Corrigé des TD de probabilités

Printemps 2009

Voici quelques corrections, indications et remarques concernant les travaux dirigés donnés par Rémi Peyre et Frédéric Simon dans le cadre du cours « Probabilités » de Christophe Sabot pour les élèves de L3 de l'ÉNS Lyon au second semestre 2008/09.

Les élèves qui le désirent peuvent suggérer aux chargés de TD leurs propres solutions, saisies avec LaTeX, à envoyer par e-mail à `rpeyre@umpa.ens-lyon.fr` . Afin de faciliter le travail de mise en page, merci aux rédacteurs éventuels de respecter les consignes que vous trouverez dans l'annexe à la fin de ce document.

Feuille 1 : Théorie de la mesure

Exercice 1

I.1.1. Notons $\mathcal{C}$ la classe de parties de B définie par l'énoncé. On vérifie que $\mathcal{C}$ satisfait les propriétés d'une tribu, chaque propriété dérivant de la propriété correspondante sur $\mathcal{F}$. Montrons par exemple la stabilité par réunion dénombrable : soient $(C_n)_{n\in\mathbb{N}}$ des éléments de $\mathcal{C}$, par définition de $\mathcal{C}$ il existe des A_n dans $\mathcal{F}$ tels que $C_n = A_n \cap B$ pour tout n. Comme $\mathcal{F}$ est une tribu, $\bigcup_n A_n$ est dans $\mathcal{F}$, donc $\bigcup_n C_n = \big(\bigcup_n A_n\big) \cap B$ est bien un élément de $\mathcal{C}$.

I.1.2. Il est immédiat, en revenant aux propriétés caractérisant une tribu, que celles-ci sont stables par intersection, et ce pour n'importe quel type d'intersection.

I.1.3. Sur $\mathbb{N}$, la tribu engendrée par $\{0\}, \{1\}, \ldots$ et $\{n\}$ est consituée des parties de $\{0, 1, \ldots, n\}$ et de leurs complémentaires dans $\mathbb{N}$. La réunion de toutes les tribus de ce type est donc la classe des parties finies de $\mathbb{N}$ et de leurs complémentaires. Or cette classe n'est pas une tribu car elle n'est pas stable par union dénombrable : par exemple, $\{0\}, \{2\}, \{4\}, \ldots$ sont tous des ensembles finis, mais leur réunion $2\mathbb{N}$ n'est ni finie, ni de complémentaire fini.

Exercice 4

I.4.1. Définissons P_n comme donnant la masse 1 à $]0,1/n]$ et une masse nulle à tous les $]1/(i+1),1/i]$ pour $i \in \{1,\dots,n-1\}$. Alors ces mesures sont compatibles, mais aucune probabilité sur $(]0,1],\mathcal{F})$ n'est compatible avec elles, car sinon elle attribuerait une mesure nulle à tous les $]1/n,1]$, donc par réunion croissante une mesure nulle à $]0,1]$ tout entier, ce qui contredit la définition d'une mesure de probabilité.

Exercice 7

I.7.1. Non : par exemple $\bigcup_{n\in\mathbb{N}^\star}]n, n+2^{-n}[$ est ouvert et non borné, mais sa mesure est $\sum_{n\geqslant 1} 2^{-n} = 1$.

I.7.2. Non : $\mathbb{R} \smallsetminus \mathbb{Q}$, par exemple, est d'intérieur vide et est de mesure non nulle (son complémentaire est même de mesure nulle).

I.7.3. Non : prendre une énumération $q_1, q_2, \dots$ des rationnels de $]0,1[$ et placer un intervalle de longueur $2^{-(i+1)}$ autour de chaque q_i. Alors la réunion de ces intervalles est un ouvert dense (car contenant tous les rationnels), mais dont la mesure est bornée par $\sum_i 2^{-(i+1)} = 1/2$.

I.7.4. Pour $n \geqslant 1$, posons

$$B_n = \bigcup_{m\in\mathbb{Z}}]2^{-n}m, 2^{-n}m + 2^{-2n}[,$$

et définissons par récurrence $A_0 = \varnothing$, et pour $n \geqslant 1$:

$$A_n = \begin{array}{ll} A_{n-1} \cup B_n & \text{si } n \text{ est impair ;} \\ A_{n-1} \smallsetminus B_n & \text{si } n \text{ est pair.} \end{array}$$

On pose alors $A = \overline{\lim}\, A_n$, càd. $A = \bigcap_{n\in\mathbb{N}} \bigcup_{p\geqslant n} A_p$ (*). Les B_n et les A_n sont clairement boréliens, donc A l'est aussi.

Nous allons maintenant monter que pour tout intervalle ouvert borné non vide I, on a $0 < \lambda(A\cap I) < \lambda(I)$. Soit donc un tel I. En fait, nous allons juste démontrer le membre de gauche de l'inégalité, sachant que le membre de droite se démontre de la même façon vu que $\mathbb{R} \smallsetminus A$ peut se définir de façon très similaire à A. On observe que I contient nécessairement un intervalle de la forme $]2^{-n}m, 2^{-n}m + 2^{-2n}[$ pour n impair, donc on peut se limiter au cas où I lui-même est de cette forme. Notons n_0 la valeur du n pour lequel I est de la forme souhaitée, alors $A_{n_0} \cap I = I$, donc $\lambda(A_{n_0} \cap I) = \lambda(I)$. On

(*). Les parties d'un ensemble sont (partiellement) ordonnées par la relation $\subset$, qui vérifie les propriétés des bornes supérieure et inférieure. La limite supérieure d'une suite de parties est alors la limite décroissante, quand l'indice i croît, des suprema des parties d'indice au moins i, et la limite inférieure est définie de même. En particulier, si une suite de parties d'un ensemble est indexée par $\mathbb{N}$, sa limite supérieure est l'ensemble des points qui apparaissent dans un nombre infini de termes de la suite, et sa limite inférieure est l'ensemble des points qui apparaissent dans tous les termes de la suite sauf un nombre fini d'entr'eux.

contaste ensuite que, pour tout $k \in \mathbb{N}$, $\lambda(B_{n_0+2k+1} \cap I) = 2^{-2k-1}\lambda(I)$, or, par définition de A :

$$(A \cap I) \supset (A_{n_0} \cap I) \setminus \left(\bigcup_{k \in \mathbb{N}} B_{n_0+2k+1} \cap I \right),$$

d'où $\lambda(A \cap I) \geqslant \lambda(I)/3 > 0$ (†).

Remarque. Cette question montre qu'il existe des boréliens qui « ne ressemblent pas trop » ni à des ouverts, ni à des fermés. Il existe un résultat qui va dans l'autre sens, qui est le *théorème de densité de Lebesgue* : celui-ci affirme que, si A est un borélien de $\mathbb{R}$, presque-tout $a \in A$ vérifie $\lambda(A \cap]a-\varepsilon, a+\varepsilon[) \underset{\varepsilon \longrightarrow 0}{\sim} 2\varepsilon$. Par conséquent, dès lors que ni A, ni $\mathbb{R} \setminus A$ ne sont de mesure nulle, l'ensemble des quotients $\lambda(A \cap I)/\lambda(I)$ possède 0 et 1 comme valeurs d'adhérence. Notamment, il n'existe aucun borélien $A \subset \mathbb{R}$ tel que, pour tout I ouvert borné non vide, on ait $\lambda(A \cap I) = \lambda(I)/2$.

I.7.5. On voit par récurrence que K_n a pour mesure $(1-d_0)(1-d_1)\cdots(1-d_{n-1})$, donc K, qui est la réunion décroissante des K_n, a pour mesure la limite décroissante de ces nombres, i.e. $\prod_{n \in \mathbb{N}}(1-d_n)$.

I.7.6. Soit K un compact comme à la question 5. Pour repérer un point de ce compact, on peut commencer par dire s'il est dans l'intervalle de gauche ou de droite de K_1, puis si, parmi les deux intervalles de K_2 issus de cet intervalle, il est dans celui de gauche ou de droite, etc. On définit ainsi une application $\varphi : K \longrightarrow \{gauche, droite\}^{\mathbb{N}^\star}$. Cette application est injective, car deux points qui sont dans le même intervalle de K_n sont nécessairement à distance moins de 2^{-n} l'un de l'autre. Elle est également surjective, car les intervalles dont les K_n sont constitués sont compacts, et la réunion décroissante de compacts non vide est non vide (‡). φ est donc une bijection. Maintenant, si K_1 et K_2 sont comme dans la question 6, $\varphi_2^{-1} \circ \varphi_1$ est une bijection entre K_1 et K_2, qui manifestement est strictement croissante, ce qui montre que c'est un homéomorphisme (§). Ainsi tous les K obtenus par le procédé de la question 5 sont homéomorphes (¶).

Maintenant, on a l'encadrement $1 - \sum d_n \leqslant \prod(1-d_n) \leqslant e^{-\sum d_n}$ (‖), donc les K construits à la question 5 peuvent être de mesure nulle (p. ex. si $d_n = 1/3$ pour tout n) ou de mesure non nulle (p. ex. si $d_n = 2^{-n-2}$). Cela conclut.

(†). On a utilisé que $\sum_{k \geqslant 0} 2^{-2k-1} = 2/3$.

(‡). Attention : si les K_n avaient été constitués d'intervalles *ouverts*, les éléments de $\{gauche, droite\}^{\mathbb{N}^\star}$ finissant par une infinité de $gauche$ ou une infinité de $droite$ n'auraient *pas* correspondu à un élément de $\bigcap K_n$!

(§). C'est un exercice de topologie que de montrer que, si ψ est une bijection strictement croissante entre deux parties A et B de $\mathbb{R}$, alors c'est un homéomorphisme. L'idée est de voir que, pour tous $a_1, a_2 \in A$, $\varphi(]a_1, a_2[\cap A) =]\varphi(a_1), \varphi(a_2)[$, et d'en déduire que φ est ouverte, et de montrer similairement que φ^{-1} est ouverte aussi.

(¶). En fait, toute partie compacte, sans point isolé et totalement discontinue d'un espace Polonais est homéomorphe à l'ensemble de Cantor (qui est le K qu'on obtient en prenant $d_n = 1/3$ pour tout n).

(‖). On peut en fait montrer que $\prod(1-d_n)$ est nul si et seulement si $\sum d_n = \infty$, en utilisant que $\ln(1-x)$ est équivalent à $-x$ au voisinage de 0.

Feuille 2 : Variables aléatoires

Exercice 4

II.4.6. L'idée est de voir que, dans les formules de changement de variables, on a un problème quand f' est nulle. On va donc essayer de construire une fonction f strictement croissante dont la dérivée s'annule sur un « gros » ensemble.

Soit Z un fermé de $\mathbb{R}$ d'intérieur vide et de mesure de Lebesgue non nulle : un tel Z, par exemple, a été construit dans l'exercice I.7. (questions 5 et 6). Soit f une fonction continue sur $\mathbb{R}$, positive, et dont l'ensemble des zéros est précisément Z — prendre par exemple pour f la fonction « distance à Z », et soit F une primitive de f, alors F est une fonction $\mathcal{C}^1$ strictement croissante. Or nous affirmons que la mesure de Lebesgue de $F(Z)$ est nulle : ce résultat est un cas particulier du théorème de Sard, qui en dimension 1 affirme que l'image des zéros de la dérivée d'une fonction $\mathcal{C}^1$ de $\mathbb{R}$ dans $\mathbb{R}$ est de mesure nulle (**). Pour le voir dans ce cas-là, on peut observer que, si U est un ouvert sur lequel $|f'| \leqslant \varepsilon$, alors $\mathrm{Leb}(f(U)) \leqslant \varepsilon \cdot \mathrm{Leb}(U)$ — simplement en revenant à la définition de la mesure de Lebesgue —, et en déduire que l'image de toute partie bornée de Z est de mesure nulle, et donc par réunion dénombrable que Z lui-même l'est aussi.

Maintenant, si X est une variable aléatoire de densité strictement positive par rapport à la mesure de Lebesgue (par exemple si X suit une loi de Cauchy), la loi de $f(X)$ attribue une mesure non nulle à $f(Z)$, qui est pourtant de mesure de Lebesgue nulle, de sorte que $f(X)$ ne peut pas être à densité.

Exercice 5

II.5.1. Pour montrer que $F^{-1}(U)$ a pour loi μ, on va montrer que la loi de $F^{-1}(U)$ a pour fonction de répartition F, i.e. que pour tout $x \in \mathbb{R}$:

$$\mathbb{P}(F^{-1}(U) \leqslant x) = F(x). \tag{1}$$

D'après la définition de F^{-1}, « $F^{-1}(U) \leqslant x$ » est équivalent à « $\forall y > x\ F(y) \geqslant U$ », et comme ici F est continue à droite cela est finalement équivalent à « $U \leqslant F(x)$ ». Or, puisque U est uniforme sur $]0,1[$, la probabilité de cet événement est $F(x)$, d'où (1).

II.5.2. Il suffit de montrer que pour tout $p \in]0,1[$, $\mathbb{P}(F(X) \leqslant p) = p$. Puisque μ est diffuse, F est continue donc tout $p \in]0,1[$ a un antécédent par F. Notons x_p le plus grand de ces antécédents (qui existe bien, car l'ensemble des antécédents de p est majoré et fermé), alors « $F(X) \leqslant p$ » est équivalent à « $X \leqslant x_p$ ». Mais comme X a pour loi μ, la probabilité de cet événement, par définition d'une fonction de répartition, est alors $F(x_p)$, càd. p.

(**). L'énoncé général du théorème de Sard est que, si f est une fonction $\mathcal{C}^k$ ($k \geqslant 1$) de $\mathbb{R}^n$ dans $\mathbb{R}^m$ avec $n \leqslant m + k - 1$, alors l'ensemble des *valeurs critiques* de f, i.e. l'image par f des points où Df n'est pas de rang maximal, est de mesure de Lebesgue nulle.

II.5.3. On va calculer F et F^{-1} pour les deux lois de l'énoncé, et alors d'après la question 1 il suffira d'appliquer F^{-1} à une variable uniforme sur $]0,1[$.

- Pour la loi exponentielle de paramètre λ, $\mathrm{d}\mathbb{P}(x) = \mathbb{1}_{x\geqslant 0}\lambda e^{-\lambda x}\mathrm{d}x$ d'où en intégrant $F(x) = 1 - e^{-\lambda x}$ pour $x \geqslant 0$ — et $F(x) = 0$ pour $x \leqslant 0$ —, et donc $F^{-1}(p) = |\ln(1-p)|/\lambda$.
- En intégrant la densité, on trouve que la fonction de répartition de la loi de Cauchy est $F(x) = \mathrm{Arctan}(x)/\pi + 1/2$, d'où $F^{-1}(p) = \tan(\pi(p - 1/2))$.

Exercice 6

II.6.1. Notons $I_k = \int x^k \mathrm{d}\sigma(x)$ et tentons de calculer les I_k. Déjà, la symétrie de la loi σ nous assure que $I_k = 0$ pour k impair. Pour k pair, on calcule alors en intégrant par parties que :

$$4I_k - I_{k+2} = \frac{1}{2\pi}\int_{-2}^{2} x^k(4-x^2)^{3/2}\mathrm{d}x = \frac{3}{2\pi(k+1)}\int_{-2}^{2} x^{k+2}\sqrt{4-x^2}\mathrm{d}x = \frac{3}{k+1}I_{k+2},$$

d'où $I_{k+2} = \frac{4(k+1)}{k+4}I_k$. Comme par ailleurs I_0 est $1/2\pi$ fois l'aire d'un demi-cercle de rayon 2, soit 1 [(††)], il s'ensuit que pour k pair :

$$I_k = \frac{4 \times 12 \times 20 \times \cdots \times (4(k-1))}{4 \times 6 \times 8 \times \cdots \times (k+2)} = \frac{C_k^{k/2}}{\frac{k}{2}+1}.$$

Remarque. D'après l'inégalité $C_n^p \leqslant 2^n$ pour tout p, on voit que $I_k \leqslant 2^k$ pour tout k.

Remarque. Pour $\ell \in \mathbb{N}$, le nombre $C_{2\ell}^{\ell}/(\ell+1)$ est toujours entier. Cette famille d'entiers, qui interviennent souvent en combinatoire, s'appellent les *nombres de Catalan*.

II.6.2. Soit μ une mesure de probabilité sur $\mathbb{R}$ à support non compact, alors pour tout $M \in \mathbb{R}_+$ il existe un $\varepsilon_M > 0$ tel que $\mu({}^{c}[-M,M]) \geqslant \varepsilon_M$, et par conséquent pour tout $k \in \mathbb{N}$:

$$\int_{\mathbb{R}} x^{2k}\mathrm{d}\mu(x) \geqslant \int_{{}^{c}[-M,M]} x^{2k}\mathrm{d}\mu(x) \geqslant M^{2k}\mu({}^{c}[-M,M]) \geqslant M^{2k}\varepsilon_M,$$

d'où $\underline{\lim}_{k\longrightarrow\infty} \left(\int_{\mathbb{R}} x^{2k}\mathrm{d}\mu(x)\right)^{1/2k} \geqslant M$. Cela état valable pour tout M, on a donc :

$$\left(\int_{\mathbb{R}} x^{2k}\mathrm{d}\mu(x)\right)^{1/2k} \underset{k\longrightarrow\infty}{\longrightarrow} +\infty. \tag{1}$$

Or pour la mesure σ :

$$\overline{\lim}_{k\longrightarrow\infty} \left(\int_{\mathbb{R}} x^{2k}\mathrm{d}\sigma(x)\right)^{1/2k} \leqslant 2^{\,(\ddagger\ddagger)},$$

(††). Au passage, il est nécessaire que I_0 vaille 1 pour que σ soit une mesure de probabilité !

de sorte qu'une mesure ayant les mêmes moments que σ ne peut pas ne pas être à support compact.

II.6.3. Si μ est une mesure de probabilité portée par $[-M, M]$, il est clair que $\int_{\mathbb{R}} x^{2k} \mathrm{d}\mu(x) \leqslant M^{2k}$, donc

$$\overline{\lim} \Big(\int_{\mathbb{R}} x^{2k} \mathrm{d}\mu(x) \Big)^{1/2k} \leqslant M,$$

ce qui est incompatible avec (1) : ainsi, une mesure ayant les mêmes moments qu'une mesure à support compact est nécessairement à support compact elle aussi. Maintenant, soient μ_1 et μ_2 deux mesures de probabilité à support compact, soit K un compact portant à la fois μ_1 et μ_2, et soit $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction continue bornée. Pour tout $\varepsilon > 0$, le théorème de Stone-Weierstrass nous assure alors l'existence d'un polynôme P_ε tel que $|\varphi(x) - P_\varepsilon(x)| \leqslant \varepsilon$ pour tout $x \in K$. On a donc pour $i \in \{1, 2\}$:

$$\Big| \int_{\mathbb{R}} \varphi(x) \mathrm{d}\mu_i(x) - \int_{\mathbb{R}} P_\varepsilon(x) \mathrm{d}\mu_i(x) \Big| \leqslant \int_K |\varphi(x) - P_\varepsilon(x)| \mathrm{d}\mu_i(x) \leqslant \varepsilon,$$

d'où

$$\int_{\mathbb{R}} \varphi(x) \mathrm{d}\mu_i(x) = \lim_{\varepsilon \longrightarrow 0} \int_{\mathbb{R}} P_\varepsilon(x) \mathrm{d}\mu_i(x). \tag{2}$$

Maintenant, si μ_1 et μ_2 ont les mêmes moments, alors par linéarité de l'intégrale ces deux mesures attribuent la même intégrale à tout polynôme, donc par (2) à toute fonction continue bornée, donc elles coïncident.

II.6.4. La valeur absolue de μ a pour densité $\mathbb{1}_{x \geqslant 0} |\sin(\sqrt{3} x^{1/3})| e^{-x^{1/3}}$, donc pour tout $k \geqslant 0$:

$$\int x^k |\mathrm{d}\mu|(x) \leqslant \int_0^\infty x^k e^{-x^{1/3}} < \infty$$

puisque $e^{-x^{1/3}}$ décroît plus vite que l'inverse de n'importe quel polynôme, en particulier plus vite que $x^{-(k+2)}$ par exemple.

II.6.5. Notons $J_k = \int_0^\infty x^k \sin(\sqrt{3} x^{1/3} e^{-x^{1/3}}) \mathrm{d}x$, et tentons de calculer les J_k. On commence par observer que :

$$J_k = \mathfrak{Im} \Big(\int_0^\infty x^k e^{(\sqrt{3}i - 1) x^{1/3}} \Big).$$

En intégrant trois fois par parties, on trouve que

$$\int_0^\infty x^k e^{(\sqrt{3}i-1)x^{1/3}} = \frac{-(\sqrt{3}i - 1)^3}{(3k+3)(3k+4)(3k+5)} \int_0^\infty x^{k+1} e^{(\sqrt{3}i-1)x^{1/3}},$$

or $(\sqrt{3}i - 1)^3 = 8 \in \mathbb{R}$, de sorte que la détermination J_0 entraînera celle de tous les J_k. Calculons maintenant J_0 : un changement de variables montre que :

$$\int_0^\infty e^{(\sqrt{3}i-1)x^{1/3}} \mathrm{d}x = 3 \int_0^\infty y^2 e^{(\sqrt{3}i-1)y} \mathrm{d}y,$$

(‡‡). En fait, que μ soit à support compact où non, le membre de gauche de (1) tend en croissant vers $\sup\{M\,;\, \mu({}^c[-M, M]) > 0\}$.

ce qui, après deux intégrations par parties, est égal à :

$$\frac{6}{(\sqrt{3}i-1)^2}\int_0^\infty e^{(\sqrt{3}i-1)y}\mathrm{d}y = -\frac{6}{(\sqrt{3}i-1)^3} = -\frac{3}{4} \in \mathbb{R},$$

de sorte que $J_0 = 0$, et donc $J_k = 0$ pour tout $k \in \mathbb{N}$.

II.6.6. Soient μ_+ et μ_- respectivement les parties positive et négative de μ, i.e. les mesures dont les densités sont resp. les parties positive et négative de la densité de μ, de sorte que μ_+ et μ_- sont des mesures positives avec $\mu_+ - \mu_- = \mu$. Alors le résultat de la question 5 nous dit que μ_+ et μ_- ont les mêmes moments, et d'après la question II.6.4. ces moments sont finis vu que μ_+ et μ_- sont toutes les deux majorées par $|\mu|$. Or μ_+ et μ_- sont des mesures distinctes—et même étrangères. Elles n'ont pas de raison d'être de probabilité, mais elles ont la même masse totale puisque $\int \mathrm{d}\mu(x) = 0$, donc en divisant μ_+ et μ_- par leur masse totale commune on obtient deux mesures de probabilités distinctes ayant tous leurs moments identiques et finis.

Exercice 7

II.7.1. $S^N(p)$ compte le nombre de X_i^N dans $[0,p]$, or les X_i sont indépendants et chacun d'eux a une probabilité p d'être dans $[0,p]$, donc $S^N(p)$ suit une binômiale de paramètres N et p — rappelons que cela signifie que

$$\mathbb{P}(S_p^N = s) = C_N^s p^s (1-p)^{N-s}.$$

II.7.2. $S^N(q) - S^N(p)$ compte le nombre de X_i^N dans $]p,q]$, donc par le même argument qu'à la question 1 cette variable suit une loi binômiale de paramètres N et $q-p$.

II.7.3. L'événement de l'énoncé signifie que n_1 exactement des X_i^N sont dans $[0,p_1]$, que n_2 exactement sont dans $]p_1, p_1+p_2]$, que n_3 exactement sont dans $]p_1+p_2, p_1+p_2+p_3]$, etc. Déjà, la probabilité d'avoir « $X_1^N, \dots, X_{n_1}^N \in [0,p_1]$ et $X_{n_1+1}^N, \dots, X_{n_1+n_2}^N \in]p_1, p_1+p_2]$ et … et $X_{N-p_k+1}^N, \dots, X_N^N \in]1-p_k, 1]$ » vaut $\prod_{j=1}^k {p_j}^{n_j}$. Mais en fait il y a plusieurs façons d'obtenir l'événement de l'énoncé, selon l'identité des indices des X_i^N qui seront dans chaque intervalle. Comme, pour chaque distribution possible des X_i^N entre les différents intervalles, la probabilité de réalisation est la même, le résultat sera donc le produit de $\prod_{j=1}^k {p_j}^{n_j}$ par le nombre de façons de diviser un ensemble de N éléments en k parties comportant respectivement $n_1, \dots, n_k$ éléments. Ce nombre est le coefficient multinômial $N!/n_1! \cdots n_k!$, comme on peut le voir en comptant de deux façons le nombre de manières de répartir N éléments sur N places numérotées. Finalement donc, le résultat recherché est :

$$\mathbb{P}[\cdots] = \frac{N!}{n_1! \cdots n_k!} \prod_{j=1}^k {p_j}^{n_j}.$$

II.7.4.

$$\begin{aligned}\mathbb{P}[S^N(x/N)=s] &= C_N^s\left(\frac{x}{N}\right)^s\left(1-\frac{x}{N}\right)^{N-s}\\ &= \frac{x^s}{s!}\underbrace{\frac{N(N-1)\cdots(N-s+1)}{N^s}}_{\longrightarrow 1}\underbrace{(1-x/N)^N}_{\longrightarrow e^{-x}}\underbrace{(1-x/N)^s}_{\longrightarrow 1}\underset{N\longrightarrow\infty}{\longrightarrow}\frac{x^s}{s!}e^{-x}.\end{aligned}$$

On vérifie que

$$\sum_{s\in\mathbb{N}}\frac{x^s}{s!}e^{-x}=e^xe^{-x}=1,$$

donc il existe bien une loi de probabilité sur $\mathbb{N}$ attribuant la masse $x^se^{-x}/s!$ à l'entier s : c'est la *loi de Poisson* de paramètre x.

II.7.5. Pour tout $x\geqslant 0$, on a l'égalité entre événements :

$$\{NX_{(1)}^N\leqslant x\}=\{X_{(1)}^N\leqslant x/N\}=\{\exists i\ \ X_i^N\leqslant x/N\}=\{S^N(x/N)\geqslant 1\},$$

donc $\mathbb{P}(NX_{(1)}^N\geqslant x)\underset{N\longrightarrow\infty}{\longrightarrow}\sum_{s\geqslant 1}\pi_s(x)=1-\pi_0(x)=1-e^{-x}$, ce qui correspond à la fonction de répartition d'une loi exponentielle de paramètre 1.

II.7.6. Par le même argument que pour la question II.7.5., on calcule que

$$\mathbb{P}(NX_{(s)}^N\leqslant x)\underset{N\longrightarrow\infty}{\longrightarrow}\sum_{t\geqslant s}\pi_t(x),$$

qui est bien l'expression d'une fonction de répartition puisque cela tend vers 1 quand $x\longrightarrow+\infty$. La loi-limite de $NX_{(s)}^N$ a alors pour densité par rapport à la mesure de Lebesgue $\frac{\mathrm{d}}{\mathrm{d}x}\sum_{t\geqslant s}\pi_t(x)$. Après avoir vérifié, grâce à l'expression explicicte des $\pi_t(x)$, qu'il est légitime d'inverser somme et dérivée, on calcule que :

$$\frac{\mathrm{d}}{\mathrm{d}x}\sum_{t\geqslant s}\pi_t(x)=\sum_{t\geqslant s}\left[\frac{x^{t-1}}{(t-1)!}e^{-x}-\frac{x^t}{t!}e^{-x}\right]=\sum_{t\geqslant s-1}\frac{x^t}{t!}e^{-x}-\sum_{t\geqslant s}\frac{x^t}{t!}e^{-x}=\frac{x^{s-1}}{(s-1)!}e^{-x}.$$

II.7.7. La fait que la loi trouvée à la question II.7.5. soit la loi exponentielle, qui correspond au temps écoulé avant qu'un événement se produise dans un processus sans mémoire, tend à indiquer qu'au voisinage de 0, quand N devient grand, la présence ou l'absence d'un point en un certain endroit n'influe pas sur la présence ou l'absence d'autres points ailleurs. Par conséquent, on devine que la loi-limite de $NX_{(s)}^N$ sera la somme de s lois exponentielles de paramètre 1 indépendantes, ce qu'on appelle aussi la *loi* Γ de paramètres s et 1. Un calcul permet alors de vérifier que c'est bien le cas.

Remarque. En fait, on peut rendre rigoureux l'argument informel que nous venons de donner. La loi de l'ensemble des $NX_{(i)}^N$, quand $N\longrightarrow\infty$, tend vers ce qu'on appelle le *processus ponctuel de Poisson* sur $\mathbb{R}_+$, qui modélise les moments successifs où un événement se produit dans un processus sans mémoire.

Feuille 3 : Indépendance

Exercice 2

III.2.1 Si A dépend des variables aléatoires X_i pour $i \in I$ et que B dépend des variables aléatoires X_i pour $i \in J$ avec et J disjoints, alors d'après l'exercice précédent A et B appartiennent à des tribus indépendantes et sont donc indépendants.

III.2.2 Soit $A \in \sigma(X_1, \ldots, X_n)$. Posons $C = \cup_{m \geqslant n+1} \sigma(X_{n+1}, \ldots, X_m)$. C est stable par intersections finies, contient Ω et la tribu engendrée par C contient $\sigma(X_{n+1}, \ldots)$.

Soit $\mathcal{D} = \{D \in \sigma(X_{n+1}, \ldots)$ tels que $\mathbb{P}(A \cap D) = \mathbb{P}(A)\mathbb{P}(D)\}$. Comme dans l'exercice précédent, $\mathcal{D}$ est une classe monotone contenant C (d'après la question 1). En appliquant le lemme des classes monotones, on en déduit alors que $\mathcal{D} = \sigma(X_{n+1}, \ldots)$. Puisque la tribu asymptotique est incluse dans $\sigma(X_{n+1}, \ldots)$, tous les évènements de la tribu asymptotique sont indépendants de A.

III.2.3 Soit B dans la tribu asymptotique. On a montré dans la question précédente que $\cup_n \sigma(X_1, \ldots, X_n)$ (qui est stable par intersection finie et qui contient Ω) était indépendante de la tribu asymptotique. En appliquant encore une fois le lemme des classes monotones à la classe monotone $\mathcal{E} = \{E \in \sigma(X_1, \ldots)$ tels que $\mathbb{P}(B \cap E) = \mathbb{P}(B)\mathbb{P}(E)\}$, on en déduit que $\sigma(\cup_n \sigma(X_1, \ldots, X_n)) = \sigma(X_1, \ldots) \subset \mathcal{E}$. Ceci montre donc que $\sigma(X_1, \ldots)$ et la tribu asymptotique sont indépendants.

III.2.4 Soit A dans la tribu asymptotique. Puisque la tribu asymptotique est incluse dans $\sigma(X_1, \ldots)$, alors tout évènement de la tribu asymptotique est indépendant de lui même. On a donc $\mathbb{P}(A \cap A) = \mathbb{P}(A)\mathbb{P}(A)$ ce qui montre que $\mathbb{P}(A) \in \{0, 1\}$.

Exercice 3

III.3.1. Par réunion finie il suffit de démontrer que n'importe quel couple de X_i d'indices distincts, par exemple X_1 et X_2, sont distincts presque-sûrement. Or par le théorème de Fubini $\mathbb{P}(X_1 = X_2) = \int_{[0,1]} \mathbb{P}(X_2 = x) \mathrm{d}\mathbb{P}(X_1 = x)^{(*)} = \int_{[0,1]} 0 \, \mathrm{d}\mathbb{P}(X_1 = x) = 0$.

Remarque. Le même raisonnement permet de montrer que, si X_1 et X_2 sont des variables indépendantes — pas forcément i.i.d.— et que la loi de X_2 est diffuse, alors X_1 et X_2 sont distinctes presque-sûrement.

III.3.2. Puisque σ ne prend qu'un ensemble discret de valeurs, il suffit de montrer que, pour tout $\tau \in \mathfrak{S}_n$, l'événement « $\sigma = \tau$ » est indépendant de $(U^{(1)}, \ldots, U^{(n)})$. Soient $\tau \in \mathfrak{S}_n$ et A un borélien de $\mathbb{R}^n$. L'événement « $\sigma = \tau$ et $(U^{(1)}, \ldots, U^{(n)}) \in A$ »

$(*)$. Attention à bien lire la formule : il n'y a pas de différentielle devant le premier symbole $\mathbb{P}$ et il y en a une devant le second, ce qui signifie qu'on intègre, *selon la loi de* X_1, la variable aléatoire (j'entends par là « une fonction de x ») qui évalue la probabilité *que* X_2 *soit exactement égal à x*.

est équivalent — à un ensemble de mesure nulle près — à « $U_{\tau(1)} > U_{\tau(2)} > \cdots > U_{\tau(n)}$ et $(U_{\tau(1)}, \ldots, U_{\tau(n)}) \in A$. » Comme la loi de $\{U_1, \ldots, U_n\}$ est symétrique en les indices des U_i, elle est invariante par toute permutation des indices, de sorte que $\mathbb{P}(\sigma = \tau$ et $(U^{(1)}, \ldots, U^{(n)}) \in A)$ est la même pour tout $\tau \in \mathfrak{S}_n$. Par conséquent, en sommant sur tous les τ possibles on voit que cette probabilité vaut $\mathbb{P}((U^{(1)}, \ldots, U^{(n)}) \in A)/n!$. Comme par ailleurs $\mathbb{P}(\sigma = \tau) = 1/n!$ (prendre $A = \mathbb{R}^n$ dans la formule précédente), cela signifie bien que « $\sigma = \tau$ » et « $(U^{(1)}, \ldots, U^{(n)}) \in A$ » sont indépendants.

III.3.3. Notre réponse à la question 2 répond aussi à la question 3.

III.3.4. Déjà, la loi de $U^{(\cdot)}$ est portée par les n-uplets $(x_1, \ldots, x_n)$ tels que $0 < x_1 < \cdots < x_n < 1$. Pour $0 < a_1 < b_1 < \cdots < a_n < b_n < 1$, déterminons la probabilité du rectangle $R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$ contenu dans cet ensemble. Pour $\tau \in \mathfrak{S}_n$, $\mathbb{P}(\sigma = \tau$ et $U^{(\cdot)} \in \mathbb{R}) = \mathbb{P}(U_{\tau(1)} \in [a_1, b_1]$ et $U_{\tau(2)} \in [a_2, b_2]$ et $\ldots$ et $U_{\tau(n)} \in [a_n, b_n]) = (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n)$, donc en sommant sur τ on voit que la loi de $U^{(\cdot)}$ est la loi à densité :

$$\mathrm{d}\mathbb{P}\big(U^{(\cdot)} = (x_1, \ldots, x_n)\big) = \mathbb{1}_{0<x_1<\cdots<x_n<1}\, n!\, \mathrm{d}x_1 \cdots \mathrm{d}x_n. \tag{1}$$

III.3.5. La loi de $U^{(i)}$ est la i-ème marginale de (1) ; sa densité est donc donnée par l'intégrale de la densité de $U^{(\cdot)}$ sur un hyperplan de i-ème coordonnée fixée. Il faut donc calculer, pour tout $x \in \mathbb{R}$:

$$n! \int_{\mathbb{R}^d} \mathbb{1}_{0<x_1<\cdots<x_{i-1}<x<x_{i+1}<\cdots<x_n<1} \mathrm{d}x_1 \ldots \mathrm{d}x_{i-1} \mathrm{d}x_{i+1} \ldots \mathrm{d}x_n. \tag{2}$$

Pour $x \notin]0, 1[$, (2) vaut clairement 0. Pour $x \in]0, 1[$, cela se factorise en :

$$n! \Big(\int_{0<x_1<\cdots<x_{i-1}<x} \mathrm{d}x_1 \ldots \mathrm{d}x_{i-1} \Big) \cdot \Big(\int_{x<x_{i+1}<\cdots<x_n<1} \mathrm{d}x_{i+1} \ldots \mathrm{d}x_n \Big).$$

Par un changement de variables, on calcule que :

$$\int_{0<x_1<\cdots<x_{i-1}<x} \mathrm{d}x_1 \ldots \mathrm{d}x_{i-1} = x^{i-1} \int_{0<x_1<\cdots<x_{i-1}<1} \mathrm{d}x_1 \ldots \mathrm{d}x_{i-1},$$

l'intégrale du membre de droite valant $1/(i-1)!$ car en prenant $n = i-1$ dans (1) on voit que $(i-1)!$ fois cette intégrale est égale à l'intégrale de la densité d'une loi de probabilité. Le même argument nous donne aussi :

$$\int_{x<x_{i+1}<\cdots<x_n<1} \mathrm{d}x_{i+1} \ldots \mathrm{d}x_n = \frac{(1-x)^{n-i}}{(n-i)!},$$

Donc finalement :

$$\mathrm{d}\mathbb{P}(U^{(i)} = x) = \frac{n!}{(i-1)!(n-i)!} \mathbb{1}_{0<x<1} x^{i-1} (1-x)^{n-i} \mathrm{d}x.$$

Exercice 6

III.6.1 On peut à chaque fois dériver sous le signe intégrale par théorème de convergence dominée / théorème de dérivation sous le signe intégrale car Z est à valeurs dans un intervalle borné. On dérive deux fois et on trouve :

$$\psi''(z) = \frac{\mathbb{E}(Z^2\exp(\lambda Z))\mathbb{E}(\exp(\lambda Z)) - \mathbb{E}(Z\exp(\lambda Z))}{(\mathbb{E}(\exp(\lambda Z)))^2}.$$

Notons f la densité de Z par rapport à la mesure μ. On définit une variable aléatoire Z_λ de densité $g(x) = \frac{\exp(\lambda x)}{\psi_Z(\lambda)} f(x)$ par rapport à la mesure μ (g est bien positive et d'intégrale 1). On peut alors regarder la variance de la variable aléatoire Z_λ. C'est exactement l'expression qu'on a trouvée ci-dessus.

III.6.2 On suppose que $I = [a, b]$. L'astuce est d'utiliser le fait que $\mathrm{Var}(Z) = \mathrm{Var}(Z - \frac{b+a}{2})$. Puisque Z est à valeurs dans I, on a $a \leqslant Z \leqslant b$ donc $-\frac{b-a}{2} \leqslant Z - \frac{b+a}{2} \leqslant \frac{b-a}{2}$. On a donc $\mathrm{Var}(Z) \leqslant \mathbb{E}((Z - \frac{b+a}{2})^2) \leqslant \frac{(b-a)^2}{4}$.

III.6.3 On a $\psi(0) = \psi'(0) = 0$ et on contrôle la dérivée seconde de ψ. On intègre deux fois et on a le résultat demandé.

III.6.4 On utilise l'indépendance des X_i pour montrer que $\mathbb{E}(\prod \exp(\lambda X_i)) = \prod \mathbb{E}(\exp(\lambda X_i))$. On a donc $\psi_{\widetilde{Z}} = \sum \psi_{X_i}$. D'après la majoration de la question 3, on a l'inégalité demandée.

III.6.5 On va utiliser l'inégalité de Markov. On a $\mathbb{P}(|\widetilde{Z}| \geqslant \varepsilon) = \mathbb{P}(\widetilde{Z} \geqslant \varepsilon) + \mathbb{P}(-\widetilde{Z} \geqslant \varepsilon)$. On a par exemple $\mathbb{P}(\widetilde{Z} \geqslant \varepsilon) = \mathbb{P}(\exp(\lambda\widetilde{Z}) \geqslant \exp(\lambda\varepsilon)) \leqslant \exp(\psi_{\widetilde{Z}}(\lambda) - \lambda\varepsilon)$ par Markov.

On a donc $\mathbb{P}(\widetilde{Z} \geqslant \varepsilon) \leqslant \exp(\frac{\lambda^2}{8}\sum(b_i - a_i)^2 - \lambda\varepsilon)$. Il suffit de minimiser cette inégalité valable pour tout λ et de traiter le deuxième terme de la manière pour obtenir l'inégalité demandée.

Feuille 4 : Lemme de Borel-Cantelli, Fonctions caractéristiques

Exercice 4

IV.4.1 Notons μ la mesure de probabilité associée à la fonction de répartition F. Soit $j < k$. X_j et X_k sont indépendantes. On peut donc calculer la fonction de répartition de $X_j - X_k$.

On a $\mathbb{P}(X_j - X_k \leqslant z) = \int F(z+y)\mathrm{d}\mu(y) = G(z)$. On peut alors utiliser le théorème de continuité sous le signe intégrale pour montrer que G est continue. La continuité de G implique alors que $\lim_\varepsilon G(0) - G(-\varepsilon) = \lim_\varepsilon \mathbb{P}(X_j - X_k \in]-\varepsilon, 0]) = \mathbb{P}(X_j - X_k = 0) = 0$.

IV.4.2 Soit π_n la permutation induite par $X_1, \ldots, X_n$. Soit σ une permutation aléatoire de $\{1, \ldots, n\}$ indépendante de $X_1, \ldots, X_n$, choisie de manière uniforme sur les $n!$ possibilités. Étudions alors la permutation induite par $X_{\sigma(1)}, \ldots, X_{\sigma(n)}$. Il est clair que cette permutation est $\pi_n \circ \sigma$ et qu'elle a la même distribution que π_n (puisque les X_i sont indépendantes de même loi). On conclut en remarquant que si π est une permutation, alors $\pi \circ \sigma$ est uniforme sur les $n!$ possibilités.

IV.4.3 On a $\mathbb{P}(A_n) = \mathbb{P}(\pi_n(n) = 1)$. Il suffit donc de compter combien de permutations fixent un point donné. Il y en a $(n-1)!$. On trouve donc le résultat annoncé.

IV.4.4 Soit σ_m une permutation aléatoire de $\{1, \ldots, m\}$ indépendante des X_i, choisie de manière uniforme sur les $m!$ possibilités. Encore une fois, $\pi_n \circ \sigma_m$ a la même distribution que π_n et puisque σ_m réarrange $\pi_n(1), \ldots, \pi_n(m)$ uniformément, sans toucher aux autres indices, on montre que $\mathbb{P}(A_m | \pi_n(j) = i_j \text{ pour } m+1 \leqslant j \leqslant n) = \frac{1}{m}$.

Si $m_1 < m_2 < \ldots < m_k$, on déduit de l'égalité précédente que $\mathbb{P}(A_{m_1} | A_{m_2} \cap \ldots \cap A_{m_k}) = \mathbb{P}(A_{m_1})$ et on montre ensuite par récurrence que les A_k sont indépendants.

IV.4.5 On pose $X_i = \mathbb{1}_{A_i}$. On a donc $R_n = X_1 + \ldots + X_n$. Les A_i sont indépendants donc les X_i le sont également. Ils sont donc non corrélés. On a donc $\text{Var}(R_n) = \text{Var}(X_1) + \ldots + \text{Var}(X_n)$. De plus, puisque $X_k \in \{0, 1\}$, on a $\text{Var}(X_k) \leqslant \mathbb{E}((X_k)^2) = \mathbb{E}(X_k)$ donc $\text{Var}(R_n) \leqslant \mathbb{E}(R_n)$.

On peut donc utiliser l'inégalité de Tchebychev :

On a $\mathbb{P}(|R_n - \mathbb{E}(R_n)| > \delta\mathbb{E}(R_n)) \leqslant \frac{\text{Var}(R_n)}{(\delta\mathbb{E}(R_n))^2} \leqslant \frac{1}{((\delta)^2\mathbb{E}(R_n))} \to 0$ car $\mathbb{E}(R_n) = \sum_{i=1}^n \frac{1}{i} \to \infty$.

On a donc $\frac{R_n}{\mathbb{E}(R_n)}$ qui tend vers 1 en probabilité.

IV.4.6 L'astuce pour pouvoir utiliser Borel-Cantelli va être d'utiliser des soussuites. Posons $n_k = \inf\{n | \mathbb{E}(R_n) \geqslant k^2\}$. On pose $T_k = R_{n_k}$.

On a alors $k^2 \leqslant \mathbb{E}(T_k) \leqslant k^2 + 1$. En réutilisant Tchebychev, on montre que $\mathbb{P}(|T_k - \mathbb{E}(T_k)| > \delta\mathbb{E}(T_k)) \leqslant \frac{1}{(\delta k)^2}$. On utilise ensuite Borel Cantelli ce qui montre que $\mathbb{P}(\overline{\lim}_k \frac{|T_k - \mathbb{E}(T_k)|}{\mathbb{E}(T_k)} > \delta) = 0$ pour tout $\delta > 0$, donc en particulier pour tout $\delta \in \mathbb{Q}^{*+}$. On a donc $\mathbb{P}(\cup_{\delta \in \mathbb{Q}^{*+}} \overline{\lim}_k \frac{|T_k - \mathbb{E}(T_k)|}{\mathbb{E}(T_k)} > \delta) = 0$. L'ensemble $\cap_{\delta \in \mathbb{Q}^{*+}} \underline{\lim}_k (\frac{|T_k - \mathbb{E}(T_k)|}{\mathbb{E}(T_k)} \leqslant \delta)$ est alors de mesure 1. Or, cet ensemble est l'ensemble des ω pour lesquels la suite $\frac{|T_k - \mathbb{E}(T_k)|}{\mathbb{E}(T_k)}$ converge vers 0. On a donc $\frac{T_k}{\mathbb{E}(T_k)}$ qui tend vers 1 presque-sûrement.

Soit maintenant ω tel que $\frac{T_k}{\mathbb{E}(T_k)}$ tend vers 1. Soit $n_k \leqslant n \leqslant n_{k+1}$. On a alors $\frac{T_k(\omega)}{\mathbb{E}(T_{k+1})} \leqslant \frac{R_n(\omega)}{\mathbb{E}(R_n)} \leqslant \frac{T_{k+1}(\omega)}{\mathbb{E}(T_k)}$. On vérifie alors que $\frac{\mathbb{E}(T_k)}{\mathbb{E}(T_{k+1})}$ tend vers 1, et on en déduit que $\frac{R_n(\omega)}{\mathbb{E}(R_n)}$ tend vers 1 presque-sûrement. Puisque $\mathbb{E}(R_n)$ est équivalent à $\ln(n)$, on a le résultat demandé.

Exercice 7

IV.7.1. Soit $\xi \in \mathbb{R}^d$ et ξ_n une suite convergeant vers ξ, alors pour tout $x \in \mathbb{R}^d$ on a $e^{i\xi_n \cdot x} \underset{n\longrightarrow\infty}{\longrightarrow} e^{i\xi\cdot x}$. Or, aussi bien $e^{i\xi_n\cdot x}$ que $e^{i\xi\cdot x}$ ont leurs valeurs absolues bornées par 1 pour tout x, et comme 1 est intégrable par rapport à μ (puisque μ est de probabilité et donc à masse totale finie), il s'ensuit par le théorème de convergence dominée que

$$\int_{\mathbb{R}^d} e^{i\xi_n\cdot x} d\mu(x) \underset{n\longrightarrow\infty}{\longrightarrow} \int_{\mathbb{R}^d} e^{i\xi\cdot x} d\mu(x),$$

càd. que $\widehat{\mu}(\xi_n) \underset{n\longrightarrow\infty}{\longrightarrow} \widehat{\mu}(\xi)$. μ est donc continue en tout point, CQFD.

IV.7.2. L'égalité de lois donnée par l'hypothèse nous donne, quand on passe aux fonctions caractéristiques :

$$\widehat{\mu}(\xi)^n = \widehat{\mu}\big(\lambda(n)\xi\big) \tag{1}$$

pour tous $n \geqslant 2$ et $\xi \in \mathbb{R}$. Nous allons en déduire la forme de $\widehat{\mu}$. Déjà, comme $\widehat{\mu}$ est continue, vu qu'elle vaut toujours 1 en 0 elle est non nulle au voisinage de 0, et donc sur $\mathbb{R}$ tout entier comme on le voit en itérant (1) — n y étant pris arbitraire. On peut donc introduire $\Phi = -\ln\circ\widehat{\mu}$ (†), qui est une certaine fonction $\Phi : \mathbb{R} \longrightarrow \mathbb{C}$ continue, nulle en 0 ; (1) se réécrit alors (‡) :

$$\Phi(\lambda(n)\xi) = n\Phi(\xi). \tag{2}$$

On observe que, comme μ n'est pas concentrée sur une masse de Dirac, $\widehat{\mu}$ est de valeur absolue strictement plus petite que 1 sur un voisinage de 0, 0 exlcus (§), et donc par (1) que $|\widehat{\mu}| < 1$ sur $\mathbb{R} \setminus \{0\}$ tout entier. On peut donc poser, pour $w \in \mathbb{R}$, $\Upsilon(w) = -\ln\Phi(\exp w)$ (¶) ; Υ est une fonction continue pour laquelle (2) devient :

$$\Upsilon(w + \ln\lambda(n)) = \Upsilon(w) + \ln n. \tag{3}$$

C'est alors un exercice d'analyse classique que de démontrer que Υ est une fonction affine de pente $\ln n/\ln\lambda(n)$, à condition d'observer d'abord qu'il y a deux $\ln\lambda(n)$ qui sont de rapport irrationnel (‖). Connaissant la forme de Υ, on en déduit celle de Φ, puis celle de

(†). On peut bien définir *le* logarithme de $\widehat{\mu}$, bien que $\widehat{\mu}$ soit *a priori* à valeurs complexes, car $\widehat{\mu}$ est continue et donc par le théorème de relèvement de l'argument elle a une unique détermination de son logarithme continue sur $\mathbb{R}$ et nulle en 0.

(‡). Il y a tout de même une petite astuce, c'est qu'il faut s'assurer que le choix que nous avons fait pour la détermination du logarithme complexe assure qu'il transforme bien les puissances entières en multiplications. On le démontre en vérifiant que c'est vrai pour $\xi = 0$ et en utilisant la continuité de $\widehat{\mu}$.

(§). Démonstration de ce lemme : prendre deux intervalles bornés disjoints à chacun desquels μ attribue une mesure non nulle, alors pour $|\xi| \neq 0$ suffisamment petit les valeurs de $e^{i\xi x}$ sur ces deux intervalles sont des arcs disjoints du cercle unité, de sorte que $\widehat{\mu}(\xi)$, qui est un barycentre de points du cercle unité attribuant une masse non nulle à chacun des ces deux intervalles, est de valeur absolue strictement plus petite que 1. On observera que ce raisonnement marche en fait pour la transformée de Fourier de n'importe quelle mesure de probabilité différente d'un Dirac, indépendamment de toute autre hypothèse.

(¶). À nouveau, on n'a pas d'ambiguïté dans la détermination du logarithme, cette fois parce que Φ est de partie réelle strictement positive. On vérifie également que cette détermination-ci du logarithme se comporte bien vis-à-vis de la multiplication par un réel.

(‖). Démontrons, par exemple, que $\ln\lambda(2)$ et $\ln\lambda(3)$ sont de rapport irrationnel : si tel n'est pas le cas, alors on trouve p et q entiers tels que $q\ln\lambda(2) = p\ln\lambda(3)$, et alors en appliquant $(p+q)$ fois (3) on trouve que $q\ln 2 = p\ln 3$, ce qui est absurde car $\ln 2/\ln 3 = \log_3 2$ est irrationnel.

$\widehat{\mu}(\xi)$ pour $\xi > 0$ (et donc pour $\xi \geqslant 0$ vu que $\widehat{\mu}(0) = 1$). Le même raisonnement marche pour $\xi < 0$, à ceci près que cette fois il faudra considérer $\widetilde{\Upsilon}(w) = -\ln \Phi(-\exp w)$.

Toutefois, les coefficents qu'on obtient pour les positifs et pour les négatifs n'ont *a priori* aucune raison d'avoir un lien : en fait, qu'ils soient conjugués découle de ce que la transformée de Fourier de n'importe quelle mesure vérifie $\widehat{\mu}(-\xi) = \overline{\widehat{\mu}(\xi)}$, comme on s'en assure en revenant à la définition.

Remarque. En relisant attentivement le raisonnement ci-dessus, on s'aperçoit qu'on n'a procédé que par équivalences, de sorte qu'une mesure de probabilité ayant un fonction caractéristique de la forme de l'énoncé sera forcément α-stable.

IV.7.3. Puisque μ a tous ses moments finis, la variance d'une variable aléatoire de loi μ est bien définie. Notons v cette variance ; vu qu'on a supposé que μ n'était pas une masse de Dirac, v est strictement positive. Comme la variance de deux variables indépendantes est la somme de leurs variances, l'hypothèse sur μ donne alors $2v/\lambda(2)^2 = v$, d'où $\lambda(2) = \sqrt{2}$ et donc $\alpha = 2$.

IV.7.4. Si nous appliquons formellement le théorème de dérivation sous le signe somme à $\widehat{\mu}(\xi) = \int e^{i\xi x} \mathrm{d}\mu(x)$, nous trouvons d'abord $\widehat{\mu}'(\xi) = i \int x e^{i\xi x} \mathrm{d}\mu(x)$, puis $\widehat{\mu}''(\xi) = -\int x^2 e^{i\xi x} \mathrm{d}\mu(x)$, d'où $\widehat{\mu}''(0) = -\int x^2 \mathrm{d}\mu(x)$, ce qui a bien un sens puisque μ a tous ses moments finis. Reste à vérifier qu'on est bien en situation d'appliquer ce théorème : pour tout ξ, les dérivées partielles $ixe^{i\xi x}$ et $-x^2 e^{i\xi x}$ sont majorées en module par $|x|$ et x^2 respectivement, lequels sont bien μ-intégrables toujours à cause de l'hypothèse de moments finis. Il s'ensuit que le raisonnement fait ci-dessus était valide, en particulier $\widehat{\mu}$ est deux fois dérivable en 0.

IV.7.5. Nous savons que $\widehat{\mu}$ est de la forme donnée par l'énoncé de la question 2 pour $\alpha = 2$ (cf. question 3), et que $\widehat{\mu}$ est dérivable en 0 (cf. question 4). En calculant la dérivée seconde de $\widehat{\mu}$ en 0 par la gauche et par la droite, on voit alors que $\overline{c} = c$, donc que c est réel. Dès lors une tranformation de Fourier inverse nous donne pour μ la densité :

$$\frac{\mathrm{d}\mu}{\mathrm{d}x}(x) = \frac{1}{2\sqrt{\pi c}} e^{-x^2/4c}.$$

On reconnaît la densité gaussienne, qui est effectivement une densité de probabilité ayant tous ses moments polynomiaux finis.

Remarque. En fait, pour répondre aux questions 3 à 5 nous avons seulement utilisé que μ avait un moment d'ordre 2. Par conséquent, une loi stable qui a un moment d'ordre 2 a automatiquement des moments de tous ordres.

IV.7.6. Pour X une variable aléatoire de loi μ, $\widehat{\mu}(-\xi) = \mathbb{E}[e^{i(-\xi)X}] = \mathbb{E}[e^{i\xi(-X)}] = \widehat{\mu}(\xi)$ puisque $-X$ a aussi pour loi μ.

IV.7.7. Nous savons que $\widehat{\mu}$ est de la forme donnée par l'énoncé de la question 2 pour $\alpha = 1$, or $\widehat{\mu}$ doit être symétrique (cf. question 6), ce qui impose à c d'être réel. Un calcul de transformée de Fourier inverse nous donne alors la densité suivante pour μ :

$$\frac{\mathrm{d}\mu}{\mathrm{d}x}(x) = \frac{1}{\pi c(1 + x^2/c^2)}.$$

On reconnaît la distribution lorentzienne, qui est bien une mesure de probabilité.

Feuille 5 : Notions de convergence

Exercice 11

V.11.1. Soit $\varphi \in \mathcal{C}_b(\mathbb{R})$, notons $\|\varphi\|_\infty$ le supremum de $|\varphi|$; on veut montrer que $\mathbb{E}[\varphi(X_n+Y_n)]$ tend vers $\mathbb{E}[\varphi(Y+x)]$, resp. que $\mathbb{E}[\varphi(X_nY_n)]$ tend vers $\mathbb{E}[\varphi(xY)]$. Soient $M<\infty$ et $\varepsilon>0$. Définissons le module de continuité suivant pour φ :

$$\omega(\varepsilon, M) = \sup_{\substack{|y|,|z|\leqslant M\\ |z-y|\leqslant\varepsilon}} |\varphi(z)-\varphi(y)|.$$

Alors, en notant que

$$\varphi(X_n+Y_n) = \varphi\big((Y_n+x)+(X_n-x)\big)$$

et en découpant selon les valeurs possibles de Y_n et de x, on trouve que

$$\big|\mathbb{E}[\varphi(X_n+Y_n)] - \mathbb{E}[\varphi(Y_n+x)]\big| \leqslant \\ \omega(\varepsilon, M) + \|\varphi\|_\infty\big(\mathbb{P}(|Y_n|\geqslant M-\varepsilon) + \mathbb{P}(|X_n-x|\geqslant\varepsilon)\big). \quad (1)$$

Or on sait d'une part que Y_n tend en loi vers Y, donc en particulier $\overline{\lim}\,\mathbb{P}(|Y_n|\geqslant M-\varepsilon) \leqslant \mathbb{P}(|Y|\geqslant M-\varepsilon)$, et d'autre part que X_n tend en loi vers x, donc, par le résultat de l'exercice 9, que $\overline{\lim}\,\mathbb{P}(|X_n-x|\geqslant\varepsilon)=0$. Du coup, en passant à la limite supérieure dans 1, on obtient :

$$\overline{\lim_{n\longrightarrow\infty}} \big|\mathbb{E}[\varphi(X_n+Y_n)] - \mathbb{E}[\varphi(Y_n+x)]\big| \leqslant \omega(\varepsilon, M) + \|\varphi\|_\infty \mathbb{P}(|Y|\geqslant M-\varepsilon).$$

On fait alors ε vers 0 pour faire disparaître le module de continuité dans le membre de droite (∗∗), puis on fait tendre M vers l'infini on fait disparaître le second terme, et on a alors prouvé que

$$\big|\mathbb{E}[\varphi(X_n+Y_n)] - \mathbb{E}[\varphi(Y_n+x)]\big| \underset{n\longrightarrow\infty}{\longrightarrow} 0. \quad (2)$$

Or $\mathbb{E}[\varphi(Y_n+x)] \longrightarrow \mathbb{E}[\varphi(Y+x)]$ à cause de la convergence faible de Y_n vers Y (††), donc en combinant cela avec (2) on a bien montré que $\mathbb{E}[\varphi(X_n+Y_n)]$ convergeait en loi vers $\mathbb{E}[\varphi(Y+x)]$. Cela conclut le cas de l'addition ; la même preuve, *mutatis mutandis*, marche pour la multiplication.

V.11.2. L'idée est d'utiliser le théorème de Skorohod. En fait, dans cette question on a juste besoin de Skorohod sur $\mathbb{R}$, qu'on a démontré dans l'exercice 10. D'après ce

(∗∗). En utilisant qu'une fonction continue sur le compact $[-M,M]$ y est automatiquement uniformément continue.

(††). Car, en introduisant la fonction continue bornée $\varphi_x(y)=\varphi(y+x)$, cela revient à dire que $\mathbb{E}[\varphi_x(Y_n)] \longrightarrow \mathbb{E}[\varphi_x(Y)]$

théorème, il existe réalisation des lois des Y_n sur un même univers Ω telle que Y_n converge vers Y *en probabilité*. Complétons cette réalisation des lois des Y_n en une réalisation des lois des (Y_n, X_n) : cela est possible en toute généralité, quitte à raffiner un peu la tribu dont est munie Ω. Cela ne *définit* pas les X_n car il y a *a priori* plein de manières d'introduire des X_n tels que les (X_n, Y_n) aient la bonne loi, mais nous choisissons arbitrairement *une* façon d'introduire ces X_n. Alors, puisque la loi des X_n n'a pas changé, ceux-ci convergent en loi vers x, donc en probabilité *car x est une constante* (cf. exercice 9). Du coup, on a obtenu que dans notre réalisation, les Y_n convergeaient en probabilité vers Y, resp. les X_n vers x, or *dans le cadre d'une convergence en probabilité* on en déduit immédiatement que (X_n, Y_n) converge en probabilité vers (x, Y), et donc *a fortiori* en loi. Mais puisque les lois des (X_n, Y_n) et de Y sont les mêmes dans la situation de départ que dans la réalisation que nous avons construite, il s'ensuit que, même dans la situation de départ, (X_n, Y_n) converge en loi vers (x, Y), CQFD.

Du coup, si f est une fonction continue de $\mathbb{R}^2$ dans $\mathbb{R}$ et φ une fonction continue bornée de $\mathbb{R}$ dans $\mathbb{R}$, $\varphi \circ f$ est continue et bornée de $\mathbb{R}^2$ dans $\mathbb{R}$, donc d'après la convergence en loi de (X_n, Y_n) vers (x, Y), $\mathbb{E}\big[\varphi\big(f(X_n, Y_n)\big)\big] \longrightarrow \mathbb{E}\big[\varphi\big(f(x, Y)\big)\big]$. Puisque c'est vrai pour toute φ, cela signifie par définition de la convergence en loi que $f(X_n, Y_n)$ converge étroitement vers $f(x, Y)$ (‡‡).

Feuille 6 : Loi des grands nombres

Exercice 4

VI.4.1. On écrit :

$$\begin{aligned}\mathbb{E}[X] = \int_\Omega X(\omega)\mathrm{d}\mathbb{P}(\omega) == \int_\Omega \mathrm{d}\mathbb{P}(\omega)\Big(\int_0^\infty \mathbb{1}_{x\leqslant X(\omega)}\mathrm{d}x\Big) \\ = \int_0^\infty \mathrm{d}x\Big(\int_\Omega \mathbb{1}_{x\leqslant X(\omega)}\mathrm{d}\mathbb{P}(\omega)\Big) = \int_0^\infty \mathbb{P}(X \geqslant x)\mathrm{d}x,\end{aligned}$$

où l'utilisation du théorème de Fubini est légitimiée par le fait que toutes les quantités écrites ici sont positives.

VI.4.2. Puisque X n'est pas L^1, cela signifie que la variable positive $|X|$ n'est pas intégrable, et donc d'après la question 1 que $\int_0^\infty \mathbb{P}(|X| \geqslant x)\mathrm{d}x = \infty$. Maintenant, comme les X_i sont indépendants, la question, par le lemme de Borel–Cantelli, est équivalente à montrer que

$$\sum_{i=1}^\infty \mathbb{P}(|X| \geqslant i) = \infty.$$

(‡‡). On a en fait démontré le résultat plus général selon lequel, pour E et F des espaces Polonais, si des v.a. Z_n à valeurs dans E convergent en loi vers Z et que f est une fonction continue des E dans F, alors les v.a. $f(Z_n)$ sur F convergent en loi vers $f(Z)$.

Or la fonction $\mathbb{P}(|X| \geqslant \cdot)$ est clairement décroissante, ce qui fait qu'on a la minoration :

$$\mathbb{P}(|X| \geqslant i) \geqslant \int_i^{i+1} \mathbb{P}(|X| \geqslant x)\mathrm{d}x,$$

et donc

$$\sum_{i=1}^{\infty} \mathbb{P}(|X| \geqslant i) \geqslant \int_1^{\infty} \mathbb{P}(|X| \geqslant x)\mathrm{d}x = \infty^{(*)},$$

CQFD.

VI.4.3. Soit $(X_i)_{i\in\mathbb{N}^\star}$ une suite de réels (déterministe), et posons $\widehat{E}_N = \frac{1}{N}\sum_{i=1}^{N} X_i$. Nous allons montrer qu'il est impossible d'avoir à la fois $\widehat{E}_N$ qui tende vers une limite finie λ et $|X_i| \geqslant i$ vérifié une infinité de fois, ce qui permettra de déduire la question 3 de la question 2. Raisonnons donc par l'absurde en supposant la condition vérifiée. Pour $i \geqslant 2$, on a par définition des $\widehat{E}_N$:

$$\widehat{E}_i = \frac{1}{i}\big((i-1)\widehat{E}_{i-1} + X_i\big),$$

d'où

$$\frac{|X_i|}{i} = \Big|\widehat{E}_i - \frac{i-1}{i}\widehat{E}_{i-1}\Big|. \tag{3}$$

Puisque nous avons supposé que $\widehat{E}_i$ convergeait vers λ, le membre de droite de (3) tend vers 0 quand i tend vers l'infini, et donc $\lim_i(|X_i|/i) = 0$. Or, puisque $|X_i| \geqslant i$ une infinité de fois, on devrait avoir aussi $\overline{\lim}_i(|X_i|/i) \geqslant 1$, ce qui est impossible.

VI.4.4. Soient μ la loi de X et ν la loi de Y. On a par la théorème de Fubini :

$$\mathbb{E}[|X+Y|] = \int |x+y|\mathrm{d}(\mu\otimes\nu)(x,y) = \int \mathrm{d}\mu(x)\Big(\int |x+y|\mathrm{d}\nu(y)\Big).$$

Or, pour tout $x\in\mathbb{R}$, $\int |x+y|\mathrm{d}\nu(y)$ est infini puisque $\int |x+y|\mathrm{d}\nu(y) \geqslant \int(|y|-|x|)\mathrm{d}\nu(y) = \int |y|\mathrm{d}\nu(y) - |x| = +\infty^{(\dagger)}$. Par conséquent, $\mathbb{E}[|X+Y|] = \int(+\infty)\mathrm{d}\mu(x) = +\infty$, ce qu'on voulait.

VI.4.5. D'après la question 4, $X_1, X_1+X_2, X_1+X_2+X_3, \ldots$ ne sont pas intégrables, donc, par linéarité de l'intégrale, les $\widehat{E}_N$, qui sont ces v.a. divisées par N, ne sont pas intégrables non plus.

VI.4.6. La loi forte (resp. L^1) des grands nombres affirme que, si X est intégrable, alors $\widehat{E}_N$ converge vers une constante $^{(\ddagger)}$ p.s. (resp. dans L^1). La question 3 (resp. 5) nous montre que la condition « X est intégrable » est en fait nécessaire au résultat (dans le cas

($*$). On sait grâce à la question 1 que $\int_0^\infty \mathbb{P}(|X| \geqslant x)$ diverge, et comme $\mathbb{P}(|X| \geqslant \cdot)$ est bornée par 1 c'est nécessairement au voisinage de $+\infty$ qu'a lieu la divergence, donc $\int_1^\infty \mathbb{P}(|X| \geqslant x)\mathrm{d}x = \infty$.

($\dagger$). Attention aux hypothèses d'intégrabilité dans ce calcul : la fonction (non intégrable) $|y|-|x|$ n'est pas positive, mais elle est néanmoins bornée négativement, ce qui permet de donner un sens à son intégrale contre la mesure de probabilité ν. Pour ceux qui ne sont pas convaincus, un raisonnement plus rigoureux consisterait à écrire $\infty = \int |y|\mathrm{d}\nu(y) \leqslant \int(|x+y|+|x|)\mathrm{d}\nu(y) = \int |x+y|\mathrm{d}\nu(y) + |x|$.

($\ddagger$). Elle affirme par ailleurs que cette constante est alors $\mathbb{E}[X]$, mais cela ne nous intéresse pas ici.

de la loi L^1, c'est parce que la différence d'une variable non intégrable et d'une constante n'est pas intégrable, donc a une norme L^1 infinie).

VI.4.7. D'après la construction de μ, $\mu({}^c[-M,M])$ est la probabilité, pour Y de loi uniforme sur $]0,1/2]$, d'avoir $1/Y\sqrt{|\ln Y|} > M$. Observons que sur $]0,1/2]$, l'application $f : Y \mapsto 1/Y\sqrt{|\ln Y|}$ est strictement décroissante. Pour $M \geqslant e$, posons $Y(M) = 1/M\sqrt{\ln M}$, alors $f(Y(M)) = M\sqrt{\ln M/|\ln Y(M)|} \leqslant M$, par conséquent $\mu({}^c[-M,M]) \leqslant 2Y(M) = o(M^{-1})$ (§).

VI.4.8. Soit $\alpha \in]1/2,1[$ et posons, pour $M > 1$, $Y^*(M) = 1/M(\ln M)^\alpha$. Alors $f(Y^*(M)) = M(\ln M)^\alpha/\sqrt{|\ln Y^*(M)|} \geqslant M$ dès que M est assez grand, et donc $\mu({}^c]-M,M[) \geqslant 2Y^*(M)$. Comme $Y^*(\cdot)$ n'est pas intégrable au voisinage de l'infini, l'absence de moment d'ordre 1 pour μ s'ensuit par le résultat de la question 1 (¶).

VI.4.9. Soit μ ayant un moment d'ordre 1, alors on a $\int_{\mathbb{R}} |x| \mathrm{d}\mu(x) < \infty$, donc par convergence dominée, pour tout $\varepsilon > 0$ on peut trouver $M(\varepsilon)$ tel que $\int_{{}^c[-M(\varepsilon),M(\varepsilon)]} |x|\mathrm{d}\mu(x) \leqslant \varepsilon$. Dans ce cas, pour tout $M \geqslant M(\varepsilon)$ l'inégalité de Markov nous donne :

$$\mu({}^c[-M,M]) \leqslant \frac{\int_{{}^c[-M,M]} |x|\mathrm{d}\mu(x)}{M} \leqslant \varepsilon M^{-1},$$

Ainsi $\mu({}^c[-M,M]) = O(\varepsilon M^{-1})$ au voisinage de l'infini, et comme cela est vrai pour tout ε, $\mu({}^c[-M,M]) = o(M^{-1})$.

VI.4.10. La loi μ est clairement symétrique par rapport à 0, donc sa transformée de Fourier est réelle puisque $\overline{\widehat{\mu}(\xi)} = \overline{\int e^{i\xi x}\mathrm{d}\mu(x)} = \int e^{-i\xi x}\mathrm{d}\mu(x) = \int e^{i\xi(-x)}\mathrm{d}\mu(x) = \int e^{i\xi y}\mathrm{d}\mu(y) = \widehat{\mu}(\xi)$.

VI.4.11. Puisque $\widehat{\mu}$ est réelle, on a que $1 - \widehat{\mu}(\varepsilon) = \mathfrak{Re}\int(1 - e^{i\varepsilon x})\mathrm{d}\mu(x) = 2\int \sin^2(\varepsilon x/2)\mathrm{d}\mu(x)$. Notre but est de borner cette intégrale. Pour cela, introduisons $M(\varepsilon) \geqslant 0$ et découpons l'intégrale au niveau de $-M(\varepsilon)$ et $M(\varepsilon)$. Pour $|x| \leqslant M(\varepsilon)$ on borne la fonction sinus par la fonction carré, et pour $|x| > M(\varepsilon)$ on la borne par 1, ce qui nous donne :

$$\big|1 - \widehat{\mu}(\varepsilon)\big| \leqslant \frac{\varepsilon^2}{2} \int_{-M(\varepsilon)}^{M(\varepsilon)} x^2 \mathrm{d}\mu(x) + 2\mu\big({}^c[-M(\varepsilon), M(\varepsilon)]\big). \tag{4}$$

Un coup de Fubini analogue à celui de la question 1 montre que

$$\int_{-M(\varepsilon)}^{M(\varepsilon)} x^2 \mathrm{d}\mu(x) = \int_0^\infty \mu\big([\pm M(\varepsilon)] \setminus [\pm\sqrt{y}]\big)\mathrm{d}y \leqslant \int_0^{M(\varepsilon)^2} \mu({}^c[-\sqrt{y},\sqrt{y}])\mathrm{d}y. \tag{5}$$

Comme $\mu({}^c[-\sqrt{y},\sqrt{y}])$ est un $o(1/\sqrt{y})$ d'après la question 7, on en déduit (‖), puisque $\int^M y^{-1/2}\mathrm{d}y$ est un $O(\sqrt{y})$ au voisinage de $+\infty$, que (5) est un $o(M(\varepsilon))$ quand $M(\varepsilon)$ tend

(§). Variante plus simple, mais qui ne montre pas le lien entre les questions 7 et 8 : quand $\varepsilon \searrow 0$, $f(\varepsilon) = o(\varepsilon^{-1})$, donc par changement de variables, quand $M \nearrow \infty$, $M = o(1/f^{-1}(M))$, càd. $Mf^{-1}(M) \longrightarrow 0$, donc $f^{-1}(M) = o(M^{-1})$. Ainsi $\mu({}^c[-M,M]) = 2f^{-1}(M) = o(M^{-1})$.

(¶). Variante plus simple, mais qui à nouveau ne montre pas le lien entre les questions 7 et 8 : en utilisant la construction de μ, $\int |x|\mathrm{d}\mu(x) = \int_{-1/2}^{1/2} \frac{\mathrm{d}y}{|y|\sqrt{|\ln|y||}} = \infty$ par les résultats sur les intégrales de Bertrand.

(‖). Cette étape du raisonnement, que je ne détaille pas, est un exercice d'analyse, pas difficile mais pas trivial non plus : dans une situation d'examen, il faudrait la rédiger !

vers l'infini. Ainsi, sous réserve que $M(\varepsilon)$ tende vers l'infini quand ε tend vers 0, (4) donne :

$$\left|1-\widehat{\mu}(\varepsilon)\right| = o\big(\varepsilon^2 M(\varepsilon) + M(\varepsilon)^{-1}\big).$$

Il suffit alors de prendre $M(\varepsilon) = \varepsilon^{-1}$ pour obtenir que, quand $\varepsilon \longrightarrow 0$, $1-\widehat{\mu}(\varepsilon)$ est un $o(\varepsilon)$, ce qui signifie bien que $\widehat{\mu}$ est dérivable en 0 et de dérivée nulle.

VI.4.12. Si nous notons ν_N la loi de $\widehat{E}_N$, la défintion de $\widehat{E}_N$ nous permet d'exprimer la fonction caractéristique de ν_N à partir de celle de μ :

$$\widehat{\nu}_N(\xi) = \widehat{\mu}(\xi/N)^N.$$

Puisque $\widehat{\mu}(\xi) = 1 + o(\xi)$ quand ξ tend vers 0 d'après la question 11, on en déduit que pour tout $\xi \in \mathbb{R}$, $\widehat{\nu}_N(\xi)$ tend vers 1 quand N tend vers l'infini. La fonction identiquement égale à 1 étant la transformée de Fourier de δ_0, cela signifie par le théorème de Lévy qu'il y a convergence en loi des $\widehat{E}_N$ vers la constante 0.

VI.4.13. Quand on converge vers une constante, la convergence en loi est équivalente à la convergence en probabilité (c'est un point de cours, voir aussi l'exercice V.9), donc la question 12 implique le résultat.

VI.4.14. Puisqu'il y a convergence en probabilité vers 0, cela signifie que les $\widehat{E}_N$ passent l'essentiel de leur temps près de 0. Puisqu'il n'y a pas convergence presque-sûre, cela signifie néanmoins que pour chaque réalisation de l'expérience aléatoire, $\widehat{E}_N$ va continuer de temps en temps à s'éloigner de 0 d'une quantité notable, mais de moins en moins souvent quand même. Enfin, puisque la loi de $\widehat{E}_N$ n'est pas intégrable, cela veut dire que les écarts des $\widehat{E}_N$ à 0 seront de temps en temps très grands.

Globalement, on peut résumer une réalisation typique de la suite des $\widehat{E}_N$ ainsi : la suite se promène pas très loin de 0, et puis de temps en temps on tire un X_i tellement grand que $\widehat{E}_N$ s'éloigne de 0, parfois même très très loin. Ensuite, les X_i suivants qu'on tire étant tout petits par rapport au X_i « catastrophique » qui a déclenché l'éloignement de 0, la suite $\widehat{E}_N$ « relaxe » vers 0 à la vitesse $1/N$. Et là il se passe très longtemps avant qu'une autre « catastrophe » se produise, tellement longtemps même qu'entretemps les $\widehat{E}_N$ se sont rapprochés de 0 d'encore plus près qu'avant.