Leçon 19

La loi des grands nombres

La loi des grands nombres est un énoncé central du calcul des probabilités, qui en particulier fait asymptotiquement émerger le déterminisme au sein d'un modèle désordonné (aléatoire). Elle illustre une intuition naturelle et commune : lancer une pièce de monnaie un grand nombre de fois, la proportion ou moyenne « empirique » (rapport entre le nombre de pile et le nombre total de lancers) tend à s'approcher de la moyenne « théorique » de la pièce ($\frac{1}{2}$ si la pièce est équilibrée).

Sa formalisation, et la démonstration de l'énoncé le plus complet, ont toutefois mis du temps historiquement à s'affirmer.

Il existe des versions innombrables de la loi des grands nombres, et cette leçon ne concerne essentiellement que la forme la plus traditionnelle pour des sommes de variables aléatoires indépendantes équidistribuées, correspondant à une expérience répétée. Dans la suite, X_n, $n \in \mathbb{N}$, est donc une famille de variables aléatoires réelles sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, et pour chaque $n \geq 1$, $S_n = X_1 + \cdots + X_n$. Il est possible de considérer plus généralement des variables aléatoires à valeurs dans $\mathbb{R}^d$.

1 Loi faible des grands nombres

Le mot « faible » n'est pas à entendre de façon péjorative ; en fait l'énoncé qui forme ce paragraphe, s'il ne produit qu'une convergence en probabilité, assure aussi cette dernière sous des hypothèses plus flexibles. Il existe bien d'autres formes de la loi faible des grands nombres.

Proposition 1. *Soit X_n, $n \in \mathbb{N}$, une suite de variables aléatoires réelles de carré intégrable, deux à deux non corrélées, telle que $\lim_{n\to\infty} \frac{1}{n^2} \sum_{k=1}^n \mathrm{Var}(X_k) = 0$; alors*

$$\frac{1}{n}\big[S_n - \mathbb{E}(S_n)\big] \text{ converge en probabilité vers } 0.$$

Cette proposition est une conséquence immédiate de la définition de la convergence en probabilité et de l'inégalité de Bienaymé-Tchebychev puisque, pour tout $\varepsilon > 0$ et tout $n \geq 1$,

$$\mathbb{P}\left(\left|\frac{1}{n}\big[S_n - \mathbb{E}(S_n)\big]\right| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2 n^2}\sum_{k=1}^{n}\mathrm{Var}(X_k).$$

Si les variables aléatoires X_n, $n \in \mathbb{N}$, sont de même loi, de carré intégrable donc, la condition

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{k=1}^{n}\mathrm{Var}(X_k) = \lim_{n\to\infty}\frac{1}{n}\mathrm{Var}(X_1) = 0$$

est automatiquement satisfaite. La conclusion de la proposition exprime dans ce cas que

$$\frac{1}{n}\big[S_n - n\,\mathbb{E}(X_1)\big] \text{ converge en probabilité vers } 0.$$

Le choix – arbitraire – de X_1 est simplement là pour indiquer un représentant de la loi commune des X_n. Autrement dit

$$\frac{S_n}{n} \text{ converge en probabilité vers } \mathbb{E}(X_1).$$

La loi forte des grands nombres renforce ce résultat en une convergence presque sûre pour des suites indépendantes.

2 Loi forte des grands nombres

L'énoncé complet de la loi forte des grands nombres fournit une équivalence pour la convergence presque sûre de $\frac{S_n}{n}$, $n \in \mathbb{N}$, vers la valeur moyenne pour une suite de variables aléatoires indépendantes équidistribuées.

Théorème 2 (Loi des grands nombres)**.** *Soit X_n, $n \in \mathbb{N}$, une suite de variables aléatoires réelles, indépendantes de même loi ; alors si $\mathbb{E}(|X_1|) < \infty$,*

$$\frac{S_n}{n} \text{ converge presque sûrement vers } \mathbb{E}(X_1).$$

Réciproquement, si la suite $\frac{S_n}{n}$, $n \in \mathbb{N}$, converge presque sûrement, alors $\mathbb{E}(|X_1|) < \infty$.

La condition $\mathbb{E}(|X_1|) < \infty$ exprime à nouveau simplement que la loi commune des variables X_n est intégrable. Dans une formulation équivalente, si X est une variable aléatoire intégrable, et si X_n, $n \in \mathbb{N}$, sont des variables aléatoires indépendantes de même loi que X, la suite $\frac{S_n}{n}$, $n \in \mathbb{N}$, converge presque sûrement vers $\mathbb{E}(X)$. Cette présentation sera favorisée au fur et à mesure.

L'énoncé est identique pour des vecteurs aléatoires dans $\mathbb{C}$ ou $\mathbb{R}^d$.

La loi des grands nombres exprime donc que la « moyenne empirique » $\frac{S_n}{n}$ est proche, en un sens presque sûr, de la « moyenne théorique » $\mathbb{E}(X_1)$, ce qui est naturellement attendu dans une expérience de type pile ou face représentée par une suite X_n, $n \in \mathbb{N}$, de variables de Bernoulli $\mathcal{B}(p)$ sur $\{0, 1\}$, dans laquelle $\frac{S_n}{n}$ représente la proportion de pile (de 1) après n lancers.

Démonstration. Elle est partielle puisque la convergence presque sûre n'est établie que sous l'hypothèse plus forte $\mathbb{E}(X_1^4) < \infty$ qui autorise un usage simple du premier critère de Borel-Cantelli. Il est commode de recentrer le problème en posant $Y_n = X_n - \mathbb{E}(X_n)$, $n \in \mathbb{N}$, qui définit de la même façon une suite de variables aléatoires indépendantes de même loi avec $\mathbb{E}(Y_1^4) < \infty$. Comme

$$\frac{S_n}{n} - \mathbb{E}(X_1) = \frac{1}{n} \sum_{k=1}^{n} \big[X_k - \mathbb{E}(X_k) \big] = \frac{1}{n} \sum_{k=1}^{n} Y_k,$$

en posant $T_n = \sum_{k=1}^{n} Y_k$, $n \in \mathbb{N}$, il faut démontrer la convergence presque sûre de $\frac{T_n}{n}$ vers 0.

La démonstration va donc s'attacher à démontrer que, pour tout $\varepsilon > 0$,

$$\sum_{n\in\mathbb{N}} \mathbb{P}\left(\left|\frac{T_n}{n}\right| \geq \varepsilon\right) \,<\, \infty.$$

En vertu de l'inégalité de Markov, pour tout $n \geq 1$,

$$\mathbb{P}\left(\left|\frac{T_n}{n}\right| \geq \varepsilon\right) \,=\, \mathbb{P}\big(|T_n| \geq \varepsilon n\big) \,=\, \mathbb{P}\big(T_n^4 \geq \varepsilon^4 n^4\big) \,\leq\, \frac{1}{\varepsilon^4 n^4}\,\mathbb{E}(T_n^4).$$

La suite de la démonstration est basée sur le développement (théorème de Fubini-Tonelli !)

$$\left(\sum_{k=1}^{n} a_k\right)^4 \,=\, \sum_{k_1,k_2,k_3,k_4=1}^{n} a_{k_1}a_{k_2}a_{k_3}a_{k_4}$$

de réels $a_1,\ldots,a_n$. Appliqué à $a_k = Y_k(\omega)$ pour chaque ω, après intégration,

$$\mathbb{E}(T_n^4) \,=\, \sum_{k_1,k_2,k_3,k_4=1}^{n} \mathbb{E}(Y_{k_1}Y_{k_2}Y_{k_3}Y_{k_4}).$$

Dans cette somme, plusieurs cas sont à considérer suivant la nature des indices, en rappelant que les Y_k, $k = 1,\ldots,n$, sont indépendantes et centrées :

– si $k_1 = k_2 = k_3 = k_4 = k$,

$$\mathbb{E}(Y_{k_1}Y_{k_2}Y_{k_3}Y_{k_4}) \,=\, \mathbb{E}(Y_k^4),$$

et il y a n termes de ce type ;

– si (k_1,k_2,k_3,k_4) est composé de trois termes égaux à k et d'un autre terme $\ell \neq k$,

$$\mathbb{E}(Y_{k_1}Y_{k_2}Y_{k_3}Y_{k_4}) \,=\, \mathbb{E}(Y_k^3 Y_\ell) \,=\, \mathbb{E}(Y_k^3)\,\mathbb{E}(Y_\ell) \,=\, 0$$

par indépendance et centrage ;

– si (k_1,k_2,k_3,k_4) est composé de deux couples égaux k,ℓ avec $k \neq \ell$,

$$\mathbb{E}(Y_{k_1}Y_{k_2}Y_{k_3}Y_{k_4}) \,=\, \mathbb{E}(Y_k^2 Y_\ell^2) \,=\, \mathbb{E}(Y_k^2)\,\mathbb{E}(Y_\ell^2)$$

par indépendance. Un comptage précis indique qu'il y a $3n(n-1)$ termes de ce type, mais il suffit pour la suite d'observer simplement qu'il y a au plus $3n^2$ termes car il y a 3 possibilités de coupler les 4 entiers ;

– si (k_1, k_2, k_3, k_4) est composé de deux termes égaux à k, et de deux autres termes ℓ, m (k, ℓ, m différents),

$$\mathbb{E}(Y_{k_1} Y_{k_2} Y_{k_3} Y_{k_4}) \,=\, \mathbb{E}(Y_k^2 Y_\ell Y_m) \,=\, \mathbb{E}(Y_k^2)\,\mathbb{E}(Y_\ell)\,\mathbb{E}(Y_m) \,=\, 0$$

par indépendance et centrage ;

– si enfin tous les k_1, k_2, k_3, k_4 sont distincts,

$$\mathbb{E}(Y_{k_1} Y_{k_2} Y_{k_3} Y_{k_4}) \,=\, \mathbb{E}(Y_{k_1})\,\mathbb{E}(Y_{k_2})\,\mathbb{E}(Y_{k_3})\,\mathbb{E}(Y_{k_4}) \,=\, 0$$

par indépendance et centrage.

En conclusion de cette analyse, et puisque les Y_k, $k = 1, \ldots, n$, sont en outre équidistribuées,

$$\mathbb{E}(T_n^4) \,\leq\, n\,\mathbb{E}(Y_1^4) + 3n^2\big[\mathbb{E}(Y_1^2)\big]^2.$$

Il s'ensuit que

$$\sum_{n \in \mathbb{N}} \frac{1}{\varepsilon^4 n^4}\,\mathbb{E}(T_n^4) \,<\, \infty,$$

ce qui conclut la première partie de la démonstration, sous l'hypothèse plus forte de quatrième moment fini.

La réciproque est intéressante, et s'appuie sur le lemme de Borel-Cantelli dans sa partie indépendante. Si la suite $\frac{S_n}{n}$, $n \in \mathbb{N}$, converge presque sûrement (vers une certaine limite, aléatoire ou non), alors la suite $\frac{X_n}{n}$, $n \in \mathbb{N}$, converge presque sûrement vers 0 car, pour tout n (≥ 2),

$$\frac{X_n}{n} \,=\, \frac{S_n - S_{n-1}}{n} \,=\, \frac{S_n}{n} - \frac{n-1}{n}\,\frac{S_{n-1}}{n-1}\,.$$

Mais la suite $\frac{X_n}{n}$, $n \in \mathbb{N}$, est composée de variables aléatoires indépendantes, donc, par le deuxième critère de Borel-Cantelli, pour tout $\varepsilon > 0$, par exemple

$\varepsilon = 1$,

$$\sum_{n\in\mathbb{N}} \mathbb{P}\Big(\Big|\frac{X_n}{n}\Big| \geq 1\Big) \;<\; \infty.$$

De plus $\mathbb{P}\big(|\frac{X_n}{n}| \geq 1\big) = \mathbb{P}(|X_1| \geq n)$ puisque les X_n ont la même loi. Il ne reste plus alors qu'à faire usage de l'Exercice 3, Leçon 3 (qui repose sur une comparaison entre série et intégrale), pour conclure que $\mathbb{E}(|X_1|) < \infty$. C'est le résultat souhaité. Il pourra être observé que la conclusion précédente a en fait lieu dès que la suite $\frac{S_n}{n}$, $n \in \mathbb{N}$, est bornée presque sûrement. □

3 Méthode de Monte-Carlo

Si X est une variable aléatoire réelle sur $(\Omega, \mathcal{A}, \mathbb{P})$ de loi $\mathbb{P}_X$, et si $\phi : \mathbb{R} \to \mathbb{R}$ est borélienne telle que $\phi(X)$ soit intégrable, le théorème de transport exprime que

$$\mathbb{E}\big(\phi(X)\big) \;=\; \int_{\mathbb{R}} \phi \, d\,\mathbb{P}_X.$$

Si X est de loi uniforme $\mathcal{U}(0,1)$ sur l'intervalle $[0,1]$, le membre de droite n'est rien d'autre que

$$\int_{[0,1]} \phi \, d\lambda.$$

La méthode de Monte-Carlo est un moyen algorithmique pour approcher la valeur de cette intégrale à travers la loi des grands nombres. En effet, si X_n, $n \in \mathbb{N}$, est une suite de variables aléatoires indépendantes de même loi (uniforme sur $[0,1]$) que X, celle-ci appliquée à la suite $\phi(X_n)$, $n \in \mathbb{N}$, assure que, presque sûrement,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \phi(X_k) \;=\; \mathbb{E}\big(\phi(X)\big) \;=\; \int_{[0,1]} \phi \, d\lambda.$$

Autrement dit, la production d'un échantillon $(X_1, \ldots, X_n)$ de variables aléatoires indépendantes de même loi uniforme sur $[0,1]$ fournit un *estimateur* $\frac{1}{n}\sum_{k=1}^{n} \phi(X_k)$ de la valeur de l'intégrale $\int_{[0,1]} \phi \, d\lambda$. Dans la terminologie de la Leçon 21, cet estimateur est sans biais de l'espérance. Il peut être quantifié à l'aide du théorème central limite (Leçon 20).

Cette méthode est fréquemment utilisée dans les calculs numériques d'intégrales.

4 Complément : Grandes déviations

Pour une suite X_n, $n \in \mathbb{N}$, de variables aléatoires réelles indépendantes de même loi qu'une variable aléatoire X, les grandes déviations dans la loi des grands nombres examinent le taux de décroissance de la probabilité pour que $\frac{S_n}{n}$ s'écarte de sa valeur moyenne $m = \mathbb{E}(X)$ à travers l'étude de la suite

$$\mathbb{P}\bigg(\frac{S_n}{n} \in B\bigg), \quad n \in \mathbb{N},$$

où B est un borélien ne contenant pas m. Typiquement, $B =]-\infty, a)$ ou $B = (b, \infty[$, avec $a < m < b$ (les intervalles étant ouverts ou fermés).

Les inégalités exponentielles issues de la transformée de Laplace, déjà examinées en Leçons 11 et 13, sont avantageusement mises à profit dans ce but. Dans la suite, il est ainsi supposé que la loi commune des X_n, $n \in \mathbb{N}$, admet une transformée de Laplace définie sur tout $\mathbb{R}$, autrement dit

$$\mathbb{E}\big(e^{uX}\big) \,<\, \infty \quad \text{pour tout } \ u \in \mathbb{R}.$$

Il sera commode de formuler diverses expressions à l'aide du logarithmique

$$\Lambda(u) \,=\, \ln\big(\mathbb{E}(e^{uX})\big), \quad u \in \mathbb{R},$$

de cette transformée de Laplace. La fonction Λ est convexe sur $\mathbb{R}$ comme conséquence de l'inégalité de Hölder (Exercice 1, Leçon 5) et vérifie $\Lambda(0) = 0$.

Soit alors, par exemple, $b > m$; d'après l'inégalité de Markov, pour tout $u \geq 0$ et tout $n \in \mathbb{N}$,

$$\mathbb{P}\Big(\frac{S_n}{n} \in [b, \infty[\Big) \,=\, \mathbb{P}(S_n \geq bn) \,\leq\, \mathbb{P}\big(e^{uS_n} \geq e^{ubn}\big) \,\leq\, e^{-ubn}\,\mathbb{E}\big(e^{uS_n}\big).$$

Par indépendance est équidistribution, $\mathbb{E}(e^{uS_n}) = \big[\mathbb{E}(e^{uX})\big]^n$, de sorte que, en rappelant le logarithmique $\Lambda(u) = \ln(\mathbb{E}(e^{uX}))$ de la transformée de Laplace,

$$\mathbb{P}\bigg(\frac{S_n}{n} \in [b, \infty[\bigg) \,\leq\, e^{-ubn + n\Lambda(u)}.$$

Cette borne ayant lieu pour tout $u \geq 0$, elle en sera d'autant meilleure en prenant l'infimum sur toutes ces valeurs de u, autrement dit, après un changement de signe, l'exposant de l'exponentielle précédente sera représenté comme

$$-n \sup_{u \geq 0} \big[ub - \Lambda(u) \big].$$

L'inégalité de Jensen montre par ailleurs que

$$\Lambda(u) \;=\; \ln \mathbb{E}(e^{uX}) \;\geq\; u\, \mathbb{E}(X) \;=\; um$$

pour tout $u \in \mathbb{R}$. Comme $b > m$, il s'ensuit que

$$\sup_{u \geq 0} \big[ub - \Lambda(u) \big] \;=\; \sup_{u \in \mathbb{R}} \big[ub - \Lambda(u) \big]$$

(en effet, si $u < 0$, $ub - \Lambda(u) \leq u(b - m) < 0$ et $\sup_{u \geq 0}[ub - \Lambda(u)] \geq 0$).

La transformée de Legendre[1] Λ^* de la log-Laplace Λ est définie par

$$\Lambda^*(x) \;=\; \sup_{u \in \mathbb{R}} \big[ux - \Lambda(u) \big], \quad x \in \mathbb{R}.$$

Elle est à valeurs dans $[0, \infty]$. Il résulte ainsi de l'analyse précédente que, pour tout entier n, si $b > m$,

$$\mathbb{P}\bigg(\frac{S_n}{n} \in [b, \infty[\bigg) \;\leq\; e^{-n\Lambda^*(b)}.$$

L'objectif suivant va consister à mettre à profit cette borne pour examiner l'asymptotique de $\mathbb{P}\big(\frac{S_n}{n} \in F\big)$ pour une partie fermée quelconque F de $\mathbb{R}$ à travers la fonctionnelle

$$\Lambda^*(F) \;=\; \inf_{x \in F} \Lambda^*(x).$$

1. Adrien-Marie Legendre, mathématicien français (1752–1833).

Il n'y aura rien à démontrer si $\Lambda^*(F) = 0$, et dans le cas contraire nécessairement $m \notin F$ (car $\Lambda^*(m) = 0$). Si tel est le cas, par exemple $F \subset]m, \infty[$, et puisque F est fermé, il est possible de déterminer $b \in F$ tel que

$$\inf_{x \in F} |x - m| \;=\; |b - m|.$$

En particulier $F \subset [b, \infty[$ et $\Lambda^*(F) \geq \Lambda^*(b)$. En conséquence, pour tout $n \in \mathbb{N}$,

$$\mathbb{P}\left(\frac{S_n}{n} \in F\right) \;\leq\; e^{-n\Lambda^*(F)}$$

et en particulier

$$\limsup_{n \to \infty} \frac{1}{n} \ln \left(\mathbb{P}\left(\frac{S_n}{n} \in F \right) \right) \;\leq\; -\Lambda^*(F).$$

Les autres positions de F par rapport à m sont analysées de façon similaire pour fournir la partie majoration du théorème suivant.

La minoration pour les ouverts est un peu plus délicate et fait appel à un changement de mesure. La conjonction donne lieu au théorème suivant.

Théorème 3 (Théorème des grandes déviations)**.** *Soit X une variable aléatoire réelle sur $(\Omega, \mathcal{A}, \mathbb{P})$ telle que $\Lambda(u) = \mathbb{E}(e^{uX}) < \infty$ pour tout $u \in \mathbb{R}$, et soit X_n, $n \in \mathbb{N}$, une suite de variables aléatoires de même loi que X ; poser $S_n = X_1 + \cdots + X_n$, $n \geq 1$. Alors, pour Λ^* la transformée de Legendre de Λ,*

$$\limsup_{n \to \infty} \frac{1}{n} \ln \left(\mathbb{P}\left(\frac{S_n}{n} \in F \right) \right) \;\leq\; -\Lambda^*(F)$$

si F est un fermé de $\mathbb{R}$, et

$$\liminf_{n \to \infty} \frac{1}{n} \ln \left(\mathbb{P}\left(\frac{S_n}{n} \in O \right) \right) \;\geq\; -\Lambda^*(O)$$

si O est un ouvert de $\mathbb{R}$.

Cet énoncé est en fait valable sans hypothèse d'intégrabilité sur la loi de X. Il est essentiellement dû à H. Cramér [2].

L'importance d'un théorème de grandes déviations est le lien étroit entre une quantité probabiliste et une expression purement analytique (issue de la transformée de Laplace).

2. Harald Cramér, mathématicien et statisticien suédois (1893–1985).

Exercices

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

Exercice 1. Si $\varphi : \mathbb{R} \to \mathbb{R}$ est une fonction convexe, démontrer que pour tous réels $x_1, \ldots, x_n$,

$$\varphi\Big(\frac{1}{n}(x_1 + \cdots + x_n)\Big) \le \frac{1}{n}\big[\varphi(x_1) + \cdots + \varphi(x_n)\big].$$

Soit à présent X une variable aléatoire réelle sur $(\Omega, \mathcal{A}, \mathbb{P})$ telle que X et $\varphi(X)$ soient intégrables ; considérer une suite X_n, $n \in \mathbb{N}$, de variables aléatoires indépendantes de même loi que X. Que peut-on déduire de la loi des grands nombres ?

Exercice 2. Soit Y_k, $k \in \mathbb{N}$, une suite de variables aléatoires deux à deux non corrélées telles que, pour chaque $k \ge 1$, $\mathbb{P}(Y_k = \sqrt{k}\,) = \mathbb{P}(Y_k = -\sqrt{k}\,) = \frac{1}{2}$; poser $S_n = Y_1 + \cdots + Y_n$, $n \ge 1$.

a) Démontrer que $\frac{S_n}{n^{\frac{3}{2}}} \to 0$ en probabilité.

b) Démontrer que $\frac{S_{n^2}}{n^3} \to 0$ presque sûrement.

Exercice 3. Si Y est une variable aléatoire de loi $\mathcal{N}(0,1)$, calculer $\mathbb{E}(Y^4)$. Soit à présent X_k, $k \in \mathbb{N}$, une suite de variables aléatoires indépendantes de même loi normale $\mathcal{N}(m, \sigma^2)$ de moyenne m et de variance σ^2 ; démontrer que

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k^4 = m^4 + 6m^2\sigma^2 + 3\sigma^4 \qquad \text{presque sûrement.}$$

Exercice 4. Soit X_n, $n \in \mathbb{N}$, une suite de variables aléatoires indépendantes de même loi normale centrée réduite $\mathcal{N}(0,1)$, et soit $S_n = X_1 + \cdots + X_n$, $n \geq 1$; expliquer pourquoi $\frac{S_n}{n} \to 0$ presque sûrement.

a) Démontrer que, pour tout $n \geq 1$, $\frac{S_n}{\sqrt{n}}$ est de loi $\mathcal{N}(0,1)$.

b) Rappeler (Exercice 1, Leçon 14) pourquoi $\mathbb{P}(X_1 \geq t) \leq \frac{1}{2}\, e^{-\frac{1}{2}t^2}$ pour tout $t \geq 0$, puis $\mathbb{P}(|X_1| \geq t) \leq e^{-\frac{1}{2}t^2}$.

c) Utiliser la question précédent pour démontrer que pour tout $\alpha > \frac{1}{2}$, $\frac{S_n}{n^\alpha} \to 0$ presque sûrement.

d*) Proposer une suite strictement croissante a_n, $n \in \mathbb{N}$, de réels positifs telle que $\frac{a_n}{\sqrt{n}} \to \infty$, $\frac{a_n}{n^\alpha} \to 0$ pour tout $\alpha > \frac{1}{2}$ et $\frac{S_n}{a_n} \to 0$ presque sûrement

Exercice 5. Démontrer que si X_n, $n \in \mathbb{N}$, est une suite de variables aléatoires de carré intégrable sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, deux à deux non correlées, telle que $\sum_{n\in\mathbb{N}} \frac{\mathrm{Var}(X_n)}{n^2} < \infty$, alors

$$\frac{1}{n}\sum_{k=1}^{n} \big[X_k - \mathbb{E}(X_k)\big] \to 0 \quad \text{dans } \mathrm{L}^2.$$

(*Indication* : utiliser le lemme de Kronecker [3] indiquant que si u_n, $n \in \mathbb{N}$, une suite de nombres réels ou à valeurs dans un espace normé, telle que la série $\sum_{n\in\mathbb{N}} \frac{u_n}{n}$ est convergente, alors $\frac{1}{n}\sum_{k=1}^{n} u_k \to 0$.)

Exercice 6 (*Khinchine's*[4] *inequality.*) Let X_n, $n \in \mathbb{N}$, be a sequence of real-valued random variables on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$, independent with the same Bernoulli distribution $\mathbb{P}(X_1 = +1) = \mathbb{P}(X_1 = -1) = \frac{1}{2}$.

a) Show that for every $u \in \mathbb{R}$, $\mathbb{E}(e^{uX_1}) \leq e^{\frac{1}{2}u^2}$.

b) Let $a_1, \ldots, a_n$ be real numbers such that $\sum_{k=1}^{n} a_k^2 = 1$; set $S = \sum_{k=1}^{n} a_k X_k$.

---

3. Leopold Kronecker, mathématicien et logicien allemand (1823–1891).
4. Alexandre Khintchine, mathématicien russe et soviétique (1894–1959).

Prove that for every $u \in \mathbb{R}$, $\mathbb{E}(e^{uS}) \leq e^{\frac{1}{2}u^2}$. Deduce that for any $t \geq 0$,

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

c*) Use the preceding question to show that, for any real numbers $a_1, \ldots, a_n$, and any $p > 0$,

$$\left[\mathbb{E}\left(\left|\sum_{k=1}^{n} a_k X_k\right|^p\right)\right]^{1/p} \leq C_p \left(\sum_{k=1}^{n} a_k^2\right)^{1/2}$$

for some constant $C_p > 0$ only depending on p.

d) Set, for $n \geq 1$, $S_n = X_1 + \cdots + X_n$. Show that for every $\rho > 1$,

$$\lim_{n \to \infty} \frac{S_n}{\sqrt{n(\ln n)^\rho}} = 0 \qquad \text{almost surely.}$$