UNIVERSITY OF THE AEGEAN

SCHOOL OF SCIENCES
DEPARTMENT OF STATISTICS AND ACTUARIAL-FINANCIAL MATHEMATICS

«Lévy Processes and Risk Theory»

Master Thesis for the Graduate Programme

This Thesis is submitted in partial fulfilment
of the requirements for the award of the
correspondingly titled master degree in

Statistics and Actuarial-Financial Mathematics

Ioannis Antoniadis

May 2023

SAMOS

Ioannis Antoniadis

Lévy Processes and Risk Theory

May 2023

Master Thesis for the Graduate Programme

Department of Statistics and Actuarial-Financial Mathematics

Author: .....

Advisor: .....

Committee Member: .....

Committee Member: .....

SAMOS

Abstract

This thesis is dedicated to the study of Lévy processes and their applications to Risk theory. A quick and accurate recount of some prerequisites and notions is provided at the start. Following that a comprehensive study regarding the Lévy processes follows. The first part ends with the showcasing of some representative Lévy processes. In the second part, the notions of Risk theory are presented and are linked to Lévy processes via a recount of the necessary theory. Everything comes together through the inclusion of some applications utilizing real world data. Additional toy examples and simulations are included, when appropriate, in the entirety of this thesis. The thesis concludes with a brief discussion focusing on the theory and material covered, as well as, personal opinions regarding avenues for future research on Lévy processes.

Keywords: Lévy processes, Lévy Jump-Diffusion models, Risk theory, Asset price modeling

Contents

List of Figures

Chapter 1

Lévy Processes

The aim of this chapter is to provide an introduction to the class of stochastic processes known as Lévy processes. At the start of the chapter, the depth of the field is briefly explored starting from the early years that the term was coined till the new advancements made throughout the recent years. Following that, some literature review was conducted, in order to determine the penetration of Lévy processes in other scientific fields. Up to this point, strict, mathematical definitions are avoided in order to quickly explore and convey the necessary ideas as quickly as possible yet concise manner.

Moving on, we deep dive in Lévy Processes, exploring the necessary requirements for a process to be classified as Lévy process. Additionally, many of the known Lévy processes are further explored in terms of their own properties, as well as providing some simulations of said processes with the resulting random realizations. Lastly, the influential theorems surrounding Lévy processes are explored, along with their proofs. Mathematically rigorous definitions are needed and used in this segment in order to successfully convey the entirety of the concepts covered in depth and breadth.

1.1 Brief history on Lévy Processes

Lévy Processes are a class of stochastic processes that satisfy certain conditions. As a concept, they were conceived during the early years of the 20th century by the French mathematician Paul Lévy. As a scientist, his goal was providing some framework or model, which would be capable of modelling complex dynamical systems or (real life) random phenomena with as much accuracy as possible.

Lévy pioneered the research around Lévy Processes with his introduction of stable distributions in probability theory (Lévy, 1925). In the upcoming years, he focused his researching efforts in Brownian motion (Brown, 1828). However, Lévy came to realize that a Brownian motion's sample path is not differentiable, thus becoming potentially limiting. The solution Lévy reached was to come up with a generalization of the Brownian motion that would incorporate jump discontinuities (Lévy, 1948). This proposed framework and its accompanying conditions are what is being used today to classify a stochastic process as a Lévy process.

A next big milestone for Lévy processes happened in 1965, when the geometric Brownian motion was introduced (Samuelson, 1965). From 1970 and on wards Lévy processes are starting to play a big role in stochastic finance (Schoutens, 2003), especially after the work of Samuelson and, consequently, Black and Scholes with their attempt to model the prices of options in financial markets using geometric Brownian motion (Black and Scholes, 1973). Using geometric Brownian motion in a finance context could be thought as the continuation of Bachelier's thesis (Bachelier, 1900).

Afterwards, Lévy processes are starting to be applied more heavily in more aspects of finance and economics, such as portfolio optimization models, econometrics and time series and option pricing models (Tankov and Cont, 2003; Bates, 1996; Bates, 1991; Mantegna and Stanley, 2000; Heston, 1993; Barndorff-Nielsen and Shephard, 2001). A mention has to be made about the application of Lévy processes in risk management (Acerbi and Tasche, 2002; Embrechts et al., 1997; Giesecke, 2003) and operational risk (Böcker and Klüppelberg, 2010).

Lastly, it is also of significance to note that Lévy processes have inspired the conception of other models in various fields of economics and finance. In asset pricing models for example, the limitations of Lévy processes have given birth to the Variance Gamma model (Madan and Seneta, 1990) or the Merton model (Merton, 1976). In time series analysis, the inability to successfully model long memory (long range dependence) has resulted in the emergence of models such as the Fractionally Integrated Lévy Motion model (Basseville and Nikiforov, 1993) or the Fractionally Integrated Geometric Brownian Motion model (FIGARCH) (Engle and Russell, 1998). Other examples include the Fractional Brownian motion (Mandelbrot and Ness, 1968), or the exploration of the connection between Lévy processes and multifractal theory (Jaffard, 1999). These can be used alongside the typical Lévy process model or even as an alternative.

However, one of the major disadvantages of adopting Lévy processes as a potential solution or part of it, has always been their need for extensive computational power in order to provide accurate estimation for the parameters of the used Lévy process. With the advancements of technology and science this disadvantage has been mitigated considerably via the introduction of new algorithms and techniques that aim to more accurately estimate the parameters of Lévy processes (Broadie and Kaya, 2006; Lewis, 2001; Fournier, 2009; Al-Saadony, 2016). Additionally, there has been on going research in order to discover simple, analytical solutions to these stochastic models, without the need for explicit computer simulations (Barndorff-Nielsen and

Shephard, 2002; Klüppelberg et al., 2004).

1.2 Literature review of some Lévy Processes' applications

While this thesis is dedicated to Lévy processes in terms of risk theory, in order to explore the breadth of their applications as much as possible, we explore various applications of Lévy processes in other scientific fields.

Lévy processes were adopted by other scientific fields as well, in order to model various complex problems of stochastic nature. A few applications spanning many different scientific fields could be the following arbitrarily picked applications. Nevertheless, it should be representative of the capabilities the Lévy processes have showcased throughout their history in modelling various phenomena with varying scale and complexity.

1.2.1 Computer Science

In computer science these processes have been used to deal with dynamic systems and networks such as client-server queuing problems (Peng and Wu, 2020). Additionally, they have been used in robotics in order to identify targets in a given environment (Katada et al., 2022). Other applications for Lévy processes can be found in many well established computational optimization algorithms (Li et al., 2022).

1.2.2 Physics

In physics they have been used as a medium to model the behavior of various physical systems (Metzler et al., 2009). Another example is the Lévy flight superdiffusion and its applications to barrier crossing problems (Dubkov et al., 2008), or the use of Lévy-driven models to successfully model the movement of particles (Baeumer and Meerschaert, 2010), or to model turbulent flow instances (Shlesinger et al., 1987), or overall transport problems (Zaburdaev et al., 2015).

1.2.3 Engineering

In engineering, Lévy processes have been used for a variety of applications such as the driving force minimizing the variables for a given engineering problem (Rather and Bala, 2021), to model the process of degradation of various systems (Shi et al., 2020). Other applications areas include general storage and reliability problems (Abdel-Hameed, 2014), as well as overall modeling queuing systems (Kyprianou, 2014).

1.2.4 Biology

Biology had them model the patterns of animals during foraging and their search for sustenance extensively (Viswanathan et al., 2011; Bartumeus and Catalan, 2009; Wosniack et al., 2017), though later research works seem to rule this theory out (James et al., 2011; Pyke, 2014). Other example applications are their use to model metastatic cancer cells' movement (Huda et al., 2018), the interaction between environmental factors and day to day movement of animals (de Jager et al., 2011). Some additional applications include modeling the behavior of human travel (Brockmann et al., 2006) and modeling the dynamics of population from a space and time perspective (Renshaw, 1991).

1.2.5 Artificial Intelligence

Machine learning is another field that has benefited by applying Lévy processes to certain problems. For example, the use of Lévy processes, specifically Gaussian processes, as priors to achieve better results in a spectral kernel learning framework (Jang et al., 2017). Additionally we can turn to the work of Zhang and Tao who study risk bounds of the learning process for time-dependent samples drawn from a Lévy process (Zhang and Tao, 2013).

1.3 Notation and prerequisites

In order for this thesis to be as self-contained as possible, we introduce and define some introductory concepts.

Various definitions, theorems and applications found in this thesis build upon these prerequisites, thus their inclusion is ought to be welcomed.

1.3.1 Sample space

Definition 1.3.1 (Sample space)**.** A sample space (symbolised as Ω) is the collection of the possible outcomes $\omega_1, \omega_2, \ldots, \omega_n \in \Omega$ of an experiment or trial ϵ, assuming the following conditions also hold:

- If ω_i and ω_j are two outcomes of the same experiment ϵ, then only one can occur per repetition of the experiment ϵ, $\forall i \neq j$.
- If $\omega_1, \omega_2, \ldots, \omega_n$ are the outcomes of a sequence of trials $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$, then $\omega_1, \omega_2, \ldots, \omega_n \in \Omega$.

Some examples follow.

Example 1.3.1 (Coin toss - Sample space)**.** *A classic example of a random trial is the coin toss. A fair coin is tossed and the outcome is recorded. We are interested in the sample space of such a trial.*

We know that tossing a coin results in 2 possible outcomes: heads or tails (which we will symbolize as H or T respectively). Then the sample space is:

$$\Omega = \{H, T\}$$

Example 1.3.2 (Die roll - Sample space)**.** *Let us imagine the classic trial of a die roll. A fair 6-sided die is rolled and the outcome is recorded. The sample space of this trial would be:*

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

Rolling a 6-sided die cannot produce an outcome different from the ones shown above. Thus, if for example $\omega = 3$, then $\omega \in \Omega$.

1.3.2 σ-Algebra

Definition 1.3.2 (σ-Algebra)**.** If $\mathcal{F}$ is a σ-algebra (or a σ-field) of a set Ω then $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ and the following conditions must hold:

- $\Omega \in \mathcal{F}$
- For any set $A \in \mathcal{F}, \text{then} A^c \in \mathcal{F}$
- For any sequence $\{A^n\} \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i$

Where $\mathcal{P}(\Omega)$ is the powerset of Ω. Some examples follow.

Example 1.3.3 (Coin toss - σ-Algebra)**.** *Based on the previous coin toss example,*

we expand the example accordingly. We are interested in an appropriate σ-Algebra for the coin toss trial. Such a σ-Algebra is:

$$\mathcal{F} = \mathcal{P}(\Omega) = 2^{\Omega} = \{\emptyset, \{H\}, \{T\}, \{H, T\}\}$$

It is obvious at first glance that this σ-Algebra satisfies all the necessary properties.

Example 1.3.4 (Die roll - σ-Algebra)**.** *Similarly we expand this example appropriately as well. Once again, we are interested in constructing an appropriate σ-Algebra for the die roll trial as described previously.*

Thus we end up with 64 items in our σ-Algebra, namely:

$$\mathcal{F} = \left\{ \begin{array}{l} \{\emptyset\}, \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}, \{1,2\}, \{1,3\}, \{1,4\}, \{1,5\}, \{1,6\}, \\ \{2,3\}, \{2,4\}, \{2,5\}, \{2,6\}, \{3,4\}, \{3,5\}, \{3,6\}, \{4,5\}, \{4,6\}, \{5,6\}, \\ \{1,2,3\}, \{1,2,4\}, \{1,2,5\}, \{1,2,6\}, \{1,3,4\}, \{1,3,5\}, \{1,3,6\}, \\ \{1,4,5\}, \{1,4,6\}, \{1,5,6\}, \{2,3,4\}, \{2,3,5\}, \{2,3,6\}, \{2,4,5\}, \\ \{2,4,6\}, \{2,5,6\}, \{3,4,5\}, \{3,4,6\}, \{3,5,6\}, \{4,5,6\}, \{1,2,3,4\}, \\ \{1,2,3,5\}, \{1,2,3,6\}, \{1,2,4,5\}, \{1,2,4,6\}, \{1,2,5,6\}, \{1,3,4,5\}, \\ \{1,3,4,6\}, \{1,3,5,6\}, \{1,4,5,6\}, \{2,3,4,5\}, \{2,3,4,6\}, \{2,3,5,6\}, \\ \{2,4,5,6\}, \{3,4,5,6\}, \{1,2,3,4,5\}, \{1,2,3,4,6\}, \{1,2,3,5,6\}, \\ \{1,2,4,5,6\}, \{1,3,4,5,6\}, \{2,3,4,5,6\}, \{1,2,3,4,5,6\} \end{array} \right\}$$

Once again, it is easily observable that our constructed σ-Algebra satisfies all necessary properties.

1.3.3 Probability measure

Definition 1.3.3 (Probability measure)**.** A probability measure, often denoted as $\mathbb{P}$, is a real-valued function on the unit interval $[0, 1]$ that is defined on events (sets of outcomes) of the sample space Ω. Or simply:

$$\mathbb{P} : \Omega \to [0, 1]$$

Naturally, for a function $\mathbb{P}$ to be a probability measure, some properties must simultaneously hold:

- $\mathbb{P}(\emptyset) = 0$ and $\mathbb{P}(\Omega) = 1$
- If $A \subseteq \Omega, 0 \leq \mathbb{P}(A) \leq 1$
- $\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$ for all pairwise disjoint subsets of Ω

Intuitively, a probability measure function like any ordinary measure function can be attributed as an attempt to quantify and provide some sense of volume for a given set or subsets of that set.

1.3.4 Probability space

Definition 1.3.4 (Probability space)**.** A probability space or probability triple is a mathematical model that is based on a sample space Ω, a σ-Algebra and an appropriate probability measure function. It is denoted as $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 1.3.5 (Probability space completeness)**.** A probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called complete if the probability measure $\mathbb{P}$ is extended in such a way that any set $A \subset B$, while $\mathbb{P}(B) = 0$, is also included on the σ-Algebra $\mathcal{F}$ and is assigned $\mathbb{P}(A) = 0$.

$\mathcal{F}^{\mathbb{P}}$ is the completion of a σ-Algebra $\mathcal{F}$ with respect to the probability measure $\mathbb{P}$ if $\mathcal{F} = \mathcal{F}^{\mathbb{P}}$ in our probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 1.3.6 (Filtered probability space)**.** A filtered probability space consists of the usual probability triple $(\Omega, \mathcal{F}, \mathbb{P})$ while adding a filtration $\mathcal{F}_t$ and is denoted as $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. A filtration is an sequentially increasing family of $\mathcal{F}$ according to the order of an ordered index set.

Usually, the time index set is used or more generally a subset of $\mathbb{R}$. In such cases, it is beneficial for the practitioner to add further conditions for a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to be considered filtered. The conditions are as follows:

- The utilized probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is complete - as defined earlier
- $\forall \mathbb{P}$-null sets $\in \mathcal{F}_t \subseteq \mathcal{F}$
- Let T be the index set of choice, then for any $t_1 < t_2 \in T \subset \mathbb{R}$, it holds that:

$$\mathcal{F}_{t_2+} \equiv \bigcap_{t_1 > t_2} \mathcal{F}_{t_1} = \mathcal{F}_{t_2}$$

A filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}^{\mathbb{P}}, \mathbb{P})$, also called stochastic basis, is considered the necessary and fundamental building block on top of which all stochastic processes may be defined and further studied.

1.3.5 Characteristic function

Definition 1.3.7 (Characteristic function)**.** The characteristic function of a $\mathbb{R}$-valued random variable X is a different way of representing the probability density function admitted by said random variable, using the Inverse Fourier Transform method. It is a $\mathbb{C}$-valued function and defined as:

$$\phi_X(\omega) = \mathbb{E}\left[e^{\iota\omega X}\right] \tag{1.1}$$

where ι and ω are the imaginary unit and real number respectively.

If any two random variables can be expressed by the same characteristic function then they are distributed according to the same probability density function.

The biggest motivation for utilizing the characteristic function of a random variable is that one and only one such function always exists for any random variable. It is sometimes the case that distributions have no defined density function or moment generating function, like the Cauchy distribution.

A characteristic function may provide a feasible solution to problems such as the one above, as it is shown on the following examples.

Example 1.3.5 (Characteristic function of $X \sim Cauchy(\chi_0, \gamma)$)**.** *We are interested in calculating the characteristic function of a standard Cauchy distribution.*

To derive the characteristic function for $Cauchy(\chi_0, \gamma)$, we first start with the characteristic function for standard $Cauchy(\chi_0 = 0, \gamma = 1)$. Let us assume its characteristic function is:

$$\phi_X(\omega) = e^{-|\omega|} \tag{1.2}$$

The next step is applying the inverse Fourier transform to equation 1.2:

$$\begin{aligned}
F^{-1}\left(\phi_X(\omega)\right) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \phi_X(\omega) e^{\iota\omega X} d\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-|\omega|} e^{\iota\omega X} d\omega \\
&= \frac{1}{2\pi}\left(\int_{-\infty}^{0} e^{\omega} e^{\iota\omega X} d\omega + \int_{0}^{+\infty} e^{-\omega} e^{\iota\omega X} d\omega\right) \\
&= \frac{1}{2\pi}\left(\int_{-\infty}^{0} e^{\omega+\iota\omega X} d\omega + \int_{0}^{+\infty} e^{-\omega+\iota\omega X} d\omega\right) \\
&= \frac{1}{2\pi}\left(\frac{e^{(1+\iota X)\omega}}{1+\iota X}\bigg|_{-\infty}^{0} + \frac{e^{-(1-\iota X)\omega}}{-1+\iota X}\bigg|_{0}^{+\infty}\right) \\
&= \frac{1}{2\pi}\left(\frac{1}{1+\iota X} - \frac{1}{-1+\iota X}\right) \\
&= \frac{1}{2\pi}\frac{-2}{(1+\iota X)(-1+\iota X)} \\
&= \frac{1}{\pi(1+\iota X)(1-\iota X)} \\
&= \frac{1}{\pi(1^2-(\iota X)^2)} \\
&= \frac{1}{\pi(1+X^2)}
\end{aligned}$$

We may now proceed to the general case of $Cauchy(\chi_0, \gamma)$ by utilizing equation

1.1:

$$\begin{aligned}\phi_X(\omega) &= \mathbb{E}\left[e^{\iota\omega X}\right] \\ &= \int_{\chi_0}^{\gamma} e^{\iota\omega X} \frac{1}{\pi\gamma\left(1+\left(\frac{X-\chi_0}{\gamma}\right)^2\right)} dX\end{aligned}$$

The following transformations are needed in order to successfully evaluate the above integral:

$$\begin{aligned}X &= \gamma y + \chi_0 \\ dX &= \gamma dy\end{aligned}$$

Thus we end up with:

$$\begin{aligned}\int_{\chi_0}^{\gamma} e^{\iota\omega X} \frac{1}{\pi\gamma\left(1+\left(\frac{X-\chi_0}{\gamma}\right)^2\right)} dX &= \int_{-\infty}^{+\infty} e^{\iota\omega(\gamma y+\chi_0)} \frac{\gamma}{\pi\gamma(1+y^2)} dy \\ &= \int_{-\infty}^{+\infty} e^{\iota\omega\gamma y+\iota\omega\chi_0} \frac{1}{\pi(1+y^2)} dy \\ &= \int_{-\infty}^{+\infty} e^{\iota\omega\gamma y} e^{\iota\omega\chi_0} \frac{1}{\pi(1+y^2)} dy \\ &= e^{\iota\omega\chi_0} \int_{-\infty}^{+\infty} e^{\iota\omega\gamma y} \frac{1}{\pi(1+y^2)} dy \\ &= e^{\iota\omega\chi_0} e^{-|\gamma\omega|} \\ &= e^{\iota\omega\chi_0 - |\gamma\omega|}\end{aligned}$$

At this point, the characteristic function of $Cauchy(\chi_0, \gamma)$ has been successfully derived.

Example 1.3.6 (Characteristic function of $X \sim Uniform(\alpha, \beta)$)**.** *We are interested in calculating the characteristic function of a Uniform distribution.*

The probability density function is:

$$U(\alpha, \beta) = \frac{1}{\beta - \alpha}$$

We continue by utilizing the 1.1 equation substituting appropriately for the con-

tinuous case with the $Uniform(\alpha, \beta)$ PDF:

$$\begin{aligned}
\phi_X(\omega) &= \mathbb{E}\left[e^{\iota\omega X}\right] \\
&= \int_\alpha^\beta e^{\iota\omega X}\frac{1}{\beta-\alpha}dX \\
&= \frac{1}{\beta-\alpha}\int_\alpha^\beta e^{\iota\omega X}dX \\
&= \frac{e^{\iota\omega\beta}}{(\beta-\alpha)\iota\omega} - \frac{e^{\iota\omega\alpha}}{(\beta-\alpha)\iota\omega} \\
&= \frac{e^{\iota\omega\beta}-e^{\iota\omega\alpha}}{(\beta-\alpha)\iota\omega}
\end{aligned}$$

Thus, we arrive at the known characteristic function for the $Uniform(\alpha, \beta)$.

Example 1.3.7 (Characteristic function of $X \sim Laplace(\mu, b)$)**.** *For our last example for this segment, we are, similarly, interested in deriving the characteristic function for $X \sim Laplace(\mu, b)$:*

$$\begin{aligned}
\phi_X(\omega) &= \mathbb{E}\left[e^{\iota\omega X}\right] \\
&= \int_{-\infty}^{+\infty} e^{\iota\omega X}\frac{1}{2b}e^{-\frac{|X-\mu|}{b}}dX \\
&= \int_{-\infty}^{\mu} e^{\iota\omega X}\frac{1}{2b}e^{\frac{X-\mu}{b}}dX + \int_{\mu}^{+\infty} e^{\iota\omega X}\frac{1}{2b}e^{-\frac{X-\mu}{b}}dX \\
&= \frac{1}{2b}\int_{-\infty}^{\mu} e^{\iota\omega X+\frac{X-\mu}{b}}dX + \frac{1}{2b}\int_{\mu}^{+\infty} e^{\iota\omega X-\frac{X-\mu}{b}}dX \\
&= \frac{1}{2b}\int_{-\infty}^{\mu} e^{\frac{\iota b\omega X-X+\mu}{b}}dX + \frac{1}{2b}\int_{\mu}^{+\infty} e^{\frac{\iota b\omega X+X-\mu}{b}}dX \\
&= \frac{1}{2b}\int_{-\infty}^{\mu} e^{\frac{X(\iota b\omega+1)-\mu}{b}}dX + \frac{1}{2b}\int_{\mu}^{+\infty} e^{\frac{-X(\iota b\omega-1)+\mu}{b}}dX \\
&= \frac{b}{2b(\iota b\omega+1)}\, e^{\frac{X(\iota b\omega+1)-\mu}{b}}\Big|_{-\infty}^{\mu} + \frac{b}{2b(\iota b\omega-1)}\, e^{\frac{X(\iota b\omega-1)+\mu}{b}}\Big|_{\mu}^{+\infty} \\
&= \frac{1}{2(\iota b\omega+1)}e^{\iota\mu\omega} - \frac{1}{2(\iota b\omega-1)}e^{\iota\mu\omega} \\
&= e^{\iota\mu\omega}\left[\frac{1}{2(\iota b\omega+1)} - \frac{1}{2(\iota b\omega-1)}\right] \\
&= \frac{e^{\iota\mu\omega}}{1+b^2\omega^2}
\end{aligned}$$

Finally, we have successfully derived the appropriate characteristic function.

1.4 Foundational conditions of Lévy processes

As was mentioned before, Lévy's proposition was a theoretical framework which naturally included some conditions and properties that must hold for the models that are a part of said framework.

This segment is dedicated to these fundamental properties. For a process to be thus labeled as a Lévy process the following are required:

- Independence of increments,
- Stationarity of increments,
- Stochastic continuity.

We continue by further discussing these properties one by one in greater detail.

1.4.1 Independence of increments

Definition 1.4.1 (Independence of increments)**.** A process $(X_t)_{t\in T}$, where $T \subset \mathbb{R}$, has independent increments if and only if:

$$\forall n \in T \quad \text{and} \quad \forall t_0, t_1, t_2, \ldots, t_{n-1}, t_n \in T$$

while:

$$t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n$$

it holds that the random variables representing the increments:

$$(X_{t_1} - X_{t_0}), (X_{t_2} - X_{t_1}), \ldots, (X_{t_n} - X_{t_{n-1}})$$

are independent of each other.

1.4.2 Stationarity of increments

Definition 1.4.2 (Stationarity of increments)**.** Any process $(X_t)_{t\geq 0}$ has stationary increments if $0 \leq t_i \leq t_j, \forall t_i, t_j \in T$, it holds that:

$$X_{t_j - t_i} \overset{\mathcal{D}}{=} X_{t_j} - X_{t_i}$$

1.4.3 Stochastic continuity

Definition 1.4.3 (Stochastic continuity)**.** Any process $(X_t)_{t\geq 0}$ is stochastically continuous if $\forall t_i, t_j \in T$ it holds that:

$$\mathbb{P}\left(\lim_{t_i \to t_j} \left|X_{t_j} - X_{t_i}\right| = 0\right) = 1$$

with $\forall i \neq j$.

1.5 Major parts of Lévy Processes' framework

In this section, we delve deeper on some accompanying, yet defining properties and theorems regarding Lévy processes. The aim is to properly showcase the various connections between the upcoming items and how they all band together to form a robust theoretical framework for studying Lévy processes successfully.

1.5.1 Infinitely Divisible Distributions

De Finetti was one of the first scientists to study infinite divisibility laws (Regazzini, 2008), though the naming scheme was coined by Bawly and Khintchine (Bawly, 1936; Khintchine, 1937).

Definition 1.5.1 (Infinitely Divisible Distributions)**.** The distribution that is admitted by a random variable X is said to be infinitely divisible if for all non-negative integers, $n \in \mathbb{N}$ it holds that:

$$X \overset{\mathcal{D}}{=} X_1^{(1/n)} + \cdots + X_n^{(1/n)} \tag{1.3}$$

given an *i.i.d.* sequence $(X_n^{(1/n)})_{0\leq n\in\mathbb{N}}$. The characteristic function of a random variable can also be used equivalently:

$$\phi_X(\omega) = \left(\phi_{X^{(1/n)}}(\omega)\right)^n \tag{1.4}$$

Below we present some examples where some fundamental distributions are shown whether they are or not infinitely divisible.

Example 1.5.1 (Normal distribution)**.** *Let us assume that the characteristic function of the Normal distribution is:*

$$\phi_X(\omega) = e^{\iota\mu\omega - \frac{\sigma^2\omega^2}{2}} \tag{1.5}$$

Additionally, we know that the following always holds:

$$\begin{aligned}
\phi_X(\omega) &= e^{\iota\mu\omega - \frac{\sigma^2\omega^2}{2}} \\
&= \left(e^{\iota\mu\omega - \frac{\sigma^2\omega^2}{2}}\right)^{\frac{n}{n}} \\
&= \left(e^{\frac{1}{n}\left(\iota\mu\omega - \frac{\sigma^2\omega^2}{2}\right)}\right)^{n} \\
&= \left(e^{\iota\frac{\mu}{n}\omega - \frac{\sigma^2\omega^2}{2n}}\right)^{n} \\
&= \left(e^{\iota\frac{\mu}{n}\omega - \frac{\frac{\sigma^2}{n}^2}{\omega}^2}\right)^{n} \\
&= \left(\phi_{X/n}(\omega)\right)^{n}
\end{aligned} \tag{1.6}$$

Thus if we sample

$$X_i \sim Normal\left(\frac{\mu}{n}, \frac{\sigma^2}{n}\right) \tag{1.7}$$

then $X = \sum_{i=1}^{n} X_i \sim Normal(\mu, \sigma^2)$.

Example 1.5.2 (Gamma distribution)**.** *Similarly, let us assume that the characteristic function of a R.V.* $X \sim Gamma(k, \theta)$ *is the following:*

$$\phi_X(\omega) = (1 - \theta\iota\omega)^{-k} \tag{1.8}$$

Then the following always holds:

$$\begin{aligned}
\phi_X(\omega) &= (1 - \theta\iota\omega)^{-k} \\
&= \left((1 - \theta\iota\omega)^{-k}\right)^{\frac{n}{n}} \\
&= \left((1 - \theta\iota\omega)^{-\frac{k}{n}}\right)^{n} \\
&= \left(\phi_{\frac{X}{n}}(\omega)\right)^{n}
\end{aligned} \tag{1.9}$$

Thus, if we sample $X_i \sim Gamma(k/n, \theta)$*, then* $X = \sum_{i=1}^{n} X_i \sim Gamma(k, \theta)$.

Example 1.5.3 (χ^2_k distribution)**.** *Let us assume that the characteristic function of* χ^2_k *distribution with* k *degrees of freedom is the following:*

$$\phi_X(\omega) = (1 - 2\iota\omega)^{-\frac{k}{2}} \tag{1.10}$$

Then, similarly, the following holds:

$$\begin{aligned}\phi_X(\omega) &= (1-2\iota\omega)^{-\frac{k}{2}} \\ &= \left((1-2\iota\omega)^{-\frac{k}{2}}\right)^{\frac{n}{n}} \\ &= \left((1-2\iota\omega)^{-\frac{k}{2n}}\right)^{n} \\ &= \left(\phi_{\frac{X}{n}}(\omega)\right)^{n}\end{aligned} \tag{1.11}$$

Finally, if we sample $X_i \sim \chi^2_{k/n}$*, then* $X = \sum_{i=1}^{n} X_i \sim \chi^2_k$.

Definition 1.5.2 (Stable Distribution)**.** A stable distribution is any non degenerate distribution for which it holds:

$$\gamma X + \delta \overset{\mathcal{D}}{=} \alpha X_1 + \beta X_2$$

where $\alpha, \beta, \gamma > 0$ and X_1, X_2 are two random realizations of the distribution through the random variable X.

Distributions in the Gaussian, Cauchy and δ-distribution family are infinitely divisible on $\mathbb{R}^d$. Similarly, distributions in the Poisson, geometric, negative binomial, exponential and Γ-distribution family are infinitely divisible on $\mathbb{R}$. One-sided strictly stable distributions with index $1/2$ on $\mathbb{R}$ are also infinitely divisible (Ken-Iti, 1999).

There are examples of distributions that are not infinitely divisible like the Uniform or the Binomial. Distributions with bounded support are not infinitely divisible, with the exception of δ-distribution (Ken-Iti, 1999).

Kyprianou (2014), correctly states that while there exist numerous accepted proofs for many probability laws being infinitely divisible (and hence their corresponding Lévy processes existing), for most of them a high degree of mathematical maturity is required to successfully follow along.

A comprehensive but non exhaustive list of proven to be infinitely divisible distributions can be found in Kyprianou's text (Kyprianou, 2014):

- Generalized inverse Gaussian (Good, 1953; Jorgensen, 1981)
- Truncated Stable (Tweedie, 1984; Hougaard, 1986; Koponen, 1995; Boyarchenko and Levendorskii, 2002; Carr et al., 2003)
- Generalized Hyperbolic (Halgreen, 1979; Bingham and Kiesel, 2004; Eberlein, 2001; Barndorff-Nielsen and Shephard, 2001)
- Meixner (Schoutens and Teugels, 1998)
- Pareto (Steutel, 1970; Thorin, 1977a)
- F-distributions (Ismail and Kelker, 1979)
- Gumbel (Johnson et al., 1994; Steutel, 1973)
- Weibull (Johnson et al., 1994; Steutel, 1970)

- Lognormal (Thorin, 1977b)
- Student t-distribution (Grosswald, 1976; Ismail, 1977)
- Lamperti Stable (Caballero et al., 2008)
- β-class (Kuznetsov, 2010)

Assuming a Lévy process L_t, we are interested in having $L_{t_1}, L_{t_2} \dots L_{t_n}$ generated by an infinitely divisible distribution. In fact, we wish for this to always be the case.

We can consider a Lévy process $L = (L_t)_{0 \leq t \leq T}$ then for any $n \in \mathbb{N}$ and $0 < t \leq T$ it holds that:

$$L_t = L_{\frac{t}{n}} + \left(L_{\frac{2t}{n}} - L_{\frac{t}{n}}\right) + \cdots + \left(L_t - L_{\frac{(n-1)t}{n}}\right) \tag{1.12}$$

The properties of stationary and independent increments lead to:

$$\left(L_{\frac{tk}{n}} - L_{\frac{t(k-1)}{n}}\right)$$

$\forall k \geq 1$ being an i.i.d. sequence of random variables. We can make use of the characteristic exponent ψ and define:

$$e^{-\psi_t(\omega)} = \mathbb{E}\left[e^{\iota\omega L_t}\right]$$

$\forall \omega \in \mathbb{R}, t \geq 0$. Now we can use 1.12 on the above equation twice $\forall m \neq n$, where $m, n \in \mathbb{Z}^+$ and thus we end up with:

$$e^{-m\psi_1(\omega)} = e^{\psi_m(\omega)} = e^{-n\psi_{\frac{m}{n}}(\omega)}$$

We have to consider two cases for t:

- $t \in \mathbb{Q} \longrightarrow e^{-\psi_t(\omega)} = e^{-t\psi_1(\omega)}$.
- $t \in \mathbb{I} \longrightarrow e^{-\psi_t(\omega)} = e^{-t\psi_1(\omega)}$. Due to right continuity of the random variable X that is guaranteed by dominated convergence.

In the end, all Lévy processes satisfy the property:

$$\mathbb{E}\left[e^{\iota\omega L_t}\right] = e^{-t\psi(\omega)} \tag{1.13}$$

and it holds $\forall t \geq 0$. Thus, all random variables $X_1, X_2, \dots, X_n$ admitted by $L_{t_1}, L_{t_2}, \dots, L_{t_n}$ are also infinitely divisible.

Finally, we have shown that any Lévy process can be associated with an infinitely divisible law and in extension, it was further shown how that claim can be proved. On the next segment, we explore the opposite case; assuming an infinitely divisible probability law (or convolution of infinitely divisible probability laws), whether it is possible for a Lévy process to be constructed.

1.5.2 Lévy-Khintchine Formula

The following theorem provides another way to link infinitely divisible distributions with Lévy processes via a certain triplet and goes a step further. More specifically, it is shown that a Lévy process, via its infinitely divisible distribution, can be further characterized and decomposed via the aforementioned triplet.

Theorem 1.5.1 (Lévy-Khintchine formula)**.** *A distribution $\mathcal{D}$ admitted via a $\mathbb{R}$-valued random variable X will be infinitely divisible with characteristic exponent ψ:*

$$\int_{\mathbb{R}} e^{\iota\omega X}\mathcal{D}dX = e^{-\psi(\omega)}$$

if and only if there exists a triplet (α, σ, Π) such that:

$$\mathbb{E}\left[e^{\iota\omega X}\right] = \exp\left[\iota\alpha\omega + \frac{1}{2}\sigma^2\omega^2 + \int_{\mathbb{R}}\left(1 - e^{\iota\omega X} + \iota\omega X\mathbb{1}_{(|X|<1)}\right)\Pi(dX)\right] \tag{1.14}$$

where $\alpha \in \mathbb{R}, \sigma \geq 0$ and a measure Π with $\Pi(\{0\}) = 0$, such that:

$$\int_{\mathbb{R}}\left(1 \wedge X^2\right)\Pi(dX) < \infty$$

always holds for $\forall\omega \in \mathbb{R}$.

Theorem 1.5.2 (Lévy-Khintchine formula for Lévy processes)**.** *Assuming a triplet (α, σ, Π) with $\alpha, \sigma \in \mathbb{R}$ and $\Pi(\{0\}) = 0$ such that:*

$$\int_{\mathbb{R}}\left(1 \wedge X^2\right)\Pi(dX) < \infty$$

then a characteristic function can be defined for all $\omega \in \mathbb{R}$:

$$\mathbb{E}\left[e^{\iota\omega X}\right] = \exp\left[\iota\alpha\omega + \frac{1}{2}\sigma^2\omega^2 + \int_{\mathbb{R}}\left(1 - e^{\iota\omega X} + \iota\omega X\mathbb{1}_{(|X|<1)}\right)\Pi(dX)\right] \tag{1.15}$$

Then given that unique triplet (α, σ, Π) a Lévy process with that characteristic function can be defined and represented in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

The triplet (α, σ, Π) described above is named Lévy or characteristic triplet and the exponent term is similarly named Lévy or characteristic exponent. On the same vein, the elements of the triplet are equally important. α is the drift term, σ is the diffusion coefficient and Π is the Lévy measure.

Additionally, it is of significance to note two things:

- Each triplet generated is unique to the corresponding distribution that is associated with the Lévy process.
- The Lévy-Khintchine Formula, as shown, enables us in deriving a unique triplet from the associated infinitely divisible distribution while the opposite also

holds.

The Lévy-Khintchine representation plays a major role in the overall theory of Lévy processes. Originally, Kolmogorov developed a formula to characterize infinitely divisible distributions, though with severe limitations (Kolmogoroff, 1931). Lévy with his work managed to generalize Kolmogorov's formula so that it can be applied to all infinitely divisible distributions (Levy, 1939).

Equations 1.14 and 1.15 as shown above are the combined contributions of Khintchine (Khinchin, 1948) and Itô (Itô, 1944; Itô, 1946) with Lévy's generalized formula as the base. They were accepted and incorporated by Lévy as evidently by their inclusion in his book (Lévy, 1948).

1.5.3 Lévy-Itô Decomposition

So far, we have discussed about various aspects of Lévy processes but one item that has been left out till now is the behavior of the paths of these processes. In this section we explore this topic alongside Lévy-Itô Decomposition.

Before rigorously defining Lévy-Itô Decomposition, let us spend a little time to showcase what it stands for from an intuitive point of view. It is relatively straightforward to express equation 1.14 in the following manner:

$$\phi_X(\omega) = \exp\left[\psi^{(1)}(\omega) + \psi^{(2)}(\omega) + \psi^{(3)}(\omega)\right] \tag{1.16}$$

The closed expressions for $\psi^{(1)}, \psi^{(2)}, \psi^{(3)}$ are the following:

$$\psi^{(1)}(\omega) = \iota\alpha\omega + \frac{1}{2}\sigma^2\omega^2 \tag{1.17}$$

$$\psi^{(2)}(\omega) = \Pi\left(\mathbb{R} \setminus (-1,1)\right) \int_{|X|\geq 1} \left(1 - e^{\iota\omega X}\right) \frac{\Pi(dX)}{\Pi\left(\mathbb{R} \setminus (-1,1)\right)} \tag{1.18}$$

$$\psi^{(3)}(\omega) = \int_{0<|X|<1} \left(1 - e^{\iota\omega X} + \iota\omega X\right) \Pi(dX) \tag{1.19}$$

$\forall \omega, \alpha, \sigma \in \mathbb{R}$ and Π being the by-now-known measure that satisfies:

$$\int_{\mathbb{R}} (1 \wedge X^2)\Pi(dX) < \infty.$$

In the Lévy-Itô decomposition the stochastic exponents $\psi^{(1)}(\omega)$, $\psi^{(2)}(\omega)$, $\psi^{(3)}(\omega)$ are of utmost importance to us. These exponents represent three distinct Lévy processes that are able to characterize any Lévy process via an appropriate manipulation of the Lévy-Khintchine formula as shown above.

Naturally, the next step is to define these processes. We can start with $\psi^{(1)}(\omega)$ which is the stochastic exponent for a linear Brownian motion. An example process

of this form could be $X^{(1)} = \{X_t^{(1)} : t \geq 0\}$ where

$$X_t^{(1)} = \sigma B_t - \alpha t, t \geq 0. \tag{1.20}$$

We can proceed with the next stochastic exponent $\psi^{(2)}(\omega)$ and the process it admits. In this case, the process that corresponds to the exponent is a compensated Poisson process. An example of said process can be $X^{(2)} = \{X_t^{(2)} : t \geq 0\}$ where

$$X_t^{(2)} = \sum_{i=1}^{N_t} \xi_i, t \geq 0. \tag{1.21}$$

In the above equation, $\{N_t : t \geq 0\}$ is a Poisson process with its rate parameter taking the value $\Pi(\mathbb{R} \setminus (-1, 1))$ and thus, $\{\xi_i : i \geq 1\}$ are identically distributed under distribution $\Pi(dX)/\Pi(\mathbb{R} \setminus (-1, 1))$ that is concentrated on $\{X : |X| \geq 1\}$.

Lastly, all that is left is to define the stochastic process for the third and last stochastic exponent $\psi^{(3)}(\omega)$. In this case however, defining the associated process is quite a bit trickier than the previous items, which were relatively straightforward. For $\psi^{(3)}(\omega)$, one way to begin is to show that closed form of the exponent can be re-written in the following manner:

$$\begin{aligned} \psi^{(3)}(\omega) &= \int_{0<|X|<1} \left(1 - e^{\iota\omega X} + \iota\omega X\right) \Pi(dX) \\ &= \sum_{n\geq 0} \left[\lambda_n \int_{2^{-(n+1)} \leq |X| < 2^{-n}} \left(1 - e^{\iota\omega X}\right) F_n(dX) \right. \\ &\qquad \left. + \iota\omega\lambda_n \left(\int_{2^{-(n+1)} \leq |X| < 2^{-n}} X F_n(dX)\right)\right] \end{aligned} \tag{1.22}$$

where:

$$\begin{aligned} \lambda_n &= \Pi\left(\left\{X : 2^{-(n+1)} \leq |X| < 2^{-n}\right\}\right) \\ F_n(dX) &= \frac{\Pi(dX)}{\lambda_n} \end{aligned}$$

Lastly, let us note that the $\psi^{(3)}(\omega)$ is restricted on the set:

$$\left\{X : 2^{-(n+1)} \leq |X| < 2^{-n}\right\}.$$

It is evident from the above that the $\psi^{(3)}(\omega)$ corresponds to a superposition of a countable number of independent compound Poisson processes. The drift parameter for these processes may be of different value. Furthermore, additional linear drift may be present as well.

Finally, we have defined the three different stochastic processes that play their role each in order to formulate any other Lévy process. The first two, as shown, are quite straightforward, the last one is not whatsoever. Thus, to delve deeper into

the Lévy-Itô decomposition, we need to study this superposition of independent compound Poisson processes in greater detail. To do so, some prerequisites need to be covered. Namely:

- Poisson random measures
- Square-integrable martingales

The reason we must turn to these two classes is because of the discontinuous nature of the paths that are derived via the expression of the last stochastic exponent. Modelling the jump behavior of a Lévy process in terms of number of jumps and frequency is not easy. However the use of those prerequisites makes it achievable as well as providing the practitioner or researcher with the added benefits that come alongside the Lévy-Itô decomposition.

Poisson Random Measures

Let us start by going over the definition for this mathematical object.

Definition 1.5.3 (Poisson random measure)**.** Let $(S, \mathcal{S}, \eta)$ be any arbitrary σ-finite measure space and $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Then, let $N : \omega \times \mathcal{S} \to \{0, 1, 2, \dots\} \cup \{\infty\}$ so that the items in $\{N(\cdot, A) : A \in \mathcal{S}\}$ are random random variables that can be defined on the the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then, N is a Poisson random measure on $(S, \mathcal{S}, \eta)$ if the following conditions are satisfied:

- $\forall A \in \mathcal{S}, N(\cdot, A)$ is a Poisson random variable characterized with intensity $0 \leq \eta(A) \leq \infty$.
- $\forall A_1, A_2, \dots, A_n \in \mathcal{S}$ that are mutually disjoint, the corresponding Poisson random variables $N(\cdot, A_1), N(\cdot, A_2), \dots, N(\cdot, A_n)$ are independent.
- $\forall \omega \in \Omega, N(\omega)$ is a measure on $(\Omega, \mathcal{S})$ $\mathbb{P}$-almost surely.

Example 1.5.4 (Compound Poisson Process - Kyprianou, 2014)**.** *Let us now showcase the above result, via utilizing a simple Compound Poisson process. Considering a simple compound Poisson process with some drift $X = \{X_t : t \geq 0\}$, specifically:*

$$X_t = \delta t + \sum_{i=1}^{N_t} \xi_i$$

with $t \geq 0, \delta \in \mathbb{R}$ and $\{\xi_i : i \geq 1\}$ are a family of independent random variables that are distributed under some distribution function F. Lastly, $N = \{N_t : t \geq 0\}$ is a Poisson process with intensity $\lambda = 0$ and the family $\{T_i : i \geq 1\}$ are the process' arrival times.

Given a set $A \in \mathcal{B}[0, \infty) \times \mathcal{B}(\mathbb{R} \setminus \{0\})$, we are interested in measuring the number

of jumps and denote them in the form of a measure $N(A)$ that has the form:

$$\begin{aligned} N(A) &= \#\left\{i \geq 1 : (T_i, \xi_i) \in A\right\} \\ &= \sum_{i=1}^{\infty} \mathbb{1}_{((T_i,\xi_i)\in A)} \end{aligned} \tag{1.23}$$

At this point one of the major insights is reached, that given any arbitrary, finite interval of time $N(A) < \infty$ $\mathbb{P}$-almost surely, since the number of jumps the process exhibits is also finite. This applies, assuming $t \geq 0$ and $A \subset \mathcal{B}[0,\infty) \times \mathcal{B}(\mathbb{R} \setminus \{0\})$.

Moving forwards, we can attempt to prove the rest of the conditions that need to be satisfied for the random Poisson measure within the bounds of a compound Poisson process. From the conditions of the theorem, we focus at the first two.

If a sequence of mutually disjointed sets $A_1, A_2, \ldots, A_j \in \mathcal{B}[0,\infty) \times \mathcal{B}(\mathbb{R} \setminus \{0\})$ exists, then $N(A_1), N(A_2), \ldots N(A_k)$ are mutually independent and Poisson distributed with corresponding rate parameters:

$$\lambda_i = \lambda \int_{A_i} dt \times F(dX)$$

given that $k \geq 1$.

From the Poisson process theory, we know that the family $\{T_1, T_2, \ldots, T_n\}$ are distributed conditionally based on the event $\{N_t = n\}]$, if $t > 0$. This distribution is the same as that of an ordered and independent sample from the Uniform distribution on the interval $[0,t]$. Additionally, we know that the family $\{\xi_i : i = 1, 2, \ldots, n\}$ are also independent and identically distributed admitted by a common distribution F.

Thus, the pairs $\{(T_i, \xi_i) : i = 1, 2, \ldots, n\}$ are identically distributed under the common probabilistic law:

$$t^{-1} ds \times F(dX)$$

with support on the interval $[0,t] \times (\mathbb{R} \setminus \{0\})$, time-ordered. Given any $A \in \mathcal{B}[0,t) \times \mathcal{B}(\mathbb{R} \setminus \{0\})$ then $N(A)$ conditioned on the event $\{N_t = n\}$ will be a binomial random variable whose probability of success is the following:

$$\mathbb{P}\left(N(A) | N_t = n\right) = \int_A t^{-1} ds \times F(dX)$$

Naturally, it is possible to adjust the proof so far so that it holds for a sequence of mutually disjointed sets $(N(A_1), N(A_2), \ldots, N(A_k)) \in \mathcal{B}[0,t) \times \mathcal{B}(\mathbb{R} \setminus \{0\})$. To do so, it is necessary to define:

$$A_0 = \left\{[0,t] \times \mathbb{R} \setminus \left\{A_1 \cup A_2 \cup \cdots \cup A_k\right\}\right\},$$

$$\begin{aligned}\lambda_0 &= \int_{A_0} \lambda ds \times F(dX) \\ &= \lambda t - \lambda_1 - \lambda_2 - \cdots - \lambda_k,\end{aligned}$$

$$n_0 = n - \sum_{i=1}^{k} n_i,$$

$$\sum_{i=1}^{k} n_i \leq n.$$

Given the above, it is possible to successfully evaluate the distribution of the sets $(N(A_1), N(A_2), \ldots, N(A_k))$

$$\begin{aligned}&\mathbb{P}\left(N(A_1) = n_1, N(A_2) = n_2, \ldots, N(A_k) = n_k | N_t = n\right) \\ &\qquad = \frac{n!}{n_0! n_1! \ldots n_k!} \prod_{i=0}^{k} \left(\frac{\lambda_i}{\lambda_t}\right)^{n_i}\end{aligned}$$

Attempting to remove the conditioning by summing it out yields the following:

$$\begin{aligned}\mathbb{P}\left(N(A_1) = n_1, N(A_2) &= n_2, \ldots, N(A_k) = n_k\right) \\ &= \sum_{n \geq \sum_{i=1}^{k} n_i} e^{-\lambda t} \frac{(\lambda t)^n}{n!} \frac{n!}{n_0! n_1! \ldots n_k!} \prod_{i=0}^{k} \left(\frac{\lambda_i}{\lambda_t}\right)^{n_i} \\ &= \sum_{n \geq \sum_{i=1}^{k} n_i} e^{\lambda_0} \frac{\lambda_0^{\left(n - \sum_{i=1}^{k} n_i\right)}}{\left(n - \sum_{i=1}^{k} n_i\right)!} \left(\prod_{i=1}^{k} e^{-\lambda_i} \frac{\lambda_i^{n_i}}{n_i!}\right) \\ &= \prod_{i=1}^{k} e^{-\lambda_i} \frac{\lambda_i^{n_i}}{n_i!}\end{aligned} \tag{1.24}$$

Finally, it is shown that the sequence $N(A_1), N(A_2), \ldots, N(A_k)$ *are independent and Poisson distributed.*

The next step is to prove the condition of the theorem whereby the sequence of $A_1, A_2, \ldots, A_k$ *are instead arbitrarily disjointed. Achieving this is a bit more straightforward than the previous case for mutually disjointed sets. It was shown above that summing a sequence of independent Poisson random variables is, again, Poisson distributed according to the sum of their specific rates. Additionally, it is further straightforward to show that:*

$$\mathbb{P}(X = \infty) = 1$$

given that the random variable $X \sim Poisson(\lambda = \infty)$.

Lastly, by writing the initial sequence $N(A_1), N(A_2), \ldots, N(A_k)$ *as a countable union of disjoint sets, within the bounds of* $\mathcal{B}[0, t') \times \mathcal{B}(\mathbb{R} \setminus \{0\})$ *and taking into*

account the two points of the previous paragraph, we can once again conclude that $N(\cdot, A_i)$ are Poisson distributed for the case of arbitrarily disjoint sets as well.

The only thing left, is to discuss about the measure N. Proving that it is indeed a measure is trivial, as this can be derived directly by referencing the measure's definition.

So far, a mathematical object capable of modelling the jump behavior of the Lévy processes was shown, including a relevant example. Logically, the next step would be to discuss whether or not such an object always exists or not. Let us assume that one such measure always exists. The proof (Kyprianou, 2014) is as follows.

Let us consider an appropriate S, so that $0 < \eta(S) < \infty$. Then a standard infinite product space $(\Omega, \mathcal{F}, \mathbb{P})$, which serves as the probability space on which the following elements are defined:

- A random variable $\mathcal{N}$
- A sequence of random variables $\{u_1, u_2, \ldots, u_n\}$

We wish that $\mathcal{N}$ is admitted by a Poisson distribution with rate $\eta(S)$. Following that, we also desire that the sequence u_i is distributed with distribution $\eta(dX)/\eta(S)$ with its support being S. Now we can define:

$$N(A) = \sum_{i=1}^{\mathcal{N}} \mathbb{1}_{(u_i \in A)}$$

$\forall A \in \mathcal{S}$, so that $\mathcal{N} = N(S)$. All of this is similar to what was shown in equation 1.23. Following the same vein, it is evident that the sequence $N(A)$ is $\mathcal{F}$-measurable since the family of random variables $\mathbb{1}_{(u_i \in A)}$ and $\mathcal{N}$ are themselves $\mathcal{F}$-measurable.

Then we must deal with the case of a sequence $A_1, A_2, \ldots, A_k$, arriving to the same equation with the same steps to evaluate it, that is equation 1.24. Then, identically once again, the measure $N(\cdot)$ fulfills all the conditions as well as the necessary requirements for it to be defined as a measure.

Following the above, in order to prove that a Poisson random measure always exists, we need to take into consideration the fact that the measure space $(S, \mathcal{S}, \eta)$ is σ-finite. The property σ-finite reveals that a countable disjoint and exhaustive sequence of sets $B_1, B_2, \ldots, B_n \in S$ so that $0 < \eta(B_i) < \infty$ with $i \geq 1$. With the above in place, we can now define a sequence of measures $\{\eta_1(\cdot), \eta_2(\cdot), \ldots, \eta_n(\cdot)\}$, where $\eta_i(\cdot) = \eta(\cdot \cap B_i)$.

Essentially the proof is split in two parts. In the first part, it is shown that $\forall i \geq 1$ there exists a probability space $(\Omega_i, \mathcal{F}_i, \mathbb{P}_i)$ on which a Poisson random measure $N_i(\cdot)$ can be defined in $(B_i, \mathcal{S} \cap B_i, \eta_i)$, where $\mathcal{S} \cap B_i = \{A \cap B_i : A \in \mathcal{S}\}$.

Alternatively, expressed in equation style mode, the concept is as follows, if:

$$N(\cdot) = \sum_{i \geq 1} N_i \left(\cdot \cap B_i\right)$$

is a Poisson random measure on S with rate η defined on the infinite product space:

$$(\Omega, \mathcal{F}, \mathbb{P}) = \prod_{i \geq 1} (\Omega_i, \mathcal{F}_i, \mathbb{P}_i) .$$

If we make use of Fubini's theorem, then we are lead to the following equation for our sequence of disjoint sets $A_1, A_2, \ldots, A_j$:

$$\begin{aligned} N\left(\bigcup_{j \geq 1} A_j\right) &= \sum_{i \geq 1} N_i \left(\bigcup_{j \geq 1} A_j \cap B_i\right) \\ &= \sum_{i \geq 1} \sum_{j \geq 1} N\left(A_j \cap B_i\right) \\ &= \sum_{j \geq 1} \sum_{i \geq 1} N\left(A_j \cap B_i\right) \\ &= \sum_{j \geq 1} N\left(A_j\right) . \end{aligned}$$

Thus, $\forall i \geq 1$ it holds that:

$$N_i\left(A \cap B_i\right) \sim Poisson\left(\lambda = \eta_i(A)\right) .$$

For the case of disjoint sets $A_1, A_2, \ldots, A_k \in \mathcal{S}$, the sequence of random variables $N\left(A_1\right), N\left(A_2\right), \ldots N\left(A_k\right)$ are independent (and Poisson distributed), since:

$$\left\{N_i\left(A_j \cap B_i\right) : i = 1, 2, \ldots, n \wedge j = 1, 2, \ldots, k\right\}$$

is an independent sequence of variables as well.

Additionally, from the process of proving the existence of a Poisson random measure, a couple extra bits of information can be deduced. Below they are briefly presented as corollaries.

Corollary 1.5.1. *If $N(\cdot)$ is a Poisson random measure, defined on the measure space $(S, \mathcal{S}, \eta)$, then $\forall A \in \mathcal{S}, N(\cdot \cap A)$ is a Poisson random measure on the space $(S \cap A, \mathcal{S} \cap A, \eta(\cdot \cap A))$. Moreover, considering an additional set $B \in \mathcal{S}$, so that:*

- $A, B \in \mathcal{S}$
- $A \cap B = \emptyset$

the random variables $N(\cdot \cap A)$ and $N(\cdot \cap B)$ are independent.

Corollary 1.5.2. *If $N(\cdot)$ is a Poisson random measure defined on the measure space $(S, \mathcal{S}, \eta)$, then we can automatically assume that the support of the measure $N(\cdot)$ is $\mathbb{P}$-almost surely countable. If η is a finite measure, then $(S, \mathcal{S}, \eta)$ is $\mathbb{P}$-almost surely finite.*

At this point we have discussed Poisson random measures in terms of definition, as well as providing a few key insights in terms of how and why they are important

to Lévy processes. Before ending this brief section, we present a couple properties and identities for Poisson random measures that will be utilized down the road.

Theorem 1.5.3. *If N is a Poisson random measure on a measure space $(S, \mathcal{S}, \eta)$ then immediately we can define*

$$X = \int_S f(x) N(dX) \tag{1.25}$$

where X is a well defined $\mathbb{R}$-valued random variable, with a measurable function $f : S \to \mathbb{R}$. X is absolutely convergent almost surely if and only if:

$$\int_S \big(1 \wedge |f(X)|\big)\eta(dX) < \infty$$

From this theorem, it is possible to derive a few more things, we express them in the form of corollaries.

Corollary 1.5.3. *Supposing that what is expressed in theorem 1.5.3 holds, then:*

$$\mathbb{E}\left[e^{\iota\omega X}\right] = exp\left[-\int_S \left(1 - e^{\iota\omega f(X)}\right)\eta(dX)\right] \tag{1.26}$$

$\forall\omega \in \mathbb{R}$.

Corollary 1.5.4. *Once again, assuming that everything in theorem 1.5.3 holds, then we can define $\mathbb{E}[X]$ and $\mathbb{E}[X^2]$ in the following manner:*

$$\mathbb{E}[X] = \int_S f(X)\eta(dX) \tag{1.27}$$

when:

$$\int_S |f(X)|\,\eta(dX) < \infty.$$

For $\mathbb{E}[X^2]$:

$$\mathbb{E}[X^2] = \int_S f(X)^2\eta(dX) + \left(\int_S f(X)\eta(dX)\right)^2 \tag{1.28}$$

when:

$$\int_S f(X)^2\eta(dX) < \infty$$

and

$$\int_S |f(X)|\,\eta(dX) < \infty.$$

A proof for the above could be the following. We can define some functions as follows:

$$f(x) = \sum_{i=1}^{n} f_i \mathbb{1}_{A_i}(X)$$

for which the following conditions hold:

- $A_1, A_2, \ldots, A_n \in \mathcal{S}$
- $\bigcap_{i=1}^{n} A_i = \emptyset$
- f_i is a constant function
- $\eta(A_1 \cup A_2 \cup \cdots \cup A_n) < \infty$

For all of the above to hold, we are lead to equation 1.25 and identically to that, we define:

$$X = \sum_{i=1}^{n} f_i N(A_i).$$

At this point we are able to safely deduce that $P(X < \infty) = 1$ almost surely. That is because each $N(A_i)$ is Poisson distributed with rate $\eta(A_i) < \infty$.

Let us agree that the moment generating function for the Poisson distribution with rate $\lambda > 0$ is the following:

$$MGF_{Poisson} = e^{-\lambda\left(1-e^{-\omega}\right)}$$

$\forall \omega \geq 0$. Then it yields:

$$\begin{aligned}
\mathbb{E}\left[e^{-\omega X}\right] &= \prod_{i=1}^{n} \mathbb{E}\left[e^{-\omega f_i N(A_i)}\right] \\
&= \prod_{i=1}^{n} exp\left[-\left(1-e^{-\omega f_i}\right)\eta(A_i)\right] \\
&= exp\left[-\sum_{i=1}^{n}\left(1-e^{-\omega f_i}\right)\eta(A_i)\right]
\end{aligned}$$

again, $\forall \omega \geq 0$. Additionally, we know that $1 - e^{-\omega f(X)} = 0$ on the interval that spans S, thus:

$$\mathbb{E}\left[e^{-\omega X}\right] = exp\left[-\int_S \left(1-e^{-\omega f(X)}\right)\eta(dX)\right].$$

The next step in the proof is re-express the last equation, under the banner of a positive measurable f only.

$$\begin{aligned}
\lim_{n\to\infty} \int_S f_n(X) N(dX) &= \int_S f(X) N(dX) \\
&= X
\end{aligned}$$

We can arrive at the above equation because N is a σ-finite measure almost surely, along with the fact that an increasing sequence of positive functions exists so that $\lim_{n\to\infty} f_n = f$.

If we apply bounded and monotone convergence in that order then we have:

$$\begin{aligned}
\mathbb{E}\left[e^{-\omega X}\right] &= \mathbb{E}\left[\exp\left[-\omega\int_S f(X)N(dX)\right]\right] \\
&= \lim_{n\to\infty}\mathbb{E}\left[\exp\left[-\omega\int_S f_n(X)N(dX)\right]\right] \\
&= \lim_{n\to\infty}\exp\left[-\int_S\left(1-e^{-\omega f_n(X)}\right)\eta(dX)\right] \\
&= \exp\left[-\int_S\left(1-e^{-\omega f(X)}\right)\eta(dX)\right]
\end{aligned}$$

$\forall\omega > 0$.

That last integral presents an interesting case, as we need to consider two cases. Specifically the integral can be finite or infinite $\forall\omega > 0$. If the integral is finite then $P(X = \infty) < 1$, if it's infinite then $P(X = \infty) = 1$. Pressing on, in the case that it is finite, the following inequality holds $\forall\omega > 0$:

$$\left(1-e^{-\omega f(X)}\right) \leq \left(1-e^{-f(X)}\right).$$

From the perspective of dominated convergence we have:

$$\lim_{\omega\to 0}\int_S\left(1-e^{-\omega f(X)}\right)\eta(dX) = 0$$

$\forall 0 < \omega < 1$. Applying again the dominated convergence yields that $P(X = \infty) = 0$.

So far we have shown that $X < \infty$ almost surely and that in turn happens if and only if:

$$\int_S (1 \vee f(X))\,\eta(dX) < \infty.$$

In order to complete the proof for theorem 1.5.3, we need to discuss what happens when the functions that we defined are not under the restriction that they are only positive measurable functions, but only measurable. At this point, it is possible to define f as $f = f^+ - f^-$, where $f^+ = f \vee 0$ and $f^- = (-f) \vee 0$ are both measurable as well. Similarly, we may define X as $X = X_+ - X_-$, where:

$$\begin{aligned}
X_+ &= \int_S f(X)N_+dX, \\
X_- &= \int_S f(X)N_-dX
\end{aligned}$$

and:

$$\begin{aligned}
N_+ &= N\left(\cdot\cap\{X\in S : f(X)\geq 0\}\right), \\
N_- &= N\left(\cdot\cap\{X\in S : f(X) < 0\}\right).
\end{aligned}$$

Furthermore, we do know a couple more thing about N_+ and N_-. Specifically, we

know that they are both Poisson random measures on the same space and with intensities:

- $\eta\left(\cdot \cap \{f \geq 0\}\right)$
- $\eta\left(\cdot \cap \{f < 0\}\right)$.

We also know that they are independent and from that the same can be derived for X_+ and X_-. Lastly, so far it is shown that X is absolutely convergent almost surely if and only if X_+ and X_- are convergent themselves.

In the case of dealing with positive-only f, then the X is absolute convergent almost surely if and only if:

$$\int_S \big(1 \wedge |f(X)|\,\big)\eta(dX)$$

and thus the theorem 1.5.3 is proven. To prove the corollaries following that theorem, we need to use some parts from the proof we just discussed. Namely the integral above and the fact that X_+ and X_- are independent. Then, $\forall \omega \in \mathbb{R}$ we have:

$$\begin{aligned}
\mathbb{E}\left[e^{\iota\omega X}\right] &= \mathbb{E}\left[e^{\iota\omega X_+}\right] + \mathbb{E}\left[e^{\iota\omega X_-}\right] \\
&= \exp\left[-\int_{\{f\geq 0\}} \left(1 - e^{\iota\omega f^+(X)}\right)\eta(dX)\right] \\
&\qquad \times \exp\left[-\int_{\{f<0\}} \left(1 - e^{\iota\omega f^-(X)}\right)\eta(dX)\right] \\
&= \exp\left[-\int_S \left(1 - e^{\iota\omega f(X)}\right)\eta(dX)\right]
\end{aligned}$$

and so the proof for corollary 1.5.3 is finished. The proof for the last corollary 1.5.4 is extremely similar to what was shown so far in the proof above. For that reason, we just briefly state the necessary steps only. Thus, the proof is this:

- define simple f,
- consider positive-only measurable f,
- proceed with the general measurable f.

We end the Poisson random measures segment here.

Square-integrable Martingales

Let us start this section by attempting to link Poisson random measures with their associated square integrable martingales. We reference the identities and properties from theorem 1.5.3 and corollaries 1.5.3 and 1.5.4. Then for a Poisson random measure $N(\cdot)$ on the space

$$\left([0,\infty) \times \mathbb{R}, \mathcal{B}[0,\infty) \times \mathcal{B}(\mathbb{R}), dt \times \Pi(dX)\right),$$

and Π is, again, a measure that is concentrated on $\mathbb{R} \setminus \{0\}$, the various integrals from those two corollaries and the theorem are transformed appropriately:

$$t \int_B \big(1 \wedge |X|\big)\Pi(dX),$$
$$t \int B \left(1 - e^{\iota\omega X}\right) \Pi(dX),$$
$$t \int_B |X|\Pi(dX),$$
$$t \int_B X^2 \Pi(dX)$$

where $B \in \mathcal{B}(\mathbb{R})$ and our previously defined $f(X) = X$. Furthermore, it is possible to define a compound Poisson process X_t assuming that a Poisson random measure $N(\cdot)$ is defined on the space as above, along with the measure Π with the same restrictions:

$$X_t = \int_{[0,t]} \int_B XN(dv \times dX)$$

given that $B \in \mathcal{B}(\mathbb{R})$ and $0 < \Pi(B) < \infty$, $\forall t \geq 0$. Then X_t, as said, will be a compound Poisson process with rate $\Pi(B)$ and the jumps are distributed under distribution

$$\Pi(B)^{-1}\Pi(dX)|_B.$$

Proving this is not too complicated, to achieve so we need to show that the process X_t satisfies the prerequisites, as defined earlier, so that it can be classified as a Lévy process. We can show that independence of increments holds since:

$$X_t - X_s = \int_{[s,t]} \int_B XN(dv \times dX)$$

are independent $\forall 0 \leq u \leq s < t < \infty$ and we can deduce this immediately, $N(\cdot)$ is independently distributed on disjoint intervals. We can also deduce that the increments are identically distributing, again, from the definition of the Poisson random measure. Specifically due to closed expression form of the intensity of the measure $dt \times \Pi(dX)$. Lastly, showing the stochastic continuity of the process is also quite straightforward, since all it requires is a closer inspection to the original restriction of $\Pi(B)$, specifically $0 < \Pi(B) < \infty$. This means that the process Xt is right-continuous with left limits, since X_t will always be finite in its summation almost surely.

If we consider the same Poisson random measure $N(\cdot)$, along with the same restrictions, but with the addition that B satisfies the condition

$$\int_B |X|\Pi(dX) < \infty$$

then:

- the resulting compound Poisson process is a $\mathbb{P}$-martingale with respect to the filtration $\mathbb{F}$ (which satisfies the usual conditions):

$$X_t = \int_{[0,t]} \int_B XN(ds \times dX) - t \int_B X\Pi(dX)$$

$\forall t \geq 0$,

- the above compound Poisson process will also be a square-integrable martingale if:

$$\int_B X^2 \Pi(dX) < \infty.$$

We can prove that the process is a $\mathbb{P}$-martingale by recalling the property of stationary and independent increments

$$\begin{aligned}
&\mathbb{E}\left[M_t - M_s | \mathcal{F}_s\right] \\
&\qquad = \mathbb{E}\left[M_{t-s}\right] \\
&\qquad = \mathbb{E}\left[\int_{[s,t]} \int_B XN(du \times dX)\right] - (t-s) \int_B X\Pi(dX) \\
&\qquad = 0
\end{aligned}$$

$\forall 0 \leq s \leq t < \infty$. We also know that:

$$\mathbb{E}\left[|M_t|\right] \leq \mathbb{E}\left[\int_{[0,t]} \int_B |X| N(ds \times dX) - t \int_B |X| \Pi(dX)\right] < \infty$$

from corollary 1.5.4.

The second part is quite similar in terms of proving it. In fact, once again due to corollary 1.5.4 it is possible to arrive to the following equation:

$$\mathbb{E}\left[\left(X_t + t \int_B X\Pi(dX)\right)^2\right] = t \int_B X^2 \Pi(dX) + t^2 \left(\int_B X\Pi(dX)\right)^2$$

considering the property $\mathbb{E}[X_t^2]$, which was shown earlier as well, we can finally conclude on the following equation:

$$\mathbb{E}\left[X_t^2\right] = t \int_B X^2 \Pi(dX) < \infty.$$

Up until this point we have successfully linked Poisson random measures with square-integrable martingales. Before proceeding with the important theorem for this segment, we need to discuss why we are interested in this item. Briefly and without technicalities, the reason lies in the fact that without square-integrable martingales the Lévy-Itô decomposition would not work as it would not be well-defined. Furthermore, based on the the corollary 1.5.4 and the finiteness of the two integrals we can further differentiate between Lévy processes. That is to say that it

is not mandatory that both of those integrals evaluate to a finite value. As we will see in a later segment, there exist special cases of Lévy processes for which those integrals are not finite.

From the Lévy triplet that we have defined previously as part of the Lévy-Khintchine formula, a lot can be derived. For example, this measure is able to disclose a lot about the moments of the process being characterized, and by extension, whether or not they are finite. Thus, it is possible to generalize the above and apply it for the p-th moment or p-th exponential moment. Assuming an identical process X_t as defined above, then X_t will have finite p-th moment or:

$$\mathbb{E}\left[|X_t|^p\right] < \infty$$

$\forall p \in \mathbb{R}^+ \cup \{0\}$, if and only if:

$$\int_B |X|^p \Pi(dX) < \infty$$

for $B \in \mathcal{B}(\mathbb{R})$. Identically, the process X_t will have a finite p-th exponential moment or:

$$\mathbb{E}\left[e^{pX_t}\right] < \infty$$

$\forall p \in \mathbb{R}$, if and only if:

$$\int_B e^{pX} \Pi(dX) < \infty.$$

A proof for all the above can be found in Sato (1999).

Talking about the finiteness or not of the integrals above cannot be complete without making a mention to the area that the aforementioned integrals are operate on. While $B \in \mathcal{B}(\mathbb{R})$, we saw how a compound Poisson process is a $\mathbb{P}$-martingale with the appropriate conditions applied. These conditions allow us to work with intervals of various forms such as:

$$B_\epsilon = (-1, \epsilon) \cup (\epsilon, 1)$$

$\forall \epsilon \in (0, 1)$. However, such a form is not absolute. What is essential whatsoever, is focusing on the behavior of the type of integrals of interest:

$$\int_{[0,t]} \int_B XN(dv \times dX)$$

and specifically the behavior of their limit when dealing with sets of the form as above, while $\epsilon \to 0$.

Summing up, the form of the interval and the area that it spans, namely B, that we integrate over largely determines whether:

$$\mathbb{E}\left[X\right] < \infty$$

and

$$\mathbb{E}\left[X^2\right] < \infty.$$

Special cases of Lévy processes arise when not both moments are finite. Next up, we write in the form of a theorem the main takeaway from this segment, which is why exactly we are interested in the associated square-integrable martingales of Poisson random measures, as well as why studying the limit of:

$$\int_{[0,t]} \int_{B_\epsilon} XN(ds \times dX) - t \int_{B_\epsilon} X\Pi(dX)$$

with $\epsilon \to 0$, while B has the form we discussed, is important.

Theorem 1.5.4. *Assuming that $N(\cdot)$ is a Poisson random measure on*

$$([0,\infty) \times \mathbb{R}, \mathcal{B}[0,\infty) \times \mathcal{B}(\mathbb{R}), dt \times \Pi(dX))$$

where Π is a measure that is concentrated on $R \setminus \{0\}$ and

$$\int_{(-1,1)} X^2 \Pi(dX) < \infty$$

then we can define the Martingale M_t^ϵ for each $\epsilon \in (0,1)$ as follows:

$$M_t^\epsilon = \int_{[0,t]} \int_{B_\epsilon} XN(ds \times dX) - t \int_{B_\epsilon} X\Pi(dX)$$

$\forall t \geq 0$. An appropriate Martingale $M = \{M_t : t \geq 0\}$ will always exist that satisfies the following properties:

- *M has stationary increments.*
- *M has independent increments.*
- *M is a càdlàg process almost surely.*
- *M is adapted to an appropriate filtration that satisfies the usual conditions.*
- *M exhibits a finite number of discontinuities on a finite interval $[0,T]$.*
- *A deterministic subsequence $\{\epsilon_n^T : n = 1, 2, \dots\}$ that converges to 0 will always exist so that:*

$$P\left(\lim_{n\to\infty} \sup_{0\leq s\leq T} \left(M_s^{\epsilon_n^T} - M_s\right)^2 = 0\right) = 1.$$

In order to prove the above theorem, we need to start by examining the list of prerequisites our square-integrable martingale must satisfy. Kyprianou (2014) in his text provides a comprehensive proof for all the conditions while treating each one separately. We do not include a proof for this theorem as, similarly to the Lévy-Khintchine formula (and its decomposition, it is quite lengthy and elaborate. Thus, escaping from the scope of this thesis.

Limiting ourselves to one comment about the above would be the following: should

one take a closer look at these conditions, it becomes clear quite quickly that a section of them are the conditions for a stochastic process to be labeled as a Lévy process. Thus, it is easy to conclude that we require the martingale M to be a Lévy process as well.

Additionally, we can notice the need for a finite number of jumps in a closed interval which, in turn, ends with the sequence of the martingales $\{M_t^\epsilon : t \leq T\}$ over the same closed interval exhibiting uniform convergence almost surely along the deterministic subsequence we defined previously.

Back to Lévy-Itô Decomposition

After going through the necessary prerequisites and building off of them in order to showcase their connection with the Lévy-Itô decomposition, it is finally time to conclude this segment. We will not include an explicit proof for the method as it is extremely lengthy, spanning multiple pages. Instead we reference Sato, which is one of the sources where a full and detailed account of the proof may be found (Ken-Iti, 1999). However, we shall re-express the Lévy-Itô decomposition in a more fitting form - as a theorem.

Theorem 1.5.5 (Lévy-Itô Decomposition)**.** *Assuming a Lévy triplet* (α, σ, Π)*, where* $\alpha \in \mathbb{R}$*,* $\sigma \in \mathbb{R}^+ \cup \{0\}$ *and* Π *is a measure that satisfies:*

- $\Pi(\{0\}) = 0$
- $\int_{\mathbb{R}} (1 \wedge |X|^2)\Pi(dX) < \infty$

then a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ *exists where* 3 *distinct Lévy processes exist, represented by the characteristic exponents 1.17, which we denote as* $X^{(1)}, X^{(2)}, X^{(3)}$*. These are a linear Brownian motion, a compensated Poisson process and a pure jump process taking the form of a square-integrable martingale with an almost surely countable number of jumps with intensity less than* 1 *on any finite time interval, respectively. Thus, a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *exists where a Lévy process* X *can be defined as:*

$$X = X^{(1)} + X^{(2)} + X^{(3)}.$$

The main takeaway from the Lévy-Itô decomposition is that it enables us to characterize and decompose any Lévy process into the sum of a linear Brownian motion, a compensated Poisson process and a pure jump process which when handled separately are easier to model than the non-decomposed process. Thus the method allows us express the Lévy process given an infinitely divisible probability law. Additionally, given a Lévy triplet directly we can use the Lévy-Itô decomposition to express the associated Lévy process.

One more important thing that should also be mentioned is that not every text uses three stochastic exponents. Some researchers and practitioners choose to decompose the Lévy-Khintchine formula differently. However, regardless of the exact exponents that are used, the fundamental aspect that should never be violated is to appropriately treat the continuous and discontinuous parts of the process. This is

the essence of the Lévy-Itô decomposition. In a future segment of this thesis, we will also present different established methodologies capable successfully decomposing a Lévy process as well.

We will end this segment by discussing the Lévy measure Π a bit. Hopefully, by now it is quite clear just how significant this measure is to the overall theoretical framework for Lévy processes. While it is undoubtedly not easy to handle, it plays a major role into shaping Lévy processes as we know them. For example let us consider the following cases, assuming a Lévy process X with Lévy triplet (α, σ, Π):

- If $\Pi(\mathbb{R}) < \infty$ then the process exhibits a finite number of jumps over any closed time interval $[t_1, t_2]$ almost surely, where $t_1 < t_2$.
- If $\Pi(\mathbb{R}) = \infty$ then the opposite holds. The process exhibits an infinite number of jumps over any closed time interval $[t_1, t_2]$ almost surely, where $t_1 < t_2$.
- If $\sigma = 0$ and:
$$\int_{|X|\leq 1} |X|\Pi(dX) < \infty$$
then the process has finite variation almost surely.
- If $\sigma \neq 0$ and:
$$\int_{|X|\leq 1} |X|\Pi(dX) = \infty$$
then the process has infinite variation almost surely.

A proof for the above can be found in Sato's book (Ken-Iti, 1999). Thus it is clear how important the measure Π is. Indeed, as we will see in another part of the thesis later, the Lévy measure Π and its various imposed restrictions may pave the way for classes of special Lévy processes.

1.6 Examples of Lévy processes

So far we have discussed the theoretical framework that encompasses the Lévy processes. We have dedicated great attention into conveying the fundamentals of the theory, along with the ways all of the mathematical items of interest band together to form this robust framework.

The aim of this segment is to go over some Lévy processes. It is necessary to mention beforehand that this is not an exhaustive list of Lévy processes and that the ongoing literature includes many more examples of processes which are not going to be covered in this segment. Still, it is mandatory to provide an account, even if limited, of the most representative Lévy processes.

1.6.1 Brownian Motion

Background of Brownian Motion

Brownian motion will be the process to discuss. It is quite a well-known Lévy process, perhaps the most well known. Its history is also a lengthy one, since it can be traced all the way back to 1827 and specifically to the experiment Robert Brown conducted. Brown, using a microscope, found out that small particles that were suspended in water exhibited some movement which seemed to be random. Brown in fact published his findings in his paper (Brown, 1828) in which he proposed that this was the result of some living organisms which he named as "molecules of the organic world".

Einstein has also dedicated his research time on this phenomenon in his paper (Einstein, 1905), Einstein provides a more compelling theory and cause for the phenomenon. He argues that the magnitude of the collision could be further used to provide insight into the size of the molecules pertaining in that collision. Later on, this refined theory of the phenomenon would be deemed extremely important, as it ended up being the cornerstone of Paul Langevin's equation. In his paper, Langevin proposed an equation that was capable of describing the motion of a particle in a fluid under the influence of random collisions (Lemons and Gythiel, 1908). This was a very significant breakthrough when it comes to Brownian motion and is in fact still in use today in many different fields as we will see later.

This observed random movement intrigued Jean Baptiste Perrin, who is the second person to make considerable advancements to the now known Brownian motion. After experimentation, Perrin concludes that this observed random movement is not caused by living organisms, directly disproving Brown's claim (Perrin, 1913). Perrin claims that the movement is the result of collision between water particles with those suspended particles. Perrin coins this phenomenon in his paper as "molecular agitation". It would later turn out that Perrin's findings played a much more pivotal role in science as it would prove the existence of atoms which was one of the topics of great debate then.

Brownian motion continued to be an active research topic with minor breakthroughs and applications being explored continuously. The next big breakthrough however arrived through the work of Metzler and Klafter (Metzler and Klafter, 2000). They were responsible for linking the anomalous diffusion, which is a phenomenon that occurs when the movement of a particle deviates significantly from its expected trajectory with many complex systems. This has in fact been observed in many physical and biological systems and has in turn been responsible for paving the way for numerous breakthroughs in the field of dynamical systems.

Applications of Brownian Motion

Let us now talk about the various applications of Brownian motion at modern times. Since the process is still an active topic of research, it has been used continuously used in many scientific fields as a consequence. Specifically, some of the more important ones are the following:

- Physics
- Materials science
- Biology
- Finance
- Engineering

When it comes to physics, one of the most important applications of the Brownian motion is the measurement of the various particles' size. Specifically, since Brownian motion is the cause for the random movement of particles suspended in a fluid, we are able to leverage that in order to measure the mass of those particles. This technique is called dynamic light scattering and is, in fact, in use in many industries (Carpenter, 1977). Another application is how Brownian motion can be utilized to measure the viscoelastic properties of complex fluids, providing insights into their structural and dynamical behavior (Gardiner and Zoller, 2004), other applications include the collective dynamics of active matter systems, such as self-propelled particles or living organisms, which exhibit Brownian-like motion (Ramaswamy, 2010).

In materials science, one major application of Brownian motion is the study of polymers' movement. For example, when considering a polymer chain's movement we can observe that it is caused by Brownian motion. Through this admission, we may calculate the shape and size of the polymer and even go as far as gleaning important insights about the nature of some of its mechanical properties (Feldman, 1989). Another application is modeling diffusion processes in various materials, including porous materials (Cussler, 1997), or the study of thermal fluctuations and Brownian motion in metals and provides insights into their role in the dielectric relaxation processes (Dutta and Horn, 1981).

In biology Brownian motion has been used to model the movement of proteins or other molecules that are found within cells, or even the movement of the cells themselves (Berg, 1993). One more interesting and modern application is its use in drug delivery systems, where it is responsible for handling the release of drug via nanoparticles (Jokerst et al., 2011). Other applications include the role of Brownian motion in the motility and chemotaxis of microorganisms (Berg, 2011) or its role as the generation and propagation of spontaneous action potentials in cerebellar Purkinje neurons (Raman and Bean, 1999).

Finance and specifically financial markets is another area where Brownian motion has been applied to. The process is capable to model the various behaviors exhibited by financial markets and in more detail, the fluctuations of interest rates, stocks and other financial variables (Black, 1976).

Lastly, in engineering some use cases for Brownian motion include the phenomenon of stochastic resonance in ring lasers (McNamara et al., 1988), as well as, the use of Brownian motion in characterizing nanoparticle dynamics, diffusion processes in nanostructures, and the behavior of nanoparticles in fluidic systems (Cao and Wang, 2011).

Mathematical introduction of Brownian motion

In order for a stochastic process to be characterized as a Brownian motion a number of conditions must hold. For example a process $X = \{X_t : t \geq 0\}$ must satisfy:

- $X_{t_0} = 0$
- X_t has continuous paths almost surely
- X_t has independent increments
- $X_t - X_s \sim \mathcal{N}(0, t-s)$, where $0 \leq s \leq t$

Besides the necessary conditions for a Brownian motion, it also has a number of interesting properties such as:

- Scaling property. If we consider for each $s > 0$ the process $\{s^{-1/2}X_{st} : t \geq 0\}$ is a Brownian motion with 0 being the origin point as the definition above. Also, for each $s > 0$, we have

$$\{X_{st} : t \geq 0\} \stackrel{\mathcal{D}}{=} \{s^{1/2}X_t : t \geq 0\}.$$

- Shifting property. For each $s > 0$ the process $\{X_{s+t} - X_s : t \geq 0\}$ is a Brownian motion starting from 0 and simultaneously it is independent of the process $\{X_u : 0 \leq u \leq s\}$.
- Time reversal property. Considering a Brownian motion $X = \{X_t : 0 \leq t \leq 1\}$, let us define another process $Z = \{Z_t : 0 \leq t \leq 1\}$ so as $Z_t = X_{1-t} - X_1$, then

$$\{Z_t : 0 \leq t \leq 1\} \stackrel{\mathcal{D}}{=} \{X_t : 0 \leq t \leq 1\}.$$

- Inversion property. Let us define a process $Z = \{Z_t : t \geq 0\}$ for which $Z_0 = 0$ and $Z_t = tX(1/t)$. Then Z is a Brownian motion as well, with 0 as its origin point.

Simulation of the process

As we mentioned earlier, we intended to include for each process covered in this segment some simulations. Using the programming language Python we simulated three random realizations of Brownian motion. Going one step further and for the sake of seeing the Lévy-Itô decomposition visually, we have also incorporated this in our code and naturally into the resulting graphs.

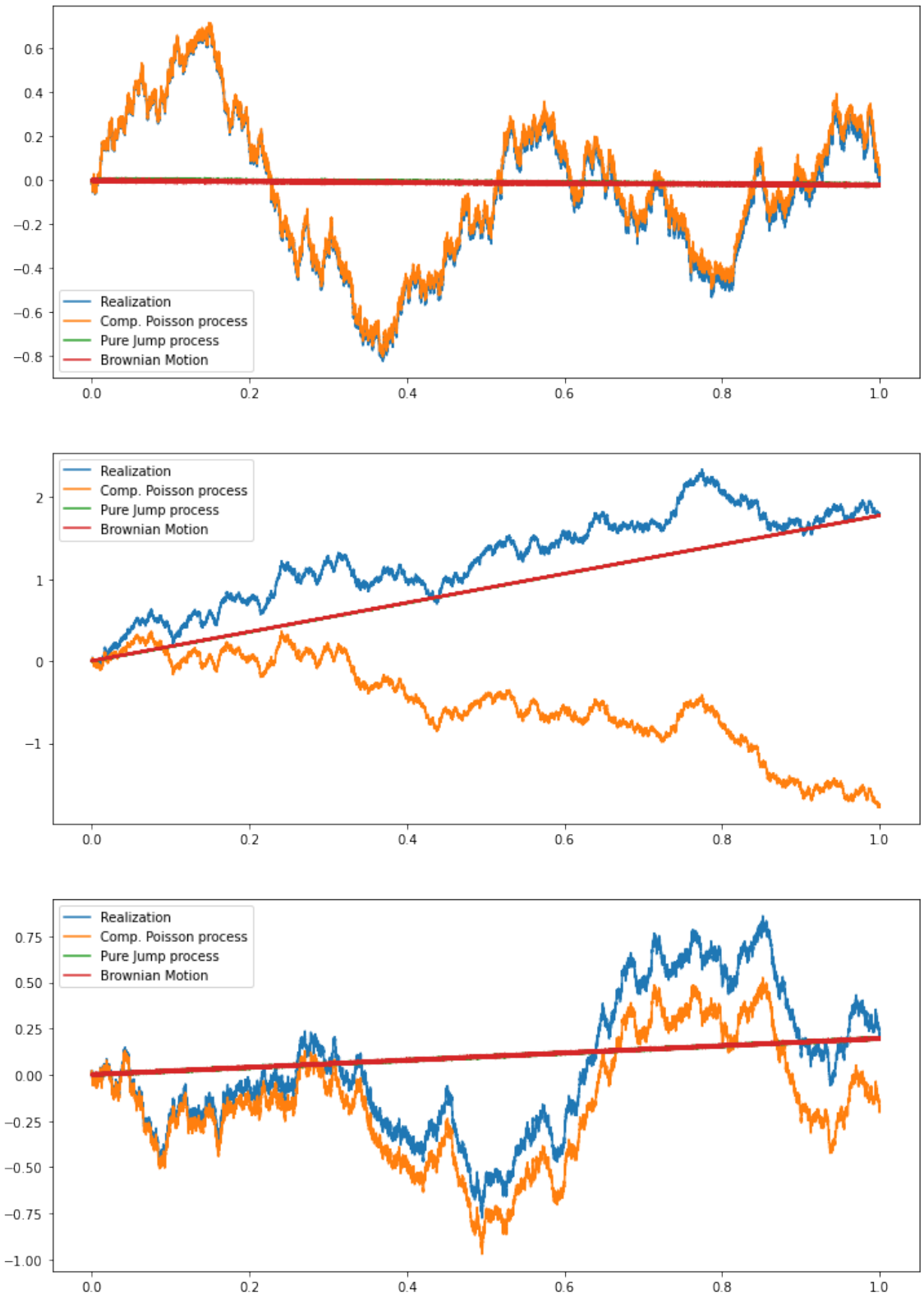

Figure 1.1: Realizations of Brownian Motion with applied Lévy-Itô Decomposition.

Taking a quick glance at the three plots, one of the most notable things is that the jump process is a straight line, similar to the drift. This is an expected outcome, as by definition the Brownian motion is the only stochastic Lévy process that is characterized by continuous paths almost surely.

One more thing to note is the goodness of fit between the original realizations of the process and the Brownian motion decomposed part. This, in part, is because the amount of linear drift used for the simulation is minor. Should a higher value of linear drift be used, the fit would become progressively worse.

1.6.2 Poisson Process

Background of Poisson process

The concept of the Poisson process dates back to the work of Poisson (Poisson, 1837) when he proposed a model capable of explaining the distribution of errors occurring during the trials of criminal and civil cases. Poisson was also able to discover a probability mass function capable of modelling the probability of occurrence of these exceedingly rare events. Furthermore, Poisson also analytically derived the the first two moments of his new discovered distribution.

In the upcoming years, one of the most major contributions for this process was the work by Cox, who introduced the non-homogeneous Poisson process, where the intensity of event occurrences varies over time (Cox, 1955). Other important contributions linked with the Poisson process is the theory of point processes (Daley and Vere-Jones, 2014).

Other interesting aspects of the theory of Poisson processes is their interplay with large deviation theory and their relation to the class of branching processes (Kingman, 1993).

Naturally, as Poisson process happens to be widely adopted and a highly researched process, there exist various extensions for it. One of which being the compound Poisson process that we have talked about in previous segments.

Applications of Poisson process

Similarly to Brownian motion, the Poisson process has been extensively used in many fields, proving to be a powerful model capable of solving multiple problems and of varying nature. For example, one of the more earlier applications of the process was by Erlang (A. Erlang, 1917) where he utilized the stochastic process to provide some solutions for various problems faced in the telecommunication sector at the time.

Another important application for the Poisson process has been the catalytic role it played for the development of queuing theory (Tocher and Takacs, 1963). Tocher

and Takacs present a lot of meaningful examples and showcase the interplay between the Poisson process and queuing theory.

Poisson process has also been used in the field of biology for multiple problems. Perhaps, one of the most important is the ability to be the basis for modeling infection rates for various viruses (Anderson and May, 1992; Radcliffe, 1977).

Another great resource for unraveling the applications of this process is the work by Barlow, Proschan and Hunter where they are applying the Poisson process among other models to problems concerning the field of reliability engineering (McCall, 1965). Of course, Poisson process is widely adopted in the field of finance as well as it was shown from the theory presented in this thesis so far.

Mathematical introduction of Poisson process

A process valued on the non-negative integers, $N = N_t : t \geq 0$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, is said to be a Poisson process with intensity $\lambda > 0$ if the following hold:

- The paths of N are $\mathbb{P}$-almost surely right-continuous with left limits.
- $\mathbb{P}(N0 = 0) = 1$.
- For $0 \leq s \leq t$, $N_t - N_s \overset{\mathcal{D}}{=} N_{t-s}$.
- For $0 \leq s \leq t$, $N_t - N_s$ is independent of $\{N_u : u \leq s\}$.
- For each $t > 0$, $N_t \overset{\mathcal{D}}{=} X \sim Poisson(\lambda t)$.

Simulation of the process

Identically to the previous example with the Brownian motion we include a simulation of three realizations of a Poisson process, along with their Lévy-Itô decomposition. One of the more noteworthy parts about the plot presented here is the interaction between the jumps of the realization of the Poisson process and the assignment of jumps from the Poisson random measure for the same time interval.

1.6.3 Gamma Process

Background of Gamma process

One of the first scientists to use the Gamma process was by Erlang (A. K. Erlang, 1909) who studied the theory of probabilities in the context of telephone conversations. It is also worth noting that in his paper, Erlang did not explicitly used the term Gamma process. Some initial statistical properties for the process were introduced by Doob (Doob, 1942) who studied the Gamma process in conjunction with Brownian motion. He also discussed the gamma process as a continuous-time

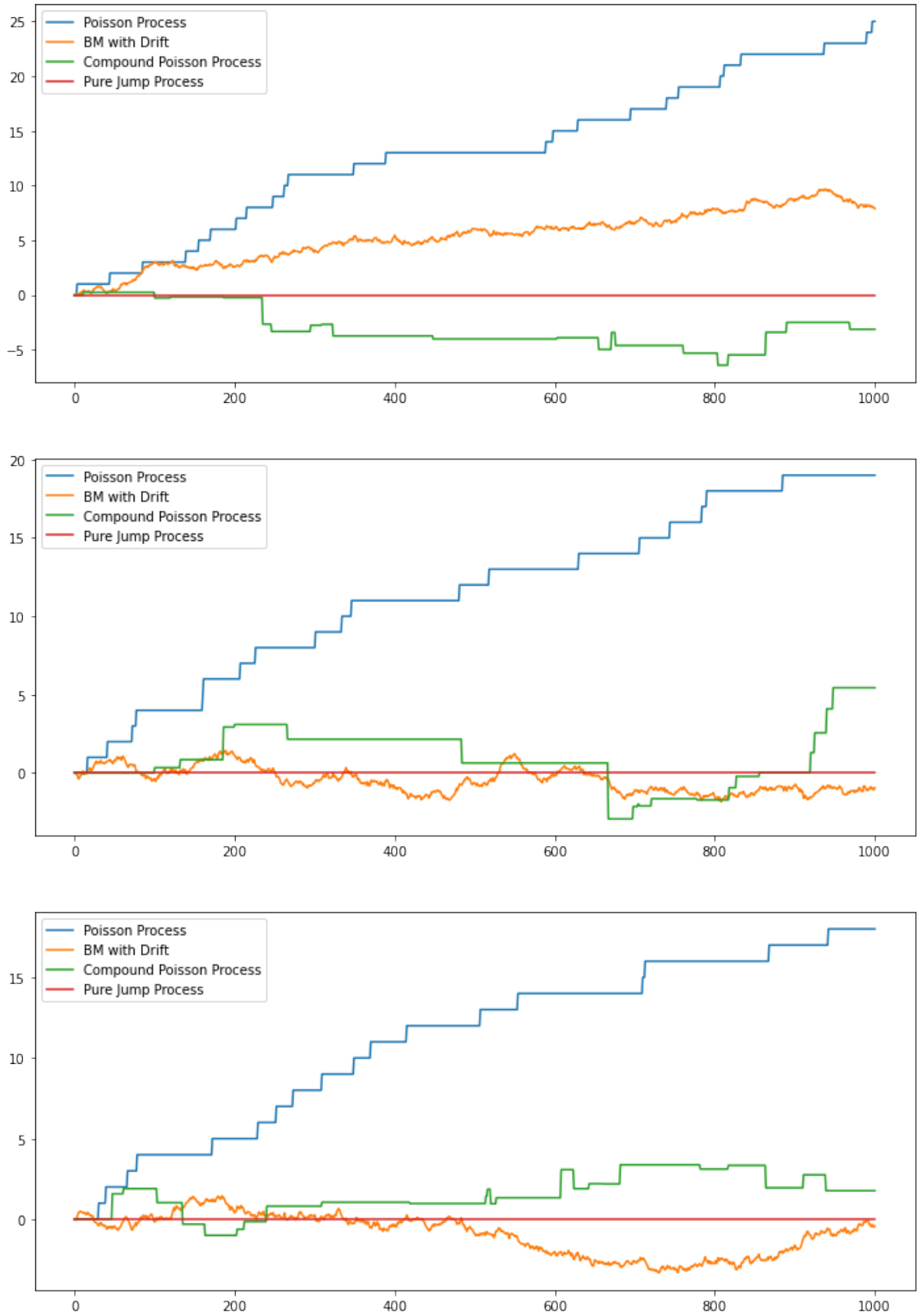

Figure 1.2: Realizations of Poisson process with applied Lévy-Itô Decomposition.

Markov process and its relation to stochastic differential equations. Besides Doob, Bertoin (Bertoin, 1998) provided a quite extensive treatment of Lévy processes which includes the Gamma process.

Another advancement can be considered the exploration of using Gamma processes in Monte Carlo simulations for pricing path-dependent options (Glasserman et al., 2001). They discuss importance sampling and stratification techniques to efficiently estimate option prices under gamma process dynamics. Lastly, it is also worth mentioning that the Gamma process has many established extensions and generalizations, with some of them being discussed in Schoutens' book (Schoutens, 2003).

Applications of Gamma process

Let us discuss a bit about the applications of Gamma process in the various fields, starting with the field of Queuing theory where it has been used to model arrival and service times in telecommunication systems (Kleinrock, 1975). In an adjacent field, reliability engineering, and specifically failure analysis it is established that Gamma process can model the failure rate and lifetime of mechanical or electrical components in systems (Crow, 2018).

In the field of finance and especially in Risk theory, it has been established that it has beneficial properties for modeling the various problems (Gerber and Shiu, 1997).

Mathematical introduction of Gamma process

A process $\Gamma = \{\Gamma_t : t \geq 0\}$ will be a Gamma process if the following conditions always hold:

- $\mathbb{P}(\Gamma_0 = 0) = 1$ almost surely.
- $\forall s < t$, $\Gamma_t - \Gamma_s \sim Gamma(\alpha, \beta(t-s))$.
- Γ_t has $i.i.d.$ increments.

Additionally, $\alpha, \beta \in \mathbb{R}^+$. One more thing of interest about the Gamma process is that its moments are undefined, thus its mean and variance are actually infinite. However, its moments are always well-defined even so. This is one of the most well known examples of Lévy processes which have infinite mean and variance and yet the Lévy-Itô decomposition can still work without issues.

Simulation of the process

Now let us take a look at a Gamma process and how it is decomposed under the Lévy-Itô Decomposition.

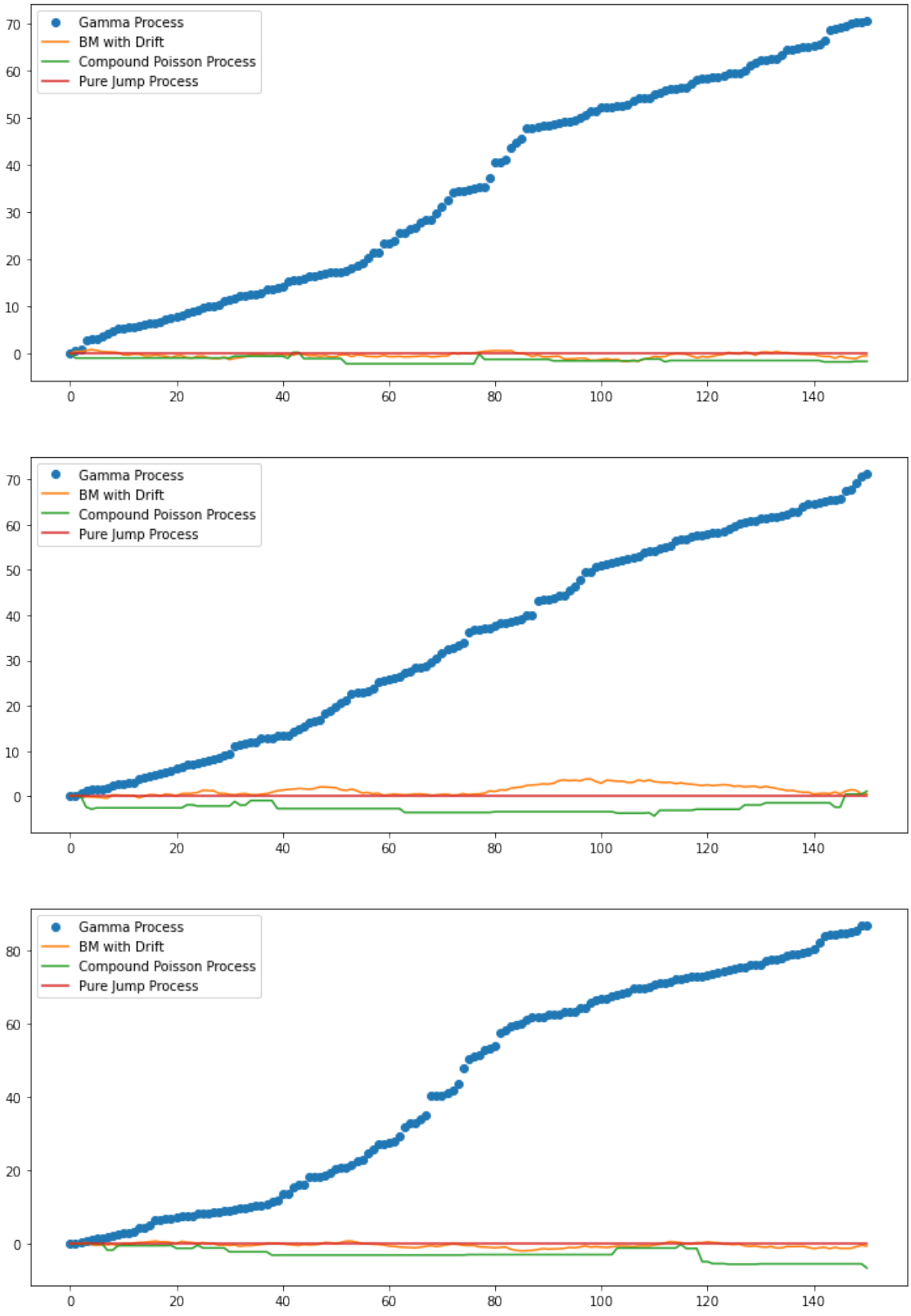

Figure 1.3: Realizations of Gamma process with applied Lévy-Itô Decomposition.

1.6.4 Variance Gamma Process

Background of Variance Gamma process

Variance Gamma process was introduced by Madan, Carr and Chang as an extension of the Brownian motion model (Madan et al., 1998). Their aim was to more accurately model the behavior of financial data, specifically their properties such as jumps, heavy tails and skewness, which they demonstrate via utilizing the proposed Variance Gamma process model for option pricing.

Applications of Variance Gamma process

The process was introduced as a model designed for finance and so far, the vast majority of applications for the Variance Gamma process are located in the field of finance. Some of them include its use in general financial modeling Barndorff-Nielsen and Shephard, 2001).

Mathematical introduction of Variance Gamma process

A Variance Gamma process is defined by the following stochastic differential equation:

$$dX(t) = \mu dt + \sigma\sqrt{V(dt)} + dJ \tag{1.29}$$

where μ is the drift of the process, σ is the volatility of the process, V is a Gamma process, J is a Compound Poisson distribution.

Simulation of the process

Below we present some simulations for the process at hand. Please note that in order to achieve negative values as well and not strictly positive, the generation for the realizations of the Variance Gamma process followed the Euler-Maruyama method instead of sampling from a Gamma distribution.

1.6.5 Inverse Gaussian Process

Background of Inverse Gaussian process

The inverse Gaussian process is a relatively newer process. A rigorous definition of the main conditions as well as some basic properties of the process was given by Wasan (Wasan, 1968). Jorgensen provides an extensive treatment to the process and its many properties (Jorgensen, 1981).

Applications of Inverse Gaussian process

The Inverse Gaussian process has been used extensively in financial applications, such as general asset pricing and risk management.

Mathematical introduction of Inverse Gaussian process

Formally, a process $X = \{X(t), t \geq 0\}$ is Inverse Gaussian if it satisfied the following conditions:

- X_t has independent and stationary increments.
- The Inverse Gaussian process is a subordinator, so it is non-negative and increasing almost surely.
- $X_t - X_s \sim InverseGaussian(\mu(t-s), \lambda(t-s))$, where $0 \leq s < t$.
- The probability density function is as follows:

$$InverseGaussian(X; \mu, \lambda) = \left(\frac{\lambda}{2\pi X^3}\right)^{\frac{1}{2}} \exp\left[\frac{-\lambda(X-\mu)^2}{2\mu^{2X}}\right]. \tag{1.30}$$

- $\mathbb{P}(X_{t_0} = 0) = 1$ almost surely.
- X_t has continuous paths and no jumps with probability one.

1.6.6 Normal Inverse Gaussian Process

Background of Normal Inverse Gaussian Process

The Normal Inverse Gaussian process is as well one of the newer Lévy processes to be studied. Some of the core literature for the process include the initial papers by Barndorff-Nielsen (Barndorff-Nielsen, 1997) and Barndorff-Nielsen and Shephard (Barndorff-Nielsen and Shephard, 2001). They introduce and study at depth the Normal Inverse Gaussian Process as well as its properties, while also evaluating the process from a asset pricing point of view.

Applications of Normal Inverse Gaussian Process

The Normal Inverse Gaussian is one of the most widely adopted models in finance and it has been the base and inspiration for the conceptualization of other other models, which can be considered in turn as a generalization of the process.

Mathematical introduction of Normal Inverse Gaussian Process

The Normal Inverse Gaussian (NIG) process is characterized by four parameters: location parameter μ, scale parameter α, asymmetry parameter β, and tail thickness parameter δ. More specifically, a stochastic process X_t is said to be a NIG process if it satisfies the following conditions:

- $\mathbb{P}(X_0 = 0) = 1$ almost surely.
- $X_t - X_s \sim NIG(\mu(t-s), \alpha(t-s), \beta, \delta)$ for all $0 \leq s \leq t$.
- The increments $X_t - X_s$ are stationary and independent.

1.6.7 Stable Process

Background of Stable Process

The history of Stable Processes can be traced back to the early 20th century, with the pioneering work of Lévy (Brookes, 1955) and Khintchine (Khinchin, 1948). They introduced the concept of stable distributions. Through their work, they showed that these distributions have the property of stable scaling, a property we discussed quite in-depth when we covered the infinitely divisible distributions in this thesis.

Samorodnitsky and Taqqu (Samorodnitsky and Taqqu, 1994) and Nolan (Nolan, 2001) provide an accurate and extensive recount of the surrounding theory as well as all relevant properties and conditions of various stable processes.

Applications of Stable process

Within the bounds of the field of finance, this process has been widely used. There are many scholarly articles regarding its use in the field. We may turn to Basegmez and Cekici who utiliza Lévy stable processes to study the implications in the Turkish stock market (Basegmez and Cekici, 2017).

Other articles and related texts feature the processes in a comprehensive manner along with rich examples covering many facets of risk theory (Samorodnitsky and Taqqu, 1994; Nolan, 2001; Embrechts et al., 1997).

Mathematical Introduction of Stable Process

A process X is a Stable process, often denoted as $X_t \sim S(\alpha, \beta, \gamma, \delta)$, where α is the scale parameter, β is the stability parameter, γ is the location parameter, and δ is the skewness parameter, if it satisfies the following conditions:

- The increments of X are admitted by a Stable probability law, or as we introduced it in this thesis, an infinitely divisible law.

- X has continuous sample paths almost surely.
- X has independent and stationary increments.

It is worth to note a Stable process exhibits finite variance if and only if $\alpha < 2$. Additionally, also note that there exist established extensions of the Stable process in the form of Tempered and Generalized Tempered Stable processes among others, which however are not Lévy processes themselves, as they do not satisfy the necessary conditions.

Simulation of the process

We introduce the following plot that visually shows three random realizations of a Lévy stable process with the same parameters.

The reason that the plot does not contain the decomposition of the process is that its decomposition is not well-defined.

1.6.8 Meixner Process

Background of the Meixner Process

The Meixner process is a special type of Lévy process which originates from the theory of orthogonal polynomials. It is related to the Meixner-Pollaczek polynomials by a martingale relation (Schoutens, 2002; Schoutens, 2003).

Applications of the Meixner Process

The Meixner process has been a widely adopted process in the field of finance in order to provide solutions to asset pricing problems with respect to many different derivatives. Again, Schoutens, 2002 and Schoutens, 2003 provide an account of the theory of the model along with examples of applying the model to financial related problems.

Mathematical Introduction of the Meixner Process

Considering a process $X = \{X_t : t \geq 0\}$, we may say that it is a Meixner process if the following conditions are satisfied.

- X has stationary and independent increments
- The increments of X are Meixner distributed, or alternatively $X_t - X_s \sim Meixner(\alpha, \beta, \delta, \mu)\ \forall t > s \geq 0$.

The probability densify function of the Meixner distribution is as follows:

$$f(x; \alpha, \beta, \delta, \mu) = C(\alpha, \beta, \delta) \exp\left[-\alpha(x - \beta)\right] M(\alpha, \delta; 2\alpha; 2\delta(x - \beta))$$

where $C(\alpha, \beta, \delta)$ is the normalization constant and $M(\alpha, \delta; 2\alpha; 2\delta(x - \beta))$ denotes the Meixner's confluent hypergeometric function.

1.6.9 Lévy subordinator process

Introduction of Lévy subordinator process

Subordinators are a special class of a Lévy processes which are especially important in the theory of local time. An accurate account of the theory of the subordinators can be given by Kallenberg in his book (Kallenberg, 2002).

Applications of Lévy subordinator process

Since this process provides great insights within the context of local time, it can and finds applications in any field which would require to solve problems in the form of first passage or exit problems.

Should we look specifically in the field of finance, we may turn to the work of Li, Li and Zhang who have used this process in the context of pricing and hedging derivatives (Li et al., 2017).

Mathematical introduction of Lévy subordinator process

We can start by defining the process. Thus, a subordinator is a $\mathbb{R}$-valued stochastic process $X = \{X_t : t \geq 0\}$ that satisfies the following conditions:

- $\mathbb{P}(X_0 = 0) = 1$ almost surely.
- $\mathbb{P}(X_t > 0) = 1$ almost surely.
- X has stationary and independent increments.
- X has càdlàg paths.

1.6.10 Spectrally one sided Lévy Process

Background of Spectrally one sided Lévy Process

The spectrally one-sided Lévy process is a class of stochastic processes that has gained prominence in the field of probability theory. It represents a subset of Lévy processes with sample paths confined to either the positive or negative half-line. The development of the spectrally one-sided Lévy process can be traced back to foundational works by several mathematicians. It is a special case of Lévy processes that is of great interest.

Applications of Spectrally one sided Lévy Process

Let us also discuss the applications of this special case of processes. Specifically, we are interested in their applications in finance. In Ruin theory, it has been used to model the positions over time of a portfolio (Asmussen and Albrecher, 2010). In credit risk theory, we can find some applications of these processes to successfully model the possibility of defaults as well as the general behavior of credit spreads (Bielecki and Rutkowski, 2013).

Mathematical introduction of Spectrally one sided Lévy Process

We can define a Spectrally one sided Lévy process as follows. Let us assume a process $X = \{X_t : t \geq 0\}$. If it satisfies the following properties then it will be a Spectrally one sided Lévy process:

- $\mathbb{P}(X_t > 0) = 1$ almost surely if it is a Spectrally positive process and $\mathbb{P}(X_t < 0) = 1$ almost surely if it is a Spectrally negative process.
- X has independent and stationary increments.
- $X_t - X_s \sim Lvy(\mu, c)$ with a probability density function as follows:

$$f(x; \mu, c) = \sqrt{\frac{c}{2\pi}} \frac{\exp\left[-\frac{c}{2(x-\mu)}\right]}{(x-\mu)^3/2}.$$

1.7 Wiener-Hopf factorization

When we were discussing the Lévy-Itô decomposition and its inner workings, we ended that segment by mentioning that there do exist alternative forms of decomposing a Lévy process in the literature. In this segment, we further explore that idea and introduce some of these techniques.

It is important to note, that we will not discuss these techniques at the same depth as the Lévy-Itô decomposition, as the aim of this segment is to provide some breadth regarding the different available methods. Additionally, each method has different advantages and disadvantages, which we will present and discuss later.

Lastly, while it might be obvious at this point, not all Lévy processes can be decomposed with all methods. Due to the nature of some processes which stems from their definition as well as the limitations of some of these methods, there do exist processes where only specific decomposing methods can be used. This is presented in the appropriate advantages and disadvantages section of each method.

Thus, the main takeaway should be that when attempting utilize Lévy processes for modeling problems of different nature, we are facing multi-dimensional problem in making the optimal choice for a decomposition method. This optimal choice is not universal, but will heavily depend on the researcher's or practitioner's priorities, such as wanting efficient computation or unique representations.

1.7.1 Introduction

Let us now briefly explain the necessary theory and background for the Wiener-Hopf factorization before looking at an example. Contrary to the Lévy-Itô decomposition, this methods aims to decompose a process into two parts, with them being the analytic and anti-analytic part respectively.

Let us consider a process X, the idea is to compute X^+ and X^- so that the following equation holds at all times:

$$X = X^+ \cdot X^- \tag{1.31}$$

where X^+ is analytic in the positive plane, while X^- is analytic in the negative plane. Thus, as it is now quite obvious, according to this method we are attempting to decompose a process into those parts, where each is analytic in its corresponding plane.

Naturally, for the decomposition to work, certain conditions must exist. Let us go over them briefly, by using the process $X(t)$ we discussed previously.

- The process has to be analytic in the positive plane in such a way so that $\mathbb{C}^+ = t \in \mathbb{C} : \mathrm{Im}(t) > 0$ while also not exhibiting any discontinuities on $\mathbb{R}$, except for at most countably finite number of simple poles.
- The process must satisfy a suitable decay condition as $|t| \to \infty$ in the positive plane.
- For any given process or function that we wish to decompose accordingly, there will exist two functions $g(t)$ and $h(t)$, where each of these function is analytic in the positive plane while also continuous in $\mathbb{R}$, in such a way so that $X(t) = g(t)h(t)$ for all $t \in \mathbb{C}^+$.

Naturally, the Wiener-Hopf factorization has much greater depth than this, but we shall draw the line here. Additionally, it is important to also state that these conditions and restrictions make up for a general case of the method. Different Lévy processes, when required, may need to call upon additional restrictions in order to apply the method successfully.

1.7.2 Example simulation

With the theory out of our way, we are interested in showing visually how this method of decomposition works when applied on a Variance Gamma process.

The content of the plot should be quite self-explanatory. We can visually see how the Wiener-Hopf Decomposition works, by taking closer look to the two different colored lines representing the positive and negative components of the process.

1.7.3 Advantages and disadvantages

Lastly, before ending the section on Wiener-Hopf decomposition we dedicate a little time to discussing the various advantages and disadvantages of the method.

Let us start with the advantages:

- The Wiener-Hopf decomposition can be used to calculate the first-passage probabilities of Lévy processes.
- The Wiener-Hopf decomposition can be used to study Lévy processes with jumps of varying sizes, as long as the jump distribution is symmetric.
- The Wiener-Hopf decomposition can be used to study the tail behavior of Lévy processes.
- The Wiener-Hopf decomposition can be applied to a wide range of Lévy processes, including those with infinite variation, non-zero drift, and non-zero correlation.

Naturally, while the method exhibits some impressive advantages concerning the study of Lévy processes, there are a number of disadvantages and limitations as well:

- The Wiener-Hopf decomposition is only applicable to Lévy processes with symmetric jump distributions. This means that the positive and negative jumps must have the same distribution. If the jump distribution is asymmetric, the decomposition cannot be used.
- In some cases, it may not be possible to calculate the Wiener-Hopf factorization of a Lévy process. This is because the factorization requires the characteristic exponent of the process to be known, and in some cases, it may not be possible to determine this exponent.
- While the Wiener-Hopf decomposition can provide insights into the behavior of Lévy processes, it may not provide a complete picture of the process. This is because the decomposition only considers the positive and negative jumps of the process and does not take into account the interactions between them.
- The Wiener-Hopf decomposition is not applicable to Lévy processes with infinite activity, as the positive and negative parts of the process may not have

finite moments.

- While the Wiener-Hopf decomposition can be useful for analyzing Lévy processes in certain applications, it may not be useful or relevant for other applications.

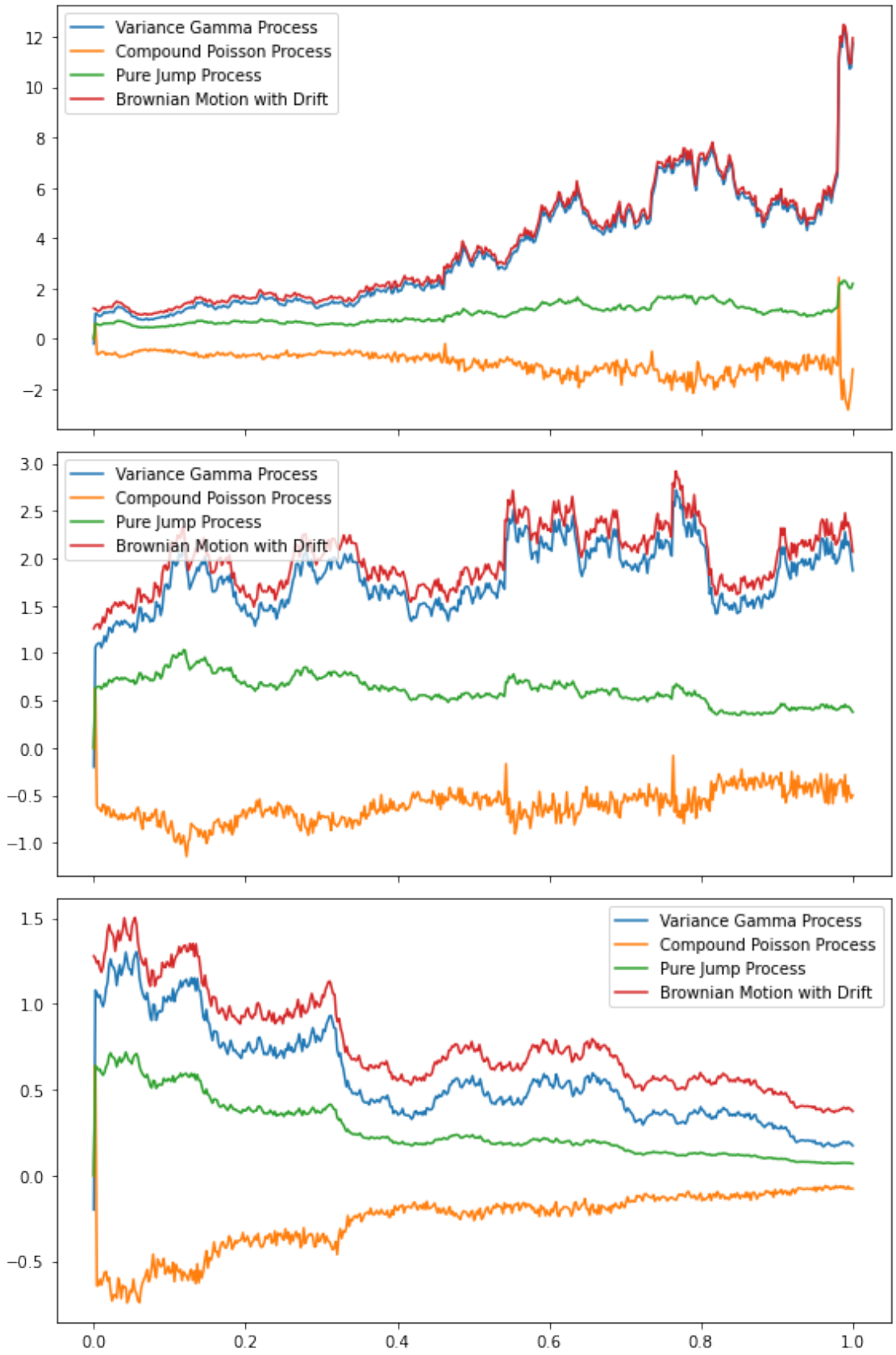

Figure 1.4: Realizations of Gamma process with applied Lévy-Itô Decomposition.

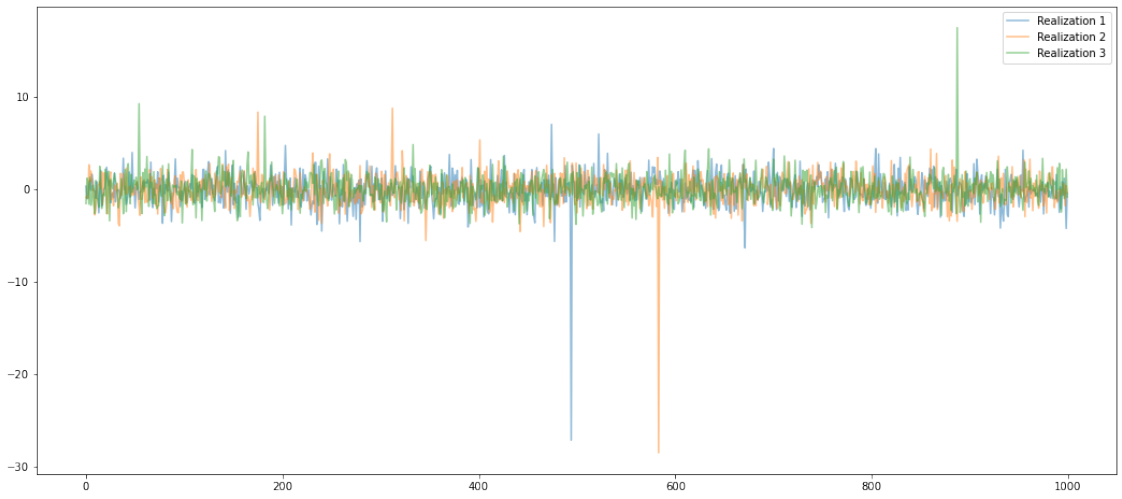

Figure 1.5: Three random realizations of a Lévy stable process.

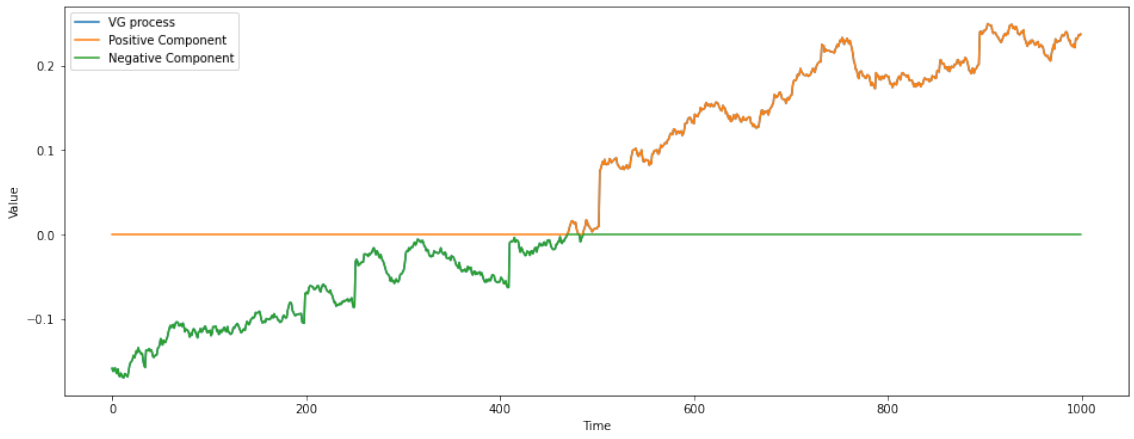

Figure 1.6: Applying the Wiener-Hopf factorization on a Variance Gamma process.

Chapter 2

Risk Theory

This will be the second part of this thesis, where the primary goal is to showcase the connection of the all the theoretical mathematical concepts regarding Lévy processes with Risk theory.

We attempt to achieve our goal by defining the concept of risk theory and the various concepts that it encompasses. Afterwards, we bridge Lévy processes and risk theory by showing how Lévy's framework can greatly assist the end goal of risk theory.

We continue by discussing some important Lévy-driven models that have made a name for themselves in the quantitative finance circles. Lastly, we end this chapter by discussing different computational methods and algorithms that simulate Lévy processes and play a catalytic role in the general risk management field.

2.1 Introduction to general Risk theory

Risk is an essential component of finance and refers to the possibility of loss or the variability of returns on an investment or portfolio. The need to manage risk arises because of the uncertainty associated with future outcomes, and financial institutions and investors have developed various techniques to mitigate and manage these risks.

The foundations of risk theory were laid in the early 20th century by actuaries and mathematicians such as Emil Gumbel and Harald Cramér who introduced the concept of risk as a probabilistic measure of uncertainty. Gumbel's work on extreme value theory and Cramér's contributions to probability theory and statistics were particularly influential in the development of risk theory.

In the context of insurance, risk theory focuses on the analysis of claims and premiums, and the calculation of various measures of risk such as the probability of ruin, the expected value of claims, and the premium load. One of the central models in risk theory is the classical risk model, which assumes that claims arrive according to a Poisson process and that the claim sizes are independent and identically distributed (i.i.d.) random variables. The classical risk model has been studied extensively in the literature, and several extensions and generalizations have been proposed to account for various real-world scenarios such as correlated claims, nonstationary claim arrival rates, and dependent claim sizes (Kaas et al., 2008; Embrechts et al., 1997).

Another important model in risk theory is the ruin model, which aims to quantify the risk of an insurance company or financial institution going bankrupt due to unforeseen events or losses. Ruin theory has been a subject of active research and several models and methods have been developed to estimate the probability of ruin, the ruin time, and the ruin probability density function (Gerber, 1979; Embrechts et al., 1997).

In the context of finance, risk theory is concerned with the measurement and management of financial risks such as market risk, credit risk, and operational risk. Market risk, in particular, has been an active area of research in risk theory, and several models and methods have been proposed to estimate various measures of market risk such as Value-at-Risk (VaR), Expected Shortfall (ES), and Conditional VaR (CVaR) (McNeil et al., 2015; Klüppelberg et al., 2004).

One of the key contributions to market risk theory is the development of Value-at-Risk (VaR) (Jorion, 1997). VaR is a widely used measure of market risk that estimates the potential loss in a portfolio over a given time horizon, at a given level of confidence.

Other important contributions to market risk theory include the development of risk measures such as expected shortfall (ES) (Artzner et al., 1999), coherent risk measures (Föllmer and Schied, 2011), and risk measures with emphasis on heavy tails (McNeil and Frey, 2000). These measures provide alternative ways of quantifying market risk, and can be used to address some of the limitations of VaR.

In addition to the development of risk measures, market risk theory has also focused on the development of risk management tools and techniques, such as risk budgets, risk limits, and risk reporting frameworks (Basel Committee on Banking Supervision, 2004). These tools are designed to help financial institutions manage their market risks more effectively, and to comply with regulatory requirements.

Mathematically, risk can be represented using different measures such as variance, covariance, expected shortfall, Value-at-Risk (VaR), Conditional Value-at-Risk (CVaR), Expected Shortfall (ES), and others. These measures provide a quantitative way of assessing and comparing risks associated with different investments and portfolios.

Lévy processes are well-suited for modeling financial returns due to their ability to capture the heavy-tailed and skewed nature of financial returns. One of the most

commonly used Lévy processes in financial risk theory is the Brownian motion with drift, which is used to model the market index. Other commonly used Lévy processes in financial risk theory include the Variance Gamma process, the Normal Inverse Gaussian process, the Meixner process and others.

Various studies have explored and established the importance of risk management and its relationship with financial performance, including Lévy process-based models. One of the main challenges in financial risk theory is the estimation of the parameters of the Lévy process, which are necessary for computing risk measures. Estimation methods include maximum likelihood estimation, Bayesian estimation, and empirical estimation based on historical data. In addition, there are various methods for modeling dependence between assets, such as copula-based models and factor models.

2.2 Ruin theory

2.2.1 Introduction

The classical ruin problem is a stochastic model used to study the probability of an insurance company or a financial institution facing financial ruin due to claims or liabilities exceeding its available capital. It is based on a continuous-time stochastic process, typically a Lévy process, that models the evolution of the insurer's capital. The process is assumed to be non-negative and to start at some initial capital level u. The insurer receives premiums at a constant rate and pays claims according to a Poisson process with intensity λ and claim size distribution F. The insurer's capital level X_t at time t evolves according to the equation:

$$X_t = u + ct = \sum_{i=1}^{N_t} Y_i \tag{2.1}$$

where N_t is a Poisson process with intensity λ representing the number of claims up to time t,Y_i is the size of the ith claim, and c is the constant premium rate. The insurer is said to be ruined if X_t falls below zero for some t.

The probability of ruin, denoted by $\psi(u)$, is the probability that the insurer's capital level ever falls below zero, or alternatively:

$$\psi(u) = \mathbb{P}(\exists 0 \leq t < \infty : X_t < 0). \tag{2.2}$$

The key quantity in ruin theory is the ruin probability $\psi(u)$. In order to calculate $\psi(u)$, one needs to solve the following boundary value problem for the Laplace transform of the ruin probability:

$$\mathcal{L}\{\psi(s)\}(u) = 1 - \frac{\mathcal{L}\{\phi(s)\}(u)}{\mathcal{L}\{R(s)\}(u)} \tag{2.3}$$

where $\mathcal{L}$ denotes the Laplace transform, $\phi(s)$ is the Lévy exponent of the Lévy process, and $R(s)$ is the Laplace transform of the ruin time density function. The Laplace transform of the ruin time density function is given by:

$$R(s) = \int_0^\infty e^{st}\mathbb{P}(T > t|X_0 = u)dt \tag{2.4}$$

The function $R(s)$ is known as the ruin time generating function.

The solution to the boundary value problem for the Laplace transform of the ruin probability is given by the Wiener-Hopf factorization theorem, which expresses the Laplace transform of the ruin probability as the product of two functions:

$$\mathcal{L}\{\psi(s)\}(u) = \frac{\psi^+(u)}{\psi^-(u)} \tag{2.5}$$

where $\psi^+(u)$ and $\psi^-(u)$ are the positive and negative solutions of the Wiener-Hopf equation:

$$\mathcal{L}\{\phi(s)(u)\psi^+(u)\} - \mathcal{L}\{\phi(-s)\}(u)\psi^-(u) = \mathcal{L}\{R(s)\}(u) \tag{2.6}$$

The ruin probability can then be obtained by inverting the Laplace transform of $\psi(s)$.

2.2.2 Ruin theory models

Cramér-Lundberg model

The Cramér-Lundberg model, also known as the classical risk model, is one of the simplest and most widely used models in ruin theory. It is a mathematical model for an insurance company's surplus process over time, taking into account both incoming premiums and outgoing claims.

Let the surplus process of an insurance company at time t be denoted by $U(t)$, with initial surplus u, and let $\lambda > 0$ and $\mu > 0$ be the constant rates of incoming premiums and outgoing claims, respectively. Then, the surplus process satisfies the following stochastic differential equation (SDE):

$$dU(t) = \lambda dt - \mu dN(t) \tag{2.7}$$

where $N(t)$ is a Poisson process with rate $\alpha > 0$ representing the number of claims

arriving in $(0, t]$.

The Poisson process $N(t)$ is assumed to be independent of the Brownian motion driving the Lévy process $X(t)$, which is used to model the company's investment returns. The Lévy process $X(t)$ is assumed to be a spectrally negative process, which means that its Laplace exponent $\psi(\omega)$ is of the form:

$$\psi(\omega) = \frac{1}{2}\sigma^2\omega - \alpha\omega - \int_{(0,\infty)} \left(e^{\omega X} - 1 - \omega X \mathbb{1}_{\{X<1\}}\right) \Pi(dX) \tag{2.8}$$

where $\alpha \geq 0$ is the constant drift, $\sigma > 0$ is the gaussian diffusion and Π is the Lévy measure.

The fundamental equation of ruin theory, known as the integro-differential equation, describes the evolution of the ruin probability in time:

$$\psi(u) = 1 - e^{-\lambda\psi(u)} \int_0^u \psi(u - X)F(X)dX \tag{2.9}$$

where $F(X)$ is the distribution function of the claim sizes, assumed to be non-negative and independent and identically distributed.

The conditions of the model are as follow:

- The claim arrival process is a Poisson process with rate λ.
- The claim sizes X_i are independent and identically distributed random variables with common distribution function F and finite mean μ.
- The premium income process is constant and equal to the expected value of the total claim amount per unit time, that is, $c = \lambda\mu$.
- The initial capital of the insurance company is u.

The Cramér-Lundberg model is a special case of the more general class of spectrally negative Lévy models. The model can be extended to include various refinements, such as stochastic interest rates, different claim sizes and frequencies, and time-dependent rates of incoming premiums and outgoing claims.

Sparre Andersen model

The Sparre Andersen model, also known as the collective risk model, is a mathematical model used in ruin theory to describe the behavior of a collective of insurance policies. In this model, each policy is assumed to have a premium rate λ and a claim size distribution F, with these parameters being independent across policies.

Let $N(t)$ be the number of policies in force at time t, and let X_i denote the claim size of the ith policy. The aggregate claims process $S(t)$ is defined as the total amount of claims up to time t, given by:

$$S(t) = \sum_{i=1}^{N_t} X_i \tag{2.10}$$

The key assumption of the Sparre Andersen model is that the number of policies in force follows a Poisson process with rate α, independent of the claim sizes and the number of policies in force at any previous time. That is, the probability of observing k new policies in the time interval $(t, t + \Delta t)$ is given by:

$$P\left(N(t + \Delta t) - N(t) = k\right) = \frac{e^{(-\alpha \Delta t)}(\alpha \Delta t)^k}{k!} \tag{2.11}$$

for $k = 0, 1, 2, \ldots, n$. Under this assumption, the process $N(t)$ is a homogeneous Poisson process with rate α.

The key quantity of interest in the Sparre Andersen model is the probability of ruin, defined as the probability that the aggregate claims process $S(t)$ exceeds the total premium income $P(t)$ at some time before or at time t:

$$\Psi(t) = P\left(S(t) > P(t)\right). \tag{2.12}$$

The conditions for the Sparre Andersen model are as follows:

- The interarrival times between claims form a Poisson process with parameter λ.
- The claim sizes are independent and identically distributed according to an exponential distribution with parameter θ.
- The arrival times and the claim sizes are independent.
- The insurer can invest in a risk-less asset with interest rate r.
- The premium rate c is constant and covers expected claims and expenses, but does not allow for profit.

Assuming that the conditions hold, then the surplus process $U(t)$ satisfies the stochastic differential equation:

$$dU(t) = \left[c - \lambda - \theta U(t)\right] dt + \sqrt{U(t)} dB(t) \tag{2.13}$$

where $B(t)$ is a standard Brownian motion.

2.3 Market Risk theory

2.3.1 Introduction

Market risk theory provides a framework for understanding the impact of market risks on asset pricing. Asset pricing is concerned with determining the fair price of financial assets such as stocks, bonds, and options. The price of an asset is influenced by various factors such as interest rates, market volatility, and market risk.

One of the key insights of market risk theory is the importance of risk measures in asset pricing. Risk measures provide a way to quantify the risk associated with a particular financial asset. In the context of Lévy processes, risk measures can be computed using the characteristics of the process. For example, the VaR of a Lévy process can be expressed in terms of its Lévy measure, which describes the distribution of jumps. The Lévy measure can be computed from the characteristic exponent of the process, which is a function of the drift, volatility, and jump characteristics of the process.

Thus, it should now be quite obvious that while the idea of managing market risk is not extremely convoluted by nature, it is however quite technical especially with its great dependency to the various risk measures. In the following subsection, let us dedicate some time to further study and explore these measures.

2.3.2 Risk measures

Value-at-Risk (VaR)

Value-at-Risk (VaR) is a widely used risk measure that estimates the maximum loss that a portfolio may experience over a given time horizon at a certain confidence level. It is a quantile-based risk measure and is defined as the negative of the $(1-\alpha)$ percentile of the loss distribution, where α is the confidence level. More specifically, let L_t denote the loss of a portfolio at time t and $F_{L_t}(l)$ denote the cumulative distribution function (CDF) of L_t. The VaR at a confidence level α is given by:

$$VaR_\alpha(L_t) = -\inf\{l \in \mathbb{R} : F_{L_t} \geq \alpha\}. \tag{2.14}$$

This can also be expressed as a function of the loss distribution L_t and the confidence level α:

$$VaR_\alpha(L_t) = -F^{-1}_{L_t}(\alpha) \tag{2.15}$$

where $F^{-1}_{L_t}$ is the inverse CDF or quantile function of L_t.

In the context of Lévy processes, VaR is often used to measure the risk of financial assets whose returns are modeled using Lévy processes. For example, if we assume that the returns of an asset follow a Lévy process X_t, then the loss at a future time horizon T can be expressed as:

$$L_T = S_0 \left(e^{X_T} - 1\right) \tag{2.16}$$

where S_0 is the initial price of the asset. The VaR of L_T at a confidence level α can then be computed using the formula above.

There are various methods for estimating VaR for Lévy processes, including Monte Carlo simulation, historical simulation, and variance-covariance method. These methods involve simulating the possible future paths of the Lévy process and computing the corresponding losses at each time step, and then using the resulting loss distribution to estimate VaR. The choice of method may depend on the specific characteristics of the Lévy process and the availability of data.

It is worth noting that VaR has some limitations as a risk measure. For instance, it does not account for the magnitude of potential losses beyond the α-quantile or the tail risk associated with extreme events. In addition, VaR assumes that the distribution of the portfolio value is stationary and does not change over time, which may not be a valid assumption in practice.

Conditional Value-at-Risk (CVaR)

Conditional Value-at-Risk (CVaR), also known as Expected Shortfall (ES), is a risk measure that extends the concept of Value-at-Risk (VaR) by taking into account the expected loss beyond the VaR level. It measures the expected loss given that the loss exceeds the VaR level.

Let X be a random variable representing the loss of a financial position over a specified period of time. The VaR at level α is defined as the α-quantile of the distribution of X as equation 2.14 shows. Then the CVaR at level α is defined as the conditional expectation of the loss X given that the loss exceeds the VaR level or more formally:

$$CVaR_a = \mathbb{E}[X|X > VaR_a] = \frac{1}{1-\alpha}\int_{\alpha}^{1} VaR_u du \tag{2.17}$$

In the context of Lévy processes, the VaR and CVaR can be computed using the characteristic exponent Ψ of the process. Specifically, for a given confidence level α, the VaR and CVaR can be obtained by solving the equations:

$$\mathbb{E}\left[\exp[-sX]\right] = \exp\left[-\Psi(s)VaR_\alpha\right] \tag{2.18}$$

$$CVaR_\alpha = \frac{1}{1-a}\int_{\alpha}^{1} VaR_u du \tag{2.19}$$

respectively, where $\Psi(s)$ is the characteristic exponent of the Lévy process.

Tail Value-at-Risk (TVaR)

Tail Value-at-Risk (TVaR) is a risk measure that quantifies the potential loss beyond a certain tail threshold. It is similar to Value-at-Risk (VaR), but unlike VaR, which only quantifies the maximum loss at a certain confidence level, TVaR considers the entire loss distribution beyond the threshold.

Formally, let X be a random variable representing the loss distribution of a portfolio, with a cumulative distribution function $F_X(x)$. Let α be the tail threshold, such that $0 < \alpha < 1$. Then, TVaR at the confidence level α is defined as the expected value of the loss distribution beyond the α-quantile:

$$TVaR_\alpha(X) = \frac{1}{1-\alpha}\int_\alpha^1 VaR_u(X)du \tag{2.20}$$

where $VaR_u(X)$ is the Value-at-Risk at the confidence level u.

Intuitively, TVaR can be interpreted as the expected loss beyond the q-quantile. It is also known as Conditional Value-at-Risk (CVaR) or Expected Shortfall (ES) when q is taken to be very small, approaching zero.

Entropic Value-at-Risk (EVaR)

Entropic Value-at-Risk (EVaR) is a risk measure that combines the concepts of entropy and Value-at-Risk (VaR) to provide a more robust measure of risk. Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let X be a random variable with distribution function F_X. Let $\alpha \in (0,1)$ be a confidence level, and let $h : \mathbb{R} \to \mathbb{R}$ be a convex function. Then the entropic Value-at-Risk (EVaR) of X at level α with respect to h is defined as:

$$EVaR^h_\alpha(X) = -\frac{1}{h(\alpha)}\ln\left(\int_{\{x:F_X(x)\leq\alpha\}} e^{-h(X-x)}dF_X(x)\right) \tag{2.21}$$

where $-\frac{1}{h(\alpha)}\ln$ denotes the entropic transformation.

Intuitively, the entropic Value-at-Risk measure quantifies the expected shortfall of X at a given confidence level α, while taking into account the shape of the distribution through the convex function h. The convexity of h implies that EVaR is more sensitive to large losses than to large gains, which makes it a useful risk measure in many financial applications.

Coherent Risk Measures (CRM)

Coherent Risk Measures (CRM) are a class of risk measures that satisfy a set of desirable properties that reflect their coherence with the axioms of expected utility theory. A coherent risk measure is a function $\rho : \mathcal{L}^1 \to \mathbb{R}$, where $\mathcal{L}^1$ is the set of all integrable random variables, that satisfies the following four axioms:

- Monotonicity: If $X \leq Y$ almost surely, then $\rho(X) \geq \rho(Y)$.

- Translation invariance: $\rho(X + c) = \rho(X) - c$ for all $c \in \mathbb{R}$.
- Positive homogeneity: $\rho(\lambda X) = \lambda\rho(X)$ for all $\lambda > 0$.
- Convexity: $\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y)$, $\forall\lambda \in [0, 1]$.

The first three axioms reflect the intuitive idea that a coherent risk measure should be insensitive to changes in the level and location of the random variable being measured, while the convexity axiom reflects the idea that the measure should be consistent with the concave preferences of an expected utility maximizer.

Coherent risk measures can be used to evaluate the risk of a portfolio of assets or a financial derivative by measuring the expected shortfall or expected utility of the portfolio or derivative with respect to a certain probability distribution. In the context of Lévy processes, coherent risk measures can be constructed using Lévy measures and Lévy-Khintchine representations, similar to the construction of other risk measures discussed earlier.

Some commonly used coherent risk measures include the Expected Shortfall, Value-at-Risk, and Tail Value-at-Risk measures which we discussed prior to this segment. These measures have the advantage of satisfying the coherence axioms and being mathematically tractable, making them popular in finance and risk management.

Spectral Risk Measures (SRM)

Spectral Risk Measures (SRMs) are a family of risk measures that utilize spectral theory to model the tail behavior of a probability distribution. They were first introduced by Acerbi and Tasche (2002) and have since become a popular choice for risk management in various industries.

For example let X be a real-valued random variable representing the loss incurred by an investment. The SRM of order α is defined as:

$$p_\alpha(X) = \inf\left\{x \in \mathbb{R} : \frac{1}{1-\alpha}\mathbb{E}\left[X - x\right]_+^{1-\alpha} \leq 1\right\} \tag{2.22}$$

where $(X - x)_+ = \max(X - x, 0)$ and $\alpha \in (0, 1)$. The SRM can also be expressed in terms of the distribution function of X as:

$$p_\alpha(X) = F^{-1}(1 - \alpha) \tag{2.23}$$

where $F(x)$ is the cumulative distribution function of X. The SRM can be interpreted as the threshold value x such that the probability of the loss exceeding x is no more than α, or more formally:

$$\mathbb{P}\left(X \geq p_\alpha(X)\right) \leq \alpha \tag{2.24}$$

The SRM is closely related to the spectral distribution function of X, which is

defined as:

$$S_X(z) = \mathbb{E}\left[\frac{1}{z - X}\right] \tag{2.25}$$

where $z \in \mathbb{C}$ is a complex number such that $\mathbb{E}[|X|^2] < |z|^2$. The spectral distribution function can be used to study the tail behavior of X and can also be used to calculate the SRM.

Risk Contribution Measures (RCM)

Risk Contribution Measures (RCM) are a class of risk measures that are used to quantify the contribution of individual assets or groups of assets to the overall risk of a portfolio. RCM can be used to identify the most significant risk drivers of a portfolio and to determine the optimal allocation of capital among these drivers.

In the context of Lévy processes, RCM can be defined in terms of the Lévy measure of the portfolio, which captures the distribution of the jumps in the portfolio value. Specifically, let $X = (X_1, \ldots, X_n)$ denote the vector of random variables representing the portfolio positions in n assets, and let S denote the portfolio value. Then, the RCM for asset i is defined as the partial derivative of the portfolio value S with respect to the position in asset i:

$$RCM_i = \frac{\partial S}{\partial X_i} \tag{2.26}$$

In other words, RCM measures the sensitivity of the portfolio value to changes in the position of a particular asset. If RCM_i is large, then small changes in the position of asset i can have a significant impact on the portfolio value, and asset i is considered to be a significant risk driver. The RCM can also be expressed in terms of the Lévy measure of the portfolio. Specifically, let Π_i denote the Lévy measure of the jump component of asset i, and let Π denote the Lévy measure of the portfolio. Then, the RCM for asset i can be written as:

$$RCM_i = \frac{1}{\Pi(S)} \int_{\mathbb{R}\setminus\{0\}} x_i \Pi_i(dx) = \frac{\Pi_i(S)}{\Pi(S)} \tag{2.27}$$

where $\Pi_i(S)$ is the jump component of asset i in the portfolio.

RCM can be used to construct optimal portfolios that balance risk contributions among the assets. One approach is to minimize the variance of the RCM, subject to a budget constraint on the total risk of the portfolio. This leads to a portfolio that minimizes the contribution of each asset to the total risk, while still maintaining the desired level of risk for the portfolio as a whole.

Tail Conditional Expectation (TCE)

Tail Conditional Expectation (TCE) is a risk measure that seeks to estimate the expected loss in the tail of the distribution beyond a certain threshold. TCE

can be seen as a generalization of the Value-at-Risk (VaR) measure, where instead of estimating the maximum loss at a certain confidence level, TCE estimates the average loss beyond the VaR threshold.

Let X be a random variable representing the returns of a financial asset. The TCE risk measure at a confidence level $\alpha \in (0, 1)$ is defined as:

$$TCE_\alpha(X) = \mathbb{E}\left[X | X \leq VaR_\alpha(X)\right] \tag{2.28}$$

where $VaR_\alpha(X)$ is the VaR at the confidence level α, defined as the α-quantile of the distribution of X. In other words, $VaR_\alpha(X)$ is the value such that the probability of X falling below it is α.

The TCE risk measure can also be expressed as the integral of the tail of the distribution of X beyond the VaR threshold:

$$TCE_\alpha(X) = \frac{1}{1-\alpha} \int_{-\infty}^{VaR_\alpha(X)} \left(x - VaR_\alpha(X)\right) f_X(x) dx \tag{2.29}$$

where $f_X(x)$ is the probability density function of X.

Expected Maximum Drawdown (MDD)

The Maximum Drawdown (MDD) is a commonly used risk measure in finance that captures the worst-case scenario for an investment or portfolio. It is defined as the maximum percentage loss from the highest point of the investment to the lowest point before a new high is reached. MDD is a non-linear risk measure, which means that it is not proportional to the size of the investment. For example, a 50% loss requires a 100% gain to recover the original value, while a 100% loss requires an infinite gain to recover.

The Expected Maximum Drawdown (MDD) is an extension of MDD that provides a probabilistic view of the maximum loss. Specifically, it is defined as the expected value of the maximum drawdown over a given time horizon, where the expectation is taken under a probability measure. Let X_t be a Lévy process representing the value of an investment at time t. We assume that $X_0 = 1$ without loss of generality. The maximum drawdown process for X_t over a time interval $[0, T]$ is defined as:

$$E[MDD] = \int_\Omega \max_{t \in [0,T]} \left(V(t,T) - \sup_{s \in [0,t]} V(s,T) \right) d\mathbb{P} \tag{2.30}$$

where $V(t, T)$ represents the value of the portfolio at time t and $\sup_{s\in[0,t]} V(s, T)$ represents the maximum value of the portfolio over the time interval $[0, t]$. The term $\max_{t\in[0,T]}(V(t, T) - \sup_{s\in[0,t]} V(s, T))$ represents the maximum drawdown of the portfolio over the time interval $[0, T]$. $\mathbb{P}$ is the probability measure over the underlying asset's returns, and Ω is the sample space.

The expected MDD is a useful risk measure because it takes into account not

only the magnitude of potential losses but also their probability of occurrence. It can be used to compare the risk of different investment strategies or portfolios, and it provides a more complete picture of the downside risk of an investment than traditional measures such as volatility or Value-at-Risk.

Sharpe Ratio

The Sharpe Ratio is a popular risk-adjusted performance measure in finance that helps investors evaluate the return of an investment relative to its risk. It is named after Sharpe, who developed it in 1994 (Sharpe, 1994).

Suppose we have a risky asset with a continuously compounded return process given by a Lévy process X_t with initial value $X_0 = 1$. We assume that the expected return of X_t is given by the drift rate μ and its volatility is given by the standard deviation σ.

The Sharpe Ratio is defined as the excess return of the risky asset over the risk-free rate (usually the rate of a Treasury bill) divided by the standard deviation of the excess return. Mathematically, we can express this as:

$$SR = \frac{(\mathbb{E}[X_t] - r_f)}{\sigma[X_t - r_f]} \tag{2.31}$$

where $\mathbb{E}[X_t]$ is the expected return of X_t, r_f is the risk-free rate, and $\sigma[X_t - r_f]$ is the standard deviation of the excess return. We can rewrite the above expression in terms of the drift rate μ and volatility σ of X_t as:

$$SR = \frac{\mu - r_f}{\sigma}. \tag{2.32}$$

The Sharpe Ratio measures the additional return per unit of risk taken on by the investor. A higher Sharpe Ratio indicates a better risk-adjusted performance of the investment. However, it is important to note that the Sharpe Ratio assumes that the returns of the risky asset are normally distributed, which may not always be the case in practice.

Sortino Ratio

The Sortino Ratio is a risk-adjusted performance measure that evaluates an investment's return relative to its downside risk. It was introduced as an improvement over the Sharpe Ratio, which measures the excess return of an investment over the risk-free rate relative to its total risk. However, the Sharpe Ratio assumes a symmetric distribution of returns, which is not always the case in practice. The Sortino Ratio addresses this by considering only the downside risk.

Let's consider a financial asset whose returns are modeled by a Lévy process with a probability distribution function F. Denote the random variable representing the asset's returns at time t by R_t. Assume that the asset's expected return and volatility are represented by the first two moments of R_t, denoted by μ and σ, respectively.

The downside risk can be measured using the semi-deviation, which is the standard deviation of the negative returns:

$$Semi - Deviation = \sqrt{\frac{\sum_{i=1}^{n}(R_i - MAR)^2}{n}} \tag{2.33}$$

where R_i is the return for period i, MAR is the minimum acceptable return, and n is the number of periods. Note that the semi-deviation only considers the negative part of the returns, capturing the downside risk of the investment. The Sortino Ratio is then defined as:

$$SortinoRatio = \frac{R_t - R_f}{Semi - Deviation} \tag{2.34}$$

where R_f is the risk-free rate. The Sortino Ratio measures the excess return of an investment over the risk-free rate relative to its downside risk. When applied alongside with Lévy processes or Lévy-driven models, the Sortino Ratio can be used to compare the risk-adjusted performance of different assets whose returns are modeled by Lévy processes with different probability distributions. For example, it can be used to compare the risk-adjusted performance of an asset whose returns follow a stable distribution with another asset whose returns follow a normal distribution.

Omega Ratio

The Omega Ratio is a risk measure that compares the expected returns of an investment to its potential downside. It was introduced by Keating and Shadwick (Keating and Shadwick, 2002) and is defined as the ratio of the probability-weighted average return above a specified threshold to the probability-weighted average loss below the same threshold.

Let X_t be a stochastic process representing the returns of an investment at time t. We assume that X_t is a Lévy process with characteristic exponent $\Psi(\theta)$, where θ is the jump size. We also assume that the process is centered, i.e., $\mathbb{E}[X_t] = 0$ for all t. Thus, the Omega ration is defined as:

$$OmegaRatio(K) = \frac{\mathbb{E}[X_t \mathbb{I}_{X_t \geq K}]}{\mathbb{P}[X_t < K]} \tag{2.35}$$

where K is a threshold value, and $\mathbb{I}_{X_t \geq K}$ is the indicator function that takes the value 1 if $X_t \geq K$ and 0 otherwise.

Intuitively, the numerator of the Omega Ratio measures the expected returns of the investment above the threshold K, while the denominator measures the probability of the investment generating a loss below the same threshold. The higher the Omega Ratio, the better the risk-return trade-off of the investment.

2.4 Computational techniques

2.4.1 PIDE methods

Introduction

The use of partial integro-differential equations (PIDEs) in the pricing and hedging of financial derivatives has a long and rich history, particularly in the context of modeling assets with Lévy processes. The development of PIDE methods for pricing financial derivatives can be traced back to the early 1990s, when the seminal paper by Heston introduced the famous Heston model, which is a stochastic volatility model with jumps (Heston, 1993).

The Heston model and other similar models with jumps require the solution of PIDEs, which can be challenging due to the presence of the integral terms. In the following years, several numerical techniques were developed to solve PIDEs arising in the context of Lévy process models, including finite difference methods (FDMs), finite element methods (FEMs), Monte Carlo simulation methods, and transform methods (Carr and Madan, 1999; Broadie and Detemple, 2004).

One of the main challenges of using PIDEs to price financial derivatives is the presence of discontinuities and singularities in the underlying Lévy process. To overcome this challenge, several authors have proposed various methods in order to combat the complexity of achieving accurate simulations (Feng and Linetsky, 2008).

In recent years, there has been a growing interest in using PIDEs to price and hedge derivatives with more complex payoff structures, such as American and Asian options, convertible bonds, and credit derivatives. Researchers have also developed new PIDE-based methods that use Fourier-cosine series expansions to solve Lévy process models with discontinuous payoffs, as proposed by Fang and Oosterlee (Fang and Oosterlee, 2009). These methods have shown promising results in terms of accuracy and efficiency, and have the potential to be applied to a wider range of problems in quantitative finance.

Toy example

Consider a European call option with strike price K, maturity T, and underlying asset price S_t that follows a geometric Lévy process (GLP) defined by the SDE:

$$dS_t = S_t^- dX_t \tag{2.36}$$

where X_t is a GLP with Lévy triplet (α, σ, Π). The option pricing problem can be formulated as the following PIDE:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{2.37}$$

for $0 < t < T$ and $0 < S < \infty$. This PIDE is subject to the condition:

$$V(T, S) = (S - K)^+ \tag{2.38}$$

where r is the risk-free interest rate.

We can solve this PIDE using the Crank-Nicolson scheme, which is a numerical method for solving partial differential equations. The Crank-Nicolson scheme discretizes both time and space and approximates the PIDE by a system of linear equations that can be solved numerically.

Let Δt and ΔS be the time and space step sizes, respectively, and let $V_{i,j}$ denote the option value at time $t_i = i\Delta t$ and asset price $S_j = j\Delta S$, where $0 \leq i \leq N$ and $0 \leq j \leq J$. Then, the Crank-Nicolson scheme can be written as:

$$\frac{1}{2}AV_{i+1} + \frac{1}{2}BV_i = C \tag{2.39}$$

where A, B, and C are matrices that depend on the time and space step sizes, the option payoff function, and the coefficients of the PIDE. The matrices A and B are tridiagonal, which means that they have non-zero entries only on the main diagonal and the two diagonals adjacent to it.

The matrices A, B and C are the following:

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\frac{1}{2}r\Delta t & 1 & 0 & \cdots & 0 \\ 0 & -\frac{1}{2}r\Delta t & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{2.40}$$

$$B = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \frac{1}{2}r\Delta t & 1 & 0 & \cdots & 0 \\ 0 & \frac{1}{2}r\Delta t & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{2.41}$$

$$C = \begin{bmatrix} 1 + \frac{\theta_1 \Delta t}{2} & -\frac{\theta_1 \Delta t}{2} & 0 & \cdots & 0 \\ -\frac{\theta_1 \Delta t}{2} & 1 + \theta_1 \Delta t & -\frac{\theta_1 \Delta t}{2} & \cdots & 0 \\ 0 & -\frac{\theta_1 \Delta t}{2} & 1 + \theta_1 \Delta t & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 + \theta_1 \Delta t \end{bmatrix} \tag{2.42}$$

We can now approximate the solution to the PIDE using the Crank-Nicolson scheme,

which is a finite difference method for solving parabolic equations. The scheme is given by:

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} = \frac{1}{2}\left(\frac{\partial^2}{\partial x^2}u_i^{n+1} + \frac{\partial}{\partial x^2}u_i^n\right) + f_i^{n+\frac{1}{2}} \tag{2.43}$$

where u_i^n is the numerical solution at time $t_n = n\Delta t$ and $x_i = i\Delta x$. We use the following boundary conditions:

$$u(t, x_0) = K \tag{2.44}$$
$$u(t, x_M) = 0 \tag{2.45}$$

where K is the strike price and M is the number of discretization points.

To solve the PIDE, we first discretize the domain of x and the time interval $[0, T]$. Then, we set up the Crank-Nicolson scheme for each time step and solve the resulting system of linear equations using matrix inversion techniques.

Finally, we obtain the numerical solution u_i^n for all n and i. The option price at time $t = 0$ is given by u_i^0, where i corresponds to the discretized value of the underlying asset's initial price.

2.4.2 Monte Carlo methods

Introduction

Monte Carlo methods are a class of computational techniques used to solve complex problems by simulating random processes. The method was first proposed by Nicholas Metroplolis and Stanislaw Ulam (Metropolis and Ulam, 1949). Monte Carlo methods have been applied to simulate and price financial derivatives, and they have proven to be particularly useful in modeling assets driven by Lévy processes.

The first Monte Carlo driven simulations of financial assets driven by Lévy processes were performed by Carr and Madan (Carr and Madan, 1999). In their paper, they proposed a numerical method for pricing options based on the Fourier transform of the characteristic function of the underlying Lévy process. In the following years, several Monte Carlo methods were developed for simulating Lévy processes like the jump-diffusion model introduced by Merton (Merton, 1976), or the variance gamma model proposed by Madan and Seneta (Madan and Seneta, 1990). These models have been widely used in the finance industry for risk management, portfolio optimization, and option pricing.

More recently, advances in computing power and numerical methods have allowed for the development of more sophisticated Monte Carlo methods for simulating Lévy processes. For example, the Monte Carlo simulation of the stochastic volatility model with Lévy jumps proposed by Barndorff-Nielsen and Shephard (Barndorff-Nielsen and Shephard, 2002) has become a popular approach for pricing

exotic options. Similarly, the Monte Carlo simulation of the Heston model with stochastic interest rates and Lévy jumps introduced by Fang and Oosterlee (Fang and Oosterlee, 2009) has become a standard tool.

Toy example

Let us also include a little toy example for the Monte Carlo methods. We shall present it in a step-by-step fashion:

- First, we need to define the Lévy process that will drive the price of the asset. For this example, we will use a geometric Lévy process with the following characteristics:
  - Initial price: $S_0 = 100$
  - Drift: $\mu = 0.05$
  - Volatility: $\sigma = 0.2$
  - Jump intensity: $\lambda = 0.1$
  - Jump size distribution: stable with $\alpha = 1.5$ and $\beta = 0$
- We need to set up the time horizon T and the number of time steps N that we want to simulate. For this example, we will set $T = 1$ and $N = 252$, which corresponds to a daily time step.
- We need to calculate the time step Δt and the drift and volatility terms that will be used in the simulation. For this example, we have:
  - Time step: $\Delta t = \frac{T}{N} = \frac{1}{252}$
  - Drift term: $r = e^{\mu \Delta t} - 1 = e^{0.05/252} - 1 \approx 0.000198$
  - Volatility term: $v = \sigma\sqrt{\Delta t} = 0.02$
- We will simulate the Lévy process using a Monte Carlo method. For each time step, we will generate a random number Z_i from the jump size distribution and add it to the process. The process can be simulated using the following formula:

$$S_{i+1} = S_i \left(1 + r\Delta t + v\sqrt{\Delta t} Z_i + \sum_{j=1}^{N_i} Y_{i,j}\right) \tag{2.46}$$

  where $Y_{i,j}$ are independent and identically distributed random variables from the jump size distribution, and N_i is a Poisson-distributed random variable with mean $\lambda \Delta t$.
- We repeat the simulation process for numerous iterations to obtain a large number of possible paths of the Lévy process.
- For each path, we need to calculate the payoff of the option at maturity. For this example, we will consider a European call option with strike price $K = 110$

and maturity $T = 1$. The payoff of the option is given by:

$$\max(S_T - K, 0) \tag{2.47}$$

- We need to discount the payoff of the option back to the present time using the risk-free interest rate. For this example, we will use an interest rate of $r_f = 0.03$. The discounted payoff is given by:

$$\frac{1}{1 + r_f} \max(S_T - K, 0) \tag{2.48}$$

- Finally, we estimate the price of the option by taking the average of the discounted payoffs across all the simulated paths.

This is just a representative recount of the steps that a practitioner or researcher should take in order to use Monte Carlo based methods in a similar setting.

2.4.3 Bayesian methods

Introduction

Bayesian methods have gained a lot of popularity in finance and were applied to various problems, such as portfolio optimization and option pricing. In the context of time series and mathematical finance, one application of Bayesian methods could be considered the work of Chan and Eisenstat who used a Bayesian approach to estimate the parameters of a time series model (Chan and Eisenstat, 2018).

Bayesian methods have been widely used in finance and related fields, and have been applied to a variety of Lévy process models. For example, Jacquier, Polson and Rossi used Bayesian methods to estimate the parameters of Lévy-driven stochastic volatility models which where additionally, characterized by fat tails and correlated errors (Jacquier et al., 2004).

More recently, Bayesian methods have been used to estimate the parameters of more complex Lévy process models, such as the rough volatility model introduced by Gatheral, Jaisson and Rosenbaum (Gatheral et al., 2014).

Toy example

Similarly, let us also assume that we wish to price an asset that is driven by a Lévy process. For the sake of this example, let us assume that we are using a jump-diffusion process with parameters $\theta = (\sigma, \mu, \gamma, \lambda)$, where σ is the diffusion coefficient, μ is the drift coefficient, γ is the jump size distribution, and λ is the jump frequency.

Naturally, when working in a bayesian framework, we shall need to choose a prior distribution for our parameters. Note that the choice of a prior distribution

is quite important and it has quite the impact on the overall performance of our attempts to price that asset. For this example, let us choose informative priors, say, a normal distribution for σ and μ, an inverse gamma distribution for γ, and a gamma distribution for λ.

Again, within the classic Bayesian framework, the next step is to define our likelihood function. Of course this needs some observations, yet once again, let us assume that we have been monitoring the asset price over time. Thus, we are lead to:

$$\log L(\theta|data) = \sum_{i=1}^{n} \log p(x_i|x_{i-1}, \Delta t, \theta) + \sum_{j=1}^{m} \log p(y_j|\theta) \tag{2.49}$$

where x is the asset price, y is the jump size, Δt is the time step, and p is the probability density function of the jump-diffusion process.

Now we have everything that we require, in order to apply the very well known Baye's formula:

$$p(\theta|data) \propto L(\theta|data)p(\theta). \tag{2.50}$$

Should one proceed with the calculations here, they shall end up with the proposed posterior density. We are interested to sample from this density, so that we may find out the posterior's parameters. To do so, we need to employ an MCMC algorithm capable of doing so. There are many options in terms of MCMC algorithms and our priorities such as accuracy or rate of convergence, should influence the choice of the algorithm. It is also possible to compute other valuable statistics or credible intervals.

At this point, we can evaluate the goodness of fit of the model by comparing the simulated data to the observed data and performing diagnostic tests on the residuals. If the practitioner or researcher is content with the fit, they may proceed by using their posterior along with its calibrated parameters to simulate future asset prices and examine the distribution of future returns or compute option prices. In the case where the fit is not good enough, one choice is to repeat the procedure by choosing different priors that should, ideally, reflect better the information at hand.

2.5 Application: Bitcoin price

2.5.1 Problem statement & Motivation

Bitcoin is a highly volatile and complex digital asset that has gained significant attention in recent years as a potential alternative investment. However, the behavior of its price is difficult to model due to its non-linearity and sensitivity to various factors such as global economic events, regulatory changes, and adoption rates. Traditional models such as Brownian motion-based models have limitations in capturing these characteristics, leading to inaccurate predictions and risk management strategies.

In contrast, Lévy processes are a class of stochastic processes that can capture a wider range of market behaviors, including jumps and heavy-tailed distributions, making them well-suited for modeling the bitcoin price. Using Lévy processes to model the bitcoin price might end up providing a more accurate representation of its dynamics.

For this reason, we set up this toy-application in the following manner. We use a Brownian motion, a Geometric Brownian motion and a general Lévy jump-diffusion model to model the evolution of returns over time for Bitcoin. Additionally, we calibrate the parameters for these models by the use of a simple iterative optimizer algorithm that attempts to minimize the squared difference between the actual returns and the simulated returns at time t, which acts as our objective function. This is different to the general PIDE or Bayesian methods that we discussed, but it should be more appropriate for the scope of this thesis.

Note that, this applications and the accompanying simulations are not incorporated up to the current standards at the time of authoring this thesis, for a number of reasons. These reasons include, the unavailability of publicly free, true high-frequency data, as well as the computational complexity involved in using more complex Lévy models and more advanced parameter calibration methods. Thus, the choice for the models and their calibration method is justified within this context.

Furthermore, with different initial parameters for the models, or different bounds for the optimizing algorithm, higher accuracy modeling may be highly feasible. However, the scope of these models is to showcase how increasing the complexity from a simple Brownian motion to a Lévy jump-diffusion model, manages to yield better modeling capabilities. This will be quite obvious to notice via the various plots we shall introduce later.

2.5.2 Simulations

Brownian Motion

Let us start by introducing the plot for the case of the Brownian motion.

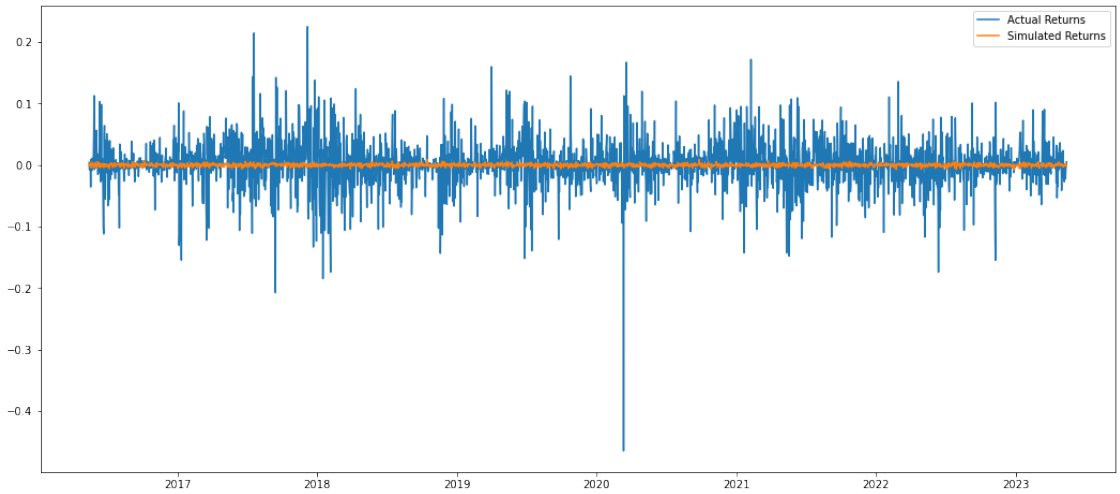

Figure 2.1: Actual Bitcoin returns versus simulated returns driven by Brownian Motion

What is important in this figure is observing the capability of the Brownian Motion at capturing the changes in returns. Is is trivial to notice that the Brownian Motion is not adapting well at all to the evolution of the returns. Due to the volatility of the asset, it is also trivial to notice the heavy tails and jumps exhibited by the returns. These appear to be even more beyond the reach of our Brownian Motion model, which was expected from our recount of theory.

Thus, to better capture the intricacies shown by the Bitcoin returns, we ought to turn to a more powerful, or alternatively, a more complex Lévy process. Doing so, should yield better results.

Geometric Brownian Motion

As a result of the following, we turn to the Geometric Brownian Motion. As we discussed in earlier parts of the thesis, this can be considered the evolution of the Brownian Motion, or from Bachelier's thesis to the Black-Scholes model (Bachelier, 1900; Black and Scholes, 1973). Let us introduce the appropriate plot.

Upon looking at the plot, it is immediately clear that the Geometric Brownian Motion is indeed much better at modeling the change of returns. We can notice that the volatility of the Bitcoin returns are better expressed and overall, it should be quite clear that we are closer to the mark with this approach. However, an important part is still missing, that is the ability to model the heavy tails of the returns. This is quite important, as we discussed previously, this behavior of financial assets can

Figure 2.2: Actual Bitcoin returns versus simulated returns driven by Geometric Brownian Motion

make or break an investment or a portfolio. Naturally, this leads to our last model and our last simulation for this application, our simple Lévy jump-diffusion model.

Lévy Jump-Diffusion process

Finally, let us present our last model. This should confirm the established theory about Lévy models being more flexible and capable of modeling heavy tails and jumps. Note that in our example here, the jump-diffusion model we use has jump magnitude that is *Laplace* distributed.

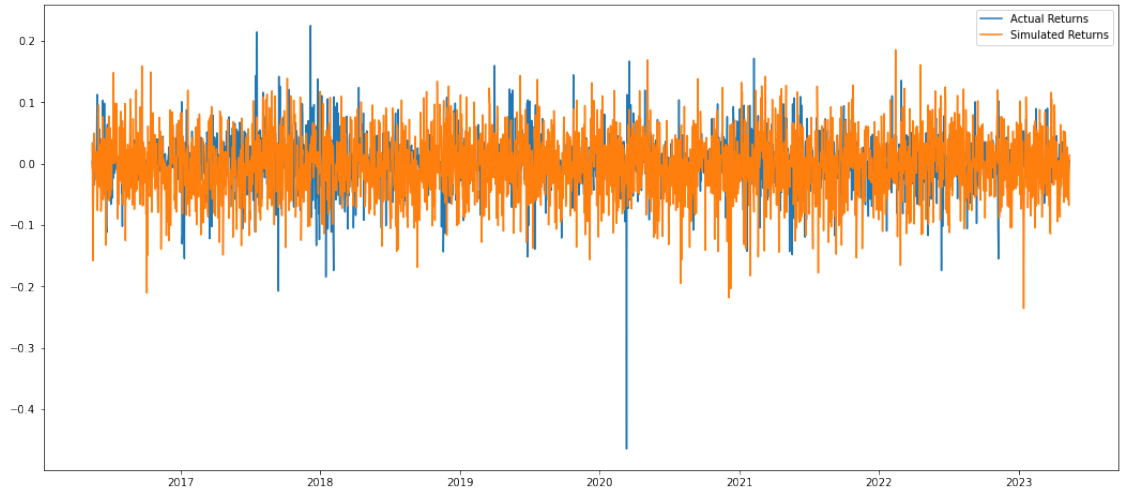

Figure 2.3: Actual Bitcoin returns versus simulated returns driven by a Lévy Jump-Diffusion process

Finally, once again, our conclusions are obvious from taking a look at our plot. It is obvious that the Lévy jump-diffusion process manages to successfully model the

heavy tails of the Bitcoin returns and quite successfully at that. Thus, we finally show that in a more practical way as well, how a Lévy model that incorporates jumps is a much more suitable choice for modeling asset prices than a model that is almost surely continuous in its paths. This, successfully confirms the established theory so far.

2.6 Application: SPLV Index price

2.6.1 Problem statement & Motivation

After demonstrating the application of Lévy processes in modeling high volatility assets such as Bitcoin in the previous application, it is crucial to extend our exploration to understand how these processes can be applied to much lower volatility assets. In this context, we turn our attention to the S&P 500 Low Volatility Index (SPLV), an index composed of the 100 least volatile stocks in the S&P 500.

The motivation behind studying low volatility assets, such as the SPLV, lies in their inherently different risk-return characteristics compared to high volatility assets. Investors typically regard these low volatility stocks as "defensive" in nature due to their resilience during market downturns and lesser price fluctuations. However, this does not mean these assets are entirely immune to abrupt price changes or 'jumps', which traditional Gaussian models fail to capture as already established by literature.

Lévy processes, with their ability to model jumps and heavy tails, should offer an improved framework for modeling such low volatility assets. A better understanding of the return dynamics of these assets would offer enhanced risk management strategies, especially for conservative investors, who primarily focus on low volatility investments. Additionally, it may also provide insights into the behavior of the broader market, as low volatility stocks can sometimes act as a barometer for overall market sentiment.

In this segment, we will investigate the suitability and effectiveness of the previous processes in modeling the returns of the SPLV index. We aim to identify the most appropriate Lévy process for this particular asset. In doing so, we maintain the previous status quo in not just the models, but the parameter estimation method as well. Let us see how the models fare this time.

2.6.2 Simulations

Brownian Motion

Without wasting time, let us immediately introduce the appropriate plot for Brownian motion, as previously.

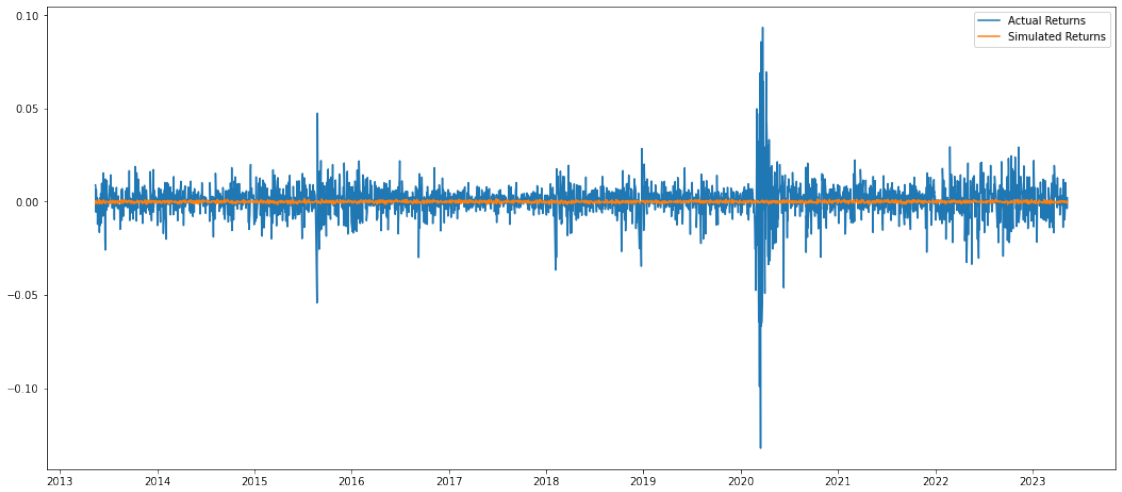

Figure 2.4: Actual SPLV returns versus simulated returns driven by Brownian Motion.

One of the first things easily deducible from the plot is that the overall volatility of the asset is indeed much more stable and lower, when compared to the case of Bitcoin. This in turn, presents an interesting phenomenon, where in the time areas of lower volatility, the Brownian motion presents a much more compelling case.

Overall, it is quite clear at first glance that even in assets with much lower volatility than Bitcoin or similar assets, a strictly Brownian motion model as the one presented is severely lacking. Of course, this completely confirms the existing theory and literature, and was an entirely expected outcome.

Geometric Brownian Motion

Similarly, after our Brownian motion model, we move to the Geometric Brownian motion as in the previous example. Let us introduce the appropriate plot once more.

In the same vein as figure 2.4 we may immediately notice that Geometric Brownian motion similarly appears to be capable of better modeling capabilities when applied to lower volatility assets. However, we can once again notice that in areas where greater volatility is involved, the model is rendered less capable, as it is unable to capture accurately those heavier tails. Of course, needless to say, it fares much better than the very simple Brownian motion model.

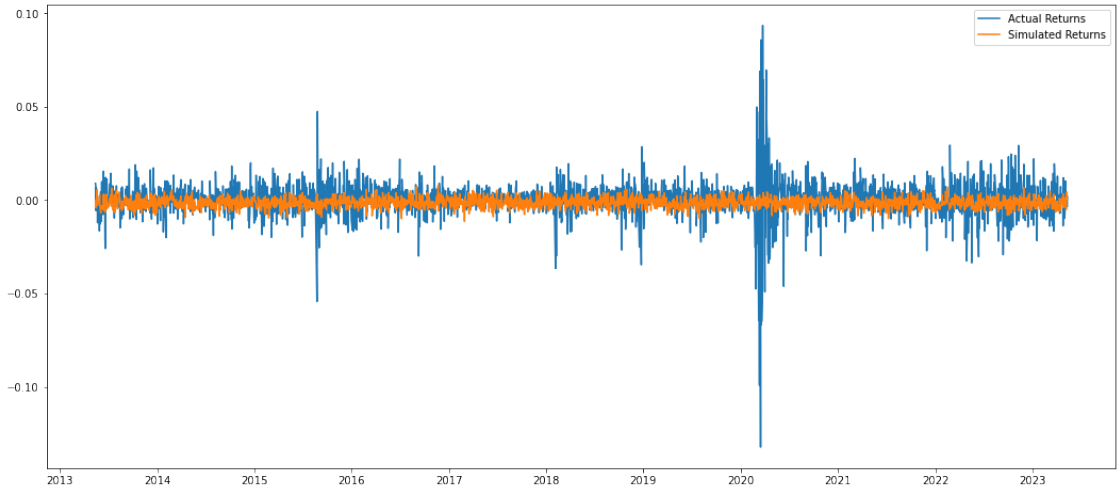

Figure 2.5: Actual SPLV returns versus simulated returns driven by Geometric Brownian Motion.

Lévy Jump-Diffusion process

Lastly, we are to introduce the last plot in the same fashion. We should expect the Lévy Jump-Diffusion model to be an additional step up in modeling ability, similar to the Bitcoin case.

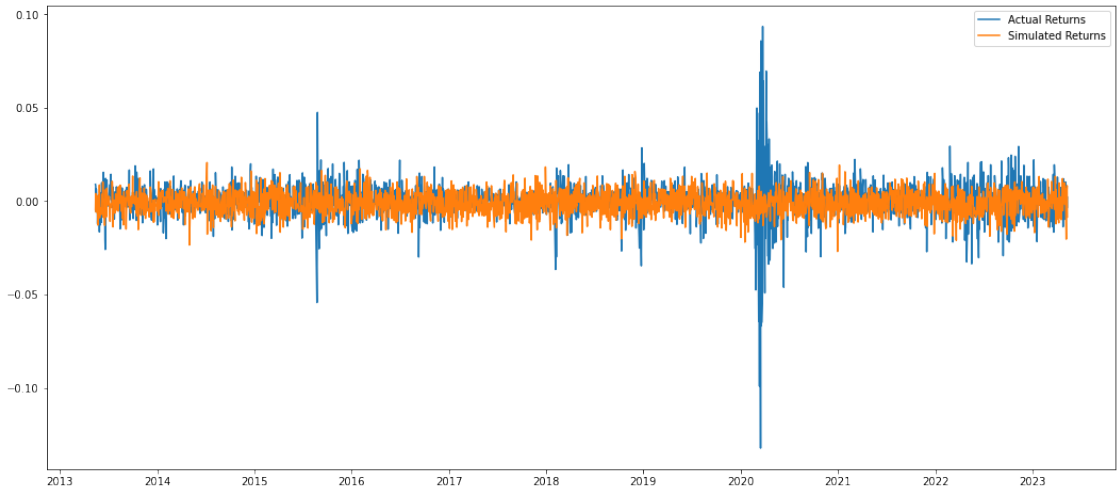

Figure 2.6: Actual SPLV returns versus simulated returns driven by a Lévy Jump-Diffusion process.

Indeed, it is just as we were expecting. The Lévy jump-diffusion process provides another layer of improvement when compared to the more simpler models of Brownian motion and Geometric Brownian motion. It is also clearly shown as well, that time intervals characterized with greater volatility are able to be modeled more accurately with this model. However, it is also important to note that the difference between the three models is lesser in the context of a low volatility asset, than in the context of a high volatility asset.

Chapter 3

Epilogue

This will be the final chapter for this thesis. In this chapter we will briefly discuss the thesis in its entirety. The aim will not only to summarize our work so far, but evaluate our work from the perspective of our initial research topic. Additionally, we also aim to discuss some key notions encompassed in the topic, while also present our conclusions from our work so far as well as the real data applications we included.

In the second part of this chapter, we briefly discuss about the field in terms of advancements, roadblocks and potentially promising avenues for future research. Naturally, this is a personal opinion and thus, subjective.

3.1 Conclusion

3.1.1 Summary

The topic of this thesis was the exploration of the theory of Lévy processes and its connection to Risk theory. At the start of the thesis, a quite comprehensive literature review took place, showcasing the breadth of Lévy processes and their penetration to various fields through a numerous amount of applications. Following this intro, we quickly presented a few necessary prerequisites as well as well as the main conditions of Lévy processes.

An account of the various theorems and contributions that form Lévy's framework, along with the various contributions over time, is given. Great care is taken to ensure that the connections between these items is adequately explored and their interplay, especially, is showcased. This includes items such as the Lévy-Khintchine formula

and its associated decomposition.

The thesis continues by providing an alternative method of decomposition for Lévy processes, exploring the said method both in an on-paper and visual fashion. The next step was the introduction and presentation of several representative Lévy processes via discussing their history, applications and mandatory conditions along with a simulation. This was the content of the first part of the thesis.

In the second part, a brief but adequate account of Risk theory is given where basic notions and definitions are presented and explained along with a literature review. Ruin theory and associated models, as well as, Market Risk theory and associated risk measures are explored both mathematically and from a literature view aspect.

Following that, two applications based on real world data follow that aim to connect the dots between theory and practice but also between the first and second part of this thesis. An account of the applications and its findings are given in the next section, which is dedicated to them.

Overall, this thesis aims to add to the growing research body showcasing the flexibility and powerful capabilities of Lévy processes when applied in the context of Risk theory, risk management and asset pricing, by providing an accurate account of the theory involved, encouraging the use of Lévy processes and presenting valid examples of their advantages when utilized to tackle real world problems.

3.1.2 Experiments and findings

A closing discussion regarding the two applications based on real world data is necessary. While a great emphasis is given to results, the various limitations and caveats should be discussed additionally.

In both examples, the data used are not high-frequency. High and ultra-frequency data are not freely and publicly available. At the time of authoring this thesis, this was the case. Thus, lower frequency data had to be used. Additionally, the high computational time and high computational resources, which characterize the vast majority of parameter estimation and simulation techniques were found to be prohibitive for this thesis. Thus, somewhat simpler methods were used. In both these cases, these caveats and limitations do, indeed, negatively impact the accuracy of the simulations somewhat.

As results are important, a brief summary follows. When attempting to price higher volatility assets with Lévy processes, models capable of jumps or heavy tail modeling have the upper hand. This was proven true with all models used in this scenario with the more complex one providing a layer of improvement over the more simple model.

When attempting to price lower volatility assets with Lévy processes a similar

result was reached. Models capable of jumps or heavy tails modeling were superior in matching the actual returns of the asset over the more simple models. The difference in the case of the lower volatility asset, was that the layer of improvement presented by the more complex models was quantitatively lesser when compared to the case of the high volatility asset.

The findings from the two applications using real world data manage to confirm and reinforce the established theory, as of the time authoring this thesis.

3.2 Avenues for future research

Lastly, it is imperative to include a section covering some avenues for future research regarding our research topic. Each of the following subsections represent one of our suggestion, where it is briefly explained and discussed. It is important to note that these suggestions are a result of personal opinions, formed during the process of authoring this thesis. Thus, it should be treated as such.

3.2.1 Jump size distribution

In all the processes and models that were represented in this thesis, the jump distribution utilized by a given model remained the same. While switching between different jump size distributions may yield better results in an application, it is important to acknowledge that in the concept of financial markets, the price fluctuations of an asset are influenced by a multitude of financial variables.

This yields the following results. There may exist time periods in the price history of a given asset where it would make much more sense to model the jump sizes under a different law. Researching how to better model these time periods as well as proposing alternative distributions for the jumps sizes in these time periods could be an interesting avenue for future research. For example, what if a model capable of altering the jump size distribution based on some proposed metric existed, where the metric was a quantitative relationship between financial variables of interest.

3.2.2 Regime switching markets

One of the interesting aspects of markets is their regime switching behavior. Essentially, this means that as time passes, the various distributional properties of the data of the markets change. This can be influenced by macroeconomic factors such as inflation. An interesting proposal could be the further research of modeling regime switching markets with Lévy driven models that take advantage of the these macroeconomic trends to yield better modeling results.

3.2.3 Long memory

This suggestion goes hand in hand with the previous one. Since it is established that various economic variables influence the price if a derivative long term then a proposed Lévy process in order to be optimal needs to be capture this phenomenon sufficiently.

It is established that Lévy processes are Markovian, thus having the memoryless property. This is contradicting with the need to have long memory, thus research to appropriately extend more Lévy processes from the Markovian framework is required to achieve the long memory property.

3.2.4 Liquidity risk and traded volume

When a trade happens in the context of a financial derivative, a 'signal' is sent which directly influences the derivative's current price. When the traded volume is low the influence to the price is also low, when the traded volume is high the influence to the price is also high. Trading without altering the current price is essential for any investor.

Lévy models capable of accurately capturing the price fluctuation of a proposed trade would be a powerful tool in order to eliminate the liquidity risk of trading.

3.2.5 Computational complexity

This is one of the most important roadblocks to the adoption of Lévy properties, as trying to simulate and calibrate these processes computationally is not only deeply rooted in advanced numerical methods, but also requires extensive computational time.

Further research and breakthroughs are needed in order to successfully present

a more compelling case for Lévy processes by making them less computationally complex and thus more approachable for practice.

3.2.6 Big data and Microstructure noise

With the improvement of technology, one of the boons is the capability of acquiring the price of an asset at smaller timer intervals. This leads to more accurate models and better predicting capabilities. The downside however, is that in this context there exist microstructure noise which acts as a de-stabilizing factor for the Lévy process.

Research could be carried out more successfully filter this noise from a Lévy process point of view. Thus, leading to overall even more accurate models.

3.2.7 Sparse data

While Big data has overall been an improvement, there exist cases where the data of an asset are sparse. This poses incredible challenges in order to successfully attempt to price it. More research could be done to provide specialized or not, Lévy processes capable of accurately capturing the price fluctuations of an asset in a sparse data context.

3.2.8 Extreme Value Theory and markets

While Lévy processes are a class of models that is quite capable of capturing heavy tails and jumps of the price of assets, it is still lacking in cases where there are extreme phenomena influencing the markets. For example, financial crisis, wars and so on, create an instant negative effect in the markets.

More research could be carried out to successfully capture potential upcoming extreme phenomen and their implications from a market point of view, by appropriately modifying Lévy processes or presenting new and specialized Lévy driven models.

3.2.9 Quantum finance and quantum Lévy processes

This is a newer and emerging field of econophysics, which incorporates quantum mechanics from the field of physics in combating problems under the field of finance.

It appears to be quite promising and may hold a lot of solutions to problems that currently do not have an exact solution.

More research could be done to illustrate and stress test the capabilities of quantum Lévy processes when applied to various asset pricing and general financial modeling problems.

3.2.10 Dependence between assets

It is often the case that it is important to examine and establish the dependency structure between different assets. This is especially important from portfolio risk management perspective. One way to achieve this is through the use of Lévy copulas. However, it still remains an open and important problem to solve.

Further research could be dedicated to extending the Lévy copulas to the Lévy processes, in order to yield more capable models to successfully capture the risk involved of managing different assets for a portfolio.

3.2.11 Pricing jumps and Market friction

While the importance of the jump size distribution was emphasized, further examining the behavior of the price of an underlying derivative around jumps is also of significance. Examining the fluctuation of the price around jumps, can yield to identifying important characteristics such as mean reversion, overshooting, or prolonged volatility.

Examining these price dynamics around jumps from the perspective of Market friction is another suggestion for future research. This can result in correlating the price dynamics with market microstructure effects such as bid-ask spreads, liquidity, and order book dynamics.

Research to incorporate results from both of those perspectives, could lead to Lévy processes and models that are far more capable of accurately capturing an underlying asset's price fluctuations than already established models.

3.2.12 Machine learning and Lévy processes

With the rise of machine learning techniques and the number of their applications equally rising, it may be interesting to further research the efficiency of 'marrying' the two approaches with respect to financial applications.

Doing so, may lead to more efficient models or provide better and more effec-

tive techniques with regards to calibrating an already proposed model to the price fluctuations of an asset.

Appendix A

Python code for simulations

The aim of this appendix is to include the various pieces of code used for the various simulations and experiments that occurred within this thesis. For reproducibility reasons the environment that was used is as follows:

- Microsoft Windows 11 64-bit Operating System,
- Python 3 version 3.9.12,
- NumPy version 1.21.5,
- Matplotlib version 3.5.1,
- PySwarms version 1.3.0,
- '123' seed for generation.

The code used for figure 1.1 is as follows:

```
import numpy as np
import matplotlib.pyplot as plt

# Define parameters
T = 1  # Time horizon
N = 100000 # Number of time steps
dt = T/N # Time step size

# Generate 3 Brownian motion realizations
np.random.seed(123)
W = np.sqrt(dt) * np.cumsum(np.random.randn(N+1, 3), axis=0)

# Compute drift parameter for each realization
gamma = W[-1, :] / T

fig, axs = plt.subplots(3, 1, figsize=(12, 18))

# Loop over realizations and plot each one along with its decomposition
for i in range(3):
```

```
    # Compute jumps
    jumps = np.diff(W[:, i])

    # Compute decomposition components
    cp = np.cumsum(jumps - np.mean(jumps))
    cp = np.concatenate(([0], cp)) - gamma[i] * np.arange(N+1) * dt
    pj = np.cumsum(gamma[i] * np.ones(N+1) * dt) + np.concatenate(([0],
        jumps - np.mean(jumps)))
    bm = W[:, i] - pj - cp

    # Plot realization and decomposition components
    axs[i].plot(np.arange(N+1) * dt, W[:, i], label='Realization')
    axs[i].plot(np.arange(N+1) * dt, cp, label='Comp. Poisson process')
    axs[i].plot(np.arange(N+1) * dt, pj, label='Pure Jump process')
    axs[i].plot(np.arange(N+1) * dt, bm, label='Brownian Motion')
    axs[i].legend()

plt.show()
```

The code used for figure 1.2 is as follows:

```
import numpy as np
import matplotlib.pyplot as plt

# Define the Poisson process intensity function
def lambda_t(t):
    return 2 + np.sin(t)

# Define the Brownian motion with drift component
def bm_drift(t, dt):
    return np.sqrt(dt) * np.random.normal()

# Define the compound Poisson process component
def compound_pp(t, dt):
    num_jumps = np.random.poisson(lambda_t(t) * dt)
    if num_jumps == 0:
        return 0
    else:
        jump_sizes = np.random.normal(size=num_jumps)
        return np.sum(jump_sizes)

# Define the pure jump process component
def pure_jump(t, dt):
    num_jumps = np.random.poisson(lambda_t(t) * dt)
    if num_jumps == 0:
        return 0
    else:
        jump_times = t + np.random.uniform(size=num_jumps) * dt
```

```
        jump_sizes = np.random.normal(size=num_jumps)
        return np.sum(jump_sizes * (jump_times <= t))

# Define the Levy-Ito decomposition function
def levy_ito_decomp(X, dt):
    BM_drift = np.zeros_like(X)
    comp_pp = np.zeros_like(X)
    pure_jump_process = np.zeros_like(X)
    for i in range(1, len(X)):
        t = i * dt
        BM_drift[i] = BM_drift[i-1] + bm_drift(t, dt)
        comp_pp[i] = comp_pp[i-1] + compound_pp(t, dt)
        pure_jump_process[i] = pure_jump_process[i-1] + pure_jump(t, dt)
    return BM_drift, comp_pp, pure_jump_process

# Define the simulation parameters
T = 10
N = 1000
dt = T / N

# Simulate 3 realizations of the Poisson process
np.random.seed(123)
X1 = np.zeros(N+1)
X2 = np.zeros(N+1)
X3 = np.zeros(N+1)
for i in range(1, N+1):
    X1[i] = X1[i-1] + np.random.poisson(lambda_t(i*dt) * dt)
    X2[i] = X2[i-1] + np.random.poisson(lambda_t(i*dt) * dt)
    X3[i] = X3[i-1] + np.random.poisson(lambda_t(i*dt) * dt)

# Perform the Levy-Ito decomposition on each realization
BM_drift1, comp_pp1, pure_jump1 = levy_ito_decomp(X1, dt)
BM_drift2, comp_pp2, pure_jump2 = levy_ito_decomp(X2, dt)
BM_drift3, comp_pp3, pure_jump3 = levy_ito_decomp(X3, dt)

# Plot the results
fig, axs = plt.subplots(3, 1, figsize=(12, 18))

axs[0].plot(X1, label='Poisson Process')
axs[0].plot(BM_drift1, label='BM with Drift')
axs[0].plot(comp_pp1, label='Compound Poisson Process')
axs[0].plot(pure_jump1, label='Pure Jump Process')
axs[0].legend()

axs[1].plot(X2, label='Poisson Process')
axs[1].plot(BM_drift2, label='BM with Drift')
axs[1].plot(comp_pp2, label='Compound Poisson Process')
axs[1].plot(pure_jump2, label='Pure Jump Process')
```

```
axs[1].legend()

axs[2].plot(X3, label='Poisson Process')
axs[2].plot(BM_drift3, label='BM with Drift')
axs[2].plot(comp_pp3, label='Compound Poisson Process')
axs[2].plot(pure_jump3, label='Pure Jump Process')
axs[2].legend()

plt.show()
```

The code used for figure 1.3 is as follows:

```
import numpy as np
import matplotlib.pyplot as plt

# Define the Gamma process intensity function
def lambda_t(t):
    return 2 + np.sin(t)

# Define the Brownian motion with drift component
def bm_drift(t, dt):
    return np.sqrt(dt) * np.random.normal()

# Define the compound Poisson process component
def compound_pp(t, dt):
    num_jumps = np.random.poisson(lambda_t(t) * dt)
    if num_jumps == 0:
        return 0
    else:
        jump_sizes = np.random.normal(size=num_jumps)
        return np.sum(jump_sizes)

# Define the pure jump process component
def pure_jump(t, dt):
    num_jumps = np.random.poisson(lambda_t(t) * dt)
    if num_jumps == 0:
        return 0
    else:
        jump_times = t + np.random.uniform(size=num_jumps) * dt
        jump_sizes = np.random.normal(size=num_jumps)
        return np.sum(jump_sizes * (jump_times <= t))

# Define the Levy-Ito decomposition function
def levy_ito_decomp(X, dt):
    BM_drift = np.zeros_like(X)
    comp_pp = np.zeros_like(X)
    pure_jump_process = np.zeros_like(X)
    for i in range(1, len(X)):
```

```
        t = i * dt
        BM_drift[i] = BM_drift[i-1] + bm_drift(t, dt)
        comp_pp[i] = comp_pp[i-1] + compound_pp(t, dt)
        pure_jump_process[i] = pure_jump_process[i-1] + pure_jump(t, dt)
    return BM_drift, comp_pp, pure_jump_process

# Define the simulation parameters
T = 10
N = 150
dt = T / N

# Simulate 3 realizations of the Gamma process
np.random.seed(123)
X1 = np.zeros(N+1)
X2 = np.zeros(N+1)
X3 = np.zeros(N+1)
for i in range(1, N+1):
    X1[i] = X1[i-1] + np.random.gamma(shape=1, scale=1/lambda_t(i*dt))
    X2[i] = X2[i-1] + np.random.gamma(shape=1, scale=1/lambda_t(i*dt))
    X3[i] = X3[i-1] + np.random.gamma(shape=1, scale=1/lambda_t(i*dt))

# Perform the Levy-Ito decomposition on each realization
BM_drift1, comp_pp1, pure_jump1 = levy_ito_decomp(X1, dt)
BM_drift2, comp_pp2, pure_jump2 = levy_ito_decomp(X2, dt)
BM_drift3, comp_pp3, pure_jump3 = levy_ito_decomp(X3, dt)

# Plot the results
fig, axs = plt.subplots(3, 1, figsize=(12, 18))

axs[0].plot(X1, label='Gamma Process', linestyle='', marker='o')
axs[0].plot(BM_drift1, label='BM with Drift')
axs[0].plot(comp_pp1, label='Compound Poisson Process')
axs[0].plot(pure_jump1, label='Pure Jump Process')
axs[0].legend()

axs[1].plot(X2, label='Gamma Process', linestyle='', marker='o')
axs[1].plot(BM_drift2, label='BM with Drift')
axs[1].plot(comp_pp2, label='Compound Poisson Process')
axs[1].plot(pure_jump2, label='Pure Jump Process')
axs[1].legend()

axs[2].plot(X3, label='Gamma Process', linestyle='', marker='o')
axs[2].plot(BM_drift3, label='BM with Drift')
axs[2].plot(comp_pp3, label='Compound Poisson Process')
axs[2].plot(pure_jump3, label='Pure Jump Process')
axs[2].legend()
```

The code used for figure 1.4 is as follows:

```
import numpy as np
import matplotlib.pyplot as plt

# Parameters
T = 1
N = 500
alpha = 1
beta = 0.5
sigma = 1
theta = -0.2
dt = T / N

# Define functions
def simulate_VG(alpha, beta, sigma, theta, T, N, seed=None):
    if seed is not None:
        np.random.seed(seed)
    t = np.linspace(0, T, N)
    S = np.zeros(N)
    S[0] = theta
    X = np.zeros(N)

    for i in range(1, N):
        # Generate Brownian motion with negative drift
        dB = -sigma**2 * dt + np.sqrt(dt) * np.random.normal()

        # Generate jump size
        dN = np.random.gamma(alpha * dt / beta, beta)

        # Update process
        X[i] = X[i-1] + dB + dN - theta * dt
        S[i] = np.exp(X[i])

    return S

def levy_ito_decomp(S):
    N = len(S)
    dS = np.diff(S)
    BM = np.cumsum(dS)
    CP = np.zeros(N)
    JP = np.zeros(N)
    for i in range(1, N):
        CP[i] = S[i] - np.exp(-alpha * dt) * S[i-1] - beta * np.sum(dS[0:i]
            * np.exp(-alpha * (i-1) * dt))
        JP[i] = S[i] - S[i-1] - CP[i]
    return BM, CP, JP

# Simulate and decompose processes
fig, axs = plt.subplots(3, 1, figsize=(8, 12), sharex=True)
```

```
for i in range(3):
    seed = i + 9
    S = simulate_VG(alpha, beta, sigma, theta, T, N, seed=seed)
    BM, CP, JP = levy_ito_decomp(S)
    axs[i].plot(np.linspace(0, T, N), S, label='Variance Gamma Process')
    axs[i].plot(np.linspace(0, T, N), CP, label='Compound Poisson Process')
    axs[i].plot(np.linspace(0, T, N), JP, label='Pure Jump Process')
    axs[i].plot(np.linspace(0, T, N-1), BM, label='Brownian Motion with
        Drift')
    axs[i].legend()

fig.tight_layout()
plt.show()
```

The code used for figure 1.5 is as follows:

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import levy_stable

# Parameters for the stable distribution
alpha = 1.9 # stability parameter
beta = 0.2 # skewness parameter
np.random.seed(123)

# Generate and plot three sample paths
plt.figure(figsize=(18,8))
for i in range(3):
    samples = levy_stable.rvs(alpha, beta, size=1000)
    plt.plot(samples, label=f"Realization {i+1}", alpha=0.5)

plt.legend()
plt.show()
```

The code used for figure 1.6 is as follows:

```
import numpy as np
import matplotlib.pyplot as plt

# Define the parameters of the Variance Gamma process
sigma = 0.1   # volatility parameter
theta = 0.1  # drift parameter
nu = 0.7      # Gamma process parameter
dt = 0.001     # time step
T = 1         # total time
N = int(T/dt) # number of time steps
```

```
np.random.seed(123)

# Generate a standard normal random variable for each time step
dW = np.random.normal(scale=np.sqrt(dt), size=N)

# Generate a Gamma-distributed random variable for each time step
dN = np.random.gamma(nu*dt/theta, theta, size=N)

# Calculate the VG process using the Wiener-Hopf factorization
VG = np.zeros(N)
for i in range(N):
    VG[i] = theta * np.sum(dW[:i]) + sigma * np.sum(np.sqrt(dN[:i])) -
        sigma * np.sum(np.sqrt(dN[i:]))
VG = VG + theta * T

# Calculate the positive and negative jump components
P = np.maximum(VG, 0)
N = np.minimum(VG, 0)

# Plot the original realization and its decompositions
fig, ax = plt.subplots(figsize=(16, 6))
ax.plot(VG, label='VG process')
ax.plot(P, label='Positive Component')
ax.plot(N, label='Negative Component')
ax.legend()
ax.set_xlabel('Time')
ax.set_ylabel('Value')
plt.show()
```

The code used for figure 2.1 is as follows:

```
import yfinance as yf
import numpy as np
import matplotlib.pyplot as plt
import pyswarms as ps

# Download historical Bitcoin data
btc = yf.Ticker("BTC-USD")
hist = btc.history(period="7y") # Get 7 years of data

# Calculate returns
returns = np.log(hist['Close'] / hist['Close'].shift(1)).dropna()

np.random.seed(123)
# Define the Brownian Motion model
def bm_model(params, returns, dt):
    sigma = np.array(params.flatten())
    n_periods = len(returns)
```

```
    simulated_returns = []
    for i in range(n_periods):
        # Calculate the return due to diffusion
        diffusion = sigma * np.sqrt(dt) * np.random.normal()

        # Calculate the total return
        total_return = diffusion
        simulated_returns.append(total_return)
    return np.array(simulated_returns)


# Define the objective function to be minimized
def objective_function(params, returns, dt):
    returns = returns.to_numpy()
    simulated_returns = bm_model(params, returns, dt)
    simulated_returns = np.transpose(simulated_returns)
    return np.transpose(np.sum((returns - simulated_returns)**2, axis=1))


# Define the bounds for the parameters
bounds = (np.array([-0.75]), np.array([0.75]))

# Define the PSO optimizer
options = {'c1': 2.05, 'c2': 2.05, 'w': 0.72984}
optimizer = ps.single.GlobalBestPSO(n_particles=30, dimensions=1,
    options=options, bounds=bounds)

# Perform the optimization
cost, opt_params = optimizer.optimize(objective_function, iters=1000,
    returns=returns, dt=1/252)

# Generate returns using the optimized parameters
simulated_returns = bm_model(opt_params, returns, dt=1/252)

# Plot the actual and simulated returns
plt.figure(figsize=(18,8))
plt.plot(returns.index, returns.values, label='Actual Returns')
plt.plot(returns.index, simulated_returns, label='Simulated Returns')
plt.legend()
plt.show()
```

The code used for figure 2.2 is as follows:

```
import yfinance as yf
import numpy as np
import matplotlib.pyplot as plt
import pyswarms as ps
```

```
# Download historical Bitcoin data
btc = yf.Ticker("BTC-USD")
hist = btc.history(period="7y") # Get 7 years of data

# Calculate returns
returns = np.log(hist['Close'] / hist['Close'].shift(1)).dropna()

np.random.seed(123)
# Define the Geometric Brownian Motion model
def gbm_model(params, returns, dt):
    if params.ndim > 1:
        mu, sigma = np.split(params, 2, axis=1)
    else:
        mu, sigma = params
    n_periods = len(returns)
    simulated_returns = []
    for i in range(n_periods):
        # Calculate the price change due to drift and diffusion
        drift = (mu - 0.5 * sigma**2) * dt
        diffusion = sigma * np.sqrt(dt) * np.random.normal()

        # Calculate the total return
        total_return = drift + diffusion
        simulated_returns.append(total_return)
    return np.array(simulated_returns)

# Define the objective function to be minimized
def objective_function(params, returns, dt):
    simulated_returns = gbm_model(params, returns, dt)
    returns = returns.to_numpy()
    simulated_returns = np.transpose(simulated_returns)
    return np.sum((returns - simulated_returns)**2)

# Define the bounds for the parameters
bounds = (np.array([-4, 0.2]), np.array([4, 0.5]))

# Define the PSO optimizer
options = {'c1': 2.05, 'c2': 2.05, 'w': 0.72984}
optimizer = ps.single.GlobalBestPSO(n_particles=30, dimensions=2,
    options=options, bounds=bounds)

# Perform the optimization
cost, opt_params = optimizer.optimize(objective_function, iters=1000,
    returns=returns, dt=1/252)

# Generate returns using the optimized parameters
simulated_returns = gbm_model(opt_params, returns, dt=1/252)
```

```
# Plot the actual and simulated returns
plt.figure(figsize=(18,8))
plt.plot(returns.index, returns.values, label='Actual Returns')
plt.plot(returns.index, simulated_returns, label='Simulated Returns')
plt.legend()
plt.show()
```

The code used for figure 2.3 is as follows:

```
import yfinance as yf
import numpy as np
import matplotlib.pyplot as plt
import pyswarms as ps
from scipy import stats

# Download historical Bitcoin data
btc = yf.Ticker("BTC-USD")
hist = btc.history(period="7y") # Get 7 years of data

# Calculate returns
returns = np.log(hist['Close'] / hist['Close'].shift(1)).dropna()

np.random.seed(123)
# Define the Geometric Brownian Motion model
def gbm_model(params, returns, dt):
    if params.ndim > 1:
        mu, sigma = np.split(params, 2, axis=1)
    else:
        mu, sigma = params
    n_periods = len(returns)
    simulated_returns = []
    for i in range(n_periods):
        # Calculate the price change due to drift and diffusion
        drift = (mu - 0.5 * sigma**2) * dt
        diffusion = sigma * np.sqrt(dt) * np.random.normal()

        # Calculate the total return
        total_return = drift + diffusion
        simulated_returns.append(total_return)
    return np.array(simulated_returns)

# Define the objective function to be minimized
def objective_function(params, returns, dt):
    simulated_returns = gbm_model(params, returns, dt)
    returns = returns.to_numpy()
    simulated_returns = np.transpose(simulated_returns)
    return np.sum((returns - simulated_returns)**2)
```

```
# Define the bounds for the parameters
bounds = (np.array([-4, 0.2]), np.array([4, 0.5]))

# Define the PSO optimizer
options = {'c1': 2.05, 'c2': 2.05, 'w': 0.72984}
optimizer = ps.single.GlobalBestPSO(n_particles=30, dimensions=2,
    options=options, bounds=bounds)

# Perform the optimization
cost, opt_params = optimizer.optimize(objective_function, iters=1000,
    returns=returns, dt=1/252)

# Generate returns using the optimized parameters
simulated_returns = gbm_model(opt_params, returns, dt=1/252)

# Plot the actual and simulated returns
plt.figure(figsize=(18,8))
plt.plot(returns.index, returns.values, label='Actual Returns')
plt.plot(returns.index, simulated_returns, label='Simulated Returns')
plt.legend()
plt.show()
```

The code used for figure 2.4 is as follows:

```
import yfinance as yf
import numpy as np
import matplotlib.pyplot as plt
import pyswarms as ps
from scipy import stats

# Download historical SPLV data
btc = yf.Ticker("SPLV")
hist = btc.history(period="10y") # Get one year of data

# Calculate returns
returns = np.log(hist['Close'] / hist['Close'].shift(1)).dropna()

np.random.seed(123)
# Define the Brownian Motion model
def bm_model(params, returns, dt):
    sigma = np.array(params.flatten())
    n_periods = len(returns)
    simulated_returns = []
    for i in range(n_periods):
        # Calculate the return due to diffusion
        diffusion = sigma * np.sqrt(dt) * np.random.normal()

        # Calculate the total return
```

```
        total_return = diffusion
        simulated_returns.append(total_return)
    return np.array(simulated_returns)


# Define the objective function to be minimized
def objective_function(params, returns, dt):
    returns = returns.to_numpy()
    simulated_returns = bm_model(params, returns, dt)
    simulated_returns = np.transpose(simulated_returns)
    return np.transpose(np.sum((returns - simulated_returns)**2, axis=1))


# Define the bounds for the parameters
bounds = (np.array([-3]), np.array([3]))

# Define the PSO optimizer
options = {'c1': 2.05, 'c2': 2.05, 'w': 0.72984}
optimizer = ps.single.GlobalBestPSO(n_particles=150, dimensions=1,
    options=options, bounds=bounds)

# Perform the optimization
cost, opt_params = optimizer.optimize(objective_function, iters=200,
    returns=returns, dt=1/252)

# Generate returns using the optimized parameters
simulated_returns = bm_model(opt_params, returns, dt=1/252)

# Plot the actual and simulated returns
plt.figure(figsize=(18,8))
plt.plot(returns.index, returns.values, label='Actual Returns')
plt.plot(returns.index, simulated_returns, label='Simulated Returns')
plt.legend()
plt.show()
```

The code used for figure 2.5 is as follows:

```
import yfinance as yf
import numpy as np
import matplotlib.pyplot as plt
import pyswarms as ps
from scipy import stats

# Download historical SPLV data
btc = yf.Ticker("SPLV")
hist = btc.history(period="10y") # Get one year of data

# Calculate returns
```

```
returns = np.log(hist['Close'] / hist['Close'].shift(1)).dropna()

np.random.seed(123)
# Define the Geometric Brownian Motion model
def gbm_model(params, returns, dt):
    if params.ndim > 1:
        mu, sigma = np.split(params, 2, axis=1)
    else:
        mu, sigma = params
    n_periods = len(returns)
    simulated_returns = []
    for i in range(n_periods):
        # Calculate the price change due to drift and diffusion
        drift = (mu - 0.5 * sigma**2) * dt
        diffusion = sigma * np.sqrt(dt) * np.random.normal()

        # Calculate the total return
        total_return = drift + diffusion
        simulated_returns.append(total_return)
    return np.array(simulated_returns)

# Define the objective function to be minimized
def objective_function(params, returns, dt):
    simulated_returns = gbm_model(params, returns, dt)
    returns = returns.to_numpy()
    simulated_returns = np.transpose(simulated_returns)
    return np.sum((returns - simulated_returns)**2)

# Define the bounds for the parameters
bounds = (np.array([-1, 0.02]), np.array([1, 0.05]))

# Define the PSO optimizer
options = {'c1': 2.05, 'c2': 2.05, 'w': 0.72984}
optimizer = ps.single.GlobalBestPSO(n_particles=150, dimensions=2,
    options=options, bounds=bounds)

# Perform the optimization
cost, opt_params = optimizer.optimize(objective_function, iters=200,
    returns=returns, dt=1/252)

# Generate returns using the optimized parameters
simulated_returns = gbm_model(opt_params, returns, dt=1/252)

# Plot the actual and simulated returns
plt.figure(figsize=(18,8))
plt.plot(returns.index, returns.values, label='Actual Returns')
plt.plot(returns.index, simulated_returns, label='Simulated Returns')
plt.legend()
```

```
plt.show()
```

The code used for figure 2.6 is as follows:

```
import yfinance as yf
import numpy as np
import matplotlib.pyplot as plt
import pyswarms as ps
from scipy import stats

# Download historical SPLV data
btc = yf.Ticker("SPLV")
hist = btc.history(period="10y") # Get one year of data

# Calculate returns
returns = np.log(hist['Close'] / hist['Close'].shift(1)).dropna()

np.random.seed(123)


# Define the jump-diffusion model
def jump_diffusion_model(params, returns, dt):

    mu, sigma, lambda_, mu_jump, sigma_jump = np.split(params, 5, axis=1)
    n_periods = len(returns)

    simulated_returns = []

    for i in range(n_periods):
        # Calculate the number of jumps
        n_jumps = np.random.poisson(lambda_ * dt)
        # Calculate the price change due to drift and diffusion
        drift = (mu - 0.5 * sigma**2) * dt
        diffusion = sigma * np.sqrt(dt) * np.random.normal()

        # Calculate the price change due to jumps
        jump_values = []
        for particle in range(len(n_jumps)):
            if n_jumps[particle][0] == 0:
                jump_values.append(0)
            else:
                jump_values.append(np.sum(np.random.laplace(mu_jump[particle][0],
                    sigma_jump[particle][0], (1, n_jumps[particle][0]))))
        jump_values = np.array(jump_values)
        # Calculate the total return
        total_return = drift[0] + diffusion[0] + jump_values
        simulated_returns.append(total_return)
```

```
    return np.array(simulated_returns)

def jump_diffusion_model_best(params, returns, dt):

    mu, sigma, lambda_, mu_jump, sigma_jump = np.split(params, 5)
    n_periods = len(returns)

    simulated_returns = []

    for i in range(n_periods):
        # Calculate the number of jumps
        n_jumps = np.random.poisson(lambda_ * dt)
        # Calculate the price change due to drift and diffusion
        drift = (mu - 0.5 * sigma**2) * dt
        diffusion = sigma * np.sqrt(dt) * np.random.normal()

        # Calculate the price change due to jumps
        jump_values = []
        if n_jumps == 0:
            jump_values.append(0)
        else:
            jump_values.append(np.sum(np.random.laplace(mu_jump,
                sigma_jump, n_jumps)))

        # Calculate the total return
        total_return = drift + diffusion + jump_values
        simulated_returns.append(total_return)

    return np.array(simulated_returns)

# Define the objective function to be minimized
def objective_function(params, returns, dt):

    simulated_returns = jump_diffusion_model(params, returns, dt)
    returns = returns.to_numpy()
    simulated_returns = np.transpose(simulated_returns)
    return np.transpose(np.sum((returns - simulated_returns)**2, axis=1))


# Define the bounds for the parameters
bounds = (np.array([-1, 0.05, 0.1, -0.01, 0.0001]), np.array([1, 0.1, 1,
    0.01, 0.01]))

# Define the PSO optimizer
options = {'c1': 2.05, 'c2': 2.05, 'w': 0.72984}
optimizer = ps.single.GlobalBestPSO(n_particles=150, dimensions=5,
    options=options, bounds=bounds)
```

```
# Perform the optimization
# print(optimizer.optimize(objective_function, iters=100, returns=returns,
    dt=1/252))
cost, opt_params = optimizer.optimize(objective_function, iters=200,
    returns=returns, dt=1/252)

# Generate returns using the optimized parameters
simulated_returns = jump_diffusion_model_best(opt_params, returns,
    dt=1/252)

# Plot the actual and simulated returns
plt.figure(figsize=(18, 8))
plt.plot(returns.index, returns.values, label='Actual Returns')
plt.plot(returns.index, simulated_returns, label='Simulated Returns')
plt.legend()
plt.show()
```

Bibliography

Abdel-Hameed, M. (2014). *Lévy processes and their applications in reliability and storage*. Springer Berlin Heidelberg.

Acerbi, C., & Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking & Finance*, *26*(7), 1487–1503.

Al-Saadony, M. (2016). A simulation study on stochastic differential equation driven by levy process.

Anderson, R. M., & May, R. M. (1992). *Infectious diseases of humans*. Oxford University Press.

Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, *9*(3), 203–228.

Asmussen, S., & Albrecher, H. (2010). *Ruin probabilities*.

Bachelier, L. (1900). Théorie de la spéculation. *Annales scientifiques de l'École normale supérieure*, *17*, 21–86.

Baeumer, B., & Meerschaert, M. M. (2010). Tempered stable lévy motion and transient super-diffusion. *Journal of Computational and Applied Mathematics*, *233*(10), 2438–2448.

Barndorff-Nielsen, O. E. (1997). Normal inverse gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics*, *24*(1), 1–13.

Barndorff-Nielsen, O. E., & Shephard, N. (2001). Non-gaussian ornstein-uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, *63*(2), 167–241.

Barndorff-Nielsen, O. E., & Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *64*(2), 253–280.

Bartumeus, F., & Catalan, J. (2009). Optimal search behavior and classic foraging theory. *Journal of Physics A: Mathematical and Theoretical*, *42*(43), 434002.

Basegmez, H., & Cekici, E. (2017). Financial applications of stable distributions: Implications on turkish stock market. *Pressacademia*, *6*(4), 364–374.

Basseville, M., & Nikiforov, I. V. (1993). *Detection of abrupt changes: Theory and application*. Prentice-Hall, Inc.

Bates, D. S. (1991). The crash of '87: Was it expected? the evidence from options markets. *The Journal of Finance*, *46*(3), 1009–1044.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies*, *9*(1), 69–107.

Bawly, G. M. (1936). Über einige Verallgemeinerungen der Grenzwertsätze der Wahrscheinlichkeitsrechnung. *Rec. Math. Moscou, n. Ser.*, *1*, 917–929.

Berg, H. C. (1993). *Random walks in biology*. Princeton University Press.

Berg, H. C. (2011). *E. coli in motion*. Springer.

Bertoin, J. (1998). *Lévy processes* (Vol. 121). Cambridge University Press.

Bielecki, T. R., & Rutkowski, M. (2013). *Credit risk: Modeling, valuation and hedging* (2004th ed.). Springer.

Bingham, N., & Kiesel, R. (2004). *Risk-neutral valuation: Pricing and hedging of financial derivatives*. Springer.

Black, F. (1976). Studies of stock price volatility changes. In *Proceedings of the business and economic statistics section meetings of the american statistical association, boston, massachusetts, august 23-26, 1976* (pp. 177–181). The Association.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, *81*(3), 637–654.

Böcker, K., & Klüppelberg, C. (2010). Multivariate models for operational risk. *Quantitative Finance*, *10*(8), 855–869.

Boyarchenko, S. I., & Levendorskii, S. Z. (2002). Perpetual american options under lévy processes. *SIAM Journal on Control and Optimization*, *40*(6), 1663–1696.

Broadie, M., & Detemple, J. B. (2004). ANNIVERSARY ARTICLE: Option pricing: Valuation models and applications. *Management Science*, *50*(9), 1145–1177.

Broadie, M., & Kaya, Ö. (2006). Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research*, *54*(2), 217–231.

Brockmann, D., Hufnagel, L., & Geisel, T. (2006). The scaling laws of human travel. *Nature*, *439*(7075), 462–465.

Brookes, B. (1955). Théorie de l'addition de variables aléatoires. by paul lévy pp. xx 385. second edition 1954. 1200f. (gauthier-villars, paris). *The Mathematical Gazette*, *39*(330), 344–344.

Brown, R. (1828). Xxvii. a brief account of microscopical observations made in the months of june, july and august 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *The Philosophical Magazine*, *4*(21), 161–173.

Caballero, M. E., Pardo, J. C., & Pérez, J. L. (2008). On the lamperti stable processes.

Cao, G., & Wang, Y. (2011). *Nanostructures and nanomaterials: Synthesis, properties, and applications*. World Scientific.

Carpenter, D. K. (1977). Dynamic light scattering with applications to chemistry, biology, and physics (berne, bruce j. pecora, robert). *Journal of Chemical Education*, *54*(10), A430.

Carr, P., Geman, H., Madan, D. B., & Yor, M. (2003). Stochastic volatility for levy processes. *Mathematical Finance*, *13*(3), 345–382.

Carr, P., & Madan, D. (1999). Option valuation using the fast fourier transform. *The Journal of Computational Finance*, *2*(4), 61–73.

Chan, J. C. C., & Eisenstat, E. (2018). Bayesian model comparison for time-varying parameter VARs with stochastic volatility. *Journal of Applied Econometrics*, *33*(4), 509–532.

Cox, D. R. (1955). Some statistical methods connected with series of events. *Journal of the Royal Statistical Society. Series B (Methodological)*, *17*(2), 129–164.

Crow. (2018). *Lognormal distributions* (E. L. Crow & K. Shimizu, Eds.). Routledge.

Cussler, E. (1997). *Diffusion: Mass transfer in fluid systems*. Cambridge University Press.

Daley, D. J., & Vere-Jones, D. (2014). *Introduction to the theory of point processes: Elementary theory and methods* (2nd, Vol. 1). Springer.

de Jager, M., Weissing, F. J., Herman, P. M. J., Nolet, B. A., & van de Koppel, J. (2011). Lévy walks evolve through interaction between movement and environmental complexity. *Science*, *332*(6037), 1551–1553.

Doob, J. L. (1942). The brownian movement and stochastic equations. *Annals of Mathematics*, *43*(2), 351–369.

Dubkov, A. A., Spagnolo, B., & Uchaikin, V. V. (2008). LéVY FLIGHT SUPERDIFFUSION: AN INTRODUCTION. *International Journal of Bifurcation and Chaos*, *18*(09), 2649–2672.

Dutta, P., & Horn, P. M. (1981). Low-frequency fluctuations in solids. *Reviews of Modern Physics*, *53*(3), 497–516.

Eberlein, E. (2001). Application of generalized hyperbolic lévy motions to finance. In *Lévy processes* (pp. 319–336). Birkhäuser Boston.

Einstein, A. (1905). Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der Physik*, *322*(8), 549–560.

Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling extremal events: For insurance and finance* (1st, Vol. volume 33). Springer.

Engle, R. F., & Russell, J. R. (1998). Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica*, *66*, 1127–1162.

Erlang, A. (1917). Solutions of some problems in the theory of probabilities of significance in automatic telephone exchanges. *Elektrotkeknikeren*, *13*, 5–13.

Erlang, A. K. (1909). Sandsynlighedsregning og telefonsamtaler. *Nyt tidsskrift for matematik*, *20*, 33–39.

Fang, F., & Oosterlee, C. W. (2009). A novel pricing method for european options based on fourier-cosine series expansions. *SIAM Journal on Scientific Computing*, *31*(2), 826–848.

Feldman, D. (1989). The theory of polymer dynamics, by m. doi and s. f. edwards, the clarendon press, oxford university press, new york, 1986, 391 pp. price: 78.50.. *Journal of Polymer Science Part C: Polymer Letters*, *27*(7), 239–240.

Feng, L., & Linetsky, V. (2008). Pricing discretely monitored barrier options and defaultable bonds in lévy process models: A fast hilbert transform approach. *Mathematical Finance*, *18*(3), 337–384.

Föllmer, H., & Schied, A. (2011). *An introduction in discrete time*. De Gruyter.

Fournier, N. (2009). Simulation and approximation of levy-driven stochastic differential equations.

Gardiner, C., & Zoller, P. (2004). Quantum noise, a handbook of markovian and non-markovian quantum stochastic methods with applications to quantum optics.

Gatheral, J., Jaisson, T., & Rosenbaum, M. (2014). Volatility is rough.

Gerber, H. U. (1979). *An introduction to mathematical risk theory*. S. S. Huebner Foundation for Insurance Education, Wharton School, University of Pennsylvania.

Gerber, H. U., & Shiu, E. S. W. (1997). The joint distribution of the time of ruin, the surplus immediately before ruin, and the deficit at ruin. *Insurance: Mathematics and Economics*, *21*(2), 129–137.

Giesecke, K. (2003). Credit risk modeling and valuation: An introduction. *SSRN Electronic Journal*.

Glasserman, P., Heidelberger, P., & Shahabuddin, P. (2001). Asymptotically optimal importance sampling and stratification for pricing path-dependent options. *Mathematical Finance*, *9*.

Good, I. J. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika*, *40*(3/4), 237.

Grosswald, E. (1976). The student t-distribution of any degree of freedom is infinitely divisible. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, *36*, 103–109.

Halgreen, C. (1979). Self-decomposability of the generalized inverse gaussian and hyperbolic distributions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, *47*(1), 13–17.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, *6*(2), 327–343.

Hougaard, P. (1986). A class of multivariate failure time distributions. *Biometrika*, *73*(3), 671.

Huda, S., Weigelin, B., Wolf, K., Tretiakov, K. V., Polev, K., Wilk, G., Iwasa, M., Emami, F. S., Narojczyk, J. W., Banaszak, M., Soh, S., Pilans, D., Vahid, A., Makurath, M., Friedl, P., Borisy, G. G., Kandere-Grzybowska, K., & Grzybowski, B. A. (2018). Lévy-like movement patterns of metastatic cancer cells revealed in microfabricated systems and implicated in vivo. *Nature Communications*, *9*(1).

Ismail, M. E. H. (1977). Bessel functions and the infinite divisibility of the student *t*- distribution. *The Annals of Probability*, *5*(4).

Ismail, M. E. H., & Kelker, D. H. (1979). Special functions, stieltjes transforms and infinite divisibility. *Siam Journal on Mathematical Analysis*, *10*, 884–901.

Itô, K. (1944). Stochastic integral. *Proceedings of the Japan Academy, Series A, Mathematical Sciences*, *20*(8).

Itô, K. (1946). On a stochastic integral equation. *Proceedings of the Japan Academy, Series A, Mathematical Sciences*, *22*(2).

Jacquier, E., Polson, N. G., & Rossi, P. (2004). Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics*, *122*(1), 185–212.

Jaffard, S. (1999). The multifractal nature of lévy processes. *Probability Theory and Related Fields*, *114*(2), 207–227.

James, A., Plank, M. J., & Edwards, A. M. (2011). Assessing lévy walks as models of animal foraging. *Journal of The Royal Society Interface*, *8*(62), 1233–1247.

Jang, P. A., Loeb, A., Davidow, M., & Wilson, A. G. (2017). Scalable levy process priors for spectral kernel learning. *Advances in Neural Information Processing Systems*, *30*.

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1994). *Continuous univariate distributions, volume 1* (2nd ed.). John Wiley & Sons.

Jokerst, J. V., Lobovkina, T., Zare, R. N., & Gambhir, S. S. (2011). Nanoparticle PEGylation for imaging and therapy. *Nanomedicine*, *6*(4), 715–728.

Jorgensen, B. (1981). *Statistical properties of the generalized inverse gaussian distribution*. Springer.

Jorion, P. (1997). *Value at risk: The new benchmark for controlling market risk*. McGraw-Hill.

Kaas, R., Goovaerts, M., Dhaene, J., & Denuit, M. (2008). *Modern actuarial risk theory*. Springer Berlin Heidelberg.

Kallenberg, O. (2002). *Foundations of modern probability* (2nd ed.). Springer.

Katada, Y., Hasegawa, S., Yamashita, K., Okazaki, N., & Ohkura, K. (2022). Swarm crawler robots using lévy flight for targets exploration in large environments. *Robotics*, *11*(4), 76.

Keating, C., & Shadwick, W. F. (2002). A universal performance measure.

Ken-Iti, S. (1999). *Lévy processes and infinitely divisible distributions*. Cambridge University Press.

Khinchin, A. (1948). *Asymptotische gesetze der wahrscheinlichkeitsrechnung*. Chelsea Publishing Company.

Khintchine, A. (1937). Zur Theorie der unbeschränkt teilbaren Verteilungsgesetze. *Rec. Math. Moscou, n. Ser.*, *2*, 79–117.

Kingman, J. F. C. (1993). *Poisson processes*. Clarendon Press.

Kleinrock, L. (1975). *Queueing systems, volume I*. John Wiley & Sons.

Klüppelberg, C., Lindner, A., & Maller, R. (2004). A continuous-time garch process driven by a lévy process: Stationarity and second-order behaviour. *Journal of Applied Probability*, *41*(3), 601–622.

Kolmogoroff, A. (1931). Über die analytischen methoden in der wahrscheinlichkeitsrechnung. *Mathematische Annalen*, *104*(1), 415–458.

Koponen, I. (1995). Analytic approach to the problem of convergence of truncated lévy flights towards the gaussian stochastic process. *Physical Review E*, *52*(1), 1197–1199.

Kuznetsov, A. (2010). Wiener-hopf factorization and distribution of extrema for a family of lévy processes.

Kyprianou, A. E. (2014). *Fluctuations of lévy processes with applications introductory lectures*. Springer.

Lemons, D. S., & Gythiel, A. (1908). Paul langevin's 1908 paper "on the theory of brownian motion" ["sur la théorie du mouvement brownien, " c. r. acad. sci. (paris) b146/b, 530–533 (1908)]. *American Journal of Physics*, *65*(11), 1079–1081.

Levy, P. (1939). L'addition des variables aléatoires définies sur une circonférence. *Bulletin de la Société Mathématique de France*, *67*, 1–41.

Lévy, P. (1925). *Calcul des probabilités*. Gauthier-Villars.

Lévy, P. (1948). *Processus stochastiques et mouvement brownien*. Gauthier-Villars.

Lewis, A. L. (2001). A simple option formula for general jump-diffusion and other exponential levy processes. *SSRN Electronic Journal*.

Li, J., Li, L., & Zhang, G. (2017). Pure jump models for pricing and hedging VIX derivatives. *Journal of Economic Dynamics and Control*, *74*, 28–55.

Li, J., An, Q., Lei, H., Deng, Q., & Wang, G.-G. (2022). Survey of lévy flight-based metaheuristics for optimization. *Mathematics*, *10*(15), 2785.

Madan, D. B., Carr, P., & Chang, E. C. (1998). The variance gamma process and option pricing. *Review of Finance*, *2*(1), 79–105.

Madan, D. B., & Seneta, E. (1990). The variance gamma (v.g.) model for share market returns. *The Journal of Business*, *63*(4), 511.

Mandelbrot, B. B., & Ness, J. W. V. (1968). Fractional brownian motions, fractional noises and applications. *SIAM Review*, *10*(4), 422–437.

Mantegna, R. N., & Stanley, H. E. (2000). An introduction to econophysics: Correlations and complexity in finance. *Physics Today*, *53*(12), 70–70.

McCall, J. J. (1965). *The Annals of Mathematical Statistics*, *36*(6), 1902–1904.

McNamara, B., Wiesenfeld, K., & Roy, R. (1988). Observation of stochastic resonance in a ring laser. *Physical Review Letters*, *60*(25), 2626–2629.

McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, *7*(3-4), 271–300.

McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management*. Princeton University Press.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, *3*(1-2), 125–144.

Metropolis, N., & Ulam, S. (1949). The monte carlo method. *Journal of the American Statistical Association*, *44*(247), 335–341.

Metzler, R., Chechkin, A. V., & Klafter, J. (2009). Levy statistics and anomalous transport: Levy flights and subdiffusion. In *Encyclopedia of complexity and systems science* (pp. 5218–5239). Springer New York.

Metzler, R., & Klafter, J. (2000). The random walk's guide to anomalous diffusion: A fractional dynamics approach. *Physics Reports*, *339*(1), 1–77.

Nolan, J. L. J. (2001). Stable distributions. models for heavy tailed data.

Peng, Y., & Wu, J. (2020). A lévy-driven stochastic queueing system with server breakdowns and vacations. *Mathematics*, *8*(8), 1239.

Perrin, J. B. (1913). Les atomes. *Nature*, *91*(2280), 473–473.

Poisson, S. (1837). *Recherches sur la probabilité des jugements en matière criminelle et en matière civile: Précédées des règles générales du calcul des probabilités*. Bachelier.

Pyke, G. H. (2014). Understanding movements of organisms: It's time to abandon the lévy foraging hypothesis (L. Giuggioli, Ed.). *Methods in Ecology and Evolution*, *6*(1), 1–16.

Radcliffe, J. (1977). The mathematical theory of infectious diseases and its applications. *Journal of The Royal Statistical Society Series C-applied Statistics*, *26*, 85–87.

Raman, I. M., & Bean, B. P. (1999). Ionic currents underlying spontaneous action potentials in isolated cerebellar purkinje neurons. *The Journal of Neuroscience*, *19*(5), 1663–1674.

Ramaswamy, S. (2010). The mechanics and statistics of active matter. *Annual Review of Condensed Matter Physics*, *1*(1), 323–345.

Rather, S. A., & Bala, P. S. (2021). Lévy flight and chaos theory-based gravitational search algorithm for mechanical and structural engineering design optimization. *Open Computer Science*, *11*(1), 509–529.

Regazzini, E. (2008). De finetti's contribution to the theory of random functions. *International Journal of Approximate Reasoning*, *47*(3), 387–393.

Renshaw, E. (1991). *Modelling biological populations in space and time*. Cambridge University Press.

Samorodnitsky, G., & Taqqu, M. S. (1994). *Stable non-gaussian random processes*. Chapman & Hall/CRC.

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. In *The world scientific handbook of futures markets* (pp. 25–38).

Schoutens, W. (2003). *Levy processes in finance: Pricing financial derivatives*. Wiley.

Schoutens, W. (2002). The meixner process : Theory and applications in finance.

Schoutens, W., & Teugels, J. L. (1998). Lévy processes, polynomials and martingales. *Communications in Statistics. Stochastic Models*, *14*(1-2), 335–349.

Sharpe, W. F. (1994). The sharpe ratio. *The Journal of Portfolio Management*, *21*(1), 49–58.

Shi, Y., Feng, Q., Shu, Y., & Xiang, Y. (2020). Multi-dimensional lévy processes with lévy copulas for multiple dependent degradation processes in lifetime analysis. *Quality Engineering*, *32*(3), 434–448.

Shlesinger, M. F., West, B. J., & Klafter, J. (1987). Lévy dynamics of enhanced diffusion: Application to turbulence. *Physical Review Letters*, *58*(11), 1100–1103.

Steutel, F. W. (1970). *Preservation of infinite divisibility under mixing and related topics*. Mathematisch Centrum.

Steutel, F. (1973). Some recent results in infinite divisibility. *Stochastic Processes and their Applications*, *1*(2), 125–143.

Tankov, P., & Cont, R. (2003). *Financial modelling with jump processes*. Chapman; Hall/CRC.

Thorin, O. (1977a). On the infinite divisbility of the pareto distribution. *Scandinavian Actuarial Journal*, *1977*(1), 31–40.

Thorin, O. (1977b). On the infinite divisibility of the lognormal distribution. *Scandinavian Actuarial Journal*, *1977*, 121–148.

Tocher, K. D., & Takacs, L. (1963). Introduction to the theory of queues. *Journal of the Royal Statistical Society. Series A (General)*, *126*(4), 598.

Tweedie, M. C. K. (1984). An index which distinguishes between some important exponential families. In J. K. Ghosh & J. Roy (Eds.), *Statistics: Applications*

and new directions. proceedings of the indian statistical institute golden jubilee international conference. calcutta: Indian statistical institute (pp. 579–604).

Viswanathan, G. M., da Luz, M. G. E., Raposo, E. P., & Stanley, H. E. (2011). *The physics of foraging*. Cambridge University Press.

Wasan, M. T. (1968). On an inverse gaussian process. *Scandinavian Actuarial Journal*, *1968*(1-2), 69–96.

Wosniack, M. E., Santos, M. C., Raposo, E. P., Viswanathan, G. M., & da Luz, M. G. E. (2017). The evolutionary origins of lévy walk foraging (F. Bartumeus, Ed.). *PLOS Computational Biology*, *13*(10), e1005774.

Zaburdaev, V., Denisov, S., & Klafter, J. (2015). Lévy walks. *Rev. Mod. Phys.*, *87*, 483–530.

Zhang, C., & Tao, D. (2013). Risk bounds of learning processes for lévy processes. *J. Mach. Learn. Res.*, *14*(1), 351–376.