Leçon 13

Somme de variables aléatoires indépendantes

La théorie de l'addition de variables aléatoires (réelles) est un thème central du calcul des probabilités. Elle intervient à tous les niveaux, notamment dans l'examen des propriétés asymptotiques de ces sommes dans la loi des grands nombres et le théorème central limite.

Il est bien entendu possible de sommer des variables aléatoires complexes ou des vecteurs aléatoires (dans $\mathbb{R}^d$, voire dans des espaces vectoriels topologiques) ; pour plus de simplicité, la leçon se limitera à l'addition de variables aléatoires réelles. Toutes les variables aléatoires sont définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

1 Loi de la somme de variables aléatoires indépendantes

Ce paragraphe décrit la loi $\mathbb{P}_{X+Y}$ de la somme $X+Y$ de deux variables aléatoires indépendantes X et Y à travers les intégrales

$$\mathbb{E}\big(\phi(X+Y)\big) = \int_{\mathbb{R}} \phi \, d\,\mathbb{P}_{X+Y}$$

contre des fonctions boréliennes $\phi : \mathbb{R} \to \mathbb{R}$, positives ou bornées.

Proposition 1. *Si X et Y sont deux variables aléatoires réelles indépendantes, de loi respective $\mathbb{P}_X$ et $\mathbb{P}_Y$, la loi $\mathbb{P}_{X+Y}$ de la somme $X+Y$ est donnée, sur toute fonction borélienne $\phi : \mathbb{R} \to \mathbb{R}$, positive ou bornée, par*

$$\mathbb{E}\big(\phi(X+Y)\big) = \int_{\mathbb{R}} \phi \, d\,\mathbb{P}_{X+Y} = \int_{\mathbb{R}} \int_{\mathbb{R}} \phi(x+y) d\,\mathbb{P}_X(x) d\,\mathbb{P}_Y(y).$$

C'est une conséquence du théorème de transport et du fait que la loi du couple (X,Y), par indépendance, est la loi produit $\mathbb{P}_X \otimes \mathbb{P}_Y$. En effet, par transport sur $\mathbb{R}^2$ pour la loi du couple (X,Y) appliqué à la fonction

$(x,y) \in \mathbb{R}^2 \mapsto \phi(x+y)$,

$$\mathbb{E}\big(\phi(X+Y)\big) \,=\, \int_{\mathbb{R}^2} \phi(x+y) d\,\mathbb{P}_{(X,Y)}(x,y).$$

Comme $\mathbb{P}_{(X,Y)} = \mathbb{P}_X \otimes \mathbb{P}_Y$, la conclusion s'ensuit par le théorème de Fubini-Tonelli.

Ainsi que l'énoncé de la proposition le montre clairement, la fonction de répartition n'est pas toujours le meilleur moyen pour décrire la loi d'une somme de variables aléatoires indépendantes, et le calcul intégral offre plus de souplesse.

L'opération qui à deux mesures (de probabilité) P et Q sur $\mathbb{R}$ associe la mesure R, sur les boréliens de $\mathbb{R}$, décrite par

$$\int_{\mathbb{R}} \phi\, dR \,=\, \int_{\mathbb{R}^2} \phi(x+y) dP(x) dQ(y)$$

pour toute fonction borélienne ϕ, positive ou bornée, est appelée *produit de convolution*, noté $R = P * Q$. La proposition se reformule ainsi en exprimant que la loi $\mathbb{P}_{X+Y}$ de la somme $X+Y$ est le produit de convolution des lois $\mathbb{P}_X$ et $\mathbb{P}_Y$ de X et Y, soit $\mathbb{P}_{X+Y} = \mathbb{P}_X * \mathbb{P}_Y$.

Si P et Q ont des densités respectives f et g par rapport à la mesure de Lebesgue λ sur $\mathbb{R}$, $P * Q$ a pour densité le produit de convolution

$$f * g(x) \,=\, \int_{\mathbb{R}} f(y) g(x-y) d\lambda(y) \,=\, \int_{\mathbb{R}} f(x-y) g(y) d\lambda(y), \quad x \in \mathbb{R},$$

des densités f et g. Il suffit en fait que l'une des deux mesures P ou Q ait une densité pour que le produit de convolution en ait une aussi. Cette observation est d'importance dans la « régularisation » des mesures, et est souvent employée avec des lois normales $\mathcal{N}(0, \sigma^2)$ (qui lorsque $\sigma \to 0$ approchent la mesure initiale).

Le produit de convolution (des mesures) représente donc pleinement la loi de l'addition de variables aléatoires indépendantes. Toutefois, le point de vue probabiliste introduit souvent plus de souplesse dans la lecture du produit de convolution. Par exemple, si P et Q sont deux masses de Dirac δ_a et δ_b, alors $\delta_a * \delta_b = \delta_{a+b}$ se vérifie sur la définition du produit de convolution. Mais δ_a est la loi de la variable constante égale à a, δ_b celle de la variable constante égale à b, les deux variables sont indépendantes (car déterministes), et leur somme est évidemment $a+b$, de loi δ_{a+b}.

De façon similaire, si X_1 et X_2 sont deux variables aléatoires indépendantes de même loi de Bernoulli $\mathcal{B}(p)$, $p \in]0,1[$, sur $\{0,1\}$, il est aisé de décrire la loi de la variable $X_1 + X_2$; en effet, cette dernière prend ses valeurs dans $\{0,1,2\}$, et, par indépendance,

$$\begin{aligned}
\mathbb{P}(X_1 + X_2 = 0) &= \mathbb{P}(X_1 = 0, X_2 = 0) = \mathbb{P}(X_1 = 0)\,\mathbb{P}(X_2 = 0) = (1-p)^2, \\
\mathbb{P}(X_1 + X_2 = 1) &= \mathbb{P}(X_1 = 1, X_2 = 0) + \mathbb{P}(X_1 = 0, X_2 = 1) \\
&= \mathbb{P}(X_1 = 1)\,\mathbb{P}(X_2 = 0) + \mathbb{P}(X_1 = 0)\,\mathbb{P}(X_2 = 1) \\
&= 2p(1-p), \\
\mathbb{P}(X_1 + X_2 = 2) &= \mathbb{P}(X_1 = 1, X_2 = 1) = \mathbb{P}(X_1 = 1)\,\mathbb{P}(X_2 = 1) = p^2.
\end{aligned}$$

Ces formules montrent que la loi de $X_1 + X_2$ est binomiale $\mathcal{B}(2,p)$. La proposition suivante s'établit par récurrence suivant le même schéma.

Proposition 2. *Si $X_1, \ldots, X_n$, $n \geq 1$, sont indépendantes de même loi de Bernoulli $\mathcal{B}(p)$, $p \in]0,1[$, sur $\{0,1\}$, alors la somme $X_1 + \cdots + X_n$ suit la loi binomiale $\mathcal{B}(n,p)$.*

L'intérêt de cette représentation de la loi binomiale par une somme de variables indépendantes de Bernoulli est par exemple mis en évidence dans le calcul de l'espérance et de la variance puisque, par linéarité,

$$\mathbb{E}(X_1 + \cdots + X_n) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n) = np$$

et par l'identité de Bienaymé (les variables sont indépendantes, donc non corrélées)

$$\mathrm{Var}(X_1 + \cdots + X_n) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n) = np(1-p).$$

Des raisonnements analogues peuvent se déployer pour d'autres lois. À titre illustratif, si X_1 et X_2 sont indépendantes de loi de Poisson $\mathcal{P}(\theta_1)$ et $\mathcal{P}(\theta_2)$ respectivement, $\theta_1, \theta_2 > 0$, alors $X_1 + X_2$ suit la loi de Poisson $\mathcal{P}(\theta_1 + \theta_2)$. Un démonstration rapide et intuitive consiste à décomposer l'événement $\{X_1 + X_2 = n\}$ pour chaque $n \geq 0$ suivant les valeurs possibles de X_2 (ou X_1), et d'user de l'indépendance pour obtenir

$$\begin{aligned}\mathbb{P}(X_1 + X_2 = n) \;&=\; \sum_{k=0}^{n} \mathbb{P}(X_1 + X_2 = n, X_2 = k) \\ &=\; \sum_{k=0}^{n} \mathbb{P}(X_1 = n - k, X_2 = k) \\ &=\; \sum_{k=0}^{n} \mathbb{P}(X_1 = n - k)\, \mathbb{P}(X_2 = k).\end{aligned}$$

Il ne reste plus qu'à exprimer les probabilités poissonniennes pour conclure que

$$\mathbb{P}(X_1 + X_2 = n) \;=\; e^{-\theta_1 - \theta_2} \sum_{k=0}^{n} \frac{\theta_1^{n-k}}{(n-k)!}\, \frac{\theta_2^k}{k!} \;=\; e^{-(\theta_1 + \theta_2)}\, \frac{(\theta_1 + \theta_2)^n}{n!}$$

d'après la formule du binôme.

L'utilisation du produit de convolution dans ces exemples discrets est plus lourd. Il est requis dans les exemples à densité qui nécessitent de décrire les produits de convolution des densités associés. Par exemple, si X et Y sont indépendantes de même loi exponentielle $\mathcal{E}(1)$, la loi de $X + Y$ a pour densité (par rapport à la mesure de Lebesgue sur $]0, \infty[$) le produit de convolution

$$\int_{\mathbb{R}} e^{-y}\, \mathbb{1}_{]0,\infty[}(y)\, e^{-(x-y)}\, \mathbb{1}_{[0,\infty[}(x - y) d\lambda(y) \;=\; e^{-x} \int_{\mathbb{R}} \mathbb{1}_{]0,x[}(y) d\lambda(y) \;=\; x\, e^{-x}$$

pour $x \in]0, \infty[$. Une récurrence simple établit que la loi de la somme de n variables aléatoires exponentielles $\mathcal{E}(1)$ indépendantes a pour densité $x \mapsto x^{n-1} e^{-x}$ par rapport à la mesure de Lebesgue sur $]0, \infty[$, soit la loi Gamma $\gamma(n, 1)$ dans les notations de l'Exercice 8, Leçon 10.

Un exercice similaire, quoique un peu plus lourd dans les détails, permet de vérifier sur le produit de convolution des densités que si X_1 et X_2 sont indépendantes de lois normales respectives $\mathcal{N}(m_1, \sigma_1^2)$ et $\mathcal{N}(m_2, \sigma_2^2)$, $m_1, m_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$, alors $X_1 + X_2$ a pour loi $\mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$.

2 Fonction caractéristique

L'outil de la fonction caractéristique est très efficace dans l'examen des lois de sommes de variables indépendantes. La proposition suivante est immédiate (et s'étend aux vecteurs aléatoires). Pour rappel, si X est une variable aléatoire réelle, $\varphi_X(u) = \mathbb{E}(e^{iuX})$, $u \in \mathbb{R}$, désigne la fonction caractéristique de sa loi.

Proposition 3. *Si X et Y sont indépendantes, $\varphi_{X+Y} = \varphi_X \, \varphi_Y$.*

Pour la démonstration, pour tout $u \in \mathbb{R}$, d'après les propriétés de l'exponentielle et par indépendance,

$$\varphi_{X+Y}(u) \;=\; \mathbb{E}(e^{iu(X+Y)}) \;=\; \mathbb{E}(e^{iuX} e^{iuY}) \;=\; \varphi_X(u) \, \varphi_Y(u).$$

Attention de ne pas confondre la fonction caractéristique

$$\varphi_{X+Y}(u) \;=\; \mathbb{E}(e^{iu(X+Y)}), \quad u \in \mathbb{R},$$

d'une somme $X + Y$ de variables aléatoires avec celle

$$\varphi_{(X,Y)}(u, v) \;=\; \mathbb{E}(e^{i(uX+vY)}), \quad (u, v) \in \mathbb{R}^2,$$

d'un couple (X, Y) (respectivement égales à $\varphi_X(u) \, \varphi_Y(u)$ et $\varphi_X(u) \, \varphi_Y(v)$ si X et Y sont indépendantes.

L'énoncé peut être testé sur quelques uns des exemples simples précédents. Si X_1 et X_2 sont indépendantes de lois de Poisson $\mathcal{P}(\theta_1)$ et $\mathcal{P}(\theta_2)$, pour tout $u \in \mathbb{R}$,

$$\varphi_{X_1+X_2}(u) \;=\; \varphi_{X_1}(u) \, \varphi_{X_2}(u) \;=\; e^{-(\theta_1+\theta_2)(e^{iu}-1)}$$

ce qui confirme que X_1+X_2 a pour loi $\mathcal{P}(\theta_1+\theta_2)$. Si X_1 et X_2 sont indépendantes de lois normales respectives $\mathcal{N}(m_1,\sigma_1^2)$ et $\mathcal{N}(m_2,\sigma_2^2)$, $m_1,m_2\in\mathbb{R}$, $\sigma_1,\sigma_2>0$, pour tout $u\in\mathbb{R}$,

$$\varphi_{X_1+X_2}(u) \;=\; \varphi_{X_1}(u)\,\varphi_{X_2}(u) \;=\; e^{i(m_1+m_2)u-\frac{1}{2}(\sigma_1^2+\sigma_2^2)u^2}$$

et donc X_1+X_2 a bien pour loi $\mathcal{N}(m_1+m_2,\sigma_1^2+\sigma_2^2)$.

3 Inégalité exponentielle

Ce paragraphe reprend le cadre déjà esquissé en Leçon 11. Il illustre, sur un exemple, une propriété de concentration assez générique obtenue par l'intermédiaire de l'inégalité de Markov avec la fonction exponentielle, bien adaptée aux sommes de variables aléatoires indépendantes.

Soient $X_1,\ldots,X_n$ des variables aléatoires réelles indépendantes, centrées et telles que $|X_k|\le 1$ presque sûrement, $k=1,\ldots,n$; poser $S_n=X_1+\cdots+X_n$.

D'après l'inégalité de Markov composée avec la fonction exponentielle, pour tout $t>0$ et tout $u>0$,

$$\mathbb{P}(S_n\ge t) \;=\; \mathbb{P}\big(e^{uS_n}\ge e^{ut}\big) \;\le\; e^{-ut}\,\mathbb{E}\big(e^{uS_n}\big).$$

L'inégalité reste vraie pour $t,u\ge 0$. Par indépendance,

$$\mathbb{E}\big(e^{uS_n}\big) \;=\; \prod_{k=1}^{n}\mathbb{E}\big(e^{uX_k}\big).$$

La convexité de la fonction exponentielle exprime que pour tous $x,y\in\mathbb{R}$ et $\theta\in[0,1]$,

$$e^{\theta x+(1-\theta)y} \;\le\; \theta e^x+(1-\theta)e^y.$$

Appliqué à $x=u$, $y=-u$ et $\theta=\frac{1}{2}(1+b)$ où $u\in\mathbb{R}$ et $|b|\le 1$, elle entraîne que

$$e^{ub} \;\le\; \mathrm{ch}(u)+b\,\mathrm{sh}(u).$$

Pour chaque $k = 1, \ldots, n$, $|X_k| \leq 1$ (presque sûrement) et $\mathbb{E}(X_k) = 0$. Par intégration de l'inégalité précédente avec $b = X_k$,

$$\mathbb{E}\big(e^{uX_k}\big) \;\leq\; \mathrm{ch}(u) + \mathbb{E}\big(X_k\,\mathrm{sh}(u)\big) \;=\; \mathrm{ch}(u).$$

Enfin, sur un développement en série, $\mathrm{ch}(u) \leq e^{\frac{1}{2}u^2}$.

En conséquence de cette étude, pour tout $u \in \mathbb{R}$,

$$\prod_{k=1}^{n} \mathbb{E}\big(e^{uX_k}\big) \;\leq\; e^{\frac{n}{2}u^2}.$$

En revenant à l'inégalité $\mathbb{P}(S_n \geq t) \leq e^{-ut}\,\mathbb{E}(e^{uS_n})$ pour tous $t, u \geq 0$, il vient

$$\mathbb{P}(S_n \geq t) \;\leq\; e^{-ut+\frac{n}{2}u^2}.$$

À $t \geq 0$ fixé, le choix de $u \geq 0$ peut être optimisé pour rendre l'exposant le plus négatif possible (et donc la majoration la plus petite) en posant $u = \frac{t}{n}$ de sorte que

$$\mathbb{P}(S_n \geq t) \;\leq\; e^{-\frac{1}{2n}t^2}.$$

Il est pertinent d'observer que la même inégalité a lieu pour les variables $-X_k$, $k = 1, \ldots, n$, qui vérifient les mêmes hypothèses, et donc par sous-additivité des probabilités,

$$\mathbb{P}\big(|S_n| \geq t\big) \;\leq\; 2\,e^{-\frac{1}{2n}t^2}$$

pour tout $t \geq 0$.

Les inégalités précédentes sont des cas particuliers d'une inégalité due à W. Hoeffding [1] qui énonce plus généralement que si les $X_1, \ldots, X_n$, indépendantes, sont telles que $X_k \in [a_k, b_k]$ ($a_k < b_k$) presque sûrement, $k = 1, \ldots, n$, pour tout $t \geq 0$,

$$\mathbb{P}\big(S_n - \mathbb{E}(S_n) \geq t\big) \;\leq\; \exp\bigg(-\frac{2t^2}{\sum_{k=1}^{n}(b_k - a_k)^2}\bigg).$$

1. Wassily Hoeffding, statisticien et probabiliste finlandais et américain (1914–1991).

Comme les précédentes, c'est un prototype d'inégalité de concentration. La somme S_n s'écarte de sa valeur moyenne avec une probabilité exponentiellement petite, dans un régime où n est grand et t de l'ordre d'au moins $\sqrt{n}$ (en accord avec le théorème central limite – voir Leçon 20). Il est intéressant de comparer l'inégalité exponentielle à ce que produit l'inégalité de Bienaymé-Tchebychev dans ce cadre. L'inégalité précédente retrouve le résultat pour les variables binomiales décrit dans la Leçon 11, sur la base de la représentation de la loi binomiale comme somme de variables aléatoires indépendantes de Bernoulli.

4 Complément : Inégalité de Bernstein

L'inégalité de Hoeffding présente l'inconvénient de ne pas tenir compte de la variance des variables considérées. Ce défaut est corrigé dans l'inégalité de Bernstein [2].

Proposition 4 (Inégalité de Bernstein)**.** *Soient $X_1, \ldots, X_n$ des variables aléatoires réelles indépendantes, de carré intégrable et telles que $X_k \leq b$ presque sûrement, $k = 1, \ldots, n$, pour un $b > 0$; soit $S_n = X_1 + \cdots + X_n$. Alors, pour tout $t \geq 0$,*

$$\mathbb{P}\big(S_n - \mathbb{E}(S_n) \geq t\big) \;\leq\; \exp\left(-\frac{t^2}{2\big(\sum_{k=1}^n \mathbb{E}(X_k^2) + \frac{1}{3}bt\big)}\right).$$

L'exemple suivant permet de mesurer le bénéfice de cette inégalité par rapport à celle de Hoeffding. Considérer des variables $X_1, \ldots, X_n$ indépendantes de même loi de Bernoulli $\mathcal{B}(p)$, $p \in]0,1[$, sur $\{0,1\}$, de sorte que $S_n = X_1 + \cdots + X_n$ suit la loi binomiale $\mathcal{B}(n,p)$, de moyenne np. L'inégalité de Hoeffding produit, pour tout $t \geq 0$,

$$\mathbb{P}(S_n - np \geq t) \;\leq\; e^{-\frac{2}{n}t^2}$$

alors que celle de Bernstein exprime que

$$\mathbb{P}(S_n - np \geq t) \;\leq\; \exp\left(-\frac{t^2}{2(np + \frac{1}{3}t)}\right).$$

Lorsque t est grand, l'inégalité de Bernstein est moins performante que celle de Hoeffding. Mais lorsque p est petit vis-à-vis de n, par exemple $p = \frac{1}{n}$, et $t \leq \sqrt{n}$, l'inégalité de Hoeffding est essentiellement sans contenu (car $e^{-\frac{2}{n}t^2}$ est d'ordre 1) alors que celle de Bernstein reste pertinente (de l'ordre de $e^{-\frac{3}{2}t}$).

2. Sergeï Natanovich Bernstein, mathématicien russe et soviétique (1880–1968).

Exercices

(*Une étoile * désignera une question de difficulté supérieure.*)

Exercice 1. Soient $X_1, X_2, \dots$ des variables aléatoires indépendantes de même loi de Bernoulli $\mathcal{B}(p)$ sur $\{0,1\}$ de paramètre $p \in]0,1[$, et soit N une variable aléatoire, indépendante des X_k, $k \geq 1$, de loi de Poisson $\mathcal{P}(\theta)$ de paramètre $\theta > 0$.

a) Calculer les fonctions génératrices des moments de X_1 et N.

b) Déterminer la loi de la variable aléatoire $\sum_{k=1}^{N} X_k$ (avec la convention $\sum_{k=1}^{0} = 0$).

Exercice 2 (*Identité de Wald*[3]). Soit X_n, $n \in \mathbb{N}$, une suite de variables aléatoires de même loi intégrable, et soit N une variable aléatoire à valeurs dans $\mathbb{N}$ intégrable, et indépendante de la suite X_n, $n \in \mathbb{N}$. Établir l'identité

$$\mathbb{E}(X_1 + \cdots + X_N) \,=\, \mathbb{E}(N)\,\mathbb{E}(X_1).$$

Exercice 3. Soient $X_1, \dots, X_n$ des variables aléatoires indépendantes de même loi de Bernoulli $\mathcal{B}(p)$ à valeurs dans $\{0,1\}$ de paramètre de succès $p \in]0,1[$; soit la fréquence $F_n = \frac{1}{n}(X_1 + \cdots + X_n)$, $n \geq 1$ (qui comptabilise la proportion de succès sur n tirages).

a) Calculer l'espérance et la variance de F_n.

b) On suppose que $p = \frac{1}{4}$. Démontrer que $\mathbb{P}\big(F_{100\,000} \in \big[\frac{24}{100}, \frac{26}{100}\big]\big) \geq \frac{95}{100}$.

3. Abraham Wald, mathématicien américain d'origine hongroise (1902–1950).

Exercice 4 (*Loi triangulaire*). Sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, soit X une variable aléatoire de loi (dite triangulaire sur $[0,2]$) de densité

$$f(x) \;=\; \begin{cases} 0 & \text{si} \quad x \in\,]-\infty, 0[, \\ x & \text{si} \quad x \in [0,1], \\ 2-x & \text{si} \quad x \in\,]1,2], \\ 0 & \text{si} \quad x \in\,]2, +\infty[, \end{cases}$$

par rapport à la mesure de Lebesgue λ sur les boréliens de $\mathbb{R}$.

a) Décrire et tracer la fonction de répartition F_X de la loi de X.

b) Démontrer que pour tout entier $n \in \mathbb{N}$,

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

Soient à présent U et V deux variables aléatoires indépendantes de même loi uniforme $\mathcal{U}(0,1)$ sur $[0,1]$. Calculer $\mathbb{E}(U^n)$ pour tout entier n.

c) Démontrer que pour tout entier $n \in \mathbb{N}$,

$$\mathbb{E}\big((U+V)^n\big) \;=\; \sum_{k=0}^{n} \frac{n!}{(k+1)!(n-k+1)!}\,.$$

d) Vérifier que

$$\sum_{k=0}^{n} \frac{n!}{(k+1)!(n-k+1)!} \;=\; \frac{1}{(n+1)(n+2)} \sum_{\ell=1}^{n+1} \binom{n+2}{\ell}$$

et en déduire une expression de $\mathbb{E}((U+V)^n)$.

e) Invoquer un théorème pour affirmer que $U+V$ a même loi que X.

Exercice 5*. Sur $(\Omega, \mathcal{A}, \mathbb{P})$, soient X et Y deux variables aléatoires indépendantes de même loi uniforme $\mathcal{U}(-1, +1)$ sur $[-1, +1]$.

a) Démontrer que $X + Y$ a une densité par rapport à la mesure de Lebesgue que l'on calculera.

b) Calculer la transformée de Fourier de la loi X. En déduire celle de $X + Y$.

c) En utilisant le théorème d'inversion de Fourier, démontrer que

$$\int_0^\infty \Big(\frac{\sin(x)}{x}\Big)^2 dx \;=\; \frac{\pi}{2}\,.$$

Exercice 6. Soit X une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $\mathbb{Z}$, dont la loi a pour fonction caractéristique φ_X ; rappeler que $\mathbb{P}(X = 0) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \varphi_X(\theta) d\theta$ (Exercice 11, Leçon 10).

Soit à présent une suite X_k, $k \geq 1$, de variables aléatoires indépendantes de même loi de Bernoulli $\mathbb{P}(X_k = +1) = \mathbb{P}(X_k = -1) = \frac{1}{2}$, et, pour tout $n \geq 1$, $S_n = X_1 + \cdots + X_n$.

a) Démontrer que pour tout $n \geq 1$,

$$\mathbb{P}(S_{2n} = 0) \;=\; \frac{1}{2\pi} \int_{-\pi}^{+\pi} \cos^{2n}(\theta) d\theta.$$

b) Déduire de la question précédente la divergence de la série de terme général $\mathbb{P}(S_{2n} = 0)$.

c) Montrer de façon probabiliste que

$$\mathbb{P}(S_{2n} = 0) \;=\; \frac{1}{2^{2n}} \binom{2n}{n}, \quad n \geq 1,$$

et démontrer que $\mathbb{P}(S_{2n} = 0) \sim \frac{1}{\sqrt{\pi n}}$ avec la formule de Stirling. (*Indication* : faire une transformation pour se ramener à une loi binomiale.)

Exercice 7* (*Loi indéfiniment divisible*). Soit X une variable aléatoire réelle de loi μ ; on dit que μ est indéfiniment divisible si pour tout entier $n \geq 1$ il existe des variables aléatoires réelles $X_{1,n}, \ldots, X_{n,n}$ indépendantes de même loi ν_n telles que la loi de la somme $X_{1,n} + \cdots + X_{n,n}$ soit μ.

a) Démontrer qu'une loi μ est indéfiniment divisible si et seulement si sa fonction caractéristique φ est, pour tout entier n, la puissance n-ième d'une fonction caractéristique.

b) μ est-elle indéfiniment divisible dans les cas suivants : i) $\mu = \delta_a$, $a \in \mathbb{R}$; ii) μ est la loi gaussienne $\mathcal{N}(m, \sigma^2)$ de moyenne m et de variance σ^2 ; iii) μ est la loi de Poisson $\mathcal{P}(\theta)$ de paramètre $\theta > 0$; iv) μ est la loi de Cauchy (la fonction caractéristique d'une loi de Cauchy est donnée par $\mathrm{e}^{-|u|}$, $u \in \mathbb{R}$) ?

c) Soit X de loi μ de Bernoulli $\mathcal{B}(p)$ sur $\{0, 1\}$ de paramètre $p \in]0, 1[$; l'objet de la question est de démontrer que μ n'est pas indéfiniment divisible. À cet effet, considérer au contraire l'existence de variables aléatoires indépendantes Y et Z de loi commune ν telles que la somme $Y + Z$ soit de loi μ. Si B est un intervalle ne contenant pas 0 et $\frac{1}{2}$, démontrer que

$$\mathbb{P}(Y + Z \in B + B) \, = \, 0$$

où $B + B = \{x + y \, ; x \in B, y \in B\}$, et en déduire que $\mathbb{P}(Y \in B) \, \mathbb{P}(Z \in B) = 0$. En tirer que Y ne peut prendre que les valeurs 0 et $\frac{1}{2}$, puis conclure à une impossibilité.

Exercice 8 (*Processus de Galton*[4]*-Watson*[5]). Soit X une variable aléatoire à valeurs dans $\mathbb{N} \cup \{0\}$; une population est modélisée par une famille $X_{k,n}$, $k, n \in \mathbb{N}$, de variables aléatoires indépendantes de même loi que X, la variable aléatoire $X_{k,n}$ représentant le nombre de descendants d'un individu k à la n-ième génération. En partant d'un individu $Z_0 = 1$, soit Z_n, $n \in \mathbb{N}$, la suite

4. Francis Galton, anthropologue, explorateur, géographe, inventeur, météorologue, écrivain, proto-généticien, psychométricien et statisticien britannique (1822–1911).

5. Henry William Watson, mathématicien anglais (1827–1903).

de variables aléatoires représentant le nombre d'individus de la population à la génération n définie à travers la relation de récurrence

$$Z_n = \sum_{k=1}^{Z_{n-1}} X_{k,n}, \quad n \in \mathbb{N}.$$

Par convention $\sum_{k=1}^{0} = 0$, de sorte que $\{Z_{n-1} = 0\} \subset \{Z_n = 0\}$, $n \in \mathbb{N}$. L'objectif de l'exercice est d'étudier la probabilité d'extinction

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \{Z_n = 0\}\right) = \lim_{n\to\infty} \mathbb{P}(Z_n = 0)$$

(comme réunion croissante d'événements). Il peut être supposé dans la suite que $\mathbb{P}(X = 0) \in]0, 1[$ car si $\mathbb{P}(X = 0) = 1$, alors $Z_n = 0$ presque sûrement et la probabilité d'extinction est égale à 1, et si $\mathbb{P}(X = 0) = 0$, alors $Z_n \geq 1$ presque sûrement et la probabilité d'extinction est nulle.

a) Si $G = G_X$ désigne la fonction génératrice de la loi de X, et $G_n = G_{Z_n}$ celle de Z_n, $n \in \mathbb{N}$, démontrer que $G_n = G_{n-1} \circ G$.

b) Si $u_n = \mathbb{P}(Z_n = 0)$, $n \in \mathbb{N}$, déduire de la question précédente que $u_n = G(u_{n-1})$, $n \in \mathbb{N}$.

c) Démontrer que la suite u_n, $n \in \mathbb{N}$, est convergente, et que $u_n \leq G(s)$, $n \in \mathbb{N}$, pour tout point fixe s de G $(G(s) = s)$.

d*) Il est supposé ici que X ne prend que trois valeurs, $0, 1, 2$, avec une probabilité strictement positive de prendre la valeur 2. Analyser la fonction génératrice G dans ce cas pour en déduire que si $\mathbb{E}(X) \leq 1$, alors $\lim_{n\to\infty} u_n = 1$, alors que si $\mathbb{E}(X) > 1$, la limite $\ell = \lim_{n\to\infty} u_n$ est l'unique point fixe de G sur $]0, 1[$.

Exercice 9 (*Inégalité de Kolmogorov*[6]). Soient $X_1, \dots, X_n$ des variables aléatoires réelles indépendantes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, centrées, de carré intégrable ; poser $S_k = X_1 + \cdots + X_k$, $k = 1, \dots, n$. L'objectif de l'exercice est de renforcer l'inégalité de Bienaymé-Tchebychev pour des variables aléatoires indépendantes, sous la forme : pour tout $t > 0$,

$$\mathbb{P}\big(\max_{1 \leq k \leq n} |S_k| \geq t\big) \leq \frac{1}{t^2} \sum_{k=1}^{n} \mathrm{Var}(X_k).$$

a) Soit $t > 0$ et soit $A_k = \{|S_k| \geq t\}$, $k = 1, \dots, n$, de sorte que

$$\big\{\max_{1 \leq k \leq n} |S_k| \geq t\big\} = \bigcup_{k=1}^{n} A_k.$$

Rappeler (de la Leçon 2) que les événements

$$B_k = A_k \cap \Big(\bigcup_{j<k} A_j\Big)^c, \quad k = 1, \dots, n,$$

sont disjoints et que leur réunion est égale à $\bigcup_{k=1}^{n} A_k$.

b) Démontrer que

$$\mathbb{E}\big(\mathbb{1}_{B_k} S_k (S_n - S_k)\big) = 0$$

pour tout $k = 1, \dots, n$.

c) Établir (à l'aide de la question précédente) que

$$\sum_{k=1}^{n} \mathbb{E}\Big(\mathbb{1}_{B_k}\big[S_n^2 - 2S_k(S_n - S_k)\big]\Big) \leq \mathbb{E}(S_n^2),$$

puis que

$$\sum_{k=1}^{n} \mathbb{E}\Big(\mathbb{1}_{B_k}\big[S_n^2 - 2S_k(S_n - S_k)\big]\Big) \geq \sum_{k=1}^{n} \mathbb{E}\big(\mathbb{1}_{B_k} S_k^2\big).$$

6. Andreï Kolmogorov, mathématicien russe et soviétique (1903–1987).

d) Prouver que, pour tout $t > 0$ et tout $k = 1, \ldots, n$,

$$\mathbb{P}(B_k) \,=\, \mathbb{P}\big(\mathbb{1}_{B_k}|S_k| \geq t\big) \,\leq\, \frac{1}{t^2}\,\mathbb{E}\big(\mathbb{1}_{B_k}S_k^2\big).$$

e) De ce qui précède, conclure à l'inégalité annoncée.