Leçon 20

Le théorème central limite

Cette leçon est une présentation sommaire et sans trop de détails d'un théorème fondamental de la théorie des probabilités, et de la mathématique en général (ainsi que son nom l'indique), *le théorème central limite.*

Son énoncé requiert un nouveau mode de convergence, la *convergence en loi*, qui est ici simplement esquissée.

1 Convergence en loi

Si $X : (\Omega, \mathcal{A}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est une variable aléatoire réelle, sa loi notée $\mathbb{P}_X$ est une mesure de probabilité sur les boréliens de $\mathbb{R}$ définie par

$$\mathbb{P}_X(B) = \mathbb{P}(X \in B), \quad B \in \mathcal{B}(\mathbb{R}).$$

La loi $\mathbb{P}_X$ peut être décrite et identifiée de diverses façons. Par exemple, par :

i) la fonction de répartition $F_X(t) = \mathbb{P}_X(\,]-\infty, t]) = \mathbb{P}(X \leq t)$, $t \in \mathbb{R}$;

ii) les intégrales $\int_{\mathbb{R}} \phi \, d\,\mathbb{P}_X = \mathbb{E}(\phi(X))$ pour toute fonction borélienne $\phi : \mathbb{R} \to \mathbb{R}$, positive ou bornée, ou même continue bornée ;

iii) la fonction caractéristique $\varphi_X(u) = \int_{\mathbb{R}} e^{iux} d\mathbb{P}_X = \mathbb{E}(e^{iuX})$, $u \in \mathbb{R}$.

La définition de convergence en loi reprend simplement ces descriptions pour la limite d'une suite de variables aléatoires.

Définition 1 (Convergence en loi)**.** *Une suite X_n, $n \in \mathbb{N}$, de variables aléatoires réelles définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ est dite converger en loi vers une variable aléatoire X si l'une des propriétés équivalentes suivantes est satisfaite :*

i) la suite des fonctions de répartition F_{X_n} des lois des variables aléatoires X_n, $n \in \mathbb{N}$, converge vers la fonction de répartition F_X de la loi de X en tout point de continuité pour F_X ;

ii) pour toute fonction $\phi : \mathbb{R} \to \mathbb{R}$ continue bornée, $\lim_{n\to\infty} \mathbb{E}(\phi(X_n)) = \mathbb{E}(\phi(X))$ *;*

iii) la suite des fonctions caractéristiques φ_{X_n} des lois des variables aléatoires X_n, $n \in \mathbb{N}$, converge vers la fonction caractéristique φ_X de la loi de X en tout point de $\mathbb{R}$.

Cette définition cache bien entendu un théorème démontrant l'équivalence entre les trois descriptions (l'équivalence avec iii) étant le point le plus délicat). Il convient d'observer la restriction dans i) aux points de continuité de F_X ; elle est tout à fait naturelle, par exemple pour une suite de variables aléatoires constantes $X_n = c + \frac{1}{n}$, $n \geq 1$, pour un $c \in \mathbb{R}$, pour laquelle il est attendu bien entendu que X_n converge en loi vers la variable constante $X = c$; mais comme $F_{X_n} = \mathbb{1}_{[c+\frac{1}{n},\infty[}$, $n \geq 1$, et $F_X = \mathbb{1}_{[c,\infty[}$, il n'y a pas convergence de $F_{X_n}(c) = 0$ vers $F_X(c) = 1$.

Il n'est pas difficile de se convaincre, par exemple sur la forme i), que si la suite X_n, $n \in \mathbb{N}$, converge en probabilité vers X, alors elle converge aussi en loi vers X. Mais cette implication cache une différence significative entre ces deux modes de convergence. Les convergences en probabilité, presque sûre ou dans L^p sont en effet des convergences de fonctions mesurables, la convergence en loi est une convergence de mesures. Par exemple, si toutes les variables X_n sont les mêmes, égales à Z par exemple, et si X est une variable aléatoire de même loi que Z mais indépendante de celle-ci, alors la suite X_n, $n \in \mathbb{N}$, converge automatiquement en loi vers X (car par exemple $F_{X_n} = F_Z = F_X$ pour tout n) mais, sauf cas dégénéré, ne peut converger en probabilité vers X. En effet, si par exemple la loi commune de Z et de X est une loi de Bernoulli $\mathcal{B}(p)$ sur $\{0, 1\}$ de paramètre de succès $p \in]0, 1[$, alors, pour tout $0 < \varepsilon \leq 1$ et tout $n \in \mathbb{N}$, par indépendance,

$$\begin{aligned}
\mathbb{P}\big(|X_n - X| \geq \varepsilon\big) &= \mathbb{P}\big(|Z - X| \geq \varepsilon\big) \\
&= \mathbb{P}(Z = 1, X = 0) + \mathbb{P}(Z = 0, X = 1) \\
&= 2p(1-p)
\end{aligned}$$

qui ne peut tendre vers 0 (sauf si $p = 0$ ou $p = 1$).

2 Théorème central limite

Le *théorème central limite* décrit les *fluctuations*, ou la vitesse de convergence dans la loi des grands nombres. Si, sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, X_n, $n \in \mathbb{N}$, est une suite de variables aléatoires réelles indépendantes équidistribuées, de même loi qu'une variable aléatoire X, et si $S_n = X_1 + \cdots + X_n$, $n \geq 1$, la loi des grands nombres exprime que si (et seulement si) $\mathbb{E}(|X|) < \infty$, alors

$$\frac{S_n}{n} \to \mathbb{E}(X) \quad \text{presque sûrement.}$$

Quelle est alors la vitesse de convergence de $\frac{S_n}{n} - \mathbb{E}(X)$ vers 0 ? Des études plus fines montrent que, sous des hypothèses de moments adaptées, la suite $n^\beta \frac{S_n}{n} = \frac{S_n}{n^{1-\beta}}$ converge toujours presque sûrement pour tout $\beta < \frac{1}{2}$ (Exercices 4 & 6, Leçon 19). Pour $\beta = \frac{1}{2}$, le mode de convergence change, et fait apparaître une limite gaussienne.

Théorème 2 (Théorème central limite)**.** *Soit, sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, X une variable aléatoire réelle et X_n, $n \in \mathbb{N}$, une suite de variables aléatoires indépendantes de même loi que X ; soit $S_n = X_1 + \cdots + X_n$, $n \geq 1$. Alors, si $0 < \mathbb{E}(X^2) < \infty$,*

$$\sqrt{n}\left[\frac{S_n}{n} - \mathbb{E}(X)\right] \to G \quad \textit{en loi}$$

où G est une variable aléatoire gaussienne de loi $\mathcal{N}(0, \sigma^2)$ avec $\sigma^2 = \mathrm{Var}(X)$.

Noter pour la formulation que

$$\sqrt{n}\left[\frac{S_n}{n} - \mathbb{E}(X)\right] = \frac{1}{\sqrt{n}}\sum_{k=1}^{n}\big[X_k - \mathbb{E}(X_k)\big].$$

Il est ainsi souvent commode de recentrer les variables. Il est aussi profitable de normaliser par la variance (écart-type plus précisément) afin d'avoir une convergence vers la loi standard $\mathcal{N}(0,1)$. La loi normale ayant une fonction de répartition continue, le théorème central limite exprime ainsi que, pour tout $t \in \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{P}\bigg(\frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{n} \big[X_k - \mathbb{E}(X_k) \big] \leq t \bigg) \;=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{1}{2}x^2} dx.$$

Le théorème central limite est une propriété de stabilité de la variance, et les lois gaussiennes sont un point fixe de l'énoncé : si les X_k, $k \geq 1$, sont indépendantes de même loi $\mathcal{N}(0,\sigma^2)$ (centrées donc pour simplifier), alors $\frac{S_n}{\sqrt{n}}$ a pour loi $\mathcal{N}(0,\sigma^2)$ pour tout $n \geq 1$. Le théorème central limite admet des versions multi-dimensionnelles avec pour limite les lois gaussiennes $\mathcal{N}(m,\Sigma)$ de même moyenne et variance que la loi commune des vecteurs aléatoires X_k. Il confirme le caractère universel de la loi normale comme attracteur de sommes de variables aléatoires indépendantes (équidistribuées), qualifiée ainsi parfois de « loi des erreurs ».

Démonstration. D'après ce qui précède, suite à une translation et une homothétie, il suffit de démontrer le théorème lorsque $\mathbb{E}(X) = 0$ et $\mathbb{E}(X^2) = 1$, et de démontrer donc que la suite $\frac{S_n}{\sqrt{n}}$, $n \in \mathbb{N}$, converge en loi vers une variable aléatoire de loi normale $\mathcal{N}(0,1)$. En acceptant la caractérisation de la convergence en loi par la transformée de Fourier, la démonstration est rapide. La fonction caractéristique de $\frac{S_n}{\sqrt{n}}$ vérifie, par indépendance et équidistribution,

$$\varphi_{\frac{S_n}{\sqrt{n}}}(u) \;=\; \mathbb{E}\big(e^{iu\frac{S_n}{\sqrt{n}}}\big) \;=\; \big[\mathbb{E}\big(e^{i\frac{u}{\sqrt{n}}X}\big)\big]^n \;=\; \big[\varphi_X\big(\tfrac{u}{\sqrt{n}}\big)\big]^n$$

pour tout $u \in \mathbb{R}$. Comme $\mathbb{E}(X) = 0$ et $\mathbb{E}(X^2) = 1$, φ_X est deux fois dérivable, et $\varphi_X(0) = 1$, $\varphi'_X(0) = 0$, $\varphi''_X(0) = -1$ (voir Leçon 10). Donc, d'après la formule de Taylor en 0,

$$\varphi_X(v) \;=\; 1 - \tfrac{1}{2}\,v^2 + v^2\varepsilon(v)$$

où $\varepsilon(v) \in \mathbb{C}$ tend vers 0 quand $v \to 0$. Ainsi, à $u \in \mathbb{R}$ fixé, quand $n \to \infty$,

$$\varphi_{\frac{S_n}{\sqrt{n}}}(u) \;=\; \Big(1 - \tfrac{1}{2n}\, u^2 + \tfrac{1}{n}\, u^2\, \varepsilon\big(\tfrac{u}{\sqrt{n}}\big)\Big)^n.$$

Il a été vu à l'Exercice 5, Leçon 3 (comme conséquence du théorème de convergence dominée), que pour tout $z \in \mathbb{C}$, $\lim_{n\to\infty} \big(1 + \frac{z}{n}\big)^n = e^z$. Il est aisé de constater que cette limite reste valide si z est remplacée par une suite z_n, $n \in \mathbb{N}$, qui converge vers z. En effet, d'après la formule du binôme et l'inégalité triangulaire,

$$\begin{aligned}\left|\Big(1 + \frac{z_n}{n}\Big)^n - \Big(1 + \frac{z}{n}\Big)^n\right| &\le \sum_{k=1}^{n} \binom{n}{k} \frac{1}{n^k}\, |z_n^k - z^k| \\ &\le \sum_{k=1}^{n} \frac{1}{k!}\, |z_n^k - z^k|\end{aligned}$$

car $\binom{n}{k} \le \frac{n^k}{k!}$. Or, si $C = \max(\sup_{n\in\mathbb{N}} |z_n|, |z|)$, qui est fini,

$$|z_n^k - z^k| \;=\; |z_n - z|\, |z_n^{k-1} + z_n^{k-2} z + \cdots + z_n z^{k-2} + z^{k-1}| \;\le\; k\, C^{k-1} |z_n - z|$$

pour tout $k \ge 1$, de sorte que

$$\left|\Big(1 + \frac{z_n}{n}\Big)^n - \Big(1 + \frac{z}{n}\Big)^n\right| \le |z_n - z| \sum_{k=1}^{\infty} \frac{C^{k-1}}{(k-1)!} \;=\; |z_n - z|\, e^C$$

qui tend vers 0 quand n tend vers l'infini. L'affirmation s'ensuit.

Il ne reste plus qu'à appliquer cette observation, pour $u \in \mathbb{R}$ fixé, à $z_n = -\frac{1}{2}u^2 + \varepsilon(\frac{u}{\sqrt{n}})$, $n \in \mathbb{N}$, qui converge vers $z = -\frac{1}{2}u^2$, afin de conclure que

$$\lim_{n\to\infty} \varphi_{\frac{S_n}{\sqrt{n}}}(u) \;=\; e^{-\frac{1}{2}u^2}.$$

Le membre de droite décrivant la fonction caractéristique d'une variable aléatoire de loi normale $\mathcal{N}(0,1)$, le théorème central limite en découle. $\square$

Il est important de noter que la convergence en loi dans le théorème central limite ne peut être renforcée en une convergence en probabilité (et a fortiori presque sûre ou dans L^p, $p > 0$). En effet, considérer le cas où $\mathbb{E}(X) = 0$ et $\mathbb{E}(X^2) = 1$ de sorte que $\frac{S_n}{\sqrt{n}}$ converge alors en loi vers une variable gaussienne G de loi $\mathcal{N}(0,1)$. Pour tout $n \geq 1$,

$$\sqrt{2}\,\frac{S_{2n}}{\sqrt{2n}} - \frac{S_n}{\sqrt{n}} \;=\; \frac{X_{n+1} + \cdots + X_{2n}}{\sqrt{n}}\,.$$

S'il y avait convergence en probabilité de la suite $\frac{S_n}{\sqrt{n}}$, $n \geq 1$, vers G, le membre de gauche tendrait, en vertu des propriétés de stabilité de la convergence en probabilité (Exercice 3, Leçon 18), vers la variable $(\sqrt{2}-1)G$. Le membre de droite quant à lui a même loi que $\frac{S_n}{\sqrt{n}}$, et converge donc (en loi) vers G. Comme G est gaussienne, $(\sqrt{2}-1)G$ ne peut avoir la même loi que G, et une contradiction apparaît.

Le premier théorème central limite a été mis en évidence par A. de Moivre [1] (à qui l'on doit la première apparition de la loi normale). Il concernait des variables de loi de Bernoulli, sur $\{0,1\}$, de paramètre $p \in]0,1[$. Comme dans ce cas $\mathbb{E}(X) = p$ et $\mathrm{Var}(X) = p(1-p)$, l'énoncé prend la forme

$$\lim_{n\to\infty} \mathbb{P}\bigg(\frac{1}{\sqrt{p(1-p)n}}\sum_{k=1}^{n}[X_k - p] \leq t\bigg) \;=\; \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-\frac{1}{2}x^2}dx$$

pour tout $t \in \mathbb{R}$, ou encore

$$\lim_{n\to\infty} \mathbb{P}\bigg(\sum_{k=1}^{n} X_k \leq pn + t\sqrt{p(1-p)n}\bigg) \;=\; \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-\frac{1}{2}x^2}dx.$$

Or $\sum_{k=1}^{n} X_k$ suit la loi binomiale $\mathcal{B}(n,p)$. Ainsi la limite précédente s'exprime explicitement comme

$$\lim_{n\to\infty} \sum_{k\in I_n} \binom{n}{k} p^k (1-p)^{n-k} \;=\; \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-\frac{1}{2}x^2}dx$$

1. Abraham de Moivre, mathématicien français (1667–1754).

où I_n est l'ensemble des entiers k de $\{0, 1, \ldots, n\}$ tels que $k \leq pn + t\sqrt{p(1-p)n}$. La démonstration s'appuie alors sur l'équivalent

$$\binom{n}{k} p^k (1-p)^{n-k} \sim \frac{1}{\sqrt{2\pi np(1-p)}} \, e^{-\frac{(k-np)^2}{2p(1-p)n}}$$

issu de la formule de Stirling, justifiant (a posteriori) l'apparition du facteur $\sqrt{2\pi}$ de la loi normale.

3 Limite de Poisson

La convergence en loi met aussi en évidence des limites vers la loi de Poisson. L'exemple le plus simple a trait à une suite de variables aléatoires de loi binomiale dont le paramètre de succès tend vers 0, conférant à la loi de Poisson sa caractéristique de « loi des événements rares ». Comme une loi binomiale est la loi de la somme de variables indépendantes de Bernoulli, cette limite est à comparer bien entendu à la limite gaussienne quand le paramètre de succès est fixé.

Proposition 3. *Soit X_n, $n \in \mathbb{N}$, une suite de variables aléatoires de loi binomiale $\mathcal{B}(n, p_n)$ où p_n vérifie $\lim_{n\to\infty} np_n = \theta > 0$; alors la suite X_n, $n \in \mathbb{N}$, converge en loi vers une variable aléatoire X de loi de Poisson $\mathcal{P}(\theta)$.*

Démonstration. L'argument de fonction caractéristique peut s'appliquer de la même façon que dans le paragraphe précédent. Une approche naïve est néanmoins instructive pour une meilleure compréhension de la convergence en loi. Pour tout $n \in \mathbb{N}$, et tout $k = 0, 1, \ldots, n$, par définition de la probabilité binomiale,

$$\mathbb{P}(X_n = k) \;=\; \binom{n}{k} p_n^k \, (1-p_n)^{n-k}.$$

D'après l'expression du coefficient binomial, et après une petite réorganisation,

$$\mathbb{P}(X_n = k) \;=\; \frac{1}{k!}\,\frac{n!}{(n-k)!}\left(\frac{p_n}{1-p_n}\right)^k (1-p_n)^n.$$

À k fixé, il convient d'examiner chacun des termes lorsque $n \to \infty$. Par définition de la factorielle, il est clair que $\frac{n!}{(n-k)!} \sim n^k$. Aussi, $\frac{p_n}{1-p_n} \sim \frac{\theta}{n}$ d'après l'hypothèse $\lim_{n\to\infty} np_n = \theta$. Ainsi

$$\lim_{n\to\infty} \frac{n!}{(n-k)!}\left(\frac{p_n}{1-p_n}\right)^k \;=\; \theta^k.$$

Enfin $\lim_{n\to\infty}(1-p_n)^n = e^{-\theta}$ puisqu'en passant au logarithme $n\ln(1-p_n) \sim -\theta$.

En résumé, pour tout entier $k \in \mathbb{N}\cup\{0\}$,

$$\lim_{n\to\infty} \mathbb{P}(X_n = k) \;=\; e^{-\theta}\,\frac{\theta^k}{k!}$$

c'est-à-dire la probabilité de Poisson sur l'entier k. Cette conclusion confirme que pour n grand, la probabilité binomiale $\mathbb{P}(X_n = k)$ est proche de $\mathbb{P}(X = k)$ où X suit la loi de Poisson $\mathcal{P}(\theta)$, et donc la convergence en loi de X_n vers X. Ce n'est toutefois pas exactement l'une des formulations équivalentes de la convergence en loi. Néanmoins, si $t \in \mathbb{R}_+$ est de continuité pour F_X, c'est-à-dire $t \geq 0$ n'est pas un entier (0 compris), il existe $k \in \mathbb{N}\cup\{0\}$ tel que $k < t < k+1$. Alors $F_{X_n}(t) = \mathbb{P}(X_n \leq k)$ et $F_X(t) = \mathbb{P}(X \leq k)$, et d'après ce qui précède

$$\lim_{n\to\infty} F_{X_n}(t) \;=\; \lim_{n\to\infty}\sum_{\ell=0}^{k}\mathbb{P}(X_n = \ell) \;=\; \sum_{\ell=0}^{k}\mathbb{P}(X = \ell) \;=\; F_X(t).$$

Il y a donc convergence des fonctions de répartition en tout point de continuité de la fonction de répartition limite, assurant la convergence en loi. La proposition est démontrée. □

L'Exercice 4 fournit une vitesse de convergence en variation totale dans cette proposition.

4 Valeurs extrêmes

Le théorème central limite peut donner l'impression que le comportement moyen d'une somme de variables aléatoires indépendantes de même loi est toujours normal (gaussien), et que toutes les variables de la somme jouent à peu près le même rôle.

Cette idée peut conduire à des erreurs d'évaluation, notamment en théorie des risques, lorsque l'une des variables, ou un petit nombre d'entre elles, est significativement plus grande que toutes les autres. La théorie des valeurs extrêmes étudie ce type de situation, et quelques exemples d'illustration sont brièvement présentés ici.

Dans ce qui suit, X_n, $n \in \mathbb{N}$, désigne une suite de variables aléatoires réelles indépendantes équidistribuées, définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. La fonction de répartition de la loi commune sera désignée plus simplement par F.

Le maximum $M_n = \max(X_1, \ldots, X_n)$, $n \geq 1$, d'un échantillon $(X_1, \ldots, X_n)$ est une première statistique d'intérêt. Par indépendance et équidistribution, la fonction de répartition F_{M_n} de M_n est, pour tout $t \in \mathbb{R}$,

$$F_{M_n}(t) \, = \, \mathbb{P}\big(\max(X_1, \ldots, X_n) \leq t \big) \, = \, \prod_{k=1}^{n} \mathbb{P}(X_k \leq t) \, = \, F(t)^n.$$

Suivant la nature de F, diverses conclusions peuvent déjà être tirées. Par exemple, si la loi commune des X_n est uniforme $\mathcal{U}(0,1)$ sur l'intervalle $[0,1]$, $F(t) = t$ si $t \in [0,1]$, et vaut 0 avant $t = 0$ et 1 après $t = 1$, de sorte que

$$\lim_{n \to \infty} F_{M_n}(t) \, = \, \begin{cases} 0 & \text{si} \quad t < 1, \\ 1 & \text{si} \quad t \geq 1. \end{cases}$$

Il en résulte que la fonction de répartition de M_n converge (en tout point) vers la fonction de répartition de la masse de Dirac au point 1 (loi de la variable constante égale à 1).

Il est possible d'examiner de la même manière $m_n = \min(X_1, \ldots, X_n)$.

Dans d'autres situations, il convient d'effectuer une translation et une homothétie sur M_n, ou m_n, de la forme $\frac{1}{a_n}[M_n - b_n]$, $a_n > 0$, $b_n \in \mathbb{R}$, afin de voir apparaître des limites dans la fonction de répartition. Par exemple, la suite $n\,m_n$, $n \in \mathbb{N}$, converge en loi vers une répartition exponentielle $\mathcal{E}(1)$ de paramètre 1. En effet, pour tout $n \geq 1$ et tout $t \in [0, n]$,

$$\begin{aligned}
\mathbb{P}(n\,m_n > t) &= \mathbb{P}\Big(\min(X_1, \ldots, X_n) > \frac{t}{n}\Big) \\
&= \prod_{k=1}^{n} \mathbb{P}\Big(X_k > \frac{t}{n}\Big) \\
&= \Big[1 - F\Big(\frac{t}{n}\Big)\Big]^n \\
&= \Big(1 - \frac{t}{n}\Big)^n.
\end{aligned}$$

À $t > 0$ fixé, quand $n \to \infty$, la condition $t \in [0, n]$ est réalisée pour n assez grand, et le membre de droite converge vers e^{-t}, justifiant la convergence en loi de $n\,m_n$, $n \in \mathbb{N}$, vers une variable aléatoire de loi exponentielle $\mathcal{E}(1)$.

Supposer à présent que la loi commune des X_n, $n \in \mathbb{N}$, est exponentielle $\mathcal{E}(1)$ de paramètre 1. La convergence en loi de la suite de variables aléatoires $M_n - \ln(n)$, $n \in \mathbb{N}$, donne lieu à une conclusion intéressante. Suivant le même schéma que précédemment, pour tout $n \geq 1$ et tout $t \in \mathbb{R}$,

$$\begin{aligned}
\mathbb{P}\big(M_n - \ln(n) \leq t\big) &= \mathbb{P}\big(\max(X_1, \ldots, X_n) \leq t + \ln(n)\big) \\
&= F(t + \ln(n))^n.
\end{aligned}$$

Si $t + \ln n \geq 0$ (qui est réalisé pour tout n assez grand), $F(t + \ln(n)) = 1 - e^{-t-\ln(n)}$, et donc

$$\mathbb{P}\big(M_n - \ln(n) \leq t\big) = \Big(1 - \frac{e^{-t}}{n}\Big)^n.$$

Le membre de droite converge cette fois vers $e^{-e^{-t}}$, c'est-à-dire la fonction de répartition de la loi de Gumbel (Exercice 5, Leçon 7).

5 Complément : Théorème de Berry-Esseen

Le théorème de Berry[2]-Esseen[3] fournit un contrôle sur la vitesse de convergence (uniforme) des fonctions de répartition vers la fonction de répartition normale dans le théorème central limite. C'est un outil important dans la mesure des approximations. Pour plus de commodité, il est énoncé pour des variables centrées.

Théorème 4. *Soit X une variable aléatoire réelle de moyenne nulle $\mathbb{E}(X) = 0$ et de variance $\mathrm{Var}(X) = \mathbb{E}(X^2) = \sigma^2 > 0$, et admettant un moment absolu d'ordre 3, $\mathbb{E}(|X|^3) < \infty$, et soit X_n, $n \in \mathbb{N}$, une suite de variables aléatoires indépendantes de même loi que X ; soit $S_n = X_1 + \cdots + X_n$, $n \geq 1$. Alors, pour tout $n \geq 1$,*

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P}\left(\frac{S_n}{\sigma\sqrt{n}} \leq t \right) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{1}{2}x^2} dx \right| \leq \frac{C\,\mathbb{E}(|X|^3)}{\sigma^3\sqrt{n}}$$

où $C > 0$ est une constante absolue.

Il existe de nombreuses versions plus générales de ce résultat. Le calcul de précision autour de la meilleure constante $0,4 < C < 0,48$ est toujours d'actualité. La démonstration originale du théorème s'appuie sur des arguments de transformation de Fourier.

2. Andrew Campbell Berry, mathématicien américain (1906–1998).
3. Carl-Gustav Esseen, mathématicien suédois (1918–2001).

Exercices

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

Exercice 1. Soient U et V deux variables aléatoires indépendantes de même loi normale centrée réduite $\mathcal{N}(0,1)$ sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$; soit la suite de variables aléatoires X_n, $n \in \mathbb{N}$, définie par $X_n = U$ si n est pair et $X_n = V$ si n est impair.

a) Que dire de la convergence de la suite X_n, $n \in \mathbb{N}$, presque sûrement, en probabilité, dans L^1 ?

b) Que dire de la convergence des suites F_{X_n}, $n \in \mathbb{N}$ (fonctions de répartition) et φ_{X_n}, $n \in \mathbb{N}$ (fonctions caractéristiques) ?

Exercice 2. Soit α_n, $n \in \mathbb{N}$, un suite de nombres appartenant à $[0,1]$; à cette suite est associée une suite X_n, $n \in \mathbb{N}$, de variables aléatoires indépendantes sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ dont les lois vérifient

$$F_{X_n}(t) \;=\; \mathbb{P}(X_n \le t) \;=\; \begin{cases} 0 & \text{si} \quad t < 0, \\ \alpha_n + (1-\alpha_n)\, t^n & \text{si} \quad t \in [0,1[, \\ 1 & \text{si} \quad t \ge 0. \end{cases}$$

a) À quelles conditions sur la suite α_n, $n \in \mathbb{N}$, la suite X_n, $n \in \mathbb{N}$, converge-t-elle en loi ?

b) Si $\lim_{n\to\infty} \alpha_n = 0$ ou 1, examiner la convergence en probabilité ou presque sûre.

Exercice 3. Soit X_θ une variable aléatoire suivant une loi de Poisson $\mathcal{P}(\theta)$ de paramètre $\theta > 0$; poser $Y_\theta = \frac{1}{\sqrt{\theta}}(X_\theta - \theta)$. Démontrer que Y_θ converge en loi quand $\theta \to \infty$ et déterminer sa limite. En déduire que

$$\lim_{n\to\infty} e^{-n} \sum_{k=0}^{n} \frac{n^k}{k!} \,=\, \frac{1}{2}\,.$$

Exercice 4*. Si μ et ν sont deux mesures de probabilité sur un espace mesurable $(E, \mathcal{B})$, la distance en variation totale (Exercice 15, Leçon 3) est rappelée comme étant définie par

$$d_{\mathrm{TV}}(\mu,\nu) \,=\, \sup_{B\in\mathcal{B}} \big|\mu(B) - \nu(B)\big|.$$

Vérifier que si X et Y sont deux variables aléatoires sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $(E, \mathcal{B})$ de lois respectives μ et ν,

$$d_{\mathrm{TV}}(\mu,\nu) \,\leq\, \mathbb{P}(X \neq Y).$$

Soient à présent Y et T deux variables aléatoires indépendantes sur $(\Omega, \mathcal{A}, \mathbb{P})$ telles que Y suit une loi de Poisson $\mathcal{P}(p)$ de paramètre $p \in]0,1[$ et T une loi de Bernoulli $\mathcal{B}(q)$ sur $\{0,1\}$ de paramètre de succès $q = 1-(1-p)e^p$ (observer que $(1-p)e^p \in]0,1[$).

a) Calculer la loi de $X = 1 - \mathbb{1}_{\{Y=T=0\}}$ et montrer que

$$\mathbb{P}(X \neq Y) \,=\, p(1-e^{-p}) \,\leq\, p^2.$$

b) Soient X_k, $k = 1,\ldots,n$, des variables aléatoires indépendantes de loi de Bernoulli $\mathcal{B}(p_k)$ de paramètres respectifs $p_k \in]0,1[$, $k = 1,\ldots,n$, et soit $S_n = X_1 + \cdots + X_n$. Démontrer qu'il existe une variable aléatoire Z de loi de Poisson $\mathcal{P}(\theta)$ de paramètre $\theta = p_1 + \cdots + p_n$ telle que

$$d_{\mathrm{TV}}(\mathbb{P}_{S_n}, \mathbb{P}_Z) \,\leq\, p_1^2 + \cdots + p_n^2.$$

c) Application. Soit Z une variable aléatoire de loi de Poisson $\mathcal{P}(\theta)$, $\theta > 0$, et soit Z_n, $n \in \mathbb{N}$, $n > \theta$, une suite de variables aléatoires de loi binomiale $\mathcal{B}(n, \frac{\theta}{n})$. Démontrer que pour tout entier $k \geq 0$, $\lim_{n\to\infty} \mathbb{P}(Z_n = k) = \mathbb{P}(Z = k)$. En désignant par ν la mesure de Poisson $\mathcal{P}(\theta)$ et par μ_n la mesure binomiale $\mathcal{B}(n, \frac{\theta}{n})$, déduire de la première partie de l'exercice que

$$d_{\mathrm{TV}}(\mu_n, \nu) \leq \frac{\theta^2}{n}.$$