion dénombrable, $2\mathbb{N} \in \mathcal{F}$, c'est-à-dire qu'il existe $n_0 \in \mathbb{N}$ tel que $2\mathbb{N} \in \mathcal{F}_{n_0}$. Or, il est immédiat que les seuls éléments de cardinal infini de $\mathcal{F}_{n_0}$ sont de la forme $\mathbb{N} \setminus A$, où A est une partie de $\{0,1,\dots,n_0-1\}$.

Exercice 2. $\liminf_n A_n$ est la notion de limite la plus restrictive, et correspond aux éléments qui sont dans tous les A_n, sauf un nombre fini d'entre eux. $\limsup_n A_n$ est une notion de limite moins restrictive, qui correspond aux éléments qui sont dans un nombre infini de A_n.

On a l'égalité $\mathbb{I}_{\liminf_n A_n} = \liminf_n \mathbb{I}_{A_n}$; en effet, si $x \in \liminf_n A_n$, alors il existe n_0 tel que pour tout $n \geq n_0$, $x \in A_n$, de sorte que $\inf_{k\geq n} \mathbb{I}_{A_n}(x) = 1$ dès que $n \geq n_0$. En prenant la limite, $\liminf_n \mathbb{I}_{A_n}(x) = 1$. De même, quand $x \notin \liminf_n A_n$, $\liminf_n \mathbb{I}_{A_n}(x) = 0$. L'égalité avec des lim sup se démontre de façon similaire ou en passant au complémentaire.

On note que

$$\mu\left(\cap_{k\geq n} A_k\right) \leq \inf_{k\geq n} \mu(A_k) \ ,$$

et ce, en utilisant le fait que $A \subseteq B$ implique $\mu(A) \leq \mu(B)$. Le résultat s'obtient ensuite en passant à la limite quand $n \to +\infty$ dans les deux membres, grâce à la remarque suivante. Notant $B_n = \cap_{k\geq n} A_k$, on remarque que $(B_n)_{n\geq 1}$ est croissante (au sens de l'inclusion), et donc

$$\mu\left(\liminf_n A_n\right) = \mu\left(\cup_{n\geq 1} B_n\right) = \lim_n \mu(B_n) \ .$$

L'inégalité pour les lim sup s'obtient en passant au complémentaire, car $(\liminf_n A_n^c)^c = \limsup_n A_n$, et en simplifiant les deux membres par la quantité finie $\mu(\cup_{n\geq 1} A_n)$. Cette dernière condition est nécessaire, considérer $A_n = \{n, n+1, \dots\}$ pour s'en persuader.

Pour le lemme de Borel-Cantelli, on a, par sous-additivité,

$$\mu\left(\cup_{k\geq n} A_k\right) \leq \sum_{k\geq n} \mu(A_k) \ .$$

Un passage à la limite donne le résultat, car le second membre tend vers 0 quand n tend vers $+\infty$ (reste d'une série convergente), et le premier membre est plus grand que $\mu(\limsup_t A_t)$ (pour tout n).

Pour l'application, on remarque que x est proche d'une fraction p/q à $1/q^{2+\varepsilon}$ près seulement si qx est proche de $\mathbb{N}$ à $1/q^{1+\varepsilon}$ près. Or, notant

$$A_q = \{x \in [0,1] \mid d(qx,\mathbb{N}) < 1/q^{1+\varepsilon}\} \ ,$$

on remarque que

$$A_q \subseteq [0,1] \cap \left(\bigcup_{p=0}^{q} \left[p/q - 1/q^{2+\varepsilon}, p/q + 1/q^{2+\varepsilon}\right] \right) \ ,$$

de sorte que la mesure de A_q est plus petite que $2/q^{1+\varepsilon}$. Par conséquent, $\sum_{q\geq 1} \lambda(A_q) < +\infty$, et le lemme de Borel-Cantelli conclut.

Pour la question subsidiaire, considérer respectivement la racine cubique de 2, et $\sum_{n\geq 1} 10^{-n!}$.

Exercice 3. Grâce au critère de Cauchy dans $\mathbb{R}$, on voit que l'ensemble considéré est

$$\bigcap_{k\in\mathbb{N}} \bigcup_{\ell\geq 1} \bigcap_{n,m\geq \ell} \{x \in E \mid |f_n(x) - f_m(x)| \leq 1/k\} \ ,$$

et ce dernier s'obtient par un nombre dénombrable d'opérations élémentaires sur des ensembles mesurables.

Exercice 4. Soit $\delta > 0$ tel que $I =]-\delta,\delta[\subseteq U$, et notons $J =]-\delta/2,\delta/2[$. Pour $r \in \mathbb{R}$, on note $A_r = f^{-1}(r + J)$. Alors $\cup_{r\in\mathbb{Q}} A_r$ est mesurable, de mesure non nulle. En effet,

$$\cup_{r\in\mathbb{Q}} A_r = f^{-1}\left(\cup_{r\in\mathbb{Q}}\, r + J\right) = f^{-1}(\mathbb{R}) = \mathbb{R} \ ,$$

et $\mathbb{R}$ est de μ-mesure non nulle. En particulier, il existe nécessairement un $r_0 \in \mathbb{Q}$ tel que $\mu(A_{r_0}) > 0$. Prenons $A = A_{r_0}$. Pour tous $x,y \in A$, $f(x),f(y) \in r + J$, et ainsi $f(x) - f(y) \in I \subseteq U$.

Quelques exercices d'approfondissement

Exercice 5. Le raisonnement se fait en trois temps. Premier temps : soit $A \in \mathcal{A}$, alors

$$A = \bigcup_{\dot{x}\subseteq A} \dot{x} \ .$$

En effet, si $x \in A$, alors par définition de $\dot{x}$, $\dot{x} \subseteq A$, et comme par ailleurs $x \in \dot{x}$, on a l'égalité proposée. Ainsi, tout élement de $\mathcal{A}$ s'écrit comme réunion d'atomes. Mais il s'agit de montrer également l'unicité de cette écriture.

Deuxième temps : à cet effet, nous allons montrer que si $x,y \in \mathcal{A}$, alors ou $\dot{x} = \dot{y}$, ou $\dot{x} \cap \dot{y} = \emptyset$ (i.e., les $\dot{x}$, $x \in \mathcal{A}$, partitionnent E). Cela se fait directement en définissant une relation d'équivalence $\mathcal{R}$ par

$$x\mathcal{R}y \quad\Longleftrightarrow\quad y \in \dot{x} \quad\Longleftrightarrow\quad (\forall A \in \mathcal{A},\quad x \in A \ \Rightarrow y \in A) \ .$$

Seule la symétrie de $\mathcal{R}$ n'est pas tout à fait évidente, elle se prouve en notant que $x \in A \ \Rightarrow y \in A$ équivaut à $y \in A^c \ \Rightarrow x \in A^c$, or $\mathcal{A}$ est stable par passage au complémentaire. Les classes de $\mathcal{R}$ étant précisément les atomes, on conclut à la partition de l'espace par les atomes.

Troisième temps : soit T l'ensemble des atomes de $\mathcal{A}$. Nous avons déjà introduit une injection de $\mathcal{A}$ dans l'ensemble des parties de T, noté $\mathcal{P}(T)$. (Cette injection associe à $A \in \mathcal{A}$ son unique décomposition en atomes.) Cela est vrai en toute généralité, mais nous allons montrer que dès que $\mathcal{A}$ est dénombrable, cette injection est en réalité une bijection. En effet, par intersection dénombrable, tout atome est élément de $\mathcal{A}$. En particulier, il y a au plus un nombre dénombrable d'atomes, de sorte que toute réunion d'atomes est encore dans $\mathcal{A}$.

En conclusion, $\mathcal{A}$ a le même cardinal que $\mathcal{P}(T)$, qui n'est dénombrable que lorsque T est fini. Ainsi n'existe-t-il pas de tribu de cardinal infini dénombrable.

Exercice 6. D'une manière générale, $\mathcal{B}(X) \otimes \mathcal{B}(Y) \subseteq \mathcal{B}(X \times Y)$. En effet,

$$\mathcal{B}(X) \otimes \mathcal{B}(Y) = \sigma\left\{A \times B,\ A \in \mathcal{B}(X) \text{ et } B \in \mathcal{B}(Y)\right\} \ ;$$

or $A \times B = (A \times Y) \cap (X \times B)$, et par symétrie il suffit de montrer que $A \times Y \in \mathcal{B}(X \times Y)$ pour tout $A \in \mathcal{B}(X)$. Considérons

$$\{C \in \mathcal{B}(X) \mid C \times Y \in \mathcal{B}(X \times Y)\} \ ;$$

c'est une tribu contenant les ouverts de X, c'est donc $\mathcal{B}(X)$ elle-même.

L'hypothèse de bases d'ouverts dénombrables nous permet de montrer l'inclusion réciproque. Ce qu'on appelle topologie produit sur $X \times Y$ est la topologie engendrée par les ensembles $U \times V$, où U et V sont des ouverts, respectivement de X et Y. Les ouverts de cette topologie sont donnés par des réunions (quelconques) d'ouverts $U \times V$. Or chacun de ces ouverts $U \times V$ s'écrit comme union dénombrable d'ouverts $U_i \times V_j$, où les U_i et V_j sont éléments des bases dénombrables d'ouverts respectives. Ainsi, les ouverts de $X \times Y$ sont finalement donnés par des réunions (dénombrables) de produits d'ouverts des deux bases dénombrables d'ouverts, de sorte que tout ouvert de $X \times Y$ est dans $\mathcal{B}(X) \otimes b(Y)$.

La topologie de la convergence uniforme sur tout compact de I est celle engendrée par les ensembles de la forme

$$B(f,R,K) = \{g \in C \mid \|f - g\|_{K,\infty} < R\} \ ,$$

où K est un compact inclus dans I, $R > 0$ et $f \in C$.

$\mathcal{C}_2 \subseteq \mathcal{C}_1$, car pour tout $x \in I$, π_x est continue (donc borélienne) pour $(C, \mathcal{C}_1)$. En effet, pour tous réels a et b, $\pi_x(]a,b[) = B((a+b)/2, |a-b|/2, \{x\})$.

Pour montrer l'inclusion réciproque, on remarque tout d'abord que la topologie de la convergence uniforme sur tout compact I est engendrée par les ouverts élémentaires (en nombre dénombrable) de la forme $B(P,r,K)$, où P est un polynôme à coefficients rationnels, $r \in \mathbb{Q}$, et K un compact à coordonnées rationnelles – par application du theorème d'approximation de Weierstrass. Or chacune des ces boules élementaires est $\mathcal{C}_2$-mesurable ; en effet, par uniforme continuité d'une application continue sur un compact,

$$\begin{aligned} B(P,r,K) &= \bigcup_{n \geq 1} \bigcap_{x \in K \cap \mathbb{Q}} \{g \in C \mid |P(x) - g(x)| \leq R - 1/n\} \\ & \bigcup_{n \geq 1} \bigcap_{x \in K \cap \mathbb{Q}} \pi_x^{-1}\left([P(x) - |R - 1/n|, P(x) + |R - 1/n|]\right) \ . \end{aligned}$$

Quelques exercices plus avancés

Cette section est corrigée très partiellement (les solutions n'ont de sens dans cette partie que cherchées pendant un temps suffisamment long), venez me voir pour que je vous corrige vos réponses ou que je vous donne des indications.

Exercice 7 [Ensembles de Cantor]. K est fermé comme intersection de fermés, et il est compact parce qu'il est en outre borné. Il est facile de construire une bijection entre $\{0,1\}^{\mathbb{N}}$ et K : K est non-dénombrable. Pour la totale disconnection (i.e., les composantes connexes de K sont réduites à des points), on rappelle que les connexes de $\mathbb{R}$ sont les intervalles, mais que comme K est *d'intérieur vide*, il ne peut contenir que des intervalles réduits à un point. (Pour voir que K est d'intérieur vide, noter que K est inclus dans la réunion d'intervalles K_n obtenue au rang n de la construction, et de même pour les intérieurs respectifs, avec la propriété supplémentaire que les composantes connexes de l'intérieur de K_n sont de longueur plus petite que 2^{-n}.) L'absence de points isolés se prouve en notant que, étant donné $x \in I$, les points de l'intersection de K et de la composante connexe de x dans K_n sont en nombre infini et sont tous à une distance de moins de $1/2^n$ de x.

Par récurrence et passage à la limite, la mesure de K est $\prod_{n\geq 0}(1-d_n)$ (ce nombre peut être nul ou strictement positif).

Exercice 8. L'ouvert $\cup_{n\geq 1}]n, n+1/2^n[$ est non borné, de mesure finie. $[0,1]\setminus\mathbb{Q}$ est de mesure 1, et d'intérieur vide. Le complémentaire d'un ensemble de Cantor de mesure non nulle est dense dans [0,1] et de mesure non pleine.

Tous les ensembles de Cantor sont homéomorphes à $\{0,1\}^{\mathbb{N}}$ (muni de la topologie-produit), de sorte qu'un Cantor de mesure nulle est homéomorphe à un Cantor de mesure strictement positive.

Soit q une numérotation de $\mathbb{Q}$. L'ouvert $\cup_{r\in\mathbb{Q}}]r, r+\varepsilon/2^{q(r)}[$ est dense dans [0,1], de mesure plus petite que ε. Le complémentaire d'un Cantor bien choisi serait dense dans [0,1], de mesure exactement égale à ε.

Intégration et Probabilités – TDs 2 et 3
Corrigé partiel

Exercice 1 (td 2). Dans un premier temps, on fixe $\eta > 0$ et on pose

$$A_{n,\eta} = \left\{x \in E \mid n^{-\alpha}\,|f(n\,x)| > \eta\right\} = \frac{1}{n}\left\{y \in E \mid n^{-\alpha}\,|f(y)| > \eta\right\} \ .$$

La mesure de $A_{n,\eta}$ est majorée par

$$\lambda(A_{n,\eta}) = \frac{1}{n}\lambda\left\{n^{-\alpha}\,|f| > \eta\right\} \leq \frac{1}{\eta n^{1+\alpha}} \int_{\mathbb{R}} |f|\,\mathrm{d}\lambda \ ,$$

de sorte que la série de terme général $\lambda(A_{n,\eta})$ est sommable. Par application du lemme de Borel-Cantelli,

$$\lambda\left(\limsup_{n\to\infty} A_{n,\eta}\right) = \lambda\left\{\limsup_{n\to\infty} \frac{|f(n\,x)|}{n^\alpha} \geq \eta\right\} = 0 \ ,$$

et on conclut par union dénombrable sur $\eta \in \{1/k,\, k \in \mathbb{N}^*\}$.

Exercice 2 (td 2). Lorsque f est intégrable $\mathbb{R} \to \mathbb{R}$, on définit les mesures positives $\mathrm{d}\nu_+ = f^+\,\mathrm{d}\lambda$ et $\mathrm{d}\nu_- = f^-\,\mathrm{d}\lambda$. Comme $f\,\mathrm{d}\lambda = \mathrm{d}\nu_+ - \mathrm{d}\nu_-$, on a $\nu_+(]a,b[) = \nu_-(]a,b[)$ pour tous a, b. Puisque ν_+ et ν_- sont boréliennes positives de masse finie, le théorème d'unicité des mesures donne $\nu_+ = \nu_-$, et ainsi,

$$\forall A \in \mathcal{B}(\mathbb{R}), \quad \int_A f\,\mathrm{d}\lambda = 0 \ .$$

On conclut à $f \leq 0$ λ-p.p. en prenant les $A_n = \{f \geq 1/n\}$, tous de mesure nulle, et en passant à l'union dénombrable croissante (et de même pour prouver que $f \geq 0$ λ-p.p.).

Pour f intégrable $\mathbb{R} \to \mathbb{C}$, on applique successivement le point précédent à $\operatorname{Re} f$ et $\operatorname{Im} f$.

Exercice 3 (td 2). Soit $\alpha \in \mathbb{C}$ de module $|\alpha| = 1$ tel que

$$\alpha \int_E f\,\mathrm{d}\mu = \left|\int_E f\,\mathrm{d}\mu\right| \ ;$$

par construction de l'intégrale, la partie réelle de l'intégrale de f est l'intégrale de la partie réelle de f, de sorte que

$$\int_E \alpha f\,\mathrm{d}\mu = \alpha \int_E f\,\mathrm{d}\mu = \left|\int_E f\,\mathrm{d}\mu\right| = \operatorname{Re}\int_E \alpha f\,\mathrm{d}\mu = \int_E \operatorname{Re}(\alpha f)\,\mathrm{d}\mu \leq \int_E |f|\,\mathrm{d}\mu \ ,$$

la dernière inégalité étant ici une égalité. Cela signifie que la fonction positive $|f| - \operatorname{Re}(\alpha f)$ est d'intégrale nulle sur E, donc elle est nulle μ-p.p., et le résultat demandé s'ensuit.

Exercice 4 (td 2). Soit $\varepsilon > 0$. Par convergence dominée (ou monotone),

$$\int_E |f|\,\mathbb{I}_{|f|\geq K}\,\mathrm{d}\mu \xrightarrow[K\to+\infty]{} 0 \ ,$$

de sorte que l'on peut prendre K_ε tel que

$$\int_E |f|\,\mathbb{I}_{|f|\geq K_\epsilon}\,\mathrm{d}\mu \leq \frac{\varepsilon}{2} \ .$$

Posons $\delta_\varepsilon = \varepsilon/(2K_\varepsilon)$. Alors, pour A de mesure strictement plus petite que δ_ε,

$$\int_A |f| \, \mathrm{d}\mu = \int_{A\cap\{|f|\geq K_\varepsilon\}} |f| \, \mathrm{d}\mu + \int_{A\cap\{|f|<K_\varepsilon\}} |f| \, \mathrm{d}\mu \leq \int_E |f| \, \mathbb{I}_{|f|\geq K_\epsilon} \, \mathrm{d}\mu + K_\varepsilon \, \mu(A) < \varepsilon \ .$$

Si f est Lebesgue-intégrable, alors, prenant $x < y$ tels que $y - x < \delta_\varepsilon$, on a $|F(y) - F(x)| \leq \int_{]x,y[} |f| \, \mathrm{d}\mu \leq \varepsilon$.

Exercice 5 (TD 2)**.** La convergence p.p. des f_n vers f implique qu'il existe un ensemble mesurable N de mesure nulle tel que pour tout $k \in \mathbb{N}$,

$$E_{k,n} = \bigcap_{j\geq n} \left\{ |f_j - f| < 2^{-k} \right\} \nearrow E \setminus N$$

lorsque $n \to \infty$. Comme $\mu(E) < +\infty$, il existe, pour chaque k, un rang n_k tel que $\mu(E_{k,n_k}) \geq \mu(E) - \varepsilon/2^k$. Notons

$$A = N \cup \bigcup_{k\geq 1} E \setminus E_{k,n_k} \ ;$$

alors $\mu(A) \leq \varepsilon$ et il y a convergence uniforme des f_n vers f sur A^c. En effet, si $x \in A^c$, alors pour tout k, $x \in E_{k,n_k}$, et $|f_j(x) - f(x)| \leq 2^{-k}$ pour tout $j \geq n_k$.

Ce résultat est faux en général quand $\mu(E) = +\infty$, comme le montre le choix $E = \mathbb{R}^+$, muni de la mesure de Lebesgue, et $f_n = \mathbb{I}_{[n,+\infty[}$, suite qui converge simplement vers 0, mais jamais uniformément sur un ensemble de mesure infinie. (On peut aussi prendre $E = \mathbb{N}$, et μ de comptage, avec les mêmes f_n.)

Exercice 6 (TD 2)**.** On considère $g_n = f + f_n - |f - f_n|$, g_n est positive, de limite μ-p.p. égale à $2f$. Le lemme de Fatou[1] montre alors que

$$2\int_E f \, \mathrm{d}\mu = \int_E \left(\liminf_{n\to\infty} g_n \right) \, \mathrm{d}\mu \leq \liminf_{n\to\infty} \int_E g_n \, \mathrm{d}\mu = 2\int_E f \, \mathrm{d}\mu - \limsup_{n\to\infty} \int_E |f_n - f| \, \mathrm{d}\mu \ ,$$

d'où le résultat.

Remarquez que g_n est en fait égale à $2\min(f_n, f)$ et qu'on aurait pu appliquer le théorème de convergence dominée à g_n pour obtenir le résultat. Ce n'est pas étonnant : le théorème de convergence dominée se prouve précisément grâce au lemme de Fatou.

Pour des f_n intégrables, la condition est $f_n \to f$ μ-p.p., ainsi que $\|f_n\|_1 \to \|f\|_1 < +\infty$ quand $n \to \infty$, et on obtient que $f_n \to f$ dans $\mathbb{L}^1(\mu)$.

Exercice 7 (TD 2)**.** *Première proposition* : $f(x) = y$ possède une solution dans $[k/2^n,(k+1)/2^n[$ si et seulement si $y \in f([k/2^n,(k+1)/2^n[)$. Or, par continuité (cf. théorème des valeurs intermédiaires), $f([k/2^n,(k+1)/2^n[)$ est un intervalle et est donc mesurable. Ainsi, soit N_n la fonction mesurable

$$N_n = \sum_{k=0}^{2^n-1} \mathbb{I}_{f([k/2^n,(k+1)/2^n[)} + \mathbb{I}_{\{f(1)\}} \ .$$

La suite des N_n croît vers N quand $n \to \infty$, et N est donc mesurable.

Seconde proposition (bien plus élégante – merci, merci, chers élèves) : fixons $n \in \mathbb{N}$; soit l'application mesurable $g : \mathbf{x} = (x_1, \ldots, x_n) \mapsto \sum_{i<j} \left(f(x_i) - f(x_j)\right)^2$, l'application (continue)

1. n'oubliez jamais : « mon ami Fatou fait tout »

de projection $\pi : \mathbf{x} = (x_1, \ldots, x_n) \mapsto x_1$, et l'ensemble mesurable $\Delta = \{\mathbf{x} = (x_1, \ldots, x_n) : \forall i \neq j,\ \ x_i \neq x_j\}$. Alors,

$$N^{-1}(\llbracket n, +\infty \rrbracket) = f\left(\pi\left(g^{-1}\{0\} \cap \Delta\right)\right) \ ,$$

Euh, et après, il y avait un peu souci (les images continues de boréliens ne sont pas des boréliens en général). Mais on peut s'en sortir autrement. Au final, c'est plus compliqué que la première méthode, cependant. Désolé...

Exercice 1 (TD 3). Pour prouver le sens direct : il existe $N \in \mathcal{A}$ de mesure nulle, et $g : E \to E'$ borélienne tels que $f(x) = g(x)$ pour $x \in N^c$. Soit $G'_1, G'_2, \ldots$ une base dénombrable d'ouverts de E'. Fixons $\varepsilon > 0$. Chaque $g^{-1}(G'_n)$ est borélien, et il existe donc un ouvert U_n tel que $g^{-1}(G'_n) \subseteq U_n$, et $\mu(U_n \setminus g^{-1}(G'_n)) < \varepsilon/2^n$ (on a utilisé ici la finitude de μ). Soit

$$A = \bigcup_{n\geq 1} \left(U_n \setminus g^{-1}(G'_n)\right) \ ;$$

A est de mesure $\mu(A) < \varepsilon$. Soit $B = A^c \setminus N = (A \cup N)^c$, $B \in \mathcal{A}$, et par régularité, il existe donc un compact $K \subseteq B$ vérifiant $\mu(B \setminus K) < \varepsilon - \mu(A)$. Par ailleurs,

$$\mu(K^c) = \mu\left(B^c \cup (B \setminus K)\right) = \mu\left((A \cup N) \cup (B \setminus K)\right) = \mu(A) + \mu(B \setminus K) < \varepsilon.$$

f coïncide avec g sur K ($K \subseteq N^c$), et il ne reste plus qu'à montrer que la restriction g_0 de g à K, ou même à A^c, est continue. Pour tout n, $g_0^{-1}(G'_n) = g^{-1}(G'_n) \cap A^c$. Comme $g^{-1}(G'_n) \subseteq U_n$ et $U_n \setminus g^{-1}(G'_n) \subseteq A$, il vient $g_0^{-1}(G'_n) = U_n \cap A^c$, ouvert de A^c. Ceci prouve la continuité de g_0 et conclut la preuve du sens direct.

Réciproquement, soit une suite $(K_n)_{n\geq 1}$ de compacts tels que $\mu(K_n^c) < 1/n$ et la restriction f_n de f à K_n est continue. On pose

$$N = \bigcap_{n\geq 1} K_n^c \ ,$$

N est ouvert, de μ-mesure nulle. On définit une application g par $g(x) = f(x)$ pour $x \in N^c$, et (pour N non-vide), $g(x) = y_0$ pour $x \in N$, où $y_0 \in E'$. $f = g$ μ-p.p., et il reste à montrer que g est borélienne. Pour ce faire, il suffit de monter que l'image réciproque par g de tout ouvert est borélienne. Soit U un ouvert de E' ; dans le cas où $y_0 \in U$, on a

$$g^{-1}(U) = \left(f^{-1}(U) \cap N^c\right) \cup N = N \cup \left(\bigcup_{n\geq 1} f^{-1}(U) \cap K_n\right) = N \cup \left(\bigcup_{n\geq 1} f_n^{-1}(U)\right) \ .$$

Or, pout tout n, $f_n^{-1}(U)$ est un ouvert de K_n, donc est de la forme $U_n \cap K_n$, où U_n est un ouvert de E. Ainsi, $g^{-1}(U)$ est bien borélien. De même dans le cas où $y_0 \notin U$.

Remarque : $E = [0,1]$, $E' = \mathbb{R}$, $\mu = \lambda$, mesure de Lebesgue. Comme vous l'avez vu dans le cours, il existe au moins un ensemble borélien N de mesure nulle contenant un sous-ensemble M non-borélien. $f = \mathbb{I}_M$ est λ-p.p. égal à la fonction borélienne $g = \mathbb{I}_N$, mais f elle-même n'est pas borélienne. Pourtant, f vérifie la seconde assertion proposée : si U est un ouvert contenant N, de mesure plus petite que $\varepsilon/2$, et si K est un compact de complémentaire de mesure plus petite que $\varepsilon/2$ et tel que la restriction de g à K est continue, alors $K' = K \cap U^c$ est un compact vérifiant que $\lambda(K')$ est plus petite que ε, et la restriction de f à K' est continue.

Exercice 2 (TD 3). On se place d'abord dans le cas $f \geq 0$. f est la limite ponctuelle d'une suite croissante $(s_n)_{n\geq 0}$ de fonctions étagées bien choisies ($s_0 = 0$). Pour $n \geq 1$, on pose

$t_n = s_n - s_{n-1}$; puisque t_n est postive et a été obtenue comme combinaison linéaire d'un nombre fini de fonctions caractéristiques, il existe des constances $c_i > 0$ et des ensembles mesurables E_i (non nécessairement disjoints), tels que

$$f = \sum_{n \geq 1} t_n = \sum_{i=1}^{+\infty} c_i \mathbb{I}_{E_i} \ .$$

Puisque, par convergence monotone,

$$\int_E f \, \mathrm{d}\mu = \sum_{i=1}^{+\infty} c_i \mu(E_i) \ ,$$

et que f est intégrable, la série du deuxième membre converge – et en particulier, chaque E_i est de mesure finie. Ainsi, par régularité de μ, pour tout i, il existe un compact K_i et un ouvert V_i tels que $K_i \subseteq E_i \subseteq V_i$ et

$$\mu\left(V_i \setminus K_i\right) < \frac{2^{-(i+1)}}{c_i} \varepsilon \ .$$

Définissons alors

$$v = \sum_{i=1}^{+\infty} c_i \mathbb{I}_{V_i} \qquad \text{et} \qquad u = \sum_{i=1}^{N} c_i \mathbb{I}_{K_i} \ ,$$

où N a été choisi de sorte que

$$\sum_{i \geq N+1} c_i \mathbb{I}_{E_i} < \frac{\varepsilon}{2} \ .$$

Alors, v est semi-continue inférieurement (comme enveloppe supérieure d'une famille de fonctions semi-continues inférieurement), et u est semi-continue supérieurement, avec $u \leq f \leq v$, et de plus

$$v - u = \sum_{i=1}^{N} c_i \left(\mathbb{I}_{V_i} - \mathbb{I}_{K_i}\right) + \sum_{i \geq N+1} c_i \mathbb{I}_{V_i} \leq \sum_{i=1}^{+\infty} c_i \mathbb{I}_{V_i \setminus K_i} + \sum_{i \geq N+1} c_i \mathbb{I}_{V_i} \ ,$$

de sorte que $v - u$ est bien d'intégrale plus petite que ε, par construction des V_i, K_i, et de N.
NB : En fait, on n'a pas eu besoin de toute la force de la régularité intérieure, on a juste besoin d'un supremum sur les fermés inclus dans l'ensemble (et pas uniquement sur les compacts).

Dans le cas général, on écrit $f = f^+ - f^-$, et l'on définit comme ci-dessus u_1 et v_1 relativement à f^+, u_2 et v_2 relativement à f^-, et l'on pose $u = u_1 - v_2$, $v = v_1 - u_2$.
NB : Pour des rappels sur les fonctions semi-continues, et la stabilité de cette classe par certaines opérations, on pourra utiliment consuler G. Choquet, *Cours de topologie*, chapitre II-8.

Exercice 3 (TD 3)**.** Seule la régularité intérieure pose problème (la régularité extérieure est vraie sous la seule hypothèse d'une mesure borélienne finie sur un métrique). Voir par exemple H. Bauer, *Probability theory and elements of measure theory*, Théorème 7.3.3.

Exercice 4 (TD 3)**.** On rappelle que pour tout $u \in \mathbb{R}$, $1 + u \geq e^u$, de sorte que, utilisant $e^{-x/n} > 0$, on a $(1-x/n)^n \mathbb{I}_{[0,n]}(x) \leq e^{-x}$, et les intégrandes sont toutes dominées par $e^{-x} x^{\alpha-1}$, qui est sommable dès que $\alpha > 0$. Par ailleurs, la suite des intégrandes converge ponctuellement vers $x \mapsto e^{-x}x^{\alpha-1}$, de sorte que par convergence dominée,

$$\int_0^n \left(1 - \frac{x}{n}\right)^n x^{\alpha-1}\,\mathrm{d}x \xrightarrow[n\to+\infty]{} \Gamma(\alpha) = \int_{\mathbb{R}^+} x^{\alpha-1}e^{-x}\,\mathrm{d}x \ .$$

Lorsque $\alpha \leq 0$, toutes les intégrales considérées valent $+\infty$.

La deuxième suite d'intégrales se traite presque de la même façon : par convergence dominée, on obtient convergence vers $1/(1-\alpha)$ lorsque $\alpha < 1$, et pour le cas $\alpha \geq 1$, le lemme de Fatou assure la divergence vers $+\infty$.

En fait, dans les deux cas, l'appel au théorème de convergence dominée peut être remplacé par un appel au théorème de convergence monotone ; en effet, la suite $((1-x/n)^n\mathbb{I}_{[0,n]}(x))_{n\geq 1}$ est croissante, de limite e^{-x}.

Exercice 5 (TD 3). Par convergence monotone,

$$\sum_{n\geq 0}\int_E |f_n|\mathrm{d}\mu < \infty \quad \text{implique} \quad \int_E \left(\sum_{n\geq 0}|f_n|\right)\,\mathrm{d}\mu < \infty \ .$$

En particulier, $\sum_{n\geq 0}|f_n|$ est finie μ-p.p., $\sum_{n\geq 0} f_n$ existe μ-p.p., et est μ-intégrable. $\sum_{n=0}^N f_n$ est dominée par la fonction intégrable $\sum_{n\geq 0}|f_n|$, et converge μ-p.p. vers $\sum_{n\geq 0} f_n$. Par convergence dominée, on a ainsi

$$\int_E \left(\sum_{n=0}^N f_n\right)\,\mathrm{d}\mu \xrightarrow[n\to+\infty]{} \int_E \left(\sum_{n\geq 0} f_n\right)\,\mathrm{d}\mu \ .$$

Par ailleurs, pour N fini,

$$\int_E \left(\sum_{n=0}^N f_n\right)\,\mathrm{d}\mu = \sum_{n=0}^N \int_E f_n\,\mathrm{d}\mu \ ,$$

et le deuxième membre est le N-ième terme d'une série absolument convergente par hypothèse. Un passage à la limite quand N tend vers l'infini conduit donc au résultat.

On écrit

$$\int_0^1 \frac{\ln x}{1-x}\,\mathrm{d}x = \int_0^1 \frac{\ln(1-x)}{x}\,\mathrm{d}x = -\int_0^1 \mathrm{d}x \sum_{k\geq 0}\frac{x^k}{k+1} \ ,$$

et ce, en utilisant le développement en série entière de $\ln(1-x)$. On pose donc $f_k(x) = x^k/(k+1)$, et les hypothèses du point précédent sont vérifiées ($f_k = |f_k|$ est d'intégrale sur $[0,1]$ égale à $1/(k+1)^2$). Ainsi, on peut échanger la somme et l'intégrale pour obtenir

$$\int_0^1 \frac{\ln(x)}{1-x}\,\mathrm{d}x = -\sum_{k\geq 0}\frac{1}{(k+1)^2} = -\frac{\pi^2}{6} \ .$$

Exercice 6 (TD 3). (Dans tout l'exercice, on utilise le fait que λ-p.p., x est différent de tous les a_n, de sorte que l'on peut diviser par $|x - a_n|$ sans se poser de questions.) Pour la première assertion, le lemme de Borel-Cantelli suffit. On pose $A_n = \{x \in \mathbb{R} : |x - a_n| < \sqrt{\alpha_n}\}$, de sorte que la mesure de A_n vaut $\lambda(A_n) = 2\alpha_n$, terme général d'une série convergente. Ainsi, pour presque tout x, $|x-a_n| > \sqrt{\alpha_n}$ à partir d'un certain rang fini, et donc $\alpha_n/|x-a_n| < \sqrt{\alpha_n}$ à partir d'un tel rang. On conclut par sommabilité de la série de terme général $\sqrt{\alpha_n}$.

Pour affiner ce résultat, on considère la fonction $f : \mathbb{R} \to \overline{\mathbb{R}}_+$ définie par

$$f(x) = \sum_{n\geq 1} \frac{\sqrt{\alpha_n}}{\sqrt{|x-a_n|}} \ .$$

Par convergence monotone, puis par intégration,

$$\begin{aligned}\int_{-k}^{k} f(x)\,\mathrm{d}x = \sum_{n\geq 1} \sqrt{\alpha_n} \int_{-k}^{k} \frac{1}{\sqrt{|x-a_n|}}\,\mathrm{d}x \ &= \ \sum_{n\geq 1} \sqrt{\alpha_n} \left[2\sqrt{|x-a_n|}\,\mathrm{sgn}\,(x-a_n)\right]_{-k}^{k} \\ &\leq \ C_k \sum_{n\geq 1} \sqrt{\alpha_n} < +\infty \ ,\end{aligned}$$

où sgn est la fonction signe et C_k est une constante postive ne dépendant que de k ; C_k est par exemple un majorant de la fonction qui à $a \in \mathbb{R}$ associe

$$2\left(\sqrt{|k-a|}\,\mathrm{sgn}\,(k-a) + \sqrt{|k+a|}\,\mathrm{sgn}\,(k+a)\right) \ ,$$

cette fonction étant majorée car continue et de limite nulle aux infinis (on le voit en introduisant les quantités conjuguées). Ainsi, on conclut facilement que f est finie p.p. sur $\mathbb{R}$, i.e. la série de terme général $\sqrt{\alpha_n}/\sqrt{|x-a_n|}$ est sommable p.p. (et a fortiori, elle est de carré sommable).

Exercice 7 (TD 3). Pour tout $t \geq 0$, $g(x,t) = \sqrt{\phi(x)^2 + t} \leq |\phi(x)| + t$, qui est sommable sur $[0,1]$. F est donc bien définie, positive, et croissante. Par convergence dominée, F est continue sur $\mathbb{R}^+$. La fonction g à intégrer est de plus dérivable en (x,t) tels que $t > 0$ ou $|\phi(x)| > 0$, de dérivée partielle par rapport à t,

$$g'(x,t) = \frac{1}{2\sqrt{\phi(x)^2 + t}} \ .$$

Si $t \geq t_0 > 0$, cette dérivée est majorée par $1/(2\sqrt{t_0})$, et est donc sommable sur $[0,1]$. Le théorème de dérivation sous le signe intégrale s'applique donc pour calculer la dérivée de F. Reste l'étude (directe) en 0.

$$\frac{F(t)-F(0)}{t} = \int_0^1 \frac{\sqrt{\phi(x)^2+t} - |\phi(x)|}{t}\,\mathrm{d}x = \int_0^1 \frac{1}{\sqrt{\phi(x)^2+t} + |\phi(x)|}\,\mathrm{d}x \ ,$$

qui, par convergence monotone, croît vers $1/(2\,|\phi(x)|)$ quand t décroît vers 0. Ainsi, F est dérivable en 0 si et seulement si $1/|\phi|$ est sommable.

De même, la fonction G est à son tour bien définie, positive et continue. On va montrer que G est dérivable en t si et seulement si $\lambda\{\phi = t\} = 0$. En effet,

$$\frac{G(t+h)-G(t)}{h} = \int_0^1 \frac{|\phi(x)-t-h| - |\phi(x)-t|}{h}\,\mathrm{d}x \ .$$

Si $h > 0$, l'accroissement sous l'intégrale est dominé par 1, et le théorème de convergence dominée permet de passer à la limite, pour obtenir la dérivée à droite de G en t :

$$G'_d(t) = \int_0^1 \left(\mathbb{I}_{\{\phi(x)\leq t\}} - \mathbb{I}_{\{\phi(x)>t\}}\right)\,\mathrm{d}x = \lambda\{\phi \leq t\} - \lambda\{\phi > t\} \ ;$$

de même, G est dérivable à gauche en t, de dérivée $G'_g(t) = \lambda\{\phi < t\} - \lambda\{\phi \geq t\}$, ce qui conclut à l'équivalence proposée.

Exercice 8 (TD 3). La fonction caractéristique de la densité gaussienne est donnée par

$$\hat{f}(t) = \int_{\mathbb{R}} e^{itx} f(x)\,\mathrm{d}x \ .$$

Comme la dérivée $x \mapsto ixe^{itx}e^{-x^2/2}$ de l'intégrande est dominée par la fonction sommable $x \mapsto xf(x)$, on peut dériver sous le signe intégrale, puis intégrer par parties, pour obtenir

$$\hat{f}'(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} ixe^{itx}e^{-x^2/2}\,\mathrm{d}x = -\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} te^{-x^2/2}e^{itx}\,\mathrm{d}x = -t\hat{f}(t) \ ,$$

d'où, en résolvant l'équation différentielle, $\hat{f}(t) = Ae^{-t^2/2}$, où $A = \hat{f}(0) = 1$.

Exercice 9 (TD 3). La Lebesgue-mesurabilité de f signifie que toute image réciproque par f d'un borélien est Lebesgue-mesurable.

Supposons, dans un premier temps, que $|g|$ est bornée par une constante M. Ainsi, pour tout $n \in \mathbb{N}$ et $k \in \mathbb{Z}$, il existe $A_{n,k}$ et $N_{n,k}$ deux boréliens disjoints, $N_{n,k}$ étant de mesure nulle, tels que

$$A_{n,k} \subseteq G_{n,k} = g^{-1}\left([k/2^n, (k+1)/2^n[\right) \subseteq A_{n,k} \cup N_{n,k} \ .$$

Les $(A_{n,k})_{k\in\mathbb{Z}}$ partitionnent $\mathbb{R}$, et par conséquent, $\mathbb{R}$ est égal à l'union (dénombrable) disjointe des $(A_{n,k})_{k\in\mathbb{Z}}$ et d'un ensemble borélien N_n, de mesure nulle car inclus dans l'union des $N_{n,k}$. On pose alors

$$f_n = \sum_{k\in\mathbb{Z}} k/2^n\, \mathbb{I}_{B_{n,k}} - M\mathbb{I}_{N_n} \qquad \text{et} \qquad h_n = \sum_{k\in\mathbb{Z}} (k+1)/2^n\, \mathbb{I}_{B_{n,k}} + M\mathbb{I}_{N_n} \ .$$

f_n et h_n sont boréliennes, et vérifient $f_n \leq g \leq h_n$ sur tout $\mathbb{R}$, et $h_n = f_n + 1/2^n$ sur $E \setminus N_n$. En particulier, notant $f = \limsup_n f_n$, $h = \liminf_n h_n$, on a que f et h sont boréliennes, que $f \leq g \leq h$ sur tout $\mathbb{R}$, et que $f = h$ λ-p.p.

Dans le cas de f quelconque (non-bornée), on considère la suite $(g_N)_{N\geq 1}$ de fonctions mesurables définies par $g_N = \min(N, \max(-N, g))$. D'après le point précédent, $|f_N|$ étant bornée par N, il existe f_N et h_N boréliennes égales p.p. telles que $f_N \leq g_N \leq h_N$. A nouveau, il suffit de prendre $f = \limsup_N f_N$, $h = \liminf_N h_N$.

Exercice 10 (TD 3). Pour une démonstration (rapide sur certains points) de l'équivalence, voir J.C. Oxtoby, *Measure and Category*, Théorème 7.5.

Exemple : la fonction $\mathbb{I}_{\mathbb{Q}}$ n'est pas Riemann-intégrable (tout point de [0,1] est point de discontinuité), mais elle est évidemment Lebesgue-intégrable, d'intégrale nulle.

Exercice 11 (TD 3). Les deux premières questions sont très proches de votre cours. Par construction de A, $E = \cup_{n\in\mathbb{Z}} T^n(A)$. Supposons A mesurable. Comme T et T^{-1} sont mesurables, chacun des $T^n(A)$, $n \in \mathbb{Z}$, est mesurable. De plus, leur réunion est disjointe, car s'il existe $n, m \in \mathbb{Z}$ et $x \in E$ tels que $x \in T^n(A) \cap T^m(A)$, alors il existe $y, z \in A$ tels que $T^n(y) = T^m(z) = x$, y et z sont dans la même orbite, et par construction de A, $y = z$, et puisque T n'a pas de point périodique, $n = m$. Mais alors,

$$\mu(E) = \sum_{n\in\mathbb{Z}} \mu(T^n(A)) = \sum_{n\in\mathbb{Z}} \mu(A) \ ,$$

car T conserve la mesure. A mesurable de mesure $\mu(A) = 0$ impliquerait alors $\mu(E) = 0$, ce qui contredit l'énoncé, mais A mesurable de mesure $\mu(A) > 0$ n'est pas plus satisfaisant, puisque cela impliquerait $\mu(E) = +\infty$.

Pour $E = [0,1]$, muni de la mesure de Lebesgue. On note $\lfloor x \rfloor$ la partie réelle de $x \in \mathbb{R}$. On fixe un élément $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, et on pose $T : x \mapsto x + \alpha - \lfloor x + \alpha \rfloor$. Cette application T convient pour l'application du point précédent.

Si μ est une mesure sur $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$ invariante par translation, il est alors facile de voir que la μ-mesure de tout intervalle dyadique $[k/2^n, (k+1)/2^n]$ est $\mu([0,1])/2^n$. Ceci détermine la mesure de tout intervalle de $\mathbb{R}$, et on a alors $\mathrm{d}\mu = \mu([0,1])\mathrm{d}\lambda$ sur les boréliens puis sur les Lebesgue-mesurables. Mais on sait alors construire des éléments de $\mathcal{P}(\mathbb{R})$ non Lebesgue-mesurables, par le point précédent.

Exercice 12 (TD 3). Je suis prêt à lire tout début de solution proprement et clairement rédigé, et à donner des indications substantielles à ceux qui auront fait l'effort de chercher un peu.

Intégration et Probabilités – TD 4
Corrigé partiel

Exercice 1. Soit $P = \{|f| > 2\}$. Par inégalité triangulaire, $P \subseteq \{|\mathbb{I}_{A_n} - f| > 1\}$, et ainsi,

$$\mu(P) \leq \int_P |\mathbb{I}_{A_n} - f| \, \mathrm{d}\mu \leq \int_E |\mathbb{I}_{A_n} - f| \, \mathrm{d}\mu \xrightarrow[n\to\infty]{} 0 \ ,$$

i.e., $\mu(P) = 0$ et $|f| \leq 2$ μ-p.p.

Que f soit de la forme $\mathbb{I}_A$, avec $A \in \mathcal{A}$ équivaut à ce que f soit mesurable, avec $f = f^2$ μ-p.p. (alors $A = \{f = 0\}$). Montrons cette dernière propriété. On calcule

$$\int_E \left| f - f^2 \right| \, \mathrm{d}\mu \leq \int_E |f - \mathbb{I}_{A_n}| \, \mathrm{d}\mu + \int_E \left| f^2 - \mathbb{I}_{A_n} \right| \, \mathrm{d}\mu \xrightarrow[n\to\infty]{} 0 \ ,$$

car

$$\int_E \left| f^2 - \mathbb{I}_{A_n} \right| \, \mathrm{d}\mu = \int_E |f - \mathbb{I}_{A_n}| \, (1_{A_n} + f) \, \mathrm{d}\mu \leq 3 \int_E |f - \mathbb{I}_{A_n}| \, \mathrm{d}\mu \ .$$

Enfin, l'hypothèse $\sum_{n\geq 1} \mu(A_n \Delta A) < \infty$ assure, par le lemme de Borel–Cantelli, que, pour μ-presque tout x, $x \in (A\Delta A_n)^c$ à partir d'un certain rang $n_0(x)$. Donc, $\mathbb{I}_{A_n}(x) = \mathbb{I}_A(x)$ pour tout $n \geq n_0(x)$, et on a la convergence demandée.

Exercice 2. La construction de ϕ découle de l'application du critère de Cauchy pour les suites convergentes. Puis, chaque g_n est par construction d'intégrale plus petite que 1, et la suite des g_n croît vers la fonction

$$g = \sum_{k\geq 1} \left| f_{\phi(k+1)} - f_{\phi(k)} \right| \ ,$$

qui est donc μ-intégrable, par le théorème de convergence monotone, et, partant, finie μ-p.p. Or, $\left| f_{\phi(n)} \right| \leq \left| f_{\phi(1)} \right| + g_{n-1} \leq h = \left| f_{\phi(1)} \right| + g$, et on a dès lors la domination que l'on cherchait. Par ailleurs, on définit légitimement la fonction mesurable

$$F = \sum_{k\geq 1} f_{\phi(k+1)} - f_{\phi(k)} \ ;$$

et on a que $f_{\phi(n)} \to F + f_{\phi(1)}$ μ-p.p. Par le théorème de convergence dominée usuel, $(f_{\phi(n)})_{n\geq 1}$ converge dans $\mathbb{L}^1(\mu)$ vers $F + f_{\phi(1)}$, et en utilisant le fait que $\| \cdot \|_1$ est une norme, il vient $f = F + f_{\phi(1)}$ μ-p.p., ce qui conclut.

Exercice 3 (td 4) **– voir aussi Exercice 8** (td 2)**.** La convergence μ-p.s. se traduit par le fait que l'ensemble mesurable

$$N = \bigcup_{\eta\in\mathbb{N}} \bigcap_{n\in\mathbb{N}} \bigcup_{p\geq n} \{|f_p - f| > 1/\eta\} = \bigcup_{\eta\in\mathbb{N}} \bigcap_{n\in\mathbb{N}} \left\{ \sup_{p\geq n} |f_p - f| > 1/\eta \right\}$$

est de mesure nulle. Or N est de mesure nulle équivaut à ce que chacun des ensembles

$$\bigcap_{n\in\mathbb{N}} \left\{ \sup_{p\geq n} |f_p - f| > 1/\eta \right\}$$

soit de mesure nulle. μ étant finie, on a par limite décroissante que

$$\mu\left(\bigcap_{n\in\mathbb{N}}\left\{\sup_{p\geq n}|f_p-f|>1/\eta\right\}\right)=\lim_{n\to\infty}\mu\left(\left\{\sup_{p\geq n}|f_p-f|>1/\eta\right\}\right)\ ,$$

ce qui conclut à l'équivalence demandée.

Pour la seconde implication, dans le cas où f_n converge vers f au sens $\mathbb{L}^1(\mu)$, il suffit d'écrire

$$\mu(|f-f_n|>\varepsilon)\leq\frac{1}{\varepsilon}\|f_n-f\|_1\to 0\ .$$

Des contre-exemples sont respectivement donnés (pour la mesure de Lebesgue sur [0,1]) par $f_n=n\,\mathbb{I}_{[0,1/n]}$ (qui converge en μ-probabilité vers 0, mais qui ne converge pas vers 0 dans $\mathbb{L}^1(\mu)$, car chaque f_n est d'intégrale 1), et par la suite (à double indice) $f_{n,m}$ suivante, qui converge vers 0 en μ-probabilité mais pas μ-p.s. Pour les entiers $n\geq 1$ et $m\in\{1,\ldots,n\}$, $f_{n,m}=\mathbb{I}_{[(m-1)/n,m/n[}$.

Si $f_n\to f$ en μ-probabilité, soit (n_k) une suite strictement croissante (donc convergeant vers $+\infty$) telle que $\mu\{|f_{n_k}-f|>2^{-k}\}\leq 2^{-k}$. Le lemme de Borel-Cantelli montre que $\limsup_{k\to\infty}2^k\,|f_{n_k}-f|\leq 1$ μ-p.s., ce qui conclut. Un sens de l'équivalence est désormais immédiat, il suffit de montrer que la seconde propriété de l'équivalence implique la première. A cet effet, on remarque que si f_n ne convergeait pas en μ-probabilité vers f, alors on pourrait trouver $\eta>0$, $\varepsilon_0>0$ et une suite strictement croissante $\phi(n)$ tels que $\mu\{\,\big|\,f_{\phi(n)}-f\,\big|\,>\varepsilon\}\geq\eta$. Or par hypothèse, on sait réextraire ψ telle que $f_{\phi(\psi(n))}$ converge vers f μ-p.s., ce qui, par convergence dominée (voir raisonnement ci-dessus), contredit la construction de ϕ.

Pour l'application, on remarque (de même!) que si le résultat n'était pas vrai, alors on aurait l'existence de $\eta>0$ et d'une sous-suite ϕ telle que $\|f_{\phi(n)}-f\|_1\geq\eta$. Extrayant ψ telle que $f_{\phi(\psi(n))}$ converge μ-p.s. vers f, on aurait par le théorème de convergence dominée classique que $\|f_{\phi(\psi(n))}-f\|_1\to 0$, ce qui contredit la construction de ϕ.

Le fait que d soit une distance sur $\mathbb{L}^0(\mu)$ est immédiat. Pour δ, on vérifie d'abord l'inégalité triangulaire. On remarque que si $\mu\{|f-g|>\varepsilon'\}\leq\varepsilon'$, alors ceci est encore vrai pour tout $\varepsilon\geq\varepsilon'$. Ainsi, soit $\delta(f,g)<a$, $\delta(f,h)<b$; il s'agit de montrer que $\delta(f,h)<a+b$. Or les hypothèses se traduisent par le fait qu'avec probabilité au moins $1-a$, $|f-g|\leq a$, et avec probabilité au moins $1-b$, $|g-h|\leq b$, de sorte qu'avec probabilité au moins $1-(a+b)$, $|f-h|\leq a+b$.

Pour montrer que $\delta(f,g)=0$ implique $f=g$ μ-p.s., ainsi que l'équivalence des distances, nous allons prouver que pour tous f et g,

$$\delta^2(f,g)\leq d(f,g)\leq 2\delta(f,g)\ .$$

Remarquons d'abord que pour tous f et g, $\delta(f,g),d(f,g)\leq 1$. Soit maintenant $\delta(f,g)\leq\varepsilon\leq 1$, alors

$$d(f,g)\leq\varepsilon\mu\left\{|f-g|\leq\varepsilon\right\}+1\,\mu\left\{\min\left(1,\,|f-g|\right)>\varepsilon\right\}\leq\varepsilon+\mu\left\{|f-g|>\varepsilon\right\}\leq 2\varepsilon\ .$$

D'autre part,

$$\begin{aligned}\mu\left\{|f-g|>\sqrt{d(f,g)}\right\} &= \mu\left\{\min\left(1,\,|f-g|\right)>\sqrt{d(f,g)}\right\}\\ &\leq \frac{1}{\sqrt{d(f,g)}}\int_E\min\left(1,\,|f-g|\right)\,\mathrm{d}\mu=\sqrt{d(f,g)}\ .\end{aligned}$$

Pour voir que ces distances métrisabilisent la convergence en probabilité, on montre que les suites convergentes sont les mêmes. Ceci s'obtient en remarquant que, comme $\mathbb{I}_{\{|f-g|>\varepsilon\}} \leq \min(1, |f-g|/\varepsilon) \leq \max(1, 1/\varepsilon) \min(1, |f-g|)$ pour $\varepsilon > 0$, alors, par intégration,

$$\mu\{|f-g| > \varepsilon\} \leq \max(1, 1/\varepsilon)\, d(f,g) \ ;$$

par ailleurs, il est clair que $d(f,g) \leq \min(1, \varepsilon) + \mu\{|f-g| > \varepsilon\}$.

La complétude se prouve de manière similaire à la solution de l'exercice 2. Si $(f_n)_{n\geq 1}$ est de Cauchy pour d, alors on extrait $(f_{\phi(n)})_{n\geq 1}$ telle que

$$\sum_{n\geq 1} d(f_{\phi(n)}, f_{\phi(n+1)}) < +\infty \ ,$$

et cela implique, μ-p.s.,

$$\sum_{n\geq 1} \min\left(1, \ \left| f_{\phi(n)} - f_{\phi(n+1)} \right|\right) < +\infty \ ,$$

d'où

$$\sum_{n\geq 1} \left| f_{\phi(n)} - f_{\phi(n+1)} \right| < +\infty \ ,$$

et $h = \sum_{n\geq 1} f_{\phi(n+1)} - f_{\phi(n)}$, $h \in \mathbb{L}^0(\mu)$, est telle que $f_{\phi(n)}$ converge vers $h + f_{\phi(1)}$ μ-p.s., donc en μ-probabilité.

La non-métrisabilité de la convergence μ-p.s. peut se voir avec l'exemple des $f_{n,m}$ ci-dessus : imaginons un instant la métrisabilité par d'. La suite $(f_{n,m})$ ne converge pas μ-ps vers 0, donc il existe une suite extraite $(f_{n',m'})$ telle que $d'(f_{n',m'},0) \geq \eta > 0$; mais de toute suite $(f_{n',m'})$ extraite de $(f_{n,m})$, il est facile de réextraire une sous-suite $(f_{n",m"})$ convergeant p.s., i.e. telle que $d'(f_{n",m"},0) \to 0$ – et on obtient une contradiction.

Exercice 4. L'uniforme intégrabilité de $\mathcal{E}_{\mathcal{H}}$ procède du fait que pour tout $x \geq 0$, $y \in \mathbb{R}^+ \mapsto y\mathbb{I}_{\{y\geq x\}}$ est croissante. Une famille $\{g_1, \ldots, g_N\}$ est bornée par $f = |g_1| + \ldots + |g_N|$ intégrable. Il suffit donc de montrer que toute fonction intégrable f forme un ensemble equi-intégrable, i.e. qu'elle vérifie

$$\lim_{x\to+\infty} \int_{\{|f|\geq x\}} |f| \,\mathrm{d}\mu = 0 \ ;$$

or, ceci est vrai par convergence dominée.

Sens direct de l'équivalence : Pour $c > 0$, $A \in \mathcal{A}$ et $h \in \mathbb{L}^1(\mu)$, on a

$$\begin{aligned} \mu\left[|h|\,\mathbb{I}_A\right] &= \mu\left[|h|\,\mathbb{I}_{A\cap\{|h|<c\}}\right] + \mu\left[|h|\,\mathbb{I}_{A\cap\{|h|\geq c\}}\right] \\ &\leq c\mu(A) + \mu\left[|h|\,\mathbb{I}_{\{|h|\geq c\}}\right] \ . \end{aligned}$$

Ainsi, pour $\mathcal{H} \subseteq \mathbb{L}^1(\mu)$,

$$\sup_{h\in\mathcal{H}} \mu\left[|h|\,\mathbb{I}_A\right] \leq c\mu(A) + \sup_{h\in\mathcal{H}} \mu\left[|h|\,\mathbb{I}_{\{|h|\geq c\}}\right] \ .$$

Lorsque $\mathcal{H}$ est uniformément intégrable, on peut associer à tout $\varepsilon > 0$ un réel $c_\varepsilon > 0$ tel que

$$\sup_{h\in\mathcal{H}} \mu\left[|h|\,\mathbb{I}_{\{|h|\geq c_\varepsilon\}}\right] < \varepsilon/2 \ .$$

Alors, $\mathcal{H}$ est bornée par n'importe lequel des $c_\varepsilon + \varepsilon/2$, et $\delta = \varepsilon/(2\,c_\varepsilon)$ convient pour vérifier la seconde assertion.

Réciproquement, on fixe $\varepsilon > 0$ et on prend $\delta > 0$ vérifiant la seconde assertion. Alors, notant

$$x = \frac{1}{\delta} \sup_{h \in \mathcal{H}} \mu\left[|h|\right] \ ,$$

on a pour tout $h \in \mathcal{H}$,

$$\mu\left[|h| \geq x\right] \leq \frac{1}{x} \mu\left[|h|\right] \leq \delta \ ,$$

et par conséquent, $\sup_{h \in \mathcal{H}} \mu\left[|h| \, \mathbb{I}_{\{|h| \geq x\}}\right] < \varepsilon$.

La bornitude seule ne suffit pas : soit $f_n = n\mathbb{I}_{[0,1/n]}$, et μ la mesure de Lebesgue sur [0,1] ; l'ensemble formé par les f_n est borné (en norme, par 1) dans $\mathbb{L}^1$, mais n'est pas équi-intégrable.

Le point (3) découle directement de la caractérisation du point (2).

Pour l'équi-intégrabilité des parties compactes : ces parties sont bornées, reste à vérifier la deuxième partie de la caractérisation introduite à la question (2). Soit $\mathcal{C}$ une telle partie et $\varepsilon > 0$. Il existe un nombre fini d'éléments $f_1, \ldots, f_N \in \mathcal{C}$ tels que $\mathcal{C}$ soit recouverte dans $\mathbb{L}^1(\mu)$ par les boules de centre un f_j et de rayon ε. On en déduit que pour tout $A \in \mathcal{A}$,

$$\sup_{g \in \mathcal{C}} \mu\left[|g| \, \mathbb{I}_A\right] \leq \varepsilon + \sup_{j=1,\ldots,N} \mu\left[|f_j| \, \mathbb{I}_A\right] < 2\varepsilon \ ,$$

pour A de mesure assez petite (on a utilisé ici l'équi-intégrabilité des familles finies).

Sens réciproque : soit G vérifiant les conditions, on note $M = \sup_{h \in \mathcal{H}} \mu\left[G(|h|)\right] < +\infty$. Soit $\varepsilon > 0$, $a = M/\varepsilon$ et c assez grand pour que $G(x)/x \geq a$ si $x > c$. Alors, pour $x \in]c, +\infty[$, $x \leq G(x)/a$, et ainsi, pour tout $h \in \mathcal{H}$,

$$\mu\left[|h| \, \mathbb{I}_{\{|h| \geq x\}}\right] \leq \frac{1}{a} \mu\left[G(|h|)\right] \leq \varepsilon \ .$$

Sens direct : on suppose $\mathcal{H}$ uniformément intégrable. Soit (c_n) une suite strictement croissante de réels, tendant vers $+\infty$, et telle que

$$\sup_{h \in \mathcal{H}} \mu\left[|h| \, \mathbb{I}_{\{|h| \geq c_n\}}\right] \leq 2^{-n} \ ;$$

choisissons $G(x) = \sum_{n \geq 1} (x - c_n)^+$; pour tout $h \in \mathcal{H}$,

$$1 \geq \sum_{n \geq 1} \mu\left[|h| \, \mathbb{I}_{\{|h| \geq c_n\}}\right] \geq \mu\left[G(|h|)\right] \ .$$

G est convexe, croissante, finie ($x < c_n \to G(x) < nx$) et vérifie bien $G'(x)/x \to +\infty$.

On applique ensuite ce critère à $G(x) = x^{p/q}$ (convexe car $p/q > 1$).

On a vu au point (1) que la domination par une fonction entraînait l'uniforme intégrabilité : ce qu'il faut prouver est bien une amélioration du théorème de convergence dominée. Il suffit que la suite (f_n) soit uniformément intégrable et converge en μ-probabilité vers f pour avoir que (f_n) converge dans $\mathbb{L}^1(\mu)$ vers f, sans avoir nécessairement de chapeau intégrable[1].

Pour la coïncidence des topologies sur $\mathcal{H}$: notons $\mathcal{B}^0$ (resp. $\mathcal{B}^1$) les boules dans $\mathcal{H}$ relatives à la topologie de $\mathbb{L}^0$ (resp. $\mathbb{L}^1$). Comme $\mu\left[\min(1, |h|)\right] \leq \mu\left[|h|\right]$, on a pour tout $h \in \mathcal{H}$ et

1. exemple : une suite (X_n) de variables aléatoires i.i.d. de loi gaussienne sur $\mathbb{R}$: $\sup_n |X_n|$ est p.s. égal à $+\infty$, mais la suite est uniformément intégrable

$r > 0$, $\mathcal{B}^1(h,r) \subseteq \mathcal{B}^0(h,r)$. Par ailleurs, $\mathcal{H} - \mathcal{H}$ est équi-intégrable. Fixons $r > 0$; il existe $a > 1$ tel que $\sup_{k \in \mathcal{H}-\mathcal{H}} \mu\left[|k|\,\mathbb{I}_{\{|k| \geq a\}}\right] < r/2$. Par suite, pour tous $h,g \in \mathcal{H}$,

$$\begin{aligned} \mu\left[|h-g|\right] &\leq \mu\left[|h-g|\,\mathbb{I}_{\{|h-g| \leq a\}}\right] + \sup_{k \in \mathcal{H}-\mathcal{H}} \mu\left[|k|\,\mathbb{I}_{\{|k| \geq a\}}\right] \\ &< a\mu\left[\min(1,\,|h-g|)\right] + \frac{r}{2}\ , \end{aligned}$$

et $\mathcal{B}^0(h, r/(2\,a)) \subseteq \mathcal{B}^1(h,r)$.

Indications pour la fin de l'exercice. Pour le lemme de Scheffé, il suffit d'appliquer la version améliorée du théorème de convergence dominée à $\min(f_n,f)$ qui converge en probabilité vers f et est dominée par f. Pour (8), on note que (i), le lemme de Fatou, et une inégalité de convexité montrent que la famille des $(|f_n - f|^p)_{n \in \bar{\mathbb{N}}}$ est uniformément intégrable ($f_\infty = f$), d'où l'on déduit (ii), via l'amélioration du théorème de convergence dominée. (ii) implique (iii) est trivial, et (iii) donne, via le lemme de Scheffé, que $(|f_n|^p)$ converge dans $\mathbb{L}^1$ vars $|f|^p$, d'où (i), via le point (4).

Intégration et Probabilités – TD 5
Corrigé partiel

Exercice 1. On applique la réciproque du théorème de convergence dominée pour avoir l'existence d'une sous-suite extraite $\left| f_{\varphi(n)} - f \right|^p$ convergeant μ-p.p. vers 0, et ainsi par unicité de la limite, $f = g$ μ-p.p.

Exercice 2. On applique l'inégalité de Cauchy-Schwarz aux fonctions $\sqrt{f}$ et $\sqrt{g}$, pour obtenir

$$\mu(E) \leq \int_E \sqrt{fg}\,\mathrm{d}\mu \leq \sqrt{\int_E f\,\mathrm{d}\mu}\sqrt{\int_E g\,\mathrm{d}\mu}\ ,$$

ce qui donne le résultat. Donc lorsque $f > 0$ est intégrable, de même que $1/f$, μ est nécessairement une mesure finie.

Exercice 3. L'indication est une conséquence immédiate de l'inégalité des accroissements finis, en tenant compte du fait que x et y sont positifs. Ainsi, en utilisant d'abord l'indication, puis l'inégalité de Hölder,

$$\begin{aligned}\int_E |f_n^r - f^r|^{p/r}\,\mathrm{d}\mu &\leq r^{p/r}\int_E |f_n - f|^{p/r}(f_n+f)^{(r-1)p/r}\,\mathrm{d}\mu \\ &\leq r^{p/r}\left(\int_E |f_n - f|^p\,\mathrm{d}\mu\right)^{1/r}\left(\int_E (f_n+f)^p\,\mathrm{d}\mu\right)^{(r-1)/r}\end{aligned} .$$

On conclut, en remarquant que le second terme de la dernière inégalité est borné, parce que convergent, et que le premier tend vers 0.

Exercice 4. Soit a et b dans I, montrons que pour tout $t \in]0,1[$, $ta + (1-t)b$ est aussi dans I. On applique pour cela l'inégalité de Hölder :

$$\int_E |f|^{ta+(1-t)b}\,\mathrm{d}\mu \leq \left(\int_E |f|^a\,\mathrm{d}\mu\right)^t\left(\int_E |f|^b\,\mathrm{d}\mu\right)^{1-t} < +\infty\ . \tag{1}$$

I n'est pas nécessairement fermé : $I =]1,+\infty[$ pour $f(x) = 1/x$ si $x \in [0,1]$ et $f(x) = 0$ sinon.

Pour $f(x) = x^{-1}(1 + |\ln x|)^{-2}$, les résultats sur les séries dites de Bertrand montrent que f^p est sommable en $+\infty$ à condition que $p \geq 1$. De même, en 0^+, le changement de variable $u = 1/x$ donne que f^p est sommable en 0^+ si et seulement si

$$u \mapsto \frac{1}{u^2}\left(\frac{u}{(1+|\ln u|)^2}\right)^p$$

est sommable en $u \sim +\infty$, ce qui est le cas si et seulement si $p \leq 1$. Donc $I = \{1\}$ dans cet exemple.

L'application du logarithme ln aux deux membres de l'inégalité (1) prouve la convexité de $\ln\circ\varphi$. exp étant convexe croissante, φ est elle-même convexe, et on en déduit qu'elle est continue sur l'intérieur de I. Reste à vérifier la continuité aux bornes, ce qui s'obtient, par exemple pour le minimum de I (dans le cas où I est fermé à gauche) en découpant l'intégrale définissant φ selon que $f < 1$ ou $f \geq 1$, et en appliquant au premier morceau le théorème de convergence monotone (lorsque l'exposant p décroît vers $\min I$), et le théorème de convergence dominée au second morceau (domination par $|f|^{\varepsilon+\min I}$ intégrable pour $\varepsilon > 0$ assez petit).

Enfin, écrivant $r = \lambda p + (1-\lambda)q$, $\lambda \geq 0$, et utilisant la convexité de $\ln\circ\varphi$, il vient $\ln\varphi(r) \leq \lambda\ln\varphi(p) + (1-\lambda)\ln\varphi(q)$, d'où $(1/r)\ln\varphi(r) \leq \max((1/p)\ln\varphi(p), (1/q)\ln\varphi(q))$.

Exercice 5. Si la suite (χ_n) convergeait dans $\mathbb{L}^1(\mathbb{R})$, elle admettrait une sous-suite convergeant p.p. vers la limite dans $\mathbb{L}^1(\mathbb{R})$. Puisque par ailleurs (χ_n) converge p.p. vers 0, le seul candidat pour la limite dans $\mathbb{L}^1(\mathbb{R})$ serait 0. Or, tous les χ_n sont d'intégrale 1 – c'est donc qu'ils ne convergent pas dans $\mathbb{L}^1(\mathbb{R})$. En revanche, il y a convergence vers 0 dans $\mathbb{L}^2(\mathbb{R})$.

$\mathbb{L}^1(\mathbb{R}) \cap \mathbb{L}^2(\mathbb{R})$ est dense dans $\mathbb{L}^2(\mathbb{R})$, car il contient les fonctions continues à support compact, elles-mêmes denses dans $\mathbb{L}^2(\mathbb{R})$. Il suffit donc de montrer que F est dense dans $\mathbb{L}^1(\mathbb{R}) \cap \mathbb{L}^2(\mathbb{R})$. A cet effet, on remarque que pour $f \in \mathbb{L}^1(\mathbb{R}) \cap \mathbb{L}^2(\mathbb{R})$, $g_n = f - (\int_{\mathbb{R}} f\,\mathrm{d}\lambda)\chi_n$ est un élement de F, et $f - g_n = (\int_{\mathbb{R}} f\,\mathrm{d}\lambda)\chi_n$ converge vers 0 dans $\mathbb{L}^2(\mathbb{R})$.

Ce résultat ne subsiste pas pour F' sur $\mathbb{L}^1([0,1]) \cap \mathbb{L}^2([0,1])$. Soit en effet $f = \mathbb{I}_{[0,1]}$. S'il existait une suite (f_n) de F' convergeant vers f dans $\mathbb{L}^2([0,1])$, alors l'inégalité de Cauchy-Schwarz (et la finitude de la mesure dans ce cas) montre que cette convergence aurait également lieu dans $\mathbb{L}^1([0,1])$. Or chaque f_n est d'intégrale 0, alors que la limite est d'intégrale 1, et donc il ne peut y avoir convergence dans $\mathbb{L}^1([0,1])$.

E est un espace vectoriel normé, vérifons qu'il est complet. Si (f_n) est de Cauchy pour $\|\cdot\|_E$, alors elle l'est aussi pour $\|\cdot\|_1$ et $\|\cdot\|_2$, et donc il existe $f \in \mathbb{L}^1$ et $g \in L^2$ telles que $f_n \to f$ dans $\mathbb{L}^1$ et $f_n \to g$ dans $\mathbb{L}^2$. En appliquant la réciproque du théorème de convergence dominée, il vient que $f = g$ μ-p.p., et $f_n \to f$ pour $\|\cdot\|_E$.

L'application qu'on considère est bien définie (par deux applications de l'inégalité de Hölder), et linéaire. Il suffit de montrer sa continuité ; par deux applications de l'inégalité de Hölder,

$$\begin{aligned}\left|\int_{\mathbb{R}} fu\,\mathrm{d}\lambda\right| \leq \left|\int_{\mathbb{R}} f_1u\,\mathrm{d}\lambda\right| + \left|\int_{\mathbb{R}} f_2u\,\mathrm{d}\lambda\right| \quad &\leq \quad \|f_1\|_\infty\|u\|_1 + \|f_2\|_2\|u\|_2 \\ &\leq \quad (\|f_1\|_\infty + \|f_2\|_2)\,\|u\|_E\ .\end{aligned}$$

Exercice 6. Le fait que τ_h soit une isométrie (i.e., qu'elle soit linéaire et préserve la norme) découle de l'invariance de la mesure de Lebesgue par translation. Pour $f \in \mathcal{C}_c(\mathbb{R})$ continue à support compact, on a la propriété requise par uniforme continuité. Si f est un élément quelconque dans L^p, soit $\varepsilon > 0$ et g une approximation de f continue à support compact, à $\varepsilon/3$ près en norme L^p. On a

$$||\tau_h f - f||_p \leq ||\tau_h f - \tau_h g||_p + ||\tau_h g - g||_p + ||g - f||_p,$$

les deux termes extrêmes du membres de droite étant chacun plus petit que $\varepsilon/3$ par choix de g, et le terme du mileu est plus petit que $\varepsilon/3$ pour $h < h_0$ (où $h_0 > 0$ est suffisamment petit) par le point précédent. Le résultat ne subsiste pas pour $p = \infty$, prendre $f = \mathbb{I}_{[0,1]}$ par exemple.

Pour l'application, on peut supposer que ce dernier est borné, de mesure strictement positive, de sorte que $f = \mathbb{I}_A$ est dans L^p (considérer les $A_n = A \cap [-n,n]$). La propriété précédente s'écrit $\mathbf{1}_{A\Delta(A+h)} \to 0$ en norme $||\cdot||_p$, soit $\lambda(A\Delta(A+h)) \to 0$ quand $h \to 0$, et en particulier, $\lambda(A \setminus (A+h)) \to 0$. Comme $\lambda(A) > 0$, cela signifie que $\lambda(A \cap (A+h)) > 0$ (et $A \cap (A+h)$ est non-vide) pour tout $h < h_0$, où $h_0 > 0$ est pris suffisamment petit. Ainsi, pour tout $h < h_0$, il existe $x, y \in A$, avec $x = y + h$, soit $h = x - y$ – et $]-h_0, h_0[$ est inclus dans $A - A$.

Exercice 7. Soit $[a, b]$ incluant le support compact de f : $a > 0$ car le support de f est compact inclus dans $\mathbb{R}_+^*$. Le support de F est contenu dans $[a, +\infty[$, et il n'y a pas de problème d'intégrabilité pour F en 0^+. De plus, F est continue, bornée et égale à $x^{-1} \int_{\mathbb{R}_+} f \, \mathrm{d}\lambda$ pour x assez grand, donc $F \in L^p$, $p > 1$.

Ensuite, la dérivée de $x \mapsto xF(x)$ est $f(x) = xF'(x) + F(x)$, de sorte qu'en intégrant par parties (les intégrales que l'on considère le sont indifféremment au sens de Riemann ou de Lebesgue – vérifiez qu'elles sont bien toutes définies) :

$$\int_0^\infty F(x)^p \, \mathrm{d}x = [xF(x)^p]_0^\infty - \int_0^\infty pxF(x)^{p-1}F'(x) \, \mathrm{d}x = p \int_0^\infty F(x)^{p-1}(F - f)(x) \, \mathrm{d}x \ ,$$

ce qui donne le premier résultat.

On applique ensuite l'inégalité de Hölder (pour les fonctions positives) au second membre de la première inégalité démontrée, pour obtenir

$$\int_0^\infty F(x)^p \, \mathrm{d}x \leq \frac{p}{p-1} ||f||_p \left(\int_0^\infty F(x)^p \, \mathrm{d}x \right)^{(p-1)/p} \ ,$$

ce qui conclut, F étant de puissance p-ième intégrable.

Pour f de signe quelconque, mais continue à support compact, on note $g = |f|$, et, avec des notations évidentes, le point précédent donne $||G||_p \leq (p-1)^{-1}p||f||_p$; comme par ailleurs $|F| \leq G$, on a bien l'extension.

Ensuite, dans le cas où $f \in L^p$ sans plus de précision, soit une suite (f_n) de fonctions continues à support compact convergeant vers f dans L^p – et étudions la convergence de la suite (F_n) qui lui est associée. On constate que

$$|F_n(x) - F(x)| \leq x^{-1} \int_0^x |f_n(t) - f(t)| \, \mathrm{d}t \leq x^{-1} \left(\int_0^x \mathrm{d}t \right)^{1/q} \left(\int_0^x |f_n(t) - f(t)|^p \, \mathrm{d}t \right)^{1/p} \ ,$$

où q est l'exposant conjugué de p. Comme $f_n \to f$ dans L^p on en déduit que $F_n(x) \to F(x)$ pour tout $x > 0$. On conclut par le lemme de Fatou ($||F||_p \leq \liminf ||F_n||_p$) et l'inégalité de Hardy appliquée aux f_n ($\liminf ||F_n||_p \leq (p-1)^{-1}p \lim ||f_n||_p = (p-1)^{-1}p||f||_p$). On peut aussi utiliser l'inégalité de Hardy pour voir que (F_n) est de Cauchy dans L^p et utiliser la complétude de $\mathbb{L}^p$.

Intégration et Probabilités – TD 6
Corrigé partiel

Exercice 1. Soit $0 < \varepsilon < 1$ et $A_\varepsilon = \{|f| > ||f||_\infty(1-\varepsilon)\}$. Par définition, $\mu(A_\varepsilon) > 0$ et on a directement $||f||_p \geq ||f||_\infty(1-\varepsilon)\mu(A_\varepsilon)^{1/p}$. Ainsi, que l'on ait $\mu(A_\varepsilon)$ fini ou infini, on obtient que $\liminf_{p\to\infty} ||f||_p \geq ||f||_\infty$. Par ailleurs, si $||f||_{p_0} < \infty$, on a pour $p > p_0$:

$$\int_E f^p \mathrm{d}\mu \leq ||f||_\infty^{p-p_0} \int_E |f|^{p_0} \,\mathrm{d}\mu < \infty,$$

et donc $\limsup_{p\to\infty} ||f||_p \leq ||f||_\infty$ (en passant aux logarithmes ci-dessus et en divisant les deux membres par p), ce qui conclut.

Constatons tout d'abord qu'il suffit de montrer que $\lim_{p\downarrow 0} p^{-1} \ln \mu\,|f|^p = \mu \ln |f|$. L'inégalité de Jensen donne immédiatement le sens $\geq$, avec une limite inférieure au lieu de la limite. Prouvons le sens $\leq$. Quand $p \downarrow 0$, $|f|^p$ converge alors vers $\mathbb{I}_{|f|>0}$ presque-partout de façon monotone croissante sur $\{|f| < 1\}$ et vers 1 de façon dominée (par $|f|^{p_0}$) sur $\{|f| \geq 1\}$. Par conséquent, $\mu\,|f|^p$ converge vers $\mu\{|f| > 0\}$. Si cette quantité est < 1, le résultat est alors acquis (la lim sup du premier membre vaut $-\infty$). Sinon, $\ln \mu\,|f|^p$ tend vers 0, donc $\ln \mu\,|f|^p \sim (\mu\,|f|^p) - 1 = \mu[|f|^p - 1]$. Il nous reste donc seulement à déterminer la limite de $p^{-1}\mu[|f|^p - 1]$. Pour ce faire, constatons que $(x^p - 1)/p = (e^{p\ln x} - 1)/p$, et donc la limite ponctuelle de $(|f|^p - 1)/p$ est bien $\ln |f|$ (on rappelle que l'on est dans le cas $f > 0$ μ-p.s.).

On voit par étude de fonction que la fonction $p \mapsto (e^{\lambda p} - 1)/p$ prolongée en 0 par λ, pour $\lambda \in \mathbb{R}$ est croissante. Par conséquent (comme on fait décroître p vers 0), ayant $1-|f|^p \geq 0$ sur $\{|f| < 1\}$, on a convergence monotone de $p^{-1}\mu[(1-|f|^p)\mathbb{I}_{\{|f|<1\}}]$ vers $\mu[(-\ln|f|)\mathbb{I}_{\{|f|<1\}}]$, tandis que l'on a convergence dominée (par exemple par $p_0^{-1}(|f|^{p_0}-1)$) de $\mu[p^{-1}(|f|^p-1)\mathbb{I}_{\{|f|>1\}}]$ vers $\mu[\ln|f|\,\mathbb{I}_{\{|f|>1\}}]$ (le cas $|f| = 1$ n'est pas important puisque toutes les intégrales correspondantes sont nulles). On conclut par soustraction (légitime, parce que l'on est sûr que les quantités provenant du raisonnement par convergence dominée sont finies).

Exercice 2. La fonction F ne pose pas de problème de définition car $L^p([0,x]) \subset L^1([0,x])$ pour tout $x > 0$. De plus,

$$|F(x+h) - F(x)| \leq \left(\int_x^{x+h} f(t)^p \mathrm{d}t\right)^{1/p} |h|^{1/q}$$

par l'inégalité de Hölder. Pour conclure, on applique la propriété d'uniforme continuité de l'intégrale (voir TD2, exercice 4).

Si g satisfait aux hypothèses de l'énoncé, en appliquant ce qui précède à $f = g'$, on obtient que $\lim_{h\to 0} |h|^{-1/q} \sup_{x\in\mathbb{R}} |g(x+h) - g(x)| = 0$. Si g ne tendait pas vers 0, on pourrait alors trouver $\varepsilon > 0$ et une suite $t_1 < t_2 < \ldots$ tendant vers ∞ tels que $|g(t_i)| > \varepsilon$ pour tout i ; de plus si $h > 0$ est tel que $h^{1/q} < \varepsilon/2$ et $|h|^{-1/q} \sup_{x\in\mathbb{R}} |g(x+h) - g(x)| \leq 1$, et en supposant, sans perte de généralité, qu'on a choisi les t_i espacés d'au moins $2h$ les uns des autres, on déduit que $\int_\mathbb{R} |g(x)|^p \mathrm{d}x \geq \infty(\varepsilon/2)^p = \infty$ – ce qui est impossible.

Exercice 3.
Il suffit de montrer que pour $A_1 \in \mathcal{A}_1^{\mu_1}$ et $A_2 \in \mathcal{A}_2^{\mu_2}$, $A_1 \times A_2 \in (\mathcal{A}_1 \otimes \mathcal{A}_2)^{\mu_1\otimes\mu_2}$. Mais

quasi par définition[1] de la négligeabilité, il existe B_1 et C_1 dans $\mathcal{A}_1$, B_2 et C_2 dans $\mathcal{A}_2$, tels que $B_1 \subseteq A_1 \subseteq C_1$ et $\mu_1[C_1 \setminus B_1] = 0$ (et de même pour B_2, C_2 et μ_2). On a ainsi $B_1 \times B_2 \subseteq A_1 \times A_2 \subseteq C_1 \times C_2$, et

$$C_1 \times C_2 \setminus B_1 \times B_2 \subseteq ((C_1 \setminus B_1) \times C_2) \cup (C_1 \times (C_2 \setminus B_2)) ,$$

et le membre de droite est bien de $\mu_1 \otimes \mu_2$-mesure nulle.

Pour montrer que l'inclusion est stricte, on prend $\mathcal{A}_1 = \mathcal{A}_2$ les boréliens sur [0,1] et $\mu_1 = \mu_2$ la mesure de Lebesgue. Soit $H = \{0\}$ et K un ensemble non-Lebesgue mesurable[2]. $H \times K$ est $\mu_1 \otimes \mu_2$-négligeable, mais non $\mathcal{A}_1^{\mu_1} \otimes \mathcal{A}_2^{\mu_2}$-mesurable (raisonner par l'absurde en utilisant une application de projection[3]).

Exercice 4. On écrit

$$g(x) = \int_0^\infty g'(t)\mathbb{I}_{\{0 \le t \le x\}}\, \mathrm{d}t .$$

Comme g est croissante, $g' \ge 0$ et le théorème de Fubini-Tonelli donne

$$\mu[g \circ f] = \int_E \mathrm{d}\mu(x) \int_0^\infty \mathrm{d}t\, g'(t)\mathbb{I}_{\{0 \le t \le f(x)\}} = \int_0^\infty g'(t)\mathrm{d}t \int_E \mathbb{I}_{\{f(x) \ge t\}}\mathrm{d}\mu,$$

ce qui conclut.

Exercice 5. Soit I l'intégrale double à calculer. Par le théorème de Fubini-Tonelli et le changement de variables linéaire $u = \sqrt{y}x$, on a

$$I = \int_{\mathbb{R}_+} \frac{\mathrm{d}y}{1+y} \int_{\mathbb{R}_+} \frac{\mathrm{d}x}{1+x^2y} = \int_{\mathbb{R}_+} \frac{\mathrm{d}y}{(1+y)\sqrt{y}} \int_{\mathbb{R}_+} \frac{\mathrm{d}u}{1+u^2} = 2\left(\int_{\mathbb{R}_+} \frac{\mathrm{d}v}{1+v^2}\right)^2 ,$$

où la dernière égalité provient du changement de variables $v = \sqrt{y}$. Ainsi, $I = \pi^2/2$. Par ailleurs, on décompose en élements simples l'intégrande (vue comme fonction de y),

$$\frac{1}{(1+y)(1+x^2y)} = \frac{1/(1-x^2)}{1+y} - \frac{x^2/(1-x^2)}{1+x^2y} \ge 0.$$

On intègre par rapport à y (sur $[1/A,A]$ puis on fait tendre A vers $+\infty$) pour conclure à

$$I = \int_{\mathbb{R}+} \frac{\ln(x^2)}{x^2-1}\, \mathrm{d}x = \frac{\pi}{4} .$$

Exercice 6. Une application du théorème de Fubini-Tonelli donne

$$\int_{\mathbb{R}^2} \mathrm{d}x\mathrm{d}y\, \mathbb{I}_{\{f(x) \ge y \ge 0\}} = \int_{\mathbb{R}} \mathrm{d}x\, f(x).$$

De même,

$$\int_{\mathbb{R}^2} \mathrm{d}x\mathrm{d}y\, \mathbb{I}_{\{f(x) = y\}} = \int_{\mathbb{R}} \mathrm{d}x\, \lambda(\{f(x)\}) = 0,$$

et donc le graphe d'une fonction mesurable est de mesure nulle, soit $\lambda\{f(x)\} = 0$ pour presque tout x. soit encore, $\{f = y\}$ est de mesure nulle pour presque tout y.

1. voir preuve la Proposition 3.2.3, page 34 du polycopié
2. pour l'existence d'un tel ensemble : polycopié, Section 3.4
3. voir Théorème 5.1.2 du polycopié, page 56

Exercice 7. On introduit la fonction $h : (x,y) \mapsto (f(x) - f(y))(g(x) - g(y))$. h est positive car f et g sont monotones de même sens. On constate que

$$\int_{\mathbb{R}} h(x,y)\,\mathrm{d}\mu(x,y) = 2\int_{\mathbb{R}} fg\,\mathrm{d}\mu - 2\int_{\mathbb{R}} f\,\mathrm{d}\mu \int_{\mathbb{R}} g\,\mathrm{d}\mu$$

en utilisant le théorème de Fubini-Tonelli et (toutes les quantités considérées étant finies) la linéarité de l'intégrale. On conclut en ré-utilisant la positivité de h.

Exercice 8. La méthode la plus naturelle pour cet exercice est un changement de variables. Néanmoins, une solution utilisant le théorème de Fubini est également instructive. On remarque d'abord que les normes sont équivalentes dans $\mathbb{R}^k$, ce qui fait que la question revient à l'appartenance de $x \to ||x||_\infty^{-p\alpha}$ à $L^1([0,1]^k)$. On décompose alors $[0,1]^k$ suivant les k régions où $x_1 = ||x||_\infty, x_2 = ||x||_\infty, \ldots$ (les bords et intersection de ces régions sont de mesure nulle), et par symétrie, l'intégrale sur chacune de ces régions est la même. Par le théorème de Fubini-Tonelli,

$$\begin{aligned}
\int_{[0,1]^k} ||x||_\infty^{-p\alpha}\mathrm{d}x &= k\int_{\{x_1=\|x\|_\infty\}} x_1^{-p\alpha}\mathrm{d}x \\
&= k\int_{\{x_1\in[0,1]\}} x_1^{-p\alpha}\mathrm{d}x_1 \int_{\{(x_2,\ldots,x_k)\in[0,x_1]^{k-1}\}} \mathrm{d}x_2\ldots\mathrm{d}x_k \\
&= \int_0^1 x_1^{-p\alpha+k-1}\mathrm{d}x_1 \ ,
\end{aligned}$$

et cette dernière expression est finie si et seulement si $p\alpha < k$.

Exercice 9. En remarquant que $x^2 - y^2 = x^2 + y^2 - 2y^2$ au numérateur, une intégration par parties, pour $x > 0$, en posant $u(y) = 1/(x^2 + y^2)$ et $v(y) = y$, donne

$$\int_0^1 \mathrm{d}y f(x,y) = \int_0^1 \frac{\mathrm{d}y}{x^2+y^2} + \frac{1}{1+x^2} - \int_0^1 \frac{\mathrm{d}y}{x^2+y^2} = \frac{1}{1+x^2} \ .$$

Donc

$$\int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}y f(x,y) = \operatorname{Arctg} 1 = \frac{\pi}{4} \ .$$

Par symétrie, on a de même que

$$\int_0^1 \mathrm{d}y \int_0^1 \mathrm{d}x f(x,y) = -\frac{\pi}{4} \ .$$

De ce fait, le théorème de Fubini-Lebesgue ne s'applique pas à f, dont on conclut qu'elle n'est pas dans $L^1([0,1])$.

Exercice 10. On écrit comme indiqué

$$I = \int_0^\infty x^{-1}\sin(x)\,\mathrm{d}x$$

(c'est une intégrale impropre définie comme étant la limite de l'intégrale I_A sur $[0,A]$ de la même expression quand $A \to \infty$) sous la forme

$$\int_0^\infty \mathrm{d}x \sin(x) \int_0^\infty \mathrm{d}t\, e^{-tx} \ .$$

On remplace la borne $+\infty$ de la première intégrale par A et on peut alors appliquer le théroème de Fubini-Lebesgue,

$$\begin{aligned} I_A = \int_0^\infty \mathrm{d}t \int_0^A \mathrm{d}x\, e^{-tx} \sin(x) &= \int_0^\infty \mathrm{d}t \int_0^A \mathrm{d}x \;\mathrm{Im}\left(e^{(-t+i)x}\right) \\ &= \int_0^\infty \mathrm{d}t \left(\frac{1}{1+t^2} + \mathrm{Im}\left(\frac{e^{(i-t)A}}{i-t}\right)\right). \end{aligned}$$

Par convergence dominée, l'intégrale du terme dépendant de A tend vers 0 quand $A \to \infty$ (dominer la partie imaginaire par le module), et on conclut que $I = \pi/2$.

Intégration et Probabilités – TD 7
Corrigé partiel

Exercice 1. Le fait que $\star$ soit une loi de composition interne sur $\mathbb{L}^1(\mathbb{R})$ a été prouvé en cours. La bilinéarité et l'associativité découlent facilement du théorème de Fubini-Lebesgue. Le point intéressant de l'exercice est que $\star$ n'admet pas d'élément neutre[1]. Raisonnons par l'absurde et supposons qu'il existe $\delta \in \mathbb{L}^1(\mathbb{R})$ telle que $\delta \star f = f$ pour tout $f \in \mathbb{L}^1(\mathbb{R})$. On considère la suite de fonctions $(f_n)_{n\geq 1}$ définie par $f_n = f(n\,\cdot)$, où $f : x \mapsto (1-|x|)^+$. Alors,

$$\delta \star f_n\,(0) = f_n(0) = 1 = \int_{\mathbb{R}} f_n(-x)\delta(x)\,\mathrm{d}x\ .$$

Par ailleurs, notons $g_n = f_n(-\,\cdot)\delta$; la suite $(g_n)_{n\geq 1}$ converge ponctuellement vers 0, et elle est dominée par $\delta \in \mathbb{L}^1(\mathbb{R})$, de sorte que le théorème de convergence dominée assure que

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n(-x)\delta(x)\,\mathrm{d}x = 0\ .$$

La contradiction conclut.

Exercice 2. On écrit, pour tout $x \in \mathbb{R}$, $g(x,\cdot)f(\cdot) \leq \|g\|_\infty\ |f| \in \mathbb{L}^1(\mathbb{R})$. $f \star g$ est donc définie en tout point, bornée ; et par application du théorème de convergence dominée (et par continuité de g), elle est continue.

Lorsque g est de classe C^1, avec pour dérivée g' continue bornée, on applique le théorème de dérivation sous le signe somme (puisque

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(g(x,\cdot)f(\cdot)\right) = g'(x,\cdot)f(\cdot) \leq \|g'\|_\infty\ |f| \in \mathbb{L}^1(\mathbb{R})$$

de même que plus haut). On a alors que $f \star g$ est dérivable, de dérivée $f \star g'$, qui est continue, bornée par le point précédent (et donc $f \star g$ est C^1).

Exercice 3. Soit une suite $(f_n)_{n\geq 1}$ (resp., $(g_n)_{n\geq 1}$) de fonctions continues à support compact tendant dans $\mathbb{L}^p(\mathbb{R})$ vers f (resp., dans $L^q(\mathbb{R})$ vers g). Alors, d'abord par inégalité triangulaire ponctuelle sur $\mathbb{R}$, puis en utilisant l'inégalité rappelée sur les normes,

$$\begin{aligned}
\|f \star g - f_n \star g_n\|_\infty &\leq \|\,|f - f_n| \star |g|\,\|_\infty + \|\,|f_n| \star |g - g_n|\,\|_\infty \\
&\leq \|f - f_n\|_p \|g\|_q + \|f_n\|_p \|g - g_n\|_q \quad \xrightarrow[n\to\infty]{} 0\ .
\end{aligned}$$

Or, un calcul direct montre que chaque $f_n \star g_n$ est encore à support compact, donc de limite nulle aux infinis. On vient de prouver que $(f_n \star g_n)_{n\geq 1}$ converge uniformément vers $f \star g$, et cette dernière est donc également de limite nulle aux infinis.

Exercice 4. Remarque générale : vu que $\bar{\omega}$ est un compact de $\mathbb{R}$, il est borné. On peut donc supposer sans perte de généralité que Ω lui-même était borné. On note M un majorant de la norme $\|\cdot\|_p$ des éléments de $\mathcal{F}$.

---

1. on peut définir la convolution pour des mesures – et à tout élement de $\mathbb{L}^1$ correspond une mesure (signée) – mais l'élément neutre de la convolution des mesures est δ_0, la masse de Dirac en 0, qui ne peut pas être représentée par un élément de $\mathbb{L}^1$!

Pour (1), on vérifie que la famille $\mathcal{H}_n$ satisfait aux hypothèses du théorème d'Ascoli. Premièrement, que les éléments de $\mathcal{H}_n$ prennent leurs valeurs dans un ensemble relativement compact (i.e. un borné) de $\mathbb{R}$. Ω étant borné, l'inégalité de Jensen (ou celle de Hölder) donne, pour tout $f \in \mathbb{L}^p(\Omega)$,

$$\|f\|_{\mathbb{L}^1(\Omega)} \leq \lambda(\Omega)^{1/q}\|f\|_{\mathbb{L}^p(\Omega)} \leq C\lambda(\Omega)^{1/q} \ ,$$

et alors, une majoration directe montre que

$$\|\rho_n \star \bar{f}\|_{\mathbb{L}^\infty(\mathbb{R})} \leq \|\rho_n\|_{\mathbb{L}^\infty(\mathbb{R})}\|\bar{f}\|_{\mathbb{L}^1(\mathbb{R})}] \leq C\lambda(\Omega)^{1/q}\|\rho_n\|_{\mathbb{L}^\infty(\mathbb{R})} = C_n < +\infty \ .$$

D'autre part, on a bien l'équicontinuité : on note $\|\rho_n\|_L$ un majorant de la norme de Lipschitz de ρ_n (elle est finie, parce que ρ_n est dérivable, de dérivée uniformément bornée) ; alors, un autre calcul direct montre que pour $x, y \in \mathbb{R}$,

$$\left| \left(\rho_n \star \bar{f}\right)(x) - \left(\rho_n \star \bar{f}\right)(y) \right| \leq |x-y| \, \|\rho_n\|_L \|\bar{f}\|_{\mathbb{L}^1(\mathbb{R})}] \leq C\|\rho_n\|_L \, |x-y| \ .$$

Ainsi, $\mathcal{H}_n$ est relativement compact dans $C(\bar{\omega})$, et donc aussi dans $\mathbb{L}^p(\omega)$. En effet, l'inclusion $(C(\bar{\omega}), \|\cdot\|_\infty) \hookrightarrow (\mathbb{L}^p(\omega), \|\cdot\|_p)$ est continue.

Pour (2), on note que comme chaque ρ_n est positive et s'intègre à 1, l'inégalité de Jensen (ou celle de Hölder) appliquée à la mesure de probabilité $\rho_n \mathrm{d}\lambda$ montre que

$$\begin{aligned} \left| \left(\rho_n \star \bar{f}\right)(x) - \bar{f}(x) \right| &\leq \int_{\mathbb{R}} \left| \bar{f}(x-y) - \bar{f}(x) \right| \rho_n(y)\mathrm{d}y \\ &\leq \left(\int_{\mathbb{R}} \left| \bar{f}(x-y) - \bar{f}(x) \right|^p \rho_n(y)\mathrm{d}y \right)^{1/p} \ , \end{aligned}$$

de sorte que

$$\left| \left(\rho_n \star \bar{f}\right)(x) - \bar{f}(x) \right|^p \leq \int_{[-1/n,\, 1/n]} \left| \bar{f}(x-y) - \bar{f}(x) \right|^p \rho_n(y)\mathrm{d}y \ ,$$

et par théorème de Fubini-Tonelli,

$$\int_\omega \left| \left(\rho_n \star \bar{f}\right)(x) - \bar{f}(x) \right|^p \mathrm{d}x \leq \int_{[-1/n,\, 1/n]} \mathrm{d}y \, \rho_n(y) \int_\omega \left| \bar{f}(x-y) - \bar{f}(x) \right|^p \mathrm{d}x \ ,$$

d'où l'on conclut aisément en utilisant la seconde hypothèse du théorème.

On conclut en remarquant, par complétude de $\mathbb{L}^p(\omega)$, qu'il suffit de montrer la précompacité de l'adhérence de $\mathcal{F}_{|\omega}$. En particulier, il suffit également de montrer que pour tout $\varepsilon > 0$, il est possible de recouvrir $\mathcal{F}_{|\omega}$ par un nombre fini de boules de rayon ε. Par (2), on prend n suffisamment grand pour que $\|\bar{f} - \rho_n \star \bar{f}\|_{\mathbb{L}^p(\omega)} < \varepsilon/2$ pour tout $\bar{f} \in \bar{\mathcal{F}}$. Comme $\mathcal{H}_n$ est relativement compact dans $\mathbb{L}^p(\omega)$, il en existe un recouvrement fini par des boules de rayon $\varepsilon/2$. Les boules correspondantes de rayon ε recouvrent alors $\mathcal{F}_{|\omega}$.

Pour l'application (4), on note que $\mathcal{F}$ est bien borné dans $\mathbb{L}^p(\mathbb{R})$, par $M\|g\|_{\mathbb{L}^1(\mathbb{R})}$. D'autre part, si $f = g \star u$, avec $u \in \mathcal{B}$, on a

$$\|\tau_h f - f\|_{\mathbb{L}^p(\mathbb{R})} \leq M\|\tau_h g - g\|_{\mathbb{L}^1(\mathbb{R})} \ ,$$

et par l'exerice 6 du TD 5, le second membre de l'inégalité tend vers 0. On conclut par théorème de Riesz-Fréchet-Kolmogorov.

Intégration et Probabilités – TD 8
Corrigé partiel

Exercice 1. Il s'agit de voir que pour tout $\lambda > 0$, l'ensemble $E_\lambda = \{M\mu > \lambda\}$ est ouvert. Soit $x \in E_\lambda$: il existe $r > 0$ et $t > \lambda$ tels que $|\mu|\,(B(x,r)) = t\,m(B(x,r))$. Nous allons montrer qu'un certain voisinage $B(x,\delta)$ de x est inclus dans E_λ (où δ est déterminé par le raisonnement ci-dessous). Soit $y \in B(x,\delta)$; alors $B(x,r) \subset B(y,r+\delta)$, de sorte que

$$|\mu|\,(B(y,r+\delta)) \geq |\mu|\,(B(x,r)) \geq t\,m(B(x,r)) = t\,\frac{r}{r+\delta}\,m(B(y,r+\delta)) > \lambda\,m(B(y,r+\delta))$$

pour, par exemple, $\delta = r\,(t/\lambda - 1) > 0$.

Remarque : dans le cas où $|\mu|$ est diffuse, $M\mu$ apparaît comme l'enveloppe supérieure d'une famille de fonctions continues et est, à ce titre, semi-continue inférieurement.

Montrons tout d'abord l'indication. Il suffit d'ordonner les boules $B_i = B(x_i,r_i)$ suivant les rayons décroissants $r_1 \geq r_2 \geq \ldots \geq r_N$. On pose $i_1 = 1$, et on élimine toutes les boules B_j, $j \geq i_1 + 1$, qui intersectent B_{i_1}. Soit i_2 le plus petit indice de ceux des boules restantes, s'il y en a. On élimine à nouveau les B_j, $j \geq i_2 + 1$, qui intersectent B_{i_2} – et on recommence un nombre fini de fois, pour obtenir $S = \{i_1, i_2, \ldots\}$. La disjonction est obtenue par construction. Et quant au recouvrement par les $B(x_s,3r_s)$, $s \in S$, il se prouve en remarquant que si une boule B_j a été éliminée par une boule $B(x_s,r_s)$, alors, vu l'ordonnancement des rayons et par inégalité triangulaire, elle est en fait incluse dans $B(x_s,3r_s)$.

Pour prouver l'inégalité (2), on fixe $\lambda > 0$, et on considère l'ouvert $E_\lambda = \{M\mu > \lambda\}$. Soit K un compact inclus dans E_λ : nous allons montrer que $m(K) \leq 3\|\mu\|/\lambda$, ce qui conclura par régularité de la mesure de Lebesgue. Tout point $x \in K$ est le centre d'une boule ouverte B pour laquelle $|\mu|(B) > \lambda\,m(B)$. Une réunion finie de telles boules recouvre K, et on en extrait une famille finie disjointe $B(z,r_z)$, $z \in \mathcal{Z}$, telle que K est recouvert par les $B(z,3r_z)$, $z \in \mathcal{Z}$. Alors,

$$m(K) \leq \sum_{z\in\mathcal{Z}} m(B(z,3r_z)) \leq 3\sum_{z\in\mathcal{Z}} m(B(z,r_z)) \leq 3\sum_{z\in\mathcal{Z}} \frac{|\mu|\,(B(z,r_z))}{\lambda} \leq 3\frac{\|\mu\|}{\lambda}\ ,$$

où la dernière inégalité est vraie par disjonction des boules considérées.

Dans le cas où la mesure μ est absolument continue par rapport à m, avec pour densité $f \in \mathbb{L}^1(m)$, $|\mu|$ est encore absolument continue par rapport à m, de densité $|f|$, et la fonction maximale vaut

$$Mf(x) = \sup_{0<r<\infty} \frac{1}{m(B(x,r))} \int_{m(B(x,r))} |f|\ \mathrm{d}m\ ,$$

tandis que $\|\mu\| = \|f\|_1 = m[|f|]$.

Exercice 2. On remarque tout d'abord que tout point de continuité de f est un point de Lebesgue. Or, on sait qu'une fonction intégrale n'est pas très éloignée d'une fonction continue... Quelques notations, pour la preuve : pour $x \in \mathbb{R}$ et $r > 0$, on note

$$T_r f(x) = \frac{1}{m(B(x,r))} \int_{B(x,r)} |f(y) - f(x)|\ \mathrm{d}m(y)\ ,$$

et on considère $Tf = \limsup_{r\to 0} T_r f$. Il s'agit de prouver que $Tf = 0$ p.p. Soit $y > 0$: montrons que $\{Tf > 2y\}$ est de mesure nulle. Pour tout entier n, il existe g continue sur $\mathbb{R}$ telle que $\|f - g\|_1 \leq 1/n$. Posons $h = f - g$; comme $T_r f \leq T_r g + T_r h$, et que g est continue (donc

$T_r g \to 0$), on a, par passage aux lim sup, $Tf \le Th$. Par ailleurs, par inégalité triangulaire, $Th \le Mh + |h|$, de sorte que

$$\{Tf > 2y\} \subset \{Th > 2y\} \subset \{Mh > y\} \cup \{|h| > y\} \ .$$

Par l'exercice 1,

$$m\,\{Mh > y\} \le \frac{3\|h\|_1}{y} \le \frac{3}{y\,n} \ ,$$

tandis que l'inégalité de Markov montre que

$$m\,\{|h| > y\} \le \frac{\|h\|_1}{y} \le \frac{1}{y\,n} \ ,$$

de sorte que, par union,

$$m\{Tf > 2y\} \le \frac{4}{y\,n} \ .$$

Cette inégalité étant vraie pour tout n, et le membre de gauche ne dépendant pas de n, on a bien que $\{Tf > 2y\}$ est de mesure nulle pour tout $y > 0$, et la conclusion s'en suit.

Exercice 3. Par définition d'une suite rétrécissant convenablement sur x,

$$\frac{\alpha}{m(E_n)} \int_{E_n} |f - f(x)| \,\mathrm{d}m \le \frac{1}{B(x,r_n)} \int_{B(x,r_n)} |f - f(x)| \,\mathrm{d}m \ ;$$

or x est de Lebesgue et $r_n \to 0$, donc le membre de droite converge vers 0. On conclut en utilisant $\alpha > 0$.

Pour prouver le théorème, il suffit de considérer un point de Lebesgue x, et une suite de nombres δ_n tendant vers 0 ; puis la suite d'ensembles $E_n = [x, x+\delta_n[$, qui rétrécit bien convenablement sur x (avec $\alpha = 1/2$). Le résultat précédent s'applique alors, et il implique en particulier la dérivabilité à droite de F en x, avec $f(x)$ pour valeur de la dérivée à droite en x. On fait de même à gauche en considérant les $E_n = [x - \delta_n, x[$.

Exercice 4. Cela procède simplement de la propriété d'uniforme continuité de l'intégrale (voir TD2, exercice 4). En effet, lorsque le résultat du Théorème 2 est vrai, et avec les notations de la définition d'AC,

$$\sum_{i=1}^{n} |f(\beta_i) - f(\alpha_i)| \le \int_A |\, f' \,| \,\mathrm{d}m \ ,$$

où A est de mesure plus petite que δ, comme l'union des intervalles $[\alpha_i,\beta_i]$, dont la somme des longueurs est plus petite que δ.

Remarque : la propriété d'uniforme continuité des intégrales se retrouve en considérant la mesure ν, définie par $\mathrm{d}\nu = f' \cdot \mathrm{d}m$, qui est bien absolument continue par rapport à m. Voir le Théorème 6.2.3. du polycopié.

Exercice 5. On remarque tout d'abord que la définition de l'absolue continuité s'étend à une réunion *dénombrable* d'intervalles disjoints dont la somme des longueurs est plus petite que δ (on a le résultat pour toutes les sommes partielles finies de la série).

Soit donc $E \subset I$, de mesure $m(E) = 0$. Fixons $\varepsilon > 0$, et soit $\delta > 0$ lui correspondant dans la définition de l'absolue continuité de f. Il existe un ouvert V' de I, contenant E, et de mesure plus petite que δ. On considère $V = V' \setminus \{a,b\}$, qui est un ouvert de $\mathbb{R}$, inclus dans

I et de mesure plus petite que δ. On l'écrit sous la forme d'une union d'intervalles disjoints $]\alpha_i,\beta_i[$: par absolue continuité de f,

$$\sum_i |f(\beta_i) - f(\alpha_i)| \leq \varepsilon \ .$$

Or, f étant en particulier continue (et croissante), le théorème des valeurs intermédiaires assure que

$$f(E) \subset f(V') \subset f(\{a,b\} \cup V) \subset \{f(a), f(b)\} \cup \bigcup_i [f(\alpha_i), f(\beta_i)] \ ,$$

et $f(E)$ est ainsi de m-mesure plus petite que ε. Ce dernier étant arbitraire, c'est que $f(E)$ est de m-mesure nulle.

g est également AC (comme somme de deux fonctions AC). Soit E un ensemble Lebesgue-mesurable : on l'écrit comme l'union entre un borélien E_1 et un ensemble de mesure nulle E_0. On a $g(E) = g(E_0) \cup g(E_1)$. Par le point précédent, g étant AC, $g(E_0)$ est encore de mesure nulle. f étant croissante (au sens large), g est strictement croissante, donc injective (et c'est pour ça qu'on l'a préférée à f). Par ailleurs, g est en particulier continue, et on sait que l'image directe d'un borélien par une application injective continue est encore un borélien. (Pour le redémontrer, considérer l'ensemble des boréliens dont l'image directe par g est borélienne : c'est une tribu[1], contenant, par le théorème des valeurs intermédiaires, tous les intervalles.) Ainsi, $g(E_1)$ est borélien, et $g(E)$ est bien Lebesgue-mesurable.

Pour prouver que Φ est une mesure, il suffit de montrer l'additivité dénombrable pour des ensembles disjoints. On réutilise ici l'injectivité de g : l'image par g d'une suite d'ensembles disjoints dans I est une suite d'ensembles disjoints (dans $\mathbb{R}$), et l'additivité dénombrable de la mesure de Lebesgue m conlut.

Puisque g est AC, l'image par g d'un ensemble de m-mesure nulle est encore de m-mesure nulle, de sorte que $\mu = \Phi$ est absolument continue par rapport à m. Le théorème de Radon-Nikodym fournit alors $h \in \mathbb{L}^1(\mathcal{M})$ (positive) telle que $\mathrm{d}\mu = h \cdot \mathrm{d}m$.

Pour conclure, on fixe $x \in I$, et on considère $E = [a,x]$. g étant croissante, on a $g(E) = [g(a),g(x)]$, et le point précédent permet d'écrire

$$g(x) - g(a) = \mu(E) = m(g(E)) = \int_E h \mathrm{d}m = \int_a^x h(t)\, \mathrm{d}t \ .$$

Par définition de g, cela se réécrit,

$$f(x) - f(a) = \int_a^x (h(t) - 1)\ \mathrm{d}t \ .$$

Puisque $h - 1 \in \mathbb{L}^1(I,m)$, le Théorème 1 (appliqué à f et $h-1$ correctement prolongées en dehors de I) assure alors que f est dérivable presque-partout, de dérivée $f' = h - 1$.

Généralisation : pour avoir le Théorème 2 pour f simplement supposée AC, il suffit de voir que toute fonction AC f peut s'écrire comme la différence $f_1 - f_2$, où f_1, f_2 sont AC et croissantes. C'est là le point un peu calculatoire et peu intéressant que j'ai sauté. (Vous savez que f à variation bornée s'écrit comme la différence entre deux fonctions croissantes ; et une fonction AC est à variation bornée ; mais on veut aussi que les deux fonctions croissantes soient encore AC.) Voir le livre de Rudin cité ci-dessus pour les détails.

1. l'injectivité de g sert pour montrer la stabilité par passage au complémentaire

Remarque : Il n'est pas vrai en général que les images continues de boréliens soient encore boréliennes, ... contrairement à ce que l'intuition suggérait à Borel (et plus modestement, à votre chargé de TDs). Suite à cette erreur de Borel, on s'est alors intéressé aux ensembles dits analytiques, images directes des boréliens par des applications continues. Ces derniers sont heureusement encore Lebesgue-mesurables (et c'est ce qu'il nous fallait !).

Cela dit, il reste vrai que, dans $\mathbb{R}$ (et plus généralement dans des espaces polonais), les images directes de boréliens par des applications bijectives mesurables sont encore boréliennes. (Les applications bijectives boréliennes sont donc bi-mesurables.)

Exercice 1. On peut prendre pour espace de probabilités l'ensemble $\Omega = \{1,\ldots,n\}^{\{1,\ldots,r\}}$ des applications de $\{1,\ldots,r\}$ dans $\{1,\ldots,n\}$ (on distingue les boules, et on leur attribue une case), muni de la tribu des parties et de la mesure de comptage renormalisée par n^r. En fait, on aurait pu considérer des boules indistinguables, et l'espace probabilisé correspondant aurait été l'ensemble des suites à n termes de somme k (chaque élément de la suite indique le nombre de boules dans chaque case), mais le cardinal de cet ensemble est moins facile à déterminer (et du coup, le facteur de renormalisation de la mesure de comptage est moins aisé à calculer).

On considère la variable aléatoire X définie par $X(\omega) = \mathrm{Card}\,\{1 \le i \le r : \omega(i) = 1\}$. Alors $\mu_{r,n}(k) = P[X = k]$ vaut $C_r^k(n-1)^{r-k}/n^r$. On le prouve en remarquant que par définition, X suit une loi binômiale $(r,1/n)$, ou par le raisonnement à la hussarde qui suit. Une configuration $[X = k]$ est donnée par le choix de k boules parmi les r (k boules que l'on place dans la case 1), et une répartition quelconque des $r-k$ boules restantes dans les $n-1$ cases restantes. Il y a donc $C_r^k(n-1)^{r-k}$ telles configurations, à comparer aux n^r configurations possibles au total. Ceci détermine la loi $\mu_{r,n}$.

Supposons alors $r/n \to \lambda \ge 0$. Dans ce cas, le terme $(n-1)^r/n^r$ tend vers $e^{-\lambda}$:

$$\log\left(\frac{(n-1)^r}{n^r}\right) = r\log\left(1-\frac{1}{n}\right) \sim \frac{r}{n} \sim \lambda\ .$$

Par ailleurs,

$$\frac{C_r^k}{(n-1)^k} = \frac{r(r-1)\ldots(r-k+1)}{k!(n-1)^k} \to \frac{\lambda^k}{k!}\ ,$$

et cela conclut au résultat demandé.

Ce qui précède est appelé approximation binômiale-Poisson, et forme une version de ce que l'on appelle la "loi des petits nombres" ou la loi des événements rares : si on prend n expériences (n grand, disons $n \ge 30$) dont une des issues a une probabilité faible p (disons np entre 1 et 5) de se réaliser (par exemple, un accident), alors le nombre des fois où cette issue intervient suit à peu près une loi de Poisson de paramètre $\lambda = np$.

Exercice 2. À la i-ième étape du dépouillement, soit n_i le nombre de voix comptées pour A et $m_i = i - n_i$ le nombre de voix pour B, de sorte que $n_{a+b} = a$. On peut alors représenter le dépouillement du scrutin par l'ensemble

$$\{(i, n_i - m_i),\ 0 \le i \le a+b\} = \{(i, 2n_i - i),\ 0 \le i \le a+b\}\ ,$$

qui correspond à une ligne brisée issue de (0,0), finissant en $(a+b,a-b)$ de pente constante $+1$ ou -1 entre deux abscisses entières.

Le nombre total de ces chemins est de C_{a+b}^a, puisqu'il suffit pour le déterminer de choisir les a entiers i où le processus saute de $+1$ entre i et $i+1$, les b autres entiers devant correspondre à des sauts de -1. Chacun de ces chemins étant équiprobable, il suffit maintenant de dénombrer les chemins qui ne croisent jamais l'axe des abscisses, qui restent toujours strictement au-dessus.

Nécessairement, ces chemins vérifient $n_1 = 1, m_1 = 0$, et le premier point de la ligne brisée est (1,1). On utilise alors le *principe de réflexion* (à retenir !) :

> Le nombre de lignes brisées de (1,1) à $(a+b,a-b)$ qui croisent l'axe des abscisses en au moins un point est égal au nombre total de lignes brisées allant de (1,1) à $(a+b,b-a)$.

(Remarquez que l'on impose rien quant au fait que les lignes doivent être au-dessus de l'axe des abscisses, ou non, dans ce principe de réflexion.) Pour justifier ceci, introduisons l'application suivante. A une ligne brisée $(i,f(i))_{1\leq i\leq a+b}$ issue de (1,1), aboutissant en $(a+b,a-b)$, et croisant au moins une fois l'axe des abscisses, on associe la ligne définie comme suit. On note j l'indice où la ligne croise l'axe des abscisses pour la première fois ; la ligne image est obtenue en remplaçant la ligne originelle par une autre définie par $(i,f(i))$ pour $i < j$, $(j,0)$, et $(i,-f(i))$ pour $i > j$. Cette ligne image va de (1,1) à $(a+b,b-a)$. Réciproquement, à toute ligne de (1,1) à $(a+b,b-a)$ coupe nécessairement l'axe des abscisses, et en considérant le premier point où cela arrive, on retrouve un chemin de (1,1) à $(a+b,a-b)$ croisant l'axe des abscisses en au moint un point.

Par le principe de réflexion, le nombre de chemins qui ne croisent jamais l'axe des abscisses est égal à la différence entre le nombre total de chemins entre (1,1) et $(a+b,a-b)$ (ou (0,0) et $((a-1)+b,(a-1)-b)$), et celui entre (1,1) et $(a+b,b-a)$ (ou (0,0) et $((b-1)+a,(b-1)-a)$) :

$$C_{a+b-1}^{a-1} - C_{a+b-1}^{b-1} = \left(\frac{a}{a+b} - \frac{b}{a+b}\right) C_{a+b}^{a} \ .$$

En comparant au nombre total C_{a+b}^{a} de chemins, on arrive à la probabilité $(a-b)/(a+b)$.

Exercice 3. On suppose que le devin annonce au hasard et que le jeu de carte est bien battu : les annonces du devin et les cartes sont deux variables aléatoires indépendantes, de lois uniformes sur leurs images (dans les deux cas, l'ensemble des arrangements sur 52 cartes). Sans perte de généralité, on fixe l'annonce du devin, et on ne considère d'aléatoire que sur l'état du jeu de cartes. L'annonce du devin correspond à un arrangement entre les 52 cartes – et seul un préfixe de cet arrangement est annoncé.

Pour qu'un préfixe de longueur au moins k soit annoncé, il faut et il suffit que les k cartes du préfixe considéré soient rangées intialement dans le jeu dans cet ordre. Ainsi, on peut voir par dénombrement que la probabilité qu'il annonce au moins k cartes est égale à $1/k!$.

On pourrait en déduire la probabilité qu'il ait à annoncer exactement k cartes, mais il est préférable de se souvenir que le théorème de Fubini (voir TD 6, exercice 4) assure que

$$\mathbb{E}[X] = \sum_{k\geq 1} P[X \geq k] \overset{\text{ici}}{=} \sum_{1\leq k\leq 52} \frac{1}{k!} \ .$$

Ceci est proche de la valeur de $e-1$, soit environ 1,7.

Cela peut sembler peu (et c'est aussi quasi-indépendant de la taille du paquet !)... Aussi, si cette expérience vous semble un peu ridicule, sachez qu'elle fut pourtant tentée (ainsi que bien d'autres plus bizarres encore) par des gens (soi-disant) sérieux, pour tenter de prouver ou infirmer l'existence de pouvoirs paranormaux !

Exercice 4. Rappelons tout d'abord la méthode (elle est générale). Soit h borélienne positive (pour éviter les problèmes d'intégrabilité). On écrit :

$$\int h(y)\, \mathrm{d}\mathbb{P}^Y(y) = \mathbb{E}[h(Y)] = \mathbb{E}[h(1/X)] = \int h(1/x) f(x)\, dx \ ,$$

et par changement de variables, on écrit

$$\int h(1/x) f(x)\, dx = \int h(y)\, g(y)\, \mathrm{d}y \ ,$$

pour une certaine fonction g, que l'on identifie alors à la densité de $\mathbb{P}^Y$ par rapport à la mesure de Lebesgue.

$\varphi : x \mapsto 1/x$ est un C^1-difféomorphisme de $\mathbb{R} \setminus \{0\}$ dans $\mathbb{R} \setminus \{0\}$. Comme la loi de X est donnée par une densité, elle ne charge pas le point $\{0\}$ (les intégrales peuvent être prises sur $\mathbb{R}^*$), et on peut appliquer le théorème de changement de variables. On obtient facilement que Y est une bien variable aléatoire (elle est bien définie, au moins p.s.), de loi de densité

$$g(y) = \frac{1}{y^2} f\left(\frac{1}{y}\right) \ .$$

En particulier, pour f gaussienne, on obtient

$$\mathbb{E}[|Y|] = \int_{\mathbb{R}} |y|\, g(y)\, \mathrm{d}y = \frac{1}{\sqrt{2\pi}} \frac{1}{|y|} \exp\left(-\frac{1}{2y^2}\right) = +\infty \ .$$

Pour f de Cauchy, $g = f$, et Y est encore de Cauchy.

Remarque : la loi de Cauchy n'admet pas d'espérance ! Mais le 0 désigne sa médiane. A fortiori, elle n'admet pas de variance, mais -1 et 1 sont ses premier et troisième quartiles.

Exercice 5. L'ordonnée du point d'impact du rayon lumineux est $Y = \tan\theta$. Comme θ est uniforme sur $]-\pi/2, \pi/2[$, on a, par le changement de variables

$$\varphi : \eta \in]-\pi/2, \pi/2[\mapsto y = \arctan\eta \in \mathbb{R} \ ,$$

que pour toute fonction borélienne,

$$E[f(Y)] = \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} f(\tan(\eta))\mathrm{d}\eta = \frac{1}{\pi} \int_{\mathbb{R}} f(y) \frac{1}{1+y^2}\, \mathrm{d}y \ ,$$

et l'on reconnaît la densité de la loi de Cauchy.

On verra à l'exercice 8 une autre méthode pour prouver que $\tan\theta$ suit une loi de Cauchy.

Exercice 6. Dans cet exercice, on ne peut pas appliquer le théorème du changement de variables, car ni $\varphi_1 : x \mapsto |x|$, ni $\varphi_2 : x \mapsto x^2$ ne sont bijectives. On utilise alors les fonctions de répartition : vu les résultats du TD 8, une loi admet une densité si et seulement si la fonction de répartition est dérivable, de dérivée cette densité.

On voit aisément que $F_{Z_1}(t) = 0$ pour $t \leq 0$, et, pour $t \geq 0$,

$$F_{Z_1}(t) = \mathbb{P}[X \leq t,\, X \geq -t] = \mathbb{P}[X \leq t,\, X > -t] = F_X(t) - F_X(-t) \ ,$$

vu que X ne charge pas les points. Par dérivation, Z_1 suit une loi de densité

$$f_{Z_1}(x) = (f(x) + f(-x))\, \mathbb{I}_{\{x \geq 0\}} \ .$$

De même, $F_{Z_2}(t) = 0$ pour $t \leq 0$, et, pour $t \geq 0$,

$$F_{Z_2}(t) = \mathbb{P}[X^2 \leq t] = \mathbb{P}[X \leq \sqrt{t},\, X > -\sqrt{t}] = F_X(\sqrt{t}) - F_X(-\sqrt{t}) \ .$$

Par dérivation, Z_2 suit une loi de densité

$$f_{Z_2}(x) = \frac{1}{2\sqrt{x}}\left(f(x) + f(-x)\right)\mathbb{I}_{\{x\geq 0\}}\ .$$

La densité de la loi du $\chi^2(1)$ est ainsi

$$f_{\chi^2(1)}(x) = \frac{1}{\sqrt{2\,\pi\,x}}e^{-x/2}\mathbb{I}_{\{x\geq 0\}}\ .$$

Exercice 7. Pour $x \geq 0$,

$$\begin{aligned} 1 - F_{X^+}(x) \ &= \ \mathbb{P}\left\{X^+ > x\right\} \ &= \ \mathbb{P}\left\{X > x\right\} \ = \ 1 - F_X(x)\ , \\ &\ \mathbb{P}\left\{X^- \geq x\right\} \ &= \ \mathbb{P}\left\{X \leq -x\right\} \ = \ F_X(-x)\ ; \end{aligned}$$

par ailleurs, le théorème de Fubini (voir TD 6, exercices 4 et 6) assure que

$$\begin{aligned} \mathbb{E}[X^+] \ &= \ \int_{\mathbb{R}^+} \mathbb{P}\left\{X^+ > x\right\} = \int_{\mathbb{R}^+} 1 - F_X(x)\ , \\ \mathbb{E}[X^-] \ &= \ \int_{\mathbb{R}^+} \mathbb{P}\left\{X^- \geq x\right\} = \int_{\mathbb{R}^+} F_X(-x)\ . \end{aligned}$$

Exercice 8. On détermine les lois via les fonctions de répartition. En effet, on rappelle que la fonction de répartition caractérise la loi : sa donnée indique la probabilité que porte la loi sur les ensembles de la forme $(-\infty,t]$ qui sont stables par intersection finie et engendrent la tribu des boréliens ; on conlut par le théorème d'unicité des mesures.

Pour $t \in \mathbb{R}$ et $u \in]0,1[$, on note tout d'abord l'équivalence

$$F^{-1}(u) \leq t \quad \Longleftrightarrow \quad F(t) \geq u\ .$$

Le sens réciproque est immédiat, par définition de $F^{-1}(u)$ comme infimum d'un ensemble auquel t appartient. Pour établir le sens direct, eu égard à la croissance de F, il suffit de montrer que $F(F^{-1}(u)) \geq u$. A cet effet, on prend une suite de points $(t_n)_n$, tendant vers $F^{-1}(u)$ et tels que $F(t_n) \geq u$ pour tout n. En utilisant la continuité à droite de F, on obtient l'inégalité recherchée. (Remarquons également que si F est continue sur $\mathbb{R}$, alors plus précisément, $F(F^{-1}(u)) = u$ pour tout u.) Ainsi,

$$P[F^{-1}(U) \leq t] = P[U \leq F(t)] = F(t),$$

et, par égalité de leur fonction de répartition, $X \overset{(d)}{=} F^{-1}(U)$.

Application. Pour X exponentielle de paramètre λ, $F : t \mapsto 1 - e^{-\lambda\,t}$, et $-(1/\lambda)\log(1 - U) \overset{(d)}{=} -(1/\lambda)\log U$ suivent également une loi exponentielle de paramètre λ. La fonction de répartition de la loi de Cauchy est $F : t \mapsto \pi/2 + \arctan t$, et $\tan(U - \pi/2)$ suit ainsi une telle loi de Cauchy. Enfin, des calculs aisés montrent que

$$\mathbb{I}_{\{U\leq 1/2\}}\log(2U) - \mathbb{I}_{\{U\geq 1/2\}}\log(2(1-U))$$

suit une loi de Laplace.

Pour X de loi sans atome, F est une fonction continue, et en particulier elle prend toutes les valeurs entre 0 et 1. On note de même que cette fois-ci, on a l'équivalence, pour tout $x \in \mathbb{R}$ et $u \in]0,1[$,

$$x \leq G(u) \quad \Longleftrightarrow \quad F(x) \leq u\ .$$

De ce fait, pour $u \in]0,1[$,

$$P[F(X) \leq u] = P[X \leq G(u)] = F(G(u)) = u,$$

la dernière égalité étant une conséquence de la continuité de F (on l'a pour G de même qu'on l'avait pour F^{-1}). Finalement la fonction de répartition de $F(X)$ est celle d'une uniforme sur $[0,1]$, d'où $F(X) \overset{(d)}{=} U$.

Exercice 9. Pour $X(\omega) = \omega^2$: pour tout borélien A,

$$X^{-1}(A) = \{\omega \in \mathbb{R} : \omega^2 \in A\} = \sqrt{A \cap \mathbb{R}_+} \cup -\sqrt{A \cap \mathbb{R}_+} \ .$$

De ce fait, la tribu $\sigma(X)$ est composée de boréliens B symétriques (au sens où $B = C \cup -C$ pour un certain borélien C inclus dans $\mathbb{R}_+$). Inversement, la tribu des boréliens symétriques est engendrée par les élements de la forme $[a,b] \cup [-b, -a]$, $0 \leq a \leq b$. Comme ces ensembles sont dans $\sigma(X)$, cette tribu est exactement égale à celle des boréliens symétriques.

Exercice 10. On a, en notant $\mathbb{P}^{|X|}$ la loi de $|X|$ sous $\mathbb{P}$,

$$x\mathbb{P}[|X| \geq x] = \int_{\mathbb{R}_+} x\, \mathbb{I}_{y \geq x} \mathrm{d}\mathbb{P}^{|X|}(y) \leq \int_{\mathbb{R}_+} y \mathbb{I}_{y \geq x} \mathrm{d}\mathbb{P}^{|X|}(y) = \mathbb{E}\left[|X| \mathbb{I}_{\{|X| \geq x\}}\right] \ .$$

Le dernier terme tend vers 0 quand $x \to \infty$, par convergence dominée.

Le changement de variables linéaire $u = v + x$ assure que

$$\Phi(x) = \int_x^{+\infty} \frac{1}{\sqrt{2\pi}}\, e^{-u^2/2}\, \mathrm{d}u = e^{-x^2/2} \int_0^{+\infty} \frac{1}{\sqrt{2\pi}}\, e^{-v\,x} e^{-v^2/2}\, \mathrm{d}v \ ,$$

et en utilisant $e^{-v\,x} \leq 1$, on obtient d'une part la majoration par $e^{-x^2/2}/2$; d'autre part, avec $e^{-v^2/2} \leq 1$, il vient

$$\Phi(x) \leq e^{-x^2/2} \int_0^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-v\,x}\, \mathrm{d}v = \frac{e^{-x^2/2}}{x\sqrt{2\pi}} \ .$$

Pour la minoration, on remarque que si $0 < x \leq u$, alors $x^2/u^2 \leq 1$, et donc (la deuxième égalité provenant d'une intégration par parties)

$$\begin{aligned} \Phi(x) \geq \int_x^{+\infty} \frac{x^2}{u^2} \frac{e^{-u^2/2}}{\sqrt{2\pi}}\, \mathrm{d}u &= \frac{x^2}{\sqrt{2\pi}} \int_x^{+\infty} \left(-\frac{1}{u}\, e^{-u^2/2}\right)' \mathrm{d}u \\ &= \frac{x^2}{\sqrt{2\pi}} \left(\frac{1}{x}\, e^{-x^2/2} - \int_x^{+\infty} e^{-u^2/2}\, \mathrm{d}u\right) \\ &= \frac{x}{\sqrt{2\pi}} e^{-x^2/2} - x^2 \Phi(x) \ , \end{aligned}$$

d'où le résultat.

Intégration et Probabilités – TD 10
Corrigé partiel

Exercice 1. Pour les fonctions génératrices, on rappelle que $G'(1)$ (dérivée à gauche) donne l'espérance, et $G''(1) = \mathbb{E}[X(X-1)]$ (à gauche), de sorte que $\operatorname{var} X = G''(1) + G'(1)\,(1-G'(1))$.

(1) $G(r) = p\,r + 1 - p$, on retrouve l'espérance p, et la variance $p\,(1-p)$.

(2) Par calcul direct, en utilisant l'identité binômiale,

$$G(r) = \sum_{k=0}^{n} C_n^k p^k (1-p)^{n-k} r^k = (p\,r + 1 - p)^n \ ,$$

ce qu'on retrouve en utilisant la définition de la loi binômiale comme somme de variables aléatoires de Bernoulli i.i.d. [indépendantes et identiquement distribuées]. Espérance np et variance $n(n-1)p^2 + np(1-np) = np(1-p)$.

(3) On rappelle que pour $k \geq 0$, $\mu\{k\} = (1-q)\,q^k$ (avec pour interprétation, le nombre de faces avant le premier pile dans une suite indépendante de Bernoulli de paramètre q). On a

$$G(r) = (1-q) \sum_{k\geq 0} (qr)^k = \frac{1-q}{1-qr} \ .$$

Espérance $q/(1-q)$ et variance $q/(1-q)^2$.

(4) On calcule

$$G(r) = e^{-\lambda} \sum_{k\geq 0} \frac{(\lambda r)^k}{k!} = e^{-\lambda(1-r)} \ .$$

Espérance λ et variance λ.

Pour les fonctions caractéristiques, on rappelle, par dérivation sous le signe $\mathbb{E}$, que la dérivée $\Phi'(0)$ en 0 donne $\mathbb{E}[X] = i\Phi'(0)$, et que la dérivée seconde $\Phi''(0)$ en 0 donne $\mathbb{E}[X^2] = -\Phi''(0)$.

(1) Pour la loi exponentielle, on calcule pour $t \in \mathbb{R}$,

$$\Phi(t) = \int_{\mathbb{R}_+} \lambda e^{(-\lambda + it)x}\, \mathrm{d}x = \frac{\lambda}{\lambda - it} \ .$$

En utilisant le rappel sur les dérivées, on en déduit que l'espérance est $1/\lambda$ et la variance $1/\lambda^2$. La loi exponentielle est la loi de durée de vie d'un appareil sans usure ni rodage : sachant que $X > t$, la loi de $X - t$ a même loi que X, ce qu'on note $P[X > t + s \,|\, X > t] = P[X > s]$ pour tous s, t (voir le cours sur l'espérance conditionnelle ou se remémorer ses cours de terminale, l'événement $X > t$ étant de probabilité strictement positive). On appelle ceci la “perte de mémoire” de la loi exponentielle.

(2) Pour la loi uniforme,

$$\Phi(t) = \int_0^1 e^{itx}\, \mathrm{d}x = \frac{e^{it} - 1}{it} = \sum_{k\geq 1} \frac{(it)^{k-1}}{k!}.$$

Par identification (voir Théorème 8.2.5 du polycopié), on en déduit que l'espérance est $1/2$, le moment d'ordre 2 est $1/3$, de sorte que la variance est $1/12$. Cette dernière valeur est à retenir.

Exercice 2. Notons F la fonction de répartition de X. Une médiane m de X est telle que $F(m-) \le 1/2 \le F(m)$. L'intervalle cherché de toutes les médianes est l'adhérence l'image réciproque par F de $F(x')$, où x' est le minimum de l'ensemble $F(\mathbb{R}) \cap [1/2,1]$. C'est bien un ensemble non-vide. On peut même l'identifier, avec les notations de l'exercice 8 du TD 9, comme étant égal à $[F^{-1}(1/2), G(1/2)]$. F est plate sur $\operatorname{med}(X)$ (éventuellement privé de ses bornes), et ne charge donc pas son intérieur. La dernière propriété découle de l'implication

$$\mathbb{P}\left[|X| \le a\right] > \frac{1}{2} \quad \Rightarrow \quad \forall c > a,\ \mathbb{P}\left[|X| \ge c\right] < \frac{1}{2}\ .$$

Pour la première inégalité de (2), on écrit, d'abord par la définition d'une médiane, puis par indépendance,

$$\begin{aligned}
\frac{1}{2}\,\mathbb{P}\left[|X-m| \ge t\right] &\le \mathbb{P}\left[\widetilde{X} \le m\right]\mathbb{P}\left[X-m \ge t\right] + \mathbb{P}\left[\widetilde{X} \ge m\right]\mathbb{P}\left[X-m \le -t\right] \\
&= \mathbb{P}\left[\widetilde{X} \le m \text{ et } X-m \ge t\right] + \mathbb{P}\left[\widetilde{X} \ge m \text{ et } X-m \le -t\right] \\
&\le \mathbb{P}\left[X \ge \widetilde{X}+t\right] + \mathbb{P}\left[X \le \widetilde{X}-t\right] = \mathbb{P}\left[\left|X-\widetilde{X}\right| \ge t\right]\ .
\end{aligned}$$

Pour la deuxième inégalité, on note que

$$\left\{\left|X-\widetilde{X}\right| \ge t\right\} = \left\{\left|(X-a)-\left(\widetilde{X}-a\right)\right| \ge t\right\} \subseteq \left\{|X-a| \ge \frac{t}{2}\right\} \cup \left\{\left|\widetilde{X}-a\right| \ge \frac{t}{2}\right\}\ ,$$

et on conclut en utilisant le fait que X et $\widetilde{X}$ ont même loi.

On utilise ensuite l'exercice 4 du TD 6, avec $g(t)=t^r$ (vu $r>0$) : il suffit de multiplier les deux membres de la première des deux inégalités précédentes par $r\,t^{r-1}$, et d'intégrer sur $\mathbb{R}_+$. Enfin, pour la dernière inégalité, pour $a,b \ge 0$, vu que $(a+b)^r \le a^r + b^r$ lorsque $r \le 1$, et, par convexité, $(a+b)^r \le 2^{r-1}(a^r+b^r)$ lorsque $r \ge 1$, on a directement

$$\left|X-\widetilde{X}\right|^r \le \left(|X-a| + \left|\widetilde{X}-a\right|\right)^r \le 2^{(r\vee 1)-1}\left(|X-a|^r + \left|\widetilde{X}-a\right|^r\right)\ ,$$

et le résultat s'en suit par intégration, X et $\widetilde{X}$ ayant même loi.

La convexité de D_X est immédiate ; la fonction D_X admet une seule valeur minimale, dont on va montrer qu'elle est prise sur $\operatorname{med}(X)$, par un calcul de dérivées à droite et à gauche. La valeur exacte des ces dernières nous donnera au passage le fait que D_X caractérise la loi de X. Comme

$$D_X(x) = \int_{\{X>x\}} (X-x)\,\mathrm{d}\mathbb{P} + \int_{\{X<x\}} (x-X)\,\mathrm{d}\mathbb{P} = \mathbb{E}[X] - x + 2\int_{\{X\le x\}} (x-X)\,\mathrm{d}\mathbb{P}\ ,$$

et par théorème de Fubini,

$$\int_{\{X\le x\}} (x-X)\,\mathrm{d}\mathbb{P} = \mathbb{E}\left[(x-X)\mathbb{I}_{[X\le x]}\right] = \mathbb{E}\left[\int_{-\infty}^{x} \mathbb{I}_{[X\le y]}\mathrm{d}y\right] = \int_{-\infty}^{x} \mathbb{P}[X \le y]\,\mathrm{d}y\ ,$$

de sorte que

$$D_X(x) = \mathbb{E}[X] - x + 2\int_{-\infty}^{x} \mathbb{P}[X\le y]\,\mathrm{d}y\ .$$

Ainsi, en utilisant la continuité à droite de la fonction de répartition, il vient, à droite,

$$D_X'(x+) = -1 + 2\mathbb{P}[X \le x]\ .$$

A gauche, on étudie la quantité

$$\frac{\int_{x'}^{x} \mathbb{P}[X \le y]\,\mathrm{d}y}{x - x'} \le \mathbb{P}[X < x]$$

lorsque $x' \to x$, $x' < x$. Par continuité à gauche de $y \mapsto \mathbb{P}[X < y]$, la limite inférieure du membre de gauche est plus grande que $\mathbb{P}[X < x]$; d'où l'on déduit

$$D'_X(x-) = -1 + 2\mathbb{P}[X < x] \ .$$

D'où l'on déduit que D_X est minimale sur $\mathrm{med}(X)$: si $\mathbb{P}[X \le x] < 1/2$, $D'_X(x+) < 0$ et D_X est décroissante au voisinage de x. Si $\mathbb{P}[X \ge x] > 1/2$, $D'_X(x-) > 0$ et D_X est croissante au voisinage de x. Les comportements asymptotiques s'obtiennent en notant que

$$D_X(x) + x = \mathbb{E}[X] + 2\int_{-\infty}^{x} \mathbb{P}[X \le y]\,\mathrm{d}y \xrightarrow[x \to -\infty]{} \mathbb{E}[X] \ ,$$

et le résultat correspondant en $+\infty$ en découle, puisque $D_X(-x) = D_{-X}(x)$.

Enfin, l'inégalité reliant médiane, espérance, et variance se prouve en notant que l'inégalité de Tchebychev assure que

$$\mathbb{P}\left[|X - \mathbb{E}[X]| \ge c\sqrt{2\,\mathrm{var}(X)}\right] \le \frac{1}{2c^2} \ ,$$

d'où, pour tout $c > 1$,

$$\mathbb{P}\left[|X - \mathbb{E}[X]| \ge c\sqrt{2\,\mathrm{var}(X)}\right] > \frac{1}{2} \ ,$$

et l'implication de (1), appliquée à la variable aléatoire $X - \mathbb{E}[X]$ conclut à

$$|\mathrm{med}(X) - \mathbb{E}[X]| \le c\sqrt{2\ \mathrm{var}(X)} \ ,$$

d'où l'assertion en faisant $c \to 1$.

Exercice 3. Il est clair qu'on a également $0 < \mathbb{E}[X_n] < +\infty$; on peut alors introduire

$$Y_n = \frac{X_n}{\mathbb{E}[X_n]} : \qquad \mathbb{E}[Y_n] = 1 \ , \quad \mathbb{E}[Y_n^2] \in [1, +\infty[\ ,$$

par l'inégalité de Cauchy-Schwarz. Soit $c \in]0,1[$: comme $\mathbb{E}[Y_n; Y_n \le c] \le c$, il vient $\mathbb{E}[Y_n; Y_n > c] \ge 1 - c$, et par l'inégalité de Cauchy-Schwarz, $\mathbb{E}[Y_n^2]\,\mathbb{P}[Y_n > c] \ge (1-c)^2$, et finalement,

$$\mathbb{P}[Y_n > c] \ge \frac{(1-c)^2}{\mathbb{E}[Y_n^2]} = (1-c)^2 \frac{(\mathbb{E}[X_n])^2}{\mathbb{E}[X_n^2]} \ .$$

On applique maintenant le lemme de Fatou pour obtenir

$$\mathbb{P}[\limsup_n \{Y_n > c\}] \ge \limsup_n \mathbb{P}[Y_n > c] \ge (1-c)^2 \frac{(\mathbb{E}[X_n])^2}{E[X_n^2]} \ .$$

Or

$$\mathbb{P}[\limsup_n Y_n \ge c] \ge \mathbb{P}[\limsup_n \{Y_n > c\}]$$

(cela est certes vrai pour $c = 0$, mais du coup on ne conclut plus à rien !). Le résultat demandé s'en déduit en faisant $c \to 0$, $c > 0$.

Pour l'application, posons

$$X_n = \sum_{1 \le k \le n} \mathbb{I}_{A_k} \quad \text{et} \quad X_\infty = \sum_{k \ge 1} \mathbb{I}_{A_k} \ ;$$

$(X_n)_{n\geq 1}$ étant croissante, et ayant $\mathbb{E}[X_n] \to +\infty$, il vient que

$$\left\{\limsup_{n\to\infty} \frac{X_n}{\mathbb{E}[X_n]} > 0\right\} \subseteq \{X_n \to \infty\} = \{X_\infty = \infty\} = \limsup_n A_n \ .$$

Pour la question subsidiaire : le point précédent donne, vu l'indépendance *et* la condition $\sum_n \mathbb{P}[A_n] = \infty$ (le membre de droite ne vaut pas identiquement 1, mais converge vers 1),

$$\mathbb{P}\left[\limsup_{n\to\infty} \frac{X_n}{\mathbb{E}[X_n]} > 0\right] = 1 \ .$$

On a donc une alternative (et une équivalence), dans le cas où les A_k sont indépendants :

- ou $\sum_k \mathbb{P}[A_k] < \infty$ et $\mathbb{P}[\limsup_n A_n] = 0$;
- ou $\sum_k \mathbb{P}[A_k] = \infty$ et $\mathbb{P}[\limsup_n A_n] = 1$.

Cela complète bien l'implication connue du lemme de Borel-Cantelli en une équivalence. C'est la loi du "tout ou rien".

Exercice 4. Dans tout l'exercice, vu que Y est à valeurs dans un compact et qu'on considère une mesure de probabilité, il n'y aura pas de problème pour dériver sous l'espérance, etc. Un calcul facile montre que

$$\begin{aligned} \psi_Y''(\lambda) &= \frac{\mathbb{E}[Y^2 e^{\lambda Y}]\mathbb{E}[e^{\lambda Y}] - \left(\mathbb{E}[Y e^{\lambda Y}]\right)^2}{\left(\mathbb{E}[e^{\lambda Y}]\right)^2} \\ &= \mathbb{E}\left[Y^2 \ e^{\lambda Y} e^{-\psi_Y(\lambda)}\right] - \left(\mathbb{E}\left[Y \ e^{\lambda Y} e^{-\psi_Y(\lambda)}\right]\right)^2 = \operatorname{var}_{\mathbb{Q}_\lambda} Y \ , \end{aligned}$$

où $\mathbb{Q}_\lambda$ est la probabilité absolument continue par rapport à $\mathbb{P}$, de densité $\omega \mapsto e^{-\psi_Y(\lambda)} e^{\lambda Y(\omega)}$ par rapport à $\mathbb{P}$. En particulier, Y étant à valeurs dans $[a,b]$, $|Y - (b+a)/2| \leq (b-a)/2$, d'où $\operatorname{var} Y \leq \mathbb{E}_{\mathbb{Q}_\lambda}[(Y-(b+a)/2)^2] \leq (b-a)^2/4$. Le résultat demandé s'en suit par intégration de $\psi_Y''(\lambda) \leq (b-a)^2/4$, en notant que $\psi_Y(0) = 0$, et aussi $\psi_Y'(0) = \mathbb{E}[Y] = 0$, car Y est centrée.

On note $X_i' = X_i - \mathbb{E}[X_i]$: soit $S = X_1' + \ldots + X_n'$; en utilisant l'indépendance des X_i', il vient $\psi_S \leq \psi_{X_1'} + \ldots + \psi_{X_n'}$, et par ce qui précède, pour tout $\lambda > 0$,

$$\psi_S(\lambda) \leq \frac{\lambda^2}{8} \sum_{i=1}^n (b_i - a_i)^2 \ .$$

On a alors, pour tout $\lambda > 0$,

$$\mathbb{P}[S \geq \varepsilon] = \mathbb{P}\left[e^{\lambda S} \geq e^{\lambda\varepsilon}\right] \leq e^{-\lambda\varepsilon}\mathbb{E}\left[e^{\lambda S}\right] = e^{-\lambda\varepsilon + \psi_S(\lambda)} \leq \exp\left(-\lambda\varepsilon + \frac{\lambda^2}{8}\sum_{i=1}^n (b_i - a_i)^2\right) \ .$$

Le résultat en découle, par optimisation sur $\lambda > 0$. Par symétrie, en considérant $-X_i$ au lieu de X_i, on en déduit la version en valeur absolue, avec les facteur 2 dans le membre de droite.

Intégration et Probabilités – TD 11
Corrigé partiel

Exercice 1. On note $\mathcal{G}_{n+1}$ la tribu engendrée par $\{\mathcal{F}_k,\ k \geq n+1\}$, et $\mathcal{H}_n$ celle engendrée par $\{\mathcal{F}_0, \ldots, \mathcal{F}_n\}$. Par la propriété[1] du regroupement par paquets, $\mathcal{H}_n$ est indépendante de $\mathcal{G}_{n+1}$, donc de $\mathcal{G}_\infty \subset \mathcal{G}_{n+1}$. Par définition de l'indépendance de deux tribus, on a alors

$$\forall\, G \in \bigcup_{n\geq 0} \mathcal{H}_n,\ \ \forall\, C \in \mathcal{G}_\infty, \quad \mathbb{P}[G \cap C] = \mathbb{P}[G]\,\mathbb{P}[C]\ .$$

Ainsi, par stabilité[2] par intersections finies, $\mathcal{G}_\infty \subset \sigma\{\mathcal{F}_k,\ k \geq 0\}$ est indépendante de $\mathcal{G}_\infty$. En particulier, l'indépendance de $\mathcal{G}_\infty$ par rapport à elle-même signifie que

$$\forall\, G \in \mathcal{G}_\infty, \quad \mathbb{P}[G] = \mathbb{P}[G \cap G] = (\mathbb{P}[G])^2\ ,$$

et chaque $\mathbb{P}[G]$ vaut donc 0 ou 1.

Soit H une variable aléatoire positive, $\mathcal{G}_\infty$-mesurable. On considère son développement dyadique

$$H = \sum_{m\in\mathbb{Z}} 2^{-m}\,\delta_m(H)\ ,$$

où pour tout m, $\delta_m(H)$ est $\mathcal{G}_\infty$-mesurable à valeurs dans $\{0,1\}$, vu une expression possible pour $\delta_m(H)$,

$$\delta_m(H) = \mathbb{I}_{\{\lfloor 2^m H\rfloor \text{ est impair}\}}\ .$$

On a donc

$$\forall\, m \in \mathbb{Z}, \quad \mathbb{P}[\delta_m(H) = 1] = 0 \text{ ou } 1,$$

et H est $\mathbb{P}$-p.s. constante, égale à

$$H = \sum_{m\in\mathbb{Z},\ \mathbb{P}[\delta_m(H)=1]=1} 2^{-m}\ .$$

On note alors qu'on peut écrire la moyenne mobile au temps n comme

$$\bar{X}_n = \frac{1}{n}\left(\sum_{1\leq j\leq k} X_j\right) + \frac{1}{n}\left(\sum_{k+1\leq j\leq n} X_j\right)\ ,$$

et que par conséquent, $\limsup_n \bar{X}_n$ et $\liminf_n \bar{X}_n$ sont $\mathcal{G}_{k+1}$-mesurables pour tout k, et partant, sont asymptotiques. Elles sont donc p.s. constantes. En particulier, l'ensemble de convergence des moyennes mobiles, égal à $\{\limsup_n \bar{X}_n = \liminf_n \bar{X}_n\}$ est asymptotique, de probabilité 1 ou 0. Nous verrons que dans le cas où les X_j sont i.i.d. réelles, il y a convergence dans $\mathbb{R}$ de la suite des moyennes mobiles si et seulement si les X_j sont intégrables (et dans ce cas, la limite est p.s. égale à $\mathbb{E}[X_j]$).

Exercice 2. On prouve d'abord l'indication. Par l'inégalité de Hölder. Ou une méthode complètement piétonne : pour tout $a > 0$, l'application

$$x \in \mathbb{R}_+^* \mapsto 1/x + ax^2$$

1. voir la Proposition 9.2.5. du polycopié
2. voir la Proposition 9.2.4. du polycopié

est minimale sur $]0, +\infty[$ lorsque $x = 1/(2a)$. Ainsi,

$$\frac{x + ax^4}{x^2} \geq (2a)^{1/3} + \frac{a}{(2a)^{2/3}} = \frac{3}{2}(2a)^{1/3} \ .$$

En remplacant x par $|X|$ et en prenant l'espérance,

$$\mathbb{E}[|X|] \geq \frac{3}{2}(2a)^{1/3}\, \mathbb{E}[X^2] - a\mathbb{E}[X^4] \ .$$

La borne inférieure, considérée comme fonction de a, est maximale lorsque

$$a = \frac{1}{2}\left(\frac{\mathbb{E}[X^2]}{\mathbb{E}[X^4]}\right)^{3/2} ,$$

et sa valeur en ce point est exactement la borne inférieure annoncée.

Il suffit maintenant d'appliquer l'indication à $X = a_1\sigma_1 + \ldots + a_n\sigma_n$, en notant que par indépendance et par centrage des σ_j, $\mathbb{E}[\sigma_i\sigma_j\sigma_k] = 0$ dès qu'au moins deux indices sont différents (et sinon, l'espérance vaut 1). Et de même pour les doubles ou les quadruples produits. Ainsi,

$$\mathbb{E}\left[\left|\sum_{i=1}^{n} a_i\, \sigma_i\right|\right] \geq \frac{\left(\sum_{i=1}^{n} a_i^2\right)^{3/2}}{\sqrt{\sum_{i=1}^{n} a_i^4 + 3\sum_{i\neq j} a_i^2 a_j^2}} \geq \frac{1}{3}\sqrt{\sum_{i=1}^{n} a_i^2} \ ,$$

en utilisant en outre que

$$\sum_{i=1}^{n} a_i^4 + 3\sum_{i\neq j} a_i^2 a_j^2 \leq 3\left(\sum_{i=1}^{n} a_i^2\right)^2 \ .$$

Exercice 3. On détermine les fonctions de répartition de Y et Z, et par dérivation, on note que les lois de Y et Z sont absolument continues par rapport à la mesure de Lebesgue, de densité g et h déterminées par la dérivée. Plus précisément, par indépendance, pour $t \in \mathbb{R}$,

$$\begin{aligned} F_Z(t) &= \mathbb{P}\left[\forall i = 1, \ldots, N,\ \ X_i \leq t\right] = \left(F_X(t)\right)^N \ , \\ F_Y(t) &= 1 - \mathbb{P}\left[\forall i = 1, \ldots, N,\ \ X_i > t\right] = 1 - \left(1 - F_X(t)\right)^N \ . \end{aligned}$$

Par dérivation, en utlisant validement le théorème fondamental du calcul, F_X étant absolument continue, et, partant, F_Z et F_Y aussi,

$$\begin{aligned} g(y) &= Nf(y)\left(1 - F_X(y)\right)^{N-1} \ , \\ h(z) &= Nf(z)\left(F_X(z)\right)^{N-1} \ . \end{aligned}$$

Exercice 4. Le plus simple est d'utiliser les fonctions de répartition. Ainsi, comme à l'exercice précédent, et par indépendance,

$$\mathbb{P}[U > t] = \mathbb{P}[X > t, Y > t] = \mathbb{P}[X > t]\,\mathbb{P}[Y > t] = e^{-(\lambda+\mu)t} \ ,$$

où l'on a utilisé le fait que la *fonction de survie* d'une loi exponentielle de paramètre λ est

$$t \in \mathbb{R}_+ \mapsto \mathbb{P}\left[\mathcal{E}(\lambda) > t\right] = e^{-\lambda t} \ .$$

La fonction de survie de U est celle d'une loi exponentielle de paramètre $\lambda+\mu$, or, les fonctions de survie caractérisent la loi (de même que les fonctions de répartition).

De même, $\mathbb{P}[V \leq t] = \mathbb{P}[X \leq t]\,\mathbb{P}[Y \leq t] = (1-e^{-\lambda t})(1-e^{-\mu t})$, et par dérivation, la densité de la loi de V par rapport à la mesure de Lebesgue est $x \in \mathbb{R}_+ \mapsto \lambda e^{-\lambda x} + \mu e^{-\mu x} - (\lambda+\mu)e^{-(\lambda+\mu)x}$.

Pour calculer la loi de (U,V), on considère h positive mesurable, et par théorème de Fubini-Tonelli, on note que

$$\begin{aligned}
\mathbb{E}[h(U,V)] &= \int_{R_+^2} h\,(x\wedge y, x\vee y)\;\lambda\,\mu\,e^{-(\lambda x+\mu y)}\,\mathrm{d}x\mathrm{d}y \\
&= \int_{0\leq x\leq y} \mathrm{d}x\mathrm{d}y\,\lambda\mu e^{-\lambda x-\mu y}h(x,y) + \int_{0\leq y<x} \mathrm{d}x\mathrm{d}y\lambda\mu e^{-\lambda x-\mu y}h(y,x) \\
&= \int_{\mathbb{R}_+^2} h(u,v)\,\lambda\mu(e^{-\lambda u-\mu v}+e^{-\mu u-\lambda v})\mathbb{I}_{\{0\leq u\leq v\}}\,\mathrm{d}u\mathrm{d}v\ .
\end{aligned}$$

(Dans la définition des ensembles d'intégration, on peut mettre des inégalité larges ou strictes puisque la loi de (X,Y) a une densité pa rapport à la mesure de Lebesgue, et ne charge pas les droites.) On vient de prouver que la loi jointe de (U,V) est la mesure de densité $(u,v) \mapsto \lambda\mu(e^{-\lambda u-\mu v}+e^{-\mu u-\lambda v})\mathbb{I}_{\{0\leq u\leq v\}}\mathrm{d}u\mathrm{d}v$ par rapport à la mesure de Lebesgue. On peut retrouver les densités des lois marginales (i.e. les lois de U et V) en prenant respectivement $f(u,v) = f(u)$ puis $f(u,v) = f(v)$ dans l'intégrale ci-dessus, c.à.d. qu'on intègre la partie en v (resp. u) de la densité de la loi jointe.

Exercice 5. J'attends des solutions élégantes avec impatience !

Exercice 6. $Z = X/Y$ est une variable aléatoire car $(x,y) \in \mathbb{R}\times\mathbb{R}^* \mapsto x/y$ est mesurable, et $Y \neq 0$ p.s. (parce que sa loi admet une densité par rapport à la mesure de Lebesgue, donnée par l'intégrale de f en sa première variable). Déterminons la loi du couple (Z,Y) : bien sûr, on connaît la loi de Y, mais pour appliquer proprement le théorème de changement de variables, il nous faut considérer un couple de variables aléatoires. Par le changement de variables

$$\varphi : (z,y) \in \mathbb{R}\times\mathbb{R}^* \mapsto (zy,y)\ ,$$

de classe C^1, bijectif, de jacobien de déterminant ne s'annulant pas sur le domaine de définition,

$$J_\varphi = \begin{bmatrix} y & z \\ 0 & 1 \end{bmatrix}\ , \qquad |\det J_\varphi| = |y|\ ,$$

on a que φ est un C^1-difféomorphisme, de sorte que le théorème de changement de variables donne que la loi de (Z,Y) admet pour densité par rapport à la mesure de Lebesgue $(z,y) \mapsto f(zy,y)\,|y|$. (Et la densité de la loi de Z, dite “première marginale”, est donnée par l'intégrale en y de la densité de la loi jointe.)

Lorsque X et Y sont normales standard indépendantes, la densité de la loi de Z est donnée par

$$g : z \in \mathbb{R} \mapsto \frac{1}{2\pi}\int_{\mathbb{R}} \exp\left(-\frac{y^2}{2}(1+z^2)\right)|y|\ \mathrm{d}y\ ,$$

et

$$g(z) = \frac{1}{\pi}\int_0^\infty \exp\left(-\frac{y^2}{2}(1+z^2)\right)y\,\mathrm{d}y = \frac{1}{\pi}\frac{1}{1+z^2}\int_0^\infty e^{-t}\mathrm{d}t = \frac{1}{\pi}\frac{1}{1+z^2}\ ,$$

où la deuxième inégalité procède du changement de variables $y^2(1+z^2)/2 = t$. On reconnaît donc que Z suit une loi de Cauchy standard.

Comme X/Y et Y/X ont même loi, qui est de Cauchy, on en déduit que l'inverse d'une variable aléatoire distribuée selon une loi de Cauchy suit encore une telle loi. Quant à $(1+Z)/(1-Z)$, pour voir que c'est encore une variable aléatoire distribuée selon une loi de Cauchy, on l'écrit sous la forme

$$\frac{1+Z}{1-Z} \overset{(d)}{=} \frac{1+Y/X}{1-Y/X} = \frac{(X+Y)/\sqrt{2}}{(X-Y)/\sqrt{2}} ,$$

où l'on remarque que $((X+Y)/\sqrt{2}, (X-Y)/\sqrt{2})$ est encore un couple de variables aléatoires i.i.d. selon une loi normale standard. Pour le voir (puisque vous n'avez pas encore de théorie générale sur ce qu'on appelle les *vecteurs gaussiens*), calculons la fonction caractéristique de la loi du couple. Par indépendance,

$$\begin{aligned}\mathbb{E}\left[e^{i\left(t_1\frac{X+Y}{\sqrt{2}}+t_2\frac{X-Y}{\sqrt{2}}\right)}\right] &= \mathbb{E}\left[e^{iX(t_1+t_2)/\sqrt{2}}\right]\mathbb{E}\left[e^{iY(t_1-t_2)/\sqrt{2}}\right] \\ &= e^{-(t_1+t_2)^2/4-(t_1-t_2)^2/4} = e^{-t_1^2/2}e^{-t_2^2/2} ,\end{aligned}$$

ce qui prouve l'assertion.

Exercice 7. Les variables aléatoires X et Y sont définies $(E,\mathcal{E},\mathbb{P}) \to (F,\mathcal{F},\mu\otimes\mu)$ (où les deux espaces considérés sont probabilisés). Elles sont indépendantes, chacune de loi μ, de sorte que la loi du couple (X,Y) est la loi produit $\mu\otimes\mu$. De ce fait, le théorème de Fubini-Tonelli donne

$$\mathbb{P}[X=Y] = \int_F \mu(\{x\})\,\mathrm{d}\mu(x) = \sum_{x\in F}\mu(\{x\})^2 ,$$

la deuxième égalité découlant de ce que la fonction $x \mapsto \mu(\{x\})$ (qui est bien mesurable par le théorème de Fubini) est non nulle en x si et seulement si μ a un atome en x. Remarquons également que de tels points sont forcément en quantité au plus dénombrable, et la somme intervenant ci-dessus est au plus dénombrable.

Pour (1) : en supposant X et Y réelles, on a une relation d'ordre... $\mathbb{P}[X>Y] = \mathbb{P}[Y>X]$ parce que les deux couples (X,Y) et (Y,X) ont même loi. Donc $\mathbb{P}[X<Y] = \mathbb{P}[X\neq Y]/2$, et

$$\mathbb{P}[X>Y] = \frac{1-\sum_{x\in\mathbb{R}}\mu(\{x\})^2}{2} .$$

Ceci n'est nul que si $\sum_x \mu(\{x\})^2 = 1$. Dans ce cas, vu $\mu(\{x\}) \in [0,1]$, et $\mu(\{x\})^2 \leq \mu(\{x\})$, on a $\sum_x \mu(\{x\}) = 1$, c'est-à-dire que μ est complètement atomique. Plus précisément, s'il existait un x tel que $0 < \mu(\{x\}) < 1$ alors on aurait $\sum_y \mu(\{y\})^2 < \sum_y \mu(\{y\}) = 1$, ce qui est exclu. Ainsi, tout x est tel que $\mu(\{x\}) = 0$ ou $\mu(\{x\}) = 1$; comme la somme est 1 on en conclut qu'il existe un unique a tel que $\mu(\{a\}) = 1$, c'est-à-dire que $X = a$ $\mathbb{P}$-p.s. ($\mu = \delta_a$, masse de Dirac en a).

Pour (2) : supposons que X_1 est non identiquement nulle, c'est-à-dire (par continuité à la limite décroissante des mesures de probabilité) qu'il existe $\varepsilon > 0$ tel que $\mathbb{P}[X_1 > \varepsilon] > 0$. Alors $\sum_n \mathbb{P}[X_n > \varepsilon] = \infty$, et les événements $\{X_n > \varepsilon\}$ étant indépendants, la réciproque du lemme de Borel-Cantelli donne que $\mathbb{P}[\limsup\{X_n > \varepsilon\}] = 1$. Ainsi, $\mathbb{P}$-p.s., une infinité de X_n est $> \varepsilon$, et $\sum_n X_n = \infty$ p.s. La seule exception est ainsi le cas trivial où la loi commune est la masse de Dirac en 0 ($X_n = 0$ p.s. pour tout n).

Pour (3) : supposons d'abord que $\ell < \infty$. Par croissance de F, définition du supremum, *etc.*, on a, pour tout $k \in \mathbb{N}$, $\mathbb{P}[X < \ell - 1/2^k] < 1$, et donc $\mathbb{P}[X \geq \ell - 1/2^k] > 0$. De ce fait, la réciproque du lemme de Borel-Cantelli (et le fait qu'une intersection dénombrable d'ensembles de mesure pleine est encore de mesure pleine) donnent que la probabilité que pour tout $k \in \mathbb{N}$, il y ait une infinité de n tels que $X_n > \ell - 1/2^k$ est égale à 1. Donc p.s. on peut trouver une extraction φ (qui dépend du point $\omega \in \Omega$ que l'on considère dans notre espace probabilisé sous-jacent) telle que $\limsup X_{\varphi(n)} \geq \ell$. Par ailleurs, $X_n \leq \ell$ pour tout n p.s., car $F(\ell) = 1$ (raisonner par l'absurde pour le voir). D'où le résultat. (Le cas $\ell = \infty$ est similaire, seules les notations changent, il faut considérer k au lieu de $\ell - 1/2^k$.)

Exercice 8. Une comparaison somme-intégrale donne, pour tout $a > 0$,

$$\int_a^\infty \mathbb{P}[X_1 \geq t]\,\mathrm{d}t = a\int_1^\infty \mathbb{P}[X_1 \geq au]\,\mathrm{d}u \leq a\sum_{n\geq 1}\mathbb{P}[X_1 \geq an] \leq a\int_0^\infty \mathbb{P}[X_1 \geq au]\mathrm{d}u = \mathbb{E}[X_1].$$

On en conclut que la somme $\sum_n \mathbb{P}[X_1 \geq an]$ est finie si et seulement si $E[X_1] < \infty$. Constatons par ailleurs que pour tout $a > 0$ et $n \in \mathbb{N}$, $\mathbb{P}[X_n \geq an] = \mathbb{P}[X_1 \geq an]$ (par identique distribution des X_j), et comme les événements $\{X_n \geq an\}$ sont indépendants, on applique selon le cas le lemme de Borel-Cantelli ou sa réciproque pour obtenir que p.s., $\limsup_{n\to\infty} X_n/n \leq a$ si $\mathbb{E}[X_1] < \infty$, et $\limsup_{n\to\infty} X_n/n \geq a$ si $\mathbb{E}[X_1] = \infty$. On conclut en faisant ensuite tendre a vers 0 dans le premier cas et vers $+\infty$ dans le second (de manière dénombrable).

Exercice 9. Les variables Z_n sont positives et $\mathbb{E}[Z_n] = 1/n^\alpha$, ce qui tend vers 0 si $n \to \infty$. Ainsi, $Z_n \to 0$ dans $\mathbb{L}^1$. En revanche, constatons que $\sum_{n\geq 1}\mathbb{P}[Z_n = 1]$ est finie si et seulement si $\alpha > 1$. Comme les événements $\{Z_n = 1\}$ sont indépendants (et que $Z_n \in \{0,1\}$), le lemme de Borel-Cantelli et sa réciproque permettent d'obtenir que, si $\alpha > 1$, alors p.s. $Z_n = 0$ à partir d'un certain rang, et si $0 < \alpha \leq 1$, il existe au contraire p.s. une infinité de n tels que $Z_n = 1$.

Exercice 10. Remarquons que l'exercice 8 donne $\lim_{n\to\infty} X_n/n = 0$, le but ici est donc de déterminer plus précisément le comportement asymptotique d'une suite de variables exponentielles.

Pour (1) : soit $a > 0$. On a (voir l'exercice 4) $\mathbb{P}[X_n \geq a\ln n] = e^{-a\ln n} = n^{-a}$, qui est sommable si et seulement si $a > 1$. Le même raisonnement que l'exercice précédent permet donc de dire d'une part (avec $a > 1$) que $\limsup_{n\to\infty} X_n/\ln n \leq 1$ p.s., puis ($a \leq 1$) que $\limsup_{n\to\infty} X_n/\ln n \geq 1$ p.s.

Pour (2) : constatons que pour tout $0 < a < 1$, par indépendance et identique distribution (voir exercice 3), $\mathbb{P}[Z_n \leq a] = \mathbb{P}[X_1 \leq a\ln n]^n = (1 - n^{-a})^n$. En prenant le logarithme, on s'aperçoit que ceci est un $o(\exp(-n^{1-a-\varepsilon})$ (pour tout $\varepsilon > 0$), c'est donc le terme général d'une série sommable. Le lemme de Borel-Cantelli donne alors que p.s., à partir d'un certain rang, $Z_n > a$ (et ce, pour tout $0 < a < 1$). De ce fait, on a p.s. $\liminf_{n\to\infty} Z_n \geq 1$.

Pour (3) : de la même façon, pour $a > 1$, on a $\mathbb{P}[Z_n \geq a] = 1 - (1 - n^{-a})^n$, et cette quantité est équivalente à n^{1-a} lorsque $n \to \infty$. De ce fait, on obtient que $\sum_{k\geq 1}\mathbb{P}[Z_{2^k} \geq a] < \infty$, et donc p.s., $\limsup_{k\to\infty} Z_{2^k} \leq 1$. Pour conclure, on majore : pour tout n il existe k tel que $2^{k-1} < n \leq 2^k$, et pour ce k,

$$Z_n \leq \frac{\max(X_1, \ldots, X_{2^k})}{\ln 2^{k-1}} = \left(\frac{\ln 2^k}{\ln 2^k - \ln 2}\right) Z_{2^k}\ .$$

En passant aux lim sup dans les deux membres, il vient $\limsup_{n\to\infty} Z_n \leq 1$ p.s., ce qui conclut.

Exercice 11. Le meilleur moyen de caractériser la loi, de calculer l'espérance, *etc.*, c'est de passer par les fonctions génératrices. Pour $r \in [0{,}1[$, le passage de la première à la deuxième ligne, puis de la deuxième à la troisième, se faisant par indépendance (et identique distribution),

$$\begin{aligned}
\mathbb{E}\left[r^Z\right] = \mathbb{E}\left[r^{X_1+\ldots+X_N}\right] &= \sum_{n\geq 0} \mathbb{E}\left[r^{X_1+\ldots+X_n} \ ; \ N=n\right] \\
&= \sum_{n\geq 0} \mathbb{P}[N=n]\,\mathbb{E}\left[r^{X_1+\ldots+X_n}\right] \\
&= \sum_{n\geq 0} \mathbb{P}[N=n]\,\left(\mathbb{E}\left[r^{X_1}\right]\right)^n \\
&= \sum_{n\geq 0} \mathbb{P}[N=n]\,(G_X(r))^n = \mathbb{E}\left[G_X(r)^N\right] = G_N(G_X(r)) \ ,
\end{aligned}$$

où l'on a noté G_N (resp. G_X) la fonction génératrice de N (resp. X). Par les résultats connus sur les séries entières, et ceux sur les séries génératrices, on voit alors que

$$\mathbb{E}\left[X_1+\ldots+X_N\right] = \lim_{r\to 1} (G_X \circ G_N)'(r) = \lim_{r\to 1} G_X'(r)\,G_N'(G_X(r)) = \mathbb{E}[N]\,\mathbb{E}[X] \ .$$

Application : Si N est de Poisson de paramètre λ et les X_j de Bernoulli de paramètre p, la somme aléatoire suit la loi de Poisson de paramètre λp.

Intégration et Probabilités – TD 12
Corrigé partiel

Exercice 1. Pour que le théorème de Kolmogorov ne s'applique pas, il suffit de ne pas avoir d'indépendance.

On peut par exemple prendre une suite p.s. constante dont seule la valeur initiale est aléatoire : par exemple $(0, 0, \ldots)$ avec probabilité $1/2$ et $(1, 1, \ldots)$ avec probabilité $1/2$. Dans ce cas, $\mathbb{P}[\limsup_n X_n = 1] = 1/2$, ce qui montre que la tribu asymptotique n'est pas triviale.

Plus généralement, on peut considérer une marche aléatoire "tuée" à la sortie d'une bande : on suppose les X_n i.i.d. selon une loi de Rademacher symétrique (i.e., les valeurs ± 1 sont prises chacune avec probabilité $1/2$), et on définit une suite $(S_n)_{n\geq 1}$ par $S_n = X_1 + \ldots + X_n$ pour $n < T_{ab} = \inf\{k \geq 1 : S_k = a \text{ ou } S_k = b\}$ avec $a < 0 < b$, et $S_n = S_{T_{ab}}$ pour $n \geq T_{ab}$. (Vous avez vu en cours pourquoi $T_{ab} < \infty$ p.s.).

Exercice 2. Les cas de la convergence p.s. et de la convergence en loi sont immédiats (pour la dernière, utiliser que la composée d'une fonction continue bornée et d'une fonction continue est encore continue bornée). Pour la convergence en probabilité, on raisonne sur des sous-suites, en utilisant le fait qu'une suite converge en probabilité si et seulement si de toute suite extraite, on peut ré-extraire une sous-suite convergeant p.s. (voir TD 4 pour la preuve de cette équivalence).

Exercice 3. Il n'est bien sûr pas vrai en général que $(Z_n, U_n) \rightsquigarrow (Z, U)$, et voici un contre-exemple assez trivial. Soit X de loi normale standard (alors X et $-X$ ont même loi). On pose $Z_n = X$, qui converge en loi vers $Z = -X$, et $U_n = X + R_n$, où $R_n \xrightarrow{\mathbb{P}} 0$. Alors $U_n \xrightarrow{\mathbb{P}} U = X$. Mais $(Z_n, U_n) = (X, X) + (0, R_n)$, où $(0, R_n) \xrightarrow{\mathbb{P}} (0, 0)$, et ainsi $(Z_n, U_n) \xrightarrow{\mathbb{P}} (X, X)$ – qui n'est pas de même loi que $(U, Z) = (X, -X)$.

En revanche, le théorème de Lévy permet de montrer que $(Z_n, U_n) \rightsquigarrow (Z, U)$ si et seulement si toute combinaison linéaire $\alpha Z_n + \beta U_n$ converge en loi vers $\alpha Z + \beta U$. (Mais nous venons de voir que les convergences en loi des marginales, i.e., les cas $\alpha = 0$ ou $\beta = 0$, ne suffisent pas à garantir la convergence des lois jointes.)

Cependant, dans le cas (1), il y a convergence jointe en probabilité, donc en loi. Dans le cas (2), par indépendance, on a la convergence suivante des fonctions caractéristiques :

$$\Phi_{(Z_n, U_n)}(\xi) = \mathbb{E}\left[e^{i\xi\cdot(Z_n, U_n)}\right] = \Phi_{Z_n}(\xi_1)\,\Phi_{U_n}(\xi_2) \to \Phi_Z(\xi_1)\,\Phi_U(\xi_2) = \Phi_{(Z,U)}(\xi) \ ,$$

et cette convergence permet, par théorème de Lévy, de conclure à la convergence en loi du couple.

Le cas (3) n'est pas complètement immédiat. On montre d'abord l'indication.

$$\left|\Phi_{Z'_n}(t) - \Phi_{Z'}(t)\right| \leq \mathbb{E}\left[\left|e^{it(Z'_n - Y'_n)} - 1\right|\right] \leq \varepsilon + \mathbb{P}\left[\,\left|\,Z'_n - Y'_n\,\right| > \delta\right] \ ,$$

où, pour la dernière indication, on a utilisé le fait que $x \mapsto e^{itx}$ est uniformément continue (pour $t \in \mathbb{R}$ fixé) ; on fixe $\varepsilon > 0$ et on considère δ qui lui est associé par uniforme continuité. En utilisant le fait que $X'_n - Y'_n$ converge en probabilité vers 0, on fait tendre n vers ∞, puis ε vers 0, et on montre ainsi que $\Phi_{Z'_n}$ et $\Phi_{Y'_n}$ ont même limite. Or on sait par théorème de Lévy que les $\Phi_{Z'_n}$ tendent ponctuellement vers $\Phi_{Z'}$, et la conclusion s'ensuit par une nouvelle application du théorème de Lévy. En particulier, dans (3), on désigne par $U = c \in \mathbb{R}$ la limite en loi et on note que, d désignant par exemple la distance de la norme du supremum, $d\left((Z_n, U_n), (Z_n, c)\right) = d(U_n, c)$ converge en probabilité vers 0, de sorte qu'il suffit de montrer, vu l'indication, que $(Z_n, c) \rightsquigarrow (Z, c)$. Mais ce dernier point est évident par théorème de Lévy.

Exercice 4. On peut raisonner comme à l'exercice 2, en passant par des sous-suites, et en utilisant le fait que si une suite converge $\mathbb{P}$-p.s., alors elle converge aussi $\mathbb{Q}$-p.s.

Ou : on note U la densité de $\mathbb{Q}$ par rapport à $\mathbb{P}$. Le fait que X_n converge en $\mathbb{P}$-probabilité vers $\mathbb{Q}$ implique en particulier la convergence en $\mathbb{P}$-probabilité de $\min\{1, |X_n - X|\}$ vers 0. Par lemme de Slutzky, il est encore vrai que

$$U \min\{1, |X_n - X|\} \xrightarrow{\mathbb{P}} 0 ,$$

et cette suite est dominée par $U \in \mathbb{L}^1(\mathbb{P})$. Par convergence dominée (voir TD 4, exercice 3),

$$\mathbb{E}_{\mathbb{P}}\left[U \min\{1, |X_n - X|\}\right] = \mathbb{E}_{\mathbb{Q}}\left[\min\{1, |X_n - X|\}\right] \to 0 ,$$

ce qui prouve la convergence en $\mathbb{Q}$-probabilité de la suite vers X.

Exercice 5. Il suffit de remarquer que la mesure m_n est la loi de la variable aléatoire $X_n = [2^n X]/2^n$, où $[x]$ désigne la partie entière de x, et où X est l'identité de $\mathbb{R}$ (ce dernier étant muni de la mesure m). Cette variable aléatoire converge ponctuellement (en tout point de $\mathbb{R}$) vers X, donc en particulier, X_n converge en loi vers X, soit m_n converge en loi vers m.

Pour l'approximation binômiale-Poisson, on utilise le théorème de Lévy, en remarquant que les fonctions caractéristiques convergent bien ponctuellement :

$$\left(1 - p_n + p_n e^{it}\right)^n \to \exp\left(\lambda\left(e^{it} - 1\right)\right) .$$

Exercice 6. On rappelle que la fonction caractéristique d'une gaussienne de moyenne m et de variance σ^2 est égale à $\Phi_{m,\sigma^2}(\xi) = e^{im\xi - \sigma^2\xi^2/2}$. De ce fait, la condition est immédiatement suffisante, par théorème de Lévy, avec convergence vers la loi normale $\mathcal{N}(m,\sigma^2)$. (Remarque : si $\sigma = 0$, la loi dégénère en δ_m, Dirac au point m.)

Pour prouver que la condition est nécessaire, commençons par observer que si Y_n converge en loi vers une variable aléatoire Y, on a en particulier, par théorème de Lévy, que le module de $\Phi_{Y_n}(\xi)$ converge pour tout ξ, et, celui-ci valant $e^{-\sigma_n^2\xi^2/2} \in]0,1]$, on conclut en prenant $\xi = 1$ que la suite des σ_n converge vers un certain $\sigma \in \mathbb{R}_+$.

Reste à régler son compte à m_n... On prouve tout d'abord, en raisonnant par l'absurde, que comme σ_n converge, $+\infty$ et $-\infty$ ne peuvent être valeurs d'adhérence de m_n. Sinon, quitte à extraire, on aurait par exemple $m_n \to \infty$. Or, par théorème de Portmanteau, $\mathbb{P}[Y < a] \leq \liminf_{n\to\infty} \mathbb{P}[Y_n < a]$. Et cette dernière limite est nulle, en effet,

$$\mathbb{P}[Y_n < a] = \int_{-\infty}^{a-m_n} \frac{e^{-x^2/(2\sigma_n^2)}}{\sqrt{2\pi\sigma_n^2}}\,\mathrm{d}x = \int_{-\infty}^{\sigma_n^{-1}(a-m_n)} \frac{e^{-x^2/2}}{\sqrt{2\pi}}\,\mathrm{d}x \to 0$$

(que σ_n tende vers 0 ou non). On aurait alors $\mathbb{P}[Y \in \mathbb{R}] = 0$, ce qui est absurde. Ainsi, $(m_n)_{n\geq 1}$ est bornée, et a une valeur d'adhérence. Soient m,m' deux valeurs d'adhérence de cette suite. Par passage à la limite dans la fonction caractéristique et en simplifiant par le module (ce qui fait partir les termes de variance), on a, par unicité de la limité, $e^{im\xi} = e^{im'\xi}$ pour tout $\xi \in \mathbb{R}$, soit $m = m'$.

Exercice 7. On utilise les fonctions de répartition (plus précisément, celles de survie). On a $\mathbb{P}[n(1-M_n) > t] = \mathbb{P}[M_n < 1-t/n] = (1-t/n)^n$, par indépendance et le fait que $\{M_n < t\} = \{X_i < t,\, 1 \le i \le n\}$. Ceci converge vers e^{-t} quand $n \to \infty$, ce qui est la fonction de survie de la loi exponentielle standard (i.e. de moyenne 1). On en conclut que $n(1-M_n)$ converge vers une telle loi.

Pour les lois de Cauchy : si $t < 0$, on a par un calcul explicite utilisant l'indépendance et le fait qu'une loi de Cauchy donne une masse $1/2$ à $\mathbb{R}_+$,

$$\mathbb{P}[n/M_n > t] \ge \mathbb{P}[M_n \ge 0] = 1 - 1/2^n \to 1\ .$$

Si maintenant $t \ge 0$, $\mathbb{P}[n/M_n > t] = \mathbb{P}[M_n < n/t,\, M_n \ge 0]$. Comme ceci se réécrit $\mathbb{P}[M_n < n/t] - \mathbb{P}[M_n \le n/t, M_n < 0]$ et que $\mathbb{P}[M_n \le 0] \to 0$ quand $n \to \infty$, le second terme tend vers 0. Quant au premier, il vaut, par indépendance,

$$\mathbb{P}[M_n < n/t] = \left(\frac{1}{\pi}\int_{-\infty}^{n/t} \frac{1}{1+x^2}\,\mathrm{d}x\right)^n = \left(\frac{1}{\pi}\left(\frac{\pi}{2} + \arctan\frac{n}{t}\right)\right)^n\ .$$

Par une égalité trigonométrique classique[1], ceci vaut $(1-\pi^{-1}\arctan(t/n))^n$, et converge vers $e^{-t/\pi}$. De même qu'avant, la considération des fonctions de survie montre donc la convergence vers une loi exponentielle de paramètre $1/\pi$ (et d'espérance π).

Exercice 8. On réécrit la probabilité considérée sous forme $\mathbb{P}[S_n - n \le 0] = \mathbb{P}[(S_n-n)/\sqrt{n} \le 0]$. Le théorème central limite donne que $(S_n-n)/\sqrt{n}$ converge en loi vers la loi gaussienne centrée réduite (la variance d'une loi exponentielle de paramètre 1 est 1, voir l'exercice 1 du TD 10). La fonction de répartition de la loi gaussienne étant continue (en 0), cette convergence en loi implique la convergence de la suite des probabilités ci-dessus vers $F(0) = 1/2$ (où F est la fonction de répartition d'une loi normale standard). Remarque : une approche erronée de cet exercice serait d'écrire la probabilité $\mathbb{P}[S_n/n \le 1]$ et d'essayer d'appliquer la loi des grands nombres : en effet S_n/n converge vers 1, et ainsi la mesure limite δ_1, masse de Dirac en 1, charge la frontière du fermé $]-\infty,1]$, de sorte que le théorème de Portmanteau ne s'applique pas. (En particulier, on ne peut pas conclure à la convergence vers $\mathbb{P}[1 \le 1] = 1$!)

Pour la seconde partie de l'exercice, on constate que la quantité considérée vaut $\mathbb{P}[S_n \le n]$, où S_n suit une loi de Poisson de paramètre n. Or on rappelle que la somme de deux variables aléatoires de Poisson indépendantes, de paramètres λ et μ, suit une loi de Poisson de paramètre $\lambda + \mu$ (cela se voit par les fonctions caractéristiques ou génératrices, voir par exemple l'exercice 1 du TD 10). Ainsi, on peut réécrire $S_n = X_1 + \ldots + X_n$, où les X_i sont indépendantes de loi de Poisson de paramètre 1 (donc de variance 1). On conclut comme ci-dessus que l'expression considérée admet une limite, et que cette dernière est égale à $1/2$.

Exercice 9. Pour la première question de l'exercice, on peut se référer à l'exercice 3 du TD 4 – et prendre la fonction "phare" par exemple, celle qui était doublement indicée en (n,m).

Pour la suite : l'événement $\{\limsup_{n\to\infty} S_n/\sqrt{n} \ge M\}$ est un événement de la tribu asymptotique, qui est $\mathbb{P}$-triviale par la loi du 0–1 (vu l'indépendance désormais supposée). Sa probabilité est 0 ou 1, montrons qu'elle est strictement positive. Le théorème de la limite centrale montre que $\mathbb{P}[S_n/\sqrt{n} \ge M] \to \Phi(M) > 0$ où Φ est la fonction de survie de la loi normale standard. Ainsi, pour tout $n \ge n_0$, $\mathbb{P}[S_n/\sqrt{n} \ge M] > \varepsilon > 0$. En particulier, $\mathbb{P}[\sup_{k\ge n} S_k/\sqrt{k} \ge M] > \varepsilon$ pour $n \ge n_0$, et par convergence monotone (décroissante),

1. cf. $1/\tan\theta = \tan(\pi/2-\theta)$

$\mathbb{P}[\limsup_n S_n/\sqrt{n} \geq M] \geq \varepsilon > 0$, et cette probabilité vaut en fait 1. On en déduit le résultat en faisant $M \to \infty$. Remarquer qu'on a aussi par symétrie $\mathbb{P}[\liminf_n S_n/\sqrt{n} = -\infty] = 1$.

Supposons alors que $S_n/\sqrt{n}$ converge en probabilité vers X. On sait alors que l'on peut extraire une sous-suite $S_{\varphi(n)}/\sqrt{\varphi(n)}$ convergeant p.s. vers X. Cependant, on prouve de la même manière que ci-dessus que $\{\limsup_{n\to\infty} S_{\varphi(n)}/\sqrt{\varphi(n)} \geq M\}$ a probabilité 1, et de même pour la limite inférieure. Ainsi, on a p.s. $\limsup S_{\varphi(n)}/\sqrt{\varphi(n)} = -\liminf S_{\varphi(n)}/\sqrt{\varphi(n)} = +\infty$, ce qui contredit toute convergence p.s.

Enfin, on a $\{S_n > 0, S_{2n} < 0\} = \{S_n > 0, S_n + (S_{2n} - S_n) < 0\}$. Or S_n et $S_{2n} - S_n$ sont indépendantes de même loi, et en particulier le théorème central limite, associé au point (2) de l'exercice 3 (ou un TCL dans $\mathbb{R}^2$) donne $(S_n/\sqrt{n}, (S_{2n} - S_n)/\sqrt{n}) \rightsquigarrow (N, N')$ où N et N' sont deux gaussiennes centrées indépendantes, de même variance (non précisée). La loi de (N, N') ne charge pas la frontière de l'octant $\{(x,y) : x < 0, x + y < 0\}$ donc, par théorème de Portmanteau, on a convergence de $\mathbb{P}[S_n > 0, S_{2n} < 0]$ vers $\mathbb{P}[N > 0, N + N' < 0] = 1/8$. La valeur précise de cette dernière probabilité s'obtient en décomposant $\mathbb{R}^2$ en 8 octants équiprobables pour la loi de (N, N'), invariante par rotation (car N, N' sont indépendantes et de même variance, raisonner par les matrices de variance-covariance, et les transformations vectorielles de vecteurs gaussiens).

Exercice 10. On a par TCL que

$$\sqrt{n}\left(\bar{X}_n - \lambda\right) \rightsquigarrow \mathcal{N}(0, \lambda) \ ,$$

où $\bar{X}_n$ désigne la moyenne empirique au temps n,

$$\bar{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k \ .$$

Par ailleurs, la loi des grands nombres assure la convergence en $\mathbb{P}$-probabilité de $\bar{X}_n$ vers λ (et aussi, par image continue, celle des $\sqrt{\bar{X}_n}$ vers $\sqrt{\lambda}$). Le lemme de Slutzky (cf. exercice 3) assure alors que

$$\sqrt{n}\,\frac{\bar{X}_n - \lambda}{\sqrt{\bar{X}_n}} \rightsquigarrow \mathcal{N}(0,1) \ .$$

Soit $u_{\alpha/2}$ le *quantile* de la loi normale standard donné par $F(u_{\alpha/2}) = 1 - \alpha/2$ (où F est la fonction de répartition de la loi normale standard). Par symétrie de cette dernière, et par théorème de Portmanteau,

$$\mathbb{P}\left[\sqrt{n}\ \left|\frac{\bar{X}_n-\lambda}{\sqrt{\bar{X}_n}}\right| \leq u_{\alpha/2}\right] \to 1-\alpha\ ,$$

soit un choix possible donné par

$$I_n(X_1^n) = \left[\bar{X}_n - \frac{u_{\alpha/2}\sqrt{\bar{X}_n}}{\sqrt{n}}, \bar{X}_n + \frac{u_{\alpha/2}\sqrt{\bar{X}_n}}{\sqrt{n}}\right]\ .$$

Exercice 11. Le modèle statistique sous-jacent est une suite de variables aléatoires $X_1, \ldots, X_n$ ($n = 2.645.756$) i.i.d. de loi commune de Bernoulli de paramètre $p \in]0,1[$ inconnu. On effectue tout d'abord un *test*. L'hypothèse est "$p = 1/2$" et on cherche à voir si on doit l'infirmer. (Les tests statistiques se contentent de mettre en évidence que certaines hypothèses sont très vraisemblablement contradictoires avec les données, ce qui ne signifie pas nécessairement qu'elles soient bonnes : les données ne les ont simplement pas encore trop gravement contredites. C'est ce qu'on appelle la dissymétrie des tests : ils ont tendance à conserver les hypothèses.)

Soit $\bar{p}_n = (X_1 + \ldots + X_n)/n$ la fréquence empirique. On a par TCL, *sous notre hypothèse*,

$$T_n = \sqrt{n}\ \frac{\bar{p}_n - 1/2}{(1/2)} \rightsquigarrow \mathcal{N}(0,1)\ .$$

En particulier, $\mathbb{P}[T_n \geq u]$ est proche de $1 - F(u)$, où F est la fonction de répartition d'une loi normale standard. Or, un calcul montre que sur les données, la valeur de la statistique de test T_n est $(1.359.670 - 1.322.878)/(2^{-1}\sqrt{2.645.756}) \geq u = 45$. Or, $1 - F(u) < e^{-(45)^2/2}/2$ (voir exercice 10 du TD 9), ce qui est ridiculement faible. L'hypothèse de départ est donc très certainement fausse. (En effet, on entend souvent dire qu'environ 51% des naissances sont des garçons...)

Une méthode plus précise consiste à trouver un intervalle dans lequel le paramètre p a 95 % de chances de se trouver. (Un test peut d'ailleurs être donné par l'indicatrice d'un tel intervalle, c'est la dualité tests-intervalles de confiance.) On a, de même qu'à l'exercice 10, $\bar{p}_n$ désignant toujours la fréquence empirique des naissances de garçons, $\bar{p}_n \to p$ en probabilité, et par TCL, théorème de l'image continue et lemme de Slutzky, quel que soit le vrai paramètre p,

$$\sqrt{n}\ \frac{\bar{p}_n - p}{\sqrt{\bar{p}_n(1-\bar{p}_n)}} \rightsquigarrow \mathcal{N}(0,1)\ .$$

On en déduit un intervalle de confiance (à niveau $1-\alpha$) donné par

$$I_n = \left[\bar{p}_n - \frac{u_{\alpha/2}\sqrt{\bar{p}_n(1-\bar{p}_n)}}{\sqrt{n}}, \bar{p}_n + \frac{u_{\alpha/2}\sqrt{\bar{p}_n(1-\bar{p}_n)}}{\sqrt{n}}\right]\ .$$

En prenant $\alpha = 0.05$ (soit une garantie asymptotique à 95 %), et $u_{\alpha/2} = 1.96$, on obtient sur les données réelles $I_n = [0.5133, 0.5145]$.