Probabilités : Exercices (corrigé)

M1 Mathématiques Fondamentales, Université Paris-Saclay, 2017–2018

Espaces de probabilité, variables aléatoires

Exercice 1. 1. $\Omega = \{1, \dots, 20\}$, $\mathrm{Card}(\Omega) = 20$, $\mathbb{P}(\omega) = 1/20\ \forall \omega \in \Omega$.

2. $\Omega = \{0,1\}^n$, où 0 : “pile”, 1 : “face” (par exemple). $\mathrm{Card}(\Omega) = 2^n$, $\mathbb{P}(\omega) = 2^{-n}\ \forall \omega \in \Omega$.
3. On a $\Omega = \mathbb{P}(S)$, et $\mathrm{Card}(\Omega) = 2^n$. Quant à $\mathbb{P}$, on n'en sait rien *a priori.*
4. On demande un espace de probabilité fini. Si on s'intéresse seulement à qui est le vainqueur, alors on pose $\Omega = \{1, 2, 0\}$, ou 1 : la première équipe gagne, 2 : la deuxième équipe gagne, 0 : match nul. Si on veut incorporer le score, alors le plus naturel est un espace infini, à savoir $\Omega = \mathbb{N}^2$. On pourrait aussi y ajouter plus d'informations, tel que $\Omega' = \mathbb{N}^2 \times \{R, P, T\}$, où R : “après temps réglementaire”, S : “après temps supplémentaire”, T : “après tir au but”. Dans tous les cas, on a (nécessairement) $\mathrm{Card}(\Omega) = \infty$. Quant à $\mathbb{P}$, on n'en sait rien *a priori.*
5. Les deux candidats sont Hollande (H) et Sarkozy (S), mais une personne peut aussi bien répondre “je ne sais pas” (N1), “je ne voterai pas” (N2), “je voterai blanc” (N3), où bien ne pas répondre (N0). Un bon sondage prend toutes ces réponses en compte (voir plus...). Un espace de probabilité possible serait $\Omega = \{H, S, N0, N1, N2, N3\}^{1047}$, avec $\mathrm{Card}(\Omega) = 6^{1047}$. Une autre possibilité serait de compter seulement le nombre de répondants pour chaque réponse possible, cela donne $\Omega' = \{0, 1, \dots, 1047\}^6$ avec $\mathrm{Card}(\Omega') = 1048^6 \ll 6^{1047}$. Cependant, le premier espace est plus commode, car il permet d'intégrer facilement l'indépendance entre les répondants (en supposant que ceux-ci ont été tirés avec remise dans la population totale). En effet, si la probabilité d'une réponse $r \in \{H, S, N0, N1, N2, N3\}$ est donnée par p_r, alors on pose $\mathbb{P}(r_1, \dots, r_{1047}) = p_{r_1} \cdots p_{r_{1047}}$.

Exercice 2. 1. Faux. Si B_r désigne la boule ouverte de rayon r autour de l'origine, alors $\cap_n B_{1/n} = \{0\}$, ce qui n'est pas un ouvert.

2. Vrai. On a
   (a) $\{0,1\}^{\mathbb{N}} = \{0,1\}^k \times \{0,1\}^{\mathbb{N}}$, $\emptyset = \emptyset \times \{0,1\}^{\mathbb{N}}$
   (b) $\{0,1\}^{\mathbb{N}} \backslash A \times \{0,1\}^{\mathbb{N}} = (\{0,1\}^k \backslash A) \times \{0,1\}^{\mathbb{N}}$
   (c) Pour $A_1, A_2, \dots \subset \{0,1\}^k$, $\bigcup_n (A_n \times \{0,1\}^{\mathbb{N}}) = (\bigcup_n A_n) \times \{0,1\}^{\mathbb{N}}$.
3. Faux (une union dénombrable croissante de tribus n'est pas toujours une tribu) Exemple : $\bigcap_n \{0\}^n \times \{0,1\}^{\mathbb{N}} = \{0\}^{\mathbb{N}}$ n'est pas de cette forme.
4. Vrai (tribu *terminale*) : Comme dans 2., on montre que pour tout n, $\{\{0,1\}^n \times A : A \subset \{0,1\}^{\mathbb{N}} \text{ mesurable}\}$ est une tribu. Une intersection de tribus est concore une tribu.
5. Faux. Par exemple, si A est une partie de $\mathbb{N}$ qui n'admet pas de densité (par exemple, $k \in A$ ssi $\lfloor \log_2 k \rfloor$ est paire), alors pour tout n, $A \cap \{0, \dots, n\}$ admet une densité (nulle), mais $A = \bigcup_n (A \cap \{0, \dots, n\})$ n'en admet pas.

Exercice 3. 1. Puisque μ est une mesure, on a,

— $\mu_A(\emptyset) = \mu(\emptyset \cap A) = \mu(\emptyset) = 0$, et

— pour tous $A_1, A_2, \dots \in \mathcal{A}$ disjoints,

$$\mu_A\Big(\bigcup_n A_n\Big) = \mu\Big(\Big(\bigcup_n A_n\Big) \cap A\Big) = \mu\Big(\bigcup_n (A_n \cap A)\Big) = \sum_n \mu(A_n \cap A) = \sum_n \mu_A(A_n).$$

Par conséquent, μ_A est également une mesure.

2. Par le point précédent, l'application $\mathbb{P}(A)\mathbb{P}(\cdot \mid A) = \mathbb{P}(\cdot \cap A)$ est une mesure, donc $\mathbb{P}(\cdot \mid A)$ l'est aussi car $\mathbb{P}(A) > 0$. De plus, on a $\mathbb{P}(\Omega \mid A) = \mathbb{P}(A)/\mathbb{P}(A) = 1$. Par conséquent, $\mathbb{P}$ est une mesure de probabilité.

Exercice 4. Première solution : Définissons notre espace de probabilité $\Omega = \{FF, FM, MF, MM\}$, où la première lettre donne le sexe de l'ainé(e) et la deuxième celui du cadet/de la cadette. On suppose que ces quatre possibilités sont équiprobables, ce qui est à notre connaissance assez proche de la réalité. L'événement "au moins une fille" est alors $A = \{FF, FM, MF\}$. Sachant cet événement, la probabilité que le deuxième enfant est une fille est alors 1/3, contre 2/3 pour un garçon. Formellement, on calcule

$$\mathbb{P}(\text{les deux enfants sont des filles} \mid A) = \mathbb{P}(FF)/\mathbb{P}(A) = 1/3.$$

Deuxième solution : La dernière solution peut ne pas être satisfaisante, car on pourrait se dire que s'il y a deux filles, la chance est plus grande de trouver des chaussures de filles devant la porte, donc nous avons oublié des corrélations importantes. Pour formaliser cela, on élargit l'espace de probabilité : on pose $\Omega = \{F, M\}^2 \times \{0, 1\}^2$, où le i-ième chiffre est 1 si le i-ième enfant a laissé ses chaussures devant la porte, et 0 sinon. L'événement A qu'il y a (exactement) une paire de chaussures de filles devant la porte est alors

$$A = \{FM10, MF01, FF10, FF01\}.$$

En ce qui concerne la probabilité sur l'espace Ω, il y a beaucoup de choix possibles. Nous n'allons pas nous restreindre à un choix précis, mais nous allons simplement supposer que la probabilité de laisser ou non ses chaussures devant la porte ne dépend pas du sexe, en particulier, nous supposons que les *éléments de l'événement A sont équiprobables*. Par conséquent, si $B = \{FF\} \times \{0, 1\}^2$ désigne l'événement que les deux enfants sont des filles, on a $\mathbb{P}(B \mid A) = 1/2$. En conclusion, la probabilité que le deuxième enfant est une fille vaut 1/2, contrairement à 1/3 dans la première solution.

Exercice 5. 1. On a pour tout $k \in \{0, \dots, 2n\}$,

$$\begin{aligned}
\mathbb{P}(X + Y = k) &= \sum_{i=0}^{k} \mathbb{P}(X = i,\, Y = k - i) \\
&= \frac{1}{(n+1)^2} \sum_{i=0}^{k} \mathbf{1}_{(i \le n,\, k-i \le n)} \\
&= \frac{1}{(n+1)^2} \begin{cases} \sum_{i=0}^{k} 1, & k \le n \\ \sum_{i=k-n}^{n} 1 & k \ge n, \end{cases} \\
&= \frac{1}{(n+1)^2} \begin{cases} k+1, & k \le n \\ 2n - k + 1 & k \ge n, \end{cases}
\end{aligned}$$

si bien que

$$\mathbb{P}(X + Y = k) = \frac{\min(k+1, 2n-k+1)}{(n+1)^2} = \frac{n+1-|k-n|}{(n+1)^2}$$

et $\mathbb{P}(X + Y = k) = 0$ pour tout $k \notin \{0, \dots, 2n\}$. Pour la loi de $X - Y$, on remarque que $(X, n - Y)$ a même loi que (X, Y), si bien que pour $k \in \{-n, \dots, n\}$,

$$\begin{aligned}
\mathbb{P}(X - Y = k) &= \mathbb{P}(X + (n - Y) = n + k) \\
&= \mathbb{P}(X + Y = n + k) \\
&= \frac{\min(n+k+1, n-k+1)}{(n+1)^2} = \frac{n+1-|k|}{(n+1)^2}
\end{aligned}$$

et $\mathbb{P}(X + Y = k) = 0$ pour tout $k \notin \{-n, \dots, n\}$.

2. Première façon. Par linéarité de l'espérance,

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y] = 2\mathbb{E}[X].$$

Or, l'espérance de X vaut

$$\mathbb{E}[X] = \frac{1}{n+1} \sum_{k=0}^{n} k = \frac{n}{2},$$

si bien que $\mathbb{E}[X+Y] = n$ (*un calcul plus astucieux est le suivant : puisque X est de même loi que $n-X$, on a* $2\mathbb{E}[X] = \mathbb{E}[X] + \mathbb{E}[n-X] = \mathbb{E}[X+n-X] = \mathbb{E}[n] = n$).

Deuxième façon. On passe par la loi de $X+Y$: par la première partie de l'exercice,

$$\mathbb{E}[X+Y] = \sum_{k=0}^{2n} k\frac{\min(k+1, 2n-k+1)}{(n+1)^2}.$$

On peut calculer cela de manière un peu fastidieuse, mais on peut aussi remarquer comme ci-dessus que $X+Y$ est égale en loi à $2n-(X+Y)$, si bien que

$$\mathbb{E}[X+Y] = \frac{1}{2}\mathbb{E}[X+Y+2n-(X+Y)] = \frac{1}{2}\mathbb{E}[2n] = n.$$

Exercice 6 (Perte de mémoire)**.** Soit X une v.a. à valeurs dans $\mathbb{N}$. Soit $n, m \in \mathbb{N}$. Alors,

$$\mathbb{P}(X \geq n+m \,|\, X \geq n) = \frac{\mathbb{P}(X \geq n+m)}{\mathbb{P}(X \geq n)},$$

et donc

$$\mathbb{P}(X \geq n+m \,|\, X \geq n) = \mathbb{P}(X \geq m) \iff \mathbb{P}(X \geq n+m) = \mathbb{P}(X \geq m)\mathbb{P}(X \geq n).$$

Posant $m = 1$, on voit que ceci implique que $\mathbb{P}(X \geq n+1) = \mathbb{P}(X \geq 1)\mathbb{P}(X \geq n)$ pour tout $n \in \mathbb{N}$, et donc $\mathbb{P}(X \geq n) = q^n$ pour un certain $q \in [0,1]$. Puisque $\lim_{n\to\infty} \mathbb{P}(X \geq n) = 0$, on a en fait $q \in [0,1)$.

Une condition nécessaire est alors que $\mathbb{P}(X \geq n) = q^n$ pour un certain $q \in [0,1)$. Cette condition est également suffisante, car on a alors $\mathbb{P}(X \geq n+m) = q^{n+m} = q^n q^m = \mathbb{P}(X \geq n)\mathbb{P}(X \geq m)$. Les lois cherchées sont alors exactement les lois géométriques $\mathcal{G}(p)$ pour $p \in (0,1]$.

Exercice 7 (Perte de mémoire 2)**.** Soit X une v.a. à valeurs dans $\mathbb{R}_+$ vérifiant les hypothèses dans l'énoncé. Pour $k \in \mathbb{N}$ posons $X_k = \lfloor 2^k X \rfloor$. On a alors pour tout $n, m \in \mathbb{N}$,

$$\mathbb{P}(X_k \geq n+m \,|\, X_k \geq n) = \mathbb{P}(X \geq 2^{-k}n + 2^{-k}m \,|\, X \geq 2^{-k}n) = \mathbb{P}(X \geq 2^{-k}m) = \mathbb{P}(X_k \geq m).$$

Par l'Exercice 1, X_k suit alors une loi géométrique de paramètre $p_k = 1 - q_k \in (0,1]$. On a alors pour tout $n \in \mathbb{N}$,

$$\mathbb{P}(X \geq 2^{-k}n) = \mathbb{P}(X_k \geq n) = q_k^n,$$

si bien que pour tout $k \in \mathbb{N}$,

$$q_k = \mathbb{P}(X \geq 2^{-k}) = \mathbb{P}(X \geq 2^{-(k+1)} \times 2) = q_{k+1}^2,$$

et donc $q_k = q^{2^{-k}}$ pour tout k, avec $q = q_0 \in [0,1)$. Par conséquent,

$$\forall k, n \in \mathbb{N} : \mathbb{P}(X \geq 2^{-k}n) = q^{2^{-k}n}.$$

Il vient que la fonction $x \mapsto \mathbb{P}(X \geq x) - q^x$ est continue à gauche en $x > 0$ et s'annule en les nombres dyadiques qui forment un ensemble dense dans $\mathbb{R}_+$. On en déduit que $\mathbb{P}(X \geq x) = q^x$ pour tout $x > 0$. De plus, cette égalité est trivialement vraie pour $x = 0$. Par conséquent, X suit la loi exponentielle de paramètre $\lambda = -\log q \in (0,\infty]$ (avec la loi exponentielle de paramètre ∞ la mesure δ_0). Ceci est alors une condition nécessaire. Cette condition est également suffisante car si $\mathbb{P}(X \geq x) = q^x$ pour un $q \in [0,1)$, alors $\mathbb{P}(X \geq x+y) = \mathbb{P}(X \geq x)\mathbb{P}(X \geq y)$, et donc la loi de X satisfait à l'hypothèse.

Exercice 8. Puisque f est positive et dérivable et $X \geq 0$, on a $f(X) = \int_0^X f'(x)\,dx = \int_0^\infty f'(x)\mathbf{1}_{x\leq X}\,dx$. De plus, puisque f est croissante, sa dérivée $f'(x) \geq 0$ pour tout $x \geq 0$. Par le théorème de Fubini–Tonelli,

$$\mathbb{E}[f(X)] = \mathbb{E}[\int_0^\infty f'(x)\mathbf{1}_{x\leq X}\,dx] = \int_0^\infty \mathbb{E}[f'(x)\mathbf{1}_{x\leq X}]\,dx = \int_0^\infty f'(x)\mathbb{P}(X \geq x)\,dx.$$

Exercice 9 (Formule du crible)**.** 1. Soient $A_1, \ldots, A_n \in \mathcal{A}$. Alors,

$$\mathbf{1}_{A_1\cup\ldots\cup A_n} = 1 - \mathbf{1}_{A_1^c\cap\ldots\cap A_n^c} = 1 - \prod_{i=1}^n \mathbf{1}_{A_i^c} = 1 - \prod_{i=1}^n (1 - \mathbf{1}_{A_i}).$$

2. Soient $A_1, \ldots, A_n \in \mathcal{A}$. Pour tout $l, n \in \mathbb{N}$, notons $\mathcal{P}_{l,n} = \{I \subset \{1, \ldots, n\} : \mathrm{Card}(I) = l\}$ et pour tout $I \subset \{1, \ldots, n\}$, $A_I = \bigcap_{i \in I} A_i$. Par la formule de la première partie de l'exercice, on a

$$\begin{aligned}
\mathbb{P}(A_1 \cup \ldots \cup A_n) &= \mathbb{E}[\mathbf{1}_{A_1 \cup \ldots \cup A_n}] \\
&= 1 - \mathbb{E}\Big[\prod_{i=1}^n (1 - \mathbf{1}_{A_i})\Big] \\
&= 1 - \mathbb{E}\Big[\sum_{l=0}^n (-1)^l \sum_{I \in \mathcal{P}_{l,n}} \mathbf{1}_{A_I}\Big] \\
&= \mathbb{E}\Big[\sum_{l=1}^n (-1)^{l-1} \sum_{I \in \mathcal{P}_{l,n}} \mathbf{1}_{A_I}\Big] \\
&= \sum_{l=1}^n (-1)^{l-1} \sum_{I \in \mathcal{P}_{l,n}} \mathbb{E}[\mathbf{1}_{A_I}] \\
&= \sum_{l=1}^n (-1)^{l-1} \sum_{I \in \mathcal{P}_{l,n}} \mathbb{P}(A_I).
\end{aligned}$$

Ceci est la formule du crible.

Exercice 10. On calcule d'abord la probabilité que Pascal aurait gagné le jeu (elle est plus facile à calculer que celle que Fermat l'aurait gagné). Au maximum, 4 lancers de pièce sont nécessaires, donc on se met sur l'éspace de probabilité $\Omega = \{P, F\}^4$, ou le i-ième P/F désigne que Pascal/Fermat gagne le i-ième lancer en question. L'événement A que Pascal gagne est alors $A = \{PPPP, PPPF, PPFP, PFPP, FPPP\}$, et sa probabilité est alors $\mathbb{P}(A) = 5/16 = 31.25\% < 46.67\% = 7/15$. On comprend alors que chacun a proposé ce qui l'arrange le plus.

Néanmoins, on peut argumenter que la proposition de Fermat est la plus juste. La continuation du jeu aurait été équivalente à une expérience de Bernoulli qui donne Pascal/Fermat gagnant avec probabilité 31.25%/68.75%. De distribuer ce pourcentage du montant total correspond à donner à chacun *l'espérance* de son gain. Ceci est communément interprété comme une solution "non biaisée" est donc la "plus juste". Notons aussi que la proposition de Pascal n'a peu à voir avec les mécanismes du jeu et donc avec ce qui a été conclu au préalable.

Exercice 11. Soit (X, Y) un vecteur aléatoire uniforme sur $[0, 1]^2$. On a alors pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ mesurable bornée,

$$\begin{aligned}
\mathbb{E}[f(X+Y)] &= \iint_{[0,1]^2} f(x+y)\, dxdy \\
&= \int_0^1 dx \int_x^{1+x} f(x+y-x)\, dy && \text{chgt de var. } y \mapsto y - x \\
&= \int_0^1 dx \int_0^2 \mathbf{1}_{(x \le y,\, 1+x \ge y)} f(y)\, dy \\
&= \int_0^2 dy\, f(y) \int_0^1 \mathbf{1}_{(x \le y,\, x \ge y-1)}\, dx && \text{thm de Fubini} \\
&= \int_0^2 f(y) \min(y, 2-y)\, dy.
\end{aligned}$$

La loi de $X + Y$ est alors celle de densité $f_{X+Y}(y) = \min(y, 2-y) = 1 - |1-y|$ sur $[0, 2]$. Puisque $X - Y$ est égale en loi à $X + Y - 1$, cette loi est alors celle de densité $f_{X-Y}(y) = f_{X+Y}(1+y) = 1 - |y|$ sur $[-1, 1]$.

Exercice 12 (Loi gamma)**.** 1. On a

$$\int_0^\infty x^{\alpha-1} e^{-\beta x}\, dx \overset{\beta x = y}{=} \beta^{-\alpha} \int_0^\infty y^{\alpha-1} e^{-y}\, dy = \frac{\Gamma(\alpha)}{\beta^\alpha}.$$

On a alors $c_{\alpha,\beta} = \frac{\beta^\alpha}{\Gamma(\alpha)}$.

2. La loi exponentielle de paramètre λ est la loi sur $[0, \infty)$ de densité $\lambda e^{-\lambda x}$. C'est donc la loi $\Gamma(1, \lambda)$.

3. Soit $N \sim \mathcal{N}(0,1)$. On a pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ mesurable bornée,
$$\mathbb{E}[f(N^2)] \propto \int_{\mathbb{R}} f(x^2) e^{-x^2/2}\, dx \overset{x^2=y}{\propto} \int_0^\infty f(y) e^{-y/2} \frac{1}{\sqrt{y}}\, dy,$$
donc la loi de N^2 est celle de densité proportionnelle à $y^{-1/2}e^{-y/2}$ sur $[0,\infty)$. Il s'agit donc de la loi $\Gamma(1/2, 1/2)$.
4. Soit $G \sim \Gamma(\alpha,\beta)$ et $b > 0$. On a alors pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ mesurable bornée,
$$\mathbb{E}[f(bG)] \propto \int_{\mathbb{R}} f(bx) x^{\alpha-1} e^{-\beta x}\, dx \overset{bx=y}{\propto} \int_{\mathbb{R}} f(y) y^{\alpha-1} e^{-(\beta/b)y}\, dy.$$
Ceci montre que $bG \sim \Gamma(\alpha, \beta/b)$.
5. La loi de $G_1 + G_2$ est la loi sur $\mathbb{R}$ de densité la convolué de celles de G_1 et G_2, c'est à dire
$$\begin{aligned}
f_{G_1+G_2}(x) &= \int_{\mathbb{R}} f_{G_1}(y) f_{G_2}(x-y)\, dy \\
&\propto \int_{\mathbb{R}} \mathbf{1}_{y\ge 0} y^{\alpha_1-1} e^{-\beta y} \mathbf{1}_{x-y\ge 0} (x-y)^{\alpha_2-1} e^{-\beta(x-y)}\, dy \\
&= e^{-\beta x} \int_0^x y^{\alpha_1-1}(x-y)^{\alpha_2-1}\, dy \\
&= e^{-\beta x} x^{\alpha_1+\alpha_2-1} \int_0^1 y^{\alpha_1-1}(1-y)^{\alpha_2-1}\, dy && y \mapsto xy, \\
&\propto e^{-\beta x} x^{\alpha_1+\alpha_2-1}.
\end{aligned}$$
Ceci montre que $G_1 + G_2 \sim \Gamma(\alpha_1+\alpha_2, \beta)$.
6. Par la partie précédente de l'exercice, il vient par récurrence que $\sum_{i=1}^n G_i \sim \Gamma(\sum_{i=1}^n \alpha_i, \beta)$.
7. Soit $G \sim \Gamma(\alpha,\beta)$, $\alpha,\beta > 0$. D'après la partie 4, on a $\beta G \sim \Gamma(\alpha, 1)$, si bien que pour tout $n \in \mathbb{N}$,
$$\mathbb{E}[(\beta G)^n] = \frac{1}{\Gamma(\alpha)} \int_0^\infty x^{\alpha-1} x^n e^{-x}\, dx = \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)} = \alpha^{(n)},$$
où $\alpha^{(n)} = \alpha \cdots (\alpha+n-1)$ (la fonction de *Pochhammer*). Ceci donne $\mathbb{E}[G^n] = \beta^{-n}\alpha^{(n)}$.
Si $N \sim \mathcal{N}(0,1)$, alors $N^2 \sim \Gamma(1/2,1/2)$ d'après la partie 3. On a alors pour tout $n \in \mathbb{N}$,
$$\mathbb{E}[N^{2n}] = \mathbb{E}[(N^2)^n] = 2^n (1/2)^{(n)} = 1 \cdots (2n-3)(2n-1).$$
De plus, puisque N et $-N$ ont même loi, on a pour tout n impaire, $\mathbb{E}[N^n] = \mathbb{E}[(-N)^n] = -\mathbb{E}[N^n]$, si bien que $\mathbb{E}[N^n] = 0$.
8. Soit $G \sim \Gamma(\alpha,\beta)$, $\alpha,\beta > 0$. On a alors pour tout $\lambda \in \mathbb{R}$,
$$\varphi_G(\lambda) = \mathbb{E}[e^{i\lambda G}] = c_{\alpha,\beta} \int_0^\infty e^{i\lambda x} x^{\alpha-1} e^{-\beta x}\, dx.$$
Comme pour la loi normale, on dérive en λ et on intègre par parties :
$$\begin{aligned}
\varphi_G'(\lambda) &= c_{\alpha,\beta} i \int_0^\infty e^{i\lambda x} x^{\alpha} e^{-\beta x}\, dx \\
&= c_{\alpha,\beta} i \frac{\alpha}{\beta - i\lambda} \int_0^\infty e^{i\lambda x} x^{\alpha-1} e^{-\beta x}\, dx \\
&= -\frac{\alpha}{\lambda + \beta i} \varphi_G(\lambda).
\end{aligned}$$
Ceci donne pour une constante $C \in \mathbb{C}$,
$$\varphi_G(\lambda) = C(\lambda + \beta i)^{-\alpha},$$
et puisque $\varphi_G(0) = 1$,
$$\varphi_G(\lambda) = \left(\frac{\beta i}{\lambda + \beta i}\right)^\alpha = \left(\frac{\beta}{\beta - i\lambda}\right)^\alpha$$
Montrons les propriétés établies dans les parties 4, 6 et 7 :

— Si $b > 0$, on a $\varphi_{bG}(\lambda) = \varphi_G(b\lambda) = \left(\frac{\beta}{\beta - ib\lambda}\right)^\alpha = \left(\frac{\beta/b}{\beta/b - i\lambda}\right)^\alpha$, et donc $bG \sim \Gamma(\alpha, \beta/b)$.

— Si $G_1, \ldots, G_n$ sont des v.a. indépendantes de lois respectives $\Gamma(\alpha_i, \beta)$, avec $\alpha_1, \ldots, \alpha_n, \beta > 0$, alors

$$\varphi_{G_1+\cdots+G_n}(\lambda) = \varphi_{G_1}(\lambda)\cdots\varphi_{G_n}(\lambda) = \prod_{i=1}^n \left(\frac{\beta}{\beta - i\lambda}\right)^{\alpha_i} = \left(\frac{\beta}{\beta - i\lambda}\right)^{\alpha_1+\cdots+\alpha_n}.$$

Donc, $G_1 + \cdots + G_n \sim \Gamma(\alpha_1 + \cdots + \alpha_n, \beta)$.

— On a pour tout $n \in \mathbb{N}$,

$$\mathbb{E}[G^n] = i^{-n}\frac{d^n}{d\lambda^n}\varphi_G(\lambda)\Big|_{\lambda=0} = \frac{\beta^\alpha \alpha(\alpha+1)\cdots(\alpha+n-1)}{(\beta - i\lambda)^{\alpha+n}}\Big|_{\lambda=0} = \beta^{-n}\alpha^{(n)},$$

avec $\alpha^{(n)} = \alpha(\alpha+1)\cdots(\alpha+n-1)$.

Exercice 13 (Méthode de premier et second moment)**.** 1. Puisque X prend ses valeurs dans $\mathbb{N}$, on a par l'inégalité de Markov,

$$\mathbb{P}(X \neq 0) = \mathbb{P}(X \geq 1) \leq \frac{\mathbb{E}X}{1} = \mathbb{E}X.$$

2. On note que $X = 0$ implique $|X - \mathbb{E}X| \geq \mathbb{E}X$. Par l'inégalité de Chebychev, on a alors

$$\mathbb{P}(X = 0) \leq \mathbb{P}(|X - \mathbb{E}X| \geq \mathbb{E}X) \leq \frac{\mathrm{Var}(X)}{(\mathbb{E}X)^2},$$

si bien que

$$\mathbb{P}(X \neq 0) = 1 - \mathbb{P}(X = 0) \geq 1 - \frac{\mathrm{Var}(X)}{(\mathbb{E}X)^2}.$$

3. Par l'inégalité de Cauchy–Schwarz, on a

$$\mathbb{E}X = \mathbb{E}[X\mathbf{1}_{X\neq 0}] \leq \sqrt{\mathbb{E}[X^2]\mathbb{E}[(\mathbf{1}_{X\neq 0})^2]} = \sqrt{\mathbb{E}[X^2]\mathbb{P}(X \neq 0)}.$$

En élevant les deux cotés de l'équation au carré, puis divisant par $\mathbb{E}[X^2]$, on obtient l'inégalité souhaitée :

$$\mathbb{P}(X \neq 0) \geq \frac{(\mathbb{E}X)^2}{\mathbb{E}[X^2]}.$$

4. On observe les équivalences suivantes :

$$\begin{aligned} & 1 - \frac{\mathrm{Var}(X)}{(\mathbb{E}X)^2} \leq \frac{(\mathbb{E}X)^2}{\mathbb{E}[X^2]} \\ \iff & -\frac{\mathrm{Var}(X)}{(\mathbb{E}X)^2} \leq \frac{(\mathbb{E}X)^2 - \mathbb{E}[X^2]}{\mathbb{E}[X^2]} \\ \iff & \frac{\mathrm{Var}(X)}{(\mathbb{E}X)^2} \geq \frac{\mathrm{Var}(X)}{\mathbb{E}[X^2]}. \end{aligned}$$

Or, puisque $\mathbb{E}[X^2] = (\mathbb{E}X)^2 + \mathrm{Var}(X) \geq (\mathbb{E}X)^2$, la dernière inégalité est toujours vérifiée. Ceci montre que $1 - \frac{\mathrm{Var}(X)}{(\mathbb{E}X)^2} \leq \frac{(\mathbb{E}X)^2}{\mathbb{E}[X^2]}$, ce qu'il fallait démontrer.

Exercice 14 (Méthode de premier et second moment : application)**.** 1. Si $I_{k,n} \geq 1$, alors il existe un $i \in \{0, \ldots, n-k\}$, tel que $X_{i+1} = \cdots = X_{i+k} = 1$, autrement dit, la suite $X_1, \ldots, X_n$ contient k uns consécutifs. Réciproquement, si la suite $X_1, \ldots, X_n$ contient k uns consécutifs, alors soit i le plus petit indice dans $\{0, \ldots, n-k\}$ tel que $X_{i+1} = \cdots = X_{i+k}$. Si $i > 0$, on a alors $X_i = 0$, sinon $i - 1$ serait également un tel indice, ce qui serait en contradiction avec le fait que i en est le plus petit. De plus, si $i = 0$, alors $X_i = 0$ par définition. On a alors $I_{k,n} \geq 1$.

2. Par linéarité de l'espérance, on a

$$\mathbb{E}[I_{k,n}] = \sum_{i=0}^{n-k} \mathbb{E}[\mathbf{1}_{B_i}] = \sum_{i=0}^{n-k} \mathbb{P}(B_i).$$

Notons que

$$\begin{aligned}\mathbb{P}(B_0) = \mathbb{P}(X_1 = \cdots = X_k) = 2^{-k},\\ \forall i \in \{1,\dots,n-k\} : \mathbb{P}(B_i) = \mathbb{P}(X_i = 0,\ X_{i+1} = \cdots = X_{i+k} = 1) = 2^{-(k+1)}.\end{aligned}$$

Il vient que

$$\begin{aligned}\mathbb{E}[I_{k,n}] &= 2^{-k} + (n-k)2^{-(k+1)}\\ &= (n-k+2)2^{-k-1}.\end{aligned}$$

3.

$$\begin{aligned}\mathbb{E}[I_{k,n}^2] &= \mathbb{E}\left[\sum_{0\le i,j\le n-k} \mathbf{1}_{B_i}\mathbf{1}_{B_j}\right]\\ &= \sum_{0\le i,j\le n-k} \mathbb{P}(B_i\cap B_j)\\ &= \sum_{0\le i\le n-k} \mathbb{P}(B_i) + \sum_{0\le i,j\le n-k,\ i\neq j} \mathbb{P}(B_i\cap B_j).\end{aligned}$$

Pour calculer $\mathbb{P}(B_i\cap B_j)$ pour $i,j\in\{0,\dots,n-k\}$, $i\neq j$, on distingue entre deux cas : si $1\le |i-j|\le k$, alors $B_i\cap B_j=\emptyset$, tandis que si $|i-j|>k$, alors B_i et B_j sont indépendants. Dans tous les cas, il vient que

$$\mathbb{P}(B_i\cap B_j)\le \mathbb{P}(B_i)\mathbb{P}(B_j).$$

Il en suit la borne suivante sur $\mathbb{E}[I_{k,n}^2]$:

$$\begin{aligned}\mathbb{E}[I_{k,n}^2] &\le \sum_{0\le i\le n-k} \mathbb{P}(B_i) + \sum_{0\le i,j\le n-k,\ i\neq j} \mathbb{P}(B_i)\mathbb{P}(B_j)\\ &\le \sum_{0\le i\le n-k} \mathbb{P}(B_i) + \left(\sum_{0\le i\le n-k} \mathbb{P}(B_i)\right)^2\\ &= \mathbb{E}[I_{k,n}] + (\mathbb{E}[I_{k,n}])^2.\end{aligned}$$

4. Par l'Exercice 14 de cette feuille et l'égalité des événemnents $A_{k,n}=\{I_{k,n}\neq 0\}=\{I_{k,n}\ge 1\}$, on a

$$\frac{(\mathbb{E}I_{k,n})^2}{\mathbb{E}[I_{k,n}^2]}\le \mathbb{P}(A_{k,n})\le \mathbb{E}I_{k,n}.$$

La borne supérieure découle alors directement de l'expression de $\mathbb{E}I_{k,n}$ obtenue dans la partie 2. De plus, par la partie 3, on a

$$\frac{(\mathbb{E}I_{k,n})^2}{\mathbb{E}[I_{k,n}^2]}\ge \frac{(\mathbb{E}I_{k,n})^2}{\mathbb{E}[I_{k,n}]+(\mathbb{E}[I_{k,n}])^2} = \frac{\mathbb{E}I_{k,n}}{1+\mathbb{E}[I_{k,n}]},$$

si bien que la borne inférieure découle encore de l'expression de $\mathbb{E}I_{k,n}$ de la partie 2.

5.

k	1	2	3	4	5	6	7	8	9	10
borne sup	25.2500	12.5000	6.1875	3.0625	1.5156	0.7500	0.3711	0.1836	0.0908	0.0449
borne inf	0.9619	0.9259	0.8609	0.7538	0.6025	0.4286	0.2707	0.1551	0.0833	0.0430

Exercice 15 (Borne de Chernoff)**.** 1. On a pour tout $\lambda\ge 0$, $\{X\ge x\}=\{e^{\lambda X}\ge e^{\lambda x}\}$, si bien que par l'inégalité de Markov,

$$\mathbb{P}(X\ge x)=\mathbb{P}(e^{\lambda X}\ge e^{\lambda x})\le \mathbb{E}[e^{\lambda X}]e^{-\lambda x}=e^{-(\lambda x-\varphi(\lambda))}.$$

En minimisant sur λ, on a

$$\mathbb{P}(X\ge x)\le \inf_{\lambda\ge 0} e^{-(\lambda x-\varphi(\lambda))}=\exp(-\sup_{\lambda\ge 0}[\lambda x-\varphi(\lambda)])=e^{-I(x)}.$$

2. Soit $\varphi_n(\lambda) = \log \mathbb{E}[e^{\lambda(X_1+\ldots+X_n)}]$ et $I_n(x) = \sup_{\lambda\geq 0}[\lambda x - \varphi_n(\lambda)]$. Par indépendance,

$$\varphi_n(\lambda) = \log \mathbb{E}[e^{\lambda(X_1+\ldots+X_n)}] = \log \mathbb{E}[e^{\lambda X}]^n = n\varphi(\lambda).$$

Par conséquent, on a pour $a \in \mathbb{R}$,

$$I_n(an) = \sup_{\lambda\geq 0}[\lambda an - n\varphi(\lambda)] = nI(a).$$

L'inégalité découle alors de la borne de Chernoff de la deuxième partie.

Indépendance de tribus, fonctions génératrices, sommes aléatoires, vecteurs gaussiens

Exercice 16. On définit les classes suivantes

$$\mathcal{C} = \{\emptyset, \{x = 0\}, \{x > 0\}\}, \quad \mathcal{D} = \{\emptyset, \{|x| = 0\}\} \cup \{\{|x| \in B\};\ B \subset]0, \infty[\text{ mesurable}\}.$$

Alors ces classes sont stables par intersection et on a $\mathcal{A} = \sigma(\mathcal{C})$, $\mathcal{B} = \sigma(\mathcal{D})$. Par un théorème du cours (conséquence du lemme des classes monotones), les tribus $\mathcal{A}$ et $\mathcal{B}$ sont alors indépendantes si et seulement si les quatre équations suivantes sont vérifiées :

$$\mathbb{P}(\{x = 0\} \cap \{|x| = 0\}) = \mathbb{P}(x = 0)\mathbb{P}(|x| = 0) \tag{1}$$
$$\mathbb{P}(\{x > 0\} \cap \{|x| = 0\}) = \mathbb{P}(x > 0)\mathbb{P}(|x| = 0) \tag{2}$$
$$\mathbb{P}(\{x = 0\} \cap \{|x| \in B\}) = \mathbb{P}(x = 0)\mathbb{P}(|x| \in B)\, \forall B \subset]0, \infty[\text{ mesurable} \tag{3}$$
$$\mathbb{P}(\{x > 0\} \cap \{|x| \in B\}) = \mathbb{P}(x > 0)\mathbb{P}(|x| \in B)\, \forall B \subset]0, \infty[\text{ mesurable.} \tag{4}$$

On a alors

$$\begin{aligned}
&(1) \iff \mathbb{P}(x = 0) = \mathbb{P}(x = 0)^2 \iff \mathbb{P}(x = 0) \in \{0, 1\}\\
&(2) \iff 0 = \mathbb{P}(x > 0)\mathbb{P}(|x| = 0) \iff \mathbb{P}(x > 0) = 0 \text{ ou } \mathbb{P}(x = 0) = 0\\
&(3) \iff 0 = \mathbb{P}(x = 0)\mathbb{P}(|x| \in B)\, \forall B \subset]0, \infty[\text{ mesurable} \iff \mathbb{P}(x = 0) = 0 \text{ ou } \mathbb{P}(|x| > 0) = 0.
\end{aligned}$$

Les équations (1),(2),(3) sont alors vérifiées ssi $\mathbb{P}(x = 0) \in \{0, 1\}$. De plus, on a

$$\begin{aligned}
(4) &\iff \mathbb{P}(x \in B) = \mathbb{P}(x > 0)\mathbb{P}(|x| \in B)\, \forall B \subset]0, \infty[\text{ mesurable}\\
&\iff \mathbb{P}(x \in B) = \mathbb{P}(x > 0)(\mathbb{P}(x \in B) + \mathbb{P}(-x \in B))\, \forall B \subset]0, \infty[\text{ mesurable}\\
&\iff (1 - \mathbb{P}(x > 0))\mathbb{P}(x \in B) = \mathbb{P}(x > 0)\mathbb{P}(-x \in B)\, \forall B \subset]0, \infty[\text{ mesurable.}
\end{aligned}$$

En conclusion, on a $\mathcal{A} \perp \mathcal{B}$ ssi $\mathbb{P}(x = 0) = 1$ ou $\mathbb{P}(x = 0) = 0$ et il existe $p \in [0, 1]$, tel que

$$(1 - p)\mathbb{P}(x \in B) = p\mathbb{P}(-x \in B)\, \forall B \subset]0, \infty[\text{ mesurable.}$$

Exercice 17.

1. Notons $p_n = \mathbb{P}(X = n)$. Puisque $g_X(z) = \mathbb{E}[z^X] = \sum_{n=0}^{\infty} p_n z^n$ et $\sum p_n = 1$, la série converge pour tout $|z| \leq 1$, et donc le rayon de convergence est d'au moins 1.
2. Puisque le rayon de convergence de la série $g_X(z)$ est au moins 1, on a pour tout $|z| < 1$,

$$\frac{d^n}{dz^n} g_X(z) = \sum_{k=0}^{n} p_k k(k-1)\cdots(k-1+1)z^{k-n} = \mathbb{E}[X(X-1)\cdots(X-n+1)z^{X-n}].$$

En faisant tendre $z \uparrow 1$, on obtient alors par convergence monotone,

$$\begin{aligned}
\frac{d^n}{dz^n} g_X(z)\Big|_{z=1} &= \lim_{z\uparrow 1} \frac{d^n}{dz^n} g_X(z)\\
&= \lim_{z\uparrow 1} \mathbb{E}[X(X-1)\cdots(X-n+1)z^{X-n}]\\
&= \mathbb{E}[X(X-1)\cdots(X-n+1)].
\end{aligned}$$

3. *Loi de Poisson :* Soit $P \sim \mathrm{Po}(\lambda)$, $\lambda > 0$. On a alors

$$g_P(z) = \sum_{n=0}^{\infty} e^{-\lambda} \frac{\lambda^n z^n}{n!} = e^{\lambda(z-1)}.$$

Par conséquent, on a pour tout $n \in \mathbb{N}^*$,

$$\mathbb{E}[P(P-1)\cdots(P-n+1)] = \frac{d^n}{dz^n} g_P(z)\Big|_{z=1} = \lambda^n.$$

En particulier, $\mathbb{E}[P] = \lambda$, $\mathbb{E}[P^2] = \mathbb{E}[P(P-1)] + \mathbb{E}[P] = \lambda^2 + \lambda$, et donc $\mathrm{Var}(P) = \mathbb{E}[P^2] - \mathbb{E}[P]^2 = \lambda$.
Loi binomiale : Soit $B \sim \mathrm{Bin}(n,p)$, $p \in [0,1]$. Puisque B peut s'écrire comme la somme de n v.a. $\mathrm{Ber}(p)$ indépendantes, on a

$$g_B(z) = g_{\mathrm{Ber}(p)}(z)^n = ((1-p) + pz)^n.$$

Par conséquent, on a pour tout $k \in \mathbb{N}^*$,

$$\mathbb{E}[B(B-1)\cdots(B-k+1)] = \frac{d^k}{dz^k} g_B(z)\Big|_{z=1} = p^k n(n-1)\cdots(n-k+1).$$

En particulier, $\mathbb{E}[B] = np$, et donc

$$\mathrm{Var}(B) = \mathbb{E}[B(B-1)] + \mathbb{E}[B] - \mathbb{E}[B]^2 = p^2 n(n-1) + np - (np)^2 = np(1-p).$$

Loi géométrique : Soit $G \sim \mathrm{Geo}(p)$, $p \in (0,1]$. On a alors

$$\begin{aligned} g_G(z) &= \sum_{n=1}^{\infty} (1-p)^{n-1} p z^n = pz \sum_{n=0}^{\infty} (1-p)^n z^n = \frac{pz}{1-(1-p)z} = \frac{\frac{p}{1-p}((1-p)z - 1) + \frac{p}{1-p}}{1-(1-p)z} \\ &= -\frac{p}{1-p} + \frac{p}{1-p} \cdot \frac{1}{1-(1-p)z} \end{aligned}$$

Par conséquent, on a pour tout $n \in \mathbb{N}^*$,

$$\mathbb{E}[G(G-1)\cdots(G-n+1)] = \frac{d^n}{dz^n} g_G(z)\Big|_{z=1} = \frac{p}{1-p} \cdot \frac{(1-p)^n n!}{(1-(1-p)z)^{n+1}}\Big|_{z=1} = \frac{(1-p)^{n-1} n!}{p^n}.$$

En particulier, $\mathbb{E}[G] = 1/p$, et donc

$$\mathrm{Var}(G) = \mathbb{E}[G(G-1)] + \mathbb{E}[G] - \mathbb{E}[G]^2 = 2(1-p)/p^2 + 1/p - 1/p^2 = (1-p)/p^2.$$

Exercice 18. On a par indépendance, pour $\lambda \in \mathbb{R}^d$,

$$\begin{aligned} \mathbb{E}[e^{i\lambda \sum_{n=1}^N X_n}] &= \mathbb{E}\Big[\sum_{n=0}^{\infty} e^{i\lambda(X_1+\cdots+X_n)} \mathbb{1}_{N=n}\Big] = \sum_{n=0}^{\infty} \mathbb{E}[e^{i\lambda(X_1+\cdots+X_n)}]\mathbb{P}(N=n) \\ &= \sum_{n=0}^{\infty} \varphi_X(\lambda)^n \mathbb{P}(N=n) = \mathbb{E}[\varphi_X(\lambda)^N] = g_N(\varphi_X(\lambda)), \end{aligned}$$

et donc $\varphi_{\sum_{n=1}^N X_n} = g_N \circ \varphi_X$.

Exercice 19. Si $N \sim \mathrm{Geo}(p)$ et $X_1 \sim \mathrm{Exp}(\beta)$, $p \in (0,1]$, $\beta > 0$, alors

$$g_N(\varphi_X(\lambda)) = \frac{p\frac{\beta}{\beta - i\lambda}}{1-(1-p)\frac{\beta}{\beta-i\lambda}} = \frac{p\beta}{\beta - i\lambda - (1-p)\beta} = \frac{p\beta}{p\beta - i\lambda}.$$

et donc $\sum_{n=1}^N X_n$ suit une loi exponentielle de paramètre $p\beta$.

De même, si $N \sim \mathrm{Geo}(p)$ et $X_1 \sim \mathrm{Geo}(q)$, $p, q \in (0,1]$, alors

$$g_N(\varphi_X(\lambda)) = \frac{p\frac{qe^{i\lambda}}{1-(1-q)e^{i\lambda}}}{1-(1-p)\frac{qe^{i\lambda}}{1-(1-q)e^{i\lambda}}} = \frac{pqe^{i\lambda}}{1-(1-q)e^{i\lambda} - (1-p)qe^{i\lambda}} = \frac{pqe^{i\lambda}}{1-(1-pq)e^{i\lambda}},$$

et donc $\sum_{n=1}^N X_n$ suit une loi géométrique de paramètre pq.

Les deux résultats se ressemblent parce que les lois géométriques et exponentielles possèdent tous les deux la propriété de *perte de mémoire*, i.e., si $G \sim \mathrm{Geo}(p)$, alors conditionnellement à $G > k$, on a encore $G-k \sim \mathrm{Geo}(p)$. De même, si $X \sim \mathrm{Exp}(\beta)$, alors conditionnellement à $X > x$, on a encore $X - x \sim \mathrm{Exp}(\beta)$. C'est cette propriété qui se cache derrière ces résultats. D'ailleurs, de la même façon qu'une suite de variables géométrique peut être construite à partir d'une suite de v.a. Bernoulli indépendantes, une suite de v.a. exponentielles peut être construite à partir un analogue continu de la suite de v.a. Bernoulli : le *processus de Poisson homogène sur* $\mathbb{R}_+$.

Exercice 20 (Loi multinomiale).

1. Notons $\vec{X}^j = (\mathbf{1}_{(Y^j=1)}, \dots, \mathbf{1}_{(Y^j=k)})$, si bien que $\vec{X} = \sum_{i=1}^n X^j$. Pour $\vec{t} = (t_1, \dots, t_k) \in \mathbb{R}^k$, on a alors,

$$\begin{aligned}
\varphi_{\vec{X}}(\vec{t}) &= \mathbb{E}[\prod_{i=1}^n e^{i\vec{t}\cdot\vec{X}^j}] \\
&= \prod_{i=1}^n \mathbb{E}[e^{i\vec{t}\cdot\vec{X}^j}] && \text{par indépendance} \\
&= (\mathbb{E}[e^{i(t_1\mathbf{1}_{(Y^1=1)}+\cdots+t_k\mathbf{1}_{(Y^1=k)})}])^n \\
&= (p_1 e^{it_1} + \cdots + p_k e^{it_k})^n
\end{aligned}$$

2. D'après l'exercice 3, la fonction caractéristique du vecteur $\vec{S}$ est égale à la composition de la fonction génératrice de la loi de Poisson, $g_\lambda(z) = e^{\lambda(z-1)}$, avec la fonction caractéristique du vecteur $(\mathbf{1}_{(Y^1=1)}, \dots, \mathbf{1}_{(Y^1=k)})$ calculée ci-dessus (poser $n = 1$). Ceci donne

$$\varphi_{\vec{S}}(\vec{t}) = e^{\lambda(p_1e^{it_1}+\cdots+p_ke^{it_k}-1)} = \prod_{l=1}^k e^{\lambda p_l(e^{it_l}-1)},$$

car $\sum_{l=1}^k p_l = 1$. Il vient que le vecteur aléatoire suit la loi $\mathrm{Po}(\lambda p_1) \otimes \cdots \otimes \mathrm{Po}(\lambda p_k)$, i.e. ses composantes sont indépendantes de lois de Poisson de paramètres respectifs $p_1\lambda, \dots, p_k\lambda$.

Exercice 21.

1. $\mathrm{Cov}(X, Y)$ est de dimension $n \times m$.
2. On a $\mathrm{Cov}(Y, X) = \mathbb{E}[(Y - \mathbb{E}[Y])(X - \mathbb{E}[X])^T] = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])^T]^T = \mathrm{Cov}(X, Y)^T$.
3. On a

$$\mathrm{Cov}(AX, Y) = \mathbb{E}[(AX - \mathbb{E}[AX])(Y - \mathbb{E}[Y])^T] = \mathbb{E}[A(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])^T] = A\,\mathrm{Cov}(X, Y).$$

Par conséquent, d'après la partie 2, $\mathrm{Cov}(X, BY) = \mathrm{Cov}(BY, X)^T = (B\,\mathrm{Cov}(Y, X))^T = \mathrm{Cov}(X, Y)B^T$. Ces deux formules donnent alors, $\Sigma_{AX} = \mathrm{Cov}(AX, AX) = A\,\mathrm{Cov}(X, X)A^T = A\Sigma_X A^T$.

4. On a d'après la partie 2, $\Sigma_X^T = \mathrm{Cov}(X, X)^T = \mathrm{Cov}(X, X) = \Sigma_X$, donc Σ_X est symétrique. De plus, on a pour tout $v \in \mathbb{R}^n$, d'après la partie 3, en remarquant que $v^T X$ est une v.a. réelle,

$$v^T\Sigma_X v = \Sigma_{v^T X} = \mathrm{Var}(v^T X) \geq 0,$$

si bien que Σ_X est positive.

La symétrie de Σ_X implique qu'elle est diagonalisable sur $\mathbb{R}$ et que ses espaces propres sont orthogonaux, en particulier, il existe une base orthonormale de vecteurs propres de Σ_X. La positivité implique en plus que les valeurs propres sont positives.

5. On a d'après la partie 3, $\mathrm{Var}(v^T X) = v^T\Sigma_X v$. Par conséquent,

$$v^T X \text{ est dégénérée} \iff \mathrm{Var}(v^T X) = 0 \iff v^T\Sigma_X v = 0 \iff v \in \ker(\Sigma_X).$$

Ici, la dernière égalité peut-être démontrée par exemple en décomposant v selon une b.o.n. de vecteurs propres de Σ_X : Si $e_1, \dots, e_n$ est une telle b.o.n., de valeurs propres $\lambda_1, \dots, \lambda_n \geq 0$, et $v = a_1e_1+\cdots+a_ne_n$, alors $v^T\Sigma_X v = \sum_{i=1}^n a_i^2\lambda_i$, et donc $v^T\Sigma_X v = 0$ si et seulement si $a_i = 0$ pour tout i tel que $\lambda_i = 0$, ce qui signifie que $v \in \ker \Sigma_X$.

Exercice 22 (Vecteur gaussien standard)**.** 1. On a $\|X\|_2^2 = X_1^2 + \cdots + X_n^2$. Les variables X_i^2 sont iid de lois $\Gamma(1/2, 1/2)$, si bien que $\|X\|_2^2 \sim \Gamma(n/2, 1/2)$. On a alors pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ mesurable bornée,

$$\begin{aligned}
\mathbb{E}[f(\|X\|_2)] &= \frac{1}{2^{n/2}\Gamma(n/2)} \int_0^\infty f(\sqrt{x}) x^{n/2-1} e^{-x/2}\, dx \\
&= \frac{1}{2^{n/2}\Gamma(n/2)} \int_0^\infty f(y) y^{n-2} e^{-y^2/2} 2y\, dy \\
&= \frac{1}{2^{n/2-1}\Gamma(n/2)} \int_0^\infty f(y) y^{n-1} e^{-y^2/2}\, dy.
\end{aligned}$$

Par conséquent, $\|X\|_2$ soit la loi de densité $\frac{1}{2^{n/2-1}\Gamma(n/2)} y^{n-1} e^{-y^2/2}$ sur $\mathbb{R}_+$.

2. Soit $O \in O(n)$. Alors OX est encore un vecteur gaussien. Notons que $\mathbb{E}[X] = 0$ et $\Sigma_X = I$, la matrice d'identité. On a $\mathbb{E}[OX] = O\mathbb{E}[X] = 0$ et $\Sigma_{OX} = OIO^T = OO^T = I$. Puisque la loi d'un vecteur gaussien est déterminée par son espérance et sa matrice de covariance, OX est alors un vecteur gaussien standard n-dimensionnel.
3. Soit $O \in O(n)$. On sait depuis le dernier exercice que $OX \overset{\text{loi}}{=} X$. On a alors,

$$O(X/\|X\|_2) = (OX)/\|X\|_2 = (OX)/\|OX\|_2 \overset{\text{loi}}{=} X/\|X\|_2.$$

4. Soit $f : \mathbb{R} \to \mathbb{R}$ mesurable bornée, $B \subset S^{n-1}$ mesurable et $O \in O(n)$. On a alors

$$\begin{aligned}
\mu_f(O^{-1}B) &= \mathbb{E}[f(\|X\|_2)\mathbf{1}_{O^{-1}B}(X/\|X\|_2)] = \mathbb{E}[f(\|OX\|_2)\mathbf{1}_B(OX/\|OX\|_2)] \\
&= \mathbb{E}[f(\|X\|_2)\mathbf{1}_B(X/\|X\|_2)] = \mu_f(B),
\end{aligned}$$

par la partie 2 de l'exercice.

5. Soit $f : \mathbb{R} \to \mathbb{R}$ mesurable bornée. On vient de montrer que μ_f est invariante par l'action de tout $O \in O(n)$. Par conséquent, il existe une constante c_f telle que $\mu_f = c_f \sigma^{n-1}$. On a

$$c_f = c_f \sigma^{n-1}(S^{n-1}) = \mu_f(S^{n-1}) = \mathbb{E}[f(\|X\|_2)].$$

Par conséquent, on a

$$\mathbb{E}[f(\|X\|_2)\mathbf{1}_B(X/\|X\|_2)] = \mu_f(B) = \mathbb{E}[f(\|X\|_2)]\sigma^{n-1}(B) = \mathbb{E}[f(\|X\|_2)]\mathbb{E}[\mathbf{1}_B(X/\|X\|_2)].$$

En prenant $f = \mathbf{1}_A$ pour tout ensemble $A \subset \mathbb{R}$ mesurable, on voit alors que $\|X\|_2$ et $X/\|X\|_2$ sont indépendantes.

Exercice 23 (Méthode Box–Muller)**.** On sait d'après le dernier exercice qu'un vecteur gaussien standard 2-dimensionnel s'ecrit comme le produit RV, où R et V sont des v.a. indépendantes, R est à valeurs dans $(0, \infty)$, $R^2 \sim \Gamma(1, 1/2) = \text{Exp}(1/2)$ et V est uniformément distribué sur le cercle S^1. Pour construire R, on remarque que si $F(x) = e^{-x}$ désigne la queue de la loi Exp(1) et W une v.a. uniforme sur l'intervalle $(0, 1)$, alors $F^{-1}(W) = -\log W \sim \text{Exp}(1)$. Par conséquent, $2(-\log W) \sim \text{Exp}(1/2)$ et donc R est égale en loi à $\sqrt{2(-\log W)}$. En ce qui concerne le vecteur aléatoire V, si $U \sim \text{Unif}(0, 1)$, alors $2\pi U \sim \text{Unif}(0, 2\pi)$, et donc V est égal en loi à $(\cos(2\pi U), \sin(2\pi U))$.

Résumant le dernier paragraphe : si on pose

$$(X, Y) = \sqrt{2(-\log W)}(\cos(2\pi U), \sin(2\pi U)),$$

alors (X, Y) est un vecteur gaussien standard 2-dimensionnel.

Cette méthode est nettement meilleure que la méthode par inversion de la fonction de répartition, car la fonction de répartition ne possède par de formule analytique. Elle devrait alors être calculée par des méthodes de quadrature de l'intégrale $\int_x^\infty e^{-y^2/2}\, dy$, ce qui est plus couteux que les opérations ci-dessus.

Exercice 24 (Extrait du partiel 2013)**.** 1. On a

$$W = 2 + \begin{pmatrix} 2 & -1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix},$$

et donc W est une transformation affine du vecteur gaussien $(X,Y)^T$ et donc une variable gaussienne. On a $\mathbb{E}[W] = 2+2+0 = 4$ et

$$\mathrm{Var}(W) = \begin{pmatrix} 2 & -1 \end{pmatrix} \Sigma \begin{pmatrix} 2 \\ -1 \end{pmatrix} = 2,$$

donc $W \sim \mathcal{N}(4,2)$.

2. $(Z,T)^T$ est un vecteur gaussien de moyenne nulle et matrice de covariance Id. Par conséquent,

$$\begin{pmatrix} 1+Z \\ Z+T \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} Z \\ T \end{pmatrix}$$

est également un vecteur gaussien, d'espérance $(1,0)^T$ et de matrice de covariance

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \Sigma.$$

Ceci montre que $(1+Z, Z+T)^T$ et $(X,Y)^T$ ont même loi.

3. D'après la dernière partie, on a

$$(X-1)^2 + (Y-X+1)^2 \overset{\text{loi}}{=} Z^2 + T^2 \sim \Gamma(1,1/2),$$

puisque Z^2 et T^2 sont iid de loi $\Gamma(1/2,1/2)$. De plus, S suit la loi $\Gamma(1,1/2)$ et est indépendante de $(X,Y)^T$ par hypothèse. Par conséquent,

$$(X-1)^2 + (Y-X+1)^2 + S \sim \Gamma(2,1/2).$$

Exercice 25. 1. Vu que les vecteurs gauches et droits sont centrés, l'équation $(X,Y)^T \overset{\text{loi}}{=} A(G_1,G_2)^T$ est équivalente au fait que les matrices de covariance à gauche et à droite coincident. On cherche donc A telle que

$$AA^T = \Sigma = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}.$$

Or, la forme quadratique associée à Σ est

$$x^2 + xy + y^2 = \left(x + \tfrac{1}{2}y\right)^2 + \left(\tfrac{\sqrt{3}}{2}y\right)^2 = \left\| \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \right\|^2.$$

Ainsi, par identification des formes quadratiques on obtient $\Sigma = AA^T$ avec

$$A = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{pmatrix}^T = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$$

On peut vérifier par calcul direct que cela marche en effet.

2. D'après la première partie,

$$\mathbb{P}(X \geq 0,\, Y \geq 0) = \mathbb{P}((G_1 \geq 0,\, \tfrac{1}{2}G_1 + \tfrac{\sqrt{3}}{2}G_2 \geq 0) = \mathbb{P}((G_1,G_2)^T \in C),$$

où C est le cône $C = \{(x,y)^T \in \mathbb{R}^2 : x \geq 0,\, \frac{1}{2}x + \frac{\sqrt{3}}{2}y \geq 0\}$. C'est donc l'intersection des deux demi-plans de vecteurs normaux intérieurs $(1,0)^T = (\cos(0), \sin(0))^T$ et $(\frac{1}{2}, \frac{\sqrt{3}}{2})^T = (\cos(\pi/3), \sin(\pi/3))^T$. L'angle au sommet du cône est alors $\pi - \pi/3 = 2\pi/3$. Par invariance par rotation du vecteur gaussien standard $(G_1,G_2)^T$, on obtient alors

$$\mathbb{P}(X \geq 0,\, Y \geq 0) = \mathbb{P}((G_1,G_2)^T \in C) = \frac{2\pi/3}{2\pi} = \frac{1}{3}.$$

Tribu terminale, convergence presque sûre, lemme de Borel–Cantelli

Exercice 26 (Tribu terminale)**.**

1. Vrai. On a pour tout $n \in \mathbb{N}$, $\limsup_{m\to\infty} X_m = \limsup_{m\to\infty} X_{n+m}$. Par conséquent, $\limsup_{m\to\infty} X_m$ est mesurable par rapport à $\sigma(X_n, X_{n+1}, \ldots)$, et donc $\{\limsup_{m\to\infty} X_m < \infty\} \in \mathcal{T}$.
2. Vrai (même raison que 1, $\lim_{m\to\infty} X_m$ est mesurable par rapport à $\sigma(X_n, X_{n+1}, \ldots)$ pour tout $n \in \mathbb{N}$).
3. Faux. Supposons qu'il existe $A \in \mathcal{A}$ tel que $A_3 = \mathbb{R} \times A$. Alors on aurait $\omega = (\omega_0, \omega_1, \ldots) \in A_3$ ssi $(0, \omega_1, \omega_2, \ldots) \in A_3$. Mais ceci est faux, car $(1, 1, \ldots) \in A_3$, mais $(0, 1, 1, \ldots) \notin A_3$.
4. Faux, car $(1, 1, \ldots) \in A_4$ mais $(0, 1, 1, \ldots) \notin A_4$. On raisonne alors comme dans 3.
5. Vrai. Posons $S_m = \sum_{k=0}^m X_k$. On a pour tout $n \in \mathbb{N}$, $\limsup_{m\to\infty} S_m < \infty$ ssi $\limsup_{m\to\infty}(S_m - S_n) < \infty$. Or, $\limsup_{m\to\infty}(S_m - S_n)$ est mesurable par rapport à $\sigma(X_{n+1}, X_{n+2}, \ldots)$ pour tout $n \in \mathbb{N}$. Par conséquent, $A_5 \in \sigma(X_n, X_{n+1}, \ldots)$ pour tout $n \in \mathbb{N}$ et donc $A_5 \in \mathcal{T}$.
6. Faux, car la première valeur influence la valeur de la somme. Formellement, on a $(0, 0, \ldots) \in A_6$, mais $(1, 0, 0, \ldots) \notin A_6$. On raisonne alors comme dans 3.

Exercice 27.

1. On a pour tout $n_0 \in \mathbb{N}$, $\bigcap_{n\ge n_0} A_n = B \cap C$ et $\bigcup_{n\ge n_0} A_n = B \cup C$. Par conséquent,
$$\liminf A_n = \bigcup_{n_0\in\mathbb{N}} \bigcap_{n\ge n_0} A_n = B \cap C \quad \text{et} \quad \limsup A_n = \bigcap_{n_0\in\mathbb{N}} \bigcup_{n\ge n_0} A_n = B \cup C.$$
2. Par les règles de De Morgan,
$$(\liminf A_n)^c = \left(\bigcup_{n_0\in\mathbb{N}} \bigcap_{n\ge n_0} A_n\right)^c = \bigcap_{n_0\in\mathbb{N}} \bigcup_{n\ge n_0} A_n^c = \limsup A_n^c.$$
3. On a pour tout $\omega \in \Omega$,
$$\begin{aligned}\liminf \mathbb{1}_{A_n}(\omega) = \sup_{n_0\in\mathbb{N}} \inf_{n\ge n_0} \mathbb{1}_{A_n}(\omega) &= \begin{cases} 1, & \text{si } \exists n_0 \in \mathbb{N}\, \forall n \ge n_0 : \mathbb{1}_{A_n}(\omega) = 1 \\ 0, & \text{sinon} \end{cases} \\ &= \begin{cases} 1, & \text{si } \exists n_0 \in \mathbb{N}\, \forall n \ge n_0 : \omega \in A_n \\ 0, & \text{sinon} \end{cases} \\ &= \mathbb{1}_{\liminf A_n}(\omega).\end{aligned}$$

   Un raisonnement analogue donne l'égalité pour lim sup, mais on peut aussi la déduire de l'égalité précédente en remarquant que
$$\begin{aligned}\mathbb{1}_{\limsup A_n} &= 1 - \mathbb{1}_{(\limsup A_n)^c} \\ &= 1 - \mathbb{1}_{\liminf A_n^c} && \text{par la dernière partie} \\ &= 1 - \liminf \mathbb{1}_{A_n^c} && \text{par l'égalité pour } \liminf \\ &= 1 - \liminf(1 - \mathbb{1}_{A_n}) \\ &= 1 - (1 - \limsup \mathbb{1}_{A_n}) \\ &= \limsup \mathbb{1}_{A_n}.\end{aligned}$$

   On a alors par le lemme de Fatou
$$\mathbb{P}(\liminf A_n) = \mathbb{E}[\mathbb{1}_{\liminf A_n}] = \mathbb{E}[\liminf \mathbb{1}_{A_n}] \le \liminf \mathbb{E}[\mathbb{1}_{A_n}] = \liminf \mathbb{P}(A_n).$$

   Avec la dernière partie de l'exercice, ceci implique,
$$\mathbb{P}(\limsup A_n) = 1 - \mathbb{P}(\liminf A_n^c) \ge 1 - \liminf \mathbb{P}(A_n^c) = 1 - (1 - \limsup \mathbb{P}(A_n)) = \limsup \mathbb{P}(A_n).$$

Exercice 28. Convergence en probabilité :

$$\begin{aligned} X_n \xrightarrow[n\to\infty]{\mathbb{P}} 0 &\iff \forall \varepsilon > 0 : \mathbb{P}(|X_n| > \varepsilon) \to 0 && \text{par définition} \\ &\iff \mathbb{P}(X_n = 1) \to 0 && \text{car } X_n \in \{0,1\}\ \forall n \in \mathbb{N}^* \\ &\iff p_n \to 0. \end{aligned}$$

Convergence presque sûre :

$$\begin{aligned} X_n \to 0 \quad \text{p.s.} &\iff \exists N \in \mathbb{N}^*\ \forall n \geq N : X_n = 0 \quad \text{p.s.} && \text{car } X_n \in \{0,1\}\ \forall n \in \mathbb{N}^* \\ &\iff \mathbb{P}(\limsup A_n) = 0 \\ &\iff \sum_{n\in\mathbb{N}^*} p_n < \infty && \text{par le lemme de Borel–Cantelli.} \end{aligned}$$

Exercice 29. On sait dejà que la convergence p.s. entraîne la convergence en probabilité.

Réciproquement, supposons que S_n converge en probabilité vers une variable S. Alors la suite de v.a. $(R_n)_{n\geq 0}$ définie par $R_n = S - S_n = X_{n+1} + X_{n+2} + \dots$ converge en probabilité vers zéro. Or, comme les variables X_n sont positives, (R_n) est une suite décroissante minorée par zéro donc elle converge presque sûrement. Sa limite en probabilité étant zéro, sa limite presque sûre est aussi zéro.

Preuve alternative : Convergence en probabilité implique l'existence d'une sous-suite qui converge presque sûrement. Utiliser ensuite le fait que la suite est croissante pour conclure.

Exercice 30. On a

$$2\sum_{1\leq i<j\leq n} X_iX_j = \sum_{1\leq i\neq j\leq n} X_iX_j = \sum_{1\leq i,j\leq n} X_iX_j - \sum_{i=1}^n X_i^2 = \left(\sum_{1\leq i\leq n} X_i\right)^2 - \sum_{i=1}^n X_i^2.$$

Par la loi forte des grands nombres, on a quand $n\to\infty$,

$$\frac{1}{n(n-1)}\left(\sum_{1\leq i\leq n} X_i\right)^2 = \frac{n}{n-1}\left(\frac{1}{n}\sum_{1\leq i\leq n} X_i\right)^2 \to \mu^2, \quad \text{p.s.}$$

De plus, puisque $\mathbb{E}[X_1^2] = \mathrm{Var}(X_1) + \mathbb{E}[X_1]^2 < \infty$ par hypothèse, on a, encore par la loi forte des grands nombres,

$$\frac{1}{n}\sum_{i=1}^n X_i^2 \to \mathbb{E}[X_1^2], \quad \text{p.s.,}$$

et donc

$$\frac{1}{n(n-1)}\sum_{i=1}^n X_i^2 \to 0, \quad \text{p.s.}$$

Ceci donne

$$\binom{n}{2}^{-1}\sum_{1\leq i<j\leq n} X_iX_j = \frac{1}{n(n-1)}\left(\left(\sum_{1\leq i\leq n} X_i\right)^2 - \sum_{i=1}^n X_i^2\right) \to \mu^2, \quad \text{p.s.}$$

Exercice 31.

1. Si E est une v.a. de loi exponentielle de paramètre λ, alors $\mathbb{P}(E > x) = e^{-\lambda x}$ pour tout $x \geq 0$. Par conséquent, $\mathbb{P}(T_k \geq 1) = e^{-\ln(k)} = 1/k$ et $\mathbb{P}(T_k \geq 1+\varepsilon) = e^{-(1+\varepsilon)\ln(k)} = 1/k^{1+\varepsilon}$.
2. Borne inférieure : Puisque $\sum_{k=1}^\infty \mathbb{P}(T_k \geq 1) = \sum_{k=1}^\infty 1/k = \infty$ et que les T_k, $k \geq 2$ sont indépendantes, la seconde partie du lemme de Borel–Cantelli implique que presque sûrement, $T_k \geq 1$ pour un nombre infini de k, et donc
$$\limsup_{k\to\infty} T_k \geq 1, \quad \text{p.s..}$$
Borne supérieure : Pour tout $\varepsilon > 0$, $\sum_{k=1}^\infty \mathbb{P}(T_k \geq 1+\varepsilon) = \sum_{k=1}^\infty 1/k^{1+\varepsilon} < \infty$. Le première partie du lemme de Borel–Cantelli implique alors que p.s., $T_k \leq 1+\varepsilon$ pour tous sauf un nombre fini de k, et donc
$$\limsup_{k\to\infty} T_k \leq 1+\varepsilon, \quad \text{p.s..}$$

Puisque les événements $\{\limsup_{k\to\infty} T_k \leq 1+1/N\}$ sont décroissants en N, il vient

$$\mathbb{P}(\limsup_{k\to\infty} T_k \leq 1) = \lim_{N\to\infty} \mathbb{P}(\limsup_{k\to\infty} T_k \leq 1+1/N) = 1.$$

En combinant les deux bornes, on obtient

$$\limsup_{k\to\infty} T_k = 1, \quad \text{p.s.}.$$

Exercice 32. 1. Soit $a>0$. On rappelle que $\mathbb{E}[X] = \int_0^\infty \mathbb{P}(X \geq x)\,dx$. Il vient que

$$\mathbb{E}[X] = \sum_{n=1}^\infty \int_{a(n-1)}^{an} \mathbb{P}(X > x)\,dx \geq \sum_{n=1}^\infty a\mathbb{P}(X \geq an),$$

puisque $\mathbb{P}(X \geq x)$ est décroissante en x. De même,

$$\mathbb{E}[X] = \sum_{n=1}^\infty \int_{a(n-1)}^{an} \mathbb{P}(X > x)\,dx \leq \sum_{n=1}^\infty a\mathbb{P}(X \geq a(n-1)) = \sum_{n=0}^\infty a\mathbb{P}(X \geq an).$$

Ce qui montre que $\sum_{n=0}^\infty \mathbb{P}(X \geq an) = \infty \iff \mathbb{E}[X] = \infty$.

2. Puisque les X_n sont indépendantes, on peut appliquer les deux partie du lemme de Borel–Cantelli et la première partie de l'exercice done alors pour tout $a>0$,

$$\mathbb{P}(\limsup\{X_n \geq an\}) = \begin{cases} 0, & \text{si } \mathbb{E}[X] < \infty \\ 1 & \text{si } \mathbb{E}[X] = \infty, \end{cases}$$

Pour finir, supposons d'abord que $\mathbb{E}[X] = \infty$, alors

$$\begin{aligned} \mathbb{P}(\limsup \frac{1}{n} X_n = \infty) &= \mathbb{P}(\bigcap_{N\in\mathbb{N}^*} \{X_n \geq Nn \text{ pour un nombre infini de } n\}) \\ &= \mathbb{P}(\bigcap_{N\in\mathbb{N}^*} \limsup_{n\to\infty}\{X_n \geq Nn\}) \\ &= 1, \end{aligned}$$

puisque l'intersection d'un nombre dénombrable d'ensembles de probabilité 1 a encore probabilité 1. De manière analogue, quand $\mathbb{E}[X] < \infty$,

$$\begin{aligned} \mathbb{P}(\limsup \frac{1}{n} X_n = 0) &= \mathbb{P}(\bigcap_{N\in\mathbb{N}^*} \{X_n < n/N \text{ pour tous sauf un nombre fini de } n\}) \\ &= \mathbb{P}(\bigcap_{N\in\mathbb{N}^*} \liminf_{n\to\infty}\{X_n < n/N\}) \\ &= 1 - \mathbb{P}(\bigcup_{N\in\mathbb{N}^*} \limsup_{n\to\infty}\{X_n \geq n/N\}) \\ &= 1, \end{aligned}$$

puisque l'union d'un nombre dénombrable d'ensembles de probabilité 0 a encore probabilité 0. CQFD.

Exercice 33 (Loi du logarithme itéré.). 1. On a pour n assez grand,

$$\begin{aligned} \mathbb{P}(\max_{0\leqslant k\leqslant K^n} S_k \geqslant Kh(K^{n-1})) &\leq 2\mathbb{P}(S_{\lfloor K^n\rfloor} \geq Kh(K^{n-1})) && \text{par l'estimée (2)} \\ &= 2\mathbb{P}(X_1 \geq Kh(K^{n-1})/\sqrt{\lfloor K^n\rfloor}) && \text{puisque } S_{\lfloor K^n\rfloor} \overset{\text{loi}}{=} \sqrt{\lfloor K^n\rfloor} X_1 \\ &\leq 2\mathbb{P}(X_1 \geq K^{1-n/2} h(K^{n-1})) \\ &\leq 2\exp(-K^{2-n}K^{n-1}\log\log K^{n-1}) && \text{par l'estimée (1)} \\ &= 2((n-1)\log K)^{-K} \end{aligned}$$

Puisque $\sum_{n=2}^{\infty}((n-1)\log K)^{-K} < \infty$, le lemme de Borel-Cantelli donne que

$$\mathbb{P}(\limsup_{n\to\infty}\{\max_{0\leqslant k\leqslant K^n}(S_k/h(K^{n-1})) \geq K\}) = 0.$$

Par conséquent,

$$\begin{aligned}
\mathbb{P}(\limsup_{n\to\infty}\frac{S_n}{h(n)} > K) &\leq \mathbb{P}\left(\limsup_{n\to\infty}\left\{\max_{K^{n-1}\leq k\leq K^n}\frac{S_k}{h(K^{n-1})} \geq K\right\}\right) && \text{par croissance de } h\\
&\leq \mathbb{P}\left(\limsup_{n\to\infty}\left\{\max_{0\leq k\leq K^n}\frac{S_k}{h(K^{n-1})} \geq K\right\}\right)\\
&= 0.
\end{aligned}$$

Ce qui implique que $\limsup_{n\to\infty}\frac{S_n}{h(n)} \leq K$ presque sûrement.

2. Par décroissance des événements, on a

$$\mathbb{P}(\limsup_{n\to\infty}\frac{S_n}{h(n)} \leq 1) = \lim_{N\to\infty}\mathbb{P}(\limsup_{n\to\infty}\frac{S_n}{h(n)} \leq 1+1/N) = 1,$$

par la partie précédente de l'exercice.

3. Montrons d'abord que les événements sont indépendants. Notons pour $k\in\mathbb{N}^*$, $\mathcal{A}_k = \sigma(X_1,\dots,X_{M^{k-1}})$ et $\overline{\mathcal{A}}_k = \sigma(X_{M^{k-1}+1},\dots,X_{M^k})$, si bien que

$$A_1,\dots,A_{k-1}\in\mathcal{A}_k \quad \text{et} \quad \mathcal{A}_k \perp \overline{\mathcal{A}}_k,$$

puisque les X_n sont indépendantes. On décompose l'événement A_k en $A_k = B_k\cap C_k$ avec

$$\begin{aligned}
B_k &= \{|S_{M^k} - S_{M^{k-1}}| \geq rh(M^k)\} \in \overline{\mathcal{A}}_k\\
C_k &= \{\operatorname{sgn}(S_{M^k} - S_{M^{k-1}}) = \operatorname{sgn}(S_{M^{k-1}})\} \in \mathcal{A}_k\vee\overline{\mathcal{A}}_k.
\end{aligned}$$

On a alors pour toute v.a. réelle Y bornée et $\mathcal{A}_k$-mesurable,

$$\begin{aligned}
\mathbb{E}[Y\mathbb{1}_{A_k}] &= \mathbb{E}[Y\mathbb{1}_{C_k}\mathbb{1}_{B_k}]\\
&= \sum_{s\in\{-1,1\}}\mathbb{E}[Y\mathbb{1}_{(\operatorname{sgn}(S_{M^{k-1}})=s)}\mathbb{1}_{(\operatorname{sgn}(S_{M^k}-S_{M^{k-1}})=s)}\mathbb{1}_{B_k}]\\
&= \sum_{s\in\{-1,1\}}\mathbb{E}[Y\mathbb{1}_{(\operatorname{sgn}(S_{M^{k-1}})=s)}]\cdot\mathbb{E}[\mathbb{1}_{(\operatorname{sgn}(S_{M^k}-S_{M^{k-1}})=s)}\mathbb{1}_{B_k}] && (\mathcal{A}_k\perp\overline{\mathcal{A}}_k)\\
&= \sum_{s\in\{-1,1\}}\mathbb{E}[Y\mathbb{1}_{(\operatorname{sgn}(S_{M^{k-1}})=s)}]\frac{1}{2}\mathbb{P}(B_k) && (\text{symétrie de } S_{M^k}-S_{M^{k-1}})\\
&= \frac{1}{2}\mathbb{P}(B_k)\mathbb{E}[Y] && (\text{linéarité de l'espérance})\\
&= \mathbb{P}(B_k)\mathbb{P}(C_k)\mathbb{E}[Y].
\end{aligned}$$

En particulier, en posant $Y\equiv 1$, on a $\mathbb{P}(A_k) = \mathbb{P}(B_k)\mathbb{P}(C_k)$, et donc

$$\mathbb{E}[Y\mathbb{1}_{A_k}] = \mathbb{E}[Y]\mathbb{P}(A_k).$$

Puisque $A_1,\dots,A_{k-1}\in\mathcal{A}_k$, ceci donne que A_k est indépendant de la famille $A_1,\dots,A_{k-1}$. Par récurrence, on obtient alors indépendance de la famille des événements A_k, $k\in\mathbb{N}^*$.

Calculons $\mathbb{P}(A_k)$. On a

$$\mathbb{P}(A_k) = \mathbb{P}(B_k)\mathbb{P}(C_k) = \frac{1}{2}\mathbb{P}(|S_{M^k}-S_{M^{k-1}}| \geq rh(M^k)) = \mathbb{P}(X_0 \geq rh(M^k)/\sqrt{M^k - M^{k-1}}),$$

puisque $S_{M^k}-S_{M^{k-1}}$ est égale en loi à $\sqrt{M^k-M^{k-1}}X_0$ et $\mathbb{P}(|X_0|\geq x) = 2\mathbb{P}(X_0\geq x)$ pour tout $x\geq 0$. Par définition de $h(n)$, on a

$$\begin{aligned}
rh(M^k)/\sqrt{M^k-M^{k-1}} &= r\sqrt{2\frac{M^k}{M^k-M^{k-1}}\log\log M^k}\\
&= \sqrt{2r^2\frac{M}{M-1}\log(k\log M)}\\
&= \sqrt{2\rho(\log k + \log\log M)}, \quad \text{avec } \rho = r^2\frac{M}{M-1} < 1.
\end{aligned}$$

Par l'estimée (1), on obtient alors pour k assez grand, en supposant d'abord que $r \geq 1/2$ et donc $\rho > 1/8$,

$$\mathbb{P}(A_k) = \mathbb{P}(X_0 \geq rh(M^k)/\sqrt{M^k - M^{k-1}}) \geq \frac{1}{\sqrt{8\pi\rho(\log k + \log\log M)}} e^{-\rho(\log k + \log\log M)},$$

et donc, puisque $e^{-\rho \log k} = k^{-\rho}$ avec $\rho < 1$,

$$\sum_{k=1}^{\infty} \mathbb{P}(A_k) = \infty.$$

Avec l'indépendance des A_k, le lemme de Borel–Cantelli donne alors $\mathbb{P}(\limsup A_k) = 1$. Si $r < 1/2$, on borne par la probabilité du même événement mais avec r remplacé par $1/2$.

4. Soit $r \in [1/2, 1)$. Choisissons M assez large tel que $r < \sqrt{\frac{M-1}{M}}$. On définit alors A_k comme dans la partie précédente. On note que sur l'événement A_k, on a

$$|S_{M^k}| = |S_{M^k} - S_{M^{k-1}}| + |S_{M^{k-1}}| \geq |S_{M^k} - S_{M^{k-1}}| \geq rh(M^k).$$

Par l'exercice précédent, on obtient,

$$\begin{aligned} \mathbb{P}\left(\limsup_{n\to\infty} \frac{|S_n|}{h(n)} \geq r\right) &\geq \mathbb{P}\left(\limsup_{k\to\infty} \frac{|S_{M^k}|}{h(M^k)} \geq r\right) \\ &\geq \mathbb{P}\left(\limsup_{k\to\infty} \left\{|S_{M^k}| \geq rh(M^k)\right\}\right) \\ &\geq \mathbb{P}(\limsup A_k) = 1. \end{aligned}$$

Ceci implique

$$\mathbb{P}\left(\limsup \frac{|S_n|}{h(n)} \geq 1\right) = \mathbb{P}\left(\bigcap_{m\in\mathbb{N}} \left\{\limsup \frac{|S_n|}{h(n)} \geq 1 - 1/m\right\}\right) = 1.$$

5. Notons $E_+ = \{\limsup \frac{S_n}{h(n)} \geq 1\}$ et $E_- = \{\limsup \frac{-S_n}{h(n)} \geq 1\}$. Par la dernière partie, on a $\mathbb{P}(E_+ \cup E_-) = 1$. Par symétrie de la loi gaussienne, on a en plus $\mathbb{P}(E_+) = \mathbb{P}(E_-)$, donc $\mathbb{P}(E_+) \geq 1/2$. Mais puisque E_+ fait partie de la tribu terminale, la loi du 0-1 de Kolmogorov donne $\mathbb{P}(E_+) \in \{0, 1\}$ et donc $\mathbb{P}(E_+) = \mathbb{P}(E_-) = 1$.

6. Les résultats des parties 2 et 5 donnent

$$\limsup_{n\to\infty} \frac{S_n}{h(n)} = 1, \quad \text{p.s..}$$

CQFD.

Exercice 34. On pose $X_n = \sum_{i=1}^n \mathbf{1}_{A_i}$. Il est alors évident que

$$\limsup A_n = \{\mathbf{1}_{A_i} = 1 \text{ infiniment souvent}\} = \{X_n \to \infty\},$$

donc $\mathbb{P}(\limsup A_n) = 1$ ssi $X_n \to \infty$ p.s. De plus, puisque $(X_n)_{n\geq 0}$ est croissante, $X_n \to \infty$ p.s. ssi $X_n \to \infty$ en probabilité, donc si $\mathbb{P}(X_n \leq M) \to 0$ quand $n \to \infty$, pour tout $M \in \mathbb{N}$ (ceci découle par exemple de l'Exercice 29 en l'appliquant aux suites $(X_n \wedge M)_{n\geq 0}$ pour tout $M \in \mathbb{N}$).

Pour montrer que $\mathbb{P}(X_n \leq M) \to 0$ quand $n \to \infty$, pour tout $M \in \mathbb{N}$, on va appliquer l'inégalité de Chebychev. On calcule le premier et second moment de X_n :

$$\begin{aligned} \mathbb{E}[X_n] &= \sum_{i=1}^n \mathbb{E}[\mathbf{1}_{A_i}] = \sum_{i=1}^n \mathbb{P}(\mathbf{1}_{A_i}) \\ \mathrm{Var}(X_n) = \mathbb{E}[(X_n)^2] - \mathbb{E}[X_n]^2 &= \sum_{i,j=1}^n (\mathbb{P}(A_i \cap A_j) - \mathbb{P}(A_i)\mathbb{P}(A_j)) \\ &= 2\sum_{1\leq i<j\leq n} (\mathbb{P}(A_i \cap A_j) - \mathbb{P}(A_i)\mathbb{P}(A_j)) + \sum_{i=1}^n \mathbb{P}(A_i)(1 - \mathbb{P}(A_i)). \end{aligned}$$

Le première hypothèse se traduit par $\mathbb{E}[X_n] \to \infty$. Par conséquent,

$$\frac{\sum_{i=1}^n \mathbb{P}(A_i)(1-\mathbb{P}(A_i))}{\mathbb{E}[X_n]^2} \le \frac{\mathbb{E}[X_n]}{\mathbb{E}[X_n]^2} \to 0, \quad n\to\infty.$$

La seconde hypothèse se traduit alors par

$$\limsup_{n\to\infty} \frac{\mathrm{Var}(X_n)}{\mathbb{E}[X_n]^2} \le 0.$$

Mais puisque $\mathrm{Var}(X_n) \ge 0$ pour tout n, cela donne

$$\lim_{n\to\infty} \frac{\mathrm{Var}(X_n)}{\mathbb{E}[X_n]^2} = 0.$$

L'inégalité de Chebychev nous donne alors

$$\begin{aligned}\mathbb{P}(X_n \le \tfrac{1}{2}\mathbb{E}[X_n]) &\le \mathbb{P}(|X_n - \mathbb{E}[X_n]| \ge \tfrac{1}{2}\mathbb{E}[X_n])\\ &\le \frac{4\,\mathrm{Var}(X_n)}{\mathbb{E}[X_n]^2}\\ &\to 0, \quad n\to\infty.\end{aligned}$$

Puisque $\mathbb{E}[X_n] \to \infty$ quand $n\to\infty$, cela montre alors que pour tout $M\in\mathbb{N}$, $\mathbb{P}(X_n \le M) \to 0$ pour $n\to\infty$ et conclut la preuve.

Remarque : Il suffit en fait un « lim inf » au lieu du « lim sup » dans l'hypothèse du théorème (pourquoi ?).

Convergences de variables aléatoires : dans L^p, en probabilité, en loi

Exercice 35. Par un théorème du cours, une famille $(Z_i)_{i\in I}$ de v.a. dans $\mathbb{R}^d$ est uniformément intégrable ssi

$$\sup_{i\in I} \mathbb{E}[\|Z_i\|] < \infty \quad \text{et} \quad \lim_{\delta\to 0} \sup_{A:\mathbb{P}(A)<\delta} \sup_{i\in I} \mathbb{E}[\|Z_i\|\mathbb{1}_A] = 0.$$

On a alors par l'inégalité triangulaire et par l'intégrabilité uniforme de $(X_i)_{i\in I}$ et $(Y_i)_{i\in I}$,

$$\sup_{i\in I} \mathbb{E}[\|X_i+Y_i\|] \le \sup_{i\in I} \mathbb{E}[\|X_i\|] + \mathbb{E}[\|Y_i\|] \le \sup_{i\in I} \mathbb{E}[\|X_i\|] + \sup_{i\in I} \mathbb{E}[\|Y_i\|] < \infty.$$

De même, on a pour tout $\delta > 0$,

$$\begin{aligned}\sup_{A:\mathbb{P}(A)<\delta} \sup_{i\in I} \mathbb{E}[\|X_i+Y_i\|\mathbb{1}_A] &\le \sup_{A:\mathbb{P}(A)<\delta} \sup_{i\in I} \mathbb{E}[\|X_i\|\mathbb{1}_A] + \mathbb{E}[\|Y_i\|\mathbb{1}_A]\\ &\le \sup_{A:\mathbb{P}(A)<\delta} \sup_{i\in I} \mathbb{E}[\|X_i\|\mathbb{1}_A] + \sup_{A:\mathbb{P}(A)<\delta} \sup_{i\in I} \mathbb{E}[\|Y_i\|\mathbb{1}_A],\end{aligned}$$

si bien que par l'intégrabilité uniforme de $(X_i)_{i\in I}$ et $(Y_i)_{i\in I}$,

$$\lim_{\delta\to 0} \sup_{A:\mathbb{P}(A)<\delta} \sup_{i\in I} \mathbb{E}[\|X_i+Y_i\|\mathbb{1}_A] = 0.$$

Le théorème ci-dessus montre alors que $(X_i+Y_i)_{i\in I}$ est uniformément intégrable.

Exercice 36. On a pour tout événement $A \in \mathcal{A}$,

$$\sup_{i\in I} \mathbb{E}[\|f(X_i)\|\mathbb{1}_A] \le \sup_{i\in I} \mathbb{E}[C(\|X_i\|+1)\mathbb{1}_A] \le C\sup_{i\in I} \mathbb{E}[\|X_i\|\mathbb{1}_A] + C\mathbb{P}(A).$$

En posant $A = \Omega$, ceci donne $\sup_{i\in I} \|f(X_i)\|_1 \le C\sup_{i\in I}\|X_i\|_1 + C < \infty$ puisque la famille $(X_i)_{i\in I}$ est uniformément intégrable. De plus, on a

$$\lim_{\delta\to 0} \sup_{A\in\mathcal{A},\,\mathbb{P}(A)<\delta)} \mathbb{E}[\|f(X_i)\|\mathbb{1}_A] \le \lim_{\delta\to 0}\left(C \sup_{A\in\mathcal{A},\,\mathbb{P}(A)<\delta)} \sup_{i\in I} \mathbb{E}[\|X_i\|\mathbb{1}_A] + C\delta\right) = 0,$$

toujours parce que la famille $(X_i)_{i\in I}$ est uniformément intégrable. Par conséquent, la famille $(f(X_i))_{i\in I}$ l'est aussi.

Exercice 37. 1. $\Rightarrow$ 2. : Convergence dans L^p implique convergence en probabilité. Pour l'intégrabilité uniforme, on remarque que pour tout n et tout événement A, on a d'après l'inégalité de Minkowski,

$$\begin{aligned}(\mathbb{E}[\|X_n\|^p \mathbb{1}_A])^{1/p} &= \|X_n \mathbb{1}_A\|_p \\ &= \|(X_n - X + X)\mathbb{1}_A\|_p \\ &\le \|(X_n - X)\mathbb{1}_A\|_p + \|X\mathbb{1}_A\|_p \\ &\le \|X_n - X\|_p + (\mathbb{E}[\|X\|^p \mathbb{1}_A)^{1/p}.\end{aligned}$$

Par hypothèse, on a $\|X_n - X\|_p \to 0$ quand $n \to \infty$. En particulier, en utilisant l'inégalité précédente avec $A = \Omega$, on a $\sup_n \mathbb{E}[\|X_n\|^p] < \infty$.

Fixons $\varepsilon > 0$. Alors il existe $N \in \mathbb{N}^*$, tel que $\|X_n - X\|_p < \varepsilon/2$ pour tout $n \ge N$. De plus, puisque $\|X\|^p$ est intégrable par hypothèse, donc uniformément intégrable, il existe $\delta' > 0$, tel que $(\mathbb{E}[\|X\|^p \mathbb{1}_A)^{1/p} < \varepsilon/2$ pour tout événement A avec $\mathbb{P}(A) < \delta'$. Pour la même raison, il existe $\delta'' > 0$, tel que $(\mathbb{E}[\|X_n\|^p \mathbb{1}_A)^{1/p} < \varepsilon$ pour tout $n < N$ et tout événement A avec $\mathbb{P}(A) < \delta''$. Avec l'inégalité en haut, cela donne avec $\delta = \min(\delta', \delta'')$,

$$\sup_{A:\mathbb{P}(A)<\delta} \sup_{n\in\mathbb{N}^*} (\mathbb{E}[\|X_n\|^p \mathbb{1}_A])^{1/p} \le \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Par un théorème du cours, ceci montre que la famille $(\|X_n\|^p)_{n\in\mathbb{N}^*}$ est uniformément intégrable.

2. $\Rightarrow$ 1. : Montrons d'abord que $X \in L^p$: Puisque $X_n \to X$ en probabilité, il existe une sous-suite $(X_{n_k})_{k\in\mathbb{N}}$ telle que X_{n_k} converge p.s. vers X. En particulier, $\|X_{n_k}\|^p \to \|X\|^p$ p.s. Par le lemme de Fatou, on a alors

$$\mathbb{E}[\|X\|^p] \le \liminf_{k\to\infty} \mathbb{E}[\|X_{n_k}\|^p] < \infty,$$

par l'intégrabilité uniforme de la famille $(\|X_n\|^p)_{n\in\mathbb{N}^*}$. Ceci montre que $X \in L^p$.

Montrons que $X_n \to X$ dans L^p. On fixe $\varepsilon > 0$. Par l'inégalité de Minkowski, on a

$$\begin{aligned}\|X_n - X\|_p &\le \|(X_n - X)\mathbb{1}_{\|X_n - X\|\le\varepsilon}\|_p + \|(X_n - X)\mathbb{1}_{\|X_n - X\|>\varepsilon}\|_p \\ &\le \varepsilon + \|X_n \mathbb{1}_{\|X_n - X\|>\varepsilon}\|_p + \|X\mathbb{1}_{\|X_n - X\|>\varepsilon}\|_p\end{aligned}$$

Posons $X_0 = X$. La famille $(\|X_n\|^p)_{n\in\mathbb{N}}$ est alors uniformément intégrable puisque l'union de deux familles uniformément intégrables l'est encore. Par un théorème du cours, il existe alors $\delta > 0$, tel que pour tout événement A tel que $\mathbb{P}(A) < \delta$,

$$\sup_{n\in\mathbb{N}} \|X_n \mathbb{1}_A\|_p = (\mathbb{E}[\|X_n\|^p \mathbb{1}_A])^{1/p} < \varepsilon.$$

De plus, par la convergence en probabilité $X_n \to X$, il existe $N \in \mathbb{N}$, tel que pour tout $n \ge N$, $\mathbb{P}(\|X_n - X\| > \varepsilon) < \delta$. Par ce qui précède, on a finalement, pour $n \ge N$,

$$\|X_n - X\|_p \le \varepsilon + \varepsilon + \varepsilon = 3\varepsilon.$$

Puisque $\varepsilon > 0$ était arbitraire, ceci montre que $\|X_n - X\|_p \to 0$ et donc $X_n \to X$ dans L^p.

Exercice 38. Montrons que $|a - b| = a + b - 2\min(a, b)$ pour tous $a, b \in \mathbb{R}$: on peut supposer que $a \ge b$, sinon on échange les rôles de a et b. On a alors $|a - b| = a - b$ et $a + b - 2\min(a, b) = a + b - 2b = a - b$. Ceci montre l'égalité.

On a alors pour tout $n \in \mathbb{N}$,

$$\mathbb{E}[|X_n - X_\infty|] = \mathbb{E}[X_n] + \mathbb{E}[X_\infty] - 2\mathbb{E}[\min(X_n, X_\infty)].$$

Puisque $0 \le \min(X_n, X_\infty) \le X_\infty$ pour tout n et X_∞ est intégrable, on a par convergence dominée,

$$\mathbb{E}[\min(X_n, X_\infty)] \to \mathbb{E}[X_\infty].$$

De plus, par hypothèse, $\mathbb{E}[X_n] \to \mathbb{E}[X_\infty]$ quand $n \to \infty$. Il en suit que $\mathbb{E}[|X_n - X_\infty|] \to \mathbb{E}[X_\infty] + \mathbb{E}[X_\infty] - 2\mathbb{E}[X_\infty] = 0$, ce qui permet de conclure.

Exercice 39. *Calcul direct :* Soit $X_n \sim \mathrm{Bin}(n, p_n)$. On a pour tout $k \in \mathbb{N}$,

$$\begin{aligned}\mathbb{P}(X_n = k) &= \binom{n}{k} p_n^k (1-p_n)^{n-k} \\ &= \frac{1}{k!} \frac{n(n-1)\cdots(n-k+1)}{n^k} (np_n)^k \left(1 - \frac{np_n}{n}\right)^{n-k} \\ &\overset{n\to\infty}{\longrightarrow} \frac{1}{k!} \lambda^k e^{-\lambda}.\end{aligned}$$

Ceci montre que $X_n \Rightarrow \mathrm{Po}(\lambda)$.

Fonctions génératrices : On a pour $|z| \le 1$,

$$g_{\mathrm{Bin}(n,p_n)}(z) = (1 - p_n + p_n z)^n = \left(1 - \frac{(z-1)np_n}{n}\right)^n \overset{n\to\infty}{\longrightarrow} e^{(z-1)\lambda}.$$

Ceci est la fonction génératrice de la loi de Poisson de paramètre λ. Par le théorème de Lévy (ou alors par le simple fait que convergence d'une suite de séries implique convergence des coefficients), on en déduit que $X_n \Rightarrow \mathrm{Po}(\lambda)$.

Exercice 40. *Calcul direct :* On a pour tout $x > 0$,

$$\mathbb{P}(pX_p > x) = \mathbb{P}(X_p \ge \lfloor x/p \rfloor + 1) = p \sum_{k=\lfloor x/p\rfloor+1}^{\infty} (1-p)^{k-1} = (1-p)^{\lfloor x/p \rfloor} p \sum_{k=0}^{\infty} (1-p)^k \overset{p\to 0}{\longrightarrow} e^{-x},$$

car $p\sum_{k=0}^\infty (1-p)^k = 1$. Ceci montre que la fonction de répartition de pX_p tend vers celle de la loi exponentielle de paramètre 1 quand $p \to 0$, ce qui entraine la convergence en loi de pX_p vers cette loi exponentielle.

Fonctions caractéristiques : On sait que la fonction génératrice de X_p s'écrit

$$\mathbb{E}[z^{X_p}] = \frac{pz}{1-(1-p)z}, \quad |z| \le 1.$$

On applique cela avec $z = e^{i\lambda p}$, $\lambda \in \mathbb{R}$, pour obtenir

$$\mathbb{E}[e^{i\lambda p X_p}] = e^{i\lambda p} \frac{p}{1-(1-p)e^{i\lambda p}} = e^{i\lambda p} \frac{p}{1-(1-p)(1+i\lambda p + o(p))} \overset{p\to 0}{\longrightarrow} \frac{1}{1-i\lambda}.$$

On identifie cela avec la fonction caractéristique de la loi exponentielle de paramètre 1. Par le théorème de Lévy, on en déduit que $pX_p \Rightarrow \mathrm{Exp}(1)$.

Exercice 41. On a pour tout $x \in \mathbb{R}$,

$$\begin{aligned}\mathbb{P}(\max(X_1,\dots,X_n) - \log n \le x) &= \mathbb{P}(\max(X_1,\dots,X_n) \le x + \log n) \\ &= \mathbb{P}(X_1 \le x + \log n)^n \\ &\overset{\log n \ge -x}{=} (1 - e^{-x-\log n})^n = \left(1 - \frac{e^{-x}}{n}\right)^n \to e^{-e^{-x}}, \quad n \to \infty.\end{aligned}$$

Puisque convergence des fonctions de répartitions implique convergence en loi, ceci montre que $\max(X_1,\dots,X_n) - \log n$ converge en loi vers la loi de fonction de répartition $x \mapsto e^{-e^{-x}}$ (appelée la loi de *Gumbel*).

Exercice 42. Soit g une fonction continue bornée. Alors la composée $g \circ f$ est encore continue et bornée. Par conséquent,

$$\mathbb{E}[g(f(X_n))] = \mathbb{E}[(g\circ f)(X_n)] \to \mathbb{E}[(g \circ f)(X)] = \mathbb{E}[g(f(X))],$$

quand $n \to \infty$, par la convergence en loi de X_n vers X. Puisque g était arbitraire, ceci implique que $f(X_n) \Rightarrow f(X)$.

Exercice 43. Par le théorème de représentation de Skorokhod, on peut construire les v.a. $X_1, X_2, \dots$ et X sur un même espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ tel que $X_n \to X$ p.s. Fixons f comme dans l'énoncé et notons A l'événement que f est continue en X. Alors par hypothèse et le théorème de transfert, on a $\mathbb{P}(A) = 1$. Notons en plus C l'événement que $X_n \to X$ quand $n \to \infty$, si bien que $\mathbb{P}(C) = 1$. On a alors $f(X_n) \to f(X)$ sur $A \cap C$. Par le théorème de convergence dominée, on a alors, quand $n \to \infty$, pour toute fonction g continue bornée,

$$\mathbb{E}[g(f(X_n))] = \mathbb{E}[g(f(X_n))\mathbb{1}_{A\cap C}] \to \mathbb{E}[g(f(X))\mathbb{1}_{A\cap C}] = \mathbb{E}[g(f(X))].$$

Ceci montre que $f(X_n) \to f(X)$ en loi.

Exercice 44. Soit $\lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{R}^k$. On a

$$\begin{aligned}
\mathbb{E}[e^{i\langle\lambda,(X_n^{(1)},\ldots,X_n^{(k)})\rangle}] &= \mathbb{E}[e^{i\lambda_1 X_n^{(1)}} \cdots e^{i\lambda_k X_n^{(k)}}] \\
&= \mathbb{E}[e^{i\lambda_1 X_n^{(1)}}] \cdots \mathbb{E}[e^{i\lambda_k X_n^{(k)}}] && \text{par indépendance} \\
&\xrightarrow{n\to\infty} \mathbb{E}[e^{i\lambda_1 X^{(1)}}] \cdots \mathbb{E}[e^{i\lambda_k X^{(k)}}] && \text{par hypothèse} \\
&= \mathbb{E}[e^{i\lambda_1 X^{(1)}} \cdots e^{i\lambda_k X^{(k)}}] && \text{par indépendance} \\
&= \mathbb{E}[e^{i\langle\lambda,(X^{(1)},\ldots,X^{(k)})\rangle}].
\end{aligned}$$

Le théorème de Lévy donne alors la convergence en loi de $(X_n^{(1)}, \ldots, X_n^{(k)})$ vers $(X^{(1)}, \ldots, X^{(k)})$ quand $n \to \infty$.

Exercice 45 (Lois stables)**.** 1. On a $|e^{i\lambda x} - 1| \le 2$ et $|i\lambda x| = \lambda|x|$ pour tout $x, \lambda \in \mathbb{R}$. Pour l'intégrabilité, il suffit alors de montrer que $\int_{\mathbb{R}}(1+|x|)f(x)\,dx < \infty$. Mais,

$$\begin{aligned}
\int_{\mathbb{R}}(1+|x|)f(x)\,dx &= 1 + \int_{\mathbb{R}} |x|f(x)\,dx && \text{car } \int_{\mathbb{R}} f(x)\,dx = 1 \\
&= 1 + 2\int_0^\infty xf(x)\,dx && f \text{ paire} \\
&\le 1 + 2(1 + \int_1^\infty \frac{c_\alpha}{x^\alpha}\,dx) && \text{car } \int_0^1 f(x)\,dx \le 1 \\
&< \infty && \text{puisque } \alpha > 1.
\end{aligned}$$

En ce qui concerne l'expression pour φ_X, on note encore que $\int_{\mathbb{R}} f(x)\,dx = 1$ et aussi $\int_{\mathbb{R}} xf(x)\,dx = 0$, puisque f est paire. Par conséquent,

$$\varphi_X(\lambda) - 1 = \int_{\mathbb{R}} e^{i\lambda x}f(x)\,dx - \int_{\mathbb{R}} f(x)\,dx - \int_{\mathbb{R}} i\lambda x f(x)\,dx = \int_{\mathbb{R}}(e^{i\lambda x} - 1 - i\lambda x)f(x)\,dx.$$

2. Soit $\lambda \neq 0$. En faisant un changement de variable $x \mapsto x/\lambda$, on obtient

$$\begin{aligned}
\varphi_X(\lambda) - 1 &= \int_{\mathbb{R}}(e^{ix} - 1 - ix)\frac{c_\alpha}{1 + |x/\lambda|^{\alpha+1}}\frac{dx}{|\lambda|} \\
&= |\lambda|^\alpha \int_{\mathbb{R}}(e^{ix} - 1 - ix)\frac{c_\alpha}{|\lambda|^{\alpha+1} + |x|^{\alpha+1}}\,dx \\
&= |\lambda|^\alpha \int_{\mathbb{R}} g_\lambda(x)\,dx, \quad \text{avec } g_\lambda(x) = (e^{ix} - 1 - ix)\frac{c_\alpha}{|\lambda|^{\alpha+1} + |x|^{\alpha+1}}.
\end{aligned}$$

Alors $g_\lambda(x) \to g_0(x) = (e^{ix} - 1 - ix)\frac{c_\alpha}{|x|^{\alpha+1}}$ quand $\lambda \to 0$. De plus, puisque $|e^{ix} - 1 - ix| = O(x^2)$ quand $x \to 0$ et $= O(x)$ quand $x \to \infty$, il existe une constante C telle que

$$\forall\lambda \in \mathbb{R} : |g_\lambda(x)| \le C\min(|x|^2, |x|)\frac{1}{|x|^{\alpha+1}} = C\min(|x|^{-(\alpha-1)}, |x|^{-\alpha}) =: G(x).$$

Puisque $\alpha \in (1,2)$, on a $\int_1^\infty x^{-\alpha}\,dx < \infty$ et $\int_0^1 x^{-(\alpha-1)}\,dx < \infty$, si bien que $G(x)$ est intégrable. Le théorème de convergence dominée donne alors

$$\frac{\varphi_X(\lambda) - 1}{|\lambda|^\alpha} = \int_{\mathbb{R}} g_\lambda(x)\,dx \to \int_{\mathbb{R}} g_0(x)\,dx, \quad \lambda \to 0.$$

Il reste à montrer que $\int_{\mathbb{R}} g_0(x)\,dx$ est un réel strictement négatif. En effet, puisque $g_0(-x) = g_0(x)^*$, on a $\int g(x)\,dx = 2\Re\int_0^\infty g(x), dx$. De plus, puisque $\Re e^{ix} \le 1$, on a $\Re g(x) \le 0$ pour tout x et même < 0 pour Lebesgue-presque tout x. Par conséquent, $\int g_0(x)\,dx$ est bien un réel strictement négatif.

3. Par indépendance des X_n, on a

$$\mathbb{E}[e^{i\lambda S_n/n^{1/\alpha}}] = \varphi(\lambda/n^{1/\alpha})^n = \exp(n\log\varphi(\lambda/n^{1/\alpha})).$$

Puisque $\varphi(\lambda) \to 1$ quand $\lambda \to 0$, on a l'équivalent

$$\log(\varphi(\lambda/n^{1/\alpha})) \sim (\varphi(\lambda/n^{1/\alpha}) - 1) \sim -\frac{C|\lambda|^\alpha}{n}, \quad n \to \infty.$$

Il vient que

$$\mathbb{E}[e^{i\lambda S_n/n^{1/\alpha}}] \to \exp(-C|\lambda|^\alpha), \quad n \to \infty.$$

4. La fonction $\lambda \to \exp(-C|\lambda|^\alpha)$ est une limite de fonctions caractéristiques de lois de probabilité sur $\mathbb{R}$. De plus, elle est continue en 0. Par le théorème de Lévy, elle est alors la fonction caractéristique d'une loi de probabilité sur $\mathbb{R}$.
5. Pour tout $a, b > 0$, on obtient
$$\mathbb{E}[e^{i\lambda(aY+bY')}] = \exp(-|a\lambda|^\alpha - |b\lambda|^\alpha) = \exp(-|(a^\alpha + b^\alpha)^{1/\alpha}\lambda|^\alpha) = \mathbb{E}[e^{i\lambda(a^\alpha+b^\alpha)^{1/\alpha}Y)}].$$
Puisque les fonctions caractéristiques déterminent les lois, cela montre l'égalité en loi
$$aY + bY' \overset{\text{loi}}{=} (a^\alpha + b^\alpha)^{1/\alpha} Y$$
Les lois 2-stables que nous connaissons sont les lois gaussiennes centrées.

Exercice 46. On a
$$\begin{aligned} A_{2n} - A_n = \frac{S_{2n}}{\sqrt{2n}} - \frac{S_n}{\sqrt{n}} &= \frac{(X_1 + \cdots + X_{2n})/\sqrt{2} - (X_1 + \cdots + X_n)}{\sqrt{n}} \\ &= \left(\tfrac{1}{\sqrt{2}} - 1\right) \frac{X_1 + \cdots + X_n}{\sqrt{n}} + \frac{1}{\sqrt{2}} \frac{X_{n+1} + \cdots + X_{2n}}{\sqrt{n}}. \end{aligned}$$
Les deux sommands de la dernière ligne sont indépendants et convergent en loi vers des gaussiennes centrées de variances respectives $\left(\frac{1}{\sqrt{2}} - 1\right)^2$ et 1/2, d'après le TCL. Par les exercices 9 et 8, $A_{2n} - A_n$ converge alors en loi vers une gaussienne de variance $\left(\frac{1}{\sqrt{2}} - 1\right)^2 + 1/2$, donc ne peut converger vers 0 en loi, ni en probabilité (car la convergence en probabilité implique convergence en loi). Ceci implique que $A_n = S_n/\sqrt{n}$ ne converge pas en probabilité, car si Y_n est une suite de v.a. qui converge en probabilité vers une v.a. Y, alors $|Y_{2n} - Y_n| \le |Y_{2n} - Y| + |Y_n - Y|$, ce qui converge vers 0 en probabilité.

Exercice 47. Par le théorème central limite, on sait que $\frac{S_n - \mu n}{\sqrt{n}} \Rightarrow \mathcal{N}(0, \mathrm{Var}(X_1))$, quand $n \to \infty$. Par l'exercice 8 de cette feuille, on en déduit que $\left(\frac{S_n - \mu n}{\sqrt{n}}\right)^2$ converge également en loi vers N^2, ou $N \sim \mathcal{N}(0, \mathrm{Var}(X_1))$. Par le théorème de représentation de Skorokhod, on peut construire un certain espace de probabilité avec des v.a. $Y_1, Y_2, \ldots$ et Y telles que Y_n est égale en loi à $\left(\frac{S_n - \mu n}{\sqrt{n}}\right)^2$ pour tout n, Y est égale en loi à N^2, et $Y_n \to Y$ presque sûrement quand $n \to \infty$. De plus, on a pour tout n,
$$\mathbb{E}[Y_n] = \mathbb{E}\left[\left(\frac{S_n - \mu n}{\sqrt{n}}\right)^2\right] = \mathrm{Var}\left(\frac{S_n}{\sqrt{n}}\right) = \mathrm{Var}(X_1) = \mathbb{E}[N^2] = \mathbb{E}[Y],$$
et donc, *a fortiori*, $\mathbb{E}[Y_n] \to \mathbb{E}[Y]$. Par le lemme de Scheffé, on a alors $Y_n \to Y$ dans L^1. En particulier, la famille $(Y_n)_{n\in\mathbb{N}^*}$ est uniformément intégrable. Puisque l'intégrabilité uniforme ne dépend que des lois marginales de la famille, on en déduit que la famille de v.a. $\left(\frac{S_n - \mu n}{\sqrt{n}}\right)^2$ est également uniformément intégrable.

Exercice 48 (Problème des allumettes de Banach). 1. Notons 0 la boîte dans la poche de gauche et l'autre, 1. On formalise le problème de manière suivante : on suppose qu'on fait une infinité de tirages, c'est-à-dire on se donne une suite iid de v.a. $(B_n)_{n\ge 1}$ de lois Ber(1/2). Ici, B_n donne la boîte de laquelle on tire la n-ième allumette. Soit N_k^0 l'index du k-ième 0 et N_k^1 l'index du k-ième 1. Notons $Y_k^i = N_k^{1-i} - N_{k-1}^{1-i}$ et notons que $(Y_k^i)_{k\ge 1}$ est une suite iid de v.a. de loi géométrique de paramètre 1/2 pour chaque $i = 0, 1$. L'événement A_n^i que le fumeur ait pris n allumettes de la boîte $1 - i$ avant d'en prendre n de la boîte i s'écrit alors
$$A_n^i = \{N_n^{1-i} < 2n\} = \left\{\sum_{k=1}^n Y_k^i < 2n\right\}, \quad i = 0, 1.$$
Notons que les deux événements sont disjoints et leur union est Ω, car
$$\begin{aligned} A_n^i &= \{N_n^{1-i} < 2n\} \\ &= \{\mathrm{Card}\{1 \le k < 2n : B_k = 1 - i\} \ge n\} \\ &= \{\mathrm{Card}\{1 \le k < 2n : B_k = i\} < n\} \\ &= \{N_n^i \ge 2n\} \\ &= (A_n^{1-i})^c. \end{aligned}$$

Puisque $N_n^{1-i} - n$ est le nombre d'allumettes prisent dans la boite i au moment ou la n-ième allumette est prise de la boîte $1-i$, la variable recherchée X_n est alors $X_n = 2n - N_n^{1-i} = 2n - \sum_{k=1}^n Y_k^i$ sur l'événement A_n^i.

2. On a pour toute fonction continue bornée $f : \mathbb{R} \to \mathbb{R}$,

$$\begin{aligned}\mathbb{E}[f(X_n)] &= \mathbb{E}[f(X_n)\mathbb{1}_{A_n^0}] + \mathbb{E}[f(X_n)\mathbb{1}_{A_n^1}] \\ &= \mathbb{E}\left[f\left(2n - \sum_{k=1}^n Y_k^0\right)\mathbb{1}_{2n-\sum_{k=1}^n Y_k^0>0}\right] + \mathbb{E}\left[f\left(2n - \sum_{k=1}^n Y_k^1\right)\mathbb{1}_{2n-\sum_{k=1}^n Y_k^1>0}\right] \\ &= 2\mathbb{E}\left[f\left(2n - \sum_{k=1}^n Y_k^0\right)\mathbb{1}_{2n-\sum_{k=1}^n Y_k^0>0}\right],\end{aligned}$$

car les suites $(Y_k^0)_{k\in\mathbb{N}}$ et $(Y_k^1)_{k\in\mathbb{N}}$ ont même loi. En particulier,

$$\mathbb{E}[f(X_n/\sqrt{n})] = 2\mathbb{E}\left[f\left(\frac{2n-\sum_{k=1}^n Y_k^0}{\sqrt{n}}\right)\mathbb{1}_{\frac{2n-\sum_{k=1}^n Y_k^0}{\sqrt{n}}>0}\right].$$

Pour établir la convergence, on utilise le TCL. On a $\mathbb{E}[Y_1^0] = 2$ et $\mathrm{Var}(Y_1^0) = (1-1/2)/(1/2)^2 = 2$. Par le TCL, $\frac{2n-\sum_{k=1}^n Y_k^0}{\sqrt{n}}$ converge alors en loi quand $n\to\infty$ vers $N \sim \mathcal{N}(0,2)$.
Notons qu'on peut supposer que $f(0) = 0$, quitte à considérer la fonction $f - f(0)$. Alors la fonction $x \mapsto f(x)\mathbb{1}_{x>0}$ est encore continue et bornée. On obtient alors

$$\lim_{n\to\infty} \mathbb{E}[f(X_n/\sqrt{n})] = 2\mathbb{E}[f(N)\mathbb{1}_{N>0}] = \mathbb{E}[f(|N|)],$$

par symétrie de la loi gaussienne. On en conclut que $X_n/\sqrt{n}$ converge en loi vers $|N|$.

3. D'après l'exercice précédent, la famille de v.a. $\left(\left(\frac{2n-\sum_{k=1}^n Y_k^{(1)}}{\sqrt{n}}\right)^2\right)_{n\in\mathbb{N}}$ est u.i., et donc, d'après la preuve de la partie précédente, la famille $\left((X_n/\sqrt{n})^2\right)_{n\in\mathbb{N}}$ également. Par l'exercice 2 de cette feuille (appliqué à $f(x) = \sqrt{x}\mathbb{1}_{x\ge1}$) la famille $(X_n/\sqrt{n})_{n\in\mathbb{N}}$ l'est encore. Par conséquent, on a

$$\lim_{n\to\infty} \mathbb{E}[X_n/\sqrt{n}] = \mathbb{E}[|N|] = \frac{2}{\sqrt{4\pi}}\int_0^\infty xe^{-x^2/4}\,dx = \frac{2}{\sqrt{\pi}}[-e^{-x^2/4}]_0^\infty = \frac{2}{\sqrt{\pi}}.$$

Donc, $\mathbb{E}[X_n] \sim (2/\sqrt{\pi})\sqrt{n}$ quand $n\to\infty$.

Exercice 49. On définit la matrice V de dimensions $k\times n$ par

$$V = (v_1|\cdots|v_k)^T.$$

On cherche alors la loi du vecteur VX. Puisque X est un vecteur gaussien, VX est également un vecteur gaussien. Son espérance vaut $\mathbb{E}[VX] = V\mathbb{E}[X] = V\mu$. Quant à la matrice de covariance, on note d'abord que

$$\Sigma_X V^T = \Sigma_X(v_1|\cdots|v_k) = (\lambda_1 v_1|\cdots|\lambda_k v_k) = V^T\Lambda,$$

où

$$\Lambda = \mathrm{diag}(\lambda_1,\dots,\lambda_k).$$

De plus, on a $VV^T = \mathrm{Id}$ puisque la famille $v_1,\dots,v_k$ est orthonormale. Par conséquent,

$$\Sigma_{VX} = V\Sigma_X V^T = VV^T\Lambda = \Lambda.$$

On conclut alors que $VX \sim \mathcal{N}(V\mu,\Lambda)$.

Exercice 50. Soient $p_1,\dots,p_k > 0$ tels que $\sum_{i=1}^k p_i = 1$. On note Y une v.a. sur $\{1,\dots,k\}$ de loi $\mathbb{P}(Y=i) = p_i$ et on pose $X = (\mathbb{1}_{Y=1},\dots,\mathbb{1}_{Y=k})^T$, si bien que X suit la loi multinomiale de paramètres 1 et $p_1,\dots,p_k$.

1. Le vecteur aléatoire X a moyenne $\mathbb{E}[X] = (p_1, \dots, p_k)^T$. Calculons sa matrice de covariance : on a pour $i = 1, \dots, k$,
$$\mathrm{Var}(\mathbb{1}_{Y=i}) = p_i(1 - p_i),$$
car $\mathbb{1}_{Y=i}$ est une v.a. de Bernoulli de paramètre p_i. De plus, on a pour $i \neq j$,
$$\mathrm{Cov}(\mathbb{1}_{Y=i}, \mathbb{1}_{Y=j}) = \mathbb{E}[\mathbb{1}_{Y=i}\mathbb{1}_{Y=j}] - \mathbb{E}[\mathbb{1}_{Y=i}]\mathbb{E}[\mathbb{1}_{Y=j}] = 0 - p_i p_j = -p_i p_j.$$
La matrice de covariance de X est alors donnée par
$$\Sigma = (\Sigma_{ij})_{i,j=1}^k, \quad \Sigma_{ij} = \begin{cases} p_i(1-p_i) & i = j \\ -p_i p_j & i \neq j \end{cases}$$
Le TCL multidimensionnel donne alors convergence en loi de $(\sum_{i=1}^n X_i - (np_1, \dots, np_k)^T)/\sqrt{n}$ vers le vecteur gaussien Z de moyenne $\mu = 0$ et de matrice de covariance Σ.
2. Le vecteur $\mathbf{1} = (1, \dots, 1)^T$ est de valeur propre zéro car pour tout i, on a $p_i - \sum_{j=1}^k p_i p_j = p_i - p_i = 0$, où on a utilisé le fait que $\sum_{j=1}^k p_j = 1$. Ceci implique que $\mathrm{Var}(\langle \mathbf{1}, Z\rangle) = \mathrm{Var}(\mathbf{1}^T Z) = \mathbf{1}^T \Sigma \mathbf{1} = 0$, et puisque $\mathbb{E}[Z] = 0$ on a $\langle \mathbf{1}, Z\rangle = 0$ p.s. Donc $Z \in \mathbf{1}^\perp$ p.s.
3. On a $\Sigma = \mathrm{diag}(p_1, \dots, p_k) - (p_i p_j)_{i,j=1}^n$, et donc avec I la matrice d'identité,
$$D\Sigma D = I - (\sqrt{p_i p_j})_{i,j=1}^n = I - \mathbf{p}\mathbf{p}^T.$$
Puisque $\mathbf{p}^T\mathbf{p} = \|\mathbf{p}\|_2^2 = 1$, on a alors $D\Sigma D\mathbf{p} = 0$. De plus $D\Sigma D v = v$ pour tout $v \in \mathbf{p}^\perp$. Par conséquent, $D\Sigma D$ est la projection orthogonale sur l'espace $\mathbf{p}^T$.
4. La loi limite de $\|D(\sum_{i=1}^n X_i - (np_1, \dots, np_k)^T)/\sqrt{n}\|_2^2$ est la loi de $\|DZ\|_2^2$ par le lemme de l'application continue. Or, DZ est un vecteur gaussien de moyenne nulle et de matrice de covariance $D\Sigma D$. On rappelle de la dernière partie que $D\Sigma D$ est la projection orthogonale sur $\mathbf{p}^\perp$. Notons $e_1, \dots, e_{k-1}$ une b.o.n. de $\mathbf{p}^\perp$. Définissons la matrice
$$O = (e_1 | \cdots | e_{k-1} | \mathbf{p})^T.$$
Alors par l'Exercice 15 de cette feuille, ODZ est un vecteur gaussien de moyenne nulle et de matrice de covariance
$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 0 \end{pmatrix}.$$
Par conséquent, $\|ODZ\|_2^2$ est la somme de $k-1$ v.a. iid de loi $\mathcal{N}(0,1)$ et donc $\|ODZ\|_2^2 \sim \Gamma(\frac{k-1}{2}, \frac{1}{2})$. Mais puisque O est une transformation orthogonale, ceci donne
$$\|DZ\|_2^2 = \|ODZ\|_2^2 \sim \Gamma(\tfrac{k-1}{2}, \tfrac{1}{2}).$$
Cette loi est également appelée la *loi du chi-deux de $k-1$ degrés de liberté* et notée $\chi^2(k-1)$.

Espérance conditionnelle

Exercice 51. On remarque d'abord que $\mathbb{E}[X|\mathcal{F}]$ satisfait l'égalité parce que $\mathbb{1}_C$ est $\mathcal{F}$-mesurable pour tout $C \in \mathcal{C}$. Supposons maintenant que Y est une v.a. intégrable et $\mathcal{F}$-mesurable telle que $\mathbb{E}[X\mathbb{1}_C] = \mathbb{E}[Y\mathbb{1}_C]$ pour tout $C \in \mathcal{C}$. Il suffit de montrer que cela implique $\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A]$ pour tout $A \in \mathcal{F}$, ce qui impliquerait que $Y = \mathbb{E}[X|\mathcal{F}]$. Pour cela, on pose
$$\mathcal{D} = \{A \in \mathcal{F} : \mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A]\}.$$
On a alors $\mathcal{D} \supset \mathcal{C}$. On montre que $\mathcal{D}$ est une classe monotone :

1. Par hypothèse, $\Omega \in \mathcal{C} \subset \mathcal{D}$.

2. Soient $A, B \in \mathcal{D}$, $A \subset B$. On a alors

$$\mathbb{E}[X\mathbb{1}_{B\setminus A}] = \mathbb{E}[X\mathbb{1}_A] - \mathbb{E}[X\mathbb{1}_B] = \mathbb{E}[Y\mathbb{1}_A] - \mathbb{E}[Y\mathbb{1}_B] = \mathbb{E}[Y\mathbb{1}_{B\setminus A}],$$

donc $B\setminus A \in \mathcal{D}$.

3. Soient $A_1, A_2, \ldots \in \mathcal{D}$, avec $A_n \subset A_{n+1}$ pour tout $n \in \mathbb{N}$. On note que $|X\mathbb{1}_{A_n}| \le |X| \in L^1$ et $|Y\mathbb{1}_{A_n}| \le |Y| \in L^1$ pour tout n. On a alors par convergence dominée,

$$\mathbb{E}[X\mathbb{1}_{\bigcup_n A_n}] = \mathbb{E}[\lim_{n\to\infty} X\mathbb{1}_{A_n}] = \lim_{n\to\infty} \mathbb{E}[X\mathbb{1}_{A_n}] = \lim_{n\to\infty} \mathbb{E}[Y\mathbb{1}_{A_n}] = \mathbb{E}[\lim_{n\to\infty} Y\mathbb{1}_{A_n}] = \mathbb{E}[Y\mathbb{1}_{\bigcup_n A_n}].$$

Donc $\bigcup_n A_n \in \mathcal{D}$.

On conclut que $\mathcal{D}$ est une classe monotone. Puisque $\mathcal{C}$ est stable par intersection finie (par hypothèse), le lemme des classes monotones donne alors $\mathcal{D} \supset \sigma(\mathcal{C}) = \mathcal{F}$. Ceci montre que $\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A]$ pour tout $A \in \mathcal{F}$. CQFD.

Exercice 52. 1. On sait que la fonction $g(x) = \mathbb{E}[f(x,Y)] = \int f(x,y)\,\mu_Y(dy)$ est mesurable, donc $g(X)$ est une v.a. $\sigma(X)$-mesurable. Il suffit alors de vérifier que $g(X)$ satisfait à la propriété caractéristique. Soit W une v.a. bornée $\sigma(X)$-mesurable. Par le lemme crucial il existe alors une fonction bornée mesurable h telle que $W = h(X)$. On a alors

$$\begin{aligned}
\mathbb{E}[h(X)g(X)] &= \int h(x)g(x)\,\mu_X(dx)\\
&= \int h(x)\left[\int f(x,y)\,\mu_Y(dy)\right]\mu_X(dx)\\
&= \int h(x)f(x,y)(\mu_X\otimes\mu_Y)(dx,dy) && \text{(par Fubini)}\\
&= \mathbb{E}[h(X)f(X,Y)].
\end{aligned}$$

Ceci montre que $g(X)$ satisfait à la propriété caractéristique et donc $g(X) = \mathbb{E}[f(X,Y)\,|\,X]$ p.s.

2. On a pour tout $y > 0$:

$$\begin{aligned}
\mathbb{E}\left[\frac{X}{X+y}\right] &= \int_0^1 \frac{x}{x+y}\,dx\\
&= 1 - y\int_0^1 \frac{1}{x+y}\,dx\\
&= 1 - y(\ln(1+y) - \ln(y))\\
&= 1 - y\ln\left(\frac{1+y}{y}\right).
\end{aligned}$$

La première partie de l'exercice et l'indépendance des deux variables aléatoires donnent alors

$$\mathbb{E}\left[\frac{X}{X+Y}\,\middle|\,Y\right] = 1 - Y\ln\left(\frac{1+Y}{Y}\right).$$

Exercice 53. On définit la classe

$$\mathcal{C} = \{B_1 \cap B_2 : B_i \in \mathcal{B}_i,\ i = 1,2\}.$$

La classe $\mathcal{C}$ est alors stable par intersection finie et engendre la tribu $\mathcal{B}_1\vee\mathcal{B}_2$. De plus, pour tout $C = B_1\cap B_2 \in \mathcal{C}$, $B_i \in \mathcal{B}_i,\ i = 1,2$, on a

$$\begin{aligned}
\mathbb{E}[Y\mathbb{1}_C] &= \mathbb{E}[Y\mathbb{1}_{B_1}\mathbb{1}_{B_2}]\\
&= \mathbb{E}[Y\mathbb{1}_{B_1}]\mathbb{E}[\mathbb{1}_{B_2}] && \text{par indépendance de } \sigma(Y)\vee\mathcal{B}_1 \text{ et } \mathcal{B}_2\\
&= \mathbb{E}[\mathbb{E}[Y|\mathcal{B}_1]\mathbb{1}_{B_1}]\mathbb{E}[\mathbb{1}_{B_2}]\\
&= \mathbb{E}[\mathbb{E}[Y|\mathcal{B}_1]\mathbb{1}_{B_1}\mathbb{1}_{B_2}] && \text{par indépendance de } \sigma(Y)\vee\mathcal{B}_1 \text{ et } \mathcal{B}_2\\
&= \mathbb{E}[\mathbb{E}[Y|\mathcal{B}_1]\mathbb{1}_C].
\end{aligned}$$

L'exercice 1 montre alors que $\mathbb{E}[Y|\mathcal{B}_1] = \mathbb{E}[Y|\mathcal{B}_1\vee\mathcal{B}_2]$.

Exercice 54. 1. Si $\mathcal{G} = \{0, \Omega\}$, cela signifie que $\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$ pour toutes fonctions f, g positives mesurables, donc que X et Y sont indépendantes. Si $\mathcal{G} = \mathcal{F}$, l'égalité est toujours vérifiée, car

$$\mathbb{E}[f(X)g(Y)|\mathcal{F}] = f(X)\mathbb{E}[g(Y)|\mathcal{F}] = \mathbb{E}[f(X)|\mathcal{F}]\mathbb{E}[g(Y)|\mathcal{F}],$$

puisque $f(X)$ et $g(Y)$ sont $\mathcal{F}$-mesurables.

2. Si $\mathbb{E}[f(X)g(Y)|\mathcal{G}] = \mathbb{E}[f(X)|\mathcal{G}]\mathbb{E}[g(Y)|\mathcal{G}]$, alors on a pour toute v.a. $Z \geq 0$, $\mathcal{G}$-mesurable,

$$\mathbb{E}[f(X)g(Y)Z] = \mathbb{E}[\mathbb{E}[f(X)g(Y)|\mathcal{G}]Z] = \mathbb{E}[\mathbb{E}[f(X)|\mathcal{G}]\mathbb{E}[g(Y)|\mathcal{G}]Z] = \mathbb{E}[f(X)\mathbb{E}[g(Y)|\mathcal{G}]Z],$$

où dans la première égalité on a appliqué la propriété caractéristique à $\mathbb{E}[f(X)g(Y)|\mathcal{G}]$, et dans la dernière égalité à $\mathbb{E}[f(X)|\mathcal{G}]$.

Inversement, pour toute v.a. $Z \geq 0$, $\mathcal{G}$-mesurable, si on a $\mathbb{E}[f(X)g(Y)Z] = \mathbb{E}[f(X)\mathbb{E}[g(Y)|\mathcal{G}]Z]$ alors on a

$$\mathbb{E}[f(X)g(Y)Z] = \mathbb{E}[f(X)\mathbb{E}[g(Y)|\mathcal{G}]Z] = \mathbb{E}[\mathbb{E}[f(X)|\mathcal{G}]\mathbb{E}[g(Y)|\mathcal{G}]Z].$$

Ceci montre que la v.a. $\mathcal{G}$-mesurable $\mathbb{E}[f(X)|\mathcal{G}]\mathbb{E}[g(Y)|\mathcal{G}]$ satisfait la propriété caractéristique pour $f(X)g(Y)$ et $\mathcal{G}$. Ceci implique que $\mathbb{E}[f(X)g(Y)|\mathcal{G}] = \mathbb{E}[f(X)|\mathcal{G}]\mathbb{E}[g(Y)|\mathcal{G}]$.

Supposons maintenant que $\mathbb{E}[g(Y)|\sigma(\mathcal{G}, \sigma(X))] = \mathbb{E}[g(Y)|\mathcal{G}]$. On a alors pour toute v.a. $Z \geq 0$, $\mathcal{G}$-mesurable, $\mathbb{E}[f(X)g(Y)Z] = \mathbb{E}[f(X)Z\mathbb{E}[g(Y)|\sigma(\mathcal{G}, \sigma(X))]] = \mathbb{E}[f(X)Z\mathbb{E}[g(Y)|\mathcal{G}]]$, ce qui donne la deuxième égalité. Inversement, supposons que pour toutes fonctions $f, g \geq 0$ mesurables et pour toute v.a. $Z \geq 0$, $\mathcal{G}$-mesurable, on ait $\mathbb{E}[f(X)g(Y)Z] = \mathbb{E}[f(X)Z\mathbb{E}[g(Y)|\mathcal{G}]]$. Par le "lemme crucial", chaque v.a. $W \geq 0$, $\sigma(X)$-mesurable, s'écrit que $f(X)$ pour une fonction $f \geq 0$ mesurable. En prenant $W = \mathbb{1}_B$ pour $B \in \sigma(X)$ et $Z = \mathbb{1}_C$ pour $C \in \mathcal{G}$, on a alors avec $A = B \cap C$, $\mathbb{E}[g(Y)\mathbb{1}_A] = \mathbb{E}[\mathbb{E}[g(Y)|\mathcal{G}]\mathbb{1}_A]$. La classe d'ensemble

$$\mathcal{C} = \{B \cap C : B \in \sigma(X),\ C \in \mathcal{G}\}$$

contient Ω et est stable par intersection finie. De plus, $\mathcal{C}$ contient $\sigma(X)$ et $\mathcal{G}$ et donc engendre $\mathcal{F}$. Par l'exercice 1, on a alors $\mathbb{E}[g(Y)\mathbb{1}_A] = \mathbb{E}[\mathbb{E}[g(Y)|\mathcal{G}]\mathbb{1}_A]$ pour tout $A \in \mathcal{F}$. De plus, $\mathbb{E}[g(Y)|\mathcal{G}]$ est $\sigma(\mathcal{G}, \sigma(X))$-mesurable. On a vu dans le cours que ceci implique que $\mathbb{E}[g(Y)|\sigma(\mathcal{G}, \sigma(X))] = \mathbb{E}[g(Y)|\mathcal{G}]$.

Exercice 55 (Théorème de la variance totale)**.** 1. Puisque $X \in L^2$, on a $\mathbb{E}[X|\mathcal{F}] = \Pi_{\mathcal{F}}(X) \in L^2$ également. Par l'inégalité de Cauchy–Schwarz, on a alors $\mathbb{E}[X|\mathcal{F}]Y \in L^1$. Puisque $\mathbb{E}[X|\mathcal{F}]$ est $\mathcal{F}$-mesurable, il vient

$$\mathbb{E}[\,\mathbb{E}[X|\mathcal{F}]Y \,|\, \mathcal{F}] = \mathbb{E}[X|\mathcal{F}]\,\mathbb{E}[Y|\mathcal{F}].$$

Par conséquent,

$$\mathbb{E}[\,\mathbb{E}[X|\mathcal{F}]Y \,|\, \mathcal{G}] = \mathbb{E}[\,\mathbb{E}[\,\mathbb{E}[X|\mathcal{F}]Y \,|\, \mathcal{F}] \,|\, \mathcal{G}] = \mathbb{E}[\,\mathbb{E}[X|\mathcal{F}]\,\mathbb{E}[Y|\mathcal{F}] \,|\, \mathcal{G}].$$

2. On a

$$\begin{aligned}
&\mathbb{E}\big[(X - \mathbb{E}[X|\mathcal{F}])(\mathbb{E}[X|\mathcal{F}] - \mathbb{E}[X|\mathcal{G}]) \,\big|\, \mathcal{G}\big] \\
&= \mathbb{E}[X\mathbb{E}[X|\mathcal{F}] - \mathbb{E}[X|\mathcal{F}]^2 - X\mathbb{E}[X|\mathcal{G}] + \mathbb{E}[X|\mathcal{F}]\mathbb{E}[X|\mathcal{G}] \,|\, \mathcal{G}] \\
&= \mathbb{E}[\mathbb{E}[X|\mathcal{F}]^2 - \mathbb{E}[X|\mathcal{F}]^2 - \mathbb{E}[X|\mathcal{G}]^2 + \mathbb{E}[X|\mathcal{G}]^2 \,|\, \mathcal{G}] && \text{(partie 1)} \\
&= 0 \quad \text{p.s.}
\end{aligned}$$

3. On a

$$\begin{aligned}
\operatorname{Var}(X|\mathcal{G}) &= \mathbb{E}[(X - \mathbb{E}[X|\mathcal{G}])^2 \,|\, \mathcal{G}] \\
&= \mathbb{E}[(X - \mathbb{E}[X|\mathcal{F}] + \mathbb{E}[X|\mathcal{F}] - \mathbb{E}[X|\mathcal{G}])^2 \,|\, \mathcal{G}] \\
&= \mathbb{E}[(X - \mathbb{E}[X|\mathcal{F}])^2 - 2(X - \mathbb{E}[X|\mathcal{F}])(\mathbb{E}[X|\mathcal{F}] - \mathbb{E}[X|\mathcal{G}]) + (\mathbb{E}[X|\mathcal{F}] - \mathbb{E}[X|\mathcal{G}])^2 \,|\, \mathcal{G}] \\
&= \mathbb{E}[\mathbb{E}[(X - \mathbb{E}[X|\mathcal{F}])^2 \,|\, \mathcal{F}] \,|\, \mathcal{G}] + \mathbb{E}[(\mathbb{E}[X|\mathcal{F}] - \mathbb{E}[\mathbb{E}[X|\mathcal{F}]|\mathcal{G}])^2 \,|\, \mathcal{G}] && \text{(partie 2)} \\
&= \mathbb{E}[\operatorname{Var}(X|\mathcal{F}) \,|\, \mathcal{G}] + \operatorname{Var}(\mathbb{E}[X|\mathcal{F}] \,|\, \mathcal{G}).
\end{aligned}$$

CQFD.

On remarque que $\mathbb{E}[X|\mathcal{F}]$ est la projection dans $L^2(\Omega, \mathbb{P})$ de X sur le sous-espace des v.a. $\mathcal{F}$-mesurables, si bien que $X - \mathbb{E}[X|\mathcal{F}]$ est orthogonale à ce sous-espace et en particulier, $X - \mathbb{E}[X|\mathcal{F}]$ et $\mathbb{E}[X|\mathcal{F}] - \mathbb{E}[X]$ sont orthogonales. Par conséquent,

$$\begin{aligned}
\mathrm{Var}(X) &= \|X - \mathbb{E}[X]\|_2^2 \\
&= \|X - \mathbb{E}[X|\mathcal{F}] + \mathbb{E}[X|\mathcal{F}] - \mathbb{E}[X]\|_2^2 \\
&= \|X - \mathbb{E}[X|\mathcal{F}]\|_2^2 + \|\mathbb{E}[X|\mathcal{F}] - \mathbb{E}[X]\|_2^2 \\
&= \mathbb{E}[\mathrm{Var}(X|\mathcal{F})] + \mathrm{Var}(\mathbb{E}[X|\mathcal{F}])).
\end{aligned}$$

Ceci est la loi de la variance totale pour $\mathcal{G} = \{\emptyset, \Omega\}$

Exercice 56. Soit $X \geq 0$ v.a. et $\mathcal{F}$ une tribu. On a pour tout $t > 0$,

$$\mathbb{P}(X \geq t \,|\, \mathcal{F}) = \mathbb{E}[\mathbb{1}_{X \geq t} \,|\, \mathcal{F}] \leq \mathbb{E}\Big[\frac{X}{t} \,\Big|\, \mathcal{F}\Big],$$

où la dernière inégalité provient de la positivité de l'intégrale conditionnelle. Par linéarité de l'espérance conditionnelle, ceci donne l'inégalité de Markov conditionnelle :

$$\mathbb{P}(X \geq t \,|\, \mathcal{F}) \leq \frac{\mathbb{E}[X|\mathcal{F}]}{t}.$$

En appliquant cette inégalité, on obtient,

$$\mathbb{P}(|X - \mathbb{E}[X|\mathcal{F}]| \geq t \,|\, \mathcal{F}) = \mathbb{P}((X - \mathbb{E}[X|\mathcal{F}])^2 \geq t^2) \leq \frac{\mathbb{E}[(X - \mathbb{E}[X|\mathcal{F}])^2 \,|\, \mathcal{F}]}{t^2},$$

ce qui donne lieu à l'inégalité de Chebychev conditionnelle :

$$\mathbb{P}(|X - \mathbb{E}[X|\mathcal{F}]| \geq t \,|\, \mathcal{F}) \leq \frac{\mathrm{Var}(X|\mathcal{F})}{t^2}.$$

Exercice 57. X^2 est $\sigma(X)$-mesurable, donc $\mathbb{E}[X^2|X] = X^2$. Pour calculer $\mathbb{E}[X|X^2]$, soit W une v.a. bornée $\sigma(X^2)$-mesurable. Par le "lemme crucial", il existe alors une fonction f mesurable bornée telle que $W = f(X^2)$. On a alors par l'hypothèse de symétrie,

$$\mathbb{E}[Xf(X^2)] = \mathbb{E}[(-X)f((-X)^2)] = -\mathbb{E}[Xf(X^2)],$$

si bien que $\mathbb{E}[Xf(X^2)] = 0$. La variable aléatoire constante égale à 0 vérifie alors

- 0 est $\sigma(X^2)$-mesurable.
- $\mathbb{E}[0W] = 0 = \mathbb{E}[XW]$ pour toute v.a. bornée W, $\sigma(X^2)$-mesurable, i.e. 0 vérifie la propriété caractéristique.

Ceci montre que $\mathbb{E}[X|X^2] = 0$.

Exercice 58. 1. Par définition de l'espérance conditionnelle, on a pour tout i et toute fonction bornée mesurable f,
- $\mathbb{E}[X_i|S_n]$ est $\sigma(S_n)$-mesurable,
- $\mathbb{E}[\mathbb{E}[X_i|S_n]f(S_n)] = \mathbb{E}[X_i f(S_n)]$.

Posons $F(x_1, \ldots, x_n) = x_1 f(x_1 + \cdots + x_n)$. On a alors pour tout i,

$$\mathbb{E}[X_1 f(S_n)] = \mathbb{E}[F(X_1, \ldots, X_n)] = \mathbb{E}[F(X_i, X_2, \ldots, X_{i-1}, X_1, X_{i+1}, \ldots, X_n)] = \mathbb{E}[X_i f(S_n)].$$

Ici, la deuxième égalité provient du fait que les vecteurs $(X_1, \ldots, X_n)$ et $(X_i, X_2, \ldots, X_{i-1}, X_1, X_{i+1}, \ldots, X_n)$ ont même loi (la loi $\mu_X^{\otimes n}$, qui est invariante par permutation des coordonnées). Ceci montre que $\mathbb{E}[X_i|S_n]$ vérifie également la propriété caractéristique pour X_1, donc $\mathbb{E}[X_i|S_n] = \mathbb{E}[X_1|S_n]$.

2. On a d'après la première partie,

$$n\mathbb{E}[X_1|S_n] = \sum_{i=1}^{n} \mathbb{E}[X_i|S_n] = \mathbb{E}[S_n|S_n] = S_n,$$

et donc

$$\mathbb{E}[X_1|S_n] = \frac{S_n}{n}.$$

3. Soit $\mathcal{G} = \sigma(X_{n+1}, X_{n+2}, \ldots)$ On a alors $\sigma(S_n, S_{n+1}, \ldots) = \sigma(S_n) \vee \mathcal{G}$. De plus, les tribus $\sigma(X_1, S_n)$ et $\mathcal{G}$ sont indépendantes. L'exercice 3 montre alors que $\mathbb{E}[X_1|S_n] = \mathbb{E}[X_1|\sigma(S_n) \vee \mathcal{G}] = \mathbb{E}[X_1|S_n, S_{n+1}, \ldots]$.

Exercice 59. 1. La fonction $x \mapsto |x|^p$, est convexe pour $p \geq 1$. Par l'inégalité de Jensen conditionnelle, on a alors

$$\mathbb{E}[|X+Y|^p] = \mathbb{E}[\mathbb{E}[|X+Y|^p \mid X]] \geq \mathbb{E}[|\mathbb{E}[X+Y \mid X]|^p] = \mathbb{E}[|X + \mathbb{E}[Y|X]|^p].$$

En prenant les racines p-ièmes, on obtient l'inégalité souhaitée.

2. Si X et Y sont indépendantes, on a $\mathbb{E}[Y|X] = \mathbb{E}[Y]$. L'inégalité découle alors de la précédente.

Exercice 60. Le fait que 4. implique 3. et que 3. implique 1. et 2. est évident. Montrons que 1. implique 3. Soient X, Y des v.a. intégrables telles que 1. est vérifiée. Alors on a

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]] \leq \mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|X]] \leq \mathbb{E}[X],$$

donc $\mathbb{E}[X] = \mathbb{E}[Y]$. Par conséquent, on a d'après 1.,

$$\mathbb{E}[X] = \mathbb{E}[Y] = \mathbb{E}[Y - \mathbb{E}[X|Y] + \mathbb{E}[X|Y]] = \mathbb{E}[(Y - \mathbb{E}[X|Y])^+] + \mathbb{E}[X],$$

et donc $\mathbb{E}[(Y - \mathbb{E}[X|Y])^+] = 0$, ce qui donne $Y \leq \mathbb{E}[X|Y]$ p.s., et donc d'après 1., $Y = \mathbb{E}[X|Y]$ p.s. De la même façon, on montre que $X = \mathbb{E}[Y|X]$ p.s. Ceci montre que 1. implique bien 3.

Pour montrer que 2. implique 3., on peut faire une preuve analogue ou alors utiliser 1. pour les v.a. $\widetilde{X} = -X$ et $\widetilde{Y} = -Y$ (on note que $\sigma(\widetilde{X}) = \sigma(X)$ car $\widetilde{X} = -X$ et $X = -\widetilde{X}$ et de même $\sigma(\widetilde{Y}) = \sigma(Y)$).

Il suffit alors de montrer que 3. implique 4. On traite d'abord le cas où $X, Y \in L^2$. En supposant que 3. est vérifiée, on a alors

$$\begin{aligned}
\mathbb{E}[(X-Y)^2] &= \mathbb{E}[X^2 + Y^2 - XY - XY] \\
&= \mathbb{E}[X^2 + Y^2 - X\mathbb{E}[Y|X] - Y\mathbb{E}[X|Y]] \\
&= \mathbb{E}[X^2 + Y^2 - X^2 - Y^2] \\
&= 0.
\end{aligned}$$

Par conséquent, $\mathbb{E}[(X-Y)^2] = 0$ et donc $X = Y$ p.s.

Remarque : Une explication géométrique de cette preuve est la suivante : En supposant 3., on a $\mathbb{E}[X-Y|X] = \mathbb{E}[X-Y|Y] = 0$, *et donc* $X - Y \in X^\perp \cap Y^\perp$. *Par conséquent,* $X - Y$ *est orthogonal au sous-espace de* L^2 *engendré par* X *et* Y, *en particulier* $X - Y \in (X-Y)^\perp$. *Par conséquent,* $\mathbb{E}[(X-Y)^2] = \|X-Y\|_2^2 = 0$.

Supposons maintenant que $X, Y \in L^1$, $X \geq 0$, $Y \geq 0$, et que 3. est vérifiée. Alors $\sqrt{X}, \sqrt{Y} \in L^2$. De plus, $\sigma(\sqrt{X}) = \sigma(X)$ et $\sigma(\sqrt{Y}) = \sigma(Y)$, puisque $X = (\sqrt{X})^2$ et $Y = (\sqrt{Y})^2$. Puisque la fonction $x \mapsto \sqrt{x}$ est concave sur $\mathbb{R}_+$ (i.e. $x \mapsto -\sqrt{x}$ est convexe), l'inégalité de Jensen conditionnelle donne

$$\mathbb{E}[\sqrt{X}|\sqrt{Y}] = \mathbb{E}[\sqrt{X}|Y] \leq \sqrt{\mathbb{E}[X|Y]} \leq \sqrt{Y}.$$

Pareil, $\mathbb{E}[\sqrt{Y}|\sqrt{X}] \leq \sqrt{Y}$. Par l'équivalence entre 1. et 4. pour des v.a. L^2, cela donne $\sqrt{X} = \sqrt{Y}$ p.s., et donc $X = Y$ p.s.

Supposons maintenant que $X, Y \in L^1$, pas nécessairement positives, et telles que 3. est vérifiée. Puisque la fonction $x \mapsto x^+ = \max(x, 0)$ est convexe, l'inégalité de Jensen conditionnelle donne $\mathbb{E}[X^+|Y] \geq \mathbb{E}[X|Y]^+ = Y^+$, et pareil, $\mathbb{E}[Y^+|X] \geq X^+$. Puisque $\sigma(X^+) \subset \sigma(X)$ et $\sigma(Y^+) \subset \sigma(Y)$, cela donne $\mathbb{E}[X^+|Y^+] = \mathbb{E}[\mathbb{E}[X^+|Y]|Y^+] \geq \mathbb{E}[Y^+|Y^+] = Y^+$, et pareil, $\mathbb{E}[Y^+|X^+] \geq X^+$. En utilisant l'équivalence entre 2. et 4. pour des v.a. positives, on obtient alors $X^+ = Y^+$ p.s. En faisant le même raisonnement avec $\widetilde{X} = -X$ et $\widetilde{Y} = -Y$, obtient $X^- = Y^-$ p.s. Ceci donne finalement $X = Y$ p.s.

Exercice 61. Calcul direct : On sait que $X_1 + X_2$ suit la loi de Poisson de paramètre $\lambda_1 + \lambda_2$. Si $n \in \mathbb{N}$, $k \in \{0, \ldots, n\}$, on a alors

$$\begin{aligned}
\mathbb{P}(X_1 = k \mid X_1 + X_2 = n) &= \frac{\mathbb{P}(X_1 = k,\, X_1 + X_2 = n)}{\mathbb{P}(X_1 + X_2 = n)} = \frac{\mathbb{P}(X_1 = k,\, X_2 = n-k)}{\mathbb{P}(X_1 + X_2 = n)} = \frac{\mathbb{P}(X_1 = k)\mathbb{P}(X_2 = n-k)}{\mathbb{P}(X_1 + X_2 = n)} \\
&= \frac{e^{-\lambda_1}\frac{\lambda_1^k}{k!}e^{-\lambda_2}\frac{\lambda_2^{n-k}}{(n-k)!}}{e^{-(\lambda_1+\lambda_2)}\frac{(\lambda_1+\lambda_2)^n}{n!}} = \binom{n}{k}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k \left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k}.
\end{aligned}$$

Il en suit que la loi de X_1 conditionnellement à $X_1 + X_2$ est la loi binomiale de paramètres $n = X_1 + X_2$ et $p = \lambda_1/(\lambda_1 + \lambda_2)$.

Approche alternative : Soient $Y^1, Y^2, \dots$ des v.a. iid de loi $\mathbb{P}(Y^1 = i) = p_i = \lambda_i/(\lambda_1 + \lambda_2)$, $i = 1, 2$. Soit P une v.a. de loi $\mathrm{Po}(\lambda_1 + \lambda_2)$, indépendante de la suite $(Y^j)_{j=1,2,\dots}$. On définit le vecteur aléatoire

$$\vec{\xi} = (\xi_1, \xi_2) = \sum_{j=1}^{P} (\mathbf{1}_{(Y^j=1)}, \mathbf{1}_{(Y^j=2)}).$$

Par l'exercice 5 de la feuille TD n^o 2, $(\xi_1, \xi_2) \overset{\text{loi}}{=} (X_1, X_2)$. Il suffit alors de montrer que conditionellement à $\xi_1 + \xi_2$, ξ_1 suit la loi binomiale de paramètres $n = \xi_1 + \xi_2$ et $p = \lambda_1/(\lambda_1 + \lambda_2)$. Or, $\xi_1 + \xi_2 = P$ et par définition, conditionnellement à P, $\vec{\xi}$ suit la loi multinomiale de paramètres P et p_1, p_2. En particulier, conditionnellement à $\xi_1 + \xi_2$, ξ_1 suit la loi binomiale de paramètres $n = X_1 + X_2$ et p_1.

Exercice 62. On rappelle que la loi de S est la loi de densité $p_1 * p_2$ par rapport à la mesure de Lebesgue sur $\mathbb{R}$. Soient $f, h : \mathbb{R} \to \mathbb{R}$ mesurables positives. On va montrer qu'on peut écrire

$$\begin{aligned}
\mathbb{E}[f(X)h(S)] &= \int \left(\int \varphi(s, x) f(x)\, dx \right) (p_1 * p_2)(s) h(s)\, ds \\
&= \mathbb{E}\left[\left(\int \varphi(S, x) f(x)\, dx \right) h(S) \right],
\end{aligned}$$

pour une certaine fonction $\varphi(s, x)$. Il en découlera que la loi de X conditionnellement à S est la loi de densité $\varphi(S, x)$ par rapport à la mesure de Lebesgue sur $\mathbb{R}$.

On écrit

$$\mathbb{E}[f(X)h(S)] = \int \left(\int p_1(x) p_2(y) f(x) h(x + y)\, dy \right) dx.$$

En changeant de variables $y \mapsto s - x$ dans l'intégrale intérieure, il vient

$$\begin{aligned}
\mathbb{E}[f(X)h(S)] &= \int \left(\int p_1(x) p_2(s - x) f(x) h(s)\, ds \right) dx \\
&= \int \left(\int p_1(x) p_2(s - x) f(x)\, dx \right) h(s)\, ds \quad \text{(par Fubini–Tonelli)}
\end{aligned}$$

Notons que si $0 = p_1 * p_2(s) = \int p_1(x) p_2(s - x)\, dx$, alors $p_1(x) p_2(s - x) = 0$ pour Lebesgue-presque tout x, par positivité de p, et donc $\int p_1(x) p_2(s - x) f(x)\, dx = 0$. Par conséquent, en posant

$$\varphi(s, x) = \begin{cases} \frac{p_1(x) p_2(s-x)}{p_1 * p_2(s)} & \text{si } p_1 * p_2(s) > 0, \\ 0 & \text{sinon,} \end{cases}$$

alors on a

$$\mathbb{E}[f(X)h(S)] = \int \left(\int \varphi(s, x) f(x)\, dx \right) (p_1 * p_2)(s) h(s)\, ds.$$

Donc la loi de X conditionnellement à S est la loi de densité $\varphi(S, x)$ par rapport à la mesure de Lebesgue sur $\mathbb{R}$.

Dans le cas où $X \sim \Gamma(\alpha_1, \beta)$, $Y \sim \Gamma(\alpha_2, \beta)$ avec $\alpha_1, \alpha_2, \beta > 0$, on a

$$\varphi(s, x) = c_s x^{\alpha_1 - 1} (s - x)^{\alpha_2 - 1} \mathbb{1}_{x \in [0, s]},$$

pour une constante c_s dépendant de s. Puisque $\int \varphi(s, x)\, dx = 1$, on a $c_s = (\int_0^s x^{\alpha - 1} (s - x)^{\alpha - 1})^{-1}$. Si Z_s est une v.a. de densité $\varphi(s, x)$ par rapport à Lebesgue sur $\mathbb{R}$, alors $s^{-1} Z_s$ est de densité

$$s\varphi(s, sx) = s^{\alpha_1 + \alpha_2 - 1} c_s x^{\alpha_1 - 1} (1 - x)^{\alpha_2 - 1} \mathbb{1}_{x \in [0, 1]}.$$

On reconnaît cette loi comme la loi Bêta $B(\alpha_1, \alpha_2)$ (cf la feuille "Recueil d'exercices").

Exercice 63. 1. Par l'Exercice 1, on a pour toute fonction bornée mesurable f,

$$\mathbb{E}[f(X + Y) \mid X] = \mathbb{E}[f(x + Y)]_{x = X},$$

et donc $X + Y$ suit la loi $\mathcal{N}(X, \sigma_Y^2)$ conditionnellement à X.

2. Première solution : On cherche $a, b \in \mathbb{R}$ tels que $aX + bY \perp X + Y$. Puisque $(X, Y)^T$ est un vecteur gaussien, $(X + Y, aX + bY)^T = \begin{pmatrix} 1 & 1 \\ a & b \end{pmatrix} (X, Y)^T$ en est un également. On a alors

$$X + Y \perp aX + bY \iff 0 = \operatorname{Cov}(X + Y, aX + bY) = a\sigma_X^2 + b\sigma_Y^2$$

Posons alors $a = \sigma_Y^2$, $b = -\sigma_X^2$. On a alors, $X = (b(X + Y) - (aX + bY))/(b - a) = U + V$, avec

$$U = \tfrac{b}{b-a}(X + Y), \quad V = -(b - a)^{-1}(aX + bY), \quad U \perp V$$

et donc

$$\begin{aligned} \sigma_U^2 &= \left(\frac{\sigma_X^2}{\sigma_X^2 + \sigma_Y^2}\right)^2 (\sigma_X^2 + \sigma_Y^2) = \frac{\sigma_X^4}{\sigma_X^2 + \sigma_Y^2} \\ \sigma_V^2 &= \frac{1}{(\sigma_X^2 + \sigma_Y^2)^2}(\sigma_Y^4\sigma_X^2 + \sigma_X^4\sigma_Y^2) = \frac{\sigma_X^2\sigma_Y^2}{\sigma_X^2 + \sigma_Y^2} \end{aligned}$$

Notons que U est $\sigma(X + Y)$-mesurable. Comme dans la première partie de l'exercice, la loi de X conditionnellement à $X + Y$ est alors la loi

$$\mathcal{N}(U, \sigma_V^2) = \mathcal{N}\left(\frac{\sigma_X^2}{\sigma_X^2 + \sigma_Y^2}(X + Y), \frac{\sigma_X^2\sigma_Y^2}{\sigma_X^2 + \sigma_Y^2}\right).$$

Deuxième solution :
On calcule d'abord $\mathbb{E}[X|X + Y]$. On sait d'après le cours qu'il existe $\alpha, \beta \in \mathbb{R}$, telles que $\mathbb{E}[X|X + Y] = \alpha + \beta(X + Y)$. En prenant des espérances, on obtient alors $0 = \alpha$. De plus, on a

$$\mathbb{E}[\mathbb{E}[X|X + Y](X + Y)] = \mathbb{E}[X(X + Y)] = \mathbb{E}[X^2] + \mathbb{E}[XY] = \sigma_X^2,$$

et

$$\mathbb{E}[\mathbb{E}[X|X + Y](X + Y)] = \mathbb{E}[\beta(X + Y)(X + Y)] = \beta\mathbb{E}[(X + Y)^2] = \beta(\sigma_X^2 + \sigma_Y^2),$$

si bien que $\beta = \sigma_X^2/(\sigma_X^2 + \sigma_Y^2)$. Par conséquent,

$$\mathbb{E}[X|X + Y] = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_Y^2}(X + Y).$$

Maintenant on note que $X - \mathbb{E}[X|X + Y]$ est une variable gaussienne indépendante de $X + Y$, car $(X - \mathbb{E}[X|X + Y], X + Y)$ est un vecteur gaussien (car image d'un vecteur gaussien par une application linéaire) et $X - \mathbb{E}[X|X + Y]$ est orthogonale à $X + Y$ dans L^2, donc $\operatorname{Cov}(X - \mathbb{E}[X|X + Y], X + Y) = 0$. Il suffit alors de déterminer la variance de $X - \mathbb{E}[X|X + Y]$. Or,

$$\begin{aligned} \operatorname{Var}(X - \mathbb{E}[X|X + Y]) &= \operatorname{Var}\left(\frac{(\sigma_X^2 + \sigma_Y^2)X - \sigma_X^2(X + Y)}{\sigma_X^2 + \sigma_Y^2}\right) \\ &= \frac{\operatorname{Var}(\sigma_Y^2 X - \sigma_X^2 Y)}{(\sigma_X^2 + \sigma_Y^2)^2} \\ &= \frac{\sigma_Y^4\sigma_X^2 + \sigma_X^4\sigma_Y^2}{(\sigma_X^2 + \sigma_Y^2)^2} \\ &= \frac{\sigma_X^2\sigma_Y^2}{\sigma_X^2 + \sigma_Y^2}. \end{aligned}$$

Donc, X est la somme de $\beta(X + Y)$ et d'une gaussienne indépendante, centrée, de variance $\frac{\sigma_X^2\sigma_Y^2}{\sigma_X^2 + \sigma_Y^2}$. Comme dans la première partie de l'exercice, la loi de X conditionnellement à $X + Y$ est alors la loi gaussienne

$$\mathcal{N}\left(\frac{\sigma_X^2}{\sigma_X^2 + \sigma_Y^2}(X + Y), \frac{\sigma_X^2\sigma_Y^2}{\sigma_X^2 + \sigma_Y^2}\right).$$

Exercice 64 (Processus de Poisson homogène sur $\mathbb{R}_+$). 1. Pour tout $n \in \mathbb{N}$, $t \geq 0$,

$$\begin{aligned}
\frac{1}{n!}\int_t^\infty x^n e^{-x}\,dx &= e^{-t}\frac{1}{n!}\int_0^\infty (x+t)^n e^{-x}\,dx && x \to x+t \\
&= e^{-t}\frac{1}{n!}\int_0^\infty \sum_{k=0}^n \frac{n!}{k!(n-k)!} x^{n-k} t^k e^{-x}\,dx && \text{formule du binôme de Newton} \\
&= e^{-t}\sum_{k=0}^n \frac{t^k}{k!}\frac{1}{(n-k)!}\int_0^\infty x^{n-k} e^{-x}\,dx \\
&= e^{-t}\sum_{k=0}^n \frac{t^k}{k!} && \text{par la définition de la fonction } \Gamma.
\end{aligned}$$

2. On a pour tout $n \in \mathbb{N}$, $\Pi([0,t)) \leq n$ si et seulement si $X_{n+1} \geq t$. Or, $X_{n+1} \sim \Gamma(n+1,\lambda)$, donc $\lambda X_{n+1} \sim \Gamma(n+1,1)$. Cela donne,

$$\begin{aligned}
\mathbb{P}(\Pi([0,t)) \leq n) &= \mathbb{P}(X_{n+1} \geq t) = \mathbb{P}(\lambda X_{n+1} \geq \lambda t) \\
&= \frac{1}{n!}\int_{\lambda t}^\infty x^n e^{-x}\,dx && \text{car } \Gamma(n+1) = n! \\
&= e^{-\lambda t}\sum_{k=0}^n \frac{(\lambda t)^k}{k!} && \text{par la première partie de l'exercice.}
\end{aligned}$$

Ceci montre que $\Pi([0,t))$ suit bien la loi de Poisson de paramètre λt.

3. Soit $n \in \mathbb{N}$. On a égalité des événements

$$\{\Pi([0,t)) = n\} = \{X_n < t,\, X_{n+1} \geq t\} = \{X_n < t,\, Y_{n+1} \geq t - X_n\}.$$

Il en suit pour tout $y \geq 0$,

$$\begin{aligned}
\mathbb{P}(X_1^{(t)} \geq y,\, \Pi([0,t)) = n) &= \mathbb{E}[\mathbf{1}_{X_1^{(t)} \geq y}\mathbf{1}_{\Pi([0,t))=n}] \\
&= \mathbb{E}[\mathbf{1}_{Y_{n+1} \geq t - X_n + y}\mathbf{1}_{X_n < t}\mathbf{1}_{Y_{n+1} \geq t - X_n}] \\
&= \mathbb{E}[\mathbb{E}[\mathbf{1}_{Y_{n+1} \geq t - X_n + y} \mid X_n]\mathbf{1}_{X_n < t}] \\
&= \mathbb{E}[e^{-\lambda y}\mathbb{E}[\mathbf{1}_{Y_{n+1} \geq t - X_n} \mid X_n]\mathbf{1}_{X_n < t}] && \text{car } Y_{n+1} \sim \mathrm{Exp}(\lambda) \text{ indép. de } X_n \\
&= e^{-\lambda y}\mathbb{E}[\mathbf{1}_{X_n < t}\mathbf{1}_{Y_{n+1} \geq t - X_n}] \\
&= e^{-\lambda y}\mathbb{P}(\Pi([0,t)) = n).
\end{aligned}$$

Ceci montre que $X_1^{(t)}$ est indépendante de $\Pi([0,t))$ et suit la loi exponentielle de paramètre λ. Avec la dernière partie de l'exercice, ceci montre que $(X_1^{(t)}, \Pi([0,t))$ suit la loi $\mathrm{Exp}(\lambda) \otimes \mathrm{Po}(\lambda)$.

4. Soit $n \in \mathbb{N}$. Soit $f : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$ bornée mesurable. On note que sur l'événement $\Pi([0,t)) = n$, on a pour tout $k \geq 2$,

$$X_k^{(t)} - X_{k-1}^{(t)} = Y_{n+k}.$$

Par conséquent,

$$\begin{aligned}
&\mathbb{E}[f((X_k^{(t)} - X_{k-1}^{(t)})_{k\geq 2}) \mid X_1^{(t)},\, \Pi([0,t))] \\
&= \sum_{n=0}^\infty \mathbb{E}[f((X_k^{(t)} - X_{k-1}^{(t)})_{k\geq 2}) \mid X_1^{(t)},\, \Pi([0,t))]\mathbf{1}_{\Pi([0,t))=n} \\
&= \sum_{n=0}^\infty \mathbb{E}[f((Y_{n+k})_{k\geq 2}) \mid X_{n+1},\, \Pi([0,t))]\mathbf{1}_{\Pi([0,t))=n} && X_1^{(t)} = X_{n+1} - t \text{ on } \{\Pi([0,t)) = n\} \\
&= \sum_{n=0}^\infty \mathbb{E}[f((Y_{n+k})_{k\geq 2})]\mathbf{1}_{\Pi([0,t))=n} && \sigma(Y_1,\ldots,Y_{n+1}) \text{ et } \sigma(Y_{n+2},\ldots) \text{ indépendantes.} \\
&= \mathbb{E}[f(Y_k)_{k\geq 2})]
\end{aligned}$$

Ceci montre l'indépendance entre $(X_k^{(t)} - X_{k-1}^{(t)})_{k\geq 2}$ et $X_1^{(t)}$, $\Pi([0,t))$ et de plus que $(X_k^{(t)} - X_{k-1}^{(t)})_{k\geq 2}$ est égale en loi à $(Y_k)_{k\geq 2}$.

5. Les deux dernières parties montrent que
$$(\Pi([0,t)), X_1^{(t)}, X_2^{(t)} - X_1^{(t)}, X_3^{(t)} - X_2^{(t)}, \dots) \sim \mathrm{Po}(\lambda) \otimes \mathrm{Exp}(\lambda)^{\otimes \mathbb{N}}.$$
Par conséquent, la suite $(X_i^{(t)})_{i\geq 1}$ est égale en loi à $(X_i)_{i\geq 1}$ et indépendante de $\Pi([0,t))$.
6. Preuve par récurrence sur k. $k=0$ est évident. Supposons que pour un certain $k \in \mathbb{N}$ l'énoncé soit vérifié pour tous les $0 = t_0 < t_1 < \ldots < t_k$. Soient $0 = t_0 < t_1 < \ldots < t_{k+1}$. Par la dernière partie de l'exercice, la suite $(X_i^{(t_1)})_{i\geq 1}$ est égale en loi à $(X_i)_{i\geq 1}$ et indépendante de $\Pi([0,t_1))$ qui suit la loi $\mathrm{Po}(\lambda t_1)$. En définissant $\Pi^{(t_1)}$ à partir de $(X_i^{(t_1)})_{i\geq 1}$ de manière analogue à Π et en utilisant l'hypothèse de récurrence pour $\Pi^{(t_1)}$, il en suit que le vecteur $(\Pi([t_1,t_2)), \dots, \Pi([t_{k-1},t_k))) = (\Pi^{(t_1)}([0,t_2-t_1)), \dots, \Pi^{(t_1)}([t_{k-1} - t_1, t_k - t_1)))$ est indépendant de $\Pi([0,t_1))$ et que ses composantes sont indépendantes et de lois respectives $\mathrm{Po}(\lambda(t_i - t_{i-1}))$, $i = 2, \dots, k$. Ceci conclut la preuve.

Processus et martingales

Exercice 65 (Temps d'arrêt)**.** 1. Soit $n \in \mathbb{N}$. Alors
$$\{S \wedge T > n\} = \{S > n\} \cap \{T > n\} \in \mathcal{F}_n,$$
car S et T sont des temps d'arrêts et $\mathcal{F}_n$ est une tribu. Par conséquent, $S \wedge T$ est un temps d'arrêt. De manière analogue,
$$\{S \vee T \leq n\} = \{S \leq n\} \cap \{T \leq n\} \in \mathcal{F}_n,$$
et donc $S \vee T$ est un temps d'arrêt. Finalement,
$$\{S + T = n\} = \bigcup_{k=0}^{n} \{S = k\} \cap \{T = n-k\} \in \mathcal{F}_n,$$
car $\{S = k\} \in \mathcal{F}_k \subset \mathcal{F}_n$ et $\{T = n-k\} \in \mathcal{F}_{n-k} \subset \mathcal{F}_n$ pour tout $k \leq n$.

2. Par définition, $\mathcal{F}_T$ consiste en les événements $A \subset \mathcal{F}$, tels que $A \cap \{T \leq n\} \in \mathcal{F}_n$ pour tout n. Or, $\{T \leq n\} = \emptyset$ if $n < p$ and $= \Omega$ if $n \geq p$. Donc, $A \in \mathcal{F}_T$ si et seulement si $A \in \mathcal{F}_n$ pour tout $n \geq p$. Cela donne $\mathcal{F}_T = \mathcal{F}_p$, puisque $\mathcal{F}_p \subset \mathcal{F}_n$ pour tout $n \geq p$.
3. Soit $A \in \mathcal{F}_S$. Alors pour tout $n \in \mathbb{N}$, puisque $\{T \leq n\} \subset \{S \leq n\}$,
$$A \cap \{T \leq n\} = (A \cap \{S \leq n\}) \cap \{T \leq n\} \in \mathcal{F}_n,$$
puisque $(A \cap \{S \leq n\}) \in \mathcal{F}_n$ (définition de $\mathcal{F}_S$) et $\{T \leq n\} \in \mathcal{F}_n$ (T est temps d'arrêt). Par conséquent, $A \in \mathcal{F}_T$.
4. Par la dernière partie, $\mathcal{F}_{S\wedge T}$ est contenue dans $\mathcal{F}_S$ et $\mathcal{F}_T$, et donc $\mathcal{F}_{S\wedge T} \subset \mathcal{F}_S \cap \mathcal{F}_T$. Il suffit alors de démontrer l'autre inclusion. Soit $A \in \mathcal{F}_S \cap \mathcal{F}_T$. Alors pour tout $n \in \mathbb{N}$,
$$A \cap \{S \wedge T \leq n\} = (A \cap \{S \leq n\}) \cup (A \cap \{T \leq n\}) \in \mathcal{F}_n,$$
puisque $A \in \mathcal{F}_S$ et $A \in \mathcal{F}_T$. Par conséquence, $A \in \mathcal{F}_{S\wedge T}$ et donc $\mathcal{F}_S \cap \mathcal{F}_T \subset \mathcal{F}_{S\wedge T}$.
5. Soit $n \in \mathbb{N}$. Alors $\{S < T\} \cap \{T = n\} = \{S \leq n-1\} \cap \{T = n\} \in \mathcal{F}_n$ et $\{S < T\} \cap \{S = n\} = \{S = n\} \cap \{T > n\} \in \mathcal{F}_n$, car S et T sont des temps d'arrêt. Par conséquent, $\{S < T\} \in \mathcal{F}_S \cap \mathcal{F}_T$.
Par symétrie, $\{T < S\} \in \mathcal{F}_S \cap \mathcal{F}_T$, et puisque $\mathcal{F}_S \cap \mathcal{F}_T$ est une tribu, $\{S = T\} = \{S < T\}^c \cap \{T < S\}^c \in \mathcal{F}_S \cap \mathcal{F}_T$.
6. Soit $n \in \mathbb{N}$. Alors puisque $T_k \in \mathbb{N}$ pour tout k,
$$\begin{aligned}\{\limsup T_k \leq n\} &= \{\text{tous sauf un nombre fini de } T_k \leq n\}\\ &= \liminf_{k\to\infty} \{T_k \leq n\} \in \mathcal{F}_n,\end{aligned}$$
car $\mathcal{F}_n$ est une tribu et $\{T_k \leq n\} \in \mathcal{F}_n$ pour tout k. Par conséquent, $\limsup T_k$ est un temps d'arrêt. De la même façon, pour tout $n \in \mathbb{N}$,
$$\begin{aligned}\{\liminf T_k > n\} &= \{\text{tous sauf un nombre fini de } T_k > n\}\\ &= \liminf_{k\to\infty} \{T_k > n\} \in \mathcal{F}_n,\end{aligned}$$
si bien que $\liminf T_k$ est un temps d'arrêt.

Exercice 66. 1. Pour tout $n \in \mathbb{N}$,

$$\{T > n\} = \bigcap_{i=1}^{n} \{X_i \leq X_0\} \in \mathcal{F}_n,$$

car $\{X_i \leq X_0\} \in \mathcal{F}_i \subset \mathcal{F}_n$ pour tout $i \in \{1, \dots, n\}$.

2. Pour tout $n \in \mathbb{N}$,

$$\begin{aligned}\mathbb{P}(T > n) = \mathbb{P}\left(\bigcap_{i=1}^{n} \{X_i \leq X_0\}\right) &= \mathbb{E}\left[\mathbb{E}\left[\prod_{i=1}^{n} \mathbf{1}_{\{X_i \leq X_0\}} \,\middle|\, X_0\right]\right] \\ &= \mathbb{E}[X_0^n] = \int_0^1 x^n \, dx = \frac{1}{n+1}.\end{aligned}$$

Par conséquent, pour tout $n \in \mathbb{N}^*$,

$$\mathbb{P}(T = n) = \mathbb{P}(T > n-1) - \mathbb{P}(T > n) = \frac{1}{n} - \frac{1}{n+1} = \frac{1}{n(n+1)}.$$

Exercice 67 (Principe de réflexion)**.** 1. Pour tout $n \in \mathbb{N}$, par la définition de l'inf,

$$\{T_x > n\} = \{\forall k \in \{1, \dots, n\} : S_k < x\} = \{\max_{k \leq n} S_k < x\},$$

et donc par passage au complémentaire, $\{T_x \leq n\} = \{\max_{k \leq n} S_k \geq x\}$. En particulier, $\{T_x \leq n\} \in \mathcal{F}_n$, puisque $\max_{k \leq n} S_k$ est $\mathcal{F}_n$-mesurable.

2. Puisque $\max_{k \leq n} S_k \geq S_n$, on a inclusion des événements $\{S_n \geq x\} \subset \{\max_{k \leq n} S_k \geq x\}$. L'inclusion énoncée suit alors de la dernière partie.

3. Par la dernière partie,

$$\begin{aligned}\mathbb{P}(S_n \geq x) &= \mathbb{P}(S_n \geq x, \, T_x \leq n) \\ &= \sum_{k=0}^{n} \mathbb{P}(S_n \geq x, \, T_x = k) \\ &= \sum_{k=0}^{n} \mathbb{P}(S_n - S_k \geq x - S_k, \, T_x = k)\end{aligned}$$

Puisque $S_k \geq x$ sur l'événement $\{T_x = k\}$, cela donne

$$\mathbb{P}(S_n \geq x) \geq \sum_{k=0}^{n} \mathbb{P}(S_n - S_k \geq 0, \, T_x = k).$$

4. Par la symétrie de la loi de X_1 et le fait que les v.a. $(X_i)_{i \geq 1}$ sont iid, les v.a. $(-X_i)_{i \geq 1}$ sont également iid et de même loi que X_1. Par conséquent,

$$S_n - S_k = X_{k+1} + \cdots + X_n \overset{\text{loi}}{=} (-X_{k+1}) + \cdots + (-X_n) = -(S_n - S_k),$$

et donc la loi de $S_n - S_k$ est également symétrique.

5. On a $S_n - S_k = X_{k+1} + \cdots + X_n$, si bien que $S_n - S_k$ est indépendante de $\mathcal{F}_k$ par l'indépendance des v.a. $(X_i)_{i \geq 1}$. Ceci donne

$$\mathbb{P}(S_n - S_k \geq 0 \,|\, \mathcal{F}_k) = \mathbb{P}(S_n - S_k \geq 0) \geq 1/2,$$

où la dernière égalité provient de la symétrie de la loi de $S_n - S_k$.

6. Pour tout $x \in \mathbb{R}$, $n \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{P}(S_n \geq x) &\geq \sum_{k=0}^{n} \mathbb{P}(S_n - S_k \geq 0,\, T_x = k) && \text{(partie 3)}\\
&= \sum_{k=0}^{n} \mathbb{E}[\mathbf{1}_{S_n - S_k \geq 0} \mathbf{1}_{T_x = k}]\\
&= \sum_{k=0}^{n} \mathbb{E}[\mathbb{E}[\mathbf{1}_{S_n - S_k \geq 0} \,|\, \mathcal{F}_k] \mathbf{1}_{T_x = k}] && (\{T_x = k\} \in \mathcal{F}_k \text{ car } T_x \text{ temps d'arrêt})\\
&\geq \sum_{k=0}^{n} \frac{1}{2} \mathbb{E}[\mathbf{1}_{T_x = k}] && \text{(partie 5)}\\
&= \frac{1}{2} \mathbb{P}(T_x \leq n)\\
&= \frac{1}{2} \mathbb{P}(\max_{k \leq n} S_k \geq x) && \text{(partie 1).}
\end{aligned}$$

Ceci donne l'inégalité souhaitée.

Exercice 68. Puisque $|M_n \vee N_n| \leq |M_n| \vee |N_n| \leq |M_n| + |N_n|$, la v.a. $M_n \vee N_n$ est intégrable pour tout $n \in \mathbb{N}$. De plus,

$$\mathbb{E}[M_{n+1} \vee N_{n+1} \,|\, \mathcal{F}_n] \geq \mathbb{E}[M_{n+1} \,|\, \mathcal{F}_n] \geq M_n,$$

car M est une sous-martingale. Le même argument donne $\mathbb{E}[M_{n+1} \vee N_{n+1} \,|\, \mathcal{F}_n] \geq N_n$. Il suit,

$$\mathbb{E}[M_{n+1} \vee N_{n+1} \,|\, \mathcal{F}_n] \geq M_n \vee N_n,$$

ce qui montre que $M \vee N$ est une sous-martingale.

Exercice 69 (Décomposition de Doob d'une sous-martingale)**.** On sait d'après le cours que pour tout processus adapté et intégrable $(M_n)_{n\geq 0}$ il existe un (p.s.) unique processus prévisible $(A_n)_{n\in\mathbb{N}}$ tel que $(M_n - A_n)_{n\geq 0}$ est une martingale et $A_0 = 0$. Ce processus est défini par $A_{n+1} - A_n = \mathbb{E}[M_{n+1} - M_n \,|\, \mathcal{F}_n]$ pour tout $n \geq 0$. Il suffit alors de vérifier que $(A_n)_{n\geq 0}$ est croissant dans notre cas, ce qui découle du fait que M est une sous-martingale :

$$A_{n+1} - A_n = \mathbb{E}[M_{n+1} - M_n \,|\, \mathcal{F}_n] = \mathbb{E}[M_{n+1} \,|\, \mathcal{F}_n] - M_n \geq M_n - M_n = 0.$$

Exercice 70 (Inégalité maximale pour les surmartingales positives)**.**

1. On a pour tout $n \in \mathbb{N}$, puisque $(H_n)_{n\geq 0}$ est prévisible et borné,

$$\begin{aligned}
\mathbb{E}[(H \cdot X)_{n+1} | \mathcal{F}_n] &= \mathbb{E}[(H \cdot X)_n + H_{n+1}(X_{n+1} - X_n) | \mathcal{F}_n]\\
&= (H \cdot X)_n + H_{n+1} \mathbb{E}[(X_{n+1} - X_n) | \mathcal{F}_n].
\end{aligned}$$

Puisque X est une surmartingale, on a $\mathbb{E}[(X_{n+1} - X_n)|\mathcal{F}_n] \leq 0$. Puisque H_n est positif, cela donne $\mathbb{E}[(H \cdot X)_{n+1}|\mathcal{F}_n] \leq (H \cdot X)_n$ pour tout $n \in \mathbb{N}$, et donc $H \cdot X$ est une surmartingale.

2. On pose $H_n = \mathbb{1}_{n \geq T}$. On a alors,

$$X_{T\wedge n} = X_0 + \sum_{k=1}^{n\wedge T} (X_k - X_{k-1}) = X_0 + \sum_{k=1}^{n} (X_k - X_{k-1}) \mathbb{1}_{k \leq T} = X_0 + (H \cdot X)_n.$$

De plus, on a pour tout $n \in \mathbb{N}^*$, $H_n = 1 - \mathbb{1}_{T \leq n-1}$, ce qui est $\mathcal{F}_{n-1}$-mesurable. Donc, H est un processus prévisible et positif. Par la dernière partie, $H \cdot X$ est alors une sur-martingale. Puisque X_0 est $\mathcal{F}_0$-mesurable, cela implique que $X_{T\wedge n} = X_0 + (H \cdot X)_n$ en est également une.

3. On remarque d'abord qu'il suffit de montrer que pour tout $a > 0$,

$$\mathbb{P}\left(\sup_{n\geq 0} X_n > a\right) \leq \frac{\mathbb{E}[X_0]}{a}.$$

Car, cela donne pour tout $0 < a' < a$,

$$\mathbb{P}\left(\sup_{n\geq 0} X_n \geq a\right) \leq \mathbb{P}\left(\sup_{n\geq 0} X_n > a'\right) \leq \frac{\mathbb{E}[X_0]}{a'},$$

et donc, par continuité du terme de droite dans la dernière inégalité,

$$\mathbb{P}\left(\sup_{n\geq 0} X_n \geq a\right) \leq \frac{\mathbb{E}[X_0]}{a}.$$

Soit $T_a = \inf\{n \in \mathbb{N} : X_n > a\}$, si bien que $\{\sup_{n\geq 0} X_n > a\} = \{T_a < \infty\}$. On a alors

$$a\mathbb{P}\left(\sup_{n\geq 0} X_n > a\right) = \mathbb{E}[a\mathbb{1}_{T_a<\infty}] \leq \mathbb{E}[X_{T_a}\mathbb{1}_{T_a<\infty}].$$

Puisque $(X_n)_{n\geq 0}$ est positive, on a par le lemme de Fatou,

$$\mathbb{E}[X_{T_a}\mathbb{1}_{T_a<\infty}] = \mathbb{E}[\lim_{n\to\infty} X_{n\wedge T_a}\mathbb{1}_{T_a<\infty}] \leq \lim_{n\to\infty} \mathbb{E}[X_{n\wedge T_a}\mathbb{1}_{T_a<\infty}].$$

D'après 2., $(X_{n\wedge T_a})_{n\geq 0}$ est encore une surmartingale. Avec la positivité de $(X_n)_{n\geq 0}$, cela donne pour tout n,

$$\mathbb{E}[X_{n\wedge T_a}\mathbb{1}_{T_a<\infty}] \leq \mathbb{E}[X_{n\wedge T_a}] \leq \mathbb{E}[X_0].$$

Ensemble, les inégalités ci-dessus donnent

$$\mathbb{P}\left(\sup_{n\geq 0} X_n > a\right) \leq \frac{\mathbb{E}[X_0]}{a}.$$

CQFD.

Exercice 71 (Martingales de carrés intégrables)**.** 1. Soit $0 \leq m < n$. On a alors

$$\mathbb{E}[(X_n - X_{n-1})(X_m - X_{m-1})] = \mathbb{E}[\mathbb{E}[X_n - X_{n-1}|\mathcal{F}_{n-1}](X_m - X_{m-1})],$$

car $X_m - X_{m-1}$ est $\mathcal{F}_{n-1}$-mesurable et $X_kX_l \in L^1$ pour tout k, l. Puisque $(X_n)_{n\geq 0}$ est une martingale, on a $\mathbb{E}[X_n - X_{n-1}|\mathcal{F}_{n-1}] = 0$ ce qui donne $\mathbb{E}[(X_n - X_{n-1})(X_m - X_{m-1})] = 0$.

2. Soit $-1 \leq m \leq n$. Par la première partie de l'exercice, on a $X_i - X_{i-1} \perp_{L^2} X_j - X_{j-1}$ pour tout $0 \leq j < i$. Par conséquent,

$$\begin{aligned}\mathbb{E}[(X_n - X_m)^2] &= \|X_n - X_m\|_2^2 \\ &= \left\|\sum_{i=m+1}^{n} (X_i - X_{i-1})\right\|_2^2 \\ &= \sum_{i=m+1}^{n} \|X_i - X_{i-1}\|_2^2 \\ &= \sum_{i=m+1}^{n} \mathbb{E}[(X_i - X_{i-1})^2].\end{aligned}$$

3. Par la dernière partie de l'exercice, on a $\mathbb{E}[X_n^2] = \sum_{i=0}^{n} \mathbb{E}[(X_i - X_{i-1})^2]$. Ceci montre que

$$\sup_{n\in\mathbb{N}} \mathbb{E}[X_n^2] < \infty \iff \sum_{n=0}^{\infty} \mathbb{E}[(X_n - X_{n-1})^2] < \infty.$$

Ceci donne l'équivalence entre a) et b).

Exercice 72 (Processus de Galton–Watson)**.** On rappelle que par définition, on a $X_{n+1} = \sum_{i=1}^{X_n} \xi_{n+1,i}$, où $(\xi_{n,i})_{n,i\in\mathbb{N}^*}$ est une famille de v.a. iid selon μ. En particulier, pour tout $n \in \mathbb{N}$, la famille $(\xi_{n+1,i})_{i\in\mathbb{N}}$ est indépendante de $\mathcal{F}_n$. On a alors

$$\mathbb{E}[W_{n+1}|\mathcal{F}_n] = m^{-n-1}\mathbb{E}[\sum_{i=1}^{X_n} \xi_{n+1,i}|\mathcal{F}_n] = m^{-n-1}\sum_{i=1}^{X_n} \mathbb{E}[\xi_{n+1,i}] = m^{-n}X_n = W_n.$$

Ici, la deuxième égalité se justifie par l'Exercice 2 de la feuille n° 5 avec $X = (X_1, \dots, X_n)$ et $Y = (\xi_{n+1,i})_{i\in\mathbb{N}}$ (dans cet exercice, la fonction f est supposée bornée, mais on peut tronquer la somme à un entier K, faire tendre $K \to \infty$ et utiliser la positivité des v.a. et convergence monotone pour conclure). L'intégrabilité de W_n en découle en remarquant que $\mathbb{E}[W_{n+1}] = \mathbb{E}[W_n]$ pour tout n (par positivité, l'espérance est bien définie), et donc $\mathbb{E}[W_n] = \mathbb{E}[W_0] = 1 < \infty$. Ceci montre que $(W_n)_{n\geq 0}$ est une martingale.

Pour montrer qu'elle est bornée dans L^2, on note que pour tout $n \in \mathbb{N}$, par indépendance,

$$\mathrm{Var}(W_{n+1}|\mathcal{F}_n) = m^{-2n-2}\,\mathrm{Var}(\sum_{i=1}^{X_n} \xi_{n+1,i}|\mathcal{F}_n) = m^{-2n-2}\sum_{i=1}^{X_n} \mathrm{Var}(\xi_{n+1,i}) = m^{-n-2}\sigma^2 W_n,$$

où la deuxième égalité se justifie encore par l'Exercice 2 de la feuille n° 5, en utilisant la définition de la variance conditionnelle en fonction de l'espérance conditionnelle. En particulier, puisque $W_n = \mathbb{E}[W_{n+1}|\mathcal{F}_n]$,

$$\mathbb{E}[(W_{n+1} - W_n)^2] = \mathbb{E}[\mathrm{Var}(W_{n+1}|\mathcal{F}_n)] = m^{-n-2}\sigma^2.$$

Puisque $m > 1$, on a alors

$$\sum_{n\geq 0} \mathbb{E}[(W_{n+1} - W_n)^2] < \infty,$$

si bien que la martingale $(W_n)_{n\geq 0}$ est bornée dans L^2 par le dernier exercice.

Exercice 73 (Variation quadratique d'une martingale)**.**

1. Puisque M est une martingale et la fonction $x \mapsto x^2$ est convexe sur $\mathbb{R}$, un théorème du cours nous dit que $(M_n^2)_{n\in\mathbb{N}}$ est une sous-martingale.
2. Par la définition du processus $\langle M\rangle$, on a pour tout $n \in \mathbb{N}$,

$$\langle M\rangle_{n+1} - \langle M\rangle_n = \mathbb{E}[M_{n+1}^2 - M_n^2 \,|\, \mathcal{F}_n] = \mathbb{E}[M_{n+1}^2 - \mathbb{E}[M_{n+1} \,|\, \mathcal{F}_n]^2 \,|\, \mathcal{F}_n] = \mathrm{Var}(M_{n+1} \,|\, \mathcal{F}_n).$$

3. Puisque $(M_n^2 - \langle M\rangle_n)_{n\in\mathbb{N}}$ est une martingale, on a pour tout $n \in \mathbb{N}$,

$$\mathbb{E}[M_n^2] - \mathbb{E}[\langle M\rangle_n] = \mathbb{E}[M_n^2 - \langle M\rangle_n] = \mathbb{E}[M_0^2] - \mathbb{E}[\langle M\rangle_0] = \mathbb{E}[M_0^2],$$

puisque $\langle M\rangle_0 = 0$ par définition. Ceci donne l'égalité souhaitée.

4. Pour tout $n \in \mathbb{N}$, on a

$$\begin{aligned}
\langle M^T\rangle_{n+1} - \langle M^T\rangle_n &= \mathbb{E}[(M^T)_{n+1}^2 - (M^T)_n^2 \,|\, \mathcal{F}_n] && \text{par définition de } \langle M^T\rangle \\
&= \mathbb{E}[M_{T\wedge(n+1)}^2 - M_{T\wedge n}^2 \,|\, \mathcal{F}_n] \\
&= \mathbb{E}[(M_{T\wedge(n+1)}^2 - M_{T\wedge n}^2)\mathbf{1}_{T>n} \,|\, \mathcal{F}_n] && \text{car } M_{T\wedge(n+1)} = M_{T\wedge n} \text{ sur } \{T \leq n\} \\
&= \mathbb{E}[M_{T\wedge(n+1)}^2 - M_{T\wedge n}^2 \,|\, \mathcal{F}_n]\mathbf{1}_{T>n} && \text{car } \{T > n\} \in \mathcal{F}_n \text{ puisque } T \text{ temps d'arrêt} \\
&= \mathbb{E}[M_{n+1}^2 - M_n^2 \,|\, \mathcal{F}_n]\mathbf{1}_{T>n} \\
&= (\langle M\rangle_{n+1} - \langle M\rangle_n)\mathbf{1}_{T>n} && \text{par définition de } \langle M\rangle.
\end{aligned}$$

De plus, on a pour tout $n \in \mathbb{N}$,

$$\begin{aligned}
\langle M\rangle_{n+1}^T - \langle M\rangle_n^T &= \langle M\rangle_{T\wedge(n+1)} - \langle M\rangle_{T\wedge n} \\
&= (\langle M\rangle_{T\wedge(n+1)} - \langle M\rangle_{T\wedge n})\mathbf{1}_{T>n} && \text{car } \langle M\rangle_{T\wedge(n+1)} = \langle M\rangle_{T\wedge n} \text{ sur } \{T \leq n\} \\
&= (\langle M\rangle_{n+1} - \langle M\rangle_n^T)\mathbf{1}_{T>n}.
\end{aligned}$$

Ceci donne

$$\langle M\rangle_{n+1}^T - \langle M\rangle_n^T = \langle M^T\rangle_{n+1} - \langle M^T\rangle_n, \quad n \in \mathbb{N}.$$

Puisque $\langle M\rangle_0^T = \langle M\rangle_0 = 0 = \langle M^T\rangle_0$, cela montre l'égalité souhaitée.

5. Pour tout $n \in \mathbb{N}$,
$$\langle M\rangle_{n+1} - \langle M\rangle_n = \mathrm{Var}(M_{n+1} \mid \mathcal{F}_n) = \mathbb{E}[(M_{n+1} - M_n)^2 \mid \mathcal{F}_n] = \mathbb{E}[X_{n+1}^2 \mid \mathcal{F}_n].$$
Puisque $\mathcal{F}_n = \sigma(M_1, \dots, M_n) = \sigma(X_1, \dots, X_n)$, la v.a. X_{n+1} est indépendante de $\mathcal{F}_n$, si bien que
$$\mathbb{E}[X_{n+1}^2 \mid \mathcal{F}_n] = \mathbb{E}[X_{n+1}^2] = \mathbb{E}[X_1^2] =: \sigma^2.$$
On obtient alors
$$\langle M\rangle_n = \sigma^2 n, \quad n \in \mathbb{N}.$$

Exercice 74 (Martingales exponentielles de la marche aléatoire)**.** On a pour $\theta \in \mathbb{R}$ et $n \in \mathbb{N}$,
$$\mathbb{E}[W_{n+1}(\theta)|\mathcal{F}_n] = \mathbb{E}\left[\frac{e^{\theta S_n}}{\cosh(\theta)^n}\frac{e^{\theta X_{n+1}}}{\cosh(\theta)}\Big|\mathcal{F}_n\right] = \frac{e^{\theta S_n}}{\cosh(\theta)^n}\mathbb{E}\left[\frac{e^{\theta X_{n+1}}}{\cosh(\theta)}\right] = W_n(\theta)\frac{e^\theta + e^{-\theta}}{2\cosh(\theta)} = W_n(\theta).$$

Exercice 75 (Processus de Poisson)**.** 1. Il suffit de montrer que pour tous $0 \le s_1 < s_2 < \dots < s_n = s < t$, $N_t - N_s$ est indépendante de $\sigma(N_{s_1}, \dots, N_{s_n})$ et de loi $\mathrm{Po}(\lambda(t-s))$. Or, presque sûrement, $N_{s_{k+1}} - N_{s_k} = \Pi([s_k, s_{k+1}))$ pour tout $k = 1, \dots, n-1$, ou alors,
$$N_{s_k} = \sum_{i=1}^{k-1} \Pi([s_i, s_{i+1})), \quad \text{p.s.}, \quad k = 1, \dots, n.$$
L'énoncé découle alors de l'Exercice 14 de la feuille n° 5 du TD.

2. Soit $0 \le s \le t$. On a $\mathbb{E}[N_t - N_s \mid \mathcal{F}_s] = \lambda(t-s)$. Par conséquent,
$$\mathbb{E}[N_t - \lambda t \mid \mathcal{F}_s] = N_s - \lambda s + \mathbb{E}[N_t - N_s \mid \mathcal{F}_s] - \lambda(t-s) = N_s - \lambda s,$$
et donc avec $f(t) = \lambda t$, $(N_t - f(t))_{t\ge 0}$ est une martingale.
De plus,
$$\begin{aligned}\mathbb{E}[(N_t - \lambda t)^2 \mid \mathcal{F}_s] &= \mathbb{E}[(N_t - N_s - \lambda(t-s) + N_s - \lambda s)^2 \mid \mathcal{F}_s]\\ &= \mathbb{E}[(N_t - N_s - \lambda(t-s))^2 \mid \mathcal{F}_s] + 2(N_s - \lambda s)\mathbb{E}[N_t - N_s - \lambda(t-s) \mid \mathcal{F}_s] + (N_s - \lambda s)^2\\ &= \mathrm{Var}(N_t - N_s) + 0 + (N_s - \lambda s)^2\\ &= \lambda(t-s) + (N_s - \lambda s)^2.\end{aligned}$$
Par conséquent, avec $g(t) = f(t) = \lambda t$, $((N_t - f(t))^2 - g(t))_{t\ge 0}$ est une martingale.
Finalement, on a pour tout $0 \le s \le t$ et tout $\theta \in \mathbb{R}$, $\mathbb{E}[e^{\theta(N_t - N_s)}] = e^{\lambda(t-s)(e^\theta - 1)} < \infty$. Par conséquent,
$$\mathbb{E}[e^{\theta N_t} \mid \mathcal{F}_s] = \mathbb{E}[e^{\theta N_s} e^{\theta(N_t - N_s)} \mid \mathcal{F}_s] = e^{\theta N_s}\mathbb{E}[e^{\theta(N_t - N_s)} \mid \mathcal{F}_s] = e^{\theta N_s} e^{\lambda(t-s)(e^\theta - 1)},$$
et donc avec $h_\theta(t) = \lambda t(e^\theta - 1)$, le processus $(e^{\theta N_t - h_\theta(t)})_{t\ge 0}$ est une martingale.

3. Si $N_t - \lambda t$ converge p.s. quand $t \to \infty$, alors $(N_n - \lambda n)/\sqrt{n}$ converge vers 0, p.s. et donc aussi en loi quand $n \to \infty$. Mais par le TCL, $(N_n - \lambda n)/\sqrt{n}$ converge en loi quand $n \to \infty$ vers la loi non-dégénérée $\mathcal{N}(0, \lambda)$. Par conséquent, $N_t - \lambda t$ ne converge pas p.s.
De manière similaire, si $(N_t - \lambda t)^2 - \lambda t$ converge p.s. quand $t \to \infty$, alors $((N_n - \lambda n)^2 - \lambda n)/t$ converge vers 0, p.s. et donc aussi en loi quand $n \to \infty$. Mais $((N_n - \lambda n)^2 - \lambda n)/n = ((N_n - \lambda n)/\sqrt{n})^2 - \lambda$ converge en loi vers la v.a. non-dégénérée $X^2 - \lambda$, où $X \sim \mathcal{N}(0, \lambda)$. Par conséquent, $(N_t - \lambda t)^2 - \lambda t$ ne converge pas p.s. quand $t \to \infty$.
En ce qui concerne la martingale $M_t^\theta = e^{\theta N_t - \lambda t(e^\theta - 1)}$, on note d'abord que si $\theta = 0$, alors $M_t^\theta = 1$ pour tout $t \ge 0$ et donc $M_t^\theta \to 1$ quand $t \to \infty$. Supposons alors que $\theta \ne 0$. Par la loi des grands nombres $N_n/(\lambda n) \to 1$ p.s. quand $n \to \infty$, et puisque $N_{\lfloor t\rfloor} \le N_t \le N_{\lceil t\rceil}$ pour tout $t \ge 0$, on a également
$$N_t/(\lambda t) \to 1 \text{ p.s. quand } t \to \infty.$$
Puisque $e^\theta - 1 > \theta$ pour $\theta \ne 0$, ceci implique que
$$\theta N_t - \lambda t(e^\theta - 1) \to -\infty \text{ p.s. quand } t \to \infty.$$
Par conséquent, $M_t^\theta \to 0$ p.s. quand $t \to \infty$.

Convergence de martingales et théorème d'arrêt

Exercice 76. Il suffit de montrer que

$$\mathbb{P}(C^c \cap \{\liminf X_n > -\infty\}) = 0,$$

car en appliquant ce résultat à la martingale $(-X_n)_{n\geq 0}$ on obtient $\mathbb{P}(C^c \cap \{\limsup X_n < +\infty\}) = 0$ et donc

$$0 = \mathbb{P}(C^c \cap \{\limsup X_n < +\infty \text{ ou } \liminf X_n > -\infty\}) = \mathbb{P}((C \cup D)^c).$$

Quitte à considerer la martingale $X_n - X_0$, on peut supposer que $X_0 = 0$. Soit $K \in \mathbb{N}$ et posons $T_K = \inf\{n : X_n \leq -K\}$. Alors $(X_{n\wedge T_K})_{n\geq 0}$ est une martingale avec $X_{n\wedge T_K} \geq -K - M$ presque sûrement. On sait d'après le cours qu'une martingale positive converge presque sûrement vers une limite finie ; en appliquant ce théorème à la martingale $(X_{n\wedge T_K} + K + M)_{n\geq 0}$ on obtient alors que $X_{n\wedge T_K}$ converge p.s. quand $n \to \infty$. Ceci implique que X_n converge p.s. sur l'événement $\{T_K = \infty\}$, ou

$$\mathbb{P}(C^c \cap \{T_K = \infty\}) = 0.$$

Par l'égalité des événements $\{\liminf X_n > -\infty\} = \lim_{K\to\infty} \uparrow \{T_K = \infty\}$, cela donne

$$\mathbb{P}(C^c \cap \{\liminf X_n > -\infty\}) = 0.$$

Exercice 77 (Lemme de Borel–Cantelli)**.** Comme $\omega \in \limsup A_n$ ssi $\sum_n \mathbf{1}_{A_n}(\omega) = \infty$ il suffit de montrer que $\sum_n \mathbf{1}_{A_n} = \infty$ ssi $\sum_n \mathbb{P}(A_n \,|\, \mathcal{F}_{n-1}) = \infty$ sur un ensemble de probabilité 1. Soit $X_n = \sum_{m=1}^n \mathbb{1}_{A_m} - \mathbb{P}(A_m|\mathcal{F}_{m-1})$ pour tout $n \geq 1$. Alors $(X_n)_{n\geq 1}$ est une martingale avec $|X_n - X_{n-1}| \leq 1$. Noter que X_n est la différence des sommes partielles pour chacune des séries ci-dessus. Soient C et D les événements de l'exercice précédant. Alors,

- sur C, $\lim_n X_n = X_\infty$ existe et donc $\sum_{n=1}^\infty \mathbb{1}_{A_n} = X_\infty + \sum_{n=1}^\infty \mathbb{P}(A_n|\mathcal{F}_{n-1})$. En particulier, $\sum_n \mathbf{1}_{A_n} = \infty$ ssi $\sum_n \mathbb{P}(A_n \,|\, \mathcal{F}_{n-1}) = \infty$.
- sur D, en fait les deux séries divergent (et donc l'équivalence est vraie), car $\limsup_n X_n = \infty$ implique $\sum_n \mathbf{1}_{A_n} = \infty$ et $\liminf_n X_n = -\infty$ implique $\sum_n \mathbb{P}(A_n \,|\, \mathcal{F}_{n-1}) = \infty$ par positivité de chacune des deux séries.

En résumant, sur $C \cup D$, $\sum_n \mathbf{1}_{A_n} = \infty$ ssi $\sum_n \mathbb{P}(A_n \,|\, \mathcal{F}_{n-1}) = \infty$. D'après le dernier exercice, $C \cup D$ est un ensemble de probabilité 1. CQFD.

Exercice 78 (Quelques (contre-)exemples)**.** 1. Par indépendance de ε_{n+1}, Y_{n+1} et $\mathcal{F}_n$, on a pour tout $n \in \mathbb{N}$,

$$\mathbb{E}[M_{n+1} - M_n \,|\, \mathcal{F}_n] = \mathbb{E}[\varepsilon_{n+1}Y_{n+1} \,|\, \mathcal{F}_n] = \mathbb{E}[\varepsilon_{n+1}]\mathbb{E}[Y_{n+1}] = 0,$$

car $\mathbb{E}[Y_n] = \frac{1}{2}(a_n - a_n) = 0$ pour tout $n \in \mathbb{N}^*$. Ceci montre que $(M_n)_{n\geq 0}$ est une martingale par rapport à la filtration $(\mathcal{F}_n)_{n\geq 1}$. Montrons qu'elle converge presque sûrement. On a

$$\sum_{n=1}^\infty \mathbb{P}(\varepsilon_n = 1) = \sum_{n=1}^\infty \frac{1}{n^2} < \infty,$$

et donc par Borel–Cantelli,

$$\mathbb{P}(\limsup_n \{\varepsilon_n = 1\}) = 0.$$

Mais puisque ε_n prend valeurs dans $\{0, 1\}$ pour tout $n \in \mathbb{N}$, cela implique

$$\mathbb{P}(\exists N \,\forall n \geq N : \varepsilon_n = 0) = \mathbb{P}(\liminf_n \{\varepsilon_n = 0\}) = 1 - \mathbb{P}(\limsup_n \{\varepsilon_n = 1\}) = 1.$$

Par conséquent,

$$\mathbb{P}(\exists N \,\forall n \geq N : M_n = M_N) = 1,$$

et donc M_n converge p.s. quand $n \to \infty$.

On souhaite trouver une condition suffisante sur $(a_n)_{n\geq 0}$ tel que $(M_n)_{n\geq 0}$ n'est pas bornée. On raisonne par contradiction. Supposons que $(M_n)_{n\geq 0}$ est bornée dans L^1. Alors $\varepsilon_n Y_n = M_n - M_{n-1}$ l'est également car $\|\varepsilon_n Y_n\|_1 \leq \|M_n\|_1 + \|M_{n-1}\|_1$ par l'inégalité triangulaire. Or,

$$\mathbb{E}[|\varepsilon_n Y_n|] = a_n \mathbb{P}(\varepsilon_n = 1) = \frac{a_n}{n^2}.$$

Cela montre que $\|M_n\|_1 \to \infty$, dès lors que $a_n/n^2 \to \infty$ quand $n \to \infty$.

2. On a $\mathcal{F}_n = \sigma(M_2, \ldots, M_n) = \sigma(\xi_2, \ldots, \xi_n)$ pour tout $n \geq 2$, car $(M_2, \ldots, M_n)$ est une fonction de $(\xi_2, \ldots, \xi_n)$ et vice versa. Il vient,

$$\mathbb{E}[M_n - M_{n-1} \,|\, \mathcal{F}_{n-1}] = \mathbb{E}[\xi_n] = \frac{-n^2}{n^2} + \frac{n^2}{n^2-1}\frac{n^2-1}{n^2} = -1 + 1 = 0.$$

On a

$$\sum_{n=1}^{\infty} \mathbb{P}(\xi_n = -n^2) = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty,$$

et donc par Borel–Cantelli,

$$\mathbb{P}(\exists N\, \forall n \geq N : \xi_n = \frac{n^2}{n^2-1}) = 1 - \mathbb{P}(\limsup_n \{\xi_n = -n^2\}) = 1.$$

Puisque $n^2/(n^2-1) \geq 1/2$ pour $n \geq 2$, cela montre que $M_n \to \infty$ p.s.

Exercice 79 (Sommes aléatoires)**.** On définit $S_n = \sum_{k=1}^n a_k X_k$ pour tout $n \in \mathbb{N}$ et on note $(\mathcal{F}_n)_{n\geq 0}$ sa filtration canonique. On a alors,

$$\mathbb{E}[S_{n+1}|\mathcal{F}_n] = \mathbb{E}[S_n|\mathcal{F}_n] + \mathbb{E}[a_{n+1}X_{n+1}|\mathcal{F}_n] = S_n + a_{n+1}\mathbb{E}[X_{n+1}] = S_n.$$

Donc, $(S_n)_{n\geq 0}$ est une martingale. On a pour tout $n \in \mathbb{N}$, par indépendance,

$$\mathbb{E}[S_n^2] = \mathrm{Var}(S_n) = \sum_{k=1}^n \mathrm{Var}(a_k X_k) = \mathrm{Var}(X_1) \sum_{k=1}^n a_k^2.$$

En particulier, puisque $\sum_{n=1}^\infty a_n^2 < \infty$, la martingale $(S_n)_{n\geq 0}$ est bornée dans L^2. Par conséquent, elle converge p.s.

Exercice 80 (Théorème de Rademacher)**.**

1. Soit $x = \sum_{k=1}^\infty b_k 2^{-k}$ comme dans l'énoncé et $n \in \mathbb{N}$. On note $r = \sum_{k=n+1}^\infty b_k 2^{-k} < 2^{-n}$. On a alors

$$\varphi_n(x) = \lfloor 2^n x \rfloor 2^{-n} = \left\lfloor \sum_{k=1}^n b_k 2^{n-k} + r 2^n \right\rfloor 2^{-n} = \left(\sum_{k=1}^n b_k 2^{n-k} \right) 2^{-n} = \sum_{k=1}^n b_k 2^{-k}.$$

Par conséquent, si $0 \leq k \leq n$, alors

$$\varphi_k(\varphi_n(x)) = \varphi_k\left(\sum_{i=1}^n b_i 2^{-i} \right) = \sum_{i=1}^k b_i 2^{-i} = \varphi_k(x).$$

2. On a d'après 1., $X_k = \varphi_k(X_n)$ pour tout $k \in n$, où les φ_k sont continues par morceaux et donc mesurables. Par conséquent, $\sigma(X_k) \subset \sigma(X_n)$ et donc $\sigma(X_0, X_1, \ldots, X_n) = \sigma(X_n)$. En ce qui concerne la second égalité, puisque $X_n = \varphi_n(X)$, on a trivialement $\sigma(X_n, X_{n+1}, \ldots) \subset \sigma(X)$ pour tout $n \in \mathbb{N}$. De plus, par 1., on a $X = \lim_{n\in\infty} \varphi_n(X) = \lim_{n\in\infty} X_n$, et donc

$$\sigma(X) \subset \sigma(X_n, X_{n+1}, \ldots).$$

pour tout $n \in \mathbb{N}$. Ceci implique la seconde égalité.

3. Soient $B_1, B_2, \ldots$ des v.a. iid de loi Ber(p). On sait d'après le cours qu'en posant $X = \sum_{k=1}^\infty B_k 2^{-k}$, X est de loi uniforme sur $(0, 1)$. De plus, on a $X_n = \sum_{k=1}^n B_k 2^{-k}$ pour tout $n \in \mathbb{N}$. On a alors,

$$\mathbb{E}[h(X_{n+1})|X_n] = \mathbb{E}[h(X_n + B_{n+1}2^{-(n+1)})|X_n] = \frac{1}{2}(h(X_n + 2^{-(n+1)}) + h(X_n)).$$

Par conséquent, avec $h(x) = 2^{n+1}(f(x + 2^{-(n+1)}) - f(x))$, on a

$$\mathbb{E}[Z_{n+1}|\mathcal{F}_n] = \mathbb{E}[h(X_{n+1})|\mathcal{F}_n] = \frac{1}{2}2^{n+1}(f(X_n + 2^{-n}) - f(X_n)) = Z_n.$$

Notons que $(\mathcal{F}_n)_{n\geq 0}$ est une filtration à cause de la partie 2. On en déduit que $(Z_n)_{n\geq 0}$ est une $(\mathcal{F}_n)_{n\geq 0}$-martingale. De plus, puisque f est Lipschitz sur $[0, 1]$ et $X_n \leq 1 - 2^{-n}$ p.s., on a

$$|Z_n| = 2^n |f(X_n + 2^{-n}) - f(X_n)| \leq 2^n L 2^{-n} = L,$$

et donc $(Z_n)_{n\geq 0}$ est bornée.

4. Puisque $(Z_n)_{n\geq 0}$ est une martingale bornée, elle converge p.s. et dans L^1 vers une v.a. Z. Cette v.a. est mesurable par rapport à $\bigcap_{n\geq 0}\sigma(X_n, X_{n+1},\ldots) = \sigma(X)$, où la dernière égalité vient de la partie 2. Par le lemme crucial, on a alors $Z = g(X)$ pour une fonction mesurable g, et puisque $(Z_n)_{n\geq 0}$ est bornée, g l'est aussi.
5. Par la décomposition dyadique de X (voir la solution de la partie 3), on a pour tout $n \in \mathbb{N}$, que $X - X_n$ est de loi uniforme sur $[0, 2^{-n})$ et indépendante de $\mathcal{F}_n$. Par conséquent,
$$\mathbb{E}[h(X)|X_n] = 2^n \int_{X_n}^{X_n+2^{-n}} h(u)\mathrm{d}u.$$
De plus, on sait que $(Z_n)_{n\geq 0}$ est une martingale fermée et donc $Z_n = \mathbb{E}[Z|\mathcal{F}_n] = \mathbb{E}[g(X)|X_n]$. Ceci donne la formule pour Z_n.
6. D'après la dernière partie, on a $f(X_n + 2^{-n}) - f(X_n) = \int_{X_n}^{X_n+2^{-n}} g(u)\mathrm{d}u$ p.s., pour tout $n \in \mathbb{N}$. Puisque $\mathbb{P}(X_n = a2^{-n}) > 0$ pour tout $a \in \{0,\ldots,2^n-1\}$, il en suit que
$$f((a+1)2^{-n}) - f(a2^{-n}) = \int_{a2^{-n}}^{(a+1)2^{-n}} g(u)\mathrm{d}u,$$
pour tout $n \in \mathbb{N}$ et $a \in \{0,\ldots,2^n-1\}$, et par conséquent, $f(x) = f(0) + \int_0^x g(u)\mathrm{d}u$ pour tout x de la forme $a2^{-n}$, $a \in \{1,\ldots,2^n\}$. Par continuité de f, on en déduit que cette formule est vrai pour tout $x \in [0,1]$.

Exercice 81 (Martingales exponentielles de la marche aléatoire)**.**

1. Soit $\theta \neq 0$. Par la loi des grands nombres, $S_n/n \to 0$ p.s. quand $n \to \infty$, si bien que presque sûrement,
$$\frac{1}{n}\log W_n(\theta) = \theta\frac{S_n}{n} - \log\cosh\theta \to -\log\cosh\theta < 0.$$
En particulier, $\log W_n(\theta) \to -\infty$ p.s. et donc $W_n(\theta) \to 0$ p.s. quand $n \to \infty$.
Si $W_n(\theta)$ était uniformément intégrable, alors elle convergerait dans L^1 vers sa limite. Or, pour tout $n \in \mathbb{N}$, $\mathbb{E}[W_n(\theta)] = \mathbb{E}[W_0(\theta)] = 1 \neq 0$. Par conséquent, la martingale $W_n(\theta)$ n'est pas u.i.
2. On sait d'après le cours que T_1 est un temps d'arrêt et que $T_1 < \infty$ p.s. La martingale arrêtée $W_n^{T_1}(\theta) = W_{n\wedge T_1}(\theta)$ converge alors p.s. vers $W_{T_1}(\theta)$. De plus, puisque $S_{n\wedge T_1} \leq 1$ pour tout $n \in \mathbb{N}$ et $\cosh(\theta) \geq 1$, on a $W_{n\wedge T_1}(\theta) \leq e^\theta$ pour tout $n \in \mathbb{N}$ et $\theta \geq 0$. Et bien sûr, $W_n \geq 0$ pour tout $n \in \mathbb{N}$. Par conséquent, la martingale arrêtée $W^{T_1}(\theta)$ est bornée et converge donc dans L^1 vers sa limite $W_{T_1}(\theta)$, d'où
$$\mathbb{E}[W_{T_1}(\theta)] = \mathbb{E}[W_0] = 1.$$
Or, $\mathbb{E}[W_{T_1}(\theta)] = \mathbb{E}[e^\theta\cosh(\theta)^{-T_1}]$, ce qui donne $\mathbb{E}[\cosh(\theta)^{-T_1}] = e^{-\theta}$.
3. On a pour tout $\theta \geq 0$,
$$\mathbb{E}[W_{T_{-1}}(\theta)] = \mathbb{E}[e^{-\theta}\cosh(\theta)^{T_{-1}}] = e^{-\theta}\mathbb{E}[\cosh(\theta)^{T_{-1}}] = e^{-2\theta},$$
d'après la dernière partie et le fait que T_{-1} et T_1 ont même loi.
4. Pour $\lambda > 0$, soit $\theta > 0$ tel que $e^\lambda = \cosh(\theta)$. En particulier, $\lambda \downarrow 0$ si et seulement si $\theta \downarrow 0$. Plus précisément, quand $\lambda \downarrow 0$,
$$e^\lambda = \cosh(\theta) = 1 + \frac{\cosh''(0)}{2}\theta^2 + o(\theta^2) = 1 + \tfrac{1}{2}\theta^2 + o(\theta^2) = e^{\frac{1}{2}\theta^2 + o(\theta^2)},$$
et donc $\theta \sim \sqrt{2\lambda}$ quand $\lambda \downarrow 0$. Avec la partie 2 de l'exercice, on a
$$1 - \mathbb{E}[e^{-\lambda T_1}] = 1 - e^{-\theta} \sim \theta \sim \sqrt{2\lambda}, \quad \text{quand } \lambda \downarrow 0.$$
5. Quand $\theta \in \mathbb{C}$, on peut montrer comme dans la partie 1 que $W_n(\theta)$ est une martingale à valeurs complexes (on admet que tous les résultats se transportent dans le complexe). En particulier, $\Re W_n(\theta)$ est encore

une martingale. Or, on a $Z_n(\gamma) = \Re W_n(i\gamma)$ pour $\gamma \in]-\frac{\pi}{2}, \frac{\pi}{2}[$, puisque $e^{i\gamma} = \cos(\gamma) + i\sin(\gamma)$ et $\cosh(i\gamma) = \cos(\gamma)$. Ceci permet de conclure.

Bien sûr, on peut se priver de cette explication élégante et rester dans le réel : on a pour $n \in \mathbb{N}$, par le théorème d'addition du cosinus,

$$\begin{aligned}\mathbb{E}[Z_{n+1}(\gamma)|\mathcal{F}_n] &= \cos(\gamma)^{-n-1}\mathbb{E}[\cos(\gamma S_n + \gamma X_{n+1})|\mathcal{F}_n] \\ &= \cos(\gamma)^{-n-1}\mathbb{E}[\cos(\gamma S_n)\cos(\gamma X_{n+1}) - \sin(\gamma S_n)\sin(\gamma X_{n+1})|\mathcal{F}_n] \\ &= \frac{Z_n(\gamma)}{\cos(\gamma)}\mathbb{E}[\cos(\gamma X_{n+1})] + \frac{\sin(\gamma S_n)}{\cos(\gamma)^{n+1}}\mathbb{E}[\sin(\gamma X_{n+1})] \\ &= \frac{Z_n(\gamma)}{\cos(\gamma)}\frac{\cos(\gamma) + \cos(-\gamma)}{2} + \frac{\sin(\gamma S_n)}{\cos(\gamma)^{n+1}}\frac{\sin(\gamma) + \sin(-\gamma)}{2} \\ &= Z_n(\gamma).\end{aligned}$$

6. On note d'abord que $T_{-a,a} \leq T_a < \infty$ p.s. par un théorème du cours. Puisque $Z_n(\frac{\pi}{2a}) = \cos(\frac{\pi}{2a}S_n)/\cos(\frac{\pi}{2a})^n$ et $\cos(\frac{\pi}{2}) = \cos(-\frac{\pi}{2}) = 0$, on a $Z_{T_{-a,a}}(\frac{\pi}{2a}) = 0$. Par conséquent
$$\mathbb{E}[Z_{T_{-a,a}}(\tfrac{\pi}{2a})] = 0 \neq 1 = \mathbb{E}[Z_0(\tfrac{\pi}{2a})].$$
Le théorème d'arrêt optionnel nous dit alors que les martingales $(Z_{n\wedge T_{-a,a}}(\frac{\pi}{2a}))_{n\geq 0}$ et $(Z_n(\frac{\pi}{2a}))_{n\geq 0}$ ne sont pas uniformément intégrables.
7. On a $0 \leq \cos(\frac{\pi}{2a}x) \leq 1$ pour tout $x \in [-a, a]$. Par conséquent, on a pour tout $n \in \mathbb{N}$,
$$0 \leq Z_{n\wedge T_{-a,a}}(\tfrac{\pi}{2a}) \leq \cos\left(\tfrac{\pi}{2a}\right)^{-(n\wedge T_{-a,a})} \leq \cos\left(\tfrac{\pi}{2a}\right)^{-T_{-a,a}}.$$
Si on avait $\mathbb{E}[\cos\left(\frac{\pi}{2a}\right)^{-T_{-a,a}}] < \infty$, la martingale $(Z_{n\wedge T_{-a,a}}(\frac{\pi}{2a}))_{n\geq 0}$ serait alors uniformément intégrable, ce qui est exclu d'après la dernière partie. Par conséquent, $\mathbb{E}[\cos\left(\frac{\pi}{2a}\right)^{-T_{-a,a}}] = \infty$.
8. Soit $\gamma \in]-\frac{\pi}{2}, \frac{\pi}{2}[$. Pour tout $n \leq T_{-a,a}$, on a $S_n \in [-a, a]$ et donc $Z_{n\wedge T_{-a,a}}(\gamma) \geq 0$ pour tout n. Par le lemme de Fatou, on a alors
$$\mathbb{E}[Z_{T_{-a,a}}(\gamma)] = \mathbb{E}[\lim_{n\to\infty} Z_{n\wedge T_{-a,a}}(\gamma)] \leq \lim_{n\to\infty}\mathbb{E}[Z_{n\wedge T_{-a,a}}(\gamma)] = \mathbb{E}[Z_0(\gamma)] = 1.$$
Par conséquent, $1 \geq \cos(\gamma a)\mathbb{E}[\cos(\gamma)^{-T_{-a,a}}]$, et donc
$$\mathbb{E}[\cos(\gamma)^{-T_{-a,a}}] \leq (\cos(\gamma a))^{-1} < \infty.$$
9. Soit λ_a la solution de $e^{\lambda_a} = \cos(\frac{\pi}{2a})^{-1}$. D'après la partie 7, on a alors $\mathbb{E}[e^{\lambda_a T_{-a,a}}] = \infty$, et donc, puisque $T_{-a,a} \geq 0$, $\mathbb{E}[e^{\lambda T_{-a,a}}] = \infty$ pour tout $\lambda \geq \lambda_a$. Par la partie 8, à cause de la continuité du cosinus, on a en plus $\mathbb{E}[e^{-\lambda T_{-a,a}}] < \infty$ pour tout $\lambda \in [0, \lambda_a[$, et on l'a trivialement pour $\lambda < 0$ puisque $T_{-a,a} \geq 0$. Ceci permet de conclure.

En ce qui concerne l'équivalent de λ_a quand $a \to \infty$, on remarque d'abord que pour $x \to 0$, on a
$$\cos(x)^{-1} = (1 - \frac{x^2}{2} + o(x^2))^{-1} = 1 + \frac{x^2}{2} + o(x^2) = e^{\frac{x^2}{2} + o(x^2)}.$$
Par conséquent, on a $\lambda_a \sim \frac{\pi^2}{8a^2}$ quand $a \to \infty$.

Exercice 82 (Modèle de Wright-Fisher). 1. Pour tout n, X_n est intégrable, car $0 \leq X_n \leq N$. Conditionnellement à $\mathcal{F}_n$, X_{n+1} suit la loi binomiale de paramètres N et $p = X_n/N$. Par conséquent,
$$\mathbb{E}[X_{n+1}\,|\,\mathcal{F}_n] = N \times \frac{X_n}{N} = X_n.$$
$(X_n)_{n\geq 0}$ est donc une martingale. Puisqu'elle est bornée (par N), elle converge alors dans L^1 et p.s. vers une limite X_∞.

2. M_n est bornée par $(N/(N-1))^n N^2$ et donc intégrable. On calcule l'espérance conditionelle :

$$\begin{aligned}
\mathbb{E}[M_{n+1}\,|\,\mathcal{F}_n] &= (\tfrac{N}{N-1})^{n+1}\sum_{i=0}^{N} i(N-i)\tbinom{N}{i}(\tfrac{X_n}{N})^i(1-\tfrac{X_n}{N})^{N-i}\\
&= (\tfrac{N}{N-1})^{n+1}\sum_{i=1}^{N-1}\tfrac{N!}{(N-i-1)!(i-1)!}(\tfrac{X_n}{N})^i(1-\tfrac{X_n}{N})^{N-i}\\
&= (\tfrac{N}{N-1})^{n+1}N(N-1)\tfrac{X_n}{N}(1-\tfrac{X_n}{N})\sum_{i=1}^{N-1}\tbinom{N-2}{i-1}(\tfrac{X_n}{N})^{i-1}(1-\tfrac{X_n}{N})^{N-i-1}\\
&= (\tfrac{N}{N-1})^{n}X_n(N-X_n)\sum_{i=0}^{N-2}\tbinom{N-2}{i}(\tfrac{X_n}{N})^{i}(1-\tfrac{X_n}{N})^{N-2-i}\\
&= M_n.
\end{aligned}$$

Ici, la dernière somme vaut 1, car c'est la somme sur les probabilités qu'une v.a. binomiale de paramètres $N-2$ et $p = X_n/N$ vaut i.

3. Puisque X_n converge dans L^1 vers X_∞, on a

$$\mathbb{E}[X_\infty] = \mathbb{E}[X_0] = k.$$

Aussi, puisque $X_n \to X_\infty$ p.s., $X_n(N-X_n) \to X_\infty(N-X_\infty)$ p.s. De plus, $0 \le X_n(N-X_n) \le N$ pour tout $n \in \mathbb{N}$. Par convergence dominée et le fait que M_n est une martingale,

$$\mathbb{E}[X_\infty(N-X_\infty)] = \lim_{n\to\infty}\mathbb{E}[X_n(N-X_n)] = \lim_{n\to\infty}(\tfrac{N-1}{N})^n\mathbb{E}[M_n] = \lim_{n\to\infty}(\tfrac{N-1}{N})^n k(N-k) = 0.$$

4. Puisque $X_\infty \in \{0,\dots,N\}$, $X_\infty(N-X_\infty) \ge 0$. Or, $\mathbb{E}[X_\infty(N-X_\infty)] = 0$ par la dernière partie, si bien que $X_\infty(N-X_\infty) = 0$ p.s., ou $X_\infty \in \{0,N\}$ p.s. Puisque $\mathbb{E}[X_\infty] = k$, cela détermine la loi de X_∞ comme étant

$$\mathbb{P}(X_\infty = 0) = \frac{N-k}{N},\quad \mathbb{P}(X_\infty = N) = \frac{k}{N}.$$

On interprète ce résultat comme suit : Presque sûrement, un des deux types a ou A envahit toute la population. La probabilité que le type a envahit la population est k/N, soit proportionelle au nombre d'individus de type a au début.

Exercice 83 (Un jeu de cartes)**.** 1. Après avoir retourné n cartes, il restent $52-n$ cartes dans le jeu. Par conséquent,

$$\mathbb{P}(A_{n+1}\,|\,R_n = j) = \frac{j}{52-n}.$$

2. On a soit $R_{n+1} = R_n$, soit $R_{n+1} = R_n - 1$. Or,

$$\begin{aligned}
\mathbb{P}(R_{n+1} = R_n\,|\,R_n) &= \mathbb{P}(A^c_{n+1}\,|\,R_n) = 1 - \frac{R_n}{52-n}\\
\mathbb{P}(R_{n+1} = R_n - 1\,|\,R_n) &= \mathbb{P}(A_{n+1}\,|\,R_n) = \frac{R_n}{52-n}.
\end{aligned}$$

Par conséquent,

$$\mathbb{E}[R_{n+1}\,|\,\mathcal{F}_n] = R_n - \mathbb{P}(R_{n+1} = R_n - 1\,|\,R_n) = R_n - \frac{R_n}{52-n} = R_n\frac{51-n}{52-n}.$$

En particulier,

$$\mathbb{E}[X_{n+1}\,|\,\mathcal{F}_n] = \frac{\mathbb{E}[R_{n+1}\,|\,\mathcal{F}_n]}{52-(n+1)} = R_n\frac{51-n}{(52-n)(51-n)} = X_n.$$

X_n est donc une martingale, et d'après la première partie de l'exercice, $X_n = \mathbb{P}(A_{n+1}\,|\,R_n) = \mathbb{P}(A_{n+1}\,|\,\mathcal{F}_n)$.

3. Il est naturel de supposer que τ est un temps d'arrêt, ce qui correspond à ce que le joueur ne connaisse que les cartes qui ont été retournées et doit donc baser sa stratégie sur cette information.

   Le joueur gagne si la $(\tau+1)$-ième carte retournée est rouge, soit, si A_{n+1} est vérifié quand $\tau = n$. Par conséquent, la probabilité de victoire est

$$\begin{aligned}\sum_{n=0}^{51} \mathbb{P}(A_{n+1},\, \tau = n) &= \sum_{n=0}^{51} \mathbb{E}[\mathbb{P}(A_{n+1} \,|\, \mathcal{F}_n)\mathbf{1}_{\tau=n}] \\ &= \sum_{n=0}^{51} \mathbb{E}[X_\tau \mathbf{1}_{\tau=n}] \\ &= \mathbb{E}[X_\tau].\end{aligned}$$

   Puisque τ est un temps d'arrêt borné, le théorème d'arrêt optionnel donne

$$\mathbb{E}[X_\tau] = \mathbb{E}[X_0] = \mathbb{E}[R_0/52] = 1/2.$$

   La probabilité de victoire est alors égale à $1/2$ pour n'importe quel temps d'arrêt τ, autrement dit, toutes les stratégies se valent.

Chaînes de Markov. Étude qualitative et quelques exemples.

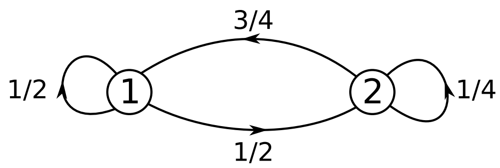

Figure 1 – Graphe décrivant la chaîne de Markov de l'Exercice 1.

Exercice 84. 1. D'après la description de $(X_n)_{n\geq 0}$, X_{n+1} ne dépend de $X_0, \dots, X_n$ qu'à travers X_n, plus précisément, on a

$$\mathbb{P}(X_{n+1} = y \,|\, X_0, \dots, X_n) = Q(X_n, y), \quad \text{où } Q = (Q(x,y))_{x,y=1,2} = \begin{pmatrix} 1/2 & 1/2 \\ 3/4 & 1/4 \end{pmatrix}.$$

   Par conséquent, $(X_n)_{n\geq 0}$ est une chaîne de Markov sur $\{1,2\}$ de matrice de transition Q.

2. On a

$$\begin{aligned} p_{n+1} &= \mathbb{P}(X_{n+1} = 1) \\ &= \mathbb{P}(X_{n+1} = 1 \,|\, X_n = 1)\mathbb{P}(X_n = 1) + \mathbb{P}(X_{n+1} = 1 \,|\, X_n = 2)\mathbb{P}(X_n = 2) \\ &= Q(1,1)p_n + Q(2,1)(1-p_n) \\ &= -p_n/4 + 3/4. \end{aligned}$$

3. On cherche p_7 quand $p_0 = 1$. Par la relation de récurrence, on a

$$p_7 = p_0(-1/4)^7 + (3/4)\sum_{i=0}^{6}(-1/4)^i = (-1/4)^7 + \frac{3(1-(-1/4)^7)}{4(1-(-1/4))} = \tfrac{3}{5} + \tfrac{2}{5}(-1/4)^7 \approx \tfrac{3}{5}.$$

4. Si $p_n = 3/5$, alors on a par la partie 2, $p_{n+1} = -(3/5)/4 + 3/4 = 3/4(1-1/5) = 3/5$. Par récurrence, on obtient alors que $p_0 = 3/5$ entraîne que $p_n = 3/5$ pour tout $n \in \mathbb{N}$.

5. Le coût moyen par jour se calcule

$$2 \times (3/5) + 4 \times (2/5) + 2 \times Q(1,2)(3/5) + 1 \times Q(2,1)(2/5) = (6+8+3+3/2)/5 = 17/5 = 3.4$$

   Il est alors de $3,40$ euros par jour.

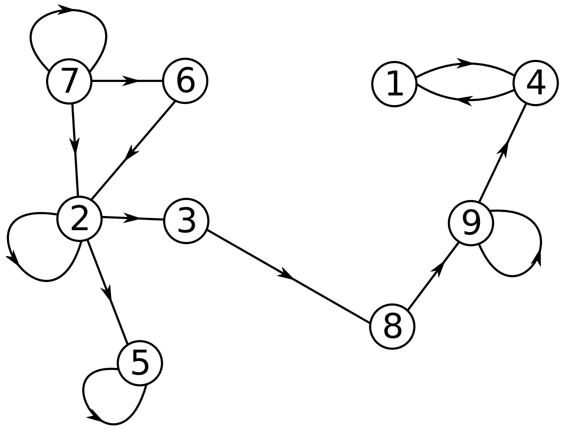

FIGURE 2 – Graphe décrivant la chaîne de Markov de l'Exercice 2. Les sommets correspondent aux états, les arêtes aux transitions entre états de probabilité non nulle.

Exercice 85. En inspectant la matrice de transition ou la Figure 2, on voit que 5 est un état absorbant et que les états 1 et 4 forment une classe de récurrence. Chacun des états restants communique avec 5 ou alors avec la classe {1,4}. Il en suit que les seules classes de récurrence sont $\{5\}$ et $\{1, 4\}$, que les états restants sont transitoires et que 5 est le seul état absorbant.

Exercice 86. On a pour tout $n \in \mathbb{N}$ et $y \in \{1, \dots, 6\}$, $X_{n+1} = \max(X_n, \xi_{n+1})$, si bien que

$$\begin{aligned}\mathbb{P}(X_{n+1} = y \,|\, X_0, \dots, X_n) &= \mathbb{P}(\max(X_n, \xi_{n+1}) = y \,|\, X_0, \dots, X_n) \\ &= \begin{cases} 1/6 & \text{si } y > X_n \\ X_n/6 & \text{si } y = X_n \\ 0 & \text{sinon} \end{cases} =: Q(X_n, y).\end{aligned}$$

L'état 6 est absorbant. Les états 1,...,5 communiquent avec 6 car $Q(x, 6) > 0$ pour tout $x \in \{1, \dots, 5\}$; par conséquent ils sont transitoires.

Pour trouver Q^n pour tout $n \in \mathbb{N}$, il est difficile de le faire directement par calcul matriciel. Mieux vaut raisonner de manière probabiliste : Pour tout $x \le y$, on a

$$\mathbb{P}_x(X_n \le y) = \mathbb{P}(\max(\xi_1, \dots, \xi_n) \le y) = (y/6)^n.$$

De plus, $\mathbb{P}_x(X_n \le x - 1) = 0$ pour tout $x = 1, \dots, 6$ et $n \in \mathbb{N}$. Par conséquent, on a

$$Q^n(x, y) = \begin{cases} \mathbb{P}_x(X_n \le y) - \mathbb{P}_x(X_n \le y - 1) = (y/6)^n - ((y-1)/6)^n & \text{si } y > x \\ \mathbb{P}_x(X_n \le x) = (x/6)^n & \text{si } y = x \\ 0 & \text{sinon.} \end{cases}$$

Exercice 87. Soit $Y_n = (X_n, n)$. Puisque Y_n et X_n sont des fonctions l'un de l'autre, on a alors $\sigma(Y_0, \dots, Y_n) = \sigma(X_0, \dots, X_n)$ pour tout $n \in \mathbb{N}$. Par conséquent, on a pour tout $n \in \mathbb{N}$,

$$\mathbb{P}(Y_{n+1} = (y, m) \,|\, Y_0, \dots, Y_n) = \mathbb{P}((X_{n+1}, n + 1) = (y, m) \,|\, X_0, \dots, X_n) = \begin{cases} Q_n(X_n, y) & \text{si } m = n + 1 \\ 0 & \text{sinon.} \end{cases}$$

Par conséquent, $(Y_n)_{n\ge0}$ est un processus de Markov sur $E \times \mathbb{N}$ de matrice de transition

$$\widetilde{Q}\Big((x, n), (y, m)\Big) = \begin{cases} Q_n(x, y) & \text{si } m = n + 1 \\ 0 & \text{sinon.} \end{cases}$$

Exercice 88. Notons $(X_n)_{n\ge0}$ la chaîne de Markov de l'énoncé et $\widetilde{T}_0 = \inf\{n \ge 1 : X_n = 0\}$. Alors

$$\mathbb{P}_0(\widetilde{T}_0 > n) = \mathbb{P}(X_i = i \,\forall i \le n) = \prod_{i=1}^{n} p_i.$$

En prenant la limite $n \to \infty$, on obtient

$$\mathbb{P}_0(\widetilde{T}_0 = \infty) = \prod_{i=1}^{\infty} p_i.$$

Par conséquent, 0 est un état récurrent si et seulement si $\prod_{i=1}^{\infty} p_i = 0$.

Exercice 89. 1. Si $X_n + \xi_{n+1} \geq 1$, alors cela signifie qu'un client sera servi au temps $n+1$, donc $X_{n+1} = X_n + \xi_{n+1} - 1 = (X_n + \xi_{n+1} - 1)^+$. En revanche, si $X_n + \xi_{n+1} = 0$, alors aucun client sera servi au temps $n+1$ et donc $X_{n+1} = 0 = (X_n + \xi_{n+1} - 1)^+$.

2. Par récurrence : pour $n = 0$, on a $X_0 = S_0 = S_0 - M_0$, car $M_0 = 0$. Supposons que c'est vrai pour un certain n. On a alors

$$X_{n+1} = (X_n + \xi_{n+1} - 1)^+ = (S_n - M_n + \xi_{n+1} - 1)^+ = (S_{n+1} - M_n)^+.$$

On a alors deux cas possibles :

(a) Si $S_{n+1} - M_n \geq 0$, alors $X_{n+1} = S_{n+1} - M_n$ et $M_{n+1} = \min(M_n, S_{n+1}) = M_n$.

(b) Si $S_{n+1} - M_n < 0$, alors $X_{n+1} = 0$ et $M_{n+1} = \min(M_n, S_{n+1}) = S_{n+1}$.

Dans les deux cas, $X_{n+1} = S_{n+1} - M_{n+1}$.

3. Pour tout n, les variables $X_0, \dots, X_n$ sont mesurables par rapport à $\sigma(X_0, \xi_1, \dots, \xi_n)$. Par conséquent, puisque ξ_{n+1} est indépendante de $\sigma(X_0, \xi_1, \dots, \xi_n)$ par hypothèse, elle est également indépendante de $\sigma(X_0, \dots, X_n)$ pour tout n. Par conséquent,

$$\mathbb{P}(X_{n+1} = y \,|\, X_0, \dots, X_n) = \mathbb{P}((X_n + \xi_{n+1} - 1)^+ = y \,|\, X_0, \dots, X_n) = Q(X_n, y),$$

avec

$$\begin{aligned} Q(x,y) &= \begin{cases} \mu(y+1-x) & \text{si } x \geq 1 \text{ ou } y \geq 1 \\ \mu(1) + \mu(0) & \text{si } x = 0 \text{ et } y = 0 \end{cases} \\ &= \mu(y+1-x) + \mu(0)\mathbb{1}_{x=y=0}. \end{aligned}$$

Par conséquent, $(X_n)_{n\geq 0}$ est une chaîne de Markov avec matrice de transition Q.

Pour montrer que la chaîne est irréductible, on remarque que $Q(x, x-1) = \mu(0) > 0$ pour tout $x \geq 1$, et donc $x \rightsquigarrow x-1$ pour tout $x \geq 1$. Par récurrence, cela donne $x \rightsquigarrow y$ pour tout $y < x$. De plus, puisque $\mu(k) > 0$ pour un $k \geq 2$, on a $Q(x, x+k-1) = \mu(k) > 0$ pour tout $x \geq 1$ et donc par récurrence $x \rightsquigarrow x + (k-1)n$ pour tout $x \geq 1$ et $n \in \mathbb{N}^*$.

Si x et y sont arbitraires, posons $n = y+1$, si bien que $x + (k-1)n > y$. Par ce qui précède, on a alors $x \rightsquigarrow x + (k-1)n \rightsquigarrow y$, si bien que $x \rightsquigarrow y$. La chaîne est alors irréductible.

4. Par la loi forte des grands nombres, on a $S_n/n \to \mathbb{E}[\xi_1 - 1]$ p.s. quand $n \to \infty$. Si $\mathbb{E}[\xi_1] > 1$, alors cela signifie que $S_n \to \infty$ p.s. Puisque $M_n \leq 0$ pour tout $n \in \mathbb{N}$, on a d'après la seconde partie : $X_n \geq S_n$ pour tout $n \in \mathbb{N}$, et donc $X_n \to \infty$ p.s. quand $n \to \infty$.

5. On a pour tout $n \leq T$, $M_n = 0$. Par la second partie de l'exo, on a alors $X_n = S_n$ pour tout $n \leq T$.

Posons $X_0 = x \geq 1$. Montrons que $T < \infty$ p.s. Quand $\mathbb{E}[\xi_1] < 1$, ceci est une conséquence immédiate de la loi forte des grands nombres. Quand $\mathbb{E}[\xi_1] = 1$, considérons la martingale arrêtée S^T. Cette martingale est positive est converge donc p.s. vers une limite. Or, puisque p.s., S_n ne converge pas quand $n \to \infty$, cela implique que T doit être finie presque sûrement. Donc $T < \infty$ p.s. dans tous les cas.

Puisque $X_n = S_n$ pour $n \leq T$ et $T < \infty$ p.s., l'état 0 est atteint p.s. pour tout $X_0 = x \geq 1$. Ceci montre que $\widetilde{T}_0 < \infty$ p.s. et donc que 0 est récurrent. Par irréductibilité, ceci est vrai pour tous les états, et donc la chaîne est récurrente.

Exercice 90. Notons $(Z_n)_{n\geq 0}$ la marche aléatoire sur l'arbre et posons $Y_n = d(\emptyset, Z_n)$, ou $d(\cdot,\cdot)$ est la distance dans l'arbre et $\emptyset$ est la racine de l'arbre. On vérifie alors aisément que $(Y_n)_{n\geq 0}$ est encore une chaîne de Markov (sur $\mathbb{N}$) de matrice de transition

$$Q_Y(x,y) = (2/3)\mathbb{1}_{y=x+1} + (1/3)\mathbb{1}_{y=x-1} + (1/3)\mathbb{1}_{x=0,y=1}.$$

Cette chaîne de Markov ressemble beaucoup à la chaîne $(X_n)_{n\geq 0}$ du dernier exercice si on prend $\mu(0) = 1/3$ et $\mu(2) = 2/3$ (si bien que $\mathbb{E}[\xi_1] = 4/3 > 1$). En effet, si on note Q_X la matrice de transition de cette chaîne, on a

$$Q_X(x, y) = (2/3)\mathbb{1}_{y=x+1} + (1/3)\mathbb{1}_{y=x-1} + (1/3)\mathbb{1}_{x=0,y=0},$$

si bien que $Q_X(x, y) = Q_Y(x, y)$ si $x \geq 1$ et quand $x = 0$, on a $Q_Y(0, 1) = 1$ et $Q_X(0, 1) = 2/3 = 1 - Q_X(0, 0)$. Ceci implique qu'on peut plonger $(Y_n)_{n\geq 0}$ dans $(X_n)_{n\geq 0}$: définissons une suite de temps aléatoires $N_0, N_1, \ldots$ par

$$N_0 = 0, \quad \text{et} \quad N_{n+1} = \begin{cases} N_n + 1 & \text{si } X_{N_n} \neq 0 \\ \inf\{k \geq N_n + 1 : X_k \neq 0\} & \text{si } X_{N_n} = 0. \end{cases}$$

On vérifie alors aisément que $(X_{N_n})_{n\geq 0}$ est une chaîne de Markov de matrice de transition Q_Y et a donc même loi que $(Y_n)_{n\geq 0}$.

Le dernier exercice nous donne maintenant que $X_n \to +\infty$ p.s. quand $n \to \infty$. Par conséquent, $Y_n \to +\infty$ p.s. également et donc $(Y_n)_{n\geq 0}$ est transitoire.

Chaînes de Markov : Mesures invariantes, périodicité, martingales, fonctions harmoniques.

Exercice 91 (Chaîne de naissance et mort).

1. Si $p_i > 0$ et $q_{i+1} > 0$ pour tout $i \geq 0$, alors on a $i \leftrightsquigarrow i+1$ pour tout $i \geq 0$ et donc par transitivité, $i \leftrightsquigarrow j$ pour tout $i < j$. La chaîne est donc irréductible. En revanche, si $p_k = 0$ pour un certain $k \geq 0$, alors on a $Q(i, j) = 0$ pour tout $i \leq k$ et $j > k$. Ceci montre que $i \not\rightsquigarrow j$ pour tout $i \leq k$ et $j > k$, et donc la chaîne n'est pas irréductible. Si $q_{i+1} = 0$ pour un $i \geq 0$, on fait un raisonnement analogue.
2. Il vient de la définition de u que $u(a) = 0$ et $u(b) = 1$. Soit $a < x < b$. Alors,

$$u(x) = \mathbb{P}_x(T_a > T_b) = \sum_y \mathbb{P}_x(T_a > T_b \,|\, X_1 = y)\mathbb{P}_x(X_1 = y).$$

Puisque $x \notin \{a, b\}$, on a $T_a \geq 1$ et $T_b \geq 1$ presque sûrement sous $\mathbb{P}_x$, et donc $T_a = T_a \circ \theta + 1$ et $T_b \circ \theta + 1$ p.s. La propriété de Markov donne alors

$$\mathbb{P}_x(T_a > T_b \,|\, X_1 = y) = \mathbb{P}_x(T_a \circ \theta > T_b \circ \theta \,|\, X_1 = y) = \mathbb{P}_y(T_a > T_b) = u(y).$$

Par conséquent,

$$u(x) = \sum_y Q(x, y)u(y) = Qu(x).$$

Dans notre cas, cela donne

$$u(x) = q_x u(x-1) + r_x u(x) + p_x u(x+1),$$

ce qui, en écrivant $u(x) = 1 \times u(x) = (q_x + r_x + p_x)u(x)$, donne

$$(q_x + p_x)u(x) = q_x u(x-1) + p_x u(x+1), \quad \text{et donc} \quad u(x+1) - u(x) = \frac{q_x}{p_x}(u(x) - u(x-1)).$$

Par récurrence, cela donne

$$\forall a \leq x < b : u(x+1) - u(x) = \frac{\gamma(x)}{\gamma(a)}(u(a+1) - u(a)) = C\gamma(x),$$

avec $C = (u(a+1) - u(a))/\gamma(a) = u(a+1)/\gamma(a)$. En sommant cette égalité, on obtient ainsi pour $a \leq x \leq b$,

$$u(x) = u(x) - u(a) = \sum_{y=a}^{x-1} u(y) - u(y+1) = C\sum_{y=a}^{x-1} \gamma(y).$$

Ceci est en particulier vrai pour $x = b$, ce qui donne

$$1 = u(b) = C \sum_{y=a}^{b-1} \gamma(y),$$

et donc $C = (\sum_{x=a}^{b-1} \gamma(x))^{-1}$. Ceci donne finalement,

$$u(x) = \frac{\sum_{y=a}^{x-1} \gamma(y)}{\sum_{y=a}^{b-1} \gamma(y)}.$$

Dans le cas particulier où $p_x = q_x$ pour tout $x \geq 1$, on a $\gamma(x) = 1$ pour tout $x \geq 0$, et donc

$$u(x) = \frac{x-a}{b-a}.$$

3. Considérons l'événement $\{T_0 = \infty\}$. On a $\mathbb{P}_1$-presque sûrement $T_n \geq n-1$ pour tout $n \geq 1$; si bien que

$$\{T_0 = \infty\} = \lim_{n\to\infty} \{T_0 > n\} \supset \lim_{n\to\infty} \{T_0 > T_n\}, \quad \mathbb{P}_1\text{-p.s.}$$

Pour l'autre inégalité, on remarque que $T_n < \infty$ p.s., car soit la chaîne est récurrente, dansquel cas chaque état est visité un nombre infini de fois (irréductibilité !), soit la chaîne est transiente, dansquel cas elle tend vers ∞ p.s. et donc doit passer par l'état n, les seules transitions étant $x \to x+1$ et $x \to x-1$. Par conséquent, on a trivialement, $\mathbb{P}_1(T_0 = \infty) \leq \mathbb{P}_1(T_0 > T_n)$ pour tout $n \in \mathbb{N}$. Avec ce qui précède, cela donne

$$\mathbb{P}_1(T_0 = \infty) = \lim_{n\to\infty} \mathbb{P}_1(T_0 > T_n).$$

La dernière partie nous donne alors

$$\mathbb{P}_1(T_0 = \infty) = \lim_{n\to\infty} \left(\sum_{y=0}^{n-1} \gamma(y)\right)^{-1} = \left(\sum_{y=0}^{\infty} \gamma(y)\right)^{-1}.$$

Par conséquent, $\mathbb{P}_1(T_0 = \infty) = 0$ si et seulement si $\sum_{y=0}^{\infty} \gamma(y) = \infty$.
Pour montrer que la chaîne est récurrente si et seulement si $\sum_{y=0}^{\infty} \gamma(y) = \infty$, il suffit par irréductibilité de montrer que 0 est un état récurrent, donc que $\mathbb{P}_0(\widetilde{T}_0 = \infty) = 0$ si et seulement si $\sum_{y=0}^{\infty} \gamma(y) = \infty$. Or, puisque $Q(0, i) = 0$ pour tout $i \geq 2$, la propriété de Markov donne,

$$\mathbb{P}_0(\widetilde{T}_0 = \infty) = \mathbb{P}_0(\widetilde{T}_0 = \infty,\, X_1 = 0) + \mathbb{P}_0(\widetilde{T}_0 = \infty,\, X_1 = 1) = 0 + p_0 \mathbb{P}_1(T_0 = \infty),$$

et puisque $p_0 > 0$, cela montre que $\mathbb{P}_0(\widetilde{T}_0 = \infty) = 0$ si et seulement si $\sum_{y=0}^{\infty} \gamma(y) = \infty$.

4. Une mesure μ est réversible si et seulement si pour tous $x, y \geq 0$,

$$\mu(x) Q(x,y) = \mu(y) Q(y,x)$$

Puisque $Q(x,y) = 0$ dès lors que $|x-y| \geq 2$ et puisque l'égalité est triviale si $x = y$ il suffit de la vérifier dans le cas $|x-y| = 1$, où, en échangeant les rôles de x et y, dans le cas $y = x+1$. On a alors

$$\mu(x) Q(x, x+1) = \mu(x+1) Q(x+1, x) \iff \mu(x+1) = \frac{p_x}{q_{x+1}} \mu(x).$$

Par récurrence, cela montre que μ est réversible si et seulement si

$$\mu(x) = \frac{p_0 \cdots p_{x-1}}{q_1 \cdots q_x} \mu(0),$$

et donc il existe une seule mesure réversible ζ avec $\zeta(0) = 1$ donnée par $\zeta(x) = \frac{p_0 \cdots p_{x-1}}{q_1 \cdots q_x}$ pour $x \geq 0$ (avec la convention qu'un produit sur un ensemble vide vaut 1).
Si $M := \sum_{x=0}^{\infty} \frac{p_0 \cdots p_{x-1}}{q_1 \cdots q_x} < \infty$, alors la mesure $\pi = \zeta / M$ est une mesure de probabilité réversible et donc invariante. Par irréductibilité, ceci montre que la chaîne de Markov est récurrente positive. Inversement,

si la chaîne est récurrente positive, alors il existe une mesure de probabilité invariante π. Par un théorème du cours, on sait qu'une chaîne de Markov récurrente et irréductible admet une unique mesure invariante (modulo multiplication par une constante), et donc $\zeta = C\pi$ pour une constante $C \geq 0$. Par conséquent, on a

$$\sum_{x=0}^{\infty} \frac{p_0 \cdots p_{x-1}}{q_1 \cdots q_x} = \sum_{x=0}^{\infty} \zeta(x) = C \sum_{x=0}^{\infty} \pi(x) = C < \infty.$$

5. Dans le cas présent, on a par l'hypothèse $p < q$,

$$\sum_{x=0}^{\infty} \frac{p_0 \cdots p_{x-1}}{q_1 \cdots q_x} = \sum_{x=0}^{\infty} (p/q)^x < \infty.$$

On est alors dans le cas récurrent positif. Soit π la mesure de probabilité invariante de la partie précédente. On a alors d'après un théorème du cours, pour tout $i \geq 0$,

$$\mathbb{E}_i[T_i] = \frac{1}{\pi(i)} = \frac{\sum_{x=0}^{\infty} (p/q)^x}{(p/q)^i} = \frac{(q/p)^i}{1 - p/q}.$$

Exercice 92 (Une chaîne périodique). 1. Pour un entier n, notons n mod m son reste pour la division par m. On a alors $i \rightsquigarrow (i+1) \bmod m$ pour tout i. Par récurrence, on a $i \rightsquigarrow (i+k) \bmod m$ pour tout $k \geq 1$. Par conséquent, $i \leftrightsquigarrow j$ pour tout $i, j \in \{0, \dots, m-1\}$ et donc la chaîne est irréductible.

2. Soit $n \in \mathbb{N}$ et $i, j \in \{0, \dots, m-1\}$. On a pour tout chemin $x_0, \dots, x_n$,

$$\prod_{k=1}^{n} Q(x_{k-1}, x_k) = \begin{cases} 1 & \text{si } x_k \equiv x_{k-1} + 1 \bmod m \text{ pour tout } k \\ 0 & \text{sinon.} \end{cases}$$

Le seul chemin $i = x_0, \dots, x_n$ tel que $x_k \equiv x_{k-1} + 1 \bmod m$ pour tout k est le chemin $x_k = i + k \bmod m$. Ce chemin satisfait $x_n = j$ si et seulement si $j = i + k \bmod m$. Il vient,

$$\begin{aligned} Q^n(i,j) &= \sum_{i = x_0, \dots, x_n = j} \prod_{k=1}^{n} Q(x_{k-1}, x_k) \\ &= \begin{cases} 1 & \text{si } y \equiv x + n \bmod m \\ 0 & \text{sinon.} \end{cases}. \end{aligned}$$

3. On a pour tout $i \in \{0, \dots, m-1\}$,

$$Q^n(i,i) > 0 \iff i \equiv i + n \mod m \iff n \equiv 0 \mod m \iff m | n.$$

Par conséquent, on a pour tout $i \in \{0, \dots, m-1\}$,

$$\text{pgcd}(\{n \in \mathbb{N}^* : Q^n(i,i) > 0\}) = m,$$

et donc la période de la chaîne est m.

4. Soit μ la probabilité uniforme sur $\{0, \dots, m-1\}$. Alors pour tout $j \in \{0, \dots, m-1\}$,

$$\mu Q(j) = \sum_{i=0}^{m-1} \mu(i) Q(i,j) = \mu(j - 1 \bmod m) = \mu(j).$$

Par irréductibilité, la probabilité stationnaire est unique et donc il n'y en a pas d'autres.

Exercice 93 (*Lazy chain*). 1. On a pour tout $n \in \mathbb{N}$ et $y \in E$,

$$\mathbb{P}_x(Y_{n+1} = y \,|\, Y_0, \dots, Y_n) = \mathbb{P}_x(Y'_{n+1} = y,\ B_{n+1} = 1 \,|\, Y_0, \dots, Y_n) + \mathbb{P}_x(Y_n = y,\ B_{n+1} = 0 \,|\, Y_0, \dots, Y_n).$$

Puisque B_{n+1} et Y'_{n+1} sont indépendantes conditionellement à $Y_0, \dots, Y_n$, de lois respectives $\mathcal{B}(p)$ et $Q(Y_n, \cdot)$, on a

$$\mathbb{P}_x(Y'_{n+1} = y,\ B_{n+1} = 1 \,|\, Y_0, \dots, Y_n) = pQ(Y_n, y).$$

De plus, puisque $\mathbb{1}_{(Y_n=y)}$ est mesurable par rapport à $Y_0, \ldots, Y_n$, on a

$$\begin{aligned}\mathbb{P}_x(Y_n = y,\ B_{n+1} = 0 \,|\, Y_0, \ldots, Y_n) &= \mathbb{1}_{(Y_n=y)} \mathbb{P}_x(B_{n+1} = 0 \,|\, Y_0, \ldots, Y_n) \\ &= (1-p)\mathbb{1}_{(Y_n=y)}.\end{aligned}$$

Par conséquent,

$$\mathbb{P}_x(Y_{n+1} = y \,|\, Y_0, \ldots, Y_n) = (pQ + (1-p)\,\mathrm{Id})(Y_n, y),$$

et donc $(Y_n)_{n\geq 0}$ est une chaîne de Markov de matrice de transition $Q_p := pQ + (1-p)\,\mathrm{Id}$.

2. Puisque $Q_p \geq pQ$, on a $Q_p^n \geq (pQ)^n = p^n Q^n$. Par conséquent, puisque $p > 0$, on a $Q_p^n(x, y) > 0$ si $Q^n(x, y) > 0$, et donc $\sum_{n=0}^{\infty} Q_p^n(x, y) > 0$ si $\sum_{n=0}^{\infty} Q^n(x, y) > 0$. Ceci montre que $(Y_n)_{n\geq 0}$ est irréductible si $(X_n)_{n\geq 0}$ l'est.
3. Soit μ une mesure sur E. Puisque $p > 0$, on a

$$\mu Q = \mu \iff p\mu Q = p\mu \iff p\mu Q + (1-p)\mu = p\mu + (1-p)\mu \iff \mu Q_p = \mu,$$

et donc μ est invariante pour $(Y_n)_{n\geq 0}$ si et seulement si μ est invariante pour $(X_n)_{n\geq 0}$.
4. Soit $x \in E$. Puisque $p < 1$, on a $Q_p(x, x) \geq (1-p)\,\mathrm{Id}(x, x) > 0$. Par conséquent,

$$\mathrm{pgcd}(\{n \in \mathbb{N}^* : Q_p^n(x, x) > 0\}) = 1,$$

et donc la chaîne est apériodique.
5. Si $(X_n)_{n\geq 0}$ est irréductible et récurrente positive, alors $(Y_n)_{n\geq 0}$ l'est également, par les deux premières parties de l'exercice. De plus, par la dernière partie de l'exercice, $(Y_n)_{n\geq 0}$ est apériodique. Par un théorème du cours, Y_n converge alors en loi vers la probabilité stationnaire quand $n \to \infty$.

   Cette convergence n'a pas forcément lieu pour $(X_n)_{n\geq 0}$ quand celle-ci est périodique, la chaîne de l'Exercice 2 en donne un exemple.

Exercice 94. Pour une fonction $f \in \mathcal{H}$, notons $Qf(x) = \sum_{y\in E} Q(x, y)f(y)$, ce qui définit un opérateur linéaire Q sur $\mathcal{H}$. Soit $n \in \mathbb{N}^*$. Alors

$$\mathbb{E}_x[(f(X_n) - f(X_{n-1})) \,|\, X_0, \ldots, X_{n-1}] = \sum_{y\in E} Q(X_{n-1}, y)f(y) - f(X_{n-1}) = (Q - \mathrm{Id})f(X_{n-1}).$$

Par conséquent, avec l'operateur linéaire $A := Q - \mathrm{Id}$, le processus M_n^f est une martingale pour la filtration naturelle de $(X_n)_{n\geq 0}$.

Exercice 95 (Fonctions harmoniques et problème de Dirichlet). 1. Puisque x est non-absorbant, l'ensemble $\{y \in E : y \neq x,\ Q(x, y) > 0\}$ n'est pas vide. Par définition de l'harmonicité, on a alors

$$\begin{aligned}(1 - Q(x, x))u(x) &= Qu(x) - Q(x, x)u(x) \\ &= \sum_{y\in E,\ y\neq x} Q(x, y)u(y) \\ &= \sum_{y\in E,\ y\neq x,\ Q(x,y)>0} Q(x, y)u(y) \\ &\leq \sup\{u(y) \,|\, y \neq x, Q(x, y) > 0\} \sum_{y\in E,\ y\neq x,\ Q(x,y)>0} Q(x, y) \\ &= \sup\{u(y) \,|\, y \neq x, Q(x, y) > 0\}(1 - Q(x, x)),\end{aligned}$$

où la dernière égalité provient du fait que $\sum_{y\in E} Q(x, y) = 1$ pour tout $x \in E$. Puisque x est non-absorbant, on a $1 - Q(x, x) > 0$, ce qui donne

$$u(x) \leq \sup\{u(y) \,|\, y \neq x, Q(x, y) > 0\}.$$

2. Si $(u(X_n))_{n\geq 0}$ est une martingale par rapport à $(\mathcal{F}_n)_{n\geq 0}$ (sous $\mathbb{P}_x$ pour tout $x \in E$), alors $\mathbb{E}_x(u(X_1)) = \mathbb{E}_x(u(X_0)) = x$ pour tout $x \in E$, et donc u est harmonique. Inversement, si u est harmonique, alors pour tout $n \in \mathbb{N}^*$, puisque $(X_n)_{n\geq 0}$ est une chaîne de Markov, pour tout $x \in E$,

$$\mathbb{E}_x[u(X_n)\,|\,\mathcal{F}_{n-1}] = \sum_{y\in E} Q(X_{n-1}, y)u(y) = Qu(X_{n-1}) = u(X_{n-1}),$$

par l'harmonicité de u. Par conséquent, $(u(X_n))_{n\geq 0}$ est une martingale par rapport à $(\mathcal{F}_n)_{n\geq 0}$ sous $\mathbb{P}_x$ pour tout $x \in E$.

3. On sait que T_{A^c} est un temps d'arrêt, et donc $\{T_{A^c} \leq n\} \in \mathcal{F}_n$ et $\{T_{A^c} > n\} \in \mathcal{F}_n$. On a pour tout $n \in \mathbb{N}$ et $x \in E$,

$$\begin{aligned}\mathbb{P}_x(X_{n+1\wedge T_{A^c}} = y\,|\,\mathcal{F}_n) &= \mathbb{P}_x(X_{n+1} = y\,|\,\mathcal{F}_n)\mathbb{1}_{(T_{A^c}>n)} + \mathbb{P}_x(X_{n\wedge T_{A^c}} = y\,|\,\mathcal{F}_n)\mathbb{1}_{(T_{A^c}\leq n)}\\ &= Q(X_{n\wedge T_{A^c}}, y)\mathbb{1}_{(T_{A^c}>n)} + \mathbb{1}_{(X_{n\wedge T_{A^c}}=y)}\mathbb{1}_{(T_{A^c}\leq n)}\end{aligned}$$

Par définition de T_{A^c}, on a $T_{A^c} > n$ si et seulement si $X_{n\wedge T_{A^c}} \in A$. Par conséquent,

$$\begin{aligned}\mathbb{P}_x(X_{n+1\wedge T_{A^c}} = y\,|\,\mathcal{F}_n) &= Q(X_{n\wedge T_{A^c}}, y)\mathbb{1}_{(X_{n\wedge T_{A^c}}\in A)} + \mathbb{1}_{(X_{n\wedge T_{A^c}}=y)}\mathbb{1}_{(X_{n\wedge T_{A^c}}\in A^c)}\\ &= Q_A(X_{n\wedge T_{A^c}}, y),\end{aligned}$$

avec

$$Q_A(x,y) = \begin{cases} Q(x,y), & x \in A \\ \mathrm{Id}(x,y), & x \in A^c \end{cases}.$$

Puisque $\sigma(X_{0\wedge T_{A^c}}, \ldots, X_{n\wedge T_{A^c}}) = \mathcal{F}_{n\wedge T_{A^c}} \subset \mathcal{F}_n$ pour tout $n \in \mathbb{N}$, on a également

$$\mathbb{P}_x(X_{n+1\wedge T_{A^c}} = y\,|\,\mathcal{F}_{n\wedge T_{A^c}}) = Q_A(X_{n\wedge T_{A^c}}, y).$$

Par conséquent, le processus $(X_{n\wedge T_{A^c}})_{n\geq 0}$ est une chaîne de Markov de matrice de transition Q_A.
Si u est harmonique sur A, alors on a pour tout $x \in A$: $Q_A u(x) = Qu(x) = u(x)$, et pour tout $x \in A^c$: $Q_A u(x) = \mathrm{Id}\, u(x) = u(x)$. Par conséquent, $Q_A u = u$ et donc u est harmonique pour la chaîne $(X_{n\wedge T_{A^c}})_{n\geq 0}$.

4. Soit u harmonique sur A. Alors d'après la dernière partie de l'exercice, u est harmonique pour la chaîne $(X_{n\wedge T_{A^c}})_{n\geq 0}$. Par la deuxième partie de l'exercice, $(u(X_{n\wedge T_{A^c}}))_{n\geq 0}$ est alors une martingale pour la filtration $(\mathcal{F}_{n\wedge T_{A^c}})_{n\geq 0}$, pour tout $x \in E$. Fixons $x \in A$. Puisque $X_{n\wedge T_{A^c}}$ prend valeurs dans $A \cup \partial A$, $\mathbb{P}_x$-p.s. pour tout $n \in \mathbb{N}$, et que $A \cup \partial A$ est un ensemble fini, la martingale $(u(X_{n\wedge T_{A^c}}))_{n\geq 0}$ est bornée dans L^∞ et donc une martingale fermée sous $\mathbb{P}_x$. Notons $X_{\infty\wedge T_{A^c}}$ sa limite (dans L^1 et p.s.) et notons que $X_{\infty\wedge T_{A^c}} = X_{T_{A^c}}$ quand $X_{T_{A^c}} < \infty$.
Par hypothèse, on a $\mathbb{P}_x(T_{A^c} < \infty) = 1$. De plus, par définition de ∂A, $\mathbb{P}_x(X_{T_{A^c}} \in \partial A,\ T_{A^c} < \infty) = 1$. Par conséquent,

$$u(x) = \mathbb{E}_x[u(X_{\infty\wedge T_{A^c}})] = \mathbb{E}_x[u(X_{T_{A^c}})\mathbb{1}_{(X_{T_{A^c}}\in\partial A,\ T_{A^c}<\infty)}].$$

Il en suit que $u(x)$ est déterminée par les valeurs de u sur ∂A. Puisque $x \in A$ était arbitraire, ceci permet de conclure.

Exercice 96 (Bonus : Fonctions harmoniques et propriété de Liouville)**.**

1. Soit u une fonction harmonique bornée. Par l'Exercice 5.2, le processus $(u(X_n))_{n\geq 0}$ est alors une martingale bornée. Fixons un élément $0 \in E$ quelconque. Soit T_0 le temps d'atteinte de 0. Puisque la chaîne de Markov est récurrente, on a $\mathbb{P}_x(T_0 < \infty) = 1$ pour tout $x \in E$. Par le théorème d'arrêt (u est bornée !), on a alors

$$\forall x \in E : u(x) = \mathbb{E}_x[u(X_0)] = \mathbb{E}_x[u(X_{T_0})] = u(0).$$

Ceci montre que u est constante et que la chaîne est Liouville.

2. Soit u une fonction harmonique bornée. Soit $x, y \in E$. Par l'Exercice 5.2, on a alors pour tout $n \in \mathbb{N}$, $u(x) = \mathbb{E}[X_n^x]$ et $u(y) = \mathbb{E}[X_n^y]$. En particulier,

$$|u(x) - u(y)| = \limsup_{n\to\infty} |\mathbb{E}[u(X_n^x)] - \mathbb{E}[u(X_n^y)]|.$$

Par hypothèse, on peut coupler X^x et X^y de telle façon que

$$\mathbb{P}(\exists N\ \forall n \geq N : X_n^x = X_n^y) = 1.$$

En particulier, on a $\mathbb{P}(X_n^x = X_n^y) \to 1$ quand $n \to \infty$, et donc

$$\begin{aligned}|u(x) - u(y)| &= \limsup_{n\to\infty} |\mathbb{E}[u(X_n^x)] - \mathbb{E}[u(X_n^y)]| \\ &\leq \limsup_{n\to\infty} \|u\|_\infty \mathbb{P}(X_n^x \neq X_n^y) \\ &= 0.\end{aligned}$$

Ceci montre que u est constante et que la chaîne est Liouville.

3. Soit u une fonction. Puisque $p > 0$, on a

$$Q_p u = u \iff pQu + (1-p)u = u \iff pQu = pu \iff Qu = u,$$

ce qui donne le résultat.

4. Nous décrivons d'abord en mots ce que nous allons faire avant de l'écrire formellement : on considère la version *lazy* de la MAS avec $p = 1/2$. Nous pouvons alors coupler les marches X^x et X^y de la façon suivante. A chaque pas, nous choisissons la même coordonnée pour les deux marches. Si X^x et X^y coïncident en cette coordonnée, on les fait bouger pareil. Sinon, on fait bouger l'une des deux au hasard, et l'autre reste sur place. En utilisant la récurrence de la MAS sur $\mathbb{Z}$, on peut ainsi faire converger chacune des d coordonnées, si bien que $X_n^x = X_n^y$ à partir d'un certain rang.

Maintenant formellement : Soit $x, y \in \mathbb{Z}^d$. Soit $(U_n, B_n, J_n)_{n\geq 0}$ une famille iid de couples aléatoires de loi $\mathrm{Unif}(\{1,\dots,d\}) \otimes \mathcal{B}(1/2) \otimes (\delta_{-1} + \delta_1)/2$. Notons $v(k)$ la k-ième coordonnée d'un vecteur v. Notons également $\mathrm{e}_1, \dots, \mathrm{e}_d$ la b.o.n. usuelle de $\mathbb{R}^d$. On définit alors

$$\begin{aligned} &X_0^x = x, \qquad \forall n \in \mathbb{N} : X_{n+1}^x = X_n^x + J_n \mathrm{e}_{U_n} \mathbb{1}_{(B_n=1)} \\ &X_0^y = y, \qquad \forall n \in \mathbb{N} : X_{n+1}^y = X_n^y + J_n \mathrm{e}_{U_n} (\mathbb{1}_{(B_n=1,\ X_n^y(U_n) = X_n^x(U_n))} + \mathbb{1}_{(B_n=0,\ X_n^y(U_n)\neq X_n^x(U_n))})\end{aligned}$$

Alors $(X_n^x)_{n\geq 0}$ est évidemment une version *lazy* (avec $p = 1/2$) de la MAS sur $\mathbb{Z}^d$. Montrons que $(X_n^y)_{n\geq 0}$ en est une aussi. Si

$$\mathcal{F}_n = \sigma((U_k, B_k, J_k),\ k = 1, \dots, n-1),$$

alors on a pour tout $n \in \mathbb{N}$,

$$\begin{aligned}&\mathbb{P}(B_n = 1,\ X_n^y(U_n) = X_n^x(U_n) \,|\, \mathcal{F}_n) + \mathbb{P}(B_n = 0,\ X_n^y(U_n) \neq X_n^x(U_n) \,|\, \mathcal{F}_n) \\ &= \mathbb{P}(B_n = 1 \,|\, \mathcal{F}_n)\mathbb{P}(X_n^y(U_n) = X_n^x(U_n) \,|\, \mathcal{F}_n) + \mathbb{P}(B_n = 0 \,|\, \mathcal{F}_n)\mathbb{P}(X_n^y(U_n) \neq X_n^x(U_n) \,|\, \mathcal{F}_n) \\ &= \frac{1}{2}\Big(\mathbb{P}(X_n^y(U_n) = X_n^x(U_n) \,|\, \mathcal{F}_n) + \mathbb{P}(X_n^y(U_n) \neq X_n^x(U_n) \,|\, \mathcal{F}_n)\Big) \\ &= \frac{1}{2}.\end{aligned}$$

Par conséquent, pour tout $n \in \mathbb{N}$ et $j \in \{\pm\mathrm{e}_1, \dots, \pm\mathrm{e}_d\}$,

$$\begin{aligned}&\mathbb{P}(X_{n+1}^y - X_n^y = j \,|\, \mathcal{F}_n) \\ &= \mathbb{P}(J_n \mathrm{e}_{U_n} = j \,|\, \mathcal{F}_n)\Big(\mathbb{P}(B_n = 1,\ X_n^y(U_n) = X_n^x(U_n) \,|\, \mathcal{F}_n) + \mathbb{P}(B_n = 0,\ X_n^y(U_n) \neq X_n^x(U_n) \,|\, \mathcal{F}_n)\Big) \\ &= \frac{1}{2}\mathbb{P}(J_n \mathrm{e}_{U_n} = j) = \frac{1}{2} \times \frac{1}{2d}.\end{aligned}$$

Cela donne aussi

$$\mathbb{P}(X_{n+1}^y - X_n^y = 0 \,|\, \mathcal{F}_n) = 1 - \sum_j \mathbb{P}(X_{n+1}^y - X_n^y = j \,|\, \mathcal{F}_n) = \frac{1}{2}.$$

Puisque $\sigma(X_0^y, \dots, X_n^y) \subset \mathcal{F}_n$, cela montre que $(X_n^y)_{n\geq 0}$ est bien une version *lazy* (avec $p = 1/2$) de la MAS sur $\mathbb{Z}^d$.

Etudions à présent le processus $(D_n)_{n\geq 0} = (X_n^x - X_n^y)_{n\geq 0}$. Notons d'abord que par construction,

$$X_{n+1}^x - X_{n+1}^y = X_n^x - X_n^y + J_n \mathrm{e}_{U_n} \mathbb{1}_{(X_n^y(U_n)\neq X_n^x(U_n))},$$

et donc

$$D_{n+1} = D_n + J_n \mathrm{e}_{U_n} \mathbb{1}_{(D_n(U_n)\neq 0)}.$$

En particulier, pour tout $k \in \{1, \ldots, d\}$, $D_{n+1}(k) - D_n(k) \in \{-1, 0, 1\}$. De plus, pour $i \in \{-1, 1\}$, on a

$$\mathbb{P}(D_{n+1}(k) - D_n(k) = i \,|\, \mathcal{F}_n) = \mathbb{P}(U_n = k)\mathbb{P}(J_n = i)\mathbb{1}_{(D_n(k)\neq 0)} = \frac{1}{2d}\mathbb{1}_{(D_n(k)\neq 0)}.$$

Le processus $(D_n(k))_{n\geq 0}$ est alors une version *lazy* (avec $p = 1/d$) de la MAS sur $\mathbb{Z}$, arrêtée quand elle atteint 0. On sait que la MAS sur $\mathbb{Z}$ est récurrente, montrons que sa version *lazy* l'est aussi. Ceci est en fait vrai en généralité, car, avec la notation de l'Exercice 3, on a

$$\begin{aligned}
\sum_n Q_p^n &= \sum_n (pQ + (1-p)\mathrm{Id})^n \\
&= \sum_n \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} Q^k \\
&= \sum_k Q^k \sum_{n=k}^\infty \binom{n}{k} p^k (1-p)^{n-k} \\
&= \frac{1}{p} \sum_k Q^k,
\end{aligned}$$

par une égalité du cours.

On déduit que presque sûrement, pour tout $k \in \{1, \ldots, d\}$, $D_n(k)$ atteint 0 en un temps fini T_k et y reste. Par conséquent, on a $X_n^x = X_n^y$ pour $n \geq \max_{k=1,\ldots,n} T_k$, et donc notre couplage est bien un couplage comme dans la partie 2 de l'exercice.

Par la partie 2 de l'exercice, on en déduit que la version *lazy* ($p = 1/2$) de la MAS sur $\mathbb{Z}^d$ est Liouville. Mais par la partie 3, cela implique que la MAS sur $\mathbb{Z}^d$ est également Liouville.

Exercices supplémentaires

Exercice 97 (Problème du collectionneur de coupons). 1. Soit $k \in \{1, \ldots, n\}$. On observe que $X_{n,k}$ représente le temps qu'il faut attendre pour tirer un coupon différent des $k-1$ collectés précédemment. Ainsi, $X_{n,k}$ suit une loi géométrique de paramètre $p_{n,k} = 1 - (k-1)/n$, c'est-à-dire que

$$\forall r \in \mathbb{N}^* \qquad \mathbb{P}(X_{n,k} = r) = (1 - p_{n,k})^{r-1} p_{n,k}.$$

En particulier, on a alors $\mathbb{E}[X_{n,k}] = 1/p_{n,k}$ et $\mathrm{Var}(X_{n,k}) = (1 - p_{n,k})/p_{n,k}^2$. De plus, à n fixé, les variables aléatoires $X_{n,k}$ sont indépendantes. Sachant que $T_n = X_{n,1} + \ldots + X_{n,n}$, on en déduit que

$$\mathbb{E}[T_n] = \sum_{k=1}^n \mathbb{E}[X_{n,k}] = \sum_{k=1}^n \frac{1}{p_{n,k}} = \sum_{k=1}^n \frac{n}{n-k+1} = n \sum_{\ell=1}^n \frac{1}{\ell} \sim n \ln n,$$

pour $n \to \infty$. De même, en utilisant en outre l'indépendance des $X_{n,k}$, on a

$$\mathrm{Var}(T_n) = \sum_{k=1}^n \mathrm{Var}(X_{n,k}) \leq \sum_{k=1}^n \frac{1}{p_{n,k}^2} = \sum_{k=1}^n \left(\frac{n}{n-k+1}\right)^2 = n^2 \sum_{\ell=1}^n \frac{1}{\ell^2} = \mathrm{O}(n^2).$$

2. Soit $\varepsilon > 0$. D'après la question précédente, il existe un entier $n_0 \geq 1$ tel que pour tout entier $n \geq n_0$,

$$\left|\frac{\mathbb{E}[T_n]}{n \ln n} - 1\right| \leq \frac{\varepsilon}{2}.$$

On en déduit que

$$\left|\frac{T_n}{n\ln n} - 1\right| > \varepsilon \qquad \Longrightarrow \qquad \left|\frac{T_n}{n\ln n} - \frac{\mathbb{E}[T_n]}{n\ln n}\right| \geq \left|\frac{T_n}{n\ln n} - 1\right| - \left|\frac{\mathbb{E}[T_n]}{n\ln n} - 1\right| > \frac{\varepsilon}{2}.$$

L'inégalité de Chebyshev conduit donc à

$$\mathbb{P}\left(\left|\frac{T_n}{n\ln n} - 1\right| > \varepsilon\right) \leq \mathbb{P}\left(|T_n - \mathbb{E}[T_n]| > \frac{\varepsilon}{2} n\ln n\right) \leq \frac{4\mathrm{Var}(T_n)}{(\varepsilon n \ln n)^2} = \mathrm{O}\left(\frac{1}{(\ln n)^2}\right).$$

On en déduit finalement que

$$\frac{T_n}{n\ln n} \xrightarrow[n\to\infty]{\mathbb{P}} 1.$$

Exercice 98 (Polynômes de Bernstein)**.** 1. Soit $S_n(x) \sim \mathrm{Bin}(n,x)$. Alors $\mathbb{P}(S_n(x) = k) = \binom{n}{k} x^k (1-x)^{n-k}$ pour tout $k \in \mathbb{N}$. Ceci donne

$$\mathbb{E}[f(S_n(x)/n)] = \sum_{k\in\mathbb{N}} f(k/n)\mathbb{P}(S_n(x) = k) == \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k(1-x)^{n-k} = B_n(x).$$

2. On a pour tout $\varepsilon > 0$, $x \in [0,1]$ et $n \in \mathbb{N}^*$,

$$\begin{aligned}|B_n(x) - f(x)| &= |\mathbb{E}[f(S_n(x)/n) - f(x)]| \\ &\leq \mathbb{E}[|f(S_n(x)/n) - f(x)|] \\ &\leq \mathbb{E}[|f(S_n(x)/n) - f(x)|\mathbb{1}_{|S_n(x)/n - x|\leq\varepsilon}] + \mathbb{E}[|f(S_n(x)/n) - f(x)|\mathbb{1}_{|S_n(x)/n-x|>\varepsilon}] \\ &=: T_1 + T_2.\end{aligned}$$

Fixons $\delta > 0$. Puisque f est continue et $[0,1]$ est compact, f est uniformément continue. Il existe alors $\varepsilon > 0$ tel que $|f(x) - f(y)| < \delta$ pour tout $|x - y| < \varepsilon$. En particulier, $T_1 < \mathbb{E}[\delta] = \delta$. De plus, f est bornée. L'inégalité de Chebychev donne alors

$$\begin{aligned}T_2 \leq \|f\|_\infty \mathbb{E}[\mathbb{1}_{|S_n(x)/n-x|>\varepsilon}] &= \|f\|_\infty \mathbb{P}(|S_n(x)/n - x| > \varepsilon) \\ &\leq \|f\|_\infty \varepsilon^{-1}\,\mathrm{Var}(S_n(x)/n) \\ &= \|f\|_\infty \varepsilon^{-1} x(1-x)/n \leq \|f\|_\infty \varepsilon^{-1}/(4n).\end{aligned}$$

Par ce qui précède, on a pour $n \geq (4\delta\varepsilon\|f\|_\infty)^{-1}$ et tout $x \in [0,1]$,

$$|B_n(x) - f(x)| \leq 2\delta.$$

Ceci montre que $B_n(x)$ tend uniformément vers $f(x)$ quand $n \to \infty$.

Exercice 99 (Variables aléatoires symétriques et fonctions caractéristiques)**.** 1. On a pour toute variable aléatoire réelle X, $\varphi_{-X}(t) = \varphi_X(-t) = \varphi_X(t)^*$, où z^* est le conjugué complexe de $z \in \mathbb{C}$. Par conséquent,

$$X \overset{\text{loi}}{=} -X \iff \varphi_{-X}(t) = \varphi_X(t)\,\forall t\in\mathbb{R} \iff \varphi_X(t) = \varphi_X(t)^*\,\forall t \in \mathbb{R} \iff \varphi_X(t) \in \mathbb{R}\,\forall t\in\mathbb{R}.$$

2. La variable Z est symétrique et indépendante de X. Donc le couple $(-Z, Y)$ a même loi que (Z, Y). Donc $-ZY$ et ZY ont même loi. De plus, on a pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ mesurable,

$$\mathbb{E}[f(X)] = \mathbb{E}[f(ZY)] = \frac{1}{2}(\mathbb{E}[f(Y)] + \mathbb{E}[f(-Y)]).$$

Par conséquent, on a $\mu_X = \frac{1}{2}(\mu_Y + \mu_{-Y})$ et en particulier, $\varphi_X = \frac{1}{2}(\varphi_Y + \varphi_{-Y}) = \frac{1}{2}(\varphi_Y + \varphi_Y^*) = \Re\varphi_Y$.

3. On a $\frac{1}{2}e^{-|x|} = \frac{1}{2}(e^{-x}\mathbf{1}_{x\geq 0} + e^x\mathbf{1}_{x<0})$. La fonction caractéristique de la loi exponentielle de paramètre 1 est $t \mapsto (1-it)^{-1}$. Par la dernière partie de l'exercice, on a

$$\varphi_X(t) = \Re\frac{1}{1-it} = \Re\frac{1+it}{|1-it|^2} = \frac{1}{1+t^2}.$$

4. Soit X comme dans la dernière partie. Comme φ_X est intégrable sur $\mathbb{R}$, on a par la formule d'inversion de Fourier,
$$\frac{1}{2}e^{-|x|} = \frac{1}{2\pi}\int_{\mathbb{R}} e^{-itx}\varphi_X(t)\,dt = \frac{1}{2}\int_{\mathbb{R}} e^{-itx}\frac{1}{\pi(1+t^2)}\,dt = \frac{1}{2}\int_{\mathbb{R}} e^{itx}\frac{1}{\pi(1+t^2)}\,dt.$$
Par conséquent, la fonction caractéristique de la loi de Cauchy de paramètre 1 est la fonction $t \mapsto e^{-|t|}$. Si X_1, X_2 sont de loi de Cauchy de paramètre 1, alors
$$\varphi_{\frac{X_1+X_2}{2}}(t) = \varphi_{X_1}(t/2)^2 = e^{-2|t/2|} = e^{-|t|} = \varphi_{X_1}(t).$$
Par conséquent, la loi de $\frac{X_1+X_2}{2}$ est celle de X_1.
5. On a :
$$\begin{aligned}
\varphi_{X_1X_2}(t) &= \int\left(\int e^{-itxy}e^{-\frac{y^2}{2}}\frac{dy}{\sqrt{2\pi}}\right)e^{-\frac{x^2}{2}}\frac{dx}{\sqrt{2\pi}}\\
&= \int \varphi_{X_2}(tx)e^{-\frac{x^2}{2}}\frac{dx}{\sqrt{2\pi}}\\
&= \int e^{-(t^2+1)\frac{x^2}{2}}\frac{dx}{\sqrt{2\pi}}\\
&= \frac{1}{\sqrt{1+t^2}}\int e^{-\frac{y^2}{2}}\frac{dy}{\sqrt{2\pi}} \qquad \text{(chgt. de var. } x \mapsto y/\sqrt{1+t^2})\\
&= \frac{1}{\sqrt{1+t^2}}.
\end{aligned}$$
Puis $\varphi_{X_1X_2+X_3X_4}(t) = \frac{1}{1+t^2}$ et la question de la partie 3 montre que $X_1X_2 + X_3X_4$ suit la loi de densité $\frac{1}{2}e^{-|x|}$ sur $\mathbb{R}$.

Exercice 100 (Loi bêta)**.** Puisque βG_1 et βG_2 suivent les lois $\Gamma(\alpha_1, 1)$ et $\Gamma(\alpha_2, 1)$, il suffit de considérer $\beta = 1$. On a alors pour une fonction f mesurable bornée,
$$\begin{aligned}
&\mathbb{E}[f(G_1/(G_1+G_2))]\\
&= \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_0^\infty\int_0^\infty f(\tfrac{x}{x+y})x^{\alpha_1-1}y^{\alpha_2-1}e^{-(x+y)}\,dxdy\\
&= \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_0^\infty\left(\int_0^\infty f(\tfrac{1}{1+z})x^{\alpha_1+\alpha_2-1}z^{\alpha_2-1}e^{-x(1+z)}\,dz\right)dx && y \to xz\\
&= \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_0^\infty f(\tfrac{1}{1+z})z^{\alpha_2-1}dz\int_0^\infty x^{\alpha_1+\alpha_2-1}e^{-x(1+z)}\,dx && \text{Fubini}\\
&= \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_0^\infty f(\tfrac{1}{1+z})z^{\alpha_2-1}(1+z)^{-\alpha_1-\alpha_2}\,dz\int_0^\infty w^{\alpha_1+\alpha_2-1}e^{-w}\,dw && x \to w/(1+z)\\
&= \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_0^\infty f(\tfrac{1}{1+z})z^{\alpha_2-1}(1+z)^{-\alpha_1-\alpha_2}\,dz\\
&= \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_0^1 f(t)(\tfrac{1-t}{t})^{\alpha_2-1}t^{\alpha_1+\alpha_2}t^{-2}\,dt && \tfrac{1}{1+z} \to t\\
&= \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_0^1 f(t)t^{\alpha_1-1}(1-t)^{\alpha_2-1}\,dt.
\end{aligned}$$
La v.a. $G_1/(G_1+G_2)$ suit alors la loi de densité sur $(0, 1)$,
$$\frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}t^{\alpha_1-1}(1-t)^{\alpha_2-1}.$$

Dans le cas $\alpha_1 = \alpha_2 = 1$ (variables exponentielles), on obtient la loi uniforme sur $(0, 1)$. Dans le cas $\alpha_1 = \alpha_2 = 1/2$ (loi de $N_1^2/(N_1^2 + N_2^2)$ pour N_1, N_2 des gaussiennes centrées indépendantes), on obtient la loi de densité $1/(\pi\sqrt{t(1-t)})$. Celle-ci s'appelle la loi de l'arc sinus, car sa fonction de répartition est
$$F(x) = \int_0^x \frac{1}{\pi\sqrt{t(1-t)}}\,dt = \frac{1}{\pi}\arcsin(2t-1) + \frac{1}{2} = \frac{2}{\pi}\arcsin(\sqrt{x}).$$

Exercice 101. 1. On a pour tout $m > n$,

$$\mathbb{E}[|Y_m - Y_n|] = \mathbb{E}\left[\sum_{k=n+1}^{m} X_1 \cdots X_k\right] = \sum_{k=n+1}^{m} \mathbb{E}[X_1 \cdots X_k] = \sum_{k=n+1}^{m} 2^{-k} \le \sum_{k=n+1}^{\infty} 2^{-k} = 2^{-n}.$$

Par conséquent, $(Y_n)_{n\ge 0}$ est une suite de Cauchy dans L^1 qui est un espace de Banach, et donc Y_n converge dans L^1 vers une limite Y.

2. On décompose

$$Y_n = \sum_{k=1}^{n} X_1 \cdots X_k = X_1 \left(1 + \sum_{k=2}^{n} X_2 \cdots X_k\right).$$

Par le même argument que dans la dernière partie, la suite $Y_n' = \sum_{k=2}^{n} X_2 \cdots X_k$ converge dans L^1 vers une limite Y', si bien que

$$Y = X_1(1 + Y').$$

Puisque Y_n' est égale en loi à Y_{n-1}, les limites Y' et Y sont également égales en loi. De plus, puisque Y_n' est indépendante de X_1 pour tout n, la limite Y' est également indépendante de X_1 (*exo : pourquoi ?*). Il vient que $Y \overset{\text{loi}}{=} X(1+Y)$ avec $X \sim \mathrm{Unif}(0,1)$ indépendante de Y.

3. Par la dernière partie, on a pour tout $t \in \mathbb{R}$,

$$\begin{aligned}
\varphi(t) &= \mathbb{E}[e^{itX(1+Y)}] \\
&= \int_{[0,1]\times\mathbb{R}} e^{itx(1+y)}\, dx \otimes \mu_Y(dy) \\
&= \int_0^1 e^{itx} \left(\int_{\mathbb{R}} e^{itxy} \mu_Y(dy)\right) dx && \text{par Fubini} \\
&= \int_0^1 e^{itx} \varphi(tx)\, dx.
\end{aligned}$$

4. On a pour $t > 0$, par un changement de variables $x \mapsto x/t$,

$$\varphi(t) = \int_0^1 e^{itx} \varphi(tx)\, dx = \frac{1}{t} \int_0^t e^{ix} \varphi(x)\, dx,$$

si bien que pour tout $t \ge 0$,

$$t\varphi(t) = \int_0^t e^{ix} \varphi(x)\, dx$$

(c'est vrai pour $t = 0$ car $0\varphi(0) = 0$). Il vient que la fonction $f(t) = t\varphi(t)$ satisfait à l'équation différentielle

$$f'(t) = \frac{e^{it}}{t} f(t),$$

qui admet comme solution générale

$$f(t) = C \exp\left(\int_1^t e^{is}/s\, ds\right).$$

Puisque $1/t = \exp(-\ln t) = \exp(-\int_1^t 1/s\, ds)$, il vient que

$$\varphi(t) = C \exp\left(\int_1^t (e^{is} - 1)/s\, ds\right).$$

Avec la condition $\varphi(0) = 1$, cela donne

$$\varphi(t) = \exp\left(\int_0^t (e^{is} - 1)/s\, ds\right).$$

Puisque $\varphi(t) = \varphi(-t)^*$, cela donne l'expression suivante pour tout $t \in \mathbb{R}$ (avec $\int_0^t := -\int_t^0$ pour $t < 0$) :

$$\varphi(t) = \exp\left(\int_0^t (e^{is} - 1)/|s|\, ds\right).$$

Exercice 102 (Théorème de Glivenko-Cantelli)**.** 1. On rappelle que

$$F_n(x) = \frac{1}{n}\mathrm{Card}\{j \in \{1,\dots,n\} \mid X_j \le x\}. \tag{5}$$

En particulier, $F_n(x)$ prend ses valeurs dans $\{0, 1/n, 2/n, \dots, (n-1)/n, 1\}$. Donc, pour tout $k \in \{0,\dots,n\}$,

$$\left\{F_n(x) = \frac{k}{n}\right\} = \bigcup_{\substack{J\subseteq\{1,\dots,n\}\\ \mathrm{Card}\,J=k}} \left(\bigcap_{j\in J}\{X_j \le x\} \cap \bigcap_{j\in\{1,\dots,n\}\setminus J}\{X_j > x\}\right).$$

Comme les $X_1,\dots,X_n$ sont des variables aléatoires, cela assure que $F_n(x)$ est aussi une variable aléatoire.

2. Notons G_n la fonction de répartition de Γ_n. Pour tout $x \in \mathbb{R}$,

$$G_n(x) = \Gamma_n((-\infty, x]) = \frac{1}{n}\sum_{j=1}^n \mathbf{1}_{X_j\le x} = \frac{1}{n}\mathrm{Card}\{j \in \{1,\dots,n\} \mid X_j \le x\} = F_n(x).$$

3. Dans le cas où $F(x) = x$ pour tout $x \in [0,1]$, les variables X_n suivent la loi uniforme sur l'intervalle $[0,1]$. De plus, à $x \in [0,1]$ fixé, en utilisant (5) et la loi forte des grands nombres, on obtient

$$\text{p.s.} \qquad F_n(x) = \frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{X_j\le x\}} \xrightarrow[n\to\infty]{} \mathbb{E}[\mathbb{1}_{\{X_1\le x\}}] = F(x).$$

En d'autres termes,

$$\forall x \in [0,1] \quad \text{p.s.} \qquad \lim_{n\to\infty} F_n(x) = F(x). \tag{6}$$

4. Il s'agit d'échanger le “pour tout $x \in [0,1]$” et le “presque sûrement” dans (6), ce qui ne peut se faire directement car x parcourt un ensemble infini non dénombrable. Cependant, on peut déduire de (6) que l'événement

$$A = \left\{\forall x \in [0,1]\cap\mathbb{Q} \quad \lim_{n\to\infty} F_n(x) = F(x)\right\}$$

est de probabilité un, car on se restreint alors à prendre x dans un ensemble dénombrable. Plaçons-nous sur l'événement A et prenons un réel $x \in [0,1]$. Par densité de $\mathbb{Q}$ dans $\mathbb{R}$, il existe une suite croissante $(a_p)_{p\in\mathbb{N}}$ de rationnels et une suite décroissante $(b_p)_{p\in\mathbb{N}}$ de rationnels qui convergent toutes deux vers x. Comme la fonction F_n est croissante, on a $F_n(a_p) \le F_n(x) \le F_n(b_p)$ pour tous entiers n et p. On obtient, en faisant tendre n vers l'infini,

$$F(a_p) \le \liminf_{n\to\infty} F_n(x) \le \limsup_{n\to\infty} F_n(x) \le F(b_p).$$

La fonction F étant continue sur $\mathbb{R}$, le membre de gauche et le membre de droite tend tous deux vers $F(x)$ quand $p \to \infty$. On en déduit que sur l'événement presque sûr A,

$$\forall x \in [0,1] \qquad \lim_{n\to\infty} F_n(x) = F(x).$$

5. Sur un événement de probabilité un (l'événement A par exemple), la suite de fonctions $(F_n)_{n\in\mathbb{N}^*}$ converge simplement vers F sur l'intervalle $[0,1]$. Comme F est continue et comme toutes les fonctions F_n sont croissantes, on peut appliquer le théorème de Dini pour conclure que la convergence est uniforme sur $[0,1]$. Elle est en fait uniforme sur $\mathbb{R}$, puisque $F_n(x) = F(x)$ pour tout réel $x \notin [0,1]$ et tout entier $n \ge 1$.

6. On a clairement pour tous réels $x \in \mathbb{R}$ et $u \in [0,1]$,

$$F^{-1}(u) \le x \qquad \Longleftrightarrow \qquad u \le F(x).$$

Donc, si U suit la loi uniforme sur $[0,1]$, on a pour tout réel $x \in \mathbb{R}$,

$$\{F^{-1}(U) \le x\} = \{U \le F(x)\} \qquad \text{et} \qquad \mathbb{P}(F^{-1}(U) \le x) = \mathbb{P}(U \le F(x)) = F(x),$$

ce qui prouve que $F^{-1}(U)$ est une variable aléatoire de fonction de répartition égale à F.

7. La question précédente assure que si $U_1, \dots, U_n$ désignent des variables aléatoires indépendantes et de loi uniforme sur $[0,1]$, alors $(F^{-1}(U_1), \dots, F^{-1}(U_n))$ a même loi que $(X_1, \dots, X_n)$. Donc, en désignant par $\stackrel{\mathcal{L}}{=}$ l'égalité en loi, on a

$$\begin{aligned}
V_n &= \sup_{x\in\mathbb{R}} \left| \frac{1}{n} \sum_{j=1}^n \mathbb{1}_{\{X_j \le x\}} - F(x) \right| \\
&\stackrel{\mathcal{L}}{=} \sup_{x\in\mathbb{R}} \left| \frac{1}{n} \sum_{j=1}^n \mathbb{1}_{\{F^{-1}(U_j) \le x\}} - F(x) \right| = \sup_{x\in\mathbb{R}} \left| \frac{1}{n} \sum_{j=1}^n \mathbb{1}_{\{U_j \le F(x)\}} - F(x) \right| = \sup_{x\in\mathbb{R}} |F_n^{\text{unif}}(F(x)) - F(x)|.
\end{aligned}$$

Pour conclure, il suffit d'observer que cette dernière quantité est égale à

$$\sup_{t\in F(\mathbb{R})} |F_n^{\text{unif}}(t) - t| \le \sup_{t\in[0,1]} |F_n^{\text{unif}}(t) - t|,$$

ce qui tend presque sûrement vers zéro quand $n \to \infty$, d'après le résultat de la question 6.

Exercice 103. 1. On note $Z_n = X_n + iY_n$. Alors $X_0 = x$, $Y_0 = y$ et

$$X_{n+1} + iY_{n+1} = X_n + iY_n + Y_n(\cos(U_{n+1}) + i\sin(U_{n+1})) = X_n + Y_n(\cos(U_{n+1}) + iY_n(1 + \sin(U_{n+1})).$$

Ceci montre que $(X_n, Y_n)_{n\ge0}$ satisfait au système donné dans l'énoncé.

2. Pour montrons d'abord que Z_n est intégrable pour tout $n \in \mathbb{N}$: Puisque $|\operatorname{Im}(Z_n)e^{iU_{n+1}}| \le |Z_n|$, on a pour tout $n \in \mathbb{N}$,

$$\mathbb{E}[|Z_{n+1}|] \le \mathbb{E}[|Z_n|] + \mathbb{E}[|Z_n|] = 2\mathbb{E}[|Z_n|].$$

Par récurrence, Z_n est intégrable pour tout $n \in \mathbb{N}$, et donc X_n et Y_n également.

Montrons que X_n et Y_n sont des martingales. Soit $U \sim \text{Unif}(0, 2\pi)$. On remarque tout d'abord que si $U \sim \text{Unif}(0, 2\pi)$, alors $\mathbb{E}[\cos(U)] = \mathbb{E}[\sin(U)] = 0$ pour tout $n \in \mathbb{N}$, car

$$\frac{1}{2\pi}\int_0^{2\pi} \cos(x)\,dx = \frac{1}{2\pi}(\sin(2\pi) - \sin(0)) = 0,$$

et pareil pour sin. Par conséquent, pour tout $n \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{E}[X_{n+1} - X_n \,|\, \mathcal{F}_n] &= \mathbb{E}[Y_n \cos(U_{n+1}) \,|\, \mathcal{F}_n] = Y_n\mathbb{E}[\cos(U_{n+1})] = 0 \\
\mathbb{E}[Y_{n+1} - Y_n \,|\, \mathcal{F}_n] &= \mathbb{E}[Y_n \sin(U_{n+1}) \,||, \mathcal{F}_n] = Y_n\mathbb{E}[\sin(U_{n+1})] = 0.
\end{aligned}$$

Il reste à montrer que $\operatorname{Im}(Z_n) > 0$ p.s. pour tout $n \in \mathbb{N}$. Or, $\operatorname{Im}(Z_n) = Y_n = y\prod_{i=1}^n(1 + \sin(U_i))$ et $1 + \sin(U_i) > 0$ p.s. pour tout $i \ge 1$ et $y > 0$ par hypothèse.

3. Pour tout $n \in \mathbb{N}$, $Y_n = y\prod_{i=1}^n(1 + \sin(U_i)) \ge 0$, si bien que $(Y_n)_{n\ge0}$ est une martingale positive. Elle converge donc p.s. vers une limite $Y_\infty \ge 0$. Supposons que $\mathbb{P}(Y_\infty > 0) > 0$. Sur l'événement $\{Y_\infty > 0\}$, $Y_{n+1}/Y_n \to 1$ p.s., et donc $\mathbb{P}(Y_{n+1}/Y_n \to 1) > 0$. Or, $Y_{n+1}/Y_n = 1 + \sin(U_{n+1})$ et $\mathbb{P}(1 + \sin(U_{n+1}) \to 1) = 0$, car les v.a. $\sin(U_n)$, $n = 1, 2, \dots$ sont iid et de loi continue. Ceci montre que $\mathbb{P}(Y_\infty > 0) = 0$ et donc $Y_\infty = 0$ p.s.

La martingale n'est pas fermée car elle ne converge pas dans L^1 : $\mathbb{E}[Y_0] = y \ne 0 = \mathbb{E}[Y_\infty]$.

4. Par indépendance de U_{n+1} et $\mathcal{F}_n$, $\mathbb{E}[\sqrt{|X_{n+1} - X_n|} \,|\, \mathcal{F}_n] = \sqrt{Y_n}\mathbb{E}[\sqrt{\cos(U_{n+1})}]$, et donc

$$\mathbb{E}[\sqrt{|X_{n+1} - X_n|}] = \mathbb{E}[\sqrt{|\cos(U_{n+1})|}]\mathbb{E}[\sqrt{Y_n}] \le y\mathbb{E}[\sqrt{1 + \sin(U)}]^n,$$

Puisque $1 + \sin(U)$ n'est pas une constante et la fonction $x \mapsto \sqrt{x}$ est strictement concave, l'inégalité de Jensen donne,

$$r := \mathbb{E}[\sqrt{1 + \sin(U)}] < \sqrt{\mathbb{E}[1 + \sin(U)]} = 1.$$

Ceci donne $\mathbb{E}[\sqrt{|X_{n+1} - X_n|}] \le yr^n$.

5. Par l'inégalité de Markov et la dernière partie,
$$\sum_n \mathbb{P}(|X_{n+1}-X_n| \geq a^n) = \sum_n \mathbb{P}(\sqrt{|X_{n+1}-X_n|} \geq a^{n/2}) \leq \sum_n yr^n/a^{n/2} = y\sum_n y(r^2/a)^{n/2}.$$
Par hypothèse, cette somme est finie. Le lemme de Borel–Cantelli nous donne alors $\mathbb{P}(\limsup\{|X_{n+1} - X_n| \geq a^n\}) = 0$. Puisque $\sum_n a^n < \infty$, ceci montre que p.s., $(X_n)_{n\geq 0}$ est une suite de Cauchy dans $\mathbb{R}$ et donc converge p.s. vers une variable X_∞.
6. Puisque $Y_n \geq 0$ pour tout n, $|M_n(\lambda)| = e^{-|\lambda|Y_n} \leq 1$ pour tout n. En particulier, $(M_n(\lambda))_{n\geq 0}$ est un processus borné. Montrons que c'est une martingale :
$$\begin{aligned}\mathbb{E}[M_{n+1}(\lambda)\,|\,\mathcal{F}_n] &= \mathbb{E}[\exp(i\lambda(X_n + Y_n\cos(U_{n+1})) - |\lambda|Y_n(1+\sin(U_{n+1})))\,|\,\mathcal{F}_n] \\ &= M_n(\lambda)\mathbb{E}[\exp(Y_n(i\lambda\cos(U_{n+1}) - |\lambda|\sin(U_{n+1})))\,|\,\mathcal{F}_n].\end{aligned}$$
Il suffit de montrer que $\varphi(y,\lambda) = \mathbb{E}[\exp(y(i\lambda\cos(U) - |\lambda|\sin(U)))] = 1$ pour tout $y \geq 0$, $\lambda \in \mathbb{R}$. Pour $\lambda \geq 0$, on a
$$\varphi(y,\lambda) = \mathbb{E}[\exp(iy\lambda(\cos(U) + i\sin(U)))] = \mathbb{E}[\exp(-iy\lambda e^{iU})] = 1,$$
Pour $\lambda < 0$, on a
$$\varphi(y,\lambda) = \mathbb{E}[\exp(iy\lambda(\cos(U) - i\sin(U)))] = \mathbb{E}[\exp(iy\lambda e^{-iU})] = \mathbb{E}[\exp(iy\lambda e^{iU})] = 1,$$
où l'avant-dernière égalité provient du fait que $e^{-iU} = e^{i(2\pi-U)} \overset{\text{loi}}{=} e^{iU}$, car $2\pi - U \overset{\text{loi}}{=} U$. Ceci montre que $M_n(\lambda)$ est une martingale bornée et converge donc p.s. et dans L^1. Puisque $Y_n \to 0$ p.s., la limite est égale à
$$\lim_n M_n(\lambda) = e^{i\lambda X_\infty}.$$
Ceci donne
$$\mathbb{E}[e^{i\lambda X_\infty}] = \mathbb{E}[M_0(\lambda)] = e^{i\lambda x - |\lambda| y}.$$
Par conséquent, $X_\infty \overset{\text{loi}}{=} x + yC$, où C suit la loi de Cauchy standard.
7. Non, dans ce cas la martingale $(X_n)_{n\geq 0}$ aurait convergé dans L^1, mais sa limite X_∞ n'est pas intégrable (il est bien connu que la loi de Cauchy ne l'est pas).

Exercice 104 (Algorithme de Metropolis). 1. Soit $x, y \in E$ distincts. Quitte à échanger le rôle de x et y, on peut supposer que $\pi(y) \leq \pi(x)$. Sachant que P est symétrique, on a alors
$$\frac{\pi(x)}{\pi(y)}Q(x,y) = \frac{Q(x,y)}{\alpha(x,y)} = P(x,y) = P(y,x) = \alpha(y,x)P(y,x) = Q(y,x),$$
de sorte que $\pi(x)Q(x,y) = \pi(y)Q(y,x)$. On en déduit que la loi π est réversible pour Q et donc invariante.

2. On observe que pour tous $x, y \in E$ distincts, $P(x,y) > 0$ si et seulement si $Q(x,y) > 0$. Une simple récurrence sur $n \geq 1$ permet de montrer que pour tous $x, y \in E$ distincts, $P^n(x,y) > 0$ si et seulement si $Q^n(x,y) > 0$. Ainsi, comme P est irréductible, pour tous $x, y \in E$ distincts, il existe un entier $n \geq 1$ pour lequel on a $P^n(x,y) > 0$. Cela implique que $Q^n(x,y) > 0$ pour ce même entier, si bien que Q est aussi irréductible.

3. (a) Par l'absurde, supposons que pour tout $x_0 \in M$ et tout $y \in E \setminus M$, on a $P(x_0,y) = 0$. On en déduit que M et $E \setminus M$ ne communiquent pas, ce qui contredit l'irréductibilité de P. Ainsi, il existe un état $x_0 \in M$ tel que $P(x_0,y) > 0$ pour un certain $y \in E \setminus M$. On observe que $\pi(x_0) > \pi(y)$, donc
$$Q(x_0,y) = \alpha(x_0,y)P(x_0,y) = \frac{\pi(y)}{\pi(x_0)}P(x_0,y) < P(x_0,y).$$
On a $Q(x_0,z) \leq P(x_0,z)$ pour tout $z \in E \setminus \{x_0\}$, avec une inégalité stricte si $z = y$. D'où
$$\sum_{z\in E\setminus\{x_0\}} Q(x_0,z) < \sum_{z\in E\setminus\{x_0\}} P(x_0,z),$$
ce qui donne $Q(x_0,x_0) > P(x_0,x_0) \geq 0$.

(b) D'après la question précédente, $Q(x_0, x_0)$ est strictement positif, donc x_0 est de période 1. Comme Q est irréductible, on en déduit que Q est apériodique.

(c) La matrice Q est irréductible et récurrente positive (puisqu'il existe une probabilité invariante). Puisque Q est apériodique, pour tout $x \in E$, $\sum_{y\in E} |\mathbb{P}_x(X_n = y) - \pi(y)|$ converge vers zéro, c'est-à-dire que la loi de X_n converge vers π.

4. Soit $X_1, \dots, X_n$ simulés par l'algorithme. Montrons que pour tout $x, y \in E$, $\mathbb{P}(X_{i+1} = y|X_i = x) = Q(x, y)$.

$$\begin{aligned}
\mathbb{P}(X_{i+1} = y|X_i = x) &= \mathbb{P}(X_{i+1} = y, U \le \alpha(X_i, V)|X_i = x) + \mathbb{P}(X_{i+1} = y, U > \alpha(X_i, V)|X_i = x)\\
&= \mathbb{P}(V = y, U \le \alpha(x, V)|X_i = x) + \mathbb{P}(X_i = y, U > \alpha(x, V)|X_i = x)\\
&= \mathbb{P}(V = y, U \le \alpha(x, y)|X_i = x) + \mathbb{1}_{\{x=y\}}\mathbb{P}(X_i = x, U > \alpha(x, V)|X_i = x)\\
&= \mathbb{P}(V = y|X_i = x)\mathbb{P}(U \le \alpha(x, y)) + \mathbb{1}_{\{x=y\}}\mathbb{P}(U > \alpha(x, V)|X_i = x)
\end{aligned}$$

car U est indépendante de V et X_i. Ainsi,

$$\begin{aligned}
\mathbb{P}(X_{i+1} = y|X_i = x) &= P(x, y)\alpha(x, y) + \mathbb{1}_{\{x=y\}} \sum_{z\in E} \mathbb{P}(V = z, U > \alpha(x, z)|X_i = x)\\
&= P(x, y)\alpha(x, y) + \mathbb{1}_{\{x=y\}} \sum_{z\in E} \mathbb{P}(V = z|X_i = x)\mathbb{P}(U > \alpha(x, z))\\
&= P(x, y)\alpha(x, y) + \mathbb{1}_{\{x=y\}} \sum_{z\in E} P(x, z)(1 - \alpha(x, z)).
\end{aligned}$$

Si $x \ne y$, cette quantité est égale à $P(x, y)\alpha(x, y) = Q(x, y)$. Si $x = y$, puisque $\alpha(x, x) = 1$, on obtient

$$\begin{aligned}
\mathbb{P}(X_{i+1} = x|X_i = x) &= P(x, x)\alpha(x, x) + \sum_{z\in E} P(x, z)(1 - \alpha(x, z)) = P(x, x) + \sum_{z\ne x} P(x, z)(1 - \alpha(x, z))\\
&= P(x, x) + \sum_{z\ne x} P(x, z) - \sum_{z\ne x} P(x, z)\alpha(x, z) = 1 - \sum_{z\ne x} Q(x, z) = Q(x, x).
\end{aligned}$$

Ainsi, dans tous les cas, $\mathbb{P}(X_{i+1} = y|X_i = x) = Q(x, y)$, donc $X_1, \dots, X_n$ admet Q pour transition.