

THÉORÈME DE LÉVY-CRAMÉR

DONATI MARTIN

RÉSUMÉ. L'objectif de cet article va être de donner la démonstration complète du théorème suivant :

Théorème (Lévy-Cramér). *Considérons X et Y deux variables aléatoires indépendantes. Si la variable $X + Y$ suit une loi normale, alors X et Y suivent des lois normales.*

Bien entendu, pour que ce résultat soit vrai, nous considérerons ici les variables de variance nulle comme des lois normales. Cette preuve est issue de [1], mais nous nous attacherons ici à donner le maximum de détail. Nous commencerons par établir quelques résultats d'analyse complexe élémentaires qui serviront à étudier les propriétés des fonctions caractéristiques des variables aléatoires.

1. RAPPELS D'ANALYSE COMPLEXE

Toute cette section est tirée de [2], auquel le lecteur est vivement encouragé à se référer. Commençons par rappeler un théorème de factorisation fondamental.

Théorème 1.1. *Soit f une fonction entière n'ayant pas de zéro, alors il existe une fonction entière g telle que $f = \exp \circ g$.*

Démonstration. La fonction exponentielle est surjective de \mathbb{C} dans \mathbb{C}^* , de sorte qu'il existe un complexe z_0 tel que $\exp(z_0) = f(0)$. Le quotient de deux fonctions holomorphes est holomorphe. De plus, toute fonction holomorphe qui ne s'annule pas sur un convexe admet une primitive. Par conséquent, la fonction $\frac{f'}{f}$ qui est entière possède une primitive entière $g(z) = z_0 + \int_{[0,z]} \frac{f'(t)}{f(t)} dt$. Soit $u = \exp \circ g$. Alors :

$$u' = g' \exp \circ g = \frac{f'}{f} u,$$

de sorte que :

$$u' f - f' u = 0.$$

En intégrant, f et u ne diffèrent donc que d'une constante sur \mathbb{C} tout entier ; or $u(0) = \exp(z_0) = f(0)$, donc $f = \exp \circ g$. \square

Nous aurons également besoin d'un autre théorème d'analyse complexe. Une fonction entière majorée en module par le module d'une fonction polynomiale est nécessairement elle-même polynomiale. Si la preuve de cet énoncé est assez élémentaire, nous allons avoir besoin d'un énoncé un peu plus fort, qui va nous contraindre à une preuve plus sophistiquée.

Théorème 1.2. *Soit f une fonction entière et P un polynôme tel que :*

$$\forall z \in \mathbb{C}, \Re(f(z)) \leq |P(z)|,$$

alors f est polynomiale de degré plus petit que celui de P .

Démonstration. La fonction f est entière, donc s'écrit :

$$f(z) = \sum_{n \geq 0} c_n z^n,$$

où $c_n \in \mathbb{C}$. Notons $U(z) = \Re(f(z))$. Les formules de Cauchy donnent, pour $n \geq 1$ et quel que soit $r > 0$, d'une part :

$$(1.1) \quad c_n r^n = \frac{1}{2\pi} \int_0^{2\pi} f(re^{i\theta}) e^{-in\theta} d\theta,$$

et d'autre part :

$$(1.2) \quad 0 = \frac{1}{2\pi} \int_0^{2\pi} f(re^{i\theta}) e^{in\theta} d\theta.$$

La somme de l'équation (1.1) et de l'équation (1.2) passée au conjugué donne :

$$c_n r^n = \frac{1}{\pi} \int_0^{2\pi} \Re(f(re^{i\theta})) e^{-in\theta} d\theta,$$

de sorte que :

$$|c_n r^n| \leq \frac{1}{\pi} \int_0^{2\pi} |U(re^{i\theta})| d\theta.$$

De plus,

$$c_0 = \frac{1}{2\pi} \int_0^{2\pi} f(re^{i\theta}) d\theta,$$

et donc :

$$2\Re(c_0) = \frac{1}{\pi} \int_0^{2\pi} U(re^{i\theta}) d\theta.$$

Finalement, on obtient :

$$|c_n r^n| + 2\Re(c_0) \leq \frac{1}{\pi} \int_0^{2\pi} |U(re^{i\theta})| + U(re^{i\theta}) d\theta,$$

et donc :

$$|c_n| r^n + 2\Re(c_0) \leq 4 \max_{|z|=r} (U(z), 0).$$

Notons $P = \sum_{k=0}^d a_k X^k$. Maintenant, l'hypothèse $U(z) \leq |P(z)|$ implique que pour tout n plus grand que 1, on a :

$$|c_n| r^n + 2\Re(c_0) \leq 4 \max_{|z|=r} \left| \sum_{k=0}^d a_k z^k \right| \underset{r \rightarrow \infty}{\sim} 4|a_d| r^d.$$

Si $n > d$, en faisant tendre r vers l'infini, on trouve $c_n = 0$. Ainsi, f est polynomiale de degré plus petit que celui de P . □

2. FONCTIONS CARACTÉRISTIQUES ET LEMME INTERMÉDIAIRE

Venons-en maintenant aux probabilités. Nous allons établir un certain nombre de propriétés intermédiaires. En premier lieu, rappelons la définition de la fonction caractéristique d'une variable aléatoire :

Définition 2.1. On appelle fonction caractéristique d'une variable aléatoire X la fonction définie par :

$$\phi_X(\xi) = \mathbb{E}(e^{i\xi X}) = \int_{\mathbb{R}} e^{i\xi x} \mathbb{P}_X(dx).$$

Proposition 2.1. La fonction caractéristique d'une variable suivant une loi normale est de la forme :

$$\phi_X(\xi) = \exp\left(i\mu\xi - \frac{1}{2}\sigma^2\xi^2\right),$$

où $\mu, \sigma \in \mathbb{R}$, et en particulier ne s'annule pas sur \mathbb{C} .

Nous passerons sur la démonstration de ce résultat bien connu.

Remarquons que l'on a :

$$\forall \xi \in \mathbb{C}, \forall (x, \eta) \in \mathbb{R}^2, \eta \neq 0, |x\xi| \leq |\eta x| |\eta^{-1}\xi| \leq \eta^2 x^2 + \eta^{-2} |\xi|^2,$$

simplement par l'identité $2|ab| \leq |a|^2 + |b|^2$, en se passant même du rapport 2. Par conséquent, l'on a la propriété suivante :

Proposition 2.2.

$$\forall \xi \in \mathbb{C}, \forall \eta \neq 0, |\phi_X(\xi)| \leq e^{\eta^{-2}|\xi|^2} \int_{\mathbb{R}} e^{-\eta^2 x^2} \mathbb{P}_X(dx).$$

Démonstration. Soit $\xi \in \mathbb{C}$:

$$\begin{aligned} |\phi_X(\xi)| &\leq \int_{\mathbb{R}} |e^{i\xi x}| \mathbb{P}_X(dx) \\ &\leq \int_{\mathbb{R}} e^{|\xi x|} \mathbb{P}_X(dx) \\ &\leq \int_{\mathbb{R}} e^{-\eta^2 x^2 + \eta^{-2} |\xi|^2} \mathbb{P}_X(dx). \end{aligned}$$

□

Posons :

$$f_X(\eta) = \int_{\mathbb{R}} e^{-\eta^2 x^2} \mathbb{P}_X(dx).$$

La propriété précédente affirme donc le résultat suivant :

Proposition 2.3. S'il existe $\eta > 0$ tel que $f_X(\eta) < \infty$, alors la fonction caractéristique ϕ_X est entière.

C'est en particulier le cas pour les lois normales, comme nous le verrons par la suite.

Le lemme suivant va reprendre tout ce qui a été fait avant, et constitue le pilier central de la preuve.

Lemme 2.1. Si X est une variable aléatoire réelle telle qu'il existe $\eta > 0$ tel que $f_X(\eta)$ soit fini, et ϕ_X ne s'annule pas, alors X suit une loi normale.

Démonstration. La propriété 2.2 donne :

$$|\phi_X(\xi)| \leq f(\eta)e^{\eta^{-2}|\xi|^2}.$$

Puisque $f(\eta) < \infty$, ϕ_X est une fonction entière, et qui par hypothèse ne s'annule pas. Ainsi, d'après le théorème 1.1, il existe une fonction entière g telle que :

$$\phi_X = \exp \circ g,$$

de sorte que :

$$|e^{g(\xi)}| \leq f(\eta)e^{\eta^{-2}|\xi|^2},$$

et donc :

$$\Re(g(\xi)) \leq \ln(f(\eta)) + \eta^{-2}|\xi|^2.$$

La théorème 1.2 s'applique donc et g est polynomiale de degré plus petit que 2. On peut donc écrire :

$$g(\xi) = \frac{a}{2}\xi^2 + i\mu\xi + b.$$

Mais puisque $\phi_X = \exp(g)$, en appliquant en 0 on obtient directement $b = 0$. Ensuite, on a d'une part :

$$\mathbb{E}(X) = -ig'(0) \exp(g(0)),$$

de sorte que :

$$\mu = \mathbb{E}(X) \in \mathbb{R},$$

et :

$$\mathbb{E}(X^2) = -1[g''(0) \exp(g(0)) + g'(0)^2 \exp(g(0))],$$

ce qui donne :

$$a = -\mathbb{E}(X^2) - \mu^2 \leq 0.$$

Finalement, on peut écrire la fonction caractéristique de X sous la forme :

$$\phi_X(\xi) = \exp\left(i\mu\xi - \frac{1}{2}\sigma^2\xi^2\right),$$

avec $\mu, \sigma \in \mathbb{R}$, ce qui est bien la fonction caractéristique associée à une loi normale. \square

Maintenant, l'objectif de la preuve se dessine : si l'on est capable de montrer toutes ces conditions pour deux variables X et Y dont la somme est normale, alors elles seront également normales. Il nous reste encore une dernière subtilité technique à traiter. Posons d'abord une formule probabiliste d'intégration par parties.

Proposition 2.4 (Intégration par parties). *Soit X une variable aléatoire réelle et \mathbb{P}_X sa loi. Soit f une fonction de classe C^1 sur \mathbb{R} . Alors on a :*

$$\forall a < b, \int_a^b f(x)\mathbb{P}_X(dx) = f(b)\mathbb{P}(X \leq b) - f(a)\mathbb{P}(X \leq a) - \int_a^b f'(t)\mathbb{P}(X \leq t)dt.$$

Démonstration. Par convention, nous noterons $\int_a^b f(x)\mathbb{P}(\mathrm{d}x) = \mathbb{E}(f(x)\mathbb{1}_{x \in [a,b]})$.
Soit $a < b$. Partons du membre de gauche :

$$\begin{aligned} \int_a^b f(x)\mathbb{P}_X(\mathrm{d}x) &= \int_a^b \left(\int_a^x f'(t)\mathrm{d}t + f(a) \right) \mathbb{P}_X(\mathrm{d}x) \\ &= \int_a^b \int_t^b f'(t)\mathbb{P}_X(\mathrm{d}x)\mathrm{d}t + f(a)\mathbb{P}(a < X \leq b) \\ &= \int_a^b \mathbb{P}(t < X \leq b)f'(t)\mathrm{d}t + f(a)\mathbb{P}(a < X \leq b) \\ &= \int_a^b [\mathbb{P}(X \leq b) - \mathbb{P}(X \leq t)]f'(t)\mathrm{d}t \\ &\quad + f(a)[\mathbb{P}(X \leq b) - \mathbb{P}(X \leq a)] \\ &= - \int_a^b f'(t)\mathbb{P}(X \leq t)\mathrm{d}t + [f(b) - f(a)]\mathbb{P}(X \leq b) \\ &\quad + f(a)[\mathbb{P}(X \leq b) - \mathbb{P}(X \leq a)] \\ &= f(b)\mathbb{P}(X \leq b) - f(a)\mathbb{P}(X \leq a) - \int_a^b f'(t)\mathbb{P}(X \leq t)\mathrm{d}t. \end{aligned}$$

□

A l'aide de l'intégration par parties, nous allons pouvoir écrire la quantité $f_X(\eta)$ sous une forme différente, de façon à faire intervenir uniquement la probabilité $\mathbb{P}(|X| \geq x)$. Cette forme nous sera beaucoup plus utile par la suite.

Proposition 2.5. *On a :*

$$f_X(\eta) = \int_0^{+\infty} 2x\eta^2 e^{\eta^2 x^2} \mathbb{P}(|X| \geq x)\mathrm{d}x + 1.$$

Démonstration. Pour pouvoir faire l'intégration par parties, il faut être sur un intervalle fini. De plus, la partie positive de l'intégrale et la partie négative vont être traitées séparément, de sorte que l'on va écrire :

$$f_X(\eta) = \lim_{b \rightarrow +\infty} \int_{-b}^0 e^{\eta^2 x^2} \mathbb{P}_X(\mathrm{d}x) + \int_0^b e^{\eta^2 x^2} \mathbb{P}_X(\mathrm{d}x).$$

Traisons le premier terme :

$$\begin{aligned} \int_{-b}^0 e^{\eta^2 x^2} \mathbb{P}_X(\mathrm{d}x) &= \mathbb{P}(X \leq 0) - e^{\eta^2 b^2} \mathbb{P}(X \leq -b) - \int_{-b}^0 2x\eta^2 e^{\eta^2 x^2} \mathbb{P}(X \leq x)\mathrm{d}x \\ &= \mathbb{P}(X \leq 0) - e^{\eta^2 b^2} \mathbb{P}(X \leq -b) + \int_0^b 2x\eta^2 e^{\eta^2 x^2} \mathbb{P}(X \leq -x)\mathrm{d}x. \end{aligned}$$

Le second terme est plus subtil à traiter :

$$\begin{aligned}
 \int_0^b e^{\eta^2 x^2} \mathbb{P}_X(dx) &= e^{\eta^2 b^2} \mathbb{P}(X \leq b) - \mathbb{P}(X \leq 0) - \int_0^b 2x\eta^2 e^{\eta^2 x^2} \mathbb{P}(X \leq x) dx \\
 &= e^{\eta^2 b^2} \mathbb{P}(X \leq b) - \mathbb{P}(X \leq 0) \\
 &\quad + \int_0^b 2x\eta^2 e^{\eta^2 x^2} (1 - \mathbb{P}(X \leq x)) dx - \int_0^b 2x\eta^2 e^{\eta^2 x^2} dx \\
 &= e^{\eta^2 b^2} \mathbb{P}(X \leq b) - \mathbb{P}(X \leq 0) \\
 &\quad + \int_0^b 2x\eta^2 e^{\eta^2 x^2} (1 - \mathbb{P}(X \leq x)) dx + 1 - e^{\eta^2 b^2} \\
 &= e^{\eta^2 b^2} (\mathbb{P}(X \leq b) - 1) - \mathbb{P}(X \leq 0) \\
 &\quad + \int_0^b 2x\eta^2 e^{\eta^2 x^2} \mathbb{P}(X \geq x) dx + 1.
 \end{aligned}$$

En combinant les deux, on a :

$$f_X(\eta) = \lim_{b \rightarrow +\infty} -e^{\eta^2 b^2} \mathbb{P}(|X| \geq b) + \int_0^b 2x\eta^2 e^{\eta^2 x^2} \mathbb{P}(|X| \geq x) dx + 1,$$

d'où :

$$(2.1) \quad f_X(\eta) \leq \int_0^{+\infty} 2x\eta^2 e^{\eta^2 x^2} \mathbb{P}(|X| \geq x) dx + 1.$$

La finitude de $f_X(\eta)$ implique que :

$$f_X(\eta) = \lim_{b \rightarrow +\infty} -e^{\eta^2 b^2} \mathbb{P}(|X| \geq b) = 0,$$

et finalement :

$$f_X(\eta) = \int_0^{+\infty} 2x\eta^2 e^{\eta^2 x^2} \mathbb{P}(|X| \geq x) dx + 1.$$

L'inégalité (2.1) donne l'égalité dans le cas où $f(\eta) = +\infty$. □

3. PREUVE DU THÉORÈME

Nous pouvons donc passer à la preuve du théorème.

Théorème (Lévy-Cramér). *Considérons X et Y deux variables aléatoires indépendantes définies sur un espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et à valeurs dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ muni de la mesure de Lebesgue sur \mathbb{R} . Si la variable $X + Y$ suit une loi normale, alors X et Y suivent toutes deux des lois normales.*

Démonstration. Notons σ l'écart type de la loi de $X + Y$. Si $\sigma = 0$, alors $\mathbb{V}(X) = \mathbb{V}(Y) = 0$ et il existe donc a et b tels que $X \sim \mathcal{N}(a, 0)$ et $Y \sim \mathcal{N}(b, 0)$. Supposons désormais que $\sigma > 0$.

La première étape consiste à montrer qu'il existe $\eta > 0$ tel que $f_{X+Y}(\eta) < \infty$. En effet :

$$f_{X+Y}(\eta) = \int_{\mathbb{R}} e^{\eta^2 x^2} e^{-\frac{\sigma^2 x^2}{2}} dx,$$

et donc si $0 < \eta < \frac{1}{\sqrt{2}}\sigma$, alors $f_{X+Y}(\eta) < \infty$.

Maintenant, penchons nous sur les variables X et Y . Étudions tout d'abord le cas où les lois de X et Y ont toutes deux un atome. Soit a et b deux tels atomes, on a :

$$\mathbb{P}(X + Y = a + b) \geq \mathbb{P}(X = a, Y = b) = \mathbb{P}(X = a)\mathbb{P}(Y = b) > 0.$$

Ainsi, si la loi de $X + Y$ est sans atome, les variables X et Y ne peuvent pas avoir toutes deux des atomes.

Par conséquent, dans toute la suite nous supposons que Y est sans atome (puisque les deux variables jouent des rôles symétriques). Il existe donc μ tel que $\mathbb{P}(Y \leq \mu) = \mathbb{P}(Y \geq \mu) = \frac{1}{2}$. La somme $X + Y - \mu$ est toujours une loi normale, ce qui nous permet de supposer sans perte de généralité que $\mu = 0$.

La seconde étape consiste à montrer la finitude de $f_X(\eta)$. On a :

$$\begin{aligned} \mathbb{P}(|X + Y| \geq x) &= \mathbb{P}(X + Y \geq x) + \mathbb{P}(X + Y \leq -x) \\ &\geq \mathbb{P}(X \geq x, Y \geq 0) + \mathbb{P}(X \leq -x, Y \leq 0) \\ &\geq \frac{1}{2}\mathbb{P}(|X| \geq x) \end{aligned}$$

D'après la propriété 2.5, pour $0 < \eta < \frac{1}{\sqrt{2}}\sigma$, il vient directement :

$$f_X(\eta) \leq 2f_{X+Y}(\eta) < \infty.$$

De plus, $\phi_X\phi_Y = \phi_{X+Y}$ donc ϕ_X ne s'annule pas puisque ϕ_{X+Y} ne s'annule pas. Par conséquent, on peut appliquer à X le lemme 2.1, et X suit une loi normale. Si X est de variance nulle, alors $Y = (X + Y) - X$ suit une loi normale, et sinon, X est sans atome et l'on peut échanger les rôles de X et Y .

Finalement, X et Y suivent deux lois normales et le résultat est démontré. \square

Évoquons la notion de stabilité. Pour plus de détails, se référer à [1].

Définition 3.1. On dit que la loi de probabilité μ est strictement stable si pour toutes variables aléatoires (X_1, \dots, X_n) indépendantes de loi μ , il existe $\alpha > 0$ tel que $n^{-\frac{1}{\alpha}} \sum_{i=1}^n X_i$ soit également de loi μ .

Selon cette définition, la loi normale centrée est 2-stable, c'est à dire (strictement) stable pour l'indice $\alpha = 2$. En particulier, la loi normale centrée est infiniment divisible, c'est à dire peut s'écrire comme une somme de longueur quelconque de lois normales indépendantes, ce qui joue un rôle central dans la convergence pour le théorème central limite. Le théorème que l'on a démontré assure donc que cette décomposition est la seule possible.

RÉFÉRENCES

- [1] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume 2. Wiley, 1971.
- [2] E. Hille. *Analytic Function Theory*, volume 1. Ginn and Company, Boston, 1959.