ESAIM: PS 17 (2013) 790–794
DOI: 10.1051/ps/2012021

ESAIM: Probability and Statistics
www.esaim-ps.org

PAUL LÉVY ET L'ARITHMÉTIQUE DES LOIS DE PROBABILITÉS

Jean Bertoin[1]

Résumé. Ce court texte reprend un exposé donné le 15 Décembre 2011 au Laboratoire de Probabilités et Modèles Aléatoires, lors d'une journée en hommage à Paul Lévy. On y rappellera comment des considérations sur l'arithmétique des lois de probabilités ont conduit Lévy à étudier les processus à accroissements indépendants.

Classification Mathématique. 60E07, 60G51, 01A60.

Reçu le juillet 5, 2012.

Ce texte repose en grande partie sur la seconde édition du livre de Paul Lévy : *théorie de l'addition des variables aléatoires* parue en 1954, la première datant de 1937 étant moins complète sur le sujet. On pourra consulter également les articles originaux de Lévy dont les oeuvres ont été rééditées par Gauthier–Villars, ainsi que des travaux relatifs de de Finetti, Kolmogorov, Khintchine, Linnik, et un mémoire de D. Dugué.

La terminologie du titre tire ses racines du théorème fondamental de l'arithmétique qui énonce que tout nombre entier $n \geq 1$ admet une décomposition comme produit fini de facteurs premiers, $n = \prod_{i=1}^{k} p_i$, et que cette décomposition est unique à une permutation de ces facteurs près. L'arithmétique des lois de probabilités s'intéresse elle aux décompositions d'une loi μ, disons pour simplifier sur $\mathbb{R}$, en produit de convolution

$$\mu = \nu_1 * \nu_2 * \dots$$

où le nombre de facteurs peut éventuellement être infini. Dans ce cas, il faut comprendre un tel produit infini comme une limite de produits finis, *i.e.* $\mu = \lim_{k\to\infty} \nu_1 * \dots * \nu_k$, au sens de la convergence étroite des lois de probabilités.

De façon équivalente, il s'agit de représenter une variable aléatoire X de loi μ comme somme de variables aléatoires indépendantes

$$X = Y_1 + Y_2 + \dots,$$

ou encore la fonction caractéristique Φ_μ de μ

$$\Phi_\mu(a) = \mathbb{E}(\mathrm{e}^{\mathrm{i}aX}) = \int \mathrm{e}^{\mathrm{i}ax}\mu(\mathrm{d}x), \qquad a \in \mathbb{R}$$

comme produit de fonctions caractéristiques

$$\Phi_\mu = \Phi_{\nu_1} \times \Phi_{\nu_2} \times \dots.$$

Mots Clés. Loi indéfiniment divisible, formule de Lévy–Khintchine.

[1] Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland. jean.bertoin@math.uzh.ch

Article published by EDP Sciences

© EDP Sciences, SMAI 2013

https://doi.org/10.1051/ps/2012021 Published online by Cambridge University Press

Cette somme de variables aléatoires, respectivement ce produit de fonctions caractéristiques, peut éventuellement contenir un nombre infini de termes. Dans ce cas, la somme infinie est prise au sens de la convergence presque sûre, *i.e.* $X = \lim_{k\to\infty}(Y_1 + \ldots + Y_k)$ p.s., et le produit infini au sens de la convergence simple, *i.e.* $\Phi_\mu(a) = \lim_{k\to\infty} \Phi_{\nu_1}(a) \times \ldots \times \Phi_{\nu_k}(a)$ pour chaque $a \in \mathbb{R}$. En effet, des résultats bien connus dûs à Paul Lévy énoncent d'une part qu'une série de variables aléatoires indépendantes converge en loi si et seulement si elle converge presque sûrement, et d'autre part que la convergence étroite des lois de probabilités est équivalente à la convergence simple des fonctions caractéristiques.

Dans ce cadre, l'analogue de la notion de nombre premier est celle de loi indécomposable, c'est-à-dire telle que les seules décompositions d'une v.a. X de loi μ en somme de v.a. indépendantes soient de la forme

$$X = (X - a) + a,$$

où a est déterministe. Par exemple la loi de Bernoulli est indécomposable, et d'une façon générale, les lois à support fini le sont souvent, mais pas toujours (évidemment, la loi binomiale de paramètre (n, p) avec $n \geq 2$ et $0 < p < 1$ ne l'est pas).

Si le parallèle avec l'arithmétique des nombres est naturel et *a priori* intéressant, en fait il tourne court assez vite. D'abord, on remarque facilement qu'une décomposition en facteurs indécomposables n'est pas nécessairement unique. Par exemple, si on note U_F une variable aléatoire de loi uniforme sur un ensemble fini F, on a

$$U_{\{0,1,2,3,4,5\}} \overset{\text{(loi)}}{=} U_{\{0,1,2\}} + U_{\{0,3\}} \overset{\text{(loi)}}{=} U_{\{0,1\}} + U_{\{0,2,4\}}$$

où l'on suppose implicitement que les sommes font intervenir des variables indépendantes. Ensuite, une décomposition peut impliquer un nombre infini de facteurs indécomposables. Par exemple, si $U_{[0,1]}$ désigne une variable de loi uniforme sur $[0, 1]$, on a

$$U_{[0,1]} \overset{\text{(loi)}}{=} \sum_{i=1}^{\infty} 2^{-i}\varepsilon_i$$

où les ε_i forment une suite i.i.d. de variables de Bernoulli à valeurs dans $\{0, 1\}$. Notons en passant qu'on a aussi

$$U_{[0,1]} \overset{\text{(loi)}}{=} \sum_{j=1}^{\infty} 3^{-j}\eta_i$$

avec cette fois η_i une suite i.i.d. de variables distribuées comme $U_{\{0,1,2\}}$. On peut évidemment trouver beaucoup d'autres décompositions similaires.

D'autre part, il existe des lois que l'on ne peut tout simplement pas décomposer en produit même infini de lois indécomposables. L'exemple le plus connu est celui de la loi gaussienne, puisque le célèbre théorème de Lévy–Cramer énonce que si X est gaussien et $X = Y + Z$ avec Y et Z indépendants, alors Y et Z sont également des variables gaussiennes. Raikoff a démontré un peu plus tard un résultat similaire pour la loi de Poisson.

Ces considérations conduisent naturellement à la notion de loi indéfiniment divisible : on dit que la loi d'une variable X est indéfiniment divisible (ou plus simplement que X est indéfiniment divisible) si pour tout $\epsilon > 0$, il existe un entier n et des variables indépendantes $Y_1, \ldots, Y_n$ telles que

$$X = Y_1 + \ldots + Y_n \quad \text{et} \quad \mathbb{P}(|Y_i| > \epsilon) \leq \epsilon \quad \text{pour } i = 1, \ldots, n.$$

Il faut noter que cette définition est en apparence plus faible que celle communément adoptée de nos jours, à savoir qu'on suppose maintenant que les variables Y_i ont toutes la même loi (on devrait parler plutôt de loi dont on peut prendre la racine à n'importe quel ordre ; noter aussi que la seconde condition est alors automatiquement remplie). Cependant, on peut montrer que les deux définitions sont en réalité équivalentes.

https://doi.org/10.1051/ps/2012021 Published online by Cambridge University Press

En dépit des différences profondes entre l'arithmétique des nombres et l'arithmétique des lois de probabilités, il existe cependant un analogue partiel du théorème fondamental de l'arithmétique des nombres, dû à Khintchine : toute loi de probabilités sur $\mathbb{R}$ peut être considérée comme le produit d'une loi indéfiniment divisible et d'un produit, éventuellement infini, de lois indécomposables. Lévy écrit dans son livre que le résultat est presque immédiat par un argument de compacité, mais personnellement l'assertion ne me paraît pas si simple. Quoiqu'il en soit, l'intérêt du résultat est limité par le caractère non-canonique de cette décomposition. En effet, Lévy a donné des exemples de lois indéfiniment divisibles qui peuvent également s'exprimer comme produit de lois indécomposables.

À l'instar de Lévy, et après des contributions fondamentales de Khintchine, mais aussi Kolmogorov et de Finetti, intéressons nous maintenant de plus près aux lois indéfiniment divisibles et à leur structure. Il y a un adage en géométrie qui dit que parfois, le plus court chemin pour résoudre un problème sur la droite consiste à travailler dans le plan ou dans l'espace. Ici, nous allons étudier un problème portant sur des lois unidimensionnelles en considérant un processus stochastique, c'est-à-dire un objet infini-dimensionnel.

Dans le chapitre VII de son livre intitulé *Intégrales à éléments aléatoires indépendants*, Lévy introduit la notion de Processus à Accroissements Indépendants (PAI), c'est-à-dire d'un processus $(X_t)_{t\geq 0}$ tel que pour tous $s, t \geq 0$, l'accroissement $X_{t+s} - X_t$ est indépendant de la famille des variables $(X_r)_{0\leq r\leq t}$. L'idée implicite est que la valeur en un temps T fixé d'un tel processus doit être une variable indéfiniment divisible, et que réciproquement toute variable indéfiniment divisible peut être construite de cette manière. De façon informelle (et Lévy ne donne d'ailleurs pas de détails), si μ est une loi indéfiniment divisible, on définit d'abord $X_0 = 0$ et X_T de loi μ, puis pour $\epsilon > 0$ fixé et n entier assez grand, on décompose μ sous la forme $\mu = \nu_1 * \ldots * \nu_n$ avec $\nu_i(\{|x| > \epsilon\}) < \epsilon$. Puis on se donne des temps $0 < t_1 < \ldots < t_n = T$ de manière arbitraire, et enfin on construit des variables $X_{t_1}, \ldots, X_{t_n}$ de sorte que les accroissements $X_{t_{i+1}} - X_{t_i}$ soient indépendants et aient pour lois ν_{i+1}. Plus rigoureusement, il faudrait probablement faire appel au théorème de Kolmogorov pour justifier l'existence d'un tel processus.

Lévy observe d'abord qu'on peut éliminer les discontinuités déterministes d'un PAI en le représentant sous la forme $X_t = X'_t + f(t)$ avec f fonction déterministe et X' processus dont les lois marginales admettent des limites à droite et à gauche en tout temps. Bien sûr X' reste un PAI. Puis dans une seconde étape, il élimine les discontinuités aléatoires qui apparaissent à temps fixes, sur un ensemble déterministe au plus dénombrable. Pour cela, il utilise de façon habile la notion de valeur médiane.

L'intuition qui guide Lévy est que c'est la régularité, ou plutôt l'irrégularité des trajectoires de X qui va révéler sa structure. Dans cette direction, il montre tout d'abord que si le processus X a des trajectoires continues p.s., alors c'est nécessairement un processus gaussien. Pour cela, il s'appuie sur un de ses propres résultats sur les sommes de variables indépendantes qu'il énonce de la manière suivante :
Pour que la somme $S = \xi_1 + \ldots + \xi_n$ de variables indépendantes dépende d'une loi très peu différente de celle de Gauss, il suffit que le plus grand de ses termes soit négligeable.

Puis, en considérant la variance $\sigma^2(t)$ de X_t, Lévy observe que X doit être un mouvement brownien changé de temps, ce qu'il écrit également sous la forme

$$X_t = \int_0^t \xi(\tau)\sqrt{\mathrm{d}\sigma^2(\tau)}$$

où $\xi(\tau)$ est une variable aléatoire de loi $\mathcal{N}(0,1)$ et pour des valeurs différentes de τ on a des variables indépendantes les unes des autres (sic). Il établit en passant un résultat désormais classique sur le module de continuité du mouvement brownien.

Lévy s'intéresse ensuite au rôle de ce qu'il appelle les discontinuités mobiles (c'est-à-dire des sauts en des temps aléatoires), et insiste sur l'importance de la loi de Poisson. Il décrit le processus de Poisson en observant que pour chaque intervalle de temps très petit $\mathrm{d}t$, la probabilité de l'existence d'un saut brusque est sensiblement $\mathrm{d}t$. Cette probabilité se répartit de manière continue sans qu'il y ait de probabilité positive en aucun point donné. Il montre que tous les sauts sont de taille 1 en notant que la probabilité d'un accroissement d'au moins 2 est de l'ordre de $\mathrm{d}t^2$, et donc la probabilité totale infiniment petite. Il propose une construction du processus de

https://doi.org/10.1051/ps/2012021 Published online by Cambridge University Press

Poisson conditionné par sa valeur au temps T, $X_T = n$, en répartissant de façon uniforme les n instants de sauts sur $[0, T]$.

Lévy construit alors par généralisations successives les PAI à partir du processus de Poisson. Une première extension simple consiste à répartir les discontinuités de manière non-homogène, puis une seconde, plus importante, à rendre variable la hauteur des sauts. Autrement dit, Lévy considère des processus de Poisson composés. Leur exposant caractéristique, c'est-à-dire le logarithme de leur fonction caractéristique, est de la forme

$$\log \mathbb{E}\left(\exp(\mathrm{i}uX_t)\right) = \int_{[0,t]\times\mathbb{R}} (\mathrm{e}^{\mathrm{i}uz} - 1)\mathrm{d}_t\mathrm{d}_u N(t,u)$$

où N est la copule, *i.e.* la fonction de distribution cumulative d'une mesure finie sur la bande $[0,t]\times\mathbb{R}$ qui décrit l'intensité des sauts du processus de Poisson composé. Si on note $n(u) = N(T,u)$, on a donc

$$\Psi_T(z) = \int_{\mathbb{R}} (\mathrm{e}^{\mathrm{i}uz} - 1)\mathrm{d}n(u)$$

où Ψ_T désigne l'exposant caractéristique de X_T.

Lévy s'attache ensuite à étendre la construction à des situations où l'intensité totale peut être infinie. Il observe que la mesure $\mathrm{d}n(u)$ doit nécessairement vérifier la condition intégrale $\int_{-1}^{+1} u^2\mathrm{d}n(u) < \infty$. Puis en utilisant des résultats de type compensation qui permettent de rendre presque-sûrement convergentes des séries de variables aléatoires indépendantes presque-sûrement divergentes par addition à chaque terme d'une quantité déterministe, il montre que si cette condition intégrale est vérifiée et de plus $n(\mathbb{R}\backslash[-1,1]) < \infty$, alors on peut associer à N un PAI dont la valeur au temps T a pour exposant caractéristique

$$\Psi_T(z) = \int_{\mathbb{R}} \left(\mathrm{e}^{\mathrm{i}uz} - 1 - \frac{\mathrm{i}uz}{1+u^2}\right)\mathrm{d}n(u).$$

Il a alors tous les outils pour énoncer le

Théorème (Formule de Lévy–Khintchine)**.** *Le logarithme de la fonction caractéristique d'une loi indéfiniment divisible est de la forme*

$$\Psi(z) = \mathrm{i}mz - \frac{\sigma^2}{2} + \int_{\mathbb{R}} \left(\mathrm{e}^{\mathrm{i}uz} - 1 - \frac{\mathrm{i}uz}{1+u^2}\right)\mathrm{d}n(u)$$

où $m \in \mathbb{R}$, $\sigma^2 \geq 0$ et n est une fonction décroissante sur $]-\infty, 0[$ et sur $]0,\infty[$, telle que $\int_{-1}^{+1} u^2\mathrm{d}n(u) < \infty$.

Les deux premiers termes de la somme correspondent à un mouvement brownien de variance σ^2 avec une dérive m. L'intégrale décrit quant à elle une composante de sauts (qui sont en partie compensés), dont l'intensité totale est représentée par la mesure $\mathrm{d}n$, qu'on appelle depuis la mesure de Lévy.

Lévy observe qu'on peut revenir maintenant sur des problèmes d'arithmétique des lois de probabilité en se restreignant au cadre des lois indéfiniment divisibles : si on ne cherche que les décompositions en facteurs indéfiniment divisibles d'une loi indéfiniment divisible μ, disons de coefficient gaussien σ^2 et de mesure de Lévy $\mathrm{d}n$, on obtient tous les facteurs en prenant un coefficient gaussien majoré par σ^2 et une mesure de Lévy $\mathrm{d}n'$ majorée par $\mathrm{d}n$. C'est une caractérisation particulièrement élégante.

Il y a un certain nombre d'éléments frappants lorsqu'on relit cette partie classique de l'oeuvre de Paul Lévy. Tout d'abord, d'un point de vue assez scolaire, le niveau de rigueur mathématique des arguments de Lévy laisse parfois à désirer ; on n'est pas toujours certain de bien comprendre les énoncés, et les preuves sont souvent loin d'être complètes. Un des mérites de K. Itô a été de reprendre les travaux de Lévy dans ce domaine et de leur donner toute la rigueur nécessaire ; voir notamment son cours à Aarhus. En revanche, on ne peut qu'être impressionné par la profondeur de l'intuition de Lévy qui explique et décrit très précisément le rôle des sauts (toutefois sans jamais avoir seulement suggéré que les trajectoires des PAI pouvaient être supposées càdlàg) et de la compensation. Même si ces éléments sont devenus classiques de nos jours, il faut se souvenir que quand Lévy a entrepris ces travaux, les outils auxquels on faisait habituellement appel en théorie des probabilités venaient

https://doi.org/10.1051/ps/2012021 Published online by Cambridge University Press

essentiellement de l'analyse de Fourier et des fonctions entières, et ce n'est pas un des moindres apports de Paul Lévy que d'avoir compris que pour résoudre un problème en apparence unidimensionnel, la bonne approche faisait appel à de l'analyse stochastique infini-dimensionnelle *via* le processus des sauts. Même si les arguments de Lévy manquent parfois de rigueur, il faut souligner que 80 ans plus tard, c'est toujours le même schéma de preuve qui est utilisé en cours de Master pour établir la formule de Lévy–Khintchine.

Pour conclure, je voudrais revenir sur la problématique qui a motivé Lévy, celle de l'arithmétique des lois de probabilités. De nos jours, disons le clairement, c'est un problème désuet, éloigné des développements modernes de la théorie des probabilités. D'autre part, nous avons vu que même si le questionnement pouvait être *a priori* naturel, le parallèle avec la théorie des nombres tourne court assez vite. Néanmoins, c'est en se laissant guider par une sorte de fil d'Ariane que Lévy dégage des concepts et des résultats qui sont toujours d'une grande modernité. Je crois qu'il s'agit là d'un élément clef de la recherche. On peut débuter un travail en espérant résoudre un problème qui au final se révélera peut être moins important que ce que l'on espérait ; mais le fait de chercher permet parfois – hélas pas toujours – de mettre la main sur ce fil d'Ariane qui conduit vers des domaines nouveaux et plus intéressants. Un dernier point enfin, une petite correction. J'ai qualifié de désuète la problématique de l'arithmétique des lois de probabilité, c'est bien sûr un peu sévère. Le résultat de Lévy sur l'arithmétique des lois indéfiniment divisibles est par exemple la clef de l'approche de Greenwood et Pitman de la factorisation de Wiener–Hopf pour les processus de Lévy. Il s'agit d'un sujet toujours d'actualité, et je recommande vivement à ceux que cela intéresse d'aller lire ce superbe article, il donne des frissons.

Références

[1] D. Dugué, Arithmétique des lois de probabilités. *Mémor. Sci. Math.* No. 137. Gauthier–Villars, Paris (1956).

[2] P. Greenwood and J. Pitman, Fluctuation identities for Lévy processes and splitting at the maximum. *Adv. Appl. Probab.* **12** (1980) 893–902.

[3] K. Itô, Stochastic Processes, *Lectures given at Aarhus University*. Springer-Verlag, Berlin (2004).

[4] P. Lévy, *Théorie de l'Addition des Variables Alatoires*. Gauthier–Villars, Paris (1937) et (1954).

[5] P. Lévy, *Oeuvres*, Vol. III. Gauthier–Villars, Paris (1976).

https://doi.org/10.1051/ps/2012021 Published online by Cambridge University Press