Université de Montpellier - Faculté des Sciences

Année universitaire 2023-2024

HAX506X - Théorie des probabilités
Corrigé des feuilles de TD

- Chaque feuille de TD est composée de trois ou quatre exercices et correspond à une séance d'1h30.
- Les exercices sont à préparer à l'avance chez vous.
- Les exercices **Pour aller plus loin** sont des exercices d'approfondissement d'un niveau de difficulté supérieur aux autres. Ils ne seront pas corrigés pendant les séances. Il n'est pas nécessaire de les préparer à l'avance. De même, la feuille **Intermède** est à traiter uniquement si vous souhaitez approfondir les notions du cours. Elle ne sera pas corrigée en TD.
- Après chaque TD, un corrigé sera disponible sur la page Moodle du cours.

Nicolas Meyer

nicolas.meyer@umontpellier.fr

Table des matières

TD 1 - Espaces probabilisés

Exercice 1

1. Soit Ω un ensemble muni d'une tribu $\mathcal{F}$ et $x \in \Omega$. Montrer que $\delta_x : A \mapsto \delta_x(A) = \mathbb{1}_A(x)$ définit une probabilité sur $(\Omega, \mathcal{F})$.

   Vérifions tout d'abord que δ_x est bien une mesure. Tout d'abord, $\delta_x(\emptyset) = \mathbb{1}_\emptyset(x) = 0$. De plus, si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments deux à deux disjoints de $\mathcal{F}$, alors

   $$\delta_x\Big(\bigcup_{n=0}^{\infty} A_n\Big) = \mathbb{1}_{\cup_{n=0}^{\infty} A_n}(x) = \sum_{n=0}^{\infty} \mathbb{1}_{A_n}(x) = \sum_{n=0}^{\infty} \delta_x(A_n)\,,$$

   où on a utilisé l'égalité $\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B$, valable pour deux ensembles A et B disjoints. Ainsi, δ_x est une mesure et comme $\delta_x(\Omega) = \mathbb{1}_\Omega(x) = 1$, c'est bien une mesure de probabilité.

2. Soit $(\mathbb{P}_n)_{n\geq 1}$ une suite de mesures de probabilité sur un espace mesurable $(\Omega, \mathcal{F})$ et $(a_n)_{n\geq 1}$ une suite de réels dans $[0,1]$ tel que $\sum_{n=1}^{\infty} a_n = 1$. Montrer que $\sum_{n=1}^{\infty} a_n \mathbb{P}_n$ est une probabilité sur $(\Omega, \mathcal{F})$.

   On sait d'après le cours d'intégration qu'une combinaison linéaire (même infinie) de mesures avec des scalaires positifs est encore une mesure. Il reste donc à vérifier que la masse totale de $\sum_{n=1}^{\infty} a_n \mathbb{P}_n$ vaut 1, ce qui résulte de l'égalité

   $$\sum_{n=1}^{\infty} a_n \mathbb{P}_n(\Omega) = \sum_{n=1}^{\infty} a_n = 1\,.$$

   En particulier, une combinaison convexe de deux probabilités est une probabilité.

3. Soit I un intervalle de $\mathbb{R}$ de mesure de Lebesgue $\lambda(I)$ finie et strictement positive. Montrer que $\mathbb{P} : A \mapsto \frac{\lambda(A)}{\lambda(I)}$ définit une probabilité sur $(I, \mathcal{B}(I))$.

   On sait que la mesure de Lebesgue sur I est une mesure. En la multipliant par $\frac{1}{\lambda(I)} > 0$ on obtient encore une mesure. Enfin, la masse totale de $\mathbb{P}$ vaut

   $$\mathbb{P}(I) = \frac{\lambda(I)}{\lambda(I)} = 1\,,$$

   donc $\mathbb{P}$ est bien une mesure de probabilité.

4. Soit $(\Omega, \mathcal{F}, \mu)$ un espace mesuré (pas forcément de probabilité) et $f : \Omega \to [0, \infty[$ une fonction mesurable telle que $\int_\Omega f(\omega)\,\mathrm{d}\mu(\omega) = 1$. Montrer que l'application $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ définie par

   $$\mathbb{P}(A) = \int_\Omega f(\omega)\mathbb{1}_A(\omega)\,\mathrm{d}\mu(\omega)$$

   est une probabilité sur $(\Omega, \mathcal{F})$.

Remarquons déjà que $\mathbb{P}$ est à valeurs dans $[0,\infty[$ puisque f est positive. Par ailleurs,

$$\mathbb{P}(\emptyset) = \int_\Omega f(\omega)\mathbb{1}_\emptyset(\omega)\,\mathrm{d}\mu(\omega) = 0 \quad \text{et} \quad \mathbb{P}(\Omega) = \int_\Omega f(\omega)\,\mathrm{d}\mu(\omega) = 1\,,$$

par hypothèse sur la fonction f. Enfin si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments deux à deux disjoints de $\mathcal{F}$, alors on sait que

$$\bigcup_{n=0}^{\infty} A_n = \sum_{n=0}^{\infty} \mathbb{1}_{A_n}\,,$$

ce qui donne

$$\mathbb{P}\Big(\bigcup_{n=0}^{\infty} A_n\Big) = \int_\Omega f(\omega)\mathbb{1}_{\bigcup_{n=0}^{\infty} A_n}(\omega)\,\mathrm{d}\mu(\omega) = \int_\Omega f(\omega)\Big(\sum_{n=0}^{\infty}\mathbb{1}_{A_n}(\omega)\Big)\,\mathrm{d}\mu(\omega)\,.$$

Le théorème de convergence monotone permet alors de conclure :

$$\mathbb{P}\Big(\bigcup_{n=0}^{\infty} A_n\Big) = \sum_{n=0}^{\infty}\int_\Omega f(\omega)\mathbb{1}_{A_n}(\omega)\,\mathrm{d}\mu(\omega) = \sum_{n=0}^{\infty}\mathbb{P}(A_n)\,.$$

Ainsi $\mathbb{P}$ est bien une mesure de probabilité sur $(\Omega,\mathcal{F})$.

Exercice 2 On considère la mesure $\mathbb{P}$ sur $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ définie par

$$\mathbb{P} = \frac{1}{3}\delta_0 + \frac{1}{3}\mathbb{1}_{]0,2[}(x)\lambda\,,$$

où λ désigne la mesure de Lebesgue sur $\mathbb{R}$.

On peut imaginer que cette mesure représente le temps d'attente à un carrefour composé de trois feux tricolores, chaque feu restant au vert pendant une minute.

1. Montrer que $\mathbb{P}$ est une probabilité sur $\mathbb{R}$.

   La mesure $\mathbb{P}$ s'écrit $\mathbb{P} = \alpha\mathbb{P}_1 + (1-\alpha)\mathbb{P}_2$ avec

   $$\alpha = \frac{1}{3}\,,\quad \mathbb{P}_1 = \delta_0\,,\quad \mathbb{P}_2 = \frac{\mathbb{1}_{]0,2[}(x)}{2}\lambda\,.$$

   La masse de Dirac $\mathbb{P}_1$ est une probabilité. Concernant $\mathbb{P}_2$, la fonction $f : x \mapsto \frac{\mathbb{1}_{]-2,0[}(x)}{2}$ est une fonction mesurable positive d'intégrale

   $$\int_\mathbb{R} f(x)\,\mathrm{d}\lambda(x) = \int_0^2 \frac{\mathbb{1}_{]0,2[}(x)}{2}\,\mathrm{d}\lambda(x) = \frac{\lambda(]0,2[)}{2} = 1\,,$$

   donc une densité de probabilité. Ainsi, $\mathbb{P}_2$ est une mesure absolument continue par rapport à la mesure de Lebesgue de densité f. La question 2. de l'exercice précédent assure alors que $\mathbb{P}$ est une mesure de probabilité.

2. Calculer $\mathbb{P}([a,b])$ pour tout $0 \leq a < b \leq 2$.

Calculons séparément l'image du segment $[a,b]$ par les deux mesures qui composent $\mathbb{P}$. D'une part,

$$\mathbb{P}_1([a,b]) = \delta_0([a,b]) = \mathbb{1}_{\{0\}}(a)\,.$$

D'autre part,

$$\mathbb{P}_2([a,b]) = \int_{\mathbb{R}} \mathbb{1}_{[a,b]}(x)\frac{\mathbb{1}_{]0,2[}(x)}{2}\mathrm{d}\lambda(x) = \frac{1}{2}\int_{\mathbb{R}} \mathbb{1}_{[a,b]}(x)\mathrm{d}\lambda(x) = \frac{b-a}{2}\,.$$

On a utilisé l'égalité $\mathbb{1}_{[a,b]}(x)\mathbb{1}_{]0,2[}(x) = \mathbb{1}_{[a,b]}(x)$ valable si $0 < a < b < 2$. Si $a = 0$ (et/ou $b = 2$) alors il faut exclure la borne a (et/ou b) dans la dernière intégrale mais cela ne change pas le résultat final. Ainsi,

$$\mathbb{P}([a,b]) = \frac{1}{3}\mathbb{1}_{\{0\}}(a) + \frac{2}{3}\frac{b-a}{2} = \frac{1}{3}\mathbb{1}_{\{0\}}(a) + \frac{b-a}{3}\,.$$

3. Déterminer $\int_{\mathbb{R}} h(x)\,\mathrm{d}\mathbb{P}(x)$ pour toute fonction $h : \mathbb{R} \to \mathbb{R}$ mesurable positive.

Si $h : \mathbb{R} \to \mathbb{R}$ est une fonction mesurable positive, alors

$$\int_{\mathbb{R}} h(x)\,\mathrm{d}\mathbb{P}(x) = \frac{1}{3}\int_{\mathbb{R}} h(x)\,\mathrm{d}\delta_0(x) + \frac{2}{3}\int_{\mathbb{R}} h(x)f(x)\,\mathrm{d}\lambda(x) = \frac{1}{3}h(0) + \frac{2}{3}\int_0^2 h(x)\,\mathrm{d}\lambda(x)\,.$$

Remarquons que la première égalité résulte de la linéarité de l'intégrale *par rapport aux mesures* $\mathbb{P}_1$ *et* $\mathbb{P}_2$, propriété qui se montre sans souci en revenant à la définition de l'intégrale.

Exercice 3 Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé.

1. Soit A et B deux événements.

   (a) Montrer que
   $$\mathbb{P}(A) + \mathbb{P}(B) - 1 \leq \mathbb{P}(A \cap B) \leq \min(\mathbb{P}(A), \mathbb{P}(B))\,.$$

   On part de la relation $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$. Puisque $\mathbb{P}(A \cup B) \leq 1$, on obtient bien $\mathbb{P}(A) + \mathbb{P}(B) - 1 \leq \mathbb{P}(A \cap B)$.

   Concernant l'inégalité de droite, il suffit de remarquer que $A \cap B \subset A$, donc par croissance de la probabilité, on en déduit $\mathbb{P}(A \cap B) \leq \mathbb{P}(A)$. Le même raisonnement avec B donne $\mathbb{P}(A \cap B) \leq \mathbb{P}(B)$, d'où le résultat.

   (b) On considère le lancer d'un dé équilibré. Proposer un exemple d'événements A et B (d'intersection non vide) pour lequel l'inégalité de gauche est une égalité. Même question pour l'inégalité de droite.

   Pour l'inégalité de gauche, il suffit de considérer deux événements de réunion l'univers entier. Par exemple, $A =$ "obtenir 5 ou moins" et $B =$ "obtenir 2 ou plus". Dans ce cas, $\mathbb{P}(A) = \mathbb{P}(B) = 5/6$ et $\mathbb{P}(A \cap B) = 4/6$, ce qui donne bien

   $$\mathbb{P}(A) + \mathbb{P}(B) - 1 = \frac{5}{6} + \frac{5}{6} - 1 = \frac{4}{6} = \mathbb{P}(A \cap B)\,.$$

Concernant l'autre inégalité, il suffit de considérer $A \subset B$. Prenons par exemple A : "obtenir un 2" et B : "obtenir un chiffre pair". Dans ce cas,

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) = \frac{1}{6} = \min(\mathbb{P}(A), \mathbb{P}(B)) .$$

2. Montrer que si $A_1, \ldots, A_n$ sont n événements, alors

$$\mathbb{P}(A_1) + \cdots + \mathbb{P}(A_n) - (n-1) \leq \mathbb{P}(A_1 \cap \cdots \cap A_n) \leq \min_{i=1,\ldots,n} \mathbb{P}(A_i) .$$

Comme il est plus aisé d'obtenir des inégalités avec des unions, on considère l'événement complémentaire

$$(A_1 \cap \cdots \cap A_n)^c = A_1^c \cup \cdots \cup A_n^c .$$

La sous-σ-additivité donne alors

$$\begin{aligned}\mathbb{P}(A_1^c \cup \cdots \cup A_n^c) &\leq \mathbb{P}(A_1^c) + \cdots + \mathbb{P}(A_n^c) \\ &= (1 - \mathbb{P}(A_1)) + \cdots + (1 - \mathbb{P}(A_n)) = n - (\mathbb{P}(A_1) + \cdots + \mathbb{P}(A_n)) .\end{aligned}$$

Ceci permet d'en déduire la première inégalité de l'énoncé :

$$\mathbb{P}(A_1 \cap \cdots \cap A_n) = 1 - \mathbb{P}((A_1 \cap \cdots \cap A_n)^c) \geq 1 - n + \mathbb{P}(A_1) + \cdots + \mathbb{P}(A_n) .$$

Concernant l'inégalité de droite, il suffit de remarquer que pour tout $i = 1, \ldots, n$, on a l'inclusion $A_1 \cap \cdots \cap A_n \subset A_i$, ce qui donne, par croissance de la probabilité :

$$\mathbb{P}(A_1 \cap \cdots \cap A_n) \leq \mathbb{P}(A_i) .$$

Ceci étant vrai pour tout i, on en déduit l'inégalité voulue :

$$\mathbb{P}(A_1 \cap \cdots \cap A_n) \leq \min_{i=1,\ldots,n} \mathbb{P}(A_i) .$$

Notons qu'on aurait également pu montrer le résultat par récurrence sur n.

Pour aller plus loin

Soit $(\Omega, \mathcal{F})$ un espace mesurable. Pour $x \in \Omega$ on définit l'atome de x par

$$A(x) = \bigcap_{A \in \mathcal{F} \,:\, x \in A} A .$$

Remarquons tout d'abord que $A(x)$ correspond à l'intersection des éléments de la tribu qui contiennent x, donc en particulier $x \in A(x)$.

1. Rappeler pourquoi l'ensemble $\mathcal{P}(\mathbb{N})$ formé des parties de $\mathbb{N}$ n'est pas dénombrable.
   Un ensemble $E \subset \mathbb{N}$ est entièrement caractérisé par sa fonction indicatrice $\mathbb{1}_E : \mathbb{N} \to \{0,1\}$. Autrement dit, on a une bijection
   $$\begin{array}{rccc} \Phi_1 : & \mathcal{P}(\mathbb{N}) & \to & \{0,1\}^{\mathbb{N}} \\ & E & \mapsto & \mathbb{1}_E\,. \end{array}$$
   Or l'ensemble $\{0,1\}^{\mathbb{N}}$ est lui-même en bijection avec $[0,1]$ via l'écriture en binaire d'un nombre décimal :
   $$\begin{array}{rccc} \Phi_2 : & \{0,1\}^{\mathbb{N}} & \to & [0,1] \\ & (u_n)_{n\geq 0} & \mapsto & 0,u_0u_1u_2\ldots\,. \end{array}$$
   Bref, l'application $\Phi_2 \circ \Phi_1$ est une bijection entre $\mathcal{P}(\mathbb{N})$ et $[0,1]$. Comme $[0,1]$ n'est pas dénombrable, on en déduit que $\mathcal{P}(\mathbb{N})$ n'est pas dénombrable.
2. Montrer que les $A(x)$ sont soit égaux soit disjoints.
   On va montrer que si $x, y \in \Omega$ sont tels que $A(x) \cap A(y) \neq \emptyset$, alors $A(x) = A(y)$. Soit donc $z \in A(x) \cap A(y)$.
   - Si B est un ensemble contenant x, alors
     $$B \supset A(x) \ni z\,.$$
     donc B contient z.
   - Si B est un ensemble ne contenant pas x, alors B^c contient x donc
     $$B^c \supset A(x) \ni z\,.$$
     donc B ne contient pas z.

   Bref, on vient de montrer que les ensembles contenant x coïncident avec les ensembles contenant z, ce qui implique que $A(x) = A(z)$. On montre de manière similaire que $A(y) = A(z)$, ce qui permet de conclure que $A(x) = A(y)$.
   Ainsi, les ensembles $A(x)$ sont soit égaux soit disjoints. En particulier, les $A(x)$ forment une partition de Ω (à condition de ne compter qu'une seule fois ceux qui sont égaux).
3. Supposons que $\mathcal{F}$ soit au plus dénombrable. Montrer que $\mathcal{F}$ contient ses atomes puis que chaque élément de $\mathcal{F}$ est une réunion au plus dénombrable d'atomes. En déduire que $\mathcal{F}$ n'est pas infini dénombrable.
   Si $\mathcal{F}$ est au plus dénombrable, alors un atome $A(x)$ s'écrit comme une intersection au plus dénombrable d'éléments de $\mathcal{F}$ donc $A(x) \in \mathcal{F}$. En particulier l'ensemble $\mathcal{A}$ des atomes est au plus dénombrable. On l'écrit alors
   $$\mathcal{A} = \bigcup_{i \in I} A_i\,.$$
   avec I fini si $\mathcal{A}$ est fini, et $I = \mathbb{N}$ si $\mathcal{A}$ est infini dénombrable.
   Par suite, comme les A_x forment une partition et contiennent x on peut écrire tout élément B de la tribu $\mathcal{F}$ sous la forme
   $$B = \bigcup_{x \in B} A(x)\,.$$

Donc on peut écrire tout élément $B \in \mathcal{F}$ comme une union d'atomes $A(x)$ et cette union est au plus dénombrable puisque l'ensemble des atomes est dénombrable. Autrement pour tout $B \in \mathcal{F}$ il existe $J \subset I$ tel que

$$B = \bigcup_{j \in J} A_j \,,$$

et cette écriture est unique. On a donc mis en évidence une bijection

$$\begin{array}{ccc} \mathcal{P}(I) & \to & \mathcal{F} \\ J & \mapsto & \bigcup_{j \in J} A_j \,. \end{array}$$

Or $\mathcal{P}(I)$ est soit fini (si I est fini), soit non dénombrable (si $I = \mathbb{N}$) d'après la question 1. On obtient donc une contradiction avec l'hypothèse "$\mathcal{F}$ est au plus dénombrable".

Ceci montre donc qu'une tribu est soit finie, soit non dénombrable.

TD 2 - Variables aléatoires

Exercice 1 On tire deux fois avec remise dans une urne contenant trois boules numérotées $1, 2, 3$. On désigne par X la somme des résultats obtenus. Montrer que X est une variable aléatoire discrète entre un espace probabilisé et un espace mesurable à déterminer. Donner la loi de X.

On considère l'univers fini $\Omega = \{1, 2, 3\}^2$ de cardinal 9 muni de la tribu pleine $\mathcal{F} = \mathcal{P}(\Omega)$ et de la probabilité uniforme $\mathbb{P}$, chaque tirage étant équiprobable. On définit également l'ensemble $\Omega' = \{2, 3, 4, 5, 6\}$ que l'on munit de la tribu pleine $\mathcal{F}' = \mathcal{P}(\Omega')$ et on considère l'application $X : \Omega \to \Omega'$ définie par

$$X(\omega_1, \omega_2) = \omega_1 + \omega_2 .$$

Alors pour tout $\omega' \in \Omega'$,

$$\{X = \omega'\} = \{(\omega_1, \omega_2) \in \Omega : \omega_1 + \omega_2 = \omega'\} \in \mathcal{F} ,$$

donc X est une variable aléatoire (en fait, si on choisit la tribu pleine sur l'espace de départ alors tous les événements $\{X = \omega'\}$ sont dans $\mathcal{F}$, donc X est nécessairement mesurable). Comme X prend ses valeurs dans l'univers fini Ω', c'est une variable aléatoire discrète.

La loi de X est donc entièrement caractérisée par les quantités

$$\mathbb{P}_X(\{\omega'\}) = \mathbb{P}(X = \omega') = \mathbb{P}(\{(\omega_1, \omega_2) \in \Omega : X(\omega_1, \omega_2) = \omega'\}) ,$$

pour tout $\omega' \in \Omega'$. On trouve alors

$$\begin{aligned}
\mathbb{P}_X(\{2\}) &= \mathbb{P}(X = 2) = \mathbb{P}(\{(1, 1)\}) = \frac{1}{9} , \\
\mathbb{P}_X(\{3\}) &= \mathbb{P}(X = 3) = \mathbb{P}(\{(2, 1), (1, 2)\}) = \frac{2}{9} , \\
\mathbb{P}_X(\{4\}) &= \mathbb{P}(X = 4) = \mathbb{P}(\{(1, 3), (3, 1), (2, 2)\}) = \frac{3}{9} , \\
\mathbb{P}_X(\{5\}) &= \mathbb{P}(X = 5) = \mathbb{P}(\{(2, 3), (3, 2)\}) = \frac{2}{9} , \\
\mathbb{P}_X(\{6\}) &= \mathbb{P}(X = 6) = \mathbb{P}(\{(3, 3)\}) = \frac{1}{9} .
\end{aligned}$$

Exercice 2 On suppose que la basketteuse française Marine Johannès a une probabilité 0.8 de marquer un lancer franc.

1. Lors d'un entraînement, elle tente une série de 10 lancers francs. On désigne par X la variable aléatoire donnant le nombre de paniers marqués.
   (a) Donner la loi de X.
   L'expérience consiste à répéter $n = 10$ fois une épreuve de Bernoulli dont la probabilité de succès est $p = 0.8$. Ainsi, X suit une loi binomiale de paramètres $n = 10$ et $p = 0.8$.

Rappelons que cela signifie que X est une variable aléatoire discrète à valeurs dans $\{0, \dots, 10\}$ avec

$$p_k = \mathbb{P}(X = k) = \binom{10}{k} 0.8^k 0.2^{10-k}\,, \quad k = 0, \dots, 10\,.$$

(b) Déterminer la probabilité que Marine Johannès marque huit paniers ou plus.

La probabilité recherchée est

$$\mathbb{P}(X \geq 8) = \mathbb{P}(X \in \{8, 9, 10\}) = p_8 + p_9 + p_{10} \approx 0.68\,.$$

(c) Donner l'espérance et la variance de X.

L'espérance d'une loi binomiale est donnée par $\mathbb{E}[X] = np = 8$, tandis que la variance vaut $\operatorname{Var}(X) = np(1-p) = 1.6$.

2. Lors d'un autre entraînement, Marine Johannès décide de tirer jusqu'à ce qu'elle inscrive un panier. On désigne par Y la variable aléatoire donnant le nombre d'essais nécessaires.

(a) Donner la loi de Y.

L'expérience consiste désormais à répéter une épreuve de Bernoulli de paramètre $p = 0.8$ jusqu'au premier succès. Ainsi, Y suit une loi géométrique de paramètre $p = 0.8$: Y prend ses valeurs dans $\mathbb{N}^*$ avec

$$q_k = \mathbb{P}(Y = k) = 0.8 \times 0.2^{k-1}\,, \quad k \in \mathbb{N}^*\,.$$

(b) Déterminer la probabilité que Marine Johannès ait besoin de strictement plus de 3 essais pour marquer son premier panier.

Cela revient à calculer

$$\mathbb{P}(Y > 3) = \mathbb{P}(Y \in \{4, 5, \dots\}) = \sum_{k=4}^{\infty} q_k = \sum_{k=4}^{\infty} 0.8 \times 0.2^{k-1} = 0.8 \times 0.2^3 \sum_{k=0}^{\infty} 0.2^k\,.$$

On reconnaît une série géométrique de raison 0.2 et on obtient

$$\mathbb{P}(Y > 3) = \frac{0.8 \times 0.2^3}{1 - 0.2} = 0.2^3 = 0.008\,.$$

(c) Donner l'espérance et la variance de Y.

L'espérance d'une loi géométrique est donnée par $\mathbb{E}[Y] = 1/p = 1.25$ et sa variance par $(1-p)/p^2 \approx 0.31$.

Exercice 3 On tire deux cartes dans un jeu de 52 cartes. On considère le vecteur aléatoire $X = (X_1, X_2)$ où X_1 donne le nombre de cartes rouges tirées et X_2 le nombre de cartes noires.

1. Quelles valeurs peut prendre le vecteur X ?
   Lors du tirage, on peut obtenir deux cartes rouges, deux cartes noires ou bien une carte rouge et une carte noire. Ainsi, le vecteur X prend les valeurs $(2,0), (0,2), (1,1)$. Sa loi est donc discrète.
2. Déterminer la loi de X.
   Il suffit de déterminer les probabilités $p_{2,0}$, $p_{0,2}$, $p_{1,1}$ avec
   $$p_{i,j} = \mathbb{P}(X = (i,j)) = \mathbb{P}(X_1 = i, X_2 = j)\,.$$
   Il y a $\binom{52}{2} = 1326$ choix de cartes possibles. Obtenir deux cartes rouges revient à choisir deux cartes parmi les 26 cartes rouges, donc
   $$p_{2,0} = \mathbb{P}(X_1 = 2, X_2 = 0) = \frac{\binom{26}{2}}{1326} = \frac{25}{102} \approx 0.245\,.$$
   De même, $p_{0,2} = \mathbb{P}(X_1 = 0, X_2 = 2) = \frac{25}{102} \approx 0.245$.
   Enfin, pour obtenir une carte rouge et une carte noire, il suffit de choisir une carte parmi les 26 rouges et une carte parmi les 26 noires, soit
   $$p_{1,1} = \mathbb{P}(X_1 = 1, X_2 = 1) = \frac{\binom{26}{1}\binom{26}{1}}{1326} = \frac{26}{51} \approx 0.51\,.$$
   On vérifie que $p_{2,0} + p_{0,2} + p_{1,1} = 1$.

Pour aller plus loin

On dit qu'une variable aléatoire réelle positive X est sans mémoire si pour tout $s, t \geq 0$,
$$\mathbb{P}(X > t + s \mid X > t) = \mathbb{P}(X > s)\,.$$
Montrer que les seules variables aléatoires sans mémoire sont les variables aléatoires de loi exponentielle et la variable aléatoire constante égale à 0.

Il est bien connu que les lois exponentielles vérifient cette propriété. En effet, si $X \sim \mathcal{E}(\theta)$, alors $\mathbb{P}(X > t) = \mathrm{e}^{-\theta t}$ pour tout $t \geq 0$, ce qui donne le résultat voulu :
$$\mathbb{P}(X > t + s \mid X > t) = \frac{\mathbb{P}(X > t + s, X > t)}{\mathbb{P}(X > t)} = \frac{\mathbb{P}(X > t + s)}{\mathbb{P}(X > t)} = \frac{\mathrm{e}^{-\theta(t+s)}}{\mathrm{e}^{-\theta t}} = \mathrm{e}^{-\theta s} = \mathbb{P}(X > s)\,.$$

Réciproquement, soit X un variable aléatoire réelle positive vérifiant la relation donnée en énoncé. Notons $g : \mathbb{R} \to \mathbb{R}$ la fonction définie par $g(t) = \mathbb{P}(X > t)$. L'hypothèse sur X se réécrit alors
$$\frac{\mathbb{P}(X > t + s, X > t)}{\mathbb{P}(X > t)} = \mathbb{P}(X > s)\,,$$
ou encore
$$\mathbb{P}(X > t + s) = \mathbb{P}(X > t)\mathbb{P}(X > s)\,,$$
c'est-à-dire $g(t + s) = g(t)g(s)$.

Posons alors $a = g(1) = \mathbb{P}(X > 1)$. Une récurrence sur la relation fonctionnelle vérifiée par g implique que pour tout $t_1, \ldots, t_n \geq 0$, on a
$$g(t_1 + \cdots + t_n) = g(t_1) \cdots g(t_n)\,.$$

- Avec $t_1 = \cdots = t_n = 1$, on obtient $g(n) = g(1)^n = a^n$ pour tout $n \geq 1$.
- Avec $t_1 = \cdots = t_n = \frac{1}{n}$, $n \geq 1$, on obtient

$$g(1) = g\Big(\frac{1}{n}\Big)^n ,$$

  donc $g\Big(\frac{1}{n}\Big) = a^{1/n}$.
- Avec $t_1 = \cdots = t_p = \frac{1}{n}$, $p \geq 1$, $n \geq 0$, on obtient

$$g\Big(\frac{p}{n}\Big) = g\Big(\frac{1}{n}\Big)^p = a^{p/n} .$$

Bref, pour tout $q \in \mathbb{Q}_+$, $g(q) = a^q$.

Maintenant, soit $x > 0$. Comme $\mathbb{Q}_+^*$ est dense dans $\mathbb{R}_+^*$, il existe une suite croissante $(x_n)_{n\geq 1}$ dans $\mathbb{Q}_+^*$ telle que $x_n \to x$. Par suite, la fonction de répartition F_X étant continue à droite, la fonction $g = 1 - F_X$ est continue à gauche, donc

$$g(x_n) \xrightarrow[n\to\infty]{} g(x) .$$

Le membre de gauche vaut a^{x_n} d'après ce qui précède et converge donc vers a^x. Par unicité de la limite on obtient $g(x) = a^x$ pour tout $x > 0$, où a^x est interprété comme 0 si $a = 0$.

Distinguons alors deux cas.

- <u>Cas 1 : $a = \mathbb{P}(X > 1) = 0$</u>
  Cela signifie que $\mathbb{P}(X > 0) = 0$. Comme X est une variable aléatoire positive, on en déduit que $\mathbb{P}(X = 0) = 1$, donc X a pour loi une masse de Dirac en 0. Notons qu'on peut confondre cette loi avec une loi exponentielle de paramètre infini.
- <u>Cas 1 : $a = \mathbb{P}(X > 1) \in]0, 1]$</u>
  Posons alors $\theta = -\ln(a) > 0$. Les calculs précédents donne que pour tout $t > 0$

$$\mathbb{P}(X > t) = g(t) = a^t = \mathrm{e}^{-\theta t} .$$

  Ceci prouve que la fonction de répartition de X est de la forme

$$F_X(t) = 1 - \mathbb{P}(X > t) = 1 - \mathrm{e}^{-\theta t} , \quad t > 0 ,$$

  Par ailleurs, F_X est nulle sur $]-\infty, 0[$ puisque X est positive par hypothèse. Par continuité à droite de F_X, on obtient $F_X(0) = 0$. Ainsi, F_X est la fonction de répartition d'une loi exponentielle de paramètre θ, donc $X \sim \mathcal{E}(\theta)$.

TD 3 - Moments d'une variable aléatoire

Exercice 1 Soit X une variable aléatoire discrète de loi $\mathbb{P}_X = \sum_{k=0}^{\infty} p_k \delta_{x_k}$, avec (x_k) une suite de réels et p_k des réels positifs qui somment à 1.

1. Déterminer, sous réserve d'existence, la valeur de $\mathbb{E}[X^p]$ pour tout entier p.

   Le calcul des moments se fait via la formule de transfert :

   $$\mathbb{E}[X^p] = \int_\Omega X(\omega)^p \, \mathrm{d}\mathbb{P}(\omega) = \int_\mathbb{R} x^p \, \mathrm{d}\mathbb{P}_X(x) = \int_\mathbb{R} x^p \sum_{k=0}^{\infty} p_k \, \mathrm{d}\delta_{x_k} = \sum_{k=0}^{\infty} x_k^p p_k \, .$$

   Rappelons au passage que $\mathbb{E}[X^p]$ est bien définie si $\mathbb{E}[|X|^p] < \infty$ ce qui signifie ici que

   $$\sum_{k=0}^{\infty} |x_k|^p p_k < \infty \, .$$

2. En déduire les valeurs de $\mathbb{E}[X]$ pour X suivant respectivement une loi de Bernoulli, binomiale, et de Poisson.

   Si X suit une loi de Bernoulli, c'est-à-dire $\mathbb{P}_X = (1-p)\delta_0 + p\delta_1$, la formule précédente donne $\mathbb{E}[X] = p$.

   Si X suit une loi binomiale, alors $\mathbb{P}_X = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \delta_k$. On obtient alors

   $$\begin{aligned}
   \mathbb{E}[X] &= \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} k \\
   &= \sum_{k=1}^{n} \frac{n!}{(n-k)!(k-1)!} p^k (1-p)^{n-k} \\
   &= np \sum_{k=1}^{n} \frac{(n-1)!}{((n-1)-(k-1))!(k-1)!} p^{k-1} (1-p)^{n-k} \, .
   \end{aligned}$$

   Le changement de variable $j = k-1$ et la formule de binôme de Newton donnent alors

   $$\sum_{j=0}^{n-1} \frac{(n-1)!}{((n-1)-j)!j!} p^j (1-p)^{n-1-j} = (p + (1-p))^{n-1} = 1 \, .$$

   Ainsi, $\mathbb{E}[X] = np$.

   Enfin, si X suit une loi de Poisson de paramètre λ, alors $\mathbb{P}_X = \sum_{k=0}^{\infty} \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!} \delta_k$, et donc

   $$\mathbb{E}[X] = \sum_{k=0}^{\infty} k \, \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!} = \mathrm{e}^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!} = \mathrm{e}^{-\lambda} \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \mathrm{e}^{-\lambda} \lambda \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = \lambda \, .$$

Exercice 2

1. Calculer les moments à tout ordre d'une variable aléatoire de loi exponentielle de paramètre $\lambda > 0$. On pourra commencer par le cas $\lambda = 1$.

   Pour $\lambda = 1$, la densité d'une variable aléatoire $X \sim \mathcal{E}(1)$ est $x \mapsto \mathrm{e}^{-x}\,\mathbb{1}_{[0,\infty[}(x)$. On remarque tout d'abord que les moments de X existent à tout ordre puisque la fonction $x \mapsto x^k\,\mathrm{e}^{-x}$ est intégrable en $+\infty$ pour tout entier k.

   Commençons par calculer l'espérance de X. Une intégration par parties donne

   $$\mathbb{E}[X] = \int_0^\infty x\,\mathrm{e}^{-x}\,\mathrm{d}x = \Big[-x\,\mathrm{e}^{-x}\Big]_0^\infty - \int_0^\infty -\mathrm{e}^{-x}\,\mathrm{d}x = 0 + \Big[-\mathrm{e}^{-x}\Big]_0^\infty = 1\,.$$

   Par suite, pour $k \geq 1$, on obtient

   $$\mathbb{E}[X^k] = \int_0^\infty x^k\,\mathrm{e}^{-x}\,\mathrm{d}x = \Big[-x^k\,\mathrm{e}^{-x}\Big]_0^\infty - \int_0^\infty -kx^{k-1}\,\mathrm{e}^{-x}\,\mathrm{d}x = k\mathbb{E}[X^{k-1}]\,.$$

   En combinant ces deux relations on en déduit que $\mathbb{E}[X^k] = k!$.

   Dans le cas général, pour $Y \sim \mathcal{E}(\lambda)$, le changement de variable $t = \lambda y$ donne

   $$\mathbb{E}[Y^k] = \int_0^\infty y^k\lambda\,\mathrm{e}^{-\lambda y}\,\mathrm{d}y = \int_0^\infty \left(\frac{t}{\lambda}\right)^k \mathrm{e}^{-t}\,\mathrm{d}t = \frac{1}{\lambda^k}\mathbb{E}[X^k] = \frac{k!}{\lambda^k}\,.$$

2. Soit X une variable aléatoire réelle positive. À l'aide du théorème de Fubini, montrer que pour tout entier k,

   $$\mathbb{E}[X^k] = k\int_0^\infty t^{k-1}\mathbb{P}(X > t)\,\mathrm{d}t\,.$$

   On part de l'égalité de droite qui peut se réécrire

   $$k\int_0^\infty t^{k-1}\mathbb{P}(X > t)\,\mathrm{d}t = k\int_0^\infty t^{k-1}\mathbb{E}[\mathbb{1}_{\{X>t\}}]\,\mathrm{d}t = k\int_0^\infty\int_\Omega t^{k-1}\mathbb{1}_{\{X(\omega)>t\}}\,\mathrm{d}\mathbb{P}(\omega)\,\mathrm{d}t\,.$$

   Le théorème de Fubini appliquée à la fonction positive $(t,\omega) \mapsto kt^{k-1}\mathbb{1}_{\{X(\omega)>t\}}$ donne

   $$\int_\Omega\int_0^\infty kt^{k-1}\mathbb{1}_{\{X(\omega)>t\}}\,\mathrm{d}t\,\mathrm{d}\mathbb{P}(\omega) = \int_\Omega\int_0^{X(\omega)} kt^{k-1}\,\mathrm{d}t\,\mathrm{d}\mathbb{P}(\omega)\,.$$

   On en déduit alors le résultat souhaité

   $$k\int_0^\infty t^{k-1}\mathbb{P}(X > t)\,\mathrm{d}t = \int_\Omega \left[t^k\right]_0^{X(\omega)}\mathrm{d}\mathbb{P}(\omega) = \int_\Omega X(\omega)^k\,\mathrm{d}\mathbb{P}(\omega) = \mathbb{E}[X^k]\,,$$

   d'où le résultat.

3. En déduire la relation $\Gamma(k) = (k-1)!$ pour tout entier $k \geq 1$, où Γ désigne la fonction Gamma d'Euler :

   $$\Gamma(x) = \int_0^\infty t^{x-1}\,\mathrm{e}^{-t}\,\mathrm{d}t\,, \quad x > 0\,.$$

   Remarquons tout d'abord que la fonction Gamma d'Euler est bien définie. En effet, on a $t^{x-1}\,\mathrm{e}^{-t} = o(t^{-2})$ quand $t \to \infty$ et $t^{x-1}\,\mathrm{e}^{-t} \sim t^{x-1}$ quand $t \to 0$ avec $x - 1 > -1$.

En appliquant la relation de la question 2. à une variable aléatoire positive X de loi exponentielle de paramètre 1 on obtient, pour tout entier $k \geq 1$,

$$k! = k \int_0^{\infty} t^{k-1} \mathbb{P}(X > t)\, \mathrm{d}t\,.$$

Il reste à calculer la probabilité $\mathbb{P}(X > t)$:

$$\mathbb{P}(X > t) = \int_t^{\infty} \mathrm{e}^{-x}\ \mathrm{d}x = \big[- \mathrm{e}^{-x} \big]_t^{\infty} = \mathrm{e}^{-t}\ .$$

L'intégrale précédente correspond donc, à un facteur k près, à la fonction Gamma d'Euler évaluée en k. On en déduit que

$$\Gamma(k) = \int_0^{\infty} t^{k-1}\, \mathrm{e}^{-t}\ \mathrm{d}t = (k-1)!\,, \quad k \geq 1\,.$$

Exercice 3 Soit X une variable aléatoire de loi $\mathcal{N}(\mu, \sigma^2)$. Comparer la probabilité

$$\mathbb{P}(|X - \mu| \geq c\sigma)$$

donnée dans la table avec les majorations obtenues par l'inégalité de Bienaymé-Tchebychev pour

$$c = 0.5\,, \quad c = 1\,, \quad c = 1.5\,, \quad c = 2\,, \quad c = 2.5\,.$$

On pose $U = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$. La probabilité à majorer s'écrit alors

$$\mathbb{P}(|X-\mu| \geq c\sigma) = \mathbb{P}(|U| \geq c) = \mathbb{P}(U \geq c) + \mathbb{P}(U \leq -c) = 2\mathbb{P}(U \geq c) = 2(1 - \Phi(c))\,,$$

où Φ est la fonction de répartition de la loi normale centrée réduite. Ces quantités s'obtiennent directement par lecture de la table.

Par ailleurs, l'inégalité de Bienaymé-Tchebychev assure que

$$\mathbb{P}(|X - \mu| \geq c\sigma) \leq \frac{\sigma^2}{(c\sigma)^2} = \frac{1}{c^2}\,.$$

On obtient alors

c	0.5	1	1.5	2	2.5
$2(1-\Phi(c))$	0.6170	0.3174	0.1336	0.0456	0.0124
$1/c^2$	4	1	0.444	0.25	0.16

On remarque que l'inégalité Bienaymé-Tchebychev est inutile pour des petites valeurs de c : elle majore des probabilités par des quantités plus grandes que 1 ! Pour des valeurs plus grandes de c, l'inégalité est plus intéressante mais elle reste grossière.

Pour aller plus loin

Le théorème de Weierstrass dit que toute fonction continue $f : [0,1] \to \mathbb{R}$ est limite uniforme de fonctions polynomiales (et l'énoncé se généralise alors à un segment $[a,b]$ quelconque). On souhaite démontrer ce résultat avec la théorie des probabilités en explicitant une suite de fonctions polynomiales.

Soit donc $f : [0,1] \to \mathbb{R}$ continue. On se fixe $x \in [0,1]$ et on considère une variable aléatoire $B_n(x)$ de loi binomiale de paramètre n et x.

1. Que vaut $\mathbb{E}[f(\frac{B_n(x)}{n})]$?

   Comme $B_n(x)$ suit une loi binomiale de paramètre n et x, la formule de transfert donne
   $$\mathbb{E}\Big[f\Big(\frac{B_n(x)}{n}\Big)\Big] = \sum_{k=0}^{n} \binom{n}{k} f\Big(\frac{k}{n}\Big) x^k (1-x)^{n-k}\,.$$
   L'objectif va être de montrer que cette fonction polynomiale en x converge uniformément vers f.

2. Soit $\epsilon > 0$. Montrer qu'il existe $\delta > 0$ tel que
   $$\Big|\mathbb{E}\Big[f\Big(\frac{B_n(x)}{n}\Big)\Big] - f(x)\Big| \leq 2\|f\|_\infty \mathbb{P}\Big(\Big|\frac{B_n(x)}{n} - x\Big| \geq \delta\Big) + \epsilon\,.$$
   Commençons par utiliser l'inégalité classique
   $$\Big|\mathbb{E}\Big[f\Big(\frac{B_n(x)}{n}\Big)\Big] - f(x)\Big| = \Big|\mathbb{E}\Big[f\Big(\frac{B_n(x)}{n}\Big) - f(x)\Big]\Big| \leq \mathbb{E}\Big[\Big|f\Big(\frac{B_n(x)}{n}\Big) - f(x)\Big|\Big]\,.$$
   La fonction f étant continue sur le compact $[0,1]$, elle est uniformément continue d'après le théorème de Heine. Ainsi, il existe $\delta > 0$ tel que pour tout $x, y \in [0,1]$,
   $$|x-y| \leq \delta \implies |f(x) - f(y)| \leq \epsilon\,.$$
   Par suite, on décompose l'espérance sous la forme
   $$\mathbb{E}\Big[\Big|f\Big(\frac{B_n(x)}{n}\Big) - f(x)\Big|\Big] = \mathbb{E}\Big[\Big|f\Big(\frac{B_n(x)}{n}\Big) - f(x)\Big|\mathbb{1}_{\left|\frac{B_n(x)}{n}-x\right|>\delta}\Big] + \mathbb{E}\Big[\Big|f\Big(\frac{B_n(x)}{n}\Big) - f(x)\Big|\mathbb{1}_{\left|\frac{B_n(x)}{n}-x\right|\leq\delta}\Big]\,.$$
   On majore alors brutalement la quantité dans la première espérance par $2\,\|f\|_\infty \mathbb{1}_{\left|\frac{B_n(x)}{n}-x\right|>\delta}$ en remarquant que $\|f\|_\infty$ existe car f est continue sur un compact. Par ailleurs, l'uniforme continuité permet de majorer la quantité dans la deuxième espérance par ϵ (et l'indicatrice est majorée par 1). En combinant toutes ces inégalités, on obtient
   $$\Big|\mathbb{E}\Big[f\Big(\frac{B_n(x)}{n}\Big)\Big] - f(x)\Big| \leq \mathbb{E}\Big[2\,\|f\|_\infty \mathbb{1}_{\left|\frac{B_n(x)}{n}-x\right|>\delta}\Big] + \epsilon = 2\,\|f\|_\infty \mathbb{P}\Big(\Big|\frac{B_n(x)}{n} - x\Big| > \delta\Big) + \epsilon\,.$$

3. Conclure grâce à l'inégalité de Bienaymé-Tchebychev.

   Comme $\mathbb{E}[B_n(x)] = nx$, on peut écrire
   $$\mathbb{P}\Big(\Big|\frac{B_n(x)}{n} - x\Big| > \delta\Big) = \mathbb{P}\Big((B_n(x) - nx)^2 > n^2\delta^2\Big) = \mathbb{P}\Big((B_n(x) - \mathbb{E}[B_n(x)])^2 > n^2\delta^2\Big)$$

L'inégalité de Bienaymé-Tchebychev assure que

$$\mathbb{P}\Big(\big|\frac{B_n(x)}{n} - x\big| > \delta\Big) = \mathbb{P}\Big((B_n(x) - \mathbb{E}[B_n(x)])^2 > n^2\delta^2\Big) \leq \frac{\mathrm{Var}(B_n(x))}{n^2\delta^2}\,.$$

Comme $\mathrm{Var}(B_n(x)) = nx(1-x)$ est majorée par $n/4$ pour $x \in [0,1]$ on en déduit que

$$\mathbb{P}\Big(\big|\frac{B_n(x)}{n} - x\big| > \delta\Big) \leq \frac{1}{4n\delta^2}\,.$$

En combinant cette inégalité avec la question précédente on obtient

$$\Big|\mathbb{E}\Big[f\Big(\frac{B_n(x)}{n}\Big)\Big] - f(x)\Big| \leq \frac{2\;\|f\|_\infty}{4n\delta^2} + \epsilon\,.$$

En notant $P_n(x)$ la fonction polynomiale $\mathbb{E}[f(\frac{B_n(x)}{n})]$, on a montré que

$$\|P_n(x) - f(x)\|_\infty \leq \frac{2\;\|f\|_\infty}{4n\delta^2} + \epsilon$$

Maintenant, si on choisit n assez grand, on peut rendre le premier terme de la majoration inférieur à ϵ, donc

$$\|P_n(x) - f(x)\|_\infty \leq 2\epsilon\,.$$

Bref, on vient de montrer que (P_n) est une suite de fonctions polynomiales qui converge uniformément vers f.

TD 4 - Fonctions associées à une variable aléatoire

Exercice 1 Rappeler la définition de la fonction caractéristique φ_X d'une variable aléatoire X et calculer φ_X dans les cas suivants.

La fonction caractéristique φ_X d'une variable aléatoire X est définie par

$$\varphi_X(t) = \mathbb{E}[\mathrm{e}^{itX}], \quad t \in \mathbb{R}.$$

1. X suit une loi uniforme sur $[a, b]$.

   La densité de X est donnée par $\frac{1}{b-a}\mathbb{1}_{[a,b]}(x)$, d'où pour $t \neq 0$,

   $$\varphi_X(t) = \int_{\mathbb{R}} \mathrm{e}^{itx} \frac{1}{b-a}\mathbb{1}_{[a,b]}(x)\,\mathrm{d}x = \frac{1}{b-a}\int_a^b \mathrm{e}^{itx}\,\mathrm{d}x = \frac{1}{b-a}\left[\frac{\mathrm{e}^{itx}}{it}\right]_a^b = \frac{\mathrm{e}^{itb} - \mathrm{e}^{ita}}{it(b-a)}.$$

   Si $t = 0$, on obtient facilement $\varphi_X(t) = 1$, qui correspond à la limite en 0 du terme précédent.

2. X suit une loi exponentielle de paramètre λ.

   La densité de X est $\lambda\,\mathrm{e}^{-\lambda x}\,\mathbb{1}_{[0,\infty[}(x)$ ce qui donne

   $$\varphi_X(t) = \int_{\mathbb{R}} \mathrm{e}^{itx}\,\lambda\,\mathrm{e}^{-\lambda x}\,\mathbb{1}_{[0,\infty[}(x)\,\mathrm{d}x = \lambda\int_0^\infty \mathrm{e}^{(it-\lambda)x}\,\mathrm{d}x = \lambda\left[\frac{\mathrm{e}^{(it-\lambda)x}}{it-\lambda}\right]_0^\infty = \frac{\lambda}{\lambda - it},$$

   où la dernière égalité est une conséquence de la limite $\mathrm{e}^{(it-\lambda)x} \xrightarrow[x\to\infty]{} 0$ pour tout réel t et tout $\lambda > 0$.

3. X suit une loi géométrique de paramètre $p \in]0, 1[$.

   Dans ce cas, la fonction caractéristique φ_X de X est

   $$\varphi_X(t) = \sum_{k=1}^{\infty} \mathrm{e}^{itk}\,p(1-p)^{k-1} = \mathrm{e}^{it}\,p\sum_{k=0}^{\infty}[\mathrm{e}^{it}(1-p)]^k = \frac{\mathrm{e}^{it}\,p}{1 - \mathrm{e}^{it}(1-p)},$$

   où on a reconnu une série géométrique de raison $\mathrm{e}^{it}(1-p)$ qui vérifie $|\,\mathrm{e}^{it}(1-p)| = 1-p < 1$.

4. X suit une loi normale centrée réduite. On pourra dériver (en justifiant) la fonction caractéristique φ_X puis, après une intégration par parties, en déduire que φ_X est solution d'une équation différentielle que l'on résolvera.

   On part de l'égalité

   $$\varphi_X(t) = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{e}^{itx}\,\mathrm{d}x.$$

   La fonction $(t, x) \mapsto \mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{e}^{itx}$ est intégrable par rapport à x, dérivable par rapport à t de dérivée vérifiant $|\,\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{e}^{itx}\,ix| \leq \mathrm{e}^{-\frac{x^2}{2}}\,|x|$ qui est intégrable par rapport à x. Ainsi, le théorème de dérivation sous le signe intégrale donne

   $$\varphi_X'(t) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} ix\,\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{e}^{itx}\,\mathrm{d}x.$$

On effectue alors une intégration par parties :

$$\varphi_X'(t) = \frac{1}{\sqrt{2\pi}}\Big[-i\,\mathrm{e}^{-\frac{x^2}{2}}\,\mathrm{e}^{itx}\,\mathrm{d}x\Big]_{-\infty}^{\infty} - \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}(-i)\,\mathrm{e}^{-\frac{x^2}{2}}\,it\,\mathrm{e}^{itx}\,\mathrm{d}x = -\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\mathrm{e}^{-\frac{x^2}{2}}\,t\,\mathrm{e}^{itx}\,\mathrm{d}x\,.$$

Ainsi, φ_X est solution de l'équation différentielle $y'(t) = -ty(t)$ avec condition initiale $y(0) = 1$. On en déduit alors que $\varphi_X(t) = \mathrm{e}^{-\frac{t^2}{2}}$.

Exercice 2 Soit $X \sim \mathcal{N}(0,1)$.

1. Calculer la fonction de répartition de $-X$ en fonction de celle de X. Qu'en déduit-on ?
   La fonction de répartition de $-X$ est donnée par
   $$F_{-X}(x) = \mathbb{P}(-X \leq x) = \mathbb{P}(X \geq -x) = \int_{-x}^{\infty}\frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t\,.$$
   Le changement de variable $y = -t$ donne alors
   $$F_{-X}(x) = -\int_{y}^{-\infty}\frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{(-y)^2}{2}}\,\mathrm{d}y = \int_{\infty}^{x}\frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\frac{y^2}{2}}\,\mathrm{d}y = \mathbb{P}(X \leq x)\,.$$
   Donc les variables aléatoires X et $-X$ ont la même fonction de répartition donc la même loi : $-X \sim \mathcal{N}(0,1)$.
   Attention, dire que ces deux variables aléatoires ont la même loi ne veut pas dire qu'elles sont égales : il est clair que X et $-X$ ne sont jamais égales (sauf si $X = 0$ mais cet événement est de probabilité nulle).
   Remarquons que la preuve de ce résultat utilise uniquement la symétrie de la densité de la loi normale. Toute variable aléatoire X admettant une densité symétrique vérifie $\mathbb{P}(-X \leq x) = \mathbb{P}(X \leq x)$, ce qui implique que X et $-X$ ont la même loi.
2. On pose $Y = X^2$. Déterminer la fonction de répartition de Y en fonction de celle de X. En déduire que Y est une variable aléatoire à densité, puis calculer $\mathbb{E}[Y]$.
   Pour $x \geq 0$, la fonction de répartition de Y s'écrit
   $$F_Y(x) = \mathbb{P}(Y \leq x) = \mathbb{P}(X^2 \leq x) = \mathbb{P}(X \in [-\sqrt{x}, \sqrt{x}]) = F_X(\sqrt{x}) - F_X(-\sqrt{x})\,.$$
   Or,
   $$F_X(-\sqrt{x}) = \mathbb{P}(X \leq -\sqrt{x}) = 1 - \mathbb{P}(X > -\sqrt{x}) = 1 - \mathbb{P}(-X \leq \sqrt{x}) = 1 - F_{-X}(\sqrt{x})\,.$$
   et par la question 1. on en déduit que alors que $F_X(-\sqrt{x}) = 1 - F_{-X}(\sqrt{x}) = 1 - F_X(\sqrt{x})$. Bref, la fonction de répartition de Y correspond à
   $$F_Y(x) = 2F_X(\sqrt{x}) - 1\,,\quad x \geq 0\,.$$
   Enfin, comme $Y = X^2$ est une variable aléatoire positive, sa fonction de répartition est nulle avant 0 : $F_Y(x) = 0$ si $x < 0$.
   On remarque que la fonction F_Y est dérivable sur $]-\infty, 0[$ et sur $]0, \infty[$. Ainsi, Y est une variable aléatoire à densité dont la densité est définie presque partout par
   $$f_Y(x) = F_Y'(x) = 2\frac{1}{2\sqrt{x}}F_X'(\sqrt{x})\mathbb{1}_{]0,\infty[}(x) = \frac{1}{\sqrt{2\pi x}}\,\mathrm{e}^{-\frac{x}{2}}\,\mathbb{1}_{]0,\infty[}(x)\,,\quad x \in \mathbb{R}^*\,.$$

Pour l'espérance, on peut appliquer directement la formule de transfert :

$$\mathbb{E}[Y] = \mathbb{E}[X^2] = \int_{\mathbb{R}} \frac{x^2}{\sqrt{2\pi}} \, \mathrm{e}^{-\frac{x^2}{2}} \, \mathrm{d}x \,.$$

Une intégration par parties donne alors

$$\mathbb{E}[Y] = \Big[- \frac{x}{\sqrt{2\pi}} \, \mathrm{e}^{-\frac{x^2}{2}} \Big]_{-\infty}^{\infty} + \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \, \mathrm{e}^{-\frac{x^2}{2}} \, \mathrm{d}x \,.$$

Le premier terme est nul tandis que le deuxième terme correspond à l'intégrale de la densité normale centrée réduite. Ainsi $\mathbb{E}[Y] = 1$.

Remarque : la loi de Y est appelée *loi du chi-deux à un degré de liberté*. Elle correspond à la loi du carré d'une variable aléatoire de loi normale.

3. Reprendre la question précédente avec cette fois-ci $Z = \exp(X)$.

Pour tout $x > 0$, on a

$$F_Z(x) = \mathbb{P}(Z \leq x) = \mathbb{P}(\exp(X) \leq x) = \mathbb{P}(X \leq \ln x) = F_X(\ln x) \,.$$

Par ailleurs, si $x \leq 0$, alors $F_Z(x) = \mathbb{P}(\exp(X) \leq x) = 0$.

La fonction F_Z est dérivable sur $]-\infty, 0[$ et sur $]0, \infty[$, avec $F_Z'(x) = 0$ sur $]-\infty, 0[$ et

$$F_Z'(x) = \frac{1}{x} F_X'(\ln x) = \frac{1}{x} \frac{1}{\sqrt{2\pi}} \, \mathrm{e}^{-\frac{(\ln x)^2}{2}} = \frac{1}{x\sqrt{2\pi}} \, \mathrm{e}^{-\frac{(\ln x)^2}{2}} \,.$$

La densité de Z est donc la fonction $f_Z : \mathbb{R} \to \mathbb{R}$ définie presque partout par

$$f_Z(t) = \frac{1}{x\sqrt{2\pi}} \, \mathrm{e}^{-\frac{(\ln x)^2}{2}} \, \mathbb{1}_{]0,\infty[}(x) \,.$$

Ici aussi, on calcule l'espérance via la formule de transfert :

$$\mathbb{E}[Z] = \mathbb{E}[\exp(X)] = \int_{\mathbb{R}} \frac{\mathrm{e}^x}{\sqrt{2\pi}} \, \mathrm{e}^{-\frac{x^2}{2}} \, \mathrm{d}x = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \, \mathrm{e}^{-\frac{(x-1)^2}{2}} \, \mathrm{e}^{\frac{1}{2}} \, \mathrm{d}x \,.$$

Le changement de variables $t = x - 1$ donne alors

$$\mathbb{E}[Z] = \sqrt{\mathrm{e}} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \, \mathrm{e}^{-\frac{t^2}{2}} \, \mathrm{d}t = \sqrt{\mathrm{e}} \,.$$

Remarque : cette loi est appelée *loi log-normale*. Elle correspond à la loi de l'exponentielle d'une variable aléatoire de loi normale.

Exercice 3 On considère une variable aléatoire X à valeurs dans $\mathbb{N}$ de fonction génératrice

$$G_X(s) = \alpha(3 + 2s^2)^3 \,, \quad s \in [0, 1] \,.$$

1. Déterminer la valeur de α.

On rappelle que la fonction génératrice d'une variable aléatoire X à valeurs dans $\mathbb{N}$ est donnée par

$$G_X(s) = \mathbb{E}[s^X] = \sum_{k=0}^{\infty} p_k s^k \,,$$

où $p_k = \mathbb{P}(X = k)$. En particulier $G_X(1) = 1$ ce qui donne $\alpha = \frac{1}{125}$.

2. Déterminer la loi de X.

   La loi d'une variable aléatoire à valeurs dans $\mathbb{N}$ est entièrement caractérisée par les quantités $p_k = \mathbb{P}(X = k)$ qui apparaissent dans la fonction génératrice. Dans notre cas on obtient

   $$G_X(s) = \frac{1}{125}(27 + 54s^2 + 36s^4 + 8s^6) = p_0 + p_2 s^2 + p_4 s^4 + p_6 s^6 \,.$$

   Par unicité du développement en série entière (ici en fait par unicité de l'écriture polynomiale), on obtient

   $$p_0 = \frac{27}{125}\,, \quad p_2 = \frac{54}{125}\,, \quad p_4 = \frac{36}{125}\,, \quad p_6 = \frac{8}{125}\,.$$

   On peut vérifier que les p_i somment bien à 1. En se souvenant qu'une loi est une mesure de probabilité, on peut aussi écrire la loi de X sous la forme

   $$\mathbb{P}_X = p_0\delta_0 + p_2\delta_2 + p_4\delta_4 + p_6\delta_6\,,$$

   avec pour p_i les valeurs précédentes.

3. À partir de G_X, donner les valeurs de l'espérance et de la variance de X.

   La fonction G_X est une fonction polynomiale donc est de classe $\mathcal{C}^\infty$ sur $[0,1]$ (en fait même sur $\mathbb{R}$). Sa dérivée est donnée par

   $$G'_X(s) = \mathbb{E}[Xs^X] = \frac{1}{125}12s(3 + 2s^2)^2\,.$$

   En évaluant en $s = 1$ on trouve, $G'_X(1) = \mathbb{E}[X] = \frac{300}{125} = 2.4$. En dérivant une seconde fois, on obtient

   $$G''_X(s) = \mathbb{E}[X(X-1)s^{X-1}] = \frac{1}{125}12(3+2s^2)^2 + \frac{1}{125}12 \times 4s(3+2s^2)\,.$$

   En évaluant en $s = 1$ on trouve $\mathbb{E}[X(X-1)] = \frac{780}{125} = 6.24$. La formule de Koenig-Huygens donne alors le résultat

   $$\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{E}[X(X-1)] + \mathbb{E}[X] - \mathbb{E}[X]^2 = 2.88\,.$$

Pour aller plus loin

Vous disposez de deux dés et vous souhaitez truquer les dés de sorte que la somme des résultats soit uniformément distribuée sur $\{2, \ldots, 12\}$. Est-ce possible ? On pourra utiliser les fonctions génératrices.

On note X_1 (resp. X_2) la variable aléatoire donnant le résultat du premier (resp. deuxième) lancer. La fonction génératrice de X_1 s'écrit alors

$$G_{X_1}(s) = \mathbb{E}[s^{X_1}] = \sum_{k=1}^{6} \mathbb{P}(X_1 = k)s^k\,,$$

et idem pour X_2. Par indépendance de X_1 et X_2 on obtient

$$G_{X_1+X_2}(s) = \mathbb{E}[s^{X_1+X_2}] = \mathbb{E}[s^{X_1}s^{X_2}] = \mathbb{E}[s^{X_1}]\mathbb{E}[s^{X_1}] = (G_{X_1}(s))^2 \,.$$

Ainsi, $G_{X_1+X_2}$ est un polynôme dont les racines sont toutes de multiplicité (au moins) deux.

Or, si U suit une loi uniforme sur $\{2, \ldots, 12\}$, alors sa fonction génératrice est donnée par

$$G_U(s) = \sum_{k=2}^{12} \mathbb{P}(U = k)s^k = \frac{1}{11}(s^2 + s^3 + \ldots + s^{12}) = \frac{1}{11}s^2(1 + s + \ldots + s^{10}) \,.$$

Les racines G_U sont donc 0 (de multiplicité 2) et les racines de $1 + s + \ldots + s^{10}$. Ces dernières correspondent aux racines 10-èmes de l'unité, sauf 1. Elles sont toutes distinctes donc sont de multiplicité 1. D'où la contradiction.

TD 5 - Calcul de lois

Exercice 1 Soit X une variable aléatoire de loi exponentielle de paramètre λ. On pose $Y = \lfloor X \rfloor + 1$, où $\lfloor x \rfloor$ désigne la partie entière d'un réel x. Déterminer la loi de Y.

La variable aléatoire Y prend ses valeurs dans $\mathbb{N}^*$, c'est donc une variable aléatoire discrète. Déterminons sa loi. Pour $k \in \mathbb{N}^*$, on a

$$\mathbb{P}(Y = k) = \mathbb{P}(\lfloor X \rfloor + 1 = k) = \mathbb{P}(k - 1 \leq X < k)\,.$$

Comme X une variable aléatoire de loi exponentielle de paramètre λ, cette probabilité vaut

$$\mathbb{P}(k-1 \leq X < k) = \int_{k-1}^{k} \lambda\,\mathrm{e}^{-\lambda t}\,\mathrm{d}t = \Big[-\mathrm{e}^{-\lambda t} \Big]_{k-1}^{k} = \mathrm{e}^{-\lambda(k-1)} - \mathrm{e}^{-\lambda k} = \mathrm{e}^{-\lambda(k-1)}(1 - \mathrm{e}^{-\lambda})\,.$$

Ainsi, pour tout entier $k \geq 1$, on a $\mathbb{P}(Y = k) = (\mathrm{e}^{-\lambda})^{k-1}(1 - \mathrm{e}^{-\lambda})$. On reconnaît une loi géométrique de paramètre $p = 1 - \mathrm{e}^{-\lambda}$.

Exercice 2

1. Déterminer la fonction caractéristique de la loi de Bernoulli, de la loi binomiale, et de la loi de Poisson. Rappeler également (sans calcul) la fonction caractéristique de la loi normale.

   On rappelle que la fonction caractéristique d'une variable aléatoire X est donnée par $\varphi_X(t) = \mathbb{E}[\mathrm{e}^{itX}]$.

   Pour une variable aléatoire $X \sim \mathcal{B}(p)$,, on obtient

   $$\varphi_X(t) = \mathbb{E}[\mathrm{e}^{itX}] = \mathrm{e}^{it\times 1} \times p + \mathrm{e}^{it\times 0} \times (1-p) = p\,\mathrm{e}^{it} + 1 - p\,.$$

   Pour une variable aléatoire $X \sim \mathcal{B}(n, p)$, on obtient

   $$\varphi_X(t) = \mathbb{E}[\mathrm{e}^{itX}] = \sum_{k=0}^{n} \mathrm{e}^{itk} \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} \binom{n}{k} (p\,\mathrm{e}^{it})^k (1-p)^{n-k} = (p\,\mathrm{e}^{it} + 1 - p)^n\,,$$

   par la formule du binôme de Newton.

   Pour une variable aléatoire $X \sim \mathcal{P}(\lambda)$, on obtient

   $$\varphi_X(t) = \sum_{k=0}^{\infty} \mathrm{e}^{itk}\,\mathrm{e}^{-\lambda} \frac{\lambda^k}{k!} = \mathrm{e}^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda\,\mathrm{e}^{it})^k}{k!} = \mathrm{e}^{-\lambda}\,\mathrm{e}^{\lambda\,\mathrm{e}^{it}} = \mathrm{e}^{\lambda(\mathrm{e}^{it} - 1)}\,.$$

   Enfin, on sait d'après le cours que la fonction caractéristique d'une variable aléatoire $X \sim \mathcal{N}(\mu, \sigma^2)$ est donnée par

   $$\varphi_X(t) = \mathrm{e}^{i\mu t - \frac{\sigma^2 t^2}{2}}\,.$$

2. Soit X et Y deux variables aléatoires indépendantes. Montrer que $\varphi_{X+Y} = \varphi_X \varphi_Y$.
Il suffit d'écrire
$$\varphi_{X+Y}(t) = \mathbb{E}[\mathrm{e}^{it(X+Y)}] = \mathbb{E}[\mathrm{e}^{itX}\,\mathrm{e}^{itY}]\,,$$
et l'indépendance de X et Y assure que $\mathbb{E}[\mathrm{e}^{itX}\,\mathrm{e}^{itY}] = \mathbb{E}[\mathrm{e}^{itX}]\mathbb{E}[\mathrm{e}^{itY}]$, ce qui permet de conclure :
$$\varphi_{X+Y}(t) = \mathbb{E}[\mathrm{e}^{itX}]\mathbb{E}[\mathrm{e}^{itY}] = \varphi_X(t)\varphi_Y(t)\,.$$

3. (a) On considère deux variables aléatoires $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ et $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ indépendantes. Déterminer la loi de $X_1 + X_2$.
On va déterminer la loi de la variable aléatoire $X_1 + X_2$ en calculant sa fonction caractéristique. L'indépendance de X_1 et X_2 et la question 2. assurent que $\varphi_{X_1+X_2}(t) = \varphi_{X_1}(t)\varphi_{X_2}(t)$. Puis la question 1. permet d'écrire
$$\varphi_{X_1+X_2}(t) = \varphi_{X_1}(t)\varphi_{X_2}(t) = \mathrm{e}^{i\mu_1 t - \frac{\sigma_1^2 t^2}{2}}\,\mathrm{e}^{i\mu_2 t - \frac{\sigma_2^2 t^2}{2}} = \mathrm{e}^{i(\mu_1+\mu_2)t - \frac{(\sigma_1^2+\sigma_2^2)t^2}{2}}\,.$$
On reconnaît la fonction caractéristique d'une loi normale $\mathcal{N}(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$. Comme la fonction caractéristique caractérise la loi, on en déduit que $X_1 + X_2$ suit une loi normale $\mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

(b) Montrer que la somme de n variables aléatoires indépendantes de loi de Bernoulli de paramètre p suit une loi binomiale de paramètres n et p.
On utilise ici aussi le fait que la fonction caractéristique caractérise la loi. Il suffit donc de montrer que la fonction caractéristique d'une somme de n variables aléatoires indépendantes de loi de Bernoulli de paramètre p est égale à la fonction caractéristique d'une loi binomiale de paramètres n et p. Notons $X_1, \ldots, X_n$ les variables aléatoires de loi de Bernoulli et X la variable aléatoire de loi binomiale. L'indépendance des X_k et les questions 1. et 2. donnent
$$\varphi_{X_1+\cdots+X_n}(t) = \varphi_{X_1}(t)\cdots\varphi_{X_n}(t) = (p\,\mathrm{e}^{it} + 1 - p)^n = \varphi_X(t)\,,$$
ce qui prouve le résultat.

(c) Montrer que la somme de deux variables aléatoires indépendantes X et Y de loi de Poisson de paramètre respectif λ et μ suit une loi de Poisson de paramètre $\lambda + \mu$.
On a vu dans la question 1. que si $X \sim \mathcal{P}(\lambda)$, alors $\varphi_X(t) = \mathrm{e}^{\lambda(\mathrm{e}^{it}-1)}$. De même, on a $\varphi_X(t) = \mathrm{e}^{\mu(\mathrm{e}^{it}-1)}$. Il suffit alors d'appliquer le résultat de la question 2. : comme X et Y sont indépendantes, la fonction caractéristique de $X + Y$ est
$$\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t) = \mathrm{e}^{\lambda(\mathrm{e}^{it}-1)}\,\mathrm{e}^{\mu(\mathrm{e}^{it}-1)} = \mathrm{e}^{(\lambda+\mu)(\mathrm{e}^{it}-1)}\,.$$
On reconnaît la fonction caractéristique d'une loi de Poisson de paramètre $\lambda + \mu$ et comme la fonction caractéristique caractérise la loi, on en déduit que $X + Y$ suit une loi de Poisson de paramètre $\lambda + \mu$.

Exercice 3 On considère deux variables aléatoires indépendantes $X \sim \mathcal{N}(0, 1)$ et $Y \sim \mathcal{N}(0, \sigma^2)$ avec $\sigma > 0$.

1. On pose $T = X^2$. Calculer $\mathbb{E}[h(T)]$ pour toute fonction $h : \mathbb{R} \to \mathbb{R}$ mesurable bornée. En déduire la loi de T.
   On pourra vérifier qu'on retrouve bien le résultat de l'exercice 2 du TD 4.
   Soit $h : \mathbb{R} \to \mathbb{R}$ mesurable bornée. La formule de transfert donne
   $$\mathbb{E}[h(T)] = \mathbb{E}[h(X^2)] = \int_{\mathbb{R}} h(x^2)\, \mathrm{d}\mathbb{P}_X(x) = \int_{\mathbb{R}} h(x^2) \frac{\mathrm{e}^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\, \mathrm{d}x\,.$$
   Par parité de l'intégrande, on en déduit que
   $$\mathbb{E}[h(T)] = 2\int_0^\infty h(x^2) \frac{\mathrm{e}^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\, \mathrm{d}x\,.$$
   Le changement de variables $t = x^2 \iff \sqrt{t} = x$ donne alors
   $$\mathbb{E}[h(T)] = 2\int_0^\infty h(t) \frac{\mathrm{e}^{-\frac{t}{2}}}{\sqrt{2\pi}} \frac{1}{2\sqrt{t}}\, \mathrm{d}t = \int_{\mathbb{R}} h(t) \mathbb{1}_{]0,\infty[}(t) \frac{\mathrm{e}^{-\frac{t}{2}}}{\sqrt{2\pi t}}\, \mathrm{d}t\,.$$
   Ceci étant valable pour toute fonction $h : \mathbb{R} \to \mathbb{R}$ mesurable bornée, cela prouve que T est une variable aléatoire de densité
   $$f_T(t) = \frac{\mathrm{e}^{-\frac{t}{2}}}{\sqrt{2\pi t}} \mathbb{1}_{]0,\infty[}(t)\,.$$
   On retrouve bien la densité obtenue via les fonctions de répartition.
2. On pose $U = \mathbb{1}_{\{X \geq 0\}} - \mathbb{1}_{\{X<0\}}$. Déterminer la loi de U.
   La variable aléatoire U est à valeurs dans $\{-1, 1\}$, c'est donc une variable aléatoire discrète. Sa loi est donnée par
   $$p_1 = \mathbb{P}(U = 1) = \mathbb{P}(X \geq 0) = 1/2 \quad \text{et} \quad p_{-1} = \mathbb{P}(U = -1) = \mathbb{P}(X < 0) = 1/2\,.$$
   Remarque : cette loi est appelée loi de Rademacher de paramètre $p_1 = 1/2$. On peut voir U comme un "signe aléatoire" : soit positif avec probabilité 1/2, soit négatif avec probabilité 1/2.
3. Montrer que $Z = UY$ suit une loi normale de paramètres à déterminer.
   Remarque : on a montré dans l'exercice 2 du TD4 que si $Y \sim \mathcal{N}(0, 1)$, alors $-Y$ suit encore une loi $\mathcal{N}(0, 1)$. C'est une conséquence de la parité de la densité d'une loi normale $\mathcal{N}(0, 1)$. Le même argument assure que si $Y \sim \mathcal{N}(0, \sigma^2)$, alors $-Y \sim \mathcal{N}(0, \sigma^2)$.
   Maintenant, si on considère $Z = UY$, alors $Z = Y$ si $U = 1$ et $Z = -Y$ si $U = -1$. Donc Z vaut Y avec probabilité 1/2 et $-Y$ avec probabilité 1/2. Bref, on considère Y et on change son signe aléatoirement avec probabilité 1/2. L'objectif de la question est de montrer qu'avec cette procédure Z suit encore une loi normale $\mathcal{N}(0, \sigma^2)$.

   On va déterminer la fonction de répartition de Z en utilisant la formule des probabilités totales avec la partition $\{U = 1\}$ et $\{U = -1\}$. Soit $t \in \mathbb{R}$. Alors
   $$\begin{aligned} F_Z(t) &= \mathbb{P}(Z \leq t) \\ &= \mathbb{P}(UY \leq t) \\ &= \mathbb{P}(UY \leq t \mid U = 1)\mathbb{P}(U = 1) + \mathbb{P}(UY \leq t \mid U = -1)\mathbb{P}(U = -1)\,. \end{aligned}$$

On obtient alors

$$F_Z(t) = \frac{\mathbb{P}(Y \leq t \mid U = 1) + \mathbb{P}(-Y \leq t \mid U = -1)}{2} = \frac{\mathbb{P}(Y \leq t) + \mathbb{P}(-Y \leq t)}{2},$$

où la deuxième égalité résulte de l'indépendance de Y et U. Il reste à remarquer que par symétrie de la densité de la loi normale $\mathcal{N}(0, \sigma^2)$, les deux probabilités précédentes sont identiques

$$\mathbb{P}(-Y \leq t) = \mathbb{P}(Y \geq -t) = \int_{]-t,\infty[} \frac{\mathrm{e}^{-\frac{u^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \,\mathrm{d}u = \int_{]-\infty,t[} \frac{\mathrm{e}^{-\frac{v^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \,\mathrm{d}v = \mathbb{P}(Y \leq t)\,,$$

où on a appliqué le changement de variables $v = -u$. Ainsi, pour tout $t \in \mathbb{R}$,

$$F_Z(t) = \mathbb{P}(Y \leq t) = F_Y(t)\,.$$

Donc Z et Y ont la même fonction de répartition donc la même loi : $Z \sim \mathcal{N}(0, \sigma^2)$.

4. Déterminer la loi de $Y + Z$.

Remarquons que $Y + Z = (1 + U)Y$ vaut 0 avec probabilité $1/2$ et $2Y$ avec probabilité $1/2$. Le plus direct consiste encore une fois à calculer la fonction de répartition de $Y + Z$: pour $t \in \mathbb{R}$, on a

$$F_{Y+Z}(t) = \mathbb{P}(Y + Z \leq t) = \mathbb{P}((1 + U)Y \leq t)\,.$$

On applique alors la même démarche qu'à la question précédente en remarquant que U et Y sont indépendantes puisque X et Y le sont et que U ne dépend que de X :

$$F_{Y+Z}(t) = \frac{\mathbb{P}(0 \leq t) + \mathbb{P}(2Y \leq t)}{2}\,.$$

La première probabilité ne fait pas intervenir le quantité dépendant de ω, elle est complètement déterministe et vaut 1 si $t \geq 0$ et 0 sinon. Bref

$$\mathbb{P}(0 \leq t) = \mathbb{1}_{[0,\infty[}(t)\,,$$

et ceci est la fonction de répartition d'une mesure de Dirac en 0.

La deuxième probabilité $\mathbb{P}(2Y \leq t) = \mathbb{P}(Y \leq t/2)$ se calcule via le changement de variable $u = 2v$:

$$\mathbb{P}(Y \leq t/2) = \int_{]-\infty,t/2]} \frac{\mathrm{e}^{-\frac{u^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \,\mathrm{d}u = \int_{]-\infty,t]} \frac{\mathrm{e}^{-\frac{(2v)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \frac{1}{2} \,\mathrm{d}v = \int_{]-\infty,t]} \frac{\mathrm{e}^{-\frac{v^2}{2(4\sigma^2)}}}{\sqrt{2\pi(4\sigma^2)}} \,\mathrm{d}v\,.$$

On reconnaît la fonction de répartition d'une loi normale $\mathcal{N}(0, 4\sigma^2)$. Ainsi, la loi de $Y + Z$ est la demi-somme d'une loi de Dirac en 0 et d'une loi normale $\mathcal{N}(0, 4\sigma^2)$:

$$\frac{1}{2} \frac{\mathrm{e}^{-\frac{x^2}{2\sigma'^2}}}{\sigma'\sqrt{2\pi}} \,\mathrm{d}x + \frac{1}{2}\delta_0\,.$$

La loi de $Y + Z$ est donc un mélange entre une loi normale centrée de variance $4\sigma_2^2$ et une masse de Dirac en 0. On remarque en particulier que $Y + Z$ ne suit pas une loi normale car ces deux variables aléatoires ne sont pas indépendantes. Notons que la masse de Dirac en 0 est cohérente : la variable aléatoire U prend la valeur -1 avec probabilité $1/2$, ce qui implique alors que $Z = -Y$, donc que $Y + Z = 0$. Ainsi, on a une chance sur deux pour que la somme $Y + Z$ soit nulle, d'où le terme $\frac{1}{2}\delta_0$.

Pour aller plus loin

On considère une succession de lancers d'une pièce ayant une probabilité p de donner pile et on note X (resp. Y) la longueur de la première (resp. deuxième) série de lancers identiques. Par exemple, si les lancers donnent $FFFFPPFP\ldots$, alors $X = 4$ et $Y = 2$.

Déterminer la loi du couple (X, Y). En déduire la loi de X, la loi de Y, puis $\mathbb{E}[X]$ et $\mathbb{E}[Y]$. Comparer ces deux espérances.

Il est clair que X et Y sont à valeurs dans $\mathbb{N}^*$. Puis l'événement $\{X = k\}$ est réalisé si et seulement si les k premiers lancers ont donné pile et le $k+1$-ème face, ou l'inverse, ces deux possibilités étant incompatibles. Par indépendance des lancers, on en déduit que

$$\mathbb{P}(X = k) = p^k(1-p) + (1-p)^k p\,.$$

Concernant la loi de Y, si $k \geq 1$, on détermine $\mathbb{P}(Y = k)$ via la formule des probabilités totales en utilisant la partition formée par les événements $\{X = n\}$, $n \geq 1$, ce qui donne

$$\mathbb{P}(Y = k) = \sum_{n=1}^{\infty} \mathbb{P}(Y = k,\, X = n)$$

Or dire, que $Y = k$ et $X = n$ signifie que le lancer a donné n piles puis k faces puis un pile, ou l'inverse. L'indépendance des lancers donne alors

$$\mathbb{P}(Y = k) = \sum_{n=1}^{\infty} (p^n(1-p)^k p + (1-p)^n p^k (1-p)) = (1-p)^k \sum_{n=1}^{\infty} p^{n+1} + p^k \sum_{n=1}^{\infty} (1-p)^{n+1}$$

On reconnaît deux séries géométriques, ce qui donne

$$\mathbb{P}(Y = k) = (1-p)^k \frac{p^2}{1-p} + p^k \frac{(1-p)^2}{p} = (1-p)^{k-1}p^2 + p^{k-1}(1-p)^2\,.$$

Pour le calcul des espérances, on utilise l'astuce classique qui consiste à considérer la fonction g définie pour tout $x \in]0,1[$ par

$$g(x) = \sum_{k=0}^{\infty} x^k = \frac{1}{1-x}\,,$$

dont la dérivée vérifie

$$g'(x) = \sum_{k=1}^{\infty} k x^{k-1} = \frac{1}{(1-x)^2}\,.$$

Par suite, l'espérance de X est donnée par

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} k\mathbb{P}(X = k) = \sum_{k=1}^{\infty} kp^k(1-p) + \sum_{k=1}^{\infty} k(1-p)^k p = p(1-p)g'(p) + (1-p)pg'(1-p)\,.$$

Ainsi,

$$\mathbb{E}[X] = p(1-p)\frac{1}{(1-p)^2} + (1-p)p\frac{1}{p^2} = \frac{p}{1-p} + \frac{1-p}{p}\,.$$

De même, on trouve

$$\mathbb{E}[Y] = \sum_{k=1}^{\infty} k\mathbb{P}(Y = k) = \sum_{k=1}^{\infty} k(1-p)^{k-1}p^2 + \sum_{k=1}^{\infty} kp^{k-1}(1-p)^2 = p^2 g'(1-p) + (1-p)^2 g'(p)\,.$$

Ainsi,

$$\mathbb{E}[Y] = p^2 \frac{1}{p^2} + (1-p)^2 \frac{1}{(1-p)^2} = 2\,.$$

En particulier, l'espérance de Y ne dépend pas du biais de la pièce !

Pour comparer ces deux espérances, il suffit de partir de remarquer que la quantité $(p-(1-p))^2$ est positive. En développant cette expression, on obtient

$$p^2 + (1-p)^2 - 2p(1-p) \geq 0\,,$$

ou encore, en divisant tout par $p(1-p)$,

$$2 \leq \frac{p}{1-p} + \frac{1-p}{p}\,.$$

Bref, on a toujours $\mathbb{E}[Y] \leq \mathbb{E}[X]$.

TD 6 - Conditionnement et indépendance

Exercice 1 On souhaite transmettre un message d'un point à un autre à travers des canaux successifs. Ce message peut prendre deux valeurs : 0 ou 1. Au passage de chaque canal, le message a une probabilité $p \in]0,1[$ d'être bruité, c'est-à-dire d'être transformé en son contraire, et $1-p$ d'être transmis fidèlement. Les canaux de transmission sont indépendants les uns des autres.

1. On considère l'événement A_n : "après le canal n, le message est identique au message initial", et on note p_n sa probabilité. Exprimer p_{n+1} en fonction de p_n et p. Que vaut p_1 ?
   Pour $n \geq 1$ fixé, le couple $(A_n, \bar{A}_n)$ forme une partition de l'univers (soit le message après le canal n est identique au message initial, soit il a été bruité). Par la formule des probabilités totales, on obtient alors
   $$p_{n+1} = \mathbb{P}(A_{n+1}) = \mathbb{P}(A_{n+1} \mid A_n)\mathbb{P}(A_n) + \mathbb{P}(A_{n+1} \mid \bar{A}_n)\mathbb{P}(\bar{A}_n)\,.$$
   La probabilité de conserver le message dans un canal vaut $1-p$, d'où $\mathbb{P}(A_{n+1} \mid A_n) = 1-p$, tandis que la probabilité que le message soit bruité vaut p, d'où $\mathbb{P}(A_{n+1} \mid \bar{A}_n) = p$. Ainsi
   $$p_{n+1} = (1-p)p_n + p(1-p_n) = (1-2p)p_n + p\,.$$
   La condition initiale est $p_1 = 1-p$, probabilité que le message n'ait pas été bruité dans le premier canal.
2. On définit la suite $(x_n)_{n\geq 1}$ par $x_n = p_n - \frac{1}{2}$. Vérifier que cette suite est géométrique. En déduire une expression pour p_n.
   Pour $n \geq 1$ un calcul immédiat donne
   $$x_{n+1} = p_{n+1} - \frac{1}{2} = (1-2p)p_n + p - \frac{1}{2} = (1-2p)(p_n - \frac{1}{2}) = (1-2p)x_n\,,$$
   ce qui prouve que la suite $(x_n)_{n\geq 1}$ est géométrique de raison $(1-2p)$, d'où
   $$x_n = (1-2p)^{n-1}x_1\,.$$
   Puis, comme $x_1 = p_1 - \frac{1}{2} = \frac{1}{2} - p$, on obtient
   $$x_n = (1-2p)^{n-1}\left(\frac{1}{2} - p\right) = \frac{1}{2}(1-2p)^n\,,$$
   d'où
   $$p_n = x_n + \frac{1}{2} = \frac{1}{2}(1-2p)^n + \frac{1}{2}\,.$$

3. Déterminer $\lim_{n\to\infty} p_n$.

   On distingue trois cas.

   - Si $p = 0$, alors $p_n = 1$ pour tout n : la transmission est fidèle, c'est-à-dire qu'on ne perd pas d'information.

   - Si $p = 1$, alors $p_{2n} = 1$ et $p_{2n+1} = 0$: la transmission change le message à chaque canal et l'information dépend de la parité du canal.

   - Si $0 < p < 1$, alors $1 - 2p \in]-1, 1[$, ce qui donne $\lim_{n\to+\infty}(1-2p)^n = 0$. Ainsi, $\lim_{n\to+\infty} p_n = 1/2$. Au bout d'un certain temps, on est incapable de retrouver le message initial.

Exercice 2 Soit X et Y deux variables aléatoires indépendantes de loi exponentielle de paramètre respectif λ et μ.

1. Rappeler la loi du couple (X, Y).

   Rappelons tout d'abord que la loi d'un vecteur (X, Y) est donnée de manière générale par

$$\mathbb{P}_{(X,Y)}(A \times B) = \mathbb{P}(X \in A, Y \in B)\,, \quad A, B \in \mathcal{B}(\mathbb{R})\,,$$

   et cela suffit pour connaître $\mathbb{P}_{(X,Y)}$ car les ensembles de la forme $A \times B$ engendrent la tribu $\mathcal{B}(\mathbb{R}^2)$.

   Ici, les variables aléatoires X et Y sont indépendantes, ce qui implique que la loi du vecteur (X, Y) correspond à la mesure produit $\mathbb{P}_X \otimes \mathbb{P}_Y$:

$$\mathbb{P}_{(X,Y)}(A \times B) = \mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)\,.$$

   Enfin, on sait que X et Y sont de densité exponentielle, donc leurs lois vérifient

$$\mathrm{d}\mathbb{P}_X = \lambda\,\mathrm{e}^{-\lambda x}\,\mathbb{1}_{[0,\infty[}(x)\,\mathrm{d}x \quad \text{et} \quad \mathrm{d}\mathbb{P}_Y = \mu\,\mathrm{e}^{-\mu y}\,\mathbb{1}_{[0,\infty[}(y)\,\mathrm{d}y\,,$$

   donc

$$\mathrm{d}\mathbb{P}_{(X,Y)} = \mathrm{d}(\mathbb{P}_X \otimes \mathbb{P}_Y) = \lambda\mu\,\mathrm{e}^{-\lambda x}\,\mathrm{e}^{-\mu y}\,\mathbb{1}_{[0,\infty[}(x)\mathbb{1}_{[0,\infty[}(y)\,\mathrm{d}x\mathrm{d}y\,,$$

   Ainsi, (X, Y) est un vecteur aléatoire à valeurs dans $\mathbb{R}^2$ de densité

$$g(x, y) = \lambda\mu\,\mathrm{e}^{-\lambda x}\,\mathrm{e}^{-\mu y}\,\mathbb{1}_{[0,\infty[}(x)\mathbb{1}_{[0,\infty[}(y)\,, \quad (x, y) \in \mathbb{R}^2\,.$$

2. Calculer $\mathbb{P}(X \leq Y)$.

En notant A l'ensemble $\{(x,y) \in \mathbb{R}^2 : x \leq y\}$), on obtient

$$\begin{aligned}
\mathbb{P}(X \leq Y) &= \mathbb{P}((X,Y) \in A) \\
&= \int_A g(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{\mathbb{R}^2} \lambda\,\mathrm{e}^{-\lambda x}\,\mu\,\mathrm{e}^{-\mu y}\,\mathbb{1}_{[0,\infty[}(x)\mathbb{1}_{[0,\infty[}(y)\mathbb{1}_{\{x\leq y\}}\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_0^\infty \int_0^y \lambda\,\mathrm{e}^{-\lambda x}\,\mu\,\mathrm{e}^{-\mu y}\;\mathrm{d}x\,\mathrm{d}y \\
&= \int_0^\infty \Big[-\mathrm{e}^{-\lambda x}\Big]_0^y \mu\,\mathrm{e}^{-\mu y}\;\mathrm{d}y \\
&= \int_0^\infty \mu\,\mathrm{e}^{-\mu y} - \mu\,\mathrm{e}^{-(\lambda+\mu)y}\;\mathrm{d}y \\
&= \Big[-\mathrm{e}^{-\mu y}\Big]_0^\infty + \Big[\mathrm{e}^{-(\lambda+\mu)y}\,\frac{\mu}{\lambda+\mu}\Big]_0^\infty \\
&= 1 - \frac{\mu}{\lambda+\mu} \\
&= \frac{\lambda}{\lambda+\mu}\,.
\end{aligned}$$

3. Déterminer $\mathbb{P}(X > t)$, pour tout réel t. En déduire la fonction de répartition puis la loi de $Z = \min(X,Y)$.

   Si $t < 0$, alors $\mathbb{P}(X > t) = 1$. Si $t \geq 0$, alors

$$\mathbb{P}(X > t) = \int_t^\infty \lambda\,\mathrm{e}^{-\lambda x}\,\mathbb{1}_{\{x\geq 0\}}\,\mathrm{d}x = \Big[-\mathrm{e}^{-\lambda x}\Big]_t^\infty = \mathrm{e}^{-\lambda t}\,.$$

   De manière condensée, $\mathbb{P}(X > t) = \mathrm{e}^{-\lambda \max(t,0)}$, et on a la même formule pour Y.
   Par suite, la fonction de répartition de $Z = \min(X,Y)$ est donnée par

$$F_Z(t) = \mathbb{P}(Z \leq t) = \mathbb{P}(\min(X,Y) \leq t) = 1 - \mathbb{P}(\min(X,Y) > t)\,.$$

   Or, dire que le minimum de X et Y est supérieur à t revient à dire que X et Y sont tous les deux supérieurs à t. Par indépendance de X et Y on trouve alors, pour $t \geq 0$,

$$F_Z(t) = 1 - \mathbb{P}(X > t, Y > t) = 1 - \mathbb{P}(X > t)\mathbb{P}(Y > t) = 1 - \mathrm{e}^{-\lambda t}\,\mathrm{e}^{-\mu t} = 1 - \mathrm{e}^{-(\lambda+\mu)t}\,.$$

   Par ailleurs, comme X et Y sont à valeurs dans $[0,\infty[$ on trouve $F_Z(t) = \mathbb{P}(\min(X,Y) \leq t) = 0$ pour $t < 0$. Ainsi, la fonction de répartition de Z est donnée par

$$F_Z(t) = \begin{cases} 1 - \mathrm{e}^{-(\lambda+\mu)t} & \text{si } t \geq 0, \\ 0 & \text{si } t < 0\,. \end{cases}$$

   On reconnaît la fonction de répartition d'une loi exponentielle de paramètre $\lambda+\mu$. Comme la fonction de répartition d'une variable aléatoire caractérise sa loi, on en déduit que Z suit une loi exponentielle de paramètre $\lambda+\mu$.
   Remarque : on vient de montrer que la loi exponentielle est stable par minimum, c'est-à-dire que le minimum de deux variables aléatoires indépendantes de loi exponentielle suit encore une loi exponentielle.

4. Soit $t \geq 0$. Montrer que les événements $\{X \leq Y\}$ et $\{Z > t\}$ sont indépendants.
   Il s'agit de montrer que $\mathbb{P}(X \leq Y, Z > t) = \mathbb{P}(X \leq Y)\mathbb{P}(Z > t)$. Les deux probabilités du membre de droite ont déjà été calculées dans les questions précédentes :
$$\mathbb{P}(X \leq Y) = \frac{\lambda}{\mu + \lambda} \quad \text{et} \quad \mathbb{P}(Z > t) = 1 - F_Z(t) = \mathrm{e}^{-(\lambda+\mu)t}\,.$$
   Calculons la probabilité $\mathbb{P}(X \leq Y, Z > t)$. Cette quantité peut se réécrire
$$\mathbb{P}(X \leq Y, Z > t) = \mathbb{P}(X \leq Y, \min(X,Y) > t) = \mathbb{P}(X \leq Y, X > t)\,,$$
   puisque l'inégalité $X \leq Y$ est équivalente à $\min(X,Y) = X$. Posons
$$B_t = \{(x,y) \in \mathbb{R}^2 : x \leq y,\, x > t\}\,.$$
   On obtient alors
$$\begin{aligned}
\mathbb{P}(X \leq Y, Z > t) &= \mathbb{P}((X,Y) \in B_t) \\
&= \int_{B_t} \lambda\,\mathrm{e}^{-\lambda x}\,\mu\,\mathrm{e}^{-\mu y}\,\mathbb{1}_{\{x\geq 0\}}\mathbb{1}_{\{y\geq 0\}}\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_0^{\infty} \lambda\,\mathrm{e}^{-\lambda x}\,\mathbb{1}_{\{x>t\}} \int_x^{\infty} \mu\,\mathrm{e}^{-\mu y}\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_t^{\infty} \lambda\,\mathrm{e}^{-\lambda x}\left[\mathrm{e}^{-\mu y}\right]_x^{\infty}\mathrm{d}x \\
&= \int_t^{\infty} \lambda\,\mathrm{e}^{-\lambda x}\,\mathrm{e}^{-\mu x}\,\mathrm{d}x \\
&= \left[\frac{\lambda}{\lambda+\mu}\,\mathrm{e}^{-(\lambda+\mu)x}\right]_t^{\infty} \\
&= \frac{\lambda}{\lambda+\mu}\,\mathrm{e}^{-(\lambda+\mu)t}\,.
\end{aligned}$$
   Ainsi, $\mathbb{P}(X \leq Y, Z > t) = \mathbb{P}(X \leq Y)\mathbb{P}(Z > t)$, ce qui prouve l'indépendance des événements $\{X \leq Y\}$ et $\{Z > t\}$.

Exercice 3 Soit X et Y deux variables aléatoires indépendantes de densité respective f_X et f_Y. Montrer que la variable aléatoire $\max(X,Y)$ admet une densité à déterminer.

La loi du maximum de deux variables aléatoires indépendantes se calcule facilement via la fonction de répartition. En effet, pour un réel t fixé, dire que $\max(X,Y) \leq t$ revient à dire que $X \leq t$ et $Y \leq t$. Par suite, comme X et Y indépendantes, la fonction de répartition de $\max(X,Y)$ s'écrit
$$\mathbb{P}(\max(X,Y) \leq t) = \mathbb{P}(X \leq t, Y \leq t) = \mathbb{P}(X \leq t)\mathbb{P}(Y \leq t) = F_X(t)F_Y(t)\,.$$
Les variables aléatoires X et Y étant à densité, les fonctions F_X et F_Y sont dérivables presque partout de dérivée respective f_X et f_Y. Ainsi, la fonction de répartition de $\max(X,Y)$ est dérivable presque partout de dérivée $f_XF_Y + F_Xf_Y$, ce qui prouve que $\max(X,Y)$ admet pour densité $f_XF_Y + F_Xf_Y$.

Pour aller plus loin

Soit X et Y deux variables aléatoires réelles indépendantes et de même loi dans $L^2(\Omega, \mathcal{F}, \mathbb{P})$ de sorte qu'il existe $c > 0$ tel que $X+Y$ et cX ont même loi.

1. Déterminer c, puis $\mathbb{E}[X]$, puis un développement limité à l'ordre 2 en 0 de φ_X.
2. En déduire que pour tout $n \geq 1$ et pour tout $t \in \mathbb{R}$,

$$\varphi_X(t) = \left[\varphi_X\left(\frac{t}{2^{n/2}}\right)\right]^{2^n},$$

puis conclure que X suit une loi normale.

On dit qu'une loi μ est stable si la somme de deux variables aléatoires indépendantes X et Y de loi μ suit encore, à un facteur multiplicatif près, la loi μ. Cet exercice montre que les lois stables de carré intégrable sont les lois normales.

1. Comme la loi de X et Y est dans $L^2(\Omega, \mathcal{F}, \mathbb{P})$, elle admet une variance. Par indépendance de X et Y on obtient

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) = 2\,\mathrm{Var}(X)\,,$$

et

$$\mathrm{Var}(cX) = c^2\,\mathrm{Var}(X)\,.$$

Comme les variables aléatoires $X+Y$ et cX ont la même loi, elles ont la même variance donc $2\,\mathrm{Var}(X) = c^2\,\mathrm{Var}(X)$.
- Si $\mathrm{Var}(X) = 0$, alors la loi de X est un Dirac en $\mathbb{E}[X]$. On confond cette loi avec une loi normale de variance 0. On exclut ce cas dans la suite.
- Si $\mathrm{Var}(X) > 0$, alors on obtient $c^2 = 2$ puis $c = \sqrt{2}$ puisque $c > 0$.
Un raisonnement similaire avec l'espérance donne $2\mathbb{E}[X] = \sqrt{2}\mathbb{E}[X]$, d'où $\mathbb{E}[X] = 0$.
Enfin, comme $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, la fonction caractéristique φ_X est dans $\mathcal{C}^2(\mathbb{R}, \mathbb{C})$ donc admet un développement limité à l'ordre deux en 0 donné par

$$\varphi_X(t) = \mathbb{E}[\mathrm{e}^{itX}] = \varphi_X(0) + \varphi_X'(0)t + \frac{\varphi_X''(0)}{2}t^2\,.$$

avec $\varphi_X^{(k)}(0) = i^k\mathbb{E}[X^k]$ pour $k = 0, 1, 2$. On obtient alors

$$\varphi_X(t) = 1 - \frac{\sigma^2}{2}t^2 + o(t^2)\,, \quad t \to 0\,,$$

où on a posé $\sigma^2 = \mathbb{E}[X^2] = \mathrm{Var}(X) > 0$.

2. Montrons le résultat par récurrence sur $n \geq 1$. Pour $n = 1$ l'égalité est trivialement vérifiée. Étudions tout de même le cas $n = 2$. En utilisant le fait que $X+Y$ et cX ont la même loi, on obtient

$$\varphi_X(t) = \mathbb{E}[\mathrm{e}^{iX}] = \mathbb{E}[\mathrm{e}^{i(t/c)(cX)}] = \varphi_{cX}(t/c) = \varphi_{X+Y}(t/c)\,, \quad t \in \mathbb{R}\,.$$

L'indépendance de X et Y implique alors que $\varphi_{X+Y} = \varphi_X\varphi_Y = (\varphi_X)^2$, d'où

$$\varphi_X(t) = (\varphi_X(t/\sqrt{2}))^2\,, \quad t \in \mathbb{R}\,,$$

ce qui prouve le résultat pour $n = 2$.
Supposons maintenant le résultat vrai pour $n \geq 1$. En reprenant les calculs précédents on peut écrire pour tout $t \in \mathbb{R}$,

$$\varphi_X(t) = (\varphi_X(t/\sqrt{2}))^2 .$$

Par ailleurs, l'hypothèse de récurrence assure que pour tout $u \in \mathbb{R}$,

$$\varphi_X(u) = \Big[\varphi_X\Big(\frac{u}{2^{n/2}}\Big)\Big]^{2^n} .$$

En combinant ces deux égalités on obtient, pour tout $t \in \mathbb{R}$,

$$\varphi_X(t) = (\varphi_X(t/\sqrt{2}))^2 = \Bigg(\Big[\varphi_X\Big(\frac{(t/\sqrt{2})}{2^{n/2}}\Big)\Big]^{2^n}\Bigg)^2 = \Big[\varphi_X\Big(\frac{t}{2^{(n+1)/2}}\Big)\Big]^{2^{n+1}} ,$$

ce qui prouve l'hérédité. Bref, l'égalité est vraie pour tout $n \geq 1$.
On réécrit alors le développement limité en $t/2^{n/2}$ pour $n \to \infty$ et $t \in \mathbb{R}$ fixé :

$$\varphi_X(t) = \Big[\varphi_X\Big(\frac{t}{2^{n/2}}\Big)\Big]^{2^n} = \Big[1 - \frac{\sigma^2}{2}\Big(\frac{t}{2^{n/2}}\Big)^2 + o\Big(\Big(\frac{t}{2^{n/2}}\Big)^2\Big)\Big]^{2^n} = \Big[1 - \frac{\sigma^2}{2}\frac{t^2}{2^n} + o\Big(\frac{1}{2^n}\Big)\Big)\Big]^{2^n} ,$$

quand $n \to \infty$. En passant à la limite quand $n \to \infty$, on en déduit que

$$\varphi_X(t) = \lim_{n\to\infty} \Big[1 - \frac{\sigma^2}{2}\frac{t^2}{2^n} + o\Big(\frac{1}{2^n}\Big)\Big)\Big]^{2^n} .$$

La limite classique $(1 - u/n)^n \to \mathrm{e}^{-u}$ quand $n \to \infty$ permet alors de conclure :

$$\varphi_X(t) = \lim_{n\to\infty} \Big[1 - \frac{\sigma^2}{2}\frac{t^2}{2^n} + o\Big(\frac{1}{2^n}\Big)\Big)\Big]^{2^n} = \exp\Big(-\frac{\sigma^2}{2}t^2\Big), \quad t \in \mathbb{R} .$$

On reconnaît la fonction caractéristique d'une loi normale $\mathcal{N}(0, \sigma^2)$ ce qui permet de conclure que $X \ \sim \mathcal{N}(0, \sigma^2)$.

TD 7 - Lemme de Borel-Cantelli et loi du 0-1

Exercice 1 On considère l'espace probabilisé $([0,1], \mathcal{B}([0,1]), \mathbb{P})$, où $\mathbb{P}$ est la mesure de Lebesgue, et on pose $A_n =]0, 1/n]$ pour tout entier $n \geq 1$.

1. Expliciter l'événement $\limsup_{n\to\infty} A_n$.
   Rappelons la définition de $\limsup_{n\to\infty} A_n$:
   $$\limsup_{n\to\infty} A_n = \bigcap_{n\geq 1} \bigcup_{k\geq n} A_k\,.$$
   Ici, pour $n \geq 1$ fixé, on a
   $$\bigcup_{k\geq n} A_k = \bigcup_{k\geq n} \left]0, \frac{1}{k}\right] = \left]0, \frac{1}{n}\right]\,.$$
   En prenant l'intersection de ces ensembles pour tout $n \geq 1$, on obtient alors
   $$\limsup_{n\to\infty} A_n = \emptyset\,.$$

2. Calculer $\sum_{n=1}^{\infty} \mathbb{P}(A_n)$. Commenter.
   La mesure de Lesbesgue de A_n est $1/n$, donc
   $$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty\,.$$
   La suite d'événements $(A_n)_{n\geq 1}$ vérifie $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$ mais $\mathbb{P}(\limsup_{n\to\infty} A_n) = 0$. C'est un contre-exemple à la réciproque du lemme de Borel-Cantelli dans le cas où les événements ne sont pas indépendants. En effet, si $n \neq k$, avec par exemple $2 \leq n < k$, alors $A_n \cap A_k = A_n$ et donc
   $$\mathbb{P}(A_n \cap A_k) = \mathbb{P}(A_k) = \frac{1}{k} \neq \frac{1}{n} \times \frac{1}{k} = \mathbb{P}(A_n)\mathbb{P}(A_k)\,.$$

Exercice 2 On lance une infinité de fois une pièce de monnaie équilibrée et on considère un entier $k \in \mathbb{N}^*$. Montrer qu'avec probabilité 1, on obtiendra une infinité de fois k Pile consécutifs.

On considère les événements

- A_1 : "on n'obtient que des Pile entre les lancers 1 et k",
- A_2 : "on n'obtient que des Pile entre les lancers $k+1$ et $2k$",
- A_3 : "on n'obtient que des Pile entre les lancers $2k+1$ et $3k$",
- etc.

Bref, A_n est l'événement "on n'obtient que des Pile entre les lancers $(n-1)k+1$ et nk".

Par indépendance des lancers, les événements A_n sont indépendants. Par ailleurs, $\mathbb{P}(A_n) = 1/2^k$, donc $\sum_{n=0}^{\infty} \mathbb{P}(A_n) = \infty$ (attention, k est fixé et on somme sur n). Ainsi, d'après le lemme de Borel-Cantelli, une infinité d'événements A_n est réalisée presque sûrement. On en déduit que presque sûrement on obtient une infinité de fois k Pile consécutifs.

Exercice 3 Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles indépendantes et $\mathcal{F}_n = \sigma(X_n)$. Montrer que l'ensemble des $\omega \in \Omega$ tels que $S_n(\omega) = \sum_{k=1}^{n} X_k(\omega)$ converge est dans la tribu terminale de la suite $(\mathcal{F}_n)_{n\geq 1}$. En déduire que cet événement est certain ou impossible.

Remarque : l'objectif de l'exercice est de montrer que la convergence d'une somme de variables aléatoires est soit presque sûre, soit impossible ; en particulier, on n'a par exemple pas une chance sur deux que la somme converge.

On souhaite donc montrer que l'événement A défini par

$$A = \{\omega \in \Omega : S_n(\omega) \text{ converge}\}$$

est dans la tribu terminale. Intuitivement, l'ensemble des ω pour lesquels A est vérifié ne dépend que du comportement asymptotique de S_n, ce qui explique pourquoi cet événement est dans la tribu asymptotique. Justifions ceci.

Soit $\omega \in \Omega$. Pour $j \geq 1$ fixé, la suite $S_n(\omega)$ converge si et seulement si la suite tronquée

$$\sum_{k=j}^{n} X_k(\omega)$$

converge. En effet, la partie manquante correspond à $S_{j-1} = \sum_{k=1}^{j-1} X_k(\omega)$ qui ne dépend pas de n donc n'a pas d'effet sur la convergence (ou non) de la série. On peut donc écrire

$$A = \left\{\omega \in \Omega : \sum_{k=j}^{n} X_k(\omega) \text{ converge}\right\},$$

ce qui implique que $A \in \bar{\mathcal{F}}_k = \sigma(\mathcal{F}_k, k \geq j)$. Ceci étant vrai pour tout $j \geq 1$, on obtient $A \in \cap_{j\geq 1} \bar{\mathcal{F}}_k = \mathcal{F}_\infty$, ce qui signifie que A appartient à la tribu terminale de la famille $(\mathcal{F}_n)_{n\geq 1}$.

Enfin, la famille de tribus $(\mathcal{F}_n)_{n\geq 1} = (\sigma(X_n))_{n\geq 1}$ est indépendante puisque les variables aléatoires X_n le sont. Donc la loi du $0-1$ assure alors que A est un événement de probabilité 0 ou 1.

Attention : si la série est convergente presque sûrement, la valeur de la somme $\sum_{k=1}^{\infty} X_k(\omega)$ n'est pas dans la tribu terminale $\mathcal{F}_\infty$. En effet, cette somme dépend évidemment des premières valeur de la série.

Exercice 4 Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles indépendantes telles que pour tout $n \geq 1$, $\mathbb{P}(X_n = 1) = p = 1 - \mathbb{P}(X_n = -1)$, avec $p \in]0,1[$. On pose $S_0 = 0$ et $S_n = X_1 + \cdots + X_n$. Enfin, on considère l'événement $A_n = \{S_n = 0\}$.

1. Déterminer la probabilité de A_n. On pourra distinguer suivant la parité de n.

L'événement $A_n = \{S_n = 0\}$ est vérifié si et seulement s'il y a eu autant de X_k prenant la valeur 1 que de X_k prenant la valeur -1. C'est impossible si n est impair : $\mathbb{P}(A_{2k+1}) = 0$. Dans le cas pair, on obtient le résultat en utilisant l'indépendance des X_k :

$$\begin{aligned}
&\mathbb{P}(A_{2n}) \\
&= \mathbb{P}(\exists 1 \leq i_1 < \cdots < i_n \leq n : X_{i_1} = \cdots = X_{i_n} = 1 \text{ et } X_i = 0 \text{ pour tout } i \notin \{i_1, \ldots, i_n\}) \\
&= \binom{2n}{n} \mathbb{P}(X_{i_1} = \cdots = X_{i_n} = 1 \text{ et } X_i = 0 \text{ pour tout } i \notin \{i_1, \ldots, i_n\}) \\
&= \binom{2n}{n} p^n (1-p)^n .
\end{aligned}$$

2. À l'aide de la formule de Stirling, déterminer un équivalent de $\mathbb{P}(A_{2n})$.
   La formule de Stirling donne

$$\binom{2n}{n} = \frac{(2n)!}{(n!)^2} \sim \frac{(2n)^{2n}\sqrt{2\pi 2n}}{\mathrm{e}^{2n}} \times \left(\frac{\mathrm{e}^n}{n^n \sqrt{2\pi n}}\right)^2 = \frac{4^n}{\sqrt{\pi n}}, \quad n \to \infty .$$

   On en déduit alors un équivalent de $\mathbb{P}(A_{2n})$ (on rappelle que $p \in]0,1[$) :

$$\mathbb{P}(A_{2n}) \sim \frac{(4p(1-p))^n}{\sqrt{\pi n}}, \quad n \to \infty .$$

3. On pose $A = \limsup_{n\to\infty} A_n$. Que représente cet événement ? Déterminer $\mathbb{P}(A)$ dans le cas $p \neq \frac{1}{2}$.
   L'événement $A = \limsup_{n\to\infty} A_n$ est réalisé si et seulement si S_n repasse une infinité de fois par 0.
   Si $p \neq \frac{1}{2}$, alors la série $\sum \mathbb{P}(A_n)$ est convergente, donc le lemme de Borel-Cantelli assure que

$$\mathbb{P}(A) = \mathbb{P}(\limsup_{n\to\infty} A_n) = 0 .$$

   Cela signifie que presque sûrement S_n ne repasse qu'un nombre fini de fois par 0.
4. On suppose maintenant que $p = \frac{1}{2}$. On va montrer que $\mathbb{P}(A) = 1$.
   (a) Expliquer pourquoi le lemme de Borel-Cantelli ne s'applique pas.
   Si $p = \frac{1}{2}$, alors $\mathbb{P}(A_{2n}) \sim \frac{1}{\sqrt{2\pi n}}$ donc la série $\sum \mathbb{P}(A_n)$ diverge. Cependant, on n'a pas l'indépendance des événements A_n donc le lemme de Borel-Cantelli ne s'applique pas.
   (b) Montrer que pour tout $n \geq 1$ et pour tout $k \geq 1$, les vecteurs aléatoires $(X_{k+1}, \ldots, X_{k+n})$ et $(X_1, \ldots, X_n)$ ont la même loi.
   Cela résulte du caractère iid de la suite $(X_n)_{n\geq 1}$. En effet, soit $n \geq 1$ fixé et $A_1, \ldots, A_n$ des éléments de $\mathcal{B}(\mathbb{R})$. Alors par indépendance de $X_1, \ldots, X_n$, on a

$$\begin{aligned}
\mathbb{P}((X_1, \ldots, X_n) \in A_1 \times \cdots \times A_n) &= \mathbb{P}(X_1 \in A_1, \ldots, X_n \in A_n) \\
&= \mathbb{P}(X_1 \in A_1) \cdots \mathbb{P}(X_n \in A_n) .
\end{aligned}$$

De même, si $k \geq 1$ fixé alors l'indépendance de $X_{k+1}, \ldots, X_{k+n}$ implique que

$$\begin{aligned}\mathbb{P}((X_{k+1}, \ldots, X_{k+n}) \in A_1 \times \cdots \times A_n) &= \mathbb{P}(X_{k+1} \in A_1, \ldots, X_{k+n} \in A_n) \\ &= \mathbb{P}(X_{k+1} \in A_1) \cdots \mathbb{P}(X_{k+n} \in A_n)\,.\end{aligned}$$

Enfin, comme les variables aléatoires X_j ont toutes la même loi, on a les égalités

$$\mathbb{P}(X_1 \in A_1) = \mathbb{P}(X_{k+1} \in A_1), \quad \ldots, \quad \mathbb{P}(X_n \in A_n) = \mathbb{P}(X_{k+n} \in A_n)\,.$$

On en déduit donc l'égalité

$$\mathbb{P}((X_1, \ldots, X_n) \in A_1 \times \cdots \times A_n) = \mathbb{P}((X_{k+1}, \ldots, X_{k+n}) \in A_1 \times \cdots \times A_n)\,.$$

Celle-ci étant valable pour tout $A_1, \ldots, A_n$ dans $\mathcal{B}(\mathbb{R})$, on en déduit que les vecteurs aléatoires $(X_{k+1}, \ldots, X_{k+n})$ et $(X_1, \ldots, X_n)$ ont la même loi.

(c) On rappelle que $A = \limsup_{n\to\infty} A_n$. Montrer que

$$A^c = \bigcup_{k=0}^{\infty} \{S_k = 0 \text{ et } \forall n \geq 1,\ S_{n+k} \neq 0\}\,.$$

et que cette union est formée d'événements disjoints.
Par définition de la limite supérieure d'une suite d'ensembles, on a

$$A^c = (\limsup_{n\to\infty} A_n)^c = \Big(\bigcap_{k=0}^{\infty} \bigcup_{n=k}^{\infty} A_n\Big)^c = \bigcup_{k=0}^{\infty} \bigcap_{n=1}^{\infty} \{S_{n+k} \neq 0\} = \bigcup_{k=0}^{\infty} \{\forall n \geq 1,\ S_{n+k} \neq 0\}\,.$$

Montrons alors que

$$\bigcup_{k=0}^{\infty} \{\forall n \geq 1,\ S_{n+k} \neq 0\} = \bigcup_{k=0}^{\infty} \{S_k = 0 \text{ et } \forall n \geq 1,\ S_{n+k} \neq 0\}\,.$$

L'inclusion $\supset$ est claire. Réciproquement, s'il existe $k \geq 0$ telle que pour tout $n \geq 1$, $S_{n+k} \neq 0$, alors on pose

$$k' = \max\{j \in \{0, \ldots, k\} : S_j = 0\}$$

(on prend le dernier moment où la suite passe par 0). Notons que k' existe car c'est le maximum d'une partie finie de $\mathbb{N}$ qui est non vide (car 0 appartient à cet ensemble). Par définition $S_{k'} = 0$ et pour tout $n \geq 1$, $S_{k'+n} \neq 0$. Ceci prouve l'autre inclusion. D'où le résultat :

$$A^c = \bigcup_{k=0}^{\infty} \{S_k = 0 \text{ et } \forall n \geq 1,\ S_{n+k} \neq 0\}\,.$$

Enfin, notons B_k les événements qui forment cette union. Soit alors $k \neq j$, disons $k > j$. Sur l'événement B_k, on a $S_k = 0$, tandis que sur B_j on a $S_k \neq 0$. Donc les événements B_k et B_j sont disjoints.

(d) En déduire que $\mathbb{P}(A) = 1$.

L'union étant formée d'événements disjoints on peut écrire

$$\mathbb{P}(A^c) = \sum_{k=0}^{\infty} \mathbb{P}(S_k = 0 \text{ et } \forall n \geq 1,\ S_{n+k} \neq 0)\,.$$

Puis, comme $S_{n+k} = S_k + X_{k+1} + \cdots + X_{k+n}$, on peut réécrire la probabilité précédente sous la forme

$$\begin{aligned}\mathbb{P}(A^c) &= \sum_{k=0}^{\infty} \mathbb{P}(S_k = 0 \text{ et } \forall n \geq 1,\ X_{k+1} + \cdots + X_{k+n} \neq 0) \\ &= \sum_{k=0}^{\infty} \mathbb{P}(S_k = 0)\mathbb{P}(\forall n \geq 1,\ X_{k+1} + \cdots + X_{k+n} \neq 0)\,,\end{aligned}$$

où on a utilisé l'indépendance de la somme S_k (qui ne dépend que des variables aléatoires $X_1, \ldots, X_k$) et de la somme $X_{k+1} + \cdots + X_{k+n}$ (qui ne dépend que des variables aléatoires $X_{k+1}, \ldots, X_{k+n}$).

En utilisant la question (b), on en déduit alors

$$\begin{aligned}\mathbb{P}(A^c) &= \sum_{k=0}^{\infty} \mathbb{P}(S_k = 0)\mathbb{P}(\forall n \geq 1,\ X_1 + \cdots + X_n \neq 0) \\ &= \mathbb{P}(\forall n \geq 1,\ X_1 + \cdots + X_n \neq 0) \sum_{k=0}^{\infty} \mathbb{P}(S_k = 0)\,.\end{aligned}$$

Cette quantité est dans $[0, 1]$. Or $\mathbb{P}(S_k = 0) = \mathbb{P}(A_k)$ est le terme général d'une série divergente, donc $\mathbb{P}(\forall n \geq 1,\ X_1 + \cdots + X_n \neq 0) = 0$. Ainsi, $\mathbb{P}(A^c) = 0$ et donc

$$\mathbb{P}(A) = 1\,.$$

Pour aller plus loin

Montrer qu'il n'existe pas de mesure de probabilité uniforme sur $\mathbb{N}$, au sens où $\mathbb{P}(n\mathbb{N}) = 1/n$ pour tout $n \geq 1$. On pourra considérer les événements $p\mathbb{N}$, où p est un nombre premier.

Supposons qu'une telle probabilité $\mathbb{P}$ existe. On considère les événements

$$A_p = p\mathbb{N} = \{0, p, 2p, 3p, \ldots\}\,,$$

pour tout nombre premier p.

Remarquons que si $p_1, \ldots, p_k$ sont des nombres premiers distincts, alors $A_{p_1} \cap \cdots \cap A_{p_k} = p_1 \cdots p_k\mathbb{N}$: en effet, un entier est un divisible par $p_1, \ldots, p_k$ si et seulement s'il est divisible par leur produit. On en déduit alors que les événements (A_p) sont indépendants puisque pour tous $p_1, \ldots, p_k$, nombres premiers distincts, on a

$$\mathbb{P}(A_{p_1} \cap \cdots \cap A_{p_k}) = \mathbb{P}(p_1 \cdots p_k\mathbb{N}) = \frac{1}{p_1 \cdots p_k} = \prod_{i=1}^{k} \mathbb{P}(A_{p_i})\,.$$

Comme $\sum_{p \text{ premiers}} \frac{1}{p} = \infty$, on en déduit alors par le deuxième point du lemme de Borel-Cantelli que $\mathbb{P}$-presque tout entier n est dans une infinité d'ensemble $p\mathbb{N}$ avec p premier, et donc est multiple d'une infinité de nombres premiers distincts. C'est clairement absurde.

TD 8 - Convergence de variables aléatoires

Exercice 1

1. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes. Montrer que (X_n) converge vers 0 presque sûrement si et seulement si pour tout $\epsilon > 0$ la série $\sum \mathbb{P}(|X_n| \geq \epsilon)$ est convergente.

   - Montrons l'implication directe.

   Supposons que $(X_n)_{n\geq 1}$ converge vers 0 presque sûrement. Cela signifie que

   $$\mathbb{P}(\forall \epsilon > 0,\, \exists k \geq 1,\, \forall n \geq k,\, |X_k| < \epsilon) = 1\,,$$

   Du point de vue des événements, on obtient $\mathbb{P}(\cap_{\epsilon>0} B_\epsilon) = 1$, avec

   $$B_\epsilon = \bigcup_{n\geq 1} \bigcap_{k\geq n} \{|X_k| < \epsilon\}\,.$$

   Soit alors $\epsilon > 0$. L'inclusion $\cap_{\alpha>0} B_\alpha \subset B_\epsilon$ assure que $\mathbb{P}(B_\epsilon) = 1$, donc que $\mathbb{P}(B_\epsilon^c) = 0$, ce qui se réécrit

   $$\mathbb{P}\Big(\bigcap_{n\geq 1} \bigcup_{k\geq n} \{|X_k| \geq \epsilon\}\Big) = 0\,.$$

   Ainsi, $\mathbb{P}(\limsup_{n\to\infty} \{|X_n| \geq \epsilon\}) = 0$. Or les événements $\{|X_n| \geq \epsilon\}$ sont indépendants car les variables aléatoires X_n le sont, donc d'après le lemme de Borel-Cantelli la série $\sum \mathbb{P}(|X_n| \geq \epsilon)$ est convergente. Ceci étant vrai pour tout $\epsilon > 0$ on a bien montré la première implication.

   - Montrons maintenant la réciproque.

   Comme la série $\sum \mathbb{P}(|X_n| \geq \epsilon)$ est convergente pour tout $\epsilon > 0$, le lemme de Borel-Cantelli implique que

   $$\mathbb{P}(\limsup_{n\to\infty} \{|X_n| \geq \epsilon\}) = \mathbb{P}\Big(\bigcap_{k\geq 1} \bigcup_{n\geq k} \{|X_n| \geq \epsilon\}\Big) = 0\,.$$

   Par passage au complémentaire, on en déduit que pour tout $\epsilon > 0$,

   $$\mathbb{P}\Big(\bigcup_{k\geq 1} \bigcap_{n\geq k} \{|X_n| < \epsilon\}\Big) = 1\,.$$

   On choisit alors des $\epsilon > 0$ rationnels et on utilise le fait qu'une intersection dénombrable d'événements presque sûrs est encore presque sûr :

   $$\mathbb{P}\Big(\bigcap_{\epsilon\in\mathbb{Q}_+^*} \bigcup_{k\geq 1} \bigcap_{n\geq k} \{|X_n| < \epsilon\}\Big) = 1\,.$$

   Ceci correspond à la convergence presque sûre $X_n \xrightarrow[n\to\infty]{p.s.} 0$.

2. On suppose maintenant X_n suit une loi de Bernoulli de paramètre p_n.
   (a) Montrer que la suite (X_n) converge vers 0 en probabilité si et seulement si $p_n \to 0$ quand $n \to \infty$.
   Comme la variable aléatoire X_n suit une loi de Bernoulli, elle prend les valeurs 0 et 1, donc pour tout $\epsilon > 0$, on a l'inégalité
   $$\mathbb{P}(|X_n| \geq \epsilon) \leq \mathbb{P}(X_n = 1) = p_n\,,$$
   avec égalité si $\epsilon \leq 1$.
   Ainsi, si $p_n \to 0$, alors $\mathbb{P}(|X_n| \geq \epsilon) \to 0$ pour tout $\epsilon > 0$, donc $X_n \xrightarrow{\mathbb{P}} 0$. Et réciproquement, si $X_n \xrightarrow{\mathbb{P}} 0$ alors avec $\epsilon = 1$, on obtient $\mathbb{P}(|X_n| \geq 1) = \mathbb{P}(X_n = 1) \to 0$, donc $p_n \to 0$.
   (b) Montrer que la suite (X_n) converge vers 0 dans L^p si et seulement si $p_n \to 0$ quand $n \to \infty$.
   Les moments d'ordre p de la variable aléatoire positive X_n sont
   $$\mathbb{E}[X_n^p] = 1^p \times p_n + 0^p \times (1 - p_n) = p_n\,.$$
   Donc $X_n \xrightarrow{L^p} 0$ si et seulement si $p_n \to 0$.
   (c) On suppose que les X_n sont indépendantes. Montrer que la suite (X_n) converge vers 0 presque sûrement si et seulement si la série $\sum p_n$ converge.
   On part de l'inégalité $\mathbb{P}(|X_n| \geq \epsilon) \leq \mathbb{P}(X_n = 1) = p_n$ établie plus haut. Si la série $\sum p_n$ converge, alors la série $\sum \mathbb{P}(|X_n| \geq \epsilon)$ converge pour tout $\epsilon > 0$ et la question 1. permet de conclure.
   De même, si la suite (X_n) converge presque sûrement vers 0, alors la question 1. assure la convergence de la série $\sum \mathbb{P}(|X_n| \geq \epsilon)$ pour tout $\epsilon > 0$. En particulier pour $\epsilon = 1$, on obtient la convergence de la série $\sum p_n$.
3. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes telles que $X_n \sim \mathcal{E}(n)$. Montrer que $X_n \to 0$ presque sûrement. Y a-t-il convergence dans L^p ?
   On applique le résultat de la question 1 de l'exercice précédent. Pour $\epsilon > 0$ et $n \geq 1$, on a
   $$\mathbb{P}(X_n \geq \epsilon) = \int_\epsilon^\infty n\,\mathrm{e}^{-nt}\,\mathrm{d}t = \mathrm{e}^{-n\epsilon}\,.$$
   La série de terme général $\mathrm{e}^{-n\epsilon}$ étant convergente pour tout $\epsilon > 0$, l'exercice précédent permet de conclure.
   On rappelle que les moments d'ordre p d'une variable aléatoire X de loi exponentielle de paramètre λ sont donnés par $\mathbb{E}[X^p] = \frac{p!}{\lambda^p}$. Ici, on obtient donc
   $$\mathbb{E}[X_n^p] = \frac{p!}{n^p} \to 0\,, \quad n \to \infty\,,$$
   ce qui implique la convergence $X_n \xrightarrow[n\to\infty]{L^p} 0$.

Exercice 2 Soit $\alpha > 0$. On considère une suite de variables aléatoires indépendantes $(Z_n)_{n\geq 1}$ de loi de Bernoulli de paramètre $n^{-\alpha}$. Montrer que $Z_n \to 0$ dans L^1 mais que

— si $\alpha \leq 1$, $\mathbb{P}(\limsup_{n\to\infty} Z_n = 1) = 1$,

— si $\alpha > 1$, $\mathbb{P}(\limsup_{n\to\infty} Z_n = 0) = 1$.

On pourra considérer les événements $A_n = \{Z_n = 1\}$.

Pour la convergence dans L^1 cela résulte de

$$\mathbb{E}[|Z_n|] = \mathbb{E}[Z_n] = n^{-\alpha} \to 0\,, \quad n \to \infty\,.$$

Remarquons tout d'abord que $\mathbb{P}(A_n) = \mathbb{P}(Z_n = 1) = n^{-\alpha}$.

Si $\alpha \leq 1$, alors la série de terme général $\mathbb{P}(A_n)$ diverge. Comme les variables aléatoires Z_n sont indépendantes, les événements A_n sont également indépendants et on peut donc appliquer le lemme de Borel-Cantelli qui donne $\mathbb{P}(\limsup_{n\to\infty} A_n) = 1$. Cela signifie que

$$\mathbb{P}(\forall n \geq 1,\ \exists k \geq n,\ Z_k = 1) = 1\,,$$

et donc $\mathbb{P}(\limsup_{n\to\infty} Z_n = 1) = 1$.

Si $\alpha > 1$, alors la série de terme général $\mathbb{P}(A_n)$ converge. Le lemme de Borel-Cantelli assure alors que $\mathbb{P}(\limsup_{n\to\infty} A_n) = 0$. En passant au complémentaire on obtient alors

$$\mathbb{P}(\exists n \geq 1,\ \forall k \geq n,\ Z_k = 0) = 1\,,$$

et donc $\limsup_{n\to\infty} Z_n = 0$ presque sûrement. En particulier, cela implique que $Z_n \xrightarrow[n\to\infty]{p.s.} 0$.

Pour aller plus loin

Soit $n \geq 1$ et $(X_k)_{k\geq 1}$ une suite de variables aléatoires indépendantes uniformément distribuées sur $\{1, \ldots, n\}$. On pose

$$T_n = \inf\{m \geq 1 : \{X_1, \ldots, X_m\} = \{1, \ldots, n\}\}\,.$$

et

$$\tau_n(k) = \inf\{m \geq 1 : \#\{X_1, \ldots, X_m\} = k\}\,.$$

L'objectif de l'exercice est de déterminer le comportement asymptotique de T_n quand n est grand. Cet exercice est appelé *problème du collectionneur*. Expliquer pourquoi (cela vous aidera en particulier à mieux appréhender les variables aléatoires T_n et $\tau_n(k)$). Pour simplifier les calculs, on pose $\tau_n(0) = 0$.

Cette situation modélise la collection d'objets numérotés de 1 à n. Le collectionneur achète à l'aveugle un objet parmi les n objets de possibles et la variable aléatoire X_k représente le k-ème objet acheté. La variable aléatoire T_n correspond quant à elle au premier instant où la collection est complète : l'ensemble des objets achetés coïncide avec l'ensemble des n objets possibles. Enfin, la variable aléatoire $\tau_n(k)$ représente le premier instant où le collectionneur a réussi à rassembler k objets distincts.

1. Montrer que les variables aléatoires $(\tau_n(k) - \tau_n(k-1))_{1\leq k\leq n}$ sont indépendantes et que $\tau_n(k) - \tau_n(k-1)$ suit une loi géométrique de paramètre $1 - \frac{k-1}{n}$.
   Remarquons tout d'abord que la variable aléatoire $\tau_n(1) - \tau_n(0) = 1$ est constante donc suit bien une loi géométrique de paramètre $1 = (n+1-1)/n$ et est indépendante des autres.

L'inégalité $\tau_n(k) > \tau_n(k-1)$ assure que les variables aléatoires $\tau_n(k) - \tau_n(k-1)$ sont à valeurs dans $\mathbb{N}^*$. Soit alors $j_1, \dots, j_n$ des entiers non nuls. L'événement

$$\{\tau_n(1) - \tau_n(0) = j_1, \dots, \tau_n(n) - \tau_n(n-1) = j_n\}$$

représente la situation où le premier objet a été obtenu lors du j_1-ème achat, le deuxième objet a été obtenu lors du $j_1 + j_2$-ème achat, etc. le n-ème objet lors du $j_1 + j_2 + \cdots + j_n$-ème achat. Le premier objet étant forcément obtenu lors du premier achat, l'événement précédent n'est réalisé que si $j_1 = 1$ (ce qui concorde avec la remarque préliminaire), ce que l'on suppose désormais. Par ailleurs, l'ordre dans lequel les objets apparaissent est libre donc correspond à une permutation de $\{1, \dots, n\}$. On en déduit alors que

$$\begin{aligned}
&\mathbb{P}(\tau_n(1) - \tau_n(0) = j_1, \tau_n(2) - \tau_n(1) = j_2, \dots, \tau_n(n) - \tau_n(n-1) = j_n) \\
&\qquad = \mathbb{P}(\exists \sigma \in \mathcal{S}_n \,:\, X_{j_1} = \cdots = X_{j_1+j_2-1} = \sigma(1), X_{j_1+j_2} = \sigma(2), \\
&\qquad\qquad X_{j_1+j_2+1} \in \{\sigma(1), \sigma(2)\}, \dots, X_{j_1+j_2+j_3-1} \in \{\sigma(1), \sigma(2)\}, X_{j_1+j_2+j_3} = \sigma(3), \\
&\qquad\qquad X_{j_1+j_2+j_3+1} \in \{\sigma(1), \sigma(2), \sigma(3)\}, \dots, X_{j_1+\cdots+j_{n-1}} = \sigma(n-1), \\
&\qquad\qquad X_{j_1+\cdots+j_{n-1}+1} \in \{\sigma(1), \dots \sigma(n-1)\}, X_{j_1+\cdots+j_n} = \sigma(n))\,.
\end{aligned}$$

Chaque permutation σ donne un autre ordre d'apparition des objets donc l'union sur $\mathcal{S}_n$ est formée d'événements disjoints. De plus, l'indépendance des variables aléatoires X_k et le fait qu'elles suivent une loi uniforme sur $\{1, \dots, n\}$ assurent alors que cette probabilité se factorise en

$$\begin{aligned}
\mathbb{P}(\tau_n(2)-&\tau_n(1) = j_1, \dots, \tau_n(n) - \tau_n(n-1) = j_n) \\
&= \sum_{\sigma \in \mathcal{S}_n} \left[\frac{1}{n} \times \left(\frac{1}{n}\right)^{j_2-1} \times \frac{1}{n} \times \left(\frac{2}{n}\right)^{j_3-1} \times \frac{1}{n} \times \cdots \left(\frac{n-1}{n}\right)^{j_n-1} \times \frac{1}{n}\right] \\
&= \sum_{\sigma \in \mathcal{S}_n} \left[\frac{1}{n^n} \prod_{k=2}^{n} \left(\frac{k-1}{n}\right)^{j_k-1}\right].
\end{aligned}$$

La somme ne dépend plus de σ donc

$$\sum_{\sigma \in \mathcal{S}_n} \left[\frac{1}{n^n} \prod_{k=2}^{n} \left(\frac{k-1}{n}\right)^{j_k-1}\right] = \frac{n!}{n^n} \prod_{k=2}^{n} \left(\frac{k-1}{n}\right)^{j_k-1} = \prod_{k=2}^{n} \left(\frac{n+1-k}{n}\right)\left(\frac{k-1}{n}\right)^{j_k-1}.$$

Bref, on vient de montrer que

$$\mathbb{P}(\tau_n(2) - \tau_n(1) = j_1, \dots, \tau_n(n) - \tau_n(n-1) = j_n) = \prod_{k=2}^{n} \left(\frac{n+1-k}{n}\right)\left(\frac{k-1}{n}\right)^{j_k-1},$$

ce qui prouve que les variables aléatoires $\tau_n(k) - \tau_n(k-1)$ sont indépendantes de loi

$$\mathbb{P}(\tau_n(k) - \tau_n(k-1) = j_k) = \left(1 - \frac{k-1}{n}\right)\left(\frac{k-1}{n}\right)^{j_k-1}.$$

Donc $\tau_n(k) - \tau_n(k-1)$ suit une loi géométrique de paramètre $1 - \frac{k-1}{n}$.

2. Exprimer T_n en fonction des différences $\tau_n(k) - \tau_n(k-1)$, puis déterminer $\mathbb{E}[T_n]$ et enfin montrer que

$$\mathrm{Var}(T_n) \leq Cn^2 \,,$$

où $C > 0$ est une constante.

Il est clair que $T_n = \tau_n(n)$: le premier instant où la collection est complète correspond au premier instant où on a obtenu n objets différents. On réécrit cela à l'aide des $\tau_n(k) - \tau_n(k-1)$:

$$T_n = \tau_n(n) = \sum_{k=1}^{n} (\tau_n(k) - \tau_n(k-1)) \,,$$

où on rappelle que $\tau_n(0) = 0$.

Rappelons que l'espérance d'une loi géométrique de paramètre p vaut $\frac{1}{p}$ et que sa variance vaut $\frac{1-p}{p^2}$. Par linéarité de l'espérance, on obtient alors

$$\mathbb{E}[T_n] = \sum_{k=1}^{n} \mathbb{E}[\tau_n(k) - \tau_n(k-1)] = \sum_{k=1}^{n} \frac{n}{n+1-k} = n \sum_{j=1}^{n} \frac{1}{j} \,,$$

après avoir utilisé le changement de variable $j = n+1-k$. Les variables aléatoires $\tau_n(k) - \tau_n(k-1)$ étant indépendantes, la variance est linéaire, d'où

$$\mathrm{Var}(T_n) = \sum_{k=1}^{n} \mathrm{Var}(\tau_n(k) - \tau_n(k-1)) = \sum_{k=1}^{n} \frac{1 - \frac{n+1-k}{n}}{\left(\frac{n+1-k}{n}\right)^2} = n \sum_{k=1}^{n} \frac{k-1}{(n+1-k)^2} \,.$$

En utilisant le changement de variable $j = n+1-k$ on en déduit que

$$\mathrm{Var}(T_n) = n \sum_{j=1}^{n} \frac{n-j}{j^2} \leq n^2 \sum_{j=1}^{n} \frac{1}{j^2} \,.$$

La somme des inverses des carrés des entiers étant convergente, la somme partielle est bornée par une constante $C > 0$. Ainsi, $\mathrm{Var}(T_n) \leq Cn^2$.

3. Soit $\epsilon > 0$. Montrer que

$$\mathbb{P}(|T_n - \mathbb{E}[T_n]| \geq \epsilon n \ln(n)) \xrightarrow[n \to \infty]{} 0 \,,$$

puis en déduire que $T_n/(n \ln(n)) \to 1$ en probabilité.

On rappelle que $\sum_{k=1}^{n} \frac{1}{k} \sim \ln(n)$ *quand* $n \to \infty$*.*

L'inégalité de Bienaymé-Techbychev donne le premier résultat :

$$\mathbb{P}(|T_n - \mathbb{E}[T_n]| \geq \epsilon n \ln(n)) \leq \frac{\mathrm{Var}(T_n)}{\epsilon^2 n^2 \ln(n)^2} \leq \frac{C}{\epsilon^2 \ln(n)^2} \xrightarrow[n \to \infty]{} 0 \,.$$

Puis, on écrit

$$\frac{T_n}{n \ln(n)} - 1 = \frac{T_n - \mathbb{E}[T_n]}{n \ln(n)} + \left(\frac{\mathbb{E}[T_n]}{n \ln(n)} - 1\right) .$$

D'après le rappel, le terme de droite tend vers 0. Pour n assez grand, cette quantité est donc inférieure à ϵ, d'où l'inégalité

$$\mathbb{P}\Big(\Big|\frac{T_n}{n\ln(n)} - 1\Big| \geq \epsilon\Big) \leq \mathbb{P}\Big(\Big|\frac{T_n - \mathbb{E}[T_n]}{n\ln(n)}\Big| \geq \frac{\epsilon}{2}\Big)\,,$$

et on a montré que cette dernière probabilité tendait vers 0. Ainsi,

$$\frac{T_n}{n\ln(n)} \xrightarrow[n\to\infty]{\mathbb{P}} 1\,.$$

Cela signifie que pour n assez grand, la collection est complète au bout de $n\ln(n)$ achats.

Intermède : convergence presque sûre

Cet exercice particulièrement long est à traiter uniquement si les notions de convergence de variables aléatoires sont bien maîtrisées. Il ne sera pas corrigé en TD. Il faut le voir comme un devoir maison (qui n'est pas à rendre) sur lequel on passe un peu de temps et qui permet de prendre du recul sur les concepts étudiés en CM et TD. Il permet aussi de se rendre compte qu'établir la convergence presque sûre est en général compliqué.

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires iid de loi exponentielle de paramètre 1. On pose $M_n = \max_{1\leq k\leq n} X_k$.

1. Déterminer la fonction de répartition de M_n.

   Les X_n étant iid, on obtient
   $$F_{M_n}(x) = \mathbb{P}(M_n \leq x) = \mathbb{P}(X_1 \leq x, \ldots, X_n \leq x) = \mathbb{P}(X_1 \leq x)^n = (1 - \mathrm{e}^{-x})^n\,,$$
   pour $x > 0$. Par ailleurs, $\mathbb{P}(M_n \leq x) = 0$ si $x < 0$.

2. En déduire que $M_n/\ln(n)$ converge vers 1 en probabilité.

   Soit $\epsilon > 0$. Il s'agit de montrer que
   $$\mathbb{P}\Big(\Big|\frac{M_n}{\ln(n)} - 1\Big| \geq \epsilon\Big) \xrightarrow[n\to\infty]{} 0\,.$$
   Cette probabilité se décompose en
   $$\begin{aligned}\mathbb{P}\Big(\Big|\frac{M_n}{\ln(n)} - 1\Big| \geq \epsilon\Big) &= 1 - \mathbb{P}\Big(\frac{M_n}{\ln(n)} - 1 < \epsilon\Big) + \mathbb{P}\Big(\frac{M_n}{\ln(n)} - 1 < -\epsilon\Big)\\ &= 1 - [1 - \exp(-(1+\epsilon)\ln(n))]^n + [1 - \exp(-(1-\epsilon)\ln(n))]^n\\ &= 1 - \Big(1 - \frac{1}{n^{1+\epsilon}}\Big)^n + \Big(1 - \frac{1}{n^{1-\epsilon}}\Big)^n \mathbb{1}_{\{\epsilon<1\}}\,.\end{aligned}$$
   Or,
   $$\Big(1 - \frac{1}{n^{1+\epsilon}}\Big)^n \sim \mathrm{e}^{-n^{-\epsilon}} \to 1 \quad \text{et} \quad \Big(1 - \frac{1}{n^{1-\epsilon}}\Big)^n \sim \mathrm{e}^{-n^{\epsilon}} \to 0\,,$$
   où la dernière équivalence est vraie pour $\epsilon < 1$, ce que l'on suppose.

   Ainsi, on a montré que
   $$\mathbb{P}\Big(\Big|\frac{M_n}{\ln(n)} - 1\Big| \geq \epsilon\Big) \xrightarrow[n\to\infty]{} 0\,,$$
   ce qui prouve la convergence en probabilité de $M_n/\ln(n)$ vers 1.

3. On veut montrer que la convergence a en fait lieu presque sûrement. Pour cela on se fixe $\epsilon \in]0,1[$ et on va montrer que presque sûrement
   $$\liminf_{n\to\infty} \frac{M_n}{\ln(n)} \geq 1 - \epsilon \quad \text{et} \quad \limsup_{n\to\infty} \frac{M_n}{\ln(n)} \leq 1 + \epsilon\,.$$
   Attention à bien comprendre la première inégalité au sens de "la limite inférieure de la suite de variables aléatoires $\frac{M_n}{\ln(n)}$ est plus grande que $1-\epsilon$" (et idem pour la deuxième inégalité). En particulier, il n'est pas question de limite inférieure d'événements, même si de tels objets vont apparaître au cours de l'exercice.

(a) Montrer que $\mathbb{P}(M_n \leq (1-\epsilon)\ln(n)) = \exp(-n^\epsilon + o(n^\epsilon))$.
Une calcul similaire à la question précédente donne

$$\mathbb{P}(M_n \leq (1-\epsilon)\ln(n)) = \Big(1 - \frac{1}{n^{1-\epsilon}}\Big)^n = \exp\Big(n \ln\big(1 - \frac{1}{n^{1-\epsilon}}\big)\Big) = \exp(-n^\epsilon + o(n^\epsilon))\,.$$

(b) En utilisant le lemme de Borel-Cantelli, en déduire que

$$\mathbb{P}\Big(\liminf_{n\to\infty} \frac{M_n}{\ln(n)} \geq 1-\epsilon\Big) = 1\,.$$

D'après la question précédente la série $\sum \mathbb{P}(M_n \leq (1-\epsilon)\ln(n))$ est convergente. Le lemme de Borel-Cantelli implique alors que

$$\mathbb{P}(\limsup_{n\to\infty}\{M_n \leq (1-\epsilon)\ln(n)\}) = 0\,.$$

En passant au complémentaire et en utilisant la relation $(\limsup A_n)^c = \liminf A_n^c$, on obtient alors

$$\mathbb{P}(\liminf_{n\to\infty}\{M_n > (1-\epsilon)\ln(n)\}) = 1\,.$$

Il s'agit alors de passer d'une limite inférieure d'événements à une limite inférieure de variables aléatoires. Par définition de la limite inférieure d'événements, l'égalité précédente se réécrit

$$\mathbb{P}\big(\exists n \geq 1,\, \forall k \geq n,\, M_k > (1-\epsilon)\ln(k)\big) = 1\,.$$

Or,

$$\begin{aligned}\mathbb{P}\big(\exists n \geq 1,\, \forall k \geq n,\, M_k > (1-\epsilon)\ln(k)\big) &\leq \mathbb{P}\Big(\exists n \geq 1,\, \inf_{k\geq n}\frac{M_k}{\ln(k)} \geq (1-\epsilon)\Big)\\ &\leq \mathbb{P}\Big(\liminf_{n\to\infty}\frac{M_n}{\ln(n)} \geq (1-\epsilon)\Big)\,,\end{aligned}$$

où la deuxième provient du fait que la suite des inf est croissante. On obtient alors

$$\mathbb{P}\Big(\liminf_{n\to\infty}\frac{M_n}{\ln(n)} > (1-\epsilon)\Big) = 1\,.$$

(c) Montrer que $\mathbb{P}(M_n > (1+\epsilon)\ln(n)) = n^{-\epsilon} + o(n^{-\epsilon})$. Est-ce suffisant pour conclure ?
On utilise la question (a) qui donne

$$\mathbb{P}(M_n > (1+\epsilon)\ln(n)) = 1 - \mathbb{P}(M_n \leq (1+\epsilon)\ln(n)) = 1 - \mathrm{e}^{-n^{-\epsilon}+o(n^{-\epsilon})}\,,$$

ce qui conduit bien à $\mathbb{P}(M_n > (1+\epsilon)\ln(n)) = n^{-\epsilon} + o(n^{-\epsilon})$.
Cependant, on ne peut pas appliquer le lemme de Borel-Cantelli directement car on ne connaît pas le comportement de la série de terme général $n^{-\epsilon} + o(n^{-\epsilon})$.

(d) Montrer que $\mathbb{P}(M_{2^k} > (1+\epsilon)\ln(2^k)) = 2^{-k\epsilon} + o(2^{-k\epsilon})$. En déduire que

$$\mathbb{P}\Big(\limsup_{k\to\infty} \frac{M_{2^k}}{\ln(2^k)} \leq (1+\epsilon)\Big) = 1\,.$$

On a donc l'inégalité voulue pour la limite supérieure mais uniquement le long d'une sous-suite de la forme 2^k.

L'égalité demandée n'est rien d'autre que celle de la question précédente avec la sous-suite 2^k. Cela implique que la série de terme général $\mathbb{P}(M_{2^k} > (1+\epsilon)\ln(2^k))$ est convergente et le lemme de Borel-Cantelli assure alors que

$$\mathbb{P}\Big(\limsup_{n\to\infty}\{M_{2^k} > (1+\epsilon)\ln(2^k)\}\Big) = 0\,.$$

On obtient alors

$$\begin{aligned}
\mathbb{P}\Big(\limsup_{k\to\infty} \frac{M_{2^k}}{\ln(2^k)} > (1+\epsilon)\Big) &= \mathbb{P}\Big(\lim_{n\to\infty}\sup_{k\geq n} M_{2^k} > (1+\epsilon)\ln(2^k)\Big) \\
&= \mathbb{P}(\forall n\geq 1,\, \exists k\geq n,\, M_{2^k} > (1+\epsilon)\ln(2^k)) \\
&= \mathbb{P}\Big(\bigcap_{n\geq 1}\bigcup_{k\geq n}\{M_{2^k} > (1+\epsilon)\ln(2^k)\}\Big) \\
&= \mathbb{P}\Big(\limsup_{n\to\infty}\{M_{2^k} > (1+\epsilon)\ln(2^k)\}\Big) \\
&= 0\,.
\end{aligned}$$

Par passage au complémentaire on obtient le résultat voulu :

$$\mathbb{P}\Big(\limsup_{k\to\infty} \frac{M_{2^k}}{\ln(2^k)} \leq (1+\epsilon)\Big) = 1\,.$$

(e) En déduire que

$$\mathbb{P}\Big(\limsup_{n\to\infty} \frac{M_n}{\ln(n)} \leq (1+\epsilon)\Big) = 1\,.$$

Pour $n\geq 1$, on choisit $k = k(n)$ tel que $2^{k-1} \leq n \leq 2^k$. Alors, la suite (M_n) étant croissante, on obtient l'inégalité

$$\frac{M_n}{\ln(n)} \leq \frac{M_{2^k}}{\ln(2^k)}\frac{\ln(2^k)}{\ln(2^{k-1})}\,,$$

puis, en passant à la limite supérieure (qui conserve les inégalités) :

$$\limsup_{n\to\infty}\frac{M_n}{\ln(n)} \leq \limsup_{k\to\infty}\frac{M_{2^k}}{\ln(2^k)}\frac{\ln(2^k)}{\ln(2^{k-1})} = \limsup_{k\to\infty}\frac{M_{2^k}}{\ln(2^k)}\,.$$

On conclut alors avec la question précédente :

$$\mathbb{P}\Big(\limsup_{n\to\infty} \frac{M_n}{\ln(n)} \leq (1+\epsilon)\Big) \geq \mathbb{P}\Big(\limsup_{k\to\infty} \frac{M_{2^k}}{\ln(2^k)} \leq (1+\epsilon)\Big) = 1\,.$$

(f) Conclure que $M_n/\ln(n) \to 1$ presque sûrement.

Dans les question précédentes, on a montré que les événements

$$\Big\{\liminf_{n\to\infty} \frac{M_n}{\ln(n)} \geq 1-\epsilon\Big\} \quad \text{et} \quad \Big\{\limsup_{n\to\infty} \frac{M_n}{\ln(n)} \leq 1+\epsilon\Big\}.$$

sont presque sûrs.

Rappelons que l'intersection d'un nombre *dénombrable* d'événements presque sûrs est encore de presque sûr. En effet, cela revient à dire que l'union d'événements de probabilité zéro est encore de probabilité zéro ce qui résulte de l'inégalité

$$\mathbb{P}\Big(\bigcup_{n=0}^{\infty} A_n) \leq \sum_{n=0}^{\infty} \mathbb{P}(A_n)\,.$$

Ainsi, l'intersection des deux événements précédents est de probabilité 1 :

$$\mathbb{P}\Big((1-\epsilon) \leq \liminf_{n\to\infty} \frac{M_n}{\ln(n)} \leq \limsup_{n\to\infty} \frac{M_n}{\ln(n)} \leq (1+\epsilon)\Big) = 1\,.$$

Ceci est vrai pour tout $\epsilon \in]0,1[$. En particulier, si on prend une suite ϵ de rationnels entre $]0,1[$ (pour avoir un nombre dénombrable d'événements), on en déduit alors que

$$\mathbb{P}\Big(\bigcap_{\epsilon\in\mathbb{Q}\cap]0,1[} \Big\{(1-\epsilon) \leq \liminf_{n\to\infty} \frac{M_n}{\ln(n)} \leq \limsup_{n\to\infty} \frac{M_n}{\ln(n)} \leq (1+\epsilon)\Big\}\Big) = 1\,.$$

Or l'intersection qui apparaît dans la probabilité précédente n'est rien d'autre que l'événement

$$\Big\{\liminf_{n\to\infty} \frac{M_n}{\ln(n)} = \limsup_{n\to\infty} \frac{M_n}{\ln(n)} = 1\Big\} = \Big\{\lim_{n\to\infty} \frac{M_n}{\ln(n)} = 1\Big\}.$$

On a donc montré que l'événement $\Big\{\lim_{n\to\infty} \dfrac{M_n}{\ln(n)} = 1\Big\}$ est de probabilité 1, ce qui signifie que $M_n/\ln(n) \xrightarrow[n\to\infty]{\text{p.s.}} 1$.

TD 9 - Loi des grands nombres

Exercice 1 Soit $(X_n)_{n\geq 1}$ une suite de variables iid de loi $\mathcal{N}(0,1)$. Montrer que la suite de terme général

$$Y_n = \frac{1}{n}\sum_{k=1}^{n} X_k \,\mathrm{e}^{X_k}$$

converge presque sûrement lorsque n tend vers l'infini vers une limite que l'on précisera.

Les variables aléatoires $X_k \,\mathrm{e}^{X_k}$ sont indépendantes et de même loi puisque les X_k le sont. De plus, la fonction

$$x \mapsto x\,\mathrm{e}^{x}\,\frac{\mathrm{e}^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$$

est intégrable en $\pm\infty$, ce qui assure l'existence de l'espérance de $X_1\,\mathrm{e}^{X_1}$. Un calcul d'intégrale donne alors

$$\mathbb{E}[X_1\,\mathrm{e}^{X_1}] = \int_{\mathbb{R}} x\,\mathrm{e}^{x}\,\frac{\mathrm{e}^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}x = \mathrm{e}^{1/2}\int_{\mathbb{R}} x\frac{\mathrm{e}^{-\frac{(x-1)^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}x = \sqrt{\mathrm{e}}\int_{\mathbb{R}} (t+1)\frac{\mathrm{e}^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}t\,,$$

où la dernière égalité résulte du changement de variable $t = x-1$. On en déduit que

$$\mathbb{E}[X_1\,\mathrm{e}^{X_1}] = \sqrt{\mathrm{e}}\int_{\mathbb{R}} t\frac{\mathrm{e}^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}t + \sqrt{\mathrm{e}}\int_{\mathbb{R}} \frac{\mathrm{e}^{-\frac{t^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}t = \sqrt{\mathrm{e}}\mathbb{E}[X_1] + \sqrt{\mathrm{e}} = \sqrt{\mathrm{e}}\,.$$

On peut alors appliquer la loi forte des grands nombres aux variables aléatoires $X_k\,\mathrm{e}^{X_k}$:

$$\frac{1}{n}\sum_{k=1}^{n} X_k\,\mathrm{e}^{X_k} \xrightarrow[n\to\infty]{p.s.} \sqrt{\mathrm{e}}\,.$$

Exercice 2 On considère une suite de lancers d'un dé équilibré et on désigne par X_k le résultat du k-ème lancer.

1. On note Y_n la variable aléatoire donnant le plus grand résultat observé au cours des n premiers lancers. Étudier la convergence de la suite $(Y_n)_{n\geq 1}$.

   On écrit $Y_n = \max_{k=1,\dots,n} X_k$. Remarquons tout d'abord que s'il existe $k \geq 1$ tel que $X_k = 6$, alors $Y_k = 6$ et la suite $(Y_n)_{n\geq 1}$ est alors constante égale à 6 à partir du rang k, donc converge presque sûrement vers 6. Ceci entraîne l'inclusion d'événements

   $$\{\exists k \geq 1 : X_k = 6\} \subset \left\{\lim_{n\to\infty} Y_n = 6\right\}$$

Il suffit donc de montrer que l'événement $A = \{\exists k \geq 1 : X_k = 6\}$ est de probabilité 1. On considère les événements $A_n = \{X_1 \neq 6, \ldots, X_n \neq 6\}$ pour $n \geq 1$. Par indépendance des X_k, on a

$$\mathbb{P}(A_n) = \mathbb{P}(X_1 \neq 6) \cdots \mathbb{P}(X_n \neq 6) = \left(\frac{5}{6}\right)^n .$$

Par ailleurs, la suite $(A_n)_{n\geq 1}$ est formée d'événements décroissants pour l'inclusion ce qui implique que

$$\mathbb{P}\Big(\bigcap_{n\geq 1} A_n\Big) = \lim_{n\to\infty} \mathbb{P}(A_n) = 0 .$$

On conclut alors en écrivant

$$\mathbb{P}(A) = 1 - \mathbb{P}\Big(\bigcap_{n\geq 1} A_n\Big) \xrightarrow[n\to\infty]{} 1 .$$

On a donc montré que la suite (Y_n) converge presque sûrement vers 6.

2. On note N_n la variable aléatoire donnant le nombre de 6 obtenus lors des n premiers lancers. Étudier la convergence de la suite $(N_n/n)_{n\geq 1}$.

   Il suffit de remarquer que

   $$N_n = \sum_{k=1}^{n} \mathbb{1}_{\{X_k=6\}} .$$

   Les variables aléatoires $\mathbb{1}_{\{X_k=6\}}$ sont indépendantes car les X_k le sont et de loi de Bernoulli de paramètre $\mathbb{P}(X_1 = 6) = \frac{1}{6}$. Leur espérance est donc finie égale à $1/6$. La loi des grands nombres assure alors la convergence presque sûre

   $$\frac{N_n}{n} \xrightarrow[n\to\infty]{p.s.} \mathbb{E}[\mathbb{1}_{\{X_1=6\}}] = \frac{1}{6} .$$

   C'est assez cohérent : si on répète un grand nombre de fois l'expérience, la fréquence d'apparition de 6 se rapproche de $\frac{1}{6}$.

Exercice 3 On suppose que le sexe d'un nouveau-né est équiréparti entre fille et garçon. Un pays propose la politique de natalité suivante : tous les couples ont des enfants jusqu'à obtenir une fille.

1. Soit X le nombre d'enfants d'un couple pris au hasard dans la population. Donner la loi de la variable aléatoire X et son espérance.

   L'expérience aléatoire consiste à répéter une épreuve de Bernoulli (ici, avoir un enfant) de paramètre $p = 1/2$ jusqu'à arriver au premier "succès" (avoir une fille). Ainsi, X suit une loi géométrique de paramètre $p = 1/2$. Cela signifie que X est à valeurs dans $\mathbb{N}^*$ avec $p_k = \mathbb{P}(X = k) = 2^{-k}$. L'espérance de X est donnée par $\mathbb{E}[X] = \frac{1}{p} = 2$.

2. On considère une génération en âge de procréer constituée de n couples et on note $X_1, \ldots, X_n$ le nombre d'enfants respectifs de chaque couple. On désigne par P_n la variable aléatoire donnant la proportion de filles issus de cette génération. Exprimer P_n en fonction de $X_1, \ldots, X_n$, puis déterminer la limite de P_n quand n tend vers l'infini. La politique de natalité du pays a-t-elle un effet ?

Le nombre total d'enfants est $X_1 + \cdots + X_n$ et chaque couple a un seul fille donc il y a n filles en tout. La proportion de filles est donc

$$P_n = \frac{n}{X_1 + \cdots + X_n}.$$

Les variables aléatoires X_k sont iid de loi géométrique donc admettent une espérance égale à 2. On peut donc appliquer la loi des grands nombres qui donne

$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow[n\to\infty]{p.s.} \mathbb{E}[X_1] = 2\,.$$

En passant à l'inverse on obtient

$$P_n = \frac{n}{X_1 + \cdots + X_n} \xrightarrow[n\to\infty]{p.s.} \frac{1}{2}\,.$$

La politique de natalité ne change rien : quand la taille de la population est grande, la proportion de filles est d'environ $1/2$, soit la même qu'en l'absence de politique de natalité.

Pour aller plus loin

On considère une suite iid $(X_n)_{n\geq 1}$ de variables aléatoires réelles et on pose $S_n = X_1 + \cdots + X_n$.

1. En utilisant le théorème de Fubini, montrer que pour toute variable aléatoire positive Y

$$\mathbb{E}[Y] = \int_0^\infty \mathbb{P}(Y \geq t)\,\mathrm{d}t\,,$$

puis en déduire que

$$\mathbb{E}[|X_1|] \leq 1 + \sum_{n=1}^{\infty} \mathbb{P}(|X_n| \geq n)\,.$$

On écrit $\mathbb{P}(Y \geq t) = \mathbb{E}[\mathbb{1}_{Y\geq t}] = \mathbb{E}[\mathbb{1}_{[t,\infty[}(Y)]$ et le théorème de Fubini assure que

$$\int_0^\infty \mathbb{P}(Y \geq t)\,\mathrm{d}t = \int_0^\infty \mathbb{E}[\mathbb{1}_{[t,\infty[}(Y)] = \mathbb{E}\Big[\int_0^\infty \mathbb{1}_{[t,\infty[}(Y)\,\mathrm{d}t\Big]\,.$$

Cette intégrale se réécrit

$$\int_0^\infty \mathbb{1}_{[t,\infty[}(Y)\,\mathrm{d}t = \int_0^\infty \mathbb{1}_{[0,Y]}(t)\,\mathrm{d}t = \int_0^Y \mathrm{d}t = Y\,,$$

ce qui donne le résultat.

On utilise alors ce résultat avec $Y = |X_1|$ qui est une variable aléatoire positive et on obtient

$$\mathbb{E}[|X_1|] = \int_0^\infty \mathbb{P}(|X_1| \geq t)\,\mathrm{d}t\,.$$

Il suffit alors de décomposer l'intégrale sous la forme

$$\int_0^\infty \mathbb{P}(|X_1| \geq t)\,\mathrm{d}t = \int_0^\infty \sum_{n=0}^\infty \mathbb{P}(|X_1| \geq t)\mathbb{1}_{[n,n+1[}(t)\,\mathrm{d}t \leq \int_0^\infty \sum_{n=0}^\infty \mathbb{P}(|X_1| \geq n)\,\mathrm{d}t$$

puisque si $t \in [n, n+1[$, alors $\mathbb{P}(|X_1| \geq t) \leq \mathbb{P}(|X_1| \geq n)$. Le théorème de convergence monotone donne alors

$$\int_0^\infty \mathbb{P}(|X_1| \geq t)\,\mathrm{d}t \leq \sum_{n=0}^\infty \left(\mathbb{P}(|X_1| \geq n) \int_0^\infty \mathbb{1}_{[n,n+1[}(t)\,\mathrm{d}t\right) = \sum_{n=0}^\infty \mathbb{P}(|X_1| \geq n)\,.$$

Enfin, on majore le premier terme de cette somme par 1 et on utilise l'égalité $\mathbb{P}(|X_1| \geq n) = \mathbb{P}(|X_n| \geq n)$ qui résulte du fait que les X_k suivent la même loi. On obtient alors

$$\sum_{n=0}^\infty \mathbb{P}(|X_1| \geq n) \leq 1 + \sum_{n=1}^\infty \mathbb{P}(|X_n| \geq n)\,.$$

Bref, on a bien montré que

$$\mathbb{E}[|X_1|] \leq 1 + \sum_{n=1}^\infty \mathbb{P}(|X_n| \geq n)\,.$$

2. En déduire que si X_1 n'est pas intégrable, alors la suite $(S_n/n)_{n\geq 1}$ diverge presque sûrement.
   Notons $A_n = \{|X_n| \geq n\}$ qui sont des événements indépendants. Si X_1 n'est pas intégrable, alors la question précédente assure que la série $\sum_{n\geq 1} \mathbb{P}(A_n)$ diverge donc d'après le lemme de Borel-Cantelli,

$$\mathbb{P}(\limsup_{n\to\infty} A_n) = 1\,.$$

Maintenant remarquons que si S_n/n converge, alors

$$\frac{S_n}{n} - \frac{S_{n+1}}{n+1} \xrightarrow[n\to\infty]{} 0 \quad \text{et} \quad \frac{S_n}{n(n+1)} \xrightarrow[n\to\infty]{} 0\,.$$

On écrit alors

$$\frac{S_n}{n} - \frac{S_{n+1}}{n+1} = \frac{(n+1)S_n - nS_{n+1}}{n(n+1)} = \frac{S_n}{n(n+1)} - \frac{|X_{n+1}|}{n+1}\,.$$

et on en déduit que la suite $(\frac{|X_{n+1}|}{n+1})$ converge vers 0. Ceci implique alors que l'événement $\{|X_n| \geq n\}$ est réalisé un nombre fini de fois. Bref, on vient de montrer les inclusions

$$\{S_n/n \text{ converge}\} \subset \{|X_n|/n \to 0\} \subset (\limsup_{n\to\infty} A_n)^c$$

On conclut alors en écrivant

$$\mathbb{P}(S_n/n \text{ converge}) \leq \mathbb{P}((\limsup_{n\to\infty} A_n)^c) = 1 - \mathbb{P}(\limsup_{n\to\infty} A_n) = 0\,.$$

Donc la S_n/n diverge presque sûrement.

TD 10 - Convergence en loi

Exercice 1

1. Rappeler la définition de la convergence en loi, puis sa caractérisation via les fonctions caractéristiques et les fonctions de répartition (dans le cas de variables aléatoires réelles).

   Soit $X, X_1, \dots, X_n, \dots$ des variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs dans $(E, \mathcal{E})$. On dit que la suite $(X_n)_{n\geq 1}$ converge en loi vers X si pour toute fonction $h : E \mapsto \mathbb{R}$ continue bornée,

   $$\mathbb{E}[h(X_n)] \to \mathbb{E}[h(X)]\,, \quad n \to \infty\,.$$

   Cette convergence est équivalente à :

   (a) $\varphi_{X_n}(t) \to \varphi_X(t)$, pour tout $t \in \mathbb{R}$, où φ_{X_k} et φ_X désignent respectivement la fonction caractéristique de X_k et de X.

   (b) Dans le cas réel : $F_{X_n}(t) \to F_X(t)$ pour tout point de continuité $t \in \mathbb{R}$ de F_X, où F_{X_k} et F_X désignent respectivement la fonction de répartition de X_k et de X.

2. On suppose que $X_n \xrightarrow[n\to\infty]{\mathcal{L}} X$, avec (X_n) et X des variables aléatoires réelles. Si $f : \mathbb{R} \to \mathbb{R}$ est une fonction continue, a-t-on $f(X_n) \xrightarrow[n\to\infty]{\mathcal{L}} f(X)$?

   Soit $h : \mathbb{R} \to \mathbb{R}$ une fonction continue bornée. Alors $h \circ f$ est continue et bornée. Comme (X_n) converge en loi vers X on a donc

   $$\mathbb{E}[h(f(X_n))] \xrightarrow[n\to\infty]{} \mathbb{E}[h(f(X))]\,.$$

   Ceci étant vrai pour toute fonction h continue bornée on en déduit que $(f(X_n))$ converge en loi vers $f(X)$.

3. Donner un exemple de variables aléatoires telles que $X_n \xrightarrow[n\to\infty]{\mathcal{L}} X$ mais $\mathbb{E}[X_n]$ ne converge pas vers $\mathbb{E}[X]$.

   Soit (X_n) une suite de variables aléatoires discrètes telle que

   $$\mathbb{P}(X_n = n) = \frac{1}{n} = 1 - \mathbb{P}(X_n = 0)\,,$$

   c'est-à-dire $\mathbb{P}_{X_n} = (1 - \frac{1}{n})\delta_{\{0\}} + \frac{1}{n}\delta_{\{n\}}$. Alors pour tout $h : \mathbb{R} \to \mathbb{R}$ continue bornée,

   $$\mathbb{E}[h(X_n)] = \Big(1 - \frac{1}{n}\Big)h(0) + \frac{1}{n}h(n)\,.$$

   Comme h est bornée, disons par une constante M, on a

   $$\Big|\frac{1}{n}h(n)\Big| \leq \frac{M}{n} \xrightarrow[n\to\infty]{} 0\,.$$

   Donc $\mathbb{E}[h(X_n)] \to \mathbb{E}[h(0)]$. Ce qui signifie que X_n converge en loi vers la variable aléatoire X constante égale à 0 (ou encore que la loi de X_n converge en loi vers une masse de Dirac en 0).

   Cependant, $\mathbb{E}[X_n] = n \times \frac{1}{n} = 1$ et $\mathbb{E}[X] = 0$. Donc $\mathbb{E}[X_n]$ ne converge pas vers $\mathbb{E}[X]$.

Exercice 2 Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires telle que pour tout $n \geq 1$, X_n suit une loi uniforme sur l'ensemble $\{k/n,\ 1 \leq k \leq n\}$. Montrer de trois manières différentes que X_n converge en loi vers une loi uniforme sur $[0,1]$.

Rappelons tout d'abord que pour tout $n \geq 1$, la loi de X_n est donnée par

$$p_{n,k} = \mathbb{P}(X_n = k/n) = \frac{1}{n}, \quad k = 1, \dots, n.$$

Avec la définition. Soit $h : \mathbb{R} \mapsto \mathbb{R}$ une fonction continue bornée. On a

$$\mathbb{E}[h(X_n)] = \sum_{k=1}^{n} h(k/n) p_{n,k} = \frac{1}{n} \sum_{k=1}^{n} h(k/n) \to \int_0^1 h(t)\,\mathrm{d}t, \quad n \to \infty.$$

d'après la propriété de convergence des sommes de Riemann. Soit X de loi uniforme sur $[0,1]$. On vient donc de montrer que

$$\mathbb{E}[h(X_n)] \to \mathbb{E}[h(X)], \quad n \to \infty.$$

Ainsi, $X_n \overset{\mathcal{L}}{\to} X$, ou encore $X_n \overset{\mathcal{L}}{\to} \mathcal{U}([0,1])$.

Avec la fonction caractéristique. La fonction caractéristique de X_n est donnée par

$$\varphi_{X_n}(t) = \mathbb{E}[\mathrm{e}^{itX_n}] = \sum_{k=1}^{n} \mathrm{e}^{itk/n} p_{n,k} = \frac{1}{n} \sum_{k=1}^{n} (\mathrm{e}^{it/n})^k = \frac{1}{n} \frac{1-\mathrm{e}^{it}}{1-\mathrm{e}^{it/n}} \mathrm{e}^{it/n}.$$

Le dénominateur vérifie $n(1-\mathrm{e}^{it/n}) \sim -it$ quand $n \to \infty$. On en déduit que pour tout $t \in \mathbb{R}$,

$$\varphi_{X_n}(t) \to \frac{\mathrm{e}^{it} - 1}{it}.$$

Or, la fonction caractéristique d'une variable aléatoire $X \sim \mathcal{U}([0,1])$ vaut

$$\varphi_X(t) = \mathbb{E}[\mathrm{e}^{itX}] = \int_0^1 \mathrm{e}^{itu}\,\mathrm{d}u = \left[\frac{\mathrm{e}^{itu}}{it}\right]_0^1 = \frac{\mathrm{e}^{it} - 1}{it}.$$

Ainsi, on a montré que pour tout $t \in \mathbb{R}$, $\varphi_{X_n}(t) \to \varphi_X(t)$ quand $n \to \infty$. Ceci implique la convergence en loi $X_n \overset{\mathcal{L}}{\to} X$.

Avec la fonction de répartition. Remarquons tout d'abord que pour tout $n \geq 1$, la fonction de répartition F_{X_n} de X_n est nulle avant 0 et vaut 1 après 1. Par ailleurs, si $t \in [0,1]$, on a

$$F_{X_n}(t) = \mathbb{P}(X_n \leq t) = \sum_{k=1}^{n} p_{n,k} \mathbb{1}_{[\frac{k}{n},\infty[}(t) = \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{[\frac{k}{n},\infty[}(t) = \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{[0,tn]}(k).$$

La somme compte le nombre d'entiers non nuls plus petits que tn, et vaut donc $\lfloor tn \rfloor$. On en déduit que

$$F_{X_n}(t) = \frac{\lfloor tn \rfloor}{n} \xrightarrow[n\to\infty]{} t.$$

On a montré que $F_{X_n}(t) \to t\mathbb{1}_{[0,1]} + \mathbb{1}_{]1,\infty[} = F_X(t)$ lorsque $n \to \infty$ ce qui prouve la convergence en loi $X_n \overset{\mathcal{L}}{\to} X$.

Exercice 3 Pour tout entier n on considère la variable aléatoire T_n de loi géométrique de paramètre λ/n, avec $\lambda > 0$ fixé. Montrer que T_n/n converge en loi vers une limite à déterminer.

On va établir la convergence en loi en utilisant la fonction caractéristique de T_n/n définie pour tout réel t par

$$\mathbb{E}[\mathrm{e}^{itT_n/n}] = \sum_{k=1}^{\infty} \mathrm{e}^{ikt/n} \frac{\lambda}{n}\Big(1-\frac{\lambda}{n}\Big)^{k-1} = \frac{\lambda}{n}\,\mathrm{e}^{it/n} \sum_{k=0}^{\infty} \Big[\mathrm{e}^{it/n}\Big(1-\frac{\lambda}{n}\Big)\Big]^k = \frac{\lambda\,\mathrm{e}^{it/n}}{n\big[1-\mathrm{e}^{it/n}\big(1-\frac{\lambda}{n}\big)\big]}\,.$$

Quand $n \to \infty$, le numérateur tend vers λ. Pour le dénominateur on part du développement limité

$$\mathrm{e}^{it/n}\Big(1-\frac{\lambda}{n}\Big) = \Big(1+\frac{it}{n}+o(1/n)\Big)\Big(1-\frac{\lambda}{n}\Big) = 1+\frac{it}{n}-\frac{\lambda}{n}+o(1/n)\,,$$

ce qui donne

$$\mathbb{E}[\mathrm{e}^{itT_n/n}] \sim \frac{\lambda}{\lambda - it}\,.$$

On reconnaît la fonction caractéristique de la loi exponentielle de paramètre λ. Redémontrons ce résultat. Soit $X \sim \mathcal{E}(\lambda)$. Pour tout réel t,

$$\mathbb{E}[\mathrm{e}^{itX}] = \int_0^{\infty} \mathrm{e}^{itx}\,\lambda\,\mathrm{e}^{-\lambda x}\,\mathrm{d}x = \int_0^{\infty} \lambda\,\mathrm{e}^{(it-\lambda)x}\,\mathrm{d}x = \Big[\frac{\lambda}{it-\lambda}\,\mathrm{e}^{(it-\lambda)x}\Big]_0^{\infty} = \frac{\lambda}{\lambda-it}$$

ce qui donne le résultat.

Bref, on a montré que $T_n/n \xrightarrow{\mathcal{L}} \mathcal{E}(\lambda)$ quand $n \to \infty$.

Exercice 4 Soit $(U_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes de loi uniforme sur $[0,1]$. On pose $M_n = \max(U_1,\dots,U_n)$ et $X_n = n(1-M_n)$.

1. Déterminer la fonction de répartition de M_n, puis celle de X_n.
   Comme les U_k sont à valeurs dans $[0,1]$, la variable aléatoire M_n aussi et donc la fonction de répartition de M_n est nulle avant 0 et vaut 1 après 1. Enfin, pour $t \in [0,1]$ on obtient

   $$\mathbb{P}(M_n \leq t) = \mathbb{P}(U_1 \leq t,\cdots,U_n \leq t) = \mathbb{P}(U_1 \leq t)\cdots\mathbb{P}(U_n \leq t) = t^n\,,$$

   où on a utilisé l'indépendance des U_k et le fait que les U_k suivent une loi uniforme sur $[0,1]$. La variable aléatoire $X_n = n(1-M_n)$ est à valeurs dans $[0,n]$ donc sa fonction de répartition est nulle avant 0 et vaut 1 après n. Pour $t \in [0,n]$, on a

   $$\mathbb{P}(X_n \leq t) = \mathbb{P}(n(1-M_n) \leq t) = \mathbb{P}(M_n \geq 1-t/n) = 1-\Big(1-\frac{t}{n}\Big)^n\,.$$

   Ainsi,

   $$\mathbb{P}(X_n \leq t) = \Big[1-\Big(1-\frac{t}{n}\Big)^n\Big]\mathbb{1}_{[0,n]}(t) + \mathbb{1}_{]n,\infty[}(t)\,.$$

2. Étudier la convergence en loi de la suite $(X_n)_{n\geq 1}$.
   La fonction de répartition de X_n vérifie pour tout $t \in \mathbb{R}$,

   $$\mathbb{P}(X_n \leq t) = \Big[1-\Big(1-\frac{t}{n}\Big)^n\Big]\mathbb{1}_{[0,n]}(t) + \mathbb{1}_{]n,\infty[}(t) \to (1-\mathrm{e}^{-x})\mathbb{1}_{[0,\infty[}(t)\,, \quad n \to \infty\,.$$

   Ceci prouve que X_n converge en loi vers une loi exponentielle de paramètre 1.

Pour aller plus loin

On considère une suite $(X_n)_{n\geq 1}$ de variables aléatoires réelles telle que pour tout $n \in \mathbb{N}$, X_n suit une loi exponentielle de paramètre $\lambda_n > 0$ avec $\lambda_n \to 0$. On pose $Z_n = X_n - \lfloor X_n \rfloor$. Étudier la convergence en loi de la suite $(Z_n)_{n\geq 1}$.

On va montrer que $(Z_n)_{n\geq 1}$ converge en loi vers une loi uniforme sur $[0,1]$ en utilisant les fonctions de répartition.

Notons F_{Z_n} la fonction de répartition de Z_n. On remarque que comme $0 \leq X_n \leq \lfloor X_n \rfloor$, la variable aléatoire Z_n prend ses valeurs dans $[0,1]$. Ainsi F_{Z_n} est nulle avant 0 et vaut 1 après 1. Enfin pour $t \in [0,1]$, on utilise la formule des probabilités totales avec la partition $\{\lfloor X_n \rfloor = k\}$ pour $k \in \mathbb{N}$, ce qui donne

$$\begin{aligned} F_{Z_n}(t) = \mathbb{P}(Z_n \leq t) &= \sum_{k=0}^{\infty} \mathbb{P}(Z_n \leq t, \lfloor X_n \rfloor = k) \\ &= \sum_{k=0}^{\infty} \mathbb{P}(X_n - \lfloor X_n \rfloor \leq t, k \leq X_n < k+1) \\ &= \sum_{k=0}^{\infty} \mathbb{P}(k \leq X_n \leq k+t)\,. \end{aligned}$$

Puis, comme X_n suit une loi exponentielle de paramètre $\lambda_n > 0$, on obtient

$$F_{Z_n}(t) = \sum_{k=0}^{\infty} \int_k^{k+t} \lambda_n \, \mathrm{e}^{-\lambda_n x} \, \mathrm{d}x = \sum_{k=0}^{\infty} (\mathrm{e}^{-\lambda_n k} - \mathrm{e}^{-\lambda_n (k+t)}) = (1 - \mathrm{e}^{-\lambda_n t}) \sum_{k=0}^{\infty} \mathrm{e}^{-\lambda_n k}\,.$$

On reconnaît une série géométrique de paramètre $\mathrm{e}^{-\lambda_n}$. Ainsi,

$$F_{Z_n}(t) = \frac{1 - \mathrm{e}^{-\lambda_n t}}{1 - \mathrm{e}^{-\lambda_n}}\,.$$

Par suite, on en déduit que si $t \in [0,1]$, alors

$$F_{Z_n}(t) = \frac{1 - \mathrm{e}^{-\lambda_n t}}{1 - \mathrm{e}^{-\lambda_n}} \sim \frac{-\lambda_n t}{-\lambda_n} = t\,,$$

si $t < 0$, alors $F_{Z_n}(t) = 0$, et si $t > 0$, alors $F_{Z_n}(t) = 1$. Ainsi,

$$F_{Z_n}(t) \xrightarrow[n\to\infty]{} t\mathbb{1}_{[0,1]}(t) + \mathbb{1}_{]1,\infty[}(t)\,.$$

On reconnaît une fonction de répartition d'une loi uniforme sur $[0,1]$. En effet, si $Z \sim \mathcal{U}([0,1])$, alors $F_Z(t)$ est nulle pour $t < 0$, égale à 1 pour $t > 1$, et vérifie

$$F_Z(t) = \int_0^t \mathbb{1}_{[0,1]}(u)\,\mathrm{d}u = t\,,$$

si $t \in [0,1]$. Bref, on vient donc de montrer que $Z_n \xrightarrow{\mathcal{L}} Z$ quand $n \to \infty$, ou encore $Z_n \xrightarrow{\mathcal{L}} \mathcal{U}([0,1])$.

TD 11 - Différents modes de convergence et théorème central limite

Exercice 1

1. Rappeler les quatre modes de convergence vus en cours, leurs implications, et réciproques (sous certaines hypothèses) éventuelles. On se restreindra au cas de variables aléatoires réelles.

   Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires réelles et X une variable aléatoire réelle. On dit que

   - $(X_n)_{n\geq 1}$ converge presque sûrement vers X si
   $$\mathbb{P}\Big(\Big\{\omega \in \Omega : X_n(\omega) \xrightarrow[n\to\infty]{} X(\omega)\Big\}\Big) = 1\,.$$
   - $(X_n)_{n\geq 1}$ converge en probabilité vers X si pour tout $\epsilon > 0$,
   $$\mathbb{P}(|X_n - X| \geq \epsilon) \xrightarrow[n\to\infty]{} 0\,.$$
   - $(X_n)_{n\geq 1}$ converge dans $L^p(\Omega, \mathcal{F}, \mathbb{P})$ vers X si
   $$\mathbb{E}[|X_n - X|^p] \xrightarrow[n\to\infty]{} 0\,.$$
   - $(X_n)_{n\geq 1}$ converge en loi vers X si pour toute fonction $h : \mathbb{R} \to \mathbb{R}$ continue bornée,
   $$\mathbb{E}[h(X_n)] \xrightarrow[n\to\infty]{} \mathbb{E}[h(X)]\,.$$

   Pour la convergence dans L^p, il faut bien évidemment supposer que les variables aléatoires considérées sont bien dans L^p. Les implications et les réciproques éventuelles sont données en Figure 1.

2. Énoncer le théorème central-limite. On écrira la convergence en loi qui apparaît dans le théorème avec la définition puis avec la caractérisation via les fonctions de répartition.

   Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes et de même loi dans $L^2(\Omega, \mathcal{F}, \mathbb{P})$. On pose $S_n = X_1 + \cdots + X_n$ et $\sigma^2 = \operatorname{Var}(X_1)$. Alors
   $$\frac{S_n - n\mathbb{E}[X_1]}{\sigma\sqrt{n}} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0,1)\,.$$
   La fonction de répartition Φ de la loi normale centrée réduite étant continue en tout point, le théorème central limite se réécrit
   $$\mathbb{P}\Big(\frac{S_n - n\mathbb{E}[X_1]}{\sigma\sqrt{n}} \leq x\Big) \xrightarrow[n\to\infty]{} \Phi(x)\,, \quad x \in \mathbb{R}\,.$$

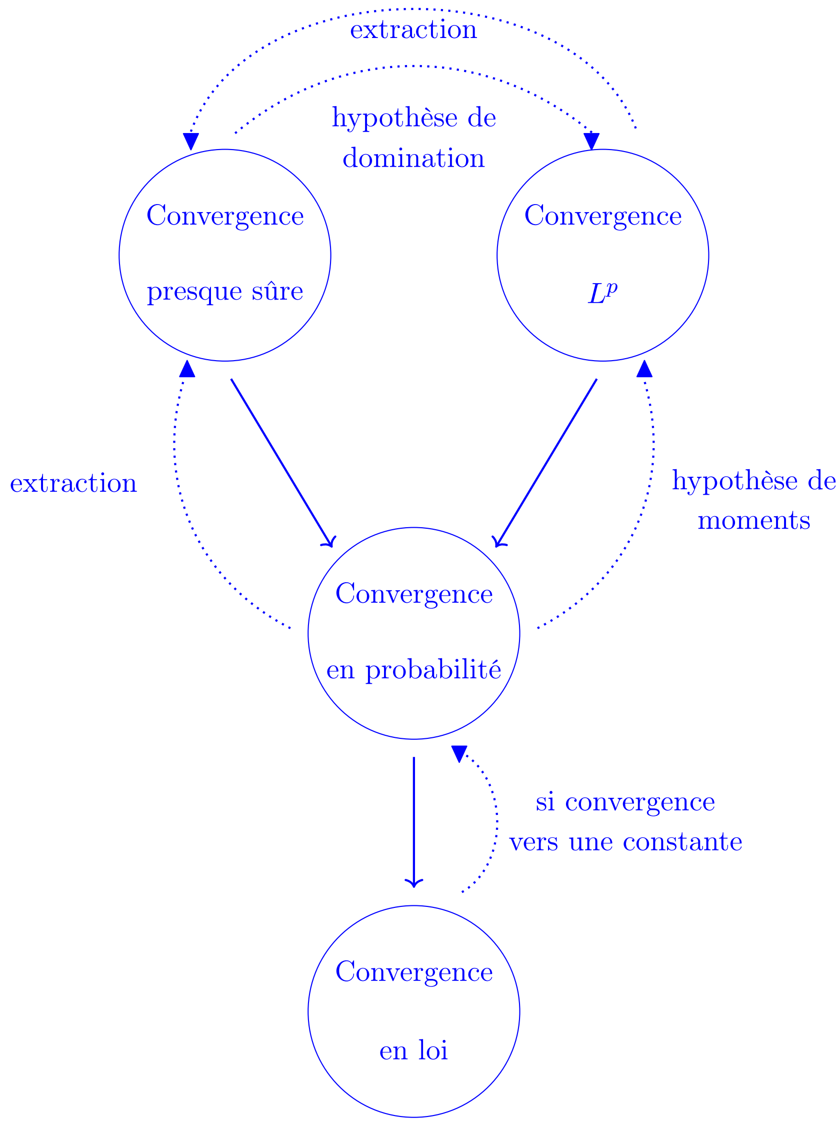


FIGURE 1 – Les différentes convergences, leurs implications et réciproques éventuelles.

EXERCICE 2 Pour tout entier $n \geq 1$ on définit la fonction f_n sur $\mathbb{R}$ par

$$f_n(x) = \frac{n}{\pi(1+n^2x^2)}, \quad x \in \mathbb{R}.$$

1. Montrer que f_n est une densité pour tout entier $n \geq 1$.
   Pour tout $n \geq 1$, f_n est une fonction mesurable (car continue), positive et vérifie
   $$\int_{\mathbb{R}} f_n(x)\,\mathrm{d}x = \int_{\mathbb{R}} \frac{n}{\pi(1+n^2x^2)}\,\mathrm{d}x = \int_{\mathbb{R}} \frac{1}{\pi(1+t^2)}\,\mathrm{d}t = \left[\frac{\arctan(t)}{\pi}\right]_{-\infty}^{\infty} = 1\,,$$
   où on a utilisé le changement de variable $t = nx$ à la deuxième égalité. Ainsi, f_n est bien une densité.

2. Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires telle que X_n admet pour densité f_n. Les variables aléatoires X_n admettent-elles des moments ?

   Pour $n \geq 1$, on a l'équivalent $xf_n(x) \sim \frac{1}{\pi n x}$ quand $x \to \infty$, et cette quantité n'est pas intégrable en $\pm\infty$. Donc X_n n'admet pas d'espérance et a fortiori pas de moment d'ordre p pour tout $p \geq 1$.

3. Étudier la convergence en loi, puis la convergence en probabilité de la suite $(X_n)_{n\geq 1}$.

   Un calcul d'intégrale similaire à celui de la question 1. donne la fonction de répartition F_n de X_n :

$$F_n(x) = \int_{-\infty}^{x} \frac{n}{\pi(1+n^2y^2)}\,\mathrm{d}y = \int_{-\infty}^{nx} \frac{1}{\pi(1+t^2)}\,\mathrm{d}t = \left[\frac{\arctan(t)}{\pi}\right]_{-\infty}^{nx} = \frac{\arctan(nx)}{\pi} + \frac{1}{2}\,.$$

   - Si $x < 0$, alors $\arctan(nx) \to -\pi/2$, donc $F_n(x) \to 0$ quand $n \to \infty$.
   - Si $x > 0$, alors $\arctan(nx) \to \pi/2$, donc $F_n(x) \to 1$ quand $n \to \infty$.
   - Enfin, $F_n(0) = 1/2$.

   Soit X variable aléatoire de loi de Dirac en 0, c'est à dire que X est identiquement nulle. Sa fonction de répartition est donnée par $F_X(x) = \mathbb{1}_{[0,\infty[}(x)$ et F_X est continue partout sauf en 0. On vient de prouve qu'en tout point de continuité de F_X, on a $F_{X_n}(x) \to F_X(x)$ quand $n \to \infty$. Ceci prouve la convergence en loi $X_n \xrightarrow{\mathcal{L}} X$. Autrement dit, la suite $(X_n)_{n\geq 1}$ converge en loi vers une masse de Dirac en 0.

   On a vu en cours que la convergence en loi vers une loi de Dirac était équivalente à la convergence en probabilité. Donc la suite $(X_n)_{n\geq 1}$ converge également en probabilité vers 0.

Exercice 3 On considère deux suites de variables aléatoires $(X_n)_{n\geq 1}$ et $(Y_n)_{n\geq 1}$ définies sur le même espace probabilisé et qui convergent en loi respectivement vers les variables aléatoires X et Y.

1. On suppose que pour tout $n \geq 1$, X_n est indépendant de Y_n et que X est indépendant de Y. Montrer que $X_n + Y_n$ converge en loi vers $X + Y$.

   Soit φ_X la fonction caractéristique de X (et de même pour Y, X_n, Y_n). On rappelle que la fonction caractéristique d'une somme de variables aléatoires est égale au produit des fonctions caractéristiques des variables aléatoires considérées. Ainsi,

$$\varphi_{X_n+Y_n} = \varphi_{X_n}\varphi_{Y_n}$$

   La convergence en loi de X_n vers X et de Y_n vers Y implique alors que

$$\varphi_{X_n}\varphi_{Y_n} \xrightarrow[n\to\infty]{} \varphi_X\varphi_Y\,.$$

   Or, X et Y sont indépendants ce qui donne $\varphi_X\varphi_Y = \varphi_{X+Y}$. On a donc prouvé que $\varphi_{X_n+Y_n} \to \varphi_{X+Y}$, ce qui implique la convergence en loi de $X_n + Y_n$ vers $X + Y$.

2. On considère deux variables aléatoires indépendantes X et Y de loi de Bernoulli de paramètre $1/2$. Pour tout entier $n \geq 1$, on pose $X_n = X + \frac{1}{n}$ et $Y_n = (1 - X) - \frac{1}{n}$. Montrer que X_n et Y_n convergent respectivement en loi vers X et Y, mais que $X_n + Y_n$ ne converge pas vers $X + Y$.

Les suites (X_n) et (Y_n) convergent presque sûrement vers X et $1-X$ respectivement, d'où la convergence en loi en remarquant que Y et $1-X$ ont la même loi (une loi de Bernoulli de paramètre $\frac{1}{2}$). Cependant, la suite (X_n+Y_n) est constante égale à 1, donc elle converge en loi vers 1 qui n'est pas la loi de $X+Y$.

Par rapport à la question 1, il manque l'hypothèse d'indépendance de X_n et Y_n. Bref, la convergence en loi n'est pas conservée via les opérations classiques.

Exercice 4 Soit $X_1, X_2, \dots$ une suite iid de variables aléatoires telle que

$$\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = 1/2\,.$$

On pose $S_n = X_1 + \cdots + X_n$. En utilisant le théorème central limite, déterminer la limite de $\mathbb{P}(S_n \geq 0)$ quand $n \to \infty$.

Un calcul direct donne $\mathbb{E}[S_n] = n\mathbb{E}[X_1] = 0$ et $\operatorname{Var}(S_n) = n\operatorname{Var}(X_1) = n$. Le théorème central limite donne alors la convergence en loi suivante :

$$\frac{S_n - n\mathbb{E}[X_1]}{\sqrt{n\operatorname{Var}(X_1)}} = \frac{S_n}{\sqrt{n}} \xrightarrow[n\to\infty]{\mathcal{L}} N\,,$$

où N désigne une variable aléatoire gaussienne centrée réduite. Comme la fonction de répartition de la loi normale centrée réduite est continue en tout point, la convergence précédente est équivalente à

$$\mathbb{P}\Big(\frac{S_n}{\sqrt{n}} \leq x\Big) \xrightarrow[n\to\infty]{} \mathbb{P}(N \leq x)\,, \quad x \in \mathbb{R}\,.$$

Ainsi, en choisissant $x = 0$, on en déduit que

$$\mathbb{P}(S_n \leq 0) = \mathbb{P}\Big(\frac{S_n}{\sqrt{n}} \leq 0\Big) \xrightarrow[n\to\infty]{} \mathbb{P}(N \leq 0) = \frac{1}{2}\,.$$

Pour aller plus loin

Soit (X_n) une suite de variables aléatoires réelles dans $L^2(\Omega, \mathcal{F}, \mathbb{P})$ ayant toutes la même espérance μ et la même variance σ^2. On suppose que $\operatorname{Cov}(X_n, X_k) \leq 0$ pour tout $n \neq k$. Montrer que la suite (X_n) vérifie la loi faible des grands nombres. On pourra étudier la convergence dans L^2 de $\bar{X}_n$.

Il suffit d'écrire

$$\mathbb{E}[(\bar{X}_n - \mu)^2] = \mathbb{E}\Big[\Big(\frac{1}{n}\sum_{k=1}^{n}(X_k - \mu)\Big)^2\Big] = \frac{1}{n^2}\mathbb{E}\Big[\Big(\sum_{k=1}^{n}(X_k - \mu)\Big)^2\Big]\,.$$

Or le carré se développe en

$$\Big(\sum_{k=1}^{n}(X_k - \mu)\Big)^2 = \sum_{1\leq i,j\leq n}(X_i - \mu)(X_j - \mu) = \sum_{i=1}^{n}(X_i - \mu)^2 + \sum_{1\leq i\neq j\leq n}(X_i - \mu)(X_j - \mu)\,.$$

En appliquant l'espérance à l'égalité précédente on obtient alors

$$\begin{aligned}
\mathbb{E}[(\bar{X}_n-\mu)^2] &= \frac{1}{n^2}\sum_{i=1}^n \mathbb{E}[(X_i-\mu)^2] + \frac{1}{n^2}\sum_{1\le i\neq j\le n}^n \mathbb{E}[(X_i-\mu)(X_j-\mu)] \\
&= \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}(X_i) + \frac{1}{n^2}\sum_{1\le i\neq j\le n}^n \mathrm{Cov}(X_i,X_j)\,.
\end{aligned}$$

L'hypothèse $\mathrm{Cov}(X_i,X_j)\le 0$ permet alors de conclure que

$$\mathbb{E}[(\bar{X}_n-\mu)^2] \le \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}(X_i) = \frac{\sigma^2}{n} \xrightarrow[n\to\infty]{} 0\,.$$

Ainsi, $X_n \xrightarrow[n\to\infty]{L^2} \mu$, donc $X_n \xrightarrow[n\to\infty]{\mathbb{P}} \mu$, ce qui prouve que la suite $(X_n)_{n\ge1}$ vérifie la loi faible des grands nombres.

TD 12 - Théorème central limite

Dans tout le corrigé et sauf mention explicite du contraire, N désigne une variable aléatoire de loi normale centrée réduite et Φ sa fonction de répartition.

Exercice 1 Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes de même loi de Poisson de paramètre 1. On pose $S_n = \sum_{k=1}^n X_k$.

1. Calculer la fonction caractéristique d'une loi de Poisson de paramètre 1.
   Soit X une variable aléatoire suivant une loi de Poisson de paramètre 1. Pour tout $t \in \mathbb{R}$,
   $$\varphi_X(t) = \mathbb{E}[\mathrm{e}^{itX}] = \sum_{k=0}^{\infty} \frac{\mathrm{e}^{-1}\,\mathrm{e}^{itk}}{k!} = \mathrm{e}^{-1} \sum_{k=0}^{\infty} \frac{(\mathrm{e}^{it})^k}{k!} = \mathrm{e}^{-1}\exp(\mathrm{e}^{it}) = \exp(\mathrm{e}^{it} - 1)\,.$$

2. En déduire que la variable aléatoire S_n suit une Poisson de paramètre n et calculer $\mathbb{E}[S_n]$.
   On sait que la fonction caractéristique d'une somme de variables aléatoires indépendantes est égale au produit des fonctions caractéristique de ces variables aléatoires. Ainsi,
   $$\varphi_{S_n}(t) = \varphi_{X_1}(t)\cdots\varphi_{X_n}(t) = (\varphi_{X_1})^n(t) = \exp(n(\mathrm{e}^{it} - 1))\,.$$
   L'espérance de S_n s'obtient par linéarité :
   $$\mathbb{E}[S_n] = \mathbb{E}[X_1 + \cdots + X_n] = \mathbb{E}[X_1] + \cdots + \mathbb{E}[X_n] = n\,.$$

3. À l'aide du théorème central limite, déterminer la limite de $\mathbb{P}(S_n/n \leq 1)$.
   Les variables aléatoires X_k étant iid de variance finie 1, on peut appliquer le théorème central limite qui donne la convergence en loi
   $$\sqrt{n}\Big(\frac{S_n}{n} - 1\Big) \xrightarrow[n\to\infty]{\mathcal{L}} N\,.$$
   Par continuité de Φ en 0, on en déduit que
   $$\mathbb{P}\Big(\sqrt{n}\Big(\frac{S_n}{n} - 1\Big) \in [0,\infty[\Big) \xrightarrow[n\to\infty]{} \mathbb{P}(N \in [0,\infty[) = \frac{1}{2}\,,$$
   ce qui se réécrit
   $$\mathbb{P}\Big(\frac{S_n}{n} \leq 1\Big) \xrightarrow[n\to\infty]{} \frac{1}{2}\,.$$

4. En déduire la limite de la suite $(u_n)_{n\geq 1}$ définie par
   $$u_n = \mathrm{e}^{-n} \sum_{k=0}^{n} \frac{n^k}{k!}\,, \quad n \geq 1\,.$$
   Comme S_n suit une loi de Poisson de paramètre n, on a l'égalité $u_n = \mathbb{P}(S_n \leq n)$. Ainsi, la question précédente donne $u_n \to \frac{1}{2}$ quand $n \to \infty$.

Exercice 2 Après une soirée arrosée, vous vous retrouvez au milieu de la rue de la Loge, en plein centre-ville de Montpellier, mais vous êtes incapable de savoir quelle direction prendre pour rentrer chez vous. Vous errez donc dans la rue en parcourant un pas d'un mètre dans un sens ou dans l'autre de manière aléatoire. Au bout de 100 pas, dans un rayon de combien de mètres autour de votre point de départ allez-vous vous trouver avec 90% de chances ?

Pour faciliter la correction, on assimile la rue de la Loge à une droite graduée et on suppose que vous partez du point d'abscisse 0. Votre déplacement peut alors être modélisé par des variables aléatoires Y_k telles que $Y_k = 1$ si vous vous déplacez dans un sens et $Y_k = -1$ si vous vous déplacez dans l'autre sens. Au bout de n pas, vous vous trouvez au point d'abscisse $S_n = Y_1 + \cdots + Y_n$. L'objectif est alors de déterminer $d > 0$ tel que

$$\mathbb{P}(-d \leq S_{100} \leq d) = 0.90\,.$$

Les X_k sont d'espérance nulle et de variance 1 donc $\mathbb{E}[S_n] = 0$ et $\mathrm{Var}(S_n) = n$. On peut alors appliquer le théorème central limite :

$$\frac{S_n}{\sqrt{n}} \xrightarrow[n\to\infty]{\mathcal{L}} N\,.$$

En particulier, pour tout $a \leq b$,

$$\mathbb{P}\Big(a \leq \frac{S_n}{\sqrt{n}} \leq b\Big) \xrightarrow[n\to\infty]{} \mathbb{P}(a \leq N \leq b)\,.$$

On considère cette convergence comme une approximation pour $n = 100$, ce qui donne

$$\mathbb{P}(-10x \leq S_{100} \leq 10x) \approx \mathbb{P}(-x \leq N \leq x) = 2\Phi(x) - 1\,.$$

Il s'agit donc de trouver $d > 0$ tel que

$$0.90 = \mathbb{P}(-d \leq S \leq d) = 2\Phi\Big(\frac{x}{10}\Big) - 1\,.$$

ce qui revient à $\Phi(\frac{x}{10}) = 0.90$. La table de la loi normale donne alors $\frac{d}{10} = 1.64$ (ou 1.65, ou bien prendre la moyenne des deux), d'où $d = 16.4$m. Ainsi, il y a 90% de chances que vous vous trouviez dans un rayon de 16.4m autour de votre point de départ.

Exercice 3 Avant une élection, un institut de sondage contacte des individus pour connaître leur intention de vote entre deux candidats A et B. On suppose que les appels sont indépendants et que chaque individu répond qu'il va voter A avec une probabilité p_A (et donc B avec une probabilité $1-p$). Le but est d'étudier le paramètre inconnu p_A du modèle. Soit n le nombre de réponses en faveur de A collectées en n appels.

1. Exprimer Z_n comme une somme de variables aléatoires iid de loi de Bernoulli dont on précisera le paramètre.

   Pour un individu k interrogé, on note X_k la variable aléatoire définie par $X_k = 1$ si l'individu a l'intention de voter pour A et $X_k = 0$ sinon. Ainsi, X_k suit une loi de Bernoulli de paramètre p_A et les variables aléatoires $X_1, X_2, \ldots$ sont indépendantes. On a alors $Z_n = \sum_{k=1}^{n} X_k$.

2. Étudier la convergence presque sûre de la suite $(Z_n/n)_{n\geq 1}$.

   Les variables aléatoires X_k sont iid d'espérance finie égale à p_A. On peut donc appliquer la loi des grands nombres qui donne

   $$\frac{Z_n}{n} = \frac{\sum_{k=1}^n X_k}{n} \xrightarrow[n\to\infty]{p.s.} p_A\,.$$

   En résumé, plus l'institut de sondage interroge de personnes, plus il se rapproche de la vraie proportion de votes pour A.

3. Soit $\delta > 0$ un paramètre (petit). En utilisant l'inégalité de Bienaymé-Tchebychev, donner un intervalle aléatoire (appelé intervalle de confiance) $I = I(Z_n, \delta)$ pour p_A, c'est-à-dire un intervalle tel que

   $$\mathbb{P}(p_A \in I) \geq 1 - \delta\,.$$

   On prendra garde au fait que I ne doit pas dépendre de p_A qui est inconnu.

   Un calcul direct donne $\mathbb{E}[Z_n] = np_A$ et $\mathrm{Var}(Z_n) = np_A(1-p_A)$. L'inégalité de Bienaymé-Tchebychev appliquée à la variable aléatoire $\frac{Z_n}{n}$ donne alors, pour tout $\epsilon > 0$,

   $$\mathbb{P}\Big(\Big|\frac{Z_n}{n} - p_A\Big| \geq \epsilon\Big) \leq \frac{p_A(1-p_A)}{n\epsilon^2}\,.$$

   Mais cette borne dépend de p_A que l'on ne connaît pas. On majore $p_A(1-p_A)$ par $\frac{1}{4}$ et on obtient alors

   $$\mathbb{P}\Big(\Big|\frac{Z_n}{n} - p_A\Big| \geq \epsilon\Big) \leq \frac{1}{4n\epsilon^2}\,.$$

   En passant au complémentaire, on peut réécrire cette inégalité sous la forme

   $$\mathbb{P}\Big(\frac{Z_n}{n} - \epsilon \leq p_A \leq \frac{Z_n}{n} + \epsilon\Big) \geq 1 - \frac{1}{4n\epsilon^2}\,.$$

   On choisit alors ϵ de sorte que $\frac{1}{4n\epsilon^2} = \delta$, c'est-à-dire $\epsilon = \frac{1}{\sqrt{4n\delta}}$. Ainsi, on a montré que

   $$\mathbb{P}\left(\frac{Z_n}{n} - \frac{1}{\sqrt{4n\delta}} \leq p_A \leq \frac{Z_n}{n} + \frac{1}{\sqrt{4n\delta}}\right) \geq 1 - \delta\,.$$

   L'intervalle de confiance recherché est donc

   $$I = I(Z_n, \delta) = \left[\frac{Z_n}{n} - \frac{1}{\sqrt{4n\delta}}, \frac{Z_n}{n} + \frac{1}{\sqrt{4n\delta}}\right].$$

   Notons que c'est un intervalle de confiance aléatoire : si on effectue un sondage auprès d'autres personnes, les bornes ne seront pas forcément les mêmes. On remarque que la longueur de l'intervalle est $\frac{2}{\sqrt{4n\delta}}$ qui diminue quand n augmente.

4. Répondre à la question en utilisant cette fois-ci le théorème central limite.

   D'après le théorème central limite,

   $$\frac{Z_n - np_A}{\sqrt{np_A(1-p_A)}} \xrightarrow[n\to\infty]{\mathcal{L}} N\,,$$

ce qui donne

$$\mathbb{P}\Big(-q \le \frac{Z_n - np_A}{\sqrt{np_A(1-p_A)}} \le q\Big) \xrightarrow[n\to\infty]{} \Phi(q) - \Phi(-q)\,.$$

ou encore, par symétrie de Φ,

$$\mathbb{P}\Big(\frac{Z_n}{n} - \frac{q\sqrt{p_A(1-p_A)}}{\sqrt{n}} \le p_A \le \frac{Z_n}{n} + \frac{q\sqrt{p_A(1-p_A)}}{\sqrt{n}}\Big) \xrightarrow[n\to\infty]{} 2\Phi(q) - 1\,.$$

Ici aussi, le paramètre p_A étant inconnu, on utilise la majoration $p_A(1-p_A) \le \frac{1}{4}$, et on en déduit alors une borne supérieure :

$$\lim_{n\to\infty} \mathbb{P}\Big(p_A \in \Big[\frac{Z_n}{n} - \frac{q}{2\sqrt{n}}, \frac{Z_n}{n} + \frac{q}{2\sqrt{n}}\Big]\Big) \ge 2\Phi(q) - 1\,.$$

Avec par exemple, $q = 1.96$, ce qui correspond à $2\Phi(q) - 1 = 0.95$, on obtient l'intervalle

$$\Big[\frac{Z_n}{n} - \frac{1.96}{2\sqrt{n}}, \frac{Z_n}{n} + \frac{1.96}{2\sqrt{n}}\Big]$$

tandis qu'avec l'inégalité de Bienaymé-Tchebvychev on obtient, avec $\delta = 0.05$,

$$\Big[\frac{Z_n}{n} - \frac{1}{\sqrt{0.2n}}, \frac{Z_n}{n} + \frac{1}{\sqrt{0.2n}}\Big]$$

Comme $\frac{1.96}{2} = 0.98 \le \frac{1}{\sqrt{0.2}} \approx 2.24$, on en déduit que l'intervalle de confiance donné par le théorème central limite est meilleur car de longueur plus petite. Cependant, il n'est valable que pour n grand puisqu'il s'obtient avec un résultat asymptotique.

Pour aller plus loin

Quel est le nombre minimum de lancers nécessaires d'une pièce équilibrée pour que la fréquence de Pile soit comprise entre 0.4 et 0.6 avec une probabilité de 0.95 ?

Soit X_k le résultat du k-ème lancer : $X_k = 1$ si on obtient Pile et $X_k = 0$ sinon. Les X_k forment une suite de variables aléatoires iid de loi de Bernoulli de paramètre $p = 1/2$. En particulier $\mathbb{E}[X_k] = p$ et $\mathrm{Var}(X_k) = p(1-p)$. La fréquence de Pile au bout de n lancers est donnée par S_n/n, où $S_n = \sum_{k=1}^n X_k$, et on cherche à déterminer la valeur minimale de n telle que

$$\mathbb{P}(0.4 \le S_n/n \le 0.6) \ge 0.95\,.$$

Cette probabilité peut se réécrire

$$\mathbb{P}(0.4 \le S_n/n \le 0.6) = \mathbb{P}(|S_n/n - 1/2| \le 0.1) = \mathbb{P}\big(-0.2\sqrt{n} \le \sqrt{n}(S_n/n - 1/2) \le 0.2\sqrt{n}\big)\,.$$

Les variables aléatoires X_k étant iid de variance finie, on peut appliquer le théorème central limite qui assure alors que

$$\sqrt{n}\frac{S_n/n - p}{\sqrt{p(1-p)}} \xrightarrow{\mathcal{L}} N\,, \quad n \to \infty\,,$$

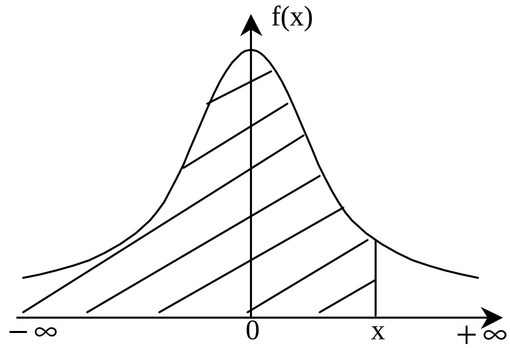
avec ici $p = 1/2$. Asymptotiquement, on est donc amené à trouver le plus petit n tel que

$$0.95 \ge \mathbb{P}\big(|N| \le 0.2\sqrt{n}\big) = 2\Phi(0.2\sqrt{n}) - 1\,.$$

La table de la loi normale donne alors $0.2\sqrt{n} \ge 1.96$, puis $n \ge 96.04$. Il faut donc au moins 97 lancers pour être sûr à 95% que la fréquence de Pile soit comprise entre 0.4 et 0.6.

Loi Normale centrée réduite

Probabilité de trouver une valeur inférieure à x.



$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2p}} e^{-\frac{u^2}{2}} du$$

X	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
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998

Table pour les grandes valeurs de x :

x	3	3,2	3,4	3,6	3,8	4	4,2	4,4	4,6	4,8
F(x)	0,99865003	0,99931280	0,99966302	0,99984085	0,99992763	0,99996831	0,99998665	0,99999458	0,99999789	0,99999921