Probabilités: Correction du DM2

Noé Blassel

Décembre 2023

Propriétés de la loi de Cauchy

On rappelle que la densité de probabilité de la loi de Cauchy standard $\mathcal{C}(1)$ est donnée par la fonction

$$p(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}, \qquad x \in \mathbb{R}. \tag{1}$$

Question 1

Soient G_1 et G_2 des variables aléatoires réelles indépendantes de même loi normale centrée réduite $\mathcal{N}(0,1)$. Montrer que G_1/G_2 est une variable aléatoire de loi $\mathcal{C}(1)$. En déduire la loi de $1/X$, où X est une variable aléatoire de loi $\mathcal{C}(1)$.

Solution. On va appliquer la méthode de la fonction muette. Soit $f : \mathbb{R}^2 \to \mathbb{R}$ mesurable et bornée. Les variables G_1 et G_2 étant indépendantes, leur densité jointe est donnée par le produit de leur densité respectives:

$$p_{(G_1,G_2)}(x,y) = \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\right)\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\right) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}. \tag{2}$$

La fonction $(x,y) \mapsto f(x/y)$ étant mesurable et bornée sur $\mathbb{R} \times \mathbb{R}^*$, et comme $G_2 \neq 0$ presque sûrement, le théorème de la fonction muette assure que

$$\mathbb{E}[f(G_1/G_2)] = \int_{\mathbb{R}\times\mathbb{R}^*} f(x/y) p_{(G_1,G_2)}(x,y)\, \mathrm{d}x\mathrm{d}y = \frac{1}{2\pi} \int_{\mathbb{R}^*\times\mathbb{R}} f(y/x) e^{-\frac{y^2+x^2}{2}}\, \mathrm{d}x\mathrm{d}y. \tag{3}$$

Dans la dernière égalité, on a échangé les noms des variables d'intégrations. On va effectuer un changement de variables en coordonnées polaires

$$(r,\theta) = \varphi(x,y) = \left(\sqrt{x^2+y^2}, 2\arctan\left(\frac{y}{x+\sqrt{x^2+y^2}}\right)\right),$$

qui est un C^1-difféomorphisme de l'ouvert $\mathcal{O} = \mathbb{R}^2 \setminus \{(x,0),\, x<0\}$ vers l'ouvert $\mathcal{O}' = \mathbb{R}_+^* \times]-\pi,\pi[$, d'inverse

$$\varphi^{-1}(r,\theta) = (r\cos\theta, r\sin\theta) = (x(r,\theta), y(r,\theta)).$$

Notons que comme

$$\mathbb{P}((G_1, G_2) \in \{(x,0),\, x<0\}) \leq \mathbb{P}(G_2 = 0) = 0,$$

nous pouvons sans crainte remplacer le domaine d'intégration dans (2) par $\mathcal{O}$. Le déterminant Jacobien de φ^{-1} est donné par

$$\mathrm{Jac}_{\varphi^{-1}}(r,\theta) = r. \tag{4}$$

De plus, on a clairement $y(r,\theta)/x(r,\theta) = \tan(\theta)$. La formule du changement de variable donne donc

$$\begin{aligned}
\mathbb{E}[f(G_1/G_2)] &= \frac{1}{2\pi}\int_{\mathcal{O}} f(y/x) e^{-\frac{y^2+x^2}{2}}\,\mathrm{d}x\mathrm{d}y \\
&= \frac{1}{2\pi}\int_{\mathcal{O}'} f(\tan\theta) e^{-\frac{r^2}{2}} r\,\mathrm{d}r\mathrm{d}\theta \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\int_0^{\infty} f(\tan\theta) e^{-\frac{r^2}{2}} r\,\mathrm{d}r\mathrm{d}\theta \\
&= \frac{1}{2\pi}\left(\int_{-\pi}^{\pi} f(\tan\theta)\,\mathrm{d}\theta\right)\left(\int_0^{\infty} re^{-\frac{r^2}{2}}\,\mathrm{d}r\right) \\
&= \frac{1}{2\pi}\left(\int_{-\pi}^{\pi} f(\tan\theta)\,\mathrm{d}\theta\right)\left[-e^{-\frac{r^2}{2}}\right]_0^{\infty} \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\tan\theta)\,\mathrm{d}\theta.
\end{aligned}$$

Pour passer de la troisième à la quatrième ligne, on a utilisé le théorème de Fubini, $f \circ \tan$ étant bornée donc intégrable sur $]-\pi,\pi[$. La fonction $f\circ\tan$ étant de surcroît π-périodique,

$$\int_{-\pi}^{\pi} f(\tan\theta)\,\mathrm{d}\theta = 2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\tan\theta)\,\mathrm{d}\theta,$$

d'où

$$\mathbb{E}[f(G_1/G_2)] = \frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} f(\tan\theta)\,\mathrm{d}\theta.$$

En utilisant de nouveau un changement de variable

$$u = \tan\theta \implies \mathrm{d}u = (1+\tan^2\theta)\mathrm{d}\theta \implies \mathrm{d}\theta = \frac{\mathrm{d}u}{1+u^2},$$

qui est un C^1-difféomorphisme de $]-\frac{\pi}{2},\frac{\pi}{2}[$ sur $\mathbb{R}$, on obtient

$$\mathbb{E}[f(G_1/G_2)] = \frac{1}{\pi}\int_{-\infty}^{\infty} f(u)\frac{\mathrm{d}u}{1+u^2} = \int_{-\infty}^{\infty} f(u)p(u)\,\mathrm{d}u.$$

On en déduit que la variable aléatoire G_1/G_2 admet la densité p: c'est donc une variable de Cauchy de loi $\mathcal{C}(1)$.

Si à présent X est une variable de Cauchy de loi $\mathcal{C}(1)$, elle a la même loi que le rapport G_1/G_2 entre deux Gaussiennes indépendantes de loi $\mathcal{N}(0,1)$. Il s'ensuit que $1/X$ à la même loi que G_2/G_1. Or, cette dernière est encore le rapport de deux Gaussiennes indépendantes. Par le calcul qui précède, $1/X$ suit donc une loi de Cauchy $\mathcal{C}(1)$. □

Question 2

On considère la densité de probabilités

$$q(x) = \frac{1}{2}e^{-|x|}, \qquad x \in \mathbb{R}. \tag{5}$$

Vérifier que q est une densité de probabilités. Étant donné une variable aléatoire Z de densité q, montrer que sa fonction caractéristique est donnée par

$$\mathbb{E}[e^{i\xi Z}] = \frac{1}{1+\xi^2}, \quad \xi \in \mathbb{R}. \tag{6}$$

On dit qu'une telle variable suit une loi de Laplace.

Solution. On vérifie facilement que l'intégrale de q sur $\mathbb{R}$ vaut 1. Comme c'est une fonction mesurable positive, c'est une densité de probabilités. Calculons

$$\begin{aligned}
\mathbb{E}[e^{i\xi Z}] &= \int_{\mathbb{R}} e^{-i\xi x} q(x)\,\mathrm{d}x \\
&= \frac{1}{2}\int_{\mathbb{R}} e^{-i\xi x} e^{-|x|} \\
&= \frac{1}{2}\int_{\mathbb{R}} (\cos(\xi x) + i\sin(\xi x)) e^{-|x|}\,\mathrm{d}x.
\end{aligned}$$

Les applications $x \mapsto i\sin(\xi x))e^{-|x|}$ et $x \mapsto \cos(\xi x)e^{-|x|}$ sont respectivement des fonctions impaires et paires de x, et sont toutes deux absolument bornées par $x \mapsto e^{-|x|}$, qui est intégrable sur $\mathbb{R}$. On en déduit, en utilisant l'imparité,

$$\int_{\mathbb{R}} i\sin(\xi x)e^{-|x|}\,\mathrm{d}x = 0,$$

puis

$$\mathbb{E}[e^{i\xi Z}] = \frac{1}{2}\int_{\mathbb{R}} \cos(\xi x)e^{-|x|}\,\mathrm{d}x = \int_0^\infty \cos(\xi x)e^{-x}\,\mathrm{d}x,$$

en utilisant cette fois la parité. En effectuant une intégration par parties avec $u = \cos(\xi x)$, $u' = -\xi\sin(\xi x)$, $v' = e^{-x} = -v$, on a

$$\mathbb{E}[e^{i\xi Z}] = \left[\cos(\xi x)\left(-e^{-x}\right)\right]_0^\infty - \xi\int_0^\infty \sin(\xi x)e^{-x}\,\mathrm{d}x = 1 - \xi\int_0^\infty \sin(\xi x)e^{-x}\,\mathrm{d}x.$$

Par une nouvelle intégration par parties, avec cette fois $u = \sin(\xi x)$, $u' = \xi \cos(\xi x)$, on obtient

$$\mathbb{E}[e^{i\xi Z}] = 1-\xi \left(\left[\sin(\xi x)(-e^{-x})\right]_0^\infty + \xi \int_0^\infty \cos(\xi x) e^{-x}\, \mathrm{d}x \right) = 1-\xi^2 \int_0^\infty \cos(\xi x) e^{-x}\, \mathrm{d}x.$$

En utilisant à nouveau

$$\mathbb{E}[e^{i\xi Z}] = \int_0^\infty \cos(\xi x) e^{-x}\, \mathrm{d}x,$$

on obtient

$$\mathbb{E}[e^{i\xi Z}] = 1 - \xi^2 \mathbb{E}[e^{i\xi Z}] \implies \mathbb{E}[e^{i\xi Z}] = \frac{1}{1+\xi^2}.$$

□

Question 3

On rappelle que pour $f : \mathbb{R} \to \mathbb{R}$ intégrable, la transformée de Fourier de f est donnée par la fonction

$$\hat{f}(\xi) = \int_{-\infty}^{+\infty} e^{-2i\pi\xi x} f(x)\, \mathrm{d}x, \quad \xi \in \mathbb{R}, \tag{7}$$

et qu'on a la propriété d'inversion de Fourier:

$$f(x) = \int_{-\infty}^{+\infty} e^{2i\pi\xi x} \hat{f}(\xi)\, \mathrm{d}\xi, \quad x \in \mathbb{R}. \tag{8}$$

À l'aide de ce résultat et de la question précédente, montrer que la fonction caractéristique d'une variable aléatoire X de densité p est donnée par

$$g_X(\xi) = e^{-|\xi|}. \tag{9}$$

Solution. Par le résultat de la question 2, on sait que

$$\mathbb{E}[e^{i\xi Z}] = \frac{1}{1+\xi^2} = \pi p(\xi),$$

si Z est une variable aléatoire de densité $q(x) = \frac{1}{2} e^{-|x|}$. Or, on peut écrire

$$\mathbb{E}[e^{i\xi Z}] = \int_{\mathbb{R}} e^{i\xi x} q(x)\, \mathrm{d}x = \pi p(\xi).$$

En échangeant les noms de ξ et x, la dernière égalité se réécrit

$$p(x) = \int_{-\infty}^{\infty} \frac{1}{\pi} e^{ix\xi} q(\xi)\, \mathrm{d}\xi.$$

Pour se ramener à notre convention pour la transformée de Fourier, on fait le changement de variable

$$\tilde{\xi} = \frac{\xi}{2\pi} \implies \mathrm{d}\xi = 2\pi \mathrm{d}\tilde{\xi}.$$

Ceci donne

$$p(x) = \int_{-\infty}^{\infty} \frac{1}{\pi} e^{2i\pi\tilde{\xi}x} q(2\pi\tilde{\xi})\, 2\pi \mathrm{d}\tilde{\xi} = \int_{-\infty}^{\infty} e^{2i\pi\xi x} 2q(2\pi\xi)\, \mathrm{d}\xi.$$

Par la propriété d'inversion de Fourier, on en déduit que

$$\hat{p}(\xi) = 2q(2\pi\xi). \tag{10}$$

Il suffit à présent d'exprimer la fonction caractéristique de X comme une transformée de Fourier:

$$g_X(\xi) = \mathbb{E}[e^{i\xi X}] = \int_{-\infty}^{\infty} e^{i\xi x} p(x)\, \mathrm{d}x = \int_{-\infty}^{\infty} e^{-2i\pi(\frac{-\xi}{2\pi})x} p(x)\, \mathrm{d}x = \hat{p}\left(-\frac{\xi}{2\pi}\right)$$

La relation (10) donne finalement

$$g_X(\xi) = 2q\left(2\pi\left(-\frac{\xi}{2\pi}\right)\right) = 2q(-\xi) = e^{-|\xi|},$$

comme voulu. □

Question 4

Vérifier que pour toute variable aléatoire X à valeurs dans $\mathbb{R}^d$, pour tous $\xi \in \mathbb{R}^d$ et $\alpha \in \mathbb{R}$, on a l'identité

$$g_{\alpha X}(\xi) = g_X(\alpha\xi), \tag{11}$$

où $g_{\alpha X}$ et g_X désignent respectivement les fonctions caractéristiques des variables aléatoires αX et X.

Solution. Soit $\xi \in \mathbb{R}^d$. On a simplement, par bilinéarité du produit scalaire,

$$g_{\alpha X}(\xi) = \mathbb{E}[e^{i\xi\cdot(\alpha X)}] = \mathbb{E}[e^{i(\alpha\xi)\cdot X}] = g_X(\alpha\xi).$$

□

Question 5

On considère à présent une suite $(X_n)_{n\geq 1}$ de variables aléatoires indépendantes et identiquement distribuées selon la loi de Cauchy $\mathcal{C}(1)$. Montrer que, pour $N \in \mathbb{N}^*$, la moyenne empirique

$$M_N = \frac{1}{N}\sum_{n=1}^{N} X_n \tag{12}$$

suit encore une loi de Cauchy standard $\mathcal{C}(1)$. La suite $(M_N)_{N\geq 1}$ converge-t-elle en loi?

Solution. La fonction caractérisitique détermine la loi. Or on a, pour $\xi \in \mathbb{R}$,

$$g_{M_N}(\xi) = g_{\frac{1}{N}\sum_{n=1}^N X_n}(\xi) = g_{\sum_{n=1}^N X_n}\left(\frac{\xi}{N}\right),$$

par la question 4. La fonction caractéristique d'une somme de variables indépendantes étant égale au produit des fonctions caractéristique de chacune des variables sommées, on en déduit

$$g_{M_N}(\xi) = \prod_{n=1}^N g_{X_n}\left(\frac{\xi}{N}\right) = g_{X_1}\left(\frac{\xi}{N}\right)^N,$$

car les X_n sont indépendantes et de même loi de Cauchy $\mathcal{C}(1)$. Or, par la question 5, on connaît la fonction caractéristique d'une variable aléatoire suivant cette loi de Cauchy. En appliquant ce résultat, on obtient

$$g_{M_N}(\xi) = \left(e^{-\left|\frac{\xi}{N}\right|}\right)^N = e^{-|\xi|} = g_{X_1}(\xi). \tag{13}$$

On en déduit que M_N suit une loi de Cauchy $\mathcal{C}(1)$. Autrement dit, la suite $(M_n)_{n\geq 1}$ est constante en loi, et elle converge évidemment en loi: en effet, pour $f : \mathbb{R} \to \mathbb{R}$ continue et bornée,

$$\mathbb{E}[f(M_N)] = \mathbb{E}[f(X_1)] \underset{N\to\infty}{\longrightarrow} \mathbb{E}[f(X_1)],$$

Puisque M_N à la même loi que X_1. □

Question 7

Montrer sans calculer d'intégrale que pour tout $N \in \mathbb{N}^*$, la variable aléatoire

$$P_N = \frac{\sum_{n=1}^N \prod_{k\neq n} X_k}{N\prod_{n=1}^N X_n} \tag{14}$$

suit aussi une loi de Cauchy standard.

Solution. On remarque simplement que

$$P_N = \sum_{n=1}^N \frac{\prod_{k\neq n} X_k}{N\prod_{n=1}^N X_n} = \frac{1}{N}\sum_{n=1}^N \frac{1}{X_n}.$$

Or, par la question 1, pour chaque $n \geq 1$, la variable aléatoire $\frac{1}{X_n}$ suit une loi de Cauchy $\mathcal{C}(1)$. De plus, la famille $\left(\frac{1}{X_n}\right)_{1\leq n\leq N}$ est indépendante. En effet, en prenant $(f_n)_{1\leq n\leq N}$ une famille de fonctions mesurables bornées, on a

$$\mathbb{E}\left[\prod_{n=1}^N f_n\left(\frac{1}{X_n}\right)\right] = \mathbb{E}\left[\prod_{n=1}^N g_n(X_n)\right] = \prod_{n=1}^N \mathbb{E}\left[g_n(X_n)\right] = \prod_{n=1}^N \mathbb{E}\left[f_n\left(\frac{1}{X_n}\right)\right],$$

par indépendance de la famille $(X_n)_{1\leq n\leq N}$, où on a noté g_n la fonction mesurable bornée $g_n(x) = f_n\left(\frac{1}{x}\right)$. Ceci implique (proposition 3.3.11 du polycopié) que la famille $(\frac{1}{X_n})_{1\leq n\leq N}$ est indépendante.

On peut alors appliquer le résultat de la question 6 pour obtenir que

$$P_N = \frac{1}{N}\sum_{n=1}^{N}\frac{1}{X_n}$$

suit une loi de Cauchy $\mathcal{C}(1)$, comme moyenne empirique de variables indépendantes suivant la loi de Cauchy $\mathcal{C}(1)$. □

Question 8

La suite $(X_n)_{n\geq 1}$ satisfait-t-elle une loi faible des grands nombres: existe-t-il $m \in \mathbb{R}$ tel que

$$M_n \xrightarrow[n\to\infty]{\mathbb{P}} m\,? \tag{15}$$

Justifier puis commenter ce résultat. Que peut-on dire a fortiori sur la loi forte?

Solution. Soit $\varepsilon > 0$, $m \in \mathbb{R}$. On a

$$\begin{aligned}\mathbb{P}(|M_n - m| > \varepsilon) &= \mathbb{P}(M_n \in]m-\varepsilon, m+\varepsilon[)\\ &= \frac{1}{\pi}\int_{m-\varepsilon}^{m+\varepsilon}\frac{\mathrm{d}x}{1+x^2} > 0,\end{aligned}$$

car M_n est une variable de Cauchy $\mathcal{C}(1)$, par la question 5. Comme ce dernier terme est strictement positif et indépendant de n, on en déduit que M_n ne converge pas en probabilité vers m. Comme m est arbitraire, la loi faible des grands nombres ne tient pas: les moyennes empiriques d'une suite de variables de Cauchy indépendantes ne convergent pas en probabilité. A fortiori, la loi forte n'est pas non plus satisfaite, puisque la convergence presque sûre d'une suite de variables aléatoires implique la convergence en probabilité. □

Question 9

Même question pour le théorème central limite: existe-t-il $m \in \mathbb{R}$, $\sigma > 0$ tels que

$$\frac{1}{\sqrt{N}}\sum_{n=1}^{N}\left(\frac{X_n - m}{\sigma}\right) \xrightarrow[n\to\infty]{\mathcal{L}} G, \tag{16}$$

où $G \sim \mathcal{N}(0,1)$ est une variable gaussienne centrée réduite?

Solution. On pouvait directement montrer la non-convergence de la fonction caractéristique de

$$\frac{\sqrt{n}}{\sigma}\left(M_n - m\right)$$

vers celle d'une Gaussienne $\mathcal{N}(0,1)$ (ni d'ailleurs celle d'une quelconque variable aléatoire, par un argument de continuité).

Dans cette correction, on va montrer un résultat un peu plus général qui montre qu'un comportement asymptotique de type TCL implique toujours une loi faible des grands nombres: pour une suite de variables aléatoires $(Z_n)_{n\geq 1}$ tels qu'ils existent $m \in \mathbb{R}$, $\sigma > 0$ vérifiant

$$G_n = \frac{\sqrt{n}}{\sigma}(Z_n - m) \xrightarrow[n\to\infty]{\mathcal{L}} G \sim \mathcal{N}(0,1),$$

on a que Z_n converge en probabilités vers m. En effet, on a, pour $\varepsilon > 0$,

$$\begin{aligned}\mathbb{P}(|Z_n - m| > \varepsilon) &= \mathbb{P}\left(|G_n| > \sqrt{n}\varepsilon/\sigma\right) \\ &= \mathbb{E}\left[\mathbb{1}_{|G_n|>\sqrt{n}\varepsilon/\sigma}\right].\end{aligned}$$

Puisque $\sqrt{n}\varepsilon/\sigma \to \infty$, pour tout $M > 0$, il existe n_0 tel que pour tout $n \geq n_0$,

$$\mathbb{1}_{|G_n|>\sqrt{n}\varepsilon/\sigma} \leq \mathbb{1}_{|G_n|>M}. \tag{17}$$

Or,

$$\mathbb{E}\left[\mathbb{1}_{|G_n|>M}\right] = \mathbb{P}(|G_n| > M) \xrightarrow[n\to\infty]{} \mathbb{P}(|G| > M),$$

en notant le fait que $\mathbb{P}(G \in \partial\{|x| > M\}) = \mathbb{P}(G = \pm M) = 0$. Il est facile de vérifier, comme $G \sim \mathcal{N}(0,1)$, que

$$\lim_{M\to\infty} \mathbb{P}(|G| > M) = 0,$$

ce qui implique que pour tout $\delta > 0$, il existe $M_\delta \geq 0$ tel que pour tout $M \geq M_\delta$,

$$\mathbb{P}(|G_n| > M) < \delta,$$

puis, en utilisant la borne (17), que pour tout $\delta > 0$, il existe n_0 tel que pour tout $n \geq n_0$,

$$\mathbb{P}\left(|G_n| > \sqrt{n}\varepsilon/\sigma\right) < \delta.$$

Autrement dit,

$$\mathbb{P}(|Z_n - m| > \varepsilon) = \mathbb{P}\left(|G_n| > \sqrt{n}\varepsilon/\sigma\right) \xrightarrow[n\to\infty]{} 0, \tag{18}$$

ce qui exprime la convergence de Z_n vers m en probabilités. Notons ce raisonnement reste valide en remplacant $\sqrt{n}$ par le terme a_n d'une suite tendant vers $+\infty$, G par n'importe quelle variable a densité, et m par une variable aléatoire à valeurs réelles.

En appliquant en particulier ce petit lemme au cas où $(Z_n)_{n\geq 1}$ est une suite de moyennes empiriques, on trouve en particulier que le TCL implique la loi faible des grands nombres. Par la contraposée, notre suite de variables de Cauchy M_n ne satisfait pas de TCL: cela entrerait en contradiction avec la conclusion de la question 8. $\square$