UNIVERSITÀ COMMERCIALE LUIGI BOCCONI

DEC-Dipartimento di Scienze delle Decisioni

PhD in Statistics - XXII cycle

# Profiling Processes of Meixner Type

Advisor: Prof.Pietro Muliere

PhD. Thesis

Emanuele Mazzola

Matr. 1217694

Academic Year 2009–2010

**DECLARATORIA SULLA TESI DI DOTTORATO**
Da inserire come prima pagina della tesi

Il/la sottoscritto/a

COGNOME MAZZOLA

NOME EMANUELE

Matr. 1217694

Titolo della tesi:

Profiling Processes of Meixner Type

Dottorato di ricerca in STATISTICA

Ciclo XXII

Tutor del dottorando Prof. P.Muliere

Anno di discussione 2010

**DICHIARA**

sotto la sua responsabilità di essere a conoscenza:

1) che, ai sensi del D.P.R. 28.12.2000, N. 445, le dichiarazioni mendaci, la falsità negli atti e l'uso di atti falsi sono puniti ai sensi del codice penale e delle Leggi speciali in materia, e che nel caso ricorressero dette ipotesi, decade fin dall'inizio e senza necessità di nessuna formalità dai benefici previsti dalla presente declaratoria e da quella sull'embargo;

2) che l'Università ha l'obbligo, ai sensi dell'art. 6, comma 11, del Decreto Ministeriale 30 aprile 1999 prot. n. 224/1999, di curare il deposito di copia della tesi finale presso le Biblioteche Nazionali Centrali di Roma e Firenze, dove sarà consentita la consultabilità, fatto salvo l'eventuale embargo legato alla necessità di tutelare i diritti di enti esterni terzi e di sfruttamento industriale/commerciale dei contenuti della tesi;

3) che il Servizio Biblioteca Bocconi archivierà la tesi nel proprio Archivio istituzionale ad Accesso Aperto e che consentirà unicamente la consultabilità on-line del testo completo (fatto salvo l'eventuale embargo);

4) che per l'archiviazione presso la Biblioteca Bocconi, l'Università richiede che la tesi sia consegnata dal dottorando alla Società NORMADEC (operante in nome e per conto dell'Università) tramite procedura on-line con contenuto non modificabile e che la Società Normadec indicherà in ogni piè di pagina le seguenti informazioni:
- tesi di dottorato (titolo *tesi)* ...Profiling Processes of Meixner Type ;
- di *(cognome e nome del dottorando)* Emanuele Mazzola;

- discussa presso l'Università commerciale Luigi Bocconi – Milano nell'anno 2010. *(anno di discussione);*
- La tesi è tutelata dalla normativa sul diritto d'autore (legge 22 aprile 1941, n.633 e successive integrazioni e modifiche). Sono comunque fatti salvi i diritti dell'Università Commerciale Luigi Bocconi di riproduzione per scopi di ricerca e didattici, con citazione della fonte;
- **solo nel caso sia stata sottoscritta apposita altra dichiarazione con richiesta di embargo**: La tesi è soggetta ad embargo della durata di .......... mesi (indicare durata embargo);

5) che la copia della tesi depositata presso la NORMADEC tramite procedura on-line è del tutto identica a quelle consegnate/inviate ai Commissari e a qualsiasi altra copia depositata negli Uffici dell'Ateneo in forma cartacea o digitale e che di conseguenza va esclusa qualsiasi responsabilità dell'Ateneo stesso per quanto riguarda eventuali errori, imprecisioni o omissioni nei contenuti della tesi;

6) che il contenuto e l'organizzazione della tesi è opera originale realizzata dal sottoscritto e non compromette in alcun modo i diritti di terzi (legge 22 aprile 1941, n.633 e successive integrazioni e modifiche), ivi compresi quelli relativi alla sicurezza dei dati personali; che pertanto l'Università è in ogni caso esente da responsabilità di qualsivoglia natura, civile, amministrativa o penale e sarà dal sottoscritto tenuta indenne da qualsiasi richiesta o rivendicazione da parte di terzi;

7) **scegliere l'ipotesi 7a o 7b indicate di seguito**:

7a) che la tesi di dottorato non è il risultato di attività rientranti nella normativa sulla proprietà industriale, non è stata prodotta nell'ambito di progetti finanziati da soggetti pubblici o privati con vincoli alla divulgazione dei risultati; non è oggetto di eventuali registrazioni di tipo brevettale o di tutela, e quindi non è soggetta a embargo;

Oppure

7b) che la tesi di Dottorato rientra in una delle ipotesi di embargo previste nell'apposita dichiarazione "**RICHIESTA DI EMBARGO DELLA TESI DI DOTTORATO**" sottoscritta a parte.

Data _22 Gennaio 2010____

F.to (indicare nome e cognome) _Emanuele Mazzola______

Emanuele Mazzola

# Acknowledgements

This work wouldn't surely have been possible without the many discussions with Prof. Pietro Muliere who was always available to listen to my doubts and to share my difficulties during this past year.
Also my deepest gratitude goes to Prof. Piercesare Secchi, for believing in me sometimes more than I did, and for giving me the chance of starting this experience.
A great thank you goes to all my teachers during these years who have contributed in making this course of study always interesting and stimulating.

A particular acknowledgement for Prof. Marc Yor, for the quick but extremely useful exchange of information during last summer, Drs. Friedrich Hubalek and Carlo Sgarra for the initial starting points of this work, and all the people I have asked for information during this period: Prof. Ernst Eberlein, Prof. Martin Schweizer, Prof. Murad Taqqu, Prof. Fabio Maccheroni and Dr. Marzia DeDonno.
A special thank goes to Dr. Souvik Ghosh, for kindly meeting me at Columbia University, New York City, and for accepting to read my work in progress, and Prof. Marco Frittelli for the useful encounter.
Thanks also to Prof. Fausta Luscia and Dr. Catia Scricciolo for their support when I was in deep doubt.

*Emanuele*

# Contents

# Preface

Lévy processes find their roots at the beginning of XX century, when at first infinitely divisible distributions were studied by Bruno DeFinetti, who first posed problems that would be solved over a seventeen-year span basically by the strong mathematical personalities of Andrei Kolmogorov, Paul Lévy and Aleksandr Yakovlevich Khinchine.

From that moment a great amount of work in this direction has been performed, both from a purely theoretical point of view, with the developement of the basic theory, and for instance with the introduction of stable processes in 1925, and with the matching of theory of stochastic processes with martingale theory in the 1940's. Lévy processes , as we are going to see, constitute a wide class containing well known examples as Poisson processes or Brownian motions, but also Markov processes and semimartingales.

In applications, mostly in mathematical finance, they provide more flexible distributions for the asset returns on which derivative securities are based; the usual model is the Black-Scholes model, dating 1973, although it basically goes back to Samuelson (1965) who improved Bachelier's introduction of Brownian motion in 1900.

Although being forced to give up some important properties of Brownian motion driven markets like completeness for instance, it has been shown both with qualitative and quantitative results that the more general Lévy processes are suitable as driving the background distributions of asset returns.

In this direction an important and realistical role have had hyperbolic distributions, introduced by Barndorff-Nielsen in the late 1970's, which also have been employed in modelization of fluid

turbulence and for instance of the distribution of particle size from aeolian sand deposits.
Meixner process is a particular case of Lévy process whose origins can be found also in the theory of orthogonal polynomials, in particular the so called Meixner-Pollaczeck polynomials, to which it is related by a martingale relation. The basics of this process have been given by W.Shoutens in 1999, who also pointed out the polynomial origin, and by B.Grigelionis the same year, who derived this process as a particular case of the generalized-$z$ processes.
The first part of this work is devoted to the general introduction of Lévy process along with the most common tools and relations for characterizing their properties. In the first chapter an historical background is also given, with the relations with infinitely divisible distributions and martingale and semimartingale theory.
The second chapter lists a non-exhaustive collection of some of the most interesting Lévy processes, from Poisson process and Brownian motion, to $\alpha$-stable processes, with details on each one's properties, and, where possible, a sample of trajectory with respect of a given choice of the parameters involved.
Part two of this work is dedicated to Meixner process.
In chapter 3 Meixner process is defined in the classical way, i.e. relying on the infinitely divisible Meixner distribution, seen as a particular case of generalized-$z$ distribution, and classical estimators of the parameters involved are given (namely moments and ML estimators); the second part of chapter 3 contains the calculations which allow the writing of Meixner process as a time-changed Brownian motion, which fact should lead to a simple way of simulating trajectories of the process, which will be shown in chapter 6, and some particular and very technical characterizations mainly due to the work of J.Pitman and M.Yor.
Chapter 4 deals with Esscher transform, the most common tool used to price financial derivatives based on Lévy processes, introducing some problems, mainly of market incompleteness, that generate when substituting the usual Brownian motion with more general Lévy processes.
Chapter 5 is devoted to the introduction of theory of orthogonal polynomials, and points out

what are the conditions that allow Meixner process to be generated from this different point of view, along with a graphical analysis of moments of Meixner process and the computation of Fisher's information for Meixner-Pollaczeck polynomials.
Last, chapter 6 shows graphically the main properties of Meixner process and its better behavior with respect to usual Brownian motion when attempting to model financial asset returns.
The simulations have been performed with R software, while most of the computing and the study of the properties has been carried on with Mathematica 7.0.0.

The present thesis is firstly a review of most of the literature available on Meixner process. The justification of this is mainly the relative dispersion of the concerning publications: Meixner process in fact appears often as an example or as a subsection of more general cases but it is not studied by itself, while for its properties, fitting properties and mathematical origins it deserves to be studied as a stand alone model.
This means a good amount of job has been performed trying to put together and give the right placement to what up to now is known about this particular element of the class of Lévy processes .
The usual theoretical introduction regarding general Lévy processes and consequent examples, constructed using the classical way (namely characteristic function) which can be found in chapters 1 and 2, convey a "classical" definition of Meixner process, first given independently by W. Schoutens and B. Grigelionis as said. Their first intention was to apply this new process as a mathematical financial model, following the fashion of the already tested hyperbolic processes introduced by Barndorff-Nielsen.
In this context our contribution stands with the proof of a particular case (corollary 1, section 4.5) of a theorem by B. Grigelionis, which provides the general Esscher martingale measure for Meixner process on which an analogous of the famous Black-Scholes pricing formula can be based.
In 2000 W. Schoutens also provides a stronger mathematical foundation of the process, by

means of orthogonal polynomials, which can be found in chapter 4.

Thus, the central chapter 3 describes the main effect of the theory contained in chapters 1 (with examples in chapter 2) and 4.

Some important details were missing from first formulation of the process: ML parameter estimation of the background Meixner distribution, which can be seen in chapter 3, and mostly a double characterization of the process in terms of subordinated Brownian motion (given by D.B.Madan and M.Yor as cited), and in terms of "process containing an hyperbolic function in the characteristic function" (given by J.Pitman and M.Yor).

The chance of writing Meixner process as a subordinated Brownian motion makes possible to simulate the process. To our knowledge, in the revised literature simulations of trajectories of Meixner process don't appear anywhere. We obtained some results in this sense, by using an original R routine based on the work of Madan and Yor: these results are shown in section 6.2.

Once defined and characterized, Meixner process can be well employed as a model for financial data, namely log returns of market indexes or assets as well. An example has been carried out in this sense in section 6.3, evaluating model performance by means of two possible distances between underlying distributions, and comparing the result with another model derived from Normal Inverse Gaussian process, which was already known.

**Note**: the notations $X_t$ or $X(t)$ will be used equivalently for a stochastic process $X = \{X_t = X(t), t \geq 0\}$.

# Part I

# Chapter 1

# Some Elements about Lévy Processes

## 1.1 Introduction

Lévy processes are a product of the mathematics of the first $35$ years of XX century. They constitute a fundamental class in the theory of stochastic processes, containing as its elements basic and already well known processes as Brownian motions and Poisson processes to name a couple, of which they provide a generalization and a more flexible example.

The importance of Lévy processes is very well acquainted both as a class of stochastic processes which for instance stands as a starting point for the study of other families of processes as Markov processes (Lévy processes actually form the class of space-time homogeneous Markov processes) or semimartingales, and as a set of model which provide more flexibility in contexts like mathematical finance.

Here, in fact, a good amount of literature has shown the inadequacy and stiffness of the usual Brownian-motion-based models as opposite to the more pitchable, fitting and general Lévy processes .

The introduction of Lévy processes in finance , though, is not painless: it rouses different kinds of problems, both practical and theoretical; the situation of incomplete market which they generate is not always easy to handle both in terms of calculations and of teoretical structures involved, as for instance the multiplicity of equivalent martingale measures on which most of the option pricing theory is based.

Since the last part of the $1990$'s different applications have ben tried for Lévy processes ranging from description of fluid turbulence to quantum fields.

## 1.2 Historical remarks

The first results on Lévy processes date back to the late $1920$'s with the study of infinitely divisible (ID) distributions.

Bruno DeFinetti was the recognized pioneer of ID distributions with his $1929 - 1931$ papers, but the term *infinitely divisible distributions* will appear only $5$ years later in the Moscow mathematical school, in the $1936$ unpublished ph.D thesis by G.M.Bawly $(1908 - 1941)$. The first formal definition of an ID distribution was given by A.Y.Khinchine in [Khi37b] and reads:

**Definition 1.** *A distribution of a random variable which for any positive integer $n$ can be represented as a sum of $n$ identically distributed independent random variables is called an infinitely divisible distribution.*

Lévy himself attributes to Khinchine the use of the name *indéfinitement divisible*. The canonical form of ID distributions is known in literature as *Lévy-Khinchine formula*, surely because it was so named by Gnedenko and Kolmogorov.

In the following it will be seen how ID distributions are intimately related to Lévy processes . By now let us recall the theorems historically significative concerning the structure of ID distributions

**Theorem 1.** *(DeFinetti) A characteristic function is ID if and only if it has the form*

$$\phi(t) = \lim_{m \to \infty} \exp[p_m(\psi_m(t) - 1)],$$

*where $p_m$ are numbers in $\mathbb{R}^+$, while $\psi_m(t)$ are characteristic functions.*

**Theorem 2.** *(DeFinetti) The limit of a sequence of finite products of Poisson-type characteristic functions is ID. The converse it is also true.*

This means that the class of ID laws coincides with the class of distribution limits of finite convolutions of distributions of Poisson type.

**Theorem 3.** *(Kolmogorov canonical representation) The function $\phi(t)$ is the characteristic function of an ID distribution with finite second moment if and only if it can be written as*

$$\log \phi(t) = i\gamma t + \int_{-\infty}^{+\infty} \left(e^{itu} - 1 - itu\right) \frac{dK(u)}{u^2},$$

*where $\gamma$ is a real constant, and $K(x)$ is a non decreasing and bounded function such that $K(-\infty) = 0$. The integrand is defined such that for $u = 0$ it is equal to $-t^2/2$.*

**Theorem 4.** *(Lévy canonical representation) The function $\phi(t)$ is the characteristic function of an ID distribution if and only if it can be written as*

$$\log \phi(t) = i\gamma t - \frac{\sigma^2}{2} t^2 + \int_{-\infty}^{0^-} \left(e^{itu} - 1 - \frac{itu}{1+u^2}\right) dM(u) + \\ + \int_{0^+}^{+\infty} \left(e^{itu} - 1 - \frac{itu}{1+u^2}\right) dN(u),$$

*where $\gamma \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$ and the functions $M(u), N(u)$ satisfy the following conditions:*

1. *$M(u)$ and $N(u)$ are non decreasing in $(-\infty, 0)$ and $(0, +\infty)$ respectively;*
2. *$M(-\infty) = N(+\infty) = 0$;*
3. *the integrals $\int_{-\varepsilon}^{0} u^2\, dM(u)$ and $\int_0^{\varepsilon} u^2\, dN(u)$ are finite for every $\varepsilon > 0$.*

**Theorem 5.** *(Lévy-Khinchine canonical representation) The function $\phi(t)$ is the characteristic function of an ID distribution if and only if it can be written as*

$$\log \phi(t) = i\gamma t + \int_{-\infty}^{+\infty} \left(e^{itu} - 1 - \frac{itu}{1+u^2}\right) \frac{1+u^2}{u^2} dG(u),$$

*where $\gamma \in \mathbb{R}$, and $G(u)$ is a non decreasing and bounded function such that $G(-\infty) = 0$. The integrand is defined such that for $u = 0$ it is equal to $-t^2/2$.*

Just after the 1928 International Congress of Mathematician held in Bologna from 3 to 10 September 1928, DeFinetti started a research regarding functions with random increments based on the theory of ID characteristic functions. His results can be summarized in a number of relevant theorems partly stated above.

The papers published in 1929 – 1931 by DeFinetti in the *Rendiconti della Reale Accademia Nazionale dei Lincei*, attracted the attention of Kolmogorov who was interested to solve the so called *DeFinetti's problem*, that is to find the general formula for the characteristic function of ID distributions.

Kolmogorov gave an exhaustive answer to the problem for the case of variables with finite second moment in two notes of 1932; his final result is known as the Kolmogorov canonical representation shown above.

The general case of DeFinetti's problem, including the case of infinite variance, was investigated in 1934 – 1935 by P.Lévy, and his result, independent from that of Kolmorogov, is the Lévy canonical representation shown above. More details are given in the paper by F.Mainardi and S.Rogosin [MR06].

A.Y.Khinchine came in 1937 to show that Lévy's result can be obtained also by an extension of Kolmogorov's method, and his final statement is the celebrated Lévy-Khinchine canonical representation formula for the ID characteristic functions. An interesting translation of the russian article by Khinchine [Khi37a] can be found as well in [MR06].

That's how the main results regarding ID distributions were born; in the following it will be shown their relationship with the concept of Lévy process .

Also, in the following part of the first chapter Lévy processes will be introduced in a general framework, describing the main results that will be somehow useful or referred to in the following. It is natural that the study of Lévy processes from a theoretical point of view has reached a real deep level of developement, but it is not the aim of this first part of the work giving account of all the details that can be for instance found in Sato [Sat99], Appelbaum [App04], and in

Jacod and Shiryaev's books [JS02] for a more accurate part of the semimartingale topic and for a more general point of observation.

Then we present the two main theorems (namely the Lévy-Khinchine representation and the Lévy-Itô decomposition) with part of the theoretical auxiliary construction to get them.

We generally will be following the approach of Appelbaum [App04], integrating with some examples taken from the lectures delivered by G.Samorodnitsky [Sam07] in 2007.

## 1.3 Definition of Lévy Process

To embed the processes we are going to introduce in the correct environment, some definitions have to be set to characterize the right spaces and the main theoretical structures which will appear in the following.

Let $(\Omega, \mathscr{F}, \mathbf{F}, P)$ be a stochastic basis, i.e. a probability space $(\Omega, \mathscr{F}, P)$ equipped with a filtration $\mathbf{F} = (\mathscr{F}_t)_{t\in\mathbb{R}_+}$; filtration here means an increasing and right-continuous family of sub-$\sigma$-algebrae of $\mathscr{F}$ (i.e. $\mathscr{F}_s \subset \mathscr{F}_t$ for $s \leq t$ and $\mathscr{F}_t = \bigcap_{s>t} \mathscr{F}_s$).

By convention let $\mathscr{F}_\infty = \mathscr{F}$ and $\mathscr{F}_{\infty^-} = \bigvee_{s\in\mathbb{R}_+} \mathscr{F}_s$.

**Definition 2.** *A stochastic process on* $(\Omega, \mathscr{F}, P)$ *is called* adapted *to the filtration* $\mathbf{F}$ *if* $X(t)$ *is* $\mathscr{F}_t$*−measurable for all* $t \geq 0$.

**Definition 3.** *A random variable* $T$ *defined on* $(\Omega, \mathscr{F}, P)$ *with values on* $[0, \infty]$ *is a* stopping time *with respect to* $\mathbf{F}$ *if the event* $\{T \leq t, t \geq 0\}$ *belongs to* $\mathscr{F}_t$.

**Definition 4.** *An adapted stochastic process* $X = \{X(t), t \geq 0\}$ *such that* $E[X(t)] \leq \infty$ *for all* $t \geq 0$ *is a* martingale *with respect to the usual filtration* $\mathbf{F}$ *if for all* $0 \leq s \leq t$ *it holds that*

$$E[X(t)|\mathscr{F}_s] = X(s) \; a.s.$$

**Definition 5.** *A stochastic process* $X = \{X(t), t \geq 0\}$ *defined on* $(\Omega, \mathscr{F}, \mathbf{F}, P)$ *with values in* $\mathbb{R}^d,\ d \in [1, \infty)$ *is called a* Lévy process *if the following conditions are satisfied:*

(L1) $X(0) = 0 \quad a.s.$

(L2) *$X$ has independent increments, i.e. $X(t) - X(s)$ is independent of $\mathscr{F}_s$ for any $0 \leq s < t \leq T$.*

(L3) *$X$ has stationary increments, i.e. for any $s, t \geq 0$ the distribution of the increment $X(t + s) - X(t)$ does not depend on $t$.*

(L4) *$X$ is stochastically continuous, i.e. for every $t \geq 0$ and $\varepsilon > 0$:*

$$\lim_{s \to t} P(|X(t) - X(s)| > \varepsilon) = 0.$$

### 1.3.1 Infinite divisibility

As we have already seen, one of the main concepts laying beneath the idea of Lévy process is the notion of infinite divisibility, on which is also based one of the main and most important results of the theory: the Lévy-Khinchine formula. The definition of infinite divisibility is given generally for random vectors, and then the link with Lévy processes is usually shown.

**Definition 6.** *Let $X$ be a random vector taking values in $\mathbb{R}^d$ with law $\mu_X$. We say that $X$ is* infinitely divisible *if, for all $n \in \mathbb{N}$, there exist i.i.d. random vectors $Y_1^{(n)}, \ldots, Y_n^{(n)}$ such that*

$$X \stackrel{d}{=} Y_1^{(n)} + \ldots + Y_n^{(n)}. \tag{1.1}$$

Let $\phi_X(u) = E[e^{i(u,X)}]$ denote the characteristic function of $X$ , where $u \in \mathbb{R}^d$.
More generally, if $\mu \in \mathcal{M}_1(\mathbb{R}^d)$, the set of all Borel probability measures on $\mathbb{R}^d$, then $\phi_\mu(u) = \int_{\mathbb{R}^d} e^{i(u,y)} \mu(dy)$.

**Proposition 1.** *The following are equivalent:*

(1) *$X$ is infinitely divisible;*

(2) *$\mu_X$ has a convolution $n^{th}$ root that is itself the law of a random vector, for each $n \in \mathbb{N}$;*

(3) *$\phi_X$ has an $n^{th}$ root that is itself the characteristic function of a random vector, for each $n \in \mathbb{N}$.*

*Proof*

(1) $\Rightarrow$ (2). The common law of the $Y_j^{(n)}$, $j = 1, \dots, n$ is the required convolution $n^{th}$ root.

(2) $\Rightarrow$ (3). Let $Y$ be a random variable with law $(\mu_X)^{1/n}$. We have , for each $u \in \mathbb{R}^d$ ,

$$\phi_X(u) = \int \dots \int e^{i(u,y_1+\dots+y_n)}(\mu_X)^{1/n}(dy_1)\dots(\mu_X)^{1/n}(dy_n) = \varphi_Y(u)^n$$

where $\varphi_Y(u) = \int_{\mathbb{R}^d} e^{i(u,y)}(\mu_X)^{1/n}(dy)$, and the required result follows.

(3)$\Rightarrow$(1). Choose $Y_1^{(n)}, \dots, Y_n^{(n)}$ to be independent copies of the given random vector; then we have

$$E[e^{i(u,X)}] = E[e^{i(u,Y_1^{(n)})}]\dots E[e^{i(u,Y_n^{(n)})}] = E[e^{i(u,Y_1^{(n)}+\dots+Y_n^{(n)})}]$$

from which we deduce (1.1) as required. □

It is possible to generalize the definition of infinite divisibility as follows:

**Definition 7.** *$\mu \in \mathcal{M}_1(\mathbb{R}^d)$ is infinitely divisible if it has a convolution $n^{th}$ root in $\mathcal{M}_1(\mathbb{R}^d)$ for each $n \in \mathbb{N}$.*

As a consequence, it can be taken as an operative characterizing definition of infinite divisibility the following

**Definition 8.** *$\mu \in \mathcal{M}_1(\mathbb{R}^d)$ is* infinitely divisible *if and only if for each $n \in \mathbb{N}$ there exists $\mu^{1/n} \in \mathcal{M}_1(\mathbb{R}^d)$ for which*

$$\phi_\mu(u) = \left[\phi_{\mu^{1/n}}(u)\right]^n$$

*for each $u \in \mathbb{R}^d$.*

Observation: the convolution $n^{th}$ root $\mu^{1/n}$ in definition 8 above is unique when $\mu$ is infinitely divisible. Moreover, in this case the complex-valued function $\phi_\mu$ always has a “distinguished” $n^{th}$ root, which we denote by $\phi_\mu^{1/n}$; this is the characteristic function of $\mu^{1/n}$ (see Sato,[Sat99] pgg. $32-4$, for details).

Trivial examples of distributions enjoying this property are the Gaussian and Poisson distributions. Other examples of infinitely divisible distributions are the compound Poisson distribution, the exponential, the $\Gamma$ distribution, the geometric, the negative binomial, the Cauchy distribution and the strictly stable distribution.

Counterexamples are the uniform and binomial distributions.

The two following examples open the path to the Lévy-Khinchine formula, as they provide a typical fashion of the characteristic function of an infinitely divisible random variable

**Example 1.** (Gaussian distribution) *Let $Y \sim N(\mu, \Sigma)$, with $\mu \in \mathbb{R}^d$ and $\Sigma$ a $d \times d$ symmetric positive definite matrix. If we consider for every $n \in \mathbb{N}$*

$$Y^{(n)} \sim N\left(\mu/n, (1/n)\Sigma\right)$$

*then it's easy to verify that $Y_1^{(n)} + \ldots + Y_n^{(n)} \overset{d}{=} Y$.*

*It is also easily computed the characteristic function of $Y$ as*

$$\phi_Y(u) = E\left[e^{i(u,Y)}\right] = \exp\left(-\frac{1}{2}u\Sigma u' + i(u,\mu)\right)$$

**Example 2.** (compound Poisson distribution) *Let $Y = \sum_{j=1}^{N} Z_j$, where $N \sim P(\lambda)$ and $Z_1, \ldots, Z_N$ i.i.d. random vectors and independent of $N$. Consider now for every $n \in \mathbb{N}$,*

$$N^{(n)} \sim P(\lambda/n)$$

*so that*

$$Y^{(n)} = \sum_{j=1}^{N^{(n)}} Z_j,$$

*where $N_1^{(n)}, \ldots, N_n^{(n)}$ are independent and identically distributed as $P(\lambda/n)$, and $\{Z_{i,j}, i \geq 1, j = 1, \ldots, n\}$ are i.i.d. sequences, each one composed by independent and identically distributed random variables independent of $N_j^{(n)}$. Then*

$$Y_j^{(n)} = \sum_{i=1}^{N_j^{(n)}} Z_{i,j}, \quad j = 1, \ldots, n$$

*and*

$$\sum_{j=1}^{n}\sum_{i=1}^{N_j^{(n)}} Z_{i,j} \stackrel{d}{=} \sum_{j=1}^{N_1^{(n)}+\cdots+N_n^{(n)}} Z_i \stackrel{d}{=} Y$$

*The characteristic function of $Y$ is*

$$\begin{aligned}\phi_Y(u) &= E[e^{i(u,Y)}] = E\left[e^{i(\sum_{j=1}^{N} u,Z_j)}\right] = E\left[E\left(e^{i\sum_{j=1}^{N}(u,Z_j)}|N\right)\right] = \\ &= E\left[E\left(e^{i(u,Z_1)}\right)^N\right]\end{aligned}$$

*By denoting with $r = E\left[e^{i(\theta,Z_1)}\right]$ and with $P_Z$ the common distribution of $Z_j, j = 1,\dots,n$, we obtain*

$$\begin{aligned}E\left[r^N\right] &= \sum_{i=1}^{\infty}\frac{r^i\lambda^i}{i!}\exp(-\lambda) = \exp(-\lambda)\sum_{i=1}^{\infty}\frac{(\lambda r)^i}{i!}\exp\{-\lambda(1-r)\} = \\ &= \exp\{-\lambda(1-E[\exp\{i(u,Z_1)\}])\} = \\ &= \exp\left\{-\lambda\int_{\mathbb{R}^d}(1-\exp\{i(u,z)\})P_Z(dz)\right\} = \\ &= \exp\left\{-\int_{\mathbb{R}^d}(1-\exp\{i(u,z)\})\lambda P_Z(dz)\right\} =\end{aligned}$$

*and setting $\lambda P_Z(dz) = \nu(dz)$ we have*

$$= \exp\left\{-\int_{\mathbb{R}^d}(1-\exp\{i(u,z)\})\nu(dz)\right\}$$

*where $\nu(\cdot)$ is a finite measure (since it is a probability measure multiplied by a finite constant).*

Also, if we consider two independent infinitely divisible random vectors $Y_1$ and $Y_2$, then it can be easily proven that the random vector $Y = Y_1 + Y_2$ is still infinitely divisible. For instance considering $Y_1$ as in example 1 and $Y_2$ as in example 2 then

$$\begin{aligned}\phi_Y(u) &= \phi_{Y_1}(u)\phi_{Y_2}(u) = \\ &= \exp\left\{-\frac{1}{2}u\Sigma u' + i(\mu,u) - \int_{\mathbb{R}^d}(1-\exp\{i(u,z)\})\nu(dz)\right\}\end{aligned}$$

This last expression is very close to the general expression of an infinitely divisible random vector which we will see hereafter.

Here we introduce two more interesting properties of distributions underlying evantual Lévy processes . Let $\phi_\mu(u)$ like in definition 6:

**Definition 9.** *Let $\mu$ be a probability measure on $\mathbb{R}^d$. It is called* selfdecomposable*, or* of class *L, if for any $b > 1$ there is a probability measure $\rho_b$ on $\mathbb{R}^d$ such that*

$$\phi_\mu(u) = \phi_\mu(u/b)\phi_{\rho_b}(u). \tag{1.2}$$

*It is called* semi selfdecomposable *if there are some $b > 1$ and some infinitely divisible measure $\rho_b$ satisfying* (1.2).

**Proposition 2.** *If $\mu$ is selfdecomposable then is infinitely divisible and for any $b > 1$ $\rho_b$ is uniquely determined and infinitely divisible.*
*If $\mu$ is semi selfdecomposable, then $\mu$ is infinitely divisible and $\rho_b$ is uniquely determined.*

*Proof*: see Sato [Sat99], pg. $93$.

The relation between possible representations of selfdecomposable distributions is given in a work by Jeanblanc, Pitman and Yor [JPY02].
A final characterization of selfdecomposable distributions is given by the following

**Proposition 3.** *A distribution $\mu$ is selfdecomposable if and only if for any fixed $H > 0$ it is the distribution of $X(1)$, where $\{X(t), t \geq 0\}$ is a process with the following properties:*

- *it is additive, i.e. (L1), (L2) and (L4) hold,*
- *it is $H-$self-similar, meaning that for each $c > 0$*

$$\{X(ct),\ t \geq 0\} \stackrel{d}{=} \{c^H X(t),\ t \geq 0\}.$$

### 1.3.2 Lévy-Khinchine formula

Here is the formula, first established by P.Lévy and A.Y.Khintchine in the late $1930$'s, which gives a characterization of infinitely divisible random variables through their characteristic functions. First a definition is needed.

**Definition 10.** *Let $\nu$ be a Borel measure defined on $\mathbb{R}^d \setminus \{0\} = \{x \in \mathbb{R}^d : x \neq 0\}$. We say that $\nu$ is a* Lévy measure *if*

$$\int_{\mathbb{R}^d \setminus \{0\}} (1 \wedge |y|^2) \nu(dy) < \infty \tag{1.3}$$

*Since $|y|^2 \wedge \varepsilon \leq |y|^2 \wedge 1$ whenever $0 < \varepsilon \leq 1$, it follows from (1.3) that*

$$\nu\{(-\varepsilon, \varepsilon)^c\} < \infty, \quad \textit{for all } \varepsilon > 0.$$

Moreover it is easy to prove that every Lévy measure on $\mathbb{R}^d \setminus \{0\}$ is $\sigma-$finite. Note also that any finite measure on $\mathbb{R}^d \setminus \{0\}$ is a Lévy measure. The result given below is usually called the Lévy-Khintchine formula and it is the cornerstone for much of what follows.

**Theorem 6.** (Lévy-Khinchine representation) $\mu \in \mathcal{M}_1(\mathbb{R}^d)$ *is infinitely divisible if there exists a vector $b \in \mathbb{R}^d$, a positive definite symmetric $d \times d$ matrix $A$ and a Lévy measure $\nu$ on $\mathbb{R}^d \setminus \{0\}$ such that, for all $u \in \mathbb{R}^d$,*

$$\begin{aligned} \phi_\mu(u) = \exp \Big\{ & i(b, u) - \frac{1}{2}(u, Au) + \\ & + \int_{\mathbb{R}^d \setminus \{0\}} \left[ e^{i(u,y)} - 1 - i(u, y) \mathbb{1}_{\hat{B}}(y) \right] \nu(dy) \Big\}, \end{aligned} \tag{1.4}$$

*where $\hat{B} = B_1(0)$, an $\mathbb{R}^d-$ball of radius $1$ around $0$.*

*Conversely, any mapping of the form* (1.4) *is the characteristic function of an infinitely divisible probability measure on $\mathbb{R}^d$.*

*Proof*: see for instance Sato, [Sat99].

Observations and remarks

1. The proof of the "only if " part is particularly elaborate. See [Sat99], pgg. $41-45$, for one way of doing this. An alternative approach will be given in the following, as a byproduct of the Lévy-Itô decomposition.

2. The choice of "cut-off" function $c(y) = y\mathbb{1}_B$ that occurs within the integral in (1.4) is arbitrary.

A replacement that is often used is

$$c(y) = \frac{y}{1 + |y|^2}.$$

The only constraint in choosing $c$ is that the function $g_c(y) = e^{i(u,y)} - 1 - i(c(y), u)$ should be $\nu-$integrable for each $u \in \mathbb{R}^d$ .

Adopting a different $c$ forces to change the vector $b$ accordingly in (1.4).

3. Relative to the choice of $c$ that we have taken, the members of the triple $(b, A, \nu)$ are called the *characteristics* of the infinitely divisible random vector $X$. Examples of these are:

   - *Gaussian case*: $b$ is the mean, $A$ is the covariance matrix, $\nu = 0$.
   - *Poisson case*: $b = 0$, $A = 0$, $\nu = c\delta_1$, with $\delta_1$ the Dirac measure with mass on $\{1\}$.
   - *Compound Poisson case*: $b = 0,\ A = 0,\ \nu = c\mu$, where $c > 0$ and $\mu$ is a probability measure on $\mathbb{R}^d$.

The characteristic function $\phi_\mu(u)$ is often written as $e^{\psi(u)}$: $\psi$ is often referred to as the *characteristic exponent* or Lévy exponent.

Here we state the relationship between Lévy processes and infinite divisibility; from this point on, where omitted, for the proofs see for instance Sato [Sat99] or Appelbaum [App04]:

**Proposition 4.** *If $X$ is a Lévy process , then $X(t)$ is infinitely divisible for each $t \geq 0$.*

*Proof: for each $n \in \mathbb{N}$ it is possible to write*

$$X(t) = Y_1^{(n)}(t) + \ldots + Y_n^{(n)}(t),$$

*where each*

$$Y_k^{(n)}(t) = X(kt/n) - X((k-1)t/n)$$

*The $Y_k^{(n)}(t)$ are i.i.d by* (L2) *and* (L3). □

By Proposition 4, we can write $\phi_{X(t)}(u) = e^{\psi(t,u)}$ for each $t \geq 0,\ u \in \mathbb{R}^d$ , where each $\psi(t,\cdot)$ is a characteristic exponent. We will see below that $\psi(t,u) = t\psi(1,u)$ for each $t \geq 0,\ u \in \mathbb{R}^d$ , but first the following lemma is needed:

**Lemma 1.** *If $X = \{X(t), t \geq 0\}$ is stochastically continuous, then the map $t \to \phi_{X(t)}(u)$ is continuous for each $u \in \mathbb{R}^d$ .*

**Theorem 7.** *If $X$ is a Lévy process , then*

$$\phi_{X(t)}(u) = e^{t\psi(u)}$$

*for each $u \in \mathbb{R}^d$, $t \geq 0$, where $\psi$ is the characteristic exponent of $X(1)$.*

We now have the Lévy-Khinchine formula for a Lévy process $X = \{X(t), t \geq 0\}$ ,

$$E[e^{i(u,X(t))}] = \exp\left(t\left[i(b,u) - \frac{1}{2}(u,Au) + \right.\right.$$
$$\left.\left. + \int_{\mathbb{R}^d\setminus\{0\}} \left[e^{i(u,y)} - 1 - i(u,y)\mathbb{I}_{\hat{B}}(y)\right] \nu(dy)\right]\right)$$

for each $t \geq 0$, $u \in \mathbb{R}^d$, where $(b, A, \nu)$ are the characteristics of $X(1)$.
We will define the characteristic exponent and the characteristics of a Lévy process $X$ to be those of the random variable $X(1)$.
It is easy to show that the sum of two independent Lévy processes is again a Lévy process

**Theorem 8.** *If $X = \{X(t), t \geq 0\}$ is a stochastic process and there exists a sequence of Lévy processes $\{X_n, n \in \mathbb{N}\}$ with each $X_n = \{X_n(t), t \geq 0\}$ such that $X_n(t)$ converges in probability to $X(t)$ for each $t \geq 0$, and*

$$\lim_{n\to\infty} \limsup_{t\to 0} P(|X_n(t) - X(t)| > a) = 0$$

*for all $a > 0$, then $X$ is a Lévy process .*

## 1.4 Subordinators

Subordinators introduce a simple procedure of transforming a stochastic process into another stochastic process through random time change by an increasing Lévy process independent of the original starting process. The idea was first introduced by Bochner in 1949. Subordination is also a good auxiliary method to reduce complex processes to known ones by means of a suitable time change.

**Definition 11.** *A* subordinator *is a one-dimensional Lévy process that is a.s. non-decreasing.*

Such processes can be thought of as a random model of time evolution, since if $T = \{T(t), t \geq 0\}$ is a subordinator we have

$$T(t) \geq 0 \; a.s., \text{for each } t > 0,$$

and

$$T(t_1) \leq T(t_2) \; a.s., \text{whenever } t_1 \leq t_2.$$

As a counterexample, the process $X = \{X(t), t \geq 0, X(t) \sim N(0, At)\}$ is such that $P(X(t) \geq 0) = P(X(t) \leq 0) = 1/2$, therefore it is clear that such a process cannot be a subordinator. More generally we have

**Theorem 9.** *If $T$ is a subordinator, then its characteristic exponent takes the form*

$$\psi(u) = ibu + \int_0^\infty (e^{iuy} - 1)\lambda(dy), \tag{1.5}$$

*where $b \geq 0$ and the Lévy measure $\lambda$ satisfies the additional requirements*

$$\lambda(-\infty, 0) = 0 \quad \textit{and} \quad \int_0^\infty (y \wedge 1)\lambda(dy) < \infty.$$

*Conversely, any mapping from $\mathbb{R}^d \to \mathbb{C}$ of the form (1.5) is the characteristic exponent of a subordinator.*

We call the pair $(b, \lambda)$ the characteristics of the subordinator $T$ .

# 1.5 Lévy measure, jumps of a Lévy process and Poisson random measures

Let us understand what the condition (1.3) say; firstly it is sufficiently clear that

$$\nu(|x| \geq 1) < \infty,$$

and moreover

$$\int_{(-1,1)} x^2 \nu(dx) < \infty \tag{1.6}$$

These two conditions are sufficient to ensure that integral (1.4) converges since the integrand is $O(1)$ for $|x| \geq 1$ and $O(x^2)$ for $|x| < 1$.

In principle (1.6) means that two possibilities could show up:

- $\nu(-1, 1) < \infty$;
- $\nu(-1, 1) = \infty$. In this latter case, it necessarily holds that $\nu(|x| \in (\varepsilon, 1)) < \infty$, but $\nu(|x| < \varepsilon) = \infty$, for $0 < \varepsilon < 1$.

The Lévy measure $\nu$ describes the size and rate of arrival of jumps of the Lévy process $X(t)$. It could be naively explained like in a small period of time $dt$ a jump of size $x$ occurs with probability $\nu(dx)dt + o(dt)$. If it were the case that $\nu(-1, 1) = \infty$ then the latter interpretation would suggest that the smaller the jump size, the greater the intensity and so the discontinuities in the path of a Lévy process are predominantly made up of arbitrarily small jumps.

## 1.5.1 Lévy measure, paths and distributional properties

The Lévy measure is responsible for the richness of the class of Lévy processes and carries useful information about the structure of the process. Also path properties can be read from the Lévy measure.

**Proposition 5.** *Let $X$ be a Lévy process on $\mathbb{R}$ with triplet $(a, b, \nu)$.*

*i) If $\nu(\mathbb{R}) < \infty$, then almost all paths of $X$ have a finite number of jumps on every compact interval. In this case, the Lévy process has* finite activity.

*ii) If $\nu(\mathbb{R}) = \infty$, then almost all paths of $X$ have an infinite number of jumps on every compact interval. In this case, the Lévy process has* infinite activity.

*Proof*: see theorem 21.3 in Sato [Sat99].

Whether a Lévy process has finite variation or not, also depends on the Lévy measure (and on the presence or absence of a Brownian part).

**Proposition 6.** *Let $X$ be a Lévy process with triplet $(a, b, \nu)$.*

*i) If $b = 0$ and $\int_{|x|\leq 1} |x| \nu(dx) < \infty$, then almost all paths of $X$ have finite variation.*

*ii) If $b \neq 0$ or $\int_{|x|\leq 1} |x| \nu(dx) = \infty$, then almost all paths of $X$ have infinite variation.*

*Proof*: see theorem 21.9 in Sato [Sat99].

The Lévy measure also carries information about the finiteness of the moments of a Lévy process. This is particularly useful information in mathematical finance, related to the existence of a martingale measure. The finiteness of the moments of a Lévy process is related to the finiteness of an integral over the Lévy measure (more precisely, the restriction of the Lévy measure to jumps larger than $1$ in absolute value, i.e. "big jumps").

**Proposition 7.** *Let $X$ be a Lévy process with triplet $(a, b, \nu)$.*

*i) $X(t)$ has finite $p$-th moment for $p \in \mathbb{R}_+$ (meaning that $E[|X(t)|^p] < \infty$) if and only if $\int_{|x|\geq 1} |x|^p \nu(dx) < \infty$*

*ii) $X(t)$ has finite $p$-th exponential moment for $p \in \mathbb{R}$ ($E[e^{pX(t)}] < \infty$) if and only if $\int_{|x|\geq 1} e^{px} \nu(dx) < \infty$*

*Proof*: the proof of these results can be found in theorem $25.3$ in Sato [Sat99].

Actually, the conclusion of this theorem holds for the general class of submultiplicative functions (cf. def. $25.1$ in Sato [Sat99]), which contains $\exp(px)$ and $|x|^p \vee 1$ as special cases.

**Definition 12.** *A function $f : (0, \infty) \to (0, \infty)$ is called* regularly varying *with exponent $\beta$ if $f(t) > 0$, for $t$ large enough and, for any $c > 0$*

$$\lim_{t\to\infty} \frac{f(ct)}{f(t)} = c^{\beta}.$$

*If $\beta = 0$, the function $f$ is said to be* slowly varying.

**Theorem 10.** *Let $Y$ be an infinitely divisible random variable with Lévy measure $\nu$. Then $Y$ has a* regularly varying tail *with exponent $\alpha$ if and only if*

$$\nu\{z : z > t\}$$

*is regularly varying with exponent $\alpha$. If this is true, moreover*

$$\lim_{t\to\infty} \frac{P(Y > t)}{\nu\{z : z > t\}} = 1.$$

A property which often appears in financial data is the following

**Definition 13.** *A distribution function $f(x)$ has* semiheavy tails *if the tails of the distribution behave like*

$$f(x) \sim \begin{cases} C_-|x|^{\rho_-} \exp(\eta_-|x|) & \textit{as } x \to -\infty, \\ C_+|x|^{\rho_+} \exp(\eta_+|x|) & \textit{as } x \to \infty \end{cases}$$

*for some $\rho_+,\ \rho_- \in \mathbb{R}$ and $C_-,\ C_+,\ \eta_-,\ \eta_+ \geq 0$.*

For a couple of observations, we introduce a very important process associated to a Lévy process $X$ .

**Definition 14.** *The* jump process $\Delta X = \{\Delta X(t), t \geq 0\}$ *is defined by*

$$\Delta X(t) = X(t) - X(t^-), \textit{ for each } t \geq 0$$

$X(t^-)$ is the left limit at the point $t$; clearly $\Delta X$ is an adapted process but it is not, in general, a Lévy process .

**Lemma 2.** *If $X$ is a Lévy process , then, for fixed $t > 0$, $\Delta X(t) = 0$ $a.s.$.*

*Proof*: let $\{t(n), n \in \mathbb{N}\}$ be a sequence in $\mathbb{R}^+$ with $t(n) \uparrow t$ as $n \to \infty$; then, since $X$ has cádlág paths, $\lim_{n\to\infty} X(t(n)) = X(t^-)$. However, by (L4) the sequence $\{X(t(n)), n \in \mathbb{N}\}$ converges in probability to $X(t)$ and so has a subsequence that converges almost surely to $X(t)$. The result follows by uniqueness of limits. □

Much of the analytic difficulty in manipulating Lévy processes arises from the fact that it is possible for them to have

$$\sum_{0\leq s\leq t} |X(s)| = \infty \quad a.s.$$

and the way these difficulties are overcome exploits the fact that we always have

$$\sum_{0\leq s\leq t} |X(s)|^2 < \infty \quad a.s.$$

But, rather than exploring $\Delta X$ itself further, it is generally more profitable to count jumps of specified size. More precisely, let $0 \leq t < \infty$, and $A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$. Define

$$N(t, A) = \sharp\{0 \leq s \leq t : \Delta X(s) \in A\} = \sum_{0\leq s\leq t} 1\!\!1_A(\Delta X(s)).$$

Note that for each $\omega \in \Omega$ and $t \geq 0$, the set function $A \to N(t, A)(\omega)$ is a counting measure on $\mathcal{B}(\mathbb{R}^d \setminus \{0\})$ and hence

$$E[N(t, A)] = \int N(t, A)(\omega) dP(\omega)$$

is a Borel measure on $\mathcal{B}(\mathbb{R}^d \setminus \{0\})$. We write $\mu(\cdot) = E[N(1, \cdot)]$ and call it the *intensity measure* associated with $X$ .

**Definition 15.** *We say that $A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ is* bounded below *if* $0 \notin \bar{A}$.

**Lemma 3.** *If $A$ is bounded below, then $N(t, A) < \infty$ $a.s.$ for all $t \geq 0$.*

*Proof*: see Appelbaum, [App04], pg. 87.

Note that if $A$ fails to be bounded below then lemma 3 may no longer hold, because of the accumulation of infinite numbers of small jumps.

**Theorem 11.**

(1) *If $A$ is bounded below, then $\{N(t,A), t \geq 0\}$ is a Poisson process with intensity $\mu(A)$.*

(2) *If $A_1, \dots, A_m \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ are disjoint, then the random variables $N(t,A_1), \dots, N(t,A_m)$ are independent.*

*Proof*: see Appelbaum, [App04], pg. 88.

## 1.6 Random measures

**Definition 16.** *Let $(S, \mathcal{A})$ be a measurable space and $(\Omega, \mathscr{F}, \mathbf{F}, P)$ be a stochastic basis. A* random measure *$M$ on $(S, \mathcal{A})$ is a collection of random variables $\{M(B),\ B \in \mathcal{A}\}$ such that:*

(1) $M(\emptyset) = 0$*;*

(2) ($\sigma-$additivity) *given any sequence $\{A_n, n \in \mathbb{N}\}$ of mutually disjoint sets in $\mathcal{A}$,*

$$M\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} M(A_n) \ a.s.;$$

(3) (independently scattered property) *for each disjoint family $\{B_1, \dots, B_n\}$ in $\mathcal{A}$, the random variables $M(B_1), \dots, M(B_n)$ are independent.*

**Definition 17.** *We say that we have a* Poisson random measure *if each $M(B)$ has a Poisson distribution whenever $M(B) < \infty$.*

In many cases of interest, we obtain a $\sigma-$finite measure $\lambda$ on $(S, \mathcal{A})$ by the prescription $\lambda(A) = E[M(A)]$ for all $A \in \mathcal{A}$. Conversely we have:

**Theorem 12.** *Given a $\sigma-$finite measure $\lambda$ on a measurable space $(S, \mathcal{A})$, there exists a Poisson random measure $M$ on a probability space $(\Omega, \mathcal{F}, P)$ such that $\lambda(A) = E[M(A)]$ for all $A \in \mathcal{A}$.*

*Proof*: See e.g. [Sat99], pg. 122. □

Suppose that $S = \mathbb{R}^+ \times U$, where $U$ is a measurable space equipped with a $\sigma-$algebra $\mathcal{C}$ , and $A = \mathcal{B}(R^+) \otimes \mathcal{C}$ . Let $p = \{p(t), t \geq 0\}$ be an adapted process taking values in $U$ such that $M$ is a Poisson random measure on $S$ , where $M([0,t) \times A) = \sharp\{0 \leq s < t : p(s) \in A\}$ for each $t \geq 0$ and $A \in \mathcal{C}$. In this case we say that $p$ is a *Poisson point process* and $M$ is its associated Poisson random measure.

The next necessary concept is a merger of the two important ideas of the *random measure* and the *martingale*.
Let $U$ be a topological space and take $\mathcal{C}$ to be its Borel $\sigma-$algebra. Let $M$ be a random measure on $S = \mathbb{R}^+ \times U$. For each $A \in \mathcal{C}$, define a process $M_A = \{M_A(t), t \geq 0\}$ by $M_A(t) = M([0,t) \times A)$.

**Definition 18.** *We say that $M$ is a* martingale-valued measure *if there exists $V \in \mathcal{C}$ such that $M_A$ is a martingale whenever $\bar{A} \cap V = \emptyset$. We call $V$ the associated* forbidden set *(which may of course itself be $\emptyset$).*

The key example of these concepts for our work is as follows:

**Example 3.** *Let $U = \mathbb{R}^d \setminus \{0\}$ and $\mathcal{C}$ be its Borel $\sigma-$algebra. Let $X$ be a Lévy process; then $\Delta X$ is a Poisson point process and $N$ is its associated Poisson random measure.*

**Definition 19.** *For each $t \geq 0$ and $A$ bounded below, we define the* compensated Poisson random measure *by*

$$\tilde{N}(t, A) = N(t, A) - t\mu(A).$$

$\{\tilde{N}(t,A), t \geq 0\}$ is a martingale and so $\tilde{N}$ extends to a martingale-valued measure with forbidden set $\{0\}$.

Here are the main properties of the Poisson random measure $N$:

(1) For each $t > 0,\ \omega \in \Omega,\ N(t,\cdot)(\omega)$ is a counting measure on $\mathcal{B}(\mathbb{R}^d \setminus \{0\})$.

(2) For each $A$ bounded below, $\{N(t,A),\ t \geq 0\}$ is a Poisson process with intensity $\mu(A) = E[N(,A)]$.

(3) $\{\tilde{N}(t,A), t \geq 0\}$ is a martingale-valued measure, where $\tilde{N}(t,A) = N(t,A) -\ t\mu(A)$, for $A$ bounded below.

### 1.6.1 Poisson integration

Let $f : \mathbb{R}^d \to \mathbb{R}^d$ be a Borel measurable function and let $A$ be bounded below; then for each $t > 0,\ \omega \in \Omega$, we may define the *Poisson integral* of $f$ as a random finite sum by

$$\int_A f(x)N(t,dx)(\omega) = \sum_{x \in A} f(x)N(t,\{x\})(\omega).$$

Note that each $\int_A f(x)N(t,dx)$ is an $\mathbb{R}^d$-valued random variable and gives rise to a cádlág stochastic process as we vary $t$ .

Now, since $N(t,\{x\}) = 0 \Leftrightarrow X(u) = x$ for at least one $0 \leq u \leq t$, we have

$$\int_A f(x)N(t,dx) = \sum_{0 \leq u \leq t} f(\Delta X(u))\mathbb{1}_A(\Delta X(u)) \tag{1.7}$$

Let $\{T_n^A, n \in \mathbb{N}\}$ be the arrival times for the Poisson process $\{N(t,A),\ t \geq 0\}$. Then another useful representation for Poisson integrals, which follows immediately from (1.7), is

$$\int_A f(x)N(t,dx) = \sum_{n \in \mathbb{N}} f(\Delta X(T_n^A))\mathbb{1}_{[0,t]}(T_n^A). \tag{1.8}$$

From this, we will sometimes use $\mu_A$ to denote the restriction to $A$ of the measure $\mu$.

**Theorem 13.** *Let $A$ be bounded below. Then:*

(1) *for each* $t \geq 0,\ \int_A f(x)N(t,dx)$ *has a compound Poisson distribution such that, for each* $u \in \mathbb{R}^d$,

$$E\left[\exp\left\{i\left(u, \int_A f(x)N(t,dx)\right)\right\}\right] = \exp\left[t\int_A (e^{i(u,x)}-1)\mu_f(dx)\right]$$

*where* $\mu_f = \mu \circ f^{-1}$;

(2) *if* $f \in L^1(A,\mu_A)$, *we have*

$$E\left[\int_A f(x)N(t,dx)\right] = t\int_A f(x)\mu(dx);$$

(3) *if* $f \in L^2(A,\mu_A)$, *we have*

$$Var\left(\left|\int_A f(x)N(t,dx)\right|\right) = t\int_A |f(x)|^2\mu(dx).$$

*Proof*: see for instance Appelbaum, [App04], pg. 92.

It follows from theorem 13 (2) that a Poisson integral will fail to have a finite mean if $f \notin L^1(A,\mu)$.

Consider the sequence of jump size random variables $\{Y_f^A(n), n \in \mathbb{N}\}$, where each

$$Y_f^A(n) = \int_A f(x)N(T_n^A, dx) - \int_A f(x)N(T_{n-1}^A, dx).$$

It follows from (1.8) and (5.20) that

$$Y_f^A(n) = f(\Delta X(T_n^A)),$$

for each $n \in \mathbb{N}$ .

**Theorem 14.**

(1) $\{Y_f^A(n), n \in \mathbb{N}\}$ *are i.i.d. with common law given by*

$$P(Y_f^A(n) \in B) = \frac{\mu(A \cap f^{-1}(B))}{\mu(A)}$$

*for each* $B \in \mathcal{B}(\mathbb{R}^d)$.

(2) $\{\int_A f(x)N(t,dx), t \geq 0\}$ *is a compound Poisson process.*

*Proof*: See e.g. [App04], pg. $93-4$.

**Definition 20.** *For each* $f \in L^1(A, \mu_A),\ t \geq 0$*, we define the* compensated Poisson integral *by*

$$\int_A f(x)\tilde{N}(t,dx) = \int_A f(x)N(t,dx) - t\int_A f(x)\mu(dx).$$

A straightforward argument shows that

$$\left\{\int_A f(x)\tilde{N}(t,dx),\ t \geq 0\right\}$$

is a martingale. By theorem 13 (1) and (3) we can easily deduce the following two important facts:

$$\begin{aligned} &E\left[\exp\left\{i\left(u, \int_A f(x)\tilde{N}(t,dx)\right)\right\}\right] = \\ &= \exp\left\{t\int_A [e^{i(u,x)} - 1 - i(u,x)]\mu_f(dx)\right\} \end{aligned} \tag{1.9}$$

for each $u \in \mathbb{R}^d$and, for $f \in L^2(A, \mu_A)$,

$$E\left[\left|\int_A f(x)\tilde{N}(t,dx)\right|^2\right] = t\int_A |f(x)|^2\mu(dx).$$

## 1.7 Lévy-Itô decomposition

Here is one of the key results in the elementary theory of Lévy process , namely the celebrated Lévy-Itô decomposition of the sample paths into continuous and jump parts. Some preliminary results are needed for the proof of the main result. Most of the auxiliary proof will be omitted; for details see [App04].

**Proposition 8.** *Let* $M_j,\ j = 1,2$ *be two cádlág centered martingales. Suppose that, for some* $j$*,* $M_j$ *is* $L^2$ *and that for each* $t \geq 0,\ E[|V(M_k(t)|^2) < \infty$ *where* $k \neq j$ *and for a cádlág mapping*

*$g : [a,b] \to \mathbb{R}^d$ with $\mathcal{P}$ a partition of interval $[a,b]$ of the form $[a = t_1 < t_2 < \ldots, < t_n < t_{n+1} = b]$ in $\mathbb{R}$, $V(g) = \sup_{\mathcal{P}} \sum_{i=1}^{n} |g(t_{i+1}) - g(t_i)|$; then*

$$E[(M_1(t), M_2(t))] = E\Big[\sum_{0\leq s\leq t} (\Delta M_1(s), \Delta M_2(s))\Big].$$

**Definition 21.** (from Schoutens, [Sch03]) *Let $\mathcal{P}$ be a partition of the interval $[a,b]$ as above introduced and $g : [a,b] \to \mathbb{R}$ cádlág. If $V(g) < \infty$ we say that $g$ has* finite variation *on $[a,b]$. If this is not the case, the function is said to be* of infinite variation.

**Definition 22.** (from Eberlein, [Ebe09]) *Let $X = \{X(t), t \geq 0\}$ be a Lévy process with Lévy measure $\nu$. Then $X$ has* finite activity *if $\nu(\mathbb{R}) < \infty$, otherwise it has* infinite activity.

The following is a special case of proposition above, which plays a major role below.

**Example 4.** *Let $A$ and $B$ be bounded below and suppose that $f \in L^2(A, \mu_A),\ g \in L^2(B, \mu_B)$. For each $t \geq 0$, let $M_1(t) = \int_A f(x)\tilde{N}(t,dx)$ and*
*$M_2(t) = \int_B g(x)\tilde{N}(t,dx)$; then*

$$V(M_1(t)) \leq V\left(\int_A f(x)N(t,dx)\right) + V\left(t\int_A f(x)\nu(dx)\right) \leq$$
$$\leq \int_A |f(x)|N(t,dx) + t\int_A f(x)\nu(dx).$$

*From this and the Cauchy-Schwarz inequality we have $E[|V(M_1(t)|^2] < \infty$, and so we can apply proposition above in this case.*

Observe that $E[M_1(t)M_2(t)] = 0$ for each $t \geq 0$ if $A \cap B = \emptyset$.
Moreover it can be shown that proposition above fails to hold when $M_1 = M_2 = B$ , where $B$ is a standard Brownian motion.

**Theorem 15.** *If $A_p,\ p = 1,2$ are disjoint and bounded below, then*
*$\left\{\int_{A_1} xN(t,dx), t \geq 0\right\}$ and $\left\{\int_{A_2} xN(t,dx), t \geq 0\right\}$ are independent stochastic processes.*

**Theorem 16.** *If X is a Lévy process with bounded jumps then we have $E[|X(t)|^m] < \infty$ for all $m \in \mathbb{N}$.*

For each $a > 0$, consider the compound Poisson process

$$\left\{\int_{|x|\geq a} xN(t,dx), t \geq 0\right\}$$

and define a new stochastic process $Y_a = \{Y_a(t),\ t \geq 0\}$ by the prescription

$$Y_a(t) = X(t) - \int_{|x|\geq a} xN(t,dx).$$

Intuitively, $Y_a$ is what remains of the Lévy process $X$ when all the jumps of size greater than $a$ have been removed. We can get more insight into its paths by considering the impact of removing each jump. Let $\{T_n, n \in \mathbb{N}\}$ be the arrival times for the Poisson process $\{N(t, B_a(0)^c), t \geq 0\}$. Then we have

$$Y_a(t) = \begin{cases} X(t) & \text{for } 0 \leq t < T_1, \\ X(T_1^-) & \text{for } t = T_1, \\ X(t) - X(T_1) + X(T_1^-) & \text{for } T_1 < t < T_2, \\ Y_a(T_2^-) & \text{for } t = T_2, \end{cases}$$

and so on recursively.

**Theorem 17.** *$Y_a$ is a Lévy process .*

*Proof*: (L1) is immediate. For (L2) and (L3) we argue as in the proof of theorem 11 and deduce that, for each $0 \leq s < t < \infty$, $Y_a(t) - Y_a(s)$ is $\mathscr{F}_{s,t}$−measurable where $\mathscr{F}_{s,t} = \sigma\{X(u) - X(v) : s \leq v \leq u < t\}$. To establish (L4), use the fact that for each $b > 0,\ t \geq 0$,

$$P(|Y_a(t)| > b\} \leq P\left(|X(t)| > b/2\right) + P\left(\left|\int_{|x|\geq a} xN(t,dx)\right| > b/2\right). \qquad \square$$

It is immediately deduced the following:

**Corollary 1.** *A Lévy process has bounded jumps if and only if it is of the form $Y_a$ for some $a > 0$.*

For each $a > 0$, we define a Lévy process $\hat{Y}_a = \{\hat{Y}_a(t),\ t \geq 0\}$ by

$$\hat{Y}_a = Y_a(t) - E[Y_a(t)].$$

It is then easy to verify that $\hat{Y}_a$ is a cádlág centred $L^2-$martingale. In the following, it will be convenient to take $a = 1$ and write the processes $Y_1, \hat{Y}_1$ simply as $Y,\ \hat{Y}$ , respectively. So $Y$ is what remains of our Lévy process when all jumps whose magnitude is larger than 1 have been removed, and $\hat{Y}$ is the centered version of $Y$. We also introduce the notation $M(t, A) = \int_A x\tilde{N}(t, dx)$ for $t \geq 0$ and $A$ bounded below.
The following is a key step towards our required result.

**Theorem 18.** *For each $t \geq 0$,*

$$\hat{Y}(t) = Y_c(t) + Y_d(t),$$

*where $Y_c$ and $Y_d$ are independent Lévy process , $Y_c$ has continuous sample paths and*

$$Y_d(t) = \int_{|x|<1} x\tilde{N}(t, dx).$$

We recall that $\mu$ is the intensity measure of the Poisson random measure $N$.

**Corollary 2.** *$\mu$ is a Lévy measure.*

*Proof*: it has been already shown that $\mu\{(-1, 1)^c\} < \infty$. We also have

$$\int_{|x|\leq 1} |x|^2 \mu(dx) = \lim_{n\to\infty} \int_{A_n} |x|^2 \mu(dx) = \lim_{n\to\infty} E[|M(t, A_n|)^2] =$$
$$= E[|Y_d|^2] < \infty,$$

and the result is established. □

**Corollary 3.** *For each* $t \geq 0$, $u \in \mathbb{R}^d$,

$$E[e^{i(u,Y_d(t))}] = \exp\left\{t\int_{|x|<1}[e^{i(u,x)} - 1 - i(u,x)]\mu(dx)\right\}.$$

*Proof*: Take limits in equation (1.9). □

**Theorem 19.** $Y_c$ *is a Brownian motion.*

**Theorem 20.** (Lévy-Itô decomposition) *If* $X$ *is a Lévy process, then there exists* $b \in \mathbb{R}^d$, *a Brownian motion* $B_A$ *with covariance matrix* $A$ *and an independent Poisson random measure* $N$ *on* $\mathbb{R}^+ \times (\mathbb{R}^d \setminus \{0\})$ *such that, for each* $t \geq 0$,

$$X(t) = bt + B_A(t) + \int_{|x|<1} x\tilde{N}(t,dx) + \int_{|x|\geq 1} xN(t,dx). \tag{1.10}$$

*Proof*: this follows from theorems (18) and (19) with

$$b = E\left[X(1) - \int_{|x|\geq 1} xN(1,dx)\right].$$

The fact that $B_A$ and $N$ are independent follows from the argument of theorem 15. □

Observe that the process $\{\int_{|x|<1} x\tilde{N}(t,dx),\ t \geq 0\}$ in (1.10) is the compensated sum of small jumps. The compensation takes care of the analytic complications in the Lévy-Khintchine formula in a probabilistically pleasing way, since it is an $L^2-$martingale.

The process $\{\int_{|x|\geq 1} xN(t,dx),\ t \geq 0\}$ describing the "large jumps" in (1.10) is a compound Poisson process by theorem (14).

Moreover, it holds the following

**Corollary 4.** *The characteristics* $(b, A, \nu)$ *of a Lévy process are uniquely determined by the process.*

*Proof*: this follows from the construction that led to theorem 20.

## 1.8 Semimartingales

Let $(\Omega, \mathbf{F}, P)$ be the usual filtered probability space:

**Definition 23.**

(1) *An adapted stochastic process* $X = \{X(t), t \geq 0\}$ *is a* local martingale *if there is a nondecreasing sequence of stopping times* $T_1 \leq T_2 \leq \ldots$ *such that for every* $n \leq 1$ stopped process

$$\{X(t \wedge T_n), t \geq 0, n \geq 1\}$$

*is a martingale.*

(2) *An adapted stochastic process* $X = \{X(t), t \geq 0\}$ *is a* semimartingale *if it can be written in the form*

$$X(t) = M(t) + A(t),\ t \geq 0$$

*where* $\{M(t),\ t \geq 0\}$ *is a local martingale, and* $\{A(t),\ t \geq 0\}$ *is an adapted stochastic process of a local bounded variation.*

The basic importance of semimartingales is given from the fact that on this family of stochastic processes it is possible to develop an unified theory of stochastic integration, i.e. to give sense of $\int_0^t F(s)X(ds)$ for a suitable class of adapted processes and itegrators $X$.
The main interesting theorem for our purposes is the following

**Theorem 21.** *Every Lévy process* $X = \{X(t), t \geq 0\}$ *is a semimartingale with respect to its natural filtration* $\mathscr{F}_t = \sigma\{X(s), 0 \leq s \leq t\}$.

*Proof*: the Lévy-Itô decomposition of $X$ is the backbone of the proof: let $X(t) = B(t) + \mu t + X_p(t)$ with $t \geq 0$ and where

$$X_p(t) = \int_{|x|\geq 1} xN(t, dx) + \int_{|x|<1} x\tilde{N}(t, dz), \quad t \geq 0 \tag{1.11}$$

as in (1.10).

By defining $Y_1(t), Y_2(t), t \geq 0$ respectively the first and second summand of (1.11), it is possible to write

$$X(t) = [B(t) + Y_2(t)] + [Y_1(t) + \mu t],\ t \geq 0,$$

and to claim that this is the semimartingale decomposition of Lévy process $X = \{X(t), t \geq 0\}$ .

Firstly $B(t) + Y_2(t),\ t \geq 0$ is a Lévy process with generating triplet $(\Sigma, 0, \nu \mathbb{1}_{\{|x|<1\}})$, and trivially a zero-mean process.

Every zero-mean Lévy process is a martingale with respect of its natural filtration; let in fact be $\tilde{X}(t)$ a zero-mean Lévy process and $\tilde{\mathscr{F}}_t = \sigma\{\tilde{X}(s), 0 \leq s \leq t\},\ t \geq 0$, hence

$$E[\tilde{X}(t)|\tilde{\mathscr{F}}_s] = E[\tilde{X}(t) - \tilde{X}(s)|\tilde{\mathscr{F}}_s] + E[\tilde{X}(s)|\tilde{\mathscr{F}}_s] =$$
$$= E[\tilde{X}(t) - \tilde{X}(s)] + \tilde{X}(s) = \tilde{X}(s).$$

Indeed it is also a martingale with respect to the filtration generated by $X$: the only extra information is in the jumps in $Y_1(t)$, but the statement follows for independence.

Then $M(t) = B(t) + Y_2(t),\ t \geq 0$ is a $\mathscr{F}_t = \sigma\{X(s),\ 0 \leq s \leq t\}-$martingale, $t \geq 0$.

Now what we need to check is that $A(t) = Y_1(t) + \mu t$ is an adapted process of bounded variation. Supposing adaptedness holds (trivially because $\mu t$ is a constant for every $t \geq 0$ and $Y_1(t)$ is adapded with respect to its own filtration, which is a subset of $\sigma\{X(s), 0 \leq s \leq t\},\ t \geq 0$), it is easy to see that

(1) $\{\mu t,\ t \geq 0\}$ is of locally bounded variation for linearity;

(2) $\{Y_1(t),\ t \geq 0\}$ is still a Lévy process with generating triplet $(0, 0, \nu \mathbb{1}_{\{|x|\geq 1\}})$, which is a compound Poisson process (of locally bounded variation because piecewise constant).

□

### 1.8.1 Characteristics of semimartingales

The notion of characteristics of a semimartingale is introduced to generalize the idea of Lévy triplet. Recall that by Lévy-Khinchine representation the Lévy triplet was identified uniquely. If $X$ is a semimartingale, the idea applies in the following intuitive way: the aim is to find two processes $\{B_t\}$ and $\{C_t\}$, and a random measure $\nu$ such that if we define the process $\phi_t(u)$ by means of the one-dimensional Lévy-Khinchine formula through

$$\log \phi_t(u) = iuB_t - \frac{1}{2}u^2 C_t + \int_{\mathbb{R}\setminus\{0\}} \left[e^{iuy} - 1 - iuh(y)\right] \nu([0,t] \times dy), \tag{1.12}$$

with $h(x)$ any bounded Borel function with compact support which "behaves like $x$" near the origin, then it holds that

$$\frac{\exp(iuX_t)}{\phi_t(u)} \tag{1.13}$$

is a martingale. It is not possible to find a triplet $(B, C, \nu)$ which is deterministic (unless $X$ is a semimartingale with independent increments), but it is possible to find an *unique* triplet for which (1.13) holds and that is predictable. A triplet of this kind is called the *characteristics* of $X$. See Jacod and Shiryaev, [JS02], pg. $75 - 76$ for che exact definition.

## 1.9 Conclusions

This first chapter set the theme of our discussion, introducing the main elements which will be specified where necessary in the following. The main references for this chapter are given by the books of Sato, [Sat99], Appelbaum, [App04], and Bertoin, [Ber96].

It is clear how the main results are linked to the Lévy-Khinchine and Itô decompositions, and as for the main definition is concerned, to the notion of infinite divisibility.

# Chapter 2

# Some Particular Examples of the Applications of Lévy Process

In this chapter we will see a quick review of some of the most interesting or important Lévy processes , describing their origins and main properties and, where possible, showing some simulated trajectories. The simulations are not always easy to obtain, unless either the increments distribution is known in a closed analytical form or a representation through subordinator can be written, which is not always the case.

So the examples introduced here, besides the classical starting points as Poisson process or Brownian motion, are relevant either for their applications, mainly in mathematical finance, or for their capability of being simulated, or both.

This in part also shows the importance of having a subordinator representation of the process: subordinators in fact provide a good way for simulation. When possible, simulations are obtained with $R$ software on a fixed time grid for the sake of simplicity. Hence most of the work in this case has been performed in creating the routines for simulating the trajectories of the processes.

The list is clearly not exhaustive, as for instance drops all the section generating from stable processes (self similar processes, fractional Lévy motions and others): the aim here is to describe the possible variety of some of these processes which will in some sense constitute the starting point for the following, and by themselves provide a wide range of models generally used to model financial markets.

The first introduced are the basic processes.

## 2.1 Poisson Process

Let $\lambda > 0$ and $\mu_\lambda$ be a probability distribution on $k = 0, 1, 2, \ldots$ and such that

$$\mu_\lambda(\{k\}) = \frac{e^{-\lambda}\lambda^k}{k}$$

i.e. the Poisson distribution $P(\lambda)$. It is easy to evaluate the characteristic function of a random variable $X \sim P(\lambda)$ as

$$\phi_P(u) = \sum_{k\geq 0} e^{iux}\mu_\lambda(\{x\}) = e^{-\lambda(1-e^{iu})} = \left[e^{-\frac{\lambda}{n}}(1-e^{iu})\right]^n,$$

the latter equality showing infinite divisibility of the Poisson distribution, since it has the characteristic function equal to the sum of the characteristic functions of $n$ independent Poisson distributions each with parameter $\lambda/n$.

This means we can define a Lévy process, the *Poisson process*, $N = \{N(t), t \geq 0\}$ with intensity parameter $\lambda$ as the process which starts at $0$, has independent and stationary increments and where the increment over a time interval of length $s > 0$ enjoys a $P(\lambda s)$ distribution.

So the Poisson process turns out to be an increasing pure jump process, with jump sizes always equal to $1$; this means that the Lévy triplet is given by $(0, 0, \lambda\delta_1)$, where $\delta_1$ is the Dirac measure at point $1$, i.e. a measure with a mass of only $1$ at point $1$.

The time between two jumps follows an $Exp(\lambda)$ distribution.

It is easy to evaluate the first moments of a Poisson distribution:

| $P(\lambda)$ | |
|---|---|
| mean | $\lambda$ |
| variance | $\lambda$ |
| skewness | $\lambda^{-1/2}$ |
| kurtosis | $3+\lambda^{-1}$ |

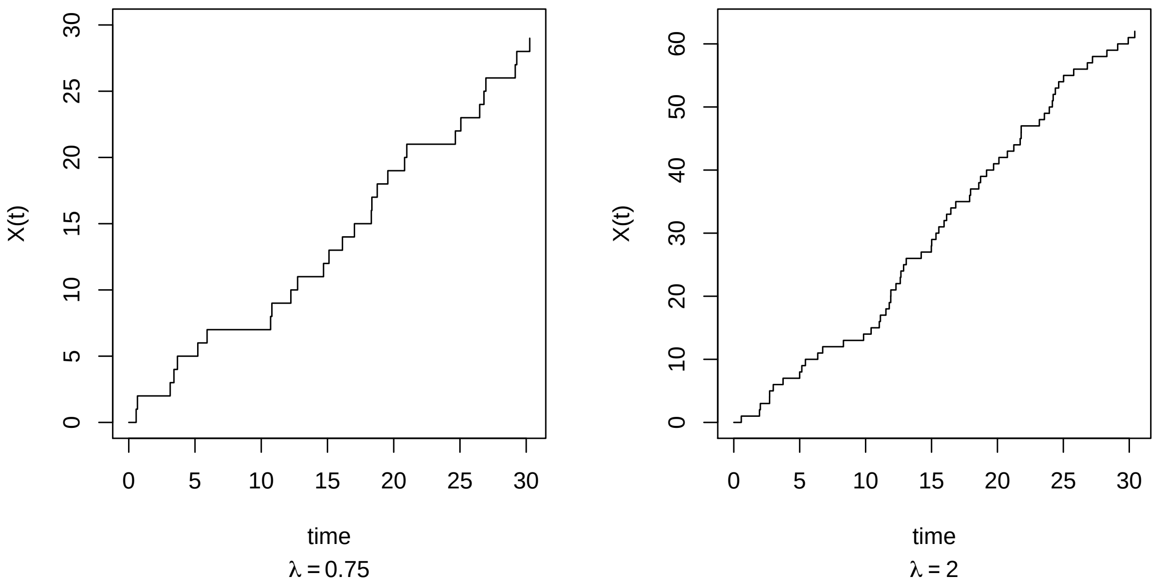


Figure 2.1: two different trajectories of a Poisson process, respectively with parameters $\lambda = 0.75$ and $\lambda = 2$

This process provides a good theoretical standing point for the construction of more complex models, besides its standard use as counting process.

## 2.2 Compound Poisson Process

Suppose now $N = \{N(t), t \geq 0\}$ is a Poisson process with intensity parameter $\lambda > 0$ and that $Z_i,\ i = 1, 2, \ldots$ is an i.i.d. sequence of random variables independent of $N$ and following a law $P_Z$ with characteristic function $\phi_Z(u)$.

Then the process

$$X(t) = \sum_{k=0}^{N(t)} Z_k,\ t \geq 0$$

is a *compound Poisson* process; the value of the process at time $t$ is the sum of a random ($\sim\ P(\lambda t)$) numbers with common distribution $P_Z$. Obviously the ordinary Poisson process corresponds to the case where $Z_i = 1$ for all $i$, i.e. where the distribution is degenerate at point 1.

From example 2 in chapter 1 we see that compound Poisson distribution is infinitely divisible and hence its associated Lévy process has characteristic function

$$E[e^{iuX(t)}] = \exp\left(t \int_{\mathbb{R}} (\exp\{iux\} - 1)\nu(dx)\right) = \exp(t\lambda(\phi_Z(u) - 1)),$$

where $\nu(dx) = \lambda P_Z(dx)$ and which leads to a Lévy triplet given by

$$\left[\int_{-1}^{1} x\nu(dx),\ 0,\ \nu(dx)\right].$$

Here are some samples of possible trajectories of this process

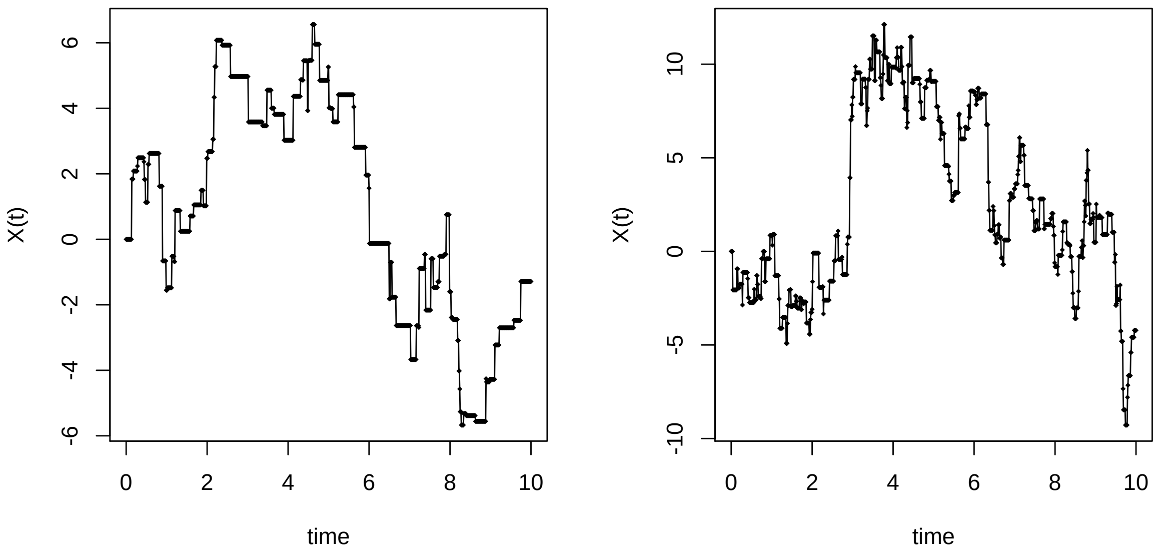


Figure 2.2: two different possible trajectories of a Compound Poisson process: intensity of jumps $\lambda = 10$, distribution of jumps $N(0,1)$ (*left*); intensity of jumps $\lambda = 30$, distribution of jumps $N(0,1)$

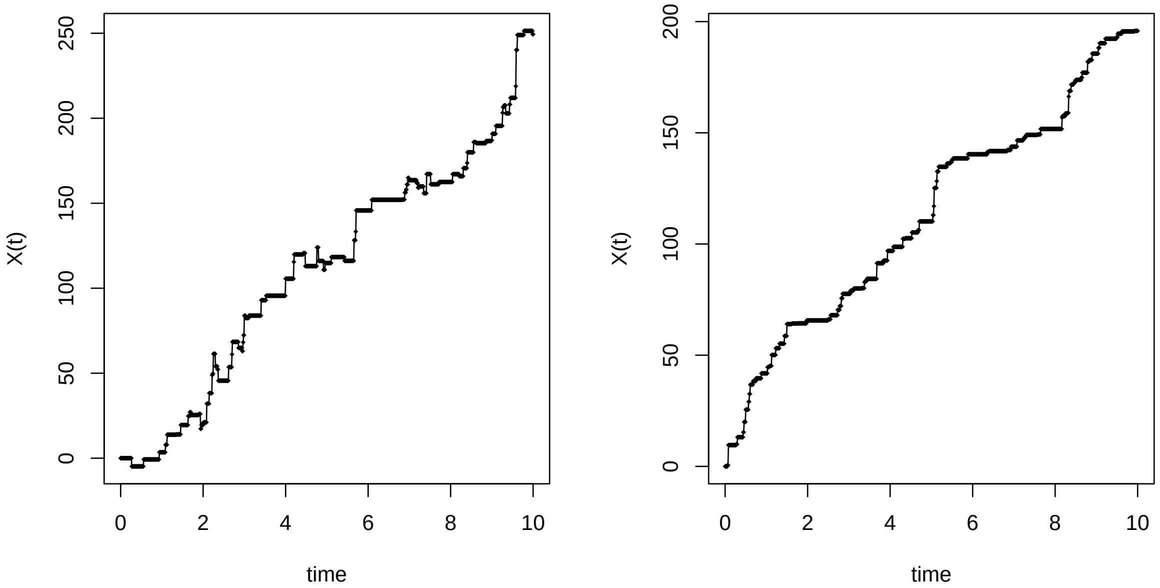


Figure 2.3: two different possible trajectories of a Compound Poisson process: intensity of jumps $\lambda = 10$, distribution of jumps $N(3,5)$ (*left*); intensity of jumps $\lambda = 10$, distribution of jumps $\Gamma(1,1/2)$

The following, and last result of this section is the proof of a slight generalization of DeFinetti's theorem 2 seen in the beginning paragraph:

**Theorem 22.** *Every infinite divisible distribution is the limit of a sequence of a compound Poisson distributions.*

*Proof*: : let $\mu$ be an infinitely divisible probability measure, and choose a real sequence $t_n \downarrow 0$. Define $\mu_n$ by its characteristic function as

$$\phi_{\mu_n}(u) := \exp\{t_n^{-1}(\phi_\mu(u) - 1)\} = \exp\left[t_n^{-1}\int_{\mathbb{R}^d\setminus\{0\}}(e^{i(u,x)} - 1)\mu^{t_n}(dx)\right].$$

The distribution of $\mu_n$ is compound Poisson. Now observe that

$$\phi_{\mu_n}(u) = \exp[t_n^{-1}(e^{t_n \log\phi_\mu(u)} - 1)] = \exp[t_n^{-1}(t_n \log\phi_\mu(u) + O(t_n^2))]$$

for each $u$ as $n \to \infty$. Therefore $\phi_{\mu_n}(u) \to e^{\log\phi_\mu(u)} = \phi_\mu(u)$. □

## 2.3 Brownian Motion

[1] A stochastic process $W(t)$ on $\mathbb{R}^d$ is a *Brownian motion* (or a *Wiener process*) if it is a Lévy process and

(1) for $t > 0$, $W(t)$ has a Gaussian distribution with mean $0$ and covariance matrix $tI$ (with $I$ the $d \times d$ identity matrix);

(2) there is $\Omega_0 \in \mathbf{F}$ with $P(\Omega_0) = 1$ such that for every $\omega \in \Omega_0$, $W(t, \omega)$ is continuous in $t$.

Details on simple properties of Brownian motion are given for instance in Sato, [Sat99], pgg. $24 - 8$.

In this standard case the probability measure underlying the process is the $N(0, t)$, which is trivially infinitely divisible.

For the general case of a one dimensional Brownian motion with drift $b \in \mathbb{R}$, the probability measure is given by

$$\mu_{b,t}(dx) = \frac{1}{\sqrt{2\pi t^2}} e^{-\frac{(x-b)^2}{2t^2}} dx;$$

[1]**A brief historical note**: the history of Brownian motion dates back to 1828 whien the scottish botanist Robert Brown observed pollen grains and the spores of mosses and Equisetum suspended in water under a microscope; he observed minute particles within vacuoles in the pollen grains executing a continuous jittery motion. He then compared the same motion in particles of dust, enabling him to rule out the hypothesis that the effect was due to pollen being "alive". In 1900 Bachelier considered Brownian motion as a possible model for stock market prices, at a time when the topic was not considered worth of studying.

In 1905 A.Einstein considered Brownian motion as a model for particles in suspension. He observed that, if the kinetic theory of fluids were correct, then the molecules of water would move at random and so a small particle would receive a random number of impacts of random strength and from random directions in any short period of time.

Such a bombardment would cause a sufficiently small particle to move in exactly the way described by Brown.

In 1923 N.Wiener defined and constructed Brownian motion rigorously for the first time.

With the work of Samuelson (1965) Brownian motion reappeared as a modeling tool for finance.

it is also well known that

$$\phi_{\mu_{b,t}}(u) = \int_{\mathbb{R}} e^{iux}\mu_{b,t}(dx) = e^{iub-\frac{1}{2}t^2u^2} = \left[e^{iu\frac{b}{n}-\frac{1}{2}\left(\frac{u}{\sqrt{n}}\right)^2 t^2}\right]^n$$

again showing that is an infinitely divisible distribution and leading then to the Lévy triplet

$$[b, t, 0]$$

It is trivial to verify that, given a standard one-dimensional Brownian motion $W(t)$, a Brownian motion with drift has the form

$$X(t) = bt + \sqrt{t}W(t)$$

The obvious moments are

| $N(b,t)$ | |
|---|---|
| mean | $b$ |
| variance | $t$ |
| skewness | 0 |
| kurtosis | 3 |

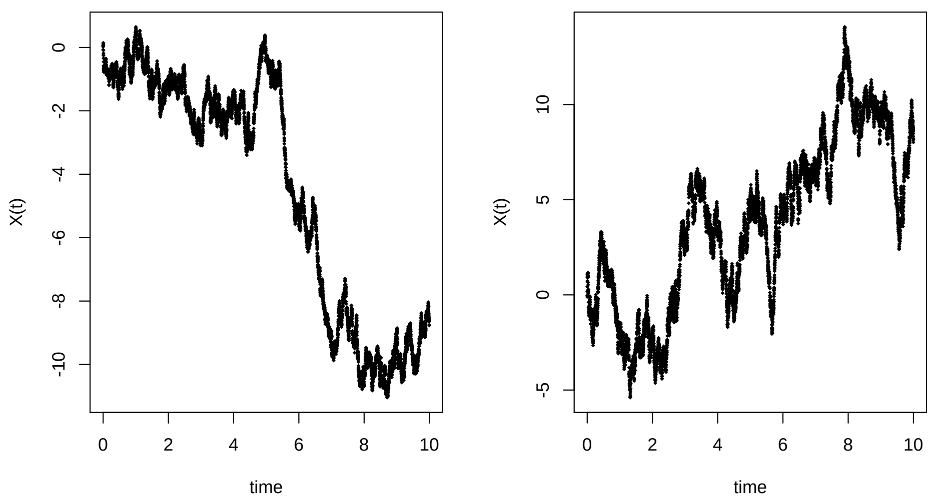


Figure 2.4: two different possible trajectories of a Brownian Motion: volatility $t = 4$, drift $b = 0$ (*left*); volatility $t = 25$, drift $b = 1.5$ (*right*)

## 2.4 Lévy jump-diffusion process

This is a process in which the jump component is given by a compound Poisson process. It can be represented in the form

$$X(t) = bt + \sqrt{c}W(t) + \sum_{i=1}^{N(t)} Y_i,$$

where $b \in \mathbb{R},\ c > 0,\ W(t)$ is the standard Brownian motion, $\{N(t), t \geq 0\}$ is a Poisson process with intensity $\lambda > 0$ and $\{Y_i, i \geq 1\}$ is a sequence of i.i.d. random variables, independent of $N(t)$.

Here is a couple of simulations from this process:

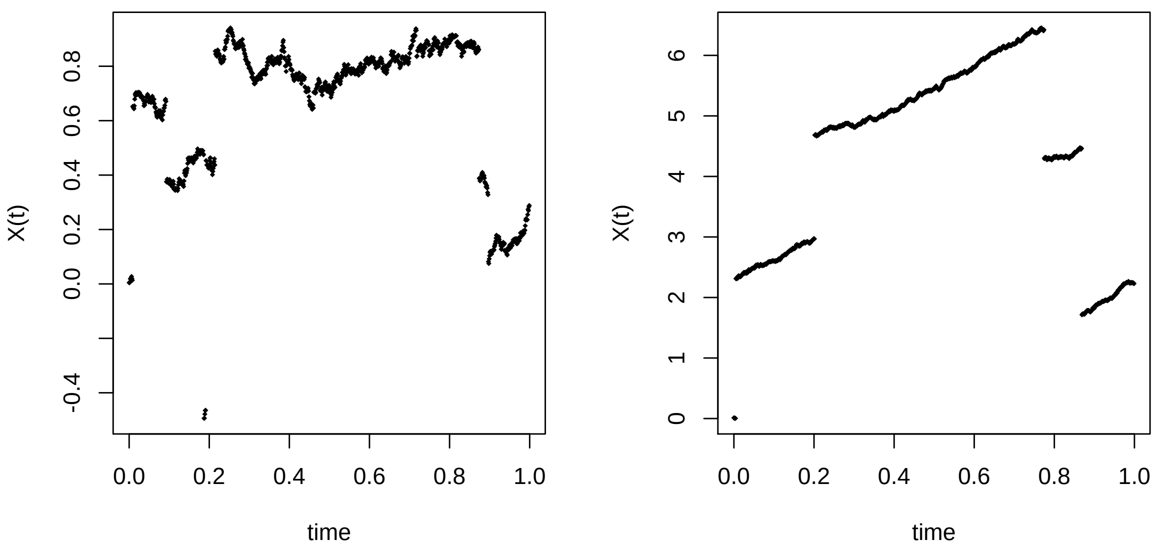


Figure 2.5: two different possible trajectories of a jump diffusion process: $\sigma^2 = 10, b = 0, Y_i \sim N(0,1),\ \lambda = 10$ (*left*); $\sigma^2 = 10, b = 3, Y_i \sim N(1,5),\ \lambda = 5$ (*right*)

For normally distributed random variables $Y_i$, Merton in [Mer76] introduces this process for asset return modeling; in a work by Kou, [Kou02], double exponentially distributed jump size variables $Y_i$ are used.

## 2.5 Gamma process

The density function of a $\Gamma(a, b)$ random variable with $a,\ b > 0$ is given by

$$f_G(x; a, b) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx},\ x > 0.$$

The characteristic function is easily obtainable as

$$\phi_G(u; a, b) = \left(1 - i\frac{u}{b}\right)^{-a} = \left[\left(1 - i\frac{u}{b}\right)^{-a/n}\right]^n,$$

which shows that it is infinitely divisible. Therefore the Gamma process is defined as the process $X_G = \{X_G(t),\ t \geq 0\}$ with parameters $a,\ b > 0$ which starts at $0$ and has stationary and independent increments distributed as a $\Gamma(a, b)$.

More precisely, time $t$ enters in the first parameter, as $X_G(t) \sim \Gamma(at, b)$, if $X$ is a Gamma process. The Lévy triplet is given by

$$\left[\frac{a(1 - \exp(-b))}{b},\ 0,\ a\exp(-bx)x^{-1}\mathbb{1}_{\{x>0\}}(dx)\right]$$

From the characteristic function, the moments are immediately derived:

| $\Gamma(a, b)$ | |
|---|---|
| mean | $a/b$ |
| variance | $a/b^2$ |
| skewness | $2a^{-1/2}$ |
| kurtosis | $3(1 + 2a^{-1})$ |

and a scaling property holds: if $X \sim \Gamma(a, b)$, then for any $c > 0$, $cX \sim \Gamma(a, b/c)$.

From the fact the increments distribution is completely analitically determined, it is easy to draw trajectories of Gamma process:

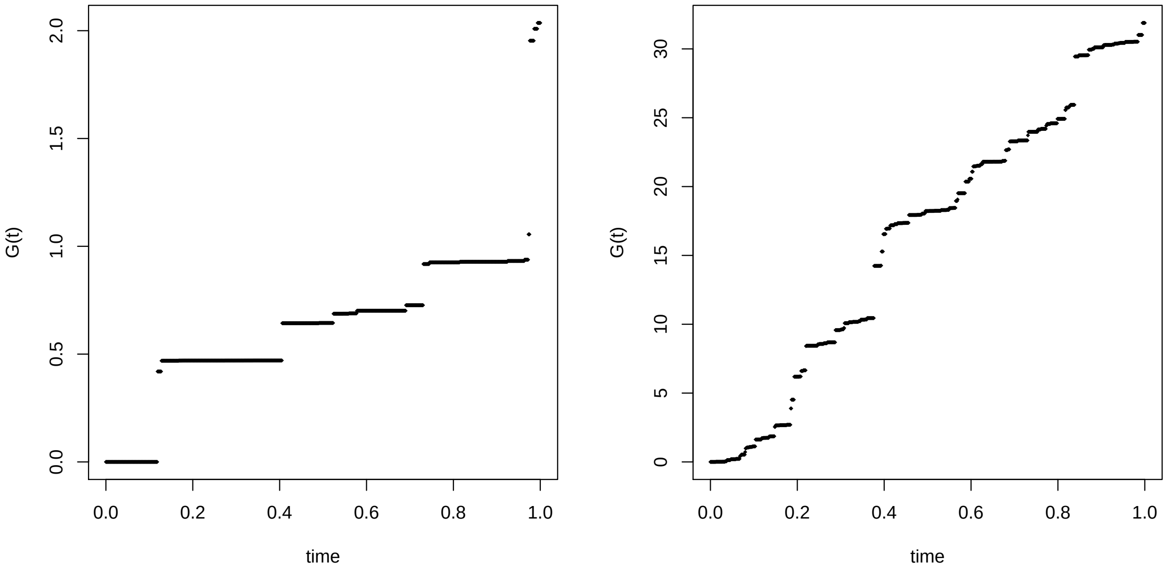


Figure 2.6: two different possible trajectories of a Gamma process $\Gamma(a,b)$ with $a = 3t$ (*left*), and $a = 30t$ (*right*), and $b = 1$ always

## 2.6 Variance Gamma process

The characteristic function of the Variance Gamma ($VG(\sigma,\nu,\theta)$) law is given by

$$\phi_{VG}(u;\sigma,\nu,\theta) = \left(1 - iu\theta\nu - \frac{1}{2}\sigma^2\nu u^2\right)^{-1/\nu}$$

which is infinitely divisible; this way a Lévy process $X$ remains defined, which starts at $0$, has independent and stationary increments and for which the increment $X(s+t) - X(s) \stackrel{d}{=} VG(\sigma\sqrt{t},\ \nu/t,\ t\theta)$ over the interval $[s, s+t]$.

For the process $X(t)$ It holds that

$$\begin{aligned}\phi_{VG}(u;\sigma,\nu,\theta) = \phi_{VG}(u;\sigma\sqrt{t},\nu/t,t\theta) = [\phi_{VG}(u;\sigma,\nu,\theta)]^t = \\ = \left(1 - iu\theta\nu + \frac{1}{2}\sigma^2\nu u^2\right)^{-t/\nu} \qquad (2.1)\end{aligned}$$

**Theorem 23.** *(Madan, Carr, Chang* [MC98]*) The $VG$ process may be expressed as the difference of two independent increasing Gamma processes, specifically*

$$X_{VG}(t;\sigma,\nu,\theta) = X_{G_1}(t;\mu_1,\nu_1) - X_{G_2}(t;\mu_2,\nu_2).$$

*Proof*: the characteristic function (2.1) may be written as the product of the following two characteristic functions,

$$\phi_{G_1}(u) = \left(\frac{1}{1 - i(\nu_1/\mu_1)u}\right)^{\left(\mu_1^2/\nu_1\right)t}$$

and

$$\phi_{G_2}(u) = \left(\frac{1}{1 - i(\nu_2/\mu_2)u}\right)^{\left(\mu_2^2/\nu_2\right)t}$$

with $\mu_1, \mu_2, \nu_1$ and $\nu_2$ satisfying

$$\frac{\mu_1^2}{\nu_1} = \frac{\mu_2^2}{\nu_2} = \frac{1}{\nu},$$
$$\frac{\nu_1\nu_2}{\mu_1\mu_2} = \frac{\sigma^2\nu}{2},$$
$$\frac{\nu_1}{\mu_1} - \frac{\nu_2}{\mu_2} = \theta\nu.$$

It follows that the $VG$ process is the difference of two Gamma processes with mean rates $\mu_1, \mu_2$ and variance rates $\nu_1, \nu_2$ respectively. □

The explicit relation between the parameters of the Gamma processes differenced in (23) and the original parameters of the $VG$ process is given by

$$\mu_1 = \frac{1}{2}\sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} + \frac{\theta}{2}, \qquad \mu_2 = \frac{1}{2}\sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} - \frac{\theta}{2}$$
$$\nu_1 = \left(\frac{1}{2}\sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} + \frac{\theta}{2}\right)^2 \nu, \qquad \nu_2 = \left(\frac{1}{2}\sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} - \frac{\theta}{2}\right)^2 \nu$$

This characterization allows the Lévy measure to be determined

$$\nu_{VG}(dx) = \begin{cases} C\exp(Gx)|x|^{-1}dx, & x < 0 \\ C\exp(-Mx)x^{-1}dx, & x > 0, \end{cases}$$

where

$$\begin{aligned} C &= 1/\nu > 0, \\ G &= \left(\sqrt{\frac{1}{4}\theta^2\nu^2 + \frac{1}{2}\sigma^2\nu} - \frac{1}{2}\theta\nu\right)^{-1} > 0, \\ M &= \left(\sqrt{\frac{1}{4}\theta^2\nu^2 + \frac{1}{2}\sigma^2\nu} + \frac{1}{2}\theta\nu\right)^{-1} > 0, \end{aligned}$$

which lead to a different possible parametrization of the process.

The Lévy measure has infinite mass, and hence $VG$ process has infinitely many jumps in any finite time interval. Since

$$\int_{-1}^{1} |x|\nu_{VG}(dx) < \infty,$$

a $VG$ process has paths of finite variation. A $VG$ process has no Brownian component and its Lévy triplet is given by

$$[\gamma,\ 0,\ \nu_{VG}(dx)],$$

where

$$\gamma = \frac{-C[G(\exp(-M) - 1)] - M[\exp(-G) - 1]}{MG}$$

### 2.6.1 Representation by subordination

Another way of defining a $VG$ process, always in [MC98] who first introduced this Lévy process , is obtaining it by evaluating Brownian motion with drift at a random time given by a gamma process. Let

$$b(t; \theta, \sigma) = \theta t + \sigma W(t)$$

where $W(t)$ is a standard Brownian motion. The process $b(t; \theta, \sigma)$ is a Brownian motion with drift $\theta$ and volatility $\sigma$. Now the $VG$ process $X_{VG}(t; \sigma, \nu, \theta)$, is defined in terms of the Brownian motion with drift $b(t;\theta,\sigma)$ and the Gamma process with unit mean rate, $\Gamma(t;1,\nu)$ as

$$X_{VG}(t;\sigma,\nu,\theta) = b(X_G(t;1,\nu);\theta,\sigma).$$

The $VG$ process has three parameters: (i) $\sigma$ the volatility of the Brownian motion, (ii) $\nu$ the variance rate of the gamma time change and (iii) $\theta$ the drift in the Brownian motion with drift. The process therefore provides two dimensions of control on the distribution over and above that of the volatility. We will observe below that control is attained over the skew via $\theta$ and over kurtosis with $\nu$.

The density function for the $VG$ process at time $t$ can be expressed conditional on the realization of the Gamma time change $g$ as a normal density function. The unconditional density may then be obtained on integrating out $g$ employing the density of the Gamma process for the time change $g$. This gives us $f_{VG}(x)$, the density for, $X_{VG}(t)$, as

$$f_{VG}(x) = \int_0^\infty \frac{1}{\sigma\sqrt{2\pi g}} \exp\left(-\frac{(x-\theta g)^2}{2\sigma^2 g}\right) \frac{g^{\frac{t}{\nu}-1}\exp\left(-g/\nu\right)}{\nu^{t/\nu}\Gamma(t/\nu)} dx.$$

Here are some simulations obtained following algorithm 6.11, pg. 192 in Cont and Tankov's work [CT03]

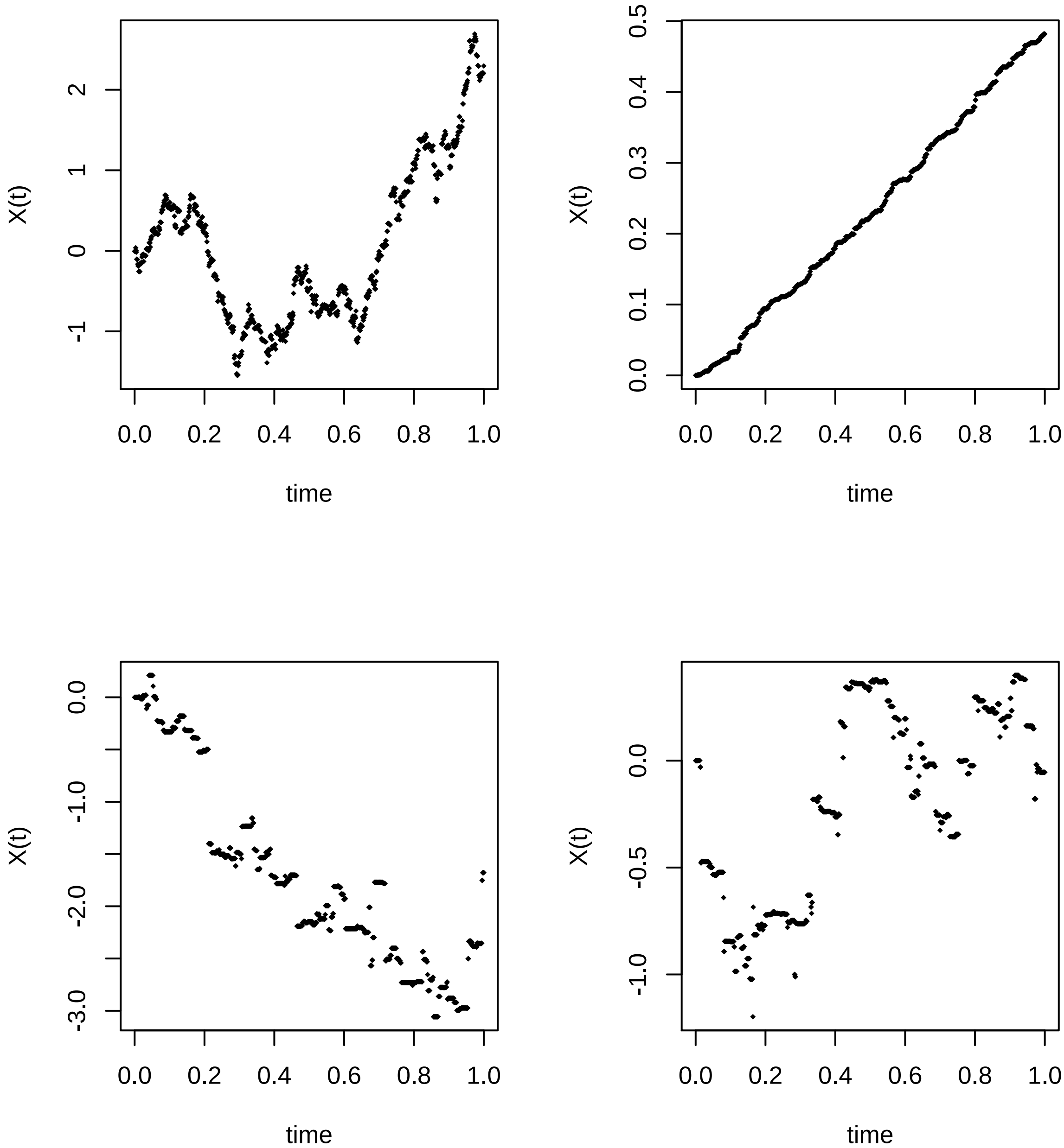


Figure 2.7: Different possible trajectories of a Variance Gamma process with $\sigma = 2,\ \theta = 0.5,\ \kappa = 1/\nu = 0.005$ (*upper left*),$\sigma = 0.02,\ \theta = 0.5,\ \kappa = 1/\nu = 0.005$ (*upper right*), $\sigma = 2,\ \theta = 0.5,\ \kappa = 1/\nu = 0.05$ (*lower left*), $\sigma = 2,\ \theta = 0.05,\ \kappa = 1/\nu = 0.05$ (*lower right*)

Let us finally see the moments with respect to different possible parametrizations introduced above:

| | $VG(\sigma,\theta,\nu)$ | $VG(\sigma,\theta,0)$ | $VG(C,G,M)$ | $VG(C,G,G)$ |
|---|---|---|---|---|
| mean | $\theta$ | $0$ | $\frac{C(G-M)}{MG}$ | $0$ |
| variance | $\sigma^2+\nu\theta^2$ | $\sigma^2$ | $\frac{C(G^2+M^2)}{(MG)^2}$ | $2CG^{-2}$ |
| skewness | $\frac{\theta\nu(3\sigma^2+2\nu\theta^2)}{(\sigma^2+\nu\theta^2)^{3/2}}$ | $0$ | $\frac{2(G^3-M^3)}{C^{1/2}(G^2+M^2)^{3/2}}$ | $0$ |
| kurtosis | $3\left[1+2\nu-\frac{\nu\sigma^4}{(\sigma^2+\nu\theta^2)^2}\right]$ | $3(1+\nu)$ | $3\left[1+\frac{2(G^4+M^4)}{C(M^2+G^2)^2}\right]$ | $3(1+\frac{1}{C})$ |

The $VG$ process can be advantageous to use in option pricing since it allows for a wider modeling of skewness and kurtosis than the Brownian motion does. As such the Variance Gamma model allows to consistently price options with different strikes and maturities using a single set of parameters. Madan and Seneta in [MS90] introduce a symmetric version of the variance gamma process is introduced. In [MC98] the authors extend the model to allow for an asymmetric form and present a formula to price European options under the variance gamma process. Hirsa and Madan in [HM03], show how to price American options under variance gamma.
Fiorani in [Fio03] presents numerical solutions for European and American barrier options under variance gamma process. He also provides computer programming code to price vanilla and barrier European and American barrier options under variance gamma process.
The variance gamma process has been successfully applied in the modeling of credit risk in structural models. The pure jump nature of the process and the possibility to control skewness and kurtosis of the distribution allow the model to price correctly the risk of default of securities having a short maturity, something that is generally not possible with structural models in which the underlying assets follow a Brownian motion.

## 2.7 Inverse Gaussian Process

Let $T^{(a,b)}$ be the first time that a standard Brownian motion with drift $b > 0$ reaches the positive level $a > 0$. It is well known that this random time follows an Inverse Gaussian distribution $IG(a, b)$ which has characteristic function

$$\phi_{IG}(u; a, b) = \exp\left[-a\left(\sqrt{-2iu + b^2} - b\right)\right]$$

Moreover this is an infinitely devisibile distribution and we can define the $IG$ process $X^{(IG)}$ with parameters $a, b > 0$ as the process starting at $0$, has independent and stationary increments such that

$$E[\exp(iuX_t^{(IG)})] = \phi_{IG}(u; at, b) =$$
$$= \exp[-at(\sqrt{-2iu + b^2} - b)].$$

The density function of the Inverse Gaussian distribution with parameters $a, b$ is explicitly known:

$$f_{IG}(x; a, b) = \frac{a}{\sqrt{2\pi}} \exp(ab) x^{-3/2} \exp\left[-\frac{1}{2}(a^2x^{-1} + b^2x)\right],\ x > 0. \tag{2.2}$$

The Lévy measure here is given by

$$\nu_{IG}(dx) = (2\pi x^3)^{-1/2} a \exp\left[-\frac{1}{2}b^2x\right] \mathbb{1}_{\{x>0\}} dx,$$

and the first element of the Lévy triplet equals

$$\gamma = \frac{a}{b}(2N(b) - 1),\ \text{where } N(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\, du.$$

It is possible to generate Inverse Gaussian random variables through a sort of rejection algorithm; this gives the chance to also draw a trajectory for an IG process. Here are four examples

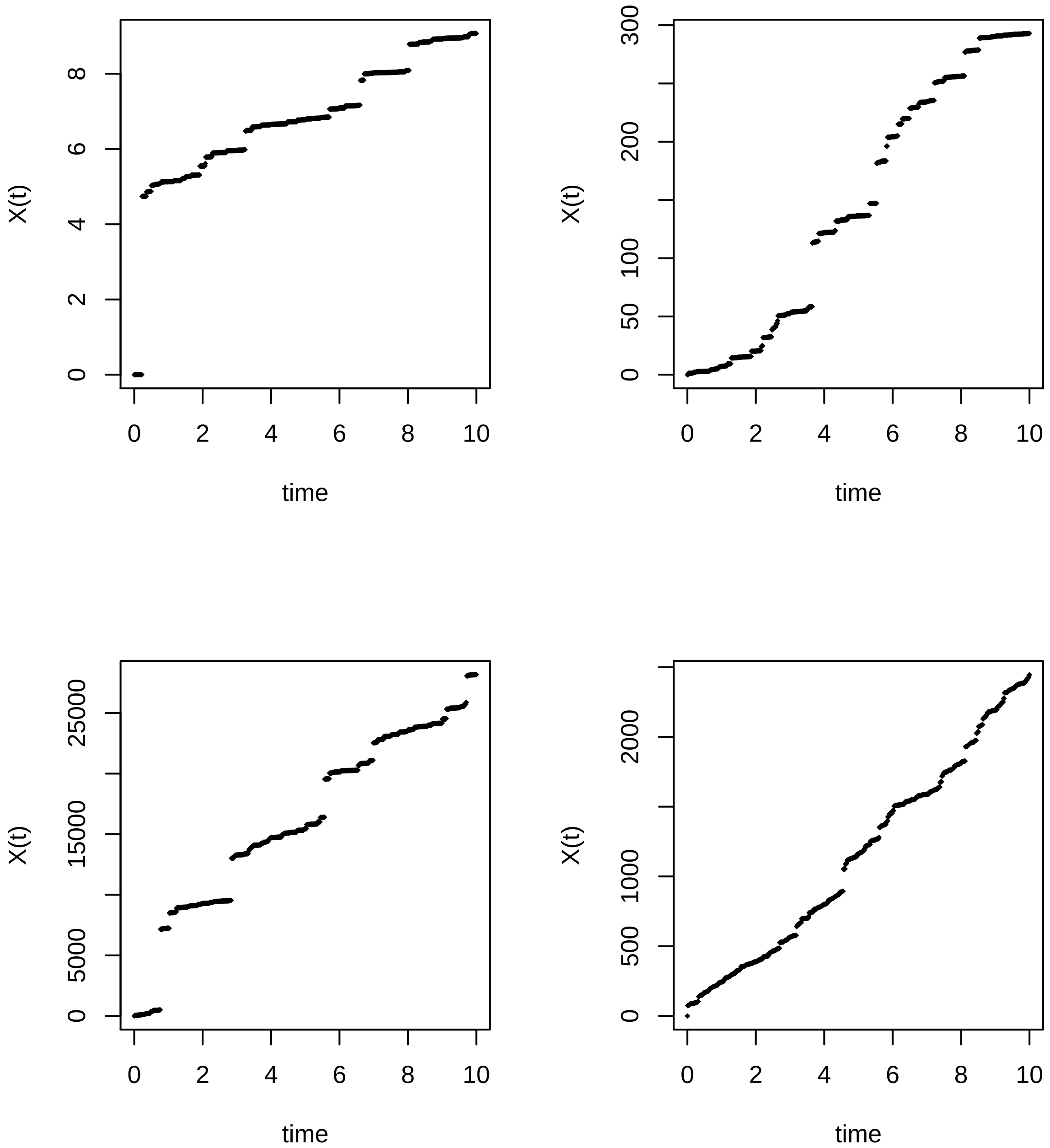


Figure 2.8: Different possible trajectories of an Inverse Gaussian process with $a = 0.005, b = 0.2$ (*upper left*), $a = 0.05, b = 0.2$ (*upper right*), $a = 0.5, b = 0.02$ (*lower left*), $a = 0.5, b = 0.2$ (*lower right*)

The Inverse Gaussian distribution is unimodal with a mode at $(\sqrt{4a^2b^2+9}-3)/(2b^2)$. We also have that:

$$E[X^{-\alpha}] = \left(\frac{b}{a}\right)^{2\alpha+1} E[X^{\alpha+1}], \quad \alpha \in \mathbb{R}.$$

Moreover

| $IG(a,b)$ | |
|---|---|
| mean | $a/b$ |
| variance | $a/b^3$ |
| skewness | $3(ab)^{-1/2}$ |
| kurtosis | $3[1+5(ab)^{-1}]$ |

and finally $IG$ distribution satisfies the scaling property $X \sim IG(a,b)$ then for a positive $c$, $cX \sim IG(a\sqrt{c}, b/\sqrt{c})$.

## 2.8 Generalized Inverse Gaussian Process

The distribution $IG(a,b)$ above can be extended to what is called the Generalized Inverse Gaussian law $GIG(\lambda, a, b)$; its density function is given by

$$f_{GIG}(x;\lambda,a,b) = \frac{(b/a)^{\lambda}}{2K_{\lambda}(ab)} x^{\lambda-1} \exp\left[-\frac{1}{2}(a^2x^{-1}+b^2x)\right], \quad x>0. \tag{2.3}$$

The parameters are such that $\lambda \in \mathbb{R}, a, b \geq 0$ and not simultaneousy $0$.

The characteristic function is

$$\phi_{GIG}(u;\lambda,a,b) = \frac{1}{K_{\lambda}(ab)} \left(1-2i\frac{u}{b^2}\right)^{\lambda/2} K_{\lambda}(ab\sqrt{1-2iub^{-2}}),$$

with $K_{\lambda}(z)$ the modified Bessel function of the third kind, i.e. one of the two linearly independent integrals of the Bessel differential equation

$$z^2\frac{d^2w}{dz^2} + z\frac{dw}{dz} - (z^2+\lambda^2)w = 0$$

It is a regular function of $z$ throughout the $z$-plane cut along the negative real axis, and for fixed $z(\neq 0)$ it is an entire function of $\lambda$. For further reference and details see Abramowitz and Stegun, [AS64], p.374.

Barndorff-Nielsen and Halgreen show in [BNH77] that this distribution is infinitely divisible, hence it is possible to define the $GIG$ process, the Lévy process whose increment over the interval $[s, s+t],\ s, t \geq 0$ has characteristic function $[\phi_{GIG}(u; \lambda, a, b)]^t$.

The Lévy measure is pretty complicated and has a density on $\mathbb{R}^+$ given by

$$\nu(x) = x^{-1} \exp\left[-\frac{1}{2}b^2 x\right] \left(a^2 \int_0^\infty \exp(-xz) g(z)\, dz + \max\{0, \lambda\}\right),$$

with

$$g(z) = \left[\pi^2 a^2 z \{J^2_{|\lambda|}(a\sqrt{2z}) + Y^2_{|\lambda|}(a\sqrt{2z})\}\right]^{-1},$$

where $J_\lambda(z)$ and $Y_\lambda(z)$ are solution of the differential equation

$$z^2 \frac{d^2 w}{dz^2} + z\frac{dw}{dz} + (z^2 - \lambda^2) w = 0$$

Each one is a regular (holomorphic) function of $z$ throughout the $z$-plane cut along the negative real axis, and for fixed $z(\neq 0)$ each is an entire (integral) function of $\lambda$. For further reference [AS64], p.358..

There is a general formula to evaluate the moments of a random variable $X$ following a $GIG(\lambda, a, b)$ distribution, and is given by

$$E[X^k] = \left(\frac{a}{b}\right)^k \frac{K_{\lambda+k}(ab)}{K_\lambda(ab)}, \quad k \in \mathbb{R},$$

from this it can be deduced that

| $GIG(\lambda, a, b)$ | |
|---|---|
| mean | $aK_{\lambda+1}(ab)/(bK_\lambda(ab))$ |
| variance | $a^2(bK_\lambda(ab))^{-2}[K_{\lambda+2}(ab)K_\lambda(ab)]$ |

### 2.8.1 Particular cases

- $IG(a, b)$ **distribution**: it is obtained for $\lambda = -1/2$ in the $GIG(\lambda, a, b)$; this descends from the fact that

$$K_{-1/2}(x) = \sqrt{\frac{\pi}{2x}} \exp(-x);$$

- $\Gamma(\tilde{a}, \tilde{b})$ **distribution**: for $a = 0,\ \lambda = \tilde{a} > 0,\ b = \sqrt{2\tilde{b}}$ in $GIG(\lambda, a, b)$, derives the $\Gamma(\tilde{a}, \tilde{b})$ distribution.

## 2.9 Tempered Stable Process

The characteristic function of the Tempered Stable distribution $TS(\kappa, a, b),\ a > 0,\ b \geq 0,\ 0 < \kappa < 1$, is given by

$$\phi_{TS}(u; \kappa, a, b) = e^{ab - a(b^{1/\kappa} - 2iu)^{\kappa}};$$

it is infinitely divisible and hence the $TS$ Lévy process can be defined from this as the process starting at $0$, having independent and stationary increments and whose increment $X_{t+s}^{(TS)} - X_s^{(TS)}$ follows a $TS(\kappa, ta, b)$ distribution over the interval $[t, t+s]$.

The Lévy measure of the process is derived from the characteristic function as

$$\nu_{TS}(dx) = \frac{a\kappa 2^{\kappa}}{\Gamma(1-\kappa)} \int_0^1 x^{-\kappa} e^{-(1/2)b^{1/\kappa}x}\, dx.$$

The main properties of this class of distributions are

| $TS(\kappa, a, b)$ | |
|---|---|
| mean | $2a\kappa b^{(\kappa-1)/\kappa}$ |
| variance | $4a\kappa(1-\kappa)b^{(\kappa-2)/\kappa}$ |
| skewness | $(\kappa-2)[ab\kappa(1-\kappa)]^{-1/2}$ |
| kurtosis | $3 + [4\kappa - 6 - \kappa(1-\kappa)][ab\kappa(1-\kappa)]^{-1}$ |

### 2.9.1 Particular cases

- $IG(a, b)$ **distribution**: for $\kappa = 1/2$;

- $\Gamma(a, b)$ **distribution**: for the limiting case $\kappa \to 0$.

## 2.10 Generalized Hyperbolic Process

Generalized hyperbolic distributions were introduced by Barndorff-Nielsen in [BN78] as a powerful five-parameter class of distributions to generate flexible Lévy processes. The Lebesgue density for these distributions is given by

$$f_{GH}(x) = a(\lambda,\alpha,\beta,\delta)\left[\delta^2+(x-\mu)^2\right]^{\frac{\lambda-1/2}{2}} + \\ + K_{\lambda-1/2}\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)\exp\{\beta(x-\mu)\} \tag{2.4}$$

where the first normalizing constant is

$$a(\lambda,\alpha,\beta,\delta) = \frac{(\alpha^2-\beta^2)^{\lambda/2}}{\alpha^{\lambda-1/2}\sqrt{2\pi}\delta^\lambda K_\lambda(\delta\sqrt{\alpha^2-\beta^2})}$$

and the parameter space is given by $\lambda\in\mathbb{R},\ \mu\in\mathbb{R}$ and

$$\begin{aligned} &\delta\geq 0, \quad \alpha>0 \quad -\alpha<\beta<\alpha \quad \text{if } \lambda>0\\ &\delta>0, \quad \alpha>0 \quad -\alpha<\beta<\alpha \quad \text{if } \lambda=0\\ &\delta>0, \quad \alpha\geq 0 \quad -\alpha\leq\beta\leq\alpha \quad \text{if } \lambda<0 \end{aligned}$$

In the same paper, the representation of the density of a generalized hyperbolic distribution as a mean-variance mixture of gaussian distribution is provided as

$$f_{GH}(x) = \int_0^\infty \frac{1}{\sqrt{2\pi v}}\exp\left\{-\frac{(x-\beta v-\mu)}{2v}\right\} g_{\gamma,\chi,\psi}(v)\,dv,$$

with $x\in\mathbb{R}$, and where the mixing density $g_{\gamma,\chi,\psi}(v)$ is the density of the generalized inverse gaussian distribution, i.e.

$$g_{\gamma,\chi,\psi}(v) = \frac{(\psi/\chi)^{\lambda/2}}{2K_\lambda(\sqrt{\chi\psi})}v^{\lambda-1}\exp\left\{-\frac{\chi v^{-1}+\psi v}{2}\right\},$$

with $v>0,\ \beta,\ \mu,\ \gamma\in\mathbb{R},\ (\chi,\psi)\in\Theta_\gamma$, as seen in (2.3), with

$$\Theta_\gamma = \begin{cases} \{(\chi,\psi):\chi\geq 0,\psi>0\}, & \text{if } \gamma>0,\\ \{(\chi,\psi):\chi>0,\psi>0\}, & \text{if } \gamma=0,\\ \{(\chi,\psi):\chi>0,\psi\geq 0\}, & \text{if } \gamma<0, \end{cases}$$

and $K_\lambda(z)$ is the usual modified Bessel function of the third kind.

Three subclasses of generalized hyperbolic distributions with semiheavy tails are used successfully in modeling observational series from finance and turbulence. Summarizing the most common cases one has

| **Parameters** | **Mixing distribution** | **Mixture distribution** |
|---|---|---|
| $\gamma = 1$ | hyperbola distribution | hyperbolic distribution |
| $\beta = \mu = \chi = 0$ <br> $\gamma = \psi > 0$ | gamma distribution | Variance Gamma distribution |

The moment generating function $M_{GH}(u)$ exists for $u$ with $|\beta + u| < \alpha$ and is given by

$$M_{GH}(u) = \exp\{\mu u\} \left[\frac{\alpha^2 - \beta^2}{\alpha^2 - (\beta+u)^2}\right]^{\lambda/2} \frac{K_\lambda(\delta\sqrt{\alpha^2 - (\beta+u)^2})}{K_\lambda(\delta\sqrt{\alpha^2-\beta^2})}.$$

As a consequence exponential moments $E[\exp(X_t)]$ are finite, which is crucial for pricing derivatives under martingale measure.

Let $X$ be an absolutely continuous real random variable with p.d.f. $f_{GH}(x)$; then the following quantities can be evaluated, taking $\zeta = \delta\sqrt{\alpha^2 - \beta^2}$:

| $GH(\lambda, \alpha, \beta, \delta)$ | |
|---|---|
| mean | $\frac{\beta\delta}{\alpha^2-\beta^2} \frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)}$ |
| variance | $\delta^2 \left[\frac{K_{\lambda+1}(\zeta)}{\zeta K_\lambda(\zeta)} + \frac{\beta^2}{\alpha^2-\beta^2}\left(\frac{K_{\lambda+2}(\zeta)}{K_\lambda(\zeta)} - \frac{K^2_{\lambda+1}(\zeta)}{K^2_\lambda(\zeta)}\right)\right]$ |

The characteristic function is obtained as $\phi_{GH}(x) = M_{GH}(ix)$; analyzing $\phi_{GH}$ in its form (1.4), obviously in the one-dimensional case, it can be observed that there is no Gaussian part, i.e. generalized hyperbolic Lévy motions are purely discontinuous processes. The Lévy measure $\nu(dx)$ is given by

$$\nu_{GH}(dx) = \frac{e^{\beta x}}{|x|} \int_0^\infty \frac{\exp\left(-|x|\sqrt{2y+\alpha^2}\right)}{\pi^2 y[J^2_{|\lambda|}(\delta\sqrt{2y}) + Y^2_{|\lambda|}(\delta\sqrt{2y})]}\, dy + \mathbb{1}_{\{\lambda \geq 0\}} \lambda e^{-\alpha|x|}.$$

where the functions $J_\lambda(z),\ Y_\lambda(z)$ the usual Bessel functions.

### 2.10.1 Particular cases

Setting $\lambda = 1/2$ in (2.4) one gets another interesting subclass, the normal inverse gaussian (NIG) distributions, used in finance for the first time in [BN98].

## 2.11 Normal Inverse Gaussian Process

The base distribution for this process is the Normal Inverse Gaussian ($NIG(\alpha, \beta, \delta)$) with parameters $\alpha > 0,\ |\beta| < \alpha$ and $\delta > 0$. The characteristic function is

$$\phi_{NIG}(u; \alpha, \beta, \delta) = \exp[-\delta(\sqrt{\alpha^2 - (\beta + iu)^2} - \sqrt{\alpha^2 - \beta^2})],$$

which is infinitely divisible (see for instance Barndorff-Nielsen, [BN97]) and so generates in the usual way a Lévy process whose variables $X_t^{(NIG)}$ follow a $NIG(\alpha, \beta, t\delta)$ law.

Moreover the Lévy measure is given by

$$\nu_{NIG}(dx) = \frac{\delta\alpha}{\pi} \frac{\exp(\beta x) K_1(\alpha|x|)}{|x|} dx,$$

where as usual $K_\lambda(x)$ denotes the modified Bessel function of the third kind.

It can be shown that a $NIG$ process has no Brownian component and Lévy triplet given by

$$\left[ \frac{2\delta\alpha}{\pi} \int_0^1 \sinh(\beta x) K_1(\alpha x)\, dx,\ 0,\ \nu_{NIG}(dx) \right].$$

The density of a $NIG(\alpha, \beta, \delta)$ is given by

$$f_{NIG}(x; \alpha, \beta, \delta) = a(\alpha, \beta, \delta) q \left(\frac{x}{\delta}\right)^{-1} K_1\left(\delta\alpha q \left(\frac{x}{\delta}\right)\right) \exp(\beta x), \tag{2.5}$$

where

$$a(\alpha, \beta, \delta) = \pi^{-1} \alpha \exp(\delta \sqrt{\alpha^2 - \beta^2}),$$

and

$$q(x) = \sqrt{1 + x^2}$$

The distribution is symmetric around 0 provided that $\beta = 0$. The parameters $\tilde{\alpha} = \delta\alpha,\ \tilde{\beta} = \delta\beta$ are invariant under location-scale changes.

Let now be $IG(a, b)$ be the Inverse Gaussian distribution with density function (2.2): because of the first two moments of $IG(a, b)$ it can be seen that $NIG(\alpha, \beta, \delta)$ distribution is a normal variance-mean mixture.

In fact it occurs as the marginal distribution of $X$ for a pair of random variables $(Z, X)$ where $Z \sim IG(\delta, \sqrt{\alpha^2 - \beta^2})$ distribution while conditional on $Z = z$, $X \sim N(\beta z, z)$. This is the reason why it is referred to (2.5) as the Normal Inverse Gaussian distribution.

It can also be observed that for $\beta = 0,\ \alpha \to \infty$ and $\delta/\alpha = \sigma^2$, the $N(0, \sigma^2)$ distribution appears, and that moreover Cauchy distribution is a particular case of $NIG(0, 0, 0)$.

### 2.11.1 Representation by subordination

As a direct consequence of the mixture representation of the $NIG$ distribution, one can see that the $NIG$ Lévy process $X(t)_{(NIG)}$ can be represented via random time change of a Brownian motion as

$$X(t)_{(NIG)} = \delta W(Z(t)) + \beta\delta^2 Z(t)$$

where $W(t)$ is the standard Brownian and $Z(t)$, stochastically independent of $W(t)$, is an $IG(1, \sqrt{\alpha^2 - \beta^2})$ process.

The variable $Z(t)$ can be interpreted as the first passage at time level $\delta t$ of a Brownian motion with drift $\sqrt{\alpha^2 - \beta^2}$ and diffusion coefficient 1. For more details see [BN97].

It holds that if $X \sim NIG(\alpha, \beta, \delta)$ then $-X \sim NIG(\alpha, -\beta, \delta)$.

The following forms for the characteristics hold:

| | $NIG(\alpha, \beta, \delta)$ | $NIG(\alpha, 0, \delta)$ |
|---|---|---|
| mean | $\delta\beta/\sqrt{\alpha^2-\beta^2}$ | 0 |
| variance | $\alpha^2\delta(\alpha^2-\beta^2)^{-3/2}$ | $\delta/\alpha$ |
| skewness | $3\beta\alpha^{-1}\delta^{-1/2}(\alpha^2-\beta^2)^{-1/4}$ | 0 |
| kurtosis | $3\left(1+\frac{\alpha^2+4\beta^2}{\delta\alpha^2\sqrt{\alpha^2-\beta^2}}\right)$ | $3(1+\frac{1}{\delta\alpha})$ |

Moreover, the $NIG$ distributions have semiheavy tails; more precisely

$$f_{NIG}(x;\alpha,\beta,\delta) \sim |x|^{-3/2}\exp[(\mp\alpha+\beta)x], \quad \text{as } x \to \pm\infty$$

The following simulations are obtained via subordination following [CT03]:

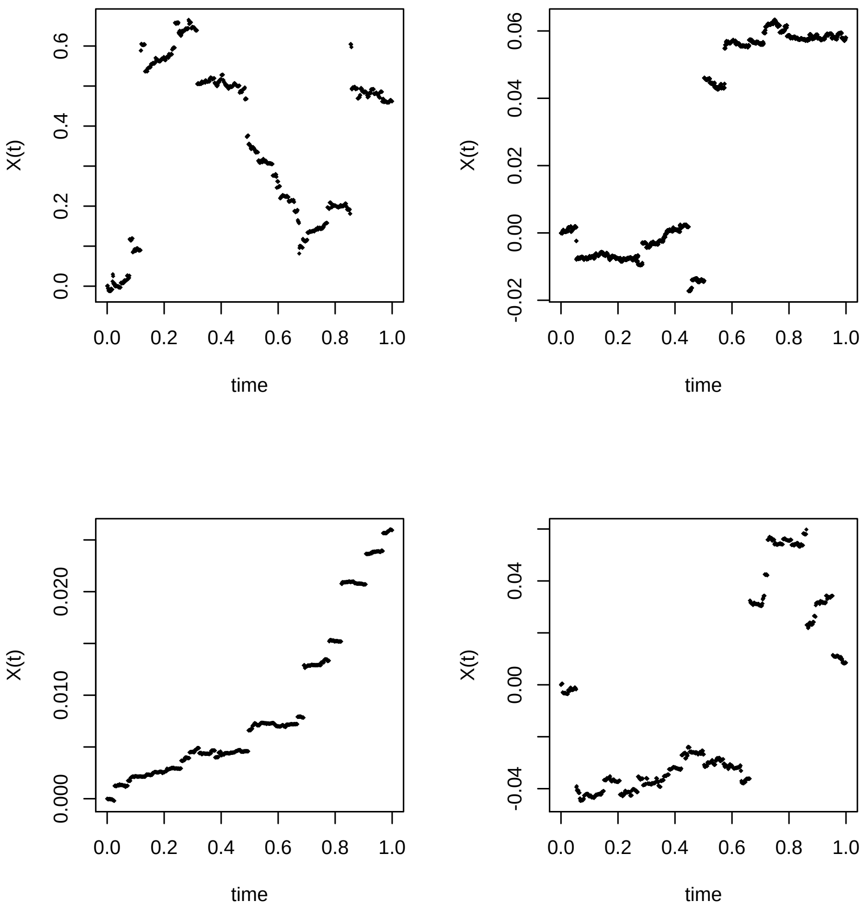


Figure 2.9: Different possible trajectories of a Normal Inverse Gaussian process, respectively with parameters $\alpha = 0.2025$, $\beta = 0.2$, $\delta = 0.5$ (upper left), $\alpha = 20$, $\beta = 20$, $\delta = 0.5$ (upper right), $\alpha = 2000$, $\beta = 2000$, $\delta = 0.005$ (lower left), $\alpha = 0.2025$, $\beta = 0.2$, $\delta = 0.05$ (lower right)

The $NIG$ distribution can approximate most hyperbolic distributions very closely but can also describe observations with considerably heavier tail behaviour than the log linear rate of decrease that characterizes the hyperbolic shape. Since, in addition, the $NIG$ distribution has more tractable probabilistic properties than the hyperbolic, it seems potentially of substantial usefulness.

The study of velocity differences in moderate and high Reynolds number turbulent wind fields is of central importance in turbulence, both theoretically and practically. Numerous and extensive observational investigations have shown that the velocity differences typically follow distributions that are close to symmetric and have tail that are either nearly log linear or somewhat heavier.

However the normal inverse Gaussian distribution seems to offer an attractive alternative starting point for parametric modelling in turbulence because of its special probabilistic properties and its ability to describe the typical tail behaviour of the velocity differences. Eberlein and Keller in [EK95], show that the hyperbolic distribution provides a very good fit to the distributions of daily returns measured on the log scale, of single stocks or portfolios of stocks from a number of German enterprises. The time series of daily returns concerned do not exhibit significant autocorrelations, nor do the derived series of squared returns. It is therefore natural to try to model the logarithmic stock price processes as Lévy processes , and for this purpose in the cited paper the hyperbolic Lévy processes are introduced, via the fact that hyperbolic distributions are infinitely divisible.

Following works have shown that the $NIG$ distribution provides an even better description of the German data than the hyperbolic, and that the data point to the $NIG$ as bein the most appropriate within the class of generalized hyperbolic distributions.

## 2.12 CGMY Processes

Carr, Geman, Madan and Yor [CGMY02] introduced a class of infinitely divisible distributions - called CGMY- which estends the Variance Gamma model above. CGMY Lévy process have purely discontinuous paths, that is it contains no Brownian part, and the Lévy density is given by

$$\nu_{CGMY}(x) = \begin{cases} C\dfrac{\exp\left(-G|x|\right)}{|x|^{1+Y}}, & x < 0, \\ C\dfrac{\exp\left(-Mx\right)}{x^{1+Y}}, & x > 0. \end{cases}$$

The parameter space is $C, G, M > 0$ and $Y \in (-\infty, 2)$; choosing the $Y$ parameter such that $Y \geq 2$, does not yield a valid Lévy measure. The process has infinite activity if and only if $Y \in [0, 2)$ and the paths have infinite variation if and only if $Y \in [1, 2)$.

### 2.12.1 Particular Cases

For $Y = 0$ one gets the three parameter Variance Gamma distributions, also a subclass of generalized hyperbolic distributions.

For $Y < 0$ the characteristic function of CGMY distribution is given by

$$\phi_{CGMY}(u) = \exp\left\{C\Gamma(-Y)\left[(M - iu)^Y - M^Y + (G + iu)^Y - G^Y\right]\right\}$$

The characteristics of the process can be evaluated as

| $CGMY(C, G, M, Y)$ | |
|---|---|
| mean | $C(M^{Y-1} - G^{Y-1})\Gamma(1 - Y)$ |
| variance | $C(M^{Y-2} - G^{Y-2})\Gamma(2 - Y)$ |
| skewness | $\frac{C(M^{Y-3} - G^{Y-3})\Gamma(3-Y)}{(C(M^{Y-2} - G^{Y-2})\Gamma(2-Y))^{3/2}}$ |
| kurtosis | $3 + \frac{C(M^{Y-4} - G^{Y-4})\Gamma(4-Y)}{(C(M^{Y-2} - G^{Y-2})\Gamma(2-Y))^{2}}$ |

## 2.13 $\alpha$-stable Processes

Stable distributions are a classical subject in probability; they constitute a four parameter class of distributions with characteristic function given by

$$\phi_S(u) = \exp\left\{\sigma^\alpha(-|u|^\alpha) + iu\omega(u,\alpha,\beta) + i\mu u\right\}$$

where

$$\omega(u,\alpha,\beta) = \begin{cases} \beta|u|^{\alpha-1}\tan\left(\pi\alpha/2\right), & \alpha \neq 1, \\ -\beta\frac{2}{\pi}\log|u|, & \alpha = 1. \end{cases}$$

The parameter space is $\alpha \in (0,2], \sigma \geq 0, \beta \in [-1,1]$ and $\mu \in \mathbb{R}$.

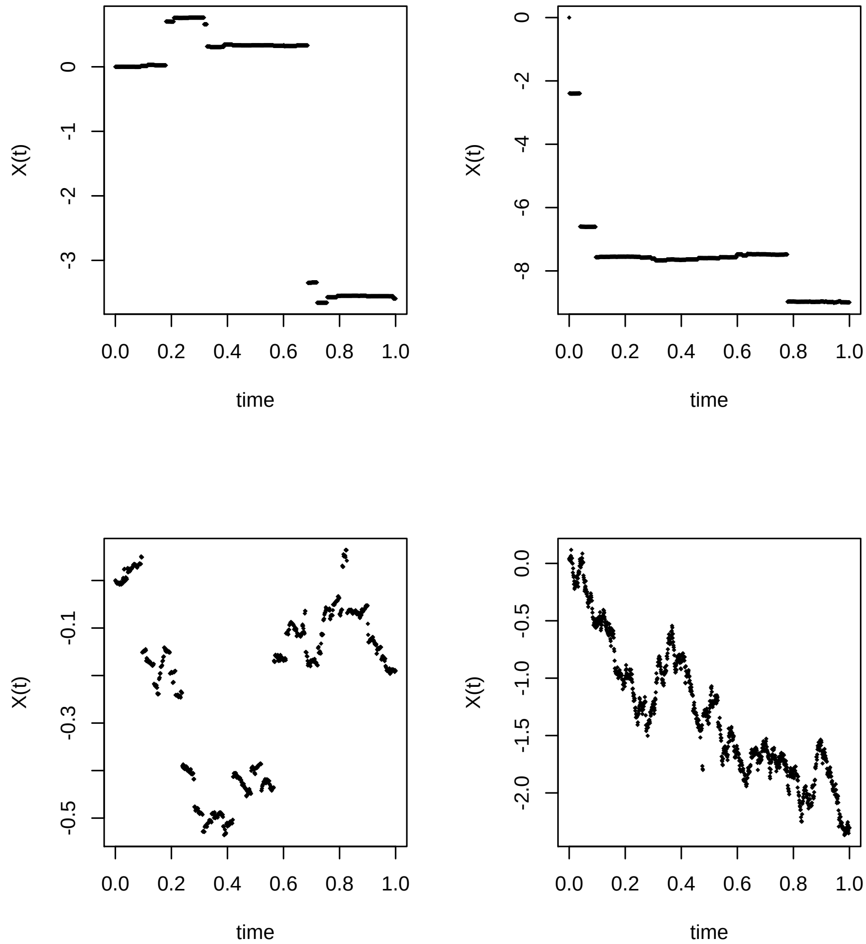


Figure 2.10: Different possible trajectories of an $\alpha-$stable process, respectively with $\alpha = 0.5, 0.75, 1, 1.9$

### 2.13.1 Particular Cases

Observe that for $\alpha = 2$ one gets gaussian distribution with mean $\mu$ and variance $2\sigma^2$. For $\alpha < 2$ there is no gaussian part which means that the paths of an $\alpha-$stable Lévy process are purely discontinuous in this case.

Summarizing, explicit densities are known in three cases only:

| **Parameters** | **Distribution** |
|---|---|
| $\alpha = 2, \beta = 0$ | Gaussian distribution |
| $\alpha = 1, \beta = 0$ | Cauchy distribution |
| $\alpha = 1/2, \beta = 1$ | Lévy distribution |

The usefulness of stable distributions in modern financial theory is limited for $\alpha \neq 2$ by the fact that the basic requirement of the existence of finite exponential moments is not satisfied.

When $\alpha$ is small, the process has very fat tails, and the trajectory is dominated by big jumps. Note how this graph resembles the trajectory of a compound Poisson process. When $\alpha$ is large, the behavior is determined by small jumps and the trajectory resembles that of a Brownian motion, although occasionally we see some jumps. The third graph (lower left) corresponds to the *Cauchy process* ($\alpha = 1$) which is between the two cases. Here both big and small jumps have a significant effect.

In the following the Meixner process will be introduced and studied, along with its alternative way of generation deriving from the theory of orthogonal polynomials.

# Part II

# Chapter 3

# Meixner Process: definition and properties

In this chapter the often neglected Meixner process will be introduced. It is possible to define this kind of process in basically two ways; the first is the classical one, through the definition of a background distribution, namely the Meixner distribution, a member of the generalized-$z$ distributions seen in the previous section, and the respective characteristic function. Since its infinite divisibility, the associated Lévy process remains defined in a completely natural fashion. From this, one can derive and evaluate all the properties, realizing an effective higher grade of flexibility of the process with respect to the usual Brownian motion for instance.
Another way of deriving this process will be described in the next chapter and is hidden in the integrals of a particular form of linear differential equation of the second order (namely an equation of hypergeometric type). This kind of approach provides a deeper insight in the structure of the process, and in general of most of Lévy processes listed up to this moment.
Also non trivial is the representation of this process as a subordinated Brownian motion; the theoretical problem is solved in a paper by Madan and Yor [MY06], and it is a good starting point for simulating the process.

## 3.1 Generalized $z$-distributions

In a paper by Prentice, [Pre75], a class of distributions having the following density

$$f_Z(x) = \frac{2\pi \exp\left\{\frac{2\pi\beta_1}{\alpha}(x-\mu)\right\}}{\alpha B(\beta_1,\beta_2)\left(1+\exp\left\{\frac{2\pi}{\alpha}(x-\mu)\right\}\right)^{\beta_1+\beta_2}},\ x\in\mathbb{R},$$

where $\alpha > 0, \beta_1 > 0, \beta_2 > 0, \mu \in \mathbb{R}$, and $B(\beta_1,\beta_2)$ is the Euler beta function is introduced. It is easy to check that the characteristic function of such a density is

$$\phi_Z(u) = \frac{B(\beta_1+\frac{i\alpha u}{2\pi},\beta_2-\frac{i\alpha u}{2\pi})}{B(\beta_1,\beta_2)}\exp\left(i\mu u\right),\ u\in\mathbb{R}.$$

Of course $z$ distributions have semiheavy tails and they are self-decomposable.

A probability distribution on $\mathbb{R}$ is called a generalized $z$ distribution ($GZD(\alpha,\beta_1,\beta_2,\delta,\mu)$) if

$$\phi_{GZ}(u) = \left(\frac{B(\beta_1+\frac{i\alpha u}{2\pi},\beta_2-\frac{i\alpha u}{2\pi})}{B(\beta_1,\beta_2)}\right)^{2\delta}\exp\left\{i\mu u\right\},\ u\in\mathbb{R},\delta>0.$$

$GZD(\alpha,\beta_1,\beta_2,\delta,\mu)$ is infinitely divisible with Lévy triplet $[a,0,\nu(dx)]$, where

$$a = \frac{\alpha\delta}{\pi}\int_0^{2\pi/\alpha}\frac{e^{-\beta_2 s}-e^{-\beta_1 s}}{1-e^{-s}}\,ds+\mu, \qquad \nu(dx)=v(x)dx,$$

and where

$$v(x) = \begin{cases} \dfrac{2\delta\exp\left\{-\frac{2\pi\beta_2}{\alpha}x\right\}}{x\left(1-\exp\left\{\frac{2\pi}{\alpha}x\right\}\right)}, & \text{if } x>0,\\[2ex] \dfrac{2\delta\exp\left\{-\frac{2\pi\beta_1}{\alpha}x\right\}}{|x|\left(1-\exp\left\{\frac{2\pi}{\alpha}x\right\}\right)}, & \text{if } x<0,\end{cases}$$

Let now $\{\kappa_n\}_{n\geq 1}$ be the sequence of the cumulants of $GZD(\alpha,\beta_1,\beta_2,\delta,\mu)$; $\gamma_1 = \kappa_3/\kappa_2^{3/2}$ the skewness, and $\gamma_2 = \kappa_4/\kappa_2^2$ the kurtosis; denote also

$$\nu_n(\beta_1,\beta_2) = \int_0^\infty s^{n-1}\frac{e^{-\beta_2 s}+(-1)^n e^{-\beta_1 s}}{1-e^{-s}}\,ds,\ n\geq 1$$

then the following formulae hold

$$\kappa_1 = \frac{\alpha\delta}{\pi}\nu_1(\beta_1,\beta_2)+\mu, \qquad \kappa_n = \frac{2\alpha^n\delta}{(2\pi)^n}\nu_n(\beta_1,\beta_2),\ n\geq 2;$$

$$\gamma_1 = \frac{\nu_3(\beta_1,\beta_2)}{(2\delta\nu_2^3(\beta_1,\beta_2))^{1/2}}, \qquad \gamma_2 = \frac{\nu_4(\beta_1,\beta_2)}{2\delta\nu_2^2(\beta_1,\beta_2)}$$

A particular case of $GZD$ distribution happens when $\beta_1 = \frac{1}{2} + \frac{\beta}{2\pi},\ \beta_2 = \frac{1}{2} - \frac{\beta}{2\pi}$, giving place to Meixner distribution $MD(\alpha, \beta, \delta, \mu)$, which we will see in the following.

**Definition 24.** *For all* $\alpha > 0, -\pi < \beta < \pi, \delta > 0,$ *and* $\mu \in \mathbb{R}$

$$MD(\alpha, \beta, \delta, \mu) = GZD\left(\alpha, \frac{1}{2} + \frac{\beta}{2\pi}, \frac{1}{2} - \frac{\beta}{2\pi}, \delta, \mu\right).$$

It can be observed that generalized hyperbolic distributions and generalized $z-$distributions are non intersecting sets; nevertheless it is known that $z-$distributions also can be characterized as variance-mean mixtures of Gaussian distributions with the mixing distribution $H(\lambda, \nu),\ \lambda > 0,\ \nu < \lambda^2/2$ having characteristic functions

$$\phi(x)_{H(\lambda,\nu)} = \prod_{k=0}^{\infty}\left(1 - \frac{ix}{\frac{1}{2}(\lambda + k)^2 - \nu}\right)^{-1},\ x \in \mathbb{R},$$

that are infinite convolutions of the exponential distributions with parameters $\lambda_k = \frac{1}{2}(\lambda + k)^2 - \nu,\ k \geq 0.$

## 3.2 Meixner distribution

It has been already pointed out in the preceding section that a probability distribution is called a Meixner distribution, denoted by $MD(\alpha, \beta, \delta, \mu)$ if

$$MD(\alpha, \beta, \delta, \mu) = GZD\left(\alpha, \frac{1}{2} + \frac{\beta}{2\pi}, \frac{1}{2} - \frac{\beta}{2\pi}, \delta, \mu\right)$$

with $GZD\left(\alpha, \beta_1, \beta_2, \delta, \mu\right)$ a generalized $z$ distribution as in the work of Grigelionis [Gri01]. Let now $(\Omega, \mathscr{F}, P)$ the usual probability space; the density of a random variable $X$ enjoying a Meixner distribution $MD(\alpha, \beta, \delta, \mu)$ is given by

$$f_M(x; \alpha, \beta, \delta, \mu) = \frac{(2\cos(\beta/2))^{2\delta}}{2\alpha\pi\Gamma(2\delta)} \exp\left\{\frac{\beta(x-\mu)}{\alpha}\right\}\left|\Gamma\left(\delta + \frac{i(x-\mu)}{\alpha}\right)\right|^2, \tag{3.1}$$

with $\alpha > 0, -\pi < \beta < \pi, \delta > 0, \mu \in \mathbb{R}$ and $\Gamma(\cdot)$ the Euler Gamma function.

$\mu$ and $\alpha$ are respectively location and scale parameters, while $\beta$ and $\delta$ decide the shape of the distribution.

By consequence, if $X \sim MD(\alpha, \beta, \delta, \mu)$, the variable $Z = (X - \mu)/\alpha$ enjoys a $MD(1, \beta, \delta, 0)$.

The characteristic function of $X$ is

$$\phi_{MD}(u) = E\left[e^{iuX}\right] = \left(\frac{\cos(\beta/2)}{\cosh\frac{\alpha u - i\beta}{2}}\right)^{2\delta} e^{i\mu u} \tag{3.2}$$

Let us now state and prove the main properties of a Meixner distribution.

**Proposition 9.** *$MD(\alpha, \beta, \delta, \mu)$ is infinitely divisible with Lévy characteristics $[a, 0, \nu(dx)]$ given by:*

$$a = \alpha\delta \tan\frac{\beta}{2} - 2\delta \int_1^{+\infty} \frac{\sinh(\beta x/\alpha)}{\sinh(\pi x/\alpha)}\,dx + \mu, \quad \nu(dx) = \delta \frac{\exp(\beta x/\alpha)}{x \sinh(\pi x/\alpha)}\,dx \tag{3.3}$$

*Proof*: (see Schoutens, [Sch03], pgg.$44 - 45$) denote

$$\kappa(u) = 2\delta \log\left(\frac{\cos(\beta/2)}{\cos\left(\frac{\alpha u + \beta}{2}\right)}\right) + \mu u, \quad u \in \left(\frac{-\pi - \beta}{\alpha}, \frac{\pi - \beta}{\alpha}\right)$$

the cumulant function of $X$. Since

$$\kappa'(u) = \alpha\delta \tan\frac{\alpha u + \beta}{2} + \mu,$$

from simple properties of the Gamma function (see [AS64], $6.1.30, 6.1.31$ in the generalized case, and [GR94], $6.421.1$) it can be obtained that:

$$\kappa''(u) = 2\alpha^2\delta\left(2\cos\frac{\alpha u + \beta}{2}\right)^{-2} = \frac{\alpha^2\delta}{\pi}\int_{-\infty}^{+\infty} |\Gamma(1 + ix)|^2 \exp\{x(\alpha u + \beta)\} =$$
$$= \delta \int_{-\infty}^{+\infty} \frac{x \exp(\beta x/\alpha)}{\sinh(\pi x/\alpha)} \exp(xu)\,dx.$$

From theorem $2.2$ in [BLBL92] we have that $\kappa(u)$ is the cumulant function of an infinitely divisible probability distribution having Lévy measure $\nu(dx)$, for which the proposed expression of $\nu(dx)$ above holds true, and the second element of the Lévy characteristics equal to zero.

Since the first Lévy characteristic is equal to

$$\kappa'(0) - \int_{|x|>1} x\pi(dx) = \alpha\delta \tan\frac{\beta}{2} - 2\delta \int_1^{+\infty} \frac{\sinh \beta x/\alpha}{\sinh \pi x/\alpha}\,dx + \mu,$$

we have the conclusion. □

In particular, for infinite divisibility of $MD(\alpha, \beta, \delta, \mu)$ it holds that

$$\phi_{MD}(u; \alpha, \beta, \delta, \mu) = [\phi_{MD}(u; \alpha, \beta, \delta/n, \mu/n)]^n,$$

for every $n \in \mathbb{N}$. A consequence of this fact is that

**Corollary 5.** *If $X_j \sim MD(\alpha, \beta, \delta_j, \mu_j),\ j = 1, \ldots, n$ and these random variables are mutually independent, then*

$$X_1 + \ldots + X_n \sim MD\left(\alpha, \beta, \sum_{j=1}^{n} \delta_j, \sum_{j=1}^{n} \mu_j\right).$$

**Proposition 10.** *$MD(\alpha, \beta, \delta, \mu)$ is self decomposable and has semiheavy tails.*

*Proof*: denote with

$$v(x) = \delta \frac{\exp(\beta x/\alpha)}{x \sinh(\pi x/\alpha)}, \quad x \in \mathbb{R}; \tag{3.4}$$

By the criterion in [Luk70], in order to prove that $MD(\alpha, \beta, \delta, \mu)$ is self decomposable, it has to be proven that $MD(\alpha, \beta, \delta, \mu)$ belongs to the Lévy class $L$, for which it suffices to check that, for all $x \in \mathbb{R}$

$$w(x) := -v(x) - xv'(x) \geq 0.$$

From (3.4) it can be easily seen that in our case

$$w(x) = \frac{\delta}{2\alpha}\left[(\pi - \beta)\exp\left(\frac{\pi + \beta}{\alpha}x\right) + (\pi + \beta)\exp\left(\frac{-\pi + \beta}{\alpha}x\right)\right]\left[\sinh\left(\frac{\pi x}{\alpha}\right)\right]^{-2},$$

which is nonnegative for all $x \in \mathbb{R}$, $\alpha > 0,\ \beta \in (-\pi, \pi),\ \mu \in \mathbb{R}$, and $\delta > 0$.

Semiheaviness of tails follows again from (3.4) using (3.2), and obtaining that

$$f_M(x; \alpha, \beta, \delta, \mu) = \begin{cases} C_-|x|^{\rho_-}\exp(-\eta_-|x|), & \text{as } x \to -\infty, \\ C_+|x|^{\rho_+}\exp(-\eta_+|x|), & \text{as } x \to \infty, \end{cases}$$

with

$$\rho_+ = \rho_- = 2\delta - 1,\ \eta_- = \frac{\pi - \beta}{\alpha},\ \eta_+ = \frac{\pi + \beta}{\alpha},\ C_\pm = \frac{e^{\pm\mu\eta_\pm}}{\Gamma(2\delta)}\left(\frac{2\pi}{\alpha B(\frac{\pi+\beta}{2\pi}, \frac{\pi-\beta}{2\pi})}\right)^{2\delta}$$

and $B(x,y)$ the Euler Beta function. □

It is easy to obtain the forms of the first moments of an $MD(\alpha,\beta,\delta,\mu)$; following the same notation introduced for generalized $z$ distributions it holds that

$$\kappa_1 = \alpha\delta\tan\left(\frac{\beta}{2}\right) + \mu, \qquad \kappa_2 = \frac{\alpha^2\delta}{1+\cos\beta},$$
$$\gamma_1 = \sin\left(\frac{\beta}{2}\right)\sqrt{\frac{2}{\delta}}, \qquad \gamma_2 = 3 + \frac{2-\cos\beta}{\delta}.$$

As it can be seen from the following graphics, both the skewness and the kurtosis of a Meixner distribution allow more flexibility with respect to an usual Gaussian one:

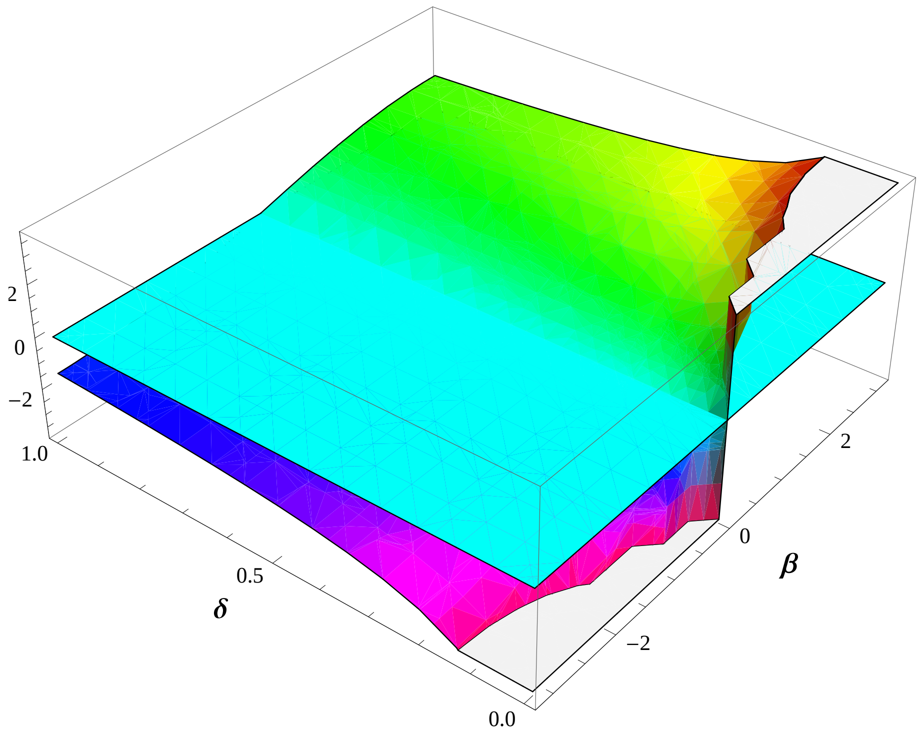


Figure 3.1: Skewness of a Meixner distribution, $-\pi < \beta < \pi$, $0 < \delta < 1$; turquoise plane $z = 0$ is the skewness of the Gaussian distribution

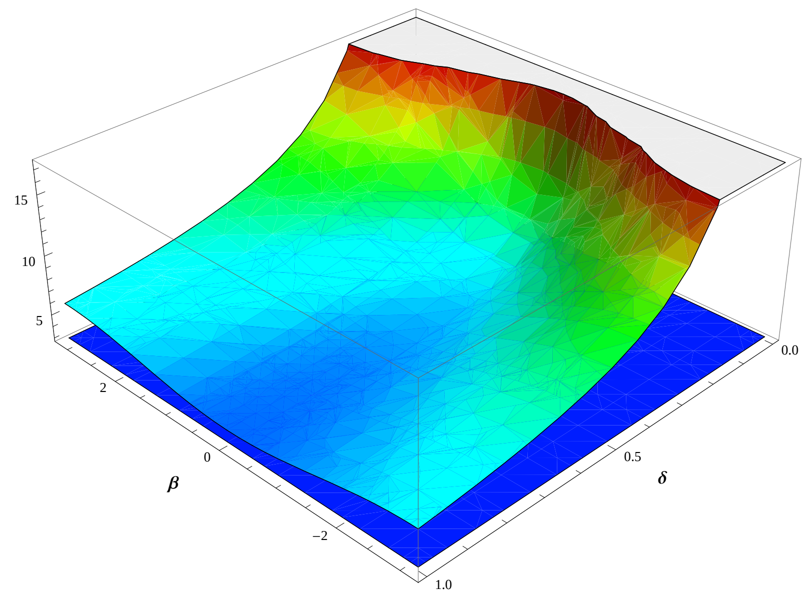


Figure 3.2: Kurtosis of a Meixner distribution, $-\pi < \beta < \pi$, $0 < \delta < 1$; blue plane $k = 3$ is the kurtosis of the Gaussian distribution

## 3.3 Estimation for the Meixner distribution

Literature regarding Meixner process usually relies on metod of moments estimated parameters, clearly for the relative simplicity of computations involved. Maximum likelihood is also possible in this case and is described for instance in a paper by Grigoletto and Provasi, [GP09], which will be followed for the section below.

### 3.3.1 Method of moments estimation

Suppose $x_1, \ldots, x_n$ a random sample drawn from $X \sim MD(\alpha, \beta, \delta, \mu)$; it is relatively simple to estimate the moments of a Meixner distribution by method of moments. Let $\bar{x}$ and $s^2$ as usual the sample mean and uncorrected variance respctively; moreover, defining $\bar{\mu}_k = n^{-1} \sum_{i=1}^{n} (x_i - \bar{x})^k$, for $k = 2, 3, 4$, let the sample skewness and kurtosis be $\bar{\gamma}_1 = \bar{\mu}_3 / \bar{\mu}_2^{3/2}$, and $\bar{\gamma}_2 = \bar{\mu}_4 / \bar{\mu}_2^2$. Then the equalities of theoretical moments with their sample counterparts leads to these relations

$$\bar{\delta} = \frac{1}{\bar{\gamma}_2 - \bar{\gamma}_1 - 3}, \qquad \bar{\beta} = \frac{\text{sgn}(\bar{\gamma}_1)}{\cos(2 - \bar{\delta}(\bar{\gamma}_2 - 3))},$$

$$\bar{\alpha} = s\sqrt{\frac{\cos\bar{\beta} + 1}{\bar{\alpha}}}, \qquad \bar{\mu} = \bar{x} - \bar{\alpha}\bar{\delta}\tan\left(\frac{\bar{\beta}}{2}\right).$$

Observe that moment estimates do not exist when $\bar{\gamma}_2 < 2\bar{\gamma}_1^2 + 3$.

### 3.3.2 Maximum Likelihood estimation

Let $x_1, \ldots, x_n$ be a random sample as above; the loglikelihood function is given by the expression

$$\begin{aligned} l_n(\alpha, \beta, \delta, \mu) =& \delta \log(2\cos(\beta/2)) - \log(2\alpha\pi) - \log(\Gamma(2\delta)) + \\ & + \beta\bar{z} + \frac{1}{n}\sum_{i=1}^{n} \log|\Gamma(\delta + iz_i)|^2, \end{aligned}$$

where

$$z_i = \frac{x_i - \mu}{\alpha}, \qquad \bar{z} = \frac{\sum_{i=1}^n z_i}{n}.$$

The MLE $\hat{\theta}_{ML}$ for the vector of parameters $\theta = (\alpha, \beta, \delta, \mu)$ is obtained by solving

$$\hat{\theta}_{ML} = \arg\max_{\theta \in \Theta} l_n(\theta)$$

with $\Theta$ the parameter space for $\theta$. For Meixner distribution is possible to compute the ML estimate $\hat{I}_n(\hat{\theta}_{ML})$ of the information matrix, since the expressions defining the first two derivatives of loglikelihood functions are explicitly available ([GP09], Appendix A), and these expressions can be used to maximize very efficiently the loglikelihood function via Newton-type algorithm based on moments estimates as starting points.

In the same paper is shown also how is analytically challenghing to verify the regularity conditions for asymptotic normality of the MLE in this framework, and is suggested to check the convergence of the estimators via Monte Carlo simulations.

By parametric bootstrap, also it can be assessed also the slow rate of convergence of the multivariate skewness and kurtosis indexes to theoretical values; in particular it resulted that the speed of convergence is inversely related to the kurtosis parameter $\delta$.
This suggest caution when performing inference based on asymptotic properties of maximum likelihood estimates for Meixner distribution, even when medium sized samples are taken into account.

## 3.4 Meixner process

Given infinite divisibility of $MD(\alpha, \beta, \delta, \mu)$, a Lévy process can be associated with it in the fashion seen in previous chapter, which is called the Meixner process.
More precisely, a Meixner process $X = \{X(t), t \geq 0\}$ is a stochastic process starting at zero, with independent and stationary increments, and with

$$X(t) \sim MD(\alpha, \beta, \delta t, \mu).$$

The characteristic function of a Meixner process will be trivially given by

$$\phi_{MP}(u; \alpha, \beta, \delta, \mu) = E\left[e^{iuX(t)}\right] = \left(\frac{\cos(\beta/2)}{\cosh\frac{\alpha u - i\beta}{2}}\right)^{2\delta t} e^{i\mu u}, \quad (3.5)$$

where $\alpha > 0, -\pi < \beta < \pi, \delta > 0, \mu \in \mathbb{R}$.
A standard notation which will be adopted sometimes from now on is $X(t) = MP(\alpha, \beta, \delta, \mu)$.
From proposition 9 it is immediately deduced that a Meixner process has no Brownian part and a pure jump part governed by the Lévy measure $\nu(dx)$, as seen in (3.3).
Because

$$\int_{-1}^{+1} |x|\nu(dx) = \infty,$$

Meixner process is of infinite variation.

So we have a Lévy process

- with no Brownian component, pure jump;

- with moments of all orders;
- of infinite variation;
- with semiheavy tails;
- selfdecomposable.

### 3.4.1 Meixner process as a subordinated Brownian Motion

As we have explained, some Lévy processes, like the Variance Gamma process and the Normal Inverse Gaussian process are known by alternative construction as time-changed Brownian motions; other processes as the CGMY process or the Meixner process are directly identified by their Lévy measure and it is not clear a priori whether it can be represented as a time changed Brownian motion. The problem has been solved in a work by Madan and Yor, [MY06], in which a complete characterization of Meixner process as a time-changed Brownian motion can be found.

**Lévy measure of a subordinated Brownian motion**

Suppose the Lévy process $X(t)$ is obtained by subordinating a Brownian motion with drift (i.e. the process $\theta u + W(u)$, for $\{W(u),\ u \geq 0\}$ a standard Brownian motion) by an independent subordinator $Y(t)$ with Lévy measure $\nu(dy)$. By a result in Sato, [Sat99], $(30.8)$, pg.198, the Lévy measure of the process $X(t)$ is given by $\mu(dx)$, where

$$\mu(dx) = \int_0^\infty \frac{1}{\sqrt{2\pi y}} \exp\left\{-\frac{(x-\theta y)^2}{2y}\right\} \nu(dy)\, dx.$$

**Absolute continuity criterion for subordinators**

Let $T_A, T_B$ be two subordinators; the law of the subordinator $T_A$ is absolutely continuous with respect to the subordinator $T_B$, on finite intervals, just if there exists a function $f(t)$ such that the Lévy measures $\nu_A(dt), \nu_B(dt)$ for the processes $T_A$ and $T_B$ respectively are related by

$$\nu_A(dt) = f(t)\nu_B(dt),$$

and furthermore, (see [Sat99], thm.33.1)

$$\int_0^\infty \left(\sqrt{f(t)} - 1\right)^2 \nu_B(dt) < \infty$$

**Explicit time change for Meixner process**

Already previously given in (3.3) and (3.5) were the Lévy measure $\nu(dx)$ and the characteristic function for the Meixner process. The proof of the chance of writing this process as a time changed Brownian motion is given in the already cited paper my Madan and Yor ([MY06], from pg.20 on): here it is clarified how it is necessary to identify the Lévy measure $l(u)$ of a subordinator such that

$$\begin{aligned}\nu(dx) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi y}} \exp\left\{-\frac{(x-Ay)^2}{2y}\right\} l(y)\,dy = \\ &= e^{Ay} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi y}} \exp\left\{-\frac{x^2}{2y} - \frac{A^2 y}{2}\right\} l(y)\,dy.\end{aligned}$$

Setting $A = \beta/\alpha$ the following must hold for a suitable $l(u)$:

$$\frac{\delta}{x \sinh\left(\frac{\pi x}{\alpha}\right)} = \int_0^{+\infty} \frac{1}{\sqrt{2\pi y}} \exp\left\{-\frac{x^2}{2y} - \frac{A^2 y}{2}\right\} l(y)\,dy$$

With somewhat delicate algebra one obtains

$$l(u) = \frac{\delta\alpha}{\sqrt{2\pi u^3}} g(u),$$

where

$$g(u) = P\left(M_1^{(3)} \geq C\sqrt{u}\right) \exp\left\{-\frac{A^2 u}{2}\right\},$$

with

$$\frac{1}{\left[M_1^{(3)}\right]^2} = T_1^{(3)} \overset{d}{=} \frac{1}{\left(\max_{t\leq 1} R_t^{(3)}\right)},$$

and $R_t^{(3)}$ the $BES(3)$ process.

For the absolute continuity of the subordinator with respect to the one sided stable $1/2$ subordinator, it is required, and easily verified that

$$\int \frac{1}{\sqrt{u^3}} \left(\sqrt{g(u)} - 1\right)^2 du < \infty$$

Also, for the simulation of Meixner process as a time changed Brownian motion it is possible to represent (see Pitman and Yor's paper [PY03])

$$P\left(M_1^{(3)} \geq C\sqrt{u}\right) = \sum_{n=-\infty}^{+\infty} (-1)^n \exp\left\{-\frac{n^2\pi^2}{2C^2u}\right\}.$$

**Simulation of the Meixner process**

The first step is to simulate the jumps of the one sided stable $1/2$ with Lévy density

$$k(x) = \frac{\delta\alpha}{\sqrt{2\pi x^3}}, \quad x > 0.$$

The small jumps of the subordinator are approximated using the drift

$$\zeta = \delta\alpha\sqrt{\frac{2\varepsilon}{\pi}}$$

while the arrival rate for the jumps above $\varepsilon$ is

$$\lambda = \delta\alpha\sqrt{\frac{2}{\pi\varepsilon}}$$

and the jump sizes for the one sided stable $1/2$ are

$$y_j = \frac{\varepsilon}{u_j^2}$$

for an independent uniform sequence $\{u_j\}$. Then the function $g(y)$ at the point $y_j$ is evaluated, and the time change variable is defined as

$$\tau = \zeta + \sum_j y_j \mathbb{I}_{\{g(y_j) > w_j\}},$$

for yet another independent uniform sequence $\{w_j\}$. It can also be observed that the function $g(y)$ only uses the parameters $\alpha, \beta$ and is independent of the parameter $\delta$.

Finally the value of the Meixner random variable or equivalently the unit time level of the process is then generated as

$$X = \frac{\beta}{\alpha}\tau + \sqrt{\tau}Z, \tag{3.6}$$

where $Z$ is an independent standard normal random variable.

### 3.4.2 Particular characterizations of Meixner process

This section is based on the already cited paper by Pitman and Yor [PY03], which particularly deals with the characterization of Lévy processes associated with hyperbolic functions; so for the sake of simplicity, remembering the presence of an hyperbolic cosine in the characteristic function of Meixner process, let us discuss about the case

$$\hat{C}(t) = \hat{C}_t = MP(2, 0, 1/2, 0)$$

as a representative example, i.e. the process identified, for $t \geq 0$ and $\theta \in \mathbb{R}$, by the characteristic function

$$E[\exp(i\theta\hat{C}_t)] = E\left[\exp\left(-\frac{1}{2}\theta^2 C_t\right)\right] = \left(\frac{1}{\cosh\theta}\right)^t;$$

this firstly shows how process $\hat{C}_t$ can be constructed from $C_t$ by Brownian subordination

$$\hat{C}_t = W(C_t) \tag{3.7}$$

for $\{W(t), t \geq 0\}$ a standard Brownian motion; the law of the subordinator $C_t$ arises in several different contexts, especially in the study of Brownian motion and Bessel process; in particular the distribution of $C_1$, for instance, is that of the hitting time of $\pm 1$ by the one-dimensional Brownian motion $W$, while the distributions of $C_t,\ t = 1, 2$ are also of significance in analytic number theory, due to the Mellin representation of the entire function

$$\xi(s) := \frac{1}{2}s(s-1)\left(\frac{1}{\pi}\right)^{s/2}\Gamma(s/2)\zeta(s),$$

where $\zeta(s) := \sum_{n=1}^{+\infty} n^{-s}$, $(Re\ s > 1)$ is the Riemann's $\zeta$ function, and the entire function

$$\xi_4(s) := \left(\frac{4}{\pi}\right)^{(s+1)/2} \Gamma\left(\frac{s+1}{2}\right) L_{\chi_4}(s),$$

where $L_{\chi_4}(s) := \sum_{n=0}^{+\infty}(-1)^n(2n+1)^{-s}$, $(Re\ s > 0)$ is the Dirichlet series associated with the quadratic character modulo $4$. The function $\xi_4(2s+1)$ appear as the Mellin transform of $(\pi/2)C_1$, and the Mellin transform of $C_2$ is simply related to $\xi$.

Here are some non trivial properties of process $\hat{C}_t$:

**Theorem 24.** *The process $\hat{C}_t$ is the unique (i.e. unique in law) Lévy process satisfying the following moment recurrence for $t > 0$ and $Re\ s > -1/2$:*

$$(t^2+t)E[|\hat{C}_{t+2}|^{2s}] = t^2E[|\hat{C}_t|^{2s}] + E[|\hat{C}_t|^{2s+2}]. \tag{3.8}$$

This fact can be proven via two relations holding for a standard Brownian motion $W_t$: from (3.7) and Brownian scaling $\hat{C}_t \overset{d}{=} W_1\sqrt{C_t}$ it holds that

$$E[|\hat{C}_t|^{2s}] = E[|W_1|^{2s}]E[C_t^s];$$

moreover, because of the identity in distribution $W_t^2 \overset{d}{=} 2y\Gamma_{1/2}$, with $\Gamma_t$ a Gamma process, and $E[\Gamma_t^s] = \Gamma(t+s)/\Gamma(t)$, for $Re\ s > -t$,

$$E[|W_t|^{2s}] = (2t)^s\frac{\Gamma(\frac{1}{2}+s)}{\Gamma(\frac{1}{2})} = 2\left(\frac{t}{2}\right)^s\frac{\Gamma(2s)}{\Gamma(s)}, \quad \text{for } Re\ s > -1/2$$

and so, by letting $\Gamma(x+1) = x\Gamma(x)$,

$$E[|W_1|^{2(s+1)}] = (2s+1)E[|W_1|^{2s}], \quad \text{for } Re\ s > -1/2.$$

Also

**Theorem 25.** *The density*

$$f_{\hat{C}_t}(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left(\frac{1}{\cosh y}\right)^t e^{iyx}\,dy$$

*satisfies the recurrence relation*

$$t(t+1)f_{\hat{C}_{t+2}}(x) = (t^2+x^2)f_{\hat{C}_t}(x) \tag{3.9}$$

Formula (3.9) for $t = 1, 2, \ldots$ was given in [Mor82]; as shown there, $f_{\hat{C}_t}(x)$ is a polynomial of degree $t$ divided by $\cosh(\pi x/2)$ if $t$ is an odd integer, and by $\sinh(\pi x/2)$ if $t$ is even. Also, for the classical representation of the Euler beta function, it holds that

$$f_{\hat{C}_t}(x) = \frac{2^{t-2}}{\pi} B\left(\frac{t+ix}{2}, \frac{t-ix}{2}\right) = \frac{2^{t-2}}{\pi\Gamma(t)} \left|\Gamma\left(\frac{t+ix}{2}\right)\right|^2.$$

A slight generalization of (3.8) can be written as

$$(t^2 + t)E[g(\hat{C}_{t+2})] = t^2 E[g(\hat{C}_t)] + E[\hat{C}_t^2 g(\hat{C}_t)], \tag{3.10}$$

where $g$ is an arbitrary bounded Borel function. This one follows from (3.8) first for symmetric $g$ by uniqueness of Mellin transform, then for general $g$ by using

$$E[g(\hat{C}_t)] = E[g(-\hat{C}_t)] = E[\tilde{g}(|\hat{C}_t|)], \quad \text{with } \tilde{g}(x) = \frac{g(x) + g(-x)}{2}.$$

Let now $M^{(a)}$ be the Meixner process whose marginal laws are derived from those of $\hat{C}$ by exponential tilting (Esscher transform, see the following chapter 4), according to the formula

$$E[g(M_t^{(a)})] = (\cos a)^t E[g(\hat{C}_t)\exp(a\hat{C}_t)], \quad t \geq 0, \ -\pi/2 < a < \pi/2$$

The functional recurrence relation (3.10) for $\hat{C}$ generalizes immediatly to show that $X = M^{(a)}$ satisfies the following functional recursion

**Theorem 26.** *A Lévy process $X$ satisfies the functional recursion*

$$c(t^2 + t)E[g(X_{t+2})] = t^2 E[g(X_t)] + E[X_t^2 g(X_t)], \tag{3.11}$$

*for all bounded Borel functions $g$ and all $t \geq 0$, for some constant $c$ if and only if $X$ is a Meixner process $M^{(a)}$ for some $a \in (-\pi/2, \pi/2)$; then $c = 1/\cos^2 a \geq 1$ and recurrence relation above holds for all Borel $g$ such that the expectations involved are well defined and finite.*

*Proof*: [PY03] suppose that process $X$ satisfies (3.11). By consideration of (3.11) for constant function $g$, it is obvious that $E(X_1^2) < \infty$ and $c = 1 + E(X_1)^2$. Hence $c \geq 1$ and $X_1$ has characteristic function $g$ with two continuous derivatives $g'$ and $g''$.

Consider now $g(x) = e^{i\theta x}$ in (3.11) to obtain the following identity of functions of $\theta$:

$$c(t^2+t)g^{t+2} = t^2 g^t - (g^t)'' = t^2 g^t - (t^2-t)g^{t-2}(g')^2 - tg^{t-1}g''$$

where all the differentiations are with respect to $\theta$ and for instance $g^t(\theta) = [g(\theta)]^t$. Cancellig the common factor of $g^t$ and equating coefficients of $t^2$ and $t$, this amounts to the pair of equalities

$$\left(\frac{g'}{g}\right)' = -cg^2 = \left(\frac{g'}{g}\right)^2 - 1.$$

The argument is completed by the following elementary result: *the unique solution $g$ of the problem*

$$\begin{cases} \left(\frac{g'}{g}\right)' = \left(\frac{g'}{g}\right)^2 - 1 \\ g(0) = 1 \\ g'(0) = i\tan(\varphi) \qquad for \quad \varphi \in (-\pi/2, \pi/2) \end{cases}$$

*is*

$$g(\theta) = \frac{\cos(\varphi)}{\cosh(\theta + i\varphi)}.$$

□

**Corollary 6.** *The process $X = \hat{C}$ is the unique Lévy process such that either*

*i) the moment recursion* (3.8) *holds for all $s = 0, 1, 2, \ldots$ and the distribution of $X_1$ is symmetric about* 0*, or*

*ii) the functional recursion* (3.11) *holds with $c = 1$ for all bounded Borel functions $g$.*

Let now be $\{\Gamma_{n,t}, t \geq 0\}$ a sequence of independent gamma processes, and consider for $\alpha > 0$ the subordinator $\{\Sigma_{\alpha,t}, t \geq 0\}$ defined by the follwing weighted sum of these processes

$$\Sigma_{\alpha,t} = \frac{2}{\pi}\sum_{n=0}^{\infty} \frac{\Gamma_{n,t}}{(\alpha+n)^2}, \quad t \geq 0.$$

The weights are chosen so by the expression of the characteristic function of a gamma process $E[\exp(-\lambda\Gamma_t)] = (1+\lambda)^{-t}$ one has

$$E\left[\exp\left(-\frac{1}{2}\theta^2\Sigma_{\alpha,t}\right)\right] = \begin{cases} \left(\dfrac{1}{\cosh\theta}\right)^t & \text{if } \alpha = 1/2, \\ \left(\dfrac{\theta}{\sinh\theta}\right)^t & \text{if } \alpha = 1. \end{cases}$$

Thus

$$C_t \stackrel{d}{=} \Sigma_{1/2,t} \tag{3.12}$$

and also holds true that

$$W_{\Sigma_{\alpha,1}} \stackrel{d}{=} \pi^{-1}\log\frac{\Gamma_\alpha}{\Gamma'_\alpha},$$

where $\Gamma_\alpha \stackrel{d}{=} \Gamma'_\alpha$ and independent. For $\alpha = 1/2$, (3.12) describes the distribution of $\hat{C}_1$.

### Lévy measures

For a Lévy process $X$ whose Lévy measure is $\Lambda_X$, let its density be $\rho_X(x) = \Lambda_X(dx)/dx$ and $\hat{X}_t = W_{X_t}$; it can be shown that the follwing formulae hold

| $\hat{X}$ | $\rho_{\hat{X}}(x) = \frac{\Lambda_{\hat{X}}(dx)}{dx}$ | $\int_{-\infty}^{+\infty} \lvert x\rvert^{2s}\rho_{\hat{X}}(x)\,dx$ | $\kappa_{2n}(\hat{X}_1) = \int_{-\infty}^{+\infty} x^{2n}\rho_{\hat{X}}(x)\,dx$ |
|---|---|---|---|
| $\hat{C}$ | $\dfrac{1}{2x\sinh(\pi x/2)}$ | $(4^s-1)\dfrac{2\Gamma(2s)}{\pi^{2s}}\zeta(2s)$ | $(4^n-1)\dfrac{4^n}{2n}\lvert B_{2n}\rvert$ |

where $B_{2n}$ are the rational Bernoulli numbers. A further non trivial characterization can be the following:

**Theorem 27.**

*i) Let $X$ be a random variable with all moments finite and all odd moments equal to $0$; then*

$$X_t \stackrel{d}{=} \hat{C}_2 \iff \kappa_{n+2}(X) = 2E[X^{2n}], \quad n \in \mathbb{N};$$

*ii) Let $X$ be a rabndom variable with all moments finite; then*

$$X_t \stackrel{d}{=} C_2 \iff \kappa_{n+1}(X) = \frac{E[X^n]}{n+1/2}, \quad n \in \mathbb{N}.$$

Thus, moreover,

**Theorem 28.** *Let $\{X(t), t \geq 0\}$ the Lévy process associated with a finite Kolmogorov measure $K_X$ via the standard Kolmogorov representation*

$$E[e^{i\theta X(t)}] = \exp(t\Psi(\theta)), \textit{ with } \quad \Psi(\theta) = i\theta c + \int (e^{i\theta x} - 1 - i\theta x)x^{-2}K(dx), \tag{3.13}$$

*$c \in \mathbb{R}$, and the integrand function defined as $-\theta^2/2$ for $x = 0$. Let also $U \sim U[0,1]$ random variable, independent of $X_2$. Then*

*i) For each fixed $t > 0$, assuming that the distribution $F_X(x)$ of $X_t$ is symmetric,*

$$X_t \stackrel{d}{=} \hat{C}_t \Longleftrightarrow K_X(dx) = F_{X_2}(x)dx \tag{3.14}$$

*ii) For each fixed $t > 0$, without the symmetry assumption,*

$$X_t \stackrel{d}{=} \hat{C}_t \Longleftrightarrow K_X(dx) = x^2 F_{U^2 X_2}(x)dx$$

**Definition 25.** *A Lévy process $\{X_t\}$ is self-generating if its Kolmogorov measure $K_X$ is a scalar multiple of the distribution of $X_u$ for some $u \geq 0$:*

$$\frac{K_X(dx)}{K_X(\mathbb{R})} = F_{X_u}(x)dx; \tag{3.15}$$

to indicate the value of $u$ and to shorthand, let us say $X$ is $SG(u)$. In particular $X$ is $SG(0)$ if and only if $\psi(\theta) = i\theta c + \sigma^2\theta^2/2$, i.e. $X$ is a Brownian motion with drift $c$ and variance $\sigma^2$. From Kolmogorov representation (3.13) $X(t)$ is $SG(u)$ if and only if

$$\frac{\psi''(\theta)}{\psi''(0)} = \exp(u\psi(\theta)) \tag{3.16}$$

To restate formula (3.14) above there is the following

**Theorem 29.** *The process $X = \hat{C}$ is the unique symmetric $SG(2)$ Lévy process with $E[X_1^2] = 1$.*

It is easily seen that for $u > 0,\ a > 0,\ b \neq 0$

$$\{X_t, t \geq 0\} \text{ is } SG(u) \Longleftrightarrow \{aX_{bt}, t \geq 0\} \text{ is } SG(u/b).$$

Also, if $X$ is $SG(u)$ and the moment generating function $M(\xi)$ is finite for some $\xi \in \mathbb{R}$, then the exponentially tilted process (Esscher transform) $\{X_t^{(\xi)}, t \geq 0\}$ with

$$P(X_t^{(\xi)} \leq x) = \frac{e^{\xi x} P(X_t \leq x)}{M^t(\xi)}$$

is easily seen to be $SG(\xi)$. The self generating Lévy process obtained from $\hat{C}$ by these operation of scaling and exponential tilting have been called generalized hyperbolic secant processes, but this is nothing but a different definition of the already introduced Meixner process. A theorem that states the exhaustion of the family of $SG$ processes is the following

**Theorem 30.** *The only Lévy processes $\{X_t\}$ with the self-generating property (3.15) for some $u \geq 0$ are Brownian motions (for $u = 0$), and Meixner and Gamma processes (for $u > 0$).*

*Proof*: the characterization for $u = 0$ is easy, so consider $X$ which is $SG(u)$ for some $u > 0$. Observe first that $X$ cannot have a Gaussian component, or equivalently that $K_X$ has no mass at $0$.

For a Gaussian component would make $X_u$ have a density, implying $P(X_u = 0) = 0$, in contradiction to (3.15).

Similarly $X$ cannot have a finite Lévy measure because then $P(X_u = 0) > 0$, which would force $K_X$ to have an atom at $0$. By use of the following scaling transformation: for $u > 0,\ a > 0,\ b \neq 0$

$$\{X_t,\ t \geq 0\} \text{ is } SG(u) \Longleftrightarrow \{aX_{bt},\ t \geq 0\} \text{ is } SG(u/b),$$

the problem of characterizing all Lévy processes which are $SG(u)$ for arbitrary $u > 0$ is reduced to the problem of characterizing all Lévy processes that are $SG(u)$ for some particular $u$, and the choice $u = 2$ is most convenient.

Also, by a suitable choice of $a$ in (3.4.2) we can reduce to (3.16) with $\psi''(0) = -1$. So it is enough to find all caracteristic exponent $\psi(\theta)$ such that

$$-\psi''(\theta) = \exp(2\psi(\theta)), \quad \text{with} \quad \psi(0) = 0. \tag{3.17}$$

Set now

$$D(\theta) = \frac{1}{E[\exp(i\theta X_1)]} = \exp(-\psi(\theta)) \tag{3.18}$$

so (3.17) is equivalent to

$$DD'' - (D')^2 = 1 \quad \text{with} \quad D(0) = 1. \tag{3.19}$$

The general solution of (3.19) is

$$D_b(\theta) = \frac{\cosh(\theta \cosh(b) + b)}{\cosh(b)}$$

for some $b \in \mathbb{C}$, including the limit case when $\cosh(b) = 0$. In particular for $b = ia$, with $a \in (-\pi/2, \pi/2)$, we find

$$D_{ia}(\theta) = \frac{\cosh(\theta \cos(a) + ia)}{\cos(a)}, \tag{3.20}$$

corresponding to Meixner process, and the limit case $a = \pm\pi/2$ corresponds to $\pm\Gamma$, for $\Gamma$ the standard Gamma process.

Other choices of $a \in \mathbb{R}$ lead to the same examples, due to symmetries of $\cosh$ and $\cos$.

To complete the proof it suffices to show that $1/D_{ia}(\theta)$ is not an infinitely divisible characteristic function if $a \notin \mathbb{R}$.

For $D$ derived by (3.18) from a Lévy process $X$ we have

$$D'(0) = -i\mu \quad \text{where} \quad \mu = E[X_1] \in \mathbb{R}$$

whereas $D'_{ia}(0) = \sinh(ia) = i\sin(a)$.

This eliminates the case when $\sin(a) \notin \mathbb{R}$ and it remains to deal with the case $\sin(a) \in \mathbb{R} \setminus [-1, 1]$.

In this case $\cos^2(a) = 1 - \sin^2(a) < 0$ implying that $\cos(a) = i\nu$ for some real $\nu \neq 0$. But then for $\cosh$ is $2i\pi-$periodic the function $D_{ia}(\theta)$ in (3.20) is $2\pi/\nu$, hence so is $1/D_{ia}(\theta)$.

If $1/D_{ia}(\theta)$ were the characteristic function of $X_1$, the Lévy measure of $X$ would be finite. But then $X$ could not be self-generating, as remarked at the beginning of the proof. □

# Chapter 4

# Esscher Transform

## 4.1 Introduction

As previously stated, Lévy processes provide a lot of flexibility in financial modeling: although financial returns increments exhibit some kind of serial dependence, many of their essential features are captured by this class of models: heavy tails, aggregational gaussianity and volatility clustering for instance, are some of their features easily described by means of models based on Lévy processes.

But introduction of jumps always rises the problem of dealing with incomplete market models; that means that there exist infinitely many martingale measures, compatible with the no-arbitrage requirement and equivalent to the physical measure describing the underlying evolution one can use to price derivative securities.

One reasonable way to solve this problem, is based on the observation that in incomplete markets the "correct" equivalent martingale measure could not be independent on the preferences of investors any more, so by guessing a suitable utility function describing these preferences, an "optimal" equivalent martingale measure should maximize the expected value of this utility. It has been proved that for many interesting cases of utility functions this problem admits a dual formulation: finding an equivalent martingale measure maximizing some class of utility functions is in fact equivalent to find an equivalent martingale measure minimizing some kind of distance (see for instance Bellini and Frittelli, [BF02]).

Another popular approach to option pricing for incomplete market models had been related to the construction of the Esscher martingale transform. As it has been already pointed out by Kallsen and Shiryaev in [KS02], two different Esscher martingale transforms exist for Lévy processes according to the choice of the parameter which defines the measure: one turns the ordinary exponential process into a martingale, and another one turns into a martingale the stochastic exponential. They have been called the *Esscher martingale transform for the exponential process* and the *Esscher martingale transform for the linear process* respectively.

It has also been shown by Esche and Schweizer in [ES05] that for exponential Lévy models the Esscher martingale transform for the linear process is also the minimal entropy martingale measure, i.e. the equivalent martingale measure which minimizes te relative entropy, and that this measure has also the property of preserving the Lévy structure of the model (see Hubalek and Sgarra, [HS06]).

Some examples of these procedures have been illustrated for instance by Fujiwara and Miyahara in [FM03]. The purpose of this part of the work is to try to fill in the gap on the same topic for the Meixner process.

## 4.2 Equivalent martingale measure: meaning.

**Definition 26.** *Let $(\Omega, \mathcal{F}_T, P)$ an usual probability space; a probability measure $Q$ defined on $(\Omega, \mathcal{F}_T)$ is an* equivalent martingale measure *if:*

1. *$Q \sim P$, i.e. if $\mathcal{N}_Q$ is the set of all null sets for measure $Q$, it holds that $\mathcal{N}_Q = \mathcal{N}_P$;*
2. *the discounted stock price process $\tilde{S}_t = \{\exp(-rt)S_t, t \geq 0\}$ is a martingale with respect to $Q$.*

The fundamental result in assessing a necessary and sufficient condition for the existence of such an equivalent probability measure, is given by Delbaen and Schachermayer in [DS94], and it is related to the absence of arbitrage in the underlying market. In particular, the existence of an equivalent martingale measure implies the absence of arbitrage, but for the converse it is required to add that it shouldn't be possible to construct an approximation to an arbitrage opportunity in some limiting sense.

The existence of such a measure is important for it allows to reduce option pricing to calculating the expected values of the discounted payoffs not with respect to the physical measure $P$ but with respect to $Q$. Working under $Q$ means, as it is usually said, working "in a risk-neutral world", since the expected return of the stock under $Q$ equals the risk-free return of the bank account:

$$E_Q[S_t|\mathcal{F}_0] = \exp(rt)S_0.$$

An equally important financial problem is that of hedging. A contingent claim can be perfectly hedged if there is a strategy which can replicate the claim, in the sense of the existence of a self-financing dynamic portfolio, investing in a bank account and a stock, whose value at any time point matches the value of the claim. Moreover, the strategy must be admissible, meaning that the value of portfolio must be bounded from below by a constant.

**Definition 27.** *A market model is called* complete *if for every integrable contingent claim there exists an admissible self-financing strategy replicating the claim.*

The issue of completeness is related to the uniqueness of the equivalent martingale measure introduced above. It is in turn linked with the so called predictable representation property of a martingale.

**Definition 28.** *A martingale $M$ is said to have the predictable representation property if for any square-integrable random variable $H(\in \mathcal{F}_T)$ it holds that*

$$H = E[H] + \int_0^T a_s \, dM_s, \tag{4.1}$$

*for some predictable process* $a = \{a_s, s \in [0, T]\}$.

Holding (4.1), the process $a$ gives the necessary self-financing admissible strategy.
The problem is that (4.1) holds just for very few martingales, among which the Brownian motion and the compensated Poisson process for instance.
The uniqueness of an equivalent martingale measure implies property (4.1) which in turn implies market completeness; vice versa not necessarily holds (see [Sch03] for further details). Property (4.1) for Brownian motion implies the completeness of the usual Black and Scholes model.

## 4.3 Esscher Transform: motivation

Both insurance and finance are interested in the fair pricing of financial products. In general the more an insurance market is liquid (situation which can be associated to many potential offers of insurance and deregulated markets, for instance) the more a "correct", "fair" price may be expected to emerge.
For example, in the case of car insurance, depending on the different characteristics of the drivers, a so-called "net premium" is evaluated which should cover the expected losses over the period of contract. To this premium, various loading factors are added (costs, market fluctuations...); the resulting gross premium is also subject to market forces which imply that a market-conform premium is finally charged.
Very important in the process of determining the above premium is the attitude of both parties involved towards risk, which can be generally described, as within the more economic literature, through the notion of utility.
Utility theory provides a method to give insight into decision making in conditions of uncertainty.
An alternative economic tool is equilibrium theory.

Depending on the adopted economic theory, different possible premiums may result: one of these is the Esscher principle. Rather than being based on the expected loss, the Esscher principle starts from the expectation of the loss under an exponentially transformed distribution, properly normalized.

Besides the pricing of individual risks, more complicated insurance producs involve time, and hence are based on specific stochastic processes. The classical insurance risk processes are of the compound Poisson type or their generalizations (like mixed and doubly stochastic compound Poisson processes); the main feature of these processes, making them distinct from typical diffusion-type models in finance, is their jump structure. Indeed, when turning to fair pricing in finance, the standard reasoning uses the so called no-arbitrage approach, basically stating that there cannot be such thing as a riskless gain, which, if precisely formulated, brings in as said, the fundamental notion of risk neutral martingale measure or equivalent martingale measure.

If the set of equivalent risk-neutral measures is not reduced to one point, then finding the hedging admissible strategies is no longer possible. The initial investment needed to reproduce the contingent claim is not defined, and in this sense there is no natural price for the claim under consideration.

Due to the jump structure of standard risk processes, we find ourselves in the so called incomplete market case.

But the introduction of jump processes has a more practical motivation.

Normality of asset returns has played a central role in financial theory; the normality of distribution has been augmented with the assumption of continuity of trajectories when Samuelson introduced in 1965 the geometric Brownian motion, then used in the first papers by Black-Scholes and Merton (1973).

As documented in a considerable number of papers written by academics and practitioners, both normality and continuity assumptions are contradicted by the data in several pieces of evidence. Return distributions are more leptokurtic than the normal one as noted by Fama as early as

$1963$; this feature is more accentuated when the holding period becomes shorter, and becomes particularly clear on high frequency data. Option prices also exhibit the so called volatility smile as well as prices higher than predicted by the Black-Scholes formula for short-dated options.
At the same time jumps may be clearly identified in equity data; in fact, the inability to trade continuously implies de facto jumps in prices. These jumps contribute or may be the source of stochastic volatility while they lead to finite variation trajectories in the absence of a diffusion term, as observed in practice. As a consequence, risk cannot be fully hedged away and in most cases there will be infinitely many such equivalent martingale measures so that pricing is directly linked to an attitude towards risk.
So the question shifts from "which premium principle to use?", for classical insurance, to "which equivalent martingale measure to use?", in the incomplete market financial context.
This is exactly where the Esscher transform enters as one of the possible pricing candidates.

## 4.4 General Theoretical overview

The Esscher transform was developed to approximate the aggregate claim amount distribution around a point of interest, say $x_0$, by applying an analitic approximation (the Edgeworth series) to the transformed distribution with the parameter $\theta$ chosen such that the new mean is equal to $x_0$.
The Esscher transform can be also readily extended to stochastic processes including those commonly used to model stock-price movements; the parameter $\theta$ is determined so that the modified probability measure is an equivalent martingale measure, with respect to which the prices of securities are expected discounted payoffs.

The starting point is the usual stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, P)$ endowed with a $d-$semimartingale $X$ with characteristics $(B, C, \nu)$; the aim is to construct another measure $P'$, locally equivalent to $P$ in such a way that $X$ has some specified properties under $P'$, like being a local martingale.

This problem transates to constructing the density process $Z_t = dP'_t/dP_t$ in such a way that the new characteristics $(B', C', \nu')$ of $X$ under $P'$ have some given properties, using the connection between the characteristics and the density process provided by Girsanov's theorem: this is in general a difficult "martingale problem".

As anticipated, a "natural" possible way of solving this problem was introduced by Esscher [Ess32] in connection with some actuarial problems.

Here follows the illustration of the method on a simple "random variable" problem, where no filtration is involved.

Let $X$ be a real-valued random variable defined on a probability space $(\Omega, \mathscr{F}, P)$ such that $P(X > 0) > 0$ and $P(X < 0) > 0$.

*Problem*: construct a measure $P'$ equivalent to $P$ such that $E_{P'}[X] = 0$.

From a practical point of view, the idea of Esscher is the following: construct a measure $Q \sim P$ by

$$Q(d\omega) = ce^{-X(\omega)^2} P(d\omega),$$

where $c$ is the normalizing constant $c = 1/E[e^{-X^2}]$; then let $\phi(\theta) = E_Q[e^{\theta X}]$ for $\theta \in \mathbb{R}$, and finally

$$Z_\theta(\omega) = \frac{e^{\theta X(\omega)}}{\phi(\theta)} \left(= e^{\theta X(\omega) - K(\theta)}, \text{ with } K(\theta) = \log \phi(\theta)\right).$$

**Definition 29.** *The map $x \mapsto e^{\theta x}/\phi(\theta)$ is called the Esscher transform.*

It is easy to conclude from the construction that $E_Q[Z_\theta] = 1$ and that $Q \sim P$.

Now the measures $P'_\theta$ are costructed via the

$$P'_\theta(d\omega) = Z_\theta(\omega) Q(d\omega) = \frac{e^{\theta X(\omega)}}{\phi(\theta)} Q(d\omega).$$

It is clear that

$$E_{P'}[X] = E_Q\left(\frac{Xe^{\theta X}}{\phi(\theta)}\right) = \frac{\phi'(\theta)}{\phi(\theta)},$$

where $\phi'$ is the derivative of $\phi$. The two assumptions $P(X > 0) > 0$ and $P(X < 0) > 0$ imply that the strictly convex function $\phi$ reaches its minimum at a unique point $\theta'$ , so the equation $\phi'(\theta) = 0$ has $\theta'$ for its unique solution.

Then defining $P' = P'_{\theta'}$, it is $P' \sim P$ and $E_{P'}[X] = 0$.

Let $S(t)$, for $t \geq 0$, denote the price of a non-dividend-paying stock or security at time $t$, and assume there is a Lévy process $X = \{X(t), t \geq 0\}$ with stationary and independent increments and $X(0) = 0$ such that

$$S(t) = S_0 e^{X(t)}, \quad t \geq 0. \tag{4.2}$$

For each $t$ the random variable $X(t)$, has an infinitely divisible distribution. Let also

$$F_t(x) = P(X(t) \leq x)$$

be its c.d.f., and

$$M_t(z) = E[e^{zX(t)}]$$

its moment generating function. By assuming that $M_t(z)$ is continuous in $t = 0$, one can easily show that

$$M_t(z) = (M_1(z))^t$$

Assuming also for the sake of simplicity that the random variable $X(t)$ has a density

$$f_t(x) = \frac{d}{dx} F_t(x), \quad t > 0,$$

it holds that

$$M_t(z) = \int_{-\infty}^{+\infty} e^{zx} f_t(x) \, dx.$$

Let now $h \in \mathbb{R}$ for which $M_t(h)$ is defined. The Esscher transform (parameter $\theta$) of the process $\{X(t)\}$ is again a process with stationary and independent increments, whereby the new p.d.f. of $X(t),\ t > 0$ is

$$f_t(x; \theta) = \frac{e^{\theta x} f_t(x)}{\int_{-\infty}^{+\infty} e^{\theta y} f_t(y) \, dy} = \frac{e^{\theta x} f_t(x)}{M_t(\theta)},$$

that is, the modified distribution of $X(t)$ is the Esscher transform of the original distribution. The corresponding moment generating function is

$$M_t(z;\theta) = \int_{-\infty}^{+\infty} e^{zx} f_t(x;\theta) = \frac{M_t(z+\theta)}{M_t(\theta)}.$$

It also still holds that

$$M_t(z;\theta) = (M_1(z;\theta))^t .$$

This concept of Esscher transform of a stochastic process appears consistently for this environment first in a work by Gerber and Shiu, [GS94]; so the measure of the process has been modified. Because the exponential function is positive, the modified probability measure is equivalent to the original probability measure, that is both measures have the same null sets. Although stock returns have been widely studied, no single distribution has emerged as a clear winner from these studies, despite the common agreement that the returns' distributions should have fatter tails than the traditional normal distribution (see also Cont's [Con01] qualitative consoderations). The cited paper [GS94] in addition to lognormal process also discusses option valuation using both the Gamma and the inverse Gaussian process. Altough the latter process allows for fatter tails, both the distributions have tails that decay exponentially. Given the fact that this rate of convergence is necessary for the existence of an Esscher transform, this should not come as any surprise. Such tail behaviour constraints, however, can be avoided by considering shifted processes and distributions that are supported for instance on $\mathbb{R}^+$.
It still remains to be answered whether the equation that defines $\theta$ (see again [GS94]) can always be solved for a general distibution.
When discussing about heavy-tailed distribution, is not unreasonable to think about stable Pareto distributions. These distributions are prominent members of the class of infinitely divisible distributions that, subsequent to the original works of Mandelbrot and Fama, have often been used to explain the stochastic behaviour of stock prices. The interest in stable distributions is largely due to the facts that only stable laws have domains of attraction (generalized central limit theorem) and that stable distributions belong to their own domain of attraction (stability).

From a practical viewpoint, stable laws are flexible, empirical models that are capable of explaining the observed leptokurtosis and skewness in return distributions. Moreover they are able to capture the essentials of probability structures when sample moments exhibit a nonstationary behaviour over time.

a stable Pareto distribution can fundemantally be described by the shape (denoted by $\alpha,\ 0 < \alpha < 2$), skewness (denoted by $\beta,\ |\beta| \leq 1$), location and scale parameters. Amongst these the mosto important is the shape parameter, which when decreased increases the tail probabilities. Two obvoius drawbacks of these distributions are the lack of second moments (also the first if $\alpha < 1$) and the absence of explicit expressions for the density functions. These disadvantages however, are not major obstacles when one considers asset pricing using the notion of risk-neutral valuation. This is due to the fact that all that is needed is the knowledge of the measure under which the discounted process is a martingale.

Suppose for instance that $X$ is $\alpha-$stable, with $0 < \alpha < 2$; then the random variable $e^X$ has no finite moments except when $X$ is totally skewed to the left (that is, $\beta = -1$). It is important to note that in this instance when $\alpha > 1$, the support of this distribution is the interval $(-\infty, +\infty)$. Thus, for $\beta = -1$ and $\alpha > 1$, all moments of $e^X$ are finite, and setting the location parameter to $0$, results in zero expectation. Hence one can consider the modeling of the stock price movement using the process $S(t) = S_0 e^{X(t)}$, with $t \geq 0$.

The value of $\beta$ that was used in deriving the above process also forces the right tails of the distribution of $X(t)$ to decay rapidly, and as a consequence, the moment generating function $E\left[e^{\gamma X}\right],\ \gamma \geq 0$, exists for all $0 < \alpha \leq 2$ and was shown to be equal to

$$\begin{array}{ll} \exp\left(-\dfrac{\sigma^\alpha\gamma^\alpha}{A}\right), & \text{if } \alpha \neq 1,\ A = \cos(\alpha\pi/2) \\ \exp\left(-\dfrac{2\sigma\gamma\log\gamma}{\pi}\right), & \text{if } \alpha = 1 \end{array}$$

Hence it is possible to consider the approach for a shifted $\alpha-$stable process $X(t) = Y(t) + \mu t$, where $Y(t)$ is a process with independent increments and an $\alpha-$stable distribution with $\beta = -1$, shift parameter equal to 0 and scale parameter equal to $(0.5\sigma^2 t)^{1/\alpha}$. So it holds that

$$M_t(z) = E\left[e^{zX(t)}\right] = \exp\left[\left(\mu z - \frac{\sigma^2 z^\alpha}{2A}\right)t\right],\ z \geq 0,$$

and $A$ as above. In particular, when $\alpha = 2$ one gets the classic lognormal distributed stock price process. However this transformed process does not have the nice properties of a Wiener process because

$$M_t(z;h) = \exp\left\{\mu z t - \frac{\sigma^2 t\left[(z+h)^\alpha - h^\alpha\right]}{2A}\right\},\ h \geq 0,\ z+h \geq 0$$

implies that when $h \neq 0$, the Esscher transform of a shifted $\alpha-$stable process is no longer an $\alpha-$stable process. Despite this drawback, one still has a process with stationary and independent increments whose expected values exist for all $h \geq 0$.

## 4.5 Esscher transform for Lévy processes

Let us state the theorems which define the Esscher transform for a Lévy process :

**Theorem 31.** *Suppose $T > 0$ and $\theta \in \mathbb{R}$, such that*

$$E\left[e^{\theta X_T}\right] < \infty$$

*Then*

$$\frac{dP^\theta}{dP} = e^{\theta X_T - \kappa(\theta) T}$$

*defines a probability measure $P^\theta$ such that $P^\theta \sim P$ and $\{X_t\}_{0 \leq t \leq T}$ is a Lévy process under $P^\theta$ with triplet $(b^\theta, c^\theta, U^\theta)$ given by*

$$\begin{aligned} b^\theta &= b + \theta c + \int (e^{\theta x} - 1) h(x)\, U(dx), \\ c^\theta &= c, \\ U^\theta(dx) &= e^{\theta x} U(dx). \end{aligned}$$

By denoting the expectation with respect to $P^\theta$ with $E^\theta$, we have $E^\theta[e^{zX_t}] = e^{\kappa^\theta(z)t}$ for $0 \leq t \leq T$, where

$$\kappa^\theta(z) = \kappa(z+\theta) - \kappa(\theta).$$

If the measure $P^\theta$ exists, it is called the Esscher transform of $P$, or Esscher measure.

**Theorem 32.** *Suppose now $T > 0$ and there exists $\theta^\sharp \in \mathbb{R}$ such that*

$$E\left[e^{\theta^\sharp X_T}\right] < \infty \qquad E\left[e^{(\theta^\sharp+1)X_T}\right] < \infty$$

*and the equation*

$$\kappa(\theta^\sharp + 1) - \kappa(\theta^\sharp) = 0$$

*holds, then*

$$\frac{dP^\sharp}{dP} = e^{\theta^\sharp X_T - \kappa(\theta^\sharp)T},$$

*defines an equivalent martingale measure for $\{S_t = S_0 e^{X_t}\}_{0\leq t\leq T}$ (exponential Lévy process). The process $\{X_t\}_{0\leq t\leq T}$ is a Lévy process under $P^\sharp$ with Lévy triplet $(b^\sharp, c^\sharp, U^\sharp)$, where*

$$\begin{aligned} b^\sharp &= b + \theta^\sharp c + \int (e^{\theta^\sharp x} - 1)h(x)\, U(dx), \\ c^\sharp &= c, \\ U^\sharp(dx) &= e^{\theta^\sharp x} U(dx). \end{aligned}$$

By denoting the expectation with respect to $P^\sharp$ with $E^\sharp$, we have $E^\sharp[e^{zX_t}] = e^{\kappa^\sharp(z)t}$ for $0 \leq t \leq T$, where

$$\kappa^\sharp(z) = \kappa(z+\theta^\sharp) - \kappa(\theta^\sharp).$$

The measure $P^\sharp$ is called Esscher martingale transform for the exponential Lévy process $e^X$.

Let us add some more definitions to section 1.8:

**Definition 30.** *A real-valued semimartingale is called* special *if it can be written as*

$$X = X_0 + M + V$$

*for some local martingale $M$ and some predictable process $V$ of finite variation, both starting at* $0$.

Alternatively $X$ is a special semimartingale if there exists a predictable process $V$ such that $X - X_0 - V$ is a local martingale.

**Definition 31.** *Process V (unique) is called the* compensator *or* drift process *of $X$ and is written* $D^X := V$.

**Definition 32.** *A real-valued semimartingale $X$ is called* exponentially special *if* $\exp(X - X_0)$ *is a special semimartingale.*

The two theorems above can be seen as a particular case of the following fundamental theorem (see Kallsen and Shiryaev, [KS02], theorem $4.1$, page $421$): let the symbol $\theta^T \cdot X$ denote the stochastic integral of $\theta$ relative to $X$, for some given Lévy process $X$ and real vector $\theta$.

**Theorem 33.** *Let $\theta \in L(X)$ be such that $\theta^T \cdot X$ is exponentially special and such that $Z^\theta$ is a uniformly integrable martingale. Define $P^\theta \sim P$ by its Radon-Nikodym density*

$$\frac{dP^\theta}{dP} := Z^\theta,$$

*and set*

$$\theta^{(i)} := (\theta^1, \ldots, \theta^{i-1}, \theta^i + 1, \theta^{i+1}, \ldots, \theta^d)^T.$$

*Then the processes $S^i = S_0^i e^{X^i}$ are $P^\theta$-local martingales if and only if $\left(\theta^{(i)}\right)^T \cdot X$ is exponentially special and*

$$K^X(\theta^{(i)}) - K^X(\theta) = 0, \textit{ for } i = 1, \ldots, d.$$

*In this case we call $P^\theta$ an Esscher martingale transform for exponential processes.*

Moreover

**Theorem 34.** *If $d = 1$ , then the Esscher martingale transform for exponential processes is unique (provided that it exists).*

For the two following proofs, which are taken from [KS02], two auxiliary preliminar results are needed.

**Theorem 35.** *A real-valued semimartingale $X$ has an exponential compensator if and only if it is exponentially special. In this case, the exponential compensator is up to indistinguishability unique.*

**Theorem 36.** *Let $\theta \in L(X)$ such that $\theta^T \cdot X$ is exponentially special. Then $K^X(\theta)$ is the exponential compensator of $\theta^T \cdot X$. More specically,*

$$\begin{aligned} Z := \exp(\theta^T \cdot X - K^X(\theta)) = \frac{\exp(\theta^T \cdot X)}{\mathcal{E}(\tilde{K}^X(\theta)} = \\ = \mathcal{E}\left(\theta^T \cdot X^c + \frac{e^{\theta^T x} - 1}{1 + \hat{W}(\theta)} * (\mu^X - \nu)\right) \in \mathcal{M}_{loc}, \end{aligned}$$

*where $\hat{W}(\theta)_t := \int (e^{\theta^T x} - 1)\nu(\{t\} \times dx)$ and $\mathcal{M}_{loc}$ is the space of all local martingales.*

*Proof of theorem* (33): it is a known result (see for instance [JS02], 3.8 pg.168) that $\exp(X^i)$ is a $P_\theta-$local martingale if and only if $\exp(X^i)Z^\theta = \exp[(\theta^{(i)})^T \cdot X - K^X(\theta)]$ is a $P_\theta-$local martingale. By theorems 35 and 36 this is the case if and only if $(\theta^{(i)})^T \cdot X$ is exponentially special and $K^X(\theta^{(i)}) = K^X(\theta)$ up to indistinguishability. □

*Proof of theorem* (34):

*Step 1*: Let $\theta, \bar{\theta} \in L(X)$ be such that

$$\theta \cdot X, \qquad (\theta + 1) \cdot X, \qquad \bar{\theta} \cdot X, \qquad (\bar{\theta} + 1) \cdot X$$

are exponentially special and such that $P_\theta$ and $P_{\bar{\theta}}$ are Esscher martingale transforms for exponential processes. Then

$$K^X(\theta + 1) - K^X(\theta) = 0 = K^X(\bar{\theta} + 1) - K^X(\bar{\theta}).$$

In particular, $\tilde{\kappa}(\theta + 1) - \tilde{\kappa}(\theta) = 0 = \tilde{\kappa}(\bar{\theta} + 1) - \tilde{\kappa}(\bar{\theta})(P \otimes A)-$ almost everywhere on the set $\{\Delta A = 0\}$.

On the set $\{\Delta A \neq 0\}$ we have

$$\Delta K^X(\theta+1) - \Delta K^X(\theta) = 0 = \Delta K^X(\bar{\theta}+1) - \Delta K^X(\bar{\theta}).$$

This implies

$$0 = b + \frac{c}{2} + c\theta + \int [(e^x - 1)e^{\theta x} - h(x)]F(dx), \qquad \text{on the set } \{\Delta A = 0\}$$
$$0 = \log \frac{1 + \int (e^{(\theta+1)x} - 1)\nu(\{t\} \times dx)}{1 + \int (e^{\theta x} - 1)\nu(\{t\} \times dx)}, \qquad \text{on the set } \{\Delta A \neq 0\}$$

Parallel statements hold for $\bar{\theta}$.

*Step 2*: Fix $(\omega, t) \in \Omega \times [0, T]$ and let $\theta \leq \bar{\theta}$ in $(\omega, t)$ without loss of generality. Firstly, suppose that $\{\Delta A_t(\omega) = 0\}$. Since $\int |(e^x - 1)e^{\theta x} - h(x)|F_t(dx) < \infty$ and likewise for $\bar{\theta}$, we have that

$$\sup_{\psi \in [\theta, \bar{\theta}]} \int |(e^x - 1)e^{\psi x} - h(x)|F_t(dx) < \infty$$

Define $v : [0, 1] \to \mathbb{R}$ by

$$v(\lambda) := b_t + (1/2)c_t + c_t(\theta + \lambda(\bar{\theta} - \theta)) + \int \left[(e^x - 1)e^{(\theta + \lambda(\bar{\theta} - \theta))x - h(x)}\right] F_t(dx).$$

Note that $v$ is a well-defined, continuous, increasing mapping. It can be then concluded that $(\bar{\theta} - \theta)c_t = 0$, $(\bar{\theta} - \theta)b_t - \int (\bar{\theta} - \theta)h(x)F_t(dx) = 0$, and $(\bar{\theta} - \theta)x = 0$ for $F_t-$almost all $x \in \mathbb{R}$.

Secondly, assume that $\{\Delta A_t(\omega) \neq 0\}$. Since $\int e^{(\theta+1)x}\nu(\{t\} \times dx) < \infty$ and $\int e^{\theta x}\nu(\{t\} \times dx) < \infty$ and likewise for $\bar{\theta}$, the same integrability conditions hold uniformly on $[\theta, \bar{\theta}]$.

This time, define $v : [0, 1] \to \mathbb{R}$ by

$$v(\lambda) := \log \frac{1 + \int (e^{(\theta + \lambda(\bar{\theta} - \theta) + 1)x} - 1)\nu(\{t\} \times dx)}{1 + \int (e^{(\theta + \lambda(\bar{\theta} - \theta))x} - 1)\nu(\{t\} \times dx)}.$$

Observe that $v$ is differentiable on $(0, 1)$ with derivative

$$v'(\lambda) = \frac{(\bar{\theta} - \theta) \int x e^{(\theta + \lambda(\bar{\theta} - \theta))x} e^x \nu(\{t\} \times dx)}{1 + \int (e^{(\theta + \lambda(\bar{\theta} - \theta))x} e^x - 1)\nu(\{t\} \times dx)} + $$
$$- \frac{(\bar{\theta} - \theta) \int x e^{(\theta + \lambda(\bar{\theta} - \theta))x} \nu(\{t\} \times dx)}{1 + \int (e^{(\theta + \lambda(\bar{\theta} - \theta))x} - 1)\nu(\{t\} \times dx)}$$

Fix $\lambda \in (0,1)$ for the moment. Define a family $(Q_\varrho)_{\varrho\in[0,1]}$ of probability measures on $\mathbb{R}$ by

$$Q_\varrho(M) := \frac{\int_M e^{\varrho x} e^{(\theta+\lambda(\bar\theta-\theta))x} P^{\Delta X_t|F_{t-}}(dx)}{\int e^{\varrho x} e^{(\theta+\lambda(\bar\theta-\theta))x} P^{\Delta X_t|F_{t-}}(dx)}$$

for $M$ a Borel set. With this notion, we have $v'(\lambda) = (\bar\theta-\theta)[E_{Q_1}(I) - E_{Q_0}(I)]$, where $I : \mathbb{R} \to \mathbb{R},\ x \mapsto x$ denotes the identity mapping. Since $(Q_\varrho)_{\varrho\in[0,1]}$ is a class with increasing likelihood ratio, it follows that $v'(\lambda) \geq 0$. Therefore $v$ is an increasing mapping on $[0,1]$.

It can now be concluded that $v(\lambda) = 0$ and $v'(\lambda) = 0$ for any $\lambda \in (0,1)$. This implies

$$0 = v'(\lambda) = (\bar\theta-\theta)\int x e^{(\theta+\lambda(\bar\theta-\theta))x}(e^x-1)\nu(\{t\}\times dx)$$

for any $\lambda \in (0,1)$, which in turn means that $(\bar\theta-\theta)x = 0$ for $\nu(\{t\}\times\cdot)-$almost all $x \in \mathbb{R}$.

*Step 3*: It follows that $\theta\cdot X = \bar\theta\cdot X$, which proves the claim. □

## 4.6 Esscher Transform for Meixner Process

We are now going to prove a simple corollary of a general theorem given by Grigelionis in [Gri99] which settles the way for the theoretical formulation of the analogue of Black and Scholes formula. Here is the general theorem with our corollary following:

**Theorem 37.** *The unique value $\theta^\star \in \mathbb{R}$ such that the discounted geometric Meixner process*

$$S_0 \exp\{X_t - rt\}$$

*with $t \geq 0, S_0 > 0, r \in \mathbb{R}$, is a martingale is given by*

$$\theta^\star = \frac{2}{\alpha}\arccos\frac{|\sin(\alpha/2)|}{\sqrt{1+\zeta^2-2\zeta\cos(\alpha/2)}} - \frac{\beta}{\alpha},$$

*where $\zeta = \exp{(\mu-r)/2\delta}$.*

*$\theta^\star$ solves the equation*

$$\cos\frac{\alpha(\theta^\star+1)+\beta}{2} = \zeta\cos\frac{\alpha\theta^\star+\beta}{2}.$$

**Corollary 1.** *The unique value $\theta \in \mathbb{R}$ such that the geometric Meixner process*

$$S_t = S_0 \exp\{X_t\}$$

*with $t \geq 0$, is a martingale is given by*

$$\theta = -\frac{1}{2} - \frac{\beta}{\alpha} + \frac{2k\pi}{\alpha}, \quad k \in \mathbb{Z}.$$

*Proof*: we first observe that Esscher transform for the Meixner process is structure-preserving, meaning that the Esscher transform applied on a $MD(\alpha,\beta,\delta,\mu)$ still produces a Meixner distribution $MD(\alpha,\alpha\theta+\beta,\delta,\mu)$.

This can be seen by noting that the exponential structure of the Esscher transform, just influences the exponential part of the density, leaving the rest unaltered (see Hubalek and Sgarra [HS06], for instance).

The corresponding Lévy measure of the transformed process is

$$\nu_\theta(x) = \delta\frac{e^{\frac{\alpha\theta+\beta x}{\alpha}}}{x\sinh\left(\frac{\pi x}{\alpha}\right)}\,dx.$$

The Esscher transform $P^\theta$ for the exponential Meixner process exists when the integral

$$\int_{\mathbb{R}} e^{\theta x}\delta\frac{e^{\frac{\beta x}{\alpha}}}{x\sinh\left(\frac{\pi x}{\alpha}\right)}\,dx$$

is finite, which happens for

$$\theta \leq \frac{\pi-\beta}{\alpha} \quad \text{in a neighborhood of } +\infty$$
$$\theta \geq \frac{-\pi-\beta}{\alpha} \quad \text{in a neighborhood of } -\infty,$$

while it turns out the integral is not finite in $x = 0$ for every $\theta$.

Moreover, the function $e^x$ is $\mathbb{R}$-integrable with respect to the measure $u_\theta(x)$ when the following conditions hold:

$$\theta \leq \frac{\pi - \alpha - \beta}{\alpha} \quad \text{in a neighborhood of} + \infty$$
$$\theta \geq \frac{-\pi - \alpha - \beta}{\alpha} \quad \text{in a neighborhood of} - \infty,$$

and, as usual, $x \neq 0$, for every $\theta$.

The two conditions combined mean that the Esscher transform $P^\theta$ for the exponential Meixner process always exists in $\mathbb{R} \setminus \{0\}$ for

$$-\frac{\pi + \beta}{\alpha} \leq \theta \leq \frac{\pi - \beta}{\alpha};$$

the Esscher-transformed Meixner process is still a Meixner process under $P^\theta$, and $e^X$ is integrable in $\mathbb{R} \setminus \{0\}$ under $P^\theta$ for

$$-\frac{\pi + \alpha + \beta}{\alpha} \leq \theta \leq \frac{\pi - \alpha - \beta}{\alpha}.$$

Now consider the cumulant difference $\kappa(\theta + 1) - \kappa(\theta)$; it holds that, for Meixner case,

$$g(\theta) = \kappa(\theta + 1) - \kappa(\theta) = 2\delta \log \left[ \frac{\cos\left(\frac{\alpha\theta + \beta}{2}\right)}{\cos\left(\frac{\alpha(\theta+1)+\beta}{2}\right)} \right]$$

Function $g(\theta)$ is $4\pi/\alpha$-periodic, and in a single period is defined for

$$\theta < -\frac{\pi + \alpha + \beta}{\alpha}, \qquad \theta > \frac{\pi - \beta}{\alpha}$$

and

$$\frac{\pi - (\alpha + \beta)}{\alpha} < \theta < -\frac{\pi + \beta}{\alpha} \qquad \text{if } \alpha > 2\pi,$$
$$-\frac{\pi + \beta}{\alpha} < \theta < \frac{\pi - (\alpha + \beta)}{\alpha} \qquad \text{else.}$$

Our interest is focused now over the $\theta \in \mathbb{R}$ such that $\kappa(\theta + 1) - \kappa(\theta) = 0$: this happens for every $\delta > 0$, when

$$\theta = -\frac{1}{2} - \frac{\beta}{\alpha} + \frac{2k\pi}{\alpha}, \quad k \in \mathbb{Z}. \qquad \square$$

By theorem 32 is now possible to evaluate the form of $U^\sharp(dx)$, the Lévy measure of transformed process, as

$$U^\sharp(dx) = \frac{\delta e^{(\beta-1/2)x}}{x\sinh(\pi x/\alpha)}\,dx$$

## 4.7 General overview on Minimal Entropy Martingale measure

The first time in literature that the topic of hedging contingent claims in incomplete market conditions in terms of minimizing risk and corresponding minimal martingale measures is issued, is in a paper by Föllmer and Schweizer, [FS91]. Such a claim in these conditions will by its own nature have an intrinsic risk. The problem is to characterize and construct the strategies which minimize the risk. In a general framework if we cosider a contingent claim at time $T$ given by a random variable

$$H \in \mathcal{L}^2(\Omega, \mathcal{F}_T, P)$$

and we let $X$ be a continuous path stochastic process on $(\Omega, \mathcal{F}, P)$ which is supposed to describe the price fluctuaton of a given stock on which $H$ relies, the problem of finding an optimal strategy is reduced to the chance of decomposing $H$ in a way similar to what the *Kunita-Watanabe decomposition* does (cf.[FS91], proposition (2.24) ), namely

$$H = H_0 + \int_0^T \xi_s^H\,dX_s + L_T^H,$$

where $H_0 \in \mathcal{L}^2(\Omega, \mathcal{F}_0, P), L_H = (L_t^H)_{0\le t\le T}$ is an $\mathcal{L}^2$-martingale orthogonal to $M$, and $\xi^H$ comes from the supposedly true Itô representation of $H$ as

$$H = H_0 + \int_0^T \xi_s^H\,dX_s, \quad P-a.s.$$

and enjoys some integrability properties ( (2.8) in [FS91]). $M = (M_t)_{0\le t\le T}$ is the local martingale component of the Doob-Meyer decomposition of the semimartingale $X$.

In a complete market situation, the optimal strategy can be computed in terms of the unique equivalent martingale measure $P^*$, in incomplete markets this is not the case: $P^*$ is no longer unique, and the choice of different martingale measures may lead to different strategies. It is shown anyway that exists a *minimal* martingale measure $\hat{P} \sim P$ such that the optimal strategy can be clculated in terms of $\hat{P}$.

The departure from the given measure $P$ can be expressed in terms of the *relative entropy*

$$H(Q|P) = \begin{cases} \int \log \frac{dQ}{dP}\, dQ & \text{if} Q \ll P, \\ +\infty & \text{else.} \end{cases}$$

In particular it can be shown that $\hat{P}$ minimizes the relative entropy $H(\cdot|P)$ among all martingale measures $P^*$ with fixed expectation.

In the work by Esche and Schweizer [ES05], the first step of the equivalence between the Esscher martingale measure for the linear Lévy process $X$ and the minimal entropy martingale measure for the exponential Lévy process $e^X$ is stated. The objective is completed in two steps. Keeping it simple for the sake of understanding, the first one is the following

**Theorem 38.** *Let $X$ be a real-valued Lévy process; if the minimal entropy martingale measure $\hat{P}$ exists for the exponential Lévy process $e^X$, then $X$ is a Lévy process under $\hat{P}$.*

The other one:

**Theorem 39.** *If the Esscher martingale measure for the linear Lévy process $\tilde{X}$ exists, then it is the minimum entropy martingale measure for the exponential Lévy process $e^X$.*

This is obviously only one half of the equivalence: the second one is found in [HS06], giving the complete characterization of the minimum entropy martingale measure for the exponential Lévy process $e^X$ as the Esscher transform for the linear Lévy process $\tilde{X}$. Namely

**Theorem 40.** *The minimal entropy martingale measure for the exponential Lévy process $e^X$ exists if and only if the Esscher martingale measure for the linear Lévy process $\tilde{X}$ exists. If both measures exist, they coincide.*

This assumption brings to a complete characterization of the minimal entropy martingale measure in terms of the Esscher transform, but unfortunately it is not comfortable to use when making practical examples.

### 4.7.1 Minimal entropy martingale measure for geometric Lévy processes

Given the process

$$\tilde{S}(t) = e^{-rt}S(t) \tag{4.3}$$

where $r \in \mathbb{R}$, and $S(t)$ is a geometric Lévy process defined as in 4.2, in [FM03] a condition is given in terms of the Lévy measure $\nu(dx)$ under which there exists a probability measure in the set of all equivalent martingale measures $EMM(P)$ where

$$\begin{aligned} EMM(P) =& \{Q \in \mathcal{P}(\Omega, \mathscr{F}) : Q \sim P \ on \ \mathscr{F} \ and \\ & \tilde{S} = (\tilde{S}(t), \mathscr{F}_t)_{t\in[0,T]} \ is \ a \ martingale \ under \ Q\} \end{aligned}$$

which is the minimum in this set.

Let now $X(t)$ be the driving Lévy process and the following fundamental condition:

There exists $\beta^* \in \mathbb{R}$ constant that satisfying both

1.
$$\int_{\{x>1\}} e^x e^{\beta^*(e^x-1)} \nu(dx) < \infty,$$

2.
$$\begin{aligned} b + \left(\frac{1}{2} + \beta^*\right)\sigma^2 + \int_{\{|x|\le 1\}} \left\{(e^x-1)e^{\beta^*(e^x-1)} - x\right\} \nu(dx) + \\ + \int_{\{|x|>1\}} (e^x-1)e^{\beta^*(e^x-1)} \nu(dx) = r. \end{aligned}$$

The main result then is given by

**Theorem 41.** *Supposing the above condition holds, then*

1. *It is possible to define a proability measure $P^*$ on $\mathscr{F}_T$ by means of the Esscher transform*

$$\left.\frac{dP^*}{dP}\right|_{\mathscr{F}_t} = \frac{e^{\beta^* \tilde{R}_t}}{E^P[e^{\beta^* \tilde{R}_t}]}$$

*for every $t \in [0,T]$, where $\{\tilde{R}_t\}_{t\in[0,T]}$ is the return process for $\{\tilde{S}_t\}_{t\in[0,T]}$ defined by*

$$\tilde{R}_t = \int_{(0,t]} \frac{d\tilde{S}_u}{\tilde{S}_u}.$$

*More concretely*

$$\left.\frac{dP^*}{dP}\right|_{\mathscr{F}_t} = \frac{e^{\beta^* \hat{X}_t}}{E^P[e^{\beta^* \hat{X}_t}]} = e^{\beta^* \hat{X}_t - b_* t}$$

*where $\hat{X}_t$ is the process defined by*

$$\hat{X}_t = X_t + \frac{1}{2}\sigma^2 t + \int_{(0,t]} \int_{\mathbb{R}\setminus\{0\}} (e^x - 1 - x) N_p(dudx),$$

*with $N_p(dudx)$ the counting measure of the point process $p_t = \Delta X_t$:*

$$N_p((0,t], A) = \sharp\{u \in D_p \cap (0,t] : p_u \in A\}$$

*for $A \in \mathcal{B}(\mathbb{R} \setminus \{0\})$, and $D_p = \{t > 0 : \Delta X_t \neq 0\}$, and*

$$b_* = \frac{\beta^*(1+\beta^*)}{2}\sigma^2 + \beta^* b + \int_{\mathbb{R}\setminus\{0\}} [e^{\beta^*(e^x - 1)} - 1 - \beta^* x \mathbb{1}_{\{|x|\leq 1\}}]\nu(dx).$$

2. *The stochastic process $X_t$ is still a Lévy process under the above probability measure $P^*$ and the Lévy triplet associated with the truncation function $h(x) = x\mathbb{1}_{\{|x|\leq 1\}}$ is given by*

$$\left[\int_{\{|x|\leq 1\}} \left\{x e^{\beta^*(e^x-1)} - 1\right\} \nu(dx), \sigma^2, \nu^*\right]$$

*where*

$$\nu^*(dx) = e^{\beta^*(e^x - 1)}\nu(dx)$$

*Furthermore the probability measure $P^*$ is in $EMM$ and attains the minimal entropy.*

*Proof*: see [FM03], pg. $518$.

In the same paper is also given an auxiliary condition for verifying the satisfaction of the fundamental condition above (see Proposition 3.3).

A similar condition is clearly not easily approachable in terms of calculations for a complex process like Meixner process. This surely is going to be part of the follow up of the research in this direction.

## 4.8 Comparisons and conclusions

A notable comparison is performed by Miyahara in [Miy04], where it is observed that the Esscher Martingale Measure (ESSMM) and the minimal entropy martingale measure (MEMM) are both obtained by Esscher transform, but they have different properties:

1. For the existence of ESSMM, the condition

$$\int_{\{|x|>1\}} |(e^x - 1)e^{h^* x}|\nu(dx) < \infty$$

   is necessary, while on the other hand, for the existence of MEMM, the corresponding condition is

$$\int_{\{|x|>1\}} |(e^x - 1)e^{\theta^*(e^x - 1)}|\nu(dx) < \infty.$$

   This condition is satised for wide class of Lévy measures, if $\theta^* < 0$. Namely, the former condition is strictly stronger than the latter condition. This means that the MEMM may be applied to the wider class of models than the ESSMM. The difference works in the stable process cases. In fact we can make sure that MEMM method can be applied to geometric stable models but ESSMM method can not be applied to this model.

2. The ESSMM is corresponding to power utility function or logarithm utility function. (See for instance always Miyahara, [Miy05], where an extensive list of possible candidates for martingale measures with their respective dual distance functions is collected). On the other hand the MEMM is corresponding to the exponential utility function. We can observe that, in the case of ESSMM, the power parameter of the utility function depends on the parameter value $h^*$ of the Esscher transform.

3. The relative entropy is very popular in the eld of information theory, and it is called Kullback-Leibler Information number or Kullback- Leibler distance. Therefore we can state that the MEMM is the nearest equivalent martingale measure to the original probability $P$ in the sense of Kullback-Leibler distance. Recently the idea of minimal distance martingale measure is studied. Göll and Ruschendorf in [GR01] mention that the relative entropy is the typical example of the distance in their theory.

4. Large deviation theory is closely related to the minimum relative entropy analysis, and well-known Sanov's theorem (or Sanov's property) provides a connection between the two fields; in fact it basically says that the MEMM is the most possible empirical probability measure of paths of price process in the class of the equivalent martingale measures. In this sense the MEMM should be considered to be an exceptional measure in the class of all equivalent martingale measures.

# Chapter 5

# Elements of theory of orthogonal polynomials

In this chapter some elements of theory of orthogonal polynomials will be exposed, as they provide an alternative but insightful approach to Lévy processes . Also, it may be understood where most of the properties of Lévy processes come from. In particular, the example of Meixner process will be discussed, together with an additional detail: the Fisher information for the corresponding family of orthogonal polynomials.

This part is important as it provides a deeper insight, different from the usual "classical" definitions, of the process of generation of the most known Lévy processes , also with interesting contacts with theory of differential equations, statistical distributions and martingales. The main reference for this part is surely the book by Schoutens, [Sch00], and the anticipating paper by Schoutens and Teugels [ST98].

## 5.1 Introduction to classical orthogonal polynomials

**Definition 33.** *A differential equation of the form*

$$s(x)y'' + \tau(x)y' + \lambda y = 0 \tag{5.1}$$

*where* $s(x), \tau(x)$ *are polynomials of at most second and first degree respectively and* $\lambda$ *is a constant, is called a differential equation of* hypergeometric type*, and its solutions functions of hypergeometric type.*

It emerges usually in many problems of applied mathematics and theoretical and mathematical physics.

If in addition

$$\lambda = \lambda_n = -n\tau' - \frac{1}{2}n(n-1)s'', \tag{5.2}$$

the equation above has a particular solution of the form $y(x) = y_n(x)$, which is a polynomial of degree $n$, called a *polynomial of hypergeometric type*. The polynomials $y_n(x)$ are the simplest solutions for (5.1).

It can be shown that these solutions of (5.1) have the orthogonal property

$$\int_a^b y_m(x)y_n(x)\rho(x)dx = d_n^2\delta_{nm},$$

with $\delta_{nm}$ the Kronecker delta, for some constants $a, b$, possibly infinite, $d_n \neq 0$, and where the weight function of orthogonality $\rho(x)$ satisfies the differential equation

$$[s(x)\rho(x)]' = \tau(x)\rho(x). \tag{5.3}$$

These polynomials of hypergeometric type $y_n(x)$ are known as the (very) *classical orthogonal polynomials of a continuous variable*.

The analogous holds for polynomials with discrete variable.

**Definition 34.** *A difference* equation of hypergeometric type *is one of the form*

$$s(x)\Delta\nabla y(x) + \tau\Delta y(x) + \lambda y(x) = 0,$$

*where* $s(x), \tau(x)$ *are polynomial of at most second and first degree respectively,* $\lambda$ *is a constant, and*

$$\Delta f(x) = f(x+1) - f(x), \quad \textit{and} \quad \nabla f(x) = f(x) - f(x-1).$$

If (5.2) holds, the difference equation above has a particular solution of the form $y(x) = y_n(x)$, which is a polynomial of degree $n$, provided moreover

$$\mu_m = \lambda + m\tau' + \frac{1}{2}m(m-1)s'' \neq 0, \quad \text{for} \quad m = 0, 1, \ldots, n-1.$$

It can be shown that the polynomials solutions of the difference equation have the orthogonal property

$$\sum_{x=a}^{b} y_m(x)y_n(x)\rho(x)dx = d_n^2\delta_{nm},$$

for some constants $a, b$, possibly infinite, $d_n \neq 0$, and where the discrete orthogonality measure $\rho(x)$ satisfies the difference equation

$$\Delta[s(x)\rho(x)] = \tau(x)\rho(x). \tag{5.4}$$

These polynomials of hypergeometric type $y_n(x)$ are known as the *classical orthogonal polynomials of a discrete variable*

### 5.1.1 Classical orthogonal polynomials of a continuous variable

There are in essence five solutions of (5.3), depending on whether the polynomial $s(x)$ is constant, linear, or quadratic, and in this last case, on whether the discriminant $D = b^2 - 4ac$ of $s(x) = ax^2 + bx + c$ is positive, negative, or zero.

These are the possible cases:

| Name | Symbol | deg $s(x)$ | $D$ |
|---|---|---|---|
| Jacobi | $P_n^{(\alpha,\beta)}(x)$ | 2 | $> 0$ |
| Bessel | | 2 | $= 0$ |
| Romanowski | | 2 | $< 0$ |
| Laguerre | $L_n^{(\alpha)}(x)$ | 1 | |
| Hermite | $H_n(x)$ | 0 | |

### Further details

**Jacobi polynomials**: if $s(x) = 1 - x^2$, and $\tau(x) = -(\alpha + \beta + 2)x + \beta - \alpha$, then

$$\rho(x) = (1 - x)^\alpha (1 + x)^\beta \qquad \text{Beta kernel}$$

Moreover $\lambda_n = n(n + \alpha + \beta + 1)$, and the orthogonality relation is satisfied for

$$a = -1, \quad b = 1, \quad \alpha, \beta > -1$$

**Bessel polynomials**: with an appropriate affine change of variable, $\rho(x)$ can be written in the form

$$\rho(x) = Cx^{-\alpha} e^{-\beta/x}, \quad C \text{ normalizing constant.}$$

Now if $\rho(x)$ is defined on $(0, +\infty)$, $\alpha > 1$ and $\beta \geq 0$ ensure that $\rho(x)$ is integrable. In this case $s(x) = x^2$ and $\tau(x) = (2 - \alpha)x + \beta$. Observe that we have a finite system of orthogonal polynomials in this case, because for this distribution only the moments of orders strictly less than $\alpha - 1$ exist.

**Romanowski polynomials**: $\rho(x)$ can be written as

$$\rho(x) = C(1 + x^2)^{-\alpha} e^{\beta \arctan x}, \quad C \text{ normalizing constant.}$$

If we assume that $\rho(x)$ is defined on $\mathbb{R}$, then $\alpha > 1/2,\ \beta \in \mathbb{R}$.

A particular case is given by

$$\alpha = \frac{n+1}{2}, \quad \beta = 0, \quad s(x) = 1 + \frac{x^2}{n}, \quad \tau(x) = -(n-1)\frac{x}{n},\ n \in \{1, 2, \ldots\}$$

then

$$\rho(x) = C\left(1 + \frac{x}{n}\right)^{-\frac{n+1}{2}}, \quad \text{with } C = \Gamma\left(\frac{n+1}{2}\right) \frac{1}{\sqrt{n\pi}\,\Gamma(n/2)}$$

from which the Student's $t$ distribution is clearly recognizable.

Because $\lambda_n = 0$, we still have a finite system of orthogonal polynomials.

**Laguerre polynomials**: for $s(x) = x$ and $\tau(x) = -x + \alpha + 1$ one has

$$\rho(x) = \frac{x^\alpha e^{-x}}{\Gamma(\alpha+1)} \qquad \text{Gamma distribution, } \Gamma(\alpha, 1).$$

Moreover $\lambda_n = 1$. Orthogonality relation is satisfied by Laguerre polynomials for

$$a = 0, \quad b = \infty, \quad \alpha > 1.$$

**Hermite polynomials**: let $s(x) = 1$ and

$$\rho(x) = \frac{e^{-x^2}}{\sqrt{\pi}}, \quad \text{Normal distribution } N(0, 1/2),$$

then $\tau(x) = -2x$ and $\lambda_n = 2n$.

Hermite polynomials are orthogonal on $\mathbb{R}$. Generally one works with rescaled Hermite polynomials $H_n(x/\sqrt{2})$, orthogonal with respect to $N(0,1)$ distribution

$$\rho(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}.$$

### 5.1.2 Orthogonal polynomials of a discrete variable

In order to find an explicit expression for $\rho(x)$ it is useful to rewrite (5.4) in the form

$$\frac{\rho(x+1)}{\rho(x)} = \frac{s(x) + \tau(x)}{s(x+1)};$$

The possible occurrences are (see Schoutens' book [Sch00] for all the details):

| Name | Symbol | deg $s(x)$ |
|---|---|---|
| Hahn | $Q_n(x; \alpha, \beta, N)$ | 2 |
| Meixner | $M_n(x; \gamma, \mu)$ | 1 |
| Krawtchouk | $K_n(x; p, N)$ | 1 |
| Charlier | $C_n(x; a)$ | 1 |

It can be observed that the Charlier polynomials are limit cases of Krawtchouk (for $p = aN^{-1}, N \to \infty$) and Meixner (for $\mu = a(a+\gamma)^{-1},\ \gamma \to \infty$) polynomials, which are themselves limit cases of the Hahn polynomials ($\alpha = pt, \beta = (1-p)t$ with $p \in (0,1)$ and $t \to \infty$ for Krawtchouk, and $\alpha = \gamma - 1,\ \beta = N(1-\gamma)\gamma^{-1}$ and $N \to \infty$ for Meixner).

It is known that all orthogonal polynomials $\{Q_n(x)\}$ on the real line satisfy a three-term recurrence relation

$$-xQ_n(x) = b_n Q_{n+1}(x) + \gamma_n Q_n(x) + c_n Q_{n-1}(x), \quad n \geq 1 \tag{5.5}$$

where in general $b_n, c_n \neq 0$ and $c_n/b_{n-1} > 0$.

For Hahn polynomials for instance one has that

$$\begin{aligned} b_n &= \frac{(n+\alpha+\beta+1)(n+\alpha+1)(N-n)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)}, \\ c_n &= \frac{n(n+\beta)(n+\alpha+\beta+N+1)}{(2n+\alpha+\beta)(2n+\alpha+\beta+1)}, \\ \gamma_n &= -(b_n + c_n) \end{aligned}$$

and moreover, the relation (5.5) can be translated into a recurrence relation which can be expressed in terms of a second-order difference equation of the form

$$a(x)Q_n(\lambda(x+1)) + b(x)Q_n(\lambda(x)) + c(x)Q_n(\lambda(x-1)) = -\lambda_n Q_n(\lambda(x)), \tag{5.6}$$

where in this case $\lambda(x) = x(x+\alpha+\beta+1)$, a quadratic function of $x$, $\lambda_n = n$, $a(x) = b_n$, $b(x) = -(a(x) + c(x))$, and $c(x) = c_n$, with $b_n, c_n$ the coefficient in the recurrence relation (5.5).

Also, it is a natural question to ask for other orthogonal polynomials to be eigenfunctions of a second order difference equation of the form (5.6): it can be shown that a set of families of orthogonal polynomials together with limit transitions between them exists and satisfies the previous request. The set is called the *Askey scheme* of hypergeometric orthogonal polynomials.

**Meixner-Pollaczeck polynomials**

As a limit case of both Hahn and dual Hahn polyomials (not introduced here, for reference see [Sch00]) stands the family of Meixner-Pollaczeck polynomials, defined by

$$P_n^{(a)}(x;\phi) = \frac{(2a)_n \exp\{in\phi\}}{n} \, {}_2F_1(-n, a+ix; 2a; 1-e^{-2i\phi}),$$

where $a > 0, 0 < \phi < \pi$, and $(a)_n$ is the *Pochhammer symbol*, defined in terms of Euler Gamma function as

$$(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)}, \quad n > 0;$$

moreover, the notation ${}_2F_1(-n, a_2; b_1; z)$ is a particular case of the generalized hypergeometric series with the first numerator parameter equal to a negative integer. Namely

$${}_pF_q(-n, \ldots, a_p; b_1, \ldots, b_q; z) = \sum_{j=0}^{n} \frac{(-n)_j \ldots (a_p)_j}{(b_1)_j \ldots (b_q)_j} \frac{z^j}{j}.$$

The orthogonality relation for these polynomials is given by the following

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{(2\phi-\pi)x} |\Gamma(a+ix)|^2 P_m^{(a)}(x;\phi) P_n^{(a)}(x;\phi)\, dx =$$
$$= \frac{\Gamma(n+2a)}{(2\sin\phi)^{2a} n!} \delta_{m,n}, \quad a > 0,\ 0 < \phi < \pi,$$

while the recurrence relation (5.5) is obtained as

$$(n+1)P_{n+1}^{(a)}(x;\phi) - 2[x\sin\phi + (n+a)\cos\phi]P_n^{(a)}(x;\phi) + (n+2a-1)P_{n-1}^{(a)}(x;\phi) = 0.$$

For further technical details see also (http://fa.its.tudelft.nl/∼koekoek/askey.html, *The Askey scheme of hypergeometric orthogonal polynomials and its q-analogue*, by R. Koekoek and R. F. Swarttouw). Observe that Laguerre polynomials introduced above are limit cases of Meixner-Pollaczeck polynomials, that is

$$\lim_{\phi \to 0} P_n^{(\frac{1}{2}\alpha + \frac{1}{2})} \left( -\frac{x}{2\phi}; \phi \right) = L_n^{(\alpha)}(x)$$

# 5.2 Connection between orthogonal polynomials and Lévy processes

## 5.2.1 Sheffer Polynomials

Let $f(t)$ and $g(t)$ be functions for which all the necessary derivatives are defined: using the classical Faá di Bruno formula (explicit $n-$th derivative formula of the composition $f(g(t))$ ), it can be shown that the equation

$$f(z)\exp\{xu(z)\} = \sum_{m=0}^{\infty} Q_m(x)\frac{z^m}{m!} \tag{5.7}$$

generates a family of polynomials $\{Q_m(x), m \geq 0\}$ when both functions $u(z)$ and $f(z)$ can be expanded in a formal power series and if $u(0) = 0$, $u'(0) \neq 0$ and $f(0) \neq 0$.

The polynomials $Q_m(x)$ so defined are of exact degree $m$ and are called *Sheffer polynomials*. Any set of such polynomials is called *Sheffer set*.

## 5.2.2 Lévy-Sheffer systems

Let now be $\tau$ the inverse functon of $u$, such that $\tau(u(z)) = z$. Then $\tau$ also can be expanded formally in a power series with $\tau(0) = 0$ and $\tau'(0) \neq 0$.

Introduce now an additional parameter $t \geq 0$ into the polynomials defined in (5.7) by replacing $f(z)$ with $f(z)^t$.

**Definition 35.** *A polynomial set $\{Q_m(x,t), m \geq 0, t \geq 0\}$ is called a Lévy-Sheffer system if it is defined by a generating function of the form*

$$f(z)^t\exp\{xu(z)\} = \sum_{m=0}^{\infty} Q_m(x,t)\frac{z^m}{m!} \tag{5.8}$$

*where*

*(i) $f(z)$ and $u(z)$ are analytic in a neighborhood of $z = 0$;*

*(ii)* $u(0) = 0, f(0) = 1$ *and* $u'(0) \neq 0$*;*

*(iii)* $1/f(\tau(i\theta))$ *is an infinitely divisible characteristic function.*

If condition $(iii)$ is satisfied there is a Lévy process $\{X_t, t \geq 0\}$ defined by the function

$$\phi(\theta) = \phi_X(\theta) = \frac{1}{f(\tau(i\theta))} \tag{5.9}$$

The basic link between the polynomials and the corresponding Lévy processes is the following martingale equality

$$E[Q_m(X_t, t)|X_s] = Q_m(X_s, s), \quad 0 \leq s \leq t, \quad m \geq 0. \tag{5.10}$$

In fact it holds, for the left hand side, that

$$\sum_{m=0}^{\infty} E[Q_m(X_t, t)|X_s] \frac{z^m}{m!} = E\left[\sum_{m=0}^{\infty} Q_m(X_t, t)\frac{z^m}{m!}\middle| X_s\right] =$$
$$= E[f(z)^t \exp\{u(z)X_t\}|X_s] =$$
$$= f(z)^t \exp\{u(z)X_s\} E[\exp\{u(z)(X_t - X_s)\}|X_s];$$

For the righthand side of (5.10)

$$\sum_{m=0}^{\infty} Q_m(X_s, s)\frac{z^m}{m!} = f(z)^s \exp\{u(z)X_s\}$$

And then the combination of the two expressions above leads to

$$E[\exp\{u(z)(X_t - X_s)\}|X_s] = f(z)^{s-t};$$

comparing this relationship with the equation determining the Lévy process

$$E[\exp\{i\theta(X_t - X_s)\}|X_s] = \phi(\theta)^{t-s},$$

it can be observed that (5.10) holds true if and only if (5.9) holds.

### 5.2.3 Meixner set of orthogonal polynomials

In his historic paper [Mei34], J.Meixner determined all sets of orthogonal polynomials that satisfy relation (5.7). Here is a sketch of Meixner's approach, as some elements will be useful in the following. From (5.7) one has

$$\tau(D)Q_m(x) = mQ_{m-1}(x), \quad m \geq 0,$$

where $D = d/dx$ the differential operator with respect to $x$. This one in turn leads to

$$\tau(D)(xQ_m(x)) = \tau'(D)Q_m(x) + mxQ_{m-1}(x), \quad m \geq 0.$$

By Favard's theorem, the monic set $\{Q_m(x), m \geq 0\}$ will be orthogonal if and only if the polynomials satisfy a three-term recurrence relation

$$Q_{m+1}(x) = (x + l_{m+1})Q_m(x) + k_{m+1}Q_{m-1}(x), \tag{5.11}$$

where $l_m \in \mathbb{R},\ k_m < 0,\ m \geq 2$. Now multiply by $m$ relation (5.11) rewritten for $Q_m(x)$, and subtract this from (5.11) with $\tau(D)$ applied to it:

$$\begin{aligned}
&(1 - \tau'(D))Q_m(x) = \\
&\quad = (l_{m+1} - l_m)mQ_{m-1}(x) + \left(\frac{k_{m+1}}{m} - \frac{k_m}{m-1}\right) m(m-1)Q_{m-2}(x);
\end{aligned}$$

after shifting $m$ to $m+1$ and applying $\tau(D)$ again, one gets:

$$\begin{aligned}
&(1 - \tau'(D))Q_m(x) = \\
&\quad = (l_{m+2} - l_{m+1})mQ_{m-1}(x) + \left(\frac{k_{m+2}}{m+1} - \frac{k_{m+1}}{m}\right) m(m-1)Q_{m-2}(x).
\end{aligned}$$

Comparing the last two formulas, it can be obtained that

$$\begin{aligned}
&l_{m+1} - l_m = \lambda, \quad m \geq 1, \\
&\frac{k_{m+1}}{m} - \frac{k_m}{m-1} = \kappa, \quad m \geq 2, \\
&(1 - \tau'(D))Q_m(x) = \lambda\tau(D)Q_m(x) + \kappa\tau^2(D)Q_m(x), \quad m \geq 0,
\end{aligned} \tag{5.12}$$

and so (5.11) becomes

$$Q_{m+1}(x) = (x + l_1 + m\lambda)Q_m(x) + m[k_2 + (m-1)\kappa]Q_{m-1}(x), \quad m \geq 0, \tag{5.13}$$

with $k_2 < 0$ and $\kappa \leq 0$. From (5.12) follows that

$$\tau'(y) = 1 - \lambda\tau(y) - \kappa\tau^2(y). \tag{5.14}$$

Moreover, from (5.13), using the fact that

$$f(z) = \sum_{m=0}^{\infty} Q_m(0)\frac{z^m}{m!},$$

the following differential equation for $f(z)$ can be obtained:

$$\frac{f'(z)}{f(z)} = \frac{k_2 z + l_1}{1 - \lambda z - \kappa z^2}. \tag{5.15}$$

Two quantities $\alpha, \beta$ are defined by the equation

$$1 - \lambda z - \kappa z^2 = (1 - \alpha z)(1 - \beta z),$$

where $\alpha\beta > 0$. Now equation (5.15) can be obviously rewritten, and from (5.14) a differential equation for $u(z)$ can be obtained:

$$u'(z) = \frac{1}{(1-\alpha z)(1-\beta z)}.$$

The explicit solutions of these equations, although a little complicated, are:

$$u(z) = \begin{cases} \frac{1}{\alpha-\beta}\log\left(\frac{1-\beta z}{1-\alpha z}\right), & \text{if } \alpha \neq \beta, \\ \frac{z}{1-\alpha z} & \text{if } \alpha = \beta; \end{cases}$$

$$\log f(z) = \begin{cases} -\frac{(k+\alpha\ell)\log(1-\alpha z)}{\alpha(\alpha-\beta)} + \frac{(k+\beta\ell)\log(1-\beta z)}{\beta(\alpha-\beta)}, & 0 \neq \alpha \neq \beta \neq 0 \\ \frac{k\log(1-\alpha z)}{\alpha^2} + \frac{k+\alpha\ell}{\alpha}\frac{z}{1-\alpha z}, & \alpha = \beta \neq 0 \\ -\frac{(k+\alpha\ell)\log(1-\alpha z)}{\alpha^2} - \frac{kz}{\alpha}, & \alpha \neq \beta = 0 \\ \frac{k}{2}z^2 + \ell z, & \alpha = \beta = 0 \end{cases}$$

and also

$$\tau(s) = \begin{cases} \frac{\exp(\beta s)-\exp(\alpha s)}{\beta\exp(\beta s)-\alpha\exp(\alpha s)}, & \text{if } \alpha \neq \beta, \\ \frac{s}{1+\alpha s}, & \text{if } \alpha = \beta \end{cases}$$

Applying the above to equation (5.8) we obtain a Lévy-Sheffer system with orthogonal polynomials. Since the explicit form of the functions $f$ and $\tau$ is now known, we can identify the ingredients in the following Kolmogorov representation (5.16): this finally determines the underlying process.

After putting $\ell = -\mu,\ k = -\sigma^2$ it can also be found that

$$\log\phi(\theta) = i\mu\theta + \sigma^2\int_0^{i\theta}\tau(z)\,dz.$$

The identification of $c$ and $K$ in the Kolmogorov representation

$$\log\phi(\theta) = ic\theta + \int_{-\infty}^{+\infty}(e^{i\theta x} - 1 - i\theta x)\frac{dK(x)}{x^2}, \tag{5.16}$$

where $c$ is a real constant and $K(y)$ is a nondecreasing and bounded function such that $K(-\infty) = 0$, is done as follows: by taking derivatives in (5.16) at $\theta = 0$ we see that $ic = \phi'(0) = iE[X_1] = -i\ell$, and hence $c = \mu$. Taking another derivative, the equation

$$\int_{-\infty}^{+\infty}\exp\{i\theta x\}dK(x) = \sigma^2\tau'(i\theta)$$

is obtained, which determines $K$ uniquely. The results of subseqent calculations lead to the explicit form

$$\int_{-\infty}^{+\infty}\exp\{i\theta x\}dK(x) = \begin{cases} \frac{\sigma^2(\alpha-\beta)^2\exp(i(\alpha+\beta)\theta)}{(\beta\exp(i\beta\theta)-\alpha\exp(i\alpha\theta))^2}, & \text{if } \alpha \neq \beta, \\ \left(\frac{\sigma}{1+i\alpha\theta}\right)^2, & \text{if } \alpha = \beta. \end{cases}$$

It can be also verified that the function $\int_{-\infty}^{+\infty}\exp\{i\theta x\}d(K(x)/K(\infty))$ is indeed a characteristic function for all values of $\alpha, \beta$ with $\alpha\beta > 0$. To simplify the further analysis, without loss of generality the following choice is made:

$$\ell = c = 0,\ k = -K(\infty) = -\sigma^2 = -1.$$

This way the following holds:

$$\psi(\theta) = \log \phi(\theta) = \begin{cases} \frac{i\theta(\alpha+\beta)+\log((\alpha-\beta)/(\alpha e^{i\alpha\theta}-\beta e^{i\beta\theta}))}{\alpha\beta}, & \text{if } 0 \neq \alpha \neq \beta \neq 0, \\ \frac{i\theta}{\alpha} - \frac{\log(1+i\alpha\theta)}{\alpha^2}, & \text{if } \alpha = \beta \neq 0, \\ \frac{i\theta}{\alpha} - \frac{(1-\exp(-i\alpha\theta))}{\alpha^2}, & \text{if } \alpha \neq \beta = 0, \\ -\frac{\theta^2}{2}, & \text{if } \alpha = \beta = 0. \end{cases}$$

### 5.2.4 Lévy-Meixner Systems

The approach now is to link all Meixner's polynomials to a unique Lévy process. The starting form for the polynomials will be (5.8) while for the Lévy process will be the forms

$$\log E[e^{i\theta X_t}] = t\psi(\theta) = t \log \phi_X(\theta) = t \int_{-\infty}^{+\infty} (e^{i\theta x} - 1 - i\theta x) \frac{dK(x)}{x^2},$$

where $K$ is a probability measure. The two elements are linked by equation

$$\psi(\theta) = -\log f(\tau(i\theta)).$$

The measure of orthogonality $\Psi_t(x)$ is also the distribution function of Lévy process $X_t$. Indeed, by taking generating functions in

$$\int_{-\infty}^{+\infty} Q_m(x,t) Q_n(x,t)\, d\Psi_t(x) = \delta_{mn} c_m(t)$$

and setting $n = 0$, one has

$$\int_{-\infty}^{+\infty} f(z)^t \exp\{xu(z)\}\, d\Psi_t(x) = c_0 = 1.$$

Putting $u(z) = i\theta$ so that $z = \tau(i\theta)$, finally produces

$$\int_{-\infty}^{+\infty} \exp(i\theta x)\, d\Psi_t(x) = \left( \frac{1}{f(\tau(i\theta))} \right)^t = E[\exp(i\theta X_t)].$$

## 5.3 Meixner process from orthogonal polynomials

### 5.3.1 Meixner process fromMeixner-Pollaczeck polynomials

When $\alpha \neq 0, \beta = \bar{\alpha}$, the fundamental relations become

$$\begin{aligned} u(z) &= \frac{1}{\alpha - \bar{\alpha}} \log\left(\frac{1 - \bar{\alpha}z}{1 - \alpha z}\right), \\ f(z) &= (1 - \alpha z)^{\frac{1}{\alpha(\alpha - \bar{\alpha})}} (1 - \bar{\alpha}z)^{-\frac{1}{\bar{\alpha}}(\alpha - \bar{\alpha})}, \\ \psi(\theta) &= i\frac{\alpha + \bar{\alpha}}{\alpha\bar{\alpha}}\theta + \frac{1}{\alpha\bar{\alpha}} \log\left(\frac{\alpha - \bar{\alpha}}{\alpha \exp(i\alpha\theta) - \bar{\alpha}\exp(i\bar{\alpha}\theta)}\right) \end{aligned} \tag{5.17}$$

and the following expression is obtained for the basic polynomials

$$\sum_{m=0}^{\infty} Q_m(x;t)\frac{z^m}{m} = (1 - \alpha z)^{\frac{t - \alpha x}{\alpha(\alpha - \bar{\alpha})}} (1 - \bar{\alpha}z)^{\frac{(\bar{\alpha}x - t)}{\bar{\alpha}(\alpha - \bar{\alpha})}}$$

Since $\beta = \bar{\alpha}$ it is natural to write $\alpha = \rho \exp(i\zeta)$; it is necessary now to identify function $\psi(\theta)$ above with a suitable variant $\psi_H(\theta)$. With the introduced expression for $\alpha$ the argument within the logarithm in the expression for $\psi(\theta)$ is rewritten in the form

$$\frac{\alpha - \bar{\alpha}}{\alpha \exp(i\alpha\theta) - \bar{\alpha}\exp(i\bar{\alpha}\theta)} = \exp(-i\theta\rho\cos\zeta)\frac{\sin\zeta}{\sin(\zeta + i\theta\rho\sin\zeta)}. \tag{5.18}$$

Hence we put $\zeta = \pi/2 + a/2$ in the expression for $\psi_H(\theta)$. Taking $\mu = 1$ it holds that

$$\psi(\theta) = i\frac{\theta}{\rho}\cos\zeta + \frac{1}{2\rho^2}\psi_H(2\rho\theta\sin\zeta). \tag{5.19}$$

Recall that a function $\psi : \mathbb{R} \to \mathbb{C}$ is the characteristic exponent of an infinitely divisible distribution if and only if there are constants $a \in \mathbb{R}$, $\sigma^2 \geq 0$ and a measure $\nu$ on $\mathbb{R} \setminus \{0\}$ with $\int_{-\infty}^{+\infty} \min\{1, x^2\}\nu(dx) < \infty$ such that

$$\psi(\theta) = ia\theta - \frac{\sigma^2}{2}\theta^2 + \int_{-\infty}^{+\infty} (\exp(i\theta x) - 1 - i\theta x \mathbb{1}_{\{|x|<1\}})\nu(dx)$$

for every $\theta$, where the measure $\nu$ is the Lévy measure.

Also remember that if we have an infinitely divisible distribution with characteristic function $\phi(\theta)$, a Lévy process $X_t$ remains defined through the relations

$$\exp(\psi_X(\theta)) = \phi_X(\theta) = E[\exp(i\theta X_1)] = \phi(\theta).$$

Now it can be observed that if $\{Y_t, t \geq 0\}$ is a Lévy process with characteristic function

$$E[e^{i\theta Y_t}] = \exp\{t\psi_Y(\theta)\},$$

then also $X_t = At + BY_{Ct}$, with $C > 0$ is a Lévy process determined by

$$\psi_X(\theta) = i\theta A + C\psi_Y(B\theta) \tag{5.20}$$

From (5.19), the identification between the processes $\{X_t, t \geq 0\}$ and Meixner process $\{H_t, t \geq 0\}$ is then achieved by choosing

$$A = \frac{1}{\rho}\cos\zeta, \quad B = 2\rho\sin\zeta, \quad C = (2\rho^2)^{-1}.$$

So

$$X_t = \frac{t}{\rho}\cos\zeta + 2\rho\sin\zeta H_{t/(2\rho^2)}$$

The Meixner-Pollaczeck polynomial is defined for $\lambda > 0$ and $0 < \zeta < \pi$ by

$$\sum_{m=0}^{\infty} P_m(y;\lambda,\zeta)\frac{w^m}{m!} = \frac{(1-\exp\{i\zeta\}w)^{-\lambda+iy}}{(1-\exp\{-i\zeta\}w)^{\lambda+iy}}$$

Here the identification is simple and leads to

$$w = z\rho, \quad \lambda = \frac{t}{2\rho^2}, \quad y = \frac{x}{2\rho\sin\zeta} - \frac{t}{2\rho^2}\cot\zeta.$$

Moreover the equality

$$Q_m(x,t) = m!\rho^m P_m\left(\frac{x}{2\rho\sin\zeta} - \frac{t}{2\rho^2}\cot\zeta,\ \frac{t}{2\rho^2},\ \zeta\right),$$

easily brings to the martingale expression

$$E\left[P_m\left(H_{\frac{t}{2\rho^2}};\ \frac{t}{2\rho^2},\ \zeta\right)\middle| H_{\frac{s}{2\rho^2}}\right] = P_m\left(H_{\frac{s}{2\rho^2}};\ \frac{s}{2\rho^2},\ \zeta\right).$$

A consequence is that the Meixner$(1,\ 2\zeta-\pi,\delta,0)$ distribution is the measure of orthogonality of the Meixner-Pollaczeck polynomials $\{P_n(x;\delta,\zeta), n = 0,1,\ldots\}$. Moreover the monic Meixner-Pollaczeck polynomials $\{\tilde{P}_n(x;\delta,\zeta), n = 0,1,\ldots\}$ are martingales for the Meixner process $(\alpha = 1, \delta = 1, \zeta = (\beta+\pi)/2)$:

$$E\left[\tilde{P}_n\left(H_t;\ t,\ \zeta\right)\middle| H_s\right] = \tilde{P}_m\left(H_s;\ s,\ \zeta\right).$$

It remains to determine $K$; using the exponential form $\alpha = \rho \exp(i\zeta)$ one gets

$$\int_{-\infty}^{+\infty} \exp(i\theta x) dK(x) = \left( \frac{\sin\zeta}{\sin(\zeta + i\theta\rho\sin\zeta)} \right)^2 .$$

A little algebra reveals that $K$ has a derivative with expression

$$\frac{dK(y)}{dy} = \frac{\sin\zeta}{\pi\rho} \left| \Gamma\left(1 - \frac{iy}{2\rho\sin\zeta}\right) \right|^2 \exp\left( - \frac{y(\pi - 2\zeta)}{2\rho\sin\zeta} \right).$$

We have now tried autonomously to obtain the expression of the moments of Meixner distribution adopting as a starting point the results for Meixner-Pollaczeck polynomials, in particular formulae (5.17), (5.18), (5.19); calculations have been performed both keeping $\zeta$ as variable, and changing it to $a$ as described above. the results are as follows. Recall that for now $\mu = 1, \delta = 1$. From (5.17), the general characteristic exponent

$$\psi(\theta) = \frac{2i\rho\theta\cos\zeta}{\rho^2} + \frac{1}{\rho^2} \log\left( \frac{e^{-i\theta\rho\cos\zeta}\sin\zeta}{\sin(\zeta + i\theta\rho\sin\zeta)} \right)$$

with $\rho > 0,\ 0 < \zeta < 2\pi$. It leads, as for (5.19) to

$$\psi_H(t) = 2\rho^2 \psi\left( \frac{t}{2\rho\sin\zeta} \right) - it\cot\zeta \tag{5.21}$$

Now the characteristic function $\phi_M(t)$ of Meixner process $H_t$ at $t = 1$ is obtainable as

$$\phi_H(t) = e^{\psi_H(t)} = -\csc^2\left( \frac{x - 2i\zeta}{2} \right) \sin^2\zeta = \cos^2\left(\frac{a}{2}\right) \sec^2\left( \frac{a + it}{2} \right). \tag{5.22}$$

Now the calculation proceeds as usual by differentiation of the characheristic function to evaluate the first two moments, the skewness and kurtosis parameter.

For the expected value we have

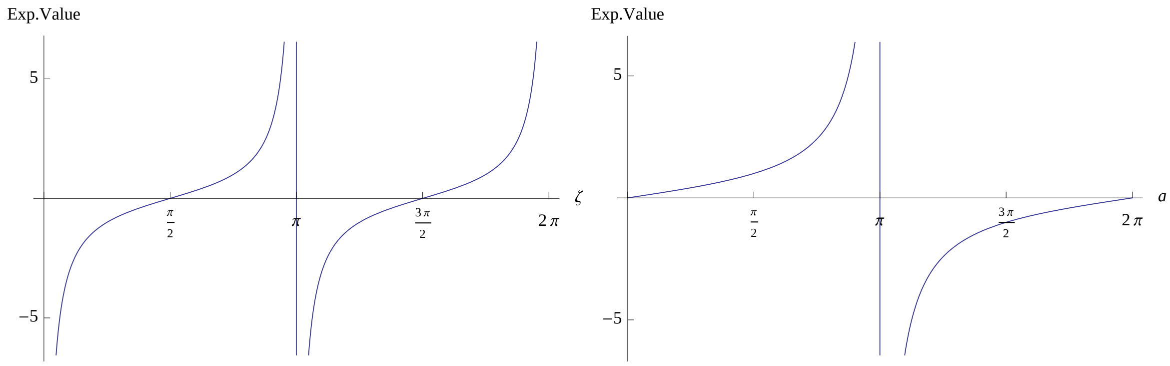


Figure 5.1: Expected value of H(1)

while for variance we obtained that

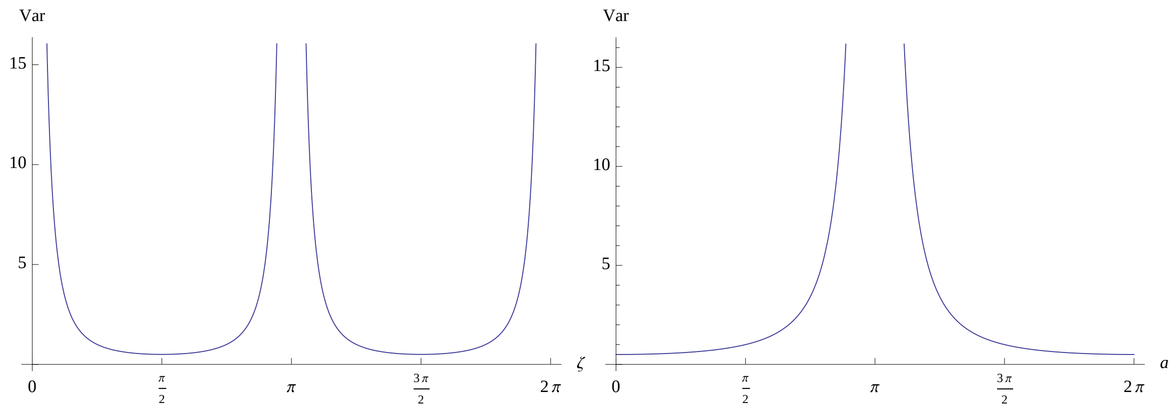


Figure 5.2: Variance of H(1)

As far as the variance, the two expression obtained are

$$Var(H(1)) = \frac{\csc^2 \zeta}{2} = \frac{1}{1+\cos(a)},$$

while for sekewness parameter it holds that

$$\kappa_3 = -\cot \zeta \sqrt{\frac{2}{\csc^2 \zeta}} = \sin(a)\sqrt{\frac{1}{1+\cos(a)}},$$

and the graphics are the following

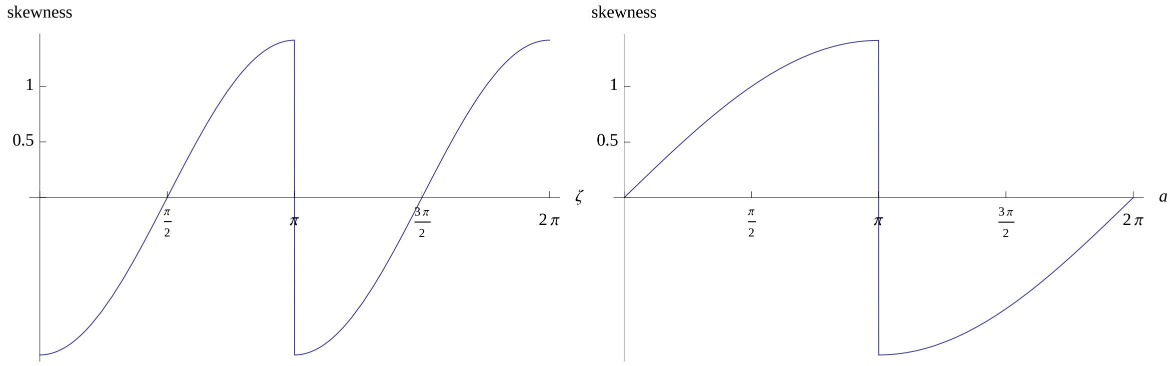


Figure 5.3: Skewness parameter of H(1)

which are obvoiusly more pitchable with respect to the constant null skewness of a Brownian motion, allowing in this case more fitting to the eventually available data. Last the kurtosis, for which similar observations dan be drawn, whose equations are

$$\kappa_4 = 5 + \cos(2\zeta) = 5 + \cos(a)$$

and whose representations

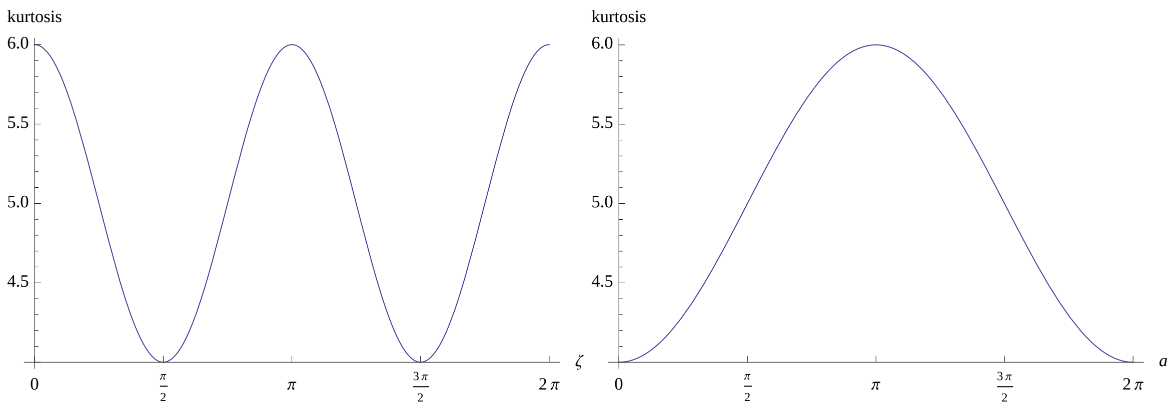


Figure 5.4: Kurtosis parameter of H(1)

From a strictly theoretical point of view, following the approach of Dominici in [Dom07] it is possible to evaluate the Fisher information of the Meixner-Pollaczeck orthogonal polynomials, a concept introduced for general orthogonal polynomials by Sanchéz-Ruiz and Dehesa in [SRD05]. They considered a sequence of real polynomials, orthogonal with respect to the weight function $\rho(x)$ on the interval $[a, b]$

$$\int_a^b P_n(x)P_m(x)\rho(d)dx = h_n\delta_{n,m}, \quad m, n = 0, 1, \dots \tag{5.23}$$

with $\deg(P_n) = n$. Introducing the normalized density functions

$$\rho_n(x) = \frac{[P_n(x)]^2\rho(x)}{h_n}, \tag{5.24}$$

they in fact defined the Fisher information corresponding to the densities (5.24) by

$$\mathcal{I}(n) = \int_a^b \frac{[\rho_n'(x)]^2}{\rho_n(x)}. \tag{5.25}$$

Applying the last formula to the classical hypergeometric polynomials, in [SRD05] $\mathcal{I}(n)$ for Jacobi, Laguerre and Hermite polynomials is evaluated.

Let us state the main theorem in [Dom07].

**Theorem 42.** *The Fisher information of the Meixner-Pollaczeck polynomials is given by*

$$I_\phi(P_n^{(a)}) = \int_{-\infty}^{+\infty} \left[\frac{\partial}{\partial\theta}\rho_n(x)\right]^2 \frac{1}{\rho_n(x)}\,dx = \frac{2[n^2 + (2n+1)a]}{\sin^2(\phi)},\ n = 0, 1, \dots \tag{5.26}$$

*Proof*: we have seen that the Meixner-Pollaczeck have the hypergeometric representation

$$P_n^{(a)}(x;\phi) = \frac{(2a)_n e^{in\phi}}{n!}\,{}_2F_1(-n, a+ix, 2a; 1 - e^{-2i\phi}), \quad a > 0,\ 0 < \phi < \pi. \tag{5.27}$$

They satisfy the orthogonality relation

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{(2\phi-\pi)x}|\Gamma(a+ix)|^2 P_m^{(a)}(x;\phi)P_n^{(a)}(x;\phi)dx = \frac{\Gamma(n+2a)\delta_{n,m}}{[2\sin(\phi)]^{2a}n!}, \tag{5.28}$$

for $m, n = 0, 1, \dots$, and the recurrence relation

$$(n+1)P_{n+1}^{(a)} - 2[x\sin(\phi) + (n+a)\cos(\phi)]P_n^{(a)} + (n+2a-1)P_{n-1}^{(a)} = 0. \tag{5.29}$$

Also, a general family of orthogonal polynomials $P_n(x)$ defined by

$$P_n(x) = {}_2F_1(-n, -x, c; z(\theta)), \quad n = 0, 1, \dots \tag{5.30}$$

is such that

$$\frac{\partial P_n}{\partial \theta} = n\frac{z'}{z}[P_n(x) - P_{n-1}(x)], \quad n = 0, 1, \dots \tag{5.31}$$

From (5.31) and (5.27) we have

$$\frac{\partial P_n^{(a)}}{\partial \phi} = n\cot(\phi)P_n^{(a)} - \frac{(n+2a-1)}{\sin(\phi)}P_{n-1}^{(a)},$$

while (5.24) and (5.28) lead to

$$\rho_n(x) = \frac{e^{(2\phi-\pi)x}|\Gamma(a+ix)|^2[2\sin(\phi)]^{2a}n![P_n^{(a)}(nx;\phi)]^2}{2\pi\Gamma(n+2a)}. \tag{5.32}$$

Observe that for $n = 0, 1, \dots$, $\int_{-\infty}^{+\infty} \rho_n(x)dx = 1$.

Differentiating (5.32) with respect to $\phi$ one can get

$$\frac{\partial \rho_n}{\partial \phi} = \frac{2\rho_n(x)}{P_n^{(a)}}\left\{[x + (n+a)\cot(\phi)]P_n^{(a)} - \frac{n+2a-1}{\sin(\phi)}P_{n-1}^{(a)}\right\} = \tag{5.33}$$

$$\overset{(5.29)}{=} \frac{\rho_n(x)}{\sin(\phi)P_n^{(a)}}\left[(n+1)P_{n+1}^{(a)} - (n+2a-1)P_{n-1}^{(a)}\right]. \tag{5.34}$$

Therefore:

$$\begin{aligned}\left[\frac{\partial}{\partial\theta}\rho_n(x)\right]^2 \frac{1}{\rho_n(x)} &= \frac{\rho_n(x)}{\sin^2(\phi)[P_n^{(a)}]^2}\Bigg[(n+1)^2[P_{n+1}^{(a)}]^2 + \\ &- 2(n+1)(n+2a-1)P_{n+1}^{(a)}P_{n-1}^{(a)} + (n+2a-1)^2[P_{n-1}^{(a)}]^2\Bigg] = \\ &= \frac{1}{\sin^2(\phi)}[(n+1)(n+2a)\rho_{n+1}(x) + n(n+2a-1)\rho_{n-1}(x) + \\ &- 2(n+1)(n+2a-1)\frac{[2\sin(\phi)]^{2a}n!}{\Gamma(n+2a)}\rho(x)P_{n+1}^{(a)}P_{n-1}^{(a)}],\end{aligned} \tag{5.35}$$

with

$$\rho(x) = \frac{e^{(2\phi-pi)x}|\Gamma(a+ix)|^2}{2\pi}.$$

By integration of (4.30) and from (5.28) one can get

$$\left[\frac{\partial}{\partial\theta}\rho_n(x)\right]^2 \frac{1}{\rho_n(x)}dx = \frac{1}{\sin^2(\phi)}[(n+1)(n+2a) + n(n+2a-1)],$$

which concludes the proof. □

# Chapter 6

# Further properties of Meixner process and some applications

Last part of our work deals with applications and graphical representations. First, the role of the parameters of Meixner distributions is established by means of graphics which show how the density responds to their variation. The strong feature of such a distribution is basically to have three independent scale parameters: $\alpha$ and $\delta$ act mostly in the same way, pitching the kurtosis of the distribution, while parameter $\beta$ is mostly responsible of the skewness. In addition, parameter $\mu$ takes account of the location of the distribution.

The main contribution in this chapter is the simulation of trajectories of Meixner process, which is as far as we know missing from relative literature. The routine was built from scratch following the theoretical way hinted by the already cited work by Madan and Yor. In this case it is also possible to understand how the variation of the parameters reflects on the trajectories of the process, besides the shape of the distribution.

The application of this model for the description of financial log returns descends basically from the works by Eberlein, [Ebe09], Eberlein and Keller [EK95], and Barndorff-Nielsen [BN78, BN97, BN98], who first tried to fit hyperbolical distributions to data of the same kind.

A similar approach has been given by Schoutens for instance in [Sch02], but with different model diagnostics from the ones we describe here.

## 6.1 Properties of Meixner distribution

This first section of the last chapter will be devoted to the graphical study of the properties of Meixner distribution. Here in fact it will e shown the action of different values of the parameters on the corresponding Meixner distribution.

Starting from equation (3.1)

$$f_M(x;\alpha,\beta,\delta,\mu) = \frac{(2\cos(\beta/2))^{2\delta}}{2\alpha\pi\gamma(2\delta)} \exp\left\{\frac{\beta(x-\mu)}{\alpha}\right\} \left|\Gamma\left(\delta + \frac{i(x-\mu)}{\alpha}\right)\right|^2, \tag{6.1}$$

it is easily observed first of all, that parameter $\mu$ is a location parameter, as the following graphics show

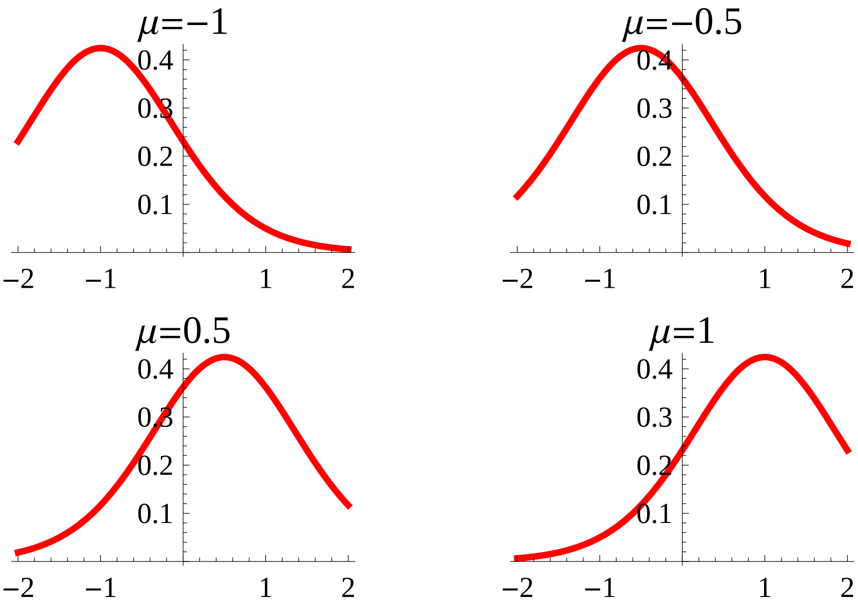


Figure 6.1: Meixner densities $MD(1, 0, 2, \mu)$, $\mu = -1, -0.5, 0.5, 1$

From the graphics the resemblance of the Meixner distribution with the usual Gaussian curve seems explicit, but every doubt is cleared with the following comparison

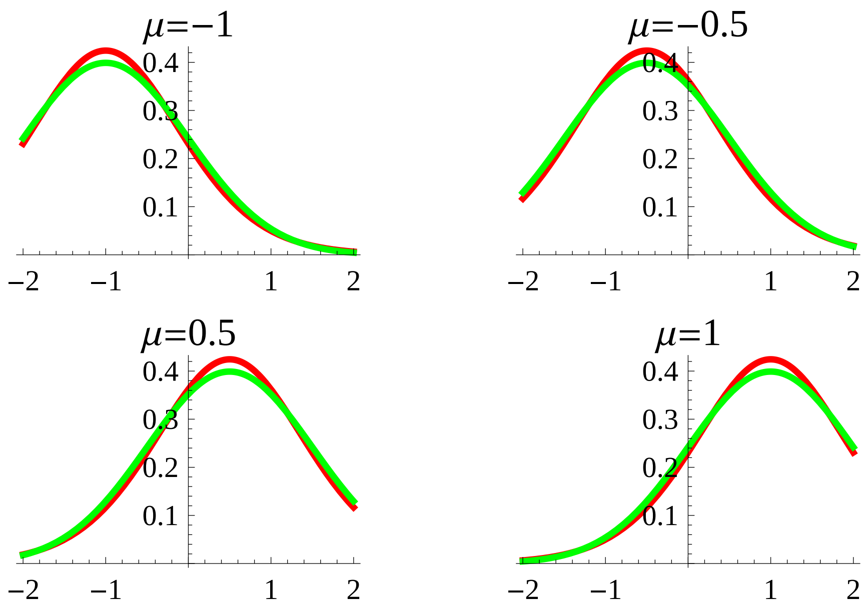


Figure 6.2: Meixner densities (*red*) $MD(1, 0, 2, \mu)$, $\mu = -1, -0.5, 0.5, 1$, compared with the Gaussian curves (*green*) having the corresponding mean and variance

As we could expect, the Gaussian curves seems ”less sharpened”, in a way that will be clearer in the following analysis. From this point on, parameter $\mu$ will be set equal to $0$ for the sake of simplicity.

Parameter $\alpha > 0$ in the distribution is the major responsible of the “excess kurtosis”, i.e. the accumulation of “more” (with respect to the Gaussian benchmark) mass above the mode of the distribution. In fact:

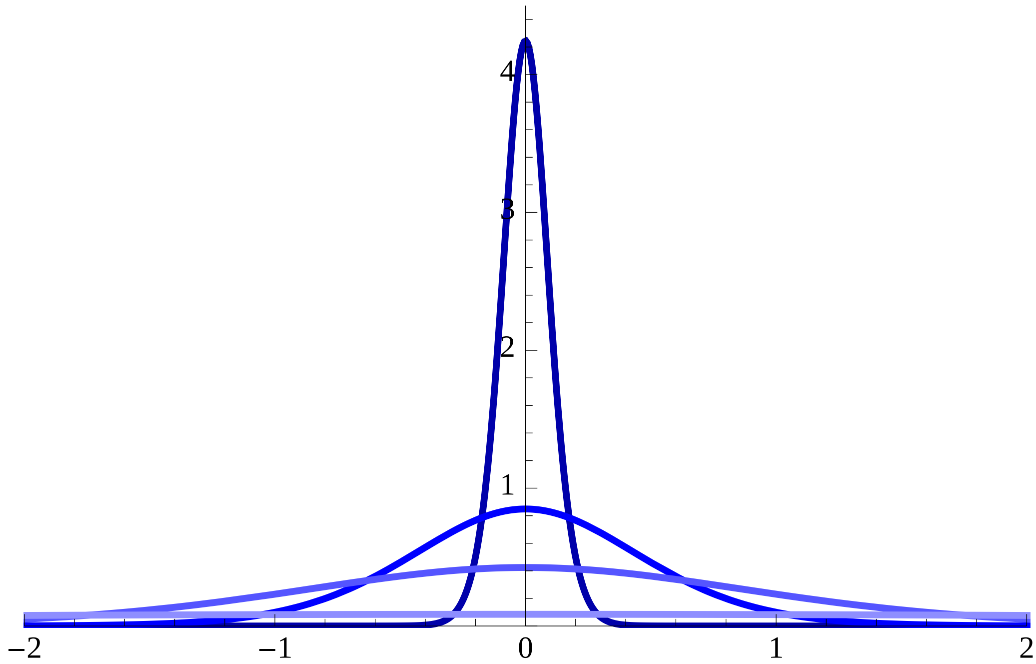


Figure 6.3: Meixner densities $MD(\alpha, 0, 2, 0)$, $\alpha = 0.1, 0.5, 1, 5$, from darker to lighter blue

A similar although "smoother" behavior induces the variation of parameter $\delta > 0$, as it is visible from the next image

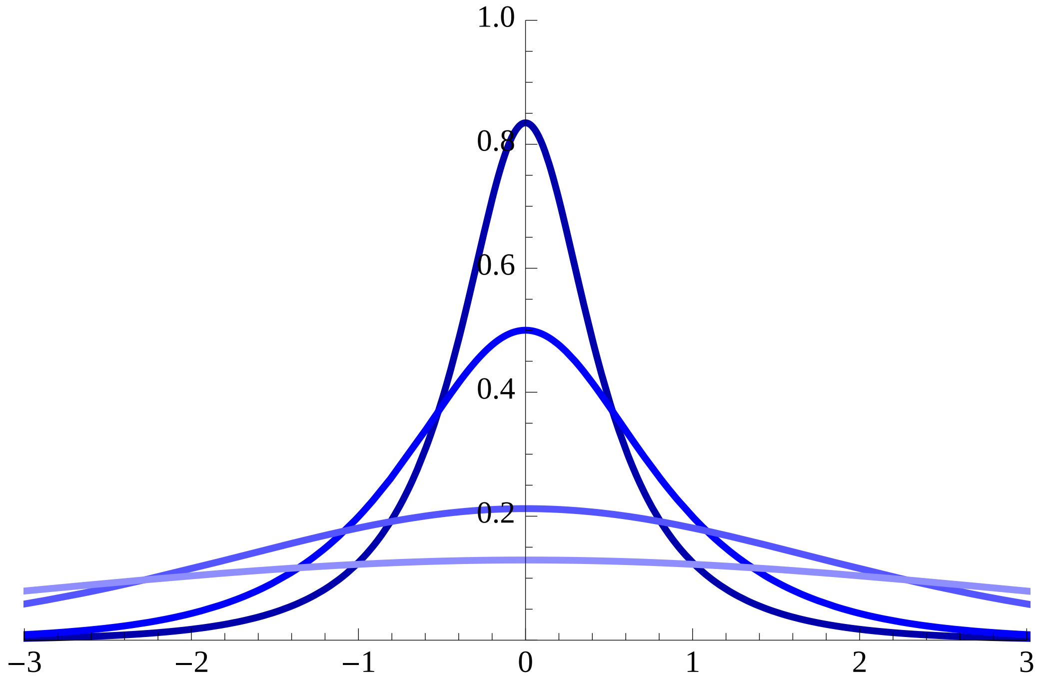


Figure 6.4: Meixner densities $MD(2, 0, \delta, 0)$, $\delta = 0.025, 0.5, 2, 5$, from darker to lighter blue

The real shape parameter in this case is $\beta \in (-\pi, \pi)$. First we observe that for $\beta \to \pm\pi$ brings to a non valid distribution function (a line flattening onto the $x$ axis); for values different from $\pm\pi$, the following behavior may be observed for Meixner distribution

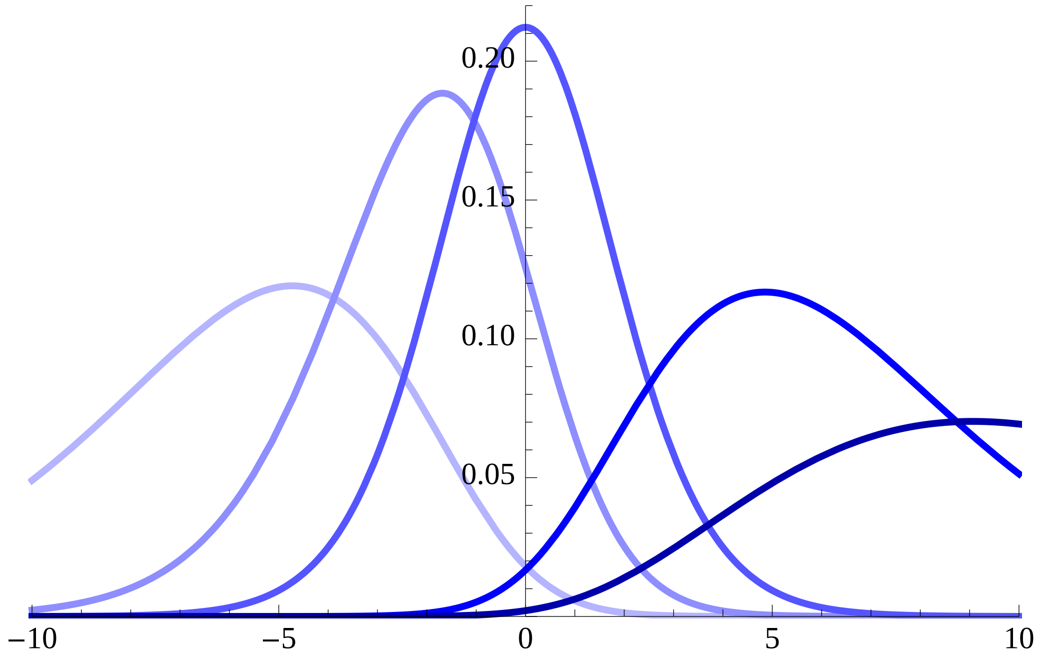


Figure 6.5: Meixner densities $MD(2, \beta, 2, 0)$, $\beta = -2, -1, 0, 2.025, 2.5$, from lighter to darker blue

While the variations of parameters $\alpha, \delta, \mu$ roughly reflect the correspondent variation of a Gaussian distribution in terms of variance and mean respectively, parameter $\beta$ is the main shape parameter of Meixner distribution, allowing an asymmetrical distribution.

As it can be seem from the comparison with a Gaussian distribution having the corresponding variance, Meixner distribution always maintains above the Gaussian curve, as firstly spotted above:

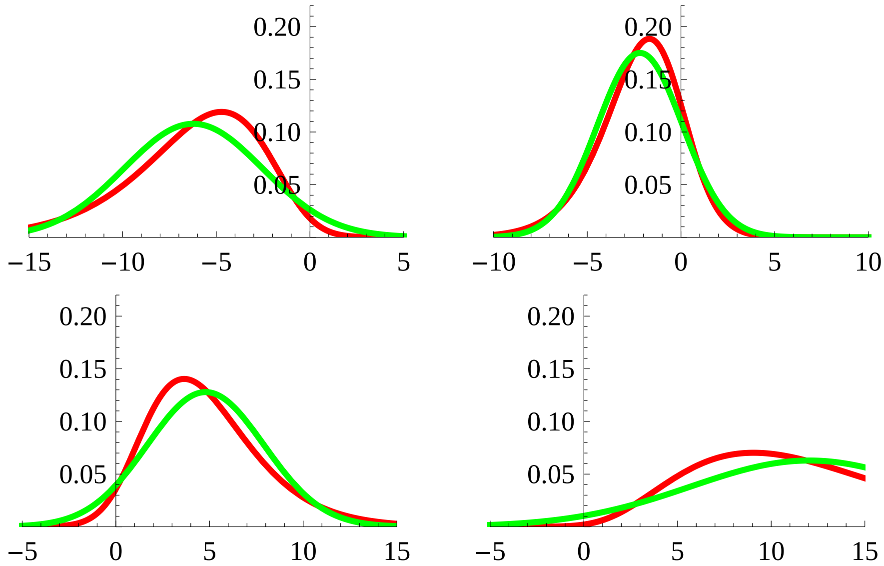


Figure 6.6: Comparison of Meixner distribution (*red*) with varying $\beta = -2, -1, 1.75, 2.5$ and corresponding Gaussian curve (*green*) with mean and variance varying accordingly

The relative stiffness of Gaussian distribution with respect to Meixner derives exacly from the lack of capability of growing in terms of variance, when needed in an asymmetrical way. This fact, combined with the action of the remaining parameters, affects heavily models based on Gaussian distribution with respect to models based on Meixner-like distributions (i.e. with 3 parameters or more).

## 6.2 Simulation of Meixner process

In this section we simulate some trajectories of Meixner process relying on the observations drawn in section 3.4.1 and in particular on equation (3.6); so the following simulation routine is suggested once again by the fundamental work by Madan and Yor [MY06].

To the knowledge of the author, there is no simulation of the trajectories of Meixner process in the examined literature.

Here follows the original R code to obtain the trajectories, always supposing $\mu = 0$:

```
#Meixner simulation from Madan-Yor "CGMY and Meixner
 Subordinators are absolutely continuous with respect to
 one sided stable subordinators"

q<-2000  #how many variables
a<-.25    #parameter a
b<-.002   #parameter b
delta<-2  #parameter delta
ep<-0.001
A<-b/a
C<-pi/a
U<-runif(q)
Y<-ep/U
W<-runif(q)
z<-a*delta*sqrt(2*ep/pi)
g<-NULL
for(n in -10:10){for(i in 1:length(Y)){
g[i]<-sum((-1)^n*exp(-n^2*pi^2/(2*C^2^Y[i])))*exp(-A^2*Y[i]/2)}

v<-NULL
for(j in 1:length(g)){
if (g[j]>W[j]) v[j]=1
else v[j]=0}
S<-sum(Y*v)

tau<-z+S
X<-A*tau+rnorm(q)*sqrt(tau)
```

```
T<-10
#s<-0.001
t<-seq(from=0, to=T, by=T/(q-1)) # time-grid
l<-length(t)
M<-NULL
for (m in 1:l){
M[m]<-cumsum(X)[m]}
plot(t, M, type='p',cex=0.35,pch=18,xlab="time",ylab="M(t)")
```

This leads to the following graphics, for different values of the parameters:

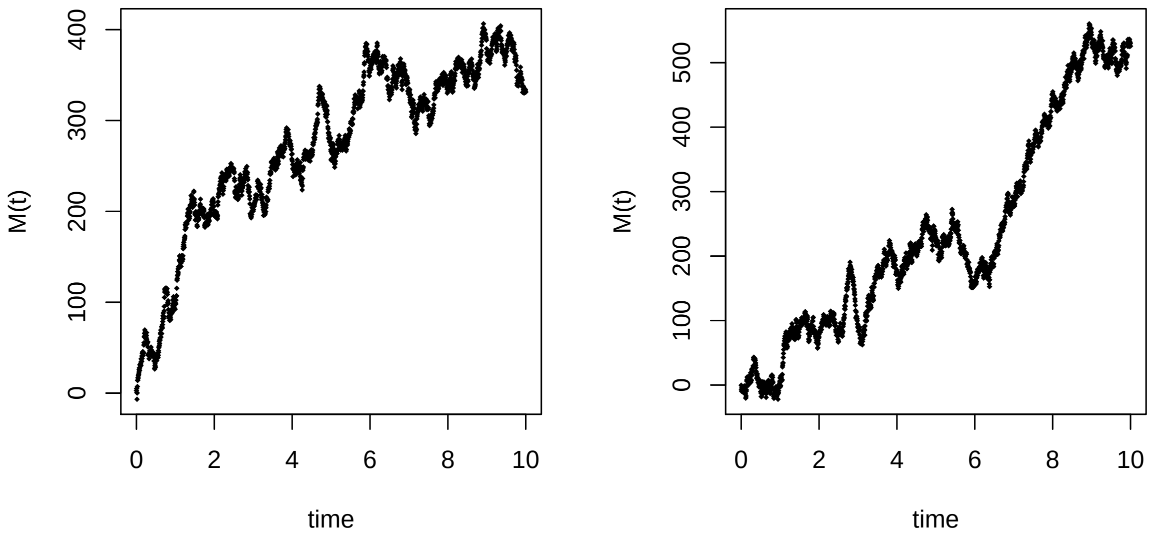


Figure 6.7: Possible trajectories of a Meixner process with the following parameter values for the triplet $(\alpha, \beta, \delta)$: $(0.25, 0.002, 0.2)$ (left), $(0.25, 0.002, 20)$ (right)

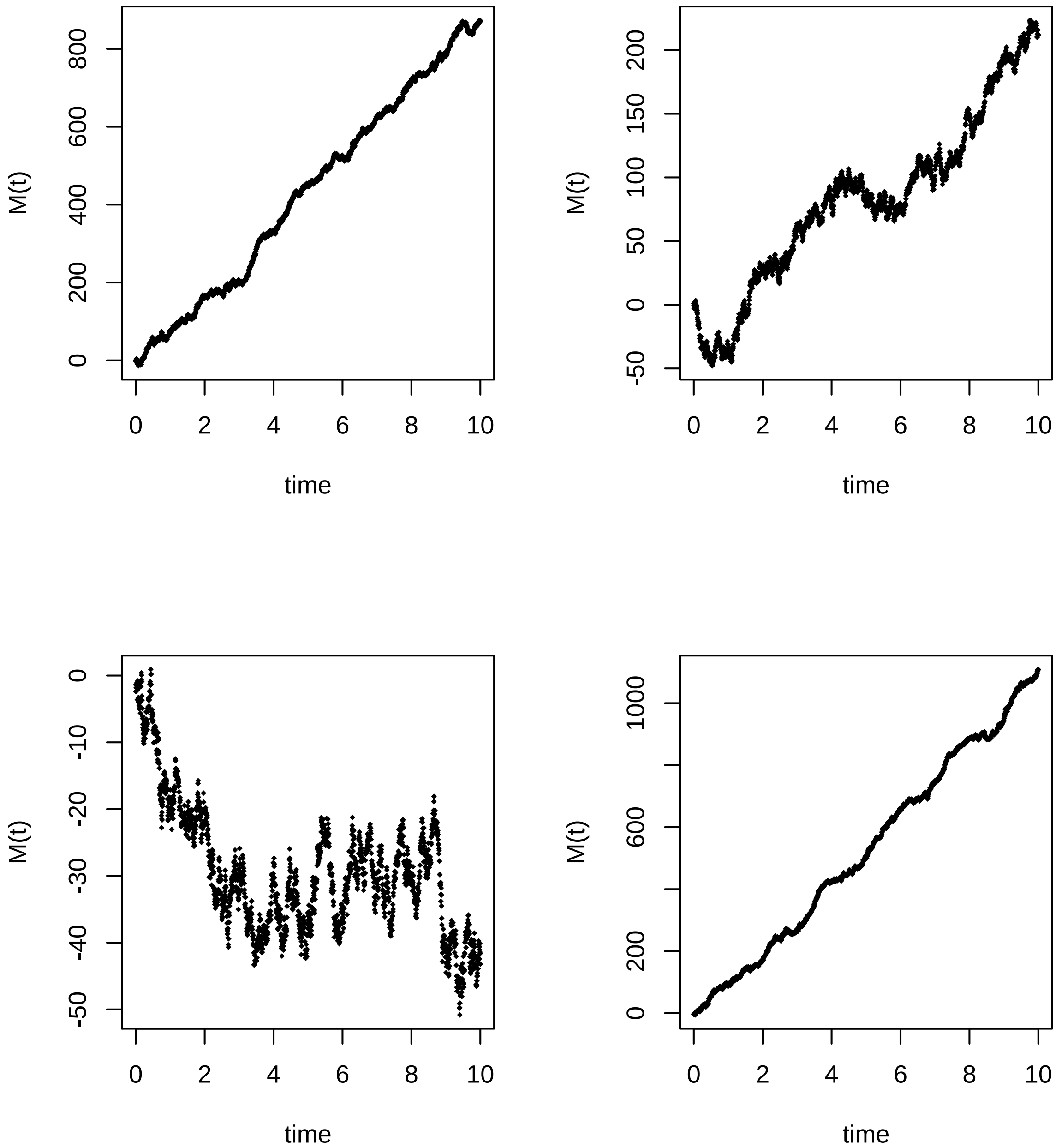


Figure 6.8: Possible trajectories of a Meixner process with the following parameter values for the triplet $(\alpha, \beta, \delta)$: $(0.25, 0.02, 2)$ (upper left), $(0.25, 0.002, 2)$ (upper right), $(25, 0.002, 2)$ (lower left), $(0.25, 0.002, 2)$ (lower right)

# 6.3 Modeling financial data via Meixner process

## 6.3.1 The stock price process

Our aim is to model the price process of an asset (a stock or an index) by a continuous-time process, often denoted with $S = \{S_t, t \geq 0\}$ as defined in (4.2), which gives asset price at time $t$. In order to allow the comparison of investments in different securities, it is usual to investigate the logarithmic returns (rates of return), defined by

$$X_t = \log S_t - \log S_{t-1}.$$

The main reason for this, is that the return over $n$ periods, for instance $n$ days, is clearly the simple sum

$$X_t + \ldots + X_{t+n-1} = \log S_{t+n-1} - \log S_{t-1};$$

another reason is that in most of the models the stock price $S_t$ will be modeled by an exponential of some stochastic process, and for continuous-time processes, returns with continuous compounding log returns are a natural choice.

It is well known that the standard continuous-time model for stock prices is the geometric Brownian motion,

$$S_t = S_0 \exp\left\{ \left( \mu - \frac{\sigma^2}{2} \right) t + \sigma W_t \right\}$$

which solves the stochastic differential equation

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

where $\sigma > 0$ and $\mu \in \mathbb{R}$ are coefficients denoting volatility and drift, and $\{W_t, t \geq 0\}$ is a standard Brownian motion. This model underlies the Black-Scholes formula.

This way the returns resulting from the gemoetric Brownian motion are increments of a Brownian motion process, and then independent and normally distributed; but a lot of literature (see for instance Cont [Con01] and related reference) show that for real data the hypotesis of normality generally fails.

An example could be given by the returns evaluated on the closing prices of the NASDAQ index for year 2009, giving back the following histogram, along with a first raw Gaussian estimation

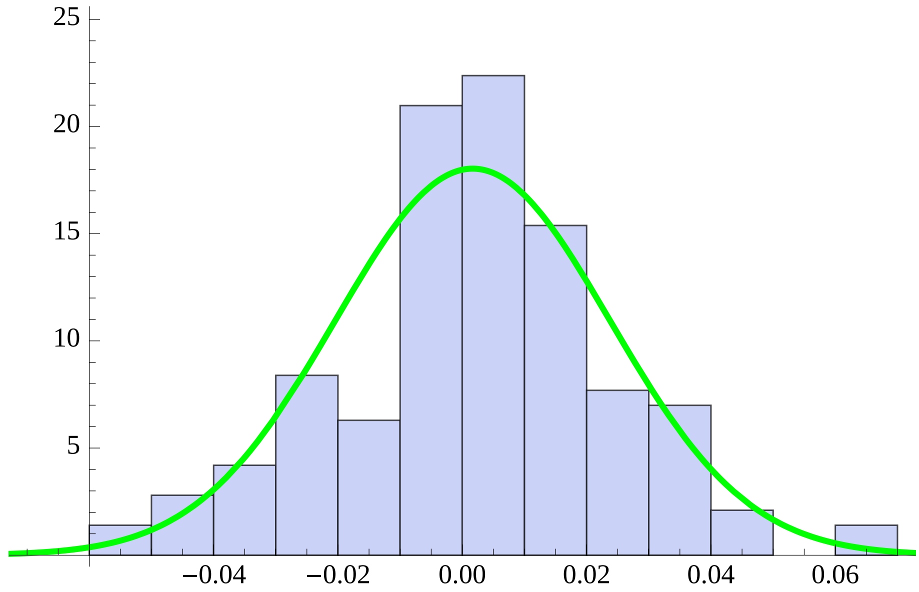


Figure 6.9: log returns of closures of NASDAQ index from 02.01.2009 to 29.07.2009 with a moment-estimated Gaussian model

The situation deteriorates if we consider the return over a wider period of time, starting from 03.05.1990, evaluating 4832 data:

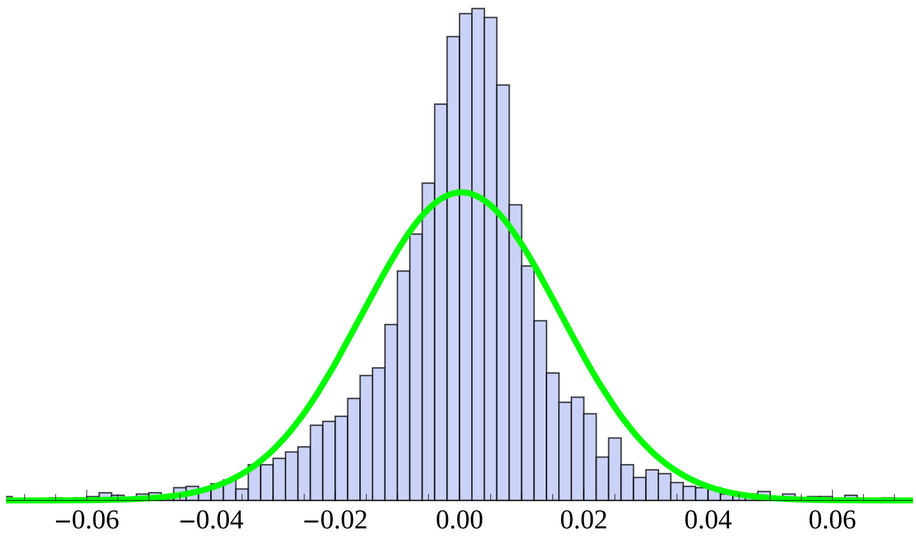


Figure 6.10: log returns of closures of NASDAQ index from 03.05.1990 to 29.07.2009 with a moment-estimated Gaussian model

A clear concentration of mass around the maximum of the empirical distribution emerges from the figures above, along with a consequent "thinning of tails", intuitively speaking. This combined facts induce to look for a model which can provide these macroscopical features, for instance Meixner distribution. In fact estimating the parameters via method of moments, the following results can be achieved:

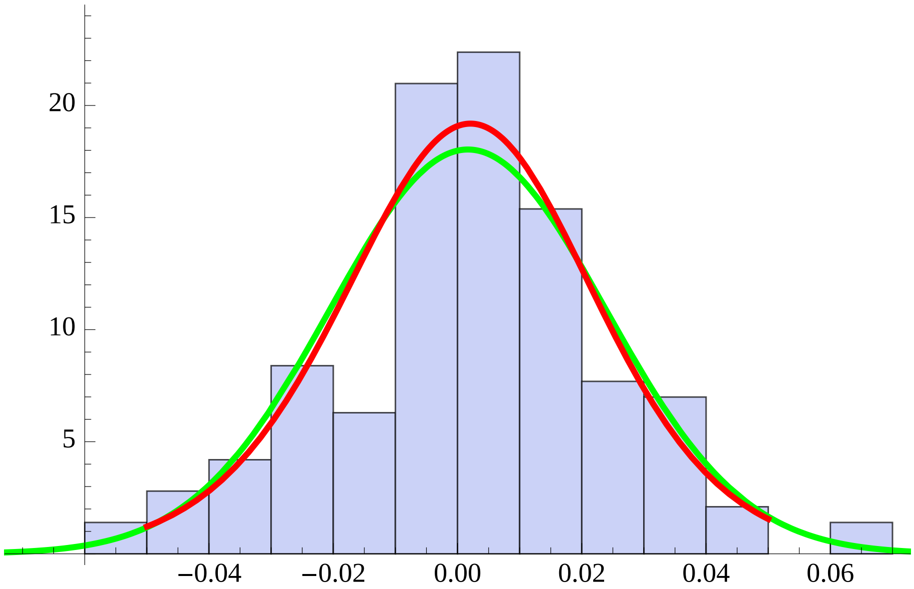


Figure 6.11: Meixner distribution added with method of moments estimated parameters: $\alpha = 0.0221138$, $\beta = -0.0834728$, $\delta = 1.99758$, $\mu = 0.0035104$

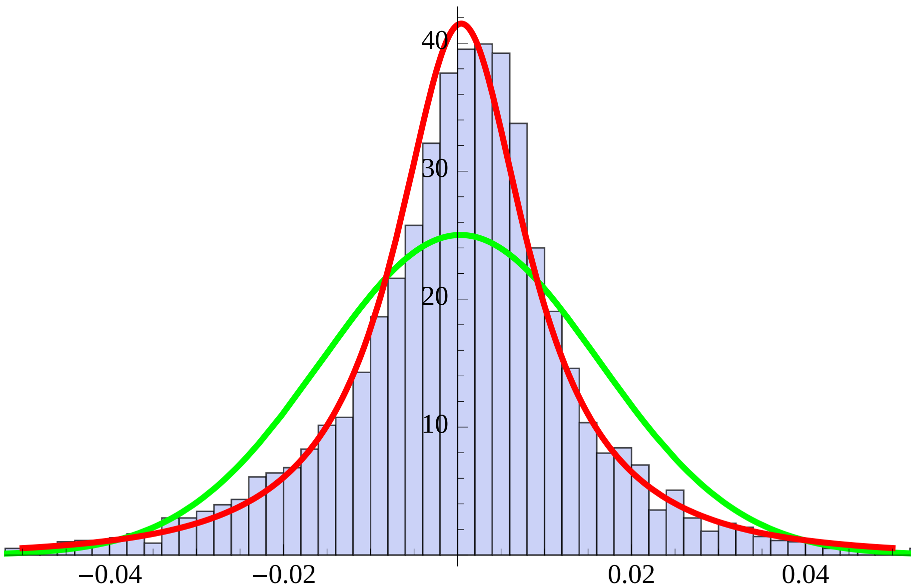


Figure 6.12: Meixner distribution added with method of moments estimated parameters: $\alpha = 0.0543601$, $\beta = -0.0299293$, $\delta = 0.172123$, $\mu = 0.000456984$

It is necessary to point out that, due to the results shown in chapter 2, for which moments estimators for many of the listed processes are available, make this estimators easier to evaluate than the maximum likelihood estimators, for instance, which are not always available in a closed form. In any case, for instance both Eberlein and Keller [EK95] and Schoutens [Sch03] make use of ML estimators obtaining basically the same results.

The same results can be obtained for stock prices.

As it can be seen, a very good fitting is also obtained with the NIG distribution (2.5), meaning that the price process follows a NIG process:

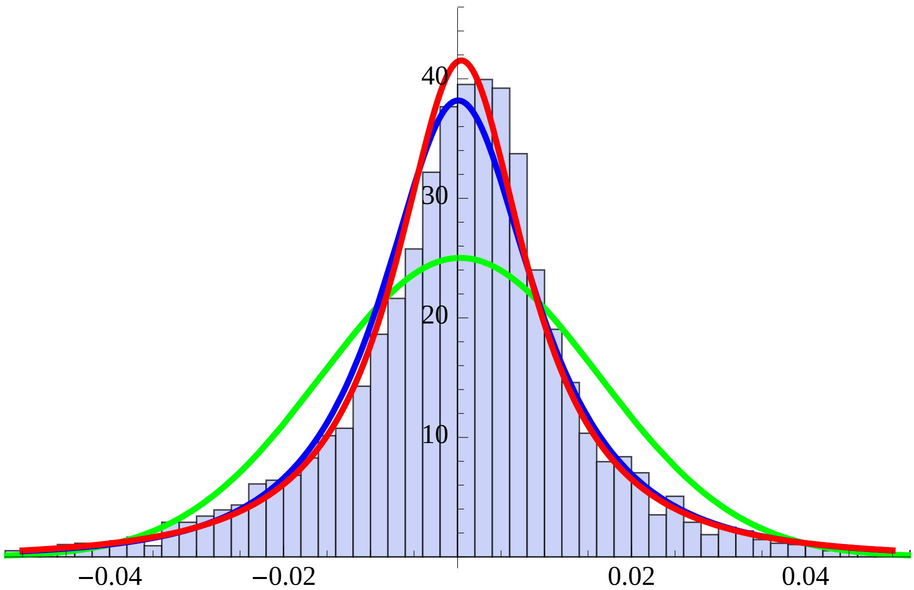


Figure 6.13: NIG distribution (blue line) added with method of moments estimated parameters: $\alpha = 45.1484$, $\beta = 1.24699$, $\delta = 0.011471$

### 6.3.2 Model Diagnostic

A simple method to assess goodness of fit for these models, is to rely upon distances between the involved distributions. A good way to perform it, that needs anyway a certain amount of programming and numerical integration, is given by two distances: the usual Kolmogorov-Smirnov distance, and the modified Anderson-Darling distance.

The *Kolmogorov-Smirnov* test, is based upon the following distance among cumulative distributions:

$$D_n = \sup_{x \in \mathbb{R}} |F_n(x) - F(x)|,$$

where $F_n(x)$ is the empirical distribution function obtained from the data, and $F(x)$ is the fitting distribution. It is also known that $\sqrt{n}D_n$ follows a Kolmogorov distribution, which is tabulated. In financial applications, though, the tails of a return distribution might be of interest, for they often happen to be deciding factors in protfolio strategies. Assuming a given probability density with distribution $F(x)$ and measuring the distance between $F$ and the empirical distribution by the Kolmogorov-Smirnov distance, it turns out that the same importance is given to tha tails as to the center of the distribution. The reason is that the Kolmogorov-Smirnov distance measures the uniform distance between the two functions, i.e. the maximum deviation regardless where it occurs.

An alternative way is provided by the following empirical variant of the *Anderson-Darling* statistic (AD), given by

$$AD = \sup_{x \in \mathbb{R}} \frac{|F_n(x) - F(x)|}{\sqrt{F(x)(1 - F(x))}};$$

in this case, the distance is rescaled by dividing the distance through the "standard deviation" of this distance, given by the denominator of the above formula. It can be observed that the denominator becomes small for very large and very small values of data: thus the same absolute deviation between $F$ and $F_n$ in the tails gets a higher weight as it occurs in the center of the distributions. A possible drawback of this choice is that the distribution of the statistic depends on the choiche of $F$ and consequently tests about the velidity of the goodness of fit assumption cannot be performed as easily as the with the Kolmogorov-Smirnov distance.

In our case, the values of distance will be enough to show the better performance of Meixner distribution in terms of goodness of fit to the data. Corresponding tests may be performed easily. The following figure shows the estimates of the distribution functions

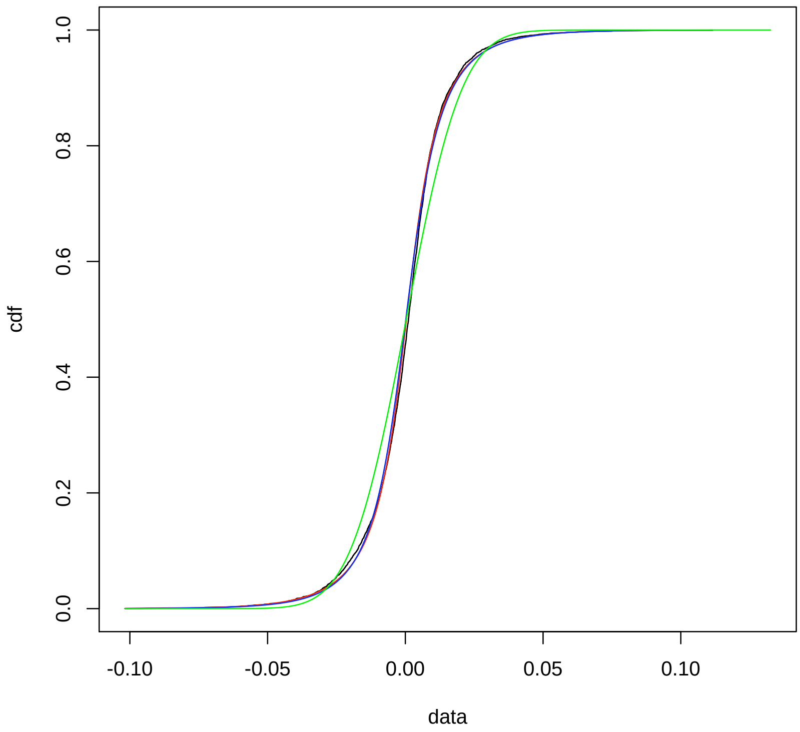


Figure 6.14: Empirical distribution function (black) with cumulative distribution functions from Meixner density (red) and NIG density (blue) obtained by numerical integration. Normal distribution function is the green line

It is obviously not clear due to the nearness of the estimates of the two Lévy process driving distribution to the original data. Here are the values of the two distances above for the examined dataset:

| NASDAQ data | KS | AD |
|---|---|---|
| Meixner | $0.0327$ | $0.0655$ |
| NIG | $0.0410$ | $0.0410$ |
| Gaussian | $0.9998$ | $0.2242$ |

The better performance of the two distribution driving the correspondend geometric Lévy model (4.2) is very clear compared to the Gaussian distribution and hence to the Brownian motion; a substantial equivalence holds for the fitting in terms of Anderson-Darling distance.

## 6.4 Conclusion

The literature about Meixner process is not very populated.

After its introduction in 1999, some papers have appeared from 2002 to 2009 dealing mainly with mathematical properties of generalized Lévy processes , i.e. for instance generalization of Chaotic Representation Property for Lévy processes , extensions of Meixner processes to "Meixner-type" processes having the parameters of the characteristic exponent state space dependent via theory of pseudo-differential operators, study of properties from the point of view of more general $q-$Lévy processes , and the chance of writing infinitesimal generator of the process through suitable integral operators. They don't deal with a simpler statistical approach.

It is also sufficiently clear that departure from classical mathematical models for finance governed by Brownian motion and Gaussian distribution is not totally painless. Theoretical properties and conditions are mostly well known, but one important problem is given by computational effort.

Nonetheless our main goal was to try to focus the interest on a process that undoubtly gives better performances when considered as a mathematical financial model. Another point is that this claim has been shown not making use of ad-hoc software and creating from scratch all the routines that in literature are often just suggested or hinted.

So one of the goals was to show that it is possible to apply this process in modeling financial data with not a great effort in terms of programming.

Despite that, many open theoretical problems remain, which can be the subject of further research, the first one being the evaluation of equivalent martingale measures different from Esscher's. Some calculations have been performed during this work, but have not given any result at the moment. Other obscure points are given by minimal martingale measure and optimal variance martingale measure.

Another interesting issue for developing research is given by the study of subordinated Meixner process, which is missing to our knowledge from analyzed literature.

Referring to a work by Aurzada and Dereich, [AD09], a small deviations problem for Meixner process can be investigated.

From an applicative point of view, it has turned out that some seismological graphics happen to be very similar to volalility clusters graphics shown for instance by Schoutens in [Sch02], Fig.6; namely, some background noise recorded on a daily basis by seismomenters and having causes depending on human activity and on natural phenomena not strictly of a seismological origin, has such a representation. It can be interesting to investigate the chance of fitting these data with Meixner-SV models as the one introduced buy Schoutens.

A possible furter employment of Meixner distribution can be spotted as a model for temperature data in a bayesian hyerarchical approach for jointly describing temperature and precipitation chianges in multiple climate model as described by Tebaldi and Sansó in [TS09].

The main impression is that in this difficult topic many results are added by very little pieces. Still by 2006 no one has been able to express Meixner process as a subordinated Brownian motion, which opens the chance to simulate the process as we have seen. Nonetheless the study of simulated trajectories as the ones we have introduced here is still missing.

The importance of these models is clear, providing a real flexible and fitting instrument mainly for financial applications; similar models such as hyperbolic models mostly, have been employed in theoretical quantum physics and in modeling natural phenomena as turbulence or sand deposits.

We have also shown how these kind of models could provide a sort of aperture towards different fields of mathematics, involving statistics indirectly, as the theory of differential equations and orthogonal polynomials.

This gives the models a sort of mathematical reliability, descending from Meixner's cited 1934 work, which was known and settled mainly for Brownian motion only up to that moment.

# Bibliography


[AD09] F. Aurzada and S. Dereich. Small deviations of general Lévy processes. *Ann.Prob*, 37(5):2066–2092, 2009.

[App04] D. Appelbaum. *Lévy Processes and Stochastic Calculus*. Cambridge Studies in Advanced Mathematics, 2004.

[AS64] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York, 9th dover printing, tenth GPO printing edition, 1964.

[Ber96] J. Bertoin. *Lévy Processes*. Number 121 in Cambridge Tracts in Mathematics. Cambridge University Press, 1996.

[BF02] F. Bellini and M. Frittelli. On the existence of minimax martingale measures. *Math.Finance*, 12(1):1–21, 2002.

[BLBL92] S.K. Bar-Lev, D. Bshouty, and G. Létac. Natural exponential families and self decomposability. *Stat.Prob.Lett.*, 13:147–152, 1992.

[BN78] O.E. Barndorff-Nielsen. Hyperbolic distributions and distributions on hyperbolae. *Scand.J.Statist*, 5:151–157, 1978.

[BN97] O.E. Barndorff-Nielsen. Normal Inverse Gaussian distributions and stochastic volatility modelling. *Scand.J.Statist*, 24:1–13, 1997.

[BN98] O.E. Barndorff-Nielsen. Processes of Normal Inverse Gaussian type. *Fin.Stoch.*, 2(1):41–68, 1998.

[BNH77] O.E. Barndorff-Nielsen and C. Halgreen. Infinite divisibility of the hyperbolic and generalized inverse gaussian distributions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 38:309–311, 1977.

[CGMY02] P. Carr, H. Geman, D.P. Madan, and M. Yor. The fine structure of asset returns: An empirical investigation. *Journal of Business*, 72(2):305–332, 2002.

[Con01] R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236, 2001.

[CT03] R. Cont and P. Tankov. *Financial modelling with jump processes*. CRC Financial Mathematics Series. Chapman & Hall, 2003.

[Dom07] D. Dominici. Fisher information of orthogonal polynomials I. arXiv:math/0708.2698v1, 2007.

[DS94] F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300:463–520, 1994.

[Ebe09] E. Eberlein. Jump-type Lévy processes. In *Handbook of Financial Time Series*, pages 439–455. Springer Verlag, 2009.

[EK95] E. Eberlein and U. Keller. Hyperbolic distributions in finance. *Bernoulli*, 1(3):281–99, 1995.

[ES05] F. Esche and M. Schweizer. Minimal entrop preserves the Lévy property: how and why. *Stoch.Proc.Appl*, 115(2):299–327, 2005.

[Ess32] F. Esscher. On the probability function in the collective theory of risk. *Skandinavisk Aktuarietidskrift*, 15:175–195, 1932.

[Fio03] F. Fiorani. *Option Pricing under the Variance Gamma Process*. PhD thesis, Università degli Studi di Trieste, 2003.

[FM03] T. Fujiwara and Y. Miyahara. The minimal entropy martingale measures for geometric Lévy processes. *Finance and Stochastics*, 7:509–531, 2003.

[FS91] H. Föllmer and M. Schweizer. Hedging contingent claims under incomplete information. In *Applied Stochastic Analysis*, pages 389–414. Gordon & Breach, 1991.

[GP09] M. Grigoletto and C. Provasi. Simulation and estimation of the Meixner distribution. *Communications in Statistics: Simulation and Computation*, 38(1):58–77, 2009.

[GR94] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, New York, corrected and enlarged edition, 1994.

[GR01] T. Göll and L. Ruschendorf. Minimax and minimal distance martingale measures and their relationship to portfolio optimization. *Finance and Stochastics*, 5(4):557–581, 2001.

[Gri99] B. Grigelionis. Processes of Meixner type. *Lithuanian Mathematical Journal*, 39(1):33–41, 1999.

[Gri01] B. Grigelionis. Generalized $z$ distributions and related stochastic processes. *Lithuanian Mathematical Journal*, 41(3):239–251, 2001.

[GS94] H.U. Gerber and E.S.W. Shiu. Option pricing by Esscher transforms. *Trans.Soc.Act.*, 46:99–191, 1994.

[HM03] A. Hirsa and D.B. Madan. Pricing american options under variance gamma, 2003.

[HS06] F. Hubalek and C. Sgarra. On the Esscher transform and entropy for exponential Lévy models. *Quantitative Finance*, 6(2):125–145, 2006.

[JPY02] M. Jeanblanc, J. Pitman, and M. Yor. Self-similar processes with independent increments associated with Lévy and Bessel processes. *Stoch.Proc.Appl.*, 100:223–231, 2002.

[JS02] J. Jacod and A.N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer, New York, 2002.

[Khi37a] A.Y. Khinchine. A new derivation of a formula by P.Lévy. *Bulletin of the Moscow State University*, 1:1–5, 1937.

[Khi37b] A.Y. Khinchine. On arithmetic of distributions. *Bulletin of the Moscow State University*, 1:6–17, 1937.

[Kou02] S.G. Kou. A jump diffusion model for option pricing. *Management Science*, 48:1086–1101, 2002.

[KS02] J. Kallsen and A.N. Shiryaev. The cumulant process and Esscher's change of measure. *Finance and Stochastics*, 6:397–428, 2002.

[Kyp07] A.E. Kyprianou. An introduction to the theory of Lévy processes, August 2007. Satellite Summer School on Lévy Processes: Theory and Applications, Sønderborg, Denmark.

[Luk70] E. Lukacs. *Characteristic functions*. Ch.Griffin, London, 2nd edition, 1970.

[MC98] P.P. Madan, D.B. andCarr and E.C. Chang. The Variance Gamma process and option pricing. *European Finance Review*, 2:79–105, 1998.

[Mei34] J. Meixner. Orthogonale polynomsysteme mit einer besonderen gestalt der erzeugende funktion. *J.London Math.Soc.*, 9:6–13, 1934.

[Mer76] R.C. Merton. Option pricing when underlying stock returns are discontinuous. *J.Fin.Econ.*, 3:125–74, 1976.

[Miy04] Y. Miyahara. A note on Esscher transformed martingale measures for geometric Lévy processes. Technical report, Nagoya City University, 2004.

[Miy05] Y. Miyahara. Martingale measures for the geometric Lévy process models. Technical report, Nagoya City University, 2005.

[Mor82] C.N. Morris. Natural exponential families with quadratic variance functions. *Ann.Stat.*, 10(1):65–80, 1982.

[MR06] F. Mainardi and S. Rogosin. The origin of infinitely divisible distributions: from de Finetti's problem to Lévy-khinchine formula, 2006.

[MS90] D.B. Madan and E. Seneta. The variance gamma (V.G.) model for share market returns. *Journal of Business*, 63(4):511–24, 1990.

[MY06] D.B. Madan and M. Yor. CGMY and Meixner subordinators are absolutely continuous with respect to one sided stable subordinators. arXiv:math/0601173v2, 2006.

[Pap05] A. Papapantoleon. An introduction to Lévy processes with applications in finance, 2005. Lecture notes.

[Pre75] R.L. Prentice. Discrimination among some parametric models. *Biometrika*, 62:607–614, 1975.

[PY03] J. Pitman and M. Yor. Infinitely divisible laws associated with hyperbolic functions. *Can.J.Math.*, 55(2):292–330, 2003.

[Sam07] G. Samorodnitsky. Lévy processes and their applications. Bocconi University Summer School, Torgnon, Italy, July 2007.

[Sat99] K.I. Sato. *Lévy processes and infinitely divisible distributions*. Cambridge University Press, 1999.

[Sch00] W. Schoutens. *Stochastic Processes and Orthogonal Polynomials*. Lecture Notes in Statistics 146. Springer-Verlag, New York, 2000.

[Sch02] W. Schoutens. The Meixner process: Theory and applications in finance. Technical report, K.U.Leuven, February 2002.

[Sch03] W. Schoutens. *Lévy Processes in Finance - Pricing Financial Derivatives*. Wiley series in probability and statistics. John Wiley & Sons, 2003.

[SRD05] J. Sánchez-Ruiz and J.S. Dehesa. Fisher information for orthogonal hypergeometric polynomials. *J. Comput. Appl. Math.*, 182(1):150–164, 2005.

[ST98] W. Schoutens and J.L. Teugels. Lévy processes; polynomials and martingales. *Commun.Statist.-Stochastic Models*, 14(1,2):335–349, 1998.

[TS09] C. Tebaldi and B. Sansó. Joint projections of temperature and precipitation change from multiple climate models: a hierarchical bayesian approach. *J.R.Statist.Soc.A*, 172:83–106, 2009.