sciendo

DOI: 10.2478/9788367405805-032

ICAMS
The International Conference on Advanced Materials and Systems

HISTOGRAMICALLY COMPARING MATHEMATICAL SIMULATION RESULTS IN MODELING A STOCHASTIC PROCESS EXHIBITING BOTH CONTINUOUS AND DISCRETE VARIATIONS

SEBASTIAN SBÎRNĂ[1], LIANA-SIMONA SBÎRNĂ[2]

[1]*DTU Compute Department, Institute of Applied Mathematics and Computer Science, Technical University of Denmark, 1 Anker Engelunds Street, Building 101A, Kongens Lyngby, Denmark, seby.sbirna@gmail.com*

[2]*Department of Chemistry, Faculty of Sciences, University of Craiova, 107i Bucharest Way, Craiova, Romania, simona.sbirna@gmail.com*

The current paper aims to present several mathematical simulation results which have been obtained by modeling a Lévy stochastic process, exhibiting two continuous variations followed by a discrete one. In modeling such a process, we surely had to decompose it into three sub-processes: the first one is a deterministic linear drift (a time-depending simple linear function); the second one is a standard (one-dimensional) Wiener process (a stochastic process respecting three essential properties, namely: independence, stationarity and continuity); the third one is a CFD Poisson stochastic process, which is discrete (as formerly stated). The comparison was made by analyzing different sets of histograms. These sets have been plotted by histogramically comparing two jump distributions, namely Pareto distribution *vs.* normal distribution. The analyzed Lévy stochastic process, exhibiting both continuous and discrete variations, has been computed upon its discrete part, by using an appropriate skeleton structure, the main advantage of such a computation being the efficiency of it, although it is known that a weakness in using "discrete skeletons" is that the location of large jumps in the simulation model cannot be precisely identified.

Keywords: Lévy stochastic process, Wiener process, histogram.

INTRODUCTION

Lately, "Lévy processes" turn out into an interesting topic in several areas, among which: finances, economics, statistics, mathematics, physics or engineering.

Indeed, Lévy processes are often used by financial traders so as to conclusively describe the volatility of the market in the real-world, but also in order to develop risk-free theoretical scenarios. They are used in economics, for developing continuous time-series models. Also, they can be used in life-table statistics and actuarial mathematics for computing various risk parameters in insurance/reinsurance, as well as by physicists and engineers, *e.g.*, in assessing parameters to characterize different kinds of turbulence.

Lévy processes' importance was also recently demonstrated in some new fields, such as mathematical finance – as by Eberlein (2020) – that they are able to properly approximate various observations in financial market, regarding jumps or spikes of the prices, their accuracy exceeding the one of the pure Wiener processes, as shown by Barndorff-Nielsen *et al.* (2001).

In the textile and leather industry, understanding stochastic processes with both continuous and discrete variations is crucial for optimizing production processes, quality control, and predictive maintenance.

For example, in dyeing and finishing, where chemical reactions and color variations are influenced by time-dependent factors, a Lévy process model can help predict and manage these variabilities, reducing waste and improving consistency.

Additionally, in leather treatment and textile wear testing, stochastic modeling can simulate and anticipate wear patterns or defects, enhancing the ability to preemptively address quality issues and extend the lifespan of materials.

Through simulations based on accurate stochastic processes, companies can make more informed decisions and improve resource efficiency.

https://doi.org/10.2478/9788367405805-032

CHOSING THE LÉVY PROCESS FOR THE STUDY

As it is well known, the term of "Lévy process" generically refers to a motion of a point that has random successive independent displacements remaining identical across the time.

The first part of the Lévy process that we have chosen for this work is the simplest deterministic process, named "linear drift", which will be denoted in our paper as "μt".

The second part of our Lévy process is the Wiener process, which will be denoted as "σB_t", where B stands for "Brownian", because the Wiener process is often called "Brownian motion".

The third (and final) part of the chosen Lévy process is a Poisson process compound upon N_t random variables: $k = X_1, X_2, \dots, X_{N_t}$.

As always, a sum of Lévy processes is also a Lévy process, we obtain the following compound Lévy process (consisting in two continuous variations followed by a discrete one):

$$Y_t = \mu t + \sigma B_t + \sum_{k=1}^{N_t} X_k \tag{1}$$

Discrete simulation is important in studying the properties of our compound Lévy processes, as the Wiener process continuously generates random movements occurring between its jumps.

Consequently, we intend to simulate the discrete parts of the compound process by using a "discrete skeleton", $\{(P_k, A_k, M_k)\}$, which corresponds to variables called: "prior", "after" and "maximum", these variables being related to valuable information about the process (before and after a random jump, as well as for the maximum value of Y_t by the time of its k^{th} jump).

INTRODUCING THE SIMULATION MODEL

As previously discussed, the simulation model will be not continuous, but discrete, as the values of the Lévy process are collected in a "discrete skeleton", $\{(P_k, A_k, M_k)\}$.

Considering the expression (1) and assuming discrete time – following Applebaum (2004):

$$T \sim \exp(\lambda) \tag{2}$$

one can further define:

$$V = \max_{0 \le t \le T} \mu t + \sigma B_t \tag{3}$$

and

$$W = \left(\max_{0 \le t \le T} \mu t + \sigma B_t\right) - (\mu T + \sigma B_t) \tag{4}$$

The discrete simulation can be performed according to an appropriate algorithm, that will be described in what follows.

ALGORITHM PROPOSED TO SIMULATE THE CHOSEN LÉVY PROCESS

The algorithm will be proposed according to the previously presented purpose.

Algorithm: Simulation of the Lévy process, assuming discrete time

Set $P_0, A_0, M_0 = 0, 0, 0$

for iteration, $k = 1, 2, \ldots, n$ *do*

Simulate V_k, W_k, X_k

Set:

$$P_k = A_{k-1} + (V_k - W_k)$$
$$A_k = A_{k-1} + (V_k - W_k) + X_k$$
$$M_k = \max\{M_{k-1}, A_{k-1} + V_k, A_{k-1} + (V_k - W_k) + X_k\}$$

end for

The code implementation in the algorithm above is written using the R programming language.

In the settings of this simulation, X_k are independent and identically distributed (i.i.d). random variables having the distribution F, which are defined as the size of the k^{th} jump, V and W are two normally distributed independent random variables, whose intensities are:

$$\phi_1 = \sqrt{\frac{\mu^2}{\sigma^4} + \frac{2\lambda}{\sigma^2}} - \frac{\mu}{\sigma^2} \tag{5}$$

and

$$\phi_2 = \sqrt{\frac{\mu^2}{\sigma^4} + \frac{2\lambda}{\sigma^2}} + \frac{\mu}{\sigma^2} \tag{5'}$$

respectively.

The function to be implemented has the following arguments: n (total number of iterations), μ, σ, λ (parameters from which it becomes possible to compute the intensities given above (ϕ_1 and ϕ_2), a (threshold to be discussed in estimating first-passage probabilities), X (argument that selects which is the type of distribution for the jump components), and "jump_size" (mean of the distribution).

There are three types of return values for this function: "final" – for the values of the "skeleton" at the iteration n, so $\{(P_n, A_n, M_n)\}$; "path" – for the values of the "skeleton" at each iteration, k and "first passage" – for a true/false value which depends on exceeding or not the value of constant a (introduced above).

As far as simulation of jumps while keeping fixed parameters, a set of 800 simulation runs will be performed, so as to later prove the variability of a simulation upon twenty simulation runs (keeping the same parameters).

First simulation is run for fixed parameters μ, σ, λ and assuming all the values for jump_size to be equal to the unity and F is chosen as a normal distribution.

LÉVY PROCESS' EVOLUTION

The values of A and M are approximately identical; the support of the normal distribution belongs to the interval $[0, \infty)$, so values of jump_size are either null or strictly positive, whereas $A_k \geq P_k$.

There cannot be identified any predominant components in terms of absolute values, so

the studied process should mostly be characterized by the positive linear drift and the values of jump_size. From the set of 800 simulation runs, the difference between the simulations runs seems to be more noticeable within the interval $[480, 500]$.

LÉVY PROCESS' DECOMPOSITION AND COMBINING THE DISCRETE PARTS

As stated before, the composed Lévy process can be decomposed into three sub-processes, the first one being a deterministic linear drift (μt) – which is continuous, the second one being a stochastic Wiener process (σB_t) – which is also continuous and the third one being a stochastic Poisson process ($\sum_{k=1}^{N_t} X_k$), – which is discrete, as presented by Costa (2023).

As far as the first part is concerned, as the name itself suggests, the linear drift is a simple linear function of the time, with coefficient $\mu \in \mathbb{R}$.

As far as the second part is concerned, t is well known that a stochastic process $W(t)$ is called Wiener process if it respects the following three properties: independence, *i.e.*, $[W(t+\Delta t) - W(t)]$ does not dependent on $W(\tau)$ for any $\tau \leq t$; stationarity, *i.e.*, the distribution of $[W(t+\Delta t) - W(t)]$ does not depend on t and continuity, *i.e.*:

$$\log_{\Delta t \to 0} \frac{P(|W(t+\Delta t) - W(t)| \geq \delta)}{\Delta t} = 0 \tag{6}$$

for $\delta > 0$, as proved by Kim and Park (2024).

As far as the third part is concerned, the Poisson process N_t – with its intensity $\lambda \geq 0$ – can be used to define processes having non-continuous increments, in terms of the number of events in non-overlapping intervals:

$$P(N(t) = N) = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \tag{7}$$

this process occurring at arrival times $S_1, S_2, \dots, S_n$ and interarrival times $T_1, T_2, \dots, T_n$, where the connection between arrival and interarrival times is that S is the sum of T, *i.e*:

$$S_n = \sum_{k=1}^{n} T_k \tag{8}$$

In order to avoid an inefficient simulation of the Wiener process, the composed Lévy process would be better described with the two discrete parts (linear drift and Wiener process) combined together.

They could be looked at as one discrete function, which would be defined as $[V_k - W_k]$, describing the movement of Y_t between jumps X_k and X_{k-1} – *i.e.*, between T_k and T_{k-1}.

At one side, as between consecutive jumps the combined component $\{\mu t + \sigma B_t\}_{t \geq 0}$ follows a volatile pathway as a consequence of the stochasticity of B_t component, the maximum value of the combined component within the interval $[T_k - T_{k-1}]$ will be denoted as V_k:

$$V_k = \max(\mu t + \sigma B_t), \qquad t \geq 0 \tag{9}$$

V_k being the value of the combined component $\{\mu t + \sigma B_t\}_{t \geq 0}$ exactly at the moment t_{max} (with $0 = T_{k-1} \leq t_{max} \leq T_k$) in which V_k reached its maximum value within the interval – *n.b.* that, for every interval, t is set to be null.

On the other side, W_k is defined as:

$$W_k = \max(\mu t + \sigma B_t) - (\mu T + \sigma B_T), \quad t \geq 0 \tag{10}$$

Consequently, the change of Y_t within the time interval $[T_k - T_{k-1}]$, which might be attributed to the continuous part of the compound process, is equal to $V_k - W_k$, which means:

$$\mu T + \sigma B_T = V_k - W_k \tag{11}$$

INTERPRETING THE "DISCRETE SKELETON"

Within the "discrete skeleton", $\{(P_k, A_k, M_k)\}$, the value P_k (prior to the next jump X_k) is obtained by adding to the value A_{k-1} (after the last jump, X_{k-1}) the change caused by the continuous process between jump X_k and X_{k-1} (represented as the difference between the values V_k and W_k):

$$P_k = \underbrace{A_{k-1}}_{Y_t \text{ after } X_{i-1}} + \underbrace{(V_k - W_k)}_{\text{linear+Wiener}} = A_{k-1} + V_k - W_k \tag{12}$$

The value A_k, obtained after X_k, is equal to the sum between the value P_k, registered before the jump, and the size of the jump, X_k:

$$A_k = P_k + X_k = A_{k-1} + (V_k - W_k) + X_k = A_{k-1} + V_k - W_k + X_k \tag{13}$$

The value M_k is the maximum value taken by Y_t during the k^{th} jump, meaning that it is equal to the maximum value of all points in the interval $[0, S_k]$ registered within the respective time period.

$$M_k = \max(M_{k-1}, A_{k-1} + V_k, A_{k-1} + V_k - W_k + X_k) \tag{14}$$

TESTING PARETO DISTRIBUTION *VS.* NORMAL DISTRIBUTION

First-passage probability is defined for a fixed value, $a \geq 0$; in this case, it is the probability that the compound Lévy process ever exceeds a; as to estimate this probability, we shall compare the value a to the maximum value of this process – *i.e.*, to the M coordinate of the "discrete skeleton" $\{(P_k, A_k, M_k)\}_{k\in\mathbb{N}}$, so if M_k does not reach level a (whatever value it can take), we can conclude that the path cannot ever go beyond a – as remarked by Kındap and Godsill (2024).

The compound Lévy process was tested for two distributions of the values of jump_size Y_k, in each case the mean jump_size being equal to the unit. Six histograms were drawn.

Namely, we histogramically have compared – by recording the values found in each small interval – the first-passage probabilities for a Pareto distribution (with a shape parameter equal to 1.04) to a normal distribution (with a standard deviation equal to 10) – *n.b.*, the normal distribution is the case in which the Wiener process is neither negatively-oriented nor dominant.

We used these two distributions to generate different values of jump_size of the Poisson process, in order for them to serve as input data in the simulation of the Lévy process.

We shall show here, in Figures 1-3, the resulting first-passage probability for each member of the "discrete skeleton" $\{(P_k, A_k, M_k)\}_{k\in\mathbb{N}}$, by means of three sets of histograms, recorded from a hundred values of $\{(P_{800}, A_{800}, M_{800})\}$, corresponding to a Pareto and, respectively, to a normal distribution, in each figure being involved just one member of the "discrete skeleton".

For simplicity, in both cases, we have kept the same value for the parameters $\{\mu, \sigma, \lambda\}$, all of them having the same value, $\mu = \sigma = \lambda = 1$.

It was easily proved that the Pareto distribution exhibits a smaller variance, whereas only the normal distribution can take negative values, meaning that in this case the Lévy process can have both positive or negative jumps, *i.e.*, just for the normal distribution, we could sometimes find cases in which $A_k < A_{k-1}$. Moreover, the values of A can be significantly different from the ones of M only in the normal distribution (and this happens, again, because this is the only one in which the jumps can be negative).

One may also observe that, whereas all the three Pareto distribution histograms are quite asymmetric (appearing in the range $[1100, 1540]$ and having the majority of the values in the range [1210, 1430] and a strong peak between 1210 and 1320, the results are more evenly distributed for the normal distribution histograms: they appear in the range [1200, 2860] for P and A and [1310, 2860] for M, being mainly represented in the range [1450, 2440].

The histograms previously presented are considered to be representative, because no other component of the Lévy process is dominant, so they have proved that variance differences are remarkable from the histograms themselves, consequently limiting all the discussion only to a single run of simulations.

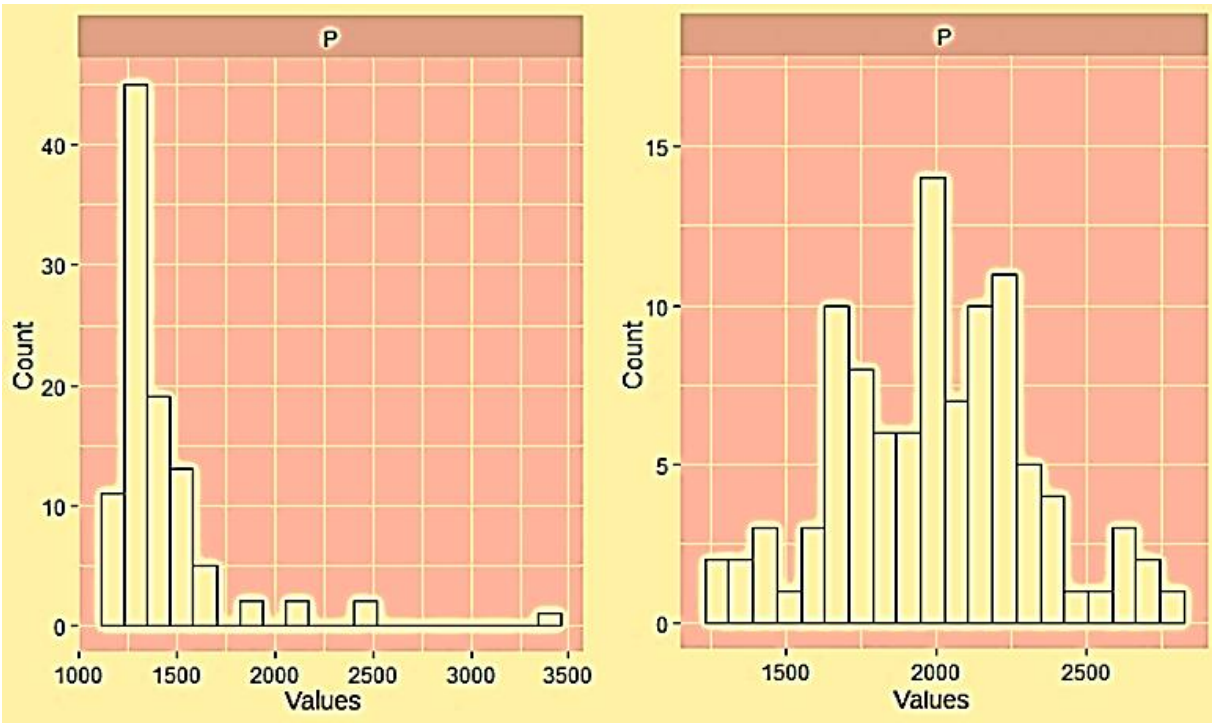

Figure 1. Histograms recorded for P from a hundred values of $\{(P_{800}, A_{800}, M_{800})\}$: Pareto distribution (left) *vs*. normal distribution (right), with all parameters presented above

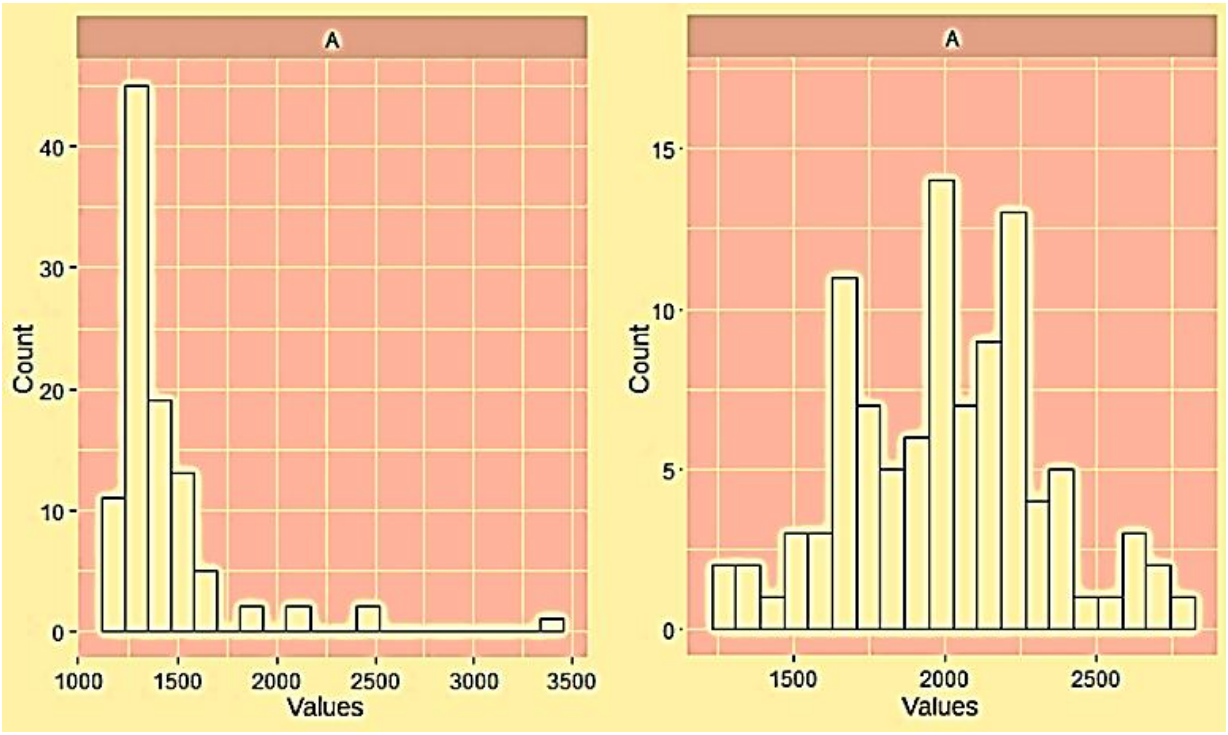

Figure 2. Histograms recorded for A from a hundred values of $\{(P_{800}, A_{800}, M_{800})\}$: Pareto distribution (left) *vs*. normal distribution (right), with all parameters presented above

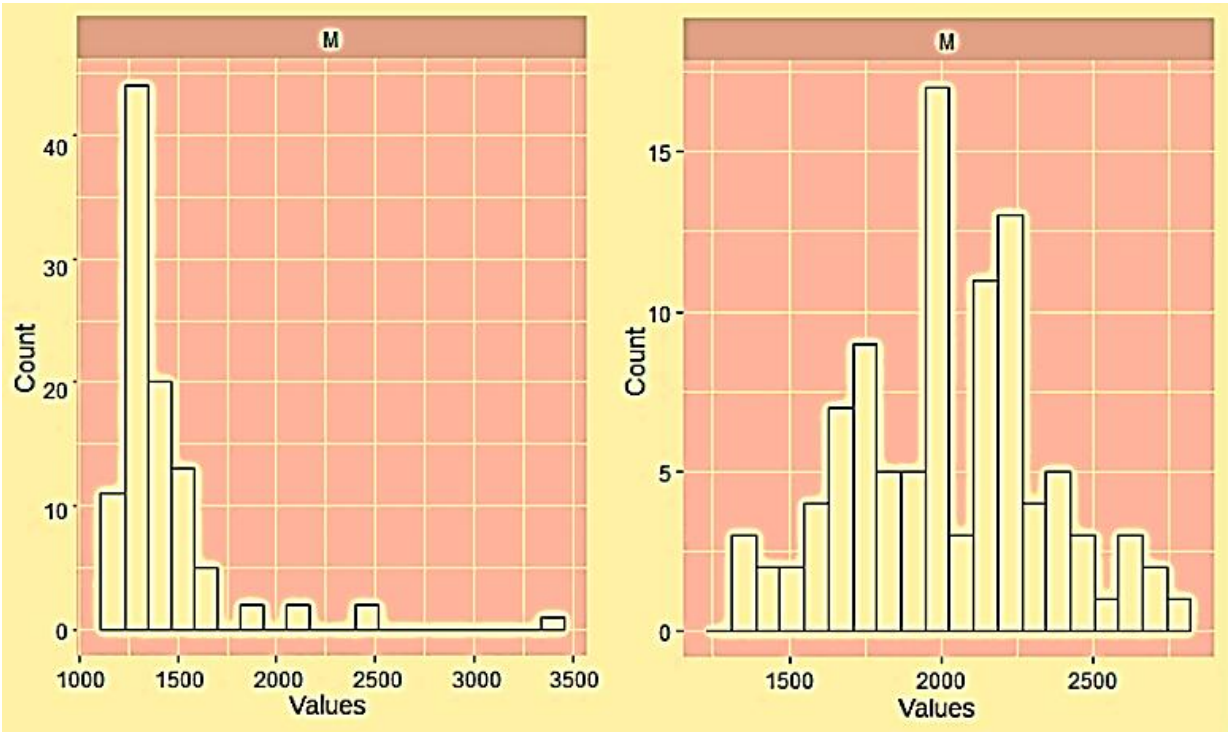

Figure 3. Histograms recorded for M from a hundred values of $\{(P_{800}, A_{800}, M_{800})\}$: Pareto distribution (left) *vs*. normal distribution (right), with all parameters presented above

CONCLUSION

Within this work, taking into account how difficult it can sometimes be to simulate the

continuous Wiener process, a compound Lévy process was computed upon its discrete parts, by using an efficient "discrete skeleton" structure.

The main benefit of this procedure is the computation efficiency, although a drawback of using "discrete skeletons" consists in the fact that the location of large jumps cannot be identified with precision within this mathematical simulation.

We have demonstrated that the Wiener process component inside the composed Lévy process is the one responsible for determining the variance of Y_t.

The values taken by the first-passage probability during the Lévy process were collected in a "discrete skeleton", $\{(P_k, A_k, M_k)\}$.

These correspond to three variables meaning – as previously mentioned: "prior", "after" and "maximum" (all of them giving useful information regarding the process at different times, either before or after a random jump of the process).

When experimenting with two different distributions of the i.i.d. variables X_k ($1 \leq X_k \leq N_t$) which can characterize the stochastic Lévy process, namely Pareto *vs*. normal distribution, we have observed some interesting aspects about the values that can be taken by the jumps in each case.

Finally, by analyzing any potential differences that could occur in variance given by these two statistical distributions, we have found that the presented histograms are representative for any of the three members of the "discrete skeleton", $\{(P_k, A_k, M_k)\}$, as no other component of the process is dominant, leading to the conclusion that variance differences are noticeable from the graphs themselves, so the discussion can be successfully restricted to a single simulation run.

REFERENCES

Applebaum, D. (2004). *Lévy Processes and Stochastic Calculus*. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511755323

Barndorff-Nielsen, O.E., Resnick, S.I. & Mikosch, T. (2001). *Lévy Processes*. Birkhäuser, Boston, Basel, Berlin. https://doi.org/10.1007/978-1-4612-0197-7_11

Costa, M.T., Kontoyiannis, I. & Godsill, S. (2023). Generalised Shot Noise Representations of Stochastic Systems Driven by Non-Gaussian Lévy Processes. arXiv. https://doi.org/10.48550/arXiv.2305.05931

Eberlein, E., Gerhart, C. & Lütkebohmert, E. (2020). A Multiple Curve Lévy Swap Market Model. *Applied Mathematical Finance*, *27*(5), 396–421. https://doi.org/10.1080/1350486X.2021.1877559

Kim, K.H. & Park, D. (2024). A Sobolev Space Theory for Time-Fractional Stochastic Partial Differential Equations Driven by Lévy Processes. *Journal of Theoretical Probability*, *37*, 671–720. https://doi.org/10.1007/s10959-023-01263-8

Kındap, Y. & Godsill, S. (2024). Point Process Simulation of Generalised Hyperbolic Lévy Processes. *Statistics and Computing*, *34*(33), 1–28. https://doi.org/10.1007/s11222-023-10344-x