Quentin Berger
Francesco Caravenna
Paolo Dai Pra

Introduction aux probabilités

Modèles et applications

Solutions des exercices

7 septembre 2021

Dunod

Note

Nous avons rassemblé dans ce document *les énoncés ainsi que les solutions détaillées de tous les exercices* qui apparaissent dans le livre. Les solutions proposées ont pour but d'illustrer les méthodes de résolution, elles ne constituent pas des « modèles de rédaction ». En particulier, il pourra arriver que certains détails de calculs ou de raisonnements soient seulement esquissés.

Nous ferons régulièrement référence aux définitions, théorèmes, propositions, équations, etc. du livre, en gardant la même numérotation que dans celui-ci ; les pages, équations et figures du document présent seront précédées de la lettre « S » (par exemple, l'équation (S1.1)).

Exercices et solutions: table des matières

Chapitre 1
Espaces probabilisés discrets : la théorie

1.1 Modèles probabilistes

Exercice 1.1. Soit (Ω, P) un espace probabilisé, et soient $A, B \subseteq \Omega$ deux événements.

(i) Montrer que si $\mathrm{P}(A) = \mathrm{P}(B) = 0$, alors $\mathrm{P}(A \cup B) = 0$.

(ii) Montrer que si $\mathrm{P}(A) = \mathrm{P}(B) = 1$, alors $\mathrm{P}(A \cap B) = 1$.

Solution 1.1. (i) Comme toute probabilité est sous-additive (Proposition 1.21 (ii)), on a

$$\mathrm{P}(A \cup B) \leq \mathrm{P}(A) + \mathrm{P}(B) = 0.$$

(ii) D'après le point précédent, $\mathrm{P}(A^c \cup B^c) = 0$. Mais comme $A \cap B = (A^c \cup B^c)^c$, on a

$$\mathrm{P}(A \cap B) = 1 - \mathrm{P}(A^c \cup B^c) = 1.$$

Exercice 1.2. Améliorons le résultat de l'exercice précédent. Soit (Ω, P) un espace probabilisé, et soit $(A_n)_{n \geq 1}$ une famille *dénombrable* d'événements.

(i) Montrer que si $\mathrm{P}(A_n) = 0$ pour tout $n \in \mathbb{N}^*$, alors $\mathrm{P}(\bigcup_{n=1}^\infty A_n) = 0$.

(ii) Montrer que si $\mathrm{P}(A_n) = 1$ pour tout $n \in \mathbb{N}^*$, alors $\mathrm{P}(\bigcap_{n=1}^\infty A_n) = 1$.

Solution 1.2. On raisonne de la même manière que dans l'exercice précédent, en remarquant que par sous-additivité dénombrable (Corollaire 1.26),

$$\mathrm{P}\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mathrm{P}(A_n) = 0\,.$$

Pour le (ii), il suffit d'utiliser que

$$\bigcap_{n=1}^{\infty} A_n = \left(\bigcup_{n=1}^{\infty} A_n^c\right)^c$$

et le point précédent.

Exercice 1.3. Soit (Ω, P) un espace probabilisé et soit $C \subseteq \Omega$ un événement.

(i) Montrer que si $\mathrm{P}(C) = 0$, alors $\mathrm{P}(A \cup C) = \mathrm{P}(A)$ pour tout événement $A \subseteq \Omega$.

(ii) Montrer que si $\mathrm{P}(C) = 1$, alors $\mathrm{P}(A \cap C) = \mathrm{P}(A)$ pour tout événement $A \subseteq \Omega$.

Solution 1.3. (i) Il suffit d'observer que d'après la formule de décomposition

$$\mathrm{P}(A) - \mathrm{P}(A \cap C) = \mathrm{P}(A \cap C^c) = 0$$

car $A \cap C^c \subseteq C^c$ et $\mathrm{P}(C^c) = 0$.

(ii) De même, on a
$$\mathrm{P}(A \cup C) - \mathrm{P}(A) = \mathrm{P}(C \cap A^c) = 0$$
car $C \cap A^c \subseteq C$ et $\mathrm{P}(C) = 0$.

Exercice 1.4 (Inégalité de Bonferroni). Soient $A_1, \ldots, A_n$ des événements d'un espace probabilisé (Ω, P). Montrer par récurrence sur $n \in \mathbb{N}^*$ que

$$\mathrm{P}\left(\bigcup_{k=1}^{n} A_k\right) \geq \sum_{k=1}^{n} \mathrm{P}(A_k) - \sum_{1 \leq i < j \leq n} \mathrm{P}(A_i \cap A_j).$$

En particulier, si $\mathrm{P}(A_i \cap A_j) = 0$ pour $i \neq j$, en déduire que $\mathrm{P}(\bigcup_{k=1}^{n} A_k) = \sum_{k=1}^{n} \mathrm{P}(A_k)$.

Solution 1.4. Pour $n = 1$ l'inégalité à démontrer est réduite à

$$\mathrm{P}(A_1) \geq \mathrm{P}(A_1),$$

qui est évidemment vérifiée.

Supposons l'inégalité vérifiée pour $n \geq 1$. En se rappelant de l'identité $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$ (Proposition 1.21 (ii)), on a

$$\mathrm{P}\left(\bigcup_{k=1}^{n+1} A_k\right) = \mathrm{P}\left(\left(\bigcup_{k=1}^{n} A_k\right) \cup A_{n+1}\right) = \mathrm{P}\left(\bigcup_{k=1}^{n} A_k\right) + \mathrm{P}(A_{n+1}) - \mathrm{P}\left(\left(\bigcup_{k=1}^{n} A_k\right) \cap A_{n+1}\right).$$

Maintenant, pour le terme $\mathrm{P}\left(\bigcup_{k=1}^{n} A_k\right)$, on utilise l'hypothèse de récurrence :

$$\mathrm{P}\left(\bigcup_{k=1}^{n} A_k\right) \geq \sum_{k=1}^{n} \mathrm{P}(A_k) - \sum_{1 \leq i < j \leq n} \mathrm{P}(A_i \cap A_j).$$

Et pour le terme $\mathrm{P}\left((\bigcup_{k=1}^{n} A_k) \cap A_{n+1}\right)$ on utilise la sous-additivité (Corollaire 1.26)

$$\mathrm{P}\left(\left(\bigcup_{k=1}^{n} A_k\right) \cap A_{n+1}\right) = \mathrm{P}\left(\bigcup_{k=1}^{n} (A_k \cap A_{n+1})\right) \leq \sum_{k=1}^{n} \mathrm{P}(A_k \cap A_{n+1}).$$

En remettant tout ensemble, on obtient

$$\begin{aligned}
\mathrm{P}\left(\bigcup_{k=1}^{n+1} A_k\right) &= \mathrm{P}\left(\bigcup_{k=1}^{n} A_k\right) + \mathrm{P}(A_{n+1}) - \mathrm{P}\left(\left(\bigcup_{k=1}^{n} A_k\right) \cap A_{n+1}\right) \\
&\geq \sum_{k=1}^{n} \mathrm{P}(A_k) - \sum_{1 \leq i < j \leq n} \mathrm{P}(A_i \cap A_j) + \mathrm{P}(A_{n+1}) - \sum_{k=1}^{n} \mathrm{P}(A_k \cap A_{n+1}) \\
&= \sum_{k=1}^{n+1} \mathrm{P}(A_k) - \sum_{1 \leq i < j \leq n+1} \mathrm{P}(A_i \cap A_j).
\end{aligned}$$

Exercice 1.5. Soient (Ω_1, P_1), (Ω_2, P_2) deux espaces probabilisés *discrets*, de densités discrètes respectives p_1 et p_2 (voir la Proposition 1.15). On définit la fonction $\mathrm{p}((\omega_1, \omega_2)) = \mathrm{p}_1(\omega_1) \cdot \mathrm{p}_2(\omega_2)$, pour tous $\omega_1 \in \Omega_1$, $\omega_2 \in \Omega_2$. Montrer que p est une densité discrète sur $\Omega = \Omega_1 \times \Omega_2$ et définit donc une probabilité *discrète* P sur Ω.

Solution 1.5. On doit vérifier que

$$\mathrm{p}(\omega) \geq 0 \quad \forall \omega \in \Omega, \qquad \sum_{\omega \in \Omega} \mathrm{p}(\omega) = 1\,. \tag{S1.1}$$

La première relation est vérifiée parce que $\mathrm{p}_1(\omega_1) \geq 0$ et $\mathrm{p}_2(\omega_2) \geq 0$ car p_1 et p_2 sont des densités discrètes, donc $\mathrm{p}(\omega) = \mathrm{p}((\omega_1, \omega_2)) = \mathrm{p}_1(\omega_1) \cdot \mathrm{p}_2(\omega_2) \geq 0$ pour tout $\omega = (\omega_1, \omega_2) \in \Omega$. Pour la deuxième relation dans (S1.1) il suffit d'utiliser la sommation par paquets :

$$\sum_{\omega \in \Omega} \mathrm{p}(\omega) = \sum_{(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2} \mathrm{p}_1(\omega_1) \cdot \mathrm{p}_2(\omega_2) = \sum_{\omega_1 \in \Omega_1} \mathrm{p}_1(\omega_1) \left(\sum_{\omega_2 \in \Omega_2} \mathrm{p}_2(\omega_2) \right) = 1\,,$$

ce qui conclut la solution.

Exercice 1.6. Avec les mêmes notations que dans l'Exemple 1.20, montrer que

$$\lim_{\beta \to +\infty} \mathrm{P}_\beta(\{\omega\}) = \frac{1}{|M|}\,, \qquad \forall \omega \in M\,. \tag{S1.2}$$

On peut donc dire que, pour $\beta \to +\infty$, la probabilité P_β *converge* (dans le sens de (S1.2)) vers la probabilité uniforme sur l'ensemble M.

Solution 1.6. Si $\omega \in M$, c'est-à-dire si $H(\omega) = m$, on a

$$\mathrm{P}_\beta(\{\omega\}) = \frac{e^{-\beta H(\omega)}}{\sum_{\sigma \in \Omega} e^{-\beta H(\sigma)}} = \frac{1}{\sum_{\sigma \in \Omega} e^{-\beta [H(\sigma) - m]}} = \frac{1}{|M| + \sum_{\sigma \notin M} e^{-\beta [H(\sigma) - m]}}\,.$$

De plus, pour $\sigma \notin M$, comme $H(\sigma) > m$ on a

$$\lim_{\beta \to +\infty} e^{-\beta [H(\sigma) - m]} = 0,$$

et on obtient facilement la conclusion.

1.2 Analyse combinatoire

Exercice 1.7. Soient Ω_1 et Ω_2 deux ensembles finis non vides, et soit P_1 (resp. P_2) la probabilité uniforme sur Ω_1 (resp. sur Ω_2). Soit P la probabilité uniforme sur $\Omega_1 \times \Omega_2$. Montrer que pour tout $A \subseteq \Omega_1$ et tout $B \subseteq \Omega_2$ on a

$$\mathrm{P}_1(A) = \mathrm{P}(A \times \Omega_2)\,, \qquad \mathrm{P}_2(B) = \mathrm{P}(\Omega_1 \times B)\,.$$

Solution 1.7. Il suffit de remarquer que $|A \times \Omega_2| = |A| \cdot |\Omega_2|$ et $|\Omega_1 \times \Omega_2| = |\Omega_1| \cdot |\Omega_2|$, donc

$$\mathrm{P}_1(A) = \frac{|A|}{|\Omega_1|} = \frac{|A| \cdot |\Omega_2|}{|\Omega_1| \cdot |\Omega_2|} = \frac{|A \times \Omega_2|}{|\Omega_1 \times \Omega_2|} = \mathrm{P}(A \times \Omega_2)\,,$$

qui est la première relation à démontrer. La deuxième relation est complètement analogue.

Exercice 1.8. On considère un paquet de 52 cartes de Poker, et on choisit 5 cartes *au hasard*. Calculer la probabilité que :

(i) parmi les 5 cartes il y a *au moins* une paire (c'est-à-dire deux cartes de même valeur) ;

(ii) parmi les 5 cartes il y a *exactement* une paire, c'est-à-dire une paire mais aucune combinaison meilleure (double paire, tierce, etc.).

Solution 1.8. Soit S l'ensemble des 52 cartes, $\Omega := \{A \subseteq S : |A| = 5\}$, P la probabilité uniforme sur Ω.

(i) Soit $E =$ « parmi les 5 cartes il y a au moins paire ». Le choix d'un élément de $E^c =$ « parmi les 5 cartes il y n'a aucune paire » peut se faire en deux étapes.

(1) On choisit 5 *valeurs* (c'est-à-dire nombre ou figure) *différentes* ; il y a $\binom{13}{5}$ possibilités pour ce choix.

(2) Une fois effectué le choix des cinq valeurs différentes dans (1), il faut choisir pour chacune des cinq valeurs l'une des quatre couleurs possibles (cœur $\heartsuit$, carreau $\diamondsuit$, trèfle $\clubsuit$, pique $\spadesuit$) ; il y a 4^5 façons de faire cela.

D'après le principe fondamental du dénombrement (Théorème 1.33), il en découle que $|E^c| = 4^5\binom{13}{5}$. Ainsi,

$$\mathrm{P}(E) = 1 - \frac{4^5\binom{13}{5}}{\binom{52}{5}} \simeq 0{,}493\,.$$

(iii) Soit F l'événement « parmi les 5 cartes il y a exactement une paire ». Le choix d'un élément de F peut se faire en trois étapes :

(1) on choisit la valeur (nombre ou figure) pour la paire : il y a 13 possibilités ;

(2) une fois le choix (1) fait, on choisit les deux couleurs pour la paire : il y a $\binom{4}{2}$ possibilités ;

(3) une fois les choix (1) et (2) effectués, on choisit 3 cartes avec des valeurs (nombre ou figure) différentes et différentes de celle utilisée pour la paire. Le nombre de possibilités est déterminé comme pour le point (i), mais avec 48 cartes au lieu de 52 et 3 cartes choisies au lieu de 5 : il y a donc $4^3\binom{12}{3}$ possibilités.

Au final, on a $|F| = 13\binom{4}{2}\,4^3\binom{12}{3}$, dont on déduit que

$$\mathrm{P}(F) = \frac{13\binom{4}{2}\,4^3\binom{12}{3}}{\binom{52}{5}} \simeq 0{,}422\,.$$

Exercice 1.9. Une classe contient 30 élèves, parmi lesquels Agathe, Camille et Mathilde. Un enseignant divise la classe, au hasard, en trois groupes de 10 élèves.

(i) Quelle est la probabilité qu'Agathe, Camille et Mathilde finissent dans trois groupes différents ? (Ne pas simplifier les coefficients binomiaux.)

(ii) Quelle est la probabilité qu'elles finissent dans le même groupe ?

Solution 1.9. (i) Soit

$$\Omega := \big\{(A_1,A_2,A_3) : A_i \subseteq \{1,2,\ldots,30\},\ |A_i| = 10,\ A_i \cap A_j = \emptyset \text{ per } i \neq j\big\}$$

et P la probabilité uniforme sur Ω. Notons que Ω est constitué de *triplets ordonnés* de sous-ensembles qui forment une partition de l'ensemble $\{1,2,\ldots,30\}$. (Il serait possible de considérer des triplets non ordonnés.) Il n'est pas restrictif de supposer qu'Agathe, Camille et Mathilde correspondent respectivement aux éléments $1,2$ et 3 de $\{1,2,\ldots,30\}$. Un élément de Ω est déterminé par les choix successifs suivants :

- on choisit A_1 : il y a $\binom{30}{10}$ possibilités ;
- on choisit A_2 dans $\{1,2,\ldots,30\}\setminus A_1$: il y a $\binom{20}{10}$ possibilités.

Évidemment, A_3 est alors complètement déterminé. Ainsi, d'après le principe fondamental du dénombrement (Théorème 1.33),

$$|\Omega| = \binom{30}{10}\binom{20}{10}.$$

Considérons maintenant l'événement $B =$ « Agathe, Camille et Mathilde finissent dans trois groupes différents ». Un élément de B est déterminé par les choix successifs suivants :

- on choisit 9 éléments pour A_1 dans $\{4,5,\ldots,30\}$: il y a $\binom{27}{9}$ possibilités ;
- on choisit 9 éléments pour A_2 dans $\{4,5,\ldots,30\}\setminus A_1$: il y a $\binom{18}{9}$ possibilité ;
- on choisit comment disposer $1,2$ et 3 dans les trois places libres des groupes : il y a $3! = 6$ possibilités.

Ainsi :

$$|B| = 6\binom{27}{9}\binom{18}{9} \quad \Rightarrow \quad \mathrm{P}(B) = \frac{6\binom{27}{9}\binom{18}{9}}{\binom{30}{10}\binom{20}{10}} \simeq 0,246\,.$$

(ii) Soit $C =$ « Agathe, Camille et Mathilde finissent dans le même groupe ». Un élément de C est déterminé par les choix successifs suivants :

- on choisit le groupe dans lequel placer Agathe, Camille et Mathilde : il y a 3 possibilités ;
- on choisit les autres éléments de ce groupe : il y a $\binom{27}{7}$ possibilités ;
- on choisit les éléments de l'un des deux autres groupes (n'importe lequel) : il y a $\binom{20}{10}$ possibilités.

Ainsi :

$$\mathrm{P}(C) = \frac{3\binom{27}{7}\binom{20}{10}}{\binom{30}{10}\binom{20}{10}} = \frac{3\binom{27}{7}}{\binom{30}{10}} \simeq 0,089\,.$$

1.3 Probabilité conditionnelle et indépendance

Exercice 1.10. Montrer, par des exemples, que les inégalités $\mathrm{P}(A\,|\,B) > \mathrm{P}(A)$ et $\mathrm{P}(A\,|\,B) < \mathrm{P}(A)$ sont toutes les deux possibles.

Solution 1.10. On considère un espace probabilisé quelconque (Ω,P) dans lequel on se donne un événement B pour lequel $\mathrm{P}(B) \in \,]0,1[$. (Par exemple, prendre l'ensemble $\Omega = \{0,1\}$ muni de la probabilité P uniforme et $B := \{0\}$, de sorte que $\mathrm{P}(B) = \frac{1}{2}$.) En posant $A = B$, on a

$$\mathrm{P}(A\,|\,B) = \mathrm{P}(B\,|\,B) = 1 > \mathrm{P}(A)\,,$$

alors qu'en posant $A = B^c$, ona

$$\mathrm{P}(A\,|\,B) = \mathrm{P}(B^c\,|\,B) = 0 < \mathrm{P}(A).$$

Exercice 1.11. Soient A, B, C trois événements d'un espace probabilisé. On suppose que A, B, C sont indépendants.

(i) Montrer que $A \cap B$ est indépendant de C.

(ii) Montrer que $A \cup B$ est indépendant de C.

Solution 1.11. (i) Il découle de la définition d'indépendance d'événements que

$$\mathrm{P}\big((A\cap B)\cap C\big) = \mathrm{P}(A\cap B\cap C) = \mathrm{P}(A)\,\mathrm{P}(B)\,\mathrm{P}(C) = \mathrm{P}(A\cap B)\,\mathrm{P}(C).$$

(ii) D'après la Proposition 1.70, c'est équivalent à démontrer que $(A\cup B)^c = A^c \cap B^c$ est indépendant de C. Comme les événements A^c, B^c, C sont indépendants (de nouveau grâce à la Proposition 1.70), on s'est ramené au cas précédent vu dans (i).

De manière alternative, on peut procéder par un calcul direct :

$$\begin{aligned}
\mathrm{P}\big((A\cup B)\cap C\big) &= \mathrm{P}((A\cap C)\cup(B\cap C)) \\
&= \mathrm{P}(A\cap C) + \mathrm{P}(B\cap C) - \mathrm{P}((A\cap C)\cap(B\cap C)) \\
&= \mathrm{P}(A\cap C) + \mathrm{P}(B\cap C) - \mathrm{P}(A\cap B\cap C) \\
&= \mathrm{P}(A)\,\mathrm{P}(C) + \mathrm{P}(B)\,\mathrm{P}(C) - \mathrm{P}(A)\,\mathrm{P}(B)\,\mathrm{P}(C) \\
&= (\mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A)\,\mathrm{P}(B))\,\mathrm{P}(C) \\
&= (\mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A\cap B))\,\mathrm{P}(C) \\
&= \mathrm{P}(A\cup B)\,\mathrm{P}(C)\,.
\end{aligned}$$

Exercice 1.12. Soient $A_1, A_2, \ldots, A_n$ des événements indépendants. On suppose que l'on a $\mathrm{P}(A_1 \cup A_2 \cup \cdots \cup A_n) = 1$. Montrer qu'il existe un $k \in \{1, 2, \ldots, n\}$ tel que $\mathrm{P}(A_k) = 1$.

Solution 1.12. En utilisant la Définition 1.68 de l'indépendance d'événements et le fait que l'indépendance de n événements implique l'indépendance de leurs complémentaires (Proposition 1.70), on obtient

$$0 = \mathrm{P}\left((A_1 \cup A_2 \cup \cdots \cup A_n)^c\right) = \mathrm{P}(A_1^c \cap \cdots \cap A_n^c) = \prod_{k=1}^{n} \mathrm{P}(A_k^c).$$

Comme un produit de nombres réels vaut zéro si et seulement si un des facteurs vaut zéro, il s'ensuit qu'il existe un $k \in \{1, 2, \ldots, n\}$ tel que $\mathrm{P}(A_k^c) = 0$, c'est-à-dire tel que $\mathrm{P}(A_k) = 1$.

Exercice 1.13. On se donne trois nombres : $\alpha_1, \alpha_2 \in [0,1]$ et $\beta \in \,]0,1[$. Montrer qu'il existe un espace probabilisé discret (Ω, P) contenant deux événements A, B tels que

$$\mathrm{P}(B) = \beta\,, \qquad \mathrm{P}(A\,|\,B) = \alpha_1\,, \qquad \mathrm{P}(A\,|\,B^c) = \alpha_2\,.$$

[*Sugg.* Considérer $\Omega = \{ab, a\bar{b}, \bar{a}b, \bar{a}\bar{b}\} = \{a, \bar{a}\} \times \{b, \bar{b}\}$, définir $A := \{ab, a\bar{b}\}$, $B := \{ab, \bar{a}b\}$ et montrer qu'il existe une unique probabilité P sur Ω qui satisfait les propriétés requises.]

Solution 1.13. Il suffit de définir la densité discrète $\mathrm{p}(\omega) := \mathrm{P}(\{\omega\})$. On doit faire en sorte que l'on ait les relations

$$\mathrm{P}(B) = \beta\,, \qquad \mathrm{P}(A\,|\,B) = \alpha_1\,, \qquad \mathrm{P}(A\,|\,B^c) = \alpha_2\,. \tag{S1.3}$$

On doit nécessairement avoir

$$\mathrm{p}(ab) \;=\; \mathrm{P}(\{ab\}) \;=\; \mathrm{P}(A \cap B) \;=\; \mathrm{P}(B)\,\mathrm{P}(A\,|\,B) \;=\; \beta\,\alpha_1\,.$$

Avec des calculs analogues, on récupère toutes les valeurs de p :

$$\begin{aligned} \mathrm{p}(ab) &= \alpha_1\beta\,, & \mathrm{p}(a\bar{b}) &= \alpha_2(1-\beta)\,, \\ \mathrm{p}(\bar{a}b) &= (1-\alpha_1)\beta\,, & \mathrm{p}(\bar{a}\bar{b}) &= (1-\alpha_2)(1-\beta)\,. \end{aligned}$$

Ainsi, la densité discrète p est complètement déterminée par les relations (S1.3) requises dans le problème. Réciproquement, il est immédiat de vérifier que la fonction p *définie* comme ci-dessus est effectivement une densité discrète (Définition 1.14), c'est-à-dire $\mathrm{p}(\omega) \geq 0$ pour tout $\omega \in \Omega$ et $\sum_{\omega\in\Omega} \mathrm{p}(\omega) = 1$, et que les propriétés requises (S1.3) sont vérifiées.

Exercice 1.14 (Paradoxe des trois prisonniers). Trois prisonniers (A, B, C) sont condamnés à être pendus. Le souverain décide d'en gracier un des trois au hasard, mais le nom de l'heureux élu ne sera communiqué qu'aux gardiens de prison. Le prisonnier A s'approche du gardien, qui connaît le nom du prisonnier qui sera gracié, et lui dit : « s'il te plaît, donne-moi un nom, entre B et C, qui sera pendu. On sait déjà que l'un des deux sera pendu à coup sûr, donc tu ne me donnes aucune information en me le disant ». Le gardien réfléchit, trouve le raisonnement censé, et répond : « B sera pendu ». À ce moment-là, A s'exclame : « Hourra ! Vu que B sera pendu, le choix n'est plus qu'entre C et moi, et j'ai donc maintenant une chance sur deux d'être gracié, alors qu'avant je n'avais qu'une chance sur trois. » Ce raisonnement est-il correct ?

Solution 1.14. Comme dans le paradoxe de Monty-Hall (Exemple 1.81) il est important de préciser la façon dont le gardien choisit le nom à donner à A. L'option la plus naturelle est que le gardien donne le nom B avec probabilité 1 si C est celui qui est gracié, et avec probabilité $\frac{1}{2}$ si A est celui qui est gracié. Si maintenant on considère les événements E_x = « x est celui qui est gracié », pour $x \in \{\text{A,B,C}\}$, et F = « le gardien annonce que B est pendu », on a

$$\mathrm{P}(E_A) = \mathrm{P}(E_B) = \mathrm{P}(E_C) = \frac{1}{3},$$

et

$$\mathrm{P}(F\,|\,E_A) = \frac{1}{2}, \quad \mathrm{P}(F\,|\,E_B) = 0, \quad \mathrm{P}(F\,|\,E_C) = 1.$$

Ainsi, sachent que le gardien annonce que B sera pendu, la probabilité (conditionnelle) que A soit gracié vaut

$$\mathrm{P}(E_A\,|\,F) = \frac{\mathrm{P}(F\,|\,E_A)\,\mathrm{P}(E_A)}{\mathrm{P}(F\,|\,E_A)\,\mathrm{P}(E_A) + \mathrm{P}(F\,|\,E_B)\,\mathrm{P}(E_B) + \mathrm{P}(F\,|\,E_C)\,\mathrm{P}(E_C)} = \frac{1}{3},$$

c'est-à-dire la même qu'au départ. Le raisonnement de A est donc faux.

Exercice 1.15 (Paradoxe des trois cartes). On place dans un sac trois cartes : une qui a deux faces rouges, une qui a deux faces noires, et une dernière qui a une face rouge et une face noire. Les yeux fermés, on tire une carte au hasard, et on la dépose sur la table sur une face au hasard. À ce moment-là, on ouvre les yeux. Si la face que l'on voit est rouge, quelle est la probabilité que l'autre face soit rouge ?

Solution 1.15. Considérons les événements : NN = « la carte tirée est celle avec deux faces noires », RR = « la carte tirée est celle avec deux faces rouges », NR = « la carte tirée est celle avec une face rouge et une face noire », R = « la face que l'on voit est rouge ». Par hypothèse, on a

$$\mathrm{P}(NN) = \mathrm{P}(RR) = \mathrm{P}(NR) = \frac{1}{3}$$

$$\mathrm{P}(R\,|\,NN) = 0, \quad \mathrm{P}(R\,|\,RR) = 1, \quad \mathrm{P}(R\,|\,NR) = \frac{1}{2}\,.$$

Ainsi, d'après la formule de Bayes (Théorème 1.55, dans la version de la formule (1.50)), la probabilité cherchée vaut

$$\mathrm{P}(RR\,|\,R) = \frac{\mathrm{P}(R\,|\,RR)\,\mathrm{P}(RR)}{\mathrm{P}(R\,|\,RR)\,\mathrm{P}(RR) + \mathrm{P}(R\,|\,NN)\,\mathrm{P}(NN) + \mathrm{P}(R\,|\,NR)\,\mathrm{P}(NR)} = \frac{2}{3}\,.$$

Exercice 1.16 (Paradoxe de Simpson). Soit (Ω, P) un espace probabilisé, et soient A_1, A_2 et B trois événements.

(i) On suppose que l'on a $\mathrm{P}(A_1\,|\,B) \geq \mathrm{P}(A_2\,|\,B)$ et $\mathrm{P}(A_1\,|\,B^c) \geq \mathrm{P}(A_2\,|\,B^c)$. Montrer que $\mathrm{P}(A_1) \geq \mathrm{P}(A_2)$.

(ii) On se donne un autre événement C, et on suppose maintenant que $\mathrm{P}(A_1\,|\,C) \geq \mathrm{P}(A_2\,|\,B)$ et $\mathrm{P}(A_1\,|\,C^c) \geq \mathrm{P}(A_2\,|\,B^c)$. A-t-on toujours $\mathrm{P}(A_1) \geq \mathrm{P}(A_2)$?

(iii) Deux hommes politiques s'affrontent lors d'un débat. Le premier affirme que le taux de réussite au baccalauréat a augmenté aussi bien chez les hommes que chez les femmes. Le deuxième répond le taux de réussite global a diminué. Est-il possible qu'aucun des deux ne mente ?

Solution 1.16. (i) D'après la formule des probabilités totales,

$$\begin{aligned}\mathrm{P}(A_1) &= \mathrm{P}(A_1\,|\,B)\,\mathrm{P}(B) + \mathrm{P}(A_1\,|\,B^c)\,\mathrm{P}(B^c) \\ &\geq \mathrm{P}(A_2\,|\,B)\,\mathrm{P}(B) + \mathrm{P}(A_2\,|\,B^c)\,\mathrm{P}(B^c) = \mathrm{P}(B_2)\,.\end{aligned}$$

(ii) En général non. Supposons par exemple que

$$\mathrm{P}(A_1\,|\,B_1) = 82\% > \mathrm{P}(A_2\,|\,B_2) = 80\%\,, \qquad \mathrm{P}(A_1\,|\,B_1^c) = 92\% > \mathrm{P}(A_2\,|\,B_2^c) = 90\%\,.$$

Si $\mathrm{P}(B_1) = 50\%$ et $\mathrm{P}(B_2) = 20\%$, on obtient

$$\mathrm{P}(A_1) = \mathrm{P}(A_1\,|\,B_1)\,\mathrm{P}(B_1) + \mathrm{P}(A_1\,|\,B_1^c)\,\mathrm{P}(B_1^c) = \frac{82}{100}\cdot\frac{1}{2} + \frac{92}{100}\cdot\frac{1}{2} = \frac{87}{100} = 87\%\,,$$

alors que

$$\mathrm{P}(A_2) = \mathrm{P}(A_2\,|\,B_2)\,\mathrm{P}(B_2) + \mathrm{P}(A_2\,|\,B_2^c)\,\mathrm{P}(B_2^c) = \frac{80}{100}\cdot\frac{1}{5} + \frac{90}{100}\cdot\frac{4}{5} = \frac{88}{100} = 88\%\,,$$

donc $\mathrm{P}(A_1) < \mathrm{P}(A_2)$.

(iii) En principe ils pourraient avoir raison tous les deux ! Par exemple, supposons que l'année dernière le baccalauréat a été passé par 20 hommes, dont 16 l'ont réussi, et 80 femmes, donc 72 l'ont réussi : le pourcentage de réussite est donc de $\frac{16}{20} = 80\%$ chez les hommes et de $\frac{72}{80} = 90\%$ chez les femmes ; de plus le pourcentage global de réussite est égal à $\frac{16+72}{20+80} = 88\%$. Supposons que cette année le baccalauréat a été passé par 50 hommes, dont 41 l'ont réussi, et 50 femmes, dont 46 l'ont réussi : le pourcentage de réussite est donc de $\frac{41}{50} = 82\%$ chez les hommes et de $\frac{46}{50} = 92\%$ chez les femmes, donc *en*

augmentation par rapport à l'année précédente; cependant, le pourcentage global de réussite est égal à $\frac{41+46}{50+50} = 87\%$, *en diminution* par rapport à l'année précédente !
On peut donner une interprétation probabiliste de cette exemple en considérant l'expérience aléatoire qui consiste à choisir au hasard deux élèves de façon aléatoire : le premier qui a passé l'examen l'année dernière, le deuxième qui l'a passé cette année. Si on introduit les événements

$$\begin{aligned} A_1 &:= \text{« le premier élève a réussi l'examen »}, \\ B_1 &:= \text{« le premier élève est un homme »}, \\ A_2 &:= \text{« le deuxième élève a réussi l'examen »}, \\ B_2 &:= \text{« le deuxième élève est un homme »}, \end{aligned}$$

le premier homme politique soutient que $\mathrm{P}(A_1 \,|\, B_1) > \mathrm{P}(A_2 \,|\, B_2)$ et $\mathrm{P}(A_1 \,|\, B_1^c) > \mathrm{P}(A_2 \,|\, B_2^c)$, le deuxième soutient que $\mathrm{P}(A_1) < \mathrm{P}(A_2)$. Comme on l'a vu plus haut, ces deux affirmations ne sont pas a priori incompatibles.

1.4 Exercices récapitulatifs

Exercice 1.17. Soient A, B deux événements. On rappelle la définition de *différence symétrique*, $A \triangle B := (A \cup B) \setminus (A \cap B)$. Montrer que

$$\mathrm{P}(A \triangle B) = \mathrm{P}(A) + \mathrm{P}(B) - 2\,\mathrm{P}(A \cap B).$$

Soient maintenant A, B, C trois événements. Montrer que

$$\mathrm{P}(A \triangle C) \leq \mathrm{P}(A \triangle B) + \mathrm{P}(B \triangle C).$$

Solution 1.17. Pour la première partie, on remarque que

$$A \cup B = (A \triangle B) \cup (A \cap B),$$

et cette dernière union est disjointe. Ainsi, par additivité des probabilités,

$$\mathrm{P}(A \cup B) = \mathrm{P}(A \triangle B) + \mathrm{P}(A \cap B),$$

c'est-à-dire $\mathrm{P}(A \triangle B) = \mathrm{P}(A \cup B) - \mathrm{P}(A \cap B)$. On conclut alors en utilisant l'identité

$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B),$$

vue dans la Proposition 1.21 (ii).

Pour la deuxième partie, il suffit de remarquer que

$$A \triangle C \subseteq (A \triangle B) \cup (B \triangle C),$$

ce que l'on vérifie facilement. L'inégalité recherchée découle alors de la sous-additivité des probabilités (de nouveau, voir la Proposition 1.21 (ii)).

Exercice 1.18. Soient A et B deux événements arbitraires d'un espace probabilisé (Ω, P).

(i) Montrer l'inégalité $\mathrm{P}(A \cap B) \geq \mathrm{P}(A) + \mathrm{P}(B) - 1$.

(ii) Montrer par récurrence que, pour tout $n \geq 2$ et tout choix d'événements $A_1, A_2, \dots, A_n$,

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) \geq \sum_{i=1}^{n} P(A_i) - (n-1).$$

Solution 1.18. Pour le cas $n = 2$, il suffit d'observer que

$$1 \geq P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Pour $n > 2$, on procède par récurrence :

$$\begin{aligned} P(A_1 \cap A_2 \cap \cdots \cap A_n) &= P\big((A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \cap A_n\big) \\ &\geq P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) + P(A_n) - 1 \\ &\geq \sum_{i=1}^{n-1} P(A_i) - (n-2) + P(A_n) - 1 \\ &= \sum_{i=1}^{n} P(A_i) - (n-1)\,, \end{aligned}$$

où on a d'abord utilisé le résultat pour $n = 2$ puis l'hypothèse de récurrence.

Exercice 1.19. On modélise une cible de fléchettes en considérant quatre cercles concentriques, de rayon $1, 2, 3, 4$ respectivement qui divise la cible en quatre zones

$$\begin{aligned} D_1 &= \{(x,y) \in \mathbb{R}^2, x^2 + y^2 \leq 1\}, & D_2 &= \{(x,y) \in \mathbb{R}^2, x^2 + y^2 \in \,]1,2]\}, \\ D_3 &= \{(x,y) \in \mathbb{R}^2, x^2 + y^2 \in \,]2,3]\}, & D_4 &= \{(x,y) \in \mathbb{R}^2, x^2 + y^2 \in \,]3,4]\}. \end{aligned}$$

Lorsque Kévin lance une fléchette, il atteint la cible avec probabilité $3/4$, et conditionnellement au fait que l'on atteint la cible la probabilité d'atteindre la zone D_i est proportionnelle à l'aire de D_i. Des lancers différents ont des résultats indépendants. Calculer la probabilité que Kévin atteigne la zone D_i, pour $i = 1, 2, 3, 4$. Calculer la probabilité qu'il atteigne la même zone de la cible deux fois d'affilée.

Solution 1.19. Introduisons les événements $A :=$ « Kévin atteint la cible » et $B_i :=$ « Kévin atteint la zone D_i », pour $i = 1, 2, 3, 4$. Comme la zone D_i possède une aire $\pi(i^2 - (i-1)^2) = \pi(2i-1)$ et que l'aire totale de la cible est $\pi 4^2 = 16\pi$, on a

$$P(A) = \frac{3}{4}\,, \qquad P(B_i \mid A) = \frac{\pi(2i-1)}{16\pi} = \frac{2i-1}{16} = \begin{cases} \frac{1}{16} & i = 1\,, \\ \frac{3}{16} & i = 2\,, \\ \frac{5}{16} & i = 3\,, \\ \frac{7}{16} & i = 4\,. \end{cases}$$

Par conséquent, comme $P(B_i \mid A^c) = 0$, la probabilité de toucher la zone D_i vaut

$$P(B_i) = P(B_i \mid A)\,P(A) + P(B_i \mid A^c)\,P(A^c) = \frac{3}{4}\,\frac{2i-1}{16} = \frac{3(2i-1)}{2^6}\,.$$

En supposant que les lancers successifs soient indépendants, la probabilité de toucher deux fois de suite une zone fixée D_i vaut $q_i := P(D_i)^2 = \frac{3^2(2i-1)^2}{2^{12}}$. Par conséquent, la probabilité que deux lancers successifs touchent la même zone (non fixée) vaut

$$q = \sum_{i=1}^{4} q_i = \frac{3^2}{2^{12}}(1+3^2+5^2+7^2) = \frac{9\cdot 84}{4096} = \frac{189}{1024} \simeq 18,5\%\,.$$

Exercice 1.20. Dans un paquet de 52 cartes de Poker, on tire au hasard trois cartes. Calculer la probabilité que

(i) parmi les trois cartes, il y ait au moins un as ;

(ii) les trois cartes tirées soient de motifs différents ;

(iii) au moins deux des cartes tirées aient la même valeur.

Solution 1.20. (i) Ω est l'ensemble des sous-ensembles de trois éléments du paquet de 52 cartes, donc $|\Omega| = \binom{52}{3}$. Le nombre de manières de choisir 3 cartes de façon à ce qu'il n'y ait aucun as est égal à $\binom{48}{3}$. Ainsi, la probabilité recherchée est

$$1 - \frac{\binom{48}{3}}{\binom{52}{3}}\,.$$

(ii) Soit A l'événement dont on doit calculer la probabilité. Choisir un élément de A signifie tout d'abord choisir 3 motifs pour les cartes (il y a $\binom{4}{3}$ possibilités) ; une fois les motifs fixés, chacune des cartes peut être choisie de 13 façons différentes. Donc, d'après le principe fondamental du dénombrement (Théorème 1.33), on a

$$|A| = \binom{4}{3} 13^3\,,$$

ce qui, divisé par $|\Omega| = \binom{52}{3}$, donne la probabilité voulue.

(iii) Si B est l'événement dont on doit calculer la probabilité, alors $B^c =$ « les trois cartes ont des valeurs différentes ». Choisir un élément de B^c signifie tout d'abord choisir trois valeurs (nombre ou figure) différentes parmi les 13 disponibles (il y a $\binom{13}{3}$ possibilités) ; une fois les valeurs fixées, chacune des cartes peut être choisie de 4 façons différentes. Donc

$$\mathrm{P}(B) = 1 - \mathrm{P}(B^c) = 1 - \frac{\binom{13}{3}4^3}{\binom{52}{3}}\,.$$

Exercice 1.21. Une loterie émet n tickets, dont $m < n$ sont gagnants. Quelle est la probabilité que quelqu'un achetant r tickets ait au moins un ticket gagnant ?

Solution 1.21. On peut choisir Ω l'ensemble des sous-ensemble de r éléments de l'ensemble des n tickets, muni de la probabilité uniforme. Si A est l'événement en question, A^c est l'ensemble des sous-ensembles de r éléments des $n-m$ tickets non gagnants. Alors

$$\mathrm{P}(A) = 1 - \mathrm{P}(A^c) = 1 - \frac{\binom{n-m}{r}}{\binom{n}{r}}\,.$$

Exercice 1.22. Nassim possède un dé à n faces et Leia un dé à m faces. En lançant leurs dés, quelle est la probabilité que Nassim obtienne un nombre strictement plus petit que celui de Leia ?

Solution 1.22. Introduisons l'événement $A :=$ « Nassim obtient un numéro strictement plus petit que celui de Leia » et les événements $B_j :=$ « le dé de Leia donne le résultat j » pour $j = 1,\ldots,m$. Alors

$$\mathrm{P}(B_j) = \frac{1}{m}, \qquad \mathrm{P}(A \mid B_j) = \frac{\min\{j-1,n\}}{n} \quad \forall j = 1,\ldots,m,$$

parce que le nombre de résultats $i \in \{1,\ldots,n\}$ du dé de Nassim pour lequel $i < j$ est égal à $\min\{j-1,n\}$ (c'est-à-dire est égal à $j-1$ si $j-1 \le n$ et est égal à n si $j-1 > n$). Donc, d'après la formule des probabilités totales

$$\mathrm{P}(A) = \sum_{j=1}^{m} \mathrm{P}(A \mid B_j)\,\mathrm{P}(B_j) = \sum_{j=1}^{m} \frac{1}{m}\frac{\min\{j-1,n\}}{n} = \frac{1}{mn}\sum_{\ell=0}^{m-1} \min\{\ell,n\}.$$

On peut rendre cette formule plus explicite en distinguant deux cas.

- Si $m \le n+1$, alors $\min\{\ell,n\} = \ell$ pour tout $\ell \le m-1$, et donc

$$\mathrm{P}(A) = \frac{1}{mn}\sum_{\ell=0}^{m-1} \ell = \frac{1}{mn}\frac{(m-1)m}{2} = \frac{m-1}{2n}.$$

- Si $m > n+1$, alors $\min\{j-1,n\}$ si $j \le n$ et $\min\{j-1,n\} = n$ si $j \ge n+1$, donc

$$\mathrm{P}(A) = \frac{1}{mn}\left(\sum_{\ell=0}^{n-1} \ell + \sum_{\ell=n+1}^{m} n\right) = \frac{1}{mn}\left(\frac{(n-1)n}{2} + (m-n)n\right) = \frac{n-1}{2m} + \frac{m-n}{m}.$$

Exercice 1.23. Pour deux entiers $i \le j$, on utilise la notation $[\![i,j]\!]$ pour désigner l'intervalle discret $\{i, i+1, \ldots, j-1, j\}$. On choisit de manière uniforme un intervalle discret de $[\![1,n]\!]$: on considère l'espace d'états $\Omega = \{[\![i,j]\!], 1 \le i \le j \le n\}$, muni de la probabilité uniforme P.

(i) Pour $1 \le k \le n$, quelle est la probabilité que l'intervalle choisi contienne exactement k éléments ?

(ii) Pour $x \in [\![1,n]\!]$, quelle est la probabilité que l'intervalle choisi contienne l'élément x ?

Solution 1.23. Commençons par déterminer le cardinal de Ω. Chaque élément $\omega = [\![i,j]\!] \in \Omega$ est identifié à un couple (i,j) avec $1 \le i \le j \le n$, ou de façon équivalente à un sous-ensemble $\{i,j\} \subseteq \{1,\ldots,n\}$ avec un ou deux éléments (suivant que l'on ait $i = j$ ou $i < j$). Le cardinal de Ω est donc égal au nombre de sous-ensembles de $\{1,\ldots,n\}$ à un ou deux éléments, c'est-à-dire

$$|\Omega| = \binom{n}{1} + \binom{n}{2} = n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2}.$$

Une autre manière d'obtenir ce résultat est par une énumération directe : pour tout $j \in \{1,\ldots,n\}$ il y a exactement j valeurs de $i \in \{1,\ldots,n\}$ telles que $i \le j$, et donc

$$|\Omega| = \sum_{j=1}^{n} j = \frac{n(n+1)}{2}.$$

Répondons maintenant aux questions de l'exercice.

(i) Soit $A := \{\omega = [\![i,j]\!] \in \Omega : \ j = i+k-1\}$ l'événement « l'intervalle choisi contient exactement k éléments ». Avec la restriction $j \le n$ on obtient $i \le n-(k-1)$, ce qui donne $|A| = n-(k-1)$ et donc

$$\mathrm{P}(A) = \frac{|A|}{|\Omega|} = 2\frac{n-(k-1)}{n(n+1)}.$$

(ii) Soit $B := \{\omega = [\![i,j]\!] \in \Omega : \ i \leq x \leq j\}$ l'événement « l'intervalle choisi contient l'élément x ». Pour déterminer tous les intervalles $\omega = [\![i,j]\!] \in B$, on note qu'il y a x choix pour i et $n-x+1$ choix pour j, ce qui donne $|B| = x(n-x+1)$ et donc

$$\mathrm{P}(B) = \frac{|B|}{|\Omega|} = 2\,\frac{x(n-x+1)}{n(n+1)}\,.$$

Exercice 1.24. On effectue une permutation aléatoire des nombres $\{1,2,\dots,n\}$. Quelle est la probabilité que 1 et 2 soient encore consécutifs après la permutation ?

Solution 1.24. Soit $\Omega = \mathfrak{S}_n$ l'ensemble des permutations des n nombres, et soit

$$A := \{\sigma \in \Omega : \sigma(2) = \sigma(1) + 1\}.$$

Chaque élément de A est déterminé par les choix successifs suivants :

(i) la valeur du couple $(\sigma(1), \sigma(2))$ peut être choisie de $n-1$ façons possibles, à savoir $(1,2)$, $(2,3)$, …, $(n-1,n)$;

(ii) les valeurs de $\sigma(i)$ pour $i > 2$ peuvent être choisies à loisir dans $\{1,2,\dots,n\} \setminus \{\sigma(1), \sigma(2)\}$, ce qui peut se faire de $(n-2)!$ façons différentes.

On a donc

$$\mathrm{P}(A) = \frac{(n-1)(n-2)!}{n!} = \frac{1}{n}\,.$$

Exercice 1.25. Soit $\mathfrak{S}_n$ l'ensemble des permutations de $\{1,2,\dots,n\}$. Pour $\sigma \in \mathfrak{S}_n$ et $I \subseteq \{1,2,\dots,n\}$, on dit que l'ensemble I est *stable* pour σ si $\sigma(i) \in I$ pour tout $i \in I$. On note $A_I \subseteq \mathfrak{S}_n$ l'ensemble des permutations pour lesquelles I est stable. Si P est la probabilité uniforme sur $\mathfrak{S}_n$, calculer $\mathrm{P}(A_I)$.

Solution 1.25. On observe que I est stable par la permutation σ si et seulement si σ est donnée par la composition d'une permutation de I avec une permutation de I^c. Ainsi, $|A_I| = |I|!(n-|I|)!$, dont on déduit

$$\mathrm{P}(A_I) = \frac{|I|!(n-|I|)!}{n!} = \frac{1}{\binom{n}{|I|}}\,.$$

Exercice 1.26. On effectue n tirages aléatoires *avec remise* dans une urne qui contient $2n$ objets distincts. Déterminer la probabilité p_n que les n objets tirés soient tous distincts. En utilisant la formule de Stirling (1.30), déterminer ensuite le comportement asymptotique de p_n quand $n \to +\infty$: montrer que $p_n \sim c\rho^n$ (dans le sens où $\lim_{n\to\infty} p_n/(c\rho^n) = 1$) et donner les valeurs de c et de ρ.

Solution 1.26. (i) Le nombre de suites de n tirages où les objets tirés sont tous différents vaut

$$2n(2n-1)\cdots(n+1) = \frac{(2n)!}{n!}\,.$$

Ainsi,

$$p_n = \frac{(2n)!}{n!(2n)^n}\,.$$

(ii) En utilisant la formule de Stirling, on obtient

$$p_n = \frac{\sqrt{4\pi n}(2n)^{2n}e^{-2n}e^{\theta(2n)/12n}}{\sqrt{2\pi n}n^n e^{-n}e^{\theta(n)/12n}(2n)^n} \sim \sqrt{2}\left(\frac{2}{e}\right)^n,$$

c'est-à-dire $c = \sqrt{2}$ et $\rho = 2/e$.

Exercice 1.27. Dans une urne qui contient n boules dont k rouges et $n-k$ vertes, avec $1 \leq k \leq n-1$, on tire une première boule, puis sans la remettre dans l'urne, on en tire une deuxième. On considère les événements décrits de manière informelle par

$$A_1 := \text{« la première boule tirée est rouge »},$$
$$A_2 := \text{« la deuxième boule tirée est rouge »}.$$

Montrer que les événements A_1 et A_2 *ne sont pas* indépendants.

Solution 1.27. Il suffit d'observer que

$$\mathrm{P}(A_1 \,|\, A_2) = \frac{k-1}{n-1} \neq \frac{k}{n} = \mathrm{P}(A_1).$$

Exercice 1.28. On souhaite éclairer une pièce avec un certain nombre d'ampoules. On suppose que la probabilité qu'une ampoule tienne au moins n jours vaut q^n, avec $q = 0,9$. On suppose aussi que les ampoules se comportent de manière indépendante. Combien d'ampoules doit-on installer pour que, avec probabilité au moins $0,99$, il reste au moins une ampoule qui fonctionne dix jours plus tard ?

Solution 1.28. La probabilité qu'une ampoule survive au moins 10 jours vaut q^{10}, donc la probabilité qu'une ampoule ait cessé de fonctionner dans les 10 premiers jours vaut $1-q^{10}$. Par conséquent, la probabilité qu'après 10 jours les N ampoules installées aient toutes cessé de fonctionner vaut $(1-q^{10})^N$. Ainsi

$$(1-q^{10})^N \leq 0,01 \iff N \geq \frac{\log(0,01)}{\log(1-q^{10})} \simeq 10,7.$$

Exercice 1.29. Montrer la *formule de décomposition pour les probabilités conditionnelles* : pour trois événements A, B, C tels que $\mathrm{P}(B \cap C) > 0$ et $\mathrm{P}(B \cap C^c) > 0$,

$$\mathrm{P}(A \,|\, B) = \mathrm{P}(C \,|\, B)\,\mathrm{P}(A \,|\, B \cap C) + \mathrm{P}(C^c \,|\, B)\,\mathrm{P}(A \,|\, B \cap C^c).$$

Solution 1.29. On observe que, par définition de la probabilité conditionnelle,

$$\mathrm{P}(A \,|\, B \cap C)\,\mathrm{P}(C \,|\, B) = \frac{\mathrm{P}(A \cap B \cap C)}{\mathrm{P}(B)},$$

et de même

$$\mathrm{P}(A \,|\, B \cap C^c)\,\mathrm{P}(C^c \,|\, B) = \frac{\mathrm{P}(A \cap B \cap C^c)}{\mathrm{P}(B)}.$$

En multipliant l'identité à démontrer par $\mathrm{P}(B)$, il s'ensuit que celle-ci est équivalente à

$$\mathrm{P}(A \cap B) = \mathrm{P}(A \cap B \cap C) + \mathrm{P}(A \cap B \cap C^c),$$

qui est bien vérifiée, par additivité des probabilités.

Exercice 1.30. Une urne contient initialement une boule noire et on ajoute dans l'urne des boules blanches, les unes après les autres. À chaque tour $n = 1, 2, 3, \ldots$, l'urne contient n

boules blanches et une boule noire : on tire alors une boule au hasard, on note sa couleur, et on la remet dans l'urne. On suppose que les tirages lors des différents tours sont effectués de manière indépendante.

(i) On note A_n l'événement « au n-ème tour on tombe sur la boule noire ». Calculer $\mathrm{P}(A_n)$.

(ii) On note $A = \bigcup_{n\geq 1} A_n$. Interpréter l'événement A et montrer que $\mathrm{P}(A) = 1$.

(iii) On note $B = \bigcap_{k\geq 1} \bigcup_{n\geq k} A_n$. Interpréter l'événement B et montrer que $\mathrm{P}(B) = 1$.

Solution 1.30. (i) Comme après n tours l'urne contient une seule boule noire sur un total de $n+1$ boules, la probabilité de tirer une boule noire vaut $\mathrm{P}(A_n) = \frac{1}{n+1}$.

(ii) L'événement A peut être interprété de la façon suivante : « la boule noire est tirée au moins une fois ». Si on définit $C_N := \bigcup_{n=1}^{N} A_n$, on obtient une suite croissante $(C_N)_{N\geq 1}$ d'événements, donc par continuité par le bas des probabilités (Proposition 1.24) on obtient $\mathrm{P}(A) = \mathrm{P}(\bigcup_{N\geq 1} C_N) = \lim_{N\to\infty} \mathrm{P}(C_N)$.

Il nous reste à calculer $\mathrm{P}(C_N)$. En passant au complémentaire, en utilisant l'indépendance des événements $(A_n)_{n\geq 1}$ et en observant que $\mathrm{P}(A_n^c) = 1 - \mathrm{P}(A_n) = 1 - \frac{1}{n+1} = \frac{n}{n+1}$, on obtient

$$\mathrm{P}(C_N) = 1 - \mathrm{P}\Big(\bigcap_{n=1}^{N} A_n^c\Big) = 1 - \prod_{n=1}^{N} \mathrm{P}(A_n^c) = 1 - \prod_{n=1}^{N} \frac{n}{n+1}\,.$$

Le produit est télescopique, donc $\mathrm{P}(C_N) = 1 - \frac{1}{N+1}$ et $\mathrm{P}(A) = \lim_{N\to\infty} \mathrm{P}(C_N) = 1$.

(iii) L'événement B peut être décrit de la manière suivante : pour tout k, la boule noire est tirée au moins une fois après le k-ème tour ; mais cela est équivalent à l'affirmation que « la boule noire est tirée une infinité de fois ».

Si on définit $B_k := \bigcup_{n\geq k} A_n$, on obtient une suite décroissante $(B_k)_{k\geq 1}$ d'événements, donc par continuité par le haut des probabilités (Proposition 1.24) on obtient $\mathrm{P}(B) = \mathrm{P}(\bigcap_{k\geq 1} B_k) = \lim_{k\to\infty} \mathrm{P}(B_k)$. Montrons que $\mathrm{P}(B_k) = 1$ pour tout $k \geq 1$, dont découle le résultat voulu $\mathrm{P}(B) = \lim_{k\to\infty} 1 = 1$ (on peut aussi utiliser le fait qu'une intersection dénombrable d'événements de probabilité 1 est de probabilité 1, voir l'Exercice 1.2).

Remarquons que B_k est l'union des événements croissants $C_{k,N} := \bigcup_{n=k}^{k+N} A_n$ pour $N \geq 1$, donc par continuité par le bas des probabilités on obtient $\mathrm{P}(B_k) = \lim_{N\to\infty} \mathrm{P}(C_{k,N})$. Comme on l'a vu plus haut, en passant au complémentaire et en utilisant l'indépendance des A_n, on obtient

$$\mathrm{P}(C_{k,N}) = 1 - \prod_{n=k}^{k+N} \mathrm{P}(A_n^c) = 1 - \prod_{n=k}^{k+N} \frac{n}{n+1} = 1 - \frac{k}{k+N+1}$$

car le produit est télescopique. Cela montre que, pour tout $k \geq 1$, on a $\lim_{N\to\infty} \mathrm{P}(C_{k,N}) = 1$, donc $\mathrm{P}(B_k) = 1$.

Exercice 1.31. Un étudiant doit répondre à une question à choix multiples, avec 5 réponses possibles. L'étudiant connaît (et donne) la bonne réponse avec probabilité p, et s'il ne la connaît pas, il choisit l'une des 5 réponses possibles au hasard.

(i) Quelle est la probabilité que l'étudiant donne la bonne réponse ?

(ii) Si la réponse donnée par l'étudiant était la bonne, avec quel degré de confiance la professeure estimera-t-elle que l'étudiant connaissait effectivement la bonne réponse ?

Solution 1.31. (i) On introduit les événements $A :=$ « l'étudiant donne la bonne réponse » et $B :=$ « l'étudiant connaît la bonne réponse ». Les données de l'énoncé permettent d'avoir que $\mathrm{P}(A\,|\,B) = 1$ e $\mathrm{P}(A\,|\,B^c) = \frac{1}{5}$, de plus $\mathrm{P}(B) = p$. Grâce à la formule des probabilités totales, on obtient

$$\begin{aligned}\mathrm{P}(A) &= \mathrm{P}(A|B)\,\mathrm{P}(B) + \mathrm{P}(A|B^c)\,\mathrm{P}(B^c)\\ &= p + \frac{1}{5}(1-p) = \frac{4p+1}{5}\end{aligned}$$

(ii) On doit calculer $\mathrm{P}(B\,|\,A)$. D'après la formule de Bayes :

$$\mathrm{P}(B|A) = \frac{\mathrm{P}(A\,|\,B)\,\mathrm{P}(B)}{\mathrm{P}(A)} = \frac{1\cdot p}{\frac{4p+1}{5}} = \frac{5p}{4p+1}\,.$$

Exercice 1.32. Anne habite à Paris et Odile à Lyon. Elles décident de se rencontrer à Paris. Au dernier moment, Odile (qui est quelqu'un de très indécis), décide de remettre au hasard la décision de partir, en lançant une pièce de monnaie. Ensuite, dans le cas où la pièce de monnaie l'invite à partir, elle lance un dé à six faces pour décider lequel des 6 trains à sa disposition elle va prendre. Si Anne va à la gare et s'aperçoit qu'Odile n'est dans aucun des 5 premiers trains, quelle est la probabilité qu'Odile arrive avec le dernier train ?

Solution 1.32. Introduisons les événements
T_i := « *Odile prend le i-ème train* »,
V := « *Odile part pour Paris* »,
N := « *Odile ne prend aucun des* 5 *premiers trains* ».
On doit calculer $\mathrm{P}(T_6\,|\,N)$. Par hypothèse on a

$$\mathrm{P}(V) = \frac{1}{2}\,, \qquad \mathrm{P}(T_i\,|\,V) = \frac{1}{6} \quad \text{pour tout } i = 1,\ldots,6\,,$$

donc

$$\begin{aligned}\mathrm{P}(T_6) &= \mathrm{P}(T_6 \cap V) = \mathrm{P}(T_6\,|\,V)\,\mathrm{P}(V) = \frac{1}{6}\frac{1}{2} = \frac{1}{12}\\ \mathrm{P}(N) &= \mathrm{P}(V^c) + \mathrm{P}(T_6) = \frac{1}{2} + \frac{1}{12} = \frac{7}{12}\,.\end{aligned}$$

Enfin, observons que $\mathrm{P}(N \mid T_6) = 1$, donc en appliquant la formule de Bayes on obtient

$$\mathrm{P}(T_6 \mid N) = \frac{\mathrm{P}(N \mid T_6)\,\mathrm{P}(T_6)}{\mathrm{P}(N)} = \frac{1\cdot\frac{1}{12}}{\frac{7}{12}} = \frac{1}{7}\,.$$

Exercice 1.33. Arthur et Léo sont gourmands et il leur arrive de se relever la nuit pour manger un carré de chocolat (ils sont raisonnables et n'en mangent alors qu'un seul). Cela leur arrive indépendamment l'un de l'autre, avec probabilité $2/5$ pour Arthur et avec probabilité $1/5$ pour Léo. Leur père remarque un matin qu'un carré de chocolat (et un seul) a été mangé pendant la nuit : quelle est la probabilité que cela soit Arthur qui se soit levé pendant la nuit ?

Solution 1.33. On introduit les événements A := « Arthur se lève et mange un carré de chocolat » et L := « Léo se lève et mange un carré chocolat » et notons que $\mathrm{P}(A) = \frac{2}{5}$ e $\mathrm{P}(L) = \frac{1}{5}$. Alors l'événement U := « Un (seul) carré de chocolat a été mangé » correspond à $U = (A \cap L^c) \cup (A^c \cap L)$ (autrement dit, $U = A \triangle L$). Par indépendance des événements A et L (et donc aussi de leurs complémentaires), on obtient

$$\begin{aligned}\mathrm{P}(U) &= \mathrm{P}(A\cap L^c)+\mathrm{P}(A^c\cap L) = \mathrm{P}(A)\,\mathrm{P}(L^c)+\mathrm{P}(A^c)\,\mathrm{P}(L)\\ &= \frac{2}{5}\left(1-\frac{1}{5}\right)+\frac{1}{5}\left(1-\frac{2}{5}\right) = \frac{2}{5}\frac{4}{5}+\frac{1}{5}\frac{3}{5} = \frac{11}{25}\,.\end{aligned}$$

Pour finir, observons que $A\cap U = A\cap L^c$: la réponse demandée est donc

$$\mathrm{P}(A\,|\,U) = \frac{\mathrm{P}(A\cap U)}{\mathrm{P}(U)} = \frac{\mathrm{P}(A\cap L^c)}{\mathrm{P}(U)} = \frac{\mathrm{P}(A)\,\mathrm{P}(L^c)}{\mathrm{P}(U)} = \frac{\frac{2}{5}\frac{4}{5}}{\frac{11}{25}} = \frac{8}{11} \simeq 73\%\,.$$

Exercice 1.34. Une compagnie d'assurance propose une police qui prévoit le paiement d'un montant forfaitaire suite à un dommage subit par le client. La compagnie classe ses assurés en trois catégories : *risque faible*, *risque modéré* et *risque élevé*. Parmi ses assurés, 75% sont à risque faible, 20% à risque modéré et les 5% restants à risque élevé.

On sait que les assurés à risque faible ont une probabilité de 2% de subir un dommage qui engendre le paiement de l'assurance, alors que cette probabilité est de 10% pour les assurés à risque modéré et de 20% pour ceux à risque élevé.

(i) Quelle est la probabilité qu'un assuré pris au hasard demande le paiement de l'assurance ?

(ii) Si un individu demande le paiement de l'assurance, quelle est la probabilité qu'il soit dans la catégorie à *risque élevé* ?

Solution 1.34. (i) Considérons les événements

- $A_1 =$ « l'individu choisi est de la catégorie *risque faible* » ;
- $A_2 =$ « l'individu choisi est de la catégorie *risque modéré* » ;
- $A_3 =$ « l'individu choisi est de la catégorie *risque élevé* » ;
- $B =$ « l'individu demande le paiement de l'assurance ».

On sait que

$$\begin{gathered}\mathrm{P}(B\,|\,A_1) = 0{,}02\,,\quad \mathrm{P}(B\,|\,A_2) = 0{,}1\,,\quad \mathrm{P}(B\,|\,A_3) = 0{,}2\,,\\ \mathrm{P}(A_1) = 0{,}75\,,\quad \mathrm{P}(A_2) = 0{,}2\,,\quad \mathrm{P}(A_3) = 0{,}05.\end{gathered}$$

D'après la formule des probabilités totales (Proposition 1.53)

$$\mathrm{P}(B) = \sum_{i=1}^{3}\mathrm{P}(B\,|\,A_i)\,\mathrm{P}(A_i) = 0{,}02\cdot 0{,}75+0{,}1\cdot 0{,}2+0{,}2\cdot 0{,}05 = 0{,}045.$$

(ii) En utilisant la formule de Bayes (Théorème 1.55)

$$\mathrm{P}(A_3\,|\,B) = \frac{\mathrm{P}(B\,|\,A_3)\,\mathrm{P}(A_3)}{\mathrm{P}(B)} = \frac{0{,}2\cdot 0{,}05}{0{,}045} = \frac{2}{9}.$$

Exercice 1.35. Le ministère de l'Éducation Nationale veut estimer la proportion $\alpha \in \,]0,1[$ des élèves de 4ème qui ont un niveau faible en mathématiques. Pour cela, on soumet à un grand nombre d'élèves une question à choix multiples avec 10 réponses possibles, dont une seule est correcte. On suppose que les étudiants avec un bon niveau en mathématiques répondent correctement à la question, alors que ceux avec un faible niveau donne une réponse au hasard (et il n'y a pas d'autres possibilités). Avec une analyse plus poussée, on découvre que parmi les élèves ayant répondu correctement à la question, seulement 80% ont un bon niveau en mathématiques. Avec ces informations, déterminer α.

Solution 1.35. Imaginons que l'on choisisse un élève au hasard et considérons les événements $A :=$ « l'élève a un niveau faible en mathématiques » et $B :=$ « l'élève répond correctement à la question ». Selon les données de l'énoncé, on a

$$\mathrm{P}(A) = \alpha, \qquad \mathrm{P}(B\,|\,A) = \frac{1}{10}, \qquad \mathrm{P}(B\,|\,A^c) = 1, \qquad \mathrm{P}(A^c\,|\,B) = \frac{80}{100} = \frac{4}{5}.$$

On trouve donc

$$\mathrm{P}(B) = \mathrm{P}(B\,|\,A)\,\mathrm{P}(A) + \mathrm{P}(B\,|\,A^c)\,\mathrm{P}(A^c) = \frac{\alpha}{10} + (1-\alpha) = 1 - \frac{9}{10}\alpha,$$

d'où

$$\mathrm{P}(B\,|\,A^c) = \frac{\mathrm{P}(A^c\,|\,B)\,\mathrm{P}(B)}{\mathrm{P}(A^c)} = \frac{\frac{4}{5}(1-\frac{9}{10}\alpha)}{1-\alpha} = \frac{\frac{4}{5}-\frac{18}{25}\alpha}{1-\alpha}.$$

Comme $\mathrm{P}(B\,|\,A^c) = 1$ on obtient finalement

$$\frac{\frac{4}{5}-\frac{18}{25}\alpha}{1-\alpha} = 1 \qquad \Rightarrow \qquad \alpha = \frac{5}{7} \simeq 71{,}4\%.$$

Exercice 1.36. Trois urnes, étiquetées α, β, γ, contiennent 10 boules chacune. Deux urnes contiennent 5 boules rouges et 5 bleues, alors que la troisième contient 3 boules rouges et 7 bleues. Cependant, on ne sait pas quelle est l'urne qui ne contient que 3 boules rouges. En l'absence d'information supplémentaire, on estime donc qu'il s'agit de l'urne α, β ou γ avec la même probabilité.

Maintenant, on tire deux boules de chacune des trois urnes. Si de l'urne α on a tiré une boule rouge et une bleue, de l'urne β deux boules rouges et de l'urne γ deux boules bleues, quelle est la probabilité que l'urne γ soit celle ne contenant que 3 boules rouges ?

Solution 1.36. Pour $x \in \{\alpha, \beta, \gamma\}$, soit A_x l'événement « l'urne avec 3 boules rouges est l'urne x ». Soit de plus B l'événement « de l'urne α on a tiré une boule rouge et une bleue, de l'urne β deux boules rouges et de l'urne γ deux boules bleues ». On a

$$\mathrm{P}(B\,|\,A_\alpha) = \frac{3\cdot 7}{\binom{10}{2}} \cdot \frac{\binom{5}{2}}{\binom{10}{2}} \cdot \frac{\binom{5}{2}}{\binom{10}{2}} = \frac{84}{3645}.$$

$$\mathrm{P}(B\,|\,A_\beta) = \frac{5\cdot 5}{\binom{10}{2}} \cdot \frac{\binom{3}{2}}{\binom{10}{2}} \cdot \frac{\binom{5}{2}}{\binom{10}{2}} = \frac{2}{243}.$$

$$\mathrm{P}(B\,|\,A_\gamma) = \frac{5\cdot 5}{\binom{10}{2}} \cdot \frac{\binom{5}{2}}{\binom{10}{2}} \cdot \frac{\binom{7}{2}}{\binom{10}{2}} = \frac{14}{243}.$$

$$\mathrm{P}(A_x) = \frac{1}{3}$$

pour $x \in \{\alpha, \beta, \gamma\}$. En conclusion, en utilisant la formule de Bayes,

$$\mathrm{P}(A_\gamma\,|\,B) = \frac{\mathrm{P}(B\,|\,A_\gamma)\,\mathrm{P}(A_\gamma)}{\sum_{x\in\{\alpha,\beta,\gamma\}} \mathrm{P}(B\,|\,A_x) P(A_x)} = \frac{35}{54}.$$

Exercice 1.37. Une urne contient M boules, dont M_1 blanches.

(i) On effectue n tirages successifs, *avec remise*. On considère les événements

$$B_j := \text{« la } j\text{-ème boule tirée est blanche »}$$
$$A_m := \text{« parmi les } n \text{ boules tirées, exactement } m \text{ sont blanches »},$$

pour $1 \le j \le n$, $0 \le m \le n$. Calculer $\mathrm{P}(B_j \,|\, A_m)$.

(ii) Calculer la probabilité conditionnelle de la question précédente, dans le cas d'un tirage *sans remise*, en supposant que m soit tel que $\mathrm{P}(A_m) > 0$.

Solution 1.37. Afin de simplifier la solution (mais on pourrait procéder autrement), faisons la remarque suivante, valable aussi bien pour le point (i) que pour le point (ii). Si on considère, dans l'ensemble Ω des suites possibles de boules tirées, la fonction qui échange la première boule avec la j-ème, alors cette fonction est une bijection de Ω dans lui-même, qui envoie A_m sur lui-même et B_j sur B_1, et donc envoie $B_j \cap A_m$ sur $B_1 \cap A_m$. Comme, aussi bien dans (i) que dans (ii), la probabilité sur Ω est celle uniforme, cette transformation ne change pas la probabilité des événements. En particulier $\mathrm{P}(B_j \cap A_m) = P(B_1 \cap A_m)$. Il n'est donc pas restrictif de supposer que $j = 1$.

(i) Dans le cas de tirages avec remise, on choisit comme espace d'états

$$\Omega = \{1, 2, \ldots, M\}^n,$$

muni de la probabilité P uniforme. On voit que $B_1 \cap A_m = B_1 \cap A'_m$ où $A'_m =$ « dans les $n-1$ tirages suivants on tire $m-1$ boules blanches ». Les éléments de $B_1 \cap A'_m$ sont déterminés par les choix successifs suivants :

- on choisit la boule blanche du premier tirage ;
- on choisit les $m-1$ autres tirages où l'on tire une boule blanche ;
- pour chacun des tirages choisis au point précédent on choisit la boule blanche à tirer ;
- pour chacun des $n-m$ tirages restants on choisit une boule « non blanche » à tirer.

On en déduit que

$$|B_1 \cap A_m| = M_1 \binom{n-1}{m-1} M_1^{m-1} (M - M_1)^{n-m}.$$

ainsi

$$\mathrm{P}(B_1 \cap A_m) = \frac{|B_1 \cap A_m|}{M^n} = \frac{M_1}{M} \binom{n-1}{m-1} \left(\frac{M_1}{M}\right)^{m-1} \left(1 - \frac{M_1}{M}\right)^{n-m}.$$

Comme

$$\mathrm{P}(A_m) = \binom{n}{m} \left(\frac{M_1}{M}\right)^{m} \left(1 - \frac{M_1}{M}\right)^{n-m},$$

on trouve facilement

$$\mathrm{P}(B_1 \,|\, A_m) = \frac{\mathrm{P}(B_1 \cap A_m)}{\mathrm{P}(A_m)} = \frac{\binom{n-1}{m-1}}{\binom{n}{m}} = \frac{m}{n}.$$

(ii) Notons que $\mathrm{P}(A_m \,|\, B_1)$ coïncide avec la probabilité que dans une urne contenant $M-1$ boules dont $M_1 - 1$ sont blanches on tire exactement $m-1$ boules blanches en $n-1$ tirages sans remise. Ainsi, en se rappelant de la formule (1.41) (ou par un calcul direct), on a

$$\mathrm{P}(A_m \,|\, B_1) = \frac{\binom{M_1-1}{m-1}\binom{M-M_1}{n-m}}{\binom{M-1}{n-1}}.$$

Comme, de manière analogue,

$$\mathrm{P}(A_m) = \frac{\binom{M_1}{m}\binom{M-M_1}{n-m}}{\binom{M}{n}}, \qquad \mathrm{P}(B_1) = \frac{M_1}{M},$$

avec des calculs directs on a

$$\mathrm{P}(B_1|A_m) = \frac{\mathrm{P}(A_m|B_1)\,\mathrm{P}(B_1)}{P(A_m)} = \frac{m}{n},$$

c'est-à-dire le même résultat obtenu que dans le cas de tirages avec remise.

Exercice 1.38. Juan m'affirme que dans son village, la météo suit les règles suivantes : s'il pleut une journée, il y a une chance sur deux qu'il ne pleuve pas le lendemain; s'il ne pleut pas une journée, il y a une chance sur trois qu'il pleuve le lendemain. Aujourd'hui, il pleut. On note p_n la probabilité qu'il pleuve dans n jours, pour tout $n \in \mathbb{N}$. Donner une relation de récurrence entre p_n et p_{n+1}, puis calculer p_n pour tout $n \in \mathbb{N}$ et enfin déterminer la valeur de $\lim_{n\to\infty} p_n$.

Solution 1.38. Introduisons l'événement $A_n :=$ « il pleuvra dans n jours ». Les hypothèses du problème donnent $\mathrm{P}(A_{n+1}^c \mid A_n) = \frac{1}{2}$ donc $\mathrm{P}(A_{n+1} \mid A_n) = 1 - \mathrm{P}(A_{n+1}^c \mid A_n) = \frac{1}{2}$; de plus $\mathrm{P}(A_{n+1} \mid A_n^c) = \frac{1}{3}$. Par conséquent, en posant $p_n := \mathrm{P}(A_n)$, d'après la formule des probabilités totales on obtient

$$\begin{aligned} p_{n+1} &= \mathrm{P}(A_{n+1}) = \mathrm{P}(A_{n+1} \mid A_n)\,\mathrm{P}(A_n) + \mathrm{P}(A_{n+1} \mid A_n^c)\,\mathrm{P}(A_n^c) \\ &= \frac{1}{2}p_n + \frac{1}{3}(1-p_n) = \frac{1}{3} + \frac{1}{6}p_n\,. \end{aligned}$$

Écrivons cette relation pour p_n, en répétant trois étapes :

$$\begin{aligned} p_n &= \frac{1}{3} + \frac{1}{6}p_{n-1}\,, \\ p_n &= \frac{1}{3} + \frac{1}{6}\left(\frac{1}{3} + \frac{1}{6}p_{n-2}\right) = \frac{1}{3} + \frac{1}{3}\frac{1}{6} + \frac{1}{6^2}p_{n-2}\,, \\ p_n &= \frac{1}{3} + \frac{1}{3}\frac{1}{6} + \frac{1}{6^2}\left(\frac{1}{3} + \frac{1}{6}p_{n-3}\right) = \frac{1}{3} + \frac{1}{3}\frac{1}{6} + \frac{1}{3}\frac{1}{6^2} + \frac{1}{6^3}p_{n-3}\,. \end{aligned}$$

En répétant cette formule pour n étapes, on obtient

$$p_n = \frac{1}{3}\left(1 + \frac{1}{6} + \frac{1}{6^2} + \cdots + \frac{1}{6^{n-1}}\right) + \frac{1}{6^n}p_0 = \frac{1}{3}\,\frac{1-\frac{1}{6^n}}{1-\frac{1}{6}} + \frac{1}{6^n}p_0 = \frac{2}{5}\left(1 - \frac{1}{6^n}\right) + \frac{1}{6^n}p_0\,,$$

où on a utilisé la somme partielle de la série géométrique, voir (0.4). Comme par hypothèse il pleut aujourd'hui, on a $p_0 = 1$ et cette formule donne la valeur de p_n pour tout n. Indépendamment de la valeur de p_0, cette formule montre que

$$\lim_{n\to\infty} p_n = \frac{2}{5}\,.$$

Il est intéressant d'observer que l'on peut calculer la limite de la suite p_n sans utiliser la formule exacte, *si l'on suppose que la limite* $\bar{p} := \lim_{n\to\infty} p_n$ *existe* : il suffit d'utiliser la relation de récurrence que l'on a obtenue

$$p_{n+1} = \frac{1}{3} + \frac{1}{6} p_n \,.$$

En effet, en prenant la limite $n \to \infty$ des deux côtés de l'égalité, et en remarquant que $\bar{p} := \lim_{n\to\infty} p_n = \lim_{n\to\infty} p_{n+1}$, on obtient

$$\bar{p} = \frac{1}{3} + \frac{1}{6}\bar{p} \qquad \text{c'est-à-dire} \qquad \left(1 - \frac{1}{6}\right)\bar{p} = \frac{1}{3} \qquad \text{donc} \qquad \bar{p} = \frac{2}{5} \,.$$

Exercice 1.39. On a deux dés réguliers : le dé α possède six faces, numérotées de 1 à 6, et le dé β possède douze faces, numérotées de 1 à 12. On choisit un des deux dés au hasard (avec la même probabilité) et on le lance n fois, où $n \geq 1$ est un nombre fixé.

(i) Quelle est la probabilité que tous les lancers donnent le chiffre 3 ?

(ii) Quelle est la probabilité que tous les lancers donnent le même numéro ?

(iii) Si tous les lancers donnent le chiffre 3, quelle est la probabilité que le dé choisi soit le dé α ? Montrer qu'une telle probabilité (conditionnelle) est toujours strictement plus grande que $1/2$, et en étudier le comportement quand $n \to \infty$.

Solution 1.39. (i) Introduisons les événements

- $A :=$ « le dé choisi est α » ; $A^c =$ « le dé choisi est β »
- $C :=$ « tous les lancers donnent pour résultat le chiffre 3 ».

Alors $\mathrm{P}(A) = \frac{1}{2}$, et de plus $\mathrm{P}(C\,|\,A) = (\frac{1}{6})^n$ et $\mathrm{P}(C\,|\,A^c) = (\frac{1}{12})^n$, dont on déduit

$$\begin{aligned}\mathrm{P}(C) \;&=\; \mathrm{P}(C\,|\,A)\,\mathrm{P}(A) + \mathrm{P}(C\,|\,A^c)\,\mathrm{P}(A^c) \;=\; \left(\frac{1}{6}\right)^n \frac{1}{2} + \left(\frac{1}{12}\right)^n \frac{1}{2} \\ &= \frac{1}{2}\left\{\left(\frac{1}{6}\right)^n + \left(\frac{1}{12}\right)^n\right\}.\end{aligned}$$

(ii) Introduisons maintenant l'événement $D :=$ « tous les lancers donnent le même numéro ». Alors $\mathrm{P}(D\,|\,A) = 6(\frac{1}{6})^n$ et $\mathrm{P}(D\,|\,A^c) = 12(\frac{1}{12})^n$, donc

$$\mathrm{P}(D) \;=\; \mathrm{P}(D\,|\,A)\,\mathrm{P}(A) + \mathrm{P}(D\,|\,A^c)\,\mathrm{P}(A^c) \;=\; \frac{1}{2}\left\{\left(\frac{1}{6}\right)^{n-1} + \left(\frac{1}{12}\right)^{n-1}\right\}.$$

(iii) D'après la formule de Bayes,

$$\mathrm{P}(A\,|\,C) \;=\; \frac{\mathrm{P}(C\,|\,A)\,\mathrm{P}(A)}{\mathrm{P}(C)} \;=\; \frac{(\frac{1}{6})^n}{(\frac{1}{6})^n + (\frac{1}{12})^n} \;=\; \frac{1}{1 + (\frac{1}{2})^n} \;=\; \frac{2^n}{2^n + 1} \,.$$

Comme $2^n > 1$ pour tout $n \geq 1$ et que $\lim_{n\to\infty} (\frac{1}{2})^n = 0$, on en déduit que $\mathrm{P}(A\,|\,C) > \frac{1}{2}$ et que $\mathrm{P}(A\,|\,C) \to 1$ quand $n \to \infty$.

Exercice 1.40. (i) Montrer que, si $\alpha_1, \alpha_2, \ldots, \alpha_n \geq 0$ et $\sum_{i=1}^n \alpha_i = 1$, alors, pour tout choix de $x_1, x_2, \ldots, x_n \in \mathbb{R}$ on a

$$\min_{i=1,2,\ldots,n} x_i \le \sum_{i=1}^{n} \alpha_i x_i \le \max_{i=1,2,\ldots,n} x_i.$$

(ii) Soient $A_1, A_2, \ldots, A_n$ des événements *disjoints* d'un espace probabilisé (Ω, P), tels que $\mathrm{P}(A_i) > 0$ pour tout $i = 1, \ldots, n$. Montrer que, pour tout événement B

$$\min_{i=1,2,\ldots,n} \mathrm{P}(B \,|\, A_i) \le \mathrm{P}(B \,|\, A_1 \cup A_2 \cup \cdots \cup A_n) \le \max_{i=1,2,\ldots,n} \mathrm{P}(B \,|\, A_i).$$

Solution 1.40. (i) Soit $m := \min_{i=1,2,\ldots n} x_i$.

$$\sum_{i=1}^{n} \alpha_i x_i \ge \sum_{i=1}^{n} \alpha_i m = m.$$

L'autre inégalité s'obtient de manière analogue.

(ii) On observe que

$$\begin{aligned} \mathrm{P}(B \,|\, A_1 \cup A_2 \cup \cdots \cup A_n) &= \frac{\mathrm{P}(B \cap (A_1 \cup A_2 \cup \cdots \cup A_n))}{\mathrm{P}(A_1 \cup A_2 \cup \cdots \cup A_n)} \\ &= \frac{\sum_{i=1}^{n} \mathrm{P}(B \cap A_i)}{\sum_{j=1}^{n} \mathrm{P}(A_j)} = \sum_{i=1}^{n} \alpha_i \mathrm{P}(B \,|\, A_i), \end{aligned}$$

où

$$\alpha_i := \frac{\mathrm{P}(A_i)}{\sum_{j=1}^{n} \mathrm{P}(A_j)}.$$

La conclusion découle alors du point précédent.

Exercice 1.41. Soient A et B deux événements de probabilité non nulle. On dit que A *est corrélé positivement à* B si $\mathrm{P}(A \,|\, B) \ge \mathrm{P}(A)$.

(i) Montrer que les trois affirmations suivantes sont équivalentes :

(a) *A est corrélé positivement à B ;*

(b) *B est corrélé positivement à A ;*

(c) *A^c est corrélé positivement à B^c.*

(ii) Donner un exemple de trois événements A, B, C d'un espace probabilisé bien choisi, qui vérifient : $\mathrm{P}(A \,|\, B) > \mathrm{P}(A)$, $\mathrm{P}(B \,|\, C) > \mathrm{P}(B)$ mais $\mathrm{P}(A \,|\, C) < \mathrm{P}(A)$. Autrement dit, A est corrélé positivement (strictement) à B, B est corrélé positivement (strictement) à C, mais A n'est pas corrélé positivement à C.

[*Sugg.* Prendre $\Omega = \{1, 2, 3\}$ avec la probabilité P uniforme sur Ω.]

Solution 1.41. (i) On observe que

$$\mathrm{P}(A \,|\, B) \ge \mathrm{P}(A) \iff \mathrm{P}(A \cap B) \ge \mathrm{P}(A)\,\mathrm{P}(B) \iff \mathrm{P}(B \,|\, A) \ge \mathrm{P}(B),$$

qui démontre (a) $\iff$ (b). De plus

$$\begin{aligned} \mathrm{P}(A|B) \ge \mathrm{P}(A) &\iff \mathrm{P}(A \cap B) \ge \mathrm{P}(A)\,\mathrm{P}(B) \\ &\iff \mathrm{P}(A) - \mathrm{P}(A \cap B^c) \ge \mathrm{P}(A)[1 - \mathrm{P}(B^c)] \iff \mathrm{P}(A \cap B^c) \le \mathrm{P}(A)\,\mathrm{P}(B^c). \end{aligned}$$

En répétant le même argument, on trouve

$$\mathrm{P}(A \cap B^c) \le \mathrm{P}(A)\,\mathrm{P}(B^c) \iff \mathrm{P}(A^c \cap B^c) \ge \mathrm{P}(A^c)\,\mathrm{P}(B^c).$$

Avec ces ceux dernières équivalences, on conclut que (a) $\iff$ (c).

(ii) Soit $\Omega = \{1,2,3\}$ muni de la probabilité P uniforme. On considère les événements $A = \{1,2\}$, $B = \{2\}$ et $C = \{2,3\}$: on a

$$\begin{aligned}
\mathrm{P}(A \mid B) &= \frac{|A \cap B|}{|B|} = 1 > \mathrm{P}(A) = \frac{2}{3}\,,\\
\mathrm{P}(B \mid C) &= \frac{|B \cap C|}{|C|} = \frac{1}{2} > \mathrm{P}(B) = \frac{1}{3}\,,\\
\mathrm{P}(A \mid C) &= \frac{|A \cap C|}{|C|} = \frac{1}{2} < \mathrm{P}(A) = \frac{2}{3}\,.
\end{aligned}$$

Exercice 1.42. Une urne contient n boules, qui peuvent être de deux couleurs, rouge ou verte. On ne connaît pas la composition de l'urne et on estime que toutes les valeurs possibles $k = 0,1,2,\ldots,n$ du nombre de boules rouges sont équiprobables.

(i) On tire une boule de l'urne et celle-ci s'avère être rouge. En sachant cela, pour quelle valeur de k la probabilité qu'il y avait k boules rouges dans l'urne est maximisée ?

(ii) Répondre à la même question, mais en supposant que l'on ait tiré (sans remise) deux boules dans l'urne, une rouge et une verte.

Solution 1.42. (i) On définit les événements $A =$ « la boule tirée est rouge » et $B_k =$ « l'urne contient k boules rouges ». On sait que $\mathrm{P}(A \,|\, B_k) = \frac{k}{n}$ et $\mathrm{P}(B_k) = \frac{1}{n+1}$. ainsi, d'après la formule de Bayes,

$$\mathrm{P}(B_k \,|\, A) = \frac{\mathrm{P}(A \,|\, B_k)\,\mathrm{P}(B_k)}{\sum_{j=0}^{n} \mathrm{P}(A \,|\, B_j)\,\mathrm{P}(B_j)} = \frac{\frac{k}{n}\frac{1}{n+1}}{\frac{1}{n(n+1)}\sum_{j=0}^{n} j} = \frac{2k}{n(n+1)},$$

où on a utilisé la première formule dans (0.8), c'est-à-dire $\sum_{j=0}^{n} j = \frac{n(n+1)}{2}$. Cette probabilité est clairement croissante en k et est donc maximisée pour $k = n$.

(ii) Soit maintenant $C =$ « les deux boules tirées sont une rouge et une verte ». Cette fois, on a

$$\mathrm{P}(C \,|\, B_k) = \frac{k(n-k)}{\binom{n}{2}}.$$

D'après la formule de Bayes on a

$$\mathrm{P}(B_k \,|\, C) = \frac{\mathrm{P}(C \,|\, B_k)\,\mathrm{P}(B_k)}{\sum_{j=0}^{n} \mathrm{P}(C \,|\, B_j)\,\mathrm{P}(B_j)}\,,$$

et en observant qu'aussi bien $\mathrm{P}(B_k)$ que $\sum_{j=0}^{n} \mathrm{P}(C \,|\, B_j)\,\mathrm{P}(B_j)$ ne dépendent pas de k, on obtient

$$\mathrm{P}(B_k \,|\, C) = \frac{k(n-k)}{Z},$$

où Z est une constante qui ne dépend pas de k. Ainsi, maximiser $\mathrm{P}(B_k \,|\, C)$ équivaut à maximiser $k(n-k)$ pour $k \in \{0,1,\ldots,n\}$. Si n est pair, $k = \frac{n}{2}$ est le seul point de maximum, mais si n est impair il y a deux points de maximum : $k = \frac{n\pm 1}{2}$.

Exercice 1.43. On considère le modèle suivant de répartition des enfants dans une famille. La probabilité qu'une famille choisie au hasard ait n enfants ($n \in \mathbb{N}$) vaut $\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!}$ (où $\lambda > 0$ est un paramètre fixé). De plus, chacun des enfants est une fille avec probabilité $1/2$, indépendamment de tous les autres enfants. On considère l'événement

$$A_k := \text{« la famille choisie (au hasard) a exactement } k \text{ filles »},$$

pour $k \geq 0$. Montrer que $\mathrm{P}(A_k) = \mathrm{e}^{-\lambda/2} \frac{(\lambda/2)^k}{k!}$.
[*Sugg.* Considérer l'événement $B_n =$ « la famille choisie a n enfants ». Déterminer tout d'abord $\mathrm{P}(A_k \,|\, B_n)$, puis calculer $\mathrm{P}(A_k)$. On pourra utiliser la série exponentielle (0.6).]

Solution 1.43. En se rappelant de la formule (1.69) relative à la répétition de n épreuves indépendantes de probabilité de succès $p = \frac{1}{2}$, il découle des hypothèses du modèle que, si $n \geq k$,

$$\mathrm{P}(A_k \,|\, B_n) = \frac{1}{2^n} \binom{n}{k}.$$

Évidemment, $\mathrm{P}(A_k \,|\, B_n) = 0$ si $k > n$. Ainsi, en appliquant la formule des probabilités totales (Proposition 1.53), on obtient

$$\begin{aligned}
\mathrm{P}(A_k) &= \sum_{n=0}^{+\infty} \mathrm{P}(A_k \,|\, B_n)\mathrm{P}(B_n) = \sum_{n=k}^{+\infty} \frac{1}{2^n} \binom{n}{k} e^{-\lambda} \frac{\lambda^n}{n!} \\
&= e^{-\lambda} \frac{1}{2^k} \frac{\lambda^k}{k!} \sum_{n=k}^{+\infty} \frac{1}{2^{n-k}} \frac{\lambda^{n-k}}{(n-k)!} = e^{-\lambda} \frac{(\lambda/2)^k}{k!} \sum_{m=0}^{+\infty} \frac{(\lambda/2)^m}{m!} \\
&= e^{-\lambda} \frac{(\lambda/2)^k}{k!} e^{\lambda/2} = e^{-\lambda/2} \frac{(\lambda/2)^k}{k!}.
\end{aligned}$$

Exercice 1.44. Soit $S = \{1, 2, \ldots, n\}$, $\Omega := \mathcal{P}(S) \times \mathcal{P}(S)$, et P la probabilité uniforme sur Ω. Ainsi, les éléments de Ω sont des couples (A, B), où A, B sont des sous-ensembles de S. On considère l'événement

$$E := \{(A, B) \in \Omega : A \subseteq B\}.$$

En outre, pour $B \subseteq S$, on définit $F_B := \{(A', B') \in \Omega :\ B' = B\} = \{(A, B) :\ A \subseteq S\}$.

(i) Déterminer $\mathrm{P}(E \,|\, F_B)$.

(ii) En utilisant la formule de décomposition $\mathrm{P}(E) = \sum_{B \subseteq S} \mathrm{P}(F_B)\mathrm{P}(E \,|\, F_B)$, montrer que $\mathrm{P}(E) = (\frac{3}{4})^n$.

[*Sugg.* Utiliser le binôme de Newton (1.39) et le fait que $|\mathcal{P}(S)| = 2^{|S|}$.]

Solution 1.44. (i) Tout d'abord,

$$\mathrm{P}(E \,|\, F_B) = \frac{|E \cap F_B|}{|F_B|}.$$

Les éléments de F_B sont au même nombre que les sous-ensemble de S, c'est-à-dire 2^n. Les éléments de $E \cap F_B$ sont au même nombre que les sous-ensemble de B, c'est-à-dire $2^{|B|}$. Donc

$$\mathrm{P}(E \,|\, F_B) = \frac{2^{|B|}}{2^n}.$$

(ii) Comme $P(F_B) = |F_B|/|\Omega| = 1/2^n$, on a

$$\mathrm{P}(E) = \sum_{B \subseteq S} \frac{2^{|B|}}{4^n} = \frac{1}{4^n} \sum_{k=0}^{n} \sum_{B \subseteq S : |B| = k} 2^k = \frac{1}{4^n} \sum_{k=0}^{n} \binom{n}{k} 2^k = \frac{3^n}{4^n},$$

où on a utilisé le fait que $\sum_{k=0}^{n}\binom{n}{k}2^k = (1+2)^n$.

Exercice 1.45. On effectue un référendum dans une population de $n \geq 1$ individus (ayant tous le droit de vote). Les individus vont voter avec probabilité $\frac{1}{2}$, indépendamment les uns des autres. De plus, si un individu va voter, il votera OUI avec probabilité $\frac{1}{2}$, indépendamment des autres.

(i) Quelle est la probabilité p qu'un individu choisi au hasard aille voter OUI ?

(ii) Quelle est la probabilité que le nombre de votes OUI soit k, pour $k \in \{0,\dots,n\}$?

(iii) En supposant que l'on ait k votes OUI, déterminer la probabilité (conditionnelle) que le nombre de votants soit m, pour $m \in \{k,\dots,n\}$. Montrer que cette probabilité vaut

$$\binom{n-k}{m-k}\left(\frac{1}{3}\right)^{m-k}\left(\frac{2}{3}\right)^{n-m}.$$

Solution 1.45. (i) Considérons les événements $A =$ « l'individu va voter » et $B =$ « l'individu vote OUI ». Alors

$$p = \mathrm{P}(A \cap B) = \mathrm{P}(A)\cdot\mathrm{P}(B\,|\,A) = \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}.$$

(ii) On doit calculer la probabilité de l'événement $C =$ « exactement k individus vont voter OUI ». Comme chaque individu, indépendamment des autres, va voter OUI avec probabilité $p = \frac{1}{4}$, on est en présence d'un schéma de répétition d'épreuves indépendantes : on a

$$\mathrm{P}(C) = \binom{n}{k}\left(\frac{1}{4}\right)^{k}\left(\frac{3}{4}\right)^{n-k}.$$

(iii) Introduisons l'événement $D =$ « exactement m individus vont voter ». On doit calculer $\mathrm{P}(D\,|\,C)$. En appliquant la formule de Bayes

$$\mathrm{P}(D\,|\,C) = \frac{\mathrm{P}(C\,|\,D)\cdot\mathrm{P}(D)}{\mathrm{P}(C)}.$$

En appliquant le schéma de répétition d'épreuves indépendantes, on a

$$\mathrm{P}(D) = \binom{n}{m}\left(\frac{1}{2}\right)^{n}, \qquad \mathrm{P}(C|D) = \binom{m}{k}\left(\frac{1}{2}\right)^{m},$$

alors que $\mathrm{P}(C)$ a été déterminée dans le point précédent. En remplaçant et en faisant les simplifications opportunes, on trouve

$$\mathrm{P}(D\,|\,C) = \frac{\binom{m}{k}\binom{n}{m}}{\binom{n}{k}}\cdot\frac{(\frac{1}{2})^m(\frac{1}{2})^n}{(\frac{1}{3})^k(\frac{3}{4})^{n-k}} = \binom{n-k}{m-k}\left(\frac{1}{3}\right)^{m-k}\left(\frac{2}{3}\right)^{n-m}.$$

Exercice 1.46. On possède une urne initialement vide et des boules numérotées par $\mathbb{N}^*$. Le premier jour, on place dans l'urne les boules n^os^ 1 et 2, après quoi on en enlève une *au hasard* (il reste donc dans l'urne une seule boule). Le deuxième jour, on place dans l'urne les boules n^os^ 3 et 4, après quoi on enlève *au hasard* une boule parmi les trois contenues dans l'urne. On répète la procédure : le i-ème jour, on place dans l'urne les boules n^os^ $2i-1$ et $2i$, après quoi on enlève *au hasard* une boule parmi les $i+1$ boules contenues dans l'urne.

On introduit, pour $i \geq 1$, l'événement

$$A_n := \text{« la boule n}^\text{o}\text{ 1 est présente dans l'urne à la fin du } n\text{-ème jour »}.$$

(i) Justifier que l'on a $A_{i+1} \subseteq A_i$ pour tout $i \in \mathbb{N}^*$, et en déduire la formule

$$\mathrm{P}(A_n) \;=\; \mathrm{P}(A_1) \cdot \mathrm{P}(A_2 \,|\, A_1) \cdots \mathrm{P}(A_n \,|\, A_{n-1}), \qquad n \in \mathbb{N}^*.$$

(ii) Montrer que $\mathrm{P}(A_n) = \frac{1}{n+1}$ pour tout $n \in \mathbb{N}^*$.

(iii) On considère l'événement $A = \bigcap_{n \geq 1} A_n$. Comment décririez-vous cet événement avec des mots ? Que vaut $\mathrm{P}(A)$?

Solution 1.46. (i) Si la boule n° 1 est tirée un certain jour i, celle-ci ne sera pas présente le jour suivant $i+1$; en passant au complémentaire, si la boule n° 1 est présente dans l'urne au jour $i+1$, celle-ci doit être présente aussi au jour i. On a donc l'inclusion d'événements $A_{i+1} \subseteq A_i$. Par conséquent, on peut écrire $A_n = A_n \cap A_{n-1} = A_n \cap \cdots \cap A_1$, dont par la formule des probabilités conditionnelles en chaîne, on a

$$\mathrm{P}(A_n) = \mathrm{P}(A_1) \cdot \mathrm{P}(A_2 \,|\, A_1) \cdots \mathrm{P}(A_n \,|\, A_{n-1}) = \mathrm{P}(A_1) \cdot \prod_{k=2}^{n} \mathrm{P}(A_k \,|\, A_{k-1}).$$

(ii) Pour tout $k \geq 2$ on a $\mathrm{P}(A_k \,|\, A_{k-1}) = \frac{k}{k+1}$: en effet, si l'événement A_{k-1} est vérifié, la boule n° 1 est présente dans l'urne à la fin du $(k-1)$-ème jour, l'urne contenant au total $k-1$ boules ; le jour suivant on ajoute deux boules, donc l'urne en contient $(k-1)+2 = k+1$ et la probabilité que la boule n° 1 ne soit pas tirée est bien $\frac{k}{k+1}$. En appliquant la formule trouvée dans le point précédent, on obtient

$$\mathrm{P}(A_n) = \prod_{k=1}^{n} \left(1 - \frac{1}{k+1}\right) = \prod_{k=1}^{n} \frac{k}{k+1} = \frac{1}{n+1}.$$

(iii) On peut interpréter A comme l'événement « la boule n° 1 n'est jamais tirée ». Les événements A_n étant décroissants, on a $\mathrm{P}(A) = \lim_{n\to\infty} \mathrm{P}(A_n) = 0$.

Exercices plus difficiles

Exercice 1.47. On reprend l'Exercice 1.46. Pour tout $k \geq 1$ fixé, on considère l'événement

$$A_n^{(k)} := \text{« la boule n}^\text{o}\ k \text{ est présente dans l'urne à la fin du } n\text{-ème jour »}.$$

(i) Montrer que $\mathrm{P}(A_n^{(k)}) = 0$ si $n < \lfloor \frac{k+1}{2} \rfloor$ et que $\mathrm{P}(A_n^{(k)}) = \frac{\lfloor \frac{k+1}{2} \rfloor}{n+1}$ pour tout $n \geq \lfloor \frac{k+1}{2} \rfloor$.

(ii) On considère l'événement $A^{(k)} = \bigcap_{n \geq \lfloor \frac{k+1}{2} \rfloor} A_n^{(k)}$. Comment décririez-vous cet événement avec des mots ? Que vaut $\mathrm{P}(A^{(k)})$?

(iii) Montrer que $\mathrm{P}\left(\bigcup_{k=1}^{\infty} A^{(k)}\right) = 0$. Interprétez ce résultat.

Solution 1.47. (i) Rappelons que les boules n$^\text{os}$ 1 et 2 sont placées dans l'urne le premier jour, les boules n$^\text{os}$ 3 et 4 sont placées dans l'urne le deuxième jour, et plus généralement les boules n$^\text{os}$ $2i-1$ et $2i$ sont placées dans l'urne le i-ème jour. Cela signifie que la boule n° k est placée dans l'urne le jour $\bar{n}_k := \lfloor \frac{k+1}{2} \rfloor$, donc pour $n < \bar{n}_k$ l'événement $A_n^{(k)}$ ne peut pas être vérifié et sa probabilité est nulle.

Considérons maintenant $n \geq \bar{n}_k$. Pour que l'événement $A_n^{(k)}$ soit vérifié, il faut que la boule nº k ne soit pas tirée de l'urne pour aucun jour i pour i allant de $i = \bar{n}_k$ jusqu'à $i = n$. Comme la probabilité que la boule nº k ne soit pas tirée de l'urne le jour i vaut $\frac{i}{i+1}$, on obtient

$$\mathrm{P}(A_n^{(k)}) = \prod_{i=\bar{n}_k}^{n} \Big(1 - \frac{1}{i+1}\Big) = \prod_{i=\lfloor \frac{k+1}{2} \rfloor}^{n} \frac{i}{i+1} = \frac{\lfloor \frac{k+1}{2} \rfloor}{n+1},$$

parce que le produit est télescopique.

(ii) L'événement $A^{(k)}$ peut être décrit comme « la boule nº k n'est jamais tirée ». Comme les événements $A_n^{(k)}$ pour $n \geq \bar{n}_k = \lfloor \frac{k+1}{2} \rfloor$ sont décroissants, on obtient par continuité par le haut des probabilités que $\mathrm{P}(A^{(k)}) = \lim_{n\to\infty} \mathrm{P}(A_n^{(k)}) = 0$.

(iii) Par sous-additivité dénombrable des probabilités on a

$$\mathbb{P}\Big(\bigcup_{k\geq 1} A^{(k)}\Big) \leq \sum_{k\geq 1} \mathbb{P}(A^{(k)}) = 0\,.$$

On peut interpréter l'événement $\bigcup_{k\geq 1} A^{(k)}$ comme « il existe un k tel que la boule nº k n'est jamais tirée », ou de manière équivalente « il y a une boule qui n'est jamais tirée ». Comme cet événement est de probabilité nulle, nous avons montré qu'*avec probabilité* 1*, chaque boule finit par être tirée à un moment où à un autre*. Autrement dit, même si le nombre de boules dans l'urne augmente tous les jours, aucune boule ne reste pour toujours dans l'urne.

Exercice 1.48. On place m boules *indistinguables* dans n boîtes de manière aléatoire. On note k_i le nombre de boules contenues dans la i-ème boîte et on décrit l'expérience aléatoire avec l'espace d'états

$$\Omega_{n,m} = \big\{(k_1,\dots,k_n) \in \mathbb{N}^n\,, k_1 + \dots + k_n = m\big\}\,,$$

que l'on munit de la probabilité uniforme P.

(i) On définit l'ensemble

$$\Omega'_{n,m} = \big\{(i_1,\dots,i_{n-1})\,,\, 1 \leq i_1 < \dots < i_{n-1} \leq m+n-1\big\}\,.$$

Montrer que l'application $\varphi : \Omega_{n,m} \to \Omega'_{n,m}$ définie par

$$\varphi(k_1,\dots,k_n) = \big(k_1+1, k_1+k_2+2, \dots, k_1+\dots+k_{n-1}+n-1\big)$$

est une bijection. En déduire que $|\Omega_{n,m}| = \binom{n+m-1}{n-1}$.

[*Sugg.* Observer que $\Omega'_{n,m}$ est en bijection avec la famille des sous-ensembles $A \subseteq \{1,\dots,n+m-1\}$ avec $|A| = n-1$.]

(ii) Soit $p_{n,m}$ la probabilité qu'il n'y ait aucune boîte vide. Montrer que

$$p_{n,m} = \prod_{j=1}^{n-1} \Big(1 - \frac{n}{m+j}\Big)\,.$$

(iii) On suppose que $m = m(n)$ dépend de n de telle sorte que l'on ait $\lim_{n\to\infty} m/(n)n^2 = x$ avec $x > 0$. Calculer la limite $\lim_{n\to\infty} p_{n,m(n)}$.

Solution 1.48. (i) Comme pour tout $(k_1,\dots,k_n) \in \Omega_{n,m}$ on a $k_i \geq 0$ pour tout $i = 1,\dots,n$ et $k_1+\cdots+k_n = m$, il s'ensuit que $k_1+\cdots+k_{n-1} \leq m$ et donc

$$1 \leq k_1+1 < k_1+k_2+2 < \cdots < k_1+\cdots+k_{n-1}+n-1 \leq m+n-1\,.$$

Cela montre que φ est effectivement une application de $\Omega_{n,m}$ dans $\Omega'_{n,m}$.
On montre maintenant que φ est bijective, en observant que sa fonction inverse est donnée par

$$\psi((i_1,\dots,i_{n-1})) := (i_1-1, i_2-i_1-1,\dots,i_{n-1}-i_{n-2}-1, n-i_{n-1}-1)\,,$$

c'est-à-dire que l'on a $\psi(\varphi((k_1,\dots,k_n))) = (k_1,\dots,k_n)$ pour tout $(k_1,\dots,k_n) \in \Omega_{n,m}$, comme on le vérifie facilement d'après la définition de φ. Cela montre que φ est bijective, de fonction inverse $\varphi^{-1} = \psi$.
Comme tout ensemble $(i_1,\dots,i_{n-1}) \in \Omega'_{n,m}$ peut être identifié avec le sous-ensemble $\{i_1,\dots,i_{n-1}\} \subseteq \{1,\dots,m+n-1\}$, on en déduit que $|\Omega_{n,m}| = |\Omega'_{n,m}|$ est égal au nombre de sous-ensembles à $n-1$ éléments de l'ensemble $\{1,\dots,m+n-1\}$, c'est-à-dire $\binom{m+n-1}{n-1}$.

(ii) L'événement $A :=$ « aucune boîte n'est vide » correspond au sous-ensemble

$$A = \{(k_1,\dots,k_n) \in \Omega_{n,m} :\ k_i \geq 1\ \forall i = 1,\dots,n\}\,.$$

Pour tout élément $(k_1,\dots,k_n) \in A$, si on pose $k'_i := k_i - 1$, on obtient un élément $(k'_1,\dots,k'_n) \in \Omega_{n,m-n}$ (notons que l'on doit avoir $m \geq n$ pour que l'événement A puisse être vérifié) ; réciproquement, tout élément $(k'_1,\dots,k'_n) \in \Omega_{n,m-n}$ détermine un élément $(k_1,\dots,k_n) \in A$ en posant $k_i := k'_i + 1$. Cela montre que l'ensemble A est en bijection avec $\Omega_{n,m-n}$, donc $|A| = |\Omega_{n,m-n}| = \binom{n+(m-n)-1}{n-1} = \binom{m-1}{n-1}$ et on obtient la formule voulue pour $p_{n,m} = \mathrm{P}(A)$:

$$p_{n,m} = \frac{|A|}{|\Omega_{n,m}|} = \frac{\binom{m-1}{n-1}}{\binom{n+m-1}{n-1}} = \frac{\frac{(m-1)!}{(m-n)!}}{\frac{(n+m-1)!}{m!}} = \frac{\prod_{j=1}^{n-1}(m-n+j)}{\prod_{j=1}^{n-1}(m+j)} = \prod_{j=1}^{n-1}\left(1-\frac{n}{m+j}\right).$$

(iii) Soit maintenant $m = m(n) \sim xn^2$ avec $x > 0$. Dans ce cas, uniformément pour $j \geq 1$, on a $\frac{n}{m+j} \leq \frac{n}{xn^2}$ et donc $(1-\frac{n}{m+j}) = (1-O(\frac{1}{n}))$ pour $n \to \infty$ avec $x > 0$ fixé. En se rappelant que $\log(1-u) = -u + O(u^2)$ pour $u \to 0$, on obtient

$$\log p_{n,m} = \sum_{j=1}^{n-1} \log\left(1-\frac{n}{m+j}\right) = \sum_{j=1}^{n-1}\left(-\frac{n}{m+j} + O\left(\frac{1}{n^2}\right)\right),$$

pour $n \to \infty$ avec $x > 0$ fixé. La contribution du deuxième terme est $(n-1)O(\frac{1}{n^2}) = O(\frac{1}{n}) = o(1)$, donc on peut écrire, pour $m = m(n) \sim xn^2$

$$\log p_{n,m(n)} = -n\sum_{j=1}^{n-1}\frac{1}{m+j} + o(1) \sim -n\sum_{j=1}^{n-1}\frac{1}{xn^2+j} + o(1)\,.$$

Pour $1 \leq j \leq n-1$ on peut estimer $\frac{1}{xn^2+n} \leq \frac{1}{xn^2+j} \leq \frac{1}{xn^2}$. Comme $\frac{1}{xn^2+n} \sim \frac{1}{xn^2}$ pour $n \to \infty$ avec $x > 0$ fixé, on obtient

$$\log p_{n,m(n)} \sim -n \sum_{j=1}^{n-1} \frac{1}{xn^2} + o(1) = -\frac{n(n-1)}{xn^2} + o(1) \xrightarrow[n\to\infty]{} -\frac{1}{x}\,.$$

Au final, on a montré que pour $m = m(n) \sim xn^2$ avec $x > 0$ fixé, on a

$$\lim_{n\to\infty} p_{n,m(n)} = \mathrm{e}^{-1/x}\,.$$

Exercice 1.49 (Problème du scrutin). Après une élection, une urne contient $n+m$ bulletins de vote, dont n sont pour la candidate A et m pour le candidat B. On ouvre les bulletins les uns après les autres et on cherche à calculer la probabilité que la candidate A soit en tête tout au long du dépouillement.

Pour décrire cette expérience aléatoire, on introduit l'espace d'états

$$\Omega_{n,m} = \left\{ \omega = (\omega_1, \dots, \omega_{n+m}) \in \{0,1\}^{n+m}, \sum_{i=1}^{n+m} \omega_i = n \right\},$$

où $\omega_i = 1$ si le i-ème bulletin est pour A et $\omega_i = 0$ si le i-ème bulletin est pour B. On munit $\Omega_{n,m}$ de la probabilité uniforme P.

(i) On définit $x_j = \sum_{i=1}^{j} \omega_i$ et $y_j = \sum_{i=1}^{j}(1-\omega_i)$ pour $\omega \in \Omega_{n,m}$ et $j \in \{1,2,\dots,n+m\}$. On observe que x_j et y_j représentent les nombres de votes obtenus respectivement par la candidate A et le candidat B après avoir dépouillé j bulletins.
On définit $s_j := (x_j, y_j)$ pour $1 \le j \le n+m$ et, par convention, $s_0 = (0,0)$. Montrer que l'application $\omega \mapsto \psi(\omega) := s := (s_j)_{1\le j\le n+m}$ est une bijection de l'ensemble $\Omega_{n,m}$ vers l'ensemble suivant

$$\begin{aligned}\tilde{\Omega}_{n,m} := \Big\{ & s := (s_j)_{1\le j\le n+m} = ((x_j,y_j))_{1\le j\le n+m}\,;\, s_0 = (0,0), s_{n+m} = (n,m), \text{ et} \\ & \forall 1 \le j \le n+m \text{ on a } (x_j,y_j) = (x_{j-1}+1, y_{j-1}) \text{ ou } (x_j,y_j) = (x_{j-1}, y_{j-1}+1) \Big\}.\end{aligned}$$

En déduire que $|\tilde{\Omega}_{n,m}| = \binom{n+m}{n}$.

(ii) Grâce à la bijection ψ, on peut utiliser l'espace d'états $\tilde{\Omega}_{n,m}$, muni de la probabilité uniforme. On définit

$$A_{n,m} = \left\{ s = (s_j)_{0\le j\le n+m} \in \tilde{\Omega}_{n,m}\,, \forall 1 \le j \le n+m\,, s_j = (x_j,y_j) \text{ vérifie } x_j \ge y_j \right\}.$$

Interprétez l'événement $A_{n,m}$. Que vaut $\mathrm{P}(A_{n,m})$ si $m > n$?

(iii) Pour $s \in A_{n,m}^c$ avec $s = ((x_j,y_j))_{0\le j\le n+m}$, on pose $j_0 = j_0(s) = \min\{j \colon x_j < y_j\}$ (avec la convention $\min\emptyset := \infty$) et on on définit $s' = (s'_j)_{0\le j\le n+m}$ en posant

$$s'_j := \begin{cases} (x_j,y_j) = s_j & \text{si } j \le j_0\,, \\ (y_j - 1, x_j + 1) & \text{si } j > j_0\,. \end{cases}$$

Montrer que si $m \le n$ l'application suivante est bijective :

$$\varphi : \begin{cases} A_{n,m}^c & \to \quad \tilde{\Omega}_{m-1,n+1} \\ s = (s_j)_{0\le j\le n+m} & \mapsto \quad s' = (s'_j)_{0\le j\le n+m} \end{cases}$$

[*Sugg.* Déterminer la bijection inverse ; on pourra noter que $y_{j_0} = x_{j_0} - 1$ et que l'application φ correspond à une symétrie de la suite $(s_j)_{j_0\le j\le n+m}$ par rapport à la droite $y = x+1$.]

(iv) En déduire que pour $m \leq n$ on a $\mathrm{P}(A_{n,m}) = \frac{n-m+1}{n+1}$.

Solution 1.49. (i) Rappelons que $x_j = \sum_{i=1}^{j} \omega_i$ et $y_j = \sum_{i=1}^{j}(1-\omega_i)$. Montrons que, pour tout $\omega \in \Omega_{n,m}$, le vecteur $\psi(\omega) = s = (s_j = (x_j, y_j))_{1 \leq j \leq n+m}$ appartient à l'ensemble $\tilde{\Omega}_{n,m}$: on a en effet $s_0 = (0,0)$ par convention ; $s_{n+m} = (n,m)$ parce que $\sum_{i=1}^{n+m} \omega_i = n$ et donc $\sum_{i=1}^{n+m}(1-\omega_i) = n+m-\sum_{i=1}^{n+m} \omega_i = m$; enfin on a $(x_j, y_j) - (x_{j-1}, y_{j-1}) = (\omega_j, 1-\omega_j) \in \{(1,0),(0,1)\}$.

Réciproquement, pour $s = (s_j = (x_j, y_j))_{0 \leq j \leq n+m} \in \tilde{\Omega}_{n,m}$ donné, on définit $\omega = (\omega_i)_{1 \leq i \leq n+m}$ en posant $\omega_i := x_i - x_{i-1}$. Alors on a $\omega_i \in \{0,1\}$ pour tout $1 \leq i \leq n+m$, parce que $(x_j, y_j) - (x_{j-1}, y_{j-1}) \in \{(1,0),(0,1)\}$ par définition de $\tilde{\Omega}_{n,m}$, et de plus $\sum_{i=1}^{n+m} \omega_i = \sum_{i=1}^{n+m} x_i - x_{i-1} = x_{n+m} - x_0 = n$, de nouveau par définition de $\tilde{\Omega}_{n,m}$. Cela montre que $\omega = (\omega_i)_{1 \leq i \leq n+m} \in \Omega_{n,m}$ et donc $\omega = \psi^{-1}(s)$.

Notons pour finir que les éléments $\omega \in \Omega_{n,m}$ sont en bijection naturelle avec les sous-ensembles de $\{1, \ldots, n+m\}$ de cardinal n, donc $|\tilde{\Omega}_{n,m}| = |\Omega_{n,m}| = \binom{n+m}{n}$

(ii) On peut décrire l'événement $A_{n,m}$ de la façon suivante : « pour tout nombre $j \in \{1, \ldots, n+m\}$ de votes dépouillés, le nombre de voix x_j obtenues par la candidate A est strictement plus grand que le nombre de voix y_j obtenues par le candidat B ».

Si $m > n$ on a $\mathrm{P}(A_{n,m}) = 0$ parce que $A_{n,m} = \emptyset$: en effet, pour tout $s \in \tilde{\Omega}_{n,m}$ on a $s_{n+m} = (x_{n+m}, y_{n+m}) = (n,m)$ et donc $x_{n+m} < y_{n+m}$, c'est-à-dire $s \notin A_{n,m}$. Intuitivement, si $m > n$ l'élection est gagnée par le candidat B, il n'est donc pas possible que la candidate A soit devant pendant tout le dépouillement.

(iii) L'application φ est bien définie pour tout vecteur $s = ((x_j, y_j))_{0 \leq j \leq n+m} \in (\mathbb{R}^2)^{n+m}$ et, par construction, satisfait $\varphi(\varphi(s)) = s$ (en effet, si on pose $(x', y') := (y-1, x+1)$, on a $(y'-1, x'+1) = (x,y)$). Cela signifie que φ est une involution : il en découle que φ est injective (en effet si $\varphi(s) = \varphi(\tilde{s})$ alors $s = \varphi(\varphi(s)) = \varphi(\varphi(\tilde{s})) = \tilde{s}$), donc φ est inversible sur tout domaine — à valeurs dans l'image du domaine — d'inverse donnée par φ elle-même, restreinte à l'image. Il reste simplement à vérifier que l'image de l'application φ du domaine $A^c_{n,m}$ est effectivement l'ensemble $\tilde{\Omega}_{m-1,n+1}$.

Utilisons la suggestion : on peut interpréter $s' = \varphi(s)$ comme la réflexion de la portion du vecteur $(s_j)_{j_0 \leq j \leq n+m}$ par rapport à la droite $y = x+1$.

Montrons tout d'abord que $s' = \varphi(s) \in \tilde{\Omega}_{m-1,n+1}$ pour tout $s \in A^c_{n,m}$.

- On a $j_0 \geq 1$ donc $s'_0 = s_0 = (0,0)$.
- Pour $s \in A^c_{n,m}$ on a $j_0 < n+m$, donc, en se rappelant que $s_{n+m} = (x_{n+m}, y_{n+m}) = (n,m)$ pour tout $s \in A^c_{n,m} \subseteq \Omega_{n,m}$, on en déduit que $s'_{n+m} = (x'_{n+m}, y'_{n+m}) = (y'_{n+m} - 1, x'_{n+m} + 1) = (m-1, n+1)$.
- La condition $(x_j, y_j) = (x_{j-1}+1, y_{j-1})$ ou $(x_j, y_j) = (x_{j-1}, y_{j-1}+1)$ — qui signifie que le chemin s avance « d'un pas vers la droite ou bien vers le haut » — est préservée par la réflexion par rapport à la droite $y = x+1$.

Montrons pour finir que φ est surjective, c'est-à-dire que pour tout $s' \in \tilde{\Omega}_{m-1,n+1}$ il existe un $s \in A^c_{n,m}$ tel que $\varphi(s) = s'$. Pour cela, on choisit $s := \varphi(s')$: on observe que $s \in A^c_{n,m}$ avec les mêmes arguments que l'on vient d'utiliser plus haut et de plus $\varphi(s) = \varphi(\varphi(s')) = s'$ car on a déjà montré que φ est une involution.

(iv) D'après le point précédent, pour $m \leq n$ on a

$$\mathrm{P}(A_{n,m}) = 1 - \mathrm{P}(A^c_{n,m}) = 1 - \frac{|A^c_{n,m}|}{|\tilde{\Omega}_{n,m}|} = 1 - \frac{|\tilde{\Omega}_{m-1,n+1}|}{|\tilde{\Omega}_{n,m}|}\,.$$

On a déjà montré que $|\tilde{\Omega}_{n,m}| = |\Omega_{n,m}| = \binom{n+m}{n}$ et donc

$$\mathrm{P}(A_{n,m}) = 1 - \frac{\binom{n+m}{n+1}}{\binom{n+m}{n}} = 1 - \frac{m}{n+1} = \frac{n-m+1}{n+1}\,.$$

Chapitre 3
Variables aléatoires discrètes : la théorie

3.1 Variables aléatoires et lois

Exercice 3.1. Soit (Ω, P) un espace probabilisé discret. Montrer que toute variable aléatoire $X : \Omega \to E$ est discrète (où E est un ensemble non vide quelconque).
[*Sugg.* Considérer dans un premier temps le cas où Ω est dénombrable.]

Solution 3.1. Supposons pour commencer que Ω est dénombrable. Si on note $\tilde{E} := \{X(\omega) : \omega \in \Omega\}$ l'image de l'application X, alors $\tilde{E}$ est de cardinal inférieur à celui de Ω (l'application $\omega \mapsto X(\omega)$ de Ω vers $\tilde{E}$ est surjective) : ainsi, $\tilde{E}$ est dénombrable. Comme

$$\mathrm{P}(X \in \tilde{E}) = \mathrm{P}(\Omega) = 1,$$

il découle de la Définition 3.12 que X est une variable aléatoire discrète. De manière plus générale, si (Ω, P) est un espace probabilisé discret, d'après la Définition 1.11 il existe $\tilde{\Omega} \subseteq \Omega$ dénombrable, pour lequel $\mathrm{P}(\tilde{\Omega}) = 1$. Si on pose $\tilde{E} := \{X(\omega) : \omega \in \tilde{\Omega}\}$, on a de nouveau que $\tilde{E}$ est dénombrable, et $\mathrm{P}(X \in \tilde{E}) = \mathrm{P}(\tilde{\Omega}) = 1$, ce qui permet de conclure grâce à la Définition 3.12.

Exercice 3.2. Soit $X : \Omega \to E$ une variable aléatoire discrète et soit $f : E \to F$ une fonction. Montrer que $f(X) : \Omega \to F$ est une variable aléatoire discrète.

Solution 3.2. Comme X est une variable aléatoire discrète, d'après la Définition 3.12 il existe $\tilde{E} \subseteq E$ dénombrable tel que $\mathrm{P}(X \in \tilde{E}) = 1$. En posant $\tilde{F} = \{f(x) : x \in \tilde{E}\}$ l'image de la restriction de f à $\tilde{E}$, on a que $\tilde{F}$ est dénombrable (car de cardinal inférieur à $\tilde{E}$). On a alors

$$\mathrm{P}(f(X) \in \tilde{F}) = \mathrm{P}(X \in \tilde{E}) = 1\,,$$

ce qui, d'après la Définition 3.12, implique que $f(X)$ est une variable aléatoire discrète.

Exercice 3.3. Soit X une variable aléatoire définie sur (Ω, P) à valeurs dans E, presque sûrement constante, c'est-à-dire telle qu'il existe $c \in E$ tel que $\mathrm{P}(X = c) = 1$. Montrer qu'il existe un *unique* élément $c \in E$ avec une telle propriété.

Solution 3.3. S'il existait deux constantes c et c' avec cette propriété, les événements $\{X = c\}$ et $\{X = c'\}$ seraient deux événements disjoints de probabilité 1, ce qui est impossible.

Exercice 3.4. Soient A, B deux événements d'un espace probabilisé (Ω, P). Pour $\lambda, \mu \in \mathbb{R}$, on considère la variable aléatoire

$$X := \lambda \mathbb{1}_A + \mu \mathbb{1}_B\,.$$

Montrer qu'il s'agit d'une variable aléatoire discrète et en déterminer la densité discrète.
[*Sugg.* Considérer dans un premier temps le cas où les quatre nombres $0, \lambda, \mu, \lambda + \mu$ sont distincts.]

Solution 3.4. Considérons pour commencer le cas où les quatre nombres $0, \lambda, \mu, \lambda+\mu$ sont distincts. On a alors $X(\Omega) = \{0, \lambda, \mu, \lambda+\mu\}$ et

$$\begin{aligned}
\{X = \lambda\} = A \cap B^c \quad &\Rightarrow \quad \mathrm{p}_X(\lambda) = \mathrm{P}(A \cap B^c) \\
\{X = \mu\} = A^c \cap B \quad &\Rightarrow \quad \mathrm{p}_X(\mu) = \mathrm{P}(A^c \cap B) \\
\{X = \lambda+\mu\} = A \cap B \quad &\Rightarrow \quad \mathrm{p}_X(\lambda+\mu) = \mathrm{P}(A \cap B) \\
\{X = 0\} = A^c \cap B^c \quad &\Rightarrow \quad \mathrm{p}_X(0) = \mathrm{P}(A^c \cap B^c)
\end{aligned}$$

Dans le cas où les quatre nombres $0, \lambda, \mu, \lambda+\mu$ ne sont pas distincts, il suffit de sommer les probabilités correspondant aux valeurs égales. Par exemple, si $\lambda = \mu \neq 0$, alors X est à valeurs dans $\{0, \lambda, 2\lambda\}$ et $\mathrm{p}_X(\lambda) = \mathrm{P}(A \cap B^c) + \mathrm{P}(A^c \cap B)$, $\mathrm{p}_X(2\lambda) = \mathrm{P}(A \cap B)$, $\mathrm{p}_X(0) = \mathrm{P}(A^c \cap B^c)$. On traite de la même manière les cas restants, c'est-à-dire $\lambda = -\mu \neq 0$, $0 = \lambda \neq \mu$, $0 = \mu \neq \lambda$, $\lambda = \mu = 0$.

3.2 Indépendance de variables aléatoires

Exercice 3.5. Soient $X : \Omega \to E$ et $Y : \Omega \to F$ deux variables aléatoires *discrètes*, définies sur le même espace probabilisé (Ω, P), à valeurs dans des ensembles qui peuvent être différents. Montrer que la variable aléatoire jointe $(X,Y)(\omega) := (X(\omega), Y(\omega))$, définie sur Ω et à valeurs dans $E \times F$, est elle-aussi une variable aléatoire discrète.

Solution 3.5. D'après la Définition 3.12, il existe $\tilde{E} \subseteq E$ et $\tilde{F} \subseteq F$ dénombrables tels que $\mathrm{P}(X \in \tilde{E}) = \mathrm{P}(Y \in \tilde{F}) = 1$. L'ensemble $\tilde{E} \times \tilde{F}$ est lui aussi dénombrable, en tant que produit cartésien d'ensembles dénombrables. De plus,

$$\mathrm{P}((X,Y) \in \tilde{E} \times \tilde{F}) = \mathrm{P}(\{X \in \tilde{E}\} \cap \{Y \in \tilde{F}\}) = 1,$$

où on a utilisé l'Exercice 1.1. Cela montre que (X,Y) est une variable aléatoire discrète.

Exercice 3.6. Soient $X : \Omega \to E$ et $Y : \Omega \to F$ deux variables aléatoires (pas nécessairement discrètes) définies sur le même espace probabilisé. Montrer que les lois marginales μ_X et μ_Y peuvent être déduites de la loi jointe $\mu_{X,Y}$ de la manière suivante :

$$\mu_X(A) = \mu_{X,Y}(A \times F), \qquad \mu_Y(B) = \mu_{X,Y}(E \times B), \qquad \forall A \subseteq E,\ \forall B \subseteq F.$$

Solution 3.6. Il suffit d'observer que, comme $\{X \in A\} = \{X \in A\} \cap \Omega = \{X \in A\} \cap \{Y \in F\}$, on a

$$\mu_X(A) = \mathrm{P}(X \in A) = \mathrm{P}(\{X \in A\} \cap \{Y \in F\}) = \mathrm{P}((X,Y) \in A \times F) = \mu_{X,Y}(A \times F).$$

La deuxième identité se démontre de la même manière, en échangeant les rôles de X et Y.

Exercice 3.7. Soient A et B deux événements dans un espace probabilisé discret (Ω, P). Déterminer la loi jointe du vecteur aléatoire $(\mathbb{1}_A, \mathbb{1}_B)$.

Solution 3.7. La densité discrète jointe $\mathrm{p}(x,y) = \mathrm{P}(\mathbb{1}_A = x, \mathbb{1}_B = y)$ est donnée par

$$\begin{aligned}
&\mathrm{p}(0,0) = \mathrm{P}(A^c \cap B^c) \qquad && \mathrm{p}(0,1) = \mathrm{P}(A^c \cap B) \\
&\mathrm{p}(1,0) = \mathrm{P}(A \cap B^c) \qquad && \mathrm{p}(1,1) = \mathrm{P}(A \cap B),
\end{aligned}$$

alors que $\mathrm{p}(x,y) = 0$ si $x, y \not\in \{0,1\}$.

Exercice 3.8. Soit X,Y deux variables aléatoires définies sur le même espace probabilisé. On suppose que X est une variable aléatoire presque sûrement constante, comme définie dans l'Exemple 3.16. Montrer que X et Y sont indépendantes.

Solution 3.8. Remarquons tout d'abord que, d'après l'Exercice 1.3, un événement de probabilité 0 ou de probabilité 1 est indépendant de n'importe quel autre événement. Montrer que X et Y sont indépendants équivaut à montrer que toute paire d'événements de la forme $\{X \in A\}$ et $\{Y \in B\}$ est formée d'événements indépendants (voir la Proposition 3.31). Cela découle du fait que tout événement de la forme $\{X \in A\}$ est de probabilité 0 ou 1. En effet, si $X : \Omega \to E$ est presque sûrement constante, cela signifie qu'il existe un $c \in E$ tel que $\mathrm{P}(X = c) = 1$. Mais alors si $A \subseteq E$, on a $\mathrm{P}(X \in A) = 1$ si $c \in A$ et $\mathrm{P}(X \in A) = 0$ si $c \notin A$.

Exercice 3.9. Soient X,Y deux variables aléatoires à valeurs respectivement dans les ensembles finis E,F. On suppose que la loi du vecteur aléatoire (X,Y) est donnée par la probabilité uniforme sur $E \times F$, introduite dans l'Exemple 1.18. Montrer que les lois marginales sont données par les probabilités uniformes respectivement sur E et F et en déduire que X et Y sont indépendantes.

Solution 3.9. Observons pour commencer que, en utilisant ce que l'on a vu dans l'Exercice 3.6, si $A \subseteq E$

$$\mu_X(A) = \mu_{X,Y}(A \times F) = \frac{|A \times F|}{|E \times F|} = \frac{|A|}{|E|},$$

c'est-à-dire μ_X est la probabilité uniforme sur E. On montre de même que μ_Y est la probabilité uniforme sur F. On a donc, pour $A \subseteq E$ et $B \subseteq F$,

$$\begin{aligned} \mathrm{P}(X \in A, Y \in B) = \mu_{X,Y}(A \times B) &= \frac{|A \times B|}{|E \times F|} = \frac{|A|}{|E|}\frac{|B|}{|F|} \\ &= \mu_X(A)\mu_Y(B) = \mathrm{P}(X \in A)\,\mathrm{P}(Y \in B) \end{aligned}$$

dont découle l'indépendance de X et Y.

Exercice 3.10. Soient $C_1, C_2, \ldots, C_n$ des événements d'un espace probabilisé discret (Ω, P). En posant $X_i := \mathbb{1}_{C_i}$, montrer l'équivalence des deux affirmations suivantes :

(i) les variables aléatoires $X_1, X_2, \ldots, X_n$ sont indépendantes ;

(ii) les événements $C_1, C_2, \ldots, C_n$ sont indépendants.

Montrer que c'est aussi le cas pour une famille arbitraire $(C_i)_{i \in I}$ d'événements.
[*Sugg.* Se rappeler de la Proposition 1.69.]

Solution 3.10. (i) $\Rightarrow$ (ii) Soit $J \subseteq \{1, 2, \ldots, n\}$. Notons que

$$\bigcap_{i \in J} C_i = \{X_i = 1 \text{ pour tout } i \in J\}.$$

Ainsi, en utilisant l'hypothèse d'indépendance des variables aléatoires $X_1, X_2, \ldots, X_n$:

$$\begin{aligned} \mathrm{P}\left(\bigcap_{i \in J} C_i\right) &= \mathrm{P}\left(X_i = 1 \text{ pour tout } i \in J, X_k \in \{0,1\} \text{ pour tout } k \notin J\right) \\ &= \prod_{i \in J} \mathrm{P}(X_i = 1) = \prod_{i \in J} \mathrm{P}(C_i). \end{aligned}$$

(ii) $\Rightarrow$ (i) Montrons que la densité discrète jointe de $X_1, X_2, \ldots, X_n$ est le produit des densités marginales. Pour $x \in \{0,1\}$, posons $C^x := C$ si $x = 1$ et $C^x = C^c$ si $x = 0$. En se rappelant du fait que l'indépendance d'événements est conservée par passage au complémentaire (Proposition 1.70) et en observant que $\{X_i = x\} = C_i^x$, on a

$$\begin{aligned}\mathrm{p}_{X_1,\ldots,X_n}(x_1,\ldots,x_n) &= \mathrm{P}(X_1 = x_1, \ldots, X_n = x_n) = \mathrm{P}(C_1^{x_1} \cap \cdots \cap C_n^{x_n}) \\ &= \prod_{i=1}^{n} \mathrm{P}(C_i^{x_i}) = \prod_{i=1}^{n} \mathrm{P}(X_i = x_i),\end{aligned}$$

où pour l'avant-dernière égalité on a utilisé la Proposition 1.69.

Pour étendre l'affirmation à une famille arbitraire, il suffit de se souvenir que l'indépendance d'une famille de variables aléatoires (resp. d'événements) équivaut par définition à l'indépendance de toute sous-famille finie.

Exercice 3.11. Lors de la répétition de n épreuves indépendantes de probabilité de succès p, on considère les variables aléatoires $S :=$ « nombre de succès parmi les n épreuves » et $T :=$ « numéro de l'épreuve où l'on a le premier succès » (on convient que $T = +\infty$ si $S = 0$). Déterminer la densité discrète jointe $\mathrm{p}_{S,T}$.

[Noter que les densités marginales p_S et p_T ont déjà été calculées dans l'Exemple 3.18, voir (3.14).]

Solution 3.11. L'événement $\{S = s, T = t\}$ peut s'exprimer comme l'intersection de trois événements correspondant aux affirmations suivantes :

- les $t-1$ premières épreuves sont des échecs ;
- la t-ème épreuve est un succès ;
- dans les $n-t$ épreuves restantes il y a exactement $s-1$ succès.

Étant donné l'indépendance d'épreuves distinctes, les trois événements ci-dessus sont indépendants. On en déduit que, pour $1 \le s,t \le n$

$$\begin{aligned}\mathrm{p}_{S,T}(s,t) = \mathrm{P}(S = s, T = t) &= (1-p)^{t-1} p \binom{n-t}{s-1} p^{s-1} (1-p)^{n-t-(s-1)} \\ &= \binom{n-t}{s-1} p^s (1-p)^{n-s},\end{aligned}$$

avec la convention $\binom{j}{i} = 0$ si $i > j$.

De plus, on a $S = 0$ et $T = +\infty$ si les n épreuves sont des échecs, d'où

$$\mathrm{p}_{S,T}(0,+\infty) = (1-p)^n .$$

3.3 Espérance et inégalités probabilistes

Exercice 3.12. Soit X une variable aléatoire discrète réelle à valeurs dans $\mathbb{N}^*$, de densité discrète donnée par $\mathrm{p}_X(x) = c_\alpha x^{-(1+\alpha)} \mathbb{1}_{\mathbb{N}^*}(x)$, où $\alpha \in]0,\infty[$ est un paramètre fixé et où la constante c_α vérifie

$$\frac{1}{c_\alpha} = \sum_{y=1}^{+\infty} \frac{1}{y^{1+\alpha}} \quad (= \zeta(1+\alpha)) .$$

Déterminer pour quelles valeurs de $p \in]0,\infty[$ on a $X \in L^p(\Omega,\mathrm{P})$.

Solution 3.12. En utilisant la définition de $L^p(\Omega, \mathrm{P})$ (Section 3.3.3) et la formule de transfert (Proposition 3.51, voir aussi (3.32)), on a $X \in L^p(\Omega, \mathrm{P})$ si et seulement si

$$\sum_{k=1}^{+\infty} k^p \, \mathrm{p}_X(k) < +\infty \iff \sum_{k=1}^{+\infty} \frac{1}{k^{1+\alpha-p}} < +\infty.$$

En se rappelant du critère de Riemann (0.3), c'est le cas si et seulement si $p < \alpha$.

Exercice 3.13. Soient X et Y des variables aléatoires indépendantes, à valeurs respectivement dans les ensembles E et F, et soient $g : E \to \mathbb{R}$ et $h : F \to \mathbb{R}$ des fonctions quelconques. Montrer que si les variables aléatoires $g(X)$ et $h(Y)$ admettent une espérance finie, alors leur produit $g(X)h(Y)$ admet aussi une espérance finie et $\mathrm{E}(g(X)h(Y)) = \mathrm{E}(g(X))\,\mathrm{E}(h(Y))$.

Solution 3.13. Il découle de la Proposition 3.39 (indépendance par transformation) que les variables aléatoires $g(X)$ et $h(Y)$ sont indépendantes. Il suffit donc d'appliquer la Proposition 3.72 à ces variables aléatoires.

Exercice 3.14. Pour $n \in \mathbb{N}^*$, soit X_n une variable aléatoire qui prend les valeurs $1, 2, \ldots, n$ avec la même probabilité. Soit $f : [0,1] \to \mathbb{R}$ une fonction continue. On pose

$$m_n = \mathrm{E}\left(f\left(\frac{X_n}{n}\right)\right).$$

Identifier la limite de m_n quand $n \to \infty$.

Solution 3.14. D'après la formule de transfert, on a

$$\mathrm{E}\left[f\left(\frac{X_n}{n}\right)\right] = \frac{1}{n}\sum_{k=1}^{n} f(k/n),$$

qui est une somme de Riemann pour l'intégrale $\int_0^1 f(x)\mathrm{d}x$. Ainsi,

$$\lim_{n\to+\infty} m_n = \int_0^1 f(x)\mathrm{d}x.$$

Exercice 3.15. Deux variables aléatoires X et Y sont définies de la manière suivante, sur la base de trois lancers indépendants d'une pièce de monnaie équilibrée. Si on obtient face au premier lancer, on pose $X = Y = 0$. Si on obtient pile au premier lancer, alors on pose $X = 1$ si le deuxième lancer donne pile et $X = -1$ si le deuxième lancer donne face. De même, si on obtient pile au premier lancer, alors on pose $Y = 1$ si le troisième lancer donne pile et $Y = -1$ si le troisième lancer donne face. Déterminer la densité discrète jointe de X et Y, et montrer que les variables aléatoires X et Y sont décorrélées mais pas indépendantes.

Solution 3.15. Notons C_i l'événement que le i-ème lancer donne pile. On a

$$\begin{aligned}
\mathrm{p}_{X,Y}(0,0) &= \mathrm{P}(C_1^c) = \frac{1}{2}\\
\mathrm{p}_{X,Y}(1,1) &= \mathrm{P}(C_1 \cap C_2 \cap C_3) = \frac{1}{8}\\
\mathrm{p}_{X,Y}(1,-1) &= \mathrm{P}(C_1 \cap C_2 \cap C_3^c) = \frac{1}{8}\\
\mathrm{p}_{X,Y}(-1,1) &= \mathrm{P}(C_1 \cap C_2^c \cap C_3) = \frac{1}{8}\\
\mathrm{p}_{X,Y}(-1,-1) &= \mathrm{P}(C_1 \cap C_2^c \cap C_3^c) = \frac{1}{8}.
\end{aligned}$$

On en déduit que

$$\mathrm{p}_X(0) = \frac{1}{2} \qquad \mathrm{p}_X(1) = \mathrm{p}_{X,Y}(1,1) + \mathrm{p}_{X,Y}(1,-1) = \frac{1}{4}$$
$$\mathrm{p}_X(-1) = \mathrm{p}_{X,Y}(1,-1) + \mathrm{p}_{X,Y}(-1,-1) = \frac{1}{4}$$

et, par symétrie, $\mathrm{p}_X = \mathrm{p}_Y$. En particulier, $\mathrm{E}(X) = \mathrm{E}(Y) = 0$. Ainsi,

$$\begin{aligned}\mathrm{Cov}(X,Y) = \mathrm{E}(XY) &= \sum_{x,y} xy\, \mathrm{p}_{X,Y}(x,y) \\ &= \mathrm{p}_{X,Y}(1,1) + \mathrm{p}_{X,Y}(-1,-1) - \mathrm{p}_{X,Y}(1,-1) - \mathrm{p}_{X,Y}(-1,1) = 0,\end{aligned}$$

donc X et Y sont décorrélées. Cependant, X et Y ne sont pas indépendantes, parce que

$$\mathrm{P}(X=0, Y=0) = \mathrm{p}_{X,Y}(0,0) = \frac{1}{2} \neq \frac{1}{4} = \mathrm{p}_X(0)\,\mathrm{p}_Y(0) = \mathrm{P}(X=0)\,\mathrm{P}(Y=0)\,.$$

Exercice 3.16. Soient A et B deux événements d'un espace probabilisé. Montrer que l'on a $\mathrm{Cov}(\mathbb{1}_A, \mathbb{1}_B) = \mathrm{P}(A \cap B) - \mathrm{P}(A)\,\mathrm{P}(B)$ et en déduire que A et B sont indépendants si et seulement si $\mathrm{Cov}(\mathbb{1}_A, \mathbb{1}_B) = 0$.

Solution 3.16. Il suffit d'observer que pour tout événement A on a $\mathrm{E}(\mathbb{1}_A) = \mathrm{P}(A)$ et que $\mathbb{1}_A \mathbb{1}_B = \mathbb{1}_{A\cap B}$. On obtient donc

$$\mathrm{Cov}(\mathbb{1}_A, \mathbb{1}_B) = \mathrm{E}(\mathbb{1}_A \mathbb{1}_B) - \mathrm{E}(\mathbb{1}_A)\,\mathrm{E}(\mathbb{1}_B) = \mathrm{P}(A \cap B) - \mathrm{P}(A)\,\mathrm{P}(B)\,,$$

dont découle la conclusion.

Exercice 3.17. Soit Y une variable aléatoire à valeurs dans $\mathbb{R}_+$, et $f :]0,+\infty[\to]0,+\infty[$ une fonction croissante. Montrer que pour tout $\varepsilon > 0$ on a

$$\mathrm{P}(Y \geq \varepsilon) \leq \frac{\mathrm{E}(f(Y))}{f(\varepsilon)}\,.$$

Solution 3.17. Comme f est croissante, on en déduit l'égalité d'événements

$$\{Y \geq \varepsilon\} = \{f(Y) \geq f(\varepsilon)\}\,.$$

En appliquant l'inégalité de Markov (Théorème 3.77) à la variable aléatoire positive $X := f(Y)$, on obtient le résultat.

Exercice 3.18 (Inégalité de Jensen stricte). Démontrer la version forte de l'inégalité de Jensen suivante. Soit X une variable aléatoire réelle et $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction *strictement* convexe, dans le sens où la relation (3.49) est une inégalité *stricte* pour tout $x \neq x_0$. Montrer que si X admet une espérance finie $\mathrm{E}(X) < \infty$ et si X *n'est pas* presque sûrement constante, alors on a l'inégalité *stricte*

$$\varphi(\mathrm{E}(X)) < \mathrm{E}(\varphi(X))\,.$$

Solution 3.18. En utilisant l'inégalité (3.49), on obtient

$$\varphi(X) - \varphi(\mathrm{E}(X)) - \lambda(\mathrm{E}(X))(X - \mathrm{E}(X)) \geq 0.$$

Posons $Y := \varphi(X) - \varphi(\mathrm{E}(X)) - \lambda(\mathrm{E}(X))(X - \mathrm{E}(X))$ et remarquons que

$$\mathrm{E}(Y) = \mathrm{E}(\varphi(X)) - \varphi(\mathrm{E}(X)).$$

Il s'agit donc de démontrer que $\mathrm{E}(Y) > 0$. En utilisant le fait que $Y \geq 0$ et la Proposition 3.57, il suffit de montrer que Y *n'est pas* presque sûrement égal à 0, c'est-à-dire que $\mathrm{P}(Y > 0) > 0$. Comme, par hypothèse, X n'est pas presque sûrement constante, on a

$$\mathrm{P}(X \neq \mathrm{E}(X)) > 0.$$

De plus, en utilisant l'inégalité (3.49) au sens strict, on a

$$\{X \neq \mathrm{E}(X)\} \subseteq \{Y > 0\}.$$

Cela montre que $\mathrm{P}(Y > 0) > 0$.

Exercice 3.19 (Inégalité de Cantelli (ou de Tchebychev unilatérale)). Soit X une variable aléatoire réelle, admettant un moment d'ordre 2 fini (c'est-à-dire $X \in L^2$).

(i) On pose $Y = X - \mathrm{E}(X)$. Montrer que pour tout $\alpha > 0$ on a les inégalités suivantes

$$\mathrm{P}(Y \geq \varepsilon) \leq \mathrm{P}\left((Y+\alpha)^2 \geq (\varepsilon+\alpha)^2\right) \leq \frac{\mathrm{E}\left((Y+\alpha)^2\right)}{(\varepsilon+\alpha)^2}, \qquad \forall \varepsilon > 0.$$

(ii) En prenant $\alpha = \frac{1}{\varepsilon}\mathrm{E}(Y^2) = \frac{1}{\varepsilon}\mathrm{Var}(X)$, conclure que l'on a l'inégalité

$$\mathrm{P}(X \geq \mathrm{E}(X) + \varepsilon) \leq \frac{\mathrm{Var}(X)}{\varepsilon^2 + \mathrm{Var}(X)}, \qquad \forall \varepsilon > 0. \tag{S3.1}$$

Solution 3.19. (i) Notons que l'on a l'inclusion d'événements

$$\{Y \geq \varepsilon\} = \{Y + \alpha \geq \varepsilon + \alpha\} \subseteq \{(Y+\alpha)^2 \geq (\varepsilon+\alpha)^2\},$$

ce qui donne la première inégalité. Pour la deuxième inégalité, il suffit d'appliquer l'inégalité de Markov (Théorème 3.77) à la variable aléatoire $(Y+\alpha)^2$.

(ii) Grâce au point précédent, il suffit de montrer que

$$\frac{\mathrm{E}\left[\left(X - \mathrm{E}(X) + \frac{1}{\varepsilon}\mathrm{Var}(X)\right)^2\right]}{\left(\varepsilon + \frac{1}{\varepsilon}\mathrm{Var}(X)\right)^2} = \frac{\mathrm{Var}(X)}{\varepsilon^2 + \mathrm{Var}(X)}.$$

On a en effet

$$\begin{aligned}
\frac{\mathrm{E}\left[\left(X - \mathrm{E}(X) + \frac{1}{\varepsilon}\mathrm{Var}(X)\right)^2\right]}{\left(\varepsilon + \frac{1}{\varepsilon}\mathrm{Var}(X)\right)^2} &= \frac{\mathrm{E}\left[(X-\mathrm{E}(X))^2\right] + \frac{1}{\varepsilon^2}\mathrm{Var}^2(X) + \frac{2}{\varepsilon}\mathrm{Var}(X)\,\mathrm{E}\left[(X-\mathrm{E}(X)\right]}{\frac{1}{\varepsilon^2}\left(\varepsilon^2 + \mathrm{Var}(X)\right)^2} \\
&= \frac{\mathrm{Var}(X)\left(\varepsilon^2 + \mathrm{Var}(X)\right)}{\left(\varepsilon^2 + \mathrm{Var}(X)\right)^2} = \frac{\mathrm{Var}(X)}{\varepsilon^2 + \mathrm{Var}(X)}.
\end{aligned}$$

Exercice 3.20 (Inégalité de Paley–Zygmund). Soit X une variable aléatoire positive, qui admet une espérance finie, c'est-à-dire $X \geq 0$, $X \in L^1$. Montrer que, pour tout $\theta \in]0,1[$, on a l'inégalité suivante

$$\mathrm{P}\left(X \geq \theta\,\mathrm{E}(X)\right) \geq (1-\theta)^2 \frac{\mathrm{E}(X)^2}{\mathrm{E}(X^2)},$$

où le terme de droite est égal à 0 si $\mathrm{E}(X^2) = +\infty$.

[*Sugg.* Montrer que $\mathrm{E}(X) \leq \theta\,\mathrm{E}(X) + \mathrm{E}(X \mathbb{1}_{\{X \geq \theta \mathrm{E}(X)\}})$ puis utiliser l'inégalité de Cauchy–Schwarz.]

Solution 3.20. Commençons par remarquer que pour tous $x, y \geq 0$ on a l'inégalité

$$x \leq y + x\mathbb{1}_{\{x \geq y\}}.$$

En posant $x = X$ et $y = \theta\,\mathrm{E}(X)$ on a donc

$$X \leq \theta\,\mathrm{E}(X) + X\mathbb{1}_{\{X \geq \theta\,\mathrm{E}(X)\}}.$$

Par monotonie de l'espérance (Proposition 3.54) on en déduit que

$$\mathrm{E}(X) \leq \theta\,\mathrm{E}(X) + \mathrm{E}(X\mathbb{1}_{\{X \geq \theta\,\mathrm{E}(X)\}}),$$

que l'on peut réécrire sous la forme

$$(1-\theta)^2\,\mathrm{E}(X)^2 \leq \mathrm{E}\left(X\mathbb{1}_{\{X \geq \theta\,\mathrm{E}(X)\}}\right)^2.$$

D'après l'inégalité de Cauchy–Schwarz (Théorème 3.82) et en utilisant le fait que $(\mathbb{1}_A)^2 = \mathbb{1}_A$

$$\mathrm{E}\left(X\mathbb{1}_{\{X \geq \theta\,\mathrm{E}(X)\}}\right)^2 \leq \mathrm{E}(X^2)\,\mathrm{E}\left(\mathbb{1}_{\{X \geq \theta\,\mathrm{E}(X)\}}\right) = \mathrm{E}(X^2)\,\mathrm{P}(X \geq \theta\,\mathrm{E}(X)).$$

On en déduit que

$$(1-\theta)^2\,\mathrm{E}(X)^2 \leq \mathrm{E}(X^2)\,\mathrm{P}(X \geq \theta\,\mathrm{E}(X)),$$

ce qui donne la conclusion voulue.

3.4 Travailler avec les lois

Exercice 3.21 (Généralisation de l'Exemple 3.25). Soient X et Y deux variables aléatoires indépendantes toutes les deux à valeurs dans $\{1, 2, \dots, n\}$, de loi donnée par $\mathrm{P}(X = k) = \mathrm{P}(Y = k) = \frac{1}{n}$ pour tout $k = 1, 2, \dots, n$.

(i) Calculer les probabilités $\mathrm{P}(X = Y)$ et $\mathrm{P}(X > Y)$.

(ii) Calculer la densité discrète de $X + Y$.

Solution 3.21. (i) Commençons par calculer

$$\mathrm{P}(X = Y) = \mathrm{P}\left(\bigcup_{k=1}^{n}\{X = k,\, Y = k\}\right) = \sum_{k=1}^{n}\mathrm{P}(X = k, Y = k) = \sum_{k=1}^{n}\frac{1}{n^2} = \frac{1}{n}.$$

Il est facile de voir que les vecteurs aléatoires (X,Y) et (Y,X) ont la même loi : on en déduit que $\mathrm{P}(X < Y) = \mathrm{P}(X > Y)$, donc $\mathrm{P}(X \neq Y) = \mathrm{P}(X < Y) + \mathrm{P}(X > Y) = 2\,\mathrm{P}(X > Y)$. Ainsi

$$\mathrm{P}(X > Y) = \frac{1}{2}\mathrm{P}(X \neq Y) = \frac{1}{2}\Big(1 - \mathrm{P}(X = Y)\Big) = \frac{1}{2} - \frac{1}{2n}.$$

(ii) Utilisons la Proposition 3.91. Dans le cas présent, on a

$$\mathrm{p}_X = \mathrm{p}_Y = \frac{1}{n}\mathbb{1}_{\{1,2,\dots,n\}}.$$

Ainsi,

$$\mathrm{p}_{X,Y}(k) = \mathrm{p}_X * \mathrm{p}_Y(k) = \frac{1}{n^2}\sum_h \mathbb{1}_{\{1,2,\dots,n\}}(h)\mathbb{1}_{\{1,2,\dots,n\}}(k-h) = \frac{1}{n^2}\sum_{h=1}^{n} \mathbb{1}_{\{1,2,\dots,n\}}(k-h)\,.$$

Remarquons que pour $h \in \{1,\dots,n\}$ on a

$$\mathbb{1}_{\{1,2,\dots,n\}}(k-h) = \begin{cases} \mathbb{1}_{\{1,2,\dots,k-1\}}(h) & \text{si } 2 \le k \le n\,, \\ \mathbb{1}_{\{k-n,k-n+1,\dots,n\}}(h) & \text{si } n < k \le 2n\,. \end{cases}$$

On en déduit

$$\mathrm{p}_{X,Y}(k) = \begin{cases} \dfrac{k-1}{n^2} & \text{si } 2 \le k \le n\,, \\ \dfrac{2n-k+1}{n^2} & \text{si } n < k \le 2n\,. \end{cases}$$

Exercice 3.22. Soit (X,Y) une variable aléatoire *discrète* à valeurs dans $\mathbb{R}^2$. On définit

$$F_{X,Y}(x,y) := \mathrm{P}(X \le x, Y \le y) \qquad \forall (x,y) \in \mathbb{R}^2.$$

Montrer que, pour tout $(x,y) \in \mathbb{R}^2$,

$$\mathrm{p}_{X,Y}(x,y) = \lim_{n\to+\infty} \left[F_{X,Y}(x,y) + F_{X,Y}\left(x-\tfrac{1}{n}, y-\tfrac{1}{n}\right) - F_{X,Y}\left(x-\tfrac{1}{n}, y\right) - F_{X,Y}\left(x, y-\tfrac{1}{n}\right)\right].$$

Solution 3.22. Remarquons pour commencer que l'on a l'égalité d'événements suivante :

$$\{X \le x, Y \le y\} = \{X \le x-\tfrac{1}{n}, Y \le y\} \cup \{X \le x, Y \le y-\tfrac{1}{n}\} \cup \{x-\tfrac{1}{n} < X \le x,\, y-\tfrac{1}{n} < Y \le y\}.$$

Dans cette dernière union, le troisième événement est disjoint des deux premiers et l'intersection des deux premiers est

$$\{X \le x-\tfrac{1}{n}, Y \le y\} \cap \{X \le x, Y \le y-\tfrac{1}{n}\} = \{X \le x-\tfrac{1}{n}, Y \le y-\tfrac{1}{n}\}.$$

Ainsi, grâce à la Proposition 1.21 (ii),

$$\begin{aligned} \mathrm{P}(X \le x, Y \le y) = {} & \mathrm{P}(X \le x-\tfrac{1}{n}, Y \le y) + \mathrm{P}\left(X \le x, Y \le y-\tfrac{1}{n}\right) \\ & - \mathrm{P}\left(X \le x-\tfrac{1}{n}, Y \le y-\tfrac{1}{n}\right) + \mathrm{P}\left(x-\tfrac{1}{n} < X \le x,\, y-\tfrac{1}{n} < Y \le y\right), \end{aligned}$$

dont on déduit immédiatement que

$$\begin{aligned} & \mathrm{P}\left(x-\tfrac{1}{n} < X \le x,\, y-\tfrac{1}{n} < Y \le y\right) \\ & \qquad = \left[F_{X,Y}(x,y) + F_{X,Y}\left(x-\tfrac{1}{n}, y-\tfrac{1}{n}\right) - F_{X,Y}\left(x-\tfrac{1}{n}, y\right) - F_{X,Y}\left(x, y-\tfrac{1}{n}\right)\right]. \end{aligned}$$

Pour conclure, il suffit maintenant de prendre la limite quand $n \to +\infty$, en observant que, par continuité par le haut des probabilités (Proposition 1.24 (iii))

$$\lim_{n\to+\infty} \mathrm{P}\left(x-\tfrac{1}{n} < X \le x,\, y-\tfrac{1}{n} < Y \le y\right) = \mathrm{P}(X=x, Y=y) = \mathrm{p}_{X,Y}(x,y).$$

Exercice 3.23. Soit X une variable aléatoire réelle et soit m une médiane de X.

(i) Montrer que, pour tous $a,b \in \mathbb{R}$, $am+b$ est une médiane de $aX+b$.

(ii) Si X est positive, montrer que $\mathrm{E}(X) \ge \frac{1}{2}m$.

Solution 3.23. (i) Si $a=0$ alors b est une médiane de $aX+b=b$, parce que $\mathrm{P}(b \ge b) = \mathrm{P}(b \le b) = 1 \ge \frac{1}{2}$. Si $a>0$ alors

$$P(aX+b \geq am+b) = P(X \geq m) \geq \frac{1}{2},$$
$$P(aX+b \leq am+b) = P(X \leq m) \geq \frac{1}{2},$$

donc $am+b$ est une médiane pour $aX+b$. Le cas $a<0$ est analogue.

(ii) Raisonnons par l'absurde et supposons qu'il existe une médiane m de X pour laquelle $E(X) < \frac{1}{2}m$. Il existe alors $\varepsilon > 0$ suffisamment petit pour que l'on ait $m > (2+\varepsilon)E(X)$. Mais alors
$$\frac{1}{2} \leq P(X \geq m) \leq P(X \geq (2+\varepsilon)E(X)) \leq \frac{1}{2+\varepsilon} < \frac{1}{2},$$
ce qui est absurde. Notons que l'inégalité $P(X \geq (2+\varepsilon)E(X)) \leq \frac{1}{2+\varepsilon}$ provient de l'inégalité de Markov (Théorème 3.77).

Exercice 3.24. Soient X,Y deux variables aléatoires à valeurs dans $\{0,1\}$, indépendantes et de lois respectives $P(X=1)=p$, $P(X=0)=1-p$ et $P(Y=1)=q$, $P(Y=0)=1-q$ avec $p,q \in [0,1]$.

(i) Calculer une médiane m_1 de X et une médiane m_2 de Y. Montrer que m_1 (resp. m_2) est unique si et seulement si $p \neq \frac{1}{2}$ (resp. $q \neq \frac{1}{2}$).

(ii) Déterminer la densité discrète de $X+Y$ et en déduire une médiane m de $X+Y$. Pour quelles valeurs de $p,q \in [0,1]$ la médiane m est-elle unique ? Montrer que tous les cas $m < m_1+m_2$, $m = m_1+m_2$ et $m > m_1+m_2$ sont possibles.

Cela montre que, à la différence de l'espérance, la médiane n'est pas linéaire.

Solution 3.24. (i) Si $p=\frac{1}{2}$ il est facile de vérifier que tout $m_1 \in [0,1]$ est une médiane. Si $p<\frac{1}{2}$ l'unique médiane est $m_1=0$: en effet $P(X \geq 0) = 1 > \frac{1}{2}$ et $P(X \leq 0) = P(X = 0) = 1-p > \frac{1}{2}$. Si $m>0$, on a en revanche $P(X \geq m) \leq P(X=1) = p < \frac{1}{2}$, donc m ne peut pas être une médiane. De manière analogue, on montre que si $p > \frac{1}{2}$ alors l'unique médiane est $m_1 = 1$.

(ii) La densité discrète de $X+Y$ est donnée par

$$\begin{aligned}
p_{X+Y}(0) &= P(X=0,Y=0) = (1-p)(1-q), \\
p_{X+Y}(2) &= P(X=1,Y=1) = pq, \\
p_{X+Y}(1) &= 1 - p_{X+Y}(2) - p_{X+Y}(0) = p(1-q)+q(1-p).
\end{aligned}$$

Observons que

- Si $(1-p)(1-q) > \frac{1}{2}$, alors $m=0$ est évidemment une médiane de $X+Y$, et elle est unique parce que si $m>0$ on a $P(X+Y>m) \leq P(X+Y>0) < \frac{1}{2}$.
- Si $(1-p)(1-q) = \frac{1}{2}$ alors tout $m \in [0,1[$ est une médiane.
- Si $(1-p)(1-q) < \frac{1}{2}$ on distingue trois sous-cas.
  - Si $pq > \frac{1}{2}$ on démontre comme ci-dessus que $m=2$ est l'unique médiane
  - Si $pq = \frac{1}{2}$ alors tout $m \in]1,2]$ est une médiane.
  - Si $pq < \frac{1}{2}$ alors $m=1$ est l'unique médiane. En effet, pour $m<1$ on a $P(X+Y \leq m) \leq P(X+Y=0) = (1-p)(1-q) < \frac{1}{2}$, et de la même manière pour $m>1$ on a $P(X+Y \geq m) < \frac{1}{2}$.

En conclusion, la médiane de $X+Y$ est unique si et seulement si $pq \neq \frac{1}{2}$ et $(1-p)(1-q) \neq \frac{1}{2}$. Notons pour finir que :

- si $p = q = \frac{3}{5}$ alors $m_1 = m_2 = 1$ mais $m = 1 < m_1 + m_2$;
- si $p = q = \frac{2}{5}$ alors $m_1 = m_2 = 0$ mais $m = 1 > m_1 + m_2$;
- si $p = q = \frac{4}{5}$ alors $m_1 = m_2 = 1$ et $m = 2 = m_1 + m_2$.

Exercice 3.25. Soit X une variable aléatoire discrète réelle et à valeurs dans $\mathbb{Z}^*$, de densité discrète donnée par $\mathrm{p}_X(x) = \frac{c_\alpha}{|x|^{1+\alpha}} \mathbb{1}_{\mathbb{Z}^*}(|x|)$, où $\alpha \in]0,+\infty[$ est un paramètre fixé et où la constante c_α vérifie

$$\frac{1}{c_\alpha} = \sum_{y \in \mathbb{Z}^*} \frac{1}{|y|^{1+\alpha}}.$$

Montrer que la fonction génératrice des moments M_X de X est donnée par $\mathrm{M}_X(0) = 1$ et $\mathrm{M}_X(t) = +\infty$ pour tout $t \neq 0$. En particulier, noter que M_X ne dépend pas de la valeur de α.

Solution 3.25. Commençons par considérer le cas $t > 0$. Alors

$$\mathrm{M}_X(t) = \sum_{x \in \mathbb{Z} \setminus \{0\}} \mathrm{e}^{tx} \frac{c_\alpha}{|x|^{1+\alpha}} \geq c_\alpha \sum_{x=1}^{+\infty} \frac{\mathrm{e}^{tx}}{|x|^{1+\alpha}} = +\infty,$$

où la dernière égalité vient du fait que

$$\lim_{x \to +\infty} \frac{\mathrm{e}^{tx}}{|x|^{1+\alpha}} = +\infty$$

(le terme général de la série ne tendant pas vers 0, la série est grossièrement divergente). De la même manière, pour $t = -s < 0$

$$\mathrm{M}_X(t) = \sum_{x \in \mathbb{Z} \setminus \{0\}} \mathrm{e}^{-sx} \frac{c_\alpha}{|x|^{1+\alpha}} \geq c_\alpha \sum_{x<0} \frac{\mathrm{e}^{-sx}}{|x|^{1+\alpha}} = c_\alpha \sum_{y=1}^{+\infty} \frac{\mathrm{e}^{sy}}{|y|^{1+\alpha}} = +\infty.$$

Exercice 3.26. Soit X une variable aléatoire réelle *positive*. Montrer que sa fonction génératrice des moments M_X est finie sur $]-\infty, 0]$ et que $\lim_{t \to -\infty} \mathrm{M}_X(t) = \mathrm{P}(X = 0)$.

[*Sugg.* Montrer que $\mathrm{P}(X=0) \leq \lim_{t \to -\infty} \mathrm{M}_X(t) \leq \mathrm{P}(X \leq 1/k)$ pour tout $k \geq 1$.]

Solution 3.26. Observons tout d'abord que si $X \geq 0$ et $t \leq 0$ alors $\mathrm{e}^{tX} \leq 1$: par monotonie de l'espérance (Proposition 3.54), il en découle que $\mathrm{M}_X(t) \leq 1$ pour tout $t \leq 0$. Pour la deuxième affirmation, on utilise les inégalités suivantes, qui sont valables pour tout $t < 0$ et $k \geq 1$, sous l'hypothèse $X \geq 0$:

$$\mathbb{1}_{\{X=0\}} \leq \mathrm{e}^{tX} = \mathbb{1}_{\{X>1/k\}} \mathrm{e}^{tX} + \mathbb{1}_{\{X \leq 1/k\}} \mathrm{e}^{tX} \leq \mathrm{e}^{t/k} + \mathbb{1}_{\{X \leq 1/k\}}.$$

De nouveau par monotonie de l'espérance, on obtient

$$\mathrm{P}(X=0) \leq \mathrm{M}_X(t) \leq \mathrm{e}^{t/k} + \mathrm{P}(X \leq 1/k).$$

Prenons maintenant la limite $t \to -\infty$ dans cette dernière expression. Soulignons qu'il s'agit d'une opération légitime car $\lim_{t \to -\infty} \mathrm{M}_X(t)$ existe, étant donné que $\mathrm{M}_X(t)$ est croissante en t, comme on peut le vérifier en remarquant que pour $s < t$ on a $\mathrm{e}^{sX} \leq \mathrm{e}^{tX}$ (en utilisant que $X \geq 0$) et en utilisant la monotonie de l'espérance. En passant à la limite, on trouve alors

$$\mathrm{P}(X=0) \le \lim_{t\to-\infty} \mathrm{M}_X(t) \le \mathrm{P}(X \le 1/k).$$

Comme cette dernière inégalité est valable pour tout $k \ge 1$, on peut prendre la limite quand $k \to +\infty$: grâce à la continuité par le haut des probabilités (Proposition 1.24 (iii)), on a

$$\lim_{k\to+\infty} \mathrm{P}(X \le 1/k) = \mathrm{P}(X=0),$$

dont la conclusion découle facilement.

Exercice 3.27. Soit X une variable aléatoire réelle à valeur dans $\mathbb{N}$. Montrer que sa fonction génératrice $s \mapsto \mathrm{G}_X(s)$ est croissante et convexe sur $[0,1]$. À quelle condition est-elle strictement croissante ? strictement convexe ?

Solution 3.27. Pour tout $n \in \mathbb{N}$ et pour tous $s,t \in [0,1]$ avec $s<t$ on a l'inégalité $s^n \le t^n$, et cette inégalité est stricte dès que $n \ge 1$. Ainsi, par monotonie de l'espérance (Proposition 3.54), on a

$$\mathrm{G}_X(s) = \mathrm{E}\left(s^X\right) \le \mathrm{E}\left(t^X\right) = \mathrm{G}_X(t).$$

De plus, en utilisant la Proposition 3.57, selon laquelle une variable aléatoire positive est d'espérance nulle si et seulement si elle est presque sûrement égale à 0, on a

$$\mathrm{G}_X(t) - \mathrm{G}_X(s) = \mathrm{E}\left(t^X - s^X\right) = 0 \iff \mathrm{P}(t^X = s^X) = 1 \iff \mathrm{P}(X=0) = 1.$$

On conclut donc que la fonction génératrice est strictement croissante dès que X *n'est pas* presque sûrement égale à 0.

Pour montrer que G_X est convexe, on doit vérifier que pour tout $0 \le s < t \le 1$ et $\alpha \in \,]0,1[$ on a

$$\mathrm{G}_X(\alpha s + (1-\alpha)t) \le \alpha \mathrm{G}_X(s) + (1-\alpha)\mathrm{G}_X(t);$$

si cette inégalité est stricte pour tout $0 \le s < t \le 1$ et $\alpha \in \,]0,1[$ alors G_X est strictement convexe. Observons que pour tout $n \in \mathbb{N}$, $0 \le s < t \le 1$ et $\alpha \in \,]0,1[$

$$(\alpha s + (1-\alpha)t)^n \le \alpha s^n + (1-\alpha)t^n,$$

et cette inégalité est stricte pour tout $n \ge 2$; cela découle du fait que la fonction $[0,1] \ni x \mapsto x^n$ est convexe pour $n \in \mathbb{N}$ et strictement convexe dès que $n \ge 2$. Ainsi,

$$\begin{aligned}\mathrm{G}_X(\alpha s + (1-\alpha)t) &= \mathrm{E}\left((\alpha s + (1-\alpha)t)^X\right) \\ &\le \alpha \mathrm{E}\left(s^X\right) + (1-\alpha)\mathrm{E}\left(t^X\right) = \alpha \mathrm{G}_X(s) + (1-\alpha)\mathrm{G}_X(t),\end{aligned}$$

ce qui démontre la convexité de G_X. De plus,

$$\alpha \mathrm{G}_X(s) + (1-\alpha)\mathrm{G}_X(t) - \mathrm{G}_X(\alpha s + (1-\alpha)t) = \mathrm{E}\left(\alpha s^X + (1-\alpha)t^X - (\alpha s + (1-\alpha)t)^X\right).$$

De nouveau, grâce à la Proposition 3.57, cette dernière espérance est nulle si et seulement si

$$\mathrm{P}\left(\alpha s^X + (1-\alpha)t^X - (\alpha s + (1-\alpha)t)^X = 0\right) = 1 \iff \mathrm{P}(X \in \{0,1\}) = 1.$$

On en conclut que G_X est strictement convexe dès que $\mathrm{P}(X \ge 2) > 0$.

Exercice 3.28 (Somme d'un nombre *aléatoire* de variables aléatoires). Soit X une variable aléatoire réelle à valeur dans $\mathbb{N}$ et soient $X_1, X_2, \ldots$ des variables aléatoires indépendantes et de même loi que X. On pose $S_0 = 0$ et $S_n := X_1 + \cdots + X_n$ pour $n \in \mathbb{N}^*$. Soit maintenant T

une variable aléatoire à valeurs dans $\mathbb{N}$, indépendante de $(X_k)_{k\geq 1}$: on considère la variable aléatoire S_T, définie par $S_T(\omega) = S_{T(\omega)}(\omega)$, c'est-à-dire

$$S_T(\omega) = 0 \quad \text{si } T(\omega) = 0\,, \qquad S_T(\omega) = \sum_{k=1}^{T(\omega)} X_k(\omega) \quad \text{si } T(\omega) \geq 1\,.$$

(i) Pour tout $n \in \mathbb{N}$, exprimer la fonction génératrice de S_n en fonction de celle de X.

(ii) Montrer que la fonction génératrice de S_T vérifie

$$\mathrm{G}_{S_T}(s) = \sum_{k\in\mathbb{N}}\sum_{n\in\mathbb{N}} \mathrm{p}_T(n)\,\mathrm{p}_{S_n}(k)s^k = \sum_{n\in\mathbb{N}} \mathrm{p}_T(n) G_{S_n}(s)\,, \qquad \forall s \in [0,1]\,.$$

En déduire que l'on a $\mathrm{G}_{S_T}(s) = \mathrm{G}_T(\mathrm{G}_X(s))$ pour tout $s \in [0,1]$.

(iii) On suppose que X et T admettent un moment d'ordre deux fini. En utilisant la Proposition 3.114, exprimer $\mathrm{E}(S_T)$ et $\mathrm{E}(S_T(S_T-1))$ en fonction de $\mathrm{E}(T)$, $\mathrm{E}(X)$, $\mathrm{E}(T^2)$ et $\mathrm{E}(X^2)$. Conclure que

$$\mathrm{E}(S_T) = \mathrm{E}(T)\,\mathrm{E}(X) \qquad \text{et} \qquad \mathrm{Var}(S_T) = \mathrm{E}(T)\,\mathrm{Var}(X) + \mathrm{Var}(T)\,\mathrm{E}(X)^2\,.$$

Solution 3.28. (i) Observons que

$$s^{S_n} = \prod_{k=1}^{n} s^{X_i}$$

est le produit de variables aléatoires indépendantes et de même loi : d'après la Proposition 3.72 (ou bien directement en utilisant la Proposition 3.116) on a

$$\mathrm{G}_{S_n}(s) = \mathrm{E}\left(s^{S_n}\right) = \prod_{k=1}^{n} \mathrm{E}\left(s^{X_i}\right) = \mathrm{G}_X(s)^n\,.$$

(ii) Écrivons, pour $s \in [0,1]$,

$$\mathrm{G}_{S_T}(s) = \mathrm{E}\left(s^{S_T}\right) = \sum_{k=0}^{+\infty} \mathrm{P}(S_T = k)s^k = \sum_{k=0}^{+\infty}\sum_{n=0}^{+\infty} \mathrm{P}(T = n, S_T = k)s^k\,,$$

et notons que $\mathrm{P}(T = n, S_T = k) = \mathrm{P}(T = n, S_n = k) = \mathrm{p}_T(n)\,\mathrm{p}_{S_n}(k)$, par indépendance de S_n et de T, pour tout $n \geq 0$. Ainsi,

$$\mathrm{G}_{S_T}(s) = \sum_{k=0}^{+\infty}\sum_{n=0}^{+\infty} \mathrm{p}_T(n)\,\mathrm{p}_{S_n}(k)s^k = \sum_{n=0}^{+\infty} \mathrm{p}_T(n)\left(\sum_{k=0}^{+\infty} \mathrm{p}_{S_n}(k)s^k\right) = \sum_{n=0}^{+\infty} \mathrm{p}_T(n)\mathrm{G}_{S_n}(s)\,,$$

où pour la deuxième égalité on a utilisé la sommation par paquet (0.13) pour des termes positifs (Fubini–Tonelli). D'après la question précédente, on a $G_{S_n}(s) = \mathrm{G}_X(s)^n$, donc

$$\mathrm{G}_{S_T}(s) = \sum_{n=0}^{+\infty} \mathrm{p}_T(n)\mathrm{G}_X(s)^n = \mathrm{G}_T\left(\mathrm{G}_X(s)\right)\,.$$

(iii) D'après la Proposition 3.114, on a $\mathrm{E}(S_T) = \lim_{t\uparrow 1} \mathrm{G}'_{S_T}(t)$ et $\mathrm{E}(S_T(S_T-1)) = \lim_{t\uparrow 1} \mathrm{G}''S_T(t)$. D'après ce que l'on a démontré au point précédent, on a

$$\mathrm{G}'_{S_T}(t) = \mathrm{G}'_T(\mathrm{G}_X(t))\,\mathrm{G}'_X(t)\,.$$

En dérivant une deuxième fois, on obtient

$$\mathrm{G}''_{S_T}(t) = \mathrm{G}''_T(\mathrm{G}_X(t))\,[\mathrm{G}'_X(t)]^2 + \mathrm{G}'_T(\mathrm{G}_X(t))\,\mathrm{G}''X(t).$$

En prenant la limite $t \uparrow 1$, en utilisant le fait que $\lim_{t\uparrow 1} \mathrm{G}_X(t)) = 1$ et de nouveau la Proposition 3.114 (pour G_T et G_X), on obtient

$$\mathrm{E}(S_T) = \lim_{t\uparrow 1} \mathrm{G}'_{S_T}(t) = \mathrm{E}(T)\,\mathrm{E}(X),$$

et

$$\begin{aligned}\mathrm{E}(S_T(S_T-1)) &= \lim_{t\uparrow 1} \mathrm{G}''S_T(t) = \mathrm{E}(T(T-1))\,\mathrm{E}(X)^2 + \mathrm{E}(T)\,\mathrm{E}(X(X-1)) \\ &= \mathrm{E}(T)\,\mathrm{Var}(X) + \mathrm{E}(T^2)\,\mathrm{E}(X)^2 - \mathrm{E}(T)\,\mathrm{E}(X)\,.\end{aligned}$$

On a donc

$$\begin{aligned}\mathrm{Var}(S_T) &= \mathrm{E}(S_T^2) - \mathrm{E}(S_T)^2 = \mathrm{E}(S_T(S_T-1)) + \mathrm{E}(S_T) - \mathrm{E}(S_T)^2 \\ &= \mathrm{E}(T)\,\mathrm{Var}(X) + \mathrm{E}(T^2)\,\mathrm{E}(X)^2 - \mathrm{E}(T)\,\mathrm{E}(X) + \mathrm{E}(T)\,\mathrm{E}(X) - \mathrm{E}(T)^2\,\mathrm{E}(X)^2 \\ &= \mathrm{E}(T)\,\mathrm{Var}(X) + \mathrm{Var}(T)\,\mathrm{E}(X)^2\,.\end{aligned}$$

3.5 Exemples importants de variables aléatoires discrètes

Exercice 3.29. Soit $X \sim \mathrm{Bin}(n,p)$. Avec un calcul direct à partir de l'expression (3.87) de la densité discrète, montrer que $\mathrm{E}(X(X-1)) = n(n-1)p^2$. En se rappelant que $\mathrm{E}(X) = np$, en déduire que l'on a $\mathrm{Var}(X) = np(1-p)$.

[*Sugg.* Appliquer la formule de transfert (3.31) et noter que $k(k-1)\binom{n}{k}p^k = p^2 \cdot \binom{n-2}{k-2}p^{k-2}$.]

Solution 3.29. D'après la formule de transfert, on a

$$\begin{aligned}\mathrm{E}(X(X-1)) &= \sum_{k=2}^{n} k(k-1)\binom{n}{k}p^k(1-p)^{n-k} \\ &= n(n-1)p^2 \sum_{k=2}^{n} \binom{n-2}{k-2}p^{k-2}(1-p)^{(n-2)-(k-2)} \\ &= n(n-1)p^2 \sum_{\ell=0}^{n-2} \mathrm{p}_{\mathrm{Bin}(n-2,p)}(\ell) = n(n-1)p^2\,.\end{aligned}$$

En écrivant $X^2 = X(X-1)+X$, on obtient par linéarité de l'espérance $\mathrm{E}(X^2) = \mathrm{E}(X(X-1)) + \mathrm{E}(X) = n^2p^2 - np^2 + np$, dont on déduit $\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = np - np^2 = np(1-p)$.

Exercice 3.30. Soit $X \sim \mathrm{Bin}(n,p)$. Avec un calcul direct à partir de l'expression (3.87) de la densité discrète, montrer que $\mathrm{M}_X(t) = (p\mathrm{e}^t + (1-p))^n$. En appliquant le Théorème 3.110, en déduire les valeurs de $\mathrm{E}(X)$ et $\mathrm{E}(X^2)$ et retrouver $\mathrm{Var}(X)$.

[*Sugg.* Appliquer la formule de transfert (3.31) et se rappeler du binôme de Newton (1.39).]

Solution 3.30. D'après la formule de transfert, on a

$$\mathrm{M}_X(t) = \sum_{k=0}^{n} \mathrm{e}^{tk}\binom{n}{k}p^k(1-p)^{n-k} = \sum_{k=0}^{n}\binom{n}{k}\big(p\mathrm{e}^t\big)^k(1-p)^{n-k} = \big(p\mathrm{e}^t + 1 - p\big)^n\,.$$

On obtient donc

$$\mathrm{M}'_X(t) = npe^t\left(pe^t + 1 - p\right)^{n-1},$$

et

$$\mathrm{M}''_X(t) = npe^t\left(pe^t + 1 - p\right)^{n-1} + n(n-1)p^2e^{2t}\left(pe^t + 1 - p\right)^{n-2},$$

dont on déduit $\mathrm{E}(X) = \mathrm{M}'_X(0) = np$, $\mathrm{E}(X^2) = \mathrm{M}''_X(0) = np(1-p) + n^2p^2$ et, pour finir, $\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = np(1-p)$.

Exercice 3.31. Soit $\lambda > 0$ un paramètre fixé. Pour $n \geq 1$, on considère X_n une variable aléatoire de loi $\mathrm{Bin}(n, \frac{\lambda}{n})$. Calculer la fonction génératrice G_{X_n} de X_n et montrer que

$$\lim_{n\to\infty} \mathrm{G}_{X_n}(s) = e^{\lambda(s-1)}, \qquad \forall s \in \mathbb{R}.$$

[*Sugg.* On pourra utiliser que $\ln(1+u) = u + o(u)$ quand $u \downarrow 0$.]

Solution 3.31. Si $X \sim \mathrm{Bin}(n,p)$, en utilisant la formule du binôme de Newton, on a

$$\mathrm{G}_X(s) = \sum_{k=0}^{n} s^k \binom{n}{k} p^k(1-p)^{n-k} = \sum_{k=0}^{n} \binom{n}{k} (sp)^k(1-p)^{n-k} = (sp+1-p)^n.$$

Ainsi, si $p = \frac{\lambda}{n}$:

$$\begin{aligned}\mathrm{G}_{X_n}(s) &= \left(1 + \frac{\lambda(s-1)}{n}\right)^n = \exp\left[n\log\left(1+\frac{\lambda(s-1)}{n}\right)\right] \\ &= \exp\left[n\left(\frac{\lambda(s-1)}{n} + o(1/n)\right)\right] = \exp\left[\lambda(s-1) + o(1)\right].\end{aligned}$$

ce qui donne la conclusion voulue.

Exercice 3.32. Soit $X \sim \mathrm{Poi}(\lambda)$. Avec un calcul direct à partir de l'expression (3.89) de la densité discrète, montrer que $\mathrm{E}(X) = \lambda$ et que $\mathrm{E}(X(X-1)) = \lambda^2$. En déduire que $\mathrm{Var}(X) = \lambda$.

Solution 3.32. Commençons par calculer $\mathrm{E}(X)$.

$$\mathrm{E}(X) = \sum_{n=0}^{+\infty} n\,\mathrm{p}_X(n) = e^{-\lambda}\sum_{n=1}^{+\infty} n\frac{\lambda^n}{n!} = \lambda e^{-\lambda}\sum_{n=1}^{+\infty}\frac{\lambda^{n-1}}{(n-1)!} = \lambda.$$

De la même manière

$$\begin{aligned}\mathrm{E}(X(X-1)) &= \sum_{n=0}^{+\infty} n(n-1)\,\mathrm{p}_X(n) = e^{-\lambda}\sum_{n=2}^{+\infty} n(n-1)\frac{\lambda^n}{n!} = \lambda^2 e^{-\lambda}\sum_{n=2}^{+\infty}\frac{\lambda^{n-2}}{(n-2)!} \\ &= \lambda^2.\end{aligned}$$

Il en découle que

$$\mathrm{E}(X^2) = \mathrm{E}(X(X-1)) + \mathrm{E}(X) = \lambda^2 + \lambda.$$

Pour finir,

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Exercice 3.33. Soient $X \sim \mathrm{Poi}(\lambda)$, $Y \sim \mathrm{Poi}(\mu)$ des variables aléatoires indépendantes.

(i) Calculer les fonctions génératrices G_X et G_Y de X et Y.

(ii) Calculer la fonction génératrice G_{X+Y} de $X+Y$. En déduire, grâce à la Proposition 3.114, que l'on a $X+Y \sim \mathrm{Poi}(\lambda+\mu)$.

Solution 3.33. (i) Si $X \sim \mathrm{Poi}(\lambda)$:

$$\mathrm{G}_X(s) = \sum_{n=0}^{+\infty} s^n \mathrm{e}^{-\lambda} \frac{\lambda^n}{n!} = \mathrm{e}^{\lambda(s-1)} .$$

De même, $\mathrm{G}_Y(s) = \mathrm{e}^{\mu(s-1)}$.

(ii) D'après la Proposition 3.116,

$$\mathrm{G}_{X+Y}(s) = \mathrm{G}_X(s)\mathrm{G}_Y(s) = \mathrm{e}^{(\lambda+\mu)(s-1)}.$$

On montre facilement par récurrence que la dérivée n-ème de G_{X+Y} est donnée par

$$\mathrm{G}_{X+Y}^{(n)}(s) = (\lambda+\mu)^n \mathrm{e}^{(\lambda+\mu)(s-1)}.$$

Alors, grâce à la Proposition 3.114,

$$\mathrm{p}_{X+Y}(n) = \frac{\mathrm{G}_{X+Y}^{(n)}(0)}{n!} = \mathrm{e}^{-(\lambda+\mu)} \frac{(\lambda+\mu)^n}{n!} ,$$

dont on déduit que $X+Y \sim \mathrm{Poi}(\lambda+\mu)$.

Exercice 3.34. Soient $X \sim \mathrm{Poi}(\lambda)$, $Y \sim \mathrm{Poi}(\mu)$ des variables aléatoires indépendantes. Pour $n \geq 0$ donné, déterminer la loi de la variable aléatoire X par rapport à la probabilité conditionnelle $\mathrm{P}(\cdot\,|X+Y=n)$, c'est-à-dire déterminer la densité discrète $\mathrm{q}(x) := \mathrm{P}(X=x|X+Y=n)$, et reconnaître une loi usuelle.

Solution 3.34. Pour $k \in \mathbb{N}$, on a

$$\mathrm{q}(k) := \mathrm{P}(X=k|X+Y=n) = \frac{\mathrm{P}(X=k, Y=n-k)}{\mathrm{P}(X+Y=n)} = \frac{\mathrm{P}(X=k)\,\mathrm{P}(Y=n-k)}{\mathrm{P}(X+Y=n)} ,$$

par indépendance de X et Y. D'après la Proposition 3.128, on a $X+Y \sim \mathrm{Poi}(\lambda+\mu)$, donc

$$\mathrm{q}(k) = \frac{\mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}\mathrm{e}^{-\mu}\frac{\mu^{n-k}}{(n-k)!}}{\mathrm{e}^{-(\lambda+\mu)}\frac{(\lambda+\mu)^n}{n!}} = \binom{n}{k} p^k (1-p)^{n-k} , \qquad \text{avec} \qquad p := \frac{\lambda}{\lambda+\mu} ,$$

où par convention $\binom{n}{k} = 0$ pour $k > n$. Ainsi, par rapport à $\mathrm{P}(\cdot\,|X+Y=n)$, la variable aléatoire X est de loi $\mathrm{Bin}(n,p)$.

Exercice 3.35. Soit X une variable aléatoire de loi $\mathrm{Géom}(p)$ où $p \in \,]0,1[$. Pour $k \in \mathbb{N}^*$, calculer la probabilité que X soit un multiple de k.

Solution 3.35. La probabilité demandée est égale à

$$\begin{aligned}\sum_{n=1}^{+\infty} \mathrm{P}(X=nk) = \sum_{n=1}^{+\infty} p(1-p)^{nk-1} &= \frac{p}{1-p} \sum_{n=1}^{+\infty} \left[(1-p)^k\right]^n \\ &= \frac{p}{1-p} \frac{(1-p)^k}{1-(1-p)^k} = \frac{p(1-p)^{k-1}}{1-(1-p)^k} ,\end{aligned}$$

où on a utilisé la série géométrique (0.5).

Exercice 3.36. Soient X, Y des variables aléatoires indépendantes de loi $\text{Géom}(p)$. Pour $n \geq 2$ donné, déterminer la loi de la variable aléatoire X par rapport à la probabilité conditionnelle $\mathrm{P}(\cdot \,|\, X+Y=n)$, c'est-à-dire déterminer la densité discrète $\mathrm{q}(x) := \mathrm{P}(X=x \,|\, X+Y=n)$, et reconnaître une loi usuelle.

Solution 3.36. Pour $k \in \mathbb{N}^*$, on a

$$\mathrm{q}(k) \;:=\; \mathrm{P}(X=k \,|\, X+Y=n) = \frac{\mathrm{P}(X=k, Y=n-k)}{\mathrm{P}(X+Y=n)} = \frac{\mathrm{P}(X=k)\,\mathrm{P}(Y=n-k)}{\mathrm{P}(X+Y=n)}\,,$$

par indépendance de X et Y. Observons que

$$\mathrm{P}(X+Y=n) = p^2 \sum_{k=1}^{n-1} (1-p)^{k-1}(1-p)^{n-k-1} = (n-1)p^2(1-p)^{n-2}\,,$$

dont on déduit

$$\mathrm{q}(k) = \frac{p(1-p)^{k-1}p(1-p)^{n-k-1}}{(n-1)p^2(1-p)^{n-2}} = \frac{1}{n-1} \qquad \text{pour} \qquad k \in \{1, \ldots, n-1\}\,,$$

alors que $q(k) = 0$ sinon. Ainsi, par rapport à $\mathrm{P}(\cdot \,|\, X+Y=n)$ la variable aléatoire X est de loi $\text{Unif}(\{1, \ldots, n-1\})$.

3.6 Exercices récapitulatifs

Exercice 3.37. Pour effectuer un contrôle de qualité, on choisit *au hasard* un échantillon de 5 objets dans un lot de 100 objets parmi lesquels 10 sont défectueux. Soit X le nombre d'objets défectueux contenus dans l'échantillon. Déterminer la densité discrète de X.

Solution 3.37. La variable aléatoire X peut seulement prendre les valeurs $0, 1, 2, 3, 4, 5$. Avoir $X = k$ signifie que dans le lot de 5 objets, k sont défectueux. On peut calculer $\mathrm{P}(X=k)$ avec la formule « cas favorables sur cas possibles » de la probabilité uniforme. *Cas possibles* : il y a $\binom{100}{5}$ façons de choisir 5 objets parmi 100. *Cas favorables* : on doit choisir 5 objets dont k sont défectueux. On choisit d'abord les k qui sont défectueux parmi les 10 défectueux, il y a $\binom{10}{k}$ façons de le faire ; puis on choisit les $5-k$ objets restants parmi les $100-10=90$ objets non défectueux restants, il y a $\binom{90}{5-k}$ façons de le faire. Au final, il y a $\binom{10}{k}\binom{90}{5-k}$ cas favorables et donc :

$$\mathrm{P}(X=k) = \frac{\binom{10}{k}\binom{90}{5-k}}{\binom{100}{5}}, \qquad k = 0, 1, \ldots, 5\,.$$

À l'aide d'une calculatrice on obtient : $\mathrm{P}(X=0) \simeq 0{,}583$, $\mathrm{P}(X=1) \simeq 0{,}340$, $\mathrm{P}(X=2) \simeq 0{,}070$, $\mathrm{P}(X=3) \simeq 0{,}007$, $\mathrm{P}(X=4) \simeq \mathrm{P}(X=5) \simeq 0$. Pour être sûr de ne pas s'être trompé dans les calculs, on peut vérifier que $\sum_{k=0}^{5} \mathrm{P}(X=k) = 1$.

Exercice 3.38. Un jeu d'argent possède un gros lot d'un montant de 512 euros. On pose à un concurrent 10 questions. À chaque réponse fausse ce montant est divisé par deux. À la première réponse correcte, le concurrent gagne le montant restant. S'il ne donne aucune bonne réponse, il ne gagne rien. Un concurrent donné répond correctement à chacune des question avec probabilité $p \in \,]0,1[$, indépendamment des réponses aux autres questions. Soit X le montant (en euros) remporté par ce concurrent. Déterminer la densité discrète p_X de X.

Solution 3.38. La variable aléatoire X prend les valeurs $0,1,2,2^2,2^3,\dots,2^9$. Pour $0 \le k \le 9$, X prend la valeur 2^{9-k} si les k premières réponses sont fausses et la $k+1$-ème est correcte. Cela arrive avec probabilité $(1-p)^k p$. Pour finir, X prend la valeur 0 si toutes les 10 réponses sont fausses, ce qui arrive avec probabilité $(1-p)^{10}$. Pour résumer, on a

$$\begin{aligned} \mathrm{p}_X(2^{9-k}) &= p(1-p)^k \qquad \text{pour } 0 \le k \le 9\,. \\ \mathrm{p}_X(0) &= (1-p)^{10}\,. \end{aligned}$$

Exercice 3.39. Alain joue à un jeu contre Marius. Ils lancent chacun (de manière indépendante) un dé équilibré à 6 faces : on note X le résultat d'Alain et Y le résultat de Marius. Si Alain obtient un résultat strictement plus grand que celui de Marius, c'est-à-dire si $X > Y$, il gagne la différence entre les deux dés (en euros). Sinon, il perd 2 euros. On note W le gain d'Alain (avec signe) au jeu. Déterminer la densité discrète de W, puis calculer $\mathrm{E}(W)$.

Solution 3.39. Le vecteur aléatoire (X,Y) est de loi uniforme sur l'ensemble $\{1,2,3,4,5,6\}^2$. En notant que

$$|\{(i,j) : 1 \le i \le j \le 6\}| = 21,$$

on a

$$\mathrm{p}_W(-2) = \mathrm{P}(X \le Y) = \mathrm{P}((X,Y) \in \{(i,j) : 1 \le i \le j \le 6\}) = \frac{21}{36}.$$

De plus

$$\mathrm{p}_W(1) = \mathrm{P}(W=1) = \mathrm{P}((X,Y) \in \{(2,1),(3,2),(4,3),(5,4),(6,5)\}) = \frac{5}{36}.$$

Avec la même méthode, on obtient $\mathrm{p}_W(2) = \frac{4}{36}$, $\mathrm{p}_W(3) = \frac{3}{36}$, $\mathrm{p}_W(4) = \frac{2}{36}$, $\mathrm{p}_W(5) = \frac{1}{36}$. Enfin,

$$\mathrm{E}(W) = -2\frac{21}{36} + \frac{5}{36} + 2\frac{4}{36} + 3\frac{3}{36} + 4\frac{2}{36} + 5\frac{1}{36} = -\frac{7}{36}.$$

Exercice 3.40. Soit X une variable aléatoire de loi $\text{Géom}(p)$ où $p \in\,]0,1[$. Calculer une médiane m de X. Pour quelles valeurs de p est-elle unique ? Montrer que $\mathrm{E}(X) - m \to +\infty$ quand $p \to 0$.

Solution 3.40. Rappelons que si $X \sim \text{Géom}(p)$ et $n \in \mathbb{N}$, alors $\mathrm{P}(X > n) = (1-p)^n$. On en déduit donc que

$$\mathrm{P}(X \le n) \le \frac{1}{2} \iff (1-p)^n \le \frac{1}{2} \iff n \ge -\frac{\log 2}{\log(1-p)},$$

Remarquons que $-\frac{\log 2}{\log(1-p)} > 0$. Si $-\frac{\log 2}{\log(1-p)} \notin \mathbb{N}$, notons m le plus petit entier supérieur à $-\frac{\log 2}{\log(1-p)}$. Alors, comme $-\frac{\log 2}{\log(1-p)} - 1 < m < -\frac{\log 2}{\log(1-p)}$,

$$\mathrm{P}(X \le m) > \frac{1}{2} \ \text{ et } \ \mathrm{P}(X \le m-1) < \frac{1}{2}.$$

Ainsi,

$$\mathrm{P}(X \ge m) = 1 - \mathrm{P}(X \le m-1) > \frac{1}{2},$$

on en déduit que m est une médiane. Si $x < m$

$$\mathrm{P}(X \le x) \le \mathrm{P}(X \le m-1) < \frac{1}{2},$$

donc x n'est pas une médiane. Si en revanche $x > m$

$$\mathrm{P}(X \geq x) \leq \mathrm{P}(X > m) = 1 - \mathrm{P}(X \leq m) < \frac{1}{2},$$

et donc x n'est pas une médiane. En conclusion, si $-\frac{\log 2}{\log(1-p)} \notin \mathbb{N}$, la médiane est unique.

Si en revanche $-\frac{\log 2}{\log(1-p)} = k \in \mathbb{N}$, on a

$$\mathrm{P}(X \leq k) = \frac{1}{2}.$$

Dans ce cas, tout $m \in [k, k+1[$ est une médiane. En effet,

$$\mathrm{P}(X \leq m) \geq \mathrm{P}(X \leq k) = \frac{1}{2},$$

et

$$\mathrm{P}(X \geq m) = 1 - \mathrm{P}(X \leq k) = \frac{1}{2}.$$

En répétant les arguments précédents, il n'est pas difficile de vérifier que si $x \notin [k, k+1[$ alors x n'est pas une médiane.

Dans tous les cas, on a montré plus haut que si m est une médiane alors $m < -\frac{\log 2}{\log(1-p)} + 1$. Ainsi, avec un développement limité,

$$\begin{aligned} \mathrm{E}(X) - m \geq \frac{1}{p} + \frac{\log 2}{\log(1-p)} - 1 &= \frac{\log(1-p) + p\log 2}{p\log(1-p)} - 1 \\ &= \frac{(\log 2 - 1)p + o(p)}{-p^2 + o(p^2)} - 1 = \frac{1 - \log 2}{p} + o\left(\frac{1}{p}\right) \end{aligned}$$

qui tend vers $+\infty$ quand $p \downarrow 0$.

Exercice 3.41. Soient X_1, X_2 des variables aléatoires indépendantes, chacune de loi uniforme discrète sur $\{1, \ldots, n\}$. pour un entier $n \in \mathbb{N}^*$ fixé. On définit la variable $Y := \min\{X_1, X_2\}$.

(i) Calculer $\mathrm{P}(Y = k)$ pour tout $k \in \mathbb{N}$.

(ii) Montrer que pour tout $t \in]0, 1[$, on a

$$\lim_{n \to \infty} \mathrm{P}(Y \leq tn) \;=\; 2t - t^2.$$

Solution 3.41. (i) Clairement, $\mathrm{P}(Y = k) = 0$ pour $k > n$. Commençons par calculer, pour $k \in \{1, \ldots, n\}$, la probabilité $\mathrm{P}(Y \geq k)$ qui est donnée par

$$\begin{aligned} \mathrm{P}(Y \geq k) \;&=\; \mathrm{P}(X_1 \geq k, X_2 \geq k) \;=\; \mathrm{P}(X_1 \geq k)\,\mathrm{P}(X_2 \geq k) \\ &=\; \left(\frac{n-k+1}{n}\right)^2 \;=\; \left(1 - \frac{k-1}{n}\right)^2. \end{aligned}$$

On a donc

$$\begin{aligned} \mathrm{P}(Y = k) \;&=\; \mathrm{P}(Y \geq k) - \mathrm{P}(Y \geq k+1) \;=\; \left(\frac{n-k+1}{n}\right)^2 - \left(\frac{n-k}{n}\right)^2 \\ &=\; \frac{1}{n^2} + \frac{2}{n}\left(1 - \frac{k}{n}\right). \end{aligned}$$

(ii) D'après la formule pour $\mathrm{P}(Y \geq k)$ calculée au point précédent, on obtient

$$\mathrm{P}(Y \leq tn) \;=\; 1-\mathrm{P}(Y > tn) \;=\; 1-\mathrm{P}(Y > \lfloor tn \rfloor) \;=\; 1-\left(1-\frac{\lfloor tn \rfloor - 1}{n}\right)^2 .$$

Comme $(\lfloor tn \rfloor - 1)/n \to t$ quand $n \to \infty$, pour tout $t \in]0,1[$ on a

$$\lim_{n\to\infty} \mathrm{P}(Y \leq tn) \;=\; 1-(1-t)^2 \;=\; 2t-t^2 .$$

Exercice 3.42. En utilisant l'approximation de Poisson (Proposition 3.127), estimer la probabilité que parmi 250 couples choisis au hasard dans la population, il y en ait au moins un dont les deux partenaires ont le même anniversaire. De même, estimer la probabilité qu'il y ait au moins deux (puis au moins trois) couples dont les deux partenaires ont le même anniversaire.

Solution 3.42. En supposant qu'une année soit composée de 365 jours, en supposant l'indépendance et l'équiprobabilité des dates d'anniversaire d'individus différents, si X et Y désignent les dates d'anniversaire des deux membres d'un couple, on a $X, Y \sim \mathrm{Unif}(\{1,2,\ldots,365\})$ avec X et Y indépendantes. Ainsi,

$$\mathrm{P}(X=Y) = \sum_{i=1}^{365} \mathrm{P}(X=i, Y=i) = 365 \cdot \frac{1}{365^2} = \frac{1}{365} .$$

En notant N le nombre de couples, parmi les 250 choisis, qui ont leur anniversaire le même jour, on a $N \sim \mathrm{Bin}(250, p)$, où $p = \frac{1}{365}$. En utilisant l'approximation de Poisson, on peut approcher la loi de N par celle d'une variable aléatoire de Poisson de paramètre $\lambda = \frac{250}{365} \simeq 0{,}685$. Ainsi,

$$\mathrm{P}(N \geq 1) = 1-\mathrm{P}(N=0) \simeq 1-\mathrm{e}^{-\lambda} \simeq 0{,}496 .$$

De manière analogue,

$$\begin{aligned}
\mathrm{P}(N \geq 2) &= 1-\mathrm{P}(N=0)-\mathrm{P}(N=1) \simeq 1-\mathrm{e}^{-\lambda}-\lambda\mathrm{e}^{-\lambda} \simeq 0{,}151 ,\\
\mathrm{P}(N \geq 3) &= 1-\mathrm{P}(N=0)-\mathrm{P}(N=1)-\mathrm{P}(N=2) \simeq 1-\mathrm{e}^{-\lambda}-\lambda\mathrm{e}^{-\lambda}-\frac{\lambda^2}{2}\mathrm{e}^{-\lambda} \simeq 0{,}032 .
\end{aligned}$$

Exercice 3.43. Deux paquets de 40 cartes sont chacun constitués de 20 cartes rouges et 20 cartes noires. On mélange chacun des deux paquets puis on les dispose l'un à côté de l'autre. On commence par dévoiler la première carte de chaque paquet. Si les deux sont rouges on gagne un euro, sinon on ne gagne rien. On poursuit en dévoilant la deuxième carte de chaque paquet : si les deux cartes sont rouges on gagne un autre euro, et ainsi de suite. Soit X le nombre d'euros gagnés après avoir dévoilé toutes les cartes. Calculer la densité discrète de X.

Solution 3.43. Pour chacun des deux paquets, on numérote les cartes de 1 à 40, et on convient que les rouges sont celles numérotées de 1 à 20. En notant $\sigma(i)$ la position de la carte numéro i du premier paquet après le mélange, la disposition des cartes du premier paquet peut être identifiée à $\sigma \in \mathfrak{S}_{40}$. De même, on note $\eta \in \mathfrak{S}_{40}$ la disposition des cartes du deuxième paquet. On peut donc prendre $\Omega := \mathfrak{S}_{40} \times \mathfrak{S}_{40}$, muni de la probabilité uniforme, et $X(\sigma,\eta)$ désigne le nombre d'euros gagnés si les paquets ont pour disposition σ, η. Pour $k = 0,1,\ldots,20$,

$$\mathrm{P}(X=k) = \frac{|\{(\sigma,\eta) :\; X(\sigma,\eta)=k\}|}{40!^2} .$$

Pour calculer $|\{(\sigma,\eta) :\; X(\sigma,\eta)=k\}|$ on utilise le schéma des choix successifs : on choisit tout d'abord σ (il y a 40! choix possibles), puis on choisit η de façon à avoir $X(\sigma,\eta)=k$.

Pour déterminer le nombre de choix possibles pour η, soit $I_\sigma := \{\sigma(1), \sigma(2), \dots, \sigma(20)\}$ l'ensemble des positions des cartes rouges du premier paquet. Une permutation η telle que $X(\sigma, \eta) = k$ peut être choisie de la manière suivante.

- On choisit k cartes rouges (du deuxième paquet) : il y a $\binom{20}{k}$ possibilités.
- On dispose ces cartes dans I_σ : il y a $\binom{20}{k}k!$ possibilités.
- On dispose les $20-k$ cartes rouges restantes dans $(I_\sigma)^c$: il y a $\binom{20}{20-k}(20-k)!$ possibilités.
- On dispose les 20 cartes noires dans les 20 positions restantes : il y a 20! possibilités.

Pour résumer

$$|\{X(\sigma,\eta)=k\}| = 40!\binom{20}{k}\binom{20}{k}k!\binom{20}{20-k}(20-k)!20! = 40!(20!)^2\binom{20}{k}^2$$

dont ont déduit que

$$\mathrm{P}(X=k) = \frac{|\{X(\sigma,\eta)=k\}|}{|\Omega|} = \frac{40!(20!)^2\binom{20}{k}^2}{(40!)^2} = \frac{\binom{20}{k}^2}{\binom{40}{20}}.$$

Calculons maintenant l'espérance $\mathrm{E}(X)$. On peut très bien le faire en utilisant l'expression de la densité discrète de X que l'on vient d'obtenir et la définition de l'espérance. Il est cependant plus astucieux d'utiliser la symétrie du problème : $\mathrm{P}(X=k) = \mathrm{P}(X=20-k)$ pour tout $k \in \{0,1,\dots,20\}$, comme on le déduit du fait que $\binom{20}{k} = \binom{20}{20-k}$. Cela peut s'exprimer par le fait que les variables X et $20-X$ ont la même loi, donc la même espérance ! On a ainsi $\mathrm{E}(X) = \mathrm{E}(20-X) = 20-\mathrm{E}(X)$, dont on déduit directement que $\mathrm{E}(X) = 10$.

Exercice 3.44. Une urne contient n boules blanches et 2 boules rouges, où $n \in \mathbb{N}^*$ est un entier fixé. On procède à un tirage *sans remise* de toutes les boules de l'urne. Introduisons la variable aléatoire $X =$ « nombre de boules blanches tirées avant d'avoir une boule rouge », dont la densité discrète sera notée $\mathrm{p}_X(k) = \mathrm{P}(X=k)$.

(i) Montrer que, pour $k \in \{0,1,\dots,n\}$, on a

$$\mathrm{p}_X(k) = \frac{2}{(n+2)(n+1)}(n-k+1).$$

(ii) Calculer $\mathrm{E}(X)$.

[*Sugg.* On pourra utiliser la formule (0.8) pour la somme des carrés des premiers entiers.]

Solution 3.44. (i) Considérons les événements $A =$ « la $k+1$-ème boule tirée est rouge », $B =$ « les k premières boules tirées sont toutes blanches ». On a

$$\begin{aligned}\mathrm{P}(X=k) &= \mathrm{P}(A\cap B) = \mathrm{P}(A\,|\,B)\,\mathrm{P}(B) \\ &= \frac{2}{n-k+2}\frac{\binom{n}{k}}{\binom{n+2}{k}} = \frac{2}{(n+2)(n+1)}(n-k+1),\end{aligned}$$

où dans le dernier passage on a effectué les simplifications adéquates.

(ii) On a

$$\begin{aligned}
\mathrm{E}(X) &= \sum_{k=0}^{n} k\, \mathrm{p}_X(k) = \frac{2}{(n+2)(n+1)} \sum_{k=0}^{n} k(n-k+1) \\
&= \frac{2}{n+2} \sum_{k=1}^{n} k - \frac{2}{(n+2)(n+1)} \sum_{k=1}^{n} k^2 = \frac{n(n+1)}{n+2} - \frac{n(2n+1)}{3(n+2)} = \frac{n}{3},
\end{aligned}$$

où dans le dernier passage on a utilisé (0.8).

Exercice 3.45. Soit X une variable aléatoire à valeurs dans $\mathbb{Z}$, de densité discrète donnée par

$$\mathrm{P}(X=k) = c_q\, q^{|k|}\,, \qquad \text{pour } k \in \mathbb{Z}$$

où $q \in \,]0,1[$ est un paramètre donné et c_q est une constante qui dépend de q.

(i) Déterminer la valeur de la constante c_q en fonction de q.

(ii) Calculer la fonction génératrice des moments de X.

(iii) Calculer $\mathrm{E}(Y)$ et $\mathrm{Var}(Y)$.

Solution 3.45. (i) En se rappelant de la série géométrique (0.5), on obtient

$$\sum_{k\in\mathbb{Z}} q^{|k|} = \sum_{k=0}^{+\infty} q^k + \sum_{k=1}^{+\infty} q^k = \frac{1}{1-q} + \frac{q}{1-q} = \frac{1+q}{1-q}\,,$$

dont il découle que $c_q = \frac{1-q}{1+q}$.

(ii) On a

$$\mathrm{M}_X(t) = \mathrm{E}\left(\mathrm{e}^{tX}\right) = c_q \sum_{k\in\mathbb{Z}} \mathrm{e}^{tk} q^{|k|} = c_q \sum_{k=0}^{+\infty} \left(q\mathrm{e}^{t}\right)^k + c_q \sum_{k=1}^{+\infty} \left(q\mathrm{e}^{-t}\right)^k .$$

En utilisant la série géométrique (0.5), on obtient que pour tout t tel que $|q\mathrm{e}^t| < 1$ et $|q\mathrm{e}^{-t}| < 1$, c'est-à-dire pour tout $t \in \,]-\log\frac{1}{q}, \log\frac{1}{q}[$

$$\mathrm{M}_X(t) = c_q \left(\frac{1}{1-q\mathrm{e}^t} + \frac{q\mathrm{e}^{-t}}{1-q\mathrm{e}^{-t}}\right) = \frac{c_q(1-q^2)}{1+q^2-2q\cosh(t)} = \frac{(1-q)^2}{1+q^2-2q\cosh(t)}\,.$$

Pour $t \notin \,]-\log\frac{1}{q}, \log\frac{1}{q}[$, il est facile de vérifier que l'une des deux sommes est infinie, donc $\mathrm{M}_X(t) = +\infty$.

(iii) On a

$$\begin{aligned}
\mathrm{M}_X'(t) &= \frac{(1-q)^2}{(1+q^2-2q\cosh(t))^2} 2q\sinh(t)\,, \\
\mathrm{M}_X''(t) &= \frac{(1-q)^2}{(1+q^2-2q\cosh(t))^3} 8q^2\sinh^2(t) + \frac{(1-q)^2}{(1+q^2-2q\cosh(t))^2} 2q\cosh(t)\,.
\end{aligned}$$

Ainsi,

$$\mathrm{E}(X) = \mathrm{M}_X'(0) = 0, \qquad \mathrm{Var}(X) = \mathrm{E}(X^2) = \mathrm{M}_X''(0) = \frac{2q}{(1-q)^2}\,.$$

Exercice 3.46. On rappelle que pour une variable aléatoire discrète X à valeurs dans $\mathbb{Z}$, la fonction génératrice des moments s'écrit sous la forme $\mathrm{M}_X(t) = \sum_{n\in\mathbb{Z}} \mathrm{e}^{nt}\, \mathrm{p}_X(n)$. Déterminer la densité discrète des variables aléatoires dont les fonctions génératrices des moments sont les suivantes :

(i) $\mathrm{M}_X(t) = \frac{1}{4}(1+\mathrm{e}^t)^2$;

(ii) $\mathrm{M}_X(t) = \dfrac{1}{2-\mathrm{e}^t}$;

(iii) $\mathrm{M}_X(t) = \cosh^2(t)$;

(iv) $\mathrm{M}_X(t) = \sum_{k=0}^{+\infty} a_k \cosh(kt)$, où on a $a_k \geq 0$ pour tout k et $\sum_{k=0}^{+\infty} a_k = 1$.

Solution 3.46. (i)

$$\mathrm{M}_X(t) = \frac{1}{4} + \frac{1}{2}\mathrm{e}^t + \frac{1}{4}\mathrm{e}^{2t} \;\Rightarrow\; \mathrm{p}_X(0) = \frac{1}{4} = \mathrm{p}_X(2),\ \mathrm{p}_X(1) = \frac{1}{2}.$$

(ii) En utilisant (0.5) :

$$\mathrm{M}_X(t) = \frac{1}{2}\frac{1}{1-\frac{\mathrm{e}^t}{2}} = \frac{1}{2}\sum_{n=0}^{+\infty}\frac{\mathrm{e}^{nt}}{2^n} \;\Rightarrow\; \mathrm{p}_X(n) = \frac{1}{2^{n+1}}.$$

(iii)

$$\mathrm{M}_X(t) = \frac{\mathrm{e}^t + \mathrm{e}^{-t}}{2} \;\Rightarrow\; \mathrm{p}_X(1) = \mathrm{p}_X(-1) = \frac{1}{2}.$$

(iv)

$$\mathrm{M}_X(t) = \left(\frac{\mathrm{e}^t+\mathrm{e}^{-t}}{2}\right)^2 = \frac{1}{4}\mathrm{e}^{2t} + \frac{1}{4}\mathrm{e}^{-2t} + \frac{1}{2} \;\Rightarrow\; \mathrm{p}_X(0) = \frac{1}{2}, \mathrm{p}_X(2) = \mathrm{p}_X(-2) = \frac{1}{4}.$$

(v)

$$\mathrm{M}_X(t) = \sum_{k=0}^{+\infty} a_k \frac{\mathrm{e}^{kt}+\mathrm{e}^{-kt}}{2}.$$

On en déduit que $\mathrm{p}_X(0) = a_0$ et, pour $k \in \mathbb{Z}\setminus\{0\}$,

$$\mathrm{p}_X(k) = \frac{a_{|k|}}{2}.$$

Exercice 3.47. Soit $\Omega_n := \{0,1\}^n$ et $\Omega := \bigcup_{n=1}^{\infty}\Omega_n$. Autrement dit, les éléments de Ω sont les suites binaires de longueur arbitraire mais finie. Si $\omega \in \Omega$, l'unique $n \geq 1$ pour lequel $\omega \in \Omega_n$ est appelée *longueur* de ω, et sera noté $\ell(\omega)$. Pour tout $\omega \in \Omega$, on définit

$$\mathrm{p}(\omega) := 2^{-2\ell(\omega)}.$$

(i) Montrer que p est une densité discrète sur Ω, au sens de la Définition 1.14. Ainsi, grâce à la relation (1.5), elle identifie une probabilité P.

(ii) Sur l'espace probabilisé (Ω, P), on définit pour tout $i \geq 1$ la variable aléatoire

$$X_i(\omega) := \begin{cases} \omega_i & \text{si } i \leq \ell(\omega) \\ 0 & \text{sinon} \end{cases}$$

Déterminer la loi de X_i.

Solution 3.47. (i) La fonction p est une densité discrète si $\mathrm{p}(\omega) \geq 0$ pour tout ω (ce qui est ici évident) et $\sum_{\omega\in\Omega}\mathrm{p}(\omega) = 1$. Mais comme $|\Omega_n| = 2^n$, on obtient

$$\sum_{\omega\in\Omega} 2^{-2\ell(\omega)} = \sum_{n\geq 1}\sum_{\omega\in\Omega_n} 2^{-2n} = \sum_{n\geq 1} 2^{-2n}|\Omega_n| = \sum_{n\geq 1} 2^{-n} = 1,$$

en ayant utilisé la formule (0.5) de la série géométrique.

(ii) Clairement, la variable aléatoire X_i prend seulement la valeur 0 ou 1. De plus,

$$\begin{aligned}\mathrm{P}(X_i = 1) &= \sum_{n\geq i}\sum_{\omega\in\Omega_n:\omega_i=1} 2^{-2\ell(\omega)} = \sum_{n\geq i} 2^{-2n}|\{\omega\in\Omega_n : \omega_i = 1\}|\\ &= \sum_{n\geq i} 2^{-2n}2^{n-1} = \frac{1}{2}\sum_{n\geq i} 2^{-n} = 2^{-i}.\end{aligned}$$

Autrement dit, $X_i \sim \mathrm{Bern}(2^{-i})$.

Exercice 3.48. Une urne contient initialement une boule noire et on ajoute des boules blanches dans l'urne, les unes après les autres, de la manière suivante. À chaque tour $n = 1,2,3,\ldots$, l'urne contient n boules blanches et une boule noire : on tire alors une boule au hasard, on note sa couleur, et on la remet dans l'urne. On suppose que les tirages lors des différents tours sont effectués de manière indépendante. On note T le numéro du tour où le joueur tire une boule noire pour la première fois.

(i) Montrer que pour tout $n\in\mathbb{N}$ on a $\mathrm{P}(T > n) = \frac{1}{n+1}$. En déduire que l'on a $\mathrm{P}(T = \infty) = 0$.

(ii) Montrer que tout $m \in [1,2]$ est une médiane de T.

(iii) Calculer la densité discrète de T. Que vaut $\mathrm{E}(T)$?

Solution 3.48. (i) La probabilité de tirer une boule blanche au k-ème tirage vaut $\frac{k}{k+1}$. Ainsi, en remarquant que l'événement $\{T > n\}$ équivaut à ne tirer que des boules blanches lors des n premiers tirages, on a (par indépendance des tirages)

$$\mathrm{P}(T > n) = \prod_{k=1}^{n} \frac{k}{k+1} = \frac{1}{n+1}.$$

Par continuité des probabilités par le haut (Proposition 1.24 iii),

$$\mathrm{P}(T = +\infty) = \lim_{n\to+\infty} \mathrm{P}(T > n) = \lim_{n\to+\infty} \frac{1}{n+1} = 0.$$

(ii) Si $m \in [1,2]$ on a

$$\mathrm{P}(T \leq m) \geq 1 - \mathrm{P}(T > 1) = \frac{1}{2},$$

et

$$\mathrm{P}(T \geq m) \geq \mathrm{P}(T > 1) = \frac{1}{2},$$

donc m est une médiane.

(iii) Notons que

$$\mathrm{p}_T(n) = \mathrm{P}(T = n) = \mathrm{P}(T > n-1) - \mathrm{P}(T > n) = \frac{1}{n} - \frac{1}{n+1} = \frac{1}{n(n+1)}$$

pour tout $n \in \mathbb{N}^*$. On en déduit que

$$\mathrm{E}(T) = \sum_{n=1}^{+\infty} n\,\mathrm{p}_T(n) = \sum_{n=1}^{+\infty} \frac{1}{n+1} = +\infty.$$

Exercice 3.49. Soient $X \sim \text{Géom}(p)$ et $Y \sim \text{Géom}(q)$ deux variables aléatoires indépendantes, où $p, q \in]0,1[$. On pose $Z = \min(X,Y)$. Montrer que Z suit une loi géométrique, dont on déterminera le paramètre.

Solution 3.49. Commençons par l'observation que $X \sim \text{Géom}(p)$ si et seulement si, pour tout $n \in \mathbb{N}$ on a $\mathrm{P}(X > n) = (1-p)^n$. Une implication découle de (3.98). Réciproquement, si on a $\mathrm{P}(X > n) = (1-p)^n$ pour tout $n \in \mathbb{N}$, alors pour tout $n \geq 1$

$$p_X(n) = \mathrm{P}(X > n-1) - \mathrm{P}(X > n) = (1-p)^{n-1} - (1-p)^n = p(1-p)^{n-1},$$

c'est-à-dire $X \sim \text{Géom}(p)$.

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

$$\mathrm{P}(Z > n) = \mathrm{P}(X > n, Y > n) = (1-p)^n(1-q)^n = (1-q)^n,$$

où $q = 1-(1-p)(1-q)$. Ainsi, $Z \sim \text{Géom}(q)$.

Exercice 3.50. Soient T, W des variables aléatoires indépendantes, de lois $T \sim \text{Bern}(\frac{1}{2})$ et $W \sim \text{Géom}(p)$, où $p \in]0,1[$ est un paramètre fixé. On définit la variable aléatoire

$$X := W\,\mathbb{1}_{\{T=0\}} + \frac{1}{W}\,\mathbb{1}_{\{T=1\}},$$

à valeurs dans les entiers et les inverses d'entiers, c'est-à-dire $X(\Omega) = \mathbb{N}^* \cup \{\frac{1}{n}\}_{n \in \mathbb{N}^*}$.

(i) Déterminer la densité discrète de X.

(ii) Montrer que la variable aléatoire $Y := 1/X$ a la même loi que X.

(iii) Calculer $\mathrm{E}(X)$.

[*Sugg.* On rappelle que $\sum_{n=1}^{\infty} \frac{1}{n} x^n = -\log(1-x)$ et $\sum_{n=1}^{\infty} n x^{n-1} = (1-x)^{-2}$, pour $|x| < 1$.]

Solution 3.50. (i) La densité discrète de X est donnée par :

$$\begin{aligned}
\mathrm{p}_X(n) &= \mathrm{P}(W = n,\, T = 0) = \frac{1}{2}\,p(1-p)^{n-1}, \quad \text{pour } n \in \mathbb{N}^*, n \geq 2,\\
\mathrm{p}_X(1) &= \mathrm{P}(W = 1) = p,\\
\mathrm{p}_X\left(\frac{1}{n}\right) &= \mathrm{P}(W = n,\, T = 1) = \frac{1}{2}\,p(1-p)^{n-1}, \quad \text{pour } n \in \mathbb{N}^*, n \geq 2.
\end{aligned}$$

(ii) On a $Y(\Omega) = \mathbb{N}^* \cup \{\frac{1}{n}\}_{n \in \mathbb{N}^*}$. D'après la réponse à la question précédente, il est clair que $\mathrm{P}(X = n) = \mathrm{P}(X = \frac{1}{n})$, pour tout $n \in \mathbb{N}^*$; par conséquent $\mathrm{p}_Y(n) = \mathrm{p}_X(\frac{1}{n}) = \mathrm{p}_X(n)$ et $\mathrm{p}_Y(\frac{1}{n}) = \mathrm{p}_X(n) = \mathrm{p}_X(\frac{1}{n})$.

(iii) On a

$$\begin{aligned}
\mathrm{E}(X) &= \sum_{x \in X(\Omega)} x\,\mathrm{p}_X(x) = \sum_{n=2}^{\infty} \frac{1}{n}\,\mathrm{p}_X\left(\frac{1}{n}\right) + 1 \cdot \mathrm{p}_X(1) + \sum_{n=2}^{\infty} n\,\mathrm{p}_X(n)\\
&= \sum_{n=2}^{\infty} \frac{1}{n}\,\frac{1}{2}\,p(1-p)^{n-1} + p + \sum_{n=2}^{\infty} n\,\frac{1}{2}\,p(1-p)^{n-1}\\
&= \frac{p}{2(1-p)} \sum_{n=1}^{\infty} \frac{1}{n}(1-p)^n + \frac{p}{2} \sum_{n=1}^{\infty} n(1-p)^{n-1}\\
&= \frac{1}{2}\left(\frac{p}{(1-p)} \log\frac{1}{1-p} + \frac{1}{p}\right).
\end{aligned}$$

Exercice 3.51. D'un paquet de 50 cartes numérotées de 1 à 50, on tire 3 cartes au hasard. On introduit les variables aléatoires suivantes

$$\begin{aligned} X &:= \text{« le nombre tiré le plus bas »;} \\ Z &:= \text{« le nombre tiré le plus haut »;} \\ Y &:= \text{« l'autre nombre tiré ».} \end{aligned}$$

(i) Déterminer les lois marginales de X, Y et Z.

(ii) Déterminer la loi de (X,Y) et montrer que $Y - X$ a la même loi que X.

Solution 3.51. (i) Notons que $X \in \{1,2,\dots,48\}$. Si $n \in \{1,2,\dots,48\}$, l'événement $\{X = n\}$ signifie « en plus de n, les deux nombres tirés sont supérieurs à n ». Il y a $\binom{50-n}{2}$ façons de choisir ces deux nombres, donc

$$\mathrm{P}(X = n) = \frac{\binom{50-n}{2}}{\binom{50}{3}}.$$

Avec des arguments analogues, pour $n \in \{3,4,\dots,50\}$ on a

$$\mathrm{P}(Z = n) = \frac{\binom{n-1}{2}}{\binom{50}{3}}$$

et pour $n \in \{2,3,\dots,49\}$

$$\mathrm{P}(Y = n) = \frac{(n-1)(50-n)}{\binom{50}{3}}.$$

(ii) Pour $n \in \{1,2,\dots,48\}$, $m \in \{2,3,\dots,49\}$ avec $n < m$, on a

$$\mathrm{P}(X = n, Y = m) = \frac{50-m}{\binom{50}{3}},$$

car $\{X = n, Y = m\}$ correspond à l'événement « en plus de n et m, l'autre nombre tiré est supérieur à m ». Remarquons que $Y - X \in \{1,\dots,48\}$. Si $k \in \{1,\dots,48\}$

$$\begin{aligned} \mathrm{P}(Y - X = k) &= \sum_n \mathrm{P}(X = n, Y = n+k) = \sum_{n=1}^{49-k} \frac{50-n-k}{\binom{50}{3}} \\ &= \frac{1}{\binom{50}{3}} \left[(50-k)(49-k) - \sum_{n=1}^{49-k} n \right] = \frac{\binom{50-k}{2}}{\binom{50}{3}}. \end{aligned}$$

Exercice 3.52. Soient $n, m \geq 0$, et soient $X \sim \mathrm{U}(\{0,\dots,n\})$ et $Y \sim \mathrm{U}(\{0,\dots,m\})$ deux variables aléatoires indépendantes. Déterminer la loi de X conditionnellement à l'événement $X + Y = m$, c'est-à-dire déterminer la densité discrète de X par rapport à la probabilité $\mathrm{P}(\cdot \,|\, X + Y = m)$.

[*Sugg.* Distinguer les deux cas $n \geq m$ et $n < m$.]

Solution 3.52. Par indépendance de X et Y, si $(h,k) \in \{0,\dots,n\} \times \{0,\dots,m\}$

$$\mathrm{p}_{X,Y}(h,k) = \mathrm{p}_X(h)\,\mathrm{p}_Y(k) = \frac{1}{mn},$$

c'est-à-dire $(X,Y) \sim \mathrm{U}(\{0,\dots,n\} \times \{0,\dots,m\})$. si $m \leq n$, on a

$$A_m := \{(h,k) \in \{0,\dots,n\} \times \{0,\dots,m\} : h+k=m\} = \{(i,m-i) : i \in \{1,\dots,m-1\}\},$$

donc

$$\mathrm{p}_{X+Y}(m) = \sum_{(h.k)\in A_m} \mathrm{p}_{X,Y}(h,k) = \frac{|A_m|}{mn} = \frac{m-1}{mn}\,.$$

Ainsi,

$$\begin{aligned}\mathrm{P}(X=h\,|\,X+Y=m) &= \frac{\mathrm{P}(X=h,\,X+Y=m)}{\mathrm{P}(X+Y=m)}\\ &= \frac{\mathrm{p}_{X,Y}(h,m-h)}{\mathrm{p}_{X+Y}(m)} = \begin{cases} \frac{1}{m-1} & \text{si } h \in \{1,2,\dots,m-1\}\\ 0 & \text{sinon.}\end{cases}\end{aligned}$$

De manière analogue, si $m > n$

$$A_m := \{(h,k) \in \{0,\dots,n\} \times \{0,\dots,m\} : h+k=m\} = \{(i,m-i) : i \in \{1,\dots,n\}\},$$

et

$$\mathrm{p}_{X+Y}(m) = \sum_{(h.k)\in A_m} \mathrm{p}_{X,Y}(h,k) = \frac{|A_m|}{mn} = \frac{n}{mn} = \frac{1}{m}\,.$$

Au final,

$$\begin{aligned}\mathrm{P}(X=h\,|\,X+Y=m) &= \frac{\mathrm{P}(X=h,\,X+Y=m)}{\mathrm{P}(X+Y=m)}\\ &= \frac{\mathrm{p}_{X,Y}(h,m-h)}{\mathrm{p}_{X+Y}(m)} = \begin{cases} \frac{1}{n} & \text{si } h \in \{1,2,\dots,n\}\\ 0 & \text{sinon.}\end{cases}\end{aligned}$$

Exercice 3.53. Soit Z une variable aléatoire réelle à valeurs dans $\{0,\pi/2,\pi\}$ et de densité discrète donnée par

$$\mathrm{p}_Z(0) = \mathrm{p}_Z(\pi/2) = \mathrm{p}_Z(\pi) = \frac{1}{3}\,.$$

On pose $X := \cos(Z)$ et $Y := \sin(Z)$.

(i) Montrer que les variables aléatoires X et Y ne sont pas indépendantes.

(ii) Montrer que $\mathrm{E}(X) = 0$, puis que $\mathrm{Cov}(X,Y) = 0$.

Solution 3.53. (i) Notons que l'on ne peut avoir simultanément $X = 1$ et $Y = 1$: ainsi,

$$\mathrm{P}(X=1,Y=1) = 0 \neq \mathrm{P}(X=1)\,\mathrm{P}(Y=1) = \frac{1}{9}\,.$$

Donc X et Y ne sont pas indépendantes.

(ii) Notons que $\mathrm{P}(X=1) = \mathrm{P}(X=0) = \mathrm{P}(X=-1) = \frac{1}{3}$ dont on déduit tout de suite que $\mathrm{E}(X) = 0$. De plus, on a $\mathrm{P}(XY=0) = 1$ et donc $\mathrm{E}(XY) = 0$. Au final,

$$\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y) = 0\,.$$

Exercice 3.54. Soit $n \geq 3$ et $\mathfrak{S}_n$ le groupe des permutations de $\{1,\dots,n\}$, muni de la probabilité P uniforme. Les éléments de $\mathfrak{S}_n$ seront notés $\sigma = (\sigma(1),\dots,\sigma(n))$. Introduisons les variables aléatoires réelles X,Y définies par

$$X(\sigma) := \sigma(1)\,, \qquad Y(\sigma) := \sigma(2)\,.$$

(i) Montrer que pour $i, j \in \{1,2,\dots,n\}$, la densité jointe de X et Y vaut

$$\mathrm{p}_{X,Y}(i,j) = \begin{cases} 1/c_n & \text{si } i \neq j\,, \\ 0 & \text{si } i = j\,, \end{cases}$$

où l'on demande de déterminer la constante c_n.

(ii) Déterminer la densité de la variable $D := Y - X$.

[*Sugg.* Il suffit de calculer $\mathrm{p}_D(m)$ pour $m > 0$ car par symétrie $\mathrm{p}_D(-m) = \mathrm{p}_D(m)$.]

Soient maintenant Z, W deux variables aléatoires réelles indépendantes, toutes les deux de loi uniforme discrète sur $\{1,\dots,n\}$, c'est-à-dire telles que $\mathrm{P}(Z = i) = \mathrm{P}(W = i) = \frac{1}{n}$, pour tout $i \in \{1,\dots,n\}$.

(iii) Calculer $\mathrm{P}(Z \neq W)$.

(iv) Montrer que $\mathrm{P}(Z = i, W = j \,|\, Z \neq W) = \mathrm{p}_{X,Y}(i,j)$ pour tout $i, j \in \{1,2,\dots,n\}$.

[Autrement dit, le vecteur aléatoire (Z,W), par rapport à la probabilité conditionnelle $\mathrm{P}(\cdot\,|\,Z \neq W)$, a la même loi jointe que le vecteur initial (X,Y).]

Solution 3.54. (i) Notons que $\{\sigma \in S_n : \sigma(1) = i; \sigma(2) = j\}$ est vide si $i = j$, et possède $(n-2)!$ éléments si $i \neq j$. Ainsi,

$$\mathrm{p}_{X,Y}(i,j) = \frac{|\{\sigma \in S_n : \sigma(1) = i; \sigma(2) = j\}|}{n!} = \begin{cases} \dfrac{1}{n(n-1)} & \text{si } i \neq j\,, \\ 0 & \text{si } i = j\,, \end{cases}$$

(ii) Soit $m \in \{1,\dots,n-1\}$. Alors

$$\mathrm{p}_D(m) = \sum_{i=1}^{n-m} \mathrm{p}_{X,Y}(i, i+m) = \frac{(n-m)}{n(n-1)}.$$

(iii) On a

$$\mathrm{P}(Z \neq W) = \sum_{i \neq j} \mathrm{P}(Z = i,\, W = j) = \frac{n(n-1)}{n^2} = \frac{n-1}{n}\,,$$

où la somme $\sum_{i \neq j}$ signifie qu'elle est restreinte aux couples $(i,j) \in \{1,2,\dots,n\}^2$ avec $i \neq j$, et où le résultat découle du fait qu'il y a $n(n-1)$ tels couples.

(iv) On a

$$\begin{aligned} \mathrm{P}(Z = i, W = j \,|\, Z \neq W) &= \frac{\mathrm{P}(\{Z = i, W = j\} \cap \{Z \neq W\})}{\mathrm{P}(Z \neq W)} \\ &= \begin{cases} \dfrac{1/n^2}{(n-1)/n} = \dfrac{1}{n(n-1)} & \text{si } i \neq j\,, \\ 0 & \text{si } i = j\,. \end{cases} \end{aligned}$$

Exercice 3.55. Soit S une variable aléatoire à valeurs dans $\bigcup_{n=1}^{\infty} \mathfrak{S}_n$ où $\mathfrak{S}_n$ est le groupe des permutations de $\{1,\dots,n\}$. Autrement dit, S est une permutation aléatoire dont la taille n est elle aussi aléatoire. On suppose que la loi de S est donnée par

$$\mathbb{P}(S = \sigma) = \frac{1}{\ell(\sigma)!}(1-p)^{\ell(\sigma)-1} p\,, \qquad \forall \sigma \in \bigcup_{n=1}^{\infty} \mathfrak{S}_n\,,$$

où $p \in]0,1[$ et où $\ell(\sigma)$ désigne la taille de σ, c'est-à-dire l'entier n tel que $\sigma \in \mathfrak{S}_n$.

(i) Vérifier que cela définit bien la loi d'une variable aléatoire discrète.

(ii) Donner la loi de $\ell(S)$. Donner la loi de S sous la probabilité conditionnelle $\mathbb{P}(\cdot \mid \ell(S) = n)$.

(iii) Calculer la probabilité de l'événement $A =$ « S n'a aucun point fixe ».

[*Sugg.* On a calculé $\mathrm{P}(A \mid \ell(S) = n)$ dans le Problème 2.4 (Section 2.1), voir en particulier (2.3).]

Solution 3.55. (i) Il s'agit de vérifier que

$$\sum_{\sigma \in \bigcup_{n=1}^{\infty} \mathfrak{S}_n} \frac{1}{\ell(\sigma)!}(1-p)^{\ell(\sigma)-1} p = 1.$$

En effet :

$$\sum_{\sigma \in \bigcup_{n=1}^{\infty} \mathfrak{S}_n} \frac{1}{\ell(\sigma)!}(1-p)^{\ell(\sigma)-1} p = \sum_{n=1}^{+\infty} \sum_{\sigma \in \mathfrak{S}_n} \frac{1}{n}(1-p)^{n-1} p = \sum_{n=1}^{+\infty} (1-p)^{n-1} p,$$

où on a utilisé le fait que $|\mathfrak{S}_n| = n!$ et que $(1-p)^{n-1}p$ est la densité discrète d'une variable aléatoire géométrique, donc est de somme égale à un.

(ii) On a

$$\begin{aligned} \mathrm{P}(\ell(S) = n) = \sum_{\sigma : \ell(\sigma) = n} \mathrm{P}(S = \sigma) &= \sum_{\sigma \in \mathfrak{S}_n} \frac{1}{n!}(1-p)^{n-1} p \\ &= (1-p)^{n-1} p, \end{aligned}$$

c'est-à-dire $\ell(S) \sim \mathrm{Géom}(p)$.

(iii) Commençons par remarquer que si $\sigma \in \mathfrak{S}_n$, en utilisant le point précédent, on a

$$\mathrm{P}(S = \sigma \mid \ell(S) = n) = \frac{\mathrm{P}(S = \sigma)}{\mathrm{P}(\ell(S) = n)} = \frac{1}{n!},$$

c'est-à-dire que la loi de S par rapport à la probabilité conditionnelle $\mathrm{P}(\cdot \mid \ell(S) = n)$ est la probabilité uniforme sur $\mathfrak{S}_n$. Comme il a été démontré dans le Problème 2.4 de la Section 2.1, on a

$$\mathrm{P}(A \mid \ell(S) = n) = \sum_{k=0}^{n} \frac{(-1)^k}{k!}.$$

D'après la formule des probabilités totales (voir la Proposition 1.53), on a

$$\begin{aligned} \mathrm{P}(A) = \sum_{n=1}^{+\infty} \mathrm{P}(A \mid \ell(S) = n) \mathrm{P}(\ell(S) = n) &= \sum_{n=1}^{+\infty} \sum_{k=0}^{n} \frac{(-1)^k}{k!}(1-p)^{n-1} p \\ &= 1 + \sum_{n=1}^{+\infty} \sum_{k=1}^{n} \frac{(-1)^k}{k!}(1-p)^{n-1} p, \end{aligned}$$

où on a séparé le terme $k = 0$ et utilisé de nouveau le fait que $\sum_{n=1}^{+\infty}(1-p)^{n-1}p = 1$. On doit maintenant intervertir les deux sommes, ce que l'on peut faire car la famille $(\frac{1}{k!}(1-p)^{n-1}p)_{1 \le k \le n}$ admet une somme finie (on le vérifie par des calculs identiques à ceux qui suivent). On obtient

$$\begin{aligned}
\mathrm{P}(A) &= 1 + \sum_{k=1}^{+\infty} \frac{(-1)^k}{k!} \sum_{n=k}^{+\infty} (1-p)^{n-1} p \\
&= 1 + \sum_{k=1}^{+\infty} \frac{(-1)^k}{k!} (1-p)^{k-1} \\
&= 1 + (1-p)^{-1} \left[\sum_{k=0}^{+\infty} \frac{(-(1-p))^k}{k!} - 1 \right] \\
&= 1 - \frac{1 - \mathrm{e}^{-(1-p)}}{1-p} \,,
\end{aligned}$$

où on a utilisé le fait que

$$\sum_{n=m+1}^{+\infty} (1-p)^{n-1} p = (1-p)^m$$

pour tout $m \in \mathbb{N}$ (voir (3.98)) et la série exponentielle.

Exercice 3.56. Pour $p \in \,]0,1[$ et $n \geq 2$ fixés, soient $Z_1, \ldots, Z_n$ des variables aléatoires indépendantes, à valeurs dans $\{-1,1\}$, avec $\mathrm{P}(Z_i = 1) = p$ pour tout $i = 1, \ldots, n$. On définit

$$X := \prod_{i=1}^{n} Z_i = Z_1 \cdot Z_2 \cdots Z_n \,.$$

(i) Déterminer $\mathrm{E}(X)$ et en déduire la loi de X.
[*Sugg.* Quelles valeurs la variable aléatoire X peut-elle prendre ?]

(ii) La variable aléatoire X est-elle indépendante du vecteur aléatoire $(Z_1, \ldots, Z_n)$?

(iii) La variable aléatoire X est-elle indépendante du vecteur aléatoire $(Z_2, \ldots, Z_n)$?

Solution 3.56. (i) Par indépendance des Z_i on a

$$\mathrm{E}(X) = \prod_{i=1}^{n} \mathrm{E}(Z_i) = \mathrm{E}(Z_1)^n = (p \cdot 1 + (1-p) \cdot (-1))^n = (2p-1)^n \,.$$

Étant donné que $X \in \{-1,1\}$, la loi de X est donnée par la densité discrète $\mathrm{p}_X(1) = q$, $\mathrm{p}_X(-1) = 1-q$ et $\mathrm{p}_X(x) = 0$ si $x \notin \{-1,+1\}$, pour un $q \in [0,1]$ opportun. Comme $\mathrm{E}(X) = q - (1-q) = 2q - 1$, de la formule trouvée ci-dessus pour $\mathrm{E}(X)$, on déduit que $q = \frac{1}{2}(1 + (2p-1)^n)$.

(ii) La variable aléatoire X est fonction de $(Z_1, \ldots, Z_n)$, donc elle ne peut pas être indépendante de $(Z_1, \ldots, Z_n)$ (à part dans le cas trivial où les Z_i, et donc X, sont presque sûrement constants). Plus formellement,

$$0 = \mathrm{P}(X = -1, (Z_1, \ldots, Z_n) = (1, \ldots, 1)) \neq \mathrm{P}(X = -1)\, \mathrm{P}((Z_1, \ldots, Z_n) = (1, \ldots, 1)) > 0 \,.$$

(iii) Il y a indépendance si et seulement si $p = \frac{1}{2}$. En effet, si $p = \frac{1}{2}$ on a

$$\mathrm{P}(X = \pm 1 \,|\, Z_2 = t_2, \ldots, Z_{n-1} = t_{n-1}) = \mathrm{P}(Z_1 = \pm \operatorname{sign}(t_2 \cdot t_3 \cdots t_{n-1})) = \frac{1}{2} = \mathrm{P}(X = \pm 1) \,,$$

donc X et $(Z_2, \ldots, Z_n)$ sont indépendants. Si en revanche $p \neq \frac{1}{2}$, on a

$$P(X=1\,|\,Z_2=1,\dots,Z_n=1)=P(Z_1=1)=p\neq\frac{1}{2}(1+(2p-1)^n)=P(X=1)\,.$$

Exercice 3.57. Soient X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$, ayant la densité jointe suivante :

$$\mathrm{p}_{X,Y}(k,n)=\begin{cases}\binom{n}{k}p^k(1-p)^{n-k}\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!} & \text{si } 0\leq k\leq n\,,\\ 0 & \text{sinon}\,,\end{cases}$$

où $p\in\,]0,1[$ et $\lambda>0$ sont des paramètres fixés. Déterminer les densités marginales de X et Y, puis calculer $\mathrm{Cov}(X,Y)$.

Solution 3.57. On a

$$\mathrm{p}_X(k)=\mathrm{e}^{-\lambda}\frac{(p\lambda)^k}{k!}\sum_{n\geq k}\frac{[\lambda(1-p)]^{n-k}}{(n-k)!}=\mathrm{e}^{-p\lambda}\frac{(p\lambda)^k}{k!}\,,$$
$$\mathrm{p}_Y(n)=\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!}\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}=\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!}\,,$$

donc $X\sim\mathrm{Poi}(p\lambda)$ et $Y\sim\mathrm{Poi}(\lambda)$. De plus,

$$\begin{aligned}\mathrm{E}(XY)&=\sum_{n,k\in\mathbb{N}}nk\,p_{X,Y}(n,k)=\sum_{n=0}^{\infty}n\,\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!}\sum_{k=0}^{n}k\binom{n}{k}p^k(1-p)^{n-k}\\&=\sum_{n=0}^{\infty}n\,p_X(n)\sum_{k=0}^{n}kp_{\mathrm{Bin}(n,p)}(k)=\sum_{n=0}^{\infty}n\,p_X(n)\,(pn)=p\,\mathrm{E}(X^2)=p(\lambda+\lambda^2)\,,\end{aligned}$$

ce qui donne $\mathrm{Cov}(X,Y)=p\lambda$.

Exercice 3.58. Jean lance un dé à six faces de manière répétée. On note X_k le résultat du k-ème lancer, pour $k\in\mathbb{N}^*$. Baptiste regarde les résultats des lancers, en attendant le premier instant T où sort un numéro $\{1,2,3,4,5\}$. Quand cela arrive, il note le numéro sorti $Y:=X_T$.

(i) Pour tout $n\in\mathbb{N}^*$, exprimer l'événement $\{T=n\}$ en termes des variables aléatoires $X_1,\dots,X_n$. En déduire la densité discrète de T et la reconnaître.

(ii) Pour $n\in\mathbb{N}^*$ et $a\in\{1,2,3,4,5\}$, exprimer l'événement $\{T=n,Y=a\}$ en termes des variables aléatoires $X_1,\dots,X_n$. En déduire la densité discrète jointe des variables T et Y.

(iii) Déterminer la loi de Y. Les variables aléatoires T et Y sont-elles indépendantes ?

Solution 3.58. (i) Notons que

$$\{T=n\}=\{X_1=6,\dots,X_{n-1}=6,X_n\neq 6\}\,.$$

Ainsi, en utilisant l'indépendance des X_k, on obtient

$$P(T=n)=\frac{5}{6}\left(\frac{1}{6}\right)^{n-1}\,,$$

c'est-à-dire $T\sim\text{Géom}(\frac{5}{6})$.

(ii) Pour tout $n\geq 1$ et $a\in\{1,2,3,4,5\}$ on a

$$\{T=n,Y=a\}=\{X_1=6,\dots,X_{n-1}=6,X_n=a\},$$

dont on déduit que

$$\mathrm{p}_{T,Y}(n,a) = \frac{1}{6^n}\,.$$

(iii) On a

$$\mathrm{p}_Y(a) = \sum_{n\geq 1} \mathrm{p}_{T,Y}(n,a) = \sum_{n\geq 1} \frac{1}{6^n} = \frac{1}{5}, \qquad \forall a \in \{1,2,3,4,5\},$$

c'est-à-dire $Y \sim \mathrm{Unif}(\{1,2,3,4,5\})$. Donc, pour tout $n \geq 1$ et $a \in \{1,2,3,4,5\}$

$$\mathrm{p}_T(n)\,\mathrm{p}_Y(a) = \frac{5}{6}\left(\frac{1}{6}\right)^{n-1}\frac{1}{5} = \frac{1}{6^n} = \mathrm{p}_{T,Y}(n,a),$$

dont découle le fait que T et Y sont indépendants.

Exercice 3.59. Soit Ω l'ensemble des sous-ensembles non vides de $\{1,2,\dots,N\}$, où N est un entier $N \geq 2$. Autrement dit,

$$\Omega := \{\omega \subseteq \{1,2,\dots,N\} : \omega \neq \emptyset\}.$$

Pour $\omega \in \Omega$ on pose $X(\omega) := \max(\omega)$ le maximum des éléments de ω et $Y(\omega) := \min(\omega)$ le minimum des éléments de ω. Enfin, soit P la probabilité uniforme sur Ω.

(i) Montrer que, pour $n \in \{1,2,\dots,N\}$,

$$\mathrm{P}(X = n) = \frac{2^{n-1}}{2^N - 1}.$$

(ii) Calculer la fonction génératrice des moments de X.

(iii) Calculer $\mathrm{E}(X)$ et $\mathrm{Var}(X)$ (plus difficile). Montrer que $\mathrm{Var}(X) \leq 2$.

(iv) Déterminer la densité discrète jointe de (X,Y).

(v) Déterminer la densité discrète de $X - Y$.

Solution 3.59. (i) Tout d'abord, on sait que $|\Omega| = 2^N - 1$, vu qu'un ensemble de N éléments possède 2^N sous-ensembles, dont l'un est l'ensemble vide. De plus, $X(\omega) = n$ si et seulement si $n \in \omega$, et $\omega \setminus \{n\} \subseteq \{1,2,\dots,n-1\}$. Ainsi

$$|\{\omega : X(\omega) = n\}| = |\{\omega' : \omega' \subseteq \{1,2,\dots,n-1\}\}| = 2^{n-1},$$

ce qui donne la formule cherchée.

(ii) En se rappelant de la série géométrique (0.4), on a, pour $2\mathrm{e}^t \neq 1$, c'est-à-dire pour $t \neq -\log 2$.

$$\begin{aligned}\mathrm{M}_X(t) = \mathrm{E}\left[\mathrm{e}^{tX}\right] &= \sum_{n=1}^{N} \mathrm{e}^{tn}\frac{2^{n-1}}{2^N-1} = \frac{\mathrm{e}^t}{2^N-1}\sum_{n=1}^{N}\left(2\mathrm{e}^t\right)^{n-1} \\ &= \frac{\mathrm{e}^t}{2^N-1}\sum_{m=0}^{N-1}\left(2\mathrm{e}^t\right)^m = \frac{\mathrm{e}^t}{2^N-1}\,\frac{\left(2\mathrm{e}^t\right)^N-1}{2\mathrm{e}^t-1} = \frac{1}{2^N-1}\,\frac{2^N\mathrm{e}^{tN}-1}{2-\mathrm{e}^{-t}}\,.\end{aligned}$$

Pour $t = -\log 2$ (c'est-à-dire $2\mathrm{e}^t = 1$), on a $\mathrm{M}_X(t) = \frac{N/2}{2^N-1}$.

(iii) On peut calculer $\mathrm{M}'_X(t)$ pour $t \neq -\log 2$: on a

$$\mathrm{M}'_X(t) = \frac{1}{2^N-1}\left(\frac{N2^N\mathrm{e}^{tN}}{2-\mathrm{e}^{-t}} - \frac{2^N\mathrm{e}^{t(N-1)}-\mathrm{e}^{-t}}{(2-\mathrm{e}^{-t})^2}\right).$$

On obtient alors que

$$\mathrm{E}(X) = \mathrm{M}'_X(0) = \frac{(N-1)2^N+1}{2^N-1} = N-1+\frac{N}{2^N-1}\,.$$

On peut aussi calculer la dérivée seconde de $\mathrm{M}_X(t)$ pour $t \neq -\log 2$: après calcul, on obtient

$$\mathrm{M}''_X(t) = \frac{1}{2^N-1}\left(\frac{N^2 2^N\mathrm{e}^{tN}}{2-\mathrm{e}^{-t}} - \frac{N2^N\mathrm{e}^{t(N-1)}}{2-\mathrm{e}^{-t}} \right.$$
$$\left. - \frac{(N-1)2^N\mathrm{e}^{t(N-1)}+\mathrm{e}^{-t}}{(2-\mathrm{e}^{-t})^2} + \frac{2^{N+1}\mathrm{e}^{t(N-2)}-2\mathrm{e}^{-2t}}{(2-\mathrm{e}^{-t})^3}\right).$$

On en déduit

$$\mathrm{E}(X^2) = \mathrm{M}''_X(0) = \frac{(N^2-2N+3)2^N-3}{2^N-1} = N^2-2N+3+\frac{N(N-2)}{2^N-1}$$

et, après simplifications,

$$\mathrm{Var}(X) = \mathrm{E}(X^2)-\mathrm{E}(X)^2 = 2-\frac{N^2 2^N}{(2^N-1)^2}\,.$$

(iv) Évidemment, pour tout $\omega \in \Omega$, on a $X(\omega) \geq Y(\omega)$. On doit donc calculer, pour $N \geq n \geq m \geq 1$

$$\mathrm{P}(X=n, Y=m).$$

Observons que l'on a $X(\omega)=n, Y(\omega)=m$ si et seulement si $n,m \in \omega$ et $\omega \setminus \{n,m\} \subseteq \{m+1,\ldots,n-1\}$. Ainsi, le nombre de ω pour lesquels $X(\omega)=n, Y(\omega)=m$ est égal au nombre de sous-ensembles de $\{m+1,\ldots,n-1\}$. Notons que ce dernier ensemble est vide si $n=m$, et possède $n-m-1$ éléments sinon. Ainsi,

$$\mathrm{P}(X=n,Y=m) = \begin{cases} \dfrac{1}{2^N-1} & \text{si } n=m \\ \dfrac{2^{n-m-1}}{2^N-1} & \text{si } n>m. \end{cases}$$

(v) L'événement $\{X-Y=k\}$, qui est non vide pour $k=0,1,\ldots,N-1$, peut s'écrire comme une union disjointe de la façon suivante :

$$\{X-Y=k\} = \bigcup_{m=1}^{N-k}\{Y=m, X=m+k\}.$$

En utilisant les calculs du point (iv), on obtient

$$\mathrm{P}(X-Y=k) = \begin{cases} \displaystyle\sum_{m=1}^{N}\frac{1}{2^N-1} = \frac{N}{2^N-1} & \text{si } k=0\,, \\ \displaystyle\sum_{m=1}^{N-k}\frac{2^{k-1}}{2^N-1} = (N-k)\frac{2^{k-1}}{2^N-1} & \text{si } 1\leq k\leq N-1. \end{cases}$$

Exercice 3.60. D'une urne contenant r boules rouges et s boules vertes, on tire successivement et sans remise k boules, avec $k \leq \min(r,s)$. Soit

$$X_i := \begin{cases} 1 & \text{si la } i\text{-ème boule tirée est rouge} \\ 0 & \text{sinon} \end{cases} \qquad \text{pour } i = 1,2,\ldots,k.$$

On définit $X = X_1 + \cdots + X_k$, qui représente le nombre de boules rouges tirées.

(i) Déterminer la loi de X.

(ii) Déterminer la loi des X_i.

(iii) Calculer $\mathrm{E}(X)$.

(iv) Montrer, par récurrence sur k, que la densité jointe des X_i est donnée par

$$\mathrm{p}_{X_1,\ldots,X_k}(x_1,\ldots,x_k) = \frac{r(r-1)\cdots(r-\sum_{i=1}^k x_i+1)s(s-1)\cdots(s-k+\sum_{i=1}^k x_i+1)}{(r+s)(r+s-1)\cdots(r+s-k+1)}.$$

Observer que, en posant $s(x) := \sum_{i=1}^k x_i$, cette formule se réécrit de la façon suivante :

$$\mathrm{p}_{X_1,\ldots,X_k}(x_1,\ldots,x_k) = \frac{\frac{r!}{(r-s(x))!}\,\frac{v!}{(v-k+s(x))!}}{\frac{(r+v)!}{(r+v-k)!}} = \frac{\binom{r}{s(x)}\binom{v}{k-s(x)}}{\binom{r+v}{k}\binom{k}{s(x)}}.$$

Solution 3.60. (i) Comme X est le nombre de boules rouges tirées lors de k tirages sans remises, on a (voir (1.41))

$$\mathrm{p}_X(n) = \binom{r}{n}\binom{v}{k-n}\Big/\binom{r+v}{n}.$$

(ii) On identifie l'ensemble des boules avec $\{1,2,\ldots,r+v\}$, en convenant que les boules rouges sont les r premières. Le schéma d'un tirage sans remise peut être modélisé en choisissant comme espace d'états Ω l'ensemble des permutations de $\{1,2,\ldots,r+v\}$, de cardinal $|\Omega| = (t+v)!$, muni de la probabilité uniforme. Alors, l'événement $\{X_i = 1\} = \{\sigma \in \Omega : \sigma(i) \in \{1,2,\ldots,r\}\}$ a pour cardinal $r(r+v-1)!$ et donc sa probabilité vaut $\frac{r}{r+v}$. En conclusion,

$$\mathrm{p}_{X_i}(1) = \frac{r}{r+v}, \qquad \mathrm{p}_{X_i}(0) = 1 - \mathrm{p}_{X_i}(1) = \frac{v}{r+v},$$

c'est-à-dire $X_i \sim \mathrm{Bern}(\frac{r}{r+v})$.

(iii) Comme on a $\mathrm{E}(X_i) = \frac{r}{r+v}$, par linéarité de l'espérance, on obtient $\mathrm{E}(X) = k\,\frac{r}{r+v}$.

(iv) Raisonnons par récurrence sur k. Pour $k=1$ la réponse est donné dans (ii). Pour $k \geq 2$, prenons $x = (x_1,\ldots,x_{k-1},x_k) \in \{0,1\}^k$, et écrivons

$$\mathrm{P}(X=x) = \mathrm{P}(X_k = x_k \,|\, X_{k-1} = x_{k-1},\ldots,X_1 = x_1)\,\mathrm{P}(X_{k-1} = x_{k-1},\ldots,X_1 = x_1).$$

D'une part, on utilise l'hypothèse de récurrence pour $\mathrm{P}(X_{k-1} = x_{k-1},\ldots,X_1 = x_1)$, et d'autre part, en utilisant (ii) avec $r+v-k-1$ boules dont $r - \sum_{i=1}^{k-1} x_i$ rouges, on a

$$\mathrm{P}(X_k = 1 \,|\, X_{k-1} = x_{k-1},\ldots,X_1 = x_1) = \frac{r - \sum_{i=1}^{k-1} x_i}{r+v-k+1}.$$

Cela permet de conclure la récurrence.

Exercice 3.61 (Urne de Pólya). Une urne contient initialement une boule blanche et une boule noire. À chaque tour, une boule est prélevée de l'urne, puis elle est replacée dans l'urne en ajoutant une autre boule de la même couleur. On souligne qu'après n tours, il y a $n+2$ boules dans l'urne. On note X_n le nombre de boules blanches dans l'urne après n tours, qui est une variable aléatoire à valeurs dans $\{1,\dots,n+1\}$.

(i) Déterminer la densité discrète de X_1.

(ii) Montrer que, pour tout $n \geq 1$, on a pour $k \in \{1,\dots,n\}$

$$\mathrm{P}(X_n = k \,|\, X_{n-1} = k) = \frac{k}{n+1}, \qquad \mathrm{P}(X_n = k+1 \mid X_{n-1} = k) = 1 - \frac{k}{n+1}.$$

(iii) En déduire par récurrence sur n que $\mathbb{P}(X_n = k) = \frac{1}{n+1}$ pour $k \in \{1,\dots,n+1\}$.

Solution 3.61. (i) Notons que $X_1 \in \{1,2\}$. De plus, $X_1 = 1$ si et seulement si la première boule tirée est la noire. Ainsi $\mathrm{P}(X_1 = 1) = \mathrm{P}(X_1 = 2) = \frac{1}{2}$.

(ii) Si $X_{n-1} = k$, le n-ème tirage est effectué avec k boules blanches et $n+1-k$ boules noires. On a alors $X_n = k+1$ si la boule tirée au n-ème tirage est blanche, ce qui arrive avec probabilité $\frac{k}{n+1}$, sinon $X_n = k$. On en déduit directement les identités suivantes :

$$\mathrm{P}(X_n = k+1 \,|\, X_{n-1} = k) = \frac{k}{n+1}, \qquad \mathrm{P}(X_n = k \mid X_{n-1} = k) = 1 - \frac{k}{n+1}.$$

(iii) L'affirmation $\mathbb{P}(X_n = k) = \frac{1}{n+1}$ est vraie pour $n = 1$, comme on l'a vu plus haut. Pour démontrer l'étape de récurrence, notons que $\mathrm{P}(X_{n+1} = k \,|\, X_n = j) = 0$ sauf si $j \in \{k-1, k\}$. Ainsi, d'après la formule des probabilités totales (Proposition 1.53) et les identités que l'on vient de démontrer, pour $k \in \{1,\dots,n+2\}$ on a :

$$\begin{aligned}\mathrm{P}(X_{n+1} = k) =& \mathrm{P}(X_{n+1} = k \mid X_n = k-1)\,\mathrm{P}(X_n = k-1) + \mathrm{P}(X_{n+1} = k \mid X_n = k)\,\mathrm{P}(X_n = k) \\ =& \frac{k-1}{n+2} \frac{\mathbb{1}_{\{2,\dots,n+2\}}(k)}{n+1} + \left(1 - \frac{k}{n+2}\right) \frac{\mathbb{1}_{\{1,\dots,n+1\}}(k)}{n+1} = \frac{1}{n+2},\end{aligned}$$

où on a utilisé l'hypothèse de récurrence. Cela conclut la démonstration.

Exercice 3.62. Un insecte dépose un nombre aléatoire $N \sim \mathrm{Poi}(\lambda)$ d'œufs. Chacun de ces œufs éclot avec probabilité $p \in \,]0,1[$, indépendamment du nombre d'œufs déposés et du fait que les autres œufs éclosent ou non. On note X le nombre (aléatoire) d'œufs qui éclosent.

(i) Que vaut la probabilité conditionnelle $\mathrm{P}(X = k \,|\, N = n)$, pour $n \in \mathbb{N}$ et $k \in \mathbb{N}$?

(ii) Déterminer la loi de X.

(iii) On pose $Y = N - X$ le nombre d'œufs qui *n'ont pas* éclot. Déterminer la densité discrète jointe de (X,Y) et en déduire que X et Y sont indépendantes.

Solution 3.62. (i) Par hypothèse, sous la probabilité conditionnelle $\mathrm{P}(\cdot \mid N = n)$, X est de loi $\mathrm{Bin}(n,p)$:

$$\mathrm{P}(X = k \,|\, N = n) = \binom{n}{k} p^k (1-p)^{n-k} \qquad \text{pour } k \in \{0,\dots,n\}.$$

(ii) Pour tout $k \in \mathbb{N}$ on a donc

$$\begin{aligned}\mathrm{P}(X=k) &= \sum_{n\in\mathbb{N}_0} \mathrm{P}(X=k|N=n)\,\mathrm{P}(N=n) = \sum_{n=k}^{\infty} \binom{n}{k} p^k (1-p)^{n-k}\, \mathrm{e}^{-\lambda}\,\frac{\lambda^n}{n!} \\ &= \mathrm{e}^{-\lambda}\,\frac{(p\lambda)^k}{k!} \sum_{n=k}^{\infty} \frac{[\lambda(1-p)]^{n-k}}{(n-k)!} = \mathrm{e}^{-\lambda}\,\frac{(p\lambda)^k}{k!}\mathrm{e}^{-\lambda(1-p)} = \mathrm{e}^{-p\lambda}\,\frac{(p\lambda)^k}{k!}\,,\end{aligned}$$

c'est-à-dire $X \sim \mathrm{Poi}(p\lambda)$.

(iii) Pour tous $k,\ell \in \mathbb{N}$, on a

$$\begin{aligned}\mathrm{P}(X=k,Y=\ell) &= \mathrm{P}(X=k,N=k+\ell) = \mathrm{P}(N=k+\ell)\mathbb{P}(X=k \mid N=k+\ell) \\ &= \mathrm{e}^{-\lambda}\,\frac{\lambda^{k+\ell}}{(k+\ell)!}\binom{k+\ell}{k} p^k (1-p)^{\ell} = \mathrm{e}^{-\lambda}\,\frac{\lambda^k p^k}{k!}\,\frac{\lambda^{\ell}(1-p)^{\ell}}{\ell!}\,.\end{aligned}$$

Il en découle que X et Y sont indépendantes, de lois $X \sim \mathrm{Poi}(\lambda p)$ et $Y \sim \mathrm{Poi}(\lambda(1-p))$.

Exercice 3.63. Considérons une famille d'urnes numérotés par les entiers naturels (zéro inclus) : la k-ème urne, avec $k \in \mathbb{N}$, contient 1 boule noire et k boules blanches. Marion tire une boule au hasard dans l'urne numéro X, où X est un nombre aléatoire de loi $\mathrm{Poi}(\lambda)$. On note A l'événement « la boule choisie par Marion est noire ».

(i) Déterminer $\mathrm{P}(A\,|\,X=k)$, pour tout $k \in \mathbb{N}$. En déduire que $\mathrm{P}(A) = \frac{1-\mathrm{e}^{-\lambda}}{\lambda}$.

(ii) Déterminer $\mathrm{q}(k) := \mathrm{P}(X=k\,|\,A)$ pour tout $k \in \mathbb{N}$.

(iii) Montrer que $\mathrm{q}(\cdot)$ coïncide avec la densité discrète de la variable aléatoire $Y := X-1$, par rapport à la probabilité $\mathrm{P}(\,\cdot\,|\,X \geq 1)$.

(iv) Calculer $\mathrm{E}(X\,|\,A)$, c'est-à-dire l'espérance de X par rapport à la probabilité $\mathrm{P}(\,\cdot\,|\,A)$.

Solution 3.63. (i) Par construction $\mathrm{P}(A\,|\,X=k) = \frac{1}{k+1}$, donc d'après la formule des probabilités totales

$$\begin{aligned}\mathrm{P}(A) &= \sum_{k\in\mathbb{N}} \mathrm{P}(A\,|\,X=k)\,\mathrm{P}(X=k) = \sum_{k=0}^{\infty} \frac{1}{k+1}\,\frac{\lambda^k}{k!}\mathrm{e}^{-\lambda} = \frac{1}{\lambda}\sum_{m=1}^{\infty}\frac{\lambda^m}{m!}\mathrm{e}^{-\lambda} \\ &= \frac{1}{\lambda}\big(1-\mathrm{P}(X=0)\big) = \frac{1-\mathrm{e}^{-\lambda}}{\lambda}\,,\end{aligned}$$

où on a fait le changement de variables $k+1=m$.

(ii) D'après la formule de Bayes

$$\mathrm{q}(k) = \frac{\mathrm{P}(A\,|\,X=k)\,\mathrm{P}(X=k)}{\mathrm{P}(A)} = \frac{\frac{1}{k+1}\,\frac{\lambda^k}{k!}\,\mathrm{e}^{-\lambda}}{\frac{1-\mathrm{e}^{-\lambda}}{\lambda}} = \frac{\mathrm{e}^{-\lambda}}{1-\mathrm{e}^{-\lambda}}\,\frac{\lambda^{k+1}}{(k+1)!}\,.$$

(iii) Pour tout $k \in \mathbb{N}$

$$\mathrm{P}(Y=k\,|\,X\geq 1) = \mathrm{P}(X=k+1\,|\,X\geq 1) = \frac{\mathrm{P}(X=k+1)}{\mathrm{P}(X\geq 1)} = \frac{\frac{\lambda^{k+1}}{(k+1)!}\mathrm{e}^{-\lambda}}{1-\mathrm{e}^{-\lambda}} = \mathrm{q}(k)\,.$$

(iv) En utilisant la formula pour $\mathrm{q}(k)$ trouvée ci-dessus, on a

$$\mathrm{E}(X\,|\,A) = \sum_{k\in\mathbb{N}} k\,\mathrm{q}(k) = \frac{\mathrm{e}^{-\lambda}}{1-\mathrm{e}^{-\lambda}}\sum_{k=0}^{\infty} k\,\frac{\lambda^{k+1}}{(k+1)!}\,,$$

et en écrivant $k=(k+1)-1$ on obtient

$$\begin{aligned}
\mathrm{E}(X\,|\,A) &= \frac{\mathrm{e}^{-\lambda}}{1-\mathrm{e}^{-\lambda}}\left(\lambda\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}-\sum_{k=0}^{\infty}\frac{\lambda^{k+1}}{(k+1)!}\right)\\
&= \frac{\mathrm{e}^{-\lambda}}{1-\mathrm{e}^{-\lambda}}\left(\lambda\mathrm{e}^{\lambda}-(\mathrm{e}^{\lambda}-1)\right) = \frac{\lambda-(1-\mathrm{e}^{-\lambda})}{1-\mathrm{e}^{-\lambda}}.
\end{aligned}$$

Exercice 3.64. On se fixe X une variable aléatoire à valeurs dans $\mathbb{N}$. Soit maintenant Y une variable aléatoire telle que pour tout $k\in\mathbb{N}$ pour lequel $\mathrm{P}(X=k)>0$, on a

$$\mathrm{P}(Y=\ell\,|\,X=k)=\binom{k}{\ell}p^{\ell}(1-p)^{k-\ell},\qquad \forall\,\ell\in\{0,\dots,k\},$$

où $p\in[0,1]$. *On dit que Y suit une loi* $\mathrm{Bin}(X,p)$.

(i) Exprimer la densité discrète jointe de (X,Y) en fonction de la densité discrète p_X de X.

(ii) Montrer que si $X\sim\mathrm{Bin}(n,q)$ où $n\in\mathbb{N}^*$, $q\in[0,1]$, alors $Y\sim\mathrm{Bin}(n,pq)$.

Solution 3.64. (i) Pour $k\in\mathbb{N}$ et $\ell\in\{0,\dots,k\}$

$$\begin{aligned}
\mathrm{p}_{X,Y}(k,\ell) &= \mathrm{P}(X=k,Y=\ell)=\mathrm{P}(Y=\ell\,|\,X=k)\,\mathrm{P}(X=k)\\
&= \binom{k}{\ell}p^{\ell}(1-p)^{k-\ell}p_X(k).
\end{aligned}$$

(ii) Dans ce cas, pour $\ell\in\{0,\dots,n\}$

$$\mathrm{p}_Y(\ell)=\sum_{k=\ell}^{n}\mathrm{p}_{X,Y}(k,\ell)=\sum_{k=\ell}^{n}\binom{k}{\ell}p^{\ell}(1-p)^{k-\ell}\binom{n}{k}q^k(1-q)^{n-k}$$

Notons que

$$\binom{k}{\ell}\binom{n}{k}=\frac{n!}{(k-\ell)!l!(n-k)!}=\binom{n}{l}\binom{n-l}{k-\ell}.$$

Donc

$$\begin{aligned}
\mathrm{p}_Y(\ell) &= \binom{n}{l}\sum_{k=\ell}^{n}\binom{n-l}{k-\ell}p^{\ell}(1-p)^{k-\ell}q^k(1-q)^{n-k}\\
&= \binom{n}{l}p^l\sum_{h=0}^{n-\ell}\binom{n-l}{h}(1-p)^h q^{h+l}(1-q)^{n-l-k}\\
&= \binom{n}{l}p^l q^l\sum_{h=0}^{n-\ell}\binom{n-l}{h}(1-p)^h q^{h}(1-q)^{n-l-k}\\
&= \binom{n}{l}p^l q^l((1-p)q+1-q)^{n-l}\\
&= \binom{n}{l}(pq)^l(1-pq)^{n-l},
\end{aligned}$$

où on a utilisé la formule du binôme de Newton (1.39).

Exercice 3.65. Soient $(X_i)_{i\geq 1}$ des variables aléatoires indépendantes, toutes de loi $\mathrm{Géom}(p)$, où $p\in\,]0,1[$. Pour $k\in\mathbb{N}^*$, on note $T_k:=X_1+\cdots+X_k$.

(i) Montrer que $|\{(n_1,\ldots,n_k)\in\mathbb{N}^k, \sum_{i=1}^k n_k = m\}| = \binom{m-1}{k-1}$ pour tout $k,m\in\mathbb{N}^*$ avec $m\geq k$. En déduire que la densité discrète de T_k est donnée par

$$p_{T_k}(m) = \binom{m-1}{k-1} p^k(1-p)^{m-k} \qquad \text{pour } m\in\{k,k+1,\ldots\}\,.$$

(ii) Soit Y une variable aléatoire géométrique Géom(q) avec $q\in\,]0,1[$, indépendante de $(X_i)_{i\geq 1}$. On note $T_Y = X_1+\cdots+X_Y$ la somme d'un nombre Y (aléatoire) de variables aléatoires X_i. Déterminer la densité discrète jointe de (Y,T_Y). En déduire que T_Y suit une loi géométrique de paramètre pq.

Solution 3.65. (i) On note $A := \{(n_1,\ldots,n_k)\in\mathbb{N}^k :\ \sum_{i=1}^k n_k = m\}$ et B l'ensemble de combinaisons de $k-1$ éléments choisis dans $\{1,\ldots,m-1\}$, pour $1\leq k\leq m$ fixé. L'identité $|\{(n_1,\ldots,n_k)\in\mathbb{N}^k :\ \sum_{i=1}^k n_k = m\}| = \binom{m-1}{k-1}$ équivaut à $|A| = |B|$. Pour le montrer, on définit l'application $\varphi : A\to B$ de la façon suivante :

$$\varphi(n_1,\ldots,n_k) = \{n_1, n_1+n_2,\ldots,n_1+\cdots+n_{k-1}\}.$$

On voit facilement qu'effectivement $\varphi(n_1,\ldots,n_k)\in B$. Montrons que φ est une bijection, en exhibant son inverse. On définit $\psi : B\to A$ de la façon suivante. Notons que si $\beta\in B$, alors β s'écrit de manière unique sous la forme

$$\beta = \{m_1,m_2,\ldots,m_{k-1}\},$$

avec $1\leq m_1 < m_2 < \cdots < m_{k-1}$; autrement dit, on considère les éléments de β rangés dans un ordre croissant. On pose alors

$$\psi(\beta) := (m_1, m_2-m_1,\ldots,m_{k-1}-m_{k-2}, m-m_{k-1}).$$

On voit immédiatement que $\psi(\beta)\in A$, que $\psi(\varphi(n_1,\ldots,n_k)) = (n_1,\ldots,n_k)$ pour tout $(n_1,\ldots,n_k)\in A$ et que $\varphi(\psi(\beta)) = \beta$ pour tout $\beta\in B$. Ainsi, $\psi = \varphi^{-1}$, donc φ est bijective et $|A| = |B| = \binom{m-1}{k-1}$.

Déterminons maintenant la densité discrète de T_k. Pour tout $m\geq k$, en utilisant l'indépendance des variables aléatoires $X_1,X_2,\ldots,X_k$, on a

$$\begin{aligned}
\mathrm{p}_{T_k}(m) &= \mathrm{P}(X_1+\cdots+X_k = m) = \mathrm{P}((X_1,X_2,\ldots,X_k)\in A)\\
&= \sum_{(n_1,\ldots,n_k)\in A}\prod_{j=1}^k \mathrm{p}_{X_j}(n_j) = \sum_{(n_1,\ldots,n_k)\in A}\prod_{j=1}^k p(1-p)^{n_j-1}\\
&= |A|p^k(1-p)^{m-k} = \binom{m-1}{k-1}p^k(1-p)^{m-k}.
\end{aligned}$$

(ii) D'après la formule de décomposition (Proposition 1.53), on a, pour tout $m\geq 1$

$$\mathrm{P}(T_Y = m) = \sum_{k=m}^{+\infty}\mathrm{P}(T_Y = m, Y = k)\,.$$

Observons que

$$\mathrm{P}(T_Y = m, Y = k) = \mathrm{P}(T_k = m, Y = k) = \mathrm{P}(T_k = m)\,\mathrm{P}(Y = k),$$

car Y est indépendante des $(X_i)_{i\geq 1}$. Ainsi, en utilisant la formule du binôme de Newton (1.39)

$$\begin{aligned}
\mathrm{P}(T_Y = m) &= \sum_{k=1}^{m} \mathrm{P}(T_k = m)\,\mathrm{p}_Y(k) = \sum_{k=1}^{m} \binom{m-1}{k-1} p^k (1-p)^{m-k} q(1-q)^{k-1} \\
&= pq \sum_{h=0}^{m-1} \binom{m-1}{h} (p(1-q))^h (1-p)^{m-1-h} \\
&= pq(p(1-q)+1-p)^{m-1} = pq(1-pq)^{m-1}\,,
\end{aligned}$$

dont il découle que $T_Y \sim \text{Géom}(pq)$.

Exercice 3.66. On teste un médicament sur n hommes et m femmes. On suppose que l'efficacité du médicament est indépendante et identique d'un individu à l'autre ; en particulier, elle est identique pour les hommes et pour les femmes. On note X_n le nombre d'hommes chez qui le médicament est jugé efficace et Y_m le nombre de femmes chez qui le médicament est jugé efficace : d'après les hypothèses précédentes, $X_n \sim \text{Bin}(n,p)$ et $Y_m \sim \text{Bin}(m,p)$ sont indépendantes, avec $p \in]0,1[$ un paramètre (non nécessairement connu) qui indique l'efficacité du médicament.

(i) En appliquant l'inégalité de Bienaymé–Tchebychev, montrer que pour tout $\varepsilon > 0$ on a

$$\mathrm{P}\left(\left|\frac{Y_m}{m} - \frac{X_n}{n}\right| \geq \varepsilon\right) \leq \frac{(n+m)}{4mn\varepsilon^2}\,.$$

Noter que la borne obtenue ne dépend pas de $p \in]0,1[$.

(ii) Le test est effectué sur $n = 15\,000$ hommes et $m = 20\,000$ femmes et le médicament a été efficace sur $14\,000$ hommes et $18\,000$ femmes. Sur la base de ces données, vous semble-t-il raisonnable de supposer que le médicament possède la même efficacité chez les hommes que chez les femmes ?

Solution 3.66. (i) La variable aléatoire $\frac{Y_m}{m} - \frac{X_n}{n}$ est d'espérance nulle, et

$$\begin{aligned}
\mathrm{Var}\left(\frac{Y_m}{m} - \frac{X_n}{n}\right) &= \mathrm{Var}\left(\frac{Y_m}{m}\right) + \mathrm{Var}\left(\frac{X_n}{n}\right) = \frac{mp(1-p)}{m^2} + \frac{np(1-p)}{n^2} \\
&= \frac{(n+m)p(1-p)}{nm}\,.
\end{aligned}$$

En observant que $p(1-p) \leq \frac{1}{4}$ pour tout $p \in]0,1[$, on obtient grâce à l'inégalité de Bienaymé–Tchebychev

$$\mathrm{P}\left(\left|\frac{Y_m}{m} - \frac{X_n}{n}\right| \geq \varepsilon\right) \leq \frac{(n+m)}{4mn\varepsilon^2}\,,$$

comme voulu.

(ii) La différence d'efficacité entre les hommes et les femmes lors du test est

$$\left|\frac{14000}{15000} - \frac{18000}{20000}\right| = \frac{1}{30}.$$

Si l'efficacité du médicament était égale entre les hommes et les femmes, la probabilité d'observer une différence égale ou supérieure à $\frac{1}{30}$ serait

$$P\left(\left|\frac{Y_m}{m}-\frac{X_n}{n}\right|\geq\varepsilon\right)$$

avec $n=15000$, $m=14000$ et $\varepsilon=\frac{1}{30}$. D'après l'inégalité du point précédent, une telle probabilité est inférieure à

$$\frac{(n+m)}{4mn\varepsilon^2}\simeq 0,02.$$

Ainsi, l'hypothèse d'une efficacité identique chez les hommes et chez les femmes semble peu vraisemblable (moins de 2%).

Exercice 3.67. On répartit $2n$ personnes, n hommes et n femmes, de manière aléatoire dans deux groupes de n personnes chacun. On note A_i l'événement « la i-ème femme est dans le premier groupe » et $X=\sum_{i=1}^n \mathbb{1}_{A_i}$ le nombre de femmes dans le premier groupe.

(i) Donner la loi de X et calculer $E(X)$.

(ii) Montrer que pour $i\neq j$, $\mathrm{Cov}(\mathbb{1}_{A_i},\mathbb{1}_{A_j})=\frac{1}{4(2n-1)}$. En déduire que $\mathrm{Var}(X)=\frac{n^2}{4(2n-1)}$.

(iii) Montrer que pour tout $\varepsilon>0$, $P(|X-\frac{n}{2}|>\varepsilon n)\leq\frac{1}{4\varepsilon^2(2n-1)}$.

Solution 3.67. (i) Le nombre X de femmes dans le premier groupe peut s'interpréter comme étant obtenu avec un tirage *sans remise* de n boules dans une urne qui en contient $2n$, avec n boules de chaque type : X suit donc une loi hypergéométrique $\mathrm{Hyper}(n,2n,n)$. Comme on l'a vu dans la Section 3.5.4, on a $E(X)=\frac{n}{2}$.

(ii) La probabilité que la i-ème femme soit dans le premier groupe vaut $\frac{1}{2}$. La probabilité qu'aussi bien la i-ème que la j-ème femme soit dans le premier groupe est le rapport entre le nombre de sous-ensembles de n personnes qui contiennent les femmes i et j, et le nombre total de sous-ensembles de n personnes, c'èst-à-dire

$$P(A_i\cap A_j)=\frac{\binom{2n-2}{n-2}}{\binom{2n}{n}}=\frac{n(n-1)}{2n(2n-1)}.$$

Ainsi

$$\begin{aligned}\mathrm{Cov}(\mathbb{1}_{A_i},\mathbb{1}_{A_j})&=E(\mathbb{1}_{A_i}\mathbb{1}_{A_j})-E(\mathbb{1}_{A_i})E(\mathbb{1}_{A_j})=P(A_i\cap A_j)-P(A_i)P(A_j)\\&=\frac{n(n-1)}{2n(2n-1)}-\frac{1}{4}=-\frac{1}{4(2n-1)}.\end{aligned}$$

On en déduit que

$$\mathrm{Var}(X)=\sum_{i=1}^n\mathrm{Var}(\mathbb{1}_{A_i})+2\sum_{i<j}\mathrm{Cov}(\mathbb{1}_{A_i},\mathbb{1}_{A_j})=\frac{n}{4}-\frac{n(n-1)}{4(2n-1)}=\frac{n^2}{4(2n-1)}.$$

(iii) D'après l'inégalité de Bienaymé–Tchebychev,

$$P\left(\left|X-\frac{n}{2}\right|>\varepsilon n\right)\leq\frac{\mathrm{Var}(X)}{\varepsilon^2n^2}=\frac{1}{4\varepsilon^2(2n-1)}.$$

Exercice 3.68. Yassine lance de manière répétée une fléchette sur une cible. Il obtient 5 points s'il touche le centre (ce qui arrive avec probabilité $\frac{1}{6}$), 2 points si on touche la zone intermédiaire (ce qui arrive avec probabilité $\frac{1}{3}$), et 1 point s'il touche la zone périphérique (ce qui arrive avec probabilité $\frac{1}{2}$). On note X_i le score obtenu lors du i-ème lancer et on suppose que les $(X_i)_{i\geq 1}$ sont des variables aléatoires indépendantes (et de même loi). On pose $S_n := \sum_{i=1}^n X_i$ le score obtenu après n lancers de fléchettes.

(i) Calculer $\mathrm{E}(S_n)$ et $\mathrm{Var}(S_n)$.

(ii) En utilisant l'inégalité de Bienaymé–Tchebychev, montrer que pour tout $1 \leq j < N/2$ on a
$$\mathrm{P}\left(S_j \geq N\right) \leq \frac{2j}{(N-2j)^2}\,.$$
En déduire que Yassine a moins d'une chance sur 10 d'obtenir un score supérieur à 200 en lançant moins de 80 fléchettes.

Solution 3.68. (i) Commençons par remarquer que
$$\mathrm{E}(X_i) = 5 \cdot \frac{1}{6} + 2 \cdot \frac{1}{3} + 1 \cdot \frac{1}{2} = 2\,,$$
et
$$\mathrm{E}(X_i^2) = 25 \cdot \frac{1}{6} + 4 \cdot \frac{1}{3} + 1 \cdot \frac{1}{2} = 6 \implies \mathrm{Var}(X_i) = \mathrm{E}(X_i^2) - \mathrm{E}(X_i)^2 = 2\,.$$
Ainsi $\mathrm{E}(S_n) = 2n = \mathrm{Var}(S_n)$.

(ii) En utilisant l'inégalité de Bienaymé–Tchebychev,
$$\begin{aligned}\mathrm{P}\left(S_j \geq N\right) = \mathrm{P}\left(S_j - 2j \geq N - 2j\right) &\leq \mathrm{P}\left(|S_j - 2j| \geq N - 2j\right)\\ &\leq \frac{\mathrm{Var}(S_j)}{(N-2j)^2} = \frac{2j}{(N-2j)^2}\,.\end{aligned}$$
En prenant $N = 200$ et $j \leq 80$ on trouve
$$\mathrm{P}\left(S_j \geq 200\right) \leq \frac{2j}{(200-2j)^2} \leq \frac{160}{(200-160)^2} = 0{,}1\,.$$

Exercices plus difficiles

Exercice 3.69. Renaud possède un lecteur de musique, où n chansons numérotées de 1 à n sont enregistrées. Le lecteur ne possède que deux boutons : *suivant*, qui passe de la chanson i à la chanson $i+1$, et *aléatoire*, qui choisit une chanson au hasard, uniformément dans $\{1,\ldots,n\}$. Lorsqu'on allume le lecteur, la chanson 1 se met en route, mais Renaud souhaite absolument écouter la chanson n. Il adopte la stratégie suivante :

(1) il se fixe un entier $k \in \{1,\ldots,n\}$;

(2) il appuie sur le bouton *aléatoire* jusqu'à ce qu'il tombe sur une chanson avec un numéro supérieur ou égal à k ;

(3) il n'appuie ensuite que sur le bouton *suivant*, jusqu'à arriver à la chanson n.

Soient $Y_1, Y_2, \ldots$ des variables aléatoires indépendantes, de même loi uniforme sur $\{1,\ldots,n\}$, qui décrivent les chansons obtenues en pressant successivement le bouton *aléatoire*. On note $T_k := \min\{i \geq 1, Y_i \in \{k, k+1, \ldots, n\}\}$ la première fois où l'on tombe sur une chanson avec un numéro supérieur ou égal à k et on note $W_k := Y_{T_k}$ le numéro de la chanson obtenue à cet instant-là.

(i) Donner la loi jointe et les lois marginales de T_k et W_k. Les variables aléatoires T_k et W_k sont-elles indépendantes ?

(ii) On note X_k le nombre de pressions nécessaires pour arriver à la chanson n avec cette stratégie. Exprimer X_k en fonction de T_k et de W_k, puis calculer $\mathrm{E}(X_k)$.

(iii) Pour quelle valeur k^* de k l'espérance $\mathrm{E}(X_k)$ est-elle minimale ? Que vaut $\mathrm{E}(X_{k^*})$? Donner le comportement asymptotique de k^* et de $\mathrm{E}(X_{k^*})$ lorsque n tend vers l'infini.

Solution 3.69. (i) Notons que T_k est à valeurs dans $\mathbb{N}^*$ et que W_k est à valeurs dans l'ensemble $\{k,k+1,\dots,n\}$. On doit calculer la densité discrète jointe de T_k et W_k, c'est-à-dire $\mathrm{P}(T_k = j, W_k = x)$ pour tout $j \in \mathbb{N}^*$ et $x \in \{k,k+1,\dots,n\}$.

Exprimons l'événement $\{T_k = j, W_k = x\}$ comme fonction des variables aléatoires $(Y_i)_{i\geq 1}$: en notant que $T_k = j$ et $W_k = Y_{T_k} = x$ si et seulement si $Y_i \notin \{k,\dots,n\}$ pour tout $i < j$ et $Y_j = x \in \{k,\dots,n\}$, on peut écrire

$$\{T_k = j, W_k = x\} = \left(\bigcap_{i=1}^{j-1} \{Y_i \notin \{k,\dots,n\}\}\right) \cap \{Y_j = x\}.$$

Comme les variables aléatoires $(Y_i)_{i\geq 1}$ sont indépendantes, on obtient

$$\mathrm{P}(T_k = j, W_k = x) = \left(\prod_{i=1}^{j-1} \mathrm{P}(Y_i \notin \{k,\dots,n\})\right) \times \mathrm{P}(Y_j = x) = \left(\frac{k-1}{n}\right)^{j-1} \times \frac{1}{n},$$

où pour la deuxième égalité on a utilisé le fait que $\mathrm{P}(Y_i = \ell) = \frac{1}{n}$ pour tout $\ell \in \{1,\dots,n\}$, d'où $\mathrm{P}(Y_i \notin \{k,\dots,n\}) = \mathrm{P}(Y_i \in \{1,\dots,k-1\}) = \sum_{\ell=1}^{k-1} \mathrm{P}(Y_i = \ell) = \frac{k-1}{n}$.

On a déterminé la densité discrète jointe de T_k et W_k. Déterminons les densités discrètes marginales à partir de celle-ci.

- Pour tout $j \in \mathbb{N}^*$,

$$\mathrm{P}(T_k = j) = \sum_{x=k}^{n} \mathrm{P}(T_k = j, W_k = x) = \sum_{x=k}^{n} \left(\frac{k-1}{n}\right)^{j-1} \times \frac{1}{n}$$
$$= \left(1 - \frac{n-k+1}{n}\right)^{j-1} \times \frac{n-k+1}{n}.$$

  Cela montre que T_k est de loi géométrique de paramètre $\frac{n-k+1}{n}$. On aurait pu arriver à la même conclusion de manière plus directe : il s'agit en effet de l'instant de premier succès lors de la répétition d'épreuves indépendantes, où un succès correspond à l'événement « le nombre est supérieur ou égal à k ».

- Pour tout $x \in \{k,\dots,n\}$, on a

$$\mathrm{P}(W_k = x) = \sum_{j=1}^{\infty} \mathrm{P}(T_k = j, W_k = x) = \sum_{j=1}^{\infty} \left(\frac{k-1}{n}\right)^{j-1} \times \frac{1}{n}$$
$$= \frac{1}{n} \times \left(\frac{n+k-1}{n}\right)^{-1} = \frac{1}{n-k+1},$$

  où on a utilisé le fait que $\sum_{j=1}^{n} p^{j-1} = (1-p)^{-1}$ avec $p = \frac{k-1}{n}$. On en déduit que W_k est de loi uniforme sur $\{k,\dots,n\}$ (qui est un ensemble de cardinal $n-k+1$).

Les deux variables aléatoires sont indépendantes, vu que la densité discrète jointe est le produit des densités discrètes marginales : pour tous $j \in \mathbb{N}^*$ et $x \in \{k, \dots, n\}$

$$\begin{aligned}\mathrm{P}(T_k = j, W_k = x) &= \Big(\frac{k-1}{n}\Big)^{j-1} \times \frac{1}{n} \\ &= \Big(1 - \frac{n-k+1}{n}\Big)^{j-1} \frac{n-k+1}{n} \times \frac{1}{n-k+1}\,.\end{aligned}$$

(ii) Pour atteindre la chanson n avec la stratégie proposée : a) il faut appuyer T_k fois sur le bouton *aléatoire* ; b) à l'instant T_k, on arrive à la position W_k : il reste à appuyer $n - W_k$ sur le bouton *suivant*. On a donc $X_k = T_k + n - W_k$.

Par linéarité de l'espérance, $\mathrm{E}(X_k) = \mathrm{E}(T_k) + n - \mathrm{E}(W_k)$. Comme T_k est une variable aléatoire géométrique de paramètre $p = \frac{n-k+1}{n}$, on a $\mathrm{E}(T_k) = \frac{n}{n-k+1}$. De plus, comme W_k est de loi uniforme sur $\{k, \dots, n\}$, on a $\mathrm{E}(W_k) = \frac{n+k}{2}$ (on peut par exemple remarquer que $W_k - k + 1$ est de loi uniforme sur $\{1, \dots, n-k+1\}$ et utiliser les calculs de la Section 3.5.1). On en conclut donc que

$$\mathrm{E}(X_k) = \frac{n}{n-k+1} + n - \frac{n+k}{2} = \frac{n}{n-k+1} + \frac{n-k}{2}\,.$$

(iii) Étudions la suite $u_k := \mathrm{E}(X_k)$ (plus précisément ses incréments) pour $1 \le k \le n$. En utilisant la formule trouvée au point précédent,

$$u_{k+1} - u_k = \frac{n}{n-k} + \frac{n-k-1}{2} - \frac{n}{n-k+1} - \frac{n-k}{2} = \frac{n}{(n-k)(n-k+1)} - \frac{1}{2}\,.$$

Ainsi, si $\frac{n}{(n-k)(n-k+1)} < \frac{1}{2}$ on a $u_{k+1} < u_k$, alors que si $\frac{n}{(n-k)(n-k+1)} > \frac{1}{2}$ on a $u_{k+1} > u_k$. Autrement dit, la suite u_k est décroissante jusqu'à $k \le k^*$ et elle est ensuite (strictement) croissante, où k^* est le plus grand entier tel que $\frac{n}{(n-k)(n-k+1)} < \frac{1}{2}$, c'est-à-dire

$$k^* = \max\big\{1 \le k \le n\,,\, (n-k)(n-k+1) \ge 2n\big\}\,.$$

En résolvant l'équation $(n-x)(n-x+1) = 2n$ en x, on trouve (après quelques simplifications) deux racines : $x_1 = n - \frac{1}{2}(\sqrt{8n+1} - 1)$ et $x_2 = n + \frac{1}{2}(\sqrt{8n+1} + 1)$. On a donc $(n-x)(n-x+1) \ge 2n$ pour $x < x_1$ et pour $x > x_2$. Comme $x_2 > n$, on obtient

$$k^* = \max\big\{k\,,\, k \le x_1\big\} = n - \big\lfloor \tfrac{1}{2}(\sqrt{8n+1} - 1)\big\rfloor\,.$$

On en déduit alors le comportement asymptotique de k^* :

$$k^* = n - \sqrt{2n} + O(1)\,, \qquad \text{quand } n \to \infty\,.$$

De plus

$$\begin{aligned}\mathrm{E}(X_{k^*}) &= \frac{n}{n-k^*+1} + \frac{n-k^*}{2} = \frac{n}{1 + \big\lfloor \frac{1}{2}(\sqrt{8n+1}-1)\big\rfloor} + \frac{1}{2}\big\lfloor \tfrac{1}{2}(\sqrt{8n+1}-1)\big\rfloor \\ &= \frac{n}{\sqrt{2n} + O(1)} + \frac{1}{2}(\sqrt{2n} + O(1)) = \sqrt{2n} + O(1)\,, \qquad \text{quand } n \to \infty\,.\end{aligned}$$

Ainsi, en choisissant k^* dans notre stratégie, on devra appuyer sur un bouton environ $\sqrt{2n}$ fois en moyenne pour atteindre la chanson n.

Exercice 3.70. Soit $(X_n)_{n\geq 1}$ des variables aléatoires à valeurs dans $\{-1,1\}$, indépendantes et de même loi : $P(X_1=+1)=P(X_1=-1)=1/2$. Pour $n\geq 1$, on pose $S_n=X_1+\cdots+X_n$.

(i) Soit $\lambda>0$. Montrer que $E(e^{\lambda S_n})=\cosh(\lambda)^n$.

(ii) Montrer que l'on a $\cosh(x)\leq \exp(\frac{1}{2}x^2)$ pour tout $x\in\mathbb{R}$. En déduire que $E(e^{\lambda S_n})\leq e^{\frac{1}{2}\lambda^2 n}$ pour tout $\lambda>0$.

[*Sugg.* Montrer que $g : x\mapsto \frac{1}{2}x^2-\log\cosh(x)$ est convexe et vérifie $g(0)=g'(0)=0$.]

(iii) Vérifier que pour tout $a>0$ et tout $\lambda>0$ on a

$$P(S_n\geq a)\leq \exp\Big(-\lambda a+\tfrac{1}{2}\lambda^2 n\Big).$$

Conclure que pour tout $a>0$, on a $P(S_n\geq a)\leq e^{-\frac{a^2}{2n}}$.

(iv) Montrer que $P(S_n\leq -a)=P(S_n\geq a)$ pour tout $a>0$ et en déduire que pour tout $x>0$ on a $P(|S_n|\geq x\sqrt{n})\leq 2e^{-\frac{1}{2}x^2}$.

Solution 3.70. (i) Il suffit d'observer que, pour tout $t\in\mathbb{R}$

$$M_{X_i}(t)=E(e^{tX_i})=\frac{1}{2}e^{+t}+\frac{1}{2}e^{-t}=\cosh(t),$$

donc

$$E(e^{\lambda S_n})=M_{S_n}(\lambda)=E\left(e^{\lambda(X_1+\cdots+X_n)}\right)=E(e^{\lambda X_1})\cdots E(e^{\lambda X_n})=(\cosh\lambda)^n$$

(ii) L'inégalité $E(e^{\lambda S_n})\leq e^{\frac{1}{2}\lambda^2 n}$ découle directement du point précédent et de l'inégalité $\cosh(x)\leq\exp(\frac{1}{2}x^2)$, que l'on déduit en montrant que

$$g(x):=\frac{1}{2}x^2-\log\cosh(x)\geq 0.$$

Le fait que g est convexe et que $g(0)=g'(0)=0$ implique que g atteint son minimum absolu pour $x=0$, donc l'inégalité ci-dessus est valable pour tout $x\in\mathbb{R}$. Ces propriétés de g se déduisent facilement des formules suivantes :

$$g'(x)=x-\tanh(x),\quad g''(x)=\tanh^2(x).$$

(iii) Pour tout $\lambda>0$, la fonction $x\mapsto e^{\lambda x}$ est strictement croissante de $\mathbb{R}_+$ dans $[1,+\infty[$. On a donc l'égalité d'événements $\{S_n\geq a\}=\{e^{\lambda S_n}\geq e^{\lambda a}\}$ pour tout $a>0$. Ainsi, en utilisant l'inégalité de Markov, on obtient

$$\begin{aligned}P(S_n\geq a)&=P\left(e^{\lambda S_n}\geq e^{\lambda a}\right)\leq\frac{1}{e^{\lambda a}}E\left(e^{\lambda S_n}\right)=e^{-a\lambda}M_{S_n}(\lambda)\\&\leq\exp\Big(-\lambda a+\frac{1}{2}\lambda^2 n\Big),\end{aligned}$$

pour tout $\lambda>0$, où pour la dernière inégalité on a utilisé le point précédent. Comme $\lambda>0$ est arbitraire, on a

$$P(S_n\geq a)\leq\exp\Big(\inf_{\lambda>0}\Big\{-a\lambda+\frac{n\lambda^2}{2}\Big\}\Big)=\exp\Big(-\frac{a^2}{2n}\Big),$$

où le minimum est obtenu pour $\lambda = \frac{a}{n}$.

(iv) Comme X_i et $-X_i$ on évidemment la même loi, on en déduit que S_n et $-S_n$ ont la même loi. Ainsi $\mathrm{P}(S_n \geq a) = \mathrm{P}(-S_n \geq a) = \mathrm{P}(S_n \leq -a)$. On a donc, pour tout $x > 0$, en utilisant l'inégalité démontrée au point précédent,

$$\mathrm{P}(|S_n| \geq x\sqrt{n}) = \mathrm{P}(S_n \geq x\sqrt{n}) + \mathrm{P}(S_n \leq -x\sqrt{n}) = 2\,\mathrm{P}(S_n \geq x\sqrt{n}) \leq 2\mathrm{e}^{-\frac{1}{2}x^2}\,.$$

Exercice 3.71. Pour $n \in \mathbb{N}^*$, on considère $(X_i)_{i\geq 1}$ une suite de variables aléatoires indépendantes, toutes de loi uniforme sur $\{1,\ldots,n\}$. On pose

$$T_n := \inf\big\{ i \geq 2 \colon X_i \in \{X_1,\ldots,X_{i-1}\} \big\}\,.$$

Il s'agit d'une reformulation du problème des anniversaires (auquel cas $n = 365$) : si on forme un groupe de personnes en ajoutant une personne à la fois, X_i représentant l'anniversaire de la i-ème personne, alors T_n représente la taille du groupe au premier instant où l'on trouve deux personnes ayant leur anniversaire le même jour.

(i) Montrer que pour tout $k \in \{2,\ldots,n\}$,

$$\mathrm{P}(T_n > k) = \prod_{j=1}^{k-1}\Big(1 - \frac{j}{n}\Big)\,.$$

(ii) Montrer que $\mathrm{P}(T_n > k) < \mathrm{e}^{-\frac{k(k-1)}{2n}}$, pour tout $k \in \mathbb{N}^*$. En déduire qu'une médiane de T_n doit être inférieure à $\tilde{m} := \lfloor \frac{1}{2}(1 + \sqrt{8n\log 2}\rfloor + 1$. Comment interprétez-vous ce résultat ?
[*Sugg.* On pourra utiliser que $1 + x < \mathrm{e}^x$ pour tout $x \in \mathbb{R}^*$.]

(iii) Montrer que $\lim_{n\to+\infty} \mathrm{P}(T_n > x\sqrt{n}) = \mathrm{e}^{-x^2/2}$, pour tout $x \geq 0$.

Solution 3.71. (i) On a que $T_n > k$ si et seulement si tous les $(X_j)_{1\leq j\leq k}$ sont distincts : en notant que $\mathrm{P}(X_{j+1} \notin \{X_1,\ldots,X_j\} \mid (X_i)_{1\leq i\leq j} \text{ tous distincts}) = 1 - \frac{j}{n}$, on obtient

$$\begin{aligned}\mathrm{P}(T_n > k) &= \mathrm{P}(X_2 \neq X_1)\,\mathrm{P}(X_3 \neq X_1, X_2 \mid X_2 \neq X_1) \\ &\qquad\qquad \cdots \mathrm{P}(X_k \neq X_1,\ldots,X_{k-1} \mid (X_i)_{1\leq i\leq k-1} \text{ tous distincts}) \\ &= \Big(1 - \frac{1}{n}\Big)\Big(1 - \frac{2}{n}\Big)\cdots\Big(1 - \frac{k-1}{n}\Big)\,.\end{aligned}$$

(ii) En utilisant l'inégalité suggérée $1 - \frac{j}{n} < \mathrm{e}^{-\frac{j}{n}}$, grâce à l'identité trouvée au point précédent, on obtient

$$\mathrm{P}(T_n > k) < \prod_{j=1}^{k-1} \mathrm{e}^{-\frac{j}{n}} = \mathrm{e}^{-\frac{k(k-1)}{2n}}\,,$$

où on a aussi utilisé le fait que $\sum_{j=1}^{k-1} j = \frac{k(k-1)}{2}$. Observons maintenant que comme T_n est une variable aléatoire à valeurs dans $\mathbb{N}$, elle admet une médiane $m \in \mathbb{N}$, qui satisfait $\mathrm{P}(T_n \geq m) \geq \frac{1}{2}$ et $\mathrm{P}(T_n \leq m) \geq \frac{1}{2}$. Si on montre que $\mathrm{P}(T_n \geq \tilde{m}+1) = \mathrm{P}(T_n > \tilde{m}) < \frac{1}{2}$, on peut alors conclure que $m < \tilde{m}+1$ et donc $m \leq \tilde{m}$. Il reste donc à montrer que $\mathrm{P}(T_n > \tilde{m}) < \frac{1}{2}$. Notons que $\mathrm{e}^{-\frac{k(k-1)}{2n}} < \frac{1}{2}$ si et seulement si $k > \frac{1}{2}(1 + \sqrt{1 + 8n\ln 2})$, donc il découle directement que

$$\mathrm{P}(T_n > \tilde{m}) < \mathrm{e}^{-\frac{\tilde{m}(\tilde{m}-1)}{2n}} < \frac{1}{2}.$$

On en déduit en particulier que $\mathrm{P}(T_n \leq \tilde{m}) \geq \frac{1}{2}$. Notons que pour $n = 365$ on a $\tilde{m} = 23$: on retrouve ce que l'on avait vu dans le paradoxe des anniversaires (Exemple 1.40).

(iii) En passant au logarithme, l'identité cherchée est équivalente à $\lim_{n\to\infty} \ln \mathrm{P}(T_n > x\sqrt{n}) = -\frac{1}{2}x^2$. Notons que l'on a $\mathrm{P}(T_n > x\sqrt{N}) = \mathrm{P}(T_n > k_n)$, où $k_n := \lfloor x\sqrt{n} \rfloor$ (remarquons que pour $x > 0$ fixé, $k_n \sim x\sqrt{n}$ quand $n \to \infty$) : d'après ce qu'on a vu dans la première question, on a donc

$$\ln \mathrm{P}(T_n > x\sqrt{n}) = \sum_{j=1}^{k_n-1} \ln\left(1 - \frac{j}{n}\right).$$

On utilise maintenant le développement limité $\ln(1-u) = -u + u^2/2 + o(u^2)$, dont il découle qu'il existe $u_0 > 0$ tel que pour tout $0 \leq u \leq u_0$ on ait $|\ln(1-u) + u| \leq u^2$. On utilisera cette estimée pour $u = \frac{j}{n}$: notons que pour tout $0 \leq j \leq k_n$ on a $0 \leq \frac{j}{n} \leq \frac{x}{\sqrt{n}}$, il existe donc un $N_0 = N_0(x)$ tel que pour tout $n \geq N_0$ on a $\frac{j}{n} \leq u_0$ pour tout $j \leq k_n$, et donc $|\ln(1-\frac{j}{n}) + \frac{j}{n}| \leq \frac{j^2}{n^2} \leq \frac{x^2}{n}$. Alors, pour $n \geq N_0$,

$$\left| \ln \mathrm{P}(T_n > x\sqrt{n}) + \sum_{j=1}^{k_n-1} \frac{j}{n} \right| \leq \sum_{j=1}^{k_n-1} \left| \ln\left(1 - \frac{j}{n}\right) - \frac{j}{n} \right| \leq k_n \frac{x^2}{n},$$

donc on déduit que

$$\left| \ln \mathrm{P}(T_n > x\sqrt{n}) + \frac{k_n(k_n-1)}{2n} \right| \leq \frac{x^3}{\sqrt{n}} \xrightarrow{n\to\infty} 0.$$

Pour conclure la démonstration, il suffit d'observer que, vu que $k_n \sim x\sqrt{n}$ comme observé plus haut, on a $\lim_{n\to\infty} \frac{k_n(k_n-1)}{2n} = \frac{x^2}{2}$.

Exercice 3.72. On considère le modèle simplifié suivant pour la propagation d'une épidémie. Au n-ème jour, on note Z_n le nombre d'individus contagieux. Chacun de ces individus contamine alors un nombre aléatoire de personnes avant d'être guéri au jour $n+1$: si X_i désigne le nombre d'individus contaminés par l'individu i lors du n-ème jour, alors on a $Z_{n+1} = \sum_{i=1}^{Z_n} X_i$; par convention, si $Z_n = 0$ alors $Z_{n+1} = 0$.

(i) On suppose que $X_i \sim \text{Géom}_0(1/2)$ (voir l'Observation 3.133) sont des variables aléatoires indépendantes, indépendantes de Z_n. Montrer que pour tout $k \in \mathbb{N}$ on a la relation suivante entre la densité discrète de Z_{n+1} et celle de Z_n

$$\mathrm{p}_{Z_{n+1}}(k) = \mathrm{p}_{Z_n}(0)\mathbb{1}_{\{0\}}(k) + \sum_{j=1}^{\infty} \mathrm{p}_{Z_n}(j) \binom{k+j-1}{j-1} 2^{-(k+j)}.$$

[*Sugg.* On pourra utiliser la question (i) de l'Exercice 3.65, que l'on pourra appliquer aux variables aléatoires $X_i' := X_i + 1$, de loi géométrique « standard » $\text{Géom}(1/2)$.]

(ii) On suppose qu'il y a un unique individu contagieux au jour 0, c'est-à-dire $Z_0 = 1$. Montrer par récurrence que la densité discrète de Z_n est donnée par

$$\mathrm{p}_{Z_n}(0) = q_n \quad \text{et} \quad \mathrm{p}_{Z_n}(k) = (1-q_n)^2 q_n^{j-1} \quad \text{pour } k \geq 1,$$

où q_n est déterminé par $q_0 = 0$ et la relation $q_{n+1} = \frac{1}{2-q_n}$ pour $n \geq 0$.

[Sugg. On pourra utiliser/démontrer que $(1-x)^{-(k+1)} = \sum_{i=0}^{\infty} \binom{k+i}{i} x^i$ pour tout $x \in [0,1[$.]

(iii) Déterminer q_n, la probabilité qu'il n'y ait aucun individu contagieux après n jours.

(iv) Déterminer la loi et l'espérance de Z_n sous la probabilité conditionnelle $P(\cdot \mid Z_n > 0)$.

Solution 3.72. (i) Commençons par remarquer que, par σ-additivité, on a

$$P(Z_{n+1} = k) = \sum_{j=0}^{+\infty} P(Z_{n+1} = k, Z_n = j).$$

Par définition de Z_{n+1} on a $P(Z_{n+1} = 0 \mid Z_n = 0) = 1$, donc l'identité précédente peut se réécrire comme

$$\begin{aligned} P(Z_{n+1} = k) &= \mathbb{1}_{\{0\}}(k) P(Z_n = 0) + \sum_{j=1}^{+\infty} P(Z_{n+1} = k, Z_n = j) \\ &= \mathbb{1}_{\{0\}}(k) P(Z_n = 0) + \sum_{j=1}^{+\infty} P\left(\sum_{i=1}^{j} X_i = k, Z_n = j\right) \\ &= \mathbb{1}_{\{0\}}(k) P(Z_n = 0) + \sum_{j=1}^{+\infty} p_{Z_n}(j) P\left(\sum_{i=1}^{j} X_i = k\right), \end{aligned}$$

où on a utilisé le fait que Z_n et $\sum_{i=1}^{j} X_i$ sont indépendantes. Pour conclure, il reste à montrer que

$$P\left(\sum_{i=1}^{j} X_i = k\right) = \binom{k+j-1}{j-1} 2^{-(k+j)}.$$

Pour le vérifier, on pose $X_i' = X_i + 1 \sim \text{Géom}(1/2)$: il suffit d'observer que

$$P\left(\sum_{i=1}^{j} X_i = k\right) = P\left(\sum_{i=1}^{j} X_i' = k - j\right),$$

et d'utiliser la question (i) de l'Exercice 3.65.

(ii) Remarquons tout d'abord que le résultat est vrai pour $n = 0$, vu que $p_{Z_0}(1) = 1$. Pour l'étape de récurrence, commençons par considérer le cas $k = 0$. D'après l'identité montrée au point précédent et l'hypothèse de récurrence, on a

$$\begin{aligned} p_{Z_{n+1}}(0) &= q_n + \sum_{j=1}^{+\infty} (1-q_n)^2 q_n^{j-1} 2^{-j} = q_n + \frac{(1-q_n)^2}{2} \sum_{j=1}^{+\infty} \left(\frac{q_n}{2}\right)^{j-1} \\ &= q_n + \frac{(1-q_n)^2}{2} \frac{1}{1 - \frac{q_n}{2}} = \frac{1}{2 - q_n}, \end{aligned}$$

où on a utilisé la somme d'une série géométrique. Cela conclut le cas $k = 0$. Pour $k \geq 1$, en utilisant l'hypothèse de récurrence, on obtient

$$
\begin{aligned}
\mathrm{p}_{Z_{n+1}}(k) &= \sum_{j=1}^{+\infty}(1-q_n)^2 q_n^{j-1}\binom{k+j-1}{j-1}2^{-(k+j)} \\
&= (1-q_n)^2 2^{-(k+1)}\sum_{j=1}^{+\infty}\binom{k+j-1}{j-1}2^{-(k+j)}\left(\frac{q_n}{2}\right)^{j-1} \\
&= (1-q_n)^2 2^{-(k+1)}\left(1-\frac{q_n}{2}\right)^{k+1} = \frac{(1-q_n)^2}{(2-q_n)^{k+1}}\,,
\end{aligned}
$$

où on a utilisé l'identité suggérée $(1-x)^{-(k+1)} = \sum_{i=0}^{\infty}\binom{k+i}{i}x^i$. Pour conclure la démonstration de l'étape de récurrence, il suffit d'observer que

$$
(1-q_{n+1})^2 q_{n+1}^{k-1} = \left(1-\frac{1}{2-q_n}\right)^2\frac{1}{(2-q_n)^{k-1}} = \frac{(1-q_n)^2}{(2-q_n)^{k+1}}\,.
$$

Enfin, on peut démontrer l'identité $(1-x)^{-(k+1)} = \sum_{i=0}^{\infty}\binom{k+i}{i}x^i$ pour $x\in[0,1[$ de plusieurs manières. Par exemple, ce qui est montré dans l'Exercice 1.48 implique que

$$
\binom{k+i}{i} = \bigl|\{(n_1,\dots,n_{k+1})\in\mathbb{N}^{k+1} : n_1+\cdots+n_{k+1}=i\}\bigr|\,.
$$

Alors, en posant $A_i := \{(n_1,\dots,n_{k+1})\in\mathbb{N}^{k+1} : n_1+\cdots+n_{k+1}=i\}$ et en utilisant la sommation par paquets, on obtient

$$
\begin{aligned}
\sum_{i=0}^{\infty}\binom{k+i}{i}x^i = \sum_{i=0}^{\infty}\sum_{(n_1,\dots,n_{k+1})\in A_i}x^i &= \sum_{(n_1,\dots,n_{k+1})\in\mathbb{N}^{k+1}}x^{n_1+\cdots+n_{k+1}} \\
&= \left(\sum_{n=0}^{+\infty}x^n\right)^{k+1} = (1-x)^{-(k+1)}\,.
\end{aligned}
$$

(iii) En utilisant la relation $q_{n+1} = \frac{1}{2-q_n}$ et le fait que $q_0 = 0$, on montre facilement par récurrence que $q_n = \frac{n}{n+1}$.

(iv) Pour $k\geq 1$ on a

$$
\begin{aligned}
\mathrm{P}(Z_n=k\,|\,Z_n>0) &= \frac{\mathrm{P}(Z_n=k,Z_n>0)}{\mathrm{P}(Z_n>0)} = \frac{\mathrm{P}(Z_n=k)}{1-\mathrm{P}(Z_n=0)} = \frac{(1-q_n)^2q_n^{k-1}}{1-q_n} \\
&= (1-q_n)q_n^{k-1}\,.
\end{aligned}
$$

Ainsi, par rapport à la probabilité conditionnelle $\mathrm{P}(\cdot\,|\,Z_n>0)$, la loi de Z_n est Géométrique de paramètre $1-q_n = \frac{1}{n+1}$, dont l'espérance vaut $n+1$.

Chapitre 6
Variables aléatoires absolument continues

6.2 Variables aléatoires à densité

Exercice 6.1. Soit X une variable aléatoire réelle à densité et soit $]a,b[$ un intervalle ouvert (borné ou non) de $\mathbb{R}$ tel que $\mathrm{P}(X \in\,]a,b[) = 1$. Soit $\varphi :]a,b[\to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ telle que $\varphi'(x) > 0$ (ou < 0) pour tout $x \in\,]a,b[$. Alors l'image de φ est un intervalle ouvert $]c,d[$. L'objectif de cet exercice est de montrer que la variable aléatoire $Y := \varphi(X)$ est à densité, de densité donnée par

$$f_Y(y) = \frac{1}{|\varphi'(\varphi^{-1}(y))|}\, f_X\left(\varphi^{-1}(y)\right) \mathbb{1}_{]c,d[}(y)\,, \tag{S6.1}$$

où $\varphi^{-1} :]c,d[\to]a,b[$ désigne la fonction inverse de φ pour la composition.

Considérons le cas $\varphi'(x) > 0$ pour tout $x \in\,]a,b[$; le cas $\varphi'(x) < 0$ est analogue.

(i) Montrer que la fonction de répartition de Y est donnée par

$$F_Y(y) = \begin{cases} 0 & \text{si } y \le c\,, \\ F_X(\varphi^{-1}(y)) & \text{si } c < y < d\,, \\ 1 & \text{si } y \ge d\,. \end{cases}$$

(ii) En supposant que F_X est $\mathcal{C}^1$ par morceaux, conclure que Y est à densité, de densité f_Y donnée par (S6.1).

(iii) Arriver à la même conclusion sans supposer que F_X est $\mathcal{C}^1$ par morceaux, en utilisant la formule de changement de variables (6.12).

Solution 6.1. Supposons que $\varphi'(x) > 0$ pour tout $x \in\,]a,b[$ (le cas $\varphi'(x) < 0$ est analogue). Ainsi φ est strictement croissante, donc l'image $\varphi(]a,b[)$ est un intervalle ouvert $]c,d[$ et la fonction φ admet un inverse $\varphi^{-1} :]c,d[\to]a,b[$, qui est aussi de classe $\mathcal{C}^1$. De plus, pour tout $y \in\,]c,d[$

$$\left(\varphi^{-1}\right)'(y) = \frac{1}{\varphi'(\varphi^{-1}(y))}. \tag{S6.2}$$

(i) Déterminons la fonction de répartition de Y. Tout d'abord $F_Y(y) = 0$ si $y \le c$ alors que $F_Y(y) = 1$ si $y \ge d$, parce que par hypothèse $\mathrm{P}(X \in\,]a,b[) = 1$, donc $\mathrm{P}(Y \in\,]c,d[) = 1$. Pour $y \in\,]c,d[$, en utilisant la monotonie stricte de φ, on a $\varphi(x) \le y$ si et seulement si $x \le \varphi^{-1}(y)$, donc

$$F_Y(y) = \mathrm{P}(\varphi(X) \le y) = \mathrm{P}(X \le \varphi^{-1}(y)) = F_X(\varphi^{-1}(y))\,. \tag{S6.3}$$

(ii) Supposons que F_X soit $\mathcal{C}^1$ par morceaux (elle est nécessairement continue). Alors F_Y est aussi continue et $\mathcal{C}^1$ par morceaux, parce que φ^{-1} est une fonction de classe $\mathcal{C}^1$. D'après la Proposition 6.17, Y est une variable aléatoire absolument continue, de densité

$$f_Y(y) = F_Y'(y) = F_X'(\varphi^{-1}(y)) \cdot (\varphi^{-1})'(y)\, \mathbb{1}_{]c,d[}(y)\,.$$

En se rappelant de (S6.2) on obtient la relation cherchée (S6.1).

(iii) Dans le cas général, pour $y \in\,]c,d[$ on peut écrire, grâce à (S6.3) et au fait que $f_X(x) = 0$ si $x \notin\,]a,b[$,

$$F_Y(y) = F_X(\varphi^{-1}(y)) = \int_a^{\varphi^{-1}(y)} f_X(x)\, \mathrm{d}x\,.$$

Comme la fonction restreinte $\varphi^{-1} :]c,y[\to]a,\varphi^{-1}(y)[$ est de classe $\mathcal{C}^1$, avec le changement de variables $x = \psi(t) := \varphi^{-1}(t)$ (pour un rappel, voir la formule (6.12)) l'intégrale peut être réécrite sous la forme

$$F_Y(y) = \int_a^{\varphi^{-1}(y)} f_X(x)\, \mathrm{d}x = \int_c^y f_X(\varphi^{-1}(t))\, (\varphi^{-1})'(t)\, \mathrm{d}t\,.$$

En se rappelant de la relation (S6.2) et en notant $g(y)$ la fonction dans le membre de droite de (S6.1), on a montré que $F_Y(y) = \int_{-\infty}^{y} g(t)\, \mathrm{d}t$, pour tout $y \in\,]c,d[$. Il est facile de vérifier que cette relation reste valable pour $y \leq c$ et $y \geq d$, parce que chacun des membres vaut 0 et 1 respectivement. En se rappelant de la Définition 6.12, il s'ensuit que Y est une variable aléatoire absolument continue, de densité $f_Y(y) = g(y)$, qui est exactement ce que l'on devait montrer.

Exercice 6.2. Soit X une variable aléatoire réelle à densité, de densité $f(x)$. Montrer que la variable aléatoire $Y := |X|$ est à densité et en déterminer la densité f_Y.

Solution 6.2. Déterminons la fonction de répartition de Y. Tout d'abord, $F_Y(y) = 0$ si $y \leq 0$, parce que $\mathrm{P}(|X| \leq 0) = \mathrm{P}(X = 0) = 0$, étant donné que X est à densité. Pour $y > 0$, en utilisant que $|X| \leq y$ si et seulement si $X \in [-y,y]$, on a

$$F_Y(y) = \mathrm{P}(|X| \leq y) = \mathrm{P}(X \in [-y,y]) = F_X(y) - F_X(-y)\,,$$

où l'on renvoie à (6.17) pour la justification de la dernière égalité.

Dans le cas où F_X est $\mathcal{C}^1$ par morceaux, alors F_Y l'est aussi : grâce à la Proposition 6.17, la variable aléatoire Y est absolument continue, de densité donnée par :

$$f_Y(y) = F_Y'(y) = \begin{cases} 0 & \text{si } y < 0\,, \\ F_X'(y) + F_X'(-y) = f_X(y) + f_X(-y) & \text{si } y > 0\,, \end{cases}$$

que l'on peut écrire sous la forme $f_Y(y) = (f_X(y) + f_X(-y))\mathbb{1}_{\mathbb{R}_+}(y)$.

Si l'on ne suppose pas que F_X est $\mathcal{C}^1$ par morceaux, on écrit, pour $y > 0$,

$$\begin{aligned} F_Y(y) = \mathrm{P}(X \in [-y,y]) &= \int_{-y}^{y} f_X(x)\mathrm{d}x \\ &= \int_{-y}^{0} f_X(x)\mathrm{d}x + \int_0^y f_X(x)\mathrm{d}x = \int_0^y f_X(-u)\mathrm{d}u + \int_0^y f_X(x)\mathrm{d}x\,, \end{aligned}$$

où on a utilisé le changement de variable $u = -x$ pour la première intégrale. En résumant, on a trouvé

$$\begin{aligned}F_Y(y) &= \begin{cases} 0 & \text{si } y \le 0 \\ \displaystyle\int_0^y (f_X(x)+f_X(-x))\mathrm{d}x & \text{si } y > 0 \end{cases} \\ &= \int_{-\infty}^{y} (f_X(x)+f_X(-x))\mathbb{1}_{\mathbb{R}_+}(x)\mathrm{d}x\,.\end{aligned}$$

En se rappelant de la Définition 6.12, on en conclut que Y est absolument continue, de densité $f_Y(y) = (f_X(y)+f_X(-y))\mathbb{1}_{\mathbb{R}_+}(y)$.

Exercice 6.3. Soit X une variable aléatoire à densité *positive*, de densité f_X que l'on suppose continue par morceaux. Montrer que, pour tout $k \ge 1$

$$\mathrm{E}(X^k) = \int_0^{\infty} k t^{k-1}\,\mathrm{P}(X > t)\mathrm{d}t\,. \tag{S6.4}$$

[Soulignons que pour $k = 1$, il s'agit de l'analogue continu de la Proposition 3.97.]

Solution 6.3. Tout d'abord, comme X est une variable aléatoire positive, on sait que $\mathrm{E}(X^k)$ est toujours bien définie (mais peut prendre la valeur $+\infty$) .

D'après la formule de transfert (6.20) (voir (6.21) pour la formule pour les moments), on a

$$\mathrm{E}(X^k) = \int_0^{+\infty} t^k f(t)\mathrm{d}t\,,$$

où l'on a restreint l'intégrale à $\mathbb{R}_+$ parce que la variable aléatoire X est positive (on peut supposer que $f_X(t) = 0$ pour $t < 0$).

Posons maintenant $H(t) := P(X > t) = 1 - F_X(t)$ pour $t > 0$: la dérivée de la fonction H est donnée par $H'(t) = -f(t)$. Procédons à une intégration par parties (pour un rappel, voir la formule (6.10)), en considérant l'intégrale sur un intervalle $[0,a]$ avec $a > 0$ pour commencer. On a

$$\begin{aligned}\int_0^a t^k f_X(t)\mathrm{d}t &= \Big[-t^k H(t)\Big]_0^a + \int_0^a k t^{k-1} H(t)\mathrm{d}t \\ &= -a^k\,\mathrm{P}(X > a) + \int_0^a k t^{k-1}\,\mathrm{P}(X > t)\mathrm{d}t\,.\end{aligned} \tag{S6.5}$$

On doit maintenant prendre la limite $a \to +\infty$. Considérons deux cas : $\mathrm{E}(X^k) = +\infty$ ou bien $\mathrm{E}(X^k) < +\infty$.

Premier cas. $\mathrm{E}(X^k) = +\infty$. Dans ce cas, grâce à (S6.5), on obtient que pour tout $a > 0$

$$\int_0^a t^k f_X(t)\mathrm{d}t \le \int_0^a k t^{k-1}\,\mathrm{P}(X > t)\mathrm{d}t\,.$$

En prenant $a \to +\infty$, le terme de gauche converge vers $\mathrm{E}(X^k) = +\infty$, ce qui donne

$$\lim_{a\to+\infty}\int_0^a k t^{k-1}\,\mathrm{P}(X > t)\mathrm{d}t = \int_0^{+\infty} k t^{k-1}\,\mathrm{P}(X > t)\mathrm{d}t = +\infty \;\; (= \mathrm{E}(X^k))\,.$$

Deuxième cas. $\mathrm{E}(X^k) < +\infty$. Dans ce cas, montrons que $\lim_{a\to+\infty} a^k\,\mathrm{P}(X > a) = 0$. En effet, pour tout $a > 0$ on a

$$\int_a^{+\infty} x^k f(x)\mathrm{d}x \ge a^k \int_a^{\infty} f(x)\mathrm{d}x = a^k\,\mathrm{P}(X > a)\,.$$

Vu que $\int_0^\infty x^k f(x)\mathrm{d}x < +\infty$, le terme de gauche tend vers 0 quand $a \to +\infty$, montrant ainsi que $\lim_{a\to+\infty} a^k \mathrm{P}(X > a) = 0$.

En prenant la limite $a \to +\infty$ dans (S6.5), on a donc

$$\mathrm{E}(X^k) = \lim_{a\to+\infty} \int_0^a t^k f_X(t)\mathrm{d}t = \lim_{a\to+\infty} \int_0^a k t^{k-1} \mathrm{P}(X > t)\mathrm{d}t = \int_0^{+\infty} k t^{k-1} \mathrm{P}(X > t)\mathrm{d}t\,.$$

Exercice 6.4. Soit X une variable aléatoire réelle dont la fonction de répartition F_X est

$$F_X(t) = \begin{cases} 0 & \text{si } t < 0\,, \\ 1 - \frac{1}{2}\mathrm{e}^{-t} & \text{si } t \geq 0\,. \end{cases}$$

Après avoir montré que F_X est bien une fonction de répartition, c'est-à-dire vérifie les propriétés (5.5), montrer que la variable aléatoire X n'est ni à densité, ni discrète.

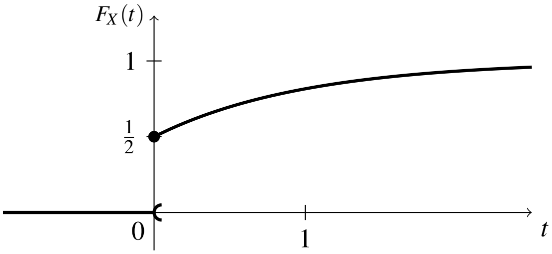
Solution 6.4. Vérifions que la fonction F_X vérifie les propriétés (5.5). Il s'agit en effet d'une fonction à valeurs dans $[0,1]$ qui est :

(i) croissante, parce qu'elle est croissante sur $]-\infty,0[$, sur $[0,+\infty[$ et $F_X(0) \geq F_X(0^-) := \lim_{t\uparrow 0} F_X(t)$;

(ii) continue à droite parce qu'elle est continue sur $]-\infty,0[$, sur $]0,+\infty[$, et $F_X(0^+) = \lim_{t\downarrow 0}(1 - \frac{1}{2}\mathrm{e}^{-t}) = 1 - \frac{1}{2} = \frac{1}{2} = F_X(0)$.

(iii) $\lim_{t\to-\infty} F_X(t) = 0$ parce que $F_X(t) = 0$ pour tout $t < 0$;
$\lim_{t\to+\infty} F_X(t) = \lim_{t\to+\infty}(1 - \frac{1}{2}\mathrm{e}^{-t}) = 1$ parce que $\lim_{t\to+\infty} \mathrm{e}^{-t} = 0$.

Le graphe de la fonction F_X est représenté ci-dessous.



Pour voir que X n'est pas une variable aléatoire absolument continue, il suffit de remarquer que la fonction de répartition n'est pas continue en $t = 0$: en se rappelant de la relation (5.8) (démontrée dans la Proposition 3.95), on a

$$\mathrm{P}(X = 0) = F_X(0) - F_X(0^-) = \frac{1}{2} > 0\,.$$

Cela montre que X ne peut pas être absolument , parce que si c'était le cas on devrait avoir $\mathrm{P}(X = x) = 0$ pour tout $x \in \mathbb{R}$, voir (6.16).

Pour voir que X n'est pas discrète, notons que F_X est continue sur $]-\infty,0[$ et sur $]0,+\infty[$. Ainsi, toujours grâce à la relation (5.8), on a

$$\mathrm{P}(X = x) = F_X(x) - F_X(x^-) = 0\,, \qquad \forall x \neq 0\,.$$

On peut donc écrire que

$$\sum_{x\in\mathbb{R}} \mathrm{P}(X = x) = \mathrm{P}(X = 0) = \frac{1}{2}\,,$$

car tous les termes autres que $x = 0$ sont nuls. Cela montre que X ne peut pas être discrète, parce que dans ce cas on aurait $\sum_{x\in\mathbb{R}} \mathrm{P}(X = x) = 1$, d'après la Proposition 3.13.

Exercice 6.5. Soient $X_1,\ldots,X_n$ des variables aléatoires réelles indépendantes et de même loi. Notons $F = F_{X_i}$ leur fonction de répartition commune et supposons que F est continue et $\mathcal{C}^1$ par morceaux, de sorte que $f(x) = F'(x)$ est la densité des X_i. Montrer que les variables aléatoires $Z := \max(X_1,\ldots,X_n)$ et $W := \min(X_1,\ldots,X_n)$ sont à densité, de densités

$$f_Z(x) = n\big(F(x)\big)^{n-1} f(x)\,, \qquad f_W(x) = n\big(1-F(x)\big)^{n-1} f(x)\,. \tag{S6.6}$$

Solution 6.5. En appliquant la Proposition 3.107 (dont la démonstration est valable, sans aucune modification, pour des variables aléatoires générales), les fonctions de répartition de W et Z sont données respectivement par

$$F_Z(x) = \prod_{k=1}^{n} F_{X_k}(x) = F(x)^n\,, \quad F_W(x) = 1 - \prod_{k=1}^{n}[1-F_{X_k}(x)] = 1 - \big(1-F(x)\big)^n\,.$$

Comme par hypothèse F est continue et $\mathcal{C}^1$ par morceaux, les fonctions F_Z et F_W le sont aussi : d'après la Proposition 6.17, Z et W sont des variables aléatoires absolument continues, de densités données par

$$\begin{aligned} f_Z(x) &= (F_Z)'(x) = n\big(F(x)\big)^{n-1} F'(x) = n\big(F(x)\big)^{n-1} f(x)\,,\\ f_W(x) &= (F_W)'(x) = n\big(1-F(x)\big)^{n-1} F'(x) = n\big(1-F(x)\big)^{n-1} f(x)\,. \end{aligned}$$

6.3 Exemples importants de variables aléatoires à densité

Exercice 6.6. Soit X une variable aléatoire $\mathrm{U}(a,b)$, avec $a < b$, et soient α,β deux réels, avec $\alpha \neq 0$. Montrer que la variable aléatoire $Y = \alpha X + \beta$ suit aussi une loi uniforme sur un intervalle dont on déterminera les extrémités.
[*Sugg.* Considérer les cas $\alpha > 0$ et $\alpha < 0$ séparément.]

Solution 6.6. En appliquant la Proposition 6.40, on obtient que $Y = \alpha X + \beta$ est une variable aléatoire absolument continue, de densité

$$f_Y(y) = \frac{1}{|\alpha|} f_X\Big(\frac{y-\beta}{\alpha}\Big) = \frac{1}{|\alpha|}\frac{1}{b-a}\mathbb{1}_{]a,b[}\Big(\frac{y-\beta}{\alpha}\Big)\,.$$

Dans le cas $\alpha > 0$, on a $\frac{y-\beta}{\alpha} \in \,]a,b[$ si et seulement si $y-\beta \in \,]\alpha a, \alpha b[$, si et seulement si $y \in \,]\alpha a+\beta, \alpha b+\beta[$. On en conclut que

$$f_Y(y) = \frac{1}{\alpha(b-a)}\mathbb{1}_{]\alpha a+\beta,\alpha b+\beta[}(y)\,,$$

c'est-à-dire $Y \sim \mathrm{U}(\alpha a+\beta, \alpha b+\beta)$.

Dans le cas $\alpha < 0$, on a $\frac{y-\beta}{\alpha} \in \,]a,b[$ si et seulement si $y-\beta \in \,]\alpha b, \alpha a[$, si et seulement si $y \in \,]\alpha b+\beta, \alpha a+\beta[$. On en conclut que

$$f_Y(y) = \frac{1}{|\alpha|(b-a)}\mathbb{1}_{]\alpha b+\beta,\alpha a+\beta[}(y) = \frac{1}{(\alpha a-\alpha b)}\mathbb{1}_{]\alpha b+\beta,\alpha a+\beta[}(y)\,,$$

c'est-à-dire $Y \sim \mathrm{U}(\alpha b+\beta, \alpha a+\beta)$.

Autrement dit, si $X \sim \mathrm{U}(a,b)$, alors la variable aléatoire $Y := \alpha X+\beta$ est uniforme continue sur l'intervalle $]c,d[$ qui est l'image de l'intervalle $]a,b[$ par la transformation affine $x \mapsto \alpha x+\beta$.

Exercice 6.7. Soit U une variable aléatoire de loi $\mathrm{U}(0,1)$ et $n \in \mathbb{N}^*$ un entier. Montrer que la variable aléatoire $Y := \lfloor nX \rfloor$, où $\lfloor x \rfloor$ désigne la partie entière de x, suit la loi uniforme (discrète) sur $\{0,\dots,n-1\}$.

Solution 6.7. Tout d'abord, notons que $nX \in]0,n[$ avec probabilité 1, donc $Y := \lfloor nX \rfloor$ est à valeurs dans $\{0,1,\dots,n-1\}$. En particulier, il s'agit d'une variable aléatoire discrète, dont on détermine maintenant la densité discrète.

Pour $k \in \{0,1,\dots,n-1\}$, remarquons que par définition de la partie entière, on a l'égalité d'événements $\{Y=k\} = \{nX \in [k,k+1[\,\} = \{X \in [\frac{k}{n}, \frac{k+1}{n}[\,\}$. Ainsi, en utilisant la densité de X, on trouve

$$\mathrm{P}(Y=k) = \mathrm{P}\left(X \in \left[\frac{k}{n}, \frac{k+1}{n}\right[\right) = \int_{\frac{k}{n}}^{\frac{k+1}{n}} \mathrm{d}x = \frac{k+1}{n} - \frac{k}{n} = \frac{1}{n}, \quad \forall k \in \{0,1,\dots,n-1\}.$$

On a déterminé la densité discrète de Y, qui correspond à celle d'une variable aléatoire uniforme discrète sur $\{0,1,\dots,n-1\}$.

Exercice 6.8. Soit U une variable aléatoire de loi $\mathrm{U}(0,1)$. Montrer que $-\log U$ est une variable aléatoire de loi $\mathrm{Exp}(1)$.

Solution 6.8. Déterminons la fonction de répartition de $Y := -\log U$. Tout d'abord, notons que $Y > 0$ avec probabilité 1, parce que $U \in]0,1[$ avec probabilité 1. Ainsi, $F_Y(t) = 0$ pour tout $t \le 0$.

Pour $t > 0$, en utilisant l'égalité d'événements

$$\{Y \le t\} = \{-\log U \le t\} = \{\log U \ge -t\} = \{U \ge \mathrm{e}^{-t}\},$$

on a

$$F_Y(t) = \mathrm{P}(U \ge \mathrm{e}^{-t}) = \mathrm{P}(U \in [\mathrm{e}^{-t}, 1]) = \int_{\mathrm{e}^{-t}}^{1} \mathrm{d}x = 1-\mathrm{e}^{-t},$$

où on a aussi utilisé la densité de U pour déterminer la valeur de $\mathrm{P}(U \ge \mathrm{e}^{-t})$ (on aurait pu utiliser directement la formule pour la fonction de répartition de U donnée dans (6.29)). On a donc montré que

$$F_Y(t) = \begin{cases} 0 & \text{si } t \le 0, \\ 1-\mathrm{e}^{-t} & \text{si } t > 0. \end{cases}$$

On reconnaît la fonction de répartition d'une variable aléatoire $\mathrm{Exp}(1)$, voir (6.44) : comme la fonction de répartition caractérise la loi d'une variable aléatoire (Proposition 5.15), on en conclut que $Y = -\log U \sim \mathrm{Exp}(1)$. De manière alternative, on peut noter que F_Y est continue et $\mathcal{C}^1$ par morceaux (il faut contrôler la continuité en 0), donc grâce à la Proposition 6.17 on obtient que Y est absolument continue, de densité $f_Y(y) = F_Y'(t) = \mathrm{e}^{-t}\mathbb{1}_{]0,\infty[}(t)$, c'est-à-dire $Y \sim \mathrm{Exp}(1)$.

Une autre manière de déterminer la densité de $Y = -\log U$ est d'utiliser l'Observation 6.26. Fixons une fonction $h : \mathbb{R} \to \mathbb{R}$ continue par morceaux et bornée continue, quelconque (une *fonction test* pour la formule de transfert (6.23)), et calculons $\mathrm{E}(h(Y)) = \mathrm{E}(h(-\log U))$. En appliquant la formule de transfert à U pour la fonction (bornée) $h(-\log u)$, on obtient

$$\mathrm{E}(h(Y)) = \int_{-\infty}^{+\infty} h(-\log u)\mathbb{1}_{]0,1[}(u)\mathrm{d}u = \int_0^1 h(-\log u)\mathrm{d}u = \int_0^{+\infty} h(y)\mathrm{e}^{-y}\mathrm{d}y,$$

où on a utilisé le changement de variables $y = -\log u$ (ou bien $u = \mathrm{e}^{-y}$).

Comme la relation $\mathrm{E}(h(Y)) = \int_{-\infty}^{+\infty} h(y)\mathrm{e}^{-y}\mathbb{1}_{]0,\infty[}(y)\mathrm{d}y$ est valable pour tout fonction $h : \mathbb{R} \to \mathbb{R}$ continue par morceaux et bornée, il découle de l'Observation 6.26 que Y est absolument continue et que sa densité est donnée par $\mathrm{e}^{-y}\mathbb{1}_{]0,\infty[}(y)$, c'est-à-dire $Y \sim \mathrm{Exp}(1)$.

Exercice 6.9. Soit X une variable aléatoire de loi $\mathrm{Exp}(\lambda)$. Pour $s \in [0, +\infty[$, on définit

$$\mathrm{P}^{(s)} := \mathrm{P}(\cdot\,|X > s), \qquad X^{(s)} := X - s.$$

Montrer que, par rapport à la probabilité (conditionnelle) $\mathrm{P}^{(s)}$, la variable aléatoire $X^{(s)}$, appelée *temps résiduel*, a pour loi $\mathrm{Exp}(\lambda)$, quel que soit $s \in [0, \infty[$.

[*Sugg.* Calculer la fonction de répartition $\mathrm{P}^{(s)}(X^{(s)} \leq x)$, en utilisant la Proposition 6.38.]

Solution 6.9. On déduit de la Proposition 6.38 que pour tout $x \geq 0$

$$\begin{aligned}\mathrm{P}^{(s)}(X^{(s)} \leq x) &= \mathrm{P}(X - s \leq x\,|X > s) = \mathrm{P}(X \leq x + s\,|X > s) \\ &= 1 - \mathrm{P}(X > x + s\,|X > s) = 1 - \mathrm{P}(X > x) = \mathrm{P}(X \leq x).\end{aligned}$$

Comme pour $x < 0$ on a $\mathrm{P}^{(s)}(X^{(s)} \leq x) = \mathrm{P}(X \leq x) = 0$, on a montré que la variable aléatoire $X^{(s)}$, par rapport à la probabilité (conditionnelle) $\mathrm{P}^{(s)}$, possède la même fonction de répartition, donc la même loi, que la variable aléatoire X par rapport à la probabilité originelle P, c'est-à-dire est de loi $\mathrm{Exp}(\lambda)$.

6.4 Vecteurs aléatoires à densité *

Exercice 6.10 (*). Soit (X,Y) un vecteur aléatoire bidimensionnel à densité. On suppose qu'il existe deux fonctions intégrables $a : \mathbb{R} \to \mathbb{R}_+$ et $b : \mathbb{R} \to \mathbb{R}_+$ telles que l'on ait

$$f_{X,Y}(x,y) = a(x)\,b(y), \qquad \forall\,(x,y) \in \mathbb{R}^2.$$

Montrer que X et Y sont des variables aléatoires indépendantes.

Solution 6.10. Par hypothèse les fonctions a et b sont positives et intégrales, donc

$$A := \int_{\mathbb{R}} a(x)\,\mathrm{d}x \in [0, \infty[, \qquad B := \int_{\mathbb{R}} b(y)\,\mathrm{d}y \in [0, \infty[.$$

De plus, en appliquant le théorème de Fubini–Tonelli, le fait que l'intégrale de la densité $f_{X,Y}$ sur $\mathbb{R}^2$ vaille 1 conduit au fait que

$$1 = \int_{\mathbb{R}^2} f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}^2} a(x)\,b(y)\,\mathrm{d}x\,\mathrm{d}y = A\,B,$$

dont on déduit en particulier que $A > 0$, $B > 0$ et $B = A^{-1}$. En appliquant la Proposition 6.49, et plus précisément la formule (6.65), on sait que X et Y sont des variables aléatoires réelles absolument continues, de densités respectives

$$f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x,y)\,\mathrm{d}y = \int_{\mathbb{R}} a(x)\,b(y)\,\mathrm{d}y = a(x)\left(\int_{\mathbb{R}} b(y)\,\mathrm{d}y\right) = B\,a(x)\,,$$
$$f_Y(x) = \int_{\mathbb{R}} f_{X,Y}(x,y)\,\mathrm{d}x = \int_{\mathbb{R}} a(x)\,b(y)\,\mathrm{d}x = b(y)\left(\int_{\mathbb{R}} a(x)\,\mathrm{d}x\right) = A\,b(y)\,.$$

Vu que $AB = 1$, il s'ensuit facilement que $f_X(x)f_Y(y) = f_{X,Y}(x,y)$ pour tout $(x,y) \in \mathbb{R}^2$. En se rappelant de la Proposition 6.50, cela démontre l'indépendance de X et Y.

Exercice 6.11 (*). Soit (X,Y) un vecteur aléatoire à densité bidimensionnel. On suppose qu'il existe un point $(x_0,y_0) \in \mathbb{R}^2$ tel que

$$f_{X,Y}(x_0,y_0) \neq f_X(x_0)\,f_Y(y_0)\,.$$

Si les fonctions $f_{X,Y}$, f_X et f_Y sont continues respectivement en (x_0,y_0), x_0 et y_0, montrer que les variables aléatoires X et Y *ne sont pas* indépendantes.

Solution 6.11. Notons $B((x_0,y_0),r)$ la boule ouverte de rayon r centrée au point (x_0,y_0) :

$$B((x_0,y_0),r) := \{(x,y) \in \mathbb{R}^2 :\ (x-x_0)^2 + (y-y_0)^2 < r^2\}\,,$$

Il suffit de montrer qu'il existe un $r > 0$ tel que

$$f_{X,Y}(x,y) \neq f_X(x)\,f_Y(y) \qquad \text{pour tout } (x,y) \in B\big((x_0,y_0),r\big)\,. \tag{S6.7}$$

En effet, comme la boule $B((x_0,y_0),r)$ est de mesure 2-dimensionnelle strictement positive (comme tout sous-ensemble ouvert de $\mathbb{R}^2$ non vide), il découle de la Proposition 6.50 que X et Y ne sont pas indépendantes.

Si la relation (S6.7) n'était pas vérifiée, pour tout $n \in \mathbb{N}^*$ il devrait exister un point $(x_n,y_n) \in B((x_0,y_0),\frac{1}{n})$ tel que

$$f_{X,Y}(x_n,y_n) = f_X(x_n)\,f_Y(y_n)\,, \qquad \forall n \in \mathbb{N}^*\,. \tag{S6.8}$$

Remarquons que, par construction, $\lim_{n\to\infty}(x_n,y_n) = (x_0,y_0)$ dans $\mathbb{R}^2$ et donc $\lim_{n\to\infty} x_n = x_0$ et $\lim_{n\to\infty} y_n = y_0$ dans $\mathbb{R}$. Par continuité de f_X en x_0, de f_Y en y_0 et de $f_{X,Y}$ en (x_0,y_0), on peut passer à la limite $n \to \infty$ dans (S6.8) et obtenir $f_{X,Y}(x_0,y_0) = f_X(x_0)\,f_Y(y_0)$, ce qui aboutit à une contradiction de l'hypothèse de l'énoncé.

Exercice 6.12 (*). Soient $]a,b[$, $]c,d[$ deux intervalles de $\mathbb{R}$. Montrer que le vecteur aléatoire (X,Y) est de loi uniforme sur le rectangle $]a,b[\times]c,d[$, dans le sens de l'Exemple 6.48, si et seulement si $X \sim \mathrm{U}(a,b)$, $Y \sim \mathrm{U}(c,d)$ et X et Y sont indépendantes.

Solution 6.12. Si X et Y sont des variables aléatoires indépendantes avec $X \sim \mathrm{U}(a,b)$, $Y \sim \mathrm{U}(c,d)$, alors grâce à la Proposition 6.50 leur densité jointe est donnée par

$$\begin{aligned} f_{X,Y}(x,y) = f_X(x)f_Y(y) &= \frac{1}{b-a}\mathbb{1}_{]a,b[}(x) \times \frac{1}{d-c}\mathbb{1}_{]c,d[}(y) \\ &= \frac{1}{(b-a)(d-c)}\mathbb{1}_{]a,b[\times]c,d[}(x,y)\,. \end{aligned}$$

On a utilisé ici que $(x,y) \in]a,b[\times]c,d[$ si et seulement si $x \in]a,b[$ et $y \in]c,d[$.

Réciproquement, si (X,Y) est de loi uniforme continue dans le rectangle $]a,b[\times]c,d[$, sa densité jointe est donnée par

$$f_{X,Y}(x,y) = \frac{1}{(b-a)(d-c)}\mathbb{1}_{]a,b[\,\times\,]c,d[}(x,y) = \frac{1}{b-a}\mathbb{1}_{(a,b)}(x)\,\frac{1}{d-c}\mathbb{1}_{(c,d)}(y)\,.$$

Ainsi (X,Y) possède la même densité jointe de $(\tilde{X},\tilde{Y})$, où $\tilde{X}\sim \mathrm{U}(a,b)$ et $\tilde{Y}\sim \mathrm{U}(c,d)$ sont indépendantes. Comme les lois marginales et l'indépendance sont des propriétés qui dépendent uniquement de la loi jointe, on en déduit que X et Y sont indépendantes, avec $X\sim \mathrm{U}(a,b)$ et $Y\sim \mathrm{U}(c,d)$.

6.5 Vecteurs aléatoires gaussiens *

Exercice 6.13. Soit $Z\sim \mathrm{N}(0,I)$ un vecteur gaussien de dimension n. Soient A et B des matrices de dimensions respectives $m\times n$ et $\ell\times n$. On pose $X := AZ$ et $Y := BZ$. Montrer que si $AB^T = 0$ (la matrice nulle), alors X et Y sont indépendants.

Solution 6.13. Posons $W = Q\begin{pmatrix} Z\\ Z\end{pmatrix}$, avec Q la matrice par blocs $Q = \begin{pmatrix} A & 0\\ 0 & B\end{pmatrix}$. Autrement dit, les n premières composantes de W sont $X_1,\dots,X_n$ et les n composantes suivantes (et les dernières) sont $Y_1,\dots,Y_n$.

Le vecteur $\begin{pmatrix} Z\\ Z\end{pmatrix}$ est un vecteur aléatoire gaussien (on note par exemple que toute combinaison linéaire de ses composantes est une combinaison linéaire des composantes de Z, donc est une variable aléatoire (réelle) normale), de matrice de covariance par blocs $V = \begin{pmatrix} I & I\\ I & I\end{pmatrix}$. Grâce à la Proposition 6.71, on en déduit que $W = QZ$ est un vecteur aléatoire normal de matrice de covariance (par blocs)

$$QVQ^T = \begin{pmatrix} A & 0\\ 0 & B\end{pmatrix}\begin{pmatrix} I & I\\ I & I\end{pmatrix}\begin{pmatrix} A^T & 0\\ 0 & B^T\end{pmatrix} = \begin{pmatrix} AA^T & AB^T\\ BA^T & BB^T\end{pmatrix} = \begin{pmatrix} AA^T & 0\\ 0 & BB^T\end{pmatrix},$$

où pour la dernière égalité on a utilisé que $AB^T = 0$ (la matrice nulle $n\times n$) et que $BA^T = (AB^T)^T = 0$ (la matrice nulle $n\times n$).

On peut appliquer la Proposition 6.73 au vecteur gaussien $W = (X_1,\dots,X_n,Y_1,\dots,Y_n)$. Comme la matrice de covariance de W est diagonale par blocs, on obtient que $\mathrm{Cov}(X_i,Y_j) = 0$ pour tous $1\le i,j\le n$, on en conclut que les vecteurs aléatoires X et Y sont indépendants (et gaussiens).

Exercice 6.14. Soit $X = \begin{pmatrix} X_1\\ X_2\\ X_3\end{pmatrix}\sim \mathrm{N}(0,V)$ un vecteur gaussien, avec $V = \begin{pmatrix} 2 & -1 & 0\\ -1 & 2 & 0\\ 0 & 0 & 1\end{pmatrix}$.

(i) Déterminer la loi du vecteur (X_1-X_2, X_2+X_3).

(ii) Le vecteur X est-il à densité ? Si oui, donner son expression.

(iii) Trouver une matrice A orthogonale telle que les composantes du vecteur AX soient indépendantes.

Solution 6.14. (i) Le vecteur $Y := \begin{pmatrix} X_1-X_2\\ X_2+X_3\end{pmatrix}$ est une transformation linéaire du vecteur X : on a en effet $Y = BX$, avec $B = \begin{pmatrix} 1 & -1 & 0\\ 0 & 1 & 1\end{pmatrix}$. D'après la Proposition 6.71, le vecteur aléatoire Y est gaussien et on peut calculer toutes les covariances :

$$\begin{aligned}
\mathrm{Var}(X_1 - X_2) &= \mathrm{Cov}(X_1 - X_2, X_1 - X_2) = \mathrm{Var}(X_1) - 2\,\mathrm{Cov}(X_1, X_2) + \mathrm{Var}(X_2) \\
&= 2 - 2 \times (-1) + 2 = 6\,, \\
\mathrm{Var}(X_2 + X_3) &= \mathrm{Cov}(X_2 + X_3, X_2 + X_3) = \mathrm{Var}(X_2) + 2\,\mathrm{Cov}(X_2, X_3) + \mathrm{Var}(X_3) \\
&= 2 + 2 \times 0 + 1 = 3\,, \\
\mathrm{Cov}(X_1 - X_2, X_2 + X_3) &= \mathrm{Cov}(X_1, X_2) + \mathrm{Cov}(X_1, X_3) - \mathrm{Var}(X_2) - \mathrm{Cov}(X_2, X_3) \\
&= -1 + 0 - 2 - 0 = -3\,.
\end{aligned}$$

De façon alternative, on peut calculer la matrice de covariance

$$AVA^T = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 6 & -3 \\ -3 & 3 \end{pmatrix}$$

(ii) Le vecteur aléatoire X est gaussien $\mathrm{N}(0, V)$ de matrice de covariance V inversible, parce que $\det V = 3 \neq 0$. D'après le Théorème 6.69, X est absolument continue, de densité

$$f_X(x) = \frac{1}{(2\pi)^{3/2}\sqrt{3}} \exp\left[-\frac{1}{2} x^T V^{-1} x\right].$$

Notons que $V^{-1} = \frac{1}{3}\begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$, donc $\frac{1}{2} x^T V^{-1} x = \frac{1}{6}(2x_1^2 - 2x_1x_2 + 2x_2^2 + x_3^2)$.

(iii) D'après la Proposition 6.71, on sait que AX est un vecteur aléatoire gaussien de matrice de covariance AVA^T. Maintenant, grâce à la Proposition 6.73, les composantes du vecteur AX seront indépendantes si la matrice de covariance AVA^T est diagonale : il suffit donc de diagonaliser la matrice V.

On trouve que les vecteurs propres de V sont $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

En normalisant les vecteurs pour avoir une matrice orthogonale, on peut donc prendre $A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (observons que $A = A^T = A^{-1}$).

Exercice 6.15. Soit $X \sim \mathrm{N}(0, I)$ un vecteur gaussien de dimension $2n$. Pour $1 \leq k \leq 2n$, on pose $Y_k := (X_1 + \cdots + X_k) - (X_{k+1} + \cdots + X_{2n})$. Les variables aléatoires $Y_1, \ldots, Y_{2n}$ sont-elles indépendantes ? Existe-t-il des indices $k, k' \in \{1, \ldots, 2n\}$ pour lesquels les variables aléatoires Y_k et $Y_{k'}$ sont indépendantes ?

Solution 6.15. On peut écrire $Y = AX$, avec

$$A = \begin{pmatrix} 1 & -1 & -1 & \cdots & -1 \\ 1 & 1 & -1 & \cdots & -1 \\ \vdots & & & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & -1 \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix}$$

D'après la Proposition 6.71, le vecteur aléatoire Y est gaussien. Pour savoir si les variables aléatoires $Y_1, \ldots, Y_{2n}$ sont indépendantes, calculons les covariances $\mathrm{Cov}(Y_k, Y_{k'})$ pour $k \neq k'$ (on prendra $k' > k$ pour simplifier).

Rappelons que $X_1,\dots,X_{2n}$ sont indépendantes grâce à la Proposition 6.73. Pour $a \le b$ entiers, posons $S_{a,b} = \sum_{i=a}^{b} X_i$; on convient que $S_a^b = 0$ si $a > b$. Avec ces notations, on peut écrire, pour $1 \le k < k' \le 2n$

$$\begin{aligned}\mathrm{Cov}(Y_k, Y_{k'}) &= \mathrm{Cov}\left(S_{1,k} + S_{k+1,k'} - S_{k'+1,2n}, S_{1,k} - S_{k+1,k'} - S_{k'+1,2n}\right)\\ &= \mathrm{Var}(S_{1,k}) - \mathrm{Var}(S_{k+1,k'}) + \mathrm{Var}(S_{k'+1,2n}),\end{aligned}$$

où on a utilisé la bilinéarité de la covariance et le fait que $S_{1,k}, S_{k+1,k'}, S_{k'+1,2n}$ sont indépendantes (grâce à l'indépendance par paquets et par transformations, voir le Corollaire 3.40) pour obtenir que $\mathrm{Cov}(S_{1,k}, S_{k+1,k'}) = \mathrm{Cov}(S_{1,k}, S_{k'+1,2n}) = \mathrm{Cov}(S_{k+1,k'}, S_{k'+1,2n}) = 0$. En utilisant le fait que $\mathrm{Var}(S_{a,b}) = \sum_{i=a}^{b} \mathrm{Var}(X_i)$ par indépendance des X_i, on obtient $\mathrm{Var}(S_{a,b}) = b - a + 1$ pour tous $a \le b$ entiers (la variance vaut 0 si $a + 1 = b$) : on en conclut que

$$\mathrm{Cov}(Y_k, Y_{k'}) = k - (k' - k) + 2n - k' = 2(n + k - k').$$

Ainsi, d'après la Proposition 6.73, Y_k et $Y_{k'}$ sont indépendantes si et seulement si $k' = k + n$.

On a donc montré que pour $n = 1$, Y_1 et Y_2 sont indépendantes ; pour $n \ge 2$, $Y_1,\dots,Y_{2n}$ ne sont pas indépendantes parce qu'on a par exemple $\mathrm{Cov}(Y_1, Y_2) = 2(n-1) > 0$, mais pour $1 \le i \le n$ les variables aléatoires Y_i et Y_{i+n} sont indépendantes.

Exercice 6.16. Soit $X \sim \mathrm{N}(\mathbf{m}, V)$ un vecteur gaussien de dimension $n \ge 2$, où $\mathbf{m} = (m,\dots,m)^T$ avec $m \in \mathbb{R}$, et où la matrice de covariance V est donnée par $V_{ii} = 1$ et $V_{ij} = \rho < 1$ pour $i \neq j$. On pose $\overline{X} := \frac{1}{n}\sum_{i=1}^{n} X_i$ et $W_n := \sum_{i=1}^{n}(X_i - \overline{X})^2$. Montrer que $\overline{X}_n$ et W_n sont indépendantes, et que $\frac{1}{1-\rho} W_n \sim \chi^2(n-1)$.

[*Sugg.* Considérer la matrice

$$A := \begin{pmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \\ \frac{1}{\sqrt{2\cdot 1}} & -\frac{1}{\sqrt{2\cdot 1}} & 0 & 0 & \cdots & 0 \\ \frac{1}{\sqrt{3\cdot 2}} & \frac{1}{\sqrt{3\cdot 2}} & \frac{-2}{\sqrt{3\cdot 2}} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \cdots & \frac{-(n-1)}{\sqrt{n(n-1)}} \end{pmatrix}.$$

Montrer que A est orthogonale, puis considérer $Y = AX$.]

Solution 6.16. Prenons la matrice A introduite dans la suggestion et notons x_k sa k-ème colonne :

$$x_k = \begin{pmatrix} \frac{1}{\sqrt{n}} \\ 0 \\ \vdots \\ 0 \\ \frac{-(k-1)}{\sqrt{k(k-1)}} \\ \frac{1}{\sqrt{(k+1)k}} \\ \vdots \\ \frac{1}{\sqrt{n(n-1)}} \end{pmatrix}.$$

Montrons maintenant que A est orthogonale, c'est-à-dire $x_k^T x_\ell = 0$ pour tous $1 \le k < \ell \le n$:

$$x_k^T x_\ell = \frac{1}{n} + 0 + \cdots + 0 - \frac{\ell-1}{\ell(\ell-1)} + \frac{1}{(\ell+1)\ell} + \cdots + \frac{1}{n(n-1)}$$
$$= \frac{1}{n} - \frac{1}{\ell} + \Big(\frac{1}{\ell} - \frac{1}{\ell+1}\Big) + \cdots + \Big(\frac{1}{n-1} - \frac{1}{n}\Big) = 0\,,$$

où on a utilisé que $\frac{1}{i(i-1)} = \frac{1}{i-1} - \frac{1}{i}$ pour tout $i \geq 2$, puis une somme télescopique. On a donc montré que A est orthogonale. De manière similaire, on a aussi

$$x_k^T x_k = \frac{1}{n} + 0 + \cdots + 0 + \frac{(k-1)^2}{k(k-1)} + \frac{1}{(k+1)k} + \cdots + \frac{1}{n(n-1)}$$
$$= \frac{1}{n} + \frac{k-1}{k} + \Big(\frac{1}{k} - \frac{1}{k+1}\Big) + \cdots + \Big(\frac{1}{n-1} - \frac{1}{n}\Big) = 1\,.$$

Ainsi $A^T A = I$, donc A est orthonormale.

Considérons maintenant le vecteur aléatoire $Y = AX$: d'après la Proposition 6.71, Y est un vecteur aléatoire gaussien, de matrice de covariance $\Sigma := AVA^T$. Notons que la matrice V peut s'écrire sous la forme $V = \rho J + (1-\rho)I$, où $J = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}$ et I est la matrice identité. On obtient que la matrice de covariance de Y est $\Sigma := \rho AJA^T + (1-\rho)I$, étant donné que $AA^T = I$. Maintenant, remarquons que

$$AJA^T = \begin{pmatrix} \sqrt{n} & \cdots & \sqrt{n} \\ 0 & \cdots & 0 \\ \vdots & \cdots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} A^T = \begin{pmatrix} n & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix}.$$

On en conclut que la matrice de covariance de Y est égale à

$$\Sigma = \begin{pmatrix} 1+\rho(n-1) & 0 & \cdots & 0 \\ 0 & 1-\rho & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 1-\rho \end{pmatrix}.$$

Notons maintenant P la projection orthogonale sur le vecteur u dont les composantes sont $u_i = 1$ pour tout $i \in \{1,\dots,n\}$, que l'on identifie avec sa matrice, égale à $P = \frac{1}{n}J$ avec J la matrice définie plus haut. En particulier, on a $PX = \bar{X}u$, et $(I-P)X = X - \bar{X} = (X_1 - \bar{X},\dots,X_n - \bar{X})$, de sorte que $W_n = \|(I-P)X\|^2$.

Considérons maintenant la matrice

$$APA^{-1} = APA^T = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix},$$

qui correspond à la projection orthogonale sur la première coordonnée. Appliquée à Y, on a

$$APA^{-1}Y = \begin{pmatrix} Y_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad (I - APA^{-1})Y = \begin{pmatrix} 0 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}.$$

Comme la matrice de covariance de Y est diagonale, les deux vecteurs calculés ci-dessus, $APA^{-1}Y$ et $(I - APA^{-1})Y$, sont indépendants (*cf.* Proposition 6.73). En se rappelant que $Y = AX$, on a $APA^{-1}Y = APX = \bar{X}Au$. D'autre part, comme A est orthonormale,

$$W_n = \|(I-P)X\|^2 = \|A(I-P)X\|^2 = \|(I - APA^{-1})Y\|^2 = \sum_{i=2}^{n} Y_i^2 .$$

On en conclut que $\bar{X}$ et W_n sont indépendantes, en tant que transformations des vecteurs aléatoires indépendants $APA^{-1}Y$ et $(I - APA^{-1})Y$. Pour finir, comme $Y_2, \ldots, Y_n$ sont des variables aléatoires indépendantes de loi $\mathrm{N}(0, 1-\rho)$, les variables aléatoires $\frac{1}{\sqrt{1-\rho}}Y_1, \ldots, \frac{1}{\sqrt{1-\rho}}Y_n$ sont indépendantes de loi $\mathrm{N}(0,1)$, donc

$$\frac{1}{1-\rho} W_n = \sum_{i=2}^{n} \Big(\frac{1}{\sqrt{1-\rho}} Y_i \Big)^2 \sim \chi^2(n-1),$$

comme on l'a vu dans la Section 6.3.5, voir notamment (6.51).

6.6 Exemples et applications

Exercice 6.17 (*). Soient $X \sim \Gamma(\alpha, \lambda)$ et $Y \sim \Gamma(\beta, \lambda)$ deux variables aléatoires indépendantes, avec $\alpha, \beta > 0$, et $\lambda > 0$. Montrer que $\frac{X}{X+Y} \sim \mathrm{Beta}(\alpha, \beta)$, où on rappelle que la densité d'une telle variable aléatoire est donnée en (6.92).

Solution 6.17. Utilisons l'Observation 6.26. Soit $h : \mathbb{R} \to \mathbb{R}$ une fonction quelconque continue par morceaux et bornée et déterminons $\mathrm{E}(h(W))$, où on a posé $W := \frac{X}{X+Y}$.

On peut appliquer la formule de transfert (Proposition 6.56) au vecteur (X,Y) avec la fonction (bornée) $g :]0, \infty[^2 \to \mathbb{R}$ définie par $g(x,y) = h(\frac{x}{x+y})$. Vu que X et Y sont indépendantes, grâce à la Proposition 6.50 le vecteur aléatoire (X,Y) a pour densité $f_{X,Y}(x,y) = f_X(x) f_Y(y)$, c'est-à-dire, en se rappelant de la formule (6.39) pour la densité de variables aléatoires Gamma,

$$f_{X,Y}(x,y) = \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \lambda^{\alpha+\beta} x^{\alpha-1} y^{\beta-1} \mathrm{e}^{-\lambda(x+y)} \mathbb{1}_{]0,+\infty[}(x) \mathbb{1}_{]0,+\infty[}(y).$$

Ainsi, la formule de transfert (6.72) donne

$$\begin{aligned} \mathrm{E}(h(W)) &= \mathrm{E}\Big(h\Big(\frac{X}{X+Y}\Big)\Big) \\ &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \lambda^{\alpha+\beta} \int_{]0,\infty[^2} h\Big(\frac{x}{x+y}\Big) x^{\alpha-1} y^{\beta-1} \mathrm{e}^{-\lambda(x+y)} \mathrm{d}x\mathrm{d}y. \end{aligned}$$

Appliquons maintenant un changement de variable bidimensionnel : $(u,v) = (\frac{x}{x+y}, x+y)$, ou bien $(x,y) = (uv, (1-u)v)$. La fonction $\psi : (u,v) \to (x,y) = (uv, (1-u)v)$ est un difféomorphisme de $U =]0,1[\times]0,\infty[$ vers $V =]0,\infty[^2$, dont la matrice jacobienne est donnée par

$$J_\psi(u,v) = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{pmatrix} = \begin{pmatrix} v & -v \\ u & 1-u \end{pmatrix}$$

de déterminant $|\det J_\psi(u,v)| = v(1-u)+uv = v$. Grâce à la formule du changement de variable (6.59), on obtient

$$\begin{aligned} \mathrm{E}(h(W)) &= \frac{1}{\Gamma(\alpha)\Gamma(\beta)}\lambda^{\alpha+\beta}\int_{]0,1[\,\times\,]0,\infty[} h(u)u^{\alpha-1}v^{\alpha-1}(1-u)^{\beta-1}v^{\beta-1}\mathrm{e}^{-\lambda v}v\,\mathrm{d}u\,\mathrm{d}v \\ &= \int_0^1 h(u)\frac{1}{\Gamma(\alpha)\Gamma(\beta)}u^{\alpha-1}(1-u)^{\beta-1}\left(\int_0^\infty \lambda^{\alpha+\beta}v^{\alpha+\beta-1}\mathrm{e}^{-\lambda v}\mathrm{d}v\right)\mathrm{d}u\,, \end{aligned}$$

où on a utilisé le théorème de Fubini–Tonelli pour la dernière égalité. En se rappelant de la définition de la fonction Gamma $\Gamma(\cdot)$, ou bien de la densité (6.39) d'une variable aléatoire Gamma(λ,α), on obtient que l'intégrale intérieure est égale à $\Gamma(\alpha+\beta)$. Pour résumer, on a montré que

$$\mathrm{E}(h(W)) = \int_0^1 h(u)\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}u^{\alpha-1}(1-u)^{\beta-1}\mathrm{d}u\,.$$

Ceci étant valable pour toute fonction h continue par morceaux et bornée, l'Observation 6.26 montre que W est absolument continue, de densité $f_W(u) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}u^{\alpha-1}(1-u)^{\beta-1}\mathbb{1}_{]0,1[}(u)$. Autrement dit, $W = \frac{X}{X+Y}$ est de loi Beta(α,β).

Exercice 6.18. Soit $(N_t)_{t\in\mathbb{R}_+}$ un processus de Poisson d'intensité λ. Pour tout intervalle borné $]a,b] \subseteq [0,\infty[$, on définit la variable aléatoire $N_{]a,b]} := N_b - N_a$. Si $I = \bigcup_{i=1}^n I_i \subseteq [0,\infty[$ est une union finie d'intervalles $I_i =]a_i,b_i]$ *disjoints*, on définit

$$N_I := \sum_{i=1}^n N_{I_i} = \sum_{i=1}^n (N_{b_i} - N_{a_i})\,.$$

On note $|]a,b]| = b-a$ la longueur d'un intervalle et de même on définit $|I| = \sum_{i=1}^n |I_i|$ la longueur totale du sous-ensemble $I = \bigcup_{i=1}^n I_i$.

Pour $k \geq 2$ et $0 = t_0 < t_1 < \cdots < t_k < \infty$ arbitraires, on définit les intervalles de temps $I_1,\ldots,I_k$ en posant $I_j =]t_{j-1},t_j]$. D'après le Théorème 6.80, les variables aléatoires $N_{I_1},\ldots,N_{I_k}$ sont indépendantes, de loi $N_{I_j} \sim \mathrm{Poi}(\lambda|I_j|)$. Soient enfin A et B deux sous-ensembles non vides et disjoints de $\{1,\ldots,k\}$; on pose $I := \bigcup_{i\in A} I_i$ et $J := \bigcup_{j\in B} I_j$.

Montrer que les variables aléatoires N_I et N_J sont indépendantes, de loi $N_I \sim \mathrm{Poi}(\lambda|I|)$ et $N_J \sim \mathrm{Poi}(\lambda|J|)$.

Solution 6.18. Comme les variables aléatoires $N_{I_1},\ldots,N_{I_k}$ sont indépendantes et que les sous-ensembles A et B sont disjoints, il découle de l'indépendance par paquets (Proposition 3.38) que les deux vecteurs aléatoires $X := (N_{I_i})_{i\in A}$ et $Y := (N_{I_j})_{j\in B}$ sont indépendants. Par conséquent, grâce à l'indépendance par transformations (Proposition 3.39), le variables aléatoires $N_I = \sum_{i\in A} N_{I_i} = f(X)$ et $N_J = \sum_{j\in B} N_{I_j} = g(Y)$ sont indépendantes. Enfin, comme $N_{I_i} \sim \mathrm{Poi}(\lambda|I_i|)$ et comme la somme de variables aléatoires de Poisson indépendantes est de Poisson de paramètre donné par la somme des paramètres (Proposition 3.128), on obtient

$$N_I = \sum_{i\in A} N_{I_i} \sim \mathrm{Poi}\left(\lambda\sum_{i\in A}|I_i|\right) = \mathrm{Poi}\big(\lambda|I|\big)\,.$$

Le même raisonnement vaut aussi pour N_J.

6.7 Exercices récapitulatifs

On rappelle que i.i.d. *signifie* indépendant·e·s et identiquement distribué·e·s. *Les exercices signalés par un astérisque requièrent l'usage de vecteurs aléatoires.*

Exercice 6.19. Pour une variable aléatoire $X \sim \mathrm{U}(-\pi/2, \pi/2)$, montrer que $Y := \cos(X)$ est une variable aléatoire à densité et en déterminer la densité.

Solution 6.19. Par hypothèse $f_X(x) = \frac{1}{\pi}\mathbb{1}_{]-\pi/2,\pi/2[}(x)$, donc

$$\mathrm{P}(X \in A) = \int_{A\cap\,]-\frac{\pi}{2},\frac{\pi}{2}[} \frac{1}{\pi}\,\mathrm{d}x = \frac{\left|A\cap\,]-\frac{\pi}{2},\frac{\pi}{2}[\,\right|}{\pi}.$$

Calculons la fonction de répartition de Y. Clairement $F_Y(y) = 0$ si $y < 0$ et $F_Y(y) = 1$ si $y > 1$, vu que $\cos(x) \in [0,1]$ pour $x \in\,]-\frac{\pi}{2},\frac{\pi}{2}[$. Notons que pour $x \in\,]-\frac{\pi}{2},\frac{\pi}{2}[$ et $y \in [0,1]$ on a $\cos(x) \le y$ si et seulement si $x \in\,]-\frac{\pi}{2}, -\arccos(y)] \cup [\arccos(y), \frac{\pi}{2}[$, donc pour tout $y \in [0,1]$ on a

$$\begin{aligned}
F_Y(y) &= \mathrm{P}(Y \le y) = \mathrm{P}(\cos(X) \le y)\\
&= \mathrm{P}\Big(X \in\,]-\tfrac{\pi}{2}, -\arccos(y)] \cup [\arccos(y), \tfrac{\pi}{2}[\,\Big)\\
&= \frac{\left|\,]-\frac{\pi}{2}, -\arccos(y)] \cup [\arccos(y), \frac{\pi}{2}[\,\right|}{\pi} = \frac{2}{\pi}\left(\frac{\pi}{2} - \arccos(y)\right).
\end{aligned}$$

La fonction F_Y étant continue et $\mathcal{C}^1$ par morceaux (sur $]-\infty,0[$, $]0,1[$ et $]0,+\infty[$), il découle de la Proposition 6.17 que Y est absolument continue, de densité

$$f_Y(y) = (F_Y)'(y) = \frac{2}{\pi}\frac{1}{\sqrt{1-y^2}}\mathbb{1}_{]0,1[}(y).$$

Exercice 6.20. Soit X un point choisi uniformément dans l'intervalle $[0,2]$. Quelle est la probabilité que le triangle équilatéral de côté X ait une aire plus grande que 1 ?

Solution 6.20. L'aire du triangle équilatéral de côté X vaut $A := \frac{\sqrt{3}}{4}X^2$, donc

$$\begin{aligned}
\mathrm{P}(A > 1) &= \mathrm{P}\left(X^2 > \frac{4}{\sqrt{3}}\right) = \mathrm{P}\left(X \in \left]-\infty, -\frac{2}{3^{1/4}}\right[\cup \left]\frac{2}{3^{1/4}}, +\infty\right[\right)\\
&= \int_{]-\infty,-\frac{2}{(3)^{1/4}}[\,\cup\,]\frac{2}{3^{1/4}},+\infty[} f_X(x)\,\mathrm{d}x\\
&= \int_{]-\infty,-\frac{2}{3^{1/4}}[\,\cup\,]\frac{2}{3^{1/4}},+\infty[} \frac{1}{2}\mathbb{1}_{]0,2[}(x)\,\mathrm{d}x\\
&= \frac{1}{2}\left(2 - \frac{2}{3^{1/4}}\right) = 1 - 3^{-1/4} \simeq 0.24.
\end{aligned}$$

Exercice 6.21. Soit $X \sim \mathrm{U}(0,1)$ et soit $Y := 4X(1-X)$.

(i) Déterminer la fonction de répartition de Y, en déduire que la variable aléatoire Y est à densité et en calculer la densité.

(ii) Justifier que $\mathrm{Cov}(X,Y)$ est bien définie et la calculer.

Solution 6.21. (i) Clairement $F_Y(y) = 0$ si $y < 0$ alors que $F_Y(y) = 1$ si $y > 1$, donc il suffit de se concentrer sur le cas $0 \leq y \leq 1$. Dans ce cas, l'inégalité $4x(1-x) \leq y$ a comme solutions $x \leq \frac{1}{2}(1-\sqrt{1-y})$ ou bien $x \geq \frac{1}{2}(1+\sqrt{1-y})$, donc

$$F_Y(y) = \mathrm{P}(Y \leq y) = \mathrm{P}\left(X \leq \frac{1}{2}(1-\sqrt{1-y})\right) + \mathrm{P}\left(X \geq \frac{1}{2}(1+\sqrt{1-y})\right).$$

Notons que $\mathrm{P}(X \leq x) = x$ et $\mathrm{P}(X \geq x) = 1-x$ si $x \in [0,1]$. Comme les points $\frac{1}{2}(1 \pm \sqrt{1-y})$ sont dans l'intervalle $[0,1]$ pour $y \in [0,1]$, il s'ensuit que

$$F_Y(y) = 1 - \frac{1}{2}(1+\sqrt{1-y}) + \frac{1}{2}(1-\sqrt{1-y}) = 1-\sqrt{1-y}.$$

Au final,

$$F_Y(y) = \begin{cases} 0 & \text{si } y < 0\,, \\ 1-\sqrt{1-y} & \text{si } y \in [0,1]\,, \\ 1 & \text{si } y > 1\,. \end{cases}$$

La fonction F_Y est continue et $\mathcal{C}^1$ par morceaux, donc la variable aléatoire Y est absolument continue. Sa densité est donnée par

$$f_Y(y) = F_Y'(y) = \frac{1}{2}\frac{1}{\sqrt{1-y}}\,1_{]0,1[}(y)\,.$$

(ii) La covariance de X et Y est bien définie car les variables aléatoires sont toutes les deux dans L^2 (elles sont en effet bornées). D'après les propriétés de l'espérance,

$$\mathrm{E}(Y) = \mathrm{E}(4X(1-X)) = 4\mathrm{E}(X) - 4\mathrm{E}(X^2)\,,$$
$$\mathrm{E}(XY) = \mathrm{E}(X \cdot 4X(1-X)) = 4\mathrm{E}(X^2) - 4\mathrm{E}(X^3)\,,$$

donc

$$\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y) = 4\mathrm{E}(X^2)(1+\mathrm{E}(X)) - 4(\mathrm{E}(X^3)+\mathrm{E}(X)^2)\,.$$

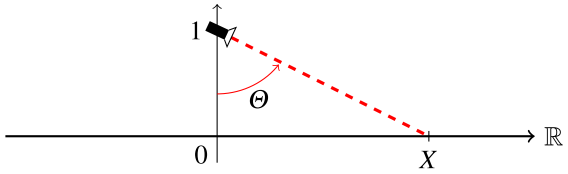
Grâce à la formule de transfert (Proposition 6.22), on calcule

$$\mathrm{E}(X^n) = \int_0^1 x^n\,\mathrm{d}x = \frac{1}{n+1}\,,$$

dont on déduit que $\mathrm{E}(X) = \frac{1}{2}$, $\mathrm{E}(X^2) = \frac{1}{3}$ et $\mathrm{E}(X^3) = \frac{1}{4}$, donc

$$\mathrm{E}(Y) = 4 \cdot \frac{1}{2} - 4 \cdot \frac{1}{3} = \frac{2}{3}\,,$$
$$\mathrm{Cov}(X,Y) = \frac{4}{3}\left(1+\frac{1}{2}\right) - 4\left(\frac{1}{4}+\frac{1}{4}\right) = 0\,.$$

Exercice 6.22. On suspend un laser à 1m au dessus du sol. L'angle qu'il forme avec la verticale est aléatoire, notée Θ, et suit la loi uniforme sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$. On note X le point marqué au sol par le laser (voir la figure ci-dessous). Donner la densité de X.



Solution 6.22. Le point X est donné par $X = \tan(\Theta)$. On peut donc facilement calculer la fonction de répartition de X :

$$F_X(x) = \mathrm{P}(X \le x) = \mathrm{P}(\tan\Theta \le x) = \mathrm{P}(\Theta \le \arctan(x)) = \frac{1}{\pi}\Big(\frac{\pi}{2} + \arctan(x)\Big).$$

Pour la dernière égalité, on a utilisé le fait que Θ est une variable aléatoire uniforme sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$, de sorte que pour $\theta \in]-\frac{\pi}{2}, \frac{\pi}{2}[$ on a $\mathrm{P}(\Theta \le \theta) = \frac{1}{\pi}(\frac{\pi}{2} + \theta)$, voir (6.29).

La fonction F_X est continue et $\mathcal{C}^1$ par morceaux, donc grâce à la Proposition 6.17, la variable aléatoire X a pour densité

$$f_X(x) = F_X'(x) = \frac{1}{\pi}\frac{1}{1+x^2}.$$

Il s'agit d'une variable aléatoire de Cauchy standard, vue dans la Section 6.3.6.

Exercice 6.23. Soit X une variable aléatoire réelle de densité donnée par

$$f_X(x) = -\log(x^c)\,\mathbb{1}_{]0,1[}(x).$$

(i) Déterminer la valeur de $c \in \mathbb{R}$ pour que f_X soit effectivement une densité, et déterminer la fonction de répartition de X.

(ii) Soit $Y = -\log X$. Montrer que $Y \sim \mathrm{Gamma}(\alpha, \lambda)$, en déterminant α et λ.

Solution 6.23. (i) On doit avoir

$$1 = \int_{\mathbb{R}} f_X(x)\,\mathrm{d}x = -c\int_0^1 \log x\,\mathrm{d}x = -c\big[x\log x - x\big]_0^1 = c,$$

donc on déduit que $c = 1$. Comme $f_X(x) = 0$ pour $x \notin]0,1[$, on a $F_X(x) = 0$ pour $x \le 0$, $F_X(x) = 1$ pour $x > 1$, et pour $0 < x \le 1$:

$$F_X(x) = -\int_0^x \log(t)\,\mathrm{d}t = x - x\log(x).$$

(ii) Notons que $-\log X$ ne prend que des valeurs positives. Pour $t \ge 0$, on a

$$\mathrm{P}(Y \le t) = \mathrm{P}(X \ge \mathrm{e}^{-t}) = 1 - F_X(\mathrm{e}^{-t}) = 1 - \mathrm{e}^{-t} - t\mathrm{e}^{-t}.$$

En dérivant par rapport à t on obtient

$$f_Y(x) = x\mathrm{e}^{-x}\,\mathbb{1}_{]0,+\infty[}(x),$$

c'est-à-dire $Y \sim \mathrm{Gamma}(2,1)$.

Exercice 6.24. Soit $X \sim \mathrm{U}(-1,1)$. Déterminer la fonction de répartition de la variable aléatoire $Y := X^+ = \max(X,0)$. En déduire que la loi de Y n'est ni discrète, ni à densité.

Solution 6.24. On a clairement $Y \ge 0$, donc pour $y < 0$ on a $F_Y(y) = \mathrm{P}(X \le y) = 0$. Pour $y \ge 0$ on a $\max(x,0) \le y$ si et seulement si $x \le y$, donc

$$F_Y(y) = F_X(y) = \int_{-\infty}^{y} \frac{1}{2} \mathbb{1}_{]-1,1[}(x)\,\mathrm{d}x = \frac{1}{2}\int_{-1}^{\min(y,1)} 1\,\mathrm{d}x = \frac{\min(y,1)+1}{2}.$$

Au final

$$F_Y(y) = \begin{cases} 0 & \text{si } y < 0\,, \\ \frac{y+1}{2} & \text{si } 0 \le y \le 1\,, \\ 1 & \text{si } y > 1\,. \end{cases}$$

Remarquons que $F_Y(0^+) = \lim_{y\downarrow 0} F_Y(y) = F_Y(0) = \frac{1}{2} \neq F_Y(0^-) = \lim_{y\uparrow 0} F_Y(y) = 0$: la fonction $F_Y(y)$ n'est donc pas continue au point $y = 0$. Par conséquent la variable aléatoire Y n'est pas à densité. De plus, la variable aléatoire Y ne peut pas non plus être discrète : si elle l'était, d'après la relation (5.7) on devrait avoir $1 = \sum_{y\in\mathbb{R}} \mathrm{P}(Y=y) = \sum_{y\in\mathbb{R}}(F_Y(y) - F_Y(y^-))$; mais $F_Y(y) = F_Y(y^-)$ pour tout $y \neq 0$ donc la dernière somme vaut $F_Y(0) - F_Y(0^-) = \frac{1}{2}$.

Exercice 6.25. Soit X une variable aléatoire réelle de loi $\mathrm{Exp}(\lambda)$, avec $\lambda > 0$. On pose $Y = \lfloor X \rfloor$, où $\lfloor x \rfloor$ désigne la partie entière de x. Notons que Y est une variable aléatoire à valeurs dans $\mathbb{N}$, donc discrète. Montrer que $Y \sim \text{Géom}_0(p)$, pour un paramètre p que l'on déterminera.

Solution 6.25. La variable aléatoire Y étant à valeurs dans $\mathbb{N}$, il suffit simplement de déterminer sa densité discrète. Pour $k \in \mathbb{N}$, en utilisant l'égalité d'événements $\{Y = k\} = \{\lfloor X \rfloor = k\} = \{X \in [k, k+1[\,\}$, l'identité (6.17) et la densité de X (voir la Section 6.3.3), on a

$$\begin{aligned} \mathrm{P}(Y=k) &= \mathrm{P}\big(X \in [k,k+1[\,\big) = \int_k^{k+1} \lambda \mathrm{e}^{-\lambda x}\mathrm{d}x \\ &= \Big[-\mathrm{e}^{-\lambda x}\Big]_k^{k+1} = \mathrm{e}^{-\lambda k}(1-\mathrm{e}^{-\lambda})\,. \end{aligned}$$

On a donc montré que $\mathrm{P}(Y = k) = (1-p)^k p$ pour tout $k \in \mathbb{N}$, avec $p = 1 - \mathrm{e}^{-\lambda}$, c'est-à-dire que $Y \sim \text{Géom}_0(p)$.

Exercice 6.26. Soit X une variable aléatoire de densité donnée par

$$f_X(x) := 3x^2\,\mathbb{1}_{[0,1]}(x)\,.$$

On considère les variables aléatoires $A := -\log X$ et $B := \min\{1, A\}$.

(i) Calculer la fonction de répartition de la variable aléatoire A et en déterminer la loi (connue).

(ii) On considère l'équation du second degré suivante, d'inconnue x, dont les coefficients sont déterminés par la variable aléatoire A :

$$x^2 + 3Ax + 2A^2 + 4 = 0\,.$$

Quelle est la probabilité que l'équation n'admette aucune solution réelle ?

(iii) Calculer la fonction de répartition de B et en déduire que B n'est pas une variable aléatoire à densité. Est-elle une variable aléatoire discrète ?

Solution 6.26. (i) Pour $t \ge 0$ on a

$$F_A(t) = \mathrm{P}(A \le t) = \mathrm{P}(X \ge \mathrm{e}^{-t}) = \int_{\mathrm{e}^{-t}}^1 3x^2\,\mathrm{d}x = 1 - \mathrm{e}^{-3t}\,,$$

alors que pour $t < 0$ on a $F_A(t) = 0$ parce que $X \in [0,1]$ et donc $A = -\log X \geq 0$. Comme F_A est continue et $\mathcal{C}^1$ par morceaux, la variable aléatoire A est absolument continue, de densité

$$f_A(t) = F_A'(t) = 3\,\mathrm{e}^{-3t}\,\mathbb{1}_{\mathbb{R}_+}(t)\,,$$

c'est-à-dire $A \sim \mathrm{Exp}(3)$.

(ii) L'équation de degré deux $ax^2 + bx + c = 0$ n'a pas de solution réelle x si et seulement si le discriminant $\Delta := b^2 - 4ac$ est strictement négatif, c'est-à-dire

$$(3A)^2 - 4(2A^2 + 4) = 9A^2 - 8A^2 - 16 = A^2 - 16 < 0\,,$$

et la probabilité de cet événement vaut

$$\mathrm{P}(A^2 - 16 < 0) = \mathrm{P}(A^2 < 16) = \mathrm{P}(A < 4) = 1 - \mathrm{e}^{-12}\,,$$

où la deuxième égalité est valable car $A \geq 0$.

(iii) Pour $t < 1$ on a $F_B(t) = \mathrm{P}(B \leq t) = \mathrm{P}(A \leq t) = F_A(t)$, alors que pour $t \geq 1$ on a $F_B(t) = \mathrm{P}(B \leq t) = 1$ parce que $B = \min\{1, A\} \leq 1$. Ainsi

$$F_B(t) = \begin{cases} 0 & \text{si } t \leq 0 \\ 1 - \mathrm{e}^{-3t} & \text{si } 0 < t < 1 \\ 1 & \text{si } t \geq 1 \end{cases}.$$

Comme on a $F_B(1^-) = 1 - \mathrm{e}^{-3}$ et $F_B(1) = 1$, on obtient que $\mathrm{P}(B = 1) = F_B(1) - F_B(1^-) = \mathrm{e}^{-3} > 0$, donc B n'est pas absolument continue. D'autre part, B n'est pas discrète, parce que dans ce cas on devrait avoir $1 = \sum_{t\in\mathbb{R}} \mathrm{P}(B = t) = \sum_{t\in\mathbb{R}}(F_B(t) - F_B(t^-))$ d'après la relation (5.7). Mais $F_B(t) - F_B(t^-) = 0$ pour tout $t \neq 1$ et on a donc $\sum_{t\in\mathbb{R}}(F_B(t) - F_B(t^-)) = F_B(1) - F_B(1^-) = \mathrm{e}^{-3} < 1$.

Exercice 6.27. Soit X une variable aléatoire de loi $\mathrm{U}(0,1)$ et soit

$$Z := \frac{1}{X^r}\,,$$

où $r \in \mathbb{R}$ est un paramètre fixé

(i) Pour quelles valeurs de $p \in \,]0,\infty[$ a-t-on $\mathrm{E}(|Z|^p) < \infty$, c'est-à-dire $Z \in L^p$?

(ii) Déterminer la loi de la variable aléatoire Z.

[*Sugg.* Calculer la fonction de répartition de Z ; séparer les cas $r < 0$, $r = 0$ et $r > 0$.]

Solution 6.27. (i) Par un calcul facile, en appliquant la formule de transfert (Proposition 6.22), on a

$$\mathrm{E}(|Z|^p) = \int_0^1 x^{-rp}\,\mathrm{d}x = \begin{cases} \frac{1}{1-rp} < \infty & \text{si } rp < 1\,, \\ +\infty & \text{si } rp \geq 1\,, \end{cases}$$

car une primitive de x^{-rp} est $\frac{1}{1-rp}x^{1-rp}$ si $rp \neq 1$ et $\log x$ si $rp = 1$. On a donc $Z \in L^p$ si et seulement si $rp < 1$. Ainsi, si $r \leq 0$, on a $Z \in L^p$ pour tout $p \in \,]0,\infty[$, alors que si $r > 0$ on a $Z \in L^p$ pour tout $p \in \,]0, \frac{1}{r}[$ et $Z \notin L^p$ si $p \geq \frac{1}{r}$.

(ii) Si $r = 0$ on a clairement $Z \equiv 1$, c'est-à-dire Z est presque sûrement constante, donc Z est une variable aléatoire discrète, de densité discrète $\mathrm{p}_Z(x) = \mathbb{1}_{\{1\}}(x)$.

Soit maintenant $r \neq 0$. Calculons la fonction de répartition F_Z de Z. Clairement $F_Z(x) = 0$ pour $x \leq 0$, parce que $Z > 0$. Pour $x > 0$ on a

$$F_Z(x) = \mathrm{P}(Z \leq x) = \mathrm{P}(X^{-r} \leq x)\,.$$

Si $r > 0$ on a $X^{-r} \leq x$ si et seulement si $X \geq x^{-1/r}$, donc

$$F_Z(x) = \mathrm{P}(X \geq \tfrac{1}{x^{1/r}}) = \begin{cases} 0 & \text{si } x \leq 1\,, \\ 1 - \frac{1}{x^{1/r}} & \text{si } x > 1\,. \end{cases}$$

Ainsi F_Z est continue et $\mathcal{C}^1$ par morceaux, donc Z est à densité, de densité

$$f_Z(x) = F_Z'(x) = \frac{1}{r}\frac{1}{x^{1+1/r}}\mathbb{1}_{]1,\infty[}(x)\,.$$

Si en revanche $r < 0$ on a $X^{-r} \leq x$ si et seulement si $X \leq x^{-1/r} = x^{1/|r|}$, donc

$$F_Z(x) = \mathrm{P}(X \leq x^{1/|r|}) = \begin{cases} 0 & \text{si } x \leq 0\,, \\ x^{1/|r|} & \text{si } 0 \leq x \leq 1\,, \\ 1 & \text{si } x \geq 1\,. \end{cases}$$

Ainsi F_Z est continue et $\mathcal{C}^1$ par morceaux, donc Z est à densité, de densité

$$f_Z(x) = F_Z'(x) = \frac{1}{|r|}x^{\frac{1}{|r|}-1}\mathbb{1}_{]0,1[}(x)\,.$$

Exercice 6.28. Soient X et Y des variables aléatoires indépendantes, uniformément distribuée dans l'intervalle $]-1,+1[$. Montrer que $Z := X + Y$ a pour densité

$$f_Z(z) = \frac{1}{4}(2 - |z|)\,\mathbb{1}_{]-2,+2[}(z)\,.$$

Solution 6.28. La variable aléatoire Z est absolument car c'est la somme de variables aléatoires indépendantes à densité, voir la Proposition 6.28, et sa densité est donnée par la convolution

$$\begin{aligned} f_Z(z) &= (f_X * f_Y)(z) = \int_{\mathbb{R}} f_X(x)\, f_Y(z-x)\,\mathrm{d}x \\ &= \frac{1}{4}\int_{\mathbb{R}} \mathbb{1}_{]-1,+1[}(x)\,\mathbb{1}_{]-1,+1[}(z-x)\,\mathrm{d}x = \frac{1}{4}\int_{\mathbb{R}} \mathbb{1}_{]-1,+1[\,\cap\,]z-1,z+1[}(x)\,\mathrm{d}x\,. \end{aligned}$$

Notons que

$$]-1,+1[\,\cap\,]z-1,z+1[= \begin{cases}]z-1,1[& \text{si } z \in [0,2[\,, \\ ]-1,z+1[& \text{si } z \in]-2,0[\,, \\ \emptyset & \text{si } z \notin]-2,+2[\,. \end{cases}$$

Ainsi $f_Z(z) = 0$ si $z \notin]-2,+2[$, alors que pour tout $z \in]-2,+2[$ la longueur de l'intervalle $]-1,+1[\,\cap\,]z-1,z+1[$ est égale à $2 - |z|$ et donc $f_Z(z) = \frac{1}{4}(2 - |z|)$.

Exercice 6.29. Soient X, Y des variables aléatoires indépendantes, toutes les deux de loi $\mathrm{Exp}(1)$, et soit $Z := X - Y$.

(i) Montrer que Z est une variable aléatoire à densité, de densité donnée par

$$f_Z(z) = \frac{1}{2}\mathrm{e}^{-|z|}\,.$$

[*Sugg.* Poser $Y' := -Y$, de sorte que $Z = X + Y'$.]

(ii) Montrer que $|Z| \sim \mathrm{Exp}(1)$.

Solution 6.29. (i) Notons que la variable aléatoire $W := -Y$ est absolument continue de densité $f_W(w) = f_Y(-w)$ et est indépendante de X (pourquoi ?). Comme $Z = X - Y = X + W$, on peut appliquer la Proposition 6.28 : comme $f_X(x) = f_Y(x) = \mathrm{e}^{-x}\,\mathbb{1}_{]0,\infty[}(x)$, pour $z \geq 0$ on a

$$f_Z(z) = \int_{\mathbb{R}} f_X(x) f_Y(x-z)\,\mathrm{d}x = \int_z^{+\infty} \mathrm{e}^{-x}\mathrm{e}^{-(x-z)}\mathrm{d}x = \frac{1}{2}\mathrm{e}^{-z}.$$

De manière analogue, pour $z \leq 0$ on a $f_Z(z) = \frac{1}{2}\mathrm{e}^{z}$. Au final, $f_Z(z) = \frac{1}{2}\mathrm{e}^{-|z|}$.

(ii) La fonction de répartition de $F_{|Z|}(t)$ est clairement nulle pour $t < 0$, alors que pour $t \geq 0$

$$F_{|Z|}(t) = \mathrm{P}(|Z| \leq t) = \mathrm{P}(-t \leq Z \leq t) = \int_{-t}^{t} f_Z(z)\,\mathrm{d}z = 2\int_0^t \frac{\mathrm{e}^{-z}}{2}\,\mathrm{d}z = 1 - \mathrm{e}^{-t}.$$

Comme $F_{|Z|}$ est continue et $\mathcal{C}^1$ par morceaux, on en déduit $|Z|$ est absolument continue, de densité donnée par

$$f_{|Z|}(t) = F'_{|Z|}(t) = \mathrm{e}^{-t}\,\mathbb{1}_{]0,\infty[}(t),$$

c'est-à-dire $|Z| \sim \mathrm{Exp}(1)$.

Exercice 6.30. Rappelons qu'une variable aléatoire réelle X est dite de *Cauchy standard* si elle a pour densité

$$f_X(x) := \frac{1}{\pi}\frac{1}{1+x^2}.$$

Montrer que la variable aléatoire $Y := 1/X$ est de Cauchy standard.

Solution 6.30. On note (ou on se rappelle du calcul (6.53)) que la fonction de répartition de X est donnée par

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} f_X(t)\,\mathrm{d}t = \int_{-\infty}^{x} \frac{1}{\pi(1+t^2)}\,\mathrm{d}t = \left[\frac{1}{\pi}\arctan(t)\right]_{t=-\infty}^{t=x} \\ &= \frac{1}{\pi}\left(\arctan(x) + \frac{\pi}{2}\right) = \frac{1}{\pi}\arctan(x) + \frac{1}{2}. \end{aligned} \tag{S6.9}$$

Calculons la fonction de répartition de Y. Pour $y < 0$ on a $\frac{1}{x} \leq y$ si et seulement si $\frac{1}{y} \leq x < 0$, donc

$$\begin{aligned} F_Y(y) &= \mathrm{P}(Y \leq y) = \mathrm{P}\left(\frac{1}{X} \leq y\right) = \mathrm{P}\left(\frac{1}{y} \leq X < 0\right) = \mathrm{P}(X \leq 0) - \mathrm{P}\left(X < \frac{1}{y}\right) \\ &= \frac{1}{2} - \mathrm{P}\left(X \leq \frac{1}{y}\right) = \frac{1}{2} - F_X\left(\frac{1}{y}\right) = -\frac{1}{\pi}\arctan\left(\frac{1}{y}\right). \end{aligned}$$

Observons que $\arctan(\frac{1}{y}) = -\arctan(-\frac{1}{y})$. En appliquant la formule $\arctan(z) + \arctan(\frac{1}{z}) = \frac{\pi}{2}$, valable pour $z > 0$, avec $z = -1/y$ on obtient donc

$$F_Y(y) = \frac{1}{\pi}\arctan\left(-\frac{1}{y}\right) = \frac{1}{\pi}\left(\frac{\pi}{2} - \arctan(-y)\right) = \frac{1}{2} + \arctan(y) = F_X(y).$$

Considérons maintenant le cas $y > 0$. Dans ce cas on a $\frac{1}{x} \leq y$ si et seulement si $x \leq 0$ ou $x \geq \frac{1}{y}$, donc

$$
\begin{aligned}
F_Y(y) &= \mathrm{P}(Y \le y) = \mathrm{P}\left(\tfrac{1}{X} \le y\right) = \mathrm{P}\left(\{X \le 0\} \cup \{X \ge \tfrac{1}{y}\}\right) \\
&= \mathrm{P}(X \le 0) + \mathrm{P}\left(X \ge \tfrac{1}{y}\right) = \frac{1}{2} + \left(1 - \mathrm{P}\left(X < \tfrac{1}{y}\right)\right) \\
&= \tfrac{3}{2} - F_X\left(\tfrac{1}{y}\right) = 1 - \tfrac{1}{\pi}\arctan\left(\tfrac{1}{y}\right) = 1 - \tfrac{1}{\pi}\left(\tfrac{\pi}{2} - \arctan(y)\right) \\
&= \tfrac{1}{2} - \arctan(y) = F_X(y)\,.
\end{aligned}
$$

Comme $F_Y(0) = \mathrm{P}(Y \le 0) = \mathrm{P}(X \le 0) = F_X(0)$, on a montré que $F_Y(y) = F_X(y)$ pour tout $y \in \mathbb{R}$. Les variables aléatoires X et Y ayant la même fonction de répartition, elles ont la même loi, donc Y est de Cauchy standard.

Exercice 6.31. Le service d'information d'une compagnie possède deux numéros de contact. Les temps d'attente T_1 et T_2 pour parler avec l'opérateur, pour les deux numéros, sont des variables aléatoires exponentielles, d'espérance $\mu = 15$ minutes. De plus T_1 et T_2 peuvent être considérées indépendantes. Philippe a à disposition deux téléphones et décide d'appeler en même temps les deux numéros, de façon à parler avec l'opérateur qui répondra en premier.

(i) Combien de temps en moyenne Philippe devra-t-il attendre pour parler avec un opérateur ?

(ii) Quelle est la probabilité qu'il attende moins de 5 minutes ?

Solution 6.31. (i) Par hypothèse, aussi bien T_1 que T_2 sont de loi $\mathrm{Exp}(\lambda)$ avec $\lambda = 1/15$ (car l'espérance d'une variable aléatoire $\mathrm{Exp}(\lambda)$ vaut $1/\lambda$). Le temps qu'il faudra attendre pour avoir une réponse est donné par $T = \min(T_1, T_2)$ qui est de loi $\mathrm{Exp}(2/15)$, comme on l'a vu dans l'Exemple 6.39. On en déduit que $\mathrm{E}(T) = 15/2$.

(ii) En se rappelant que pour $X \sim \mathrm{Exp}(\lambda)$ on a $\mathrm{P}(X > t) = \mathrm{e}^{-\lambda t}$ pour tout $t \ge 0$, on obtient

$$\mathrm{P}(T \le 5) = 1 - \mathrm{P}(T > 5) = 1 - \mathrm{e}^{-\frac{2}{15}5} = 1 - \mathrm{e}^{-2/3} \simeq 0{,}49.$$

Exercice 6.32. Soit X une variable aléatoire de loi $\mathrm{Exp}(1)$.

(i) Pour $n \in \mathbb{N}^*$, calculer la médiane m_n de X^n. Montrer que $\lim_{n\to\infty} m_n = 0$ alors que $\lim_{n\to\infty} \mathrm{E}(X^n) = +\infty$.

(ii) Soit $Y \sim \mathrm{Exp}(1)$ une variable aléatoire indépendante de X. Calculer la fonction de répartition de $X+Y$ et montrer que le quantile d'ordre $\alpha \in\,]0,1[$ de $X+Y$ est strictement inférieur à la somme des quantiles d'ordre α de X et Y.

Solution 6.32. (i) On peut trouver une médiane m_n de X^n en calculant la fonction de répartition de X^n, voir la Proposition 3.99. On a $F_{X^n}(t) = 0$ pour $t < 0$. Pour $t > 0$ on a

$$F_{X^n}(t) = \mathrm{P}(X^n \le t) = \mathrm{P}(X \le t^{1/n}) = 1 - \mathrm{e}^{-t^{1/n}}\,,$$

où on a utilisé la fonction de répartition de X, donnée en (6.44). Autrement dit, on a $F_{X^n}(t) = F_X(t^{1/n})$ pour tout $t \in \mathbb{R}$.

La fonction de répartition F_{X^n} étant continue et strictement croissante sur $]0,+\infty[$, elle est bijective de $]0,+\infty[$ vers $]0,1[$: d'après le Lemme 6.33 (voir (3.64)), la médiane m_n et unique et coïncide avec $F_{X^n}^{-1}(\frac{1}{2})$, c'est-à-dire est l'unique solution de $F_{X^n}(t) = \frac{1}{2}$. On doit donc résoudre

$$1 - \mathrm{e}^{-t^{1/n}} = \frac{1}{2} \quad \Longleftrightarrow \quad \mathrm{e}^{t^{1/n}} = 2 \quad \Longleftrightarrow \quad t = (\log 2)^n\,.$$

Au final, la médiane vaut $m_n = (\log 2)^n$ et on observe que $\lim_{n\to\infty} m_n = 0$ parce que $\log 2 \approx 0{,}693 < 1$.

D'autre part, d'après la formule de transfert, l'espérance de X^n est égale à

$$\mathrm{E}(X^n) = \int_0^{+\infty} x^n \mathrm{e}^{-x} \mathrm{d}x = \Gamma(n+1) = n!\,,$$

où on a utilisé la Définition (6.34) de la fonction $\Gamma(\cdot)$ et la relation (6.36) (que l'on démontre par récurrence). En particulier $\lim_{n\to+\infty} \mathrm{E}(X^n) = +\infty$.

(ii) D'après la Proposition 6.36, $X+Y$ est une variable aléatoire Gamma$(2,1)$, donc elle admet pour densité

$$f_{X+Y}(x) = x\mathrm{e}^{-x}\mathbb{1}_{]0,+\infty[}(x)\,.$$

La fonction de répartition de $X+Y$ est donnée par $F_{X+Y}(t) = 0$ pour $t \le 0$ et pour $t > 0$

$$F_{X+Y}(t) = \int_0^t x\mathrm{e}^{-x}\mathrm{d}x = \left[-x\mathrm{e}^{-x}\right]_0^t + \int_0^t \mathrm{e}^{-x}\mathrm{d}x = 1-(1+t)\mathrm{e}^{-t}\,,$$

où on a utilisé une intégration par parties.
Les deux fonctions de répartitions F_X et F_Y étant continues et strictement croissantes sur $]0,+\infty[$, elles sont bijectives : en se rappelant de la Proposition 3.105, on a grâce au Lemme 6.33 que le quantile q_α de X (ou de Y) d'ordre $\alpha \in (0,1)$ est égal à $F_X^{-1}(\alpha)$, c'est-à-dire est la solution de

$$F_X(q_\alpha) = \alpha \quad \Longleftrightarrow \quad 1-\mathrm{e}^{-q_\alpha} = \alpha \quad \Longleftrightarrow \quad \mathrm{e}^{-q_\alpha} = 1-\alpha\,,$$

ce qui donne $q_\alpha = -\log(1-\alpha)$.
De manière analogue, le quantile $\tilde{q}_\alpha$ de $X+Y$ est égal à $F_{X+Y}^{-1}(\alpha)$, c'est-à-dire est la solution de

$$F_{X+Y}(\tilde{q}_\alpha) = \alpha \quad \Longleftrightarrow \quad (1+\tilde{q}_\alpha)\mathrm{e}^{-\tilde{q}_\alpha} = 1-\alpha\,.$$

Cette équation ne peut pas être résolue de manière exacte, mais on peut comparer $\tilde{q}_\alpha$ avec $2q_\alpha$, c'est-à-dire avec la somme des quantiles de X et Y d'ordre α. Il suffit en effet d'estimer $F_{X+Y}(2q_\alpha)$ et de le comparer avec α : si l'on montre que $F_{X+Y}(2q_\alpha) > \alpha$, vu que F_{X+Y} est strictement croissante, cela signifie que $\tilde{q}_\alpha < 2q_\alpha$.
Comme $q_\alpha = -\log(1-\alpha)$ (donc $\mathrm{e}^{-q_\alpha} = 1-\alpha$), on a

$$\begin{aligned} F_{X+Y}(2q_\alpha) - \alpha &= (1-\alpha) - (1-2q_\alpha)\mathrm{e}^{-2q_\alpha} \\ &= (1-\alpha) - (1+2\log(1-\alpha))(1-\alpha)^2\,. \end{aligned}$$

Pour montrer que $F_{X+Y}(2q_\alpha) > \alpha$, il suffit donc de montrer que pour tout $\alpha \in\,]0,1[$

$$1-(1+2\log(1-\alpha))(1-\alpha) > 0\,.$$

On étudie pour cela la fonction définie par $g(x) = 1-(1+2\log x)x$ pour tout $x \in\,]0,1]$. On a que $g'(x) = -3+2\log x$ est strictement positive si $x < \mathrm{e}^{-3/2}$ et est strictement négative si $x > \mathrm{e}^{-3/2}$: ainsi g est strictement croissante sur $]0,\mathrm{e}^{-3/2}[$ puis strictement décroissante sur $]\mathrm{e}^{-3/2},1[$. Comme $\lim_{x\downarrow 0} g(x) = 1$ et $g(1) = 0$, on obtient que $g(x) > 0$ pour tout $x \in\,]0,1[$, ce qui conclut la démonstration.

Exercice 6.33. Soit X une variable aléatoire de loi $\mathrm{N}(0,1)$ et soit $Y = \mathrm{e}^X$.

(i) Montrer que Y est une variable aléatoire positive à densité, et en donner la densité $f_Y(y)$. *La loi de Y est appelée* log-normale.

(ii) Calculer $\mathrm{E}(Y^n)$ pour tout $n \in \mathbb{N}$.

(iii) Montrer que, pour tout $n \in \mathbb{N}$, on a $\mathrm{E}(\sin(2\pi X)\mathrm{e}^{nX}) = 0$.

(iv) On considère maintenant la variable aléatoire positive W, de densité donnée par

$$f_W(t) = (1 + \sin(2\pi \log t)) f_Y(t) \qquad \forall t \in \mathbb{R}.$$

Vérifier que f_W est bien une densité de probabilité et montrer que $\mathrm{E}(W^n) = \mathrm{E}(Y^n)$ pour tout $n \in \mathbb{N}$.

On a montré que Y et W, qui ont des densités différentes, ont les mêmes moments. Ainsi, la loi log-normale *n'est pas déterminée par ses moments.*

Solution 6.33. (i) Utilisons l'Observation 6.26. Soit h une fonction continue par morceaux et bornée quelconque. D'après la formule de transfert on a

$$\mathrm{E}(h(\mathrm{e}^X)) = \int_{-\infty}^{+\infty} h(\mathrm{e}^x) \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}x^2} \mathrm{d}x = \int_0^{\infty} h(y) \frac{1}{\sqrt{2\pi}} y^{-1} \mathrm{e}^{-\frac{1}{2}(\log y)^2} \mathrm{d}y,$$

où on a utilisé le changement de variable $y = \mathrm{e}^x$ ($x = \log y$). Comme cette relation est vérifiée pour toute fonction h continue par morceaux et bornée, d'après l'Observation 6.26 on en déduit que Y est absolument continue, de densité donnée par

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} y^{-1} \mathrm{e}^{-\frac{1}{2}(\log y)^2} \mathbb{1}_{]0,\infty[}(y).$$

(ii) On a $\mathrm{E}(Y^n) = \mathrm{E}(\mathrm{e}^{nX}) = \mathrm{e}^{\frac{1}{2}n^2}$, où on a utilisé que la fonction génératrice des moments de $X \sim \mathrm{N}(0,1)$ est donnée par $\mathrm{M}_X(t) = \mathrm{e}^{\frac{1}{2}t^2}$ pour tout $t \in \mathbb{R}$, voir (6.49).

(iii) En utilisant la formule de transfert, on a

$$\mathrm{E}(\sin(2\pi X)\,\mathrm{e}^{nX}) = \int_{-\infty}^{+\infty} \sin(2\pi x)\,\mathrm{e}^{nx}\, \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{-\frac{1}{2}x^2}\, \mathrm{d}x.$$

On peut « compléter le carré » en écrivant $\mathrm{e}^{nx}\,\mathrm{e}^{-\frac{1}{2}x^2} = \mathrm{e}^{\frac{1}{2}n^2}\,\mathrm{e}^{-\frac{1}{2}(x-n)^2}$. Avec le changement de variable $z := x - n$ on obtient

$$\begin{aligned}
\mathrm{E}(\sin(2\pi X)\,\mathrm{e}^{nX}) &= \mathrm{e}^{\frac{1}{2}n^2} \int_{-\infty}^{+\infty} \sin(2\pi x)\, \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{-\frac{1}{2}(x-n)^2}\, \mathrm{d}x \\
&= \mathrm{e}^{\frac{1}{2}n^2} \int_{-\infty}^{+\infty} \sin(2\pi(z+n))\, \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{-\frac{1}{2}z^2}\, \mathrm{d}z.
\end{aligned}$$

En observant que $\sin(2\pi(z+n)) = \sin(2\pi z)$ par périodicité, on en conclut que

$$\mathrm{E}(\sin(2\pi X)\,\mathrm{e}^{nX}) = \mathrm{e}^{\frac{1}{2}n^2} \int_{-\infty}^{+\infty} \sin(2\pi z)\, \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{-\frac{1}{2}z^2}\, \mathrm{d}z = 0,$$

parce que la fonction intégrée est (intégrable et) impaire : $\sin(-2\pi z) = -\sin(2\pi z)$.

(iv) La fonction f_W est positive, parce que $\sin(2\pi \log t) \le 1$, et continue sur $]0, +\infty[$. De plus, en utilisant la formule de transfert on a

$$\begin{aligned}
\int_{-\infty}^{+\infty} f_W(t)\mathrm{d}t = \int_0^{\infty} (1 + \sin(2\pi \log t)) f_Y(t)\mathrm{d}t &= \mathrm{E}(1 + \sin(2\pi \log Y)) \\
&= 1 + \mathrm{E}(\sin(2\pi X)) = 1,
\end{aligned}$$

où on a utilisé la linéarité de l'espérance et la question précédente pour voir que $\mathrm{E}(\sin(2\pi X)) = 0$.

De la même manière, pour toute fonction $h :]0,+\infty[\to \mathbb{R}$ continue par morceaux et positive, on a

$$\mathrm{E}(h(W)) = \int_0^\infty (1+\sin(2\pi\log t))h(t)f_Y(t)\mathrm{d}t = \mathrm{E}(h(Y)) + \mathrm{E}(h(Y)\sin(2\pi\log Y).$$

Ainsi, pour tout $n \in \mathbb{N}$, $\mathrm{E}(W^n) = \mathrm{E}(Y^n) + \mathrm{E}(Y^n\sin(2\pi\log Y))$, avec

$$\mathrm{E}(Y^n\sin(2\pi\log Y)) = \mathrm{E}(\mathrm{e}^{nX}\sin(2\pi X)) = 0,$$

grâce à la question précédente.

Exercice 6.34. Pour une variable aléatoire réelle X de densité f_X, on définit l'entropie de (la loi de) X par

$$H_X := -\int_{\mathbb{R}^n} \log(f_X(x))f_X(x)\mathrm{d}x = -\mathrm{E}\big(\log f_X(X)\big),$$

si $f_X|\log f_X|$ est intégrable (on utilise la convention $0\log 0 = 0$).

(i) Exprimer H_{aX+b} en fonction de H_X, pour $a \neq 0$, $b \in \mathbb{R}$.

(ii) Soit $X \sim \mathrm{N}(0,1)$. Montrer que $H_X = \frac{1}{2}\log(2\pi\mathrm{e})$.

(iii) Soit Y une variable aléatoire à densité telle que $\mathrm{E}(Y) = 0$ et $\mathrm{Var}(Y) = 1$. Soit $X \sim \mathrm{N}(0,1)$.

(a) Montrer que $H_X = -\mathrm{E}(\log f_X(Y))$, puis que $H_X - H_Y = \mathrm{E}\big(\frac{f_Y(X)}{f_X(X)}\log\frac{f_Y(X)}{f_X(X)}\big)$.

(b) Montrer que $x \mapsto x\log x$ est convexe sur $\mathbb{R}_+$, et en déduire que l'on a $H_X \geq H_Y$.

Autrement dit, la variable aléatoire $\mathrm{N}(0,1)$ *est celle qui maximise l'entropie parmi les variables aléatoires d'espérance* 0 *et de variance* 1.

Solution 6.34. (i) D'après la Proposition 6.19, la densité de $aX+b$ est donnée par $f_{aX+b}(x) = \frac{1}{|a|}f_X(\frac{x-b}{a})$. On obtient donc

$$\begin{aligned}
H_{aX+b} &= -\int_{-\infty}^{+\infty} \log\Big(\frac{1}{|a|}f_X(\tfrac{x-b}{a})\Big)\frac{1}{|a|}f_X(\tfrac{x-b}{a})\mathrm{d}x \\
&= \log|a|\int_{-\infty}^{+\infty}\frac{1}{|a|}f_X(\tfrac{x-b}{a})\mathrm{d}x - \int_{-\infty}^{+\infty}\log\big(f_X(\tfrac{x-b}{a})\big)\frac{1}{|a|}f_X(\tfrac{x-b}{a})\mathrm{d}x \\
&= \log|a| - \int_{-\infty}^{+\infty}\log\big(f_X(t)\big)f_X(t)\mathrm{d}t = \log|a| + H_X,
\end{aligned}$$

où on a opéré un changement de variable $t = (x-b)/a$ et utilisé le fait que $\int_{-\infty}^{+\infty} f_X(x) = 1$.

(ii) Si $X \sim \mathrm{N}(0,1)$, comme $f_X(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}$, on a

$$\begin{aligned}
H_X = -\mathrm{E}(\log f_X(X)) &= -\mathrm{E}\Big(\log\big(\tfrac{1}{\sqrt{2\pi}}\big) - \frac{1}{2}X^2\Big) \\
&= \frac{1}{2}\big(\log(2\pi)+1\big) = \frac{1}{2}\log(2\pi\mathrm{e}).
\end{aligned}$$

(iii) (a) Notons que si $X \sim \mathrm{N}(0,1)$, alors on a encore $f_X(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}$. Si $\mathrm{E}(Y) = 0$ et $\mathrm{E}(Y^2) = 1$, on obtient de la même manière que dans la question précédente

$$-\mathrm{E}\big(\log f_X(Y)\big) = -\mathrm{E}\Big(\log\big(\tfrac{1}{\sqrt{2\pi}}\big) - \frac{1}{2}Y^2\Big) = \frac{1}{2}\big(\log(2\pi)+1\big) = H_X\,.$$

On peut alors écrire (rappelons que l'on utilise la convention $0\log 0 = 0$)

$$\begin{aligned} H_X - H_Y &= -\mathrm{E}(\log f_X(Y)) + \mathrm{E}(\log f_Y(Y)) = \mathrm{E}\left(\log\frac{f_Y(Y)}{f_X(Y)}\right) \\ &= \int_{\mathbb{R}}\left(\log\frac{f_Y(y)}{f_X(y)}\right) f_Y(y)\mathrm{d}y \\ &= \int_{\mathbb{R}}\frac{f_Y(y)}{f_X(y)}\left(\log\frac{f_Y(y)}{f_X(y)}\right) f_X(y)\mathrm{d}y = \mathrm{E}\left(\frac{f_Y(X)}{f_X(X)}\log\frac{f_Y(X)}{f_X(X)}\right). \end{aligned}$$

(b) Soit $g : x \mapsto x\log x$ la fonction définie sur $\mathbb{R}_+$ (par convention $g(0)=0$). On a que $g'(x) = 1+\log x$ est (strictement) croissante sur $]0,+\infty[$, donc g est une fonction convexe. Grâce à l'inégalité de Jensen (Proposition 3.80), on obtient

$$H_X - H_Y = \mathrm{E}\left(g\left(\frac{f_Y(X)}{f_X(X)}\right)\right) \geq g\left(\mathrm{E}\left(\frac{f_Y(X)}{f_X(X)}\right)\right).$$

Maintenant, par un simple calcul,

$$\mathrm{E}\left(\frac{f_Y(X)}{f_X(X)}\right) = \int_{-\infty}^{+\infty}\frac{f_Y(x)}{f_X(x)}\cdot f_X(x)\mathrm{d}x = \int_{-\infty}^{+\infty} f_Y(x)\mathrm{d}x = 1\,.$$

On a donc montré que $H_X - H_Y \geq g(1) = 1\log 1 = 0$.

Exercice 6.35. Soit X un point aléatoire de l'intervalle $]0,1[$, pas nécessairement uniformément distribué. Celui-ci divise l'intervalle $]0,1[$ en deux segments. Soit $Y \geq 1$ le rapport entre le segment le plus long et le segment le plus court.

(i) Exprimer Y en fonction de X.
[*Sugg.* On pourra trouver utile d'utiliser les événements $\{X < \frac{1}{2}\}$ et $\{X \geq \frac{1}{2}\}$.]

(ii) On suppose que $X \sim \mathrm{U}(0,1)$. Déterminer la fonction de répartition et la densité de Y. Montrer que Y n'admet pas d'espérance.

(iii) On suppose maintenant que X est une variable aléatoire à densité, à valeurs dans $]0,1[$, dont la densité f_X satisfait la relation :

$$f_X(x) + f_X(1-x) = 2\,, \qquad \text{pour tout } x \in\,]0,1[\,. \tag{S6.10}$$

Montrer que la loi de Y est égale à celle trouvée dans la question précédente.
[*Sugg.* Déduire de la relation (S6.10) que $F_X(z) - F_X(1-z) = 2z-1$ pour $z \in\,]0,1[$.]

(iv) Donner une densité f_X qui satisfait la relation (S6.10), mais qui est *différente* de la densité d'une variable aléatoire de loi $\mathrm{U}(0,1)$.

Solution 6.35. (i) Par définition on a

$$Y = \left(\frac{X}{1-X}\right)\mathbb{1}_{\{X\geq\frac{1}{2}\}} + \left(\frac{1-X}{X}\right)\mathbb{1}_{\{X<\frac{1}{2}\}}\,.$$

(ii) Comme $Y \geq 1$, on a $F_Y(y) = 0$ pour $y < 1$, alors que pour $y \geq 1$, en utilisant le point précédent

$$
\begin{aligned}
F_Y(y) &= \mathrm{P}(Y \le y) = \mathrm{P}(Y \le y, X \ge \tfrac{1}{2}) + \mathrm{P}(Y \le y, X < \tfrac{1}{2}) \\
&= \mathrm{P}(\tfrac{X}{1-X} \le y, X \ge \tfrac{1}{2}) + \mathrm{P}(\tfrac{1-X}{X} \le y, X < \tfrac{1}{2}) \\
&= \mathrm{P}(X \le \tfrac{y}{1+y}, X \ge \tfrac{1}{2}) + \mathrm{P}(X \ge \tfrac{1}{1+y}, X < \tfrac{1}{2}) \\
&= \mathrm{P}(\tfrac{1}{2} \le X \le \tfrac{y}{1+y}) + \mathrm{P}(\tfrac{1}{1+y} \le X < \tfrac{1}{2}) = \mathrm{P}(\tfrac{1}{1+y} \le X \le \tfrac{y}{1+y}) .
\end{aligned}
\tag{S6.11}
$$

Comme $X \sim \mathrm{Unif}(0,1)$, pour $0 \le a \le b \le 1$ on a $\mathrm{P}(a \le X \le b) = |[a,b]| = b - a$, donc

$$
F_Y(y) = \begin{cases} 0 & \text{si } y < 1\,, \\ \frac{y-1}{y+1} & \text{si } y \ge 1\,. \end{cases} \tag{S6.12}
$$

On a ainsi que F_Y est continue et $\mathcal{C}^1$ par morceaux, donc Y est absolument continue de densité

$$
f_Y(y) = F_Y'(y) = \frac{2}{(1+y)^2}\, \mathbb{1}_{[1,\infty[}(y)\,.
$$

L'espérance de Y (qui est bien définie car Y est positive) est donc donnée par

$$
\mathrm{E}(Y) = \int_{\mathbb{R}} y\, f_Y(y)\, \mathrm{d}y = \int_1^\infty \frac{2y}{(1+y)^2}\, \mathrm{d}y = +\infty,
$$

parce que $\frac{2y}{(1+y)^2} \sim \frac{2}{y}$ quand $y \to +\infty$.

(iii) En appliquant la formule (S6.11) obtenue au point précédent, pour $y \ge 1$ on a

$$
F_Y(y) = \mathrm{P}(\tfrac{1}{1+y} \le X \le \tfrac{y}{1+y}) = \int_{\frac{1}{1+y}}^{\frac{y}{1+y}} f_X(x)\, \mathrm{d}x\,.
$$

Avec le changement de variables $x = 1 - z$ dans l'intégrale, en notant que $1 - \frac{1}{1+y} = \frac{y}{1+y}$, on obtient

$$
F_Y(y) = \int_{\frac{1}{1+y}}^{\frac{y}{1+y}} f_X(1-z)\, \mathrm{d}z\,.
$$

En changeant la variable d'intégration $z \to x$ et en faisant la moyenne des deux expressions, on obtient

$$
F_Y(y) = \int_{\frac{1}{1+y}}^{\frac{y}{1+y}} \frac{1}{2}\big(f_X(x) + f_X(1-x)\big)\, \mathrm{d}x\,.
$$

Par hypothèse $\frac{1}{2}(f_X(x) + f_X(1-x)) = 1$, ainsi $F_Y(y) = \frac{y}{1+y} - \frac{1}{1+y} = \frac{y-1}{1+y}$ pour tout $y \ge 1$. Comme on a clairement $F_Y(y) = 0$ pour $y < 1$, la fonction de répartition de Y, et donc sa loi, est la même que celle déterminée dans le point précédent.

(iv) Il suffit de considérer par exemple, la densité donnée par

$$
f(x) := \begin{cases} \frac{3}{2} & \text{si } x \in (0, \frac{1}{2})\,, \\ 1 & \text{si } x = \frac{1}{2}\,, \\ \frac{1}{2} & \text{si } x \in (\frac{1}{2}, 1)\,, \\ 0 & \text{sinon}\,. \end{cases}
$$

(La valeur de la densité $f(x)$ pour $x = \frac{1}{2}$ n'est pas importante, mais elle a été définie explicitement pour faire en sorte que la relation (S6.10) soit satisfaite.)

Exercice 6.36. Soient $(X_n)_{n\geq 1}$ des variables aléatoires i.i.d. de loi $X_i \sim \mathrm{U}(0,1)$.

(i) On pose $Y_n := -\log(X_n)$ pour $n \in \mathbb{N}^*$. Déterminer la loi des Y_n et expliquer pourquoi les variables aléatoires $(Y_n)_{n\geq 1}$ sont indépendantes.

(ii) Déterminer la loi de $S_n := Y_1 + \cdots + Y_n$.

(iii) Calculer la fonction de répartition de $Z_n := X_1 \cdot X_2 \cdots X_n$ et en déduire que Z_n est une variable aléatoire à densité. En calculer la densité.

[*Sugg.* Utiliser les questions précédentes.]

Solution 6.36. (i) Comme

$$\mathrm{P}(Y_n \leq y) = \mathrm{P}(-\log(X_n) \leq y) = \mathrm{P}(X_n \geq \mathrm{e}^{-y}) = 1 - F_{X_n}(\mathrm{e}^{-y}),$$

on en déduit que Y_n est absolument continue, de densité

$$f_{Y_n}(y) = f_{X_n}(\mathrm{e}^{-y})\mathrm{e}^{-y} = \mathrm{e}^{-y}\,\mathbb{1}_{[0,\infty[}(y),$$

c'est-à-dire $Y_n \sim \mathrm{Exp}(1) = \mathrm{Gamma}(1,1)$. Les variables aléatoires $(Y_n)_{n\geq 1}$ sont indépendantes car fonctions individuellement des variables aléatoires indépendantes $(X_n)_{n\geq 1}$, grâce à la Proposition 3.39.

(ii) Il découle des propriétés notables des variables aléatoires Gamma (Proposition 6.36) que $S_n \sim \mathrm{Gamma}(n,1)$, c'est-à-dire $f_{S_n}(s) = \frac{1}{(n-1)!}\, s^{n-1}\,\mathrm{e}^{-s}\,\mathbb{1}_{[0,\infty[}(s)$.

(iii) Comme $Z_n \in [0,1]$, on a que $F_{Z_n}(z) = 0$ si $z < 0$ et $F_{Z_n}(z) = 1$ si $z > 1$, alors que pour $z \in \,]0,1[$

$$F_{Z_n}(z) = \mathrm{P}(X_1 \cdots X_n \leq z) = \mathrm{P}(\mathrm{e}^{-S_n} \leq z) = \mathrm{P}(S_n \geq -\log z) = 1 - F_{S_n}(-\log z).$$

On en déduit que

$$f_{Z_n}(z) = F'_{Z_n}(z) = \frac{1}{z}\, f_{S_n}(-\log z) = \frac{1}{(n-1)!}(-\log z)^{n-1}\,\mathbb{1}_{]0,1[}(z).$$

Exercice 6.37. Soient $(X_n)_{n\geq 1}$ des variables aléatoires i.i.d. de loi $\mathrm{U}(0,2)$ et soit

$$Y_n := \min\{Y_1, \ldots, Y_n\}, \qquad \forall n \in \mathbb{N}^*.$$

Soit $T \sim \mathrm{Géom}(p)$ une variable aléatoire indépendante des $(X_n)_{n\geq 1}$, et on pose

$$Z := Y_T, \qquad \text{c'est-à-dire} \quad Z(\omega) := Y_{T(\omega)}(\omega).$$

Déterminer la fonction de répartition de Z et montrer qu'il s'agit d'une variable aléatoire à densité.

[*Sugg.* Déterminer tout d'abord $\mathrm{P}(Z \leq x, T = n)$ pour tout $n \in \mathbb{N}^*$.]

Solution 6.37. Comme sur l'événement $\{T = n\}$ on a $Z = Y_n$, on peut écrire

$$\mathrm{P}(Z \leq z, T = n) \,=\, \mathrm{P}(Y_n \leq z, T = n) \,=\, \mathrm{P}(Y_n \leq z)\,\mathrm{P}(T = n).$$

en ayant utilisé l'indépendance de Y_n et T. En se rappelant de la Proposition 3.107, ou par un calcul direct, on obtient

$$\begin{aligned} \mathrm{P}(Y_n \leq z) &= \mathrm{P}(\min\{X_1, \ldots, X_n\} \leq z) \,=\, 1 - \mathrm{P}(\min\{X_1, \ldots, X_n\} > z) \\ &= 1 - \mathrm{P}(X_1 > z)\,\mathrm{P}(X_2 > z) \cdots \mathrm{P}(X_n > z) \,=\, 1 - (1-z)^n. \end{aligned}$$

Observons que $P(T=n) = p(1-p)^{n-1}\mathbb{1}_{\mathbb{N}^*}(n)$. On peut alors calculer la fonction de répartition de Z : vu que Z est à valeurs dans $[0,1]$, on a $F_Z(z)=0$ si $z<0$ et $F_Z(z)=1$ si $z>1$, alors que pour $z\in[0,1]$, en remarquant que les événements $\{T=n\}$ pour $n\in\mathbb{N}^*$ forment une partition de l'espace probabilisé, on obtient

$$\begin{aligned} F_Z(z) &= P(Z\le z) = \sum_{n\in\mathbb{N}^*} P(Z\le z, T=n) = p\sum_{n=1}^{\infty}(1-(1-z)^n)(1-p)^{n-1} \\ &= \sum_{n=1}^{\infty} p(1-p)^{n-1} - p(1-z)\sum_{n=1}^{\infty}[(1-z)(1-p)]^{n-1} \\ &= 1-\frac{p(1-z)}{1-(1-z)(1-p)} = \frac{z}{p+(1-p)z}, \end{aligned}$$

où l'on a utilisé la somme connue (0.5) de la série géométrique. On obtient que F_Z est continue et $\mathcal{C}^1$ par morceaux, dont on déduit que Z est à densité, de densité donnée par

$$f_Z(z) = F_Z'(z) = \frac{p}{(p+(1-p)z)^2}\mathbb{1}_{]0,1[}(z).$$

Exercice 6.38. Soit $(X_n)_{n\ge1}$ une suite de variables aléatoires réelles i.i.d. de loi $U(0,1)$. On introduit la variable aléatoire $T:\Omega\to\mathbb{N}\cup\{+\infty\}$ et pour $k\in\mathbb{N}^*$ l'événement A_k définis par

$$T(\omega) := \inf\left\{k\ge1:\ X_k(\omega)\le\frac{1}{3}\right\},\qquad A_k=\left\{X_k\le\frac{1}{3}\right\}.$$

On définit alors la variable aléatoire

$$Y := X_T\,\mathbb{1}_{\{T<\infty\}},$$

c'est-à-dire $Y(\omega):=X_{T(\omega)}(\omega)$ si $T(\omega)<\infty$, et $Y(\omega):=0$ sinon.

(i) Pour tout $n\in\mathbb{N}^*$, exprimer l'événement $\{T=n\}$ en terme des événements $(A_k)_{k\ge1}$. En déduire la loi de T.

(ii) Déterminer la loi de Y.

[*Sugg.* Calculer $P(Y\le x, T=n)$ pour $n\in\mathbb{N}^*$.]

Solution 6.38. (i) Par définition, on a l'égalité d'événements

$$\{T=n\} = \left(A_1^c\cap A_2^c\cap\ldots\cap A_{n-1}^c\right)\cap A_n. \tag{S6.13}$$

Les événements $(A_k)_{k\ge1}$ sont indépendants, parce que les variables aléatoires $(X_k)_{k\ge1}$ sont indépendantes. Comme $P(A_k)=\frac{1}{3}$ (pourquoi ?), il s'ensuit que pour tout $n\in\mathbb{N}^*$

$$P(T=n) = P(A_1^c)\,P(A_2^c)\cdots P(A_{n-1}^c)\,P(A_n) = \frac{1}{3}\left(\frac{2}{3}\right)^{n-1},$$

c'est-à-dire $T\sim \text{Géom}(\frac{1}{3})$.

(ii) Notons que pour $x\in[0,1]$

$$P(Y\le x, T=n) = P(X_T\le x, T=n) = P(X_n\le x, T=n).$$

En appliquant la relation (S6.13) on peut écrire

$$\begin{aligned}\mathrm{P}(Y \le x, T = n) &= \mathrm{P}\left(X_1 > \tfrac{1}{3}, X_2 > \tfrac{1}{3}, \ldots, X_{n-1} > \tfrac{1}{3}, X_n \le \tfrac{1}{3}, X_n \le x\right)\\ &= \mathrm{P}\left(X_1 > \tfrac{1}{3}, X_2 > \tfrac{1}{3}, \ldots, X_{n-1} > \tfrac{1}{3}, X_n \le \min\{\tfrac{1}{3}, x\}\right).\end{aligned}$$

En utilisant le fait que les variables aléatoires $(X_k)_{k\ge1}$ sont i.i.d., on obtient

$$\mathrm{P}(Y \le x, T = n) = \mathrm{P}\left(X_1 > \tfrac{1}{3}\right)^{n-1} \mathrm{P}\left(X_1 \le \min\{\tfrac{1}{3}, x\}\right) = \left(\frac{2}{3}\right)^{n-1} \min\left\{\frac{1}{3}, x\right\}.$$

Par conséquent, pour $x \in [0,1]$, on peut écrire

$$\begin{aligned}F_Y(x) = \mathrm{P}(Y \le x) &= \sum_{n\in\mathbb{N}^*} \mathrm{P}(Y \le x, T = n) = \min\left\{\frac{1}{3}, x\right\} \times \sum_{n\in\mathbb{N}^*} \left(\frac{2}{3}\right)^{n-1}\\ &= \min\left\{\frac{1}{3}, x\right\} \times 3 = \begin{cases} 3x & \text{si } 0 \le x < \frac{1}{3}, \\ 1 & \text{si } \frac{1}{3} \le x \le 1. \end{cases}\end{aligned}$$

Comme $F_Y(x) = 0$ si $x < 0$ et $F_Y(x) = 1$ si $x > 1$, la fonction F_Y est continue et $\mathcal{C}^1$ par morceaux, donc Y est absolument continue, de densité

$$f_Y(x) = F_Y'(x) = 3\,\mathbb{1}_{]0,\frac{1}{3}[}(x),$$

c'est-à-dire $Y \sim \mathrm{U}(0, \frac{1}{3})$. Autrement dit, on a montré qu'une variable aléatoire de loi $\mathrm{U}(0, \frac{1}{3})$ peut être obtenue à partir d'une suite de variables aléatoires i.i.d. $\mathrm{U}(0,1)$, en considérant la première de ces variables aléatoires qui prend une valeur inférieure ou égale à $\frac{1}{3}$.

Exercice 6.39. Soient X et Y des variables aléatoires réelles indépendantes. On suppose que X est à densité, de densité f_X, alors que Y est discrète, de densité discrète p_Y. On suppose de plus que Y ne prend qu'un nombre fini de valeurs, c'est-à-dire que l'ensemble $Y(\Omega) = \{x \in \mathbb{R} : \mathrm{p}_Y(x) > 0\} = \{x_1, \ldots, x_n\}$ est fini.

(i) Exprimer la fonction de répartition de $Z := X + Y$ en termes de f_X et p_Y.

(ii) Montrer que Z est une variable aléatoire à densité, et en déterminer la densité en fonction de f_X et p_Y.

Solution 6.39. (i) Soit $\{x_1, \ldots, x_n\}$ l'ensemble des valeurs prises par Y. Comme les événements $\{Y = x_k\}$ pour $k \in \{1, \ldots, n\}$ forment une partition de l'espace probabilisé, pour tout $t \in \mathbb{R}$ on peut écrire

$$\begin{aligned}\mathrm{P}(X + Y \le t) &= \sum_{k=1}^{n} \mathrm{P}(X \le t - x_k, Y = x_k) = \sum_{k=1}^{n} \mathrm{p}_Y(x_k) \int_{-\infty}^{t-x_k} f_X(x)\,\mathrm{d}x\\ &= \sum_{k=1}^{n} \mathrm{p}_Y(x_k) \int_{-\infty}^{t} f_X(x - x_k)\,\mathrm{d}x \qquad \text{(S6.14)}\\ &= \int_{-\infty}^{t} \left(\sum_{k=1}^{n} f_X(x - x_k)\,\mathrm{p}_Y(x_k)\right) \mathrm{d}x.\end{aligned}$$

(ii) La relation (S6.14) montre que la fonction de répartition F_{X+Y} de la variable aléatoire $X + Y$ est donnée par

$$F_{X+Y}(t) = \int_{-\infty}^{t} f(x)\,\mathrm{d}x, \qquad \text{où} \qquad f(x) := \sum_{k=1}^{n} f_X(x-x_k)\,\mathrm{p}_Y(x_k).$$

Cela signifie que $X+Y$ est absolument continue, de densité donnée par $f_{X+Y}(x) = f(x)$.

Exercice 6.40. Un circuit électrique est constitué de n composants en série : le circuit cesse de fonctionner dès que l'un des composants est cassé. Les temps de vie $T_1, T_2, \dots, T_n$ des n composants sont des variables aléatoires réelles indépendantes et de même loi absolument continue, dont on note la densité f. Clairement, $f(x) = 0$ pour $x < 0$ (les temps sont des quantités positives !). On suppose que f est une fonction continue sur $\mathbb{R}_+$, avec $f(0) > 0$. En notant X_n le temps de vie du dispositif entier, montrer que, pour tout $\varepsilon > 0$,

$$\lim_{n\to+\infty} \mathrm{P}(X_n > \varepsilon) = 0$$

Solution 6.40. Par construction on a $X_n = \min(T_1, T_2, \dots, T_n)$, donc

$$\mathrm{P}(X_n > \varepsilon) = \mathrm{P}(T_1 > \varepsilon, T_2 > \varepsilon, \dots, T_n > \varepsilon) = \mathrm{P}(T_1 > \varepsilon)^n, \tag{S6.15}$$

où on a utilisé le fait que les variables aléatoires $T_1, T_2, \dots, T_n$ sont i.i.d.. On observe que

$$\mathrm{P}(T_1 > \varepsilon) = 1 - \mathrm{P}(T_1 \le \varepsilon) = 1 - \int_0^{\varepsilon} f(x)\,\mathrm{d}x. \tag{S6.16}$$

Vu que $f(0) > 0$ et que f est continue, il existe $\delta > 0$ tel que $f(x) > \frac{1}{2} f(0)$ pour tout $x \in [0,\delta]$. En particulier, pour tout $\varepsilon \in \,]0,\delta[$ on a que

$$\int_0^{\varepsilon} f(x)\,\mathrm{d}x \ge \int_0^{\varepsilon} \frac{1}{2} f(0)\,\mathrm{d}x = \frac{1}{2} f(0)\,\varepsilon > 0.$$

L'intégrale est une fonction croissante de ε, on a donc aussi $\int_0^{\varepsilon} f(x)\,\mathrm{d}x > 0$ pour $\varepsilon \ge \delta$. Au final, $\int_0^{\varepsilon} f(x)\,\mathrm{d}x > 0$ pour tout $\varepsilon > 0$; on déduit ainsi de (S6.16) que $\mathrm{P}(T_1 > \varepsilon) < 1$ et donc $\lim_{n\to+\infty} \mathrm{P}(X_n > \varepsilon) = 0$, grâce à (S6.15).

Exercice 6.41 (Biais par la taille, I). Soit X une variable aléatoire *positive* de densité $f(x)$, qui admet une espérance $\mathrm{E}(X)$ *finie*. On considère X^* une variable aléatoire (positive) de densité

$$f_{X^*}(x) = \frac{1}{\mathrm{E}(X)}\, x f(x).$$

La variable aléatoire X^* est appelée *version biaisée par la taille* de X.

(i) Montrer que pour toute fonction $g : \mathbb{R} \to \mathbb{R}$ continue par morceaux et bornée on a

$$\mathrm{E}(g(X^*)) = \frac{1}{\mathrm{E}(X)}\, \mathrm{E}(X g(X)).$$

(ii) Montrer que les fonctions de répartitions de X et X^* vérifient $F_{X^*}(t) \le F_X(t)$, $\forall t \in \mathbb{R}$.

[*Sugg.* Noter que $x \mapsto x$ est croissante alors que $x \mapsto \mathbb{1}_{]-\infty,t]}(x)$ est décroissante, puis utiliser l'inégalité FKG (Proposition 3.83).]

(iii) (*) Soit $U \sim \mathcal{U}(0,1)$ une variable aléatoire indépendante de X^*. On pose $Y = UX^*$. Montrer que Y est à densité, de densité donnée par

$$f_Y(y) = \frac{1}{\mathrm{E}(X)}\, \mathrm{P}(X > y)\mathbb{1}_{\mathbb{R}_+}(y).$$

Solution 6.41. (i) D'après la formule de transfert (Proposition 6.22) appliquée à X^*, on a

$$\mathrm{E}(g(X^*)) = \int_0^{+\infty} g(x)\frac{1}{\mathrm{E}(X)}xf(x)\mathrm{d}x$$

où l'intégrale est seulement sur $\mathbb{R}_+$ car, X étant positive, on peut supposer que $f(x) = 0$ pour $x < 0$. Le fait que g soit bornée nous assure que l'intégrale est finie. En appliquant de nouveau la formule de transfert (Proposition 6.22), cette fois à X avec la fonction $x \mapsto xg(x)$, on a

$$\mathrm{E}(g(X^*)) = \frac{1}{\mathrm{E}(X)}\,\mathrm{E}\left(Xg(X)\right).$$

(ii) Fixons $t \in \mathbb{R}$, et notons g_t la fonction définie par $g_t(x) = \mathbb{1}_{]-\infty,t]}(x)$, c'est-à-dire $g_t(x) = x$ si $x \leq t$ et $g_t(x) = 0$ si $x > t$. Notons que pour toute variable aléatoire réelle Z on a $g_t(Z) = \mathbb{1}_{\{Z\leq t\}}$ et en particulier $\mathrm{E}(g_t(Z)) = \mathrm{P}(Z \leq t) = F_Z(t)$. Grâce à la question précédente, g_t étant continue par morceaux et bornée, on obtient

$$F_{X^*}(t) = \mathrm{E}(g_t(X^*)) = \frac{1}{\mathrm{E}(X)}\,\mathrm{E}(Xg_t(X)).$$

Maintenant, remarquons que $x \mapsto g_t(x)$ est décroissante, donc $x \mapsto -g_t(x)$ est croissante : vu que $x \mapsto x$ est aussi croissante, grâce à l'inégalité FKG (Proposition 3.83) on obtient

$$-\mathrm{E}(Xg_t(X)) \geq -\mathrm{E}(g_t(X))\,\mathrm{E}(X) \quad\Longrightarrow\quad \mathrm{E}(Xg_t(X)) \leq \mathrm{E}(X)\,\mathrm{E}(g_t(X)).$$

Appliqué dans l'identité précédente, cela donne

$$F_{X^*}(t) \leq \frac{1}{\mathrm{E}(X)}\,\mathrm{E}(X)\,\mathrm{E}(g_t(X)) = \mathrm{E}(g_t(X)) = F_X(t).$$

(iii) Utilisons l'Observation 6.26. Soit h une fonction continue par morceaux et bornée quelconque. Notons que, U et X^* étant indépendants, la densité jointe de (U,X^*) est donnée par $f_{U,X^*}(u,x) = f_U(u)f_{X^*}(x) = \mathbb{1}_{]0,1[}(u)\frac{1}{\mathrm{E}(X)}xf(x)\mathbb{1}_{]0,\infty[}(x)$, d'après la Proposition 6.50. En utilisant la formule de transfert (Proposition 6.56) avec la fonction $(u,x) \mapsto h(ux)$, on obtient

$$\begin{aligned}\mathrm{E}(h(Y)) = \mathrm{E}(h(UX^*)) &= \int_{]0,1[\,\times\,]0,+\infty[} h(ux)\frac{1}{\mathrm{E}(X)}xf(x)\,\mathrm{d}u\,\mathrm{d}x \\ &= \frac{1}{\mathrm{E}(X)}\int_0^\infty \left(\int_0^1 h(ux)x\mathrm{d}u\right)f(x)\mathrm{d}x \\ &= \frac{1}{\mathrm{E}(X)}\int_0^\infty \left(\int_0^x h(y)\mathrm{d}y\right)f(x)\mathrm{d}x\end{aligned}$$

où on a utilisé le théorème de Fubini–Tonelli, puis un changement de variable $y = ux$ dans l'intégrale interne. En utilisant de nouveau le théorème de Fubini–Tonelli,

$$\begin{aligned}\int_0^\infty \left(\int_0^x h(y)\mathrm{d}y\right)f(x)\mathrm{d}x &= \int_{\mathbb{R}^2} h(y)f(x)\mathbb{1}_{\{(x,y),0<y<x<+\infty\}}\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_0^\infty \left(\int_y^\infty f(x)\mathrm{d}x\right)h(y)\mathrm{d}y = \int_0^\infty \mathrm{P}(X > y)h(y)\mathrm{d}y.\end{aligned}$$

On en conclut que

$$\mathrm{E}(h(Y)) = \int_0^\infty h(y)\, \frac{1}{\mathrm{E}(X)}\, \mathrm{P}(X > y)\mathrm{d}y\,.$$

Comme la fonction h est arbitraire (continue par morceaux et bornée), on obtient grâce à l'Observation 6.26 que Y est absolument continue de densité donnée par $f_Y(y) = \frac{1}{\mathrm{E}(X)}\,\mathrm{P}(X > y)\mathbb{1}_{]0,+\infty[}(y)$.

Exercice 6.42. Soient $X_1, X_2, \dots, X_n$ des variables aléatoires i.i.d., de loi $\mathrm{U}(0,1)$, et on définit

$$L_n := \min\{X_1, X_2, \dots, X_n\}\,, \qquad Z_n := n L_n\,.$$

Montrer que la fonction de répartition $F_{Z_n}(t)$ converge quand $n \to \infty$ vers une limite $F(t)$, qui est la fonction de répartition d'une variable aléatoire $\mathrm{Exp}(1)$.

Solution 6.42. Pour $x \in [0,1]$ on a $\mathrm{P}(L_n > x) = \mathrm{P}(X_1 > x)^n = (1-x)^n$, donc

$$F_{L_n}(x) = 1 - (1-x)^n\,,$$

alors que l'on a clairement $F_{L_n}(x) = 0$ si $x \le 0$ et $F_{L_n}(x) = 1$ si $x \ge 1$.

Pour tout $t \ge 0$ fixé, on a $t/n \in [0,1]$ pour n suffisamment grand, donc

$$F_{Z_n}(t) = \mathrm{P}(Z_n \le t) = \mathrm{P}(L_n \le \tfrac{t}{n}) = 1 - (1 - \tfrac{t}{n})^n \xrightarrow[n\to\infty]{} 1 - \mathrm{e}^{-t}\,.$$

Pour $t < 0$ on a $F_{Z_n}(t) = 0$ pour tout $n \in \mathbb{N}^*$ et donc $\lim_{n\to\infty} F_{Z_n}(t) = 0$. Au final,

$$F(t) := \lim_{n\to\infty} F_{Z_n}(t) = \begin{cases} 0 & \text{si } t < 0\,, \\ 1 - \mathrm{e}^{-t} & \text{si } t \ge 0\,, \end{cases}$$

qui est bien la fonction de répartition de la loi $\mathrm{Exp}(1)$, voir (6.44).

Exercice 6.43. Soit X une variable aléatoire de densité donné par

$$f(x) = \Big(1 + \frac{x}{\beta}\Big)^{-(1+\beta)} \mathbb{1}_{\mathbb{R}_+}(x)\,,$$

avec $\beta \in\,]1, +\infty[$.

(i) Calculer l'espérance $\mu := \mathrm{E}(X)$.

(ii) Calculer la fonction de répartition $F_X(t)$ de X, puis déterminer les quantiles q_α d'ordre $\alpha \in\,]0,1[$ de X.

(iii) Déterminer les limites de μ quand $\beta \to 1$ et quand $\beta \to \infty$; de même pour q_α.

Solution 6.43. (i) D'après la formule de transfert, on a

$$\begin{aligned} \mathrm{E}\Big(1 + \frac{X}{\beta}\Big) &= \int_0^{+\infty} \Big(1 + \frac{x}{\beta}\Big) f_X(x)\mathrm{d}x = \int_0^{+\infty} \Big(1 + \frac{x}{\beta}\Big)^{-\beta} \mathrm{d}x \\ &= \left[\frac{\beta}{1-\beta}\Big(1 + \frac{x}{\beta}\Big)^{1-\beta}\right]_0^\infty = \frac{\beta}{\beta - 1}\,, \end{aligned}$$

où on a utilisé le fait que $\lim_{x\to+\infty} \big(1 + \frac{x}{\beta}\big)^{1-\beta} = 0$ pour $\beta > 1$. Par linéarité de l'espérance, on a $1 + \frac{1}{\beta}\,\mathrm{E}(X) = \frac{\beta}{\beta-1}$, donc on obtient $\mu = \mathrm{E}(X) = \frac{2\beta-1}{\beta-1}$.

(ii) La fonction de répartition de X est donnée par $F_X(t) = 0$ pour $t \leq 0$ (parce que X est positive) et pour $t > 0$

$$F_X(t) = \int_0^t \left(1+\frac{x}{\beta}\right)^{-(1+\beta)} \mathrm{d}x = \left[\left(1+\frac{x}{\beta}\right)^{-\beta}\right]_0^t = 1 - \left(1+\frac{t}{\beta}\right)^{-\beta}.$$

La fonction de répartition, étant continue et strictement croissante, c'est une bijection de $]0,+\infty[$ vers $]0,1[$. En se rappelant de la Proposition 3.105 et du Lemme 6.33, on a que le quantile q_α de X d'ordre α est $F_X^{-1}(\alpha)$, c'est-à-dire est la solution de

$$\begin{aligned} F_X(q_\alpha) = \alpha \quad &\Longleftrightarrow \quad \left(1+\frac{q_\alpha}{\beta}\right)^{-\beta} = 1-\alpha \\ &\Longleftrightarrow \quad q_\alpha = \beta\left((1-\alpha)^{-1/\beta} - 1\right). \end{aligned}$$

(iii) Comme $\mu = \frac{2\beta-1}{\beta-1}$, on a facilement

$$\lim_{\beta\to 1} \mu = +\infty, \qquad \lim_{\beta\to+\infty} \mu = 1.$$

En utilisant la formule $q_\alpha = \beta\left((1-\alpha)^{-1/\beta} - 1\right)$, on obtient que

$$\lim_{\beta\to 1} q_\alpha = (1-\alpha)^{-1} - 1 = \frac{\alpha}{1-\alpha}, \qquad \lim_{\beta\to+\infty} q_\alpha = -\log(1-\alpha),$$

où pour la dernière égalité on a utilisé que

$$(1-\alpha)^{-1/\beta} - 1 = \mathrm{e}^{-\frac{1}{\beta}\log(1-\alpha)} - 1 \sim -\frac{1}{\beta}\log(1-\alpha) \qquad \text{quand } \beta \to +\infty.$$

Exercice 6.44. Soit $\varphi : \mathbb{R}_+ \to [0,1[$ une fonction continue, croissante, telle que $\varphi(0) = 0$, $\lim_{t\to+\infty} \varphi(t) = 1$, et dont le comportement asymptotique quand $x \downarrow 0$ est donné par

$$\varphi(x) = \alpha x^k + o(x^k), \qquad \text{avec } \alpha, k > 0.$$

(i) Montrer que la fonction $F(t) := (1-\varphi(1/t))\mathbb{1}_{]0,\infty[}(t)$ est une fonction de répartition, c'est-à-dire qu'elle vérifie les propriétés (5.5).

(ii) Considérons une suite $(X_n)_{n\geq 1}$ de variables aléatoires i.i.d. dont la fonction de répartition est la fonction F définie plus haut. On pose

$$Y_n := \frac{\max(X_1, X_2, \ldots, X_n)}{n^{1/k}}.$$

Montrer que, pour tout $y > 0$, on a $\lim_{n\to+\infty} \mathrm{P}(Y_n \leq y) = \mathrm{e}^{-\alpha/y^k}$.

Solution 6.44. (i) Comme φ est croissante, la fonction $]0,\infty[\ \ni t \mapsto \varphi(1/t)$ est décroissante, donc la fonction $F(t) = 1-\varphi(1/t)$ est croissante sur la demi-droite $]0,\infty[$; elle est de plus continue, en tant que composée de fonctions continues. Étant donné que $F(0^+) = \lim_{t\downarrow 0} F(t) = 1 - \lim_{s\uparrow+\infty} \varphi(s) = 1-1 = 0$ et comme $F(t) = 0$ pour $t \leq 0$, on en déduit que F est croissante et continue sur $\mathbb{R}$. Enfin, on a clairement $\lim_{t\to-\infty} F(t) = 0$ alors que $\lim_{t\to+\infty} F(t) = 1 - \lim_{s\downarrow 0} \varphi(s) = 1-0 = 1$, donc toutes les propriétés (5.5) d'une fonction de répartition sont vérifiées.

(ii) Pour $y > 0$ on a

$$\begin{aligned}
\mathrm{P}(Y_n \leq y) &= \mathrm{P}(X_1 \leq n^{1/k}y, X_2 \leq n^{1/k}y, \ldots, X_n \leq n^{1/k}y) = \mathrm{P}(X_1 \leq n^{1/k}y)^n \\
&= F_X(n^{1/k}y)^n = \left(1 - \varphi\left(\frac{1}{n^{1/k}y}\right)\right)^n = \left(1 - \frac{\alpha}{ny^k} + o\left(\frac{1}{n}\right)\right)^n \\
&= \exp\left\{n \log\left(1 - \frac{\alpha}{ny^k} + o\left(\frac{1}{n}\right)\right)\right\} \xrightarrow{n\to\infty} \exp\left\{-\frac{\alpha}{y^k}\right\},
\end{aligned}$$

où on a utilisé le développement limité $\log(1+x) = x + o(x)$ quand $x \to 0$.

Exercice 6.45. Un lanceur de javelot effectue n lancers, $n \in \mathbb{N}^*$. On note X_i la distance obtenue lors du i-ème lancer, et on suppose que $X_1, \ldots, X_n$ sont des variables aléatoires i.i.d. avec $X_n \sim \mathrm{Exp}(\lambda)$, où $\lambda \in \,]0, \infty[$. On note M_n la distance maximale à laquelle le javelot a été lancé.

(i) Soit $W_n := \frac{M_n}{\log n}$ et F_{W_n} sa fonction de répartition. Montrer que

$$\lim_{n\to+\infty} F_{W_n}(x) = \begin{cases} 0 & \text{si } x < 1/\lambda\,, \\ \mathrm{e}^{-1} & \text{si } x = 1/\lambda\,, \\ 1 & \text{si } x > 1/\lambda\,. \end{cases}$$

(ii) En déduire que pour tout $\varepsilon > 0$

$$\lim_{n\to+\infty} \mathrm{P}\left(\left|\frac{M_n}{\log n} - \frac{1}{\lambda}\right| > \varepsilon\right) = 0.$$

(iii) On définit maintenant $Z_n := M_n - \frac{1}{\lambda}\log n$, et on note $F_{Z_n}(t)$ sa fonction de répartition. Montrer que la limite $F(t) := \lim_{n\to\infty} F_{Z_n}(t)$ existe, et la déterminer, pour tout $t \in \mathbb{R}$. Observer que F est la fonction de répartition d'une variable aléatoire à densité, dont on déterminera la densité.

Solution 6.45. (i) Comme $f_{X_1}(t) = \lambda \mathrm{e}^{-\lambda t} \mathbb{1}_{]0,\infty[}(t)$, on obtient

$$F_{X_1}(x) = \mathrm{P}(X_1 \leq x) = \int_{-\infty}^{x} f_{X_1}(t)\,\mathrm{d}t = \left(1 - \mathrm{e}^{-\lambda x}\right) \mathbb{1}_{]0,\infty[}(x)\,. \tag{S6.17}$$

Par conséquent $F_{W_n}(w) = 0$ si $w \leq 0$, alors que pour $w > 0$, en utilisant le développement limité $\log(1+x) = x + o(x)$ quand $x \to 0$, on obtient

$$\begin{aligned}
F_{W_n}(w) &= \mathrm{P}(W_n \leq w) = \mathrm{P}(M_n \leq w \log n) = \mathrm{P}(X_1 \leq w\log n, \ldots, X_n \leq w \log n) \\
&= \mathrm{P}(X_1 \leq w \log n)^n = F_{X_1}(w \log n)^n = \left(1 - \mathrm{e}^{-\lambda w \log n}\right)^n \\
&= \left(1 - \frac{1}{n^{\lambda w}}\right)^n = \exp\left\{n \log\left(1 - \frac{1}{n^{\lambda w}}\right)\right\} \\
&= \exp\left\{n\left(-\frac{1}{n^{\lambda w}} + o\left(\frac{1}{n^{\lambda w}}\right)\right)\right\} \xrightarrow{n\to\infty} \begin{cases} 0 & \text{si } w < \frac{1}{\lambda}\,, \\ \mathrm{e}^{-1} & \text{si } w = \frac{1}{\lambda}\,, \\ 1 & \text{si } w > \frac{1}{\lambda}\,. \end{cases}
\end{aligned}$$

(ii) On peut écrire

$$\begin{aligned}
\mathrm{P}\left(\left|\frac{M_n}{\log n}-\frac{1}{\lambda}\right|>\varepsilon\right) &= \mathrm{P}\left(\frac{M_n}{\log n}>\frac{1}{\lambda}+\varepsilon\right)+\mathrm{P}\left(\frac{M_n}{\log n}<\frac{1}{\lambda}-\varepsilon\right)\\
&= 1-F_{W_n}\left(\frac{1}{\lambda}+\varepsilon\right)+F_{W_n}\left(\frac{1}{\lambda}-\varepsilon\right)\\
&\xrightarrow{n\to\infty} 1-1+0=0\,.
\end{aligned}$$

(iii) En utilisant (S6.17) on a

$$\begin{aligned}
F_{Z_n}(t) &= \mathrm{P}(Z_n\le t) = \mathrm{P}\left(M_n\le\frac{1}{\lambda}\log n+t\right) = F_{X_1}\left(\frac{1}{\lambda}\log n+t\right)^n\\
&= \begin{cases}\left(1-\mathrm{e}^{-\lambda(\frac{1}{\lambda}\log n+t)}\right)^n = \left(1-\frac{1}{n}\mathrm{e}^{-\lambda t}\right)^n & \text{si } t>-\frac{1}{\lambda}\log n\,,\\ 0 & \text{si } t\le -\frac{1}{\lambda}\log n\,.\end{cases}
\end{aligned}$$

Pour tout $t\in\mathbb{R}$ fixé, on a $t>-\frac{1}{\lambda}\log n$ pour n suffisamment grand. Comme on a $\lim_{n\to\infty}(1-\frac{c}{n})^n=\mathrm{e}^{-c}$, par exemple grâce au développement de Taylor du logarithme que l'on a rappelé plus haut, on trouve la limite

$$F(t) := \lim_{n\to\infty}F_{Z_n}(t) = \lim_{n\to\infty}\left(1-\frac{1}{n}\mathrm{e}^{-\lambda t}\right)^n = \mathrm{e}^{-\mathrm{e}^{-\lambda t}}\,,$$

pour tout $t\in\mathbb{R}$.
Observons que la fonction F que l'on vient de déterminer est croissante et continue sur tout $\mathbb{R}$ et de plus $\lim_{t\to-\infty}F(t)=0$, $\lim_{t\to+\infty}F(t)=1$. Les propriétés (5.5) sont donc vérifiées, donc Fest la fonction de répartition d'une variable aléatoire réelle. De plus, F étant de classe $\mathcal{C}^1$ (même $\mathcal{C}^\infty$), il découle de la Proposition 6.17 que la variable aléatoire dont F est la fonction de répartition est absolument continue, de densité donnée par

$$f(t) = F'(t) = \lambda\,\mathrm{e}^{-\lambda t}\,\mathrm{e}^{-\mathrm{e}^{-\lambda t}}\,.$$

Exercice 6.46. Soit $(X_i)_{i\ge1}$ des variables aléatoires i.i.d. de loi $\mathrm{Exp}(1)$. Pour tout $n\ge1$ on pose $Y_n:=\prod_{i=1}^n X_i$.

(i) Montrer que $\mathrm{E}(Y_n)=1$ pour tout $n\in\mathbb{N}^*$.

(ii) Montrer que $\mathrm{E}(\sqrt{Y_n})=\left(\frac{\sqrt{\pi}}{2}\right)^n$ pour tout $n\in\mathbb{N}^*$.

[*Sugg.* On pourra utiliser l'intégrale de Gauss $\int_{-\infty}^{\infty}e^{-x^2}\mathrm{d}x=\sqrt{\pi}$.]

(iii) En appliquant l'inégalité de Markov, montrer que pour tout $n\in\mathbb{N}^*$ et tout $\theta>0$

$$\mathrm{P}\left(Y_n>\theta^n\right)\le\left(\frac{\sqrt{\pi}}{2\sqrt{\theta}}\right)^n.$$

En conclure que si $\theta\in\,]\frac{\pi}{4},1[$, on a $\lim_{n\to\infty}\mathrm{P}(Y_n>\theta^n)=0$.

(iv) En utilisant la question précédente, montrer que si $\theta\in\,]\frac{\pi}{4},1[$ alors

$$\lim_{N\to\infty}\mathrm{P}\Big(\bigcup_{n\ge N}\{Y_n>\theta^n\}\Big)=0\,.$$

En utilisant la continuité des probabilités par le haut (Proposition 1.24), en déduire que $\mathrm{P}\left(\bigcap_{N\ge1}\bigcup_{n\ge N}\{Y_n>\theta^n\}\right)=0$. En passant au complémentaire, conclure que l'événement « la suite $(Y_n)_{n\ge1}$ converge vers 0 » est de probabilité 1.

Solution 6.46. (i) Grâce à l'indépendance des variables aléatoires $X_1, \ldots, X_n$, on a (voir la Proposition 3.72)

$$\mathrm{E}(Y_n) = \prod_{i=1}^{n} \mathrm{E}(X_i) = 1\,,$$

car $\mathrm{E}(X_i) = 1$ pour tout i.

(ii) De même, on a

$$\mathrm{E}(\sqrt{Y_n}) = \mathrm{E}\Big(\prod_{i=1}^{n} \sqrt{X_i}\Big) = \prod_{i=1}^{n} \mathrm{E}\big(\sqrt{X_i}\big)\,.$$

Il suffit donc de calculer, grâce à la formule de transfert,

$$\mathrm{E}(\sqrt{X_i}) = \int_0^\infty \sqrt{x}\,\mathrm{e}^{-x}\mathrm{d}x = \int_0^\infty 2u^2\mathrm{e}^{-u^2}\mathrm{d}u\,,$$

où on a fait un changement de variable $u = \sqrt{x}$ ($x = u^2$). Avec une intégration par parties, on trouve

$$\mathrm{E}(\sqrt{X_i}) = \Big[-u\mathrm{e}^{-u^2}\Big]_0^\infty + \int_0^\infty \mathrm{e}^{-u^2}\mathrm{d}u = \frac{1}{2}\sqrt{\pi}\,,$$

car par symétrie $\int_0^\infty \mathrm{e}^{-u^2}\mathrm{d}u = \frac{1}{2}\int_{-\infty}^\infty \mathrm{e}^{-u^2}\mathrm{d}u = \frac{1}{2}\sqrt{\pi}$. Autrement, on aurait pu observer que $\int_0^\infty \sqrt{x}\,\mathrm{e}^{-x}\mathrm{d}x = \Gamma(3/2)$, dont la valeur est connue, voir (6.38).

Pour résumer, on a montré que $\mathrm{E}(\sqrt{Y_n}) = \prod_{i=1}^n \mathrm{E}(\sqrt{X_i}) = \big(\frac{\sqrt{\pi}}{2}\big)^n$.

(iii) En utilisant l'égalité d'événements $\{Y_n > \theta^n\} = \{\sqrt{Y_n} > (\sqrt{\theta})^n\}$ et en appliquant l'inégalité de Markov à la variable aléatoire positive $\sqrt{Y_n}$, on obtient

$$\mathrm{P}(Y_n > \theta^n) = \mathrm{P}\big(\sqrt{Y_n} > (\sqrt{\theta})^n\big) \leq \frac{1}{(\sqrt{\theta})^n}\mathrm{E}(\sqrt{Y_n}) = \left(\frac{1}{\sqrt{\theta}}\,\frac{\sqrt{\pi}}{2}\right)^n.$$

Pour $\theta > \frac{\pi}{4}$ on a $\frac{\sqrt{\pi}}{2\sqrt{\theta}} < 1$, donc l'inégalité ci-dessus nous permet de conclure que $\lim_{n\to+\infty} \mathrm{P}(Y_n > \theta^n) = 0$.

(iv) Par sous-additivité des probabilités, pour tout $N \geq 1$ on a

$$\mathrm{P}\Big(\bigcup_{n\geq N}\{Y_n > \theta^n\}\Big) \leq \sum_{n\geq N} \mathrm{P}(Y_n > \theta^n) \leq \sum_{n\geq N}\left(\frac{\sqrt{\pi}}{2\sqrt{\theta}}\right)^n,$$

où on a utilisé le point précédent. Comme $\frac{\sqrt{\pi}}{2\sqrt{\theta}} \in\,]0,1[$, en utilisant la série géométrique (0.5), on obtient

$$\sum_{n\geq N}\left(\frac{\sqrt{\pi}}{2\sqrt{\theta}}\right)^n = \left(\frac{\sqrt{\pi}}{2\sqrt{\theta}}\right)^N \sum_{k=0}^{+\infty}\left(\frac{\sqrt{\pi}}{2\sqrt{\theta}}\right)^k = \left(\frac{\sqrt{\pi}}{2\sqrt{\theta}}\right)^N \times \frac{1}{1-\frac{\sqrt{\pi}}{2\sqrt{\theta}}}\,,$$

ce qui tend vers 0 quand $N \to +\infty$, parce que $\frac{\sqrt{\pi}}{2\sqrt{\theta}} < 1$.

Notons que la suite d'événements $A_N := \bigcup_{n\geq N}\{Y_n > \theta^n\}$ est décroissante pour l'inclusion : d'après la continuité par le haut des probabilités (Proposition 1.24) on a $\mathrm{P}(\bigcap_{N\geq 1} A_N) = \lim_{N\to+\infty}\mathrm{P}(A_N)$. D'après ce qui précède, on a donc $\lim_{N\to+\infty}\mathrm{P}(A_N) = 0$, c'est-à-dire

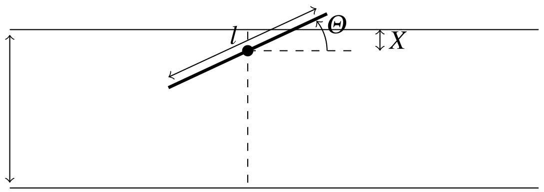
$$\mathrm{P}\Big(\bigcap_{N\geq 1}\bigcup_{n\geq N}\{Y_n > \theta^n\}\Big) = 0\,.$$

En passant au complémentaire,

$$\mathrm{P}\Big(\bigcup_{N\geq 1}\bigcap_{n\geq N}\{Y_n \leq \theta^n\}\Big) = 1\,.$$

Cela démontre que, avec probabilité 1, il existe un $N \geq 1$ tel que pour tout $n \geq N$ on a $Y_n \leq \theta^n$: cela signifie qu'avec probabilité 1, la suite $(Y_n)_{n\geq 1}$ est majoré par θ^n à partir d'un certain rang. Comme $\theta \in \,]\frac{\pi}{4}, 1[$, on a $\theta^n \to 0$. En particulier, l'événement « la suite $(Y_n)_{n\geq 1}$ est majorée par θ^n à partir d'un certain rang » est inclus dans l'événement « la suit $(Y_n)_{n\geq 1}$ tend vers 0 » : le premier événement étant de probabilité 1, le deuxième est lui aussi de probabilité 1.

Exercice 6.47 (L'aiguille de Buffon *). Voici une expérience réalisée par le comte Buffon, un naturaliste, mathématicien et philosophe du 18-ème siècle : on lance une aiguille sur un parquet composé de planches parallèles, et on se demande quelle est la probabilité que l'aiguille tombe à cheval sur (au moins) une rainure du parquet. On suppose que les rainures du parquet sont à une distance d les unes des autres. On note X la distance du centre de l'aiguille à la rainure la plus proche, et Θ l'angle que fait la l'aiguille avec la direction des rainures, voir la figure ci-dessous.



L'hypothèse que l'on fait est que X est uniforme sur l'intervalle $[0, d/2]$ (si la distance est plus grande que $d/2$ alors l'aiguille est plus proche d'une autre rainure), que Θ est uniforme sur l'intervalle $[0, \pi/2]$ (par symétrie du problème, cela suffit), et que X et Θ sont indépendantes. L'événement que l'aiguille soit à cheval sur la rainure du parquet correspond donc à l'événement $\{X < \frac{l}{2}\sin\Theta\}$, $\frac{l}{2}$ étant la longueur de la « demi-aiguille ».

(i) Donner la densité jointe de (X, Θ).

(ii) On suppose que $l \leq d$. Montrer que $\mathrm{P}(X < \frac{l}{2}\sin\Theta) = \frac{2l}{\pi d}$.

En répétant cette expérience un très grand nombre de fois, et en notant la proportion de fois où l'aiguille tombe à cheval sur une rainure, Buffon en déduit une valeur approchée de cette probabilité, donc de π ! Il s'agit de la méthode de Monte-Carlo*, sur laquelle nous reviendrons dans le Paragraphe 7.1.2.*

Solution 6.47. (i) Les variables aléatoires étant indépendantes, la densité jointe du vecteur (X, Θ) est donnée par

$$f_{X,\Theta}(x,t) = f_X(x) f_\Theta(t) = \frac{4}{\pi d}\mathbb{1}_{[0,d/2]}(x)\mathbb{1}_{[0,\pi/2]}(t)\,,$$

d'après la Proposition 6.50.

(ii) Calculons la probabilité de l'événement $\{X < \frac{l}{2}\sin\Theta\}$: en utilisant la densité jointe $f_{X,\Theta}(x,t)$ du vecteur (X, Θ), on obtient (en utilisant (6.63))

$$\begin{aligned}\mathrm{P}(X < \tfrac{l}{2}\sin\Theta) &= \int_{(x,t)\,:\,x<\frac{1}{2}l\sin t} f_{X,\Theta}(x,t)\mathrm{d}x\mathrm{d}t = \int_0^{\pi/2}\left(\int_0^{\frac{1}{2}l\sin t}\frac{4}{\pi d}\mathrm{d}x\right)\mathrm{d}t\\ &= \frac{2l}{\pi d}\int_0^{\pi/2}\sin t\,\mathrm{d}t = \frac{2l}{\pi d},\end{aligned}$$

où pour la deuxième identité on a utilisé que $\frac{1}{2}l\sin t \le d/2$ pour tout $t \in [0,\pi/2]$.

Exercice 6.48 (*). On place deux points de manière uniforme et indépendante sur un intervalle $[0,L]$. Cela découpe l'intervalle en trois segments. Quelle est la probabilité que l'on puisse former un triangle avec ces trois segments ?

Solution 6.48. Notons A l'événement : « il est possible de former un triangle avec les trois segments ». Tout d'abord, observons que A^c correspond à l'événement « il existe un segment de longueur supérieure à $L/2$ ». Si on note X,Y les deux points choisis uniformément et indépendamment dans l'intervalle $[0,L]$, on a donc

$$\begin{aligned}\mathrm{P}(A^c) = \mathrm{P}\big(X\in[0,L/2], Y\in[0,L/2]\big) + \mathrm{P}\big(X\in[L/2,L], Y\in[L/2,L]\big)\\ + \mathrm{P}\big(X\in[0,L/2], Y\ge X+L/2\big) + \mathrm{P}\big(X\in[L/2,L], X\ge Y+L/2\big).\end{aligned}$$

Comme X et Y sont des variables aléatoires indépendantes et uniformes continues sur $[0,L]$, les deux première probabilités sont égales, et valent

$$\mathrm{P}\big(X\in[0,L/2], Y\in[0,L/2]\big) = \mathrm{P}\big(X\in[0,L/2]\big)\,\mathrm{P}\big(Y\in[0,L/2]\big) = \left(\frac{L/2}{L}\right)^2 = \frac{1}{4}.$$

En utilisant que la densité jointe de (X,Y) est $\frac{1}{L^2}\mathbb{1}_{[0,L]}(x)\mathbb{1}_{[0,L]}(y)$ (par indépendance, voir la Proposition 6.50) et en se rappelant de (6.63), on a

$$\begin{aligned}&\mathrm{P}\big(X\in[0,L/2], Y\ge X+L/2\big)\\ &\quad= \int_{\mathbb{R}^2}\frac{1}{L^2}\mathbb{1}_{[0,L]}(x)\mathbb{1}_{[0,L]}(y)\mathbb{1}_{\{(x,y),x\in[0,L/2],y\in[L/2+x,L]\}}(x,y)\,\mathrm{d}x\,\mathrm{d}y\\ &\quad= \int_0^{L/2}\int_{x+L/2}^{L}\frac{1}{L^2}\,\mathrm{d}x\,\mathrm{d}y = \frac{1}{L^2}\int_0^{L/2}(L/2-x)\mathrm{d}x = \frac{1}{L^2}\times\frac{1}{2}(L/2)^2 = \frac{1}{8},\end{aligned}$$

et de même pour la dernière probabilité. Au final, on obtient $\mathrm{P}(A^c) = \frac{1}{4}+\frac{1}{4}+\frac{1}{8}+\frac{1}{8} = \frac{3}{4}$, donc $\mathrm{P}(A) = \frac{1}{4}$.

Exercice 6.49 (*). Soit Θ une variable aléatoire de loi $\mathrm{U}(0,2\pi)$. On pose $X=\cos(\Theta)$ et $Y=\sin(\Theta)$.

(i) Calculer l'espérance de X et Y, et la matrice de covariance du vecteur (X,Y).

(ii) Le vecteur aléatoire (X,Y) est-il absolument continu ?

Solution 6.49. Rappelons que la densité de Θ est donnée par $\frac{1}{2\pi}\mathbb{1}_{]0,2\pi[}(x)$.

(i) D'après la Proposition 6.21 (la formule de transfert (6.20)), on a

$$\mathrm{E}(X) = \mathrm{E}(\cos(\Theta)) = \int_0^{2\pi}\cos(t)\frac{1}{2\pi}\mathrm{d}t = \frac{1}{2\pi}\big[\sin(t)\big]_0^{2\pi} = 0.$$

De la même manière

$$\mathrm{E}(Y) = \mathrm{E}(\sin(\Theta)) = \int_0^{2\pi} \sin(t)\frac{1}{2\pi}\mathrm{d}t = \frac{1}{2\pi}\big[-\cos(t)\big]_0^{2\pi} = 0\,.$$

Calculons maintenant la matrice de covariance. Commençons par $\mathrm{Cov}(X,X) = \mathrm{Var}(X) = \mathrm{E}(X^2)$ (car $\mathrm{E}(X) = 0$) et $\mathrm{Cov}(Y,Y) = \mathrm{Var}(Y) = \mathrm{E}(Y^2)$ (car $\mathrm{E}(Y) = 0$). Notons que $X^2 = \sin(\Theta)^2 = \frac{1}{2} + \frac{1}{2}\cos(2\Theta)$: par linéarité de l'espérance on a $\mathrm{E}(X^2) = \frac{1}{2} + \frac{1}{2}\mathrm{E}(\cos(2\Theta))$. Ensuite, grâce à la formule de transfert (6.20), on obtient

$$\mathrm{E}(\cos(2\Theta)) = \int_0^{2\pi} \cos(2t)\frac{1}{2\pi}\mathrm{d}t = \frac{1}{2\pi}\big[\tfrac{1}{2}\sin(2t)\big]_0^{2\pi} = 0\,.$$

On en conclut que $\mathrm{Var}(X) = \mathrm{E}(X^2) = \frac{1}{2}$.
Un calcul analogue donne $\mathrm{E}(Y^2) = \frac{1}{2}$. Autrement, on aurait pu observer que $X^2+Y^2 = \cos(\Theta)^2 + \sin(\Theta)^2 = 1$ de sorte que $Y^2 = 1 - X^2$: par linéarité de l'espérance, on trouve $\mathrm{E}(Y^2) = 1 - \mathrm{E}(X^2) = \frac{1}{2}$.
Calculons maintenant la covariance $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X) = \mathrm{E}(XY)$ (car $\mathrm{E}(X) = \mathrm{E}(Y) = 0$). Toujours grâce à la formule de transfert (6.20), on a

$$\begin{aligned}\mathrm{E}(XY) = \mathrm{E}(\cos(\Theta)\sin(\Theta)) &= \int_0^{2\pi} \cos(t)\sin(t)\frac{1}{2\pi}\mathrm{d}t \\ &= \frac{1}{2\pi}\left[\frac{1}{2}\cos(t)^2\right]_0^{2\pi} = 0\,.\end{aligned}$$

On en conclut que la matrice de covariance du vecteur (X,Y) est $\frac{1}{2}I$, où I est la matrice identité.

(ii) Le vecteur (X,Y) n'est pas à densité. En effet, on a déjà remarqué que $X^2+Y^2 = 1$: si on définit $C_1 := \{(x,y), x^2+y^2 = 1\}$, on a donc

$$\mathrm{P}((X,Y) \in C_1) = 1\,.$$

Mais si (X,Y) était à densité, de densité $f_{X,Y}$, comme C_1 est de mesure 2-dimensionnelle nulle, en appliquant la relation (6.63) on devrait avoir

$$\mathrm{P}((X,Y) \in C_1) = \int_{C_1} f_{X,Y}(x,y)\mathrm{d}x\mathrm{d}y = 0\,,$$

ce qui aboutit à une contradiction.

Exercice 6.50 (*). Un signal est transmis à un instant aléatoire X. On allume le récepteur à un instant aléatoire Y, et on laisse le récepteur allumé pendant un intervalle de temps aléatoire Z. On suppose que X,Y,Z sont des variables aléatoires indépendantes avec $X \sim \mathrm{U}(0,2)$ et $Y,Z \sim \mathrm{U}(0,1)$. Quelle est la probabilité que le signal soit reçu ?

Solution 6.50. Le récepteur reste allumé durant l'intervalle de temps (aléatoire) $[Y, Y+Z]$, il faut donc calculer

$$\mathrm{P}(X \in [Y,Y+Z]) = \mathrm{P}(X \geq Y, X \leq Y+Z) = \mathrm{P}((X,Y,Z) \in C)\,,$$

où

$$C := \{(x,y,z) \in \mathbb{R}^3 :\ x \geq y,\ x \leq y+z\} = \{(x,y,z) \in \mathbb{R}^3 :\ x \in [y,y+z]\}\,.$$

La densité du vecteur aléatoire (X,Y,Z) est

$$f_{X,Y,Z}(x,y,z) = f_X(x)\,f_Y(y)\,f_Z(z) = \frac{1}{2}\,\mathbb{1}_{[0,2]}(x)\,\mathbb{1}_{[0,1]}(y)\,\mathbb{1}_{[0,1]}(z)\,.$$

En appliquant le théorème de Fubini–Tonelli, la probabilité demandée est égale à

$$\begin{aligned}
\mathrm{P}(X \in [Y, Y+Z]) &= \int_{\mathbb{R}^3} \mathbb{1}_C(x,y,z)\,f_{X,Y,Z}(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \\
&= \frac{1}{2}\int_0^1 \left(\int_0^1 \left(\int_0^2 \mathbb{1}_{[y,y+z]}(x)\mathrm{d}x\right)\mathrm{d}y\right)\mathrm{d}z \\
&= \frac{1}{2}\int_0^1 \left(\int_0^1 z\,\mathrm{d}y\right)\mathrm{d}z = \frac{1}{2}\int_0^1 z\,\mathrm{d}z = \frac{1}{4}\,.
\end{aligned}$$

Exercice 6.51 (*). Soit (X,Y) un vecteur aléatoire bidimensionnel de densité

$$f_{X,Y}(x,y) = \frac{\alpha(\alpha+1)}{(1+x+y)^{2+\alpha}}\,\mathbb{1}_{[0,\infty[}(x)\,\mathbb{1}_{[0,\infty[}(y)\,,$$

où $\alpha \in\,]0,\infty[$ est une constante fixée.

(i) Sans faire de calculs, expliquer pourquoi les composantes X et Y ont la même densité.

(ii) Montrer que la fonction de répartition de X (et de Y) est donnée par

$$F_X(t) = \left(1 - \frac{1}{(1+t)^\alpha}\right)\mathbb{1}_{[0,\infty[}(t)\,.$$

(iii) Pour quelles valeurs de $p \in [1,\infty[$ a-t-on $X \in L^p$?

Solution 6.51. (i) La fonction $f_{X,Y}(x,y)$ est symétrique en x,y, donc les densités marginales de X et Y, obtenues en intégrant la fonction $f_{X,Y}$ par rapport à chacune des variables, sont égales. On a de plus

$$\begin{aligned}
f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x,y)\,\mathrm{d}y = \alpha\left[\frac{-1}{(1+x+y)^{1+\alpha}}\right]_{y=0}^{y=\infty} \mathbb{1}_{[0,\infty[}(x) \\
= \frac{\alpha}{(1+x)^{1+\alpha}}\,\mathbb{1}_{[0,\infty[}(x)\,.
\end{aligned}$$

(ii) Par une intégration directe, à partir de la densité f_X que l'on vient d'obtenir, on obtient la formule pour $F_X(x)$.

(iii) On a

$$\mathrm{E}(|X|^p) = \int_{\mathbb{R}} |x|^p\,f_X(x)\,\mathrm{d}x = \int_0^\infty \alpha\,\frac{x^p}{(1+x)^{1+\alpha}}\,\mathrm{d}x\,.$$

Comme la fonction intégrée est asymptotiquement équivalente à $\alpha\,\frac{1}{x^{1+\alpha-p}}$ quand $x \to +\infty$, l'intégrale est finie — donc $X \in L^p$ — si et seulement si $p < \alpha$.

Exercice 6.52 (*). Soit (X,Y) un un vecteur aléatoire à valeurs dans $\mathbb{R}^2$, de densité

$$f_{X,Y}(x,y) := \begin{cases} c\,\mathrm{e}^{-x} & \text{si } 0 < x < y < x+1\,, \\ 0 & \text{sinon}\,, \end{cases}$$

où $c \in \mathbb{R}$ est une constante bien choisie.

(i) Montrer que X est une variable aléatoire de loi $\mathrm{Exp}(1)$ et calculer la valeur de la constante c.

(ii) Montrer que $Z := \log(X)$ est une variable aléatoire à densité, et en déterminer la densité. Pour quelles valeurs de p a-t-on $Z \in L^p$?

(iii) Déterminer la densité de Y. Calculer $\mathrm{E}(\mathrm{e}^{X-Y})$.

Solution 6.52. (i) On a

$$f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x,y)\,\mathrm{d}y = \int_x^{x+1} c\,\mathrm{e}^{-x}\,\mathrm{d}y = c\,\mathrm{e}^{-x}\int_x^{x+1} 1\,\mathrm{d}y = c\,\mathrm{e}^{-x},$$

dont on déduit que $c = 1$ et $X \sim \mathrm{Exp}(1)$.

(ii) En se rappelant que $F_X(x) = 1-\mathrm{e}^{-x}$ pour $x \geq 0$, on obtient

$$F_Z(t) = \mathrm{P}(\log(X) \leq t) = \mathrm{P}(X \leq \mathrm{e}^t) = F_X(\mathrm{e}^t) = 1-\mathrm{e}^{-\mathrm{e}^t}.$$

Ainsi, F_Z est une fonction continue et $\mathcal{C}^1$ par morceaux, la variable aléatoire Z est absolument continue, de densité

$$f_Z(z) = F_Z'(z) = \mathrm{e}^{z-\mathrm{e}^z}.$$

On a

$$\mathrm{E}(|Z|^p) = \int_{\mathbb{R}} |z|^p f_Z(z)\,\mathrm{d}z < \infty,$$

pour tout $p > 0$: en effet, la fonction intégrée est intégrable en $-\infty$ parce que $f_Z(z) \leq \mathrm{e}^z$ et elle est intégrable en $+\infty$ parce que quand z est grand et positif on a $\mathrm{e}^z - z > z$, donc $f_Z(z) \leq \mathrm{e}^{-z}$.

(iii) Pour $y > 1$ on a

$$f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x,y)\,\mathrm{d}x = \int_{y-1}^{y} f_{X,Y}(x,y)\,\mathrm{d}x = \mathrm{e}^{-(y-1)} - \mathrm{e}^{-y} = (\mathrm{e}-1)\mathrm{e}^{-y}.$$

Pour $0 < y \leq 1$ on a

$$f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x,y)\,\mathrm{d}x = \int_0^y f_{X,Y}(x,y)\,\mathrm{d}x = 1-\mathrm{e}^{-y}.$$

Enfin, pour $y \leq 0$ on a $f_Y(y) = 0$.
De plus,

$$\begin{aligned}\mathrm{E}(\mathrm{e}^{X-Y}) &= \int_{\mathbb{R}^2} \mathrm{e}^{x-y} f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y = \int_0^{+\infty}\left(\int_x^{x+1} \mathrm{e}^{-y}\,\mathrm{d}y\right)\mathrm{d}x \\ &= \int_0^{+\infty} (\mathrm{e}^{-x}-\mathrm{e}^{-x-1})\mathrm{d}x = 1-\mathrm{e}^{-1}.\end{aligned}$$

Exercice 6.53 (*). Soit (X,Y) un vecteur aléatoire bidimensionnel de densité f donnée par

$$f(x,y) = c\,y\,\mathrm{e}^{-xy}\,\mathbb{1}_{[0,\infty[\times[0,2]}(x,y).$$

(i) Déterminer la valeur de $c \in \mathbb{R}$ pour que f soit effectivement une densité.

(ii) Déterminer les densités marginales de X et Y et reconnaître la loi de Y. Les variables aléatoires X et Y sont-elles indépendantes ?

(iii) Montrer que $V := \max(X,Y)$ est une variable aléatoire réelle à densité et en déterminer la densité.

(iv) On pose $U := X + Y$. Dire si U et V sont indépendantes.

[*Sugg.* Il n'est pas nécessaire de calculer la loi jointe de (U,V).]

Solution 6.53. (i) Pour $y > 0$ fixé, la fonction $h(x) := y\mathrm{e}^{-xy}\mathbb{1}_{[0,\infty[}(x)$ est la densité d'une variable aléatoire $\mathrm{Exp}(y)$, donc $\int_{\mathbb{R}} h(x)\,\mathrm{d}x = 1$. Par conséquent

$$\begin{aligned}\int_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y &= c\int_0^2 \left(\int_0^\infty y\mathrm{e}^{-xy}\,\mathrm{d}x\right)\mathrm{d}y = c\int_0^2 \left(\int_{\mathbb{R}} h(x)\,\mathrm{d}x\right)\mathrm{d}y \\ &= c\int_0^2 1\,\mathrm{d}y = 2c,\end{aligned}$$

dont on déduit que $c = \frac{1}{2}$.

(ii) Pour $x < 0$ on a clairement $f_X(x) = 0$. Pour $x > 0$ on a

$$\begin{aligned}f_X(x) &= \int_{\mathbb{R}} f(x,y)\,\mathrm{d}y = \frac{1}{2}\int_0^2 y\mathrm{e}^{-xy}\,\mathrm{d}y = \frac{1}{2}\left[y\,\frac{-\mathrm{e}^{-xy}}{x}\right]_{y=0}^{y=2} + \frac{1}{2}\int_0^2 \frac{\mathrm{e}^{-xy}}{x}\,\mathrm{d}y \\ &= -\frac{\mathrm{e}^{-2x}}{x} + \frac{1}{2}\frac{1}{x^2}(1-\mathrm{e}^{-2x}) = \frac{1-\mathrm{e}^{-2x}-2x\mathrm{e}^{-2x}}{2x^2}.\end{aligned}$$

(La valeur $f_X(0)$ n'est pas importante.)

Notons que f_X est continue sur $]0,\infty[$ et $f_X(x) \sim 1/(2x^2)$ quand $x \to +\infty$, alors que quand $x \downarrow 0$, en développant le numérateur on a

$$\begin{aligned}1-\mathrm{e}^{-2x}-2x\mathrm{e}^{-2x} &= 1-\left(1-2x+\frac{(2x)^2}{2}+O(x^3)\right)-2x(1-2x+O(x^2)) \\ &= 2x-2x^2+O(x^3)-2x+4x^2+O(x^3) = 2x^2+O(x^3),\end{aligned}$$

dont on déduit que $f_X(x) \to 1$ quand $x \downarrow 0$.

Passons à Y. On a $f_Y(y) = 0$ pour $y \notin [0,2]$, alors que pour $y \in]0,2]$,

$$f_Y(y) = \int_{\mathbb{R}} f(x,y)\,\mathrm{d}x = \frac{1}{2}\int_0^\infty y\mathrm{e}^{-xy}\,\mathrm{d}x = \frac{1}{2}.$$

On a donc $f_Y(y) = \frac{1}{2}\mathbb{1}_{]0,2]}(y)$, c'est-à-dire $Y \sim \mathrm{U}(0,2)$.

Les variables aléatoires X et Y ne sont pas indépendantes parce que la densité jointe $f(x,y)$ ne coïncide pas avec le produit des densités marginales $g(x,y) := f_X(x)\,f_Y(y)$ (hors d'un ensemble de $(x,y) \in \mathbb{R}^2$ de mesure nulle). Pour le montrer précisément, notons que les fonctions $f(x,y)$ et $g(x,y)$ sont toutes les deux continues sur l'ouvert $]0,\infty[\times]0,2[$, donc il suffit de montrer qu'il existe un point dans cet ouvert où elles diffèrent (en effet, dans ce cas, elles diffèrent alors au voisinage de ce point). On a $f(x,1) = \frac{1}{2}\mathrm{e}^{-x}$ et

$$g(x,1) = f_X(x)\,f_Y(1) = \frac{1}{2}f_X(x) \sim \frac{1}{4x^2} \qquad \text{quand } x \to +\infty.$$

Par conséquent $\lim_{x\to+\infty} f(x,1)/g(x,1) = 0$ et donc $f(x,1)/g(x,1) < \frac{1}{2}$ pour x suffisamment grand ; en particulier, pour de tels x on a $f(x,1) \neq g(x,1)$.

(iii) La fonction de répartition de V est donnée par $F_V(v) = 0$ si $v \leq 0$, et pour $v > 0$

$$\begin{aligned}
F_V(v) &= \mathrm{P}(V \le v) = \mathrm{P}(X \le v, Y \le v) = \int_{\mathbb{R}^2} \mathbb{1}_{\{x \le v,\ y \le v\}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \frac{1}{2}\int_0^{\min(2,v)} \left(\int_0^v y\mathrm{e}^{-xy}\,\mathrm{d}x\right)\mathrm{d}y = \frac{1}{2}\int_0^{\min(2,v)} \mathrm{P}(\mathrm{Exp}(y) \le v)\mathrm{d}y \\
&= \frac{1}{2}\int_0^{\min(2,v)} (1-\mathrm{e}^{-vy})\mathrm{d}y = \frac{\min(2,v)}{2} - \frac{1-\mathrm{e}^{-v\cdot\min(2,v)}}{2v}\,.
\end{aligned}$$

Au final,

$$F_V(v) = \begin{cases} 0 & \text{si } v \le 0\,, \\ \dfrac{v}{2} - \dfrac{1-\mathrm{e}^{-v^2}}{2v} & \text{si } 0 < v < 2\,, \\ 1 - \dfrac{1-\mathrm{e}^{-2v}}{2v} & \text{si } v \ge 2\,. \end{cases}$$

On vérifie facilement que F_V est continue et $\mathcal{C}^1$ par morceaux, donc V est absolument continue, de densité donnée par $f_V(v) = F_V'(v)$, c'est-à-dire

$$f_V(v) = \begin{cases} 0 & \text{si } v \le 0\,, \\ \dfrac{1}{2} + \dfrac{1-(2v^2+1)\,\mathrm{e}^{-v^2}}{2v^2} & \text{si } 0 < v < 2\,, \\ \dfrac{1-(1+2v)\,\mathrm{e}^{-2v}}{2v^2} & \text{si } v \ge 2\,. \end{cases}$$

Notons que pour $v > 2$ on a $f_V(v) = f_X(v)$, comme cela doit être le cas car $F_V(v) = \mathrm{P}(X \le v, Y \le v) = \mathrm{P}(X \le v) = F_X(v)$ si $v > 2$, vu que $Y \le 2$.

(iv) Les variables aléatoires U et V *ne sont pas* indépendantes. Intuitivement, connaître la valeur de $V := \max\{X,Y\}$ donne des informations sur la somme $U := X+Y$. Plus précisément, si $V \le 1$ on a $X \le 1$ et $Y \le 1$, donc forcément $U \le 2$; autrement dit, on a l'inclusion d'événements $\{V \le 1\} \subseteq \{U \le 2\}$ et donc

$$\mathrm{P}(V \le 1, U > 2) = 0\,.$$

Si on montre que $\mathrm{P}(V \le 1) > 0$ et $\mathrm{P}(U > 2) > 0$, on en déduira que $\mathrm{P}(V \le 1, U > 2) \neq \mathrm{P}(V \le 1)\,\mathrm{P}(U > 2)$ et donc, par définition, U et V ne sont pas indépendantes.

D'après la question précédente $\mathrm{P}(V \le 1) = F_V(1) = \frac{1}{2} - \frac{1-\mathrm{e}^{-1}}{2} = \frac{\mathrm{e}^{-1}}{2} > 0$. De plus, comme la densité de $f_X(x)$ de X est strictement positive pour $x > 0$, on a

$$\mathrm{P}(X > 2) = \int_2^\infty f_X(x)\,\mathrm{d}x > 0\,.$$

Comme $Y \ge 0$, on a $\mathrm{P}(U > 2) \ge \mathrm{P}(X > 2) > 0$.

Exercice 6.54 (*). Soit $Z := (X,Y)$ un vecteur aléatoire bidimensionnel de loi uniforme sur l'ensemble $C := \big([0,\frac{1}{2}] \times [0,\frac{1}{2}]\big) \cup \big([\frac{1}{2},1] \times [\frac{1}{2},1]\big)$. Déterminer les lois des variables aléatoires réelles X et Y. Sont-elles indépendantes ?

Solution 6.54. L'ensemble C est représenté dans la Figure S6.1.
Par hypothèse, (X,Y) est un vecteur aléatoire absolument continu de densité

$$f_{X,Y}(x,y) = 2\,\mathbb{1}_C(x,y)\,.$$

On en déduit que les variables aléatoires X et Y sont absolument continues de densités

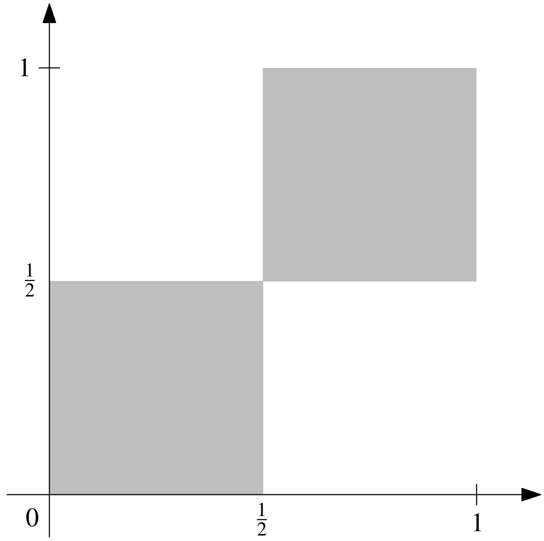


Fig. S6.1 L'ensemble $C = [0, \frac{1}{2}] \times [0, \frac{1}{2}]$ de l'Exercice 6.54.

$$f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x,y)\,\mathrm{d}y\,, \qquad f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x,y)\,\mathrm{d}x\,.$$

Si $x \in \,]0, \frac{1}{2}[$ on a $f_{X,Y}(x,y) = 2\mathbb{1}_{[0,\frac{1}{2}]}(y)$, donc

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

De même, si $x \in \,]\frac{1}{2}, 1[$ on a $f_{X,Y}(x,y) = 2\mathbb{1}_{[\frac{1}{2},1]}(y)$, donc

$$f_X(x) = \int_{\frac{1}{2}}^{1} 2\,\mathrm{d}y = 1\,.$$

Comme on a $f_{X,Y}(x,y) = 0$ si $x \notin [0,1]$ et que la valeur de f_X au point $x = \frac{1}{2}$ n'est pas importante, on peut écrire $f_X(x) = \mathbb{1}_{[0,1]}(x)$, donc $X \sim \mathrm{U}(0,1)$. Avec des calculs analogues, on a $Y \sim \mathrm{U}(0,1)$.

Vu que $f_{X,Y}(x,y) = 0$ pour tout (x,y) dans le carré ouvert $]0, \frac{1}{2}[\, \times \,]\frac{1}{2}, 1[$, qui est de mesure égale à $\frac{1}{4}$, alors que $f_X(x)f_Y(y) = 1$ pour ces valeurs de (x,y), on en déduit que $f_{X,Y}(x,y)$ ne coïncide pas avec $f_X(x)f_Y(y)$ hors d'un ensemble de mesure nulle, donc X et Y ne sont pas indépendantes. Autrement, il suffit de noter que

$$0 = \mathrm{P}\big(X \in \,]0, \tfrac{1}{2}[\,, Y \in \,]\tfrac{1}{2}, 1[\,\big) \neq \mathrm{P}\big(X \in \,]0, \tfrac{1}{2}[\,\big)\,\mathrm{P}\big(Y \in \,]\tfrac{1}{2}, 1[\,\big) = \frac{1}{4}\,.$$

Exercice 6.55 (*). Soit (X,Y) un vecteur aléatoire bidimensionnel de loi uniforme sur le disque de rayon unitaire centré en l'origine. Déterminer, si possible sans calculs, les probabilités conditionnelles suivantes :

$$\mathrm{P}\left(\max(|X|,|Y|) \leq \frac{1}{2\sqrt{2}} \;\middle|\; X^2 + Y^2 \leq \frac{1}{4}\right), \;\; \mathrm{P}\left(\max(|X|,|Y|) \leq \frac{1}{2} \;\middle|\; |X| + |Y| \leq 1\right).$$

[*Sugg.* Faire un dessin.]

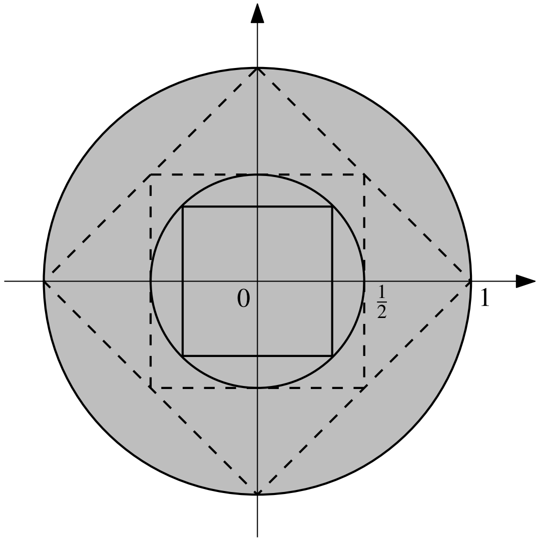


Fig. S6.2 Les disques $D_1 := \{x^2+y^2 \le 1\}$ et $D_2 := \{x^2+y^2 \le \frac{1}{4}\}$ et les carrés $C_1 := \{|x|+|y| \le 1\}$, $C_2 := \{\max(|x|,|y|) \le \frac{1}{2}\}$ et $C_3 := \{\max(|x|,|y|) \le \frac{1}{2\sqrt{2}}\}$ (Exercice 6.55)

Solution 6.55. Si on note $D_1 := \{(x,y) \in \mathbb{R}^2 :\ x^2+y^2 \le 1\}$ le disque de rayon unitaire centré en l'origine de $\mathbb{R}^2$, le vecteur aléatoire (X,Y) est de loi $\mathrm{U}(D_1)$, c'est-à-dire est absolument continu de densité

$$f_{X,Y}(x,y) = \frac{1}{\mathrm{mes}(D_1)}\,\mathbb{1}_{D_1}(x,y)\,.$$

Ainsi, pour tout sous-ensemble $A \subseteq \mathbb{R}^2$ (dont l'indicatrice $\mathbb{1}_A$ est intégrable) on a

$$\begin{aligned}\mathrm{P}((X,Y)\in A) &= \int_{\mathbb{R}^2} \mathbb{1}_A(x,y)\,f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y = \frac{1}{\mathrm{mes}(D_1)}\int_{\mathbb{R}^2} \mathbb{1}_A(x,y)\,\mathbb{1}_{D_1}(x,y)\,\mathrm{d}x\,\mathrm{d}y\\ &= \frac{\mathrm{mes}(A\cap D_1)}{\mathrm{mes}(D_1)}\,.\end{aligned}$$

Par conséquent, pour tout choix de A,B on a

$$\mathrm{P}\big((X,Y)\in A \,|\, (X,Y)\in B\big) = \frac{\mathrm{mes}(A\cap D_1)}{\mathrm{mes}(B\cap D_1)}\,.$$

En introduisant le disque $D_2 := \{(x,y)\in\mathbb{R}^2 :\ x^2+y^2 \le \frac{1}{4}\}$ et les carrés $C_1 := \{(x,y)\in\mathbb{R}^2 : |x|+|y| \le 1\}$, $C_2 := \{(x,y)\in\mathbb{R}^2 :\ \max(|x|,|y|) \le \frac{1}{2}\}$ et $C_3 := \{(x,y)\in\mathbb{R}^2 :\ \max(|x|,|y|) \le \frac{1}{2\sqrt{2}}\}$, les probabilités demandées sont

$$\begin{aligned}\mathrm{P}\left(\max(|X|,|Y|) \le \frac{1}{2\sqrt{2}} \,\middle|\, X^2+Y^2 \le \frac{1}{4}\right) &= \mathrm{P}\big((X,Y)\in C_3 \,|\, (X,Y)\in D_2\big)\\ &= \frac{\mathrm{mes}(C_3)}{\mathrm{mes}(D_2)} = \frac{\frac{1}{2}}{\pi\frac{1}{4}} = \frac{2}{\pi}\,,\end{aligned}$$

$$\mathrm{P}\left(\max(|X|,|Y|) \leq \frac{1}{2} \;\middle|\; |X|+|Y| \leq 1\right) = \mathrm{P}\big((X,Y) \in C_2 \,|\, (X,Y) \in C_1\big)$$
$$= \frac{\mathrm{mes}(C_2)}{\mathrm{mes}(C_1)} = \frac{1}{2}\,.$$

Exercice 6.56 (*). Soient $(X_n)_{n\in\mathbb{N}}$ des variables aléatoires réelles i.i.d. de loi $\mathrm{Exp}(1)$. On définit, pour $n \geq 1$,

$$U_n := \frac{X_0}{X_1+\cdots+X_n}\,.$$

(i) Montrer que la fonction de répartition de U_n est donnée, pour $t \geq 0$, par

$$F_{U_n}(t) = 1 - \frac{1}{(1+t)^n}\,.$$

[*Sugg.* Observer que $Y_n := X_1+\cdots+X_n$ suit la loi ... et est indépendante de ...]

(ii) En déduire que la variable aléatoire U_n est à densité. Pour tout n fixé, déterminer pour quelles valeurs de $p > 0$ on a $U_n \in L^p$.

Solution 6.56. (i) Clairement $F_{U_n}(t) = 0$ pour $t \leq 0$. Introduisons la variable aléatoire $Y_n := X_1+\cdots+X_n$, qui est de loi $\mathrm{Gamma}(n,1)$, grâce à la Proposition 6.36. On note que l'on peut écrire

$$F_{U_n}(t) = \mathrm{P}(U_n \leq t) = \mathrm{P}\left(\frac{X_0}{Y_n} \leq t\right) = \mathrm{P}\big(X_0 \leq tY_n\big)$$
$$= \mathrm{P}\big((X_0,Y_n) \in \{(x,y) \in \mathbb{R}^2 :\, x \leq ty\}\big)\,.$$

Comme les variables aléatoires X_0 et Y_n sont indépendantes, le vecteur aléatoire (X_0,Y_n) est absolument continu, de densité

$$f_{X_0,Y_n}(x,y) = f_{X_0}(x)\, f_{Y_n}(y) = \mathrm{e}^{-x}\,\mathbb{1}_{]0,\infty[}(x)\,\frac{y^{n-1}}{(n-1)!}\,\mathrm{e}^{-y}\,\mathbb{1}_{]0,\infty[}(y)\,.$$

Pour $t > 0$ on peut donc écrire

$$\begin{aligned}
F_{U_n}(t) &= \int_{\mathbb{R}^2} \mathbb{1}_{\{(x,y)\in\mathbb{R}^2:\, x\leq ty\}}\, f_{X_0,Y_n}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_0^\infty \int_0^\infty \mathbb{1}_{\{x\leq ty\}}\,\mathrm{e}^{-x}\,\frac{y^{n-1}\,\mathrm{e}^{-y}}{(n-1)!}\,\mathrm{d}x\,\mathrm{d}y = \int_0^\infty \frac{y^{n-1}\,\mathrm{e}^{-y}}{(n-1)!}\left(\int_0^{ty}\mathrm{e}^{-x}\,\mathrm{d}x\right)\mathrm{d}y \\
&= \int_0^\infty \frac{y^{n-1}\,\mathrm{e}^{-y}}{(n-1)!}\big(1-\mathrm{e}^{-ty}\big)\,\mathrm{d}y = 1 - \frac{1}{(1+t)^n}\int_0^\infty (1+t)^n\,\frac{y^{n-1}\,\mathrm{e}^{-(1+t)y}}{(n-1)!}\,\mathrm{d}y \\
&= 1 - \frac{1}{(1+t)^n}\,,
\end{aligned}$$

où on a fait apparaître dans la dernière intégrale la densité d'une variable aléatoire $\mathrm{Gamma}(n,1+t)$.

(ii) Comme F_{U_n} est une fonction continue et $\mathcal{C}^1$ par morceaux, il s'ensuit que U_n est une variable aléatoire à densité, de densité donnée par

$$f_{U_n}(t) = F'_{U_n}(t) = \frac{1}{(1+t)^{n+1}}\,\mathbb{1}_{[0,\infty[}(t)\,.$$

Par conséquent

$$\mathrm{E}(|U_n|^p) = \int_0^\infty \frac{t^p}{(1+t)^{n+1}}\,\mathrm{d}t < \infty \qquad \Longleftrightarrow \qquad p < n\,,$$

parce que $t^p/(1+t)^{n+1} \sim t^{p-n-1}$ quand $t \to +\infty$.

Exercice 6.57 (*). Soient X et Y des variables aléatoires réelles indépendantes, où Y suit la loi $\mathrm{U}(-\frac{\pi}{2}, \frac{\pi}{2})$ et où X a pour densité

$$f_X(x) = x\mathrm{e}^{-\frac{x^2}{2}}\,\mathbb{1}_{]0,+\infty[}(x).$$

Déterminer la loi jointe et les densités marginales des variables aléatoires Z et W, définies par

$$Z := X\cos Y\,, \qquad W := X\sin Y\,.$$

[*Sugg.* Observer que $X = \sqrt{Z^2+W^2}$ et $Y = \arctan(W/Z)$.]

Solution 6.57. Considérons la fonction $\varphi :]0,+\infty[\, \times\,]-\pi/2,\pi/2[\,\to\,]0,+\infty[\, \times \mathbb{R}$ définie par $\varphi(x,y) = (x\cos y, x\sin y)$. On a $\varphi^{-1}(z,w) = (\sqrt{z^2+w^2}, \arctan(w/z))$, donc

$$J_{\varphi^{-1}}(z,w) = \begin{pmatrix} \frac{z}{\sqrt{z^2+w^2}} & \frac{w}{\sqrt{z^2+w^2}} \\ -\frac{w}{z^2}\frac{1}{1+(w/z)^2} & \frac{1}{z}\frac{1}{1+(w/z)^2} \end{pmatrix}$$

d'où

$$\det J_{\varphi^{-1}}(z,w) = \frac{1}{\sqrt{z^2+w^2}}.$$

Comme X et Y sont des variables aléatoires indépendantes absolument continues, le vecteur aléatoire (X,Y) est absolument continu, de densité

$$f_{X,Y}(x,y) \;=\; f_X(x)\, f_Y(y) \;=\; x\mathrm{e}^{-\frac{1}{2}x^2}\,\mathbb{1}_{]0,\infty[}(x)\,\frac{1}{\pi}\,\mathbb{1}_{]-\pi/2,\pi/2[}(y)\,.$$

Il découle alors de la Proposition 6.60 que (Z,W) est un vecteur aléatoire absolument continu, de densité

$$f_{Z,W}(z,w) = f_{X,Y}(\varphi^{-1}(z,w))|\det J_{\varphi^{-1}}(z,w)| = \frac{1}{\pi}\mathrm{e}^{-\frac{1}{2}(z^2+w^2)}1_{]0,+\infty[}(z).$$

On en déduit que Z et W sont indépendantes, avec $W \sim \mathrm{N}(0,1)$ et Z de densité

$$f_Z(z) = \sqrt{\tfrac{2}{\pi}}\,\mathrm{e}^{-\frac{1}{2}z^2}\,1_{]0,+\infty[}(z)\,.$$

Exercice 6.58 (Biais par la taille, II *). Soit X une variable aléatoire réelle de densité $f_X(x)$. On suppose que X est positive, donc $f_X(x) = 0$ pour $x \le 0$, et qu'elle est d'espérance finie : $\mathrm{E}(X) < \infty$. Soient maintenant σ, τ deux variables aléatoires réelles, de densité jointe

$$f_{\sigma,\tau}(a,b) := c\, f_X(a+b)\,\mathbb{1}_{]0,\infty[}(a)\,\mathbb{1}_{]0,\infty[}(b)\,,$$

où $c \in \mathbb{R}$ est une constante opportune.

(i) Montrer que σ et τ ont la même loi, absolument continue de densité

$$f_\tau(t) = c\,\mathrm{P}(X > t)\,\mathbb{1}_{]0,\infty[}(t)\,.$$

En déduire que l'on a $c = \frac{1}{\mathrm{E}(X)}$.

(ii) Montrer que les variables aléatoires X^* et U définies par

$$X^* := \sigma + \tau \qquad \text{et} \qquad U := \frac{\tau}{\sigma+\tau}$$

sont indépendantes, avec de plus $U \sim \mathrm{U}(0,1)$ et X^* de densité $f_{X^*}(x) = \frac{1}{\mathrm{E}(X)}\, x\, f_X(x)$. Comme on l'a vu dans l'Exercice 6.41, X^* est appelée *version biaisée par la taille* de X.

Solution 6.58. (i) Par symétrie ($f_{\sigma,\tau}(a,b) = f_{\sigma,\tau}(b,a)$), les densités marginales données par

$$f_\sigma(s) = \int_{-\infty}^{+\infty} f_{\sigma,\tau}(s,t)\,\mathrm{d}t \qquad f_\tau(t) = \int_{-\infty}^{+\infty} f_{\sigma,\tau}(s,t)\,\mathrm{d}s$$

sont égales. Clairement, on a $f_\tau(t) = 0$ pour $t \leq 0$. Pour $t > 0$

$$f_\tau(t) = \int_{-\infty}^{+\infty} f_{\sigma,\tau}(a,t)\,\mathrm{d}a = \int_0^{\infty} c f_X(a+t)\,\mathrm{d}a = c\int_t^{\infty} f(u)\mathrm{d}u\,,$$

où on a fait le changement de variable $u = a+t$. On trouve donc

$$f_\tau(t) = c\,\mathrm{P}(X > t)\mathbb{1}_{]0,\infty[}(t)\,.$$

Vu que $\int_{-\infty}^{+\infty} f_\tau(t)\,\mathrm{d}t = 1$, on a

$$\frac{1}{c} = \int_0^{+\infty} \mathrm{P}(X > t)\mathrm{d}t = \mathrm{E}(X)\,,$$

où pour la dernière égalité on a utilisé l'Exercice 6.3 (avec $k = 1$).

(ii) Considérons la fonction $\varphi :]0,+\infty[^2 \to]0,+\infty[\times]0,1[$ définie par $\varphi(s,t) = (s+t, \frac{s}{s+t})$. On a $\varphi^{-1}(x,u) = (ux, (1-u)x)$, donc

$$J_{\varphi^{-1}}(x,u) = \begin{pmatrix} u & x \\ 1-u & -x \end{pmatrix}$$

d'où $\det J_{\varphi^{-1}}(x,u) = -x$. Comme le vecteur aléatoire (σ,τ) est absolument continu, on déduit de la Proposition 6.60 que le vecteur aléatoire $(X^*,U) = \varphi(\sigma,\tau)$ est absolument continu, de densité

$$f_{X^*,U}(x,u) = f_{\sigma,\tau}(\varphi^{-1}(x,u))\big|\det J_{\varphi^{-1}}(x,u)\big| = \frac{1}{\mathrm{E}(X)} f_X(x) x \mathbb{1}_{]0,+\infty[}(x)\mathbb{1}_{]0,1[}(u)\,.$$

De manière alternative, on peut considérer une fonction $h : \mathbb{R}^2 \to \mathbb{R}$ bornée et continue sauf sur un ensemble de mesure 2-dimensionnelle nulle quelconque, et calculer

$$\begin{aligned}
\mathrm{E}(h(X^*,U)) &= \mathrm{E}\left(h\left(\sigma+\tau, \frac{\tau}{\sigma+\tau}\right)\right) \\
&= \int_{]0,+\infty[^2} h\left(s+t, \frac{t}{s+t}\right) \frac{1}{\mathrm{E}(X)} f_X(s+t)\mathrm{d}s\mathrm{d}t \\
&= \int_{]0,+\infty[\times]0,1[} h(x,u) \frac{1}{\mathrm{E}(X)} f_X(x)\, x\,\mathrm{d}x\,\mathrm{d}u\,,
\end{aligned}$$

où on a appliqué le même changement de variable que ci-dessus pour obtenir la dernière égalité. Grâce à l'Observation 6.58, cela signifie que (X^*,U) a pour densité jointe $\frac{1}{\mathrm{E}(X)} f_X(x) x \mathbb{1}_{]0,+\infty[}(x)\mathbb{1}_{]0,1[}(u)$.

De la densité jointe, on déduit que X^* et U sont indépendantes, de densités $f_{X^*}(x) = \frac{1}{\mathrm{E}(X)} f_X(x) x \mathbb{1}_{]0,+\infty[}(x)$ et $f_U(u) = \mathbb{1}_{]0,1[}(u)$, c'est-à-dire $U \sim \mathrm{U}(0,1)$.

Exercices plus difficiles

Exercice 6.59. Olivier est un lanceur de javelot. Après son premier lancer, qui envoie le javelot à une distance X_0, il se lance dans une succession de lancers : le n-ème lancer de javelot tombe à une distance X_n. Oliver se demande combien de lancers T il devra effectuer pour améliorer sont résultat initial, c'est-à-dire

$$T := \min\{n \geq 1 : X_n > X_0\}.$$

On suppose que les $(X_n)_{n\in\mathbb{N}}$ sont des variables aléatoires indépendantes et de même loi (en faisant abstraction de la fatigue...), que l'on suppose à densité. On va montrer que T *suit une loi* universelle*, qui ne dépend pas de la loi des* X_n, de densité discrète

$$\mathrm{p}_T(k) = \frac{1}{k(k+1)} \mathbb{1}_{\mathbb{N}^*}(k), \tag{S6.18}$$

dont il découle que le nombre de lancers nécessaires pour s'améliorer possède une espérance $\mathrm{E}(T) = +\infty$! Avoir ignoré la fatigue rend ce résultat encore plus surprenant.

(i) On définit les événements

$$A_k^{(n)} := \left\{X_k = \max_{0\leq i\leq n} X_i\right\}, \qquad \text{pour } n \in \mathbb{N}^*,\ k \in \{0,\dots,n\}.$$

Montrer, avec un argument de symétrie, que $\mathrm{P}\left(A_0^{(n)}\right) = \mathrm{P}\left(A_1^{(n)}\right) = \cdots = \mathrm{P}\left(A_n^{(n)}\right)$.

(ii) Montrer que, pour tout $i \neq j$ on a $\mathrm{P}(X_i \neq X_j) = 1$.

[*Sugg.* Remarquer que $X_i - X_j$ est une variable aléatoire à densité (pourquoi ?).]

(iii) Expliquer l'inclusion d'événements $A_i^{(n)} \cap A_j^{(n)} \subseteq \{X_i = X_j\}$ et en déduire que

$$\mathrm{P}\left(A_i^{(n)} \cap A_j^{(n)}\right) = 0, \qquad \forall i \neq j.$$

(iv) Expliquer pourquoi $\bigcup_{k=0}^n A_k^{(n)} = \Omega$ et en déduire que

$$\mathrm{P}(A_k^{(n)}) = \frac{1}{n+1}, \qquad \forall n \in \mathbb{N}^*,\ k \in \{0,\dots,n\}.$$

[*Sugg.* On pourra utiliser l'Exercice 1.4.]

(v) Expliquer pourquoi on a $\{T > n\} = A_0^{(n)}$ et donc $\mathrm{P}(T > n) = \frac{1}{n+1}$. En déduire la formule (S6.18) et le fait que $\mathrm{E}(T) = +\infty$.

Solution 6.59. (i) Intuitivement, les événements $A_k^{(n)}$ pour $k \in \{0,1,\dots,n\}$ sont simplement obtenus en permutant les variables aléatoires $X_0, X_1, \dots, X_n$; mais toutes les permutations des X_i ont la même loi, parce que les variables sont i.i.d., donc les événements $A_k^{(n)}$ ont la même probabilité.

De manière plus formelle, en définissant $Z_k := X_k - \max_{0\leq i\leq n} X_i$, on peut écrire $A_k^{(n)} = \{Z_k = 0\}$, donc $\mathrm{P}(A_k^{(n)}) = \mathrm{P}(Z_k = 0) = \mu_{Z_k}(\{0\})$, en notant μ_{Z_k} la loi de Z_k. Il suffit alors de montrer que les variables aléatoires Z_k et $Z_{k'}$, pour tout choix de $k,k' \in \{0,1,\dots,n\}$, ont la même loi. En posant $Y := \max_{0\le i\le n} X_i$ et $f(x,y) := x-y$, on peut écrire $Z_k = f(X_k,Y)$ et $Z_{k'} = f(X_{k'},Y)$: il suffit donc de montrer que les vecteurs aléatoires bidimensionnels (X_k,Y) et $(X_{k'},Y)$ ont la même loi, grâce à la Proposition 3.11 (conservation des lois). En introduisant la fonction $g : \mathbb{R}^{n+1} \to \mathbb{R}^2$ définie par $g(x_0,x_1,\dots,x_n) := (x_0, \max_{0\le i\le n} x_i)$ et la permutation (transposition) $\pi_k : \{0,1,\dots,n\} \to \{0,1,\dots,n\}$ définie par $\pi_k(0) := k$, $\pi_k(k) = 0$ et $\pi_k(i) = i$ si $i \neq 0$, $i \neq k$, on peut écrire $(X_k,Y) = g(\pi_k(X_0,\dots,X_n))$ et $(X_{k'},Y) = g(\pi_{k'}(X_0,\dots,X_n))$; encore par conservation des lois, il suffit de montrer que les vecteurs aléatoires $(n+1)$-dimensionnels $\pi_k(X_0,\dots,X_n)$ et $\pi_{k'}(X_0,\dots,X_n)$ ont la même loi, mais cela découle du fait que les vecteurs ont tous les deux des composantes indépendantes et de même densité.

(ii) Pour $i \neq j$, les variables aléatoires X_i et $-X_j$ sont indépendantes et absolument continues, donc leur somme $X_i - X_j$ est une variable aléatoire absolument continue (Proposition 6.28). Par conséquent, $\mathrm{P}(X_i - X_j = 0) = 0$.

(iii) Si $\omega \in A_i^{(n)}$, alors par définition $X_i(\omega) = \max_{0\le \ell\le n} X_\ell(\omega)$; de même, si $\omega \in A_j^{(n)}$, on a $X_j(\omega) = \max_{0\le \ell\le n} X_\ell(\omega)$. En particulier, si $\omega \in A_i^{(n)} \cap A_j^{(n)}$, on a $X_i(\omega) = X_j(\omega)$. Cela montre l'inclusion d'événements $A_i^{(n)} \cap A_j^{(n)} \subseteq \{X_i = X_j\}$, dont on déduit que

$$\mathrm{P}\left(A_i^{(n)} \cap A_j^{(n)}\right) \le \mathrm{P}(X_i = X_j) = 0 .$$

Ainsi $\mathrm{P}\left(A_i^{(n)} \cap A_j^{(n)}\right) = 0$ pour $i \neq j$.

(iv) Pour tout $\omega \in \Omega$, les nombres réels $\{X_0(\omega), X_1(\omega), \dots, X_n(\omega)\}$ possèdent un maximum, que l'on note $X_{\bar{k}}(\omega)$, où l'indice $\bar{k} = \bar{k}(\omega) \in \{0,1,\dots,n\}$ est défini comme le plus petit indice pour lequel le maximum est atteint (dans le cas où il est atteint pour plusieurs indices). En particulier, $X_{\bar{k}}(\omega) = \max_{0\le i\le n} X_i(\omega)$, ce qui signifie que $\omega \in A_{\bar{k}}^{(n)}$. On a donc montré que tout $\omega \in \Omega$ appartient à un ensemble $A_k^{(n)}$, pour un choix opportun de $k \in \{0,\dots,n\}$, c'est-à-dire $\bigcup_{k=0}^n A_k^{(n)} = \Omega$. De cette observation et de la sous-additivité, on obtient

$$1 = \mathrm{P}(\Omega) \le \sum_{k=0}^n \mathrm{P}\left(A_k^{(n)}\right) .$$

En appliquant l'inégalité de Bonferroni (Exercice 1.4), on obtient

$$1 = \mathrm{P}(\Omega) \ge \sum_{k=0}^n \mathrm{P}\left(A_k^{(n)}\right) - \sum_{0\le i<j\le n} \mathrm{P}\left(A_i^{(n)} \cap A_j^{(n)}\right) .$$

Mais on a déjà montré que $\mathrm{P}\left(A_i^{(n)} \cap A_j^{(n)}\right) = 0$ si $i < j$. On déduit donc des relations précédentes que

$$1 = \sum_{k=0}^n \mathrm{P}\left(A_k^{(n)}\right) .$$

Enfin, comme les événements $\mathrm{P}\left(A_k^{(n)}\right)$ ont tous la même probabilité quel que soit k, il s'ensuit que $\mathrm{P}\left(A_k^{(n)}\right) = \frac{1}{n+1}$, comme demandé.

(v) Par définition, on a $\omega \in \{T > n\}$ si et seulement si $T(\omega) > n$, c'est-à-dire $X_k(\omega) \le X_0(\omega)$ pour tout $k = 1,2,\dots,n$; autrement dit, $\omega \in \{T > n\}$ si et seulement si $X_0(\omega) =$

$\max_{0\le i\le n} X_i(\omega)$, c'est-à-dire si et seulement si $\omega \in A_0^{(n)}$. Cela montre que $\{T > n\} = A_0^{(n)}$, dont on déduit que $\mathrm{P}(T > n) = \mathrm{P}(A_0^{(n)}) = \frac{1}{n+1}$ et donc, pour tout $k \in \mathbb{N}^*$,

$$\mathrm{P}(T = k) = \mathrm{P}(T > k-1) - \mathrm{P}(T > k) = \frac{1}{k} - \frac{1}{k+1} = \frac{1}{k(k+1)},$$

ce qui donne la formule (S6.18). En particulier, on obtient

$$\mathrm{E}(T) = \sum_{k\in\mathbb{N}^*} k\,\mathrm{P}(T = k) = \sum_{k\in\mathbb{N}^*} \frac{1}{k+1} = +\infty.$$

Exercice 6.60 (*). Reprenons le contexte de l'Exercice 6.59 : soient $X_0, X_1, X_2, \dots$ des variables aléatoires réelles indépendantes et de même loi à densité, et soit

$$T := \min\{n \ge 1 : X_n > X_0\}$$

le nombre de tentatives nécessaires pour améliorer le résultat initial. On s'intéresse dans cet exercice au résultat que l'on obtient à l'instant T : on note $Y := X_T$, c'est-à-dire $Y(\omega) := X_{T(\omega)}(\omega)$. On notera f (resp. F) la densité (resp. la fonction de répartition) commune des X_i et on supposera que f est continue par morceaux.

(i) Montrer que, pour tout $k \ge 1$ et tout $t \in \mathbb{R}$

$$\mathrm{P}(Y \le t, T = k) = \mathrm{P}(X_1, \dots, X_{k-1} \le X_0 < X_k \le t).$$

(ii) Montrer que, pour tout $k \ge 1$ et $t \in \mathbb{R}$

$$\mathrm{P}(Y \le t, T = k) = \int_{-\infty}^{t} \left(\int_{-\infty}^{x_k} F(x_0)^{k-1} f(x_0)\,\mathrm{d}x_0 \right) f(x_k)\,\mathrm{d}x_k,$$

et en déduire que $\mathrm{P}(Y \le t, T = k) = \frac{1}{k(k+1)} F(t)^{k+1}$. Retrouver la conclusion de l'Exercice 6.59, c'est-à-dire $\mathrm{P}(T = k) = \frac{1}{k(k+1)}$ pour tout $k \in \mathbb{N}^*$.

(iii) Montrer que la fonction de répartition de Y est donnée par

$$\mathrm{P}(Y \le t) = F(t) + (1 - F(t)) \log(1 - F(t)).$$

En déduire que Y est à densité et exprimer cette densité en fonction de f et F.

On cherche maintenant à savoir *de combien* on améliore le résultat initial : on considère la variable aléatoire $W := Y - X_0$. On va d'abord déterminer la densité jointe de (X_0, Y) : pour cela, on se donne une *fonction test* quelconque $h : \mathbb{R}^2 \to \mathbb{R}$ bornée et continue sauf au plus sur un ensemble de mesure 2-dimensionnelle nulle, et on va exprimer $\mathrm{E}(h(X_0, Y))$.

(iv) En utilisant l'égalité d'événements $\{T = k\} = \{X_1, \dots, X_{k-1} \le X_0 < X_k\}$, montrer que pour tout $k \in \mathbb{N}^*$

$$\mathrm{E}\left[h(X_0, Y)\mathbb{1}_{\{T=k\}}\right] = \int_{\mathbb{R}^2} h(x_0, x_k) F(x_0)^{k-1} f(x_0) f(x_k) \mathbb{1}_{\{x_0 < x_k\}} \mathrm{d}x_0\,\mathrm{d}x_k.$$

(v) Admettons que l'on puisse échanger les sommes (infinies) suivantes avec l'espérance ou l'intégrale † :

†. On peut le justifier avec des outils d'analyse plus avancés, comme le théorème de Fubini–Tonelli.

$$\sum_{k=1}^{n} \mathrm{E}\left[h(X_0,Y)\mathbb{1}_{\{T=k\}}\right] = \mathrm{E}\left[\sum_{k=1}^{n} h(X_0,Y)\mathbb{1}_{\{T=k\}}\right],$$

$$\sum_{k=1}^{n} \int_{\mathbb{R}^2} h(x,y)F(x)^{k-1}f(x)f(y)\mathbb{1}_{\{x<y\}}\mathrm{d}x\,\mathrm{d}y$$
$$= \int_{\mathbb{R}^2} \sum_{k=1}^{n} h(x,y)F(x)^{k-1}f(x)f(y)\mathbb{1}_{\{x<y\}}\mathrm{d}x\,\mathrm{d}y.$$

Montrer alors que la densité jointe de (X_0,Y) est donnée par

$$f_{X_0,Y}(x,y) = \frac{1}{1-F(x)}f(x)f(y)\mathbb{1}_{\{x<y\}},$$

et retrouver la densité marginale de Y.

(vi) Déterminer la densité jointe de $(X_0,W) = (X_0,Y-X_0)$, puis exprimer la densité marginale de W sous forme d'intégrale. Vérifier que dans le cas où $X \sim \mathrm{Exp}(\lambda)$ avec $\lambda > 0$, c'est-à-dire $f(x) = \lambda \mathrm{e}^{-\lambda x}\mathbb{1}_{\mathbb{R}_+}(x)$, alors X_0 et W sont indépendantes, et $W \sim \mathrm{Exp}(\lambda)$.

Solution 6.60. (i) Remarquons que l'on a l'égalité d'événements $\{Y = X_T \leq t, T = k\} = \{X_k \leq t, T = k\}$ et que cet événement correspond au fait que les $k-1$ premières tentatives pour surpasser X_0 ont échoué (c'est-à-dire $X_i \leq X_0$ pour tout $1 \leq i \leq k-1$) et que la k-ème tentative pour surpasser X_0 est un succès ($X_k > X_0$), sans que X_k dépasse t ($X_k \leq t$). Pour résumer,

$$\{X_k \leq t, T = k\} = \{X_i \leq X_0\,, \forall\, 1 \leq i \leq k-1\} \cap \{X_0 < X_k \leq t\},$$

et donc

$$\mathrm{P}(Y \leq t) = \mathrm{P}(X_1,\ldots,X_{k-1} \leq X_0 < X_k \leq t).$$

(ii) On définit $C_t := \{(x_0,\ldots,x_k) :\ x_1,\ldots,x_{k-1} \leq x_0 < x_k \leq t\}$, de sorte que

$$\begin{aligned}\mathrm{P}(Y \leq t, T = k) &= \mathrm{P}\left((X_0,X_1,\ldots,X_k) \in C_t\right)\\ &= \int_{C_t} f(x_0)f(x_1)\ldots f(x_{k-1})f(x_k)\mathrm{d}x_0\ldots\mathrm{d}x_k\end{aligned}$$

où on a utilisé le fait que la densité jointe de $(X_0,X_1,\ldots,X_k)$ est donnée par le produit $f(x_0)f(x_1)\ldots f(x_k)$, par indépendance des X_i. D'après le théorème de Fubini–Tonelli, on peut commencer par intégrer par rapport à $x_1,\ldots,x_{k-1}$, puis x_0, puis x_k :

$$\begin{aligned}&\mathrm{P}(Y \leq t, T = k)\\ &= \int_{-\infty}^{t}\left(\int_{-\infty}^{x_k}\left(\int_{]-\infty,x_0]^{k-1}} f(x_1)\ldots f(x_{k-1})\mathrm{d}x_1\ldots\mathrm{d}x_{k-1}\right)f(x_0)\mathrm{d}x_0\right)f(x_k)\mathrm{d}x_k\\ &= \int_{-\infty}^{t}\left(\int_{-\infty}^{x_k} F(x_0)^{k-1}f(x_0)\mathrm{d}x_0\right)f(x_k)\mathrm{d}x_k\end{aligned}$$

où on a remarqué que l'intégrale interne valait $\mathrm{P}(X_i \leq x_0\,, \forall\, 1 \leq i \leq k-1)$, ce qui par indépendance des X_i est égal à $\mathrm{P}(X \leq x_0)^{k-1} = F(x_0)^{k-1}$.
Maintenant, en utilisant le fait que $F'(x_0) = f(x_0)$, on peut intégrer

$$\int_{-\infty}^{x_k} F(x_0)^{k-1}f(x_0)\mathrm{d}x_0 = \left[\frac{1}{k}F(x_0)^k\right]_{-\infty}^{x_k} = \frac{1}{k}F(x_k)^k,$$

où on a utilisé le fait que $\lim_{x\to-\infty} F(x) = 0$. On obtient finalement

$$\begin{aligned}
\mathrm{P}(Y \le t, T = k) &= \int_{-\infty}^{t} \frac{1}{k} F(x_k)^k f(x_k) \mathrm{d}x_k \\
&= \left[\frac{1}{k(k+1)} F(x_k)^{k+1}\right]_{-\infty}^{t} = \frac{1}{k(k+1)} F(t)^{k+1}.
\end{aligned}$$

Fixons maintenant $k \in \mathbb{N}^*$. Notons que les événements $(\{Y \le n, T = k\})_{n\ge1}$ sont croissants pour l'inclusion, donc par continuité par le bas des probabilités (Proposition 1.24)

$$\lim_{n\to+\infty} \mathrm{P}(Y \le n, T = k) = \mathrm{P}\Big(\bigcup_{n\ge1} \{Y \le n\} \cap \{T = k\}\Big) = \mathrm{P}\big(\{Y < +\infty\} \cap \{T = k\}\big).$$

Comme $\mathrm{P}(Y < +\infty) = 1$, on obtient que $\mathrm{P}(\{Y < +\infty\} \cap \{T = k\}) = \mathrm{P}(T = k)$ (voir l'Exercice 1.1). On en conclut que

$$\mathrm{P}(T = k) = \lim_{n\to+\infty} \mathrm{P}(Y \le n, T = k) = \frac{1}{k(k+1)} \lim_{n\to+\infty} F(n)^{k+1} = \frac{1}{k(k+1)},$$

parce que $\lim_{t\to+\infty} F(t) = 1$.

(iii) En sommant sur $k \in \mathbb{N}^*$ et en utilisant que $\{T = k\}_{k\in\mathbb{N}^*}$ est une partition de Ω, on obtient pour $t \in \mathbb{R}$

$$\mathrm{P}(Y \le t) = \sum_{k=1}^{+\infty} \mathrm{P}(Y \le t, T = k) = \sum_{k=1}^{+\infty} \frac{1}{k(k+1)} F(t)^{k+1}.$$

Si $F(t) = 1$, alors $\mathrm{P}(Y \le t) = 1$, parce que $\sum_{k=1}^{+\infty} \frac{1}{k(k+1)} = 1$, grâce à une somme télescopique basée sur le fait que $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$.

Si $F(t) < 1$, on a

$$\mathrm{P}(Y \le t) = \sum_{k=1}^{+\infty} \Big(F(t)\frac{F(t)^k}{k} - \frac{F(t)^{k+1}}{k+1}\Big) = F(t) \sum_{k=1}^{+\infty} \frac{F(t)^k}{k} - \sum_{k=1}^{+\infty} \frac{F(t)^{k+1}}{k+1}$$

(on a utilisé la linéarité pour les sommes infinies et aussi que les deux sommes sont finies car elles sont majorées par $\sum_{k=1}^{\infty} F(t)^{k+1} < +\infty$). En se rappelant de la série de Taylor de $\log(1+x)$, voir (0.7), on a

$$\begin{aligned}
\sum_{k=1}^{+\infty} \frac{F(t)^k}{k} &= -\log\big(1 - F(t)\big) \\
\sum_{k=1}^{+\infty} \frac{F(t)^{k+1}}{k+1} &= \sum_{k=0}^{+\infty} \frac{F(t)^{k+1}}{k+1} - F(t) = -\log\big(1 - F(t)\big) - F(t).
\end{aligned}$$

Pour résumer, on trouve que

$$\begin{aligned}
F_Y(t) = \mathrm{P}(Y \le t) &= -F(t)\log\big(1 - F(t)\big) + F(t) + \log\big(1 - F(t)\big) \\
&= F(t) + (1 - F(t))\log\big(1 - F(t)\big).
\end{aligned}$$

Comme la fonction f est continue par morceaux, on a que F est continue et $\mathcal{C}^1$ par morceaux, donc c'est aussi le cas pour $F_Y(t)$: on déduit alors de la Proposition 6.17 que la variable aléatoire Y est absolument continue, de densité donnée par

$$f_Y(y) = F_Y'(y) = F'(y) - F'(y)(1+\log(1-F(y))) = -f(y)\log\big(1-F(y)\big)$$

(la dérivée de $x \mapsto (1-x)\log(1-x)$ est $-1-\log(1-x)$). Par convention, on pose $f_Y(y) = 0$ si $F(y) = 1$ (parce que $F_Y(t) = 1$ pour tout $t \geq y$).

(iv) On définit $C := \{(x_0,\dots,x_k) :\ x_1,\dots,x_{k-1} \leq x_0 < x_k\}$. Grâce à la formule de transfert, appliquée au vecteur aléatoire $X_0, X_1, \dots, X_k$, on a

$$\begin{aligned}\mathrm{E}\big[h(X_0,Y)\mathbb{1}_{\{T=k\}}\big] &= \mathrm{E}\big[h(X_0,X_k)\mathbb{1}_{\{(X_0,X_1,\dots,X_k)\in C\}}\big] \\ &= \int_C h(x_0,x_k)f(x_0)f(x_1)\dots f(x_{k-1})f(x_k)\mathrm{d}x_0\dots\mathrm{d}x_k\,.\end{aligned}$$

D'après le théorème de Fubini–Tonelli, on peut intégrer d'abord par rapport à $x_1,\dots,x_{k-1}$: de façon analogue à ce qu'on a vu plus haut, on obtient

$$\begin{aligned}&\mathrm{E}\big[h(X_0,Y)\mathbb{1}_{\{T=k\}}\big] \\ &= \int_{\mathbb{R}^2}\left(\int_{(-\infty,x_0]^{k-1}} f(x_1)\dots f(x_{k-1})\mathrm{d}x_1\dots\mathrm{d}x_{k-1}\right)h(x_0,x_k)f(x_0)f(x_k)\mathbb{1}_{\{x_0<x_k\}}\mathrm{d}x_0\mathrm{d}x_k \\ &= \int_{\mathbb{R}^2} F(x_0)^{k-1}h(x_0,x_k)f(x_0)f(x_k)\mathbb{1}_{\{x_0<x_k\}}\mathrm{d}x_0\mathrm{d}x_k\end{aligned}$$

(v) Notons que comme $\{T=k\}_{k\in\mathbb{N}^*}$ forme une partition de Ω, on a

$$\begin{aligned}\mathrm{E}\big[h(X_0,Y)\big] &= \mathrm{E}\left[\sum_{k=1}^{\infty} h(X_0,Y)\mathbb{1}_{\{T=k\}}\right] = \sum_{k=1}^{\infty}\mathrm{E}\big[h(X_0,Y)\mathbb{1}_{\{T=k\}}\big] \\ &= \sum_{k=1}^{\infty}\int_{\mathbb{R}^2} F(x)^{k-1}h(x,y)f(x)f(y)\mathbb{1}_{\{x<y\}}\mathrm{d}x\mathrm{d}y \\ &= \int_{\mathbb{R}^2}\sum_{k=1}^{\infty} h(x,y)F(x)^{k-1}f(x)f(y)\mathbb{1}_{\{x<y\}}\mathrm{d}x\mathrm{d}y\,,\end{aligned}$$

où on a interverti la somme (infinie) avec l'espérance puis avec l'intégrale, comme suggéré. Notons que si $F(x) = 1$ alors $\mathrm{P}(X \in [x,+\infty[\,) = 0$ (rappelons que $\mathrm{P}(X = x) = 0$) et on peut donc supposer que $f(x) = 0$: la somme est par convention nulle. Dans le cas où $F(x) < 1$, d'après la série géométrique (0.5) on obtient que $\sum_{k=1}^{\infty} F(x)^{k-1} = \frac{1}{1-F(x)}$.
On obtient donc

$$\mathrm{E}\big[h(X_0,Y)\big] = \int_{\mathbb{R}^2} h(x,y)\frac{1}{1-F(x)}f(x)f(y)\mathbb{1}_{\{x<y\}}\mathrm{d}x\mathrm{d}y\,.$$

Comme la fonction h est arbitraire (bornée et continue sauf au plus sur un ensemble de mesure 2-dimensionnelle nulle), de l'Observation 6.58 on déduit que le vecteur aléatoire (X_0,Y) est absolument continu, de densité

$$f_{X_0,Y}(x,y) = \frac{1}{1-F(x)}f(x)f(y)\mathbb{1}_{\{x<y\}}\,.$$

(Par convention, $f_{X_0,Y}(x,y) = 0$ si $F(x) = 1$, comme mentionné plus haut).
La densité de Y est alors donnée par

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{+\infty} f_{X_0,Y}(x,y)\mathrm{d}x = f(y)\int_{-\infty}^{y} \frac{f(x)}{1-F(x)}\mathrm{d}x \\ &= f(y)\Big[-\log\big(1-F(x)\big)\Big]_{-\infty}^{y} = -f(y)\log(1-F(y)), \end{aligned}$$

où on a utilisé le fait que $F'(x) = f(x)$ et que $\lim_{x\to-\infty} F(x) = 0$. De nouveau, par convention, on a $f_Y(y) = 0$ si $F(y) = 1$ (dans ce cas on a $f(y) = 0$).

(vi) Considérons la fonction $\varphi : \{(x,y) \in \mathbb{R}^2 : \; x < y\} \to \mathbb{R}\times\,]0,+\infty[$ définie par $\varphi(x,y) = (x, y-x)$. On a $\varphi^{-1}(u,w) = (v, v+w)$, d'où

$$J_{\varphi^{-1}}(v,w) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

et $\det J_{\varphi^{-1}}(v,w) = 1$. Comme le vecteur aléatoire (X_0,Y) est absolument continu, d'après la Proposition 6.60 le vecteur aléatoire $(X_0,W) = \varphi(X_0,Y)$ est lui aussi absolument continu, de densité

$$f_{X_0,W}(x,w) = f_{X_0,Y}(\varphi^{-1}(x,w)) = \frac{1}{1-F(x)} f(x) f(x+w) \mathbb{1}_{]0,+\infty[}(w).$$

La densité de W est donc donnée par

$$f_W(w) = \int_{-\infty}^{+\infty} \frac{1}{1-F(x)} f(x) f(x+w)\mathrm{d}x \; \mathbb{1}_{]0,+\infty[}(w).$$

Dans le cas particulier où $X_i \sim \mathrm{Exp}(\lambda)$, c'est-à-dire $f(x) = \lambda \mathrm{e}^{-\lambda x}\mathbb{1}_{]0,\infty[}(x)$, on a $F(x) = (1-\mathrm{e}^{-\lambda x})\mathbb{1}_{]0,\infty[}(x)$. On trouve ainsi

$$\begin{aligned} f_{X_0,W}(x,w) &= \frac{1}{\mathrm{e}^{-\lambda x}} \lambda \mathrm{e}^{-\lambda x} \mathbb{1}_{]0,+\infty[}(x) \lambda \mathrm{e}^{-\lambda(x+w)} \mathbb{1}_{]0,+\infty[}(w) \\ &= \lambda \mathrm{e}^{-\lambda x} \mathbb{1}_{]0,+\infty[}(x) \lambda \mathrm{e}^{-\lambda w} \mathbb{1}_{]0,+\infty[}(w), \end{aligned}$$

dont on déduit que X_0 et Y sont indépendantes, avec $X_0 \sim \mathrm{Exp}(\lambda)$, $W \sim \mathrm{Exp}(\lambda)$.

Exercice 6.61 (Concentration de la mesure, I). Soient $X_1,\dots,X_n$ des variables aléatoires indépendantes et de même loi, de moment d'ordre deux $\mathrm{E}(X^2) = 1$, et de moment d'ordre quatre fini $m_4 := \mathrm{E}(X^4) < +\infty$. On considère le vecteur aléatoire

$$V_n := (X_1,\dots,X_n) \in \mathbb{R}^n,$$

et on note $Z_n := \|V_n\|^2 = \sum_{i=1}^n X_i^2$ le carré de sa norme euclidienne.

(i) Calculer $\mathrm{E}(Z_n)$ et $\mathrm{Var}(Z_n)$.

(ii) Déduire de l'inégalité de Jensen que $\mathrm{E}(\|V_n\|) \le \sqrt{n}$.

(iii) Soit $R > 0$ un réel. Montrer que pour tout $n \ge 1$ on a

$$\mathrm{P}\big(|\|V_n\| - \sqrt{n}| \ge R\big) \le \mathrm{P}\big(|Z_n - n| \ge R\sqrt{n}\big) \le \frac{m_4 - 1}{R^2}. \tag{S6.19}$$

[*Sugg.* Distinguer les cas $\{\|V_n\| \ge \sqrt{n}+R\}$ et $\{\|V_n\| \le \sqrt{n}-R\}$ (dans le deuxième cas on a $\sqrt{n} \ge R$).]

(iv) Supposons que $X_i \sim \mathrm{N}(0,1)$ pour tout $i \geq 1$. Calculer m_4 et montrer l'affirmation suivante : *pour tout $n \in \mathbb{N}^*$, le vecteur V_n est de norme comprise entre $\sqrt{n}-10$ et $\sqrt{n}+10$ avec une probabilité supérieure à 98%.* Ainsi, avec grande probabilité, le vecteur V_n se trouve dans « l'écorce » entre deux sphères de rayon $\sqrt{n} \pm 5$ (le rayon diverge, mais l'épaisseur reste constante !).

Solution 6.61. (i) Par linéarité de l'espérance, on a

$$\mathrm{E}(Z_n) = \sum_{i=1}^{n} \mathrm{E}(X_i^2) = n\,.$$

D'autre part, comme les variables aléatoires $X_1, \ldots, X_n$ sont indépendantes, les variables $X_1^2, \ldots, X_n^2$ le sont elles aussi : grâce aux propriétés de la variance (voir notamment le Corollaire 3.75), on obtient

$$\mathrm{Var}(Z_n) = \sum_{i=1}^{n} \mathrm{Var}(X_i^2) = n(m_4 - 1)\,,$$

où on a utilisé que $\mathrm{Var}(X_i^2) = \mathrm{E}(X_i^4) - \mathrm{E}(X_i^2)^2 = m_4 - 1$.

(ii) Comme la fonction $x \mapsto \sqrt{x}$ est concave, d'après l'inégalité de Jensen on a

$$\mathrm{E}(\|V_n\|) = \mathrm{E}(\sqrt{Z_n}) \leq \sqrt{\mathrm{E}(Z_n)} = \sqrt{n}\,.$$

(Notons que si φ est concave, en supposant que $\mathrm{E}(X) < +\infty$, il suffit d'appliquer l'inégalité de Jensen à la fonction convexe $-\varphi$ pour obtenir que $\mathrm{E}(\varphi(X)) \leq \varphi(\mathrm{E}(X))$.)

(iii) Observons que l'on a l'égalité d'événements

$$\left\{\left|\|V_n\| - \sqrt{n}\right| \geq R\right\} = \left\{\|V_n\| \geq \sqrt{n} + R\right\} \cup \left\{\|V_n\| \leq \sqrt{n} - R\right\}.$$

Maintenant, notons que

$$\left\{\|V_n\| \geq \sqrt{n} + R\right\} = \left\{Z_n \geq (\sqrt{n} + R)^2\right\} \subseteq \left\{Z_n - n \geq R\sqrt{n}\right\},$$

car $(\sqrt{n}+R)^2 = n + 2R\sqrt{n} + R^2 \geq n + R\sqrt{n}$. Notons aussi que l'événement $\{\|V_n\| \leq \sqrt{n} - R\}$ est vide si $\sqrt{n} - R < 0$ et que si $\sqrt{n} - R \geq 0$

$$\left\{\|V_n\| \leq \sqrt{n} - R\right\} = \left\{Z_n \leq (\sqrt{n} - R)^2\right\} \subseteq \left\{Z_n - n \leq -R\sqrt{n}\right\},$$

car $(\sqrt{n}-R)^2 = n - 2R\sqrt{n} + R^2 \leq n - R\sqrt{n}$ (parce que $R \leq \sqrt{n}$).
Pour résumer, on a

$$\begin{aligned}&\left\{\left|\|V_n\| - \sqrt{n}\right| \geq R\right\} \\ &\qquad \subseteq \left\{Z_n - n \geq R\sqrt{n}\right\} \cup \left\{Z_n - n \leq -R\sqrt{n}\right\} = \left\{|Z_n - n| \geq R\sqrt{n}\right\}.\end{aligned}$$

On en conclut que

$$\mathrm{P}\left(\left|\|V_n\| - \sqrt{n}\right| \geq R\right) \leq \mathrm{P}\left(|Z_n - n| \geq R\sqrt{n}\right) \leq \frac{\mathrm{Var}(Z_n)}{R^2 n}\,,$$

où on a utilisé l'inégalité de Bienaymé–Tchebychev (Proposition 3.77), en notant que $n = \mathrm{E}(Z_n)$. D'après le calcul de la question précédente on a $\mathrm{Var}(Z_n) = n(m_4 - 1)$, dont découle l'inégalité annoncée.

(iv) Si $X_i \sim \mathrm{N}(0,1)$, on a en effet $\mathrm{E}(X_i^2) = 1$ et on trouve $m_4 = 3$.

En effet, la fonction génératrice des moments de X_i est $\mathrm{M}_{X_1}(t) = \mathrm{e}^{\frac{1}{2}t^2}$ pour tout $t \in \mathbb{R}$ (voir (6.49)), donc $m_4 = \mathrm{E}(X_i^4) = \mathrm{M}_{X_1}^{(4)}(0)$, où $\mathrm{M}_{X_1}^{(4)}(t) = (t^4 + 6t^2 + 3)\mathrm{e}^{\frac{1}{2}t^2}$ est la dérivée quatrième de M_{X_1}.

Une autre manière de trouver $m_4 = 3$ est de calculer directement, avec une intégration par parties :

$$\mathrm{E}(X_1^4) = \int_{-\infty}^{+\infty} x^4 \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x = \left[x^3 \frac{-1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}}\right]_{-\infty}^{+\infty} + 3\int_{-\infty}^{+\infty} x^2 \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{d}x = 3$$

où on a remarqué que la dernière intégrale est égale à $\mathrm{E}(X_1^2) = 1$.

Ainsi, d'après la question précédente, on a

$$\mathrm{P}\Big(\|V_n\| \notin \,]\sqrt{n} - 10, \sqrt{n} + 10[\,\Big) = \mathrm{P}\big(\big|\|V_n\| - \sqrt{n}\big| \geq 10\big) \leq \frac{2}{10^2}\,.$$

En passant au complémentaire, on obtient

$$\mathrm{P}\Big(\|V_n\| \in \,]\sqrt{n} - 10, \sqrt{n} + 10[\,\Big) \geq 1 - \frac{2}{100} = 98\%\,.$$

Exercice 6.62 (Concentration de la mesure, II). Soient $(X_n)_{n\geq 1}$ des variables aléatoires indépendantes de loi uniforme sur $[0,1]$. On considère le vecteur aléatoire

$$W_n = (U_1, \ldots, U_n) \in \mathbb{R}^n\,,$$

qui est de loi uniforme sur l'hypercube $[0,1]^n$.

(i) Pour $\delta \in \,]0, \frac{1}{2}[$, on définit « l'écorce de l'hypercube d'épaisseur δ » :

$$\Gamma(\delta) := [0,1]^n \setminus [\delta, 1-\delta]^n\,.$$

Montrer que $\mathrm{P}(W_n \in \Gamma(\delta)) \geq 1 - \mathrm{e}^{-2\delta n}$, pour tout $n \in \mathbb{N}^*$, et en déduire que *le vecteur W_n appartient à $\Gamma(3/n)$ avec probabilité supérieure à* $99,7\%$

(ii) En utilisant l'Exercice 6.61, montrer que pour tout $n \in \mathbb{N}^*$, *le vecteur W_n est de norme comprise entre $\sqrt{n/3} - 6$ et $\sqrt{n/3} + 6$ avec probabilité supérieure à* $99,2\%$.

[*Sugg.* Observer que $X_i := \sqrt{3}U_i$ a pour moment d'ordre deux $\mathrm{E}(X_i^2) = 1$, puis appliquer l'estimée (S6.19) au vecteur $V_n := (X_1, \ldots, X_n)$.]

(iii) Conclure qu'avec grande probabilité ($> 99\%$) un point choisi uniformément dans l'hypercube $[0,1]^n$ *se trouve dans l'intersection de « l'écorce » d'épaisseur $3/n$ de l'hypercube et « l'écorce » comprise entre deux sphères de rayon* $\sqrt{n/3} \pm 6$.

Solution 6.62. (i) Estimons la probabilité du complémentaire :

$$\begin{aligned}
\mathrm{P}\big(W_n \notin \Gamma(\delta)\big) &= \mathrm{P}\big((U_1, \ldots, U_n) \in \,]\delta, 1-\delta[^n\big) \\
&= \mathrm{P}\big(U_i \in \,]\delta, 1-\delta[\ \forall 1 \leq i \leq n\big) \\
&= \mathrm{P}\big(U_1 \in \,]\delta, 1-\delta[\,\big)^n = \big(1-2\delta\big)^n,
\end{aligned}$$

où on a utilisé le fait que $U_1, \ldots, U_n$ sont indépendantes et de loi $\mathrm{U}(0,1)$. On utilise maintenant que $1 - x \leq \mathrm{e}^{-x}$ pour tout $x \in \mathbb{R}$ pour obtenir

$$\mathrm{P}\big(W_n \notin \Gamma(\delta)\big) \leq \mathrm{e}^{-2\delta n}\,.$$

En passant au complémentaire, on obtient l'estimée annoncée. En l'appliquant à $\delta = 3/n$, on trouve $\mathrm{e}^{-2\delta n} = \mathrm{e}^{-6} \approx 0,00248$:

$$\mathrm{P}\big(W_n \in \Gamma(3/n)\big) \geq 1-\mathrm{e}^{-6} \geq 99,7\%\,.$$

(ii) On a $\mathrm{E}(U_i^2) = \int_0^1 x^2 \mathrm{d}x = \frac{1}{3}$, donc $X_i = \sqrt{3}\,U_i$ a pour moment d'ordre deux $\mathrm{E}(X_i^2) = 3\,\mathrm{E}(U_i^2) = 1$. Le moment d'ordre quatre de X_i vaut $m_4 = \mathrm{E}(X_i^4) = 9\,\mathrm{E}(U_i^4)$ avec $\mathrm{E}(U_i^4) = \int_0^1 x^4 \mathrm{d}x = \frac{1}{5}$. En appliquant le résultat de l'Exercice 6.61 à $V_n = (X_1,\dots,X_n)$ ($m_4 = \frac{9}{5}$), on obtient

$$\mathrm{P}\Big(\big|\|V_n\| - \sqrt{n}\big| > R\Big) \leq \frac{4}{5R^2}\,.$$

Comme $\|W_n\| = \|V_n\|/\sqrt{3}$, on en déduit que

$$\mathrm{P}\Big(\big|\|W_n\| - \sqrt{n/3}\big| \geq R/\sqrt{3}\Big) \leq \frac{4}{5R^2}\,.$$

En utilisant cette inégalité avec $R = 6\sqrt{3}$, on a

$$\mathrm{P}\Big(\big|\|W_n\| - \sqrt{n/3}\big| \geq 6\Big) \leq \frac{1}{135} \approx 0,00741\,.$$

En passant au complémentaire on en conclut que

$$\begin{aligned}\mathrm{P}\Big(\|W_n\| \in \big]\sqrt{n/3}-6, \sqrt{n/3}+6\big[\Big) &= \mathrm{P}\Big(\big|\|W_n\| - \sqrt{n/3}\big| < 6\Big)\\ &\geq 1-\frac{1}{135} \geq 99,2\%\,.\end{aligned}$$

(iii) Notons que pour toute paire d'événements A,B,

$$\mathrm{P}(A\cap B) = 1-\mathrm{P}(A^c\cup B^c) \geq 1-\mathrm{P}(A^c)-\mathrm{P}(B^c) = \mathrm{P}(A)+\mathrm{P}(B)-1\,.$$

En mettant ensemble les deux points précédents, on obtient

$$\begin{aligned}&\mathrm{P}\Big(W_n \in \Gamma(n/3), \|W_n\| \in \big]\sqrt{n/3}-6, \sqrt{n/3}+6\big[\Big)\\ &\quad\geq P\Big(W_n \in \Gamma(n/3)\Big) + \mathrm{P}\Big(\|W_n\| \in \big]\sqrt{n/3}-6, \sqrt{n/3}+6\big[\Big) - 1\\ &\quad\geq 1-\mathrm{e}^{-6}+1-\frac{1}{135}-1 = 1-\mathrm{e}^{-6}-\frac{1}{135} \approx 99,01\%\,,\end{aligned}$$

ce qui correspond à l'affirmation que l'on souhaitait démontrer.

Exercice 6.63 (Une loi du 0-1). Soit $(X_i)_{i\geq 1}$ une suite de variables aléatoires réelles indépendantes et de même loi. On pose $S_0 = 0$ et pour $n \geq 1$, $S_n = \sum_{i=1}^n X_i$. On pourra interpréter la suite $(S_n)_{n\geq 0}$ comme la suite des positions d'un marcheur aléatoire, la variable aléatoire X_i donnant la distance (algébrique) parcourue lors du i-ème pas ; cela généralise la marche aléatoire simple considérée dans la Section 4.4, où les pas étaient supposés de longueur 1. On s'intéresse aux propriétés de la suite $(S_n)_{n\geq 0}$: le but de l'exercice est de montrer que les probabilités $\mathrm{P}(\sup_{n\geq 0} S_n = +\infty)$ et $\mathrm{P}(\inf_{n\geq 0} S_n = -\infty)$ ne peuvent valoir que 0 ou bien 1.

Pour $k \in \mathbb{Z}$, on introduit la variable aléatoire $T_k := \min\{n\,, S_n \geq k\}$, avec par convention $\min\emptyset = +\infty$.

(i) Montrer que l'on a l'égalité d'événements

$$\left\{\sup_{n\geq 0} S_n = +\infty\right\} = \bigcap_{k=1}^{\infty} \{T_k < +\infty\}\,.$$

En déduire que $\mathrm{P}(\sup_{n\geq 0} S_n = +\infty) = \lim_{k\to\infty} \mathrm{P}(T_k < +\infty)$.

(ii) Soit $k \in \mathbb{N}^*$ fixé. Montrer que l'on a l'inclusion d'événements suivante :

$$\left\{\sup_{n\geq 0} S_n = +\infty\right\} \subseteq \bigcup_{m=0}^{\infty} \{T_k = m\} \cap \left\{\sup_{n\geq m}(S_n - S_m) = +\infty\right\}.$$

En déduire

$$\mathrm{P}\left(\sup_{n\geq 0} S_n = +\infty\right) \leq \sum_{m=0}^{\infty} \mathrm{P}(T_k = m)\,\mathrm{P}\left(\sup_{n\geq m}(S_n - S_m) = +\infty\right).$$

(iii) Montrer que $\mathrm{P}(\sup_{n\geq m}(S_n - S_m) = +\infty) = \mathrm{P}(\sup_{n\geq 0} S_n = +\infty)$. En déduire que pour $k \in \mathbb{N}^*$ fixé on a l'inégalité

$$\mathrm{P}\left(\sup_{n\geq 0} S_n = +\infty\right) \leq \mathrm{P}(T_k < +\infty)\,\mathrm{P}\left(\sup_{n\geq 0} S_n = +\infty\right).$$

(iv) Conclure que l'on a la dichotomie suivante :

- si $\exists k \in \mathbb{N}^*$ tel que $\mathrm{P}(T_k < +\infty) < 1$, alors $\mathrm{P}(\sup_{n\geq 0} S_n = +\infty) = 0$;
- si $\forall k \in \mathbb{N}^*$ on a $\mathrm{P}(T_k < +\infty) = 1$, alors $\mathrm{P}(\sup_{n\geq 0} S_n = +\infty) = 1$.

(v) Montrer de même que la probabilité $\mathrm{P}(\inf_{n\geq 0} S_n = -\infty)$ ne peut valoir que 0 ou 1.

Solution 6.63. (i) Notons que $\sup_{n\geq 0} S_n \geq k$ si et seulement s'il existe un entier n tel que $S_n \geq k$, c'est-à-dire si et seulement si $T_k := \min\{n : S_n \geq k\} < +\infty$. On a donc l'égalité d'événements

$$\left\{\sup_{n\geq 0} S_n \geq k\right\} = \{T_k < +\infty\}\,.$$

En prenant l'intersection pour $k \in \mathbb{N}^*$, on obtient

$$\bigcap_{k=1}^{\infty} \{T_k < +\infty\} = \bigcap_{k=1}^{\infty} \left\{\sup_{n\geq 0} S_n \geq k\right\} = \left\{\sup_{n\geq 0} S_n = +\infty\right\}.$$

Maintenant, observons que la suite d'événements $(\{T_k < +\infty\})_{k\geq 1}$ est décroissante pour l'inclusion, parce que si $T_{k+1} < +\infty$ alors nécessairement $T_k < +\infty$: grâce à la continuité par le haut des probabilités (Proposition 1.24), on obtient

$$\mathrm{P}\left(\bigcap_{k=1}^{\infty} \{T_k < +\infty\}\right) = \lim_{k\to+\infty} \mathrm{P}(T_k < +\infty)\,.$$

De l'identité précédente, on en déduit que $\mathrm{P}(\sup_{n\geq 0} S_n = +\infty) = \lim_{k\to\infty} \mathrm{P}(T_k < +\infty)$.

(ii) On sait du point précédent que $\{\sup_{n\geq 0} S_n = +\infty\} = \bigcap_{j=1}^{\infty} \{T_j < +\infty\}$: on en déduit que pour tout $k \in \mathbb{N}^*$ fixé

$$\begin{aligned}\Big\{\sup_{n\geq 0} S_n = +\infty\Big\} &= \{T_k < +\infty\} \cap \Big\{\sup_{n\geq 0} S_n = +\infty\Big\} \\ &= \bigcup_{m=0}^{\infty} \bigg(\{T_k = m\} \cap \Big\{\sup_{n\geq 0} S_n = +\infty\Big\}\bigg).\end{aligned}$$

Maintenant, remarquons que, pour tout m fixé, avoir $\sup_{n\geq 0} S_n = +\infty$ est équivalent à avoir $\sup_{n\geq m} S_n = +\infty$ (parce que $\sup_{n<m} S_n < +\infty$), ce qui est aussi équivalent à avoir $\sup_{n\geq m}(S_n - S_m) = +\infty$ (parce que $S_m < +\infty$). On a donc

$$\Big\{\sup_{n\geq 0} S_n = +\infty\Big\} = \bigcup_{m=0}^{\infty} \bigg(\{T_k = m\} \cap \Big\{\sup_{n\geq m}(S_n - S_m) = +\infty\Big\}\bigg).$$

Par sous-additivité, on obtient

$$\mathrm{P}\Big(\sup_{n\geq 0} S_n = +\infty\Big) \leq \sum_{m=0}^{\infty} \mathrm{P}\Big(T_k = m\,,\ \sup_{n\geq m}(S_n - S_m) = +\infty\Big).$$

Maintenant, notons que $\{T_k = m\} = \{S_1 < k, \ldots, S_{m-1} < k, S_m \geq k\}$ est un événement qui ne dépend que de $X_1, \ldots, X_m$ et que d'autre part on a $S_n - S_m = \sum_{i=m+1}^{n} X_i$, de sorte que l'événement $\{\sup_{n\geq m}(S_n - S_m) = +\infty\}$ dépend seulement de $X_{m+1}, X_{m+2}, \ldots$. Par indépendance par paquets, ces deux événements sont indépendants, donc

$$\mathrm{P}\Big(\sup_{n\geq 0} S_n = +\infty\Big) \leq \sum_{m=0}^{\infty} \mathrm{P}(T_k = m)\,\mathrm{P}\Big(\sup_{n\geq m}(S_n - S_m) = +\infty\Big).$$

(iii) En posant $\tilde{S}_j := S_{m+j} - S_m$, c'est-à-dire $\tilde{S}_0 = 0$ et $\tilde{S}_j = \sum_{i=1}^{j} X_{m+i}$, il est intuitivement clair que « $(\tilde{S}_j)_{j\geq 0}$ a la même loi que $(S_j)_{j\geq 0}$ », dont on pourrait déduire que l'on a $\mathrm{P}(\sup_{j\geq 0} \tilde{S}_j = +\infty) = \mathrm{P}(\sup_{j\geq 0} S_n = +\infty)$. Le problème est que l'on n'a pas défini la notion de loi pour une suite infinie.

On procède de la façon suivante. En définissant $M_n := \max_{0\leq j\leq n} S_j$, on a

$$\mathrm{P}\Big(\sup_{n\geq 0} S_n = +\infty\Big) = \mathrm{P}\bigg(\bigcap_{k\geq 1}\bigcup_{n\geq 1}\{M_n \geq k\}\bigg) = \lim_{k\to+\infty}\lim_{n\to+\infty} \mathrm{P}\big(M_n \geq k\big)$$

où on a successivement utilisé la continuité par le haut des probabilités pour la suite décroissante d'événements $A_k := \bigcup_{n\geq 1}\{M_n \geq k\}$ en $k \geq 1$, puis la continuité par le bas de la probabilité pour les suites croissantes d'événements $(\{M_n \geq k\})_{n\geq 1}$ pour $k \in \mathbb{N}^*$ fixé (voir la Proposition 1.24).

On définit maintenant $\tilde{M}_n := \max_{0\leq j\leq n} \tilde{S}_j$ (rappelons que $\tilde{S}_j := S_{m+j} - S_m$), et notons que $\tilde{M}_n$ a la même loi que M_n, pour tout n. Cela découle du fait que le vecteur aléatoire $(S_0, \ldots, S_n)$ a la même loi que le vecteur aléatoire $(\tilde{S}_0, \ldots, \tilde{S}_n)$ et que l'on a appliqué la même fonction $f : (x_1, \ldots, x_n) \mapsto (0, x_1, x_1 + x_2, \ldots, x_1 + \cdots + x_n)$ aux deux vecteurs aléatoires. On en déduit que $\mathrm{P}(M_n \geq k) = \mathrm{P}(\tilde{M}_n \geq k)$ pour tout $n \in \mathbb{N}$ et tout $k \geq 1$, donc

$$\begin{aligned}\mathrm{P}\Big(\sup_{n\geq 0} S_n = +\infty\Big) &= \lim_{k\to+\infty}\lim_{n\to+\infty} \mathrm{P}\big(M_n \geq k\big) \\ &= \lim_{k\to+\infty}\lim_{n\to+\infty} \mathrm{P}\big(\tilde{M}_n \geq k\big) = \mathrm{P}\Big(\sup_{n\geq 0} \tilde{S}_n = +\infty\Big),\end{aligned}$$

qui est l'identité recherchée.

Par conséquent, d'après la question précédente, on obtient que

$$\begin{aligned}\mathrm{P}\left(\sup_{n\geq 0} S_n = +\infty\right) &\leq \left(\sum_{m=0}^{\infty} \mathrm{P}(T_k = m)\right) \mathrm{P}\left(\sup_{n\geq 0} S_n = +\infty\right) \\ &= \mathrm{P}(T_k < +\infty)\, \mathrm{P}\left(\sup_{n\geq 0} S_n = +\infty\right).\end{aligned}$$

(iv) Partons de l'inégalité démontrée à la question précédente, qui est valide pour tout $k \in \mathbb{N}^*$ fixé.

• S'il existe un $k \in \mathbb{N}^*$ tel que $\mathrm{P}(T_k < +\infty) < 1$, alors si $\mathrm{P}(\sup_{n\geq 0} S_n = +\infty) > 0$, l'inégalité précédente donne

$$\mathrm{P}\left(\sup_{n\geq 0} S_n = +\infty\right) < \mathrm{P}\left(\sup_{n\geq 0} S_n = +\infty\right),$$

ce qui est une contradiction. On en conclut que $\mathrm{P}(\sup_{n\geq 0} S_n = +\infty) = 0$ dans ce cas.

• Si $\mathrm{P}(T_k < +\infty) = 1$ pour tout $k \in \mathbb{N}^*$, la première question de l'exercice nous assure que

$$\mathrm{P}\left(\sup_{n\geq 0} S_n = +\infty\right) = \lim_{k\to+\infty} \mathrm{P}\left(T_k < +\infty\right) = 1\,.$$

(v) Définissons $\hat{X}_i := -X_i$ et $\hat{S}_n = \sum_{i=1}^n \hat{X}_i = -S_n$. On a $\inf_{n\geq 0} S_n = -\sup_{n\geq 0} \hat{S}_n$, donc

$$\mathrm{P}\left(\inf_{n\geq 0} S_n = -\infty\right) = \mathrm{P}\left(\sup_{n\geq 0} \hat{S}_n = +\infty\right),$$

dont on déduit le résultat voulu, en appliquant le point précédent à $\hat{S}_n$ au lieu de S_n.

Chapitre 7
Théorèmes limites

7.1 La loi des grands nombres

Exercice 7.1. Soit $(X_i)_{i\geq 1}$ une suite de variables aléatoires d'espérance μ finie, qui satisfait la loi faible des grands nombres (voir la Définition 7.1). On note $\overline{X}_n := \frac{1}{n}\sum_{i=1}^n X_i$ la moyenne empirique.

(i) Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction *continue au point* μ. Montrer que pour tout $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathrm{P}\left(\left|f(\overline{X}_n) - f(\mu)\right| > \varepsilon\right) = 0\,.$$

(ii) On suppose de plus que f est *bornée*, c'est-à-dire qu'il existe $M < \infty$ tel que $|f(x)| \leq M$ pour tout $x \in \mathbb{R}$. Déduire du point précédent que

$$\lim_{n\to+\infty} \mathrm{E}\left(f(\overline{X}_n)\right) = f(\mu)\,.$$

Solution 7.1. (i) Par définition de la continuité de f au point μ, pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que pour $|x-\mu| < \delta$ on a $|f(x) - f(\mu)| < \varepsilon$; Par conséquent, si $|f(x) - f(\mu)| > \varepsilon$ on a $|x-\mu| \geq \delta$, ce qui implique que l'on a l'inclusion d'événements

$$\{|f(\overline{X}_n) - f(\mu)| > \varepsilon\} \subseteq \{|\overline{X}_n - \mu| > \delta\}\,.$$

On a alors, d'après la loi des grand nombres,

$$\lim_{n\to\infty} \mathrm{P}\left(\left|f(\overline{X}_n) - f(\mu)\right| > \varepsilon\right) \leq \lim_{n\to\infty} \mathrm{P}\left(\left|\overline{X}_n - \mu\right| > \delta\right) = 0\,.$$

(ii) Soit donc $M \in \,]0,\infty[$ tel que

$$|f(x)| \leq M\,, \qquad \forall x \in \mathbb{R}\,.$$

Fixons $\varepsilon > 0$ et décomposons

$$\begin{aligned} \mathrm{E}\left[f(\overline{X}_n)\right] - f(\mu) &= \mathrm{E}\left[f(\overline{X}_n) - f(\mu)\right] \\ &= \mathrm{E}\left[\left(f(\overline{X}_n) - f(\mu)\right)\mathbb{1}_{\{|f(\overline{X}_n)-f(\mu)|\leq\varepsilon\}}\right] \\ &\qquad + \mathrm{E}\left[\left(f(\overline{X}_n) - f(\mu)\right)\mathbb{1}_{\{|f(\overline{X}_n)-f(\mu)|>\varepsilon\}}\right]\,, \end{aligned}$$

donc

$$\begin{aligned} \left|\mathrm{E}\left[f(\overline{X}_n)\right] - f(\mu)\right| \leq \mathrm{E}\left[\left|f(\overline{X}_n) - f(\mu)\right|\mathbb{1}_{\{|f(\overline{X}_n)-f(\mu)|\leq\varepsilon\}}\right] \\ + \mathrm{E}\left[\left|f(\overline{X}_n) - f(\mu)\right|\mathbb{1}_{\{|f(\overline{X}_n)-f(\mu)|>\varepsilon\}}\right]\,. \end{aligned}$$

Considérons les deux espérances dans le membre de droite : dans le premier, on peut utiliser l'estimée $|f(\overline{X}_n) - f(\mu)| \leq \varepsilon$ (grâce à la fonction indicatrice) et dans le deuxième on peut utiliser la majoration $|f(\overline{X}_n) - f(\mu)| \le |f(\overline{X}_n)| + |f(\mu)| \le 2M$ (car f est bornée : $|f(x)| \le M$ pour tout $x \in \mathbb{R}$). Par conséquent

$$\left|\mathrm{E}\left[f(\overline{X}_n)\right] - f(\mu)\right| \le \varepsilon\, \mathrm{P}\left(|f(\overline{X}_n) - f(\mu)| \le \varepsilon\right) + 2M\, \mathrm{P}\left(|f(\overline{X}_n) - f(\mu)| > \varepsilon\right).$$

En majorant $\mathrm{P}\left(|f(\overline{X}_n) - f(\mu)| \le \varepsilon\right) \le 1$, grâce au point précédent on obtient

$$\limsup_{n\to\infty} \left|\mathrm{E}\left[f(\overline{X}_n)\right] - f(\mu)\right| \le \varepsilon + 2M \lim_{n\to\infty} \mathrm{P}\left(|f(\overline{X}_n) - f(\mu)| > \varepsilon\right) = \varepsilon.$$

Comme $\varepsilon > 0$ est arbitraire, on a $\limsup_{n\to\infty} |\mathrm{E}[f(\overline{X}_n)] - f(\mu)| = 0$ ce qui équivaut à $\lim_{n\to\infty} \mathrm{E}[f(\overline{X}_n)] = f(\mu)$.

Exercice 7.2. Soit $(X_i)_{i\ge 1}$ une suite de variables aléatoires (pas nécessairement de même loi) de même espérance μ et dans L^2. On suppose qu'il existe une fonction $\rho : \mathbb{N} \to \mathbb{R}_+$ telle que, pour tout $i, j \in \mathbb{N}^*$ on ait $|\mathrm{Cov}(X_i, X_j)| \le \rho(|j-i|)$. Montrer que si $\lim_{k\to\infty} \rho(k) = 0$, alors la suite $(X_i)_{i\ge 1}$ satisfait la loi faible des grands nombres, c'est-à-dire (7.2).
[*Sugg.* On pourra majorer la variance $\mathrm{Var}(\bar{X}_n)$ par $\frac{1}{n}(\rho(0) + 2\sum_{k=1}^n \rho(k))$.]

Solution 7.2. Rappelons que $\overline{X}_n = \frac{S_n}{n}$ où $S_n := X_1 + \cdots + X_n$. D'après l'inégalité de Bienaymé–Tchebychev, comme $\mathrm{E}(\overline{X}_n) = \mu$, on peut écrire

$$\mathrm{P}(|\overline{X}_n - \mu| \ge \varepsilon) \le \frac{\mathrm{Var}(\overline{X}_n)}{\varepsilon^2}.$$

Il suffit donc de montrer que $\lim_{n\to\infty} \mathrm{Var}(\overline{X}_n) = 0$.

Avec la formule

$$\mathrm{Var}(S_n) = \sum_{i=1}^n \mathrm{Var}(X_i) + 2 \sum_{1\le i<j\le n} \mathrm{Cov}(X_i, X_j),$$

en notant que $\mathrm{Var}(X_i) = \mathrm{Cov}(X_i, X_i) \le \rho(0)$ par hypothèse, on obtient l'estimée

$$\begin{aligned}
\mathrm{Var}(S_n) &\le \sum_{i=1}^n \mathrm{Var}(X_i) + 2 \sum_{1\le i<j\le n} |\mathrm{Cov}(X_i, X_j)| \\
&\le \sum_{i=1}^n \rho(0) + 2 \sum_{1\le i<j\le n} \rho(j-i) = n\rho(0) + 2\sum_{i=1}^{n-1}\sum_{k=1}^{n-i} \rho(k) \\
&\le n\left(\rho(0) + 2\sum_{k=1}^{n-1} \rho(k)\right),
\end{aligned}$$

où pour l'égalité on a utilisé le changement de variables $k = j - i$ et pour la dernière inégalité on a étendu la somme sur k. Par conséquent,

$$\mathrm{Var}(\overline{X}_n) = \frac{\mathrm{Var}(S_n)}{n^2} \le \frac{\rho(0)}{n} + \frac{2}{n}\sum_{k=1}^{n-1} \rho(k).$$

Le premier terme du membre de droite tend vers 0 quand $n \to \infty$ et il en va de même pour le deuxième terme, grâce au lemme de Cesàro (on le redémontre ci-dessous).

Grâce à l'hypothèse $\lim_{k\to\infty} \rho(k) = 0$, pour tout $\varepsilon > 0$ il existe $\bar{k}_\varepsilon \in \mathbb{N}$ tel que pour tout $k > \bar{k}_\varepsilon$ on a $\rho(k) < \varepsilon$. Si on définit $C_\varepsilon := \max_{k=1,\ldots,\bar{k}_\varepsilon} \rho(k)$, on peut écrire, pour $k \ge \bar{k}_\varepsilon$,

$$\sum_{k=1}^{n-1} \rho(k) = \sum_{k=1}^{\bar{k}_\varepsilon} \rho(k) + \sum_{k=\bar{k}_\varepsilon+1}^{n-1} \rho(k) \leq \bar{k}_\varepsilon C_\varepsilon + n\varepsilon\,,$$

donc

$$\limsup_{n\to\infty} \frac{2}{n} \sum_{k=1}^{n-1} \rho(k) \leq \limsup_{n\to\infty} \frac{2}{n} \big(k_\varepsilon C_\varepsilon + n\varepsilon\big) = 2\varepsilon\,.$$

Comme $\varepsilon > 0$ est arbitraire, cela conclut la démonstration.

Exercice 7.3. Soit X une variable aléatoire réelle *positive*, d'espérance $\mathrm{E}(X) = +\infty$, et soit $(X_i)_{i\geq 1}$ une suite de variables aléatoires indépendantes et de même loi que X. On pose $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ et on se fixe un $A > 0$ (grand) arbitraire.

(i) Montrer qu'il existe un réel $B > 0$ (qui dépend de A) tel que $\mathrm{E}(X\mathbb{1}_{\{X\leq B\}}) \geq 2A$, en supposant que X est soit discrète, soit à densité.

[*La démonstration pour une variable aléatoire réelle X générale requiert des outils plus avancés.*]

(ii) On pose $Y_i := X_i \mathbb{1}_{\{X_i\leq B\}}$ pour tout $i \geq 1$. Observer que Y_i admet un moment d'ordre deux fini (pourquoi ?) et déduire de la loi des grands nombres que

$$\lim_{n\to+\infty} \mathrm{P}\left(\frac{1}{n}\sum_{i=1}^n Y_i \geq A\right) = 1\,.$$

(iii) En comparant X_i et Y_i, en déduire que $\lim_{n\to+\infty} \mathrm{P}(\overline{X}_n \geq A) = 1$. Se convaincre qu'il s'agit d'un moyen de formaliser l'affirmation « la moyenne empirique tend vers $+\infty$ ».

Solution 7.3. (i) Il suffit de montrer que l'on a

$$\mathrm{E}(X) = \lim_{B\in\mathbb{N},\, B\to+\infty} \mathrm{E}(X\,\mathbb{1}_{\{X\leq B\}})\,. \tag{S7.1}$$

Ainsi, si $\mathrm{E}(X) = +\infty$, alors pour tout $A > 0$ il existe $B \in \mathbb{N}$ tel que $\mathrm{E}(X\,\mathbb{1}_{\{X\leq B\}}) \geq 2A$.

Démontrons (S7.1) lorsque X est discrète, de densité discrète p_X. Comme X est positive, on a $\mathrm{E}(X) = \sum_{x>0} x\,\mathrm{p}_X(x)$. Grâce à la propriété de *sommation par paquets*, on peut écrire

$$\mathrm{E}(X) = \sum_{n=1}^{\infty} a_n \qquad \text{où} \qquad a_n := \sum_{x\in\,]n-1,n]} x\,\mathrm{p}_X(x)\,.$$

Par définition des séries, on peut alors écrire

$$\mathrm{E}(X) = \lim_{B\to+\infty} \alpha_B \qquad \text{où} \qquad \alpha_B := \sum_{n=1}^{B} a_n\,.$$

On peut enfin réécrire

$$\alpha_B = \sum_{n=1}^{B} \sum_{x\in\,]n-1,n]} x\,\mathrm{p}_X(x) = \sum_{x\in\mathbb{R}} x\,\mathbb{1}_{\,]0,B]}(x)\,\mathrm{p}_X(x) = \mathrm{E}(X\,\mathbb{1}_{\{X\leq B\}})\,,$$

où on a utilisé la formule de transfert $\mathrm{E}(\varphi(X)) = \sum_{x\in\mathbb{R}} \varphi(x)\,\mathrm{p}_X(x)$.

Démontrons (S7.1) quand X est absolument continue, de densité f_X. Comme X est positive, on a $\mathrm{E}(X) = \int_0^\infty x\,f_X(x)\,\mathrm{d}x$ et par définition de l'intégrale de Riemann sur le domaine non borné $]0,\infty[$ on a

$$\mathrm{E}(X) = \lim_{B\to\infty} \int_0^B x\, f_X(x)\, \mathrm{d}x = \lim_{B\to\infty} \int_{-\infty}^{+\infty} x\, \mathbb{1}_{]0,B]}(x)\, f_X(x)\, \mathrm{d}x = \lim_{B\to\infty} \mathrm{E}(X\, \mathbb{1}_{\{X\leq B\}})\,,$$

où on a utilisé la formule de transfert $\mathrm{E}(\varphi(X)) = \int_{-\infty}^{+\infty} \varphi(x)\, f_X(x)\, \mathrm{d}x$.

(ii) Les variables aléatoires $(Y_i)_{i\geq 1}$ sont indépendantes (par conservation de l'indépendance) et identiquement distribuées (par conservation de la loi), parce que l'on peut écrire $Y_i = \varphi(X_i)$ avec $\varphi(x) := x\, \mathbb{1}_{]-\infty,B]}(x)$. En notant $\mu' := \mathrm{E}(Y_i)$ l'espérance de Y_i, on a $\mu' \geq 2A$, d'après le point précédent. De plus, les Y_i admettent un moment d'ordre deux fini, simplement parce qu'elles sont bornées : $|Y_i| \leq B$ par construction. On peut alors appliquer la loi faible des grands nombres à la suite $(Y_i)_{i\geq 1}$, ce qui donne

$$\forall \varepsilon > 0: \qquad \lim_{n\to+\infty} \mathrm{P}\left(\left| \frac{1}{n} \sum_{i=1}^n Y_i - \mu' \right| \geq \varepsilon \right) = 0\,.$$

En prenant $\varepsilon := A$ et en se rappelant que $\mu' \geq 2A$, on obtient l'inclusion d'événements

$$\left\{ \frac{1}{n} \sum_{i=1}^n Y_i < A \right\} \subseteq \left\{ \left| \frac{1}{n} \sum_{i=1}^n Y_i - \mu' \right| \geq \varepsilon \right\}.$$

Par monotonie des probabilités on obtient donc $\lim_{n\to+\infty} \mathrm{P}(\frac{1}{n}\sum_{i=1}^n Y_i < A) = 0$, ce qui, en passant au complémentaire, donne

$$\lim_{n\to+\infty} \mathrm{P}\left(\frac{1}{n} \sum_{i=1}^n Y_i \geq A \right) = 1\,.$$

(iii) Comme les variables aléatoires X_i sont positives et que $Y_i = X_i\, \mathbb{1}_{\{X_i \leq B\}}$, on a $X_i \geq Y_i$. Par conséquent, on peut majorer $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i \geq \frac{1}{n}\sum_{i=1}^n Y_i$, dont on déduit l'inclusion d'événements $\{\frac{1}{n}\sum_{i=1}^n Y_i \geq A\} \subseteq \{\overline{X}_n \geq A\}$. Par monotonie des probabilités, on déduit du point précédent que

$$\lim_{n\to\infty} \mathrm{P}(\overline{X}_n \geq A) \geq \lim_{n\to\infty} \mathrm{P}\left(\frac{1}{n} \sum_{i=1}^n Y_i \geq A \right) = 1\,,$$

donc $\lim_{n\to\infty} \mathrm{P}(\overline{X}_n \geq A) = 1$ (une probabilité ne peut pas être strictement supérieure à 1).

7.2 Le théorème central limite

Exercice 7.4. Montrer que la limite (7.33) découle de la Proposition 7.22.

Solution 7.4. Soit $Z \sim \mathrm{N}(0,1)$. En introduisant la quantité

$$\delta_n := \sup_{x\in\mathbb{R}} \big| \mathrm{P}(Z_n \leq x) - \mathrm{P}(Z \leq x) \big|\,,$$

on sait grâce à la Proposition 7.22 que l'on a $\lim_{n\to\infty} \delta_n = 0$. Par définition,

$$\big| \mathrm{P}(Z_n \leq x) - \mathrm{P}(Z \leq x) \big| \leq \delta_n\,, \qquad \forall n \in \mathbb{N}^*,\ \forall x \in \mathbb{R}\,,$$

et de même $|\mathrm{P}(Z_n < x) - \mathrm{P}(Z < x)| \leq \delta_n$, parce que l'on a $\mathrm{P}(Z_n < x) = \lim_{y\uparrow x} \mathrm{P}(Z_n \leq y)$ et $\mathrm{P}(Z \leq x) = \mathrm{P}(Z < x)$. Notons que pour tout intervalle $I =]a,b] \subseteq \mathbb{R}$ on a

$$\begin{aligned}\big|\mathrm{P}(Z_n \in I) - \mathrm{P}(Z \in I)\big| &= \big|\mathrm{P}(Z_n \le b) - \mathrm{P}(Z_n \le a) - \mathrm{P}(Z \le b) + \mathrm{P}(Z \le a)\big| \\ &\le \big|\mathrm{P}(Z_n \le b) - \mathrm{P}(Z \le b)\big| + \big|\mathrm{P}(Z_n \le a) - \mathrm{P}(Z \le a)\big| \le 2\delta_n \,.\end{aligned}$$

Avec des arguments analogues, on montre la même inégalité pour des intervalles du type $]a,b[$, $[a,b[$ ou $[a,b]$. On en déduit la limite (7.33).

Exercice 7.5. En référence au Théorème 7.23, démontrer que $\rho \ge \sigma^3$.

Solution 7.5. Rappelons que $\rho = \mathrm{E}(|X_1 - \mu|^3)$ et $\sigma^2 = \mathrm{E}(|X_1 - \mu|^2)$. Comme $\varphi(x) := |x|^{3/2}$ est une fonction convexe, en appliquant l'inégalité de Jensen à la variable aléatoire $Y := |X_1 - \mu|^2$ on obtient $\mathrm{E}(\varphi(Y)) \ge \varphi(\mathrm{E}(Y))$, c'est-à-dire $\rho \ge \sigma^3$.

Exercice 7.6. Montrer que les quantités E_n et $\hat{E}_n$, définies dans (7.38) et (7.40), sont toutes les deux plus petites que $\|F_{Z_n} - \Phi\|_\infty$, définie en (7.34).

Solution 7.6. En se rappelant que $Z_n := (S_n - \mu n)/(\sigma\sqrt{n})$, la relation (7.38) peut se réécrire sous la forme

$$\begin{aligned}E_n := &\sup_{m\in\mathbb{Z}} \left| \mathrm{P}\left(S_n \le m\right) - \Phi\left(\frac{m-\mu n}{\sigma\sqrt{n}}\right)\right| \\ = &\sup_{m\in\mathbb{Z}} \left| \mathrm{P}\left(Z_n \le \frac{m-\mu n}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{m-\mu n}{\sigma\sqrt{n}}\right)\right| \\ = &\sup_{m\in\mathbb{Z}} \left| F_{Z_n}\left(\frac{m-\mu n}{\sigma\sqrt{n}}\right) - \Phi\left(\frac{m-\mu n}{\sigma\sqrt{n}}\right)\right| \\ \le &\sup_{x\in\mathbb{R}} \left| F_{Z_n}(x) - \Phi(x)\right| =: \|F_{Z_n} - \Phi\|_\infty \,.\end{aligned}$$

Un raisonnement complètement analogue est valable pour la quantité $\hat{E}_n$ dans (7.40), en notant que

$$\mathrm{P}\left(S_n \le m\right) = \mathrm{P}\left(S_n \le m + \frac{1}{2}\right) = \mathrm{P}\left(Z_n \le \frac{m + \frac{1}{2} - \mu n}{\sigma\sqrt{n}}\right).$$

7.3 Exercices récapitulatifs

Exercice 7.7. Soit $f : [0,1] \to \mathbb{R}$ une fonction continue. Montrer que l'on a

$$\lim_{n\to\infty} \int_0^1 \cdots \int_0^1 f\Big(\frac{x_1 + \cdots + x_n}{n}\Big) \mathrm{d}x_1 \ldots \mathrm{d}x_n = f\Big(\frac{1}{2}\Big).$$

[*Sugg.* Interpréter l'intégrale comme une espérance $\mathrm{E}(f(\overline{X}_n))$ et utiliser l'Exercice 7.1.]

Solution 7.7. Rappelons la formule de transfert : si $X = (X_1, \ldots, X_n)$ est un vecteur aléatoire absolument continu, de densité $f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$, pour toute fonction $g : \mathbb{R}^n \to \mathbb{R}$ continue et bornée on a

$$\mathrm{E}(g(X)) = \int_{\mathbb{R}^n} g(x_1, \ldots, x_n)\, f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)\, \mathrm{d}x_1 \ldots \mathrm{d}x_n \,.$$

En particulier, si les $X_1, \ldots, X_n$ sont des variables aléatoires i.i.d. de loi uniforme continue sur $[0,1]$, on a

$$f_{X_1,\dots,X_n}(x_1,\dots,x_n) = f_{X_1}(x_1)\cdots f_{X_n}(x_n) = \mathbb{1}_{[0,1]}(x_1)\cdots\mathbb{1}_{[0,1]}(x_n)\,.$$

En choisissant $g(x_1,\dots,x_n) := f(\frac{x_1+\dots+x_n}{n})$, avec $f : [0,1] \to \mathbb{R}$ continue (sur un intervalle fermé, donc f est bornée), on obtient donc

$$\mathrm{E}(f(\overline{X}_n)) = \mathrm{E}(g(X)) = \int_0^1 \cdots \int_0^1 f\left(\frac{x_1+\dots+x_n}{n}\right) \mathrm{d}x_1 \dots \mathrm{d}x_n\,.$$

En appliquant l'Exercice 7.1, on obtient finalement $\lim_{n\to\infty} \mathrm{E}(f(\overline{X}_n)) = f(\mu)$ avec $\mu = \mathrm{E}(X_1) = \frac{1}{2}$, qui est la conclusion voulue.

Exercice 7.8. Soit $U \sim \mathrm{U}(0,2\pi)$. Pour $k \in \mathbb{N}^*$, on définit $X_k = \sin(kU)$ et pour $n \in \mathbb{N}^*$ on pose $S_n = \sum_{k=1}^n X_k$.

(i) Montrer que les variables aléatoires X_k sont de même loi, mais qu'elles ne sont pas indépendantes.

[*Sugg.* Montrer que les événements $\{X_1 > \frac{\sqrt{3}}{2}\}$ et $\{X_2 > \frac{\sqrt{3}}{2}\}$ sont disjoints.]

(ii) Montrer que $\mathrm{E}(X_k) = 0$, $\mathrm{Var}(X_k) = \frac{1}{2}$ et $\mathrm{Cov}(X_j, X_k) = 0$ pour $j \neq k$.

(iii) Montrer que pour tout $\varepsilon > 0$ on a

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{1}{n}|S_n| > \varepsilon\right) = 0\,.$$

Solution 7.8. (i) Montrons que tous les X_k ont la même loi que X_1. Pour cela, il suffit de montrer que pour toute fonction $\varphi : \mathbb{R} \to \mathbb{R}$ continue on a

$$\mathrm{E}(\varphi(\sin(kU))) = \mathrm{E}(\varphi(\sin(U))) \qquad \forall k \in \mathbb{N}^*\,. \tag{S7.2}$$

D'après la formule de transfert, on peut écrire

$$\mathrm{E}(\varphi(\sin(kU))) = \int_{\mathbb{R}} \varphi(\sin(ku))\, f_U(u)\,\mathrm{d}u = \frac{1}{2\pi}\int_0^{2\pi} \varphi(\sin(ku))\,\mathrm{d}u\,,$$

et avec le changement de variable $x = ku$ on obtient

$$\mathrm{E}(\varphi(\sin(kU))) = \frac{1}{2\pi}\frac{1}{k}\int_0^{2\pi k} \varphi(\sin(x))\,\mathrm{d}x = \frac{1}{2\pi}\frac{1}{k}\sum_{i=1}^k \int_{2\pi(i-1)}^{2\pi i} \varphi(\sin(x))\,\mathrm{d}x\,.$$

Vu que $\sin(x)$ est une fonction périodique, de période 2π, toutes les intégrales qui apparaissent dans la dernière somme sont égales : on obtient donc

$$\mathrm{E}(\varphi(\sin(kU))) = \frac{1}{2\pi}\int_0^{2\pi} \varphi(\sin(x))\,\mathrm{d}x = \mathrm{E}(\varphi(\sin(U)))\,,$$

qui est la relation cherchée.

Montrons que par exemple X_1 et X_2 ne sont pas indépendantes, en montrant que

$$\mathrm{P}(X_1 > t, X_2 > t) \neq \mathrm{P}(X_1 > t)\,\mathrm{P}(X_2 > t)$$

pour $t = \frac{\sqrt{3}}{2}$. En effet, dans le domaine $x \in [0,2\pi]$, on a $\sin(x) = \frac{\sqrt{3}}{2}$ pour $x = \frac{1}{3}\pi$ ou bien pour $x = \frac{2}{3}\pi$; de la même manière on a $\sin(x) > \frac{\sqrt{3}}{2}$ si et seulement si $x \in \,]\frac{\pi}{3}, \frac{2\pi}{3}[$. Par conséquent, comme $X_1 = \sin(U)$,

$$\left\{X_1 > \frac{\sqrt{3}}{2}\right\} = \left\{U \in \left]\frac{\pi}{3}, \frac{2\pi}{3}\right[\right\}.$$

De façon analogue, pour $x \in [0, 2\pi]$ on a $\sin(2x) > \frac{\sqrt{3}}{2}$ si et seulement si $2x \in \left]\frac{\pi}{3}, \frac{2\pi}{3}\right[\cup \left]2\pi + \frac{\pi}{3}, 2\pi + \frac{2\pi}{3}\right[$ ou bien $x \in \left]\frac{\pi}{6}, \frac{\pi}{3}\right[\cup \left]\frac{7\pi}{6}, \frac{4\pi}{3}\right[$. Comme $X_2 = \sin(2U)$, on obtient

$$\left\{X_2 > \frac{\sqrt{3}}{2}\right\} = \left\{U \in \left]\frac{\pi}{6}, \frac{\pi}{3}\right[\right\} \cup \left\{U \in \left]\frac{7\pi}{6}, \frac{4\pi}{3}\right[\right\}.$$

Cela montre que l'on ne peut pas avoir simultanément $X_1 > \frac{\sqrt{3}}{2}$ et $X_2 > \frac{\sqrt{3}}{2}$, c'est-à-dire $\{X_1 > \frac{\sqrt{3}}{2}\} \cap \{X_2 > \frac{\sqrt{3}}{2}\} = \emptyset$ et donc $\mathrm{P}(X_1 > \frac{\sqrt{3}}{2}, X_2 > \frac{\sqrt{3}}{2}) = 0$. D'autre part, on a $\mathrm{P}(X_1 > \frac{\sqrt{3}}{2}) = \mathrm{P}(X_2 > \frac{\sqrt{3}}{2}) = \mathrm{P}(U \in \left]\frac{\pi}{3}, \frac{2\pi}{3}\right[) = \frac{1}{6}$, donc

$$\mathrm{P}\left(X_1 > \tfrac{\sqrt{3}}{2}, X_2 > \tfrac{\sqrt{3}}{2}\right) \neq \mathrm{P}\left(X_1 > \tfrac{\sqrt{3}}{2}\right) \mathrm{P}\left(X_2 > \tfrac{\sqrt{3}}{2}\right)$$

ce qui montre que X_1 et X_2 ne sont pas indépendantes.
De façon générale, on peut montrer que pour tous $i \neq j$, X_i et X_j ne sont pas indépendantes, mais cela devient plus technique.

(ii) Comme la loi de X_k ne dépend pas de k, on a

$$\mathrm{E}(X_k) = \mathrm{E}(X_1) = \mathrm{E}(\sin(U)) = \frac{1}{2\pi} \int_0^{2\pi} \sin(x)\,\mathrm{d}x = \frac{1}{2\pi} \big[-\cos(x)\big]_0^{2\pi} = 0,$$

$$\mathrm{E}(X_k^2) = \mathrm{E}(X_1^2) = \mathrm{E}(\sin(U)^2) = \frac{1}{2\pi} \int_0^{2\pi} \sin(x)^2\,\mathrm{d}x = \frac{1}{2},$$

où pour la dernière égalité on peut remarquer que $\int_0^{2\pi} \sin(x)^2\,\mathrm{d}x = \int_0^{2\pi} \cos(x)^2\,\mathrm{d}x$, par symétrie, dont on déduit que

$$\int_0^{2\pi} \sin(x)^2\,\mathrm{d}x = \frac{1}{2} \int_0^{2\pi} (\sin(x)^2 + \cos(x)^2)\,\mathrm{d}x = \int_0^{2\pi} 1\,\mathrm{d}x = 2\pi.$$

Par conséquent, $\mathrm{Var}(X_k) = \mathrm{E}(X_k^2) - \mathrm{E}(X_k)^2 = \frac{1}{2}$. Calculons pour finir

$$\mathrm{E}(X_j X_k) = \mathrm{E}(\sin(jU)\sin(kU)) = \frac{1}{2\pi} \int_0^{2\pi} \sin(ju)\,\sin(ku)\,\mathrm{d}u.$$

Comme $\sin(x)\sin(y) = \frac{1}{2}(\cos(x-y) - \cos(x+y))$, on obtient, pour $j \neq k$

$$\begin{aligned} \mathrm{E}(X_j X_k) &= \frac{1}{\pi} \int_0^{2\pi} \big(\cos((j-k)u) - \cos((j+k)u)\big)\,\mathrm{d}u \\ &= \frac{1}{\pi} \left[\frac{\sin(j-k)u}{j-k} - \frac{\sin(j+k)u}{j+k}\right]_0^{2\pi} = 0, \end{aligned}$$

dont on déduit que $\mathrm{Cov}(X_j, X_k) = 0$ pour tous $j \neq k$.

(iii) On peut calculer $\mathrm{E}(S_n) = \sum_{i=1}^n \mathrm{E}(X_i) = 0$ et

$$\mathrm{Var}(S_n) = \sum_{i=1}^n \mathrm{Var}(X_i) + \sum_{\substack{j,k \in \{1,\dots,n\} \\ j \neq k}} \mathrm{Cov}(X_j, X_k) = \frac{n}{2}.$$

Ainsi, d'après l'inégalité de Bienaymé–Tchebychev, pour tout $\varepsilon > 0$,

$$\mathrm{P}\left(\frac{1}{n}|S_n| > \varepsilon\right) = \mathrm{P}\big(|S_n - \mathrm{E}(S_n)| > \varepsilon n\big) \le \frac{\mathrm{Var}(S_n)}{\varepsilon^2 n^2} = \frac{\frac{1}{2}n}{\varepsilon^2 n^2} = \frac{1}{2\varepsilon^2 n} \xrightarrow[n\to\infty]{} 0\,.$$

Exercice 7.9. Soit $(X_i)_{i\ge 1}$ une suite de variables aléatoires i.i.d. de loi $\mathrm{U}(0, \frac{12}{5})$ et soit

$$Y_n := \prod_{i=1}^n X_i\,, \qquad \text{pour } n \in \mathbb{N}^*\,.$$

(i) Montrer que $\mathrm{E}(\log X_i) = \log(\frac{12}{5}) - 1 \approx -0,12$. En déduire que

$$\lim_{n\to\infty} \mathrm{P}\left(\sum_{i=1}^n \log X_i > -\frac{n}{10}\right) = 0\,.$$

(ii) Montrer que $\lim_{n\to\infty} \mathrm{P}(Y_n \ge \mathrm{e}^{-n/10}) = 0$.

(iii) Montrer que pour tout $\varepsilon > 0$ on a $\lim_{n\to\infty} \mathrm{P}(Y_n > \varepsilon) = 0$. En notant m_n la médiane de Y_n, en déduire que $\lim_{n\to+\infty} m_n = 0$.

[*Sugg.* On rappelle que la médiane m_n est définie par $\mathrm{P}(Y_n \ge m_n) \ge \frac{1}{2}$ et $\mathrm{P}(Y_n \le m_n) \ge \frac{1}{2}$.]

(iv) Calculer $\mathrm{E}(Y_n)$ et montrer que $\lim_{n\to\infty} \mathrm{E}(Y_n) = +\infty$.

Solution 7.9. (i) D'après la formule de transfert,

$$\begin{aligned}\mathrm{E}(\log X_i) &= \int_{\mathbb{R}} (\log x)\, f_{X_i}(x)\,\mathrm{d}x = \frac{5}{12}\int_0^{\frac{12}{5}} \log x\,\mathrm{d}x = \frac{5}{12}\big[x(\log x - 1)\big]_0^{\frac{12}{5}} \\ &= \log\frac{12}{5} - 1 \approx -0,12\,.\end{aligned}$$

Les variables aléatoires $(\log X_i)_{i\ge 1}$ sont indépendantes et de même loi (par conservation de l'indépendance et de la loi) et dans L^1 (même dans L^2). D'après la loi des grands nombres, si on pose $\mu = \mathrm{E}(\log X_1)$, pour tout $\varepsilon > 0$ on a

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{1}{n}\sum_{i=1}^n \log X_i > \mu + \varepsilon\right) \le \lim_{n\to\infty} \mathrm{P}\left(\left|\frac{1}{n}\sum_{i=1}^n \log X_i - \mu\right| > \varepsilon\right) = 0\,.$$

Comme $\mu = \mathrm{E}(\log X_1) \approx -0,12 < -0,1 = -\frac{1}{10}$, en choisissant $\varepsilon := -\frac{1}{10} - \mu > 0$ on obtient le résultat voulu.

(ii) On peut observer que

$$Y_n = \prod_{i=1}^n X_i \ge \mathrm{e}^{-n/10} \qquad \Longleftrightarrow \qquad \log Y_n = \sum_{i=1}^n \log X_i \ge -\frac{n}{10}\,,$$

donc $\mathrm{P}(Y_n \ge \mathrm{e}^{-n/10})$ coïncide avec la probabilité estimée dans le point précédent, dont on a montré qu'elle tendait vers zéro quand $n \to \infty$.

(iii) Pour $\varepsilon > 0$, si on choisit $n \in \mathbb{N}^*$ suffisamment grand pour que $\mathrm{e}^{-n/10} > \varepsilon$, on peut écrire l'inclusion d'événements $\{Y_n > \varepsilon\} \subseteq \{Y_n \ge \mathrm{e}^{-n/10}\}$, donc par monotonie des probabilités

$$\lim_{n\to\infty} \mathrm{P}(Y_n > \varepsilon) \le \lim_{n\to\infty} \mathrm{P}(Y_n \ge \mathrm{e}^{-n/10}) = 0\,,$$

comme on l'a vu dans le point précédent.

Par définition de la limite, il existe $\bar{n} \in \mathbb{N}$ tel que pour tout $n \geq \bar{n}$ on a $\mathrm{P}(Y_n > \varepsilon) < \frac{1}{2}$, donc $m_n \leq \varepsilon$ par définition de la médiane. Comme on a clairement $m_n \geq 0$ (vu que $Y_n \geq 0$), on a montré que pour tout $\varepsilon > 0$ il existe $\bar{n}$ tel que $0 \leq m_n \leq \varepsilon$ pour tout $n \geq \bar{n}$. Cela montre que $\lim_{n\to\infty} m_n = 0$.

(iv) Les variables aléatoires X_i étant indépendantes et identiquement distribuées, on peut calculer

$$\mathrm{E}(Y_n) = \prod_{i=1}^{n} \mathrm{E}(X_i) = \mathrm{E}(X_1)^n .$$

Comme $\mathrm{E}(X_1) = \frac{5}{12}\int_0^{\frac{12}{5}} x\,\mathrm{d}x = \frac{1}{2}\frac{12}{5} = \frac{6}{5} > 1$, il s'ensuit que $\lim_{n\to\infty}\mathrm{E}(Y_n) = +\infty$.

Exercice 7.10. Soit $(X_i)_{i\geq 1}$ une suite de variables aléatoires i.i.d. admettant une espérance finie $\mu := \mathrm{E}(X_i) \in \mathbb{R}$. Pour $n \in \mathbb{N}^*$ pose $W_n = \sum_{i=1}^{2n} X_i X_{i+1}$.

(i) En général, les variables aléatoires $(Y_i := X_i X_{i+1})_{i\geq 1}$ sont-elles indépendantes ?
[*Sugg.* Considérer le cas $\mu \neq 0$, avec $\sigma^2 := \mathrm{Var}(X_i) \in \,]0,+\infty[$, et calculer $\mathrm{Cov}(Y_i, Y_{i+1})$.]

(ii) On pose $W_n^{(1)} = \sum_{j=1}^{n} X_{2j-1}X_{2j}$ et $W_n^{(2)} = \sum_{j=1}^{n} X_{2j}X_{2j+1}$. En appliquant la loi faible des grands nombres, montrer que pour tout $\varepsilon > 0$ on a

$$\lim_{n\to\infty} \mathrm{P}\left(\left|\frac{1}{n}W_n^{(1)} - \mu^2\right| > \varepsilon\right) = 0 \quad \text{e} \quad \lim_{n\to\infty} \mathrm{P}\left(\left|\frac{1}{n}W_n^{(2)} - \mu^2\right| > \varepsilon\right) = 0 .$$

(iii) En observant que $W_n = W_n^{(1)} + W_n^{(2)}$, montrer que pour tout $\varepsilon > 0$ on a

$$\lim_{n\to\infty} \mathrm{P}\left(\left|\frac{1}{2n}W_n - \mu^2\right| > \varepsilon\right) = 0 .$$

(iv) Retrouver ce résultat en utilisant l'Exercice 7.2, en supposant que les X_i admettent un moment d'ordre deux fini.

Solution 7.10. (i) Intuitivement, deux variables d'indices consécutifs $Y_i = X_i X_{i+1}$ et $Y_{i+1} = X_{i+1}X_{i+2}$ sont fonctions de la même variable aléatoire X_{i+1} et donc, *en générale*, elles ne sont pas indépendantes. En effet, si les variables aléatoires $(Y_i)_{i\geq 1}$ étaient indépendantes, elles devraient être décorrélées et en particulier on devrait avoir $\mathrm{Cov}(Y_1, Y_2) = 0$. On peut calculer $\mathrm{E}(Y_i) = \mathrm{E}(X_i)\,\mathrm{E}(X_{i+1}) = \mu^2$. Si on pose $\sigma^2 := \mathrm{Var}(X_i)$, on a de plus

$$\mathrm{E}(Y_1Y_2) = \mathrm{E}(X_1X_2^2X_3) = \mathrm{E}(X_1)\,\mathrm{E}(X_2^2)\,\mathrm{E}(X_3) = \mu^2(\mu^2 + \sigma^2) .$$

Comme $\mathrm{E}(X_2^2) = \mathrm{E}(X_2)^2 + \mathrm{Var}(X_2) = \mu^2 + \sigma^2$, on a

$$\mathrm{Cov}(Y_1, Y_2) = \mathrm{E}(Y_1Y_2) - \mathrm{E}(Y_1)\,\mathrm{E}(Y_2) = \mu^2\sigma^2 .$$

Cela montre que, dans le cas où $\mu \neq 0$ et $\sigma^2 \neq 0$, les variables aléatoires Y_i ne sont pas décorrélées et à plus forte raison ne sont pas indépendantes.

Avec une analyse plus poussée on peut enlever la condition $\mu \neq 0$. En revanche, la condition $\sigma^2 \neq 0$ est essentielle : en effet si $\sigma^2 = \mathrm{Var}(X_i) = 0$, les variables aléatoires X_i sont presque sûrement constantes, donc les Y_i sont aussi presque sûrement constantes et sont par conséquent indépendantes (voir l'Exercice 3.8).

(ii) Définissons les variables aléatoires $Z_j := X_{2j-1}X_{2j}$ pour $j \in \mathbb{N}^*$ et observons qu'elles sont i.i.d., grâce aux propriétés d'indépendance par paquets et de conservation de l'indépendance (en effet Z_1 est fonction de (X_1, X_2), alors que Z_2 est fonction de (X_3, X_4), etc.). De plus, Z_j est d'espérance finie : $\mathrm{E}(Z_j) = \mathrm{E}(X_{2j-1})\mathrm{E}(X_{2j}) = \mu^2$. D'après la loi des grands nombres, on a alors, pour tout $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathrm{P}\left(\left|\frac{Z_1 + \cdots + Z_n}{n} - \mathrm{E}[Z_1]\right| > \varepsilon\right) = 0\,,$$

qui est la première relation à démontrer. La deuxième relation est complètement analogue.

(iii) Pour $x,y \in \mathbb{R}$, montrons que si $|\frac{x+y}{2} - \mu^2| > \varepsilon$, alors on doit avoir $|x-\mu^2| > \varepsilon$ ou bien $|y-\mu^2| > \varepsilon$. En effet, si on avait à la fois $|x-\mu^2| \le \varepsilon$ et $|y-\mu^2| \le \varepsilon$, c'est-à-dire $\mu^2 - \varepsilon \le x \le \mu^2 + \varepsilon$ et $\mu^2 - \varepsilon \le y \le \mu^2 + \varepsilon$, alors en sommant ces inégalités on obtiendrait $2(\mu^2 - \varepsilon) \le x + y \le 2(\mu^2 + \varepsilon)$, c'est-à-dire $|\frac{x+y}{2} - \mu^2| \le \varepsilon$.

En choisissant $x = \frac{1}{n}W_n^{(1)}$ et $y = \frac{1}{n}W_n^{(2)}$, On obtient l'inclusion d'événements

$$\left\{\left|\frac{1}{2n}W_n - \mu^2\right| > \varepsilon\right\} \subseteq \left\{\left|\frac{1}{n}W_n^{(1)} - \mu^2\right| > \varepsilon\right\} \cup \left\{\left|\frac{1}{n}W_n^{(2)} - \mu^2\right| > \varepsilon\right\},$$

dont découle le fait que

$$\mathrm{P}\left(\left|\frac{1}{2n}W_n - \mu^2\right| > \varepsilon\right) \le \mathrm{P}\left(\left|\frac{1}{n}W_n^{(1)} - \mu^2\right| > \varepsilon\right) + \mathrm{P}\left(\left|\frac{1}{n}W_n^{(2)} - \mu^2\right| > \varepsilon\right).$$

En prenant la limite $n \to \infty$ on obtient donc le résultat.

(iv) On a déjà observé que $\mathrm{Cov}(Y_i, Y_{i+1}) \neq 0$. Observons maintenant que $\mathrm{Cov}(Y_i, Y_j) = 0$ si $|j-i| \ge 2$, parce que dans ce cas Y_i et Y_j sont des variables aléatoires indépendantes grâce aux propriétés d'indépendance par paquets et de conservation de l'indépendance. En effet, Y_i et Y_j sont fonctions respectivement de (X_i, X_{i+1}) et (X_j, X_{j+1}) et la condition $|j-i| \ge 2$ garantit que les variables aléatoires X_n pour $n \in \{i, i+1, j, j+1\}$ sont distinctes (c'est-à-dire ont des indices distincts).

On peut donc majorer $|\mathrm{Cov}(Y_i, Y_j)| \le \rho(|j-i|)$, où l'on a défini $\rho(k) := 0$ pour $k \ge 2$ et $\rho(0) := \mathrm{Var}(Y_1)$ et $\rho(1) := |\mathrm{Cov}(Y_1, Y_2)|$. Clairement $\lim_{k\to\infty} \rho(k) = 0$, donc d'après l'Exercice 7.2 la loi des grands nombres est vérifiée pour les variables aléatoires $(Y_i)_{i\ge 1}$, c'est-à-dire

$$\forall \varepsilon > 0 : \qquad \lim_{N\to\infty} \mathrm{P}\left(\left|\frac{Y_1 + \cdots + Y_N}{N} - \mathrm{E}(Y_1)\right| > \varepsilon\right) = 0\,.$$

Comme $\mathrm{E}(Y_1) = \mu^2$, on a montré la relation voulue (avec $N = 2n$).

Exercice 7.11. En France, il y a eu 736 000 naissances en 2020, dont 377 000 garçons et 359 000 filles. Vous semble-t-il raisonnable de soutenir l'idée qu'un nouveau-né ait autant de chance d'être un garçon qu'une fille ?

[*Sugg.* Calculer, dans l'hypothèse où les sexes sont équiprobables, la probabilité d'avoir un écart entre le nombre de garçons et le nombre de filles *égal ou supérieur* à celui observé.]

Solution 7.11. Pour $i = 1, 2, \ldots, N = 736\,000$ on désigne par X_i le sexe du i-ème enfant né, disons $X_i = 1$ si c'est un garçon et $X_i = 0$ si c'est une fille. Si on fait l'hypothèse que les sexes sont équiprobables, on peut supposer que les $(X_i)_{1\le i\le N}$ sont des variables aléatoires i.i.d. de loi $\mathrm{Bern}(p)$ avec $p = \frac{1}{2}$, c'est-à-dire $\mathrm{P}(X_i = 1) = \mathrm{P}(X_i = 0) = \frac{1}{2}$; en particulier $\mu := \mathrm{E}[X_i] = p = \frac{1}{2}$ et $\sigma^2 := \mathrm{Var}[X_i] = p(1-p) = \frac{1}{4}$. Définissons $S_N := X_1 + X_2 + \cdots + X_N$, qui représente le nombre de garçons, et calculons la probabilité que l'écart entre le nombre de garçons et le nombre de filles soit égal ou supérieur à celui observé, c'est-à-dire $\mathrm{P}(S_N \ge 377\,000)$.

D'après le théorème central limite, comme $N = 736\,000$ est « grand », la variable aléatoire $Z_N := (S_N - \mu N)/(\sigma\sqrt{N})$ est de loi approximativement normale standard $Z \sim \mathrm{N}(0,1)$. On peut donc calculer approximativement

$$\begin{aligned}\mathrm{P}(S_N \geq 213\,000) &= \mathrm{P}\left(\frac{S_N - \mu N}{\sigma\sqrt{N}} \geq \frac{377\,000 - \frac{1}{2}736\,000}{\frac{1}{2}\sqrt{736\,000}}\right) = \mathrm{P}\left(Z_N \geq \frac{9000}{40\sqrt{115}}\right) \\ &\approx \mathrm{P}\big(Z \geq 20{,}98\big) = 1 - \mathrm{P}(Z < 20{,}98) = 1 - \Phi(20{,}98)\,,\end{aligned}$$

où on a utilisé le fait que $\Phi(x) = \mathrm{P}(Z \leq x) = \mathrm{P}(Z < x)$, parce que $\mathrm{P}(Z = x) = 0$ vu que $Z \sim \mathrm{N}(0,1)$ est absolument continue. La table page S-225 reporte les valeurs de $\Phi(x)$ pour $x \leq 3{,}99$, parce que pour $x > 4$ on a $\Phi(x) \approx 1$. En particulier $\Phi(20{,}98) \approx 1$ et donc $\mathrm{P}(S_N \geq 377\,000) \approx 1 - \Phi(20{,}98) \approx 0$ (en utilisant la fonction erf, un ordinateur donne la valeur $1 - \Phi(20{,}98) \approx 5 \cdot 10^{-98}$!). Donc l'hypothèse que les naissances ont autant de chance d'être des filles que des garçons n'est pas en accord avec les données observées.

Exercice 7.12. Un groupe promouvant un référendum estime que 60% de la population est disposée à signer une pétition pour appuyer la proposition de référendum. On suppose que les personnes à qui l'on demande de signer sont choisies au hasard. Si l'on doit recueillir 30 000 signatures, combien de personnes est-il nécessaire de contacter pour que le seuil des 30 000 signatures soit atteint avec une probabilité d'au moins 0,95 ?

Solution 7.12. On note X_i la variable aléatoire qui vaut 1 si la i-ème personne interpellée est disposée à signer la pétition et 0 sinon. Supposons que les $X_1, X_2, \ldots$ sont des variables aléatoires i.i.d. avec $X_i \sim \mathrm{Bern}(p)$ où $p = 0{,}6$. Elles ont donc une espérance μ et un écart-type σ donnés par

$$\mu = p = 0{,}6\,, \qquad \sigma = \sqrt{p(1-p)} \approx 0{,}49\,. \tag{S7.3}$$

Si on interpelle n personnes, le nombre de personnes disposées à signer la pétition est donné par la variable aléatoire $S_n := X_1 + \cdots + X_n$. La condition à satisfaire est donc

$$\mathrm{P}(S_n \geq N_0) \geq 0{,}95\,, \qquad \text{où} \quad N_0 := 30\,000\,. \tag{S7.4}$$

Pour appliquer la méthode de l'approximation normale, on se ramène à la variable aléatoire $Z_n := (S_n - \mu n)/(\sigma\sqrt{n}) \approx Z$, où Z désigne comme d'habitude une variable aléatoire $\mathrm{N}(0,1)$. En notant Φ le fonction de répartition de Z, on a

$$\begin{aligned}\mathrm{P}(S_n \geq N_0) &= \mathrm{P}\left(\frac{S_n - \mu n}{\sigma\sqrt{n}} \geq \frac{N_0 - \mu n}{\sigma\sqrt{n}}\right) \approx \mathrm{P}\left(Z \geq \frac{N_0 - \mu n}{\sigma\sqrt{n}}\right) \\ &= 1 - \Phi(z)\,, \qquad \text{où on a posé} \quad z := \frac{N_0 - \mu n}{\sigma\sqrt{n}}\,,\end{aligned} \tag{S7.5}$$

et la relation (S7.4) devient (approximativement)

$$1 - \Phi(z) \geq 0{,}95 \quad \Longleftrightarrow \quad z \leq \Phi^{-1}(0{,}05)\,. \tag{S7.6}$$

La valeur de $\Phi^{-1}(0{,}05)$ ne peut pas être directement trouvée dans la table page S-225, où n'apparaissent que les nombres supérieurs à $0{,}5$. Cependant, en se rappelant que $1 - \Phi(z) = \Phi(-z)$ pour tout $z \in \mathbb{R}$ (*cf.* la relation (7.26)), on peut écrire

$$1 - \Phi(z) \geq 0{,}95 \quad \Longleftrightarrow \quad \Phi(-z) \geq 0{,}95 \quad \Longleftrightarrow \quad (-z) \geq \Phi^{-1}(0{,}95)\,,$$

et on peut facilement trouver la valeur de $\Phi^{-1}(0,95)$. En effet, dans la table page S-225 on lit $\Phi(1,64) = 0,9495$ et $\Phi(1,65) = 0.9505$, dont on déduit que

$$\Phi^{-1}(0,95) \approx 1,645\,. \tag{S7.7}$$

En revenant à (S7.5), la condition $(-z) \geq \Phi^{-1}(0,95)$ se traduit en

$$\frac{N_0 - \mu n}{\sigma\sqrt{n}} \leq -\Phi^{-1}(0,95) \quad \Longleftrightarrow \quad \mu(\sqrt{n})^2 - \sigma\Phi^{-1}(0,95)\sqrt{n} - N_0 \geq 0\,.$$

Il s'agit d'une inégalité de degré deux en $\sqrt{n}$, dont les solutions positives sont données par la formule suivante :

$$\sqrt{n} \geq \frac{1}{2\mu}\Big(\sigma\Phi^{-1}(0,95) + \sqrt{\sigma^2\Phi^{-1}(0,95)^2 - 4\mu \cdot N_0}\Big) \approx 224,28\,, \tag{S7.8}$$

où on a utilisé les valeurs numériques pour μ, σ, N_0 et $\Phi^{-1}(0.95)$ données dans (S7.3), (S7.4) et (S7.7). Au final, le nombre de personnes nécessaires est donné par

$$n \geq (224,28)^2 \approx 50\,302\,.$$

(Si on voulait appliquer la correction de continuité, il suffirait de remplacer N_0 par $29\,999,5$, mais le résultat changerait très peu.)

Exercice 7.13. Un grand studio photo reçoit une commande qui nécessite l'utilisation de lampes à très haute luminosité. La durée de vie de ces lampes suit la loi exponentielle d'espérance égale à 100 heures, et coûtent 100 euros l'une. Les durées de vie de lampes distinctes peuvent être considérées comme indépendantes. Pour la commande, on prévoit que l'on aura besoin de 10 000 heures de lumière produite par ces lampes. De plus, à cause des coûts de transport élevés, il vaut mieux acheter les lampes nécessaires en une seule fois.

(i) En utilisant l'approximation normale, déterminer le nombre minimal de lampes qu'il faut acheter afin que l'on puisse avoir 10 000 heures de lumière, garanties avec une probabilité d'au moins $0,95$.

(ii) Une autre compagnie propose des lampes dont la durée de vie est exponentielle d'espérance 200 heures, au coût de 190 euros par lampe. Vaut-il mieux acheter les lampes à cette compagnie plutôt qu'à la précédente ? On achète aussi dans ce cas le nombre minimal de lampes nécessaires pour garantir 10 000 heures de lumière avec une probabilité d'au moins $0,95$.

Solution 7.13. (i) On note X_i la durée de vie de la i-ème ampoule et on suppose que $X_1, X_2, \dots$ sont des variables aléatoires i.i.d. de loi $\mathrm{Exp}(\lambda)$, où $\lambda := 1/\mu = 1/100$ (on rappelle que $\mathrm{E}(\mathrm{Exp}(\lambda)) = 1/\lambda$). En se rappelant de la relation (6.45), l'écart-type des X_i vaut

$$\sigma = \frac{1}{\lambda} = \mu = 100\,.$$

En posant $S_n := X_1 + \cdots + X_n$, la condition à vérifier est donc

$$\mathrm{P}(S_n \geq N_0) \geq 0,95\,, \qquad \text{où} \quad N_0 := 10\,000\,.$$

On remarque que le problème posé est formellement identique à celui de la Solution 7.12 : seules les valeurs de μ, σ, N_0 changent (et la loi de X_i, les détails de la loi autres que μ et σ ne jouent aucun rôle dans l'approximation normale). Appliquons directement la formule finale (S7.8), sans répéter les calculs : le nombre minimum d'ampoules à acheter est donné par

$$\sqrt{n} \geq \frac{1}{2\mu}\left(\sigma\Phi^{-1}(0,95) + \sqrt{\sigma^2\Phi^{-1}(0,95)^2 - 4\mu \cdot N_0}\right) \approx 10,86\,, \tag{S7.9}$$

c'est-à-dire

$$n \geq (10.86)^2 \approx 118\,.$$

(ii) Comme chaque ampoule considérée dans le premier point coûte 100 euros, le prix total pour acheter 118 lampes est égal à 11 800 euros. Déterminons maintenant le nombre d'ampoules de la deuxième compagnie qu'il serait nécessaire d'acheter : en appliquant la formule (S7.9) avec les nouvelles données $\mu = 200$, $\sigma = 200$ (et toujours $N_0 = 10\,000$) on obtient

$$\sqrt{n} \geq 7,94\,, \qquad \text{c'est-à-dire} \qquad n \geq (7,94)^2 \approx 64\,,$$

pour un coût total de $64 \cdot 190 = 12\,160$ euros. Les ampoules de la deuxième compagnie sont donc moins avantageuses, même si elle durent en moyenne deux fois plus longtemps tout en coûtant (strictement) moins du double !

Exercice 7.14. Une compagnie aérienne offre à ses passager le choix entre deux types d'en-cas : salé ou sucré. De par son expérience, la compagnie suppose que les deux choix sont équiprobables. Supposons que lors d'un vol donnée il y ait 150 passagers, qui choisissent leur en-cas indépendamment les uns des autres. Calculer, en utilisant l'approximation normale avec correction de continuité :

(i) la probabilité qu'au moins 60 passagers choisissent l'en-cas salé ;

(ii) s'il y a à bord 90 unités de chaque type d'en-cas, la probabilité que l'un des passagers ne puisse pas avoir l'en-cas qu'il souhaite.

Solution 7.14. On note la préférence de chacun des passagers $i = 1,\ldots,N := 150$ par une variable aléatoire X_i qui vaut 1 si le passager choisit l'en-cas salé et qui vaut 0 si le passager choisit l'en-cas sucré. On suppose que $X_1, X_2, \ldots, X_N$ sont des variables aléatoires i.i.d. $\mathrm{Bern}(p)$ avec $p = \frac{1}{2}$, de sorte que $\mu := \mathrm{E}(X_i) = p = \frac{1}{2}$ et $\sigma^2 := \mathrm{Var}(X_i) = p(1-p) = \frac{1}{4}$. On définit pour finir la variable aléatoire $S_N := X_1 + X_2 + \cdots + X_N$ qui compte le nombre de passagers qui choisissent les en-cas salés. D'après le théorème central limite, la variable aléatoire standardisée $Z_N := (S_N - N\mu)/(\sigma\sqrt{N})$ est de loi approximativement égale à celle d'une variable aléatoire $Z \sim \mathrm{N}(0,1)$.

(i) On doit calculer $\mathrm{P}(S_N \geq 60) = \mathrm{P}(S_N \geq 59,5)$, où l'égalité correspond à l'astuce de la correction de continuité. D'après le théorème central limite, on a

$$\begin{aligned}\mathrm{P}(S_N \geq 59,5) = \mathrm{P}\left(\frac{S_N - \mu N}{\sigma\sqrt{N}} \geq \frac{59,5 - \frac{1}{2}150}{\frac{1}{2}\sqrt{150}}\right) &\approx \mathrm{P}\left(Z_N \geq \frac{-15,5}{6,12}\right)\\ &\approx \mathrm{P}\big(Z \geq -2,53\big) = 1 - \mathrm{P}\big(Z < -2,53\big) = 1 - \Phi(-2,53)\,.\end{aligned}$$

En se rappelant que $\Phi(x) = 1 - \Phi(-x)$ et en utilisant la table page S-225, on obtient finalement $\mathrm{P}(S_N \geq 60) \approx \Phi(2,53) \approx 99,43\%$. (Pour comparer, la probabilité « exacte » vaut $\approx 99,44\%$.)

(ii) Il y aura un passager qui n'aura pas l'en-cas qu'il souhaite s'il arrive que $S_N > 90$. On doit donc calculer

$$\mathrm{P}(\{S_N > 90\} \cup \{S_N < 60\}) = \mathrm{P}(S_N > 90) + \mathrm{P}(S_N < 60)\,.$$

D'après le point précédent

$$\mathrm{P}(S_N < 60) = 1 - \mathrm{P}(S_N \geq 60) \approx 1 - 99{,}43\% = 0{,}57\%\,.$$

Observons que, par symétrie

$$\mathrm{P}(S_N > 90) = \mathrm{P}(S_N < 60) \approx 0{,}57\%\,,$$

parce que la variable aléatoire $S_N = X_1 + \cdots + X_N$ possède la même loi que $N - S_N = (1 - X_1) + \cdots + (1 - X_N)$, vu que les $(1 - X_i)_{1 \leq i \leq N}$ sont des variables aléatoires i.i.d. $\mathrm{Bern}(\frac{1}{2})$. Au final, la probabilité demandée vaut $\approx 2 \cdot 0{,}57\% = 1{,}14\%$.

Exercice 7.15. Lors d'un jeu de hasard, on gagne 2 euros en cas de succès, et on perd 1 euro en cas d'échec. La probabilité de succès est $p = 0{,}4$.

(i) Soit X le nombre de succès lors de n répétitions du jeu. Quelle est la densité discrète de X ?

(ii) Soit Y le nombre (relatif) d'euros gagnés lors de n répétitions du jeu. Après avoir déterminé une relation qui relie X et Y, déterminer l'espérance et la variance de Y.

(iii) Calculer une valeur approchée de la probabilité de gagner au moins 50 euros lors de 220 répétitions du jeu.

(iv) Combien de fois, au minimum, est-il nécessaire de répéter le jeu pour que la probabilité de gagner au moins 50 euros soit supérieure à $0{,}99$?

[*Sugg.* Pour les points (iii) et (iv) utiliser la correction de continuité.]

Solution 7.15. (i) Il s'agit d'un schéma de répétition d'épreuves indépendantes, donc le nombre X de succès en n épreuves est de loi $\mathrm{Bin}(n, p)$ avec $p = 0{,}4$. Sa densité discrète est

$$\mathrm{p}_X(k) = \mathrm{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad \text{pour } k \in X(\Omega) = \{0, 1, \ldots, n\}\,.$$

(ii) Le nombre de succès vaut X, donc le nombre d'échecs vaut $n - X$. Comme tout succès apporte un gain de 2 euros et tout échec une perte de 1 euro, le nombre (relatif) d'euros gagnés Y est donné par

$$Y = 2 \cdot X - 1 \cdot (n - X) = 3X - n\,.$$

Rappelons que $\mathrm{E}(X) = np$ et $\mathrm{Var}(X) = np(1-p)$. D'après les propriétés de l'espérance et de la variance, on obtient donc

$$\mathrm{E}(Y) = 3\,\mathrm{E}(X) - n = (3p - 1)n\,, \qquad \mathrm{Var}(Y) = 9\,\mathrm{Var}(X) = 9p(1-p)n\,.$$

(iii) On doit calculer $\mathrm{P}(Y \geq 50)$ dans le cas où $n = 220$ et $p = 0{,}4$. Ramenons-nous à la variable aléatoire X, en notant que

$$\mathrm{P}(Y \geq 50) = \mathrm{P}(3X - 220 \geq 50) = \mathrm{P}(X \geq 90) = \mathrm{P}(X \geq 89{,}5)\,,$$

où la dernière égalité correspond à la correction de continuité. Comme X est une somme de n variables aléatoires i.i.d. $\mathrm{Bern}(p)$, d'espérance $\mu = p$ et de variance $\sigma^2 = p(1-p)$, on peut appliquer le théorème central limite. En introduisant la variable aléatoire $Z_n := (X-\mu n)/(\sigma\sqrt{n})$, qui est de loi approximativement égale à celle d'une variable aléatoire $Z \sim \mathrm{N}(0,1)$, on obtient

$$\begin{aligned}\mathrm{P}(X \geq 89{,}5) &= \mathrm{P}\left(\frac{X-\mu n}{\sigma\sqrt{n}} \geq \frac{89{,}5-0{,}4\cdot 220}{\sqrt{0{,}4\cdot 0{,}6\cdot 220}}\right) \approx \mathrm{P}\left(Z_n \geq \frac{1{,}5}{7{,}27}\right)\\ &\approx \mathrm{P}(Z \geq 0{,}21) = 1-\mathrm{P}(Z<0{,}21) = 1-\Phi(0{,}21)\,.\end{aligned}$$

En consultant la table page S-225, on obtient $\mathrm{P}(Y \geq 50) \approx 1-0{,}5832 \approx 41{,}68\%$. (Pour comparer, la probabilité « exacte » vaut $\approx 41{,}65\%$.)

Exercice 7.16. Un jeu consiste à tirer au hasard deux cartes d'un paquet (52 cartes, 4 motifs); on gagne si aucune des cartes n'est un carreau $\diamondsuit$.

(i) Déterminer la probabilité de succès dans ce jeu.

(ii) Pour $n \geq 1$, soit p_n la probabilité que lors de $2n$ répétitions du jeu le nombre de succès soit au moins n. Déterminer, approximativement, la valeur de p_{50}.

(iii) Déterminer la plus petite valeur de n pour laquelle $p_n \geq 0{,}99$.

Solution 7.16. (i) Comme il y a 39 cartes qui ne sont pas des carreaux, la probabilité de succès vaut

$$p = \frac{\binom{39}{2}}{\binom{52}{2}} = \frac{19}{34} \approx 0{,}559\,.$$

(ii) On note X_i la variable aléatoire qui vaut 1 si la i-ème répétition du jeu est un succès et qui vaut 0 sinon. Les variables aléatoires $X_1, X_2, \ldots, X_n$ sont i.i.d. de lois marginales $\mathrm{Bern}(p)$, donc l'espérance et l'écart-type valent

$$\mu = p \approx 0{,}559\,, \qquad \sigma = \sqrt{p(1-p)} \approx 0{,}497\,.$$

En posant $S_n := X_1 + \cdots + X_n \sim \mathrm{Bin}(n,p)$, en utilisant l'approximation normale on obtient

$$\begin{aligned}p_n = \mathrm{P}(S_{2n} \geq n) &= \mathrm{P}\left(\frac{S_{2n}-2n\mu}{\sigma\sqrt{2n}} \geq \frac{n-2n\mu}{\sigma\sqrt{2n}}\right) \approx \mathrm{P}\left(Z \geq \frac{n-2n\mu}{\sigma\sqrt{2n}}\right)\\ &= 1-\Phi\left(\frac{n-2n\mu}{\sigma\sqrt{2n}}\right) = \Phi\left(\frac{2n\mu-n}{\sigma\sqrt{2n}}\right),\end{aligned} \tag{S7.10}$$

où on a utilisé le fait que $1-\Phi(z) = \Phi(-z)$, voir la relation (7.26). Pour $n = 50$, en substituant les valeurs de μ, σ, n et en consultant la table page S-225, on obtient

$$p_n \approx \Phi\left(\frac{100\cdot 0{,}559-50}{0{,}497\sqrt{100}}\right) \approx \Phi(1{,}19) \approx 0{,}88\,.$$

(iii) En utilisant la formule (S7.10) on obtient

$$p_n \geq 1-10^{-3} \iff \frac{2n\mu-n}{\sigma\sqrt{2n}} \geq \Phi^{-1}(0{,}999) \iff$$

$$\frac{2\mu-1}{\sigma\sqrt{2}}\sqrt{n} \geq \Phi^{-1}(0{,}999) \iff n \geq \left(\frac{\sqrt{2}\sigma}{2\mu-1}\Phi^{-1}(0{,}999)\right)^2.$$

De la table page S-225 on peut lire les valeurs $\Phi(3{,}07) \approx 0{,}9989$ et $\Phi(3{,}08) \approx 0{,}9990$, donc $\Phi^{-1}(0{,}999) \approx 3{,}075$. En substituant les valeurs numériques de toutes les constantes, on obtient

$$n \geq n_0 := \left(\frac{\sqrt{2} \cdot 0{,}497}{2 \cdot 0{,}559 - 1} 3{,}075\right)^2 \approx 336\,.$$

Exercice 7.17. Un dispositif est constitué d'un composant électrique qui est remplacé au moment où il s'arrête de fonctionner. Ainsi, si $T_1, T_2, \dots, T_n$ sont les temps de vie des n composants que l'on a à notre disposition, le temps de vie total du dispositif est $T = T_1 + T_2 + \dots + T_n$. On suppose que $T_i \sim \mathrm{Exp}(1)$ pour tout $i \geq 1$, et que les T_i sont indépendantes. En utilisant l'approximation normale, calculer :

(i) si $n = 100$ la probabilité $\mathrm{P}(T < 90)$;

(ii) la valeur minimale de n pour avoir $\mathrm{P}(T < 90) \leq 0{,}05$.

Solution 7.17. (i) En se rappelant que l'espérance et l'écart-type d'une variable aléatoire $\mathrm{Exp}(\lambda)$ valent tous les deux $1/\lambda$, les variables aléatoires T_i ont pour espérance et variance

$$\mu = 1\,, \qquad \sigma = 1\,.$$

En utilisant l'approximation normale, on obtient

$$\begin{aligned}\mathrm{P}(T < 90) &= \mathrm{P}\left(\frac{T - \mu n}{\sigma\sqrt{n}} < \frac{90 - \mu n}{\sigma\sqrt{n}}\right) \approx \mathrm{P}\left(Z < \frac{90 - \mu n}{\sigma\sqrt{n}}\right) = \Phi\left(\frac{90 - \mu n}{\sigma\sqrt{n}}\right)\\ &= \Phi\left(\frac{90 - 100}{\sqrt{100}}\right) = \Phi(-1) = 1 - \Phi(1) \approx 1 - 0{,}8413 \approx 0{,}1587\,.\end{aligned}$$

(ii) En utilisant les calculs du point précédent, on obtient

$$\begin{aligned}&\mathrm{P}(T < 90) \approx \Phi\left(\frac{90 - \mu n}{\sigma\sqrt{n}}\right) \leq 0{,}05 \iff 1 - \Phi\left(\frac{90 - n}{\sqrt{n}}\right) \geq 0{,}95\\ \iff\ &\Phi\left(\frac{n - 90}{\sqrt{n}}\right) \geq 0{,}95 \iff \frac{n - 90}{\sqrt{n}} \geq \Phi^{-1}(0{,}95) \approx 1{,}645\,,\end{aligned}$$

où on observe que la valeur $\Phi^{-1}(0{,}95) \approx 1{,}645$ a déjà été trouvée dans (S7.7), en utilisant la table page S-225. On obtient donc une inégalité de degré deux en la variable $\sqrt{n}$:

$$(\sqrt{n})^2 - 1{,}645\sqrt{n} - 90 \geq 0\,,$$

dont les solutions positives sont données par

$$\sqrt{n} \geq \frac{1}{2}\left(1{,}645 + \sqrt{1{,}645^2 + 4 \cdot 90}\right) \approx 10{,}345\,,$$

c'est-à-dire $n \geq (10{,}345)^2 \approx 108$.

Exercice 7.18. Un million de personnes votent lors d'une élection, et doivent choisir entre deux candidats A et B. Le vote d'un nombre n d'électeurs est sous le contrôle d'une organisation mafieuse, qui garantie que ceux-ci voteront pour A. Tous les autres électeurs votent *au hasard*, en choisissant avec une probabilité égale l'un des deux candidats, indépendamment les uns des autres.

(i) Supposons que l'organisation contrôle $n = 2\,000$ votes. Quelle est la probabilité (approchée) que le candidat A gagne l'élection ?

(ii) Quel est le nombre minimum n d'individus que l'organisation doit contrôler pour garantir que la probabilité de victoire de A soit au moins de 99% ?

Solution 7.18. (i) Soit X le nombre de votes reçus par le candidat A parmi les $998\,000$ électeurs non contrôlés. Par hypothèse $X \sim \mathrm{Bin}(998\,000, \frac{1}{2})$. Notons que le candidat A gagne si et seulement si $X > 498\,000$. En utilisant l'approximation normale (sans correction de continuité car les nombre sont très grands et cela n'a pas beaucoup d'influence) on obtient

$$\mathrm{P}(X > 498\,000) = \mathrm{P}\left(\frac{X - 998\,000 \cdot \frac{1}{2}}{\sqrt{998\,000 \cdot \frac{1}{2} \cdot \frac{1}{2}}} > \frac{498\,000 - 998\,000 \cdot \frac{1}{2}}{\sqrt{998\,000 \cdot \frac{1}{2} \cdot \frac{1}{2}}}\right)$$
$$\approx \mathrm{P}\left(Z > \frac{-1\,000}{499,5}\right) \approx 1 - \Phi(-2) = \Phi(2) \approx 0,9772\,.$$

(ii) Soit maintenant X le nombre de votes reçus par A parmi les $1\,000\,000 - n$ électeurs non contrôlés. Par hypothèse $X \sim \mathrm{Bin}(1\,000\,000 - n, \frac{1}{2})$ et le candidat A gagne si et seulement si $X > 500\,000 - n$, donc avec probabilité

$$\mathrm{P}(X > 500\,000 - n)$$
$$= \mathrm{P}\left(\frac{X - (1\,000\,000 - n) \cdot \frac{1}{2}}{\sqrt{(1\,000\,000 - n) \cdot \frac{1}{2} \cdot \frac{1}{2}}} > \frac{500\,000 - n - (1\,000\,000 - n) \cdot \frac{1}{2}}{\sqrt{(1\,000\,000 - n) \cdot \frac{1}{2} \cdot \frac{1}{2}}}\right)$$
$$\approx \mathrm{P}\left(Z > \frac{-n}{\sqrt{1\,000\,000 - n}}\right) \approx \Phi\left(\frac{n}{\sqrt{1\,000\,000 - n}}\right).$$

En imposant que $\mathrm{P}(X > 500\,000 - n) \geq 0,99$ on obtient donc

$$\frac{n}{\sqrt{1\,000\,000 - n}} \geq \Phi^{-1}(0,99) \approx 2,325\,,$$

où on a utilisé la table page S-225 pour trouver la valeur de $\Phi^{-1}(0,99)$ (on note en effet que $\Phi(2,32) = 0,9898$ et $\Phi(2,33) = 0,9901$). En élevant au carré la relation précédente, on obtient

$$\frac{n^2}{1\,000\,000 - n} \geq (2,325)^2 \approx 5,406\,,$$

c'est-à-dire $n^2 + 5,406n - 5\,406\,000 \geq 0$, dont les solutions positives sont

$$n \geq \frac{1}{2}\left(-5,406 + \sqrt{(5,406)^2 + 4 \cdot 5\,406\,000}\right) \approx 2\,323\,.$$

Exercice 7.19. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires i.i.d. de lois marginales $X_n \sim \mathrm{U}(-1,1)$. Déterminer, pour tout $t \in \mathbb{R}$ la limite

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{(X_1 + \cdots + X_n)^2}{n} > t\right).$$

Solution 7.19. Notons que $\mu := \mathrm{E}(X_n) = 0$ et $\sigma^2 := \mathrm{Var}(X_n) = \frac{2^2}{12} = \frac{1}{3}$, donc $\sigma = \sqrt{1/3} \approx 0,577$ (on a appliqué les formules pour l'espérance et la variance d'une variable aléatoire uniforme continue, voir la Section 6.3.1). On peut alors réécrire la probabilité d'intérêt en faisant apparaître la moyenne empirique standardisée

$$Z_n := \frac{(X_1 + \cdots + X_n) - \mu n}{\sigma\sqrt{n}} = \frac{X_1 + \cdots + X_n}{\sigma\sqrt{n}},$$

de la façon suivante :

$$\mathrm{P}\left(\frac{(X_1 + \cdots + X_n)^2}{n} > t\right) = \mathrm{P}\left(\frac{X_1 + \cdots + X_n}{\sqrt{n}} > \sqrt{t}\right) = \mathrm{P}\big(Z_n > \sigma\sqrt{t}\big).$$

D'après le théorème central limite, en notant Z une variable aléatoire normale standard, on obtient

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{(X_1 + \cdots + X_n)^2}{n} > t\right) = \mathrm{P}\big(Z > \sigma\sqrt{t}\big) = 1 - \Phi(\sigma\sqrt{t}) \approx 1 - \Phi(0,577\sqrt{t}).$$

Notons que la limite est exprimée en termes de la fonction « connue » Φ, dont les valeurs peuvent être trouvée en utilisant la table page S-225.

Exercice 7.20. En utilisant le théorème central limite de manière opportune, calculer, pour tout $t \in \mathbb{R}$

$$\lim_{n\to+\infty} \mathrm{e}^{-n} \sum_{k=0}^{n+t\sqrt{n}} \frac{n^k}{k!}. \tag{S7.11}$$

[*Sugg.* Soient $(X_n)_{n\geq 1}$ des variables aléatoires i.i.d. de loi $\mathrm{Poi}(1)$: exprimer la somme (S7.11) en fonction de ces variables aléatoires.]

Solution 7.20. Pour $(X_n)_{n\geq 1}$ des variables aléatoires i.i.d. de loi $\mathrm{Poi}(1)$, la somme $S_n := X_1 + \cdots + X_n$ est de loi $\mathrm{Poi}(n)$ d'après la Proposition 3.128, donc

$$\mathrm{P}(S_n = k) = \mathrm{e}^{-n}\frac{n^k}{k!}, \qquad \forall k \in \mathbb{N}.$$

Par conséquent, on peut écrire

$$\mathrm{e}^{-n} \sum_{k=0}^{n+\sqrt{n}} \frac{n^k}{k!} = \sum_{k=0}^{n+\sqrt{n}} \mathrm{P}(S_n = k) = \mathrm{P}(S_n \leq n + \sqrt{n}) = \mathrm{P}\left(\frac{S_n - n}{\sqrt{n}} \leq 1\right).$$

D'après le théorème central limite, la loi de la variable aléatoire $Z_n := \frac{S_n - n}{\sqrt{n}}$ converge vers la loi $\mathrm{N}(0,1)$ (dans le sens de la convergence ponctuelle de la fonction de répartition), donc, en notant Z une variable normale standard,

$$\lim_{n\to+\infty} \mathrm{e}^{-n} \sum_{k=0}^{n+\sqrt{n}} \frac{n^k}{k!} = \lim_{n\to+\infty} \mathrm{P}(Z_n \leq 1) = \mathrm{P}(Z \leq 1) = \Phi(1) \approx 0,8413,$$

où on a trouvé la valeur de $\Phi(1)$ dans la table page S-225.

Exercice 7.21. Soit $(X_i)_{i\geq 1}$ une suite variables aléatoires i.i.d. de lois marginales données par $\mathrm{P}(X_i = +1) = \mathrm{P}(X_i = -1) = \frac{1}{2}$. On pose $S_0 = 0$ et $S_n = \sum_{i=1}^n X_i$ pour $n \geq 1$. Il s'agit de la définition de la marche aléatoire simple symétrique de la Section 4.4 : S_n représente la position d'un marcheur aléatoire après n pas.

(i) Estimer la probabilité $P(|S_n| \geq \sqrt{n})$ qu'après un grand nombre n de pas, le marcheur soit à une distance supérieure à $\sqrt{n}$ de son point de départ.

(ii) Soit $\alpha \in]0,1[$. Montrer que $\lim_{n\to\infty} P(S_n^2 \geq n^\alpha) = 1$.

Solution 7.21. (i) Observons que $\mu := E(X_i) = -1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 0$ alors que $E(X_i^2) = 1$, car on a $X_i^2 = 1$, donc $\sigma^2 = \mathrm{Var}(X_i) = E(X_i^2) - E(X_i)^2 = 1$. D'après le théorème central limite, la variable aléatoire $Z_n := (S_n - \mu n)/(\sigma\sqrt{n}) = S_n/\sqrt{n}$ possède approximativement la même fonction de répartition que $Z \sim N(0,1)$, donc

$$\begin{aligned} P(|S_n| \geq \sqrt{n}) &= P(|Z_n| \geq 1) = P(Z_n \geq 1) + P(Z_n \leq -1) \\ &\approx P(Z \geq 1) + P(Z \leq -1) = 1 - \Phi(1) + \Phi(-1)\,. \end{aligned}$$

Comme on a $\Phi(x) = 1 - \Phi(-x)$ (voir la relation (7.26)), on obtient $P(|S_n| \geq \sqrt{n}) \approx 2(1-\Phi(1))$. On trouve $\Phi(1) \approx 0,8413$ dans la table page S-225, dont on déduit l'estimée $P(|S_n| \geq \sqrt{n}) \approx 2(1-0,8413) = 31,74\%$.

(ii) Rappelons la relation (7.31), qui découle du théorème central limite : pour toute suite $(z_n)_{n\geq 1}$ qui admet une limite $z := \lim_{n\to\infty} z_n \in \mathbb{R}$, on a

$$\lim_{n\to\infty} P(Z_n \leq z_n) = \lim_{n\to\infty} P(Z_n < z_n) = \Phi(z)\,.$$

En se rappelant que $Z_n = S_n/\sqrt{n}$, comme on a $\lim_{n\to\infty} n^{\alpha-\frac{1}{2}} = 0$ pour $\alpha < \frac{1}{2}$, on obtient

$$\begin{aligned} \lim_{n\to\infty} P(|S_n| \geq n^\alpha) &= \lim_{n\to\infty} P(|Z_n| \geq n^{\alpha-\frac{1}{2}}) \\ &= \lim_{n\to\infty} \big\{ P(Z_n \geq n^{\alpha-\frac{1}{2}}) + P(Z_n \leq -n^{\alpha-\frac{1}{2}}) \big\} \\ &= \lim_{n\to\infty} \big(1 - P(Z_n < n^{\alpha-\frac{1}{2}}) + P(Z_n \leq -n^{\alpha-\frac{1}{2}})\big) \\ &= 1 - \Phi(0) + \Phi(0) = 1\,. \end{aligned}$$

Exercice 7.22. La longueur des trombones produits par une certaine entreprise sont des variables aléatoires i.i.d. de loi inconnue, dont l'espérance et la variance sont notées μ et σ^2. La valeur de σ^2 est connue et égale à $0,25\,\mathrm{mm}^2$, alors que la valeur de μ (exprimé en mm) est inconnue, et on cherche à l'estimer empiriquement.

Pour cela, on mesure la longueur $X_1,\dots,X_n$ de n trombones choisis au hasard, et on en note la moyenne arithmétique $\overline{X}_n := \frac{1}{n}(X_1 + \cdots + X_n)$. Si n est grand, on sait d'après la loi des grands nombres que $\overline{X}_n$ sera proche de μ. Pour rendre cette affirmation plus quantitative, on choisit un réel $\delta > 0$ et on considère l'intervalle I_δ de largeur δ centré en $\overline{X}_n$, à savoir

$$I_\delta := \big]\overline{X}_n - \delta\,, \overline{X}_n + \delta\big[\,.$$

Déterminer δ_n de façon à ce que la probabilité que l'intervalle I_{δ_n} contienne μ vale approximativement $0,95$ quand n est grand.

[*Sugg.* Exprimer l'événement $\{\mu \in I_{\delta_n}\}$ sous la forme $\{a < \overline{X}_n < b\}$ avec a,b bien choisis.]

Solution 7.22. L'événement $\{\mu \in I_{\delta_n}\}$ = « L'intervalle I_{δ_n} contient μ » peut être écrit de la façon suivante :

$$\{\mu \in I_{\delta_n}\} = \{\overline{X}_n - \delta_n < \mu < \overline{X}_n + \delta_n\} = \{\mu - \delta_n < \overline{X}_n < \mu + \delta_n\}\,.$$

En introduisant la variable aléatoire $Z_n := \frac{\sqrt{n}}{\sigma}(\overline{X}_n - \mu)$, on peut écrire

$$\{\mu \in I_{\delta_n}\} = \left\{ -\frac{\sqrt{n}}{\sigma}\delta_n < Z_n < \frac{\sqrt{n}}{\sigma}\delta_n \right\},$$

donc

$$\mathrm{P}(\mu \in I_{\delta_n}) = \mathrm{P}\left(-\frac{\sqrt{n}}{\sigma}\delta_n < Z_n < \frac{\sqrt{n}}{\sigma}\delta_n \right) = \mathrm{P}\left(Z_n < \frac{\sqrt{n}}{\sigma}\delta_n \right) - \mathrm{P}\left(Z_n \leq -\frac{\sqrt{n}}{\sigma}\delta_n \right).$$

Si on choisit $\delta_n := t\frac{\sigma}{\sqrt{n}}$ avec $t > 0$, d'après le théorème central limite, on obtient

$$\lim_{n\to\infty} \mathrm{P}(\mu \in I_{\delta_n}) = \mathrm{P}(Z < t) - \mathrm{P}(Z \leq -t) = \Phi(t) - \Phi(-t) = 2\Phi(t) - 1,$$

où $Z \sim \mathrm{N}(0,1)$ et où on a utilisé la relation $\Phi(-t) = 1 - \Phi(t)$, grâce à (7.26). Cherchons maintenant $t \in \mathbb{R}$ tel que $2\Phi(t) - 1 = 0,95$ c'est-à-dire $\Phi(t) = 0,975$: on trouve dans la table page S-225 que $t \approx 1,96$. La réponse cherchée est donc $\delta_n \sim 1,96\frac{\sqrt{n}}{\sigma}$.

Exercice 7.23. Soit $(X_i)_{i\geq 1}$ et $(Y_i)_{i\geq 1}$ deux suites indépendantes de variables aléatoires i.i.d. dans L^2, d'espérances $\mu := \mathrm{E}(X_i)$, $\tilde{\mu} := \mathrm{E}(Y_i)$ et de variances $\sigma^2 := \mathrm{Var}(X_i)$ et $\tilde{\sigma}^2 := \mathrm{Var}(Y_i)$. On note

$$Z_n := \frac{\sqrt{n}}{\sigma}(\overline{X}_n - \mu), \qquad \tilde{Z}_n := \frac{\sqrt{n}}{\tilde{\sigma}}(\overline{Y}_n - \tilde{\mu}).$$

(i) Montrer que pour tous $\alpha, \beta \in \mathbb{R}$, dont au moins l'un est non nul, on a

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{1}{\sqrt{\alpha^2+\beta^2}}(\alpha Z_n + \beta\tilde{Z}_n) \leq x \right) = \mathrm{P}(Z \leq x) \qquad \forall x \in \mathbb{R},$$

où $Z \sim \mathrm{N}(0,1)$. Plus généralement, montrer que si $\lim_{n\to\infty} x_n = x \in \mathbb{R}$, alors

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{\alpha Z_n + \beta\tilde{Z}_n}{\sqrt{\alpha^2+\beta^2}} \leq x_n \right) = \mathrm{P}(Z \leq x).$$

(ii) On considère maintenant trois suites $(\alpha_n)_{n\geq 1}$, $(\beta_n)_{n\geq 1}$, $(x_n)_{n\geq 1}$ telles que

$$\lim_{n\to\infty} \alpha_n = \alpha \in \mathbb{R}, \qquad \lim_{n\to\infty} \beta = \beta \in \mathbb{R}, \qquad \lim_{n\to\infty} x_n = x \in \mathbb{R}.$$

Montrer, en utilisant l'inégalité de Bienaymé–Tchebychev, que pour tout $\varepsilon > 0$

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{|(\alpha_n - \alpha)Z_n|}{\sqrt{\alpha^2+\beta^2}} > \varepsilon \right) = 0, \qquad \lim_{n\to\infty} \mathrm{P}\left(\frac{|(\alpha_n - \alpha)\tilde{Z}_n|}{\sqrt{\alpha^2+\beta^2}} > \varepsilon \right) = 0,$$

et en conclure que

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{\alpha_n Z_n + \beta_n\tilde{Z}_n}{\sqrt{\alpha^2+\beta^2}} \leq x_n \right) = \mathrm{P}(Z \leq x).$$

[*Sugg.* Pour obtenir la dernière relation, montrer pour commencer que si U, V, W sont des variables aléatoires réelles, alors pour tout $x \in \mathbb{R}$ et $\varepsilon > 0$ on a les inégalités suivantes

$$\mathrm{P}(U+V+W \leq x) \leq \mathrm{P}(U \leq x + 2\varepsilon) + \mathrm{P}(|V| > \varepsilon) + \mathrm{P}(|W| > \varepsilon),$$
$$\mathrm{P}(U+V+W \leq x) \geq \mathrm{P}(U \leq x - 2\varepsilon) - \mathrm{P}(|V| > \varepsilon) - \mathrm{P}(|W| > \varepsilon).$$

Appliquer ensuite ces inégalités pour un choix de U, V, W opportun.]

Solution 7.23. (i) Observons que l'on peut écrire

$$Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{X}_i \quad \text{avec} \quad \hat{X}_i := \frac{X_i - \mu}{\sigma}, \qquad \tilde{Z}_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \hat{Y}_i \quad \text{avec} \quad \hat{Y}_i := \frac{Y_i - \tilde{\mu}}{\tilde{\sigma}},$$

par conséquent, pour tous $\alpha, \beta \in \mathbb{R}$,

$$\frac{\alpha Z_n + \beta \tilde{Z}_n}{\sqrt{\alpha^2 + \beta^2}} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} W_i \qquad \text{où} \qquad W_i := \frac{\alpha \hat{X}_i + \beta \hat{Y}_i}{\sqrt{\alpha^2 + \beta^2}}. \tag{S7.12}$$

Il découle des hypothèses que les variables aléatoires $\hat{X}_i$ et $\hat{Y}_i$ sont d'espérance nulle et de variance 1 et sont indépendantes ; par conséquent, d'après les propriétés de l'espérance et de la variance, les variables aléatoires $\hat{W}_i$ sont aussi d'espérance nulle et de variance 1 :

$$\mathrm{E}(W_i) = \frac{\alpha\,\mathrm{E}(\hat{X}_i) + \beta\,\mathrm{E}(\hat{Y}_i)}{\sqrt{\alpha^2 + \beta^2}} = 0, \qquad \mathrm{Var}(W_i) = \frac{\alpha^2\,\mathrm{Var}(\hat{X}_i) + \beta^2\,\mathrm{Var}(\hat{Y}_i)}{\alpha^2 + \beta^2} = 1,$$

où on a utilisé l'indépendance de X_i et Y_i pour le calcul de la variance. De l'indépendance des suites $(X_i)_{i \geq 1}$ et $(Y_i)_{i \geq 1}$ et du fait que ces suites de variables aléatoires sont indépendantes, on en déduit que les variables aléatoires $(\hat{W}_i)_{i \geq 1}$ sont aussi indépendantes. D'après le théorème central limite, on a alors

$$\lim_{n \to \infty} \mathrm{P}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} W_i \leq x \right) = \mathrm{P}(Z \leq x) \qquad \text{avec } Z \sim \mathrm{N}(0,1).$$

Plus généralement, en appliquant la relation (7.31), si $\lim_{n \to \infty} x_n = x$ on a

$$\lim_{n \to \infty} \mathrm{P}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} W_i \leq x_n \right) = \mathrm{P}(Z \leq x) \qquad \text{avec } Z \sim \mathrm{N}(0,1).$$

Combinée à la définition (S7.12), on obtient la conclusion voulue.

(ii) Observons que, Z_n étant une moyenne empirique standardisée, elle possède une espérance $\mathrm{E}(Z_n) = 0$ et une variance $\mathrm{Var}(Z_n) = 1$, voir (7.21). En utilisant l'inégalité de Bienaymé–Tchebychev et les propriétés de la variance, on obtient donc

$$\mathrm{P}\left(\frac{|(\alpha_n - \alpha) Z_n|}{\sqrt{\alpha^2 + \beta^2}} > \varepsilon \right) \leq \frac{1}{\varepsilon^2}\,\mathrm{Var}\left(\frac{(\alpha_n - \alpha) Z_n}{\sqrt{\alpha^2 + \beta^2}} \right) = \frac{(\alpha_n - \alpha)^2}{\varepsilon^2(\alpha^2 + \beta^2)},$$

qui tend vers 0 quand $n \to +\infty$. La relation avec $\tilde{Z}_n$ à la place de Z_n est analogue.

Considérons maintenant trois variables aléatoires réelles U, V, W et fixons $x \in \mathbb{R}$, $\varepsilon > 0$. On observe que l'événement $\{U + V + W \leq x, |V| \leq \varepsilon, |W| \leq \varepsilon\}$ est inclus dans l'événement $\{U \leq x + 2\varepsilon\}$, donc en posant $\mathcal{D} := \{|V| \leq \varepsilon, |W| \leq \varepsilon\}$ on peut écrire

$$\begin{aligned} \mathrm{P}(U + V + W \leq x) &= \mathrm{P}(U + V + W \leq x, \mathcal{D}) + \mathrm{P}(U + V + W \leq x, \mathcal{D}^c) \\ &\leq \mathrm{P}(U \leq x + 2\varepsilon) + \mathrm{P}(\mathcal{D}^c) \\ &\leq \mathrm{P}(U \leq x + 2\varepsilon) + \mathrm{P}(|V| > \varepsilon) + \mathrm{P}(|W| > \varepsilon), \end{aligned} \tag{S7.13}$$

car on a $\mathcal{D}^c = \{|V| > \varepsilon\} \cup \{|W| > \varepsilon\}$. De manière analogue, comme l'événement $\{U + V + W \leq x, |V| \leq \varepsilon, |W| \leq \varepsilon\}$ contient $\{U \leq x - 2\varepsilon, |V| \leq \varepsilon, |W| \leq \varepsilon\}$, on peut écrire

$$\begin{aligned}
\mathrm{P}(U+V+W \le x) &\ge \mathrm{P}(U+V+W \le x, \mathcal{D}) \\
&\ge \mathrm{P}(U \le x-2\varepsilon, \mathcal{D}) = \mathrm{P}(U \le x-2\varepsilon) - \mathrm{P}(U \le x-2\varepsilon, \mathcal{D}^c) \\
&\ge \mathrm{P}(U \le x-2\varepsilon) - \mathrm{P}(\mathcal{D}^c) \\
&\ge \mathrm{P}(U \le x-2\varepsilon) - \mathrm{P}(|V| > \varepsilon) - \mathrm{P}(|W| > \varepsilon)\,.
\end{aligned} \tag{S7.14}$$

Appliquons maintenant ces inégalités aux variables aléatoires suivantes :

$$U := \frac{\alpha Z_n + \beta \tilde{Z}_n}{\sqrt{\alpha^2+\beta^2}}\,, \qquad V := \frac{(\alpha_n - \alpha) Z_n}{\sqrt{\alpha^2+\beta^2}}\,, \qquad W := \frac{(\alpha_n - \alpha) \tilde{Z}_n}{\sqrt{\alpha^2+\beta^2}}\,.$$

Grâce à (S7.13), on peut écrire

$$\begin{aligned}
\limsup_{n\to\infty} \mathrm{P}\left(\frac{\alpha_n Z_n + \beta_n \tilde{Z}_n}{\sqrt{\alpha^2+\beta^2}} \le x_n \right) &= \limsup_{n\to\infty} \mathrm{P}\big(U+V+W \le x_n\big) \\
&\le \limsup_{n\to\infty} \big\{ \mathrm{P}\big(U \le x_n + 2\varepsilon\big) + \mathrm{P}\big(|V| > \varepsilon\big) + \mathrm{P}\big(|W| > \varepsilon\big) \big\} \\
&= \mathrm{P}(Z \le x + 2\varepsilon)\,,
\end{aligned}$$

où on a utilisé les relations démontrées plus haut. Avec des arguments analogues, en utilisant (S7.14) on obtient

$$\liminf_{n\to\infty} \mathrm{P}\left(\frac{\alpha_n Z_n + \beta_n \tilde{Z}_n}{\sqrt{\alpha^2+\beta^2}} \le x_n \right) \ge \mathrm{P}(Z \le x - 2\varepsilon)\,.$$

Comme $\varepsilon > 0$ est arbitraire et que la fonction Φ définie par $\Phi(x) = \mathrm{P}(Z \le x)$ est continue, on a montré la relation finale.

Exercice 7.24. Soit $(X_i)_{i\ge1}$ une suite de variables aléatoires i.i.d. telles que $\log X_i$ admette une espérance finie. Pour tout $n \in \mathbb{N}^*$, on pose $Y_n := \prod_{i=1}^n X_i$.

(i) En appliquant la loi des grands nombres à la suite $(\log X_i)_{i\ge1}$, montrer que pour tout $\varepsilon > 0$ on a

$$\lim_{n\to\infty} \mathrm{P}\Big(\big| (Y_n)^{1/n} - \mathrm{e}^{\mathrm{E}(\log X_1)} \big| > \varepsilon \Big) = 0\,.$$

[*Sugg.* On pourra utiliser l'Exercice 7.1.]

(ii) En déduire que :
- si $\mathrm{E}(\log X_1) < 0$, alors il existe $\theta < 1$ tel que $\lim_{n\to\infty} \mathrm{P}(Y_n \le \theta^n) = 1$;
- si $\mathrm{E}(\log X_1) > 0$, alors il existe $\theta > 1$ tel que $\lim_{n\to\infty} \mathrm{P}(Y_n \ge \theta^n) = 1$;

(iii) Supposons maintenant que $\mathrm{E}(\log X_1) = 0$ et que $\sigma^2 = \mathrm{E}((\log X_1)^2)$ est finie et non nulle. En appliquant le théorème central limite à la suite $(\log X_i)_{i\ge1}$, montrer que

$$\lim_{n\to\infty} \mathrm{P}\Big(Y_n \le \mathrm{e}^{-n^{1/4}} \Big) = \frac{1}{2} \quad \text{et} \quad \lim_{n\to\infty} \mathrm{P}\Big(Y_n \ge \mathrm{e}^{n^{1/4}} \Big) = \frac{1}{2}\,.$$

Solution 7.24. (i) On introduit les variables aléatoires $W_i := \log X_i$, qui par hypothèse admettent une espérance finie, et on pose $\overline{W}_n := \frac{1}{n}\sum_{i=1}^n W_i$ leur moyenne empirique. En considérant la fonction continue $f(y) := \mathrm{e}^y$, on observe que

$$f(\overline{W}_n) = \mathrm{e}^{\frac{1}{n}\sum_{i=1}^n \log X_i} = \prod_{i=1}^n (X_i)^{1/n} = (Y_n)^{1/n}\,, \qquad f(\mathrm{E}(W_1)) = \mathrm{e}^{\mathrm{E}(\log X_1)}\,.$$

En appliquant l'Exercice 7.1, on obtient la conclusion voulue.

(ii) Considérons le cas $\mathrm{E}(\log X_1) < 0$. Alors $\mathrm{e}^{\mathrm{E}(\log X_1)} < 1$ et on peut choisir $\varepsilon > 0$ suffisamment petit pour que $\theta := \mathrm{e}^{\mathrm{E}(\log X_1)} + \varepsilon < 1$. Comme

$$\mathrm{P}(Y_n > \theta^n) = \mathrm{P}\left((Y_n)^{1/n} > \mathrm{e}^{\mathrm{E}(\log X_1)} + \varepsilon\right) \leq \mathrm{P}\left(|(Y_n)^{1/n} - \mathrm{e}^{\mathrm{E}(\log X_1)}| > \varepsilon\right),$$

d'après le point précédent, on a

$$\lim_{n\to\infty} \mathrm{P}(Y_n > \theta^n) \leq \lim_{n\to\infty} \mathrm{P}\left(|(Y_n)^{1/n} - \mathrm{e}^{\mathrm{E}(\log X_1)}| > \varepsilon\right) = 0$$

et donc $\lim_{n\to\infty} \mathrm{P}(Y_n \leq \theta^n) = 1$.

Considérons maintenant le cas $\mathrm{E}(\log X_1) > 0$. Alors $\mathrm{e}^{\mathrm{E}(\log X_1)} > 1$ et on choisit $\varepsilon > 0$ suffisamment petit pour que $\theta := \mathrm{e}^{\mathrm{E}(\log X_1)} - \varepsilon > 1$. Comme

$$\mathrm{P}(Y_n < \theta^n) = \mathrm{P}\left((Y_n)^{1/n} < \mathrm{e}^{\mathrm{E}(\log X_1)} - \varepsilon\right) \leq \mathrm{P}\left(|(Y_n)^{1/n} - \mathrm{e}^{\mathrm{E}(\log X_1)}| > \varepsilon\right),$$

d'après le point précédent on a $\lim_{n\to\infty} \mathrm{P}(Y_n < \theta^n) \leq \mathrm{P}(|(Y_n)^{1/n} - \mathrm{e}^{\mathrm{E}(\log X_1)}| > \varepsilon) = 0$ et donc $\lim_{n\to\infty} \mathrm{P}(Y_n \geq \theta^n) = 1$.

(iii) On peut écrire

$$\mathrm{P}\left(Y_n \leq \mathrm{e}^{-n^{1/4}}\right) = \mathrm{P}\left(\log Y_n \leq -n^{1/4}\right) = \mathrm{P}\left(\sum_{i=1}^{n} \log X_i \leq -n^{1/4}\right).$$

Par hypothèse, les variables aléatoires $(\log X_i)_{i\geq 1}$ on une espérance nulle et une variance $\sigma^2 \in]0,\infty[$, donc d'après le théorème central limite la loi de la variable aléatoire $Z_n := \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{n} \log X_i$ converge vers la loi de $Z \sim \mathrm{N}(0,1)$ quand $n \to \infty$ (au sens de la convergence ponctuelle de la fonction de répartition). Par conséquent, pour toute suite convergente $x_n \to x \in \mathbb{R}$, on a

$$\lim_{n\to\infty} \mathrm{P}(Z_n \leq x_n) = \lim_{n\to\infty} \mathrm{P}(Z_n < x_n) = \Phi(x),$$

grâce à la relation (7.31). Cela permet de conclure que

$$\lim_{n\to\infty} \mathrm{P}\left(Y_n \leq \mathrm{e}^{-n^{1/4}}\right) = \lim_{n\to\infty} \mathrm{P}\left(Z_n \leq -\frac{1}{\sigma\sqrt{n}} n^{1/4}\right) = \Phi(0) = \frac{1}{2}.$$

Avec des arguments analogues on obtient

$$\begin{aligned}\lim_{n\to\infty} \mathrm{P}\left(Y_n \geq \mathrm{e}^{n^{1/4}}\right) &= \lim_{n\to\infty} \mathrm{P}\left(Z_n \geq \frac{1}{\sigma\sqrt{n}} n^{1/4}\right) \\ &= 1 - \lim_{n\to\infty} \mathrm{P}\left(Z_n < \frac{1}{\sigma\sqrt{n}} n^{1/4}\right) = 1 - \Phi(0) = \frac{1}{2}.\end{aligned}$$

Exercice 7.25 (Marche aléatoire). Soit $(X_i)_{i\geq 1}$ une suite de variables aléatoires i.i.d. On suppose que $\mathbb{E}(X_i) = 0$ et que $\mathrm{Var}(X_i) = \sigma^2$ est finie et non nulle. On pose $S_0 = 0$ et pour $n \geq 1$, $S_n = \sum_{i=1}^{n} X_i$. On va s'intéresser à la suite $(S_n)_{n\geq 0}$, interprétée comme la suite des positions d'un marcheur aléatoire dont le i-ème pas est donné par X_i.

(i) Montrer que $\lim_{k\to\infty} \mathrm{P}(S_k \geq k^{1/4}) = 1/2$.

(ii) En utilisant le fait que pour $k \in \mathbb{N}^*$ on a $\mathrm{P}(\sup_{n\geq 0} S_n \geq k^{1/4}) \geq \mathrm{P}(S_k \geq k^{1/4})$, montrer que $\mathrm{P}(\sup_{n\geq 0} S_n = +\infty) \geq 1/2$.

(iii) Montrer que l'on a de même $P(\inf_{n\geq 0} S_n = -\infty) \geq 1/2$.

(iv) D'après l'Exercice 6.63, les probabilités $P(\sup_{n\geq 0} S_n = +\infty)$ et $P(\inf_{n\geq 0} S_n = -\infty)$ ne peuvent valoir que 0 ou 1. En déduire que l'on a

$$P\left(\sup_{n\geq 0} S_n = +\infty \ \text{ et } \ \inf_{n\geq 0} S_n = -\infty\right) = 1\,.$$

Que peut-on en conclure sur les trajectoires du marcheur ?

Solution 7.25. (i) D'après le théorème central limite, on a

$$\begin{aligned}\lim_{k\to\infty} P(S_k \geq k^{1/4}) &= \lim_{k\to\infty} P\left(\frac{S_k}{\sigma\sqrt{k}} \geq \frac{k^{1/4}}{\sigma\sqrt{k}}\right)\\ &= 1 - \lim_{k\to\infty} P\left(\frac{S_k}{\sigma\sqrt{k}} < \frac{k^{1/4}}{\sigma\sqrt{k}}\right) = 1 - \Phi(0) = \frac{1}{2}\end{aligned}$$

car on a $\lim_{k\to\infty} \frac{k^{1/4}}{\sigma\sqrt{k}} = 0$, voir la relation (7.31).

(ii) Pour tout $k \in \mathbb{N}^*$ on a l'inclusion d'événements $\{\sup_{n\geq 0} S_n \geq k^{1/4}\} \supseteq \{S_k \geq k^{1/4}\}$, donc par monotonie des probabilités on a $P(\sup_{n\geq 0} S_n = +\infty) \geq P(S_k \geq k^{1/4})$.
Observons maintenant que pour toute variable aléatoire W à valeurs dans $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ on a l'égalité d'événements $\{W = +\infty\} = \bigcap_{k=1}^{\infty} \{W \geq k^{1/4}\}$, en particulier

$$\left\{\sup_{n\geq 0} S_n = +\infty\right\} = \bigcap_{k=1}^{\infty} \left\{\sup_{n\geq 0} S_n \geq k^{1/4}\right\}.$$

Comme les événements $\{W \geq k^{1/4}\}$ sont décroissants en k (pour l'inclusion), par continuité par le haut des probabilités on obtient

$$P\left(\sup_{n\geq 0} S_n = +\infty\right) = \lim_{k\to\infty} P\left(\sup_{n\geq 0} S_n \geq k^{1/4}\right) \geq \lim_{k\to\infty} P(S_k \geq k^{1/4}) = \frac{1}{2}\,.$$

(iii) En appliquant le point précédent aux variables aléatoires $(-X_i)_{i\geq 1}$, qui sont i.i.d. d'espérance nulle et de variance σ^2, on obtient $P(\sup_{n\geq 0}(-S_n) = +\infty) \geq 1/2$. Comme $\sup_{n\geq 0}(-S_n) = -\inf_{n\geq 0} S_n$, on en déduit que $P(\inf_{n\geq 0} S_n = -\infty) \geq 1/2$.

(iv) Les probabilités $P(\sup_{n\geq 0} S_n = +\infty)$ et $P(\inf_{n\geq 0} S_n = -\infty)$ peuvent seulement valoir 0 ou 1, d'après l'Exercice 6.63. On a montré dans les points précédents que celles-ci valent au moins $\frac{1}{2}$, par conséquent $P(\sup_{n\geq 0} S_n = +\infty) = 1$ et $P(\inf_{n\geq 0} S_n = -\infty) = 1$. D'après l'Exercice 1.1, l'intersection (dénombrable) d'événements de probabilité 1 est encore de probabilité 1, donc $P(\sup_{n\geq 0} S_n = +\infty \ \text{ et } \ \inf_{n\geq 0} S_n = -\infty) = 1$.
Cela signifie que les trajectoires de la marche aléatoire $(S_n)_{n\geq 1}$ ont un caractère oscillant : pour les temps grands elles prennent des valeurs arbitrairement grandes, aussi bien positives que négatives.

Exercices plus difficiles

Exercice 7.26 (Concentration de la mesure, III). Soit $(X_i)_{i\geq 1}$ des variables aléatoires indépendantes et de même loi, de deuxième moment $E(X_i^2) = 1$ et de moment d'ordre 4 fini : $m_4 := E(X_i^4) < +\infty$. On fait l'hypothèse que $m_4 > 1$.

On considère le vecteur aléatoire

$$V_n := (X_1, \dots, X_n),$$

et soit $\|V_n\|$ sa norme euclidienne. On souhaite montrer que la fonction de répartition de la variable aléatoire $\|V_n\| - \sqrt{n}$, c'est-à-dire

$$F_n(t) := \mathrm{P}(\|V_n\| - \sqrt{n} \le t) \qquad \text{pour } t \in \mathbb{R}, \tag{S7.15}$$

converge quand $n \to \infty$ vers la fonction de répartition d'une variable aléatoire normale.

(i) Expliquer pourquoi les variables aléatoires $(X_i^2)_{i \ge 1}$ sont i.i.d., d'espérance 1 et de variance $m_4 - 1 > 0$.

(ii) On pose $Z_n = \|V_n\|^2 = \sum_{i=1}^n X_i^2$. En appliquant la loi des grands nombres à Z_n, montrer que pour tout $\varepsilon > 0$ on a

$$\lim_{n\to\infty} \mathrm{P}\Big(\frac{1}{\sqrt{n}}\|V_n\| \in [1-\varepsilon, 1+\varepsilon]\Big) = 1.$$

(iii) Considérons maintenant une version centrée et normalisée de Z_n :

$$\widehat{Z}_n = \frac{Z_n - n}{\sqrt{n}}.$$

Montrer que pour toute suite $(t_n)_{n\ge1}$ telle que $\lim_{n\to\infty} t_n = z \in \mathbb{R}$ on a

$$\lim_{n\to\infty} \mathrm{P}(\widehat{Z}_n \le t_n) = \mathrm{P}(\widehat{Z} \le z)$$

où $\widehat{Z} \sim \mathrm{N}(0, \tilde{\sigma}^2)$, pour une constante $\tilde{\sigma}^2$ dont on précisera la valeur.

[*Sugg.* On pourra utiliser la relation (7.31).]

(iv) En se rappelant de (S7.15), observer que $F_n(t) = \mathrm{P}\big((\frac{1}{n}Z_n)^{1/2} - 1 \le \frac{t}{\sqrt{n}}\big)$. En déduire que pour tout $t \in \mathbb{R}$

$$F_n(t) = \mathrm{P}\big(\widehat{Z}_n \le t_n\big), \qquad \text{avec } t_n := \sqrt{n}\Big(\big(1 + \tfrac{t}{\sigma\sqrt{n}}\big)^2 - 1\Big).$$

Conclure que

$$\lim_{n\to\infty} F_n(t) = \mathrm{P}\big(\tfrac{1}{2}\widehat{Z} \le t\big), \qquad \text{pour tout } t \in \mathbb{R}.$$

Quelle est la loi de $\frac{1}{2}\widehat{Z}$?

Solution 7.26. (i) Les variables aléatoires $(X_n^2)_{n\ge1}$ sont indépendantes et identiquement distribuées, par conservation de l'indépendance et de la loi. Par hypothèse $\mathrm{E}(X_i^2) = 1$ et $\mathrm{Var}(X_i^2) = \mathrm{E}((X_i^2)^2) - \mathrm{E}(X_i^2)^2 = \mathrm{E}(X_i^4) - 1 = m_4 - 1 > 0$.

(ii) On peut supposer que $\varepsilon \in\,]0,1[$. Alors

$$\mathrm{P}\Big(\frac{1}{\sqrt{n}}\|V_n\| \in [1-\varepsilon, 1+\varepsilon]\Big) = \mathrm{P}\Big(\frac{1}{n}\|V_n\|^2 \in [(1-\varepsilon)^2, (1+\varepsilon)^2]\Big).$$

Comme $(1-\varepsilon)^2 < 1$ et $(1+\varepsilon)^2 > 1$, on peut fixer $\delta > 0$ petit de sorte que $(1-\varepsilon)^2 < 1-\delta$ et $(1+\varepsilon)^2 > 1+\delta$. Ainsi, on peut écrire

$$\lim_{n\to\infty} \mathrm{P}\Big(\frac{1}{\sqrt{n}}\|V_n\| \in [1-\varepsilon, 1+\varepsilon]\Big) \ge \lim_{n\to\infty} \mathrm{P}\Big(\frac{1}{n}\sum_{i=1}^n X_i^2 \in [1-\delta, 1+\delta]\Big) = 1,$$

d'après la loi des grands nombres appliquée aux variables aléatoires $(X_n^2)_{n\geq 1}$.

(iii) En introduisant les variables aléatoires $W_i := X_i^2$, qui ont une espérance $\mu := \mathrm{E}(W_i) = 1$ et une variance $\sigma^2 := \mathrm{Var}(W_i) = m_4 - 1$, on peut écrire

$$\widehat{Z}_n = \frac{Z_n - n}{\sqrt{n}} = \frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n} W_i - n\right) = \sqrt{n}\,(\overline{W}_n - 1)\,.$$

D'après le théorème central limite, en introduisant une variable aléatoire $Z \sim \mathrm{N}(0,1)$ et en utilisant la relation (7.31), si $\lim_{n\to\infty} t_n = z$ on obtient

$$\lim_{n\to\infty} \mathrm{P}(\widehat{Z}_n \leq t_n) = \lim_{n\to\infty} \mathrm{P}\left(\frac{\sqrt{n}}{\sigma}(\overline{W}_n - 1) \leq \frac{t_n}{\sigma}\right) = \mathrm{P}\left(Z \leq \frac{z}{\sigma}\right) = \mathrm{P}(\widehat{Z} \leq z)\,,$$

où $\widehat{Z} := \sigma Z \sim \mathrm{N}(0,\sigma^2)$, où on rappelle que $\sigma^2 = m_4 - 1$.

(iv) De la définition (S7.15) de F_n, en se rappelant que $Z_n := \|V_n\|^2$, on en déduit que

$$F_n(t) = \mathrm{P}(\|V_n\| - \sqrt{n} \leq t) = \mathrm{P}\left(\left(\frac{Z_n}{n}\right)^{1/2} - 1 \leq \frac{t}{\sqrt{n}}\right) = \mathrm{P}\left(\frac{Z_n}{n} \leq \left(1 + \frac{t}{\sqrt{n}}\right)^2\right)$$
$$= \mathrm{P}\left(\frac{Z_n - n}{\sqrt{n}} \leq \sqrt{n}\left(\left(1 + \frac{t}{\sqrt{n}}\right)^2 - 1\right)\right) = \mathrm{P}\left(\widehat{Z}_n \leq t_n\right),$$

où on rappelle que $\widehat{Z}_n = \frac{Z_n - n}{\sqrt{n}}$ et où on a défini $t_n := \sqrt{n}\left(\left(1 + \frac{t}{\sqrt{n}}\right)^2 - 1\right)$. Comme on a $t_n = \sqrt{n}\left(2\frac{t}{\sqrt{n}} + \frac{t^2}{n}\right) = 2t + \frac{t^2}{\sqrt{n}}$ on obtient $\lim_{n\to\infty} t_n = 2t$, donc d'après le point précédent on obtient la relation voulue :

$$\lim_{n\to\infty} F_n(t) = \lim_{n\to\infty} \mathrm{P}\left(\widehat{Z}_n \leq t_n\right) = \mathrm{P}(\widehat{Z} \leq 2t) = \mathrm{P}(\tfrac{1}{2}\widehat{Z} \leq t)\,.$$

La loi de $\frac{1}{2}\widehat{Z}$ est $\mathrm{N}(0, \frac{1}{4}\sigma^2) = \mathrm{N}(0, \frac{m_4 - 1}{4})$.

Exercice 7.27. Soit $(X_i)_{i\geq 1}$ une suite de variables aléatoires i.i.d. *positives*, d'espérance $\mu := \mathrm{E}(X_i) > 0$ et de variance $\sigma^2 := \mathrm{Var}(X_i)$ finie et non nulle. On note $\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i$ et on pose $Z_n := \frac{\sqrt{n}}{\sigma}(\overline{X}_n - \mu)$.

(i) Montrer que $\log \overline{X}_n = \log \mu + \log\left(1 + \frac{\sigma}{\mu\sqrt{n}} Z_n\right)$.

(ii) Montrer que pour tout $t \in \mathbb{R}$, on a

$$\mathrm{P}\left(\log \overline{X}_n - \log \mu \leq t\right) = \mathrm{P}\left(Z_n \leq \frac{\mu\sqrt{n}}{\sigma}(\mathrm{e}^t - 1)\right).$$

(iii) En déduire que pour tout $x \in \mathbb{R}$ on a

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{\mu\sqrt{n}}{\sigma}(\log \overline{X}_n - \log \mu) \leq x\right) = \mathrm{P}(Z \leq x)\,,$$

où $Z \sim \mathrm{N}(0,1)$.

[*Sugg.* On pourra utiliser la relation (7.31).]

Considérons maintenant une autre suite $(Y_i)_{i\geq 1}$ de variables aléatoires i.i.d. *positives*, indépendantes des $(X_i)_{i\geq 1}$, d'espérance $\tilde{\mu} := \mathrm{E}(Y_i) > 0$ et de variance $\tilde{\sigma}^2 := \mathrm{Var}(Y_i)$ finie et non nulle. On pose $\overline{Y}_n := \frac{1}{n}\sum_{i=1}^{n} Y_i$ et $\tilde{Z}_n := \frac{\sqrt{n}}{\tilde{\sigma}}(\overline{Y}_n - \tilde{\mu})$. On étudie le rapport

$$R_n := \frac{X_1 + \cdots + X_n}{Y_1 + \cdots + Y_n} = \frac{\overline{X}_n}{\overline{Y}_n}\,.$$

(iv) Montrer que $\log R_n = \log(\mu/\tilde{\mu}) + \log\big(1 + \frac{\sigma}{\mu\sqrt{n}} Z_n\big) - \log\big(1 + \frac{\tilde{\sigma}}{\tilde{\mu}\sqrt{n}} \tilde{Z}_n\big)$.

(v) En déduire que pour tout $t \in \mathbb{R}$ on a

$$\mathrm{P}\big(\log R_n - \log(\mu/\tilde{\mu}) \le t\big) = \mathrm{P}\left(\frac{\sigma}{\mu} Z_n - \frac{\tilde{\sigma}}{\tilde{\mu}} \mathrm{e}^t \tilde{Z}_n \le \sqrt{n}(\mathrm{e}^t - 1)\right).$$

(vi) Conclure que si on pose $\widehat{\sigma} := \sqrt{(\frac{\sigma}{\mu})^2 + (\frac{\tilde{\sigma}}{\tilde{\mu}})^2}$, alors on a

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{\sqrt{n}}{\widehat{\sigma}}\left(\log R_n - \log\frac{\mu}{\tilde{\mu}}\right) \le x\right) = \mathrm{P}(Z \le x)\,, \qquad \text{pour tout } x \in \mathbb{R}\,,$$

où $Z \sim \mathrm{N}(0,1)$.

[*Sugg.* On pourra utiliser l'Exercice 7.23.]

Solution 7.27. (i) On peut écrire

$$\overline{X}_n = \mu + (\overline{X}_n - \mu) = \mu\left(1 + \frac{1}{\mu}(\overline{X}_n - \mu)\right) = \mu\left(1 + \frac{\sigma}{\mu\sqrt{n}} Z_n\right),$$

dont il découle que $\log \overline{X}_n = \log\mu + \log(1 + \frac{\sigma}{\mu\sqrt{n}} Z_n)$.

(ii) D'après le point précédent

$$\begin{aligned}\mathrm{P}\big(\log \overline{X}_n - \log\mu \le t\big) = \mathrm{P}\left(\log\left(1 + \frac{\sigma}{\mu\sqrt{n}} Z_n\right) \le t\right) &= \mathrm{P}\left(1 + \frac{\sigma}{\mu\sqrt{n}} Z_n \le \mathrm{e}^t\right)\\ &= \mathrm{P}\left(Z_n \le \frac{\mu\sqrt{n}}{\sigma}(\mathrm{e}^t - 1)\right).\end{aligned}$$

(iii) D'après le point précédent avec $t = x\frac{\sigma}{\mu\sqrt{n}}$, on peut écrire

$$\begin{aligned}\lim_{n\to\infty} \mathrm{P}\left(\frac{\mu\sqrt{n}}{\sigma}(\log \overline{X}_n - \log\mu) \le x\right) &= \lim_{n\to\infty} \mathrm{P}\left(\log \overline{X}_n - \log\mu \le x\frac{\sigma}{\mu\sqrt{n}}\right)\\ &= \lim_{n\to\infty} \mathrm{P}\left(Z_n \le \frac{\mu\sqrt{n}}{\sigma}(\mathrm{e}^{x\frac{\sigma}{\mu\sqrt{n}}} - 1)\right) = \mathrm{P}(Z \le x)\,,\end{aligned}$$

où pour la dernière égalité on a appliqué le théorème central limite, plus précisément la relation (7.31), grâce au fait que $\lim_{n\to\infty} \frac{\mu\sqrt{n}}{\sigma}(\mathrm{e}^{x\frac{\sigma}{\mu\sqrt{n}}} - 1) = x$.

(iv) Comme $\log R_n = \log \overline{X}_n - \log \overline{Y}_n$, en appliquant deux fois le point (i) on obtient la relation $\log R_n = \log(\mu/\tilde{\mu}) + \log\big(1 + \frac{\sigma}{\mu\sqrt{n}} Z_n\big) - \log\big(1 + \frac{\tilde{\sigma}}{\tilde{\mu}\sqrt{n}} \tilde{Z}_n\big)$.

(v) En appliquant le point précédent, avec quelques manipulations on obtient

$$\begin{aligned}
\mathrm{P}\left(\log R_n - \log(\mu/\tilde{\mu}) \le t\right) &= \mathrm{P}\left(\log\left(1+\frac{\sigma}{\mu\sqrt{n}}Z_n\right) \le \log\left(1+\frac{\tilde{\sigma}}{\tilde{\mu}\sqrt{n}}\tilde{Z}_n\right)+t\right)\\
&= \mathrm{P}\left(1+\frac{\sigma}{\mu\sqrt{n}}Z_n \le \left(1+\frac{\tilde{\sigma}}{\tilde{\mu}\sqrt{n}}\tilde{Z}_n\right)\mathrm{e}^t\right)\\
&= \mathrm{P}\left(\frac{\sigma}{\mu}Z_n - \frac{\tilde{\sigma}}{\tilde{\mu}}\mathrm{e}^t\tilde{Z}_n \le \sqrt{n}(\mathrm{e}^t-1)\right).
\end{aligned}$$

(vi) Grâce au point précédent avec $t := x\frac{\widehat{\sigma}}{\sqrt{n}}$, on peut écrire

$$\begin{aligned}
\mathrm{P}\left(\frac{\sqrt{n}}{\widehat{\sigma}}\left(\log R_n - \log\frac{\mu}{\tilde{\mu}}\right) \le x\right) &= \mathrm{P}\left(\log R_n - \log\frac{\mu}{\tilde{\mu}} \le x\frac{\widehat{\sigma}}{\sqrt{n}}\right)\\
&= \mathrm{P}\left(\frac{\sigma}{\mu}Z_n - \frac{\tilde{\sigma}}{\tilde{\mu}}\mathrm{e}^{x\frac{\widehat{\sigma}}{\sqrt{n}}}\tilde{Z}_n \le \sqrt{n}(\mathrm{e}^{x\frac{\widehat{\sigma}}{\sqrt{n}}}-1)\right).
\end{aligned}$$

Si on définit $\alpha = \frac{\sigma}{\mu}$ et $\beta_n = -\frac{\tilde{\sigma}}{\tilde{\mu}}\mathrm{e}^{x\frac{\widehat{\sigma}}{\sqrt{n}}}$, on peut écrire

$$\mathrm{P}\left(\frac{\sqrt{n}}{\widehat{\sigma}}\left(\log R_n - \log\frac{\mu}{\tilde{\mu}}\right) \le x\right) = \mathrm{P}\left(\alpha Z_n + \beta_n\tilde{Z}_n \le \sqrt{n}(\mathrm{e}^{x\frac{\widehat{\sigma}}{\sqrt{n}}}-1)\right).$$

Notons que $\lim_{n\to\infty}\beta_n = \beta := \frac{\tilde{\sigma}}{\tilde{\mu}}$ et définissons $x_n := \frac{1}{\sqrt{\alpha^2+\beta^2}}\sqrt{n}(\mathrm{e}^{x\frac{\widehat{\sigma}}{\sqrt{n}}}-1)$. Alors on a $\lim_{n\to\infty}x_n = \frac{x\widehat{\sigma}}{\sqrt{\alpha^2+\beta^2}} = x$, donc grâce à l'Exercice 7.23 on obtient

$$\begin{aligned}
\lim_{n\to\infty}\mathrm{P}\left(\frac{\sqrt{n}}{\widehat{\sigma}}\left(\log R_n - \log\frac{\mu}{\tilde{\mu}}\right) \le x\right) &= \lim_{n\to\infty}\mathrm{P}\left(\frac{1}{\sqrt{\alpha^2+\beta^2}}\left(\alpha Z_n + \beta\tilde{Z}_n\right) \le x_n\right)\\
&= \mathrm{P}(Z \le x),
\end{aligned}$$

avec $Z \sim \mathrm{N}(0,1)$.

Exercice 7.28. On considère une variable aléatoire X qui suit la loi *de Pareto* de paramètre $\alpha \in \,]0,1[$: cela signifie que X a pour densité

$$f_X(x) = \frac{\alpha}{x^{1+\alpha}}\mathbb{1}_{[1,+\infty[}(x).$$

(i) Montrer que f_X est bien une densité de probabilité, et que l'on a $\mathrm{E}(X) = +\infty$.

(ii) Calculer la fonction de répartition F_X de X.

Soit $(X_i)_{i\ge 1}$ une suite de variables aléatoires i.i.d., de loi de Pareto de paramètre $\alpha \in \,]0,1[$. Pour $n \in \mathbb{N}^*$, on pose

$$M_n = \max\{X_1,\ldots,X_n\} \qquad \text{et} \qquad S_n = \sum_{i=1}^n X_i.$$

(iii) Calculer la fonction de répartition de de M_n, que l'on notera F_{M_n}.

(iv) Soit $(x_n)_{n\in\mathbb{N}}$ une suite de réels positifs telle que $x_n \ll n^{1/\alpha}$, dans le sens où $\lim_{n\to\infty} x_n/n^{1/\alpha} = 0$. Montrer que l'on a $\lim_{n\to\infty} F_{M_n}(x_n) = 0$ et en déduire que

$$\lim_{n\to\infty}\mathrm{P}(S_n \le x_n) = 0.$$

[*Sugg.* Observer que les variables aléatoires X_i sont *positives*.]

On considère maintenant une suite $(x_n)_{n\geq 0}$ de réels positifs telle que $x_n \gg n^{1/\alpha}$, dans le sens où $\lim_{n\to\infty} x_n/n^{1/\alpha} = +\infty$.

(v) Observer que l'on peut écrire $S_n = S_n^{(1)} + S_n^{(2)}$, avec

$$S_n^{(1)} = \sum_{i=1}^{n} X_i \mathbb{1}_{\{X_i \geq \frac{1}{2}x_n\}}\,, \qquad S_n^{(2)} = \sum_{i=1}^{n} X_i \mathbb{1}_{\{X_i < \frac{1}{2}x_n\}}\,,$$

puis montrer que $\mathrm{P}(S_n \geq x_n) \leq \mathrm{P}\big(S_n^{(1)} \geq \frac{1}{2}x_n\big) + \mathrm{P}\big(S_n^{(2)} \geq \frac{1}{2}x_n\big)$.

(vi) Montrer l'égalité d'événements $\{S_n^{(1)} \geq \frac{1}{2}x_n\} = \bigcup_{i=1}^{n}\{X_i \geq \frac{1}{2}x_n\}$, puis en déduire

$$\mathrm{P}\big(S_n^{(1)} \geq \tfrac{1}{2}x_n\big) \leq n\,\mathrm{P}\big(X_1 \geq \tfrac{1}{2}x_n\big)\,.$$

Conclure que l'on a $\lim_{n\to\infty} \mathrm{P}\big(S_n^{(1)} \geq \frac{1}{2}x_n\big) = 0$.

(vii) Montrer que l'on a

$$\mathrm{P}\big(S_n^{(2)} \geq \tfrac{1}{2}x_n\big) \leq \frac{2n}{x_n}\,\mathrm{E}\big[X_1 \mathbb{1}_{\{X_1 < \frac{1}{2}x_n\}}\big]\,.$$

Calculer $\mathrm{E}\big(X_1 \mathbb{1}_{\{X_1 < \frac{1}{2}x_n\}}\big)$ et conclure que l'on a $\lim_{n\to\infty} \mathrm{P}\big(S_n^{(2)} \geq \frac{1}{2}x_n\big) = 0$.

(viii) Conclure que si $(x_n)_{n\geq 0}$ vérifie $\lim_{n\to\infty} x_n/n^{1/\alpha} = +\infty$ alors

$$\lim_{n\to\infty} \mathrm{P}(S_n \geq x_n) = 0\,.$$

Les questions (iv) et (viii) montrent qu'en un certain sens, S_n *prend des valeurs de l'ordre de* $n^{1/\alpha}$*, avec grande probabilité*. Un résultat (beaucoup) plus difficile montre que, de manière analogue au théorème central limite, la fonction de répartition de $\frac{1}{n^{1/\alpha}}S_n$ converge vers la fonction de répartition d'une variable aléatoire réelle (non normale), dont la loi est appelée *loi stable positive d'indice* α.

Solution 7.28. (i) Observons que f est positive et que

$$\int_{\mathbb{R}} f_X(x)\,\mathrm{d}x = \int_1^\infty \frac{\alpha}{x^{1+\alpha}}\,\mathrm{d}x = \Big[-\frac{1}{x^\alpha}\Big]_1^\infty = 1\,.$$

De plus

$$\mathrm{E}(X) = \int_{\mathbb{R}} x\,f_X(x)\,\mathrm{d}x = \int_1^\infty x\,\frac{\alpha}{x^{1+\alpha}}\,\mathrm{d}x = +\infty\,,$$

parce que, quand $x \to +\infty$, la fonction intégrée est asymptotiquement équivalente à $\frac{\alpha}{x^\alpha}$, qui n'est pas intégrable pour $\alpha < 1$.

(ii) On a $F_X(x) = \mathrm{P}(X \leq x) = 0$ si $x \leq 1$, alors que pour $x > 1$ on a

$$F_X(x) = \mathrm{P}(X \leq x) = \int_{-\infty}^{x} f_X(z)\,\mathrm{d}z = \int_1^x \frac{\alpha}{z^{1+\alpha}}\,\mathrm{d}z = \Big[-\frac{1}{z^\alpha}\Big]_1^x = 1 - \frac{1}{x^\alpha}\,.$$

(iii) On a $F_{M_n}(t) = 0$ si $t \leq 1$, alors que pour $t > 1$ on a

$$F_{M_n}(t) = \mathrm{P}(M_n \leq t) = \mathrm{P}(X_1 \leq t, \dots, X_n \leq t) = \mathrm{P}(X_1 \leq t)^n = F_{X_1}(t)^n = \Big(1 - \frac{1}{t^\alpha}\Big)^n\,.$$

(iv) Si $x_n \ll n^{1/\alpha}$, dans le sens où $\lim_{n\to\infty} x_n/n^{1/\alpha} = 0$, alors

$$\lim_{n\to\infty} F_{M_n}(x_n) = \lim_{n\to\infty}\left(1-\frac{1}{x_n^\alpha}\right)^n \leq \lim_{n\to\infty}\exp\left(-\frac{n}{x_n^\alpha}\right) = \lim_{n\to\infty}\exp\left(-\frac{1}{(\frac{x_n}{n^{1/\alpha}})^\alpha}\right) = 0\,,$$

où on a utilisé la majoration $1-t \leq \exp(-t)$, valable pour tout $t\in\mathbb{R}$.
Comme les variables aléatoires X_i sont positives, on a $S_n \leq M_n$ et on a donc l'inclusion d'événements $\{S_n \leq x_n\} \subseteq \{M_n \leq x_n\}$. On en déduit que pour $x_n \ll n^{1/\alpha}$ on a $\lim_{n\to\infty} \mathrm{P}(S_n \leq x_n) \leq \lim_{n\to\infty}\mathrm{P}(M_n \leq x_n) = 0$.

(v) Si $a\geq 0$ et $b \geq 0$ sont tels que $a+b\geq x$, on doit avoir $a\geq \frac12 x$ ou bien $b\geq\frac12 x$ (dans le cas contraire on aurait $a+b<x$). Comme $S_n = S_n^{(1)} + S_n^{(2)}$, on en déduit l'inclusion d'événements $\{S_n \geq x_n\} \subseteq \{S_n^{(1)} \geq \frac12 x_n\} \cup \{S_n^{(2)} \geq \frac12 x_n\}$, donc l'inégalité $\mathrm{P}(S_n\geq x_n) \leq \mathrm{P}(S_n^{(1)} \geq \frac12 x_n) + \mathrm{P}(S_n^{(2)} \geq \frac12 x_n)$.

(vi) La variable aléatoire $S_n^{(1)}$ est la somme de n variables aléatoires $X_i \mathbb{1}_{\{X_i \geq \frac12 x_n\}}$, qui prennent la valeur 0 ou bien une valeur $\geq \frac12 x_n$. Si on a $S_n^{(1)} \geq \frac12 x_n$, alors on doit avoir $X_i \geq \frac12 x_n$ au moins pour un $i=1,\ldots,n$ (sinon on aurait $S_n^{(1)}=0$) ; réciproquement, si $X_i\geq \frac12 x_n$ pour au moins un $i=1,\ldots,n$, alors clairement $S_n^{(1)} \geq \frac12 x_n$. Cela montre que l'on a l'égalité d'événements $\{S_n^{(1)} \geq \frac12 x_n\} = \bigcup_{i=1}^n \{X_i \geq \frac12 x_n\}$, donc par sous-additivité

$$\mathrm{P}(S_n^{(1)} \geq \tfrac12 x_n) \leq \sum_{i=1}^n \mathrm{P}(X_i \geq \tfrac12 x_n) = n\,\mathrm{P}(X_1 \geq \tfrac12 x_n)\,.$$

Enfin, en observant que

$$\mathrm{P}(X_1 \geq \tfrac12 x_n) = 1-\mathrm{P}(X_1 < \tfrac12 x_n) = 1 - F_{X_1}(\tfrac12 x_n) = \frac{1}{(\frac12 x_n)^\alpha}\,,$$

on obtient que, si $x_n \gg n^{1/\alpha}$ quand $n\to\infty$,

$$\mathrm{P}(S_n^{(1)} \geq \tfrac12 x_n) \leq n\,\mathrm{P}(X_1 \geq \tfrac12 x_n) = \frac{n}{(\frac12 x_n)^\alpha} = \frac{2^\alpha}{(\frac{x_n}{n^{1/\alpha}})^\alpha} \xrightarrow[n\to\infty]{} 0\,.$$

(vii) D'après l'inégalité de Markov, on obtient

$$\mathrm{P}\left(S_n^{(2)} \geq \tfrac12 x_n\right) \leq \frac{\mathrm{E}(S_n^{(2)})}{\frac12 x_n} = \frac{2n}{x_n}\,\mathbb{E}\left(X_1 \mathbb{1}_{\{X_1 < \frac12 x_n\}}\right)\,.$$

Calculons maintenant, grâce à la formule de transfert,

$$\begin{aligned}\mathbb{E}\left(X_1 \mathbb{1}_{\{X_1<\frac12 x_n\}}\right) &= \int_{\mathbb{R}} x\mathbb{1}_{\{x<\frac12 x_n\}} f_X(x)\,\mathrm{d}x = \int_1^{\frac12 x_n} \frac{\alpha}{x^\alpha}\,\mathrm{d}x = \frac{\alpha}{1-\alpha}\left[x^{1-\alpha}\right]_1^{\frac12 x_n} \\ &= \frac{\alpha}{1-\alpha}\left((\tfrac12 x_n)^{1-\alpha} - 1\right) \leq \frac{\alpha\,2^{\alpha-1}}{1-\alpha} x_n^{1-\alpha}\,.\end{aligned}$$

En appliquant l'inégalité obtenue juste avant, on obtient enfin

$$\mathrm{P}\left(S_n^{(2)} \geq \tfrac12 x_n\right) \leq \frac{2n}{x_n}\,\mathbb{E}\left(X_1 \mathbb{1}_{\{X_1<\frac12 x_n\}}\right) \leq \frac{\alpha\,2^\alpha}{1-\alpha}\,\frac{n}{x_n^\alpha} = \frac{\alpha\,2^\alpha}{1-\alpha}\,\frac{1}{(\frac{x_n}{n^{1/\alpha}})^\alpha} \xrightarrow[n\to\infty]{} 0\,.$$

(viii) Grâce à ce que l'on a montré dans les trois derniers points, on a

$$\lim_{n\to\infty} \mathrm{P}(S_n \geq x_n) \leq \lim_{n\to\infty} \mathrm{P}(S_n^{(1)} \geq \tfrac{1}{2}x_n) + \lim_{n\to\infty} \mathrm{P}(S_n^{(2)} \geq \tfrac{1}{2}x_n) = 0\,.$$

On en déduit que $\lim\limits_{n\to\infty} \mathrm{P}(S_n \geq x_n) = 0$.

Chapitre 8
Applications à la statistique

8.1 Modèles statistiques paramétriques

Exercice 8.1. Montrer que $Y_n := \overline{X}_n$ fournit un estimateur sans biais et efficace pour θ dans les modèles statistiques donnés par :

(i) $\mathrm{p}(x;\theta) := \mathrm{e}^{-\theta}\frac{\theta^x}{x!}$, avec $x \in \mathbb{N}$ et $\theta \in \Theta =]0,+\infty[$;

(ii) $f(x;\theta) := \frac{1}{\sqrt{2\pi\sigma^2}}\exp\big(-\frac{(x-\theta)^2}{2\sigma^2}\big)$, où $\theta \in \mathbb{R}$ et $\sigma > 0$ est une constante que l'on suppose connue.

Solution 8.1. (i) Si $\mathrm{p}(x;\theta)$ est la densité discrète des X_n, alors $X_n \sim \mathrm{Poi}(\theta)$. En particulier $\mathrm{E}_\theta(X_n) = \theta$, donc on a aussi $\mathrm{E}_\theta\left(\overline{X}_n\right) = \theta$. Ainsi, Y_n est un estimateur sans biais. Rappelons qu'un estimateur sans biais est efficace si l'inégalité (8.5) est en réalité une égalité. Dans le cas présent, $h(\theta) = \theta$ et

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\log\mathrm{p}(x;\theta) = \frac{x}{\theta} - 1,$$

dont on déduit que

$$\mathrm{E}_\theta\left[\Big(\frac{\mathrm{d}}{\mathrm{d}\theta}\log\mathrm{p}(X_1;\theta)\Big)^2\right] = \frac{1}{\theta^2}\,\mathrm{Var}_\theta(X_1) = \frac{1}{\theta}.$$

Ainsi $V_n(\theta) = \frac{\theta}{n}$, ce qui coïncide donc avec $\mathrm{Var}_\theta(Y_n)$.

(ii) Le fait que l'estimateur est sans biais est similaire au cas précédent, car $X_n \sim \mathrm{N}(\theta,\sigma^2)$. De plus,

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\log f(x;\theta) = \frac{x-\theta}{\sigma^2},$$

dont on déduit

$$\mathrm{E}_\theta\left[\left(\frac{\mathrm{d}}{\mathrm{d}\theta}\log f(X_1;\theta)\right)^2\right] = \frac{1}{\sigma^4}\,\mathrm{Var}_\theta(X_1) = \frac{1}{\sigma^2}.$$

Ainsi $V_n(\theta) = \frac{\sigma^2}{n}$, ce qui coïncide donc avec $\mathrm{Var}_\theta(Y_n)$.

Exercice 8.2. On considère le modèle statistique dans lequel

$$f(x;\theta) := \frac{1}{\sqrt{2\pi\theta}}\exp\left(-\frac{x^2}{2\theta}\right), \qquad \theta \in \Theta =]0,+\infty[\,.$$

On pose $Y_n := \frac{1}{n}\sum_{i=1}^n X_i^2$.

(i) Montrer que Y_n est un estimateur sans biais pour θ.

(ii) Montrer que $\mathrm{E}_\theta(X_i^4) = 3\theta^2$, puis que $\mathrm{Var}_\theta(Y_n) = \frac{2}{n}\theta^2$. En déduire que Y_n est un estimateur efficace pour θ.

Solution 8.2. (i) On a $X_i \sim \mathrm{N}(0,\theta)$, donc $\mathrm{E}_\theta(X_i^2) = \theta$. Par linéarité de l'espérance on en déduit que $\mathrm{E}_\theta(Y_n) = \theta$.

(ii) Avec la formule de transfert, on obtient

$$\begin{aligned}\mathrm{E}_\theta(X_i^4) &= \int_{-\infty}^{+\infty} \frac{x^4}{\sqrt{2\pi\theta}} \mathrm{e}^{-\frac{x^2}{2\theta}}\,\mathrm{d}x \\ &= \left[\frac{-x^3\theta}{\sqrt{2\pi\theta}}\mathrm{e}^{-\frac{x^2}{2\theta}}\,\mathrm{d}x\right]_{-\infty}^{+\infty} + 3\theta \int_{-\infty}^{+\infty} \frac{x^2}{\sqrt{2\pi\theta}}\mathrm{e}^{-\frac{x^2}{2\theta}}\,\mathrm{d}x \\ &= 0 + 3\theta\,\mathrm{E}_\theta(X_i^2) = 3\theta^2,\end{aligned}$$

où on a utilisé une intégration par parties pour la première égalité, puis le fait que $X_n \sim \mathrm{N}(0,\theta)$.

Comme les $(X_i)_{i\geq 1}$ sont indépendants, les $(X_i^2)_{i\geq 1}$ le sont aussi, donc on déduit des propriétés de la variance que

$$\mathrm{Var}_\theta(Y_n) = \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}_\theta(X_i^2) = \frac{1}{n}\Big(\mathrm{E}_\theta(X_i^4) - \mathrm{E}_\theta(X_i^2)^2\Big) = \frac{2}{n}\theta^2.$$

De plus,

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\log f(x;\theta) = -\frac{1}{2\theta} + \frac{x^2}{2\theta^2} = \frac{1}{2\theta^2}(x^2-\theta),$$

d'où, en utilisant le fait que $\mathrm{E}_\theta(X_1^2) = \theta$, $\mathrm{E}_\theta(X_1^4) = 3\theta^2$,

$$\mathrm{E}_\theta\left[\Big(\frac{\mathrm{d}}{\mathrm{d}\theta}\log f(X_1;\theta)\Big)^2\right] = \frac{1}{4\theta^4}(\mathrm{E}_\theta(X_1^4) - 2\theta\,\mathrm{E}_\theta(X_1^2) + \theta^2) = \frac{1}{2\theta^2}.$$

Ainsi $V_n(\theta) = \frac{2\theta^2}{n}$, ce qui coïncide avec $\mathrm{Var}_\theta(Y_n)$.

Exercice 8.3. On considère le modèle statistique dans lequel

$$f(x;\theta) := \frac{\theta}{x^{1+\theta}}\,\mathbb{1}_{]1,+\infty[}(x),$$

pour $\theta \in \Theta :=]1,+\infty[$. Montrer que $Y_n := \overline{X}_n$ est un estimateur sans biais, mais pas efficace, pour $h(\theta) = \frac{\theta}{\theta-1}$. Quel serait le problème si on avait $\Theta =]0,+\infty[$?

Solution 8.3. Tout d'abord, pour $\theta > 1$, on a

$$\mathrm{E}_\theta(X_i) = \int_1^{+\infty} \theta x^{-\theta}\mathrm{d}x = \left[\frac{\theta}{1-\theta}x^{1-\theta}\right]_1^{+\infty} = \frac{\theta}{\theta-1}.$$

Par linéarité de l'espérance, on a $\mathrm{E}_\theta(Y_n) = \frac{\theta}{\theta-1}$: il s'agit donc d'un estimateur sans biais pour $h(\theta)$. On peut déjà remarquer que pour $\theta \in]0,1]$ on a $\mathrm{E}_\theta(X_i) = +\infty$; en particulier $Y_n = \overline{X}_n$ ne peut pas être sans biais dans ce cas.

Pour vérifier l'efficacité de l'estimateur, il faut calculer $\mathrm{Var}_\theta(Y_n)$. Mais pour $\theta \in]1,2]$, on a $\mathrm{E}_\theta(X_i^2) = \int_1^{+\infty}\theta x^{1-\theta}\mathrm{d}x = +\infty$. Donc X_i n'admet pas de moment d'ordre 2 fini (ou une variance finie) et ce n'est pas non plus le cas pour Y_n. L'estimateur Y_n ne peut donc pas être efficace.

Vérifions tout de même que l'estimateur ne peut pas être efficace même si on se restreint à $\theta \in]2,+\infty[$. Pour $\theta > 2$, après quelques calculs on trouve $\mathrm{E}_\theta(X_i^2) = \frac{\theta}{\theta-2}$, donc $\mathrm{Var}_\theta(X_i) = \frac{\theta}{(\theta-2)(1-\theta)^2}$, donc par indépendance des X_i,

$$\mathrm{Var}_\theta(Y_n) = \frac{1}{n}\frac{\theta}{(\theta-2)(\theta-1)^2}\,.$$

(Si on veut, on peut poser $\mathrm{Var}_\theta(Y_n) = +\infty$ pour $\theta \in]1,2]$.) De plus, pour $x \in]1,+\infty[$,

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\log f(x;\theta) = \frac{1}{\theta} - \log x,$$

dont on déduit que (pour tout $\theta > 0$)

$$\mathrm{E}_\theta\left[\left(\frac{\mathrm{d}}{\mathrm{d}\theta}\log f(X_1;\theta)\right)^2\right] = \frac{1}{\theta^2} - \frac{2}{\theta}\mathrm{E}_\theta(\log X_1) + \mathrm{E}_\theta\left((\log X_1)^2\right) = \frac{1}{\theta^2}\,,$$

où on a utilisé que, grâce à une intégration par parties,

$$\begin{aligned}\mathrm{E}_\theta(\log X_1) &= \int_1^{+\infty}\theta x^{-(1+\theta)}\log x\mathrm{d}x = \Big[-x^{-\theta}\log x\Big]_1^{+\infty} + \int_1^{+\infty}x^{-(1+\theta)}\mathrm{d}x = \frac{1}{\theta}\,,\\ \mathrm{E}_\theta\left((\log X_1)^2\right) &= \int_1^{+\infty}\theta x^{-(1+\theta)}(\log x)^2\mathrm{d}x\\ &= \Big[-x^{-\theta}(\log x)^2\Big]_1^{+\infty} + 2\int_1^{+\infty}x^{-(1+\theta)}\log x\mathrm{d}x = \frac{2}{\theta}\mathrm{E}_\theta(\log X_1) = \frac{2}{\theta^2}\,.\end{aligned}$$

En notant que $h'(\theta) = \frac{-1}{(\theta-1)^2}$, on trouve finalement (pour tout $\theta > 1$) que $V_n(\theta) = \frac{1}{n}\frac{\theta^2}{(\theta-1)^4}$, qui est différent de $\mathrm{Var}_\theta(Y_n)$; on a $\frac{\mathrm{Var}_\theta(Y_n)}{V_n(\theta)} = \frac{(\theta-1)^2}{\theta(\theta-2)}$ pour $\theta > 2$.

Exercice 8.4. On considère le modèle statistique dans lequel

$$f(x;\theta) := \theta\,\mathrm{e}^{-\theta x}\,\mathbb{1}_{]0,+\infty[}(x)\,, \qquad \theta \in \Theta =]0,+\infty[\,.$$

(i) Montrer que $Y_n := \overline{X}_n$ est un estimateur sans biais et efficace pour $h(\theta) = 1/\theta$.

(ii) Montrer que $Y_n := \frac{n-1}{\sum_{i=1}^n X_i}$ est un estimateur sans biais, mais pas efficace, pour θ. Montrer en revanche qu'il est *asymptotiquement efficace*, dans le sens où $\lim_{n\to\infty}\frac{\mathrm{Var}_\theta(Y_n)}{V_n(\theta)} = 1$, où $V_n(\theta)$ est la borne inférieure de Cramér–Rao, qui apparaît dans (8.5).

Solution 8.4. (i) En notant que $X_1 \sim \mathrm{Exp}(\theta)$, on a $\mathrm{E}_\theta(X_1) = \frac{1}{\theta}$, dont on déduit immédiatement que Y_n est sans biais. De plus, comme $\mathrm{Var}_\theta(X_1) = \frac{1}{\theta^2}$, on a tout de suite

$$\mathrm{Var}_\theta(Y_n) = \frac{1}{n\theta^2}.$$

Calculons maintenant la borne inférieure de Cramér–Rao $V_n(\theta)$, voir (8.5). En observant que

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\log f(x;\theta) = \frac{1}{\theta} - x,$$

un calcul facile donne

$$\mathrm{E}_\theta\left[\left(\frac{\mathrm{d}}{\mathrm{d}\theta}\log f(X_1;\theta)\right)^2\right] = \mathrm{E}_\theta\left[\left(X - \frac{1}{\theta}\right)^2\right] = \mathrm{Var}_\theta(X_1) = \frac{1}{\theta^2}.$$

De plus, comme $h'(\theta) = -\frac{1}{\theta^2}$, on a

$$V_n(\theta) = \frac{1/\theta^4}{n/\theta^2} = \frac{1}{n\theta^2} = \mathrm{Var}_\theta(Y_n),$$

ce qui démontre l'efficacité de Y_n.

(ii) Notons que $\sum_{i=1}^n X_i \sim \mathrm{Gamma}(n,\theta)$ (voir la Proposition 6.36). Ainsi, grâce à la Proposition 6.20,

$$\begin{aligned}\mathrm{E}_\theta(Y_n) &= (n-1)\,\mathrm{E}_\theta\left(\frac{1}{\sum_{i=1}^n X_i}\right) = (n-1)\int_0^{+\infty}\frac{1}{y}\frac{\theta^n}{(n-1)!}y^{n-1}\mathrm{e}^{-\theta y}dy \\ &= \frac{\theta^n}{(n-2)!}\int_0^{+\infty} y^{n-2}\mathrm{e}^{-\theta y}dy = \frac{\theta^n}{(n-2)!}\frac{(n-2)!}{\theta^{n-1}} = \theta,\end{aligned}$$

où on a utilisé l'identité

$$\Gamma(\alpha) = \int_0^{+\infty}\lambda^\alpha t^{\alpha-1}\,\mathrm{e}^{-\lambda t}\,\mathrm{d}t\,.$$

Donc Y_n est sans biais. De manière analogue, on calcule

$$\begin{aligned}\mathrm{E}_\theta(Y_n^2) &= (n-1)^2\,\mathrm{E}_\theta\left[\left(\frac{1}{\sum_{i=1}^n X_i}\right)^2\right] = (n-1)^2\int_0^{+\infty}\frac{1}{y^2}\frac{\theta^n}{(n-1)!}y^{n-1}\mathrm{e}^{-\theta y}dy \\ &= \frac{(n-1)\theta^n}{(n-2)!}\int_0^{+\infty} y^{n-3}\mathrm{e}^{-\theta y}dy = \frac{(n-1)\theta^n}{(n-2)!}\frac{(n-3)!}{\theta^{n-2}} = \frac{(n-1)\theta^2}{n-2},\end{aligned}$$

dont on déduit

$$\mathrm{Var}_\theta(Y_n) = \frac{\theta^2}{n-2}.$$

De plus,

$$V_n(\theta) = \frac{1}{n\,\mathrm{E}_\theta\left[\left(\frac{\mathrm{d}}{\mathrm{d}\theta}\log f(X_1;\theta)\right)^2\right]} = \frac{1}{n\,\mathrm{E}_\theta\left[\left(X-\frac{1}{\theta}\right)^2\right]} = \frac{\theta^2}{n}.$$

Ainsi, on a $\mathrm{Var}_\theta(Y_n) > V_n(\theta)$, ce qui montre que Y_n n'est pas un estimateur efficace pour θ.

D'autre part, avec les calculs que l'on vient de faire, on trouve que

$$\frac{\mathrm{Var}_\theta(Y_n)}{V_n(\theta)} = \frac{n}{n-2} \longrightarrow 1\,, \qquad \text{quand } n \to +\infty\,,$$

ce qui montre que l'estimateur Y_n est asymptotiquement efficace.

8.2 Intervalles de confiance

Exercice 8.5. On considère le modèle statistique des Exemples 8.3 et 8.6, c'est-à-dire où $\mathrm{p}(x;\theta)$ est la densité d'une variable de Bernoulli de paramètre $\theta \in \Theta =]0,1[$. Montrer que l'estimateur $Y_n := \overline{X}_n$ est asymptotiquement normal, et déterminer un intervalle de confiance asymptotique pour $h(\theta) = \theta$, de niveau de confiance $1-\alpha$.

Solution 8.5. D'après le théorème central limite, $\sqrt{n}\frac{Y_n-\theta}{\sqrt{\theta(1-\theta)}} \xrightarrow{\mathcal{L}} \mathrm{N}(0,1)$. Ainsi, la condition (8.19) est vérifiée avec $\sigma(\theta) := \sqrt{\theta(1-\theta)} \le \overline{\sigma} = 1/2$, pour tout $\theta \in \Theta$; on a noté que $x \mapsto x(1-x)$ atteint son maximum en $x = \frac{1}{2}$. On a donc la normalité asymptotique de Y_n, donc grâce à (8.21), un intervalle de confiance asymptotique pour θ, de niveau de confiance $1-\alpha$, est donné par

$$\left[Y_n - \frac{1}{2\sqrt{n}} z_{\alpha/2}, Y_n + \frac{1}{2\sqrt{n}} z_{\alpha/2}\right].$$

Exercice 8.6. On considère le modèle statistique des Exemples 8.23 et 8.26 : la densité $f(x;\theta)$ est donnée par

$$f(x;\theta) = \frac{1}{\theta} \mathbb{1}_{]0,\theta[}(x), \qquad \theta \in \Theta =]0,+\infty[\,.$$

Soit $Y_n = \max(X_1,\dots,X_n)$ l'estimateur pour θ introduit dans l'Exemple 8.23.

(i) Soit $\theta \in \Theta$. Montrer que pour tout $x > \theta$,

$$\mathrm{P}_\theta\left(Y_n \in\ [x^{-1}\theta,\theta]\right) = 1 - x^{-n}.$$

(ii) En déduire que pour tout $\alpha \in]0,1[$ l'intervalle $[Y_n, Y_n + \frac{Y_n}{n} g_n(\alpha)]$, où on a posé $g_n(\alpha) := n(\alpha^{-1/n} - 1)$, est un intervalle de confiance pour θ de niveau $1-\alpha$.

Solution 8.6. (i) Comme $f(x;\theta)$ est la densité des X_n, on a $X_n \sim \mathrm{U}(0,\theta)$. Donc $Y_n \in]0,\theta[$ et pour tout $x > \theta$ (de sorte que $x^{-1}\theta \in]0,\theta[$) on a

$$\begin{aligned}\mathrm{P}_\theta\left(Y_n \in [x^{-1}\theta,\theta]\right) &= 1 - \mathrm{P}_\theta(Y_n < x^{-1}\theta) = 1 - \mathrm{P}_\theta(X_i < x^{-1}\theta\,, \forall 1 \le i \le n)\\ &= 1 - \mathrm{P}_\theta\left(X_1 < x^{-1}\theta\right)^n = 1 - x^{-n},\end{aligned}$$

où on a utilisé le fait que $\mathrm{P}_\theta(X_1 < t) = \frac{1}{\theta} t$ pour tout $t \in]0,\theta[$.

(ii) Notons que l'on a $Y_n \in [x^{-1}\theta,\theta]$ si et seulement si $\theta \in [Y_n, xY_n]$. Avec la question précédente, on obtient

$$\mathrm{P}_\theta\Big(\theta \in [Y_n, xY_n]\Big) = 1 - x^{-n}.$$

En prenant $x = \alpha^{-1/n}$ et utilisant la notation $g_n(\alpha) = n(\alpha^{-1/n} - 1)$, on obtient

$$\mathrm{P}_\theta\left(\theta \in \left[Y_n, Y_n + \frac{Y_n}{n} g_n(\alpha)\right]\right) = \mathrm{P}_\theta\left(\theta \in [Y_n, \alpha^{-1/n} Y_n]\right) = 1 - \alpha,$$

ce qui montre que $[Y_n, Y_n + \frac{Y_n}{n} g_n(\alpha)]$ est un intervalle de confiance pour θ de niveau de confiance $1-\alpha$.

En ce qui concerne la limite $\lim_{n\to+\infty} g_n(\alpha)$, on obtient avec un développement limité que pour tout $\alpha \in]0,1[$,

$$g_n(\alpha) = n\left(\mathrm{e}^{\frac{1}{n}\log\frac{1}{\alpha}} - 1\right) = (1 + o(1))\, n \times \frac{1}{n} \log\frac{1}{\alpha} \longrightarrow \log\frac{1}{\alpha},$$

quand $n \to +\infty$.

Exercice 8.7. On considère le modèle statistique de l'Exercice 8.4 : on a

$$f(x;\theta) = \theta \mathrm{e}^{-\theta x} \mathbb{1}_{]0,+\infty[}(x), \qquad \theta \in \Theta =]1,+\infty[.$$

On pose $Y_n = \overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$.

(i) Montrer que Y_n est un estimateur asymptotiquement normal pour $1/\theta$, de variance asymptotique $\sigma^2(\theta) = 1/\theta^2$.

(ii) En déduire un intervalle de confiance asymptotique de niveau de confiance $1-\alpha$ pour $1/\theta$, puis pour θ.

Solution 8.7. (i) Les variables aléatoires X_i sont i.i.d. de loi $\mathrm{Exp}(\theta)$, donc en particulier elles sont d'espérance $\mathrm{E}_\theta(X_i) = 1/\theta$ et de variance $1/\theta^2$. D'après le théorème central limite, $\sqrt{n}\frac{Y_n - 1/\theta}{1/\theta} \xrightarrow{\mathcal{L}} \mathrm{N}(0,1)$, donc Y_n est un estimateur asymptotiquement normal pour $1/\theta$, de variance asymptotique $\sigma^2(\theta) = 1/\theta^2$.

(ii) Vu que $\sigma(\theta) := 1/\theta \leq \overline{\sigma} = 1$ pour tout $\theta \in \Theta =]1,+\infty[$, on obtient de (8.21) qu'un intervalle de confiance asymptotique pour $1/\theta$ de niveau de confiance $1-\alpha$ est donné par

$$\left[Y_n - \frac{1}{\sqrt{n}} z_{\alpha/2}, Y_n + \frac{1}{\sqrt{n}} z_{\alpha/2}\right].$$

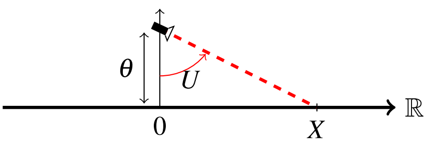
Vu que $1/\theta > 0$, on a en réalité l'intervalle de confiance suivant :

$$\left[\left(Y_n - \frac{1}{\sqrt{n}} z_{\alpha/2},\right)^+, Y_n + \frac{1}{\sqrt{n}} z_{\alpha/2}\right],$$

où on rappelle que $x^+ = \max(x,0)$ désigne la partie positive. Maintenant, notons que pour $0 \leq a < b$, on a $\theta \in [a,b]$ si et seulement si $\frac{1}{\theta} \in [\frac{1}{b}, \frac{1}{a}]$; par convention, il s'agit de l'intervalle $[\frac{1}{b}, +\infty[$ si $a = 0$. On obtient donc un intervalle de confiance asymptotique pour θ de niveau de confiance $1-\alpha$:

$$\left[\frac{1}{Y_n + \frac{1}{\sqrt{n}} z_{\alpha/2}}, \frac{1}{\left(Y_n - \frac{1}{\sqrt{n}} z_{\alpha/2}\right)^+}\right].$$

Exercice 8.8. Quelqu'un a suspendu un laser à une hauteur inconnue $\theta \in \Theta =]0,+\infty[$. L'angle que le laser forme avec la verticale est aléatoire, noté U, de loi $\mathrm{U}(-\frac{\pi}{2}, \frac{\pi}{2})$. On note X le point marqué au sol par le laser (voir la figure ci-dessous). On observe un échantillon de points $(X_n)_{n\geq 1}$ correspondant à une succession d'angles $(U_n)_{n\geq 1}$ indépendants, de même loi $\mathrm{U}(-\frac{\pi}{2}, \frac{\pi}{2})$, et on essaie d'estimer la hauteur θ à laquelle est accrochée le laser.



(i) Montrer que, sous P_θ, c'est-à-dire si le laser est accroché à une hauteur θ, alors les X_i ont pour densité $f(x;\theta) = \frac{\theta}{\pi(\theta^2 + x^2)}$.

(ii) Montrer que $\mathrm{E}_\theta(|X_i|) = +\infty$, et justifier que $\overline{X}_n := \frac{1}{n}\sum_{i=1}^n X_i$ ne permet pas d'estimer θ.

(iii) Montrer que $\mathrm{P}_\theta(|X_i| \leq 1) = \frac{2}{\pi}\arctan(\frac{1}{\theta})$. En déduire que $Y_n := \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{|X_i|\leq 1\}}$ est un estimateur sans biais et consistant de $\frac{2}{\pi}\arctan(\frac{1}{\theta})$.

(iv) Montrer que Y_n est un estimateur asymptotiquement normal pour $\frac{2}{\pi}\arctan(\frac{1}{\theta})$. En déduire un intervalle de confiance pour $\frac{2}{\pi}\arctan(\frac{1}{\theta})$, puis pour θ.

Solution 8.8. (i) Si le laser est suspendu à une hauteur θ, alors on a $X_i = \theta \tan U_i$. On peut donc calculer la densité de X_i par rapport à P_θ, de la même manière que dans l'Exercice 6.22. Calculons la fonction de répartition

$$\begin{aligned} F_\theta(x) &= \mathrm{P}_\theta(X \le x) = \mathrm{P}(\theta \tan U_i \le x) \\ &= \mathrm{P}(U_i \le \arctan(x/\theta)) = \frac{1}{\pi}\Big(\frac{\pi}{2} + \arctan(x/\theta)\Big)\,. \end{aligned}$$

Dans la dernière égalité, on a utilisé le fait que U_i est une variable aléatoire uniforme continue sur $]-\frac{\pi}{2},\frac{\pi}{2}[$, de sorte que pour $t \in]-\frac{\pi}{2},\frac{\pi}{2}[$ on a $\mathrm{P}(\Theta \le t) = \frac{1}{\pi}(\frac{\pi}{2}+t)$, voir (6.29). La fonction F_θ est continue et $\mathcal{C}^1$ par morceaux, donc grâce à la Proposition 6.17, la variable aléatoire X a pour densité

$$f(x;\theta) = F'_\theta(x) = \frac{1}{\pi}\,\frac{1}{\theta}\,\frac{1}{1+(x/\theta)^2} = \frac{1}{\pi}\,\frac{\theta}{\theta^2+x^2}\,.$$

Il s'agit donc d'une variable aléatoire $\mathrm{Cauchy}(0,\theta)$, voir la Section 6.3.6.

(ii) On sait qu'une variable aléatoire de Cauchy n'admet pas d'espérance finie. Plus précisément, pour tout $\theta > 0$

$$\mathrm{E}_\theta(|X_i|) = \frac{\theta}{\pi}\int_{-\infty}^{\infty} \frac{|x|}{\theta^2+x^2}\mathrm{d}x = \frac{2\theta}{\pi}\int_0^{\infty}\frac{x}{\theta^2+x^2}\mathrm{d}x$$

par symétrie. De plus, $\frac{x}{\theta^2+x^2} \sim \frac{1}{x}$ quand $x \to +\infty$ et n'est donc pas intégrable. On a donc $\mathrm{E}_\theta(|X_i|) = +\infty$.

Premièrement, $\frac{1}{n}\sum_{i=1}^n X_i$ n'est pas un estimateur sans biais, parce qu'il n'admet pas une espérance finie. Deuxièmement, comme on ne peut pas appliquer la loi des grands nombres à $(X_i)_{i\ge1}$, l'estimateur $\frac{1}{n}\sum_{i=1}^n X_i$ ne peut pas être consistant (voir aussi l'Observation 7.9).

(iii) Comme arctan est une fonction strictement croissante et impaire, on a

$$\mathrm{P}_\theta(|X_i| \le 1) = \mathrm{P}(\theta\tan U_i \in [-1,1]) = \mathrm{P}\Big(U_i \in \big[-\arctan(\tfrac{1}{\theta}), \arctan(\tfrac{1}{\theta})\big]\Big)\,.$$

En utilisant le fait que $U_i \sim \mathrm{U}(-\frac{\pi}{2},\frac{\pi}{2})$, on en déduit que

$$\mathrm{P}_\theta(|X_i| \le 1) = \frac{1}{\pi}\,2\arctan(\frac{1}{\theta})\,.$$

Par linéarité de l'espérance, on obtient $\mathrm{E}_\theta(Y_n) = \mathrm{E}_\theta(\mathbb{1}_{\{|X_1|\le1\}}) = \mathrm{P}_\theta(|X_1| \le 1) = \frac{2}{\pi}\arctan(\frac{1}{\theta})$, donc Y_n est un estimateur sans biais pour $\frac{2}{\pi}\arctan(\frac{1}{\theta})$.

(iv) Notons que $(\mathbb{1}_{\{|X_i|\le1\}})_{i\ge1}$ sont des variables aléatoires i.i.d. de loi $\mathrm{Bern}(p_\theta)$ avec $p_\theta := \frac{2}{\pi}\arctan(\frac{1}{\theta})$. D'après le théorème central limite, on a $\sqrt{n}\frac{Y_n - p_\theta}{\sqrt{p_\theta(1-p_\theta)}} \xrightarrow{\mathcal{L}} \mathrm{N}(0,1)$. Cela montre que Y_n est un estimateur asymptotiquement normal pour $p_\theta = \frac{2}{\pi}\arctan(\frac{1}{\theta})$, de variance asymptotique $\sigma^2(\theta) = p_\theta(1-p_\theta)$ qui vérifie $\sigma(\theta) \le \overline{\sigma} = 1/2$ pour tout $\theta \in \Theta$. Grâce à (8.21), un intervalle de confiance asymptotique pour $p_\theta := \frac{2}{\pi}\arctan(\frac{1}{\theta})$ de niveau de confiance $1-\alpha$ est donné par

$$\left[Y_n - \frac{1}{2\sqrt{n}} z_{\alpha/2}, Y_n + \frac{1}{2\sqrt{n}} z_{\alpha/2}\right].$$

Vu que $p_\theta \in [0,1]$, on obtient en réalité l'intervalle de confiance suivant :

$$\left[\max\left(Y_n - \frac{1}{2\sqrt{n}} z_{\alpha/2}, 0\right), \min\left(Y_n + \frac{1}{2\sqrt{n}} z_{\alpha/2}, 1\right)\right].$$

Maintenant, pour tous $0 \le a < b \le 1$, en utilisant que arctan est strictement croissante, on obtient que $p_\theta = \frac{2}{\pi}\arctan(\frac{1}{\theta}) \in [a,b]$ si et seulement si $\frac{1}{\theta} \in [\tan(\frac{\pi}{2}a), \tan(\frac{\pi}{2}b)]$, c'est-à-dire si et seulement si $\theta \in [\frac{1}{\tan(\frac{\pi}{2}b)}, \frac{1}{\tan(\frac{\pi}{2}a)}]$, où par convention la borne inférieure de l'intervalle vaut 0 si $b = 1$ et la borne supérieure de l'intervalle vaut $+\infty$ si $a = 0$. On obtient finalement l'intervalle de confiance asymptotique suivant pour θ, de niveau de confiance $1-\alpha$

$$\left[\tan\left(\frac{\pi}{2}\max\left(Y_n - \frac{1}{2\sqrt{n}} z_{\alpha/2}, 0\right)\right)^{-1}, \tan\left(\frac{\pi}{2}\min\left(Y_n + \frac{1}{2\sqrt{n}} z_{\alpha/2}, 1\right)\right)^{-1}\right].$$

Exercice 8.9. Deux géologues, Arthur et Luca, étudient l'âge des dépôts sédimentaires d'une zone géographique donnée, et souhaitent notamment estimer l'âge du dépôt le plus ancien et l'âge du plus récent. Pour cela, ils effectuent des prélèvements de roche de manière aléatoire, qu'ils sont capables de dater : on note $X_1, \ldots, X_n$ les âges des n morceaux de roche qu'ils ont prélevés. Ils considèrent le modèle statistique dans lequel

$$f(x;\theta) = \frac{1}{b-a}\mathbb{1}_{]a,b[}(x), \qquad \theta = (a,b) \in \Theta = \{(x,y) \colon 0 < x < y < +\infty\}.$$

Autrement dit, ils supposent que les âges sont répartis uniformément dans l'intervalle $]a,b[$, où a et b sont deux paramètres à déterminer. Arthur propose les estimateurs $Y_n = \min(X_1, \ldots, X_n)$ et $W_n = \max(X_1, \ldots, X_n)$ pour a et b respectivement.

(i) Montrer que Y_n et W_n sont des estimateurs consistants pour a et b, mais qu'ils ne sont pas sans biais. Montrer qu'ils sont en revanche asymptotiquement sans biais.

(ii) Luca propose de considérer $\tilde{Y}_n = Y_n - \frac{1}{n-1}(W_n - Y_n)$ et $\tilde{W}_n = W_n + \frac{1}{n-1}(W_n - Y_n)$. Montrer que $\tilde{Y}_n$ et $\tilde{W}_n$ sont des estimateurs sans biais et consistants pour a et b.

(iii) Kristel arrive en renfort, et fait les deux remarques suivantes : pour tout $t > 0$, et pour tout $\varepsilon > 0$

$$\lim_{n\to+\infty} \mathrm{P}_\theta\left(Y_n - a > \frac{t}{n}(b-a)\right) = \mathrm{e}^{-t}, \quad \lim_{n\to+\infty} \mathrm{P}_\theta\left(|(W_n - Y_n) - (b-a)| > \varepsilon\right) = 0.$$

Montrer que Kristel a raison et en déduire que, pour tout $t > 0$,

$$\lim_{n\to+\infty} \mathrm{P}_\theta\left(Y_n - a > \frac{t}{n}(W_n - Y_n)\right) = \mathrm{e}^{-t}.$$

Donner un intervalle de confiance asymptotique pour a, de niveau $1-\alpha$.

Solution 8.9. (i) Comme $X_i \sim \mathrm{U}(a,b)$ par rapport à P_θ, on a $\mathrm{P}_\theta(Y_n > a) = 1$. Ainsi, pour tout $\varepsilon > 0$ on a

$$\begin{aligned}\mathrm{P}_\theta(|Y_n - a| > \varepsilon) &= \mathrm{P}_\theta(Y_n > a + \varepsilon) \\ &= \mathrm{P}_\theta(X_i > a + \varepsilon, \forall 1 \le i \le n) = \mathrm{P}_\theta(X_1 > a + \varepsilon)^n,\end{aligned}$$

où on a utilisé le fait que les $(X_i)_{i\geq 1}$ sont indépendants pour la dernière égalité. Pour $\varepsilon \in]0, b-a[$ on a $\mathrm{P}_\theta(X_1 > a+\varepsilon) = 1 - \frac{\varepsilon}{b-a}$; si $\varepsilon \geq b-a$ la probabilité est nulle. Ainsi, pour $\varepsilon \in]0, b-a[$ on a

$$\mathrm{P}_\theta(|Y_n - a| > \varepsilon) = \Big(1 - \frac{\varepsilon}{b-a}\Big)^n \xrightarrow{n\to\infty} 0\,,$$

ce qui démontre que $(Y_n)_{n\geq 1}$ est un estimateur consistant pour a. De la même manière, pour $\varepsilon \in]0, b-a[$ on a

$$\mathrm{P}_\theta(|W_n - b| > \varepsilon) = \Big(1 - \frac{\varepsilon}{b-a}\Big)^n \xrightarrow{n\to\infty} 0\,,$$

ce qui démontre que $(W_n)_{n\geq 1}$ est un estimateur consistant pour b.
On peut de plus déterminer les espérances $\mathrm{E}_\theta(Y_n)$ et $\mathrm{E}_\theta(W_n)$. Calculons la fonction de répartition de $Y_n - a$: on sait que $Y_n - a \in]0, b-a[$ et avec les mêmes calculs que ci-dessus, pour $t \in]0, b-a[$

$$\mathrm{P}_\theta(Y_n - a \leq t) = 1 - \mathrm{P}_\theta(Y_n > a+t) = 1 - \Big(1 - \frac{t}{b-a}\Big)^n .$$

La fonction de répartition de $Y_n - a$ est continue et $\mathcal{C}^1$ par morceaux, donc en dérivant on trouve que $Y_n - a$ a pour densité $\frac{n}{(b-a)}(1 - \frac{y}{b-a})^{n-1}\mathbb{1}_{]0,b-a[}(y)$. On obtient ainsi

$$\begin{aligned}
\mathrm{E}_\theta(Y_n - a) &= \int_0^{b-a} y\,\frac{n}{b-a}\Big(1 - \frac{y}{b-a}\Big)^{n-1}\mathrm{d}y \\
&= (b-a)\int_0^1 nu(1-u)^{n-1}\mathrm{d}u \\
&= (b-a)\int_0^1 u^n\mathrm{d}u = \frac{b-a}{n+1}\,,
\end{aligned}$$

où on a fait un changement de variable $u = y/(b-a)$ puis utilisé une intégration par parties. Par linéarité de l'espérance, on en conclut que $\mathrm{E}_\theta(Y_n) = a + \frac{1}{n+1}(b-a) \neq a$. Ainsi $(Y_n)_{n\geq 1}$ n'est pas un estimateur sans biais pour a, mais comme $\lim_{n\to+\infty}\mathrm{E}_\theta(Y_n) = a$, il est asymptotiquement sans biais.
De la même manière, $b - W_n$ a pour densité $\frac{n}{(b-a)}(1 - \frac{w}{b-a})^{n-1}\mathbb{1}_{]0,b-a[}(w)$, dont on peut déduire que $\mathrm{E}_\theta(W_n) = b - \frac{1}{n+1}(b-a) \neq b$. Donc $(W_n)_{n\geq 1}$ n'est pas un estimateur sans biais pour b, mais comme $\lim_{n\to+\infty}\mathrm{E}_\theta(W_n) = b$, il est asymptotiquement sans biais.

(ii) Par linéarité de l'espérance, on a

$$\mathrm{E}_\theta(\tilde{Y}_n) = \mathrm{E}_\theta(Y_n) - \frac{1}{n-1}(\mathrm{E}_\theta(W_n) - \mathrm{E}_\theta(Y_n))\,.$$

Avec les calculs du point précédent, on a

$$\mathrm{E}_\theta(Y_n) = a + \frac{1}{n+1}(b-a) \qquad \text{et} \qquad \mathrm{E}_\theta(W_n) = b - \frac{1}{n+1}(b-a)\,,$$

donc

$$\mathrm{E}_\theta(Y_n) - \mathrm{E}_\theta(W_n) = b - a - \frac{2}{n+1}(b-a) = \frac{n-1}{n+1}(b-a)\,.$$

Cela montre que $\mathrm{E}_\theta(\tilde{Y}_n) = a$. Le fait que $\mathrm{E}_\theta(\tilde{W}_n) = b$ est complètement analogue.

On peut noter que $0 \leq \frac{1}{n-1}(W_n - Y_n) \leq \frac{b-a}{n-1}$. Ainsi, pour tout $\varepsilon > 0$, il existe n_0 suffisamment grand tel que pour $n \geq n_0$ on ait $0 \leq Y_n - \tilde{Y}_n = \frac{1}{n-1}(W_n - Y_n) \leq \varepsilon/2$ et donc $|\tilde{Y}_n - a| \leq |Y_n - a| + \varepsilon/2$. Pour $\varepsilon > 0$, on obtient que pour $n \geq n_0$

$$\mathrm{P}_\theta\left(|\tilde{Y}_n - a| > \varepsilon\right) \leq \mathrm{P}_\theta\left(|Y_n - a| + \varepsilon/2 > \varepsilon\right) = \mathrm{P}_\theta\left(|Y_n - a| > \varepsilon/2\right).$$

Cette dernière probabilité tend vers 0 quand $n \to +\infty$, comme on l'a vu précédemment. Cela démontre que $(\tilde{Y}_n)_{n\geq 2}$ est un estimateur consistant pour a ; de manière complètement analogue, $(\tilde{W}_n)_{n\geq 2}$ est un estimateur consistant pour b.

(iii) Avec les calculs précédents, on obtient que pour tout $t > 0$

$$\begin{aligned}
\mathrm{P}_\theta\left(Y_n > a + \frac{t}{n}(b-a)\right) &= \mathrm{P}_\theta\left(X_i > a + \frac{t}{n}(b-a), \forall 1 \leq i \leq n\right) \\
&= \begin{cases} 0 & \text{si } \frac{t}{n} \geq 1\,, \\ \left(1 - \frac{t}{n}\right)^n & \text{si } \frac{t}{n} < 1\,. \end{cases}
\end{aligned}$$

Comme pour tout $t > 0$ fixé on a $\frac{t}{n} < 1$ pour n suffisamment grand, on obtient

$$\lim_{n\to\infty} \mathrm{P}_\theta\left(Y_n - a > \frac{t}{n}(b-a)\right) = \lim_{n\to\infty}\left(1 - \frac{t}{n}\right)^n = \mathrm{e}^{-t}.$$

De plus, en utilisant que $|(W_n - Y_n) - (b-a)| \leq |Y_n - a| + |W_n - b|$, on obtient

$$\begin{aligned}
\mathrm{P}_\theta\left(|(W_n - Y_n) - (b-a)| > \varepsilon\right) &\leq \mathrm{P}_\theta\left(|Y_n - a| + |W_n - b| > \varepsilon\right) \\
&\leq \mathrm{P}_\theta\left(\{|Y_n - a| > \varepsilon/2\} \cup \{|W_n - b| > \varepsilon/2\}\right) \\
&\leq \mathrm{P}_\theta\left(|Y_n - a| > \varepsilon/2\right) + \mathrm{P}_\theta\left(|W_n - b| > \varepsilon/2\right),
\end{aligned}$$

où on a utilisé le fait que si $|Y_n - a| + |W_n - b| > \varepsilon$ alors au moins un des deux termes $|Y_n - a|$ ou $|W_n - b|$ doit être supérieur à $\varepsilon/2$. Grâce à la question (i), pour tout $\varepsilon > 0$ les deux probabilités tendent vers 0 quand $n \to +\infty$, ce qui démontre que

$$\lim_{n\to+\infty} \mathrm{P}_\theta\left(|(W_n - Y_n) - (b-a)| > \varepsilon\right) = 0.$$

Posons $Z_n := (W_n - Y_n) - (b-a)$, de sorte que l'on doit estimer la probabilité

$$\mathrm{P}_\theta\left(Y_n - a > \frac{t}{n}(W_n - Y_n)\right) = \mathrm{P}_\theta\left(Y_n - a > \frac{t}{n}(b - a + Z_n)\right).$$

Fixons $\varepsilon > 0$ (quelconque). On a

$$\begin{aligned}
\mathrm{P}_\theta\left(Y_n - a > \frac{t}{n}(b-a+Z_n)\right) &\geq \mathrm{P}_\theta\left(Y_n - a > \frac{t}{n}(b-a+Z_n), Z_n \leq \varepsilon(b-a)\right) \\
&\geq \mathrm{P}_\theta\left(Y_n - a > (1+\varepsilon)\frac{t}{n}(b-a), Z_n \leq \varepsilon(b-a)\right) \\
&\geq \mathrm{P}_\theta\left(Y_n - a > (1+\varepsilon)\frac{t}{n}(b-a)\right) - \mathrm{P}_\theta\left(Z_n > \varepsilon(b-a)\right),
\end{aligned}$$

où on a utilisé dans la deuxième inégalité que $\{Y_n - a > x\} \supseteq \{Y_n - a > x'\}$ pour $x \leq x'$, alors que pour la dernière inégalité on a observé que pour toute paire d'événements A, B on a $\mathrm{P}(A \cap B) = \mathrm{P}(A) - \mathrm{P}(A \cap B^c) \geq \mathrm{P}(A) - \mathrm{P}(B^c)$. En prenant la limite $n \to +\infty$, et en utilisant les résultats que l'on vient de démontrer, on obtient que

$$\liminf_{n\to+\infty} \mathrm{P}_\theta\left(Y_n - a > \frac{t}{n}(b-a+Z_n)\right) \geq \mathrm{e}^{-(1+\varepsilon)t} .$$

D'autre part, on a que

$$\begin{aligned}
&\mathrm{P}_\theta\Big(Y_n - a > \frac{t}{n}(b-a+Z_n)\Big) \\
&\quad \leq \mathrm{P}_\theta\left(Y_n - a > \frac{t}{n}(b-a+Z_n), Z_n \geq -\varepsilon(b-a)\right) + \mathrm{P}_\theta(Z_n < -\varepsilon(b-a)) \\
&\quad \leq \mathrm{P}_\theta\left(Y_n - a > (1-\varepsilon)\frac{t}{n}(b-a)\right) + \mathrm{P}_\theta(Z_n < -\varepsilon(b-a)) .
\end{aligned}$$

En prenant la limite quand $n \to +\infty$ et en utilisant les résultats que l'on vient de démontrer (notamment que $\mathrm{P}_\theta(Z_n < -\varepsilon(b-a)) \leq \mathrm{P}_\theta(|Z_n| > \varepsilon(b-a)) \to 0$), on obtient

$$\limsup_{n\to+\infty} \mathrm{P}_\theta\left(Y_n - a > \frac{t}{n}(b-a+Z_n)\right) \leq \mathrm{e}^{-(1-\varepsilon)t} .$$

Vu que $\varepsilon > 0$ a été choisi arbitrairement, on peut rendre $\mathrm{e}^{-(1+\varepsilon)t}$ et $\mathrm{e}^{-(1-\varepsilon)t}$ arbitrairement proche de e^{-t}, ce qui démontre que

$$\lim_{n\to\infty} \mathrm{P}_\theta\left(Y_n - a > \frac{t}{n}(W_n - Y_n)\right) = \mathrm{e}^{-t} .$$

Vu que $\mathrm{P}_\theta(Y_n > a) = 1$, en passant au complémentaire dans le résultat que l'on vient de démontrer, on obtient

$$\lim_{n\to\infty} \mathrm{P}_\theta\left(Y_n \in \left[a, a + \frac{t}{n}(W_n - Y_n)\right]\right) = 1 - \mathrm{e}^{-t} ,$$

ou encore

$$\lim_{n\to\infty} \mathrm{P}_\theta\left(a \in \left[Y_n - \frac{t}{n}(W_n - Y_n), Y_n\right]\right) = 1 - \mathrm{e}^{-t} ,$$

En prenant $t = \log \frac{1}{\alpha}$ de sorte que $1 - \mathrm{e}^{-t} = 1 - \alpha$, on a donc construit un intervalle de confiance asymptotique pour a, de niveau de confiance $1 - \alpha$, donné par

$$\left[Y_n - (\log \tfrac{1}{\alpha}) \frac{W_n - Y_n}{n}, Y_n\right] .$$

Même si cela n'était pas demandé, un intervalle de confiance asymptotique pour b, de niveau de confiance $1 - \alpha$, est donné de manière analogue par $[W_n, W_n + (\log \frac{1}{\alpha}) \frac{W_n - Y_n}{n}]$.

Exercice 8.10. Pour tester l'efficacité d'un vaccin, on utilise deux groupes de n personnes chacun : on injecte le vaccin aux membres du premier groupe ; on injecte un placebo (une substance inactive, qui n'a aucun effet) aux membres du deuxième groupe. Après quelques mois on récolte les données. On note X_i la variable aléatoire de Bernoulli qui vaut 1 si la i-ème personne du premier groupe a été malade et 0 sinon ; de même, on note Y_i la variable aléatoire de Bernoulli qui vaut 1 si la i-ème personne du deuxième groupe a été malade et 0 sinon.

Notre modèle statistique P_θ contient deux paramètres : la probabilité p qu'une personne non vaccinée tombe malade ; la probabilité q qu'une personne vaccinée soit protégée du virus (c'est le paramètre que l'on cherche à estimer). Par rapport à la probabilité P_θ avec $\theta = (p,q) \in \Theta =]0,1[^2$, les variables $(X_i, Y_i)_{i\geq 1}$ sont indépendantes, avec $Y_i \sim \mathrm{Bern}(p)$ et $X_i \sim \mathrm{Bern}((1-q)p)$ (car afin qu'un individu vacciné tombe malade il faut qu'il soit en contact avec le virus *et* que le vaccin ne le protège pas). On considère l'estimateur

$$R_n := \frac{X_1 + \cdots + X_n}{Y_1 + \cdots + Y_n} = \frac{\overline{X}_n}{\overline{Y}_n},$$

où $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ et $\overline{Y}_n = \frac{1}{n}\sum_{i=1}^n Y_i$.

(i) En reprenant les conclusions de l'Exercice 7.27, montrer que $(\log R_n)_{n\geq 0}$ est un estimateur asymptotiquement normal pour $\log(1-q)$, de variance asymptotique donnée par $\sigma^2(\theta) = \frac{2-q}{(1-q)p} - 2$.

(ii) En déduire que $1 - R_n$ est un estimateur consistant pour q.

(iii) Donner un intervalle de confiance asymptotique pour $\log(1-q)$ de niveau $1-\alpha$ (qui dépendra de $\sigma(\theta)$). En déduire un intervalle de confiance pour q, de niveau $1-\alpha$.

(iv) Un laboratoire a testé le vaccin avec deux groupes de $n = 20000$ personnes : 550 d'entre elles sont tombées malades, dont seulement 50 du premier groupe. En utilisant ces données, et en acceptant le fait que $\sigma^2(\theta) \approx \frac{1+R_n}{\overline{X}_n} - 2$, donner un intervalle de confiance pour l'efficacité q du vaccin, de niveau de confiance 95%.

Solution 8.10. (i) On est exactement dans le contexte de l'Exercice 7.27 : les $(X_i)_{i\geq 1}$ sont des variables aléatoires i.i.d. d'espérance $\mathrm{E}_\theta(X_i) = (1-q)p > 0$ et de variance $\mathrm{Var}_\theta(X_i) = (1-q)p(1-(1-q)p)$, les $(Y_i)_{i\geq 1}$ sont des variables aléatoires i.i.d. d'espérance $\mathrm{E}_\theta(Y_i) = p > 0$ et de variance $\mathrm{Var}_\theta(Y_i) = p(1-p)$; de plus $(X_i)_{i\geq 1}$ et $(Y_i)_{i\geq 1}$ sont indépendantes. La dernière question de l'Exercice 7.27 démontre que

$$\frac{\sqrt{n}}{\sigma(\theta)}\Big(\log R_n - \log\frac{\mathrm{E}_\theta(X_i)}{\mathrm{E}_\theta(Y_i)}\Big) = \frac{\sqrt{n}}{\sigma(\theta)}\Big(\log R_n - \log(1-q)\Big) \xrightarrow{\mathcal{L}} \mathrm{N}(0,1),$$

avec

$$\sigma^2(\theta) = \frac{\mathrm{Var}_\theta(X_i)}{\mathrm{E}_\theta(X_i)^2} + \frac{\mathrm{Var}_\theta(Y_i)}{\mathrm{E}_\theta(Y_i)^2} = \frac{1-(1-q)p}{(1-q)p} + \frac{1-p}{p} = \frac{1}{(1-q)p} + \frac{1}{p} - 2.$$

Ainsi, $(\log R_n)_{n\geq 1}$ est un estimateur asymptotiquement normal pour $\log(1-q)$, de variance asymptotique $\sigma^2(\theta) = \frac{2-q}{(1-q)p} - 2$.

(ii) Il découle directement de la Proposition 8.29 que $(\log R_n)_{n\geq 1}$, étant asymptotiquement normal, est un estimateur consistant pour $\log(1-q)$: autrement dit, pour tout $\varepsilon > 0$ on a

$$\lim_{n\to+\infty} \mathrm{P}_\theta\left(\big|\log R_n - \log(1-q)\big| > \varepsilon\right) = 0.$$

Maintenant, pour $\varepsilon > 0$, considérons la probabilité

$$\begin{aligned}\mathrm{P}_\theta\left(|(1-R_n)-q| > \varepsilon\right) &= \mathrm{P}_\theta\left((1-R_n)-q > \varepsilon\right) + \mathrm{P}_\theta\left((1-R_n)-q < -\varepsilon\right)\\ &= \mathrm{P}_\theta\left(R_n < (1-\varepsilon')(1-q)\right) + \mathrm{P}_\theta\left(R_n > (1+\varepsilon')(1-q)\right)\end{aligned}$$

où on a posé $\varepsilon' := \frac{\varepsilon}{1-q} > 0$. On obtient donc

$$\begin{aligned}\mathrm{P}_\theta\left(|(1-R_n)-q| > \varepsilon\right) &= \mathrm{P}_\theta\left(\log R_n < \log(1-q) - \log(\tfrac{1}{1-\varepsilon'})\right)\\ &\qquad + \mathrm{P}_\theta\left(\log R_n > \log(1-q) + \log(1+\varepsilon')\right)\\ &\leq \mathrm{P}_\theta\left(|\log R_n - \log(1-q)| > \log(\tfrac{1}{1-\varepsilon'})\right)\\ &\qquad + \mathrm{P}_\theta\left(|\log R_n - \log(1-q)| > \log(1+\varepsilon')\right).\end{aligned}$$

Comme on l'a vu plus haut, les deux termes tendent vers 0 quand $n \to +\infty$, parce que $\log(\frac{1}{1-\varepsilon'}) > 0$ et $\log(1+\varepsilon') > 0$. Pour conclure, pour tout $\varepsilon > 0$ on a

$$\lim_{n\to+\infty} \mathrm{P}_\theta\left(|(1-R_n)-q| > \varepsilon\right) = 0,$$

ce qui démontre que $1-R_n$ est un estimateur consistant pour q.

(iii) De la normalité asymptotique de l'estimateur $\log R_n$, on en déduit directement qu'un intervalle de confiance asymptotique pour $\log(1-q)$, de niveau de confiance $1-\alpha$, est donné par

$$\left[\log R_n - \frac{\sigma(\theta)}{\sqrt{n}} z_{\alpha/2}, \log R_n + \frac{\sigma(\theta)}{\sqrt{n}} z_{\alpha/2}\right].$$

Vu que $\log(1-q) \in [a,b]$ si et seulement si $q \in [1-\mathrm{e}^b, 1-\mathrm{e}^a]$, on obtient l'intervalle de confiance asymptotique suivant pour q, de niveau de confiance $1-\alpha$:

$$\left[1-R_n \mathrm{e}^{\frac{\sigma(\theta)}{\sqrt{n}} z_{\alpha/2}}, 1-R_n \mathrm{e}^{-\frac{\sigma(\theta)}{\sqrt{n}} z_{\alpha/2}}\right].$$

(iv) En utilisant les données fournies dans l'énoncé, on obtient $\overline{X}_n = \frac{50}{20000} = \frac{1}{400}$ et $\overline{Y}_n = \frac{500}{20000} = \frac{1}{40}$, donc $R_n = \frac{1}{10}$ et $\sigma^2(\theta) \approx 438$. Pour obtenir un intervalle de confiance à 95%, il faut prendre $\alpha = 0,05$, pour lequel on a $z_{\alpha/2} \approx 1,96$, voir la table page S-225 de la fonction de répartition Φ d'une variable aléatoire $\mathrm{N}(0,1)$. Pour $z_{\alpha/2} \approx 1,96$, avec $\sigma^2(\theta) \approx 438$, on trouve

$$\frac{\sigma(\theta)}{\sqrt{n}} z_{\alpha/2} \approx \frac{20,93}{141,42} 1,96 \approx 0,29.$$

On obtient donc l'intervalle de confiance asymptotique suivant pour q, de niveau de confiance 95% :

$$\left[1-\frac{1}{10}\mathrm{e}^{0,29}, 1-\frac{1}{10}\mathrm{e}^{-0,29}\right] \approx [0,866, 0,926].$$

8.3 Estimateur du maximum de vraisemblance

Exercice 8.11. On considère le modèle statistique suivant :

$$f(x;\theta) = (\theta+1)x^\theta \mathbb{1}_{]0,1[}(x), \qquad \theta \in \Theta =]-1,+\infty[.$$

Déterminer l'estimateur du maximum de vraisemblance, montrer qu'il est asymptotiquement normal et en calculer la variance asymptotique $\sigma^2(\theta)$.

Solution 8.11. Vérifions que les hypothèses (A)–(G) sont vérifiées et calculons la variance asymptotique avec la formule (8.31). Il nous sera utile pour la suite d'observer que si X est une variable aléatoire a densité, de densité donnée par $f(x;\theta)$, alors $Y := \log X \sim \mathrm{Exp}(\theta+1)$. En effet, pour $y > 0$,

$$\mathrm{P}(Y \le y) = \mathrm{P}(X \ge \mathrm{e}^{-y}).$$

En dérivant cette égalité, on obtient $f_Y(y) = \mathrm{e}^{-y} f(\mathrm{e}^{-y};\theta)$, dont découle facilement que $Y \sim \mathrm{Exp}(\theta+1)$.

Considérons maintenant les hypothèses (A)–(G). Les hypothèses (A) et (B) sont trivialement vérifiées. En ce qui concerne (C), on observe que pour $x_1, \ldots, x_n \in]0,1[$,

$$L_n(\underline{x}, \theta) = \frac{1}{n} \sum_{i=1}^{n} \log f(x_i; \theta) = \log(\theta+1) + \theta \frac{1}{n} \sum_{i=1}^{n} \log x_i .$$

En dérivant, on obtient

$$\frac{\mathrm{d} L_n(\underline{x}, \theta)}{\mathrm{d}\theta} = \frac{1}{\theta+1} - \frac{1}{n} \sum_{i=1}^{n} (-\log x_i),$$

donc on déduit que $\theta \mapsto L_n(\underline{x}, \theta)$ admet un unique point de maximum

$$\hat{\theta}(x_1, \ldots, x_n) = -1 + \frac{n}{\sum_{i=1}^{n} (-\log x_i)}.$$

L'estimateur MV est donc donné par

$$Y_n = -1 + \frac{n}{\sum_{i=1}^{n} (-\log X_i)}.$$

La condition (D) découle du fait que $-\log X_1 \sim \mathrm{Exp}(\theta+1)$. En changeant la valeur de $\theta \in \Theta$, la loi de $-\log X_1$ (et donc de X_1) change, comme on le voit par exemple en remarquant que $\mathrm{E}_\theta(-\log X_1) = 1/(\theta+1)$ est une fonction injective de θ.

Pour l'hypothèse (E), on note que $\log f(X_1; t) = \log(\theta+1) + \theta \log(X_1)$. Donc (E) découle directement du fait que $-\log(X_1)$ admet un moment d'ordre deux fini, en tant que variable aléatoire exponentielle.

Passons maintenant à l'hypothèse (F). La (F) (i) est immédiate. Pour la (F) (ii), il suffit d'observer que

$$B(x, \theta) := \frac{\mathrm{d}^2 \log f(x; \theta)}{\mathrm{d}\theta^2} = -\frac{1}{(\theta+1)^2}.$$

Cette fonction est continue en θ et ne dépend pas de x, donc elle est nécessairement continue en θ uniformément en x. Pour la (F) (iii), vérifions-la pour $k = 1$:

$$\int_0^1 \frac{\mathrm{d}}{\mathrm{d}\theta} f(x; \theta) \mathrm{d}x = \int_0^1 \left[x^\theta + (\theta+1) x^\theta \log x \right] \mathrm{d}x = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}x} \left(x^{\theta+1} \log x \right) \mathrm{d}x = 0,$$

où pour le dernier passage on a utilisé le théorème fondamental du calcul intégral. Le cas $k = 2$ est analogue, au vu de l'identité suivante, obtenue par un calcul direct

$$\frac{\mathrm{d}^2}{\mathrm{d}\theta^2} f(x; \theta) = 2x^\theta \log x + (\theta+1) x^\theta (\log x)^2 = \frac{\mathrm{d}}{\mathrm{d}x} \left(x^{\theta+1} \log^2 x \right).$$

Il faut noter que la vérification de (F) (iii) peut être faite en utilisant un résultat classique sur la dérivation d'intégrales dépendant d'un paramètre, voir par exemple le Théorème 2.27 dans [25]. Cependant, cette méthode présuppose des connaissances plus avancées d'analyse réelle.

Enfin, l'hypothèse (G) est aussi vérifié car, comme on l'a déjà vu, $B(x, \theta)$ ne dépend pas de x.

La normalité asymptotique de l'estimateur MV Y_n se déduit donc du Théorème 8.40, qui de plus donne

$$\frac{1}{\sigma^2(\theta)} := -\mathrm{E}_\theta[B(X_1,\theta)] = \frac{1}{(\theta+1)^2},$$

d'où $\sigma^2(\theta) = (\theta+1)^2$.

Exercice 8.12. Déterminer l'estimateur du maximum de vraisemblance dans les modèles statistiques suivants :

(i) $\mathrm{p}(x;\theta) = \theta^x(1-\theta)^{1-x}$ pour $x \in \{0,1\}$ est la densité discrète d'une variable aléatoire de Bernoulli de paramètre θ, pour $\theta \in \Theta =]0,1[$;

(ii) $f(x;\theta)$ est la densité d'une variable aléatoire normale $\mathrm{N}(0,\theta^2)$, pour $\theta \in \Theta =]0,+\infty[$;

(iii) $f(x;\theta) = \theta x^{-(1+\theta)} \mathbb{1}_{]1,+\infty[}$, pour $\theta \in \Theta =]0,+\infty[$.

Dans les trois cas, déterminer si les hypothèses (A)–(G) sont vérifiées. Si l'estimateur est asymptotiquement normal, en donner la variance asymptotique $\sigma^2(\theta)$.

Solution 8.12. (i) Dans ce cas, les hypothèses (A) et (B) sont trivialement vérifiées, avec $I = \{0,1\}$. Pour l'hypothèse (C), on note que pour $x_1,\ldots,x_n \in \{0,1\}$,

$$\begin{aligned} L_n(\underline{x},\theta) &= \frac{1}{n}\sum_{i=1}^n \log \mathrm{p}(x_i;\theta) = \frac{1}{n}\sum_{i=1}^n x_i \log\theta + \frac{1}{n}\sum_{i=1}^n (1-x_i)\log(1-\theta) \\ &= \overline{x}_n \log\theta + (1-\overline{x}_n)\log(1-\theta), \end{aligned}$$

où on a posé $\overline{x}_n := \frac{1}{n}\sum_{i=1}^n x_i$. En dérivant, on trouve

$$\frac{\mathrm{d}L_n(\underline{x},\theta)}{\mathrm{d}\theta} = \frac{\overline{x}_n - \theta}{\theta(1-\theta)},$$

dont on déduit que $\theta \mapsto L_n(\underline{x},\theta)$ admet un unique point de maximum, égal à

$$\hat{\theta}(x_1,\ldots,x_n) = \overline{x}_n = \frac{1}{n}\sum_{i=1}^n x_i .$$

L'estimateur MV est donc donné par

$$Y_n = \overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i ,$$

c'est-à-dire que l'on retrouve la moyenne empirique. On sait déjà, grâce au théorème central limite, que $(Y_n)_{n\geq 1}$ est un estimateur sans biais et asymptotiquement normal pour $\theta = \mathrm{E}_\theta(X_1)$, de variance asymptotique $\sigma^2(\theta) = \theta(1-\theta) = \mathrm{Var}_\theta(X_1)$.

Pour être complet, vérifions si les conditions (D)–(F) sont vérifiées. La condition (D) découle du fait que $X_1 \sim \mathrm{Bern}(\theta)$: la loi de X_1 change si on change la valeur de $\theta \in \Theta$, comme on le voit par exemple en observant que $\mathrm{E}_\theta(X_1) = \theta$. Pour l'hypothèse (E), on note que

$$\log \mathrm{p}(X_1;\theta) = X_1 \log\theta + (1-X_1)\log(1-\theta),$$

et admet donc un moment d'ordre deux fini parce que X_1 admet un moment d'ordre deux fini, en tant que variable de Bernoulli.

Passons maintenant à l'hypothèse (F). La (F) (i) est immédiate. Pour la (F) (ii), on a que

$$B(x,\theta) := \frac{\mathrm{d}^2 \log \mathrm{p}(x;\theta)}{\mathrm{d}\theta^2} = \frac{x(2\theta-1)-\theta^2}{\theta^2(1-\theta)^2}.$$

On a donc que $B(x,\theta)$ est continue pour $x \in I \in \{0,1\}$. Vu que l'ensemble I est fini, on obtient que $B(x,t)$ est continue en θ uniformément en $x \in I$. En effet, pour tout x fixé, on a que pour $\theta \in \Theta$ et $\varepsilon > 0$ il existe $\delta_\varepsilon = \delta_\varepsilon(x)$ tel que $|t-\theta| < \delta_\varepsilon$ implique $|B(x,t) - B(x,\theta)| < \varepsilon$: le choix $\delta_\varepsilon := \min_{x\in I} \delta_\varepsilon(x)$ permet alors de vérifier la condition (F) (ii). Pour la (F) (iii), il découle directement du fait que I est fini que l'on peut intervertir les sommes et les dérivées.
Pour finir, l'hypothèse (G) est aussi vérifiée, parce que $B(X_1,t) = \frac{X_1(2\theta-1)-\theta^2}{\theta^2(1-\theta)^2}$, avec X_1 qui admet un moment d'ordre deux fini.
La normalité asymptotique de l'estimateur MV Y_n découle alors du Théorème 8.40 (en plus du théorème central limite), qui de plus nous donne

$$\frac{1}{\sigma^2(\theta)} := -\mathrm{E}_\theta[B(X_1,\theta)] = -\frac{\theta(2\theta-1)-\theta^2}{\theta^2(1-\theta)^2} = \frac{1}{\theta(1-\theta)},$$

d'où $\sigma^2(\theta) = \theta(1-\theta)$.

(ii) On a dans ce cas, pour $x \in \mathbb{R}$

$$f(x;\theta) = \frac{1}{\sqrt{2\pi\theta}} \mathrm{e}^{-\frac{x^2}{2\theta}},$$

qui est le modèle statistique de l'Exercice 8.2. Les hypothèse (A) et (B) sont trivialement vérifiées, avec $I = \mathbb{R}$. En ce qui concerne l'hypothèse (C), pour $x_1,\dots,x_n \in \mathbb{R}$, on a

$$L_n(\underline{x},\theta) = \frac{1}{n}\sum_{i=1}^n \log f(x_i;\theta) = -\frac{1}{2}\log(2\pi\theta) - \frac{1}{n}\sum_{i=1}^n x_i^2 \frac{1}{2\theta}.$$

En dérivant, on trouve

$$\frac{\mathrm{d}L_n(\underline{x},\theta)}{\mathrm{d}\theta} = \frac{1}{2\theta^2}\left(\frac{1}{n}\sum_{i=1}^n x_i^2 - \theta\right),$$

dont on déduit que $\theta \mapsto L_n(\underline{x},\theta)$ admet un unique point de maximum, égal à

$$\hat{\theta}(x_1,\dots,x_n) = \frac{1}{n}\sum_{i=1}^n x_i^2.$$

L'estimateur MV est donc donné par

$$Y_n = \overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i^2.$$

On sait déjà, grâce au théorème central limite, que $(Y_n)_{n\geq 1}$ est un estimateur sans biais et asymptotiquement normal pour $\theta = \mathrm{E}_\theta(X_1^2)$, de variance asymptotique

$$\sigma^2(\theta) = \mathrm{Var}_\theta(X_1^2) = \mathrm{E}_\theta(X_1^4) - \mathrm{E}_\theta(X_1^2)^2 = 2\theta^2$$

(voir les calculs de l'Exercice 8.2, en particulier $\mathrm{E}_\theta(X_1^4) = 3\theta^2$).
Vérifions quand même si les hypothèses (D)–(F) sont vérifiées. La condition (D) découle du fait que $X_1 \sim \mathrm{N}(0,\theta)$: la loi de X_1 change si on change la valeur de $\theta \in \Theta$, comme on le voit par exemple en notant que $\mathrm{E}_\theta(X_1^2) = \theta$. Pour l'hypothèse (E), on observe que

$$\log f(X_1;\theta) = -\frac{1}{2}\log(2\pi\theta) - \frac{X_1^2}{2\theta},$$

qui admet un moment d'ordre deux fini parce que X_1^2 admet un moment d'ordre deux fini, X_1 étant une variable normale.

Passons maintenant à l'hypothèse (F). La (F) (i) est immédiate. En revanche, l'hypothèse (F) (ii) n'est pas vérifiée : en effet, on a

$$B(x,\theta) := \frac{\mathrm{d}^2 \log f(x;\theta)}{\mathrm{d}\theta^2} = \frac{\theta - 2x^2}{2\theta^3}.$$

Pour voir que (F) (ii) n'est pas vérifiée, notons que pour $t,\theta \in]0,+\infty[$

$$\begin{aligned} B(x,t) - B(x,\theta) &= \frac{\theta^2 - t^2}{2t^2\theta^2} + \frac{(t^3-\theta^3)x^2}{t^3\theta^3} \\ &= \frac{\theta+t}{2t^2\theta^2}(\theta - t) + \frac{t^2+t\theta+\theta^2}{t^3\theta^3}(t-\theta)x^2 \end{aligned}$$

et que $(t-\theta)x^2$ ne peut pas être rendu arbitrairement petit uniformément en $x \in \mathbb{R}$. En revanche, on peut facilement vérifier que $B(x,\theta)$ vérifie les hypothèses modifiées de l'Observation 8.39, voir (8.30).

Pour la (F) (iii), comme dans l'Exercice 8.11, on pourrait utiliser le Théorème 2.27 dans [25]. Vérifions (F) (iii) par des calculs directs. Pour $k = 1$, avec un calcul direct, on a

$$\frac{\mathrm{d}}{\mathrm{d}\theta} f(x;\theta) = -\frac{1}{2\sqrt{2\pi}\theta^{3/2}} \mathrm{e}^{-\frac{x^2}{2\theta}} + \frac{1}{\sqrt{2\pi\theta}} \frac{x^2}{2\theta^2} \mathrm{e}^{-\frac{x^2}{2\theta}} = \frac{x^2-\theta}{2\theta^2} f(x;\theta),$$

donc

$$\int_{-\infty}^{+\infty} \frac{\mathrm{d}}{\mathrm{d}\theta} f(x;\theta)\mathrm{d}x = \int_{-\infty}^{+\infty} \frac{x^2-t}{2\theta^2} f(x;\theta)\mathrm{d}x = \mathrm{E}_\theta\left(\frac{X_1^2-\theta}{2\theta^2}\right) = 0,$$

où dans le dernier passage on a utilisé la linéarité de l'espérance et le fait que $\mathrm{E}_\theta(X_1^2) = \theta$. Le cas $k = 2$ est analogue, parce que

$$\frac{\mathrm{d}^2}{\mathrm{d}\theta^2} f(x;\theta) = \left(\left(\frac{x^2-\theta}{2\theta^2}\right)^2 + \frac{\theta - 2x^2}{2\theta^3}\right) f(x;\theta)$$

de sorte que

$$\int_{-\infty}^{+\infty} \frac{\mathrm{d}^2}{\mathrm{d}\theta^2} f(x;\theta)\mathrm{d}x = \mathrm{E}_\theta\left(\left(\frac{X_1^2-\theta}{2\theta^2}\right)^2 + \frac{\theta - 2X_1^2}{2\theta^3}\right) = \frac{2\theta^2}{4\theta^4} - \frac{1}{2\theta^2} = 0,$$

où on a utilisé le fait que $\mathrm{E}_\theta(X_1^2) = \theta$ et $\mathrm{E}_\theta(X_1^4) = 3\theta^2$.

Enfin, l'hypothèse (G) est aussi vérifiée, parce que $B(X_1,t) = \frac{\theta - 2X_1^2}{2\theta^3}$, avec X_1^2 qui admet un moment d'ordre deux fini.

La normalité asymptotique de l'estimateur Y_n découle alors du Théorème 8.40 (en plus du théorème central limite), qui de plus donne

$$\frac{1}{\sigma^2(\theta)} := -\mathrm{E}_\theta[B(X_1,\theta)] = -\frac{\theta - 2\theta}{2\theta^3} = \frac{1}{2\theta^2},$$

d'où $\sigma^2(\theta) = 2\theta^2$.

(iii) Les hypothèses (A) et (B) sont trivialement vérifiées, avec $I =]1, +\infty[$. Pour la (C), on observe que pour $x_1, \ldots, x_n \in]1, +\infty[$,

$$L_n(\underline{x}, \theta) = \frac{1}{n} \sum_{i=1}^{n} \log f(x_i; \theta) = \log \theta - (1+\theta) \frac{1}{n} \sum_{i=1}^{n} \log x_i \,.$$

En dérivant, on trouve

$$\frac{\mathrm{d} L_n(\underline{x}, \theta)}{\mathrm{d}\theta} = \frac{1}{\theta} - \frac{1}{n} \sum_{i=1}^{n} \log x_i \,,$$

dont on déduit que $\theta \mapsto L_n(\underline{x}, \theta)$ admet un unique point de maximum, égal à

$$\hat{\theta}(x_1, \ldots, x_n) = \frac{n}{\sum_{i=1}^{n} \log x_i} \,.$$

L'estimateur MV est donc donné par

$$Y_n = \frac{n}{\sum_{i=1}^{n} \log X_i} .$$

Pour la condition (D), avec un calcul direct on trouve que les fonctions de répartition sont égales à

$$\mathrm{P}_\theta \left(X_1 \leq x \right) = \begin{cases} 0 & \text{si } x \leq 1 \,, \\ 1 - (1-x)^{-\theta} & \text{si } x > 1 \,. \end{cases}$$

Ainsi, vu que pour $x > 1$ on a $(1-x)^{-\theta} = (1-x)^{-t}$ si et seulement si $\theta = t$, on obtient que la fonction de répartition de X_1 (donc sa loi) change si on change le paramètre θ, montrant que la condition (D) est vérifiée.

Pour l'hypothèse (E), on observe que $\log f(X_1; t) = \log \theta - (1+\theta) \log(X_1)$. Notons que par une double intégration par parties,

$$\begin{aligned} \mathrm{E}_\theta \left((\log X_1)^2 \right) &= \int_1^\infty (\log x)^2 \theta x^{-(1+\theta)} \mathrm{d}x \\ &= 2 \int_1^\infty (\log x) x^{-(1+\theta)} \mathrm{d}x = \frac{2}{\theta} \int_1^\infty x^{-(1+\theta)} \mathrm{d}x = \frac{2}{\theta^2} < +\infty \,. \end{aligned}$$

Donc $\log f(X_1; t)$ admet un moment d'ordre deux fini et (E) est vérifiée.

Passons maintenant à l'hypothèse (F). La (F) (i) est immédiate. Pour la (F) (ii), on a

$$B(x, \theta) := \frac{\mathrm{d}^2 \log f(x; \theta)}{\mathrm{d}\theta^2} = -\frac{1}{\theta^2} .$$

Cette fonction est continue en θ et ne dépend pas de x, donc elle est nécessairement continue en θ uniformément en x. Pour la (F) (iii), vérifions-la pour $k = 1$:

$$\begin{aligned} \int_1^\infty \frac{\mathrm{d}}{\mathrm{d}\theta} f(x; \theta) \mathrm{d}x &= \int_1^{+\infty} \left[x^{-(1+\theta)} - \theta x^{-(1+\theta)} \log x \right] \mathrm{d}x \\ &= \int_1^\infty \frac{\mathrm{d}}{\mathrm{d}x} \left(x^{-\theta} \log x \right) \mathrm{d}x = 0 \,, \end{aligned}$$

où dans le dernier passage on a utilisé le théorème fondamental du calcul intégral. Le cas $k = 2$ est analogue, parce que l'on a, grâce à un calcul direct

$$\frac{\mathrm{d}^2}{\mathrm{d}\theta^2} f(x;\theta) = -2x^{-(1+\theta)} \log x + \theta x^{-(1+\theta)} (\log x)^2 = \frac{\mathrm{d}}{\mathrm{d}x} \left(-x^{-\theta} \log^2 x \right).$$

Enfin, l'hypothèse (G) est aussi vérifiée car, comme on l'a déjà observé, $B(x,\theta)$ ne dépend pas de x.

La normalité asymptotique de l'estimateur MV Y_n découle alors du Théorème 8.40, qui de plus nous donne

$$\frac{1}{\sigma^2(\theta)} := -\mathrm{E}_\theta[B(X_1,\theta)] = \frac{1}{\theta^2},$$

d'où $\sigma^2(\theta) = \theta^2$.

Exercice 8.13. Montrer que le modèle statistique

$$f(x;\theta) = \frac{\sqrt{\theta}}{\pi} \frac{1}{1+\theta x^2}, \qquad \theta \in \Theta =]0,+\infty[,$$

satisfait les hypothèses (A)–(G), et calculer la variance asymptotique $\sigma^2(\theta)$ de l'estimateur du maximum de vraisemblance correspondant.

Solution 8.13. Les hypothèses (A) et (B) sont facilement vérifiées, avec $I := \mathbb{R} \setminus \{0\}$ (la raison de ce choix pour I sera plus claire d'ici peu). La condition (C) est moins banale. Montrons que pour tout $\underline{x} = (x_1, x_2, \ldots, x_n) \in I^n$ fixé, la fonction

$$L_n(\underline{x},\theta) = \frac{1}{n} \sum_{i=1}^{n} \log f(x_i;\theta) = \frac{1}{2} \log \theta - \log \pi - \frac{1}{n} \sum_{i=1}^{n} \log\left(1+\theta x_i^2\right)$$

admet un unique maximum local dans Θ. Notons que $\theta \mapsto L_n(\underline{x},\theta)$ est une fonction dérivable, avec

$$\frac{\mathrm{d}}{\mathrm{d}\theta} L_n(\underline{x},\theta) = \frac{1}{2\theta} - \frac{1}{n} \sum_{i=1}^{n} \frac{x_i^2}{1+\theta x_i^2} = \frac{1}{\theta} \left[\frac{1}{2} - \frac{1}{n} \sum_{i=1}^{n} \frac{\theta x_i^2}{1+\theta x_i^2} \right].$$

L'application $\theta \mapsto \alpha(\theta) := \frac{1}{2} - \frac{1}{n} \sum_{i=1}^{n} \frac{\theta x_i^2}{1+\theta x_i^2}$ est continue et telle que $\alpha(0) = \frac{1}{2}$. De plus, en utilisant l'hypothèse que $x_i \neq 0$, par définition de I, la fonction α est strictement décroissante et $\lim_{\theta \to +\infty} \alpha(\theta) = -\frac{1}{2}$. Par continuité, il existe un unique $\hat{\theta}$ pour lequel $\frac{\mathrm{d}}{\mathrm{d}\theta} L_n(\underline{x},\hat{\theta}) = 0$. De plus, $\frac{\mathrm{d}}{\mathrm{d}\theta} L_n(\underline{x},\theta) > 0$ pour $\theta < \hat{\theta}$ et $\frac{\mathrm{d}}{\mathrm{d}\theta} L_n(\underline{x},\theta) < 0$ pour $\theta > \hat{\theta}$, ce qui implique que $\hat{\theta}$ est un maximum local pour $\theta \mapsto L_n(\underline{x},\theta)$ et que c'est le seul.

Considérons maintenant l'hypothèse (D). Elle découle de l'observation que

$$\mathrm{P}_\theta(X_1 \leq x) = \frac{1}{\pi} \left(\arctan(\sqrt{\theta} x) + \frac{\pi}{2} \right)$$

qui, pour $x \neq 0$, est une fonction injective de θ.

Pour la condition (E), notons que

$$\mathrm{E}_\theta \left[(\log \mathrm{p}(X_1,t))^2 \right] = \int_{-\infty}^{+\infty} \frac{\sqrt{\theta}}{\pi} \frac{1}{1+\theta x^2} \log^2 \left(\frac{\sqrt{\theta}}{\pi} \frac{1}{1+\theta x^2} \right) \mathrm{d}x < +\infty,$$

où la finitude de cette dernière intégrale vient par exemple du fait que la fonction intégrée tend vers 0 quand $|x| \to +\infty$ plus rapidement que la fonction $1/(1+|x|)^{3/2}$, qui est intégrable sur $\mathbb{R}$.

Passons maintenant à l'hypothèse (F). La (F) (i) est immédiate. Pour la (F) (ii), on obtient

$$B(x,\theta) := \frac{\mathrm{d}^2 \log f(x;\theta)}{\mathrm{d}\theta^2} = -\frac{1}{2\theta^2} + \frac{x^4}{(1+\theta x^2)^2}\,. \tag{S8.1}$$

Notons que

$$\left|\frac{\mathrm{d}}{\mathrm{d}\theta}B(x,\theta)\right| = \left|\frac{1}{\theta^3} - \frac{2x^6}{(1+\theta x^2)^3}\right| \leq \frac{1}{\theta^3} + \frac{2x^6}{(1+\theta x^2)^3} \leq \frac{3}{\theta^3},$$

où on a minoré le dénominateur $(1+\theta x^2)^3 \geq (\theta x^2)^3 = \theta^3 x^6$. Ainsi, pour $0 < \theta_1 < \theta_2 < \infty$, en appliquant l'inégalité de Taylor–Lagrange (dans le cas présent réduite à l'inégalité des accroissements finis) on obtient

$$|B(x,\theta_2) - B(x,\theta_1)| \leq \sup_{\theta\in[\theta_1,\theta_2]} \left|\frac{\mathrm{d}}{\mathrm{d}\theta}B(x,\theta)\right| (\theta_2-\theta_1) \leq \frac{3(\theta_2-\theta_1)}{\theta_1^2},$$

dont on déduit que $B(x,\theta)$ est continue en θ uniformément en x.

Vérifions maintenant la condition (F) (iii). Comme dans l'Exercice 8.11, on pourrait utiliser le Théorème 2.27 de [25]. On utilise ici une vérification directe du fait que

$$\int_{-\infty}^{+\infty} \frac{\mathrm{d}^k}{\mathrm{d}\theta^k} f(x;\theta)\mathrm{d}x = 0 \tag{S8.2}$$

pour $k = 1, 2$. Par commodité, on introduit la notation $h(x) := f(x;1)$, de sorte que

$$f(x;\theta) = \sqrt{\theta}h(\sqrt{\theta}x)\,,$$

On voit alors que

$$\frac{\mathrm{d}}{\mathrm{d}\theta} f(x;\theta) = \frac{1}{2\sqrt{\theta}}h(\sqrt{\theta}x) + \frac{x}{2}h'(\sqrt{\theta}x)\,. \tag{S8.3}$$

On vérifie facilement que cette dernière fonction est intégrable sur $\mathbb{R}$ (elle décroît vers 0 quand $x \to \pm\infty$ comme $1/x^2$) et que

$$\frac{\mathrm{d}}{\mathrm{d}\theta} f(x;\theta) = \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{x}{2\sqrt{\theta}}h(\sqrt{\theta}x)\right].$$

Ainsi, en appliquant le théorème fondamental du calcul intégral,

$$\int_{-\infty}^{+\infty} \frac{\mathrm{d}}{\mathrm{d}\theta} f(x;\theta)\mathrm{d}x = \lim_{x\to+\infty} \frac{x}{2\sqrt{\theta}}h(\sqrt{\theta}x) - \lim_{x\to-\infty} \frac{x}{2\sqrt{\theta}}h(\sqrt{\theta}x) = 0,$$

ce qui démontre (S8.2) pour $k = 1$. Le cas $k = 2$ est un peu plus compliqué. En dérivant une deuxième fois (S8.3), on obtient

$$\frac{\mathrm{d}^2}{\mathrm{d}\theta^2} f(x;\theta) = -\frac{1}{4\theta^{3/2}}h(\sqrt{\theta}x) + \frac{x}{4\theta}h'(\sqrt{\theta}x) + \frac{x^2}{4\sqrt{\theta}}h''(\sqrt{\theta}x). \tag{S8.4}$$

Considérons maintenant, un par un, les trois termes du membre de droite de (S8.4). On vérifie facilement que tous les trois décroissent vers 0 quand $x \to \pm\infty$ comme $1/x^2$, et sont donc intégrables sur $\mathbb{R}$. En utilisant le fait que

$$\int_{-\infty}^{+\infty} h(\sqrt{\theta}x)\mathrm{d}x = \frac{1}{\sqrt{\theta}} \int_{-\infty}^{+\infty} f(x;\theta)\mathrm{d}x = \frac{1}{\sqrt{\theta}},$$

on a

$$-\int_{-\infty}^{+\infty} \frac{1}{4\theta^{3/2}} h(\sqrt{\theta}x)\mathrm{d}x = -\frac{1}{4\theta^2}.$$

De plus, en intégrant par parties,

$$\int_{-\infty}^{+\infty} \frac{x}{4\theta} h'(\sqrt{\theta}x)\mathrm{d}x = -\int_{-\infty}^{+\infty} \frac{1}{4\theta^{3/2}} h(\sqrt{\theta}x)\mathrm{d}x = -\frac{1}{4\theta^2}.$$

Enfin, avec une double intégration par parties, on obtient

$$\int_{-\infty}^{+\infty} \frac{x^2}{4\sqrt{\theta}} h''(\sqrt{\theta}x)\mathrm{d}x = \frac{1}{2}\int_{-\infty}^{+\infty} \frac{1}{\theta^{3/2}} h(\sqrt{\theta}x)\mathrm{d}x = \frac{1}{2\theta^2}.$$

Avec ces trois résultats, en revenant à (S8.4) on obtient

$$\int_{-\infty}^{+\infty} \frac{\mathrm{d}^2}{\mathrm{d}\theta^2} f(x;\theta)\mathrm{d}x = 0,$$

ce qui complète la démonstration de (F) (iii).

Enfin, la condition (G) découle du fait que, comme on le voit dans (S8.1), $B(x;\theta)$ est une fonction bornée en x pour tout $\theta \in \Theta$ fixé.

Il ne reste donc plus qu'à calculer la variance asymptotique $\sigma(\theta)$, pour laquelle on utilise la formule (8.31) dans le Théorème 8.40. Grâce à (S8.1) ci-dessus,

$$-\frac{1}{\sigma^2(\theta)} = \mathrm{E}_\theta[B(X_1,\theta)] = -\frac{1}{2\theta^2} + \int_{-\infty}^{+\infty} \frac{x^4}{(1+\theta x^2)^2} f(x;\theta)\mathrm{d}x. \tag{S8.5}$$

D'autre part, en utilisant d'abord un changement de variable $y = \sqrt{\theta}x$ puis deux intégrations par parties, on obtient

$$\begin{aligned}
\int_{-\infty}^{+\infty} \frac{x^4}{(1+\theta x^2)^2} f(x;\theta)\mathrm{d}x &= \int_{-\infty}^{+\infty} \frac{x^4}{(1+\theta x^2)^2} \frac{\sqrt{\theta}}{\pi} \frac{1}{1+\theta x^2}\mathrm{d}x = \frac{1}{\pi\theta^2}\int_{-\infty}^{+\infty} \frac{y^4}{(1+y^2)^3}\mathrm{d}y \\
&= -\frac{1}{\pi\theta^2}\frac{1}{4}\int_{-\infty}^{+\infty} y^3 \frac{\mathrm{d}}{\mathrm{d}y}\left(\frac{1}{(1+y^2)^2}\right)\mathrm{d}y = \frac{1}{\pi\theta^2}\frac{3}{4}\int_{-\infty}^{+\infty} \frac{y^2}{(1+y^2)^2}\mathrm{d}y \\
&= -\frac{1}{\pi\theta^2}\frac{3}{8}\int_{-\infty}^{+\infty} y\frac{\mathrm{d}}{\mathrm{d}y}\left(\frac{1}{1+y^2}\right)\mathrm{d}y = \frac{1}{\pi\theta^2}\frac{3}{8}\int_{-\infty}^{+\infty} \frac{1}{1+y^2}\mathrm{d}y = \frac{3}{8\theta^2}.
\end{aligned}$$

En insérant ce résultat dans (S8.5), on obtient

$$-\frac{1}{\sigma^2(\theta)} = -\frac{1}{2\theta^2} + \frac{3}{8\theta^2} = -\frac{1}{8\theta^2},$$

d'où

$$\sigma^2(\theta) = 8\theta^2.$$

Exercice 8.14. On considère le modèle statistique de l'Exemple 8.23 où $X_1 \sim \mathrm{U}(0,\theta)$, c'est-à-dire

$$f(x;\theta) = \frac{1}{\theta}\mathbb{1}_{[0,\theta]}(x), \qquad \theta \in \Theta =]0,+\infty[.$$

(i) Parmi les hypothèses (A)–(E), lesquelles ne sont pas vérifiées ?

(ii) Montrer cependant que pour tout $n \geq 1$ et $\underline{x} = (x_1, \ldots, x_n) \in]0, +\infty[^n$, la fonction $\theta \mapsto f(x_1; \theta) f(x_1; \theta) \cdots f(x_n; \theta)$ admet un unique maximum sur $]0, +\infty[$, en un point $\hat{\theta}_n(x_1, x_2, \ldots, x_n)$ à déterminer.

(iii) En somme, l'estimateur $Y_n = \hat{\theta}_n(X_1, X_2, \ldots, X_n)$ pour $n \geq 1$ correspond à l'estimateur du maximum de vraisemblance. L'estimateur $(Y_n)_{n \geq 1}$ est-il consistant ? Est-il asymptotiquement normal ?

Solution 8.14. (i) L'hypothèse (A) est satisfaite, mais l'hypothèse (B) ne l'est pas. En effet, par rapport à P_θ on a $X_1 \sim \mathrm{U}(0, \theta)$, donc si I est tel que $\mathrm{P}_\theta(X_1 \in I) = 1$ pour tout $\theta \in \Theta$, on a nécessairement $]0, \theta[\subseteq I$ pour tout $\theta \in \Theta$, c'est-à-dire $]0, +\infty[\subseteq I$. Mais on n'a pas $f(x; \theta) > 0$ pour tout $x \in]0, +\infty[$ (pour aucun $\theta \in \Theta$).

Vu que l'hypothèse (B) n'est pas satisfaite, cela n'a pas de sens de considérer l'hypothèse (C) et cela s'avère aussi un problème pour (E) parce que $f(X_1, t) = 0$ si $X_1 > t$ (ce qui arrive par rapport à P_θ avec probabilité positive si $\theta > t$). De même, cela n'a pas de sens de considérer les hypothèses (F)-(G) (l'hypothèse (F) (i) n'est pas vérifiée).

En revanche, l'hypothèse (D) est vérifiée, parce que $X_1 \sim \mathrm{U}(0, \theta)$ par rapport à P_θ, donc en particulier $\mathrm{E}_\theta(X_1) = \theta$ est différente pour des valeurs de θ distinctes.

(ii) Pour $\underline{x} \in]0, +\infty[^n$, considérons la fonction

$$\begin{aligned} f(x_1; \theta) f(x_1; \theta) \cdots f(x_n; \theta) &= \theta^n \mathbb{1}_{[0,\theta]}(x_1) \cdots \mathbb{1}_{[0,\theta]}(x_n) \\ &= \theta^n \mathbb{1}_{[0,\theta]}(\max(x_1, \ldots, x_n)), \end{aligned}$$

où on a utilisé que $x_i \in [0, \theta]$ pour tout $1 \leq i \leq n$ si et seulement si $\max(x_1, \ldots, x_n) \in [0, \theta]$. Ainsi, la fonction $\theta \mapsto f(x_1; \theta) f(x_1; \theta) \cdots f(x_n; \theta)$ vaut θ^n pour $\theta \leq \max(x_1, \ldots, x_n)$ et vaut 0 pour $\theta > \max(x_1, \ldots, x_n)$: elle admet un unique point de maximum sur $]0, +\infty[$, égal à

$$\hat{\theta}(x_1, \ldots, x_n) = \max(x_1, \ldots, x_n).$$

(iii) On pose donc $Y_n := \max(X_1, \ldots, X_n)$: on retrouve le modèle statistique et l'estimateur des Exemples 8.23-8.26. Le fait que $(Y_n)_{n \geq 1}$ est un estimateur consistant pour θ a été démontré dans l'Exemple 8.26 ; voir aussi l'Exercice 8.6, où un intervalle de confiance pour θ basé sur Y_n est déterminé.

Le fait que $(Y_n)_{n \geq 1}$ *ne soit pas* un estimateur asymptotiquement normal pour θ a été démontré dans l'Exemple (8.32). Cette affirmation découle facilement du fait que Y_n est toujours inférieur à θ : pour n'importe quelle valeur de $\sigma(\theta)$, $\frac{\sqrt{n}}{\sigma(\theta)}(Y_n - \theta)$ ne prend que des valeurs négatives, donc

$$\lim_{n \to +\infty} \mathrm{P}_\theta \left(\frac{\sqrt{n}}{\sigma(\theta)}(Y_n - \theta) \leq 0 \right) = 1 \neq \mathrm{P}(Z \leq 0) = \frac{1}{2},$$

où $Z \sim \mathrm{N}(0, 1)$. On en conclut que l'on ne peut pas avoir $\frac{\sqrt{n}}{\sigma(\theta)}(Y_n - \theta) \xrightarrow{\mathcal{L}} \mathrm{N}(0, 1)$.

Exercice 8.15. On considère le modèle statistique suivant :

$$f(x; \theta) = \theta x^{\theta - 1} \mathrm{e}^{-x^\theta} \mathbb{1}_{]0, +\infty[}, \qquad \theta \in \Theta =]0, +\infty[.$$

(i) Montrer que les hypothèses (A)–(B) et (D)–(E) sont vérifiées pour $I =]0, +\infty[\setminus \{1\}$.

(ii) Pour $n \geq 1$ et $\underline{x} = (x_1, \ldots, x_n) \in I^n$, on définit $L_n(\underline{x}, \theta) = \frac{1}{n} \sum_{i=1}^n \log f(x_i; \theta)$, comme dans (8.22). Montrer que pour tout $\underline{x} \in I^n$ la fonction $\theta \mapsto L_n(\underline{x}, \theta)$ est strictement concave (c'est-à-dire $\frac{\mathrm{d}^2}{\mathrm{d}\theta^2} L_n(\underline{x},\theta) < 0$ pour tout θ) et que l'on a $\lim_{\theta\downarrow 0} L_n(\underline{x},\theta) = \lim_{\theta\uparrow+\infty} L_n(\underline{x},\theta) = -\infty$. En déduire que l'hypothèse (C) est vérifiée, et que l'estimateur MV est donc consistant.

(iii) Montrer que les hypothèses (F)-(G) sont vérifiées, exceptée (F)-(ii). *On peut cependant montrer que l'estimateur MV est asymptotiquement normal et vérifie la conclusion du Théorème 8.40.* Montrer que la variance asymptotique vaut

$$\sigma(\theta)^2 = \theta^2 \left(1 + \int_0^\infty u\,(\log u)^2 \mathrm{e}^{-u}\mathrm{d}u\right)^{-1}.$$

Solution 8.15. (i) Les hypothèses (A) et (B) sont satisfaites pour $I =]0,+\infty[\,\setminus\{1\}$; notons que l'on peut enlever le point 1 (la raison de ce choix sera claire plus bas) parce que $\mathrm{P}_\theta(X_1 = 1) = 0$, pour tout $\theta > 0$.

Pour l'hypothèse (D), par un calcul direct (notons que $\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^{-x^\theta} = -\theta x^{\theta-1}\mathrm{e}^{-x^\theta}$), on trouve que les fonctions de répartition sont égales à

$$\mathrm{P}_\theta\left(X_1 \le x\right) = \begin{cases} 0 & \text{si } x \le 1\,, \\ 1-\mathrm{e}^{-x^\theta} & \text{si } x > 1\,. \end{cases}$$

Ainsi, comme pour $x > 1$ on a $(1-x)^{-\theta} = (1-x)^{-t}$ si et seulement si $\theta = t$, on obtient que les fonctions de répartition de X_1 (donc les lois) sont différentes pour des valeurs de θ distinctes, ce qui permet de vérifier la condition (D).

Pour la (E), notons que

$$\log f(X_1,t) = \log t + (t-1)\log X_1 - (X_1)^t\,,$$

donc il suffit de montrer que $\mathrm{E}_\theta((\log X_1)^2) < +\infty$ et $\mathrm{E}_\theta((X_1)^{2t}) < +\infty$ pour tout $t > 0$. Notons que l'on a

$$\mathrm{E}_\theta\left((\log X_1)^2\right) = \int_0^\infty \theta x^{\theta-1}(\log x)^2\mathrm{e}^{-x^\theta}\mathrm{d}t\,,\quad \mathrm{E}_\theta\left((X_1)^{2t}\right) = \int_0^\infty \theta x^{2t+\theta-1}\mathrm{e}^{-x^\theta}\mathrm{d}t\,,$$

et les deux intégrales sont finies : il faut simplement voir que $x^{\theta-1}(\log x)^2$ est intégrable en 0 car $\theta-1 > -1$ (de même pour $x^{2t+\theta-1}$), et que $x^{2t+\theta-1}\mathrm{e}^{-x^\theta}$ est intégrable en $+\infty$ car $x^{2t+\theta-1}\mathrm{e}^{-x^\theta} \le \mathrm{e}^{-x^\theta/2}$ pour x suffisamment grand (de même pour $x^{\theta-1}(\log x)^2\mathrm{e}^{-x^\theta}$).

(ii) Tout d'abord, on a

$$L_n(\underline{x},\theta) = \log\theta + (\theta-1)\frac{1}{n}\sum_{i=1}^n \log x_i - \frac{1}{n}\sum_{i=1}^n x_i^\theta\,,$$

donc $\frac{\mathrm{d}}{\mathrm{d}\theta}L_n(\underline{x},\theta) = \frac{1}{\theta} + \frac{1}{n}\sum_{i=1}^n \log x_i - \frac{1}{n}\sum_{i=1}^n x_i^\theta \log x_i$, et

$$\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}L_n(\underline{x},\theta) = -\frac{1}{\theta^2} - \frac{1}{n}\sum_{i=1}^n x_i^\theta(\log x_i)^2 < 0\,.$$

Cela montre que $\theta \mapsto L_n(\underline{x},\theta)$ est strictement concave.

Pour montrer que $\lim_{\theta\downarrow 0} L_n(\underline{x},\theta) = \lim_{\theta\uparrow+\infty} L_n(\underline{x},\theta) = -\infty$, il suffit de montrer que les mêmes limites sont valables pour $\log f(x;\theta) = \log\theta + (\theta-1)\log x - x^\theta$, pour tout $x \in I =]0,+\infty[\,\setminus\{1\}$. Quand $\theta\downarrow 0$, on a $x^\theta \to 1$, donc

$$\lim_{\theta\downarrow 0}\log f(x;\theta)=\lim_{\theta\downarrow 0}\big(\log\theta-\log x-1\big)=-\infty\,.$$

Pour la limite quand $\theta\to+\infty$, on sépare les cas $x\in\,]0,1[$ et $x\in\,]1,+\infty[$. Pour $x\in\,]0,1[$, on a $\lim_{\theta\to+\infty}x^{\theta}=0$, donc

$$\lim_{\theta\uparrow+\infty}\log f(x;\theta)=\lim_{\theta\uparrow+\infty}\big(\log\theta+(\theta-1)\log x\big)=-\infty\,,$$

vu que $\log\theta+(\theta-1)\log x\sim\theta\log x$ quand $t\to+\infty$, avec $\log x<0$. Pour $x\in\,]1,+\infty[$, on a $\log\theta=o(x^{\theta})$ et $(\theta-1)\log x=o(x^{\theta})$ quand $\theta\to+\infty$ ($x^{\theta}=\mathrm{e}^{t\log x}$ croît exponentiellement vite), donc

$$\lim_{\theta\uparrow+\infty}\log f(x;\theta)=\lim_{\theta\uparrow+\infty}-x^{\theta}=-\infty\,.$$

Comme $\theta\mapsto L_n(\underline{x},\theta)$ est strictement concave et que l'on a $\lim_{\theta\downarrow 0}L_n(\underline{x},\theta)=-\infty$, $\lim_{\theta\uparrow+\infty}L_n(\underline{x},\theta)=-\infty$, la fonction $L_n(\underline{x},t)$ doit être strictement croissante jusqu'à un point $\hat{\theta}$ puis strictement décroissante. Cela montre que $\theta\mapsto L_n(\underline{x},\theta)$ admet un unique point de maximum, noté $\hat{\theta}(x_1,\dots,x_n)$, donc (C) est vérifiée. Du Théorème 8.37 on déduit que l'estimateur MV $(Y_n=\hat{\theta}(X_1,\dots,X_n))_{n\geq 1}$ est consistant pour θ.

(iii) Notons que

$$\log f(x;\theta)=\log\theta+(\theta-1)\log x-x^{\theta}\,,$$

donc l'hypothèse (F) (i) est trivialement vérifiée.

Pour (F) (ii), on a

$$B(x,\theta):=\frac{\mathrm{d}^2\log f(x;\theta)}{\mathrm{d}\theta^2}=-\frac{1}{\theta^2}-(\log x)^2x^{\theta}\,,$$

qui n'est pas continue uniformément en $x\in I$; en effet, pour avoir $(\log x)^2|x^{\theta}-x^{t}|<\varepsilon$, il faut prendre $|\theta-t|\leq\delta_{\varepsilon}(x)$, avec $\delta_{\varepsilon}(x)$ qui tend vers 0 quand $x\to+\infty$.

On commence par vérifier (F) (iii) pour $k=1$: notons que (rappelons que $\frac{\mathrm{d}}{\mathrm{d}\theta}x^{\theta-1}=(\log x)x^{\theta-1}$ et $\frac{\mathrm{d}}{\mathrm{d}\theta}x^{\theta}=(\log x)x^{\theta}$)

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\theta}f(x;\theta)&=x^{\theta-1}\mathrm{e}^{-x^{\theta}}+\theta(\log x)x^{\theta-1}\mathrm{e}^{-x^{\theta}}-\theta x^{\theta-1}(\log x)x^{\theta}\mathrm{e}^{-x^{\theta}}\\&=\frac{\mathrm{d}}{\mathrm{d}x}\Big[(\log x)x^{\theta}\mathrm{e}^{-x^{\theta}}\Big]\,,\end{aligned}$$

donc

$$\int_0^{\infty}\frac{\mathrm{d}}{\mathrm{d}\theta}f(x;\theta)\mathrm{d}x=\int_0^{+\infty}\frac{\mathrm{d}}{\mathrm{d}x}\Big[(\log x)x^{\theta}\mathrm{e}^{-x^{\theta}}\Big]\mathrm{d}x=0\,,$$

où on a utilisé le théorème fondamental du calcul intégral pour la dernière égalité. De manière analogue, pour $k=2$, on a

$$\begin{aligned}\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}f(x;\theta)=2(\log x)x^{\theta-1}\mathrm{e}^{-x^{\theta}}&+\theta x^{\theta-1}(\log x)^2-2(\log x)x^{2\theta-1}\mathrm{e}^{-x^{\theta}}\\&-3\theta(\log x)^2x^{2\theta-1}\mathrm{e}^{-x^{\theta}}+\theta(\log x)^2x^{3\theta-1}\mathrm{e}^{-x^{\theta}}\,,\end{aligned}$$

et d'autre part

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[(\log x)^2 x^{\theta} \mathrm{e}^{-x^{\theta}}\right] = 2(\log x)x^{\theta-1}\mathrm{e}^{-x^{\theta}} + \theta x^{\theta-1}(\log x)^2 - \theta(\log x)^2 x^{2\theta-1}\mathrm{e}^{-x^{\theta}}$$
$$\frac{\mathrm{d}}{\mathrm{d}x}\left[(\log x)^2 x^{2\theta} \mathrm{e}^{-x^{\theta}}\right] = 2(\log x)x^{2\theta-1}\mathrm{e}^{-x^{\theta}} + 2\theta(\log x)^2 x^{2\theta-1}\mathrm{e}^{-x^{\theta}} - \theta(\log x)^2 x^{3\theta-1}\mathrm{e}^{-x^{\theta}}$$

dont découle que

$$\frac{\mathrm{d}^2}{\mathrm{d}\theta^2} f(x;\theta) = \frac{\mathrm{d}}{\mathrm{d}x}\left[(\log x)^2 x^{\theta}\mathrm{e}^{-x^{\theta}} - (\log x)^2 x^{2\theta}\mathrm{e}^{-x^{\theta}}\right].$$

On en déduit comme plus haut que $\int_0^\infty \frac{\mathrm{d}^2}{\mathrm{d}\theta^2} f(x;\theta)\mathrm{d}x = 0$.
Il reste à vérifier la condition (G). Rappelons que $B(x,\theta) = -\frac{1}{\theta^2} - (\log x)^2 x^{\theta}$, donc il suffit de vérifier que $(\log X_1)^2 (X_1)^{\theta}$ admet un moment d'ordre deux par rapport à P_θ. Notons que l'on a

$$\mathrm{E}_\theta\left((\log X_1)^4 (X_1)^{2\theta}\right) = \int_0^\infty \theta(\log x)^4 x^{3\theta-1}\mathrm{e}^{-x^{\theta}}\mathrm{d}x.$$

Cette dernière intégrale est finie, parce que $(\log x)^4 x^{3\theta-1}$ est intégrable en 0 (car $3\theta - 1 > -1$) et $(\log x)^4 x^{3\theta-1}\mathrm{e}^{-x^{\theta}}$ est intégrable en $+\infty$ car majoré par $\mathrm{e}^{-x^{\theta}/2}$ pour x suffisamment grand.
La variance asymptotique est donnée par la conclusion (8.31) du Théorème 8.40 :

$$\frac{1}{\sigma^2(\theta)} = -\mathrm{E}_\theta[B(X_1,\theta)] = \frac{1}{\theta^2} + \mathrm{E}_\theta\left((\log X_1)^2 (X_1)^{\theta}\right),$$

avec

$$\mathrm{E}_\theta\left((\log X_1)^2 (X_1)^{\theta}\right) = \int_0^{+\infty} (\log x)^2 x^{\theta}\theta x^{\theta-1}\mathrm{e}^{-x^{\theta}}\mathrm{d}x = \int_0^{+\infty} \frac{1}{\theta^2}(\log u)^2\mathrm{e}^{-u}\mathrm{d}u.$$

où la dernière égalité est due à un changement de variable $u = x^{\theta}$. Pour finir, on a obtenu

$$\frac{1}{\sigma^2(\theta)} = -\mathrm{E}_\theta[B(X_1,\theta)] = \frac{1}{\theta^2}\left(1 + \int_0^{+\infty} (\log u)^2\mathrm{e}^{-u}\mathrm{d}u\right),$$

dont on déduit le résultat voulu.

Chapitre 9
Simulations

Rappelons que dans les programmes qui suivent, par souci de brévité, *nous n'écrirons jamais le préambule lié à l'utilisation des différentes bibliothèques nécessaires*. Écrivons-le ici :

```
import numpy.random as rd
from math import *
import matplotlib.pyplot as plt
import time
```

Nous utiliserons la bibliothèque `matplotlib.pyplot` pour dessiner le graphe de fonctions et des histogrammes et la bibliothèque `time` pour déterminer le temps d'exécution d'un programme.

9.2 Simulation de variables aléatoires discrètes

Exercice 9.1. On considère l'expérience qui consiste à lancer deux dés réguliers à six faces et à faire la somme des deux résultats obtenus. Écrire un algorithme qui permet de simuler cette expérience.

Solution 9.1. Pour le programme, il suffit de se rappeler que si $U \sim \mathrm{U}(0,1)$, alors $\lfloor 6U \rfloor + 1$ est une variable aléatoire de loi uniforme discrète sur $\{1,2,3,4,5,6\}$, ce qui permet de simuler un lancer de dé régulier à six faces. Il suffit de répéter deux fois cette opération (un deuxième appel à la fonction `random` renverra une valeur indépendante) et de sommer les résultats.

On donne ci-dessous une fonction Python qui simule la somme des résultats de deux dés régulier.

```
def DeuxDes():
    x=int(6*rd.random())+1
    y=int(6*rd.random())+1
    return x+y
```

Exercice 9.2. Écrire un algorithme qui permet de simuler une variable aléatoire à valeurs dans $\{-1,0,1,2\}$ de densité discrète donnée par $\mathrm{p}_x(-1) = \frac{1}{2}$ et $\mathrm{p}_X(0) = \mathrm{p}_X(1) = \mathrm{p}_X(2) = \frac{1}{6}$.

Solution 9.2. Il suffit d'utiliser la méthode décrite dans la Section 9.2 :

- on divise l'intervalle $[0,1[$ en sous-intervalles $[0,\frac{1}{2}[$, $[\frac{1}{2},\frac{2}{3}[$, $[\frac{2}{3},\frac{5}{6}[$ et $[\frac{5}{6},1[$, de longueurs respectives $\frac{1}{2}$, $\frac{1}{6}$, $\frac{1}{6}$ et $\frac{1}{6}$;
- on simule une variable aléatoire $U \sim \mathrm{U}(0,1)$ grâce à la fonction `random`;
- on détermine dans quel intervalle tombe U, si elle tombe dans le premier on renvoie -1, si elle tombe dans le deuxième on renvoie 0, si elle tombe dans le troisième on renvoie 1, si elle tombe dans le dernier on renvoie 2.

Un algorithme possible écrit en Python est donné ci-dessous.

```
def Variable():
    U=rd.random()
    if U<1/2:
        return -1
    elif U<2/3:
        return 0
    elif U<5/6:
        return 1
    else:
        return 2
```

9.3 Simulation de variables aléatoires à densité

Exercice 9.3 (Weibull). Soit X une variable aléatoire réelle à densité, de densité donnée par

$$f(x) = \alpha x^{\alpha-1} \mathrm{e}^{-x^\alpha} \mathbb{1}_{]0,\infty[}(x),$$

où $\alpha > 0$ est un paramètre fixé. Une telle variable aléatoire est dite *de Weibull* de paramètre α.

(i) Vérifier que f_X est effectivement une densité et calculer la fonction de répartition de X.

(ii) En déduire une méthode pour simuler une variable aléatoire de Weibull de paramètre α.

Solution 9.3. (i) On a

$$\int_{-\infty}^{+\infty} f_X(x)\mathrm{d}x = \int_0^\infty \alpha x^{\alpha-1} \mathrm{e}^{-x^\alpha} \mathrm{d}x = \Big[-\mathrm{e}^{-x^\alpha} \Big]_0^\infty = 0 - (-1) = 1\,,$$

donc f_X est effectivement une densité.
La fonction de répartition de X vaut $F_X(t) = 0$ si $t \leq 0$ et pour $t > 0$

$$F_X(t) = \int_0^y \alpha x^{\alpha-1} \mathrm{e}^{-x^\alpha} \mathrm{d}x = \Big[-\mathrm{e}^{-x^\alpha} \Big]_0^t = 1 - \mathrm{e}^{-t^\alpha}\,.$$

(ii) On utilise la méthode d'inversion : la fonction F_X est bijective de $]0,+\infty[$ sur $]0,1[$. On peut calculer son inverse, en résolvant $F_X(t) = u$, pour tout $u \in]0,1[$ fixé :

$$1 - \mathrm{e}^{-t^\alpha} = u \iff t^\alpha = -\log(1-u) \iff t = \big(-\log(1-u)\big)^{1/\alpha}\,.$$

Grâce à la méthode d'inversion, si $U \sim \mathrm{U}(0,1)$, alors $F_X^{-1}(U) = \big(-\log(1-U)\big)^{1/\alpha}$ est de loi de Weibull de paramètre α. Observons que l'on a aussi $1-U \sim \mathrm{U}(0,1)$, donc $\big(-\log(U)\big)^{1/\alpha}$ est aussi de loi de Weibull de paramètre α.
Un algorithme Python pour simuler une variable aléatoire de Weibull de paramètre α est donné ci-dessous.

```
def Weibull(a):
    return (-log(rd.random()))**(1/a)
```

Exercice 9.4. On rappelle que si X,Y sont deux variables aléatoires indépendantes de loi $\mathrm{N}(0,1)$ alors $X/Y \sim \mathrm{Cauchy}(0,1)$, d'après l'Exemple 6.59. En déduire un algorithme qui permet de simuler une variable aléatoire de loi de Cauchy standard.

Solution 9.4. Il suffit donc de simuler deux variables $\mathrm{N}(0,1)$ indépendantes et de calculer le rapport.

On peut utiliser la méthode de Box–Müller de la Section 9.3.2. Rappelons que si $U \sim \mathrm{U}(0,1)$, $V \sim \mathrm{U}(0,1)$ sont indépendantes, alors les variables aléatoires $X = \sqrt{-2\log V} \times \cos(2\pi U)$ et $Y = \sqrt{-2\log V} \times \sin(2\pi U)$ sont indépendantes de loi $\mathrm{N}(0,1)$: le rapport est donc donné par $\cos(2\pi U)/\sin(2\pi U)$. Un algorithme Python pour simuler une variable aléatoire de Cauchy standard basée sur cette méthode est donc le suivant (noter que l'on n'a besoin que d'une seule variable $U \sim \mathrm{U}(0,1)$).

```
def Cauchy1():
        U=rd.random()
    return cos(2*pi*U)/sin(2*pi*U)
```

Une autre méthode est d'utiliser la fonction `normal` de la bibliothèque `numpy.random` pour simuler deux variables aléatoires X et Y indépendantes de loi $\mathrm{N}(0,1)$, pour ensuite faire le rapport X/Y. Un algorithme Python pour simuler une variable aléatoire de Cauchy standard basée sur cette méthode est donc le suivant.

```
def Cauchy2():
    return rd.normal(0,1)/rd.normal(0,1)
```

Exercice 9.5 (Gamma$(\alpha,1)$, $\alpha < 1$). On propose maintenant une méthode pour simuler une variable aléatoire $X \sim \mathrm{Gamma}(\alpha,1)$, avec $\alpha \in]0,1[$.

(i) Montrer que pour tout $x \in]0,+\infty[$ on a

$$g(x) := x^{\alpha-1}\mathrm{e}^{-x} \leq C_\alpha x^{\alpha-1}\mathrm{e}^{-x^\alpha} =: h(x)\,, \quad \text{où } C_\alpha = \exp\left((1-\alpha)\alpha^{\alpha/(1-\alpha)}\right).$$

(ii) En observant que $\frac{\alpha}{C_\alpha}h(x)\mathbb{1}_{]0,\infty[}(x)$ est la densité d'une variable aléatoire de Weibull (voir l'Exercice 9.3), appliquer la méthode de rejet pour simuler une variable aléatoire de loi $\mathrm{Gamma}(\alpha,1)$.

(iii) Donner la probabilité de rejet dans cette méthode.

Solution 9.5. (i) Considérons la fonction $k(x) := x^\alpha - x$ pour $x \geq 0$. Sa dérivée est $k'(x) = \alpha x^{\alpha-1} - 1$, qui est positif si $x \leq \alpha^{1/(1-\alpha)}$ et négatif si $x \geq \alpha^{1/(1-\alpha)}$. On obtient donc que $k(x)$ est maximale pour $x = \alpha^{1/(1-\alpha)}$ et que le maximum vaut $(1-\alpha)\alpha^{\alpha/(1-\alpha)}$. Comme $\frac{g(x)}{x^{\alpha-1}\mathrm{e}^{-x^\alpha}} = \mathrm{e}^{k(x)}$, on obtient l'inégalité voulue.

(ii) Appliquons la méthode de rejet décrite dans la Section 9.3.3 :

- on simule des variables aléatoires indépendantes Z_i, U_i de lois respectives de Weibull de paramètre α (grâce à l'Exercice 9.3, pour $V \sim \mathrm{U}(0,1)$, $(-\log V)^{1/\alpha}$ est de loi de Weibull de paramètre α) et $\mathrm{U}(0,1)$;
- on rejette (Z_i, U_i) jusqu'au premier instant T où $U_T\, h(Z_T) \leq g(Z_T)$ ou de manière équivalente $C_\alpha U_T \leq \exp(Z_T^\alpha - Z_T)$;
- on renvoie la valeur de Z_T, qui est une variable aléatoire à densité, de densité donnée par $f = \frac{1}{A}g$ avec $A = \Gamma(\alpha) = \int_0^\infty g(x)\mathrm{d}x$.

Le programme Python peut s'écrire de la façon suivante.

```
def Gamma1(a):
    C=exp((1-a)*a**(1/(1-a)))
    Z=(-log(rd.random()))**(1/a)
    while C*rd.random()>exp(Z**a-Z):
        Z=(-log(rd.random()))**(1/a)
    return X
```

(iii) La probabilité d'acceptation dans cette méthode est donnée par le rapport

$$p = \frac{\int_0^\infty g(x)\mathrm{d}x}{\int_0^\infty h(x)\mathrm{d}x} = \frac{\Gamma(\alpha)}{C_\alpha/\alpha} = \alpha\Gamma(\alpha)\exp\Big(-(1-\alpha)\alpha^{\alpha/(1-\alpha)}\Big),$$

où on a utilisé que $\int_0^\infty h(x)\mathrm{d}x = \frac{C_\alpha}{\alpha}$ car $\frac{\alpha}{C_\alpha}h(x)\mathbb{1}_{]0,\infty[}(x)$ est une densité (d'une variable de Weibull). La probabilité de rejet est donc $1-p = 1-\alpha\Gamma(\alpha)\exp(-(1-\alpha)\alpha^{\alpha/(1-\alpha)})$.

Exercice 9.6 (Méthode polaire de Marsaglia pour des variables normales *). En pratique, le méthode de Box–Müller n'est pas celle utilisé pour simuler une variable aléatoire normale, parce que le calcul d'un sinus ou d'un cosinus possède un coût computationnel élevé. Donnons maintenant une méthode alternative, appelée *méthode polaire de Marsaglia* : il s'agit d'une variation de la méthode de Box–Müller qui permet d'éviter l'utilisation de fonctions trigonométriques. Soit (X,Y) un vecteur aléatoire de loi uniforme continue sur le disque unitaire $\mathrm{D}_1 = \{(x,y), x^2+y^2<1\}$.

(i) Écrire le vecteur (X,Y) en coordonnées polaires, c'est-à-dire soient (R,Θ) tels que $X = R\cos\Theta$ et $Y = R\sin\Theta$, avec $R\in[0,1[$ et $\Theta\in[0,2\pi[$. Montrer que R et Θ sont des variables aléatoires indépendantes, avec R de densité $f_R(r) = 2r\mathbb{1}_{(0,1)}(r)$ et $\Theta\sim\mathrm{U}(0,2\pi)$.

(ii) Montrer que $W := \sqrt{-4\log R}$ a pour densité $f_W(w) = w\mathrm{e}^{-w^2/2}\mathbb{1}_{(0,\infty)}(w)$. En utilisant la Proposition 9.5, montrer que $\sqrt{-4\log R}\,\frac{X}{R}$ e $\sqrt{-4\log R}\,\frac{Y}{R}$ sont des variables aléatoires indépendantes de loi normale standard $\mathrm{N}(0,1)$.

(iii) En se rappelant comment on peut simuler le vecteur aléatoire (X,Y) grâce à la méthode de rejet (Exemple 9.15), en déduire un algorithme pour simuler une variable aléatoire (en fait deux variables aléatoires indépendantes...) de loi $\mathrm{N}(0,1)$ *sans utiliser de sinus ou cosinus* (noter que $R^2 = X^2+Y^2$).

Solution 9.6. (i) Le vecteur aléatoire (X,Y) est absolument continu, à valeurs dans $\mathrm{D}_1 := \{(x,y)\,;\,x^2+y^2<1\}$, de densité $\frac{1}{\pi}\mathbb{1}_{\mathrm{D}_1}(x,y)$. D'après la Proposition 6.60, on en déduit que le vecteur (R,Θ) est absolument continu, parce que $\psi : (r,t)\mapsto(x,y) = (r\cos t, r\sin t)$ est un difféomorphisme de $V =]0,1[\times]0,2\pi[$ dans $U = \mathrm{D}_1\setminus C$ où $C = \{(x,y)\in\mathrm{D}_1, x\geq 0, y=0\}$, avec $|\det J_\psi(r,t)| = r$; dans la Proposition 6.60 on utilise son inverse $\varphi = \psi^{-1} : U\to V$. La densité jointe de (R,Θ) est donnée par

$$\begin{aligned} f_{R,\Theta}(r,t) &= f_{X,Y}(\psi(r,t))|\det J_\psi(r,t)|\mathbb{1}_{]0,1[\times]0,2\pi[}(r,t) \\ &= 2r\mathbb{1}_{]0,1[}(r)\times\frac{1}{2\pi}\mathbb{1}_{]0,2\pi[}(t). \end{aligned}$$

On en déduit que R et Θ sont indépendantes avec R de densité $r\mathbb{1}_{]0,1[}(r)$ et $\Theta\sim\mathrm{U}(0,2\pi)$.

(ii) Calculons la fonction de répartition de $W := \sqrt{-4\log R}$. Tout d'abord, W est strictement positive, donc $F_W(t) = 0$ si $t\leq 0$. Pour $t>0$,

$$\begin{aligned} F_W(t) &= \mathrm{P}\big(\sqrt{-4\log R}\leq t\big) = \mathrm{P}\Big(R\geq \mathrm{e}^{-t^2/4}\Big) \\ &= \int_{\mathrm{e}^{-t^2/4}}^1 2r\mathrm{d}r = 1-\mathrm{e}^{-t^2/2}. \end{aligned}$$

On a donc obtenu que F_W est continue et $\mathcal{C}^1$ par morceaux, donc grâce à la Proposition 6.17 la variable aléatoire W est absolument continue, de densité

$$f_W(w) = F'_W(w) = we^{-t^2/2}\mathbb{1}_{]0,\infty[}(w).$$

Grâce à la Proposition 9.5, comme W et Θ sont indépendantes avec W absolument continue de densité $we^{-t^2/2}\mathbb{1}_{]0,\infty[}(w)$ et $\Theta \sim \mathrm{U}(0,2\pi)$, on en déduit que $Z_1 = W\cos(\Theta) = \sqrt{-4\log R}\,\frac{X}{R}$ et $Z_2 = W\sin(\Theta) = \sqrt{-4\log R}\,\frac{Y}{R}$ sont indépendantes, de loi $\mathrm{N}(0,1)$.

(iii) On utilise la méthode de rejet de l'Exemple 9.15 pour simuler un vecteur aléatoire (X,Y) de loi uniforme continue sur D_1 ; puis on calcule $S = X^2+Y^2$ ($R = \sqrt{S}$) et on renvoie $X\sqrt{-2(\log S)/S}$, qui est de loi $\mathrm{N}(0,1)$. Un programme Python correspondant est donné ci-dessous.

```python
def Normale1():
    X=2*rd.random()-1
    Y=2*rd.random()-1
    while X**2+Y**2>1:
        X=2*rd.random()-1
        Y=2*rd.random()-1
    S=X**2+Y**2
    return sqrt(-2*log(S)/S)*X
```

9.6 Exercices récapitulatifs

Note. Pour déterminer le temps d'exécution d'un programme en Python, on peut utiliser la fonction `time` de la bibliothèque `time`, qui permet d'obtenir le temps (en seconde) au début et à la fin de l'exécution du programme. Par exemple, pour connaître le temps utilisé par l'ordinateur pour générer 10^6 variables aléatoires i.i.d. de loi $\mathrm{U}(0,1)$, on peut utiliser le code suivant (après avoir ajouté `import time` dans le préambule) :

```python
t=time.time()
for k in range(1000000):
    rd.random()
print(time.time()-t)
```

Exercice 9.7. Écrire un algorithme qui permet de simuler une variable aléatoire X de densité discrète $\mathrm{p}_X(i) = \frac{1}{\log 10}\log(1+\frac{1}{i})$ pour $i \in \{1,2,\dots,9\}$. On vérifiera par ailleurs qu'il s'agit bien d'une densité discrète.

[*Sugg.* Observer que si $U \sim \mathrm{U}(0,1)$, alors $\lfloor 10^U \rfloor$ possède la densité discrète voulue.]

Solution 9.7. Tout d'abord, vérifions qu'il s'agit d'une densité discrète. On a en effet $\mathrm{p}_X(i) \geq 0$ pour tout $i \in \{1,2,\dots,9\}$, et

$$\begin{aligned}\sum_{i=1}^{9}\mathrm{p}_X(i) &= \frac{1}{\log 10}\sum_{i=1}^{9}\log\left(\frac{i+1}{i}\right)\\ &= \frac{1}{\log 10}\sum_{i=1}^{9}\big(\log(i+1)-\log(i)\big) = \frac{1}{\log 10}\Big(\log(10)-\log 1\Big) = 1,\end{aligned}$$

où on a utilisé une somme télescopique.

Observons que si $U \sim \mathrm{U}(0,1)$, alors en posant $X := \lfloor 10^U \rfloor$, on a $X \in \{1,\dots,9\}$ (avec probabilité 1) et pour tout $i \in \{1,2,\dots,9\}$

$$\begin{aligned}\mathrm{P}(X=i)&=\mathrm{P}\left(10^U\in[i,i+1[\right)=\mathrm{P}\left(U\in\left[\frac{1}{\log 10}\log(i),\frac{1}{\log 10}\log(i+1)\right[\right)\\&=\frac{1}{\log 10}\big(\log(i+1)-\log(i)\big)=\frac{1}{\log 10}\log\left(1+\frac{1}{i}\right).\end{aligned}$$

Ainsi, pour simuler la variable aléatoire X, il suffit de simuler une variable aléatoire $U\sim\mathrm{U}(0,1)$ et calculer $\lfloor 10^U\rfloor$. Un programme Python correspondant est donné ci-dessous.

```
def X():
    return int(10**(rd.random()))
```

Exercice 9.8. Écrire un algorithme qui permet de simuler les variables aléatoires suivantes :

(i) Y de densité discrète $\mathrm{p}_Y(1)=\mathrm{p}_Y(0)=\mathrm{p}_Y(-1)=\frac{1}{3}$;

(ii) Z de densité discrète $\mathrm{p}_Z(k)=\frac{1}{3}\cdot 2^{-|k|}$ pour $k\in\mathbb{Z}$.

[*Sugg.* Pour simuler Z, il peut être utile de générer une variable aléatoire géométrique de paramètre p bien choisi, et la combiner à la variable aléatoire Y.]

Solution 9.8. (i) On peut observer que Y est de loi uniforme discrète sur $\{-1,0,1\}$. Si $U\sim\mathrm{U}(0,1)$, on a vu dans la Section 9.2 que $\lfloor 3U\rfloor$ est de loi uniforme discrète sur $\{0,1,2\}$ et il est facile de vérifier que $\lfloor 3U\rfloor-1$ est de loi uniforme discrète sur $\{-1,0,1\}$. Pour simuler Y, il suffit donc de simuler une variable $U\sim\mathrm{U}(0,1)$, de prendre la partie entière de $3U$ et de soustraire 1. Un programme Python correspondant est donné ci-dessous.

```
def Y():
    return int(3*rd.random())-1
```

(ii) Si $W\sim\text{Géom}(\frac{1}{2})$, on a $\mathrm{P}(W=k)=2^{-k}$ pour tout $k\in\mathbb{N}^*$. Si W et Y sont indépendantes, alors $Y\times W$ est une variable aléatoire à valeurs dans $\mathbb{Z}$: on a $\mathrm{P}(YW=0)=\mathrm{P}(Y=0)=\frac{1}{3}$ et pour $k\in\mathbb{N}$ on a $\mathrm{P}(YW=k)=\mathrm{P}(Y=1,W=k)=\frac{1}{3}2^{-k}$ et $\mathrm{P}(YW=-k)=\mathrm{P}(Y=-1,W=k)=\frac{1}{3}2^{-k}$. Pour résumer, on a que YW possède la même densité discrète que Z : pour simuler Z, il suffit donc de simuler une variable aléatoire $W\sim\text{Géom}(\frac{1}{2})$ (voir la Section 9.2, ou mieux, déduire de l'Observation 9.4 que $\lfloor-\log(U)/\log(2)\rfloor+1$ est de loi $\text{Géom}(\frac{1}{2})$) et de la multiplier par Y. Un programme Python correspondant est donné ci-dessous.

```
def Z():
    W=int(-log(rd.random())/log(2))+1
    Y=int(3*rd.random())-1
    return Y*W
```

Exercice 9.9. Comparer les temps pris par l'ordinateur pour simuler 10^6 variables aléatoires i.i.d. de loi $\mathrm{Bin}(n,p)$ pour $n=2$, $n=10$ et $n=50$ dans les cas $p=0,2$ et $p=0,7$:

(i) en utilisant la méthode proposée dans la Section 9.2 (voir page 376) ;

(ii) en utilisant la fonction `binomial` de la bibliothèque `numpy.random`.

[*Sugg.* On pourra utiliser la fonction `time` de la bibliothèque *time* pour avoir le temps au début et à la fin de l'exécution de l'algorithme.]

Solution 9.9. (i) Rappelons le programme proposé dans la Section 9.2 pour simuler une variable aléatoire de loi $\mathrm{Bin}(n,p)$ (on rappelle aussi le programme pour simuler une variable de Bernoulli)

```
def Bern(p):
    if rd.random()<p:
        return 1
    else:
        return 0
def Binom(n,p):
    X=0
    for i in range(n):
        X=X+Bern(p)
    return X
```

Répétons maintenant 10^6 appels à `Binom` pour $p = 0.2$ et $p = 0.7$ et différentes valeurs de n.

```
for p in [0.2,0.7]:
  for n in [2,10,50]:
    t=time.time()
    for k in range(10**6):
      Binom(n,p)
    t=time.time()-t
    print("Pour n={} et p={} le temps d'execution
                                        vaut {}".format(n,p,t))
```

Le programme ci-dessus a renvoyé :

```
Pour n=2 et p=0.2 le temps d'execution vaut 2.862373113632202
Pour n=10 et p=0.2 le temps d'execution vaut 8.448970079421997
Pour n=50 et p=0.2 le temps d'execution vaut 36.73152470588684
Pour n=2 et p=0.7 le temps d'execution vaut 2.8175301551818848
Pour n=10 et p=0.7 le temps d'execution vaut 8.374844312667847
Pour n=50 et p=0.7 le temps d'execution vaut 36.88935399055481
```

(ii) On répète maintenant le même programme, en remplaçant l'appel à `Binom` par un appel à `rd.binomial`.

```
for p in [0.2,0.7]:
  for n in [2,10,50]:
    t=time.time()
    for k in range(10**6):
      rd.binomial(n,p)
    t=time.time()-t
    print("Pour n={} et p={} le temps d'execution
                                        vaut {}".format(n,p,t))
```

Le programme ci-dessus a renvoyé :

```
Pour n=2 et p=0.2 le temps d'execution vaut 5.242967128753662
Pour n=10 et p=0.2 le temps d'execution vaut 5.101454734802246
Pour n=50 et p=0.2 le temps d'execution vaut 5.204687833786011
Pour n=2 et p=0.7 le temps d'execution vaut 5.12961483001709
Pour n=10 et p=0.7 le temps d'execution vaut 5.199053049087524
Pour n=50 et p=0.7 le temps d'execution vaut 5.229662179946899
```

On voit donc que pour des petites valeurs de n il n'y a pas de grande différence entre le programme `Binom` proposé dans la Section 9.2 et la fonction `binomial` de la bibliothèque `numpy.random` (le programme `Binom` semble en fait être plus rapide !). En revanche, lorsque n devient un peu plus grand, la fonction `binomial` de la bibliothèque `numpy.random` devient *beaucoup* plus rapide.

Exercice 9.10. En utilisant la méthode d'inversion, écrire un algorithme qui permet de simuler les variables aléatoires X dont les densités sont les suivantes :

(i) $f(x) = \alpha x^{-(1+\alpha)} \mathbb{1}_{]1,+\infty[}(x)$, avec $\alpha > 0$;

(ii) $f(x) = \mathrm{e}^{-x} \mathrm{e}^{-\mathrm{e}^{-x}}$ (pour tout $x \in \mathbb{R}$) ;

(iii) $f(x) = \frac{1}{\pi\sqrt{x(1-x)}} \mathbb{1}_{]0,1[}(x)$; on rappelle que $\frac{\mathrm{d}}{\mathrm{d}x} \arcsin(x) = \frac{1}{\sqrt{1-x^2}}$.

Solution 9.10. Dans chacun des cas, on doit calculer la fonction de répartition $F_X(x)$ et trouver son inverse (ou pseudo-inverse) $F_X^{-1} :]0,1[\to \mathbb{R}$.

(i) Un calcul rapide donne $F_X(x) = \int_1^x \alpha t^{-(1+\alpha)} \mathrm{d}t = 1 - x^{-\alpha}$ pour $x > 1$ alors que $F_X(x) = 0$ pour $x \leq 1$. Donc F_X est bijective de $]1,+\infty[$ dans $]0,1[$ et pour $u \in]0,1[$

$$F_X(x) = u \quad \Longleftrightarrow \quad t^{-\alpha} = 1 - u \quad \Longleftrightarrow \quad t = (1-u)^{-1/a}.$$

Donc l'inverse de F_X est $F_X^{-1}(u) = (1-u)^{-1/\alpha}$. Si $U \sim \mathrm{U}(0,1)$, en se rappelant que l'on a aussi $1 - U \sim \mathrm{U}(0,1)$, d'après la méthode d'inversion on en déduit que $U^{-1/\alpha}$ est une variable aléatoire de fonction de répartition F_X.

Ainsi, un programme Python qui permet de simuler une variable aléatoire de densité $f(x) = \alpha x^{-(1+\alpha)} \mathbb{1}_{]1,\infty[}(x)$ (une telle variable aléatoire est dite de *Pareto*) peut s'écrire de la façon suivante :

```
def Pareto(a):
    return (rd.random())**(-1/a)
```

(ii) Avec un calcul rapide, on obtient $F_X(x) = \int_{-\infty}^x \mathrm{e}^{-t} \mathrm{e}^{-\mathrm{e}^{-t}} \mathrm{d}t = \mathrm{e}^{-\mathrm{e}^{-x}}$ pour tout $x \in \mathbb{R}$. Donc F_X est bijective de $\mathbb{R}$ dans $]0,1[$ et on peut calculer son inverse, en résolvant $F_X(x) = \mathrm{e}^{-\mathrm{e}^{-x}} = u$ pour $u \in]0,1[$: on obtient $F_X^{-1}(u) = -\ln(-\ln u)$.

Ainsi, un programme Python qui permet de simuler une variable aléatoire de densité $f(x) = \mathrm{e}^{-x} \mathrm{e}^{-\mathrm{e}^{-x}}$ (une telle variable aléatoire est dite de *Gumbel*) peut s'écrire de la façon suivante :

```
def Gumbel():
    return -log(-log(rd.random()))
```

(iii) Avec un calcul rapide, on obtient, pour tout $x \in]0,1[$

$$\begin{aligned} F_X(x) = \int_0^x \frac{1}{\pi\sqrt{t(1-t)}} \mathrm{d}t &= \int_0^{\sqrt{x}} \frac{2}{\pi\sqrt{1-u^2}} \mathrm{d}u \\ &= \left[\frac{2}{\pi} \arcsin(u)\right]_0^{\sqrt{x}} = \frac{2}{\pi} \arcsin(\sqrt{x}), \end{aligned}$$

où on a utilisé un changement de variable $\sqrt{t} = u$ pour faire apparaître la dérivée de arcsin. D'autre part, on a $F_X(x) = 0$ pour $x \leq 0$ et $F_X(x) = 1$ pour $x \geq 1$.

On peut maintenant trouver l'inverse de F_X : en résolvant $F_X(x) = u$ pour $u \in]0,1[$, on obtient $F_X^{-1}(u) = \sin(\frac{\pi}{2}u)^2$ pour $u \in]0,1[$.

Ainsi, un programme Python qui permet de simuler une variable aléatoire de densité $f(x) = \frac{1}{\pi\sqrt{x(1-x)}} \mathbb{1}_{]0,1[}(x)$ (une telle variable aléatoire est dite de l'*Arcsinus*) peut s'écrire de la façon suivante :

```
def Arcsinus():
    return sin(pi*rd.random()/2)**2
```

Exercice 9.11. Trouver une méthode pour simuler une variable aléatoire absolument continue, de densité donnée par $f(x) = \frac{1}{2}\mathrm{e}^{-|x|}$ pour $x \in \mathbb{R}$.

[*Sugg.* Montrer que si $Y \sim \mathrm{Exp}(1)$ et $Z \sim \mathrm{Bern}(\frac{1}{2})$ sont indépendantes, alors $ZY - (1-Z)Y$ possède la densité voulue.]

Solution 9.11. Utilisons la suggestion : l'idée est que la variable considérée est une version « symétrisée » d'une variable aléatoire Exp(1). Soient $Y \sim \mathrm{Exp}(1)$ et $Z \sim \mathrm{Bern}(\frac{1}{2})$ deux variables aléatoires indépendantes et posons $W = ZY - (1-Z)Y$.

En utilisant l'Observation 6.26 on peut déterminer la loi de W. Soit $h : \mathbb{R} \to \mathbb{R}$ une fonction continue par morceaux et bornée (quelconque). Calculons $\mathrm{E}(h(W))$ en séparant l'espérance selon que $Z = 1$ (auquel cas $W = Y$) ou que $Z = 0$ (auquel cas $W = -Y$) :

$$\mathrm{E}(h(W)) = \mathrm{E}\left(h(Y)\mathbb{1}_{\{Z=1\}}\right) + \mathrm{E}\left(h(-Y)\mathbb{1}_{\{Z=0\}}\right) = \frac{1}{2}\mathrm{E}\left(h(Y)\right) + \frac{1}{2}\mathrm{E}\left(h(-Y)\right),$$

où on a utilisé l'indépendance de Y et Z pour la deuxième égalité. Maintenant, en utilisant la formule de transfert (Proposition 6.21) appliquée à Y, on obtient

$$\begin{aligned}\mathrm{E}(h(W)) &= \frac{1}{2}\int_0^\infty h(y)\mathrm{e}^{-y}\mathrm{d}y + \frac{1}{2}\int_0^\infty h(-y)\mathrm{e}^{-y}\mathrm{d}y \\ &= \frac{1}{2}\int_0^\infty h(y)\mathrm{e}^{-y}\mathrm{d}y + \frac{1}{2}\int_{-\infty}^0 h(y)\mathrm{e}^{y}\mathrm{d}y = \int_{-\infty}^{+\infty} h(y)\frac{1}{2}\mathrm{e}^{-|y|}\mathrm{d}y.\end{aligned}$$

Cette identité étant valide pour toute fonction h continue par morceaux et bornée, d'après l'Observation 6.26 on obtient que W est absolument continue, de densité donnée par $f(x) = \frac{1}{2}\mathrm{e}^{-|x|}$ pour $x \in \mathbb{R}$.

Ainsi, un programme qui permet de simuler une variable aléatoire de densité $f(x) = \frac{1}{2}\mathrm{e}^{-|x|}$ peut être écrit de la manière suivante (on renvoie à la Section 9.3.1 pour la méthode d'inversion pour simuler une variable aléatoire Exp(1)).

```
def ExpoSymetrique():
    Y=-log(rd.random())
    if rd.random()<0.5:
        return Y
    else:
        return -Y
```

On peut aussi noter que, d'après l'Exercice 6.29, si $X \sim \mathrm{Exp}(1)$ et $Y \sim \mathrm{Exp}(1)$ sont indépendantes, alors $Y - X$ a pour densité $f(x) = \frac{1}{2}\mathrm{e}^{-|x|}$. Ainsi, un programme alternatif qui permet de simuler une telle variable aléatoire est le suivant.

```
def ExpoSymetrique2():
    return log(rd.random())-log(rd.random())
```

Exercice 9.12. Proposer deux méthodes pour générer une variable aléatoire khi-deux à n degrés de liberté, $X \sim \chi^2(n)$ avec $n \in \mathbb{N}^*$, c'est-à-dire $X \sim \mathrm{Gamma}\left(\frac{n}{2}, \frac{1}{2}\right)$ avec $n \in \mathbb{N}^*$, introduite dans la Section 6.3.5.

Solution 9.12. Pour une première méthode, on peut utiliser le fait que si $Z_1, Z_2, \ldots, Z_n$ sont des variables aléatoires indépendantes de loi $\mathrm{N}(0,1)$, alors $X = Z_1^2 + \cdots + Z_n^2$ est de loi $\chi^2(n)$, voir la Section 6.3.5 (en particulier (6.51)). Ainsi, un programme qui permet de simuler une variable aléatoire $\chi^2(n)$ peut s'écrire de la manière suivante :

```
def Chi2(n):
    X=0
```

```
    for i in range(n):
        X+=(rd.normal(0,1))**2
    return X
```

Pour une deuxième méthode, on peut utiliser le fait que grâce à l'identité (6.40), si $Y \sim \mathrm{Gamma}(\frac{n}{2},1)$ alors $2Y \sim \mathrm{Gamma}(\frac{n}{2},\frac{1}{2})$. Pour simuler Y, on peut utiliser la méthode de rejet de l'Exemple 9.14 si $n \geq 3$ et de l'Exercice 9.5 si $n=1$; pour $n=2$ on utilise que $Y \sim \mathrm{Exp}(1)$. Un programme alternatif à celui proposé ci-dessus pour simuler une variable aléatoire $\chi^2(n)$, peut donc s'écrire de la façon suivante (on renvoie à la solution de l'Exercice 9.5 et à l'Exemple 9.14 ; rappelons de plus que si $U \sim \mathrm{U}(0,1)$ alors $-\log U \sim \mathrm{Exp}(1)$, donc $-2\log U \sim \mathrm{Exp}(\frac{1}{2})$).

```
def Chi2alt(n):
    if n==1:
        C=exp(1/4)
        Y=(-log(rd.random()))**2
        while C*rd.random()>exp(Y**(0.5)-Y):
            X=(-log(rd.random()))**2
        return 2*Y
    elif n==2:
        return -2*log(rd.random())
    else:
        a=n/2
        C=(2*(a-1))**(a-1)*exp(1-a)
        Y=-2*log(rd.random())
        while C*rd.random()>Y**(a-1)*exp(-0.5*Y):
            Y=-2*log(rd.random())
        return 2*Y
```

Exercice 9.13. Proposer deux méthodes pour générer une variable aléatoire de loi $\mathrm{Beta}(a,b)$ avec $a,b \in [1,\infty[$, introduite dans la Section 6.6.2, voir (6.92).
[*Sugg.* Utiliser d'une part une méthode de rejet, d'autre part l'Exercice 6.17.]

Solution 9.13. Commençons avec la méthode utilisant l'Exercice 6.17. Si $X \sim \mathrm{Gamma}(a,1)$ et $Y \sim \mathrm{Gamma}(b,1)$ sont deux variables aléatoires indépendantes, alors $\frac{X}{X+Y}$ est de loi $\mathrm{Beta}(a,b)$. En utilisant la fonction `gamma` de la bibliothèque `numpy.random`, un programme qui permet de simuler une variable aléatoire $\mathrm{Beta}(a,b)$ peut donc s'écrire de la façon suivante :

```
def Beta(a,b):
    X=rd.gamma(a,1)
    Y=rd.gamma(b,1)
    return X/(X+Y)
```

Une autre méthode consiste à observer que pour $a,b \geq 1$ on a

$$x^{a-1}(1-x)^{b-1} \leq C_{a,b} := \frac{(a-1)^{a-1}(b-1)^{b-1}}{(a+b-2)^{a+b-2}} \leq 1 \qquad \text{pour tout } x \in]0,1[$$

(par convention $C_{a,b}=1$ si $a=1$ ou $b=1$). On trouve la constante $C_{a,b}$ en étudiant la fonction $x \mapsto x^{a-1}(1-x)^{b-1}$, dont le maximum est atteint pour $x=\frac{a-1}{a+b-2}$. On peut donc utiliser la méthode de rejet décrite dans la Section 9.3.3 :

- on simule des variables aléatoires indépendantes U_i, V_i de loi $U_i \sim \mathrm{U}(0,1)$ et $V_i \sim \mathrm{U}(0,C_{a,b})$;
- on rejette (U_i,V_i) jusqu'au premier instant T où $V_T \leq U_T^{a-1}(1-U_T)^{b-1}$;
- on renvoie la valeur de U_T, dont la densité est proportionnelle à $x^{a-1}(1-x)^{b-1}\mathbb{1}_{]0,1[}(x)$.

Un programme Python correspondant peut s'écrire de la façon suivante :

```python
def Beta2(a,b):
    if a==1 or b==1:
        C=1
    else:
        C=(a-1)**(a-1)*(b-1)**(b-1)/(a+b-2)**(a+b-2)
    U=rd.random()
    while C*rd.random()>U**(a-1)*(1-U)**(b-1):
        U=rd.random()
    return U
```

Exercice 9.14. On considère une variable aléatoire réelle W de densité donnée par

$$f(x) = \frac{1}{2\Gamma(5/4)} \mathrm{e}^{-x^4}.$$

Vérifier qu'il s'agit bien d'une densité de probabilité. En utilisant la méthode de rejet générale, donner un algorithme permettant de simuler la variable aléatoire W.

[*Sugg.* On pourra commencer par montrer que $\mathrm{e}^{-x^4} \leq c\mathrm{e}^{-x^2}$ pour tout $x \in \mathbb{R}$, avec $c = \mathrm{e}^{1/4}$.]

Solution 9.14. Commençons par observer que $x^4 - x^2 = (x^2 - \frac{1}{2})^2 - \frac{1}{4} \geq -\frac{1}{4}$ pour tout $x \in \mathbb{R}$. On en déduit que $\mathrm{e}^{-x^4} \leq \mathrm{e}^{-x^2+1/4}$ pour tout $x \in \mathbb{R}$, dont découle l'inégalité annoncée.

On sait simuler une variable aléatoire normale $\mathrm{N}(0, \frac{1}{2})$, dont la densité est donnée par $\frac{1}{\sqrt{\pi}}\mathrm{e}^{-x^2}$ et est donc proportionnelle à e^{-x^2}. On peut alors utiliser la méthode de rejet décrite dans la Section 9.3.3 :

- on simule des variables aléatoires indépendantes $Z_i \sim \mathrm{N}(0, \frac{1}{2})$ et $U_i \sim \mathrm{U}(0,1)$;
- on rejette les points (Z_i, U_i) jusqu'au premier instant T où $U_T\, \mathrm{e}^{1/4}\mathrm{e}^{-Z_T^2} \leq \mathrm{e}^{-Z_T^4}$ ou de manière équivalente $U_T \leq \exp(Z_T^2 - Z_T^4 - \frac{1}{4})$;
- on renvoie la valeur de Z_T, qui est une variable aléatoire de densité proportionnelle à e^{-x^4}.

En utilisant la fonction `normal` de la bibliothèque `numpy.random` pour simuler une variable aléatoire $\mathrm{N}(0, \frac{1}{2})$, un programme qui permet de simuler la variable aléatoire W peut s'écrire de la manière suivante :

```python
def W():
    Z=rd.normal(0,0.5)
    while rd.random()>exp(Z**2-Z**4-1/4):
        Z=rd.normal(0,0.5)
    return Z
```

Exercice 9.15. On cherche à écrire un programme qui génère une variable aléatoire de loi $\mathrm{Gamma}(n,1)$, où $n \in \mathbb{N}^*$.

(i) D'après la Proposition 6.36, si $(X_i)_{i\geq 1}$ sont des variables aléatoires i.i.d. de loi $\mathrm{Exp}(1)$, alors $X_1 + \cdots + X_n \sim \mathrm{Gamma}(n,1)$. En déduire un algorithme pour générer une variable aléatoire de loi $\mathrm{Gamma}(n,1)$.

(ii) Générer 10^6 variables aléatoires i.i.d. de loi $\mathrm{Gamma}(n,1)$ pour $n=2$, $n=5$ et $n=10$, en utilisant la méthode de la question précédente puis la méthode de rejet présentée dans la Section 9.3.3, voir l'Exemple 9.14 (pour simuler une variable aléatoire $\mathrm{Gamma}(\alpha,1)$ avec $\alpha > 1$). Comparer les temps d'exécution des deux méthodes.

Solution 9.15. (i) Il suffit de sommer n variables aléatoires indépendantes de loi $\mathrm{Exp}(1)$ (rappelons que l'on décrit dans la Section 9.3.1 la méthode d'inversion pour simuler une variable aléatoire $\mathrm{Exp}(1)$: si $U \sim \mathrm{U}(0,1)$, alors $-\log U \sim \mathrm{Exp}(1)$). Un programme correspondant peut s'écrire de la manière suivante :

```
def Gamma2(n):
    X=0
    for i in range(n):
        X+=-log(rd.random())
    return X
```

(ii) Pour faciliter la lecture, reproduisons ici le programme de l'Exemple 9.14 pour simuler une variable aléatoire $\mathrm{Gamma}(\alpha,1)$ avec $\alpha > 1$.

```
def Gamma(a):
    C=(2*(a-1))**(a-1)*exp(1-a)
    Z=-2*log(rd.random())
    while C*rd.random()>Z**(a-1)*exp(-0.5*Z):
        Z=-2*log(rd.random())
    return Z
```

On peut alors comparer les temps d'exécution pour les deux méthodes :

- Pour la première méthode :

```
for n in [2,5,10]:
    t=time.time()
    for k in range(10**6):
        Gamma2(n)
    t=time.time()-t
    print("Pour n={} le temps d'execution vaut {}".format(n,t))
```

```
Pour n=2 le temps d'execution vaut 3.259047269821167
Pour n=5 le temps d'execution vaut 5.888744831085205
Pour n=10 le temps d'execution vaut 10.534276962280273
```

- Pour la deuxième méthode :

```
for n in [2,5,10]:
    t=time.time()
    for k in range(10**6):
        Gamma2(n)
    t=time.time()-t
    print("Pour n={} le temps d'execution vaut {}".format(n,t))
```

```
Pour n=2 le temps d'execution vaut 5.3578901290893555
Pour n=5 le temps d'execution vaut 15.995364904403687
Pour n=10 le temps d'execution vaut 340.28289318084717
```

On réalise que le temps d'exécution devient très élevé lorsque n est grand : cela est dû au fait que dans la méthode de rejet la probabilité d'acceptation devient très faible quand α est grand.

Exercice 9.16. Appliquons la méthode de rejet pour simuler une loi de Cauchy standard.

(i) Montrer que $\frac{1}{1+x^2} \leq \frac{2}{(1+x)^2}$ pour tout $x \in]0,\infty[$.

(ii) En utilisant la méthode d'inversion, donner un algorithme qui permet de générer une variable aléatoire de densité $\frac{1}{(1+x)^2}\mathbb{1}_{]0,\infty[}(x)$.

(iii) En déduire, en utilisant la méthode de rejet, un algorithme qui permet de générer une variable aléatoire de densité $\frac{2}{\pi(1+x^2)}\mathbb{1}_{]0,\infty[}(x)$.

(iv) Conclure en donnant une méthode (basée sur la question précédente) pour générer une variable aléatoire de loi de Cauchy standard.

[*Sugg.* On pourra utiliser la même idée que dans l'Exercice 9.11 ; on peut aussi utiliser l'Exercice 6.29.]

Solution 9.16. (i) Il suffit de voir que $(x-1)^2 \geq 0$ pour tout $x \in \mathbb{R}$, d'où $1+2x+x^2 \leq 2+2x^2$, ou encore $(1+x)^2 \leq 2(1+x^2)$, dont on déduit l'inégalité voulue.

(ii) Avec un calcul rapide, la fonction de répartition d'une telle variable aléatoire est donnée par $F(t)=0$ si $t \leq 0$ et $F(t)=\int_0^t (1+x)^{-2}\mathrm{d}x = 1-(1+t)^{-1}$ si $t>0$. L'inverse de la fonction de répartition se trouve en résolvant $F(t)=u$ pour $u \in \,]0,1[$: on obtient $F^{-1}(u)=\frac{u}{1-u}$ pour $u \in \,]0,1[$. D'après la méthode d'inversion décrite dans la Section 9.3.1 on en déduit que si $U \sim \mathrm{U}(0,1)$, alors la variable aléatoire $\frac{U}{1-U}$ a pour fonction de répartition F, c'est-à-dire est une variable aléatoire absolument continue de densité $f(x)=\frac{1}{(1+x)^2}\mathbb{1}_{]0,\infty[}(x)$.

```
def Intermediaire1():
    U=rd.random()
    return U/(1-U)
```

(iii) On peut utiliser la méthode de rejet décrite dans la Section 9.3.3 :

- on simule des variables aléatoires indépendantes X_i, U_i, où X_i a pour densité $f(x)=\frac{1}{(1+x)^2}\mathbb{1}_{]0,\infty[}(x)$ et $U_i \sim \mathrm{U}(0,1)$;
- on rejette les points (X_i,U_i) jusqu'au premier instant T où $U_T\,\frac{2}{(1+X_T)^2} \leq \frac{1}{1+X_T^2}$ ou de manière équivalente $U_T \leq \frac{1}{2}+\frac{X_T}{1+X_T^2}$;
- on renvoie la valeur de X_T, qui est une variable aléatoire de densité proportionnelle à $\frac{1}{1+x^2}\mathbb{1}_{]0,\infty[}$, c'est-à-dire avec la loi voulue.

Un programme Python correspondant peut s'écrire de la façon suivante :

```
def Intermediaire2():
    X=Intermediaire1()
    while rd.random()>0.5+X/(1+X**2):
        X=Intermediaire1()
    return X
```

(iv) Soient Y et Z deux variables aléatoires indépendantes, avec Y absolument continue de densité $f_Y(y) := \frac{2}{\pi(1+y^2)}\mathbb{1}_{]0,\infty[}$ et $Z \sim \mathrm{Bern}(\frac{1}{2})$. On pose $X = ZY+(1-Z)Y$, c'est-à-dire $X=Y$ si $Z=1$ et $X=-Y$ si $Z=0$. On peut alors montrer, exactement comme dans l'Exercice 9.11, que X est absolument continue, de densité $\frac{1}{2}f_Y(|y|)=\frac{1}{\pi(1+y^2)}$ pour $y \in \mathbb{R}$, c'est-à-dire X est une variable aléatoire de Cauchy standard.

Un programme Python pour simuler une variable de Cauchy standard peut donc s'écrire de la façon suivante :

```
def Cauchy3():
    Y=Intermediaire2()
    if rd.random()<0.5:
        return Y
    else:
        return -Y
```

Exercice 9.17. Générer 10^6 variables aléatoires i.i.d. de loi de Cauchy standard, en utilisant : (i) la méthode d'inversion de la fonction de répartition (voir page 381) ; (ii) la méthode de l'Exercice 9.4; (iii) la méthode de l'Exercice 9.16; (iv) la fonction `standard_cauchy` de la bibliothèque `numpy.random`. Comparer les temps d'exécution des différentes méthodes.

Solution 9.17. Pour faciliter la lecture, on réécrit ci-dessous les différents programmes pour simuler une variable aléatoire de Cauchy standard.

(i) Méthode d'inversion :

```
def Cauchy():
    return tan(pi*(rd.random()-0.5))
```

Pour estimer le temps d'exécution :

```
t=time.time()
for k in range(10**6):
    Cauchy()
t=time.time()-t
print("Le temps d'execution vaut {}".format(t))
```

```
Le temps d'execution vaut 1.1975128650665283
```

(ii) La méthode de l'Exercice 9.4 :

```
def Cauchy2():
    return rd.norma(0,1)/rd.normal(0,1)
```

Pour estimer le temps d'exécution :

```
t=time.time()
for k in range(10**6):
    Cauchy2()
t=time.time()-t
print("Le temps d'execution vaut {}".format(t))
```

```
Le temps d'execution vaut 12.47413682937622
```

(iii) La méthode de rejet de l'Exercice 9.16.

```
def Cauchy3():
    U=rd.random()
    X=U/(1-U)
    while rd.random()>0.5+X/(1+X^2):
        U=rd.random()
            X=U/(1-U)
    if rd.random()<0.5:
        return X
    else:
        return -X
```

Pour estimer le temps d'exécution :

```
t=time.time()
for k in range(10**6):
    Cauchy3()
t=time.time()-t
print("Le temps d'execution vaut {}".format(t))
```

```
Le temps d'execution vaut 3.692305326461792
```

(iv) En utilisant la fonction `standard_cauchy` de la bibliothèque `numpy.random` :

```
t=time.time()
for k in range(10**6):
    rd.standard_cauchy()
t=time.time()-t
print("Le temps d'execution vaut {}".format(t))
```

```
Le temps d'execution vaut 0.6622400283813477
```

La méthode la plus efficace est celle de la fonction `standard_cauchy` de la bibliothèque `numpy.random`, qui semble être environ deux fois plus rapide que la méthode d'inversion. La méthode du rapport de deux variables aléatoires $N(0,1)$ indépendantes semble être la moins efficace.

Exercice 9.18. Générer 10^6 variables aléatoires i.i.d. de loi $N(0,1)$, en utilisant : (i) la méthode de Box–Müller de la Section 9.3.2 ; (ii) la méthode de Marsaglia de l'Exercice 9.6 ; (iii) la fonction `normal` de la bibliothèque `numpy.random`, qui utilise la méthode dite *zigourrat*. Comparer les temps d'exécution des différentes méthodes.

Solution 9.18. Pour faciliter la lecture, on réécrit ci-dessous les différents programmes Python pour simuler une variable aléatoire normale standard.

(i) Méthode de Box–Müller du Paragraphe 9.3.2 :

```
def Normale():
    return sqrt(-2*log(rd.random()))*cos(2*pi*rd.random())
```

Pour estimer le temps d'exécution :

```
t=time.time()
for k in range(10**6):
    Normale()
t=time.time()-t
print("Le temps d'execution vaut {}".format(t))
```

```
Le temps d'execution vaut 2.918368101119995
```

(ii) La méthode de Marsaglia de l'Exercice 9.6 :

```
def Normale1():
    X,Y=2*rd.random()-1,2*rd.random()-1
    while X**2+Y**2>1:
        X,Y=2*rd.random()-1,2*rd.random()-1
    R=sqrt(X**2+Y**2)
    return sqrt(-4*log(R))*X/R
```

Pour estimer le temps d'exécution :

```
t=time.time()
for k in range(10**6):
    Normale1()
t=time.time()-t
print("Le temps d'execution vaut {}".format(t))
```

```
Le temps d'execution vaut 4.177669286727905
```

(iii) En utilisant la fonction `normal` de la bibliothèque `numpy.random` :

```
t=time.time()
for k in range(10**6):
    rd.normal(0,1)
t=time.time()-t
print("Le temps d'execution vaut {}".format(t))
```

```
Le temps d'execution vaut 5.216531991958618
```

Il semble que la fonction `normal` de la bibliothèque `numpy.random` soit la méthode la moins efficace ! En réalité, on peut ajouter un paramètre dans les fonctions de la bibliothèque `numpy.random` pour simuler plusieurs réalisations d'une variable aléatoire : par exemple, `rd.random(10)` génère 10 variables aléatoires indépendantes de loi $\mathrm{U}(0,1)$. Les fonctions de la bibliothèque `numpy.random` sont optimisées pour pouvoir simuler un grand nombre de variables aléatoires. Dans le cas de variables aléatoires $\mathrm{N}(0,1)$, par exemple, on peut simuler 10^6 variables aléatoires $\mathrm{N}(0,1)$ en écrivant `rd.normal(0,1,10**6)` ; on peut estimer le temps d'exécution :

```
t=time.time()
rd.normal(0,1,10**6)
t=time.time()-t
print("Le temps d'execution vaut {}".format(t))
```

```
Le temps d'execution vaut 0.05921292304992676
```

Cette fois, le temps d'exécution est extrêmement bas !

Exercice 9.19. En utilisant la méthode de Monte-Carlo, donner une valeur approchée de la probabilité que si l'on prend 3 points indépendamment et uniformément dans le disque $\mathrm{D}_1 = \{(x,y) \in \mathbb{R}^2, x^2+y^2 \leq 1\}$, le triangle qu'ils forment contienne l'origine $(0,0)$. Avez-vous une conjecture pour la vraie valeur de cette probabilité ? Êtes-vous capable de démontrer cette conjecture ?
[*Sugg.* Utiliser l'Exemple 9.15 pour générer des points indépendants de loi uniforme continue sur D_1.]

Solution 9.19. Étant donnés trois points $X,Y,Z \in \mathbb{R}^2$, l'origine $0 \in \mathbb{R}^2$ est contenue dans le triangle formé par ces points si et seulement si 0 est un barycentre de X,Y,Z, c'est-à-dire si et seulement s'il existe $a,b,c \in \,]0,1[$ avec $a+b+c=1$ tels que $aX+bY+cZ=0$. En divisant par a, on obtient que $0 \in \mathbb{R}^2$ est contenu dans le triangle formé par les points $X,Y,Z \in \mathbb{R}^2$ si et seulement s'il existe $u,v>0$ tels que $X+uY+vZ=0$. En écrivant $X=(x_1,x_2)$, $Y=(y_1,y_2)$ et $Z=(z_1,z_2)$, on obtient deux équations en u,v :

$$\begin{cases} x_1+uy_1+vz_1=0 \\ x_2+uy_2+vz_2=0 \end{cases} \iff \begin{cases} u(y_1z_2-y_2z_1)=x_2z_1-x_1z_2 \\ v(z_1y_2-z_2y_1)=x_2y_1-x_1y_2 \end{cases}$$

donc $u,v>0$ si et seulement si $(y_1z_2-y_2z_1)(x_2z_1-x_1z_2)>0$ et $(z_1y_2-z_2y_1)(x_2y_1-x_1y_2)>0$.

Écrivons maintenant un programme Python qui prend en argument trois points $X=[x_1,x_2]$, $Y=[y_1,y_2]$, $Z=[z_1,z_2]$ et renvoie `True` si 0 est dans le triangle formé par ces points et `Fasle` sinon.

```
def DansTriangle(X,Y,Z):
    terme1=Y[0]*Z[1]-Y[1]*Z[0]
    terme2=X[1]*Z[0]-X[0]*Z[1]
    terme3=X[1]*Y[0]-X[0]*Y[1]
    if terme1*terme2>0 and terme1*terme3<0:
        return True
    else:
        return False
```

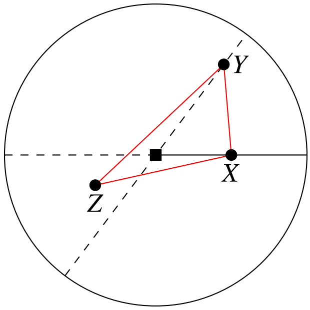
Il suffit maintenant de simuler un grand nombre n de vecteurs aléatoires X_i,Y_i,Z_i indépendants de loi uniforme sur D_1 (en utilisant la méthode de rejet décrite dans l'Exemple 9.15, le programme est réécrit ci-dessous pour faciliter la lecture), de compter le nombre de fois où la fonction `DansTriangle`(X_i,Y_i,Z_i) renvoie `True` (c'est-à-dire le nombre de fois où l'événement $A_i =$ « l'origine est dans le triangle formé par les points X_i,Y_i,Z_i » est vérifié) et de diviser par le nombre de répétitions de l'expérience. On renvoie à la Section 9.4 pour les détails de cette méthode. Un programme peut donc s'écrire de la manière suivante :

```python
def Disque():
    x,y=2*rd.random()-1,2*rd.random()-1
    while x**2+y**2>1:
        x,y=2*rd.random()-1,2*rd.random()-1
    return x,y
n=10**6
somme=0
for i in range(n):
    X,Y,Z=Disque(),Disque(),Disque()
    if DansTriangle(X,Y,Z):
        somme+=1
print("La probabilite vaut environ {}".format(somme/n))
```

```
La probabilite vaut environ 0.25026
```

La probabilité semble très proche de $0,25 = \frac{1}{4}$. On peut en fait démontrer ce résultat.

Notons X,Y,Z trois points indépendants, de loi uniforme continue sur D_1. Par symétrie, en opérant une rotation, on peut toujours supposer que X est de la forme $(x,0)$ avec $x>0$, c'est-à-dire qu'il est sur l'axe horizontal positif. Faisons un dessin, avec cette simplification :



L'origine $0 \in \mathrm{D}_1$ est à l'intérieur du triangle formé par X,Y,Z si et seulement si le point Z est dans le secteur symétrique par rapport à celui déterminé par X et Y. En posant $Y = R_1 \cos\Theta_1$ et $Z = R_2 \cos\Theta_2$, l'événement $A =$ « l'origine est dans le triangle formé par les points X,Y,Z » est donné par

$$A = \{\Theta_1 \in]0,\pi[\,, \Theta_2 \in]\pi,\pi+\Theta_1[\,\} \cup \{\Theta_2 \in]0,\pi[\,, \Theta_1 \in]\pi,\pi+\Theta_2[\,\}.$$

Observons que l'union est disjointe et que l'événement A dépend seulement des angles Θ_1,Θ_2.

Grâce à l'Exercice 9.6 (première question), Θ_1,Θ_2 sont des variables aléatoires de loi uniforme continue sur $]0,2\pi[$ (ce fait est relativement clair) et elles sont indépendantes. On a donc

$$\mathrm{P}\left(\Theta_1 \in]0,\pi[\,, \Theta_2 \in]\pi,\pi+\Theta_1[\,\right) = \int_0^{\pi} \int_{\pi}^{\pi+t_1} \frac{1}{(2\pi)^2}\,\mathrm{d}t_2\,\mathrm{d}t_1 = \frac{1}{4\pi^2}\int_0^{\pi} t_1 \mathrm{d}t_1 = \frac{1}{8},$$

donc on conclut que

$$\mathrm{P}\left(\text{« l'origine est dans le triangle formé par les points } X,Y,Z \text{ »}\right) = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}.$$

Si on veut éviter d'utiliser la symétrie pour se réduire au cas où X et sur l'axe horizontal, on peut écrire $X = R_0 \cos\Theta_0$ et reprendre les calculs précédents en remplaçant Θ_1 par $\Theta_1 - \Theta_0$ et Θ_2 par $\Theta_2 - \Theta_0$.

Exercice 9.20. On considère la fonction $f(x) = x^{-1/4}|\cos(1/x)|$, définie sur $\mathbb{R}^*$, et on cherche à estimer l'intégrale $\int_0^{2/\pi} f(x)\mathrm{d}x$.

(i) Montrer que f et f^2 sont intégrables sur $]0, \frac{2}{\pi}[$.

(ii) Écrire un algorithme qui permet d'estimer l'intégrale $\int_0^{2/\pi} f(x)\mathrm{d}x$, en utilisant la méthode de Monte-Carlo.

(iii) Supposons que le programme qui calcule $\frac{1}{n}\sum_{i=1}^{n} f(X_i)$, où X_i sont des variables aléatoires indépendantes de loi $\mathrm{U}(0, \frac{2}{\pi})$, ait renvoyé pour $n = 10^8$ la valeur $0,56915$. Donner un intervalle de confiance pour la valeur de l'intégrale $\int_0^{2/\pi} f(x)\mathrm{d}x$, de niveau de confiance 95%.

[*Sugg.* Observer que $\mathrm{Var}(f(X_i)) \leq \mathrm{E}(f(X_i)^2) \leq \int_0^{2/\pi} x^{-1/2}\,\mathrm{d}x = 2\sqrt{2/\pi}$.]

Solution 9.20. (i) On a $0 \leq f(x) \leq x^{-1/4}$ et $0 \leq f(x)^2 \leq x^{-1/2}$ pour tout $x \in]0, \frac{2}{\pi}[$, donc f et f^2 sont intégrables sur l'intervalle $]0, \frac{2}{\pi}[$ parce que $x \mapsto x^{-1/2}$ et $x \mapsto x^{-1/4}$ le sont.

(ii) Si $X \sim \mathrm{U}(0, \frac{2}{\pi})$, on a $\mathrm{E}(f(X)) = \frac{\pi}{2}\int_0^{2/\pi} f(x)\mathrm{d}x$. La méthode de Monte-Calro décrite dans la Section 9.4 permet d'estimer $\mathrm{E}(f(X))$ grâce à la moyenne empirique $\frac{1}{n}\sum_{i=1}^{n} f(X_i)$, où $(X_i)_{i\geq 1}$ sont des variables aléatoires indépendantes telles que $X_i \sim \mathrm{U}(0, \frac{2}{\pi})$. Un programme qui prend en argument un entier n et qui renvoie la valeur (aléatoire) $\frac{1}{n}\sum_{i=1}^{n} f(X_i)$ peut s'écrire de la façon suivante

```
def f(x):
    return x**(-1/4)*abs(cos(1/x))

def Integrale(n):
    somme=0
    for i in range(n):
        somme+=f(2*rd.random()/pi)
    return 2*somme/(pi*n)
```

(iii) Pour $n = 10^8$, le programme renvoie la valeur $0,56915$. En posant $Y_n := \frac{1}{n}\sum_{i=1}^{n} f(X_i)$, on obtient un intervalle de confiance asymptotique de niveau de confiance $1-\alpha$ grâce à (9.5) :

$$\left[Y_n - \bar{\sigma}\frac{z_{\alpha/2}}{\sqrt{n}} \,,\, Y_n + \bar{\sigma}\frac{z_{\alpha/2}}{\sqrt{n}} \right],$$

où $z_{\alpha/2}$ est le quantile d'ordre $1-\alpha/2$ d'une variable aléatoire $\mathrm{N}(0,1)$ et $\bar{\sigma}^2$ est une constante, qui doit être plus grande que $\mathrm{Var}(f(X_i))$. Dans le cas présent, on prend $1-\alpha = 0,95$ de sorte que $z_{\alpha/2} \approx 1,96$ (grâce à la table page S-225) ; $\bar{\sigma} = (2\sqrt{2/\pi})^{1/2}$ (grâce à la suggestion) et $n = 10^8$. On obtient l'intervalle de confiance suivant pour $\int_0^{2/\pi} f(x)dx$, de niveau de confiance 95% : $[0,56884, 0,56947]$.

Exercice 9.21 (Père Noël secret). On souhaite écrire un algorithme qui permet d'effectuer un tirage du père Noël du problème 2.4, en évitant qu'une personne tire son propre nom. Autrement dit, on souhaite générer une permutation σ de $\mathfrak{S}_n$ choisie de manière uniforme dans l'ensemble $\mathfrak{S}_n^\star = \{\sigma \in \mathfrak{S}_n, \sigma(i) \neq i \ \forall i \in \{1,\dots,n\}\}$ des permutations qui n'ont aucun point fixe. En utilisant une méthode de rejet, écrire un tel algorithme. Quelle est la probabilité de rejet ?

[*Sugg.* On pourra utiliser la fonction `permutation` de la bibliothèque `numpy.random` pour générer une permutation aléatoire choisie uniformément dans $\mathfrak{S}_n$.]

Solution 9.21. On applique la méthode de rejet :

- on simule des permutations indépendantes $(\sigma_i)_{i\geq 1}$ de loi uniforme sur $\mathfrak{S}_n$;

- on rejette σ_i jusqu'au premier instant T où $\sigma_T \in \mathfrak{S}_n^\star$;
- on renvoie σ_T, qui est de loi uniforme (discrète) sur $\mathfrak{S}_n^\star$.

Pour appliquer cette méthode de rejet, on a besoin d'une fonction qui vérifie si une permutation possède ou non un point fixe : il faut vérifier si $\sigma(0) \neq 0$, $\sigma(1) \neq 1$, etc. ; on peut s'arrêter dès qu'on trouve un i tel que $\sigma(i) = i$. On donne ci-dessous un programme `Accepte` qui prend comme argument une permutation σ et renvoie `True` si σ ne possède aucun point fixe et `False` sinon.

```
def Accepte(sigma):
    n=len(sigma)
    AucunPointFixe=True
    i=0
    while AucunPointFixe and i<n:
        if sigma[i]==i:
            AucunPointFixe=False
        i+=1
    return AucunPointFixe
```

Un programme qui met en pratique la méthode de rejet peut alors s'écrire de la façon suivante :

```
def Permutation1(n):
    sigma=rd.permutation(n)
    while not Accepte(sigma):
        sigma=rd.permutation(n)
    return sigma
```

La probabilité p de rejet est la probabilité qu'une permutation aléatoire σ de loi uniforme sur $\mathfrak{S}_n$ possède un point fixe : il s'agit de la probabilité calculée dans le Problème 2.4 (Section 2.1), voir (2.3). La probabilité vaut $p = 1 - \sum_{k=0}^{n} \frac{(-1)^k}{k!}$, c'est-à-dire $p \approx 1 - \mathrm{e}^{-1} \approx 0,632$ si n est grand (même pour $n = 6$ l'approximation est excellente).

Exercice 9.22 (Problème des anniversaires). On considère la variante suivante du problème des anniversaires : on sélectionne au hasard n personnes nées une année non bissextile, que l'on numérote de 1 à n, et on cherche à estimer la probabilité qu'*au moins* trois *d'entre elles aient leur anniversaire le même jour*.

(i) Écrire un programme `Anniv` qui prend en argument un entier n et qui renvoie une liste de n dates anniversaires choisies aléatoirement dans $\{1, \dots, 365\}^n$.

(ii) Écrire un programme `Trois` qui prend une liste `L` et qui renvoie `True` s'il existe trois indices distincts i, j, k tel que `L[`i`]=L[`j`]=L[`k`]`, et `False` sinon.

[*Sugg.* On pourra utiliser la fonction `count` de Python : `L.count(`i`)` renvoie le nombre d'occurrences de i dans la liste `L`.]

(iii) À l'aide de la méthode de Monte-Carlo (avec 10 000 répétitions), estimer la probabilité dans un groupe de n personnes trois d'entre elles aient le même anniversaire pour $n = 88$, $n = 100$, $n = 150$. Donner des intervalles de confiance pour ces probabilités, de niveau de confiance 95%.

Solution 9.22. (i) Il suffit de simuler n variables aléatoires de loi uniforme discrète sur $\{1, 2, \dots, 365\}$; d'après la méthode décrite dans la Section 9.2, on peut utiliser que si $U \sim \mathrm{U}(0,1)$ alors $\lfloor 365U \rfloor + 1$ est uniforme discrète sur $\{1, 2, \dots, 365\}$.

```
def Anniv(n):
    return [int(365*rd.random())+1 for i in range(n)]
```

(ii) Il suffit de regarder combien de fois l'élément `L[i]` apparaît dans la liste `L` pour $i = 0, 1, \ldots$; on peut s'arrêter dès que ce nombre est supérieur à 3.

```
def Trois(L):
    n=len(L)
    Trouve=False
    i=0
    while i<n and not Trouve:
        x=L[i]
        if L.count(x)>=3:
            Trouve=True
        i+=1
    return Trouve
```

(iii) On peut appliquer la méthode de Monte-Carlo pour estimer la probabilité de l'événement $A =$ « dans un groupe de n personnes, trois d'entre elles ont le même anniversaire ». Pour $k = 10^6$, on répète k fois la simulation de l'expérience aléatoire : `Trois(Anniv(n))` renvoie `True` si dans les anniversaires de n personnes il y en a trois qui sont identiques (c'est-à-dire si A est vérifié) et `False` si non ; il suffit de compter le nombre de fois où on obtient `True` et de diviser par k.

```
k=10**5
for n in [88,100,150]:
  somme=0
  for i in range(k):
    if Trois(Anniv(n)):
      somme+=1
  p=somme/k
  print("Pour n={} la probabilite vaut environ {}".format(n,p))
```

```
Pour n=88 la probabilite vaut environ 0.51051
Pour n=100 la probabilite vaut environ 0.64694
Pour n=150 la probabilite vaut environ 0.96507
```

En appliquant la formule (9.4), on obtient des intervalles de confiance asymptotiques pour $\mathrm{P}(A)$ de niveau de confiance $1-\alpha$. En notant que $z_{\alpha/2} \approx 1,96$ pour $1-\alpha = 95\%$ (grâce à la table page S-225) et $k = 10^5$, de sorte que $\frac{z_{\alpha/2}}{2\sqrt{k}} \approx 0,0031$, on obtient les intervalles de confiance suivants de niveau de confiance 95%, pour la probabilité p_n que dans un groupe de n personnes trois d'entre elles aient le même anniversaire :

- Pour $n = 88$, $[0,5074, 0,5136]$;
- Pour $n = 100$, $[0,6438, 0,6501]$;
- Pour $n = 150$, $[0,9619, 0,9681]$.

Exercice 9.23. Un jeu se déroule de la manière suivante. On tire un chiffre au hasard dans $\{0, 1, \ldots, 9\}$, que l'on note X_1 : si $X_1 = 0$ le jeu s'arrête et le score est $S = 0$; si $X_1 \geq 1$ le jeu continue. On tire alors un chiffre au hasard dans $\{0, \ldots, X_1\}$, que l'on note X_2 : si $X_2 = 0$ le jeu s'arrête et le score est $S = 0, X_1$; si $X_2 \geq 1$ le jeu continue Après n tours, si $X_n \geq 1$, on tire un chiffre au hasard dans $\{0, \ldots, X_n\}$, que l'on note X_{n+1} : si $X_{n+1} = 0$, le jeu s'arrête et le score est $S := 0, X_1 X_2 \cdots X_n$; si $X_{n+1} \geq 1$, le jeu continue...

(i) Écrire un programme qui permet de simuler ce jeu, c'est-à-dire qui renvoie le score (aléatoire) d'une partie. Estimer le score moyen $\mathrm{E}(S)$ en simulant 100 000 parties.

Obtenons maintenant des résultats rigoureux sur ce jeu.

(ii) On note $T = \min\{i \geq 1, X_i = 0\}$ le nombre de tours que dure le jeu. Montrer que l'on a $\mathrm{P}(T > n) = \mathrm{P}(X_1 \geq 1, \ldots, X_n \geq 1) \leq (\frac{9}{10})^n$. Conclure que $\mathrm{P}(T < +\infty) = 1$.

(iii) Pour tout $k \in \{0,\ldots,9\}$, on définit une variante du jeu où au premier tour on choisit X_1 uniformément dans $\{0,\ldots,k\}$ plutôt que dans $\{0,\ldots,9\}$; le reste du jeu se déroule de la même manière. On note alors s_k l'espérance du score obtenu. Montrer que $s_k := \sum_{i=0}^{k} \frac{1}{k+1}\left(\frac{i}{10} + \frac{1}{10}s_i\right)$, pour tout $k \in \{0,\ldots,9\}$, puis (par récurrence) que $s_k = \frac{k}{19}$. Conclure que $\mathrm{E}(S) = \frac{9}{19}$.

Solution 9.23. (i) Il suffit d'appliquer les instructions, avec une boucle `while` : à chaque tour on met à jour la valeur de X_n (entre 0 et X_{n-1}) et le score ; noter que le jeu s'arrête lorsque $X_n = 0$. Observons que si S_n est le score après n tours, avec par convention $S_0 = 0$, alors dans le cas où $X_n > 0$ on a la relation $S_n = S_{n-1} + \frac{X_n}{10^n}$; notons qu'il faut donc aussi garder en mémoire le numéro du tour. Rappelons que si $U \sim \mathrm{U}(0,1)$, alors $\lfloor kU \rfloor$ est de loi uniforme discrète sur $\{0,1,\ldots,k-1\}$.

```python
def Jeu():
    score=0
    n=1
    X=int(10*rd.random())
    while X>0:     ## le jeu s'arrete quand X=0
        score+=X/10**n
        X=int((X+1)*rd.random())
        n=n+1
    return score
```

Pour estimer le score moyen $\mathrm{E}(S)$, on répète $n = 100\,000$ parties, on fait la somme des scores et on divise par n. Cela correspond à la moyenne empirique $\frac{1}{n}\sum_{i=1}^{n} \tilde{S}_i$, où $(\tilde{S}_i)_{i\geq 1}$ sont des variables aléatoires indépendantes et de même loi que S : d'après la loi des grands nombres, cette moyenne empirique est, avec grande probabilité, proche de $\mathrm{E}(S)$.

```python
n=100000
somme=0
for i in range(n):
    somme+=Jeu()
print("Le score moyen vaut environ {}".format(somme/n))
```

```
Le score moyen vaut environ 0.47429
```

(ii) On a l'égalité d'événements $\{T > n\} = \{X_i \neq 0, \forall 1 \leq i \leq n\}$, dont on déduit l'identité $\mathrm{P}(T > n) = \mathrm{P}(X_1 \geq 1, \ldots, X_n \geq 1)$. Notons maintenant que pour tous $i_1 \geq \cdots \geq i_n \geq 1$, on a

$$\mathrm{P}(X_1 = i_1, X_2 = i_2, \ldots, X_n = i_n) = \frac{1}{10} \cdot \frac{1}{i_1+1} \cdots \frac{1}{i_{n-1}+1}$$

parce que $\mathrm{P}(X_k = i_k \mid X_{k-1} = i_{k-1}) = \frac{1}{i_{k-1}+1}$, pour tout $i_k \leq i_{k-1}$. On en déduit que pour tous $i_1 \geq \cdots \geq i_{n-1} \geq 1$, en sommant sur $i_n \in \{1,\ldots,i_{n-1}\}$,

$$\begin{aligned}
&\mathrm{P}(X_1 = i_1, X_2 = i_2, \ldots, X_{n-1} = i_{n-1}, X_n \geq 1) \\
&\qquad = \frac{1}{10} \cdot \frac{1}{i_1+1} \cdots \frac{1}{i_{n-2}+1} \times \frac{i_{n-1}}{i_{n-1}+1} \leq \frac{1}{10} \cdot \frac{1}{i_1+1} \cdots \frac{1}{i_{n-2}+1} \times \frac{9}{10},
\end{aligned}$$

parce que $i_{n-1} \leq 9$. En répétant cette procédure, on obtient facilement par récurrence que

$$\mathrm{P}(X_1 \geq 1, \ldots, X_n \geq 1) \leq \left(\frac{9}{10}\right)^n.$$

Il en découle que $\mathrm{P}(T > n) \leq \left(\frac{9}{10}\right)^n \to 0$ quand $n \to \infty$, donc $\mathrm{P}(T = +\infty) = 0$.

(iii) Notons $S^{(k)}$ le score obtenu si le jeu commence avec X_1 uniforme sur $\{0,\ldots,k\}$. Clairement, on a $S^{(0)} = 0$. Pour $k = 1$, on peut faire le raisonnement suivant : si

$X_1 = 0$, le score vaut $S^{(1)} = 0$; si $X_1 = 1$, le score vaut $S^{(1)} = 0,1 + \frac{1}{10}\tilde{S}^{(1)}$, où $\tilde{S}^{(1)}$ est le score d'un jeu « décalé de 1 », c'est-à-dire du jeu où $\tilde{X}_1 = X_2, \tilde{X}_2 = X_3, \ldots$ — qui est le même jeu, en commençant avec $\tilde{X}_1$, qui est de loi uniforme discrète sur $\{0,\ldots,X_1\} = \{0,1\}$ — et possède la même loi que $S^{(1)}$. Par linéarité de l'espérance, on a $\mathrm{E}(S^{(1)}) = \mathrm{E}(S^{(1)}\mathbb{1}_{\{X_1=1\}}) = \frac{1}{2}(\frac{1}{10} + \frac{1}{10}\mathrm{E}(\tilde{S}^{(1)}))$, en notant que $\mathrm{P}(X_1 = 1) = \frac{1}{k+1} = \frac{1}{2}$ pour $k = 1$. On obtient donc $s_1 = \frac{1}{20}(1+s_1)$, d'où $s_1 = \frac{1}{19}$.

De manière générale, on peut écrire pour $k \in \{1,\ldots,9\}$,

$$S^{(k)} = \sum_{i=0}^{k} \mathbb{1}_{\{X_1=i\}}\Big(\frac{i}{10} + \frac{1}{10}\tilde{S}^{(i)}\Big),$$

où $\tilde{S}^{(i)}$ est le score d'un jeu « décalé de 1 », c'est-à-dire du jeu où $\tilde{X}_1 = X_2, \tilde{X}_2 = X_3, \ldots$ — qui est le même jeu, en commençant avec $\tilde{X}_1$, qui est de loi uniforme discrète sur $\{0,\ldots,X_1 = i\}$ — et possède la même loi que $S^{(i)}$. Par linéarité de l'espérance, et en observant que X_1 et $\tilde{S}^{(i)}$ sont indépendantes, on obtient

$$s_k = \mathrm{E}(S^{(k)}) = \sum_{i=0}^{k} \mathrm{P}(X_1 = i)\Big(\frac{i}{10} + \frac{1}{10}\mathrm{E}(\tilde{S}^{(i)})\Big) = \sum_{i=0}^{k} \frac{1}{k+1}\Big(\frac{i}{10} + \frac{1}{10}s_i\Big).$$

On peut écrire cette relation de la façon suivante :

$$10(k+1)s_k = \sum_{i=0}^{k} i + s_0 + s_1 + \cdots + s_k = \frac{k(k+1)}{2} + s_0 + s_1 + \cdots + s_k\,,$$

ou encore $(10k+9)s_k = \frac{k(k+1)}{2} + s_0 + \cdots + s_{k-1}$. Montrons par récurrence que $s_k = \frac{k}{19}$. On a déjà vu que $s_0 = 0$, $s_1 = \frac{1}{19}$. Pour $k \geq 1$,

$$\begin{aligned}(10(k+1)+9)s_{k+1} &= \frac{(k+1)(k+2)}{2} + \frac{0}{19} + \cdots + \frac{k}{19}\\ &= \frac{(k+1)(k+2)}{2} + \frac{1}{19}\frac{k(k+1)}{2} = \frac{k+1}{19}(10k+19)\,,\end{aligned}$$

ce qui conclut la récurrence. On obtient donc $E(S) = s_9 = \frac{9}{19} \approx 0,4737$.

Exercice 9.24 (Urne de Pólya). On considère l'expérience de l'Exercice 3.61. Une urne contient initialement une boule blanche et une boule noire. À chaque tour, une boule est prélevée de l'urne, puis elle est replacée dans l'urne en ajoutant une boule de la même couleur. On note X_n le nombre de boules blanches dans l'urne après n tours, qui est une variable aléatoire à valeurs dans $\{1,\ldots,n+1\}$. Pour tout $n \in \mathbb{N}$, on pose $Y_n := \frac{1}{n+2}X_n$ la proportion de boules blanches dans l'urne après n tours.

(i) Écrire un programme `Polya` qui prend en argument un entier n, et renvoie la liste $[X_0, X_1, \ldots, X_n]$ des variables aléatoires X_i lors d'une réalisation de cette expérience.

(ii) Représenter le graphe de 10 réalisations de $(Y_n)_{0\leq n\leq N}$ pour $N = 1000$, $N = 10000$. Quelles conjectures seriez-vous prêts à formuler ?

Solution 9.24. (i) Il suffit de remarquer qu'une fois que $X_0, X_1, \ldots, X_k$ sont déterminés, on sait que $X_{k+1} = X_k + 1$ ou bien $X_{k+1} = X_k$ suivant qu'au $(k+1)$-ème tirage on ait tiré une boule blanche (avec probabilité $\frac{X_k}{k+2}$ parce qu'il y a $k+2$ boules dans l'urne, dont X_k blanches) ou une boule noire (avec probabilité $1 - \frac{X_k}{k+2}$). Un programme qui répète cette

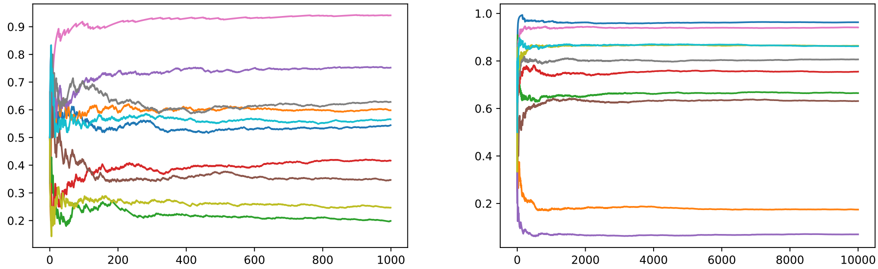


Fig. S9.1 (De l'Exercice 9.24) Repprésentation graphique de 10 réalisations de $(Y_n)_{0\leq n\leq N}$ pour $N = 1\,000$ (à gauche) et $N = 10\,000$ (à droite).

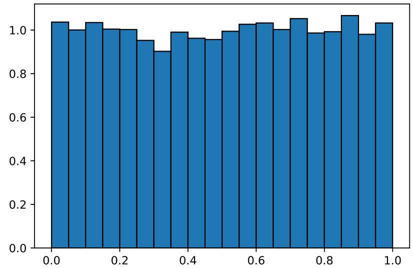


Fig. S9.2 (De l'Exercice 9.24) Histogramme de $k = 10^4$ réalisations de $Y_N = \frac{X_N}{N+2}$ pour $N = 5000$. Cela semble confirmer l'idée que $Y_\infty = \lim_{n\to\infty} Y_n$ (si la limite existe) est de loi $\mathrm{U}(0,1)$.

opération n fois, en mettant à jour la valeur de $X_0,\dots,X_k$ à chaque tour, peut s'écrire de la façon suivante :

```
def Polya(n):
    X=[1]
    for k in range(n):
        j=X[k]
        if rd.random()<j/(k+2):
            X.append(j+1)
        else:
            X.append(j)
    return X
```

(ii) Représentons le graphe de 10 réalisations de $(Y_n)_{0\leq n\leq N}$ pour $N = 1\,000$ et $N = 10\,000$. Les graphes obtenus sont présentés dans la Figure S9.1.

```
for N in [1000,10000]:
    for i in range(10):
        X=Polya(N)
        Y=[X[k]/(k+2) for k in range(N+1)]
        plt.plot(Y)
    plt.show()
```

Les graphes semblent indiquer que la suite $(Y_n)_{n\geq 0}$ *converge toujours*, c'est-à-dire que la limite $Y_\infty = \lim_{n\to\infty} Y_n$ existe avec probabilité 1, mais que la limite n'est jamais la même, c'est-à-dire Y_∞ est une variable aléatoire. Ce résultat est en effet vrai, mais des instruments d'analyse plus avancés sont nécessaires pour le démontrer.

D'autre part, d'après l'Exercice 3.61, on a que pour tout $n \in \mathbb{N}^*$ la variable aléatoire X_n est de loi uniforme discrète sur $\{1, 2, \ldots, n+1\}$, c'est-à-dire que Y_n est de loi uniforme discrète sur $\{\frac{1}{n+2}, \frac{2}{n+2}, \ldots, \frac{n+1}{n+2}\}$. Il semble donc naturel de s'attendre à ce que la limite Y_∞ soit de loi uniforme sur $]0, 1[$, ce qui est en effet le cas. Pour tester cette hypothèse, on peut faire un histogramme de $k = 10^4$ réalisations de $Y_N = \frac{X_N}{N+2}$, disons pour $N = 5000$ (rappelons que `Polya(N)[N]` permet de simuler une réalisation de X_N), que l'on présente dans la Figure S9.2.

```
k=10**4
N=5000
plt.hist([Polya(N)[N]/(N+2) for i in range(k)],density=True)
```

Exercice 9.25 (Modèle d'Ising à champ moyen). On considère l'évolution de l'opinion de $N \geq 2$ personnes à propos d'une question donnée. Plus précisément, pour un vecteur $x = (x_1, \ldots, x_N) \in \{-1, +1\}^N$ on interprète $x_i \in \{-1, +1\}$ comme l'opinion de la i-ème personne. On notera $H(x) := \sum_{k=1}^N x_k$ la « somme des opinions » et $H_i(x) = \sum_{k \neq i} x_k = H(x) - x_i$ la « somme des opinions observée par la personne i ».

On considère alors une suite $(X_n)_{n \geq 0}$ de variables aléatoires à valeurs dans $\{-1, +1\}^N$. La valeur de X_0 est donnée, et on définit X_n de manière récursive : pour $n \geq 0$, une fois X_n connu, on définit X_{n+1} en choisissant un individu i au hasard uniformément dans $\{1, \ldots, N\}$ et en changeant la i-ème coordonnée de X_n de manière aléatoire, en la posant égale à :

- $+1$ avec probabilité $\dfrac{e^{\beta H_i(X_n)}}{e^{\beta H_i(X_n)} + e^{-\beta H_i(X_n)}} = (1 + e^{-2\beta H_i(X_n)})^{-1}$;
- -1 avec probabilité $\dfrac{e^{-\beta H_i(X_n)}}{e^{\beta H_i(X_n)} + e^{-\beta H_i(X_n)}} = (1 + e^{2\beta H_i(X_n)})^{-1}$.

Dans ces formules, $\beta \geq 0$ est un paramètre qui mesure « l'influence » de l'opinion $H_i(x)$ des autres personnes sur la personne i (plus β est grand, plus la personne i a tendance à s'aligner avec l'opinion majoritaire). On cherche à comprendre l'évolution de X_n, plus particulièrement de l'*opinion générale* $H(X_n)$, notamment en fonction du paramètre β.

(i) Écrire un programme `Update` qui prend en argument un entier $N \geq 2$, un vecteur $x = (x_1, \ldots, x_n) \in \{-1, +1\}^N$ et un paramètre $\beta \geq 0$, et qui renvoie un vecteur (aléatoire) $x' \in \{-1, +1\}^N$ obtenu à partir de x selon les règles décrites plus haut.

(ii) Écrire un programme `Chaine`, qui prend en argument un entier $N \geq 2$ (le nombre de personnes), un paramètre $\beta \geq 0$ et un entier m (l'horizon de temps considéré), et qui renvoie l'évolution de l'opinion générale $Y_n := H(X_n)$ pour $0 \leq n \leq m$, c'est-à-dire la liste $[Y_0, Y_1, \ldots, Y_m]$ des valeurs de Y_n lors d'une réalisation de l'expérience. On prendra $X_0 := ((-1)^i)_{1 \leq i \leq N}$ comme point de départ.

[*Sugg.* On gardera en mémoire la valeur de $X_n \in \{-1, +1\}^N$ pour $0 \leq n \leq m$; le programme `Update` de la question précédente sera utile (on pourra l'adapter pour éviter de calculer $H(X_n)$ à chaque étape).]

(iii) Représenter les graphes d'une réalisation de $(Y_n)_{0 \leq n \leq m}$ avec $N = 25$ et $m = 100\,000$, pour les valeurs de $\beta = 0{,}02; 0{,}03; 0{,}04; 0{,}05; 0{,}06; 0{,}07$. Qu'observez-vous ?

Il s'avère qu'il y a une transition de phase *pour l'évolution de* $(Y_n)_{n \geq 0}$*, autour de la valeur* $\beta \sim 1/N$.

[Pour rendre le phénomène plus clair, on peut représenter les graphes d'une réalisation de $(Y_n)_{0 \leq n \leq m}$ pour $N = 100$ personnes, avec $m = 1\,000\,000$ et pour les valeurs $\beta = 0{,}01 \pm 0{,}003$, voir la Figure S9.3.]

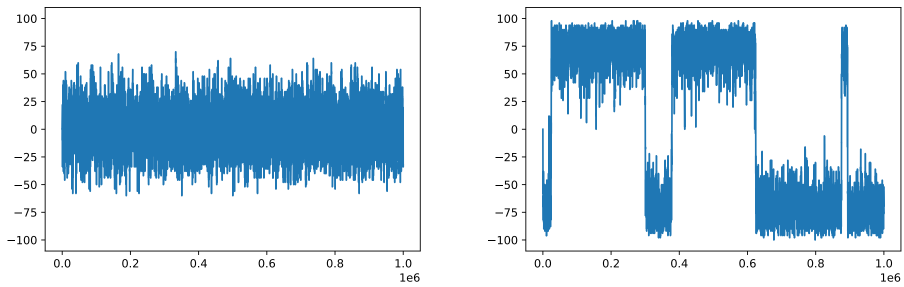


Fig. S9.3 (De l'Exercice 9.25) Représentation graphique de l'évolution de l'opinion générale $(Y_n)_{0\leq n\leq m}$ pour $N = 100$ personnes, avec $m = 10^6$: à gauche dans le cas $\beta = 0{,}007$ (l'opinion générale Y_n semble osciller autour de 0, c'est-à-dire qu'il n'y a jamais de majorité claire qui se dégage) ; à droite dans le cas $\beta = 0{,}013$ (l'opinion générale Y_n semble avoir une majorité de $+1$ ou de -1, en passant de temps en temps d'un cas à l'autre, de façon aléatoire).

Solution 9.25. (i) Il suffit de choisir un indice $i \in \{1,\dots,N\}$ de façon aléatoire et uniforme, de calculer $H_i(x) = H(x) - x_i$, et de déterminer si changer x_i en $+1$ ou -1 selon les règles décrites (c'est-à-dire $x_i = +1$ avec probabilité $(1+\mathrm{e}^{-2\beta H_i(x)})^{-1}$).

```
def Update(N,x,beta):
    i=rd.randint(0,N)
    Hi=sum(x)-x[i]
    if rd.random()<1/(1+exp(-2*beta*Hi)):
        x[i]=1
    else:
        x[i]=-1
    return x
```

(ii) Partons de $X_0 = ((-1)^i)_{1\leq i\leq N}$ et $Y_0 = H(X_0)$. En faisant évoluer X_n pour $0 \leq n \leq m$, on met à jour X_n et on garde en mémoire tous les $Y_n = H(X_n)$ pour $0 \leq n \leq m$. Au n-ème tour on met à jour X_n, en utilisant la même idée que ci-dessus : on choisit un i au hasard uniformément dans $\{1,\dots,N\}$, on calcule $H_i(X_n) = Y_n - X_n^{(i)}$ (où $X_n^{(i)}$ est la i-ème composante de X_n) et on met à jour la i-ème coordonnée de X_n, en la posant égale à $+1$ avec probabilité $1/(1+\mathrm{e}^{-2\beta H_i(X_n)})$ (dans ce cas on a $Y_{n+1} = H(X_{n+1}) = Y_n - X_n^{(i)} + 1$) ou à -1 avec probabilité $1/(1+\mathrm{e}^{2\beta H_i(X_n)})$ (dans ce cas on a $Y_{n+1} = H(X_{n+1}) = Y_n - X_k^{(i)} - 1$).

Un programme correspondant peut s'écrire de la manière suivante (la liste `Y` garde en mémoire les valeurs de $Y_n = H(X_n)$ pour $0 \leq n \leq m$).

```
def Chaine(N,beta,m):
    X=[(-1)**i for i in range(N)]
    h=sum(X)
    Y=[h]
    for n in range(m+1):
        i=rd.randint(0,N)
        h=Y[n]-X[i]
        if rd.random()< 1/(1+exp(-2*beta*h)):
            X[i]=1
            h+=1
        else:
            X[i]=-1
            h-=1
        Y.append(h)
    return Y
```

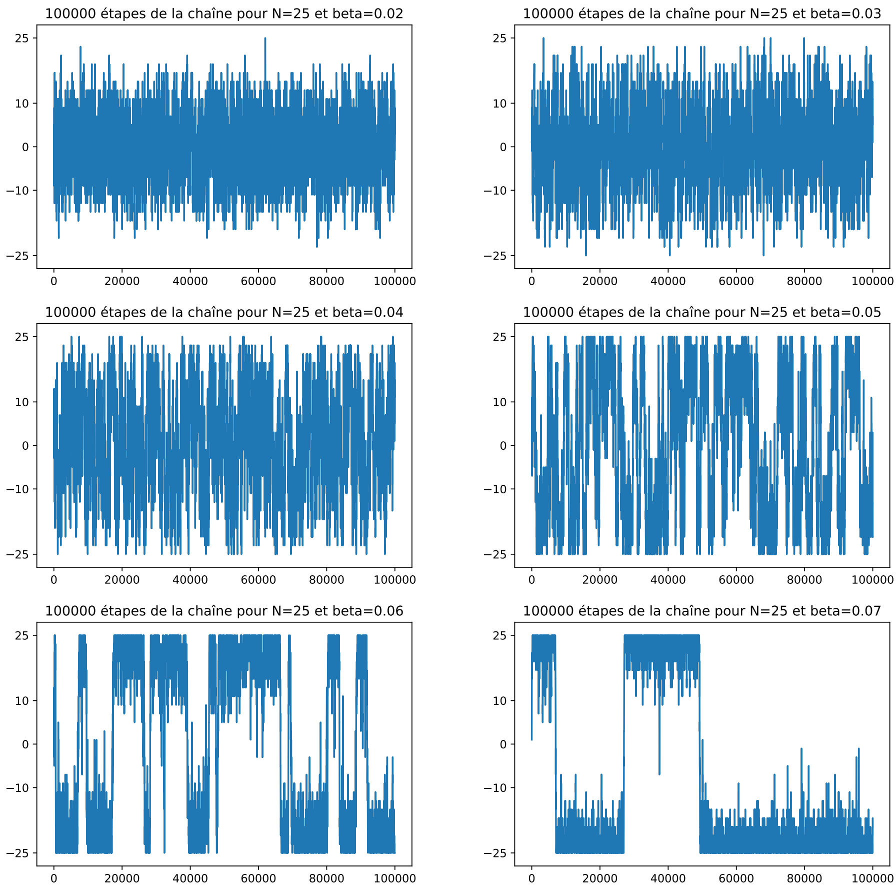


Fig. S9.4 (De l'Exercice 9.25) Représentation graphique de l'opinion générale $(Y_n)_{0\leq n\leq m}$ avec $N = 25$ personnes, pour $m = 10^5$, dans les cas $\beta = 0,02; 0,03; 0,04; 0,05,\ 0,06, 0,07$. Pour $\beta = 0,02; 0,03$, l'opinion générale Y_n semble osciller autour de 0, dans le sens où il n'y a jamais de majorité claire ; pour $\beta = 0,06; 0,07$, l'opinion générale Y_n semble avoir une majorité de $+1$ ou de -1, en passant de temps en temps d'un cas à l'autre, de façon aléatoire.

(iii) On donne ci-dessous un moyen pour représenter graphiquement les réalisations de $(Y_n)_{0\leq n\leq m}$. On force l'échelle verticale à être dans tous les cas entre -28 et $+28$ pour pouvoir comparer les différents graphes. Les graphes obtenus sont présentés dans la Figure S9.4.

```
N=25
m=10**5
for beta in [0.02+0.01*i for i in range(6)]:
  plt.axes()
  plt.ylim([-28,28])
  plt.yticks([-25,-10,0,10,25])
  plt.plot(Chaine(N,beta,m))
  plt.title("{} pas de Yn pour N={} et beta={}".format(m,N,beta))
  plt.show()
```

Table de la loi normale

La table suivante donne les valeurs de la fonction de répartition $\Phi(z)$ de la loi normale standard $\mathrm{N}(0,1)$, pour $0 \leq z \leq 3.5$. Rappelons que

$$\Phi(z) := \int_{-\infty}^{z} \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}\,\mathrm{d}x\,.$$

Les valeurs de $\Phi(z)$ pour $z < 0$ peuvent être retrouvées grâce à la formule

$$\Phi(z) = 1 - \Phi(-z)\,.$$

z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

Récapitulatif des lois usuelles

Lois usuelles discrètes

	Densité discrète $\mathrm{p}_X(k)$	*Espérance* $\mathrm{E}(X)$	*Variance* $\mathrm{Var}(X)$	*Fonction génératrice des moments* $\mathrm{M}_X(t) = \mathrm{E}(e^{tX})$
Binomiale $\mathrm{Bin}(n,p)$ $n \in \{1,2,\dots\}$ $p \in [0,1]$	$\binom{n}{k} p^k (1-p)^{n-k}$ $k \in \{0,1,\dots,n\}$	np	$np(1-p)$	$\big(pe^t + (1-p)\big)^n$
Bernoulli $\mathrm{Bern}(p)$ $\mathrm{Bin}(1,p)$ $p \in [0,1]$	$\begin{cases} p & \text{si } k=1 \\ 1-p & \text{si } k=0 \end{cases}$ $k \in \{0,1\}$	p	$p(1-p)$	$pe^t + (1-p)$
Poisson $\mathrm{Poi}(\lambda)$ $\lambda \in]0,\infty[$	$e^{-\lambda} \dfrac{\lambda^k}{k!}$ $k \in \{0,1,\dots\}$	λ	λ	$e^{\lambda(e^t-1)}$
Géométrique $\mathrm{Géom}(p)$ $p \in]0,1]$	$p(1-p)^{k-1}$ $k \in \{1,2,\dots\}$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$	$\begin{cases} \dfrac{p}{e^{-t}-(1-p)} & \text{si } t < \log\frac{1}{1-p} \\ +\infty & \text{si } t \geq \log\frac{1}{1-p} \end{cases}$

Lois usuelles à densité

	Densité $f_X(x)$	*Espérance* $\mathrm{E}(X)$	*Variance* $\mathrm{Var}(X)$	*Fonction génératrice des moments* $\mathrm{M}_X(t) = \mathrm{E}(e^{tX})$
Uniforme continue $\mathrm{U}(a,b)$ $a,b \in \mathbb{R},\ a<b$	$\dfrac{1}{b-a}$ $x \in]a,b[$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$	$\dfrac{e^{tb}-e^{ta}}{t(b-a)}$
Exponentielle $\mathrm{Exp}(\lambda)$ $\mathrm{Gamma}(1,\lambda)$ $\lambda \in]0,\infty[$	$\lambda e^{-\lambda x}$ $x \in]0,\infty[$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$	$\begin{cases} \dfrac{\lambda}{\lambda-t} & \text{si } t<\lambda \\ +\infty & \text{si } t \geq \lambda \end{cases}$
Gamma $\mathrm{Gamma}(\alpha,\lambda)$ $\alpha,\lambda \in]0,\infty[$	$\dfrac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$ $x \in]0,\infty[$	$\dfrac{\alpha}{\lambda}$	$\dfrac{\alpha}{\lambda^2}$	$\begin{cases} \left(\dfrac{\lambda}{\lambda-t}\right)^\alpha & \text{si } t<\lambda \\ +\infty & \text{si } t \geq \lambda \end{cases}$
Normale $\mathrm{N}(\mu,\sigma^2)$ $\mu \in \mathbb{R},\ \sigma \in]0,\infty[$	$\dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $x \in \mathbb{R}$	μ	σ^2	$e^{\mu t + \frac{\sigma^2}{2}t^2}$