Basics of Lévy processes*

Ole E. Barndorff-Nielsen
The T.N. Thiele Centre for Mathematics in Natural Science,
Department of Mathematical Sciences,
University of Aarhus, Ny Munkegade, DK-8000 Aarhus C, Denmark
& CREATES, University of Aarhus
oebn@imf.au.dk

Neil Shephard
Nuffield College, New Road, Oxford OX1 1NF, UK,
Department of Economics, University of Oxford
& Oxford-Man Institute, University of Oxford,
neil.shephard@economics.ox.ac.uk

June 9, 2012

Abstract

This is a draft Chapter from a book by the authors on "Lévy Driven Volatility Models".

Contents

*We are grateful to the many colleagues and students who have over many years, given us comments on this chapter. Of course we welcome further comments as we push it to completion.

Keywords: compound process; cumulant function; inverse Gaussian; jumps; Lévy-Khintchine representation; Lévy measure; Lévy process; Poisson process; quadratic variation; Skellam process; stable process; time deformation.

JEL Classifications: C22, C63, G10

1 What is this Chapter about?

In this Chapter we provide a first course on Lévy processes in the context of financial economics. The focus will be on probabilistic and econometric issues; understanding the models and assessing their fit to returns on speculative assets. The Chapter will delay the discussion of some of the technical aspects of this material until Barndorff-Nielsen and Shephard (2012a). Our stochastic analysis primer Barndorff-Nielsen and Shephard (2012b) may be of help to readers without a strong background in probability theory. Throughout we hope our treatment will be as self-contained as possible.

This long Chapter has 12 other Sections, whose goals are to:

- Build continuous time random walk models using the concept of infinite divisibility.
- Introduce Lévy processes with non-negative increments.
- Extend the analysis to Lévy processes with real increments.
- Introduce time change, where we replace calendar time by a random clock.
- Use statistical methods to fit some common Lévy processes to financial data.
- Introduce quadratic variation, a central concept in econometrics and stochastic analysis.

- Briefly discuss stochastic analysis in the context of Lévy processes.
- Introduce various methods for building multivariate Lévy processes.
- Discuss the relationship between stochastic volatility and general semimartingales.
- Draw conclusions to the Chapter.
- Provide some exercises.
- Discuss the literature associated with Lévy processes.

Lévy processes can only provide a rather partial description of asset prices, for they have independent increments and so ignore volatility clustering. However, later developments in this book will extend the setting to deal with that issue.

2 What is a Lévy process?

2.1 The random walk

The most basic model of the logarithm of the price of a risky asset is a discrete time random walk. It is built by summing independent and identically distributed (i.i.d.) random variables C_1, C_2, ... to deliver

$$Y_t = \sum_{j=1}^{t} C_j, \quad \text{with} \quad Y_0 = 0, \quad t = 0, 1, 2, \ldots .$$

The process is moved by the i.i.d. increments

$$Y_{t+1} - Y_t = C_{t+1}. \tag{1}$$

Hence future changes in a random walk are unpredictable.

What is the natural continuous time version of this process? There are at least two strong answers to this question.

2.2 Brownian motion

The first approach is based on a central limit type result. Again suppose that C is an i.i.d. sequence whose first two moments exist. Then, for a given $n > 0$, define the partial sum

$$Y_t^{(n)} = \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor tn \rfloor} \left\{ C_j - \mathrm{E}(C_1) \right\}, \qquad t \in \mathbb{R}_{\geq 0}, \tag{2}$$

where t represents time and $\lfloor a \rfloor$ denotes the integer part of a. It means that over any fixed interval for t the process is made up of centred and normalised sums of i.i.d. events. We then allow n, the number of these events in any fixed interval of time of unit length, to go off to infinity. This is often labelled "in-fill" asymptotics. As a result $Y_t^{(n)}$ obeys a central limit theory and becomes Gaussian. Further, this idea can be extended to show that the whole partial sum, as a random function, converges to a scaled version of Brownian motion, as n goes to infinity. At first sight this suggests the only reasonable continuous time version of a random walk, which will sum up many small events, is Brownian motion. This insight is, however, incorrect.

2.3 Infinite divisibility

Our book follows a second approach. Suppose that the goal is to design a continuous time process Y_t such that Y_1, its value at time 1, has a given distribution D. We can divide the time from 0 until 1 into n pieces of equal length. The corresponding increments are assumed to be independent from a common distribution $D^{(n)}$ such that the sum

$$Y_t^{(n)} = \sum_{j=1}^{\lfloor tn \rfloor} C_j^{(n)}, \qquad \text{where} \qquad C_j^{(n)} \overset{i.i.d.}{\sim} D^{(n)},$$

has the distribution D when $t = 1$. Then as n increases the division of time between zero and one becomes ever finer. In response, the increments and their distribution $D^{(n)}$ also change, but by construction D, the distribution of the sum, is left unchanged. A simple example of this is where $Y_1 \sim Po(1)$, a Poisson random variable with mean 1, then if

$$Y_t^{(n)} = \sum_{j=1}^{\lfloor tn \rfloor} C_j^{(n)}, \qquad \text{where} \qquad C_j^{(n)} \overset{i.i.d.}{\sim} Po(1/n),$$

this produces a random walk with independent Poisson increments which sum to a Poisson random variable. Hence this process makes sense even as n goes to infinity and so this type of construction can be used to introduce a continuous time model — the *Poisson process*. A second example is where

$$Y_t^{(n)} = \sum_{j=1}^{\lfloor tn \rfloor} C_j^{(n)}, \qquad \text{where} \qquad C_j^{(n)} \overset{i.i.d.}{\sim} N(0, 1/n),$$

then as $n \to \infty$ this process converges to Brownian motion.

The class of distributions for which this general construction is possible is that for which D is *infinitely divisible*. The resulting processes are called Lévy processes. Examples of infinitely divisible distributions include, focusing for the moment only on non-negative random variables, the Poisson, gamma, reciprocal gamma, inverse Gaussian, reciprocal inverse Gaussian, F and positive stable distributions. A detailed technical discussion of infinite divisibility will be given in Barndorff-Nielsen and Shephard (2012a).

2.4 Lévy processes and semimartingales

The natural continuous time version of the discrete time increment given in (1) is, for any value of $s > 0$,

$$Y_{t+s} - Y_t, \qquad t \in [0, \infty) .$$

Increments play a crucial role in the formal definition of a Lévy process.

Definition 1 *Lévy process. A càdlàg stochastic process $Y = \{Y_t\}_{t\geq 0}$ with $Y_0 = 0$ is a Lévy process if and only if it has independent and (strictly) stationary increments.*

The càdlàg assumption is an important one which is discussed in detail below. For the moment let us focus on the independence and stationarity assumption. This means that the shocks to the process are independent over time and that they are summed, while the stationarity assumption specifies that the distribution of $Y_{t+s} - Y_t$ may change with s but does not depend upon t. The independence and stationarity of the increments of the Lévy process implies that the cumulant function of Y_t

$$\begin{aligned} \mathrm{C}\{\theta \ddagger Y_t\} &= \log\left[\mathrm{E}\exp\{i\theta Y_t\}\right] \\ &= t\log\left[\mathrm{E}\exp\{i\theta Y_1\}\right] \\ &= t\mathrm{C}\{\theta \ddagger Y_1\}, \end{aligned}$$

so the distribution of Y_t is completely governed by the cumulant function of Y_1, the value of the process at time one.

If a Lévy process is used as a model for the log-price of an underlying asset then the increments can be thought of as returns. Consequently Lévy based models provide a potentially flexible framework with which to model the marginal distribution of returns. However, the returns will be independent and identically distributed when measured over time intervals of fixed length. Thus important serial dependencies such as volatility clustering are ignored.

A feature of some stochastic processes, including Brownian motion, is that they have continuous sample paths with probability one. We need to relax this assumption. We will allow jumps but require the process Y to be, with probability one, right continuous

$$\lim_{s\downarrow t} Y_s = Y_t$$

and have limits from the left

$$Y_{t-} = \lim_{s\uparrow t} Y_s.$$

For such processes the jump at time t is written as

$$\Delta Y_t = Y_t - Y_{t-},$$

and the processes are said to be càdlàg (continu à droite, limites à gauche). We might, in particular, expect these types of jumps to appear in financial economics due to dividend payments, microcrashes due to short-term liquidity challenges or news, such as macroeconomic announcements. The similarly named property càglàd (continu à gauche, limites à droite), which has

$$\lim_{s \downarrow t} Y_s = Y_{t+} \qquad \text{and} \qquad Y_t = \lim_{s \uparrow t} Y_s,$$

plays an important role in our stochastic analysis primer.

Semimartingales have a central role in modern stochastic analysis. They are càdlàg (by definition) and provide, in particular, a basis for the definition of a stochastic integral. Consequently it is important to note that all Lévy processes Y are semimartingales (written $Y \in \mathcal{SM}$). They also have the property that

$$\Pr(\Delta Y_t > 0) = \Pr(\Delta Y_t < 0) = 0,$$

i.e. there are no fixed discontinuities.

For $Y \in \mathcal{SM}$, if H is càdlàg then H_- is locally bounded and the *stochastic integral*

$$X_t = \int_0^t H_{u-} \mathrm{d}Y_u,$$

is well defined. This is often written in the more abstract notation as the process

$$X = H \bullet Y.$$

More details of stochastic integrals are given in Barndorff-Nielsen and Shephard (2012b) and also in Section 8.

3 Non-negative Lévy processes — subordinators

We start our more detailed discussion of Lévy processes by considering Lévy processes with non-negative increments. Such processes are often called *subordinators* — the reason for this nomenclature will become clear in Section 5. This is our focus for two reasons: (i) they are mathematically considerably simpler, (ii) many of the models we build in this book will have components which are Lévy processes with non-negative increments. The discussion of processes with arbitrary real increments will be given in the next Section. As the processes are positive, it is natural to work with the *kumulant function* in the form

$$\overline{\mathrm{K}} \{\theta \ddagger Y_1\} = \log \{\mathrm{E} \exp(-\theta Y_1)\}, \quad \text{where} \quad \theta \geq 0.$$

Occasionally the more standard cumulant function (or log Laplace transform)

$$\mathrm{K}\{\theta \ddagger Y_1\} = \log\{\mathrm{E}\exp(\theta Y_1)\},$$

where $\theta \in \{\phi : \mathrm{E}\exp(\phi Y_1) < \infty\}$, will be used.

3.1 Examples

3.1.1 Poisson process

Suppose we count the number of events which have occurred from time 0 until $t \geq 0$. The very simplest continuous time model for this type of observation is a Lévy process with independent Poisson increments, so that

$$Y_1 \sim Po(\psi), \quad \psi > 0,$$

with density

$$f_{Y_1}(y) = \frac{e^{-\psi}\psi^y}{y!}, \qquad y = 0, 1, 2, \ldots.$$

The process Y is called a (homogeneous) *Poisson process* and the mean of Y_1, ψ, is often called the intensity of this counting process. A simulated sample path of this process, when $\psi = 1$, is given in Figure 1(a). It shows a jumping process, where the jumps are irregularly spaced in time and are of equal size. The times at which events occur are called *arrival times*, and are written as $\tau_1, \tau_2, \ldots, \tau_{Y_t}$.

For the Poisson process

$$\begin{aligned} \overline{\mathrm{K}}\{\theta \ddagger Y_1\} &= \log[\mathrm{E}\exp\{-\theta Y_1\}] \\ &= \psi(e^{-\theta} - 1) \\ &= -\psi(1 - e^{-\theta}). \end{aligned}$$

Now we have that $-t\psi(1-e^{-\theta})$, the kumulant function of Y_t, corresponds to the kumulant function for a $Po(t\psi)$.

The Poisson process as a semimartingale Let Y be a Poisson process with $Y_1 \sim Po(\psi)$. Then Y is of finite variation and hence Y is a semimartingale. We can write it in the form of a semimartingale, as

$$Y = Y_0 + A + M, \tag{3}$$

where $Y_0 = 0$,

$$A_t = \mathrm{E}(Y_t | \mathcal{F}_0) = \psi t \qquad \text{and} \qquad M_t = Y_t - \psi t.$$

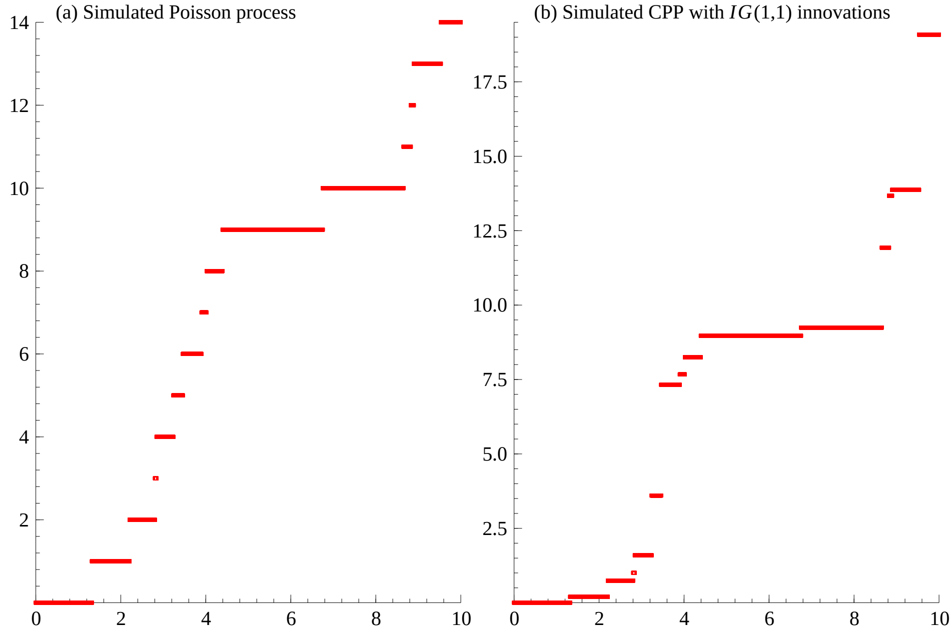

Figure 1: *Horizontal axis is time* t*, vertical is* Y_t*. (a) Sample path of a homogeneous Poisson process with intensity* $\psi = 1$*. (b) Corresponding compound Poisson process with* $C_j \sim IG(1,1)$*. code:* `levy_graphs.ox`.

The M process, which is often called a compensated Poisson process, is a mean-0 martingale, for $\mathrm{E}(M_{t+s}|\mathcal{F}_t) = M_t$ and $\mathrm{E}(M_t) = 0$ for $t, s \geq 0$. The A process is of finite variation.

As Y is a Poisson process, any stochastic integral $X = H \bullet Y$ simplifies to the random sum

$$X_t = \sum_{j=1}^{Y_t} H_{\tau_j -}$$

where $\tau_1, \tau_2, ..., \tau_{Y_t}$ are the arrival times of Y.

3.1.2 Compound Poisson process

Suppose N is a Poisson process and C is an i.i.d. sequence of strictly positive random variables. Then define a *compound Poisson process* as (for $t \geq 0$)

$$Y_t = \sum_{j=1}^{N_t} C_j, \quad \text{where} \qquad Y_0 = 0. \tag{4}$$

That is, Y_t is made up of the addition of a random number N_t of i.i.d. random variables. This is a Lévy process for the increments of this process

$$Y_{t+s} - Y_t = \sum_{j=N_t+1}^{N_{t+s}} C_j$$

are independent and are stationary as the increments of the Poisson process are independent and stationary.

At this point it is important to note that there is no added flexibility if the distribution of the C is allowed to have a positive probability of being exactly zero for this would, in effect, just knock out or *thin* some of the Poisson process arrivals. This point will recur in our later exposition. It is informative to note that for non-negative C

$$\begin{aligned}
\overline{\mathrm{K}}\left\{\theta \ddagger Y_1\right\} &= \log\left[\mathrm{E}\exp\left\{-\theta Y_1\right\}\right] \\
&= \log\left[\mathrm{E}_{N_1}\mathrm{E}\exp\left\{-\theta Y_1\right\}|N_1\right] \\
&= \log\left(\mathrm{E}\exp\left[N_1\overline{\mathrm{K}}\left\{\theta \ddagger C_1\right\}\right]\right) \\
&= \mathrm{K}\left\{\overline{\mathrm{K}}\left(\theta \ddagger C_1\right) \ddagger N_1\right\} \\
&= -\psi\left\{1-\exp\overline{\mathrm{K}}\left(\theta \ddagger C_1\right)\right\}. \qquad (5)
\end{aligned}$$

Example 1 *Figure 1(b) gives a simulation using* $C_j \overset{i.i.d.}{\sim} IG(1,1)$, *taking exactly the same Poisson process draws as used in Figure 1(a). IG(1,1) is a special case of the inverse Gaussian law IG(*δ,γ*), with density*

$$\frac{\delta}{\sqrt{2\pi}}e^{\delta\gamma}y^{-3/2}e^{-\frac{1}{2}\left(\delta^2y^{-1}+\gamma^2y\right)}, \quad y \in \mathbb{R}_{>\mathbf{0}}.$$

The resulting $\overline{\mathrm{K}}\left\{\theta \ddagger Y_1\right\}$ *is*

$$-\psi\left(1-\exp\left[\delta\left\{\gamma-\left(\gamma^2+2\theta\right)^{1/2}\right\}\right]\right).$$

3.1.3 Negative binomial process

A negative binomial Lévy process Y requires the process at time one to follow a negative binomial distribution $Y_1 \sim NB(r,p)$, where $r>0$, $p\in(0,1)$, with probability function

$$\Pr(Y_1=y)=\frac{\Gamma\left(r+y\right)}{y!\Gamma\left(r\right)}p^r\left(1-p\right)^y, \quad y=0,1,2,....$$

This implies

$$\overline{\mathrm{K}}\left\{\theta \ddagger Y_1\right\}=r\log p-r\log\left\{1-\left(1-p\right)e^{-\theta}\right\}. \qquad (6)$$

So this is infinitely divisible, so supports a Lévy process. The form of the cumulant function implies $Y_t \sim NB(tr,p)$ — so this has a finite number of jumps over a finite time interval. Such processes are called "finite activity" and in this case it looks somwhat like a Poisson process. However, the spaces in times between arrivals are not uniformly distributed and the variance of the number of arrivals in a fixed time interval will be bigger than the corresponding expected value. We will discuss the relationship between these two processes in the Section 5 on time-change.

3.1.4 Infinite activity subordinators

Gamma process The Poisson process and the compound Poisson process are by far the most well-known non-negative Lévy processes in economics. The jumps happen, typically, rather rarely. Consequently increments to these processes are often exactly zero, even when measured over quite large time intervals. This feature of the process is fundamentally different from the gamma Lévy process. A *gamma Lévy process* Y makes Y_1 obey a gamma law

$$Y_1 \sim \Gamma(\nu, \alpha), \quad \nu, \alpha > 0,$$

with density

$$f_{Y_1}(y) = \frac{\alpha^{\nu}}{\Gamma(\nu)} y^{\nu-1} \exp(-\alpha y), \quad y \in \mathbb{R}_{>\mathbf{0}}. \tag{7}$$

Here $2\nu > 0$ is thought of as a degrees of freedom parameter, controlling the skewness of the distribution. The other parameter, α, is a scale parameter.

The kumulant function of the gamma distribution is

$$\overline{\mathrm{K}}\{\theta \ddagger Y_1\} = \nu \log\left(1 + \frac{\theta}{\alpha}\right),$$

implying

$$\overline{\mathrm{K}}\{\theta \ddagger Y_t\} = t\nu \log\left(1 + \frac{\theta}{\alpha}\right),$$

and so $Y_t \sim \Gamma(\nu t, \alpha)$. The gamma process has the useful property that it has increments which are strictly positive whatever small time interval has elapsed. Such Lévy processes are said to have *infinite activity* ($\mathcal{IA}$). This feature puts them apart from a compound Poisson process.

A sample path of a gamma process is drawn in Figure 2(a). The path is dominated by large jumps and shows informally that the path of the process is not continuous (anywhere). It was drawn by splitting time into intervals of length 1/2000 and sampling from the implied random walk with increments taken from the $\Gamma(\nu/2000, \alpha)$ distribution. Very similar paths are produced by using smaller time intervals. The process is a rough upward trend with occasional large shifts.

In the special case of $\nu = 1$, then

$$f_{Y_1}(y) = \alpha \exp(-\alpha y), \quad y \in \mathbb{R}_{>0},$$

which is the exponential distribution $\Gamma(1, \alpha)$. Thus the exponential Lévy process has $Y_t \sim \Gamma(t, \alpha)$.

Inverse Gaussian process An *inverse Gaussian* (IG) Lévy process Y requires the process at time one to follow an inverse Gaussian distribution $Y_1 \sim IG(\delta, \gamma)$, where $\delta > 0, \gamma \geq 0$, with density

$$f_{Y_1}(y) = \frac{\delta}{\sqrt{2\pi}} e^{\delta\gamma} y^{-3/2} \exp\left\{-\frac{1}{2}\left(\delta^2 y^{-1} + \gamma^2 y\right)\right\}, \quad y \in \mathbb{R}_{>0}.$$

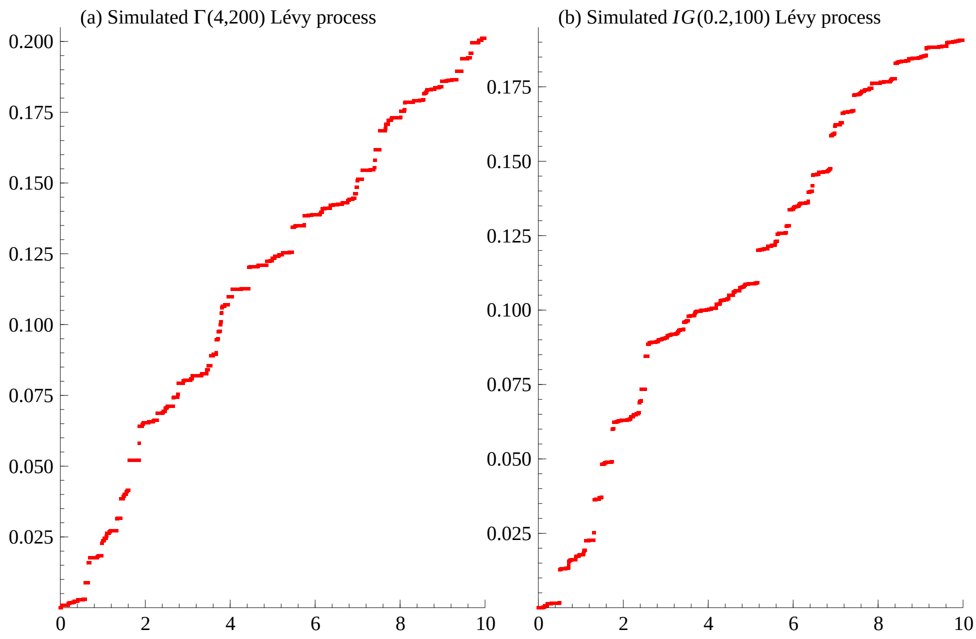

Figure 2: *Simulated* Γ *and* IG *Lévy processes, using intervals of length 1/2000.* *Code* `levy_graphs.ox`.

This implies

$$\overline{\mathrm{K}}\left\{\theta \ddagger Y_1\right\} = \delta\left\{\gamma - \left(\gamma^2 + 2\theta\right)^{1/2}\right\}.$$

Like the gamma process, the IG process has an infinite number of jumps in any small interval of time. The form of the cumulant function implies $Y_t \sim IG(t\delta, \gamma)$.

A sample path of an IG Lévy process is drawn in Figure 2(b). The parameters were selected to have the same mean and variance of Y_1 as that used to draw the path of the gamma process given in Figure 2(a). Again the process is a rough upward trend with occasional large shifts.

Some other non-negative processes A *reciprocal* (or inverse) *gamma* ($R\Gamma$) Lévy process Y requires the process at time one to be a reciprocal gamma variable $Y_1 \sim R\Gamma(\nu, \alpha)$, $\alpha, \nu > 0$, so that $Y_1^{-1} \sim \Gamma(\nu, \alpha)$, with density

$$f_{Y_1}(y) = \frac{\alpha^\nu}{\Gamma(\nu)} y^{-\nu-1} \exp\left(-\alpha y^{-1}\right), \qquad y \in \mathbb{R}_{>0}.$$

Only the moments of order less than ν exist for this distribution.

Sums of independent reciprocal gamma variables are not distributed as reciprocal gamma. However, the reciprocal gamma is infinitely divisible (and so yields a Lévy process), although we do not know the distribution of Y_t in closed form. This makes simulation of this process non-trivial.

A *lognormal* (LN) Lévy process Y requires the process at time one to be a lognormal variable

$Y_1 \sim LN(\mu, \sigma^2)$, $\sigma^2 \geq 0$, with infinitely divisible density

$$f_{Y_1}(y) = \frac{1}{\sqrt{2\pi}} y^{-1} \exp\left\{-\frac{1}{2\sigma^2}(\log y - \mu)^2\right\}, \quad y \in \mathbb{R}_{>0}.$$

The proof that the lognormal is infinitely divisible is probabilistically challenging (cf. Thorin (1977)) and not discussed here.

A *reciprocal* (inverse) Gaussian (RIG) Lévy process Y requires the process at time one to be a reciprocal inverse Gaussian variable $Y_1 \sim RIG(\delta, \gamma)$, $\delta > 0$, $\gamma > 0$, with density

$$f_{Y_1}(y) = \frac{\gamma}{\sqrt{2\pi}} e^{\delta\gamma} y^{-1/2} \exp\left\{-\frac{1}{2}\left(\delta^2 y^{-1} + \gamma^2 y\right)\right\}, \quad y \in \mathbb{R}_{>0}.$$

The corresponding kumulant function is

$$\overline{\mathrm{K}}\left\{\theta \ddagger Y_1\right\} = -\frac{1}{2}\log\left(1 + 2\theta/\gamma^2\right) + \delta\gamma\left\{1 - \left(1 + 2\theta/\gamma^2\right)^{1/2}\right\}.$$

By construction, $Y_1^{-1} \sim IG(\gamma, \delta)$ provided $\delta > 0$. Again sums of independent RIG variables are not distributed as RIG, but the RIG distribution is infinitely divisible.

Generalised inverse Gaussian Lévy process Many of the above infinite activity processes are special cases of the *generalized inverse Gaussian* (GIG) Lévy process. This puts

$$Y_1 \sim GIG(\nu, \delta, \gamma),$$

with GIG density

$$f_{Y_1}(y) = \frac{(\gamma/\delta)^\nu}{2K_\nu(\delta\gamma)} y^{\nu-1} \exp\left\{-\frac{1}{2}(\delta^2 y^{-1} + \gamma^2 y)\right\}, \quad y \in \mathbb{R}_{>0}, \tag{8}$$

where $K_\nu(\cdot)$ is a modified Bessel function of the third kind (cf. Barndorff-Nielsen and Shephard (2012c)). This density has been shown (Barndorff-Nielsen and Halgreen (1977)) to be infinitely divisible. Prominent special cases are achieved in the following ways:

$$\begin{array}{ll}
IG(\delta,\gamma) = GIG(-\tfrac{1}{2},\delta,\gamma), & PH(\delta,\gamma) = GIG(1,\delta,\gamma), \\
R\Gamma(\nu,\delta^2/2) = GIG(-\nu,\delta,0), & \Gamma(\nu,\gamma^2/2) = GIG(\nu>0,0,\gamma), \\
RIG(\delta,\gamma) = GIG(\tfrac{1}{2},\delta,\gamma), & PHA(\delta,\gamma) = GIG(0,\delta,\gamma) \\
RPH(\delta,\gamma) = GIG(-1,\delta,\gamma). &
\end{array}$$

Here all these distributions have been introduced above except for

- the positive hyperbolic distribution (PH);
- the positive hyperbola distribution (PHA);
- reciprocal positive hyperbolic (RPH);

In order to obtain these results we have to allow δ or γ to be zero 0. In these cases the GIG's density has to be interpreted in the limiting sense, using the well-known results that for $y \downarrow 0$ we have

$$K_\nu(y) \sim \begin{cases} -\log y & \text{if} \quad \nu = 0 \\ \\ \Gamma(|\nu|)2^{|\nu|-1}y^{-|\nu|} & \text{if} \quad \nu \neq 0. \end{cases}.$$

In general, the density of the increments to this process is unknown in explicit form and we cannot directly simulate from it without using computationally intensive methods.

3.2 Lévy measures for non-negative processes

It should be clear by now that the kumulant function of Y_1 plays an important role in Lévy processes. In this subsection this observation will be further developed in order to build towards the vital Lévy-Khintchine representation which shows us the form characteristic functions of Lévy processes must obey. As this representation is so important, and is also mathematically involved, its development will be carried out in stages. At first sight this looks unnecessary from a modelling viewpoint, however we will see that practical modelling will sometimes be carried out directly via some of the terms which make up the Lévy-Khintchine representation. Hence a good understanding of this Section is essential for later developments.

To start off with, think of a Poisson process $Y_t \sim Po(\psi t)$. Then, letting $\delta_1(y)$ be the Dirac delta centred at $y = 1$, we write

$$\begin{aligned} \overline{\mathrm{K}}\{\theta \ddagger Y_1\} &= -\psi(1 - e^{-\theta}) \\ &= -\psi \int_0^\infty (1 - e^{-\theta y})\delta_1(\mathrm{d}y) \\ &= -\psi \int_0^\infty (1 - e^{-\theta y})P(\mathrm{d}y), \end{aligned}$$

where P is the Dirac delta probability measure centred at one. The introduction of the probability measure is entirely expository in this context; however, expressing kumulant functions in this type of way will become essential later. Before proceeding another level of abstraction has to be introduced. Instead of working with probability measures we will have to use more general measures v concentrated on $\mathbb{R}_{>0}$. An important point is that some of the measures that will be used later will not be integrable (that is $\int_0^\infty v(\mathrm{d}y) = \infty$) and so probability measures are insufficient for a discussion of Lévy processes. In the simple Poisson case the measure is introduced by expressing

$$\overline{\mathrm{K}}\{\theta \ddagger Y_1\} = -\int_0^\infty (1 - e^{-\theta y})v(\mathrm{d}y) \tag{9}$$

where $v = \psi\delta_1$ is called the *Lévy measure*. Of course this measure is integrable, indeed it integrates to ψ.

Let us now generalise the above setup to the compound Poisson process (4), but still requiring C to be strictly positive — ruling out the possibility that C can be exactly zero with non-zero probability. Then, writing the distribution of C_1 as $P(\mathrm{d}y \ddagger C_1)$,

$$\begin{aligned}\overline{\mathrm{K}}\{\theta \ddagger Y_1\} &= -\psi\left\{1 - \exp\overline{\mathrm{K}}\left(\theta \ddagger C_1\right)\right\} \\ &= -\psi \int_0^\infty (1 - e^{-\theta y}) P(\mathrm{d}y \ddagger C_1) \\ &= -\int_0^\infty (1 - e^{-\theta y}) v(\mathrm{d}y), \end{aligned} \tag{10}$$

again, but now with $v(\mathrm{d}y) = \psi P(\mathrm{d}y \ddagger C_1)$. Again this measure is integrable as it is proportional to the probability measure of C_1. In the simple case where C_1 has a density we write

$$v(\mathrm{d}y) = u(y)\mathrm{d}y$$

and call $u(y)$ (which is ψ times the density of C_1) the *Lévy density*. In such cases the kumulant function becomes

$$\overline{\mathrm{K}}\{\theta \ddagger Y_1\} = -\int_0^\infty (1 - e^{-\theta y}) u(y) \mathrm{d}y.$$

A simple example of this is where $C_j \overset{i.i.d.}{\sim} \Gamma(\nu, \alpha)$. Then the Lévy density is

$$u(y) = \psi \frac{\alpha^\nu}{\Gamma(\nu)} y^{\nu-1} \exp(-\alpha y). \tag{11}$$

Of course this Lévy density integrates to ψ — not one.

Remark 1 *If they exist, the cumulants of* Y_1 *satisfy*

$$\kappa_j = \left.\frac{\partial^j \mathrm{K}\{\theta \ddagger Y_1\}}{\partial \theta^j}\right|_{\theta=0} = \int_0^\infty y^j v(\mathrm{d}y),$$

i.e. they equal the moments of v.

Although Poisson and compound Poisson processes have integrable Lévy measures, for v is proportional to a probability measure which integrates to one, theoretically more general Lévy processes can be constructed without abandoning the form (9). The non-integrable measures v will not correspond to compound Poisson processes. To ensure that they yield a valid kumulant function we require that $\int_0^\infty (1 - e^{-\theta y}) v(\mathrm{d}y)$ exists for all $\theta > 0$, while continuing to rule out the possibility that v has an atom at zero. It is simple to prove that a necessary and sufficient condition for existence is that

$$\int_0^\infty \min(1, y)\, v(\mathrm{d}y) < \infty.$$

If the Lévy measure is absolutely continuous then we can define $u(y)$ as the Lévy density where $v(\mathrm{d}y) = u(y)\mathrm{d}y$. However, as some Lévy measures are not finite, it follows that Lévy densities are not necessarily integrable. This is at first sight confusing. This point comes up in the following two examples.

Example 2 *It turns out that the Lévy density of $Y_1 \sim IG(\delta, \gamma)$ is*

$$u(y) = (2\pi)^{-1/2}\,\delta y^{-3/2}\exp(-\gamma^2 y/2), \quad y \in \mathbb{R}_{>0}. \tag{12}$$

This Lévy density is not integrable as it goes off to infinity too rapidly as y goes to zero. This implies an IG process is not a compound Poisson process — it has an infinite number of jumps in any finite time period. Although the Lévy density is not integrable it does satisfy the finiteness condition on $\int_0^\infty \min(1, y)\, v(\mathrm{d}y)$ for the addition of the y factor regularises the density near zero.

Example 3 *It can also be shown that the Lévy density of $Y_1 \sim \Gamma(\nu, \alpha)$ is*

$$u(y) = \nu y^{-1}\exp(-\alpha y), \quad y \in \mathbb{R}_{>0}. \tag{13}$$

Again this is not an integrable Lévy density although it is slower to go off to infinity than the inverse Gaussian case. This means in practice that it will have fewer very small jumps than the IG process.

These two results are special cases of the result for the $GIG(\nu, \delta, \gamma)$ distribution (8). The corresponding Lévy density is then

$$u(y) = y^{-1}\left[\frac{1}{2}\int_0^\infty e^{-\frac{1}{2}\delta^{-2}y\xi} g_\nu(\xi)\mathrm{d}\xi + \max\{0, \nu\}\lambda\right]\exp\left(-\gamma^2 y/2\right) \tag{14}$$

where

$$g_\nu(y) = \frac{2}{y\pi^2}\left\{J^2_{|\nu|}(\sqrt{y}) + N^2_{|\nu|}(\sqrt{y})\right\}^{-1}$$

and J_ν and N_ν are Bessel functions. This result is derived in Barndorff-Nielsen and Shephard (2012a). Although this looks forbidding, when ν is a half integer these functions are analytically simple.

Example 4 *A case outside the GIG class is the positive stable $PS(\kappa, \delta)$ process. Although the probability density of this variable is in general unknown in simple form (see Barndorff-Nielsen and Shephard (2012a) for details), the kumulant function is*

$$\overline{\mathrm{K}}\{\theta \ddagger Y_1\} = -\delta\,(2\theta)^\kappa, \qquad 0 < \kappa < 1, \quad \delta > 0,$$

which implies it does not possess moments of order κ and above. The Lévy density for the positive stable Lévy process is given by

$$u(y) = cy^{-1-\kappa}, \quad \textit{where} \quad c = \delta 2^\kappa \frac{\kappa}{\Gamma(1-\kappa)}, \tag{15}$$

while $Y_t \sim PS(\kappa, t\delta)$. Feller (1971a) discusses many basic facts and examples of stable laws, while Samorodnitsky and Taqqu (1994) give a very comprehensive account of these laws and the associated Lévy and other processes, in particular fractional Brownian motion. See also Sato (1999).

3.3 Lévy-Khintchine representation for non-negative processes

3.3.1 Representation

Having allowed the Lévy measure not to be integrable, a single extra step is required in order to produce a general setup. We allow a drift $a > 0$ to be added to the cumulant function. This is carried out in the following fundamental theorem.

Theorem 2 *Lévy-Khintchine representation for non-negative Lévy processes. Suppose Y is a Lévy process with non-negative increments. Then the kumulant function can be written as*

$$\overline{\mathrm{K}}\{\theta \ddagger Y_1\} = -a\theta - \int_0^\infty \left(1 - e^{-\theta y}\right) v(\mathrm{d}y), \tag{16}$$

where $a \geq 0$ and v is a measure on $\mathbb{R}_{>0}$ such that

$$\int_0^\infty \min\{1, y\} v(\mathrm{d}y) < \infty. \tag{17}$$

Conversely, any pair (a, v) with these properties determines a non-negative Lévy process Y such that Y_1 has kumulant function determined by (16).

The importance of this representation is that the kumulant function of all non-negative Lévy processes can be written in this form. In other words, non-negative Lévy processes are completely determined by a and the Lévy measure v (which has to satisfy (17)).

In the special case when $\int_0^\infty v(\mathrm{d}y) < \infty$ we say that Y is of *finite activity* $(\mathcal{FA})$ — indeed all such processes can be written as a compound Poisson process. In cases where this does not hold, Y is said to be an *infinite activity* $(\mathcal{IA})$ process as it has an infinite number of very small jumps in any finite time interval.

Importantly, when we move on to Lévy processes on the real line the condition (17) has to be strengthen to $\int_{|y|>0} \min\{1, y^2\} v(\mathrm{d}y) < \infty$. This will be discussed in some detail in Section 4.3.

3.3.2 Positive tempered stable process

An important implication of the Lévy-Khintchine representation is that Lévy processes can be built by specifying a and v directly, implying the probability distribution of Y_1. An important example of this is the positive tempered stable, $PTS(\kappa, \delta, \gamma)$, class which exponentially tilts the Lévy density $p_S(y; \kappa, \delta)$ of the $PS(\kappa, \delta)$, to deliver

$$p(y; \kappa, \delta, \gamma) = e^{\delta\gamma} \exp\left(-\frac{1}{2}\gamma^2 y\right) p_S(y; \kappa, \delta), \quad y \in \mathbb{R}_{>0}.$$

The resulting Lévy density is

$$u(y) = \delta\gamma^{-2\kappa}\frac{\kappa}{\Gamma(\kappa)\Gamma(1-\kappa)}y^{-\kappa-1}\exp\left(-\frac{1}{2}\gamma^2 y\right), \quad y,\delta>0,\ 0<\kappa<1,\ \gamma\geq 0, \tag{18}$$

which means the process has infinite activity. The density of Y_1 is not generally known in simple form, however the Lévy density is simple and plays such a crucial role that this difficulty is not overly worrying. Special cases of this structure include the IG Lévy density (12) and the Γ Lévy density (13), which is the limiting case of $\kappa \downarrow 0$. Notice that the constraint $\kappa < 1$ is essential in order to satisfy the condition (17) in the Lévy-Khintchine representation. The corresponding kumulant function can be calculated by solving (16) implying

$$\overline{\mathrm{K}}\{\theta \ddagger Y_1\} = \delta\gamma - \delta\left(\gamma^{1/\kappa}+2\theta\right)^{\kappa},$$

while for $\gamma > 0$ all cumulants of Y_1 exist, the first two cumulants being

$$2\kappa\delta\gamma^{(\kappa-1)/\kappa} \quad \text{and} \quad 4\kappa(1-\kappa)\delta\gamma^{(\kappa-2)/\kappa}.$$

Finally the kumulant function implies the convenient property that $Y_t \sim PTS(\kappa, t\delta, \gamma)$.

3.4 Simulation for non-negative Lévy processes*

3.4.1 Simulating paths via the Lévy density

For subordinators the paths of Lévy processes can be simulated directly off the Lévy density $u(y)$, $y \in \mathbb{R}_{>0}$. Define the tail mass function

$$v^+(y) = \int_y^\infty u(x)\mathrm{d}x,$$

which is a decreasing function for all $y \in \mathbb{R}_{>0}$. Denote the inverse function of v^+ by $v^{\leftarrow}$, i.e.

$$v^{\leftarrow}(y) = \inf\left\{x > 0 : v^+(x) \leq y\right\}.$$

Then the desired result, called the *series representation*, is that

$$Y_t^{(m)} = \sum_{i=1}^{m} v^{\leftarrow}(b_i/T)I(u_i \leq t), \quad \text{for} \quad 0 \leq t \leq T, \tag{19}$$

converges uniformly to Y on $[0,T]$ as $m \to \infty$. Here $u_i \overset{i.i.d.}{\sim} U(0,T)$, where $U(0,T)$ is the uniform distribution on $(0,T)$, is independent of $b_1 < ... < b_i < ...$ which are the arrival times of a Poisson process with intensity 1. Clearly the computational speed of these techniques will depend upon how expensive it is to compute $v^{\leftarrow}$ and how quickly $v^{\leftarrow}(y)$ falls as y increases.

Example 5 *Compound Poisson process. Let the CPP have intensity ψ and probability density $f(y)$ for the positive jumps, then the Lévy density and tail mass function are $u(y) = \psi f(y)$ and $v^+(y) = \psi\{1 - F(y)\}$, implying*

$$v^{\leftarrow}(y) = \begin{cases} F^{-1}\left\{1 - \left(\frac{y}{\psi}\right)\right\}, & y < \psi \\ 0, & y \geq \psi. \end{cases}$$

Hence the inverse only involves computing the quantiles of the jumps. Overall this implies

$$Y_t^{(m)} = \sum_{b_i \leq T\psi}^{m} F^{-1}\left\{1 - \left(\frac{b_i}{\psi T}\right)\right\} I(u_i \leq t).$$

Clearly if $b_i > T\psi$, then there is no approximation by using this series representation. This method has a simple interpretation. If we sample from

$$F^{-1}\left\{1 - \left(\frac{b_i}{\psi T}\right)\right\} \quad \text{until} \quad b_i > T\psi,$$

then an ordered sequence from $f(y)$ of size $Po(\psi T)$ is produced. The effect of the $I(u_i \leq t)$ term is to sample randomly from this ordered sequence a random share of the draws. So the infinite series representation samples compound Poisson processes rather effectively.

As a special case, suppose $u(y) = \nu\alpha\exp(-\alpha y)$ so that $v^+(y) = \nu e^{-\alpha y}$, which has the convenient property that it can be analytically inverted:

$$v^{\leftarrow}(y) = \max\left\{0, -\frac{1}{\alpha}\log\left(\frac{y}{\nu}\right)\right\}.$$

Hence as soon as $y > \nu$ then $v^{\leftarrow}(y) = 0$, implying

$$Y_t^{(m)} = -\frac{1}{\alpha}\sum_{b_i \leq t\nu}^{m} \log\left(\frac{b_i}{t\nu}\right) I(u_i \leq t).$$

For some types of subordinators special methods have been devised to simulate their paths without inverting $v^{\leftarrow}$.

Example 6 *Rosinski's method for dealing with $PTS(\kappa, \delta, \gamma)$ processes. This method approximates the path over $[0, T]$ by the sum of non-negative terms*

$$Y_t^{(m)} = \sum_{i=1}^{m} \min\left\{\left(\frac{AT}{b_i\kappa}\right)^{1/\kappa}, B^{-1}e_i v_i^{1/\kappa}\right\} I(u_i \leq t), \quad \text{for} \quad 0 \leq t \leq T, \tag{20}$$

where

$$A = \delta 2^{\kappa}\frac{\kappa}{\Gamma(1-\kappa)}, \quad B = \frac{1}{2}\gamma^{1/\kappa},$$

$I(\cdot)$ is an indicator function, $\{e_i\}$, $\{v_i\}$, $\{b_i\}$, $\{u_i\}$ are independent of one another and over i except for the $\{b_i\}$ process. Here $u_i \overset{i.i.d.}{\sim} U(0,T)$, $v_i \overset{i.i.d.}{\sim} U(0,1)$, the $\{e_i\}$ are exponential with mean 1. Further the $b_1 < ... < b_i < ...$ are the arrival times of a Poisson process with intensity 1. Then as $m \to \infty$ the process $Y_t^{(m)}$ converges uniformly on $[0,T]$ to a sample path of the tempered stable Lévy process.

3.4.2 Simulation via a small jump approximation

We will break up the subordinator into

$$Y = Y^1 + Y^2,$$

where Y^2 corresponds to a compound Poisson process with jumps of absolute size greater than $\varepsilon > 0$, while Y^1 will contain the small jumps. A first order approximation is to take ε as very small and to neglect Y^1 entirely and simply simulate the straightforward Y^2.

A second order approximation to the path of Y^1 is by using a Brownian motion, based on the central limit result

$$\left\{\sigma_\varepsilon^{-1}\left(Y_t^1 - \mu_\varepsilon t\right)\right\}_{t\geq 0} \to \{W_t\}_{t\geq 0},$$

where $\mu_\varepsilon = \int_{-\varepsilon}^{\varepsilon} y v(\mathrm{d}y)$ and $\sigma_\varepsilon^2 = \int_{-\varepsilon}^{\varepsilon} y^2 v(\mathrm{d}y)$, which holds under certain conditions. A more detailed discussion of this is given in Barndorff-Nielsen and Shephard (2012a).

4 Processes with real increments

4.1 Examples of Lévy processes

4.1.1 Motivation

In this Section the focus will be on Lévy processes with innovations which are on the real line. Many of them play important roles in financial economics as direct models of financial assets.

4.1.2 Brownian motion

In Brownian motion we write

$$Y_1 \sim N(0,1),$$

with density

$$f_{Y_1}(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y^2\right), \qquad y \in \mathbb{R},$$

while

$$\mathrm{K}\left\{\theta \ddagger Y_1\right\} = \log\left[\mathrm{E}\exp\left\{\theta Y_1\right\}\right] = \frac{1}{2}\theta^2.$$

The implication of this is that marginally $Y_t \sim N(0, t)$, while the increments are independent as usual with

$$Y_{t+s} - Y_t \sim N(0, s).$$

A standard Brownian motion, written W, can be generalised to allow for the increments to have a non-zero mean and a different scale than one. A *drift* μ and a *volatility* term σ can be introduced to deliver the Lévy process

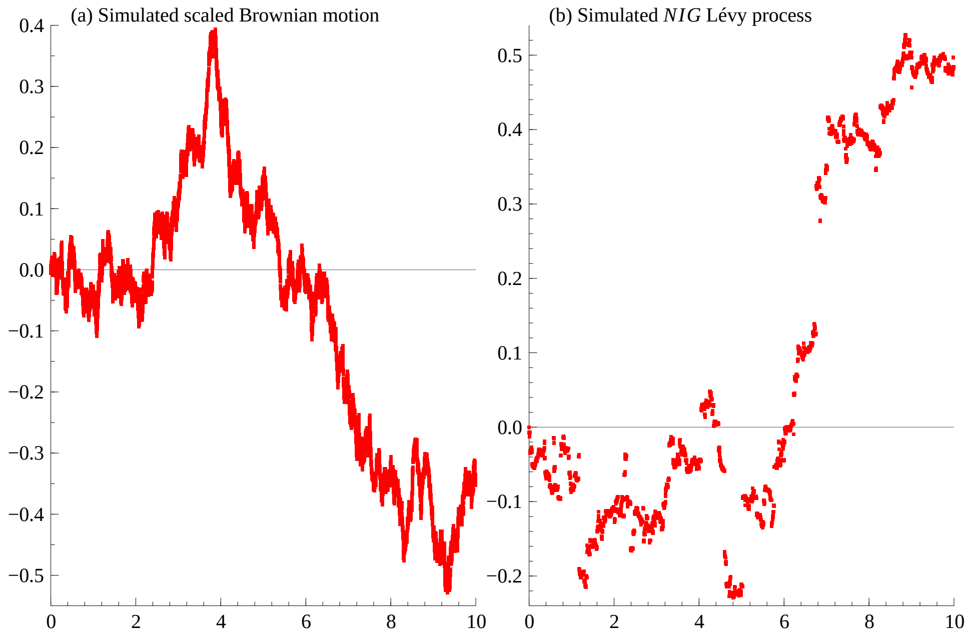

Figure 3: *(a) Sample path of* $\sqrt{0.02}$ *times standard Brownian motion. (b) Sample path of a* $NIG(0.2, 0, 0, 10)$ *Lévy process. Thus the increments of the processes have a common variance. Code:* `levy_graphs.ox`.

$$Y_t = \mu t + \sigma W_t,$$

with increments

$$Y_{t+s} - Y_t \sim N(\mu s, \sigma^2 s).$$

The associated kumulant function for Y_1 is $\mu\theta + \frac{1}{2}\theta^2\sigma^2$.

A graph of a sample path from standard Brownian motion is displayed in Figure 3(a). It illustrates that the path is continuous. In a moment we will see that, except for the pure linear drift case $Y_t = \mu t$, Brownian motion is the only Lévy process with this property — all other Lévy processes have jumps.

4.1.3 Compound Poisson process

Compound Poisson processes were introduced in (4), but there we required the shocks $\{C_j\}$ to be strictly positive. Here this condition is relaxed, just ruling out that they have an atom at zero. In this case, again,

$$\begin{aligned} \mathrm{K}\{\theta \ddagger Y_1\} &= \log\left[\mathrm{E}\exp\{\theta Y_1\}\right] \\ &= \psi\{\exp \mathrm{K}\{\theta \ddagger C_1\} - 1\} \end{aligned}$$

(so long as $\mathrm{K}\{\theta \ddagger C_1\}$ exists).

Example 7 *Suppose that* $C_j \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$*, then*

$$\mathrm{K}\{\theta \ddagger Y_1\} = \psi\left\{\exp\left(\mu\theta + \frac{1}{2}\theta^2\sigma^2\right) - 1\right\}.$$

So here the Lévy process is constant until there is a new arrival from the Poisson process. The arrival then moves the Lévy process by a Gaussian variable. This variable can have a non-zero mean and a non-unit variance. Quite a lot of effort has been expended on working on the derivative pricing theory associated with this simple structure.

4.1.4 Skellam process

Let

$$Y_t = N_t^{(1)} - N_t^{(2)},$$

where $N_t^{(i)}$ are independent Poisson processes with $\mathrm{E}(N_1^{(i)}) = \psi_i > 0$. This is a Skellam process and has the attractive feature that it can be scaled to have a fixed tick size which may be helpful for certain types of high frequency financial data.

The essential nature of this discreteness at the microscopic financial level is shown in Figure 4 which reports the first 80 best bid (squares) and best ask (crosses) rescaled prices from a Euro/Dollar futures contract for 10th November 2008. For this contract the tick size is 0.0001 of a unit, i.e. prices move from, for example, 1.2768 to 1.2767 U.S. Dollar to the Euro.

Now for the Skellam process

$$\mathrm{K}\{\theta \ddagger Y_1\} = \mathrm{K}\left\{\theta \ddagger N_1^{(1)}\right\} + \mathrm{K}\left\{-\theta \ddagger N_1^{(2)}\right\} = \left(\psi_1 e^{\theta} + \psi_2 e^{-\theta} - \psi_1 - \psi_2\right),$$

so is infinitely divisible. The Y_t is distributed as a Skellam $Sk(t\psi_1, t\psi_2)$ variable — the distribution of the difference between two independent Poisson variables with means $\psi_1 t$ and $\psi_2 t$. A $Sk(\psi_1, \psi_2)$ probability function is

$$p_k = e^{-\psi_1-\psi_2}\sum_{n=0}^{\infty}\frac{\psi_1^k\psi_2^{k+n}}{k!(k+n)!} = e^{-\psi_1-\psi_2}\left(\frac{\psi_1}{\psi_2}\right)^{k/2} I_k(2\sqrt{\psi_1\psi_2}), \quad k = 0, \pm1, \pm2, ...,$$

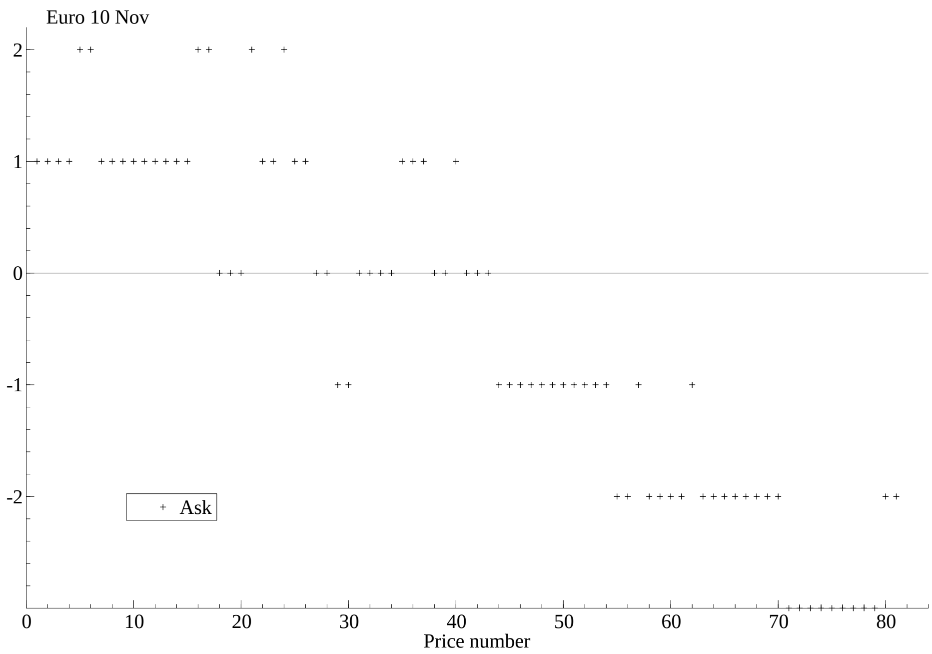

Figure 4: *From* Barndorff-Nielsen, Pollard, and Shephard (2012)

where $I_k(x)$ is a modified Bessel function of the first kind. Importantly $\mathrm{E}(Y_t) = (\psi_1 - \psi_2)\, t$ and $\mathrm{Var}(Y_t) = (\psi_1 + \psi_2)\, t$. Hence if $\psi_1 = \psi_2$ the process is a martingale.

4.2 Normal variance-mean mixture processes

A rather general way of building densities on the real line is by using a normal variance-mean mixture

$$Y = \mu + \beta\sigma^2 + \sigma U, \tag{21}$$

where $U \sim N(0,1)$ and $U \perp\!\!\!\perp \sigma^2$ ($\perp\!\!\!\perp$ denoting stochastic independence). In this mixture the constraint $\beta = 0$ would imply Y would be symmetrically distributed, while $\beta < 0$ typically delivers a negatively skewed density. Further, $\mathrm{E}(Y) = \mu + \beta\mathrm{E}(\sigma^2)$ so long as $\mathrm{E}(\sigma^2)$ exists and

$$\mathrm{C}\left\{\zeta \ddagger Y_1\right\} = i\mu\zeta + \mathrm{K}\left\{i\beta\zeta - \frac{1}{2}\zeta^2 \ddagger \sigma^2\right\},$$

where for complex z,

$$\mathrm{K}\left\{z \ddagger \sigma^2\right\} = \log\left\{\mathrm{E}\exp\left(z\sigma^2\right)\right\}.$$

Provided σ^2 is infinitely divisible this normal variance-mean mixture has an elegant time deformation intepretation which will be an important common theme to the threads of this book. Many of the commonly used parametric densities used in financial economics fall within this class by specific choices for the density for σ^2. Here we discuss some of these important special cases.

4.2.1 Normal inverse Gaussian

If we assume $\sigma^2 \sim IG(\delta, \gamma)$ then $Y \sim NIG(\alpha, \beta, \mu, \delta)$, where $\alpha = \sqrt{\beta^2 + \gamma^2}$, has a normal inverse Gaussian (NIG) distribution. The NIG Lévy process puts

$$Y_1 \sim NIG(\alpha, \beta, \mu, \delta), \quad \mu \in \mathbb{R}, \delta \in \mathbb{R}_{>0}, \quad |\beta| < \alpha$$

which has the density

$$f_{Y_1}(y) = a(\alpha, \beta, \mu, \delta) q \left(\frac{y - \mu}{\delta} \right)^{-1} K_1 \left\{ \delta \alpha q \left(\frac{y - \mu}{\delta} \right) \right\} \exp \left\{ \beta \left(y - \mu \right) \right\},$$

where $q(y) = \sqrt{1 + y^2}$, $K_1(\cdot)$ is a modified Bessel function of the third kind and

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

Sometimes it is convenient to reparameterise this model, noting that

$$\overline{\alpha} = \delta \alpha, \quad \overline{\beta} = \delta \beta,$$

are invariant under changes of the location and scale parameters μ and δ. A popular location-scale invariant choice is achieved by defining

$$\xi = \left(1 + \delta \sqrt{\alpha^2 - \beta^2} \right)^{-1/2} \quad \text{and} \quad \chi = \frac{\beta}{\alpha} \xi, \tag{22}$$

where ξ, the steepness parameter, and χ, the asymmetry parameter, obey a triangular constraint

$$\left\{ (\chi, \xi) : -1 < \chi < 1, \quad |\chi| < \xi < 1 \right\}. \tag{23}$$

The flexibility of the model is shown in Figure 5, which displays the log-density for a variety of values of χ, ξ. Such a plot is called a *shape triangle*. As $\xi \to 0$ so the log-density becomes more quadratic, while for values of ξ around 0.5 the tails are approximately linear. For larger values of ξ the tails start decaying at a rate which looks appreciably slower than linear. In the limit as $\xi \to 1$ while $\chi = 0$ the density becomes that of a Cauchy variable.

Note that it is sometimes more convenient to think of $\rho = \beta/\alpha = \chi/\xi$, rather than χ, as an asymmetry parameter. A fixed value of ρ corresponds to a straight proportionality line in the shape triangle.

One of the NIG's attractions is that the cumulant function has the simple form

$$\mathrm{K} \left\{ \theta \ddagger Y_1 \right\} = \delta \left\{ \sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + \theta)^2} \right\} + \mu \theta,$$

which means that

$$Y_t \sim NIG(\alpha, \beta, t\mu, t\delta).$$

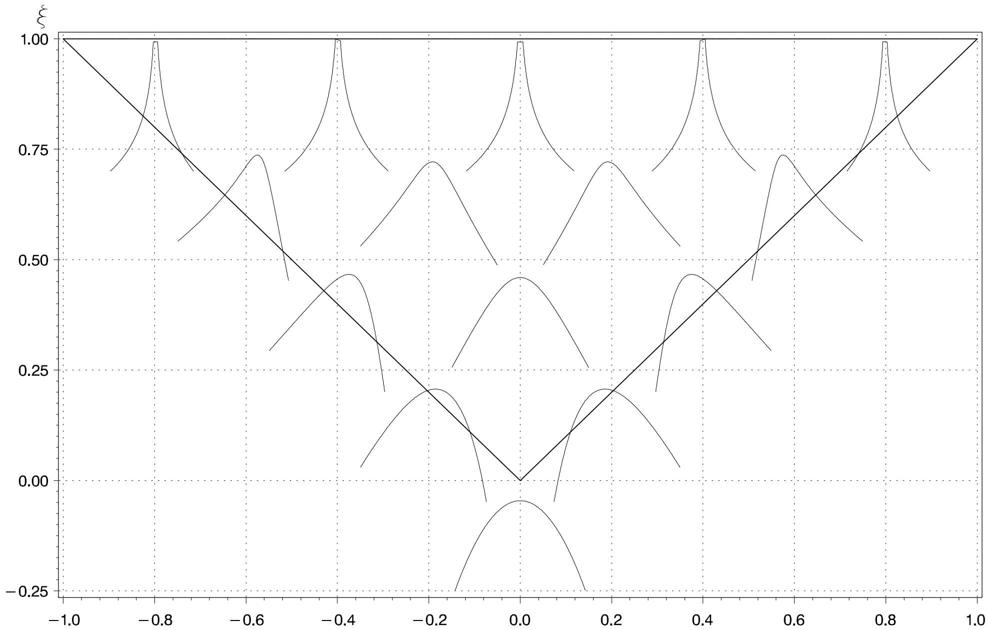

Figure 5: *Shape triangle for the NIG model. That is we graph the shape of the log-density for the NIG model for a variety of values of the steepness parameter ξ and the asymmetry parameter χ. The representative distributions shown all have variance of one. This graph was kindly made available to us by Preben Blæsild.*

In particular this implies that the increments of the NIG Lévy process are non-zero NIG with probability one. A sample path of a NIG Lévy process is drawn in Figure 3(b). For this case we have chosen there is symmetry and no drift (since $\mu = \beta = 0$) and the variance per unit of time is the same as for the Brownian motion given in Figure 3(a). The process moves only by jumps, with infinitely many jumps in any time interval, however small. The irregular size of the jumps implies the very jagged shape of the picture.

4.2.2 Normal gamma process

If we assume $\sigma^2 \sim \Gamma(\nu, \gamma^2/2)$ then $Y \sim N\Gamma(\nu, \gamma, \beta, \mu)$, which we call the normal gamma (written $N\Gamma$) distribution. From the cumulant function

$$\mathrm{K}\left\{\theta \ddagger Y_1\right\} = \mu\theta - \nu \log\left(1 - \frac{\theta\beta + \theta^2/2}{\gamma^2/2}\right), \tag{24}$$

it follows that $Y_t \sim N\Gamma(t\nu, \gamma, \beta, t\mu)$. This means this process is analytically simple to handle. The density of the process at time one is

$$f_{Y_1}(y) = \frac{\gamma^{2\nu}\alpha^{1-2\nu}}{\sqrt{2\pi}\Gamma(\nu)2^{\nu-1}} \bar{K}_{\nu-1/2}\left(\frac{\gamma^2}{2}\left|y-\mu\right|\right) \exp\left\{\beta\left(y-\mu\right)\right\}, \quad \text{where} \quad \bar{K}_\nu(x) = x^\nu K_\nu(x).$$

Note that γ^{-2} is a scale parameter and that ν and β/γ are invariant under location-scale changes.

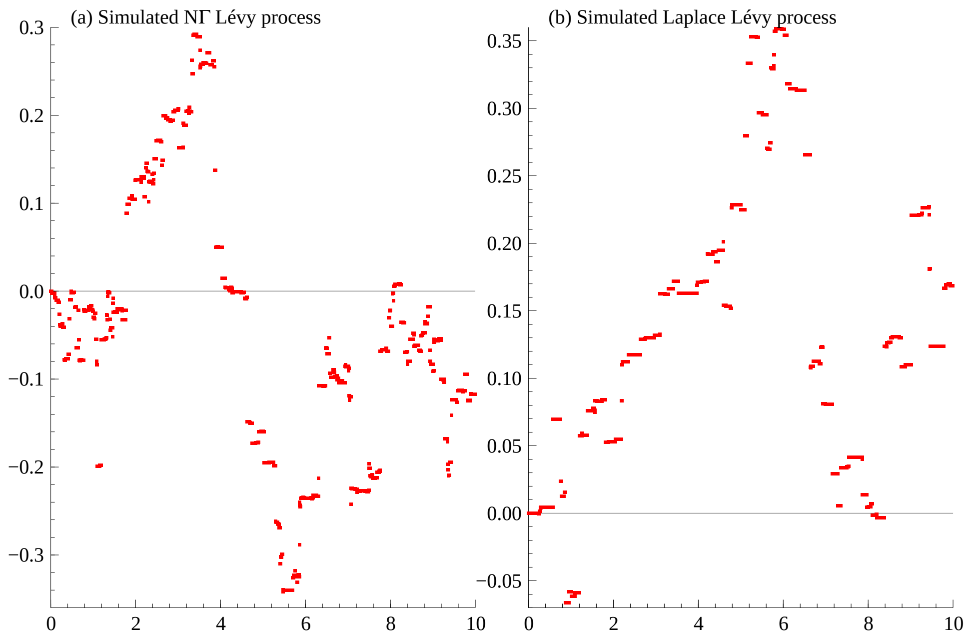

Figure 6: (a) *Sample path of a* $N\Gamma$*(4,200,0,0) Lévy process. Such processes are often called variance gamma processes in the literature. (b) Sample path of a La(0.2,0,0) Lévy process. Code:* `levy_graphs.ox`.

The special case of $\beta = 0$ (symmetry) has been used extensively in the finance literature where it is often called the *variance gamma* (VG) process. It was introduced by Madan and Seneta (1990), while the $\beta \neq 0$ case was developed by Madan, Carr, and Chang (1998). Later we will mention a generalisation of the normal gamma process, the extended Koponen (or KoBol) class, which is often referred to in the finance literature as the CMGY Lévy process.

Figure 6(a) graphs a simulated path from a $N\Gamma(4, 200, 0, 0)$ process. As we would expect, the sample path has some of the features of the NIG process we drew in Figure 3(b). In particular both of these infinite activity processes are very jagged. More detailed mathematical analysis of the corresponding Lévy measures shows that the NIG process has more very small jumps than the $N\Gamma$ process. In particular near zero the Lévy density $u(y)$ of the NIG process behaves like $y^{-3/2}$, while the corresponding result for the $N\Gamma$ process is y^{-1}.

An interesting feature of the normal gamma process is that it can be written in the form $Y_+ - Y_-$ where Y_+ and Y_- are independent Γ subordinators (see Exercise 8).

For comparison with the NIG log-density (see Figure 5), note that NΓ log-density either has a cusp at zero, when $\nu \in (0, 1]$, or is concave.

4.2.3 Hyperbolic, Laplace and skewed Student's t processes

If we assume $\sigma^2 \sim PH(\delta, \gamma)$ then $Y \sim H(\alpha, \beta, \mu, \delta)$, where $\alpha = \sqrt{\beta^2 + \gamma^2}$, has the hyperbolic distribution. This distribution can be shown to be infinitely divisible, although the proof of this is difficult — see Barndorff-Nielsen and Shephard (2012a). The hyperbolic process puts $Y_1 \sim H(\alpha, \beta, \mu, \delta)$, where the density is

$$f_{Y_1}(y) = \frac{\gamma}{2\sqrt{\beta^2 + \gamma^2}\delta K_1(\delta\gamma)} \exp\left\{-\alpha\sqrt{\delta^2 + (y-\mu)^2} + \beta(y-\mu)\right\}, \quad y \in \mathbb{R}. \tag{25}$$

All the moments of this process exist so long as $\gamma > 0$, while the cumulant function is

$$\mathrm{K}\{\theta \ddagger Y_1\} = \frac{1}{2}\log\left\{\frac{\gamma^2}{\alpha^2 - (\beta+\theta)^2}\right\} + \log\left\{\frac{K_1\left\{\delta\sqrt{\alpha^2 - (\beta+\theta)^2}\right\}}{K_1(\delta\gamma)}\right\} + \theta\mu. \tag{26}$$

We can compare the hyperbolic model to the NIG density using the shape triangle. In particular reparameterise into the location-scale invariant parameters given in (22), then Figure 7 shows the log-densities for this model. We see that again as $\xi \to 0$ we get the normal quadratic log-density. For higher values the log-density gets increasingly linear decay in the tails as $\xi \to 1$. Indeed in the limit we get the Laplace densities, see below. This contrasts with the NIG density which has the ability to have thicker tails. Hyperbolic laws have the interesting and important feature that the

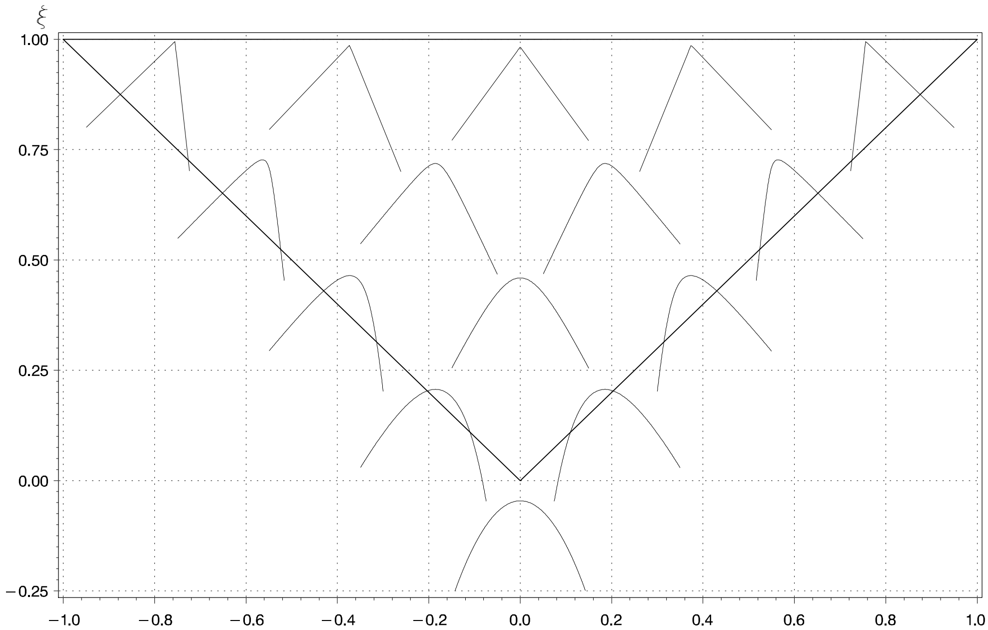

Figure 7: *Shape triangle for the hyperbolic model. That is we graph the shape of the log-density for the hyperbolic model for a variety of values of the steepness parameter ξ and the asymmetry parameter χ. This graph was kindly made available to us by Preben Blæsild.*

log-density is a hyperbola (hence their name) and so behaves approximately linearly in the tails of the distribution.

Hyperbolic Lévy processes have the disadvantage that we do not have an exact expression for the density of Y_t for $t \neq 1$, nor can we simulate from the process in a non-intensive manner. Both of these properties are inherited from the fact that this is also the case for the positive hyperbolic process we discussed in the previous Section.

The Laplace distributions (symmetric and asymmetric) occur as limiting cases of (25) for α, β and μ fixed and $\delta \downarrow 0$. We write this as $La(\alpha, \beta, \mu)$. The corresponding density is

$$\frac{\alpha^2 - \beta^2}{2\alpha} \exp\left\{-\alpha\left|y-\mu\right| + \beta\left(y-\mu\right)\right\}, \quad \text{where} \quad \alpha = \sqrt{\beta^2 + \gamma^2}, \tag{27}$$

which is achieved by $\sigma^2 \sim E(\gamma^2/2) = \Gamma(1, \gamma^2/2)$. One of the main features of this model is that $Y_t \sim N\Gamma(t, \gamma, \beta, t\mu)$.

Finally, if we assume $\sigma^2 \sim R\Gamma(\nu, \delta^2/2)$ then $Y \sim T(\nu, \delta, \beta, \mu)$, a skewed Student's t distribution, which is infinitely divisible and so can be used as the basis of a Lévy process. The skewed Student's t Lévy process puts $Y_1 \sim T(\nu, \delta, \beta, \mu)$, where the density is

$$\frac{\delta^{-2\nu}\left|\beta\right|^{1-2\nu}}{\sqrt{2\pi}\Gamma(\nu)2^{\nu-1}} \bar{K}_{\nu-1/2}\left\{\left|\beta\right|\sqrt{\delta^2 + \left(y-\mu\right)^2}\right\} \exp\left\{\beta\left(y-\mu\right)\right\},$$

where, for any $\nu \in \mathbb{R}$, $\overline{K}_\nu(x) = x^\nu K_\nu(x)$. The more familiar Student's t distribution is found when we let $\beta \to 0$, then the density becomes

$$\frac{\Gamma\left(\nu + 1/2\right)}{\sqrt{\delta^2\pi}\Gamma(\nu)}\left\{1 + \left(\frac{y-\mu}{\delta}\right)^2\right\}^{-\nu-1/2}.$$

In the symmetric case only moments of order less than ν will exist — at any time horizon. However, we do not know the distribution of Y_t for this process in any explicit form, while simulation has to be carried out in quite an involved manner. Hence this process is not as easy to handle as the NIG or normal gamma Lévy processes.

4.2.4 Generalized hyperbolic process

If we assume $\sigma^2 \sim GIG(\nu, \delta, \gamma)$ then $Y \sim GH(\nu, \alpha, \beta, \mu, \delta)$, where $\alpha = \sqrt{\beta^2 + \gamma^2}$, the *generalised hyperbolic* (GH) distribution. Its density is

$$\frac{\gamma^{2\nu}\alpha^{1-2\nu}}{\overline{K}_\nu\left(\delta\gamma\right)\sqrt{2\pi}} \overline{K}_{\nu-\frac{1}{2}}\left\{\alpha\sqrt{\delta^2 + \left(y-\mu\right)^2}\right\} \exp\left\{\beta\left(y-\mu\right)\right\}, \tag{28}$$

where K_ν is the modified Bessel function of the third kind and $\overline{K}_\nu(x) = x^\nu K_\nu(x)$. It is helpful to recall that $K_\nu(y) = K_{-\nu}(y)$. This distribution includes as special cases many of the above distributions in the following way:

$$\begin{array}{ll} N(\mu,\sigma^2)=\lim_{\gamma\to\infty}GH(\nu,\gamma,0,\mu,\sigma^2\gamma), & NIG(\alpha,\beta,\mu,\delta)=GH\left(-\frac{1}{2},\alpha,\beta,\mu,\delta\right), \\ NRIG(\alpha,\beta,\mu,\delta)=GH\left(\frac{1}{2},\alpha,\beta,\mu,\delta\right), & H(\alpha,\beta,\mu,\delta)=GH(1,\alpha,\beta,\mu,\delta), \\ T(\nu,\delta,\beta,\mu)=GH(-\nu,\beta,\beta,\mu,\delta), & La(\alpha,\beta,\mu)=GH(1,\alpha,\beta,\mu,0), \\ N\Gamma(\nu,\delta,\beta,\mu)=GH(\nu,\alpha,\beta,\mu,0), & RH(\alpha,\beta,\mu,\delta)=GH(-1,\alpha,\beta,\mu,\delta), \end{array}$$

for $\nu > 0$. The GH distribution is infinitely divisible (a proof is given in Barndorff-Nielsen and Shephard (2012a)) and so can be used as the basis of a rather general Lévy process. The special cases not introduced above are

- normal reciprocal inverse Gaussian distribution ($NRIG$), which happens when $\sigma^2 \sim RIG(\delta,\gamma)$.
- reciprocal hyperbolic (RH), which happens when $\sigma^2 \sim RPH(\delta,\gamma)$.

The cumulant function of the GH is

$$\mathrm{K}\left\{\theta \ddagger Y_1\right\} = \frac{\nu}{2}\log\left\{\frac{\gamma}{\alpha^2-(\beta+\theta)^2}\right\} + \log\left\{\frac{K_\nu\left\{\delta\sqrt{\alpha^2-(\beta+\theta)^2}\right\}}{K_\nu\left\{\delta\sqrt{\alpha^2-\beta^2}\right\}}\right\} + \theta\mu, \quad |\beta+\theta|<\alpha,$$

while the first two moments (when they exist) are

$$\begin{aligned} \mathrm{E}(Y_1) &= \mu+\beta\frac{\delta}{\gamma}\frac{K_{\nu+1}(\delta\gamma)}{K_\nu(\delta\gamma)} \quad \text{and} \\ \mathrm{Var}(Y_1) &= \delta^2\left(\frac{K_{\nu+1}(\delta\gamma)}{\delta\gamma K_\nu(\delta\gamma)}+\frac{\beta^2}{\gamma^2}\left[\frac{K_{\nu+2}(\delta\gamma)}{K_\nu(\delta\gamma)}-\left\{\frac{K_{\nu+1}(\delta\gamma)}{K_\nu(\delta\gamma)}\right\}^2\right]\right). \end{aligned}$$

Not surprisingly, in general we do not know the GH density of Y_t for $t \neq 1$, nor can we simulate from the process in a non-intensive manner. This model is so general that it is typically difficult to manipulate mathematically and so is not often used empirically. Instead special cases are usually employed.

4.2.5 Normal positive stable and symmetric stable processes

If we assume $\sigma^2 \sim PS(\alpha/2,\delta)$ then Y has a normal stable distribution, $S(\alpha,\beta,\mu,\delta)$, which is infinitely divisible and so supports a Lévy process. The important special case where $\mu=\beta=0$ is the well known symmetric stable Lévy process with index α. Except for the boundary case of $\alpha=2$, the symmetric stable distribution has the empirically unappealing feature that the variance of Y_1 is infinity. The density of this variable is unknown in general, with exceptions being the Gaussian variable ($\alpha=2$) and the Cauchy variable ($\alpha=1$).

Despite the complexity of the density of a stable random variable the cumulant function is simply

$$\mathrm{C}\left\{\zeta \ddagger Y_1\right\} = \delta\left|\zeta\right|^\alpha,$$

which implies $Y_t \sim S(\alpha, 0, 0, t\delta)$. Note that for this process the mean exists only if $\alpha > 1$. The Lévy density for a symmetric stable process is given by

$$u(y) = \delta \left|y\right|^{-1-\alpha}, \quad 0 < \alpha < 2. \tag{29}$$

Stable Lévy processes have the remarkable property that for $\lambda > 0$

$$\{Y_{\lambda t}\}_{t \geq 0} \stackrel{L}{=} \left\{\lambda^{1/\alpha} Y_t\right\}_{t \geq 0}. \tag{30}$$

Thus, in particular, increments of the Lévy process over time λt are, in distribution, just scaled versions of increments over time t. This fractal like property is called *self-similarity*, and stable Lévy processes (symmetric or not) are the only Lévy processes which possess this feature.

Although stable processes have received considerable attention in financial economics since their introduction into that subject in the early 1960s, it has been known since the late 1960s that they provide a poor fit to the empirical data we usually see in practice. This is because returns over long time intervals, which are sums of returns over finer time intervals, tend to be more Gaussian than the ones over short horizons. Hence our interest in this type of process will usually be to provide theoretical illustrations, rather than as practical models.

4.2.6 Normal tempered stable process

If we assume $\sigma^2 \sim PTS(\kappa, \delta, \gamma)$ then $Y \sim NTS(\kappa, \delta, \gamma, \beta, \mu)$, a normal tempered stable distribution. The NTS distribution is infinitely divisible and so can be used to generate a Lévy process. Special cases of the normal tempered stable process is the NIG Lévy process and the normal gamma Lévy process (when $\kappa \downarrow 0$). This process will be discussed in more detail in Example 5.2.3.

4.3 Lévy-Khintchine representation

The Lévy-Khintchine representation for positive variables given in (16) can be generalised to cover Lévy processes with increments on the real line. Four basic developments are needed. First, the Lévy measure must be allowed to have support on the real line, not just the positive half-line, but still excluding the possibility that the measure has an atom at zero. Second, the parameter a needs to be allowed to be a real, not just positive. Third, we imagine that an independent Brownian motion component is added to the process. Fourth, the below technical condition on the Lévy measure (32) must be satisfied. The result is the celebrated Lévy-Khintchine representation for Lévy processes.

Theorem 3 *Lévy-Khintchine representation. Suppose Y is a Lévy process. Then*

$$\mathrm{C}\{\zeta \ddagger Y_1\} = ai\zeta - \frac{1}{2}\sigma^2\zeta^2 + \int_{\mathbb{R}} \left\{e^{i\zeta y} - 1 - i\zeta y \mathbf{1}_{[-1,1]}(y)\right\} v(\mathrm{d}y), \tag{31}$$

where $a \in \mathbb{R}$, $\sigma \geq 0$, and the Lévy measure v must satisfy

$$\int_{\mathbb{R}} \min\{1, y^2\} v(\mathrm{d}y) < \infty \tag{32}$$

and v has no atom at 0.

Lévy processes are completely determined by the *characteristic triplet* (a, σ^2, v): the drift a, the variance σ^2 of the Brownian motion and the Lévy measure v (which has to satisfy (32)). Importantly only processes with $v = 0$ do not have jumps — but in that case Y is a scaled Brownian motion with drift a. The representation implies that all Lévy processes can be decomposed into

$$Y_t = at + \sigma W_t + L_t^d,$$

a drift, a scaled Brownian motion W and an independent pure jump process L_t^d.

There are many interesting additional features of the Lévy-Khintchine representation which we bring out in Barndorff-Nielsen and Shephard (2012a). Important points are that the centring function $\mathbf{1}_{[-1,1]}(y)$ is one choice amongst many and that many interesting properties of the distribution of Y_1 can be deduced from direct inspection of v (e.g. existence of moments and unimodality).

4.4 Blumenthal-Getoor index

A key measure of the degree of variation of Lévy processes is the Blumenthal-Getoor index α defined by

$$\alpha = \inf\left\{\beta : \int_{\mathbb{R}} \left(1 \wedge |y|^\beta\right) v(\mathrm{d}y) < \infty\right\}.$$

This index has the property that

$$p - \lim_{n \to \infty} \sum_{i=1}^{n} \left| Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}} \right|^\delta < \infty$$

for any $\delta > \alpha$.

Clearly for any CPP we have $\alpha = 0$; but what about infinite activity processes? In general we know $\alpha \leq 2$ as the limit exists for quadratic variation, but more generally? The Blumenthal-Getoor index is determined by the behaviour of the Lévy measure near zero, with larger indexes arising when the Lévy density goes to infinity quickly as $|y| \to 0$. Informally the higher the index the larger the frequency of small jumps in the Lévy process. Some examples of the index are given in Table 1.

4.5 Power variation

Related to the Blumenthal-Getoor index is power variation, which is

$$\{Y\}^{[r]} = p - \lim_{n \to \infty} n^{r/2-1} \sum_{i=1}^{n} \left| Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}} \right|^r.$$

name	w(y)	B-G index
IG	$cy^{-3/2}\exp(-y)$	1/2
Γ	$cy^{-1}\exp(-y)$	0
PTS	$cy^{-1-\alpha}\exp(-y)$	α
NIG	$\approx c\,\lvert y\rvert^{-2}$ for small y	1
NΓ	$\approx c\,\lvert y\rvert^{-1}$ (for small y)	0
PS	$c\,\lvert y\rvert^{-1-\alpha}$	α

Table 1: *Blumenthal-Getoor index for various infinitely divisible distributions.*

In a moment we will see its most famous case where $r = 2$, which is quadratic variation

$$\{Y\}^{[r]} = \sigma^2 + \sum_{s \le 1} (\Delta Y_s)^2,$$

but the other cases are interesting too. In particular when $r < 2$ we have that

$$\{Y\}^{[r]} = \mu_r \sigma^r, \quad \mu_r = \mathrm{E}\,|U|^r, \quad U \sim N(0,1),$$

which reveals the volatility of the Brownian motion component of the Lévy process, while when $r > 2$ then power variation goes off to infinity. This suggests that sums of power of absolute returns maybe useful ways of assessing the importance of jumps in financial economics. This was first formalised in finance by Barndorff-Nielsen and Shephard (2004), although the mathematics behind this goes back further (see, for example, Jacod and Protter (2012)).

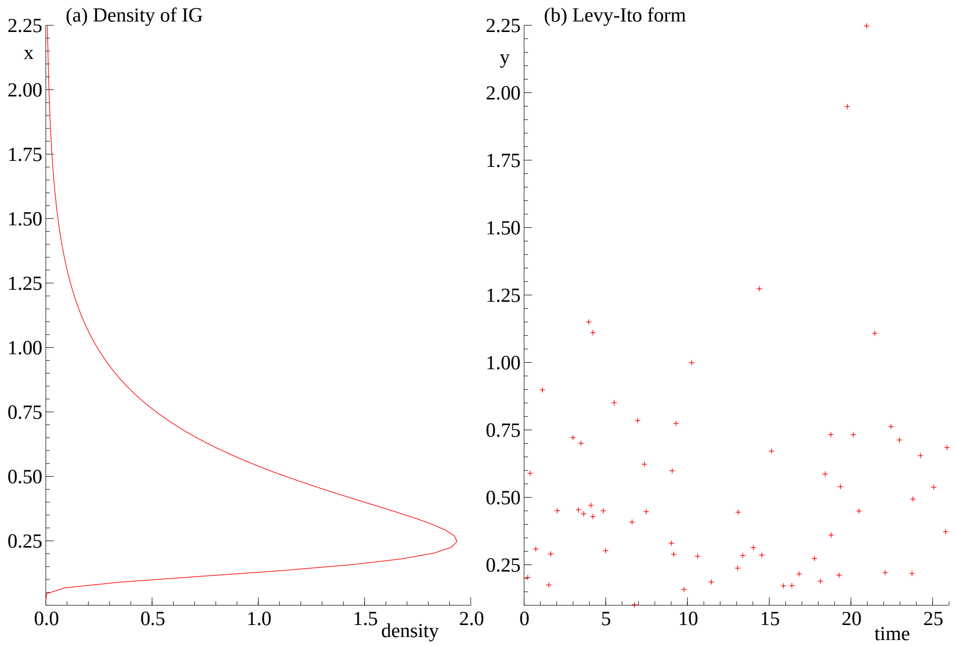

Figure 8: *(a) rotates the IG(1, 2) density for* χ_1*. (b) displaying the Poisson basis* $N(\mathrm{d}y, \mathrm{d}t)$*. Filename is* `levy_graphs.ox`*.*

4.6 Lévy-Ito representation

An important application of Lévy bases is in the Lévy-Ito representation of Lévy processes.

Theorem 4 Lévy-Ito representation[1]. *Let L be a Lévy process with Lévy measure v. Then L has the representation*

$$L_t = at + \sqrt{b}W_t + \int_0^t \int_{\mathbb{R}} y\{N(\mathrm{d}y, \mathrm{d}s) - \mathbf{1}_{[-1,1]}(y)v(\mathrm{d}y)\mathrm{d}s\}. \tag{33}$$

where W is a Brownian motion and N is a Poisson basis on $\mathbb{R}_{>0} \times \mathbb{R}$, independent of W and with mean measure $v(\mathrm{d}y)\mathrm{d}t$.

That N is a Poisson basis means that it is a random measure which attaches a random number $N(A)$ to any (bounded Borel) subset of the underlying space, in this case $\mathbb{R}_{>0} \times \mathbb{R}$, and that the numbers are independent for disjoint subsets. Furthermore, since we are talking about a Poisson basis the law of $N(A)$ is Poisson with mean equal to the integral over A of the measure $v(dy)\mathrm{d}t$.

The representation may be given the alternative form

$$\begin{aligned} L_t &= at + \sqrt{b}W_t \\ &\quad + \int_{\{|y|<\varepsilon\}} y\{N_t(\mathrm{d}y) - tv(\mathrm{d}y)\} \\ &\quad + \int_{\{|y|\geq\varepsilon\}} yN_t(\mathrm{d}y), \end{aligned}$$

where

$$N_t(\mathrm{d}y) = \int_0^t N_{\mathrm{d}y,\mathrm{d}s}$$

and ε is an arbitrarily chosen positive number (here a will generally depend on ε.).

This is insightful in a number of ways. For example, it demonstrates that Lévy processes are always semimartingales for we can decompose

$$A_t = at + \sum_{0<s\leq t} \mathbf{1}_{\{|\Delta L_s|\geq\varepsilon\}}\Delta L_s \qquad \text{and} \qquad M_t = \sqrt{b}W_t + \int_{\{|y|<\varepsilon\}} y\{N_t(\mathrm{d}y) - tv(\mathrm{d}y)\}.$$

The latter is obviously a martingale, while A is of finite variation (since the number of jumps ΔL_s of absolute size $\geq \varepsilon$ is locally finite, as follows from the fact that $\int_{\{|y|\geq\varepsilon\}} v(\mathrm{d}y) < \infty$.)

The mean value of L_1 exists if and only if (cf. Sato (1999, p. 39))

$$\int_{|y|>1} |y|v(\mathrm{d}y) < \infty.$$

In this case the representation (33) can be recast in the form

$$L_t = \mathrm{E}(L_t) + \sqrt{b}W_t + \int_0^t \int_{\mathbb{R}} y\{N(\mathrm{d}y, \mathrm{d}s) - v(\mathrm{d}y)\mathrm{d}s\}. \tag{34}$$

[1]For a proof see, for instance, Sato (1999, Ch. 4).

If, moreover, the Lévy measure is a member of the infinite activity finite variation (denoted $\mathcal{IAFV}$) class then

$$L_t = a_0 t + \int_0^t \int_{\mathbb{R}} yN(\mathrm{d}y, \mathrm{d}s), \tag{35}$$

where the *drift* a_0 is given by

$$a_0 = \mathrm{E}(L_1) - \int_{\mathbb{R}} yv(\mathrm{d}y).$$

Recall that finite variation means here that the sums of the absolute values of the infinitesimal increments of the process are bounded with probability one.

Any non-negative Lévy process L without drift is representable in the Lévy-Ito form

$$L_t = \int_0^t \int_0^\infty yN(\mathrm{d}y, \mathrm{d}s) \tag{36}$$

with v satisfying

$$\int_0^\infty \min\{1, y\}v(\mathrm{d}y) < \infty.$$

5 Time deformation and time-change

5.1 Basic framework

Financial markets sometimes seem to move more rapidly than in other periods. One way of starting to model this is to allow the relationship between calendar time and the pace of the market to be random. We call a stochastic process which models the random clock a *time-change*, while the resulting process is said to be *time deformed*.

Definition 5 *A time-change is any non-decreasing random process T with $T_0 = 0$. The special case where the time-change has independent and stationary increments is called a subordinator.*

The requirement that the time-change is non-decreasing rules out the chance that time can go backwards. A special case of a time-change is a subordinator, while subordinators are special cases of Lévy processes (e.g. Poisson or IG Lévy processes are subordinators). All subordinators are pure upward jumping processes plus non-negative drift. We should note here that the finance literature often labels time-changes subordinators, while in the probability literature the term subordinator is reserved for positive Lévy processes and we stick to the latter convention in this book.

In this Section we will study what happens when a subordinator is used to change the clock, that is deform, a stochastically independent Lévy process L. The result is

$$Y_t = L_{T_t},$$

which is often written as

$$Y_t = L \circ T_t,$$

or

$$Y = L \circ T.$$

The increments of this process are

$$Y_{t+s} - Y_t = L \circ T_{t+s} - L \circ T_t,$$

which are independent and stationary and so Y is a Lévy process.

Brownian motion is the only Lévy process with continuous sample paths, however this property does not survive being deformed by a pure subordinator for such a process has to be a pure jump process.

5.2 Examples

5.2.1 Compound Poisson process

Let T be a Poisson process, independent of a scaled Brownian motion W and so that

$$Y = \beta T + \sigma W \circ T,$$

then $Y_1|T_1 \sim N(\beta T_1, \sigma^2 T_1)$. This is a compound Poisson process

$$Y_t = \sum_{j=1}^{T_t} C_j, \quad C_j \overset{i.i.d.}{\sim} N(\mu, \sigma^2), \tag{37}$$

with shocks which are Gaussian. The implication is that when the process jumps, the jumps are independent of the time we have waited until the jump.

5.2.2 Normal inverse Gaussian process

Suppose T is an $IG(\delta, \gamma)$ Lévy process and $Y_t = \mu t + \beta T_t + W_{T_t}$, where W is Brownian motion, then

$$Y_1|T_1 \sim N(\mu + \beta T_1, T_1)$$

and so unconditionally the increments are independent with

$$Y_1 \sim NIG(\alpha, \beta, \mu, \delta), \quad \alpha^2 = \beta^2 + \gamma^2.$$

Hence this deformed Brownian motion is the NIG Lévy process, which we simulated in Figure 2(b). Likewise the $N\Gamma$ Lévy process can be obtained by subordinating Brownian motion plus drift by a gamma process.

5.2.3 Normal tempered stable process

Suppose T is a $PTS(\kappa, \delta, \gamma)$ Lévy process. If $Y_t = \mu t + \beta T_t + W_{T_t}$, then Y is called a *normal tempered stable* (NTS) Lévy process. We write

$$Y_1 \sim NTS(\kappa, \alpha, \beta, \mu, \delta), \quad \text{where} \quad \alpha = \sqrt{\beta^2 + \gamma^2},$$

but the corresponding probability density is generally unknown (except for an infinite series representation, see Feller (1971b, p. 583)). The cumulant function, on the other hand, is rather simple

$$\mathrm{K}\{\theta \ddagger Y_1\} = \mu\theta + \delta\gamma - \delta\left\{\alpha^2 - (\beta + \theta)^2\right\}^{\kappa}.$$

The form of this function implies $Y_t \sim NTS(\kappa, \alpha, \beta, \mu t, \delta t)$. It can be shown (after some considerable work), using the cumulant function of the NTS process, that the Lévy density is

$$u(y) = \frac{\delta}{\sqrt{2\pi}} \frac{\kappa 2^{\kappa+1}}{\Gamma(1-\kappa)} \alpha^{\kappa+\frac{1}{2}} |y|^{-\left(\kappa+\frac{1}{2}\right)} K_{\kappa+\frac{1}{2}}(\alpha |y|) \exp\{\beta y\}.$$

The direct use of this Lévy density is obviously going to be difficult due to its complexity. The deformation interpretation will mean that we can usually sidestep this, instead employing the simple Lévy density of the $PTS(\kappa, \delta, \gamma)$ Lévy process given in (18).

5.2.4 Type G and P Lévy processes

Definition 6 *We call Lévy processes which can be written as $Y_t = \mu t + \beta T_t + W_{T_t}$, for some subordinator T, type G Lévy processes.*

This is the subset of Lévy processes for which there is a deformation of Brownian motion with drift interpretation.

Definition 7 *We call Lévy processes which can be written as $Y_t = N_{T_t}$, for some subordinator T and where N is a Poisson process, type P Lévy processes.*

This is the subset of Lévy processes for which there is a deformation of Poisson process representation. We will now give some examples of this.

5.2.5 Negative binomial process

Suppose T is a $\Gamma(\nu, \alpha)$ Lévy process and let $Y_t = N_{T_t}$, where N is a standard Poisson process with unit intensity and $N \perp\!\!\!\perp T$. Then Y is a negative binomial Lévy process with Y_t having a negative binomial distribution

$$NB\left(t\nu, \frac{1}{1+\alpha}\right),$$

which follows from the well known Poisson-gamma mixture. Of course if $\nu = \lambda\alpha$ and $\alpha \to \infty$ then $T_t \to \lambda t$ and so Y_t becomes like a Poisson process with intensity λ. For this model

$$\begin{aligned}\overline{\mathrm{K}}\left\{\theta \ddagger Y_1\right\} &= \log \mathrm{E}\left[\exp\left\{\psi(e^{-\theta}-1)\right\}\right], \quad \text{where} \quad \psi \sim \Gamma(\nu,\alpha), \\ &= \mathrm{K}\left\{\left(e^{-\theta}-1\right) \ddagger \psi\right\} \\ &= -\nu \log\left(1+\frac{1-e^{-\theta}}{\alpha}\right),\end{aligned}$$

which is a reparameterisation of (6). Then $\mathrm{E}(Y_1) = \nu/\alpha$.

5.2.6 Discrete Laplace process

A special case of this is where T is an exponential $\Gamma(1,\alpha)$ Lévy process. Then Y_1 has a geometric probability function

$$\Pr(Y_1 = y) = p\,(1-p)^y\,, \quad y = 0,1,2,..., \quad p = \frac{1}{1+\alpha},$$

while generally $Y_t \sim NB\left(t, \frac{1}{1+\alpha}\right)$. Further let $Y_t = N^{(1)}_{T^{(1)}_t} - N^{(2)}_{T^{(2)}_t}$, where $N^{(1)}, N^{(2)}, T^{(1)}$ and $T^{(2)}$ are independent processes, the first two of which are Poisson processes and the last two exponential processes. Then Y_t is a discrete Laplace Lévy process while

$$\Pr(Y_1 = y) = \frac{1-p}{1+p}p^{|y|}, \quad y = 0, \pm 1, \pm 2, ... \tag{38}$$

This has a zero mean and variance of $2p/(1-p)^2$. It is a heavier tailed alternative to the Skellam distribution discussed in Section 4.1.4, but again it only has one parameter.

5.2.7 Poisson-IG process

More generally if T is a GIG Lévy process and we let $Y_t = N_{T_t}$, then we call the result a Poisson-GIG Lévy process. The most important special case is the gamma one just discussed and the $IG(\delta,\gamma)$ case. Writing $\gamma^{2*} = \gamma^2 + 2$, the P-IG$(\delta,\gamma)$ variable has

$$\begin{aligned}\Pr(Y_1 &= y) = \int_0^\infty \frac{e^{-\psi}\psi^y}{y!}\frac{\delta}{\sqrt{2\pi}}e^{\delta\gamma}\psi^{-3/2}\exp\left\{-\frac{1}{2}\left(\delta^2\psi^{-1}+\gamma^2\psi\right)\right\}\mathrm{d}\psi, \\ &= \frac{\delta}{y!\sqrt{2\pi}}e^{\delta\gamma}\int_0^\infty \psi^{(y-\frac{1}{2})-1}\exp\left\{-\frac{1}{2}\left(\delta^2\psi^{-1}+\gamma^{2*}\psi\right)\right\}\mathrm{d}y, \\ &= \frac{2\delta\sqrt{\gamma^*/\delta}}{\sqrt{2\pi}}e^{\delta\gamma}\frac{K_{y-\frac{1}{2}}\left(\delta\gamma^*\right)\left(\delta/\gamma^*\right)^y}{y!}, \quad y = 0,1,2,...,\end{aligned}$$

where $K_\nu(\cdot)$ is a modified Bessel function of the third kind. This follows from using the properties of the GIG distribution. The Y_t is marginally P-IG$(t\delta,\gamma)$.

Of course this is a finite activity process with

$$\overline{\mathrm{K}}\left\{\theta \ddagger Y_1\right\} = \overline{\mathrm{K}}\left\{\left(1-e^{-\theta}\right) \ddagger \psi\right\}, \quad \text{where} \quad \psi \sim IG(\delta,\gamma),$$

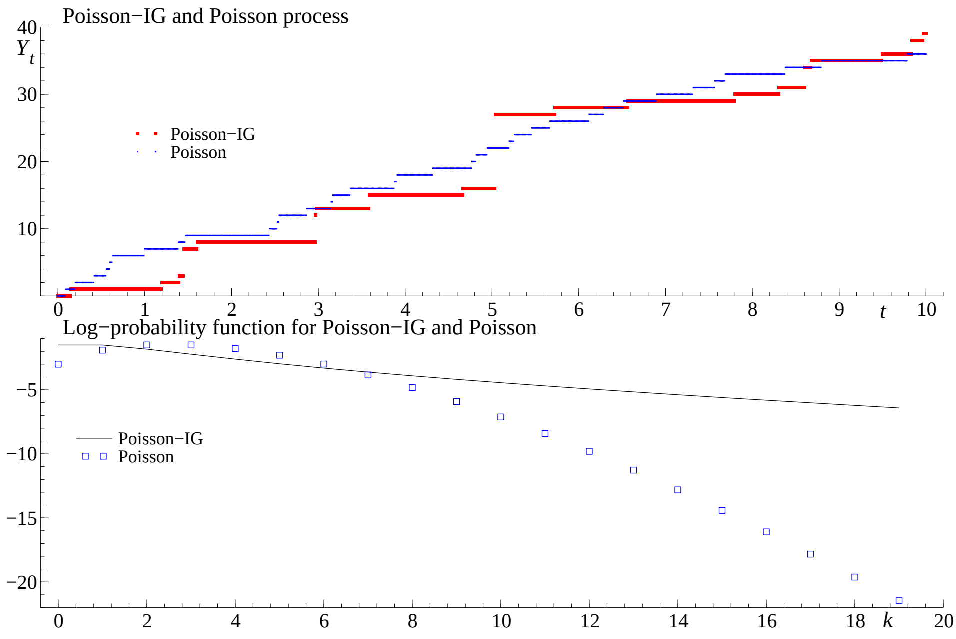

Figure 9: *Top graph: Simulation from a Poisson and Poisson-IG*$(1.5, 0.5)$ *process with mean intensity of 3. It shows the larger gaps in the P-IG process and larger jumps. Bottom graph: log of the probability function for* Y_1 *for the Poisson and P-IG distributions. It shows the approximate quadratic tails of the Poisson distribution and the approximate linear tails of the P-IG distribution. Code:* `poisson.ox`.

$$= \delta\left[\gamma - \left\{\gamma^2 + 2\left(1 - e^{-\theta}\right)\right\}^{1/2}\right].$$

Now $\mathrm{E}(Y_1) = \delta/\gamma$. For $\gamma \geq 1/2$ this can be thought of as a compound Poisson process with intensity δ and a discrete mixing distribution (5) for C_1 whose cumulant function is

$$\exp\left\{\overline{\mathrm{K}}\left(\theta \ddagger C_1\right)\right\} = 1 + \gamma - \left\{\gamma^2 + 2\left(1 - e^{-\theta}\right)\right\}^{1/2}, \quad \gamma \geq 1/2.$$

From this $\mathrm{E}(C_1) = 1/\gamma$.

The top part of Figure 9 shows a simulated sample path from a Poisson-IG process, it illustrates the longer gaps between arrivals and the occasional very large new arrival, compared to a Poisson process. The two processes are simulated to have the same mean, but the Poisson-IG process has $\delta = 1.5$ and $\gamma = 0.5$. The bottom part of the Figure show the log of the probability function for the Poisson-IG and Poisson distributions for Y_1, which has equal means of 3. It shows that the Poisson distribution decays approximately quadratically in the tails, while the Poisson-IG's decay is more like linear.

5.2.8 Time-changed Skellam processes

Suppose that

$$Y_t = \left(N^{(1)} - N^{(2)}\right)_{T_t} = N^{(1)}_{T_t} - N^{(2)}_{T_t},$$

where $N^{(i)}$ are independent Poisson processes with unit intensity and T_t is a gamma process. Now $\left(N^{(1)} - N^{(2)}\right)_t$ is marginally a Skellam variable with

$$\Pr(N_t^{(1)} - N_t^{(2)} = k) = e^{-2t} I_{|k|}(2t), \quad k = 0, \pm 1, \pm 2, ...,$$

whose cumulant function is

$$\mathrm{K}\left\{\theta \ddagger Y_t\right\} = t\left(e^{\theta} + e^{-\theta} - 2\right).$$

This means that for a random time change T independent of Y we have that

$$\mathrm{K}\left\{\theta \ddagger Y_{T_1}\right\} = \mathrm{K}\left\{e^{\theta} + e^{-\theta} - 2 \ddagger T_1\right\},$$

which has an analytic solution for all cases where the cumulant function of T_1 is known. Leading examples are the gamma and inverse Gaussian cases.

Although the cumulant function is simple the probability function is less easy to work with. For example in the Skellam-gamma case

$$\begin{aligned}\Pr(N_{T_1}^{(1)} - N_{T_1}^{(2)} &= k) = \frac{\alpha^{\nu}}{\Gamma(\nu)} \int_0^{\infty} e^{-2\psi} I_{|k|}(2\psi) \psi^{\nu-1} e^{-\alpha\psi} \mathrm{d}\psi \\ &= \frac{\alpha^{\nu}}{\Gamma(\nu)} \int_0^{\infty} I_{|k|}(2\psi) \psi^{\nu-1} e^{-(\alpha+2)\psi} \mathrm{d}\psi.\end{aligned}$$

An analytic expression for this integral does not seem to be available.

6 Empirical estimation and testing of GH Lévy processes

6.1 A likelihood approach

6.1.1 Estimation of GH Lévy processes

Here we will assess how well Lévy processes fit the marginal distribution of financial returns. Their flexibility allows important improvements over conventional Brownian motion, however they clearly neglect the dynamics of returns.

Throughout we assume that Y is observed at unit time intervals which we will think of as representing a day. Then inference will be based on the increments

$$y_i = Y_i - Y_{i-1}, \quad i = 1, 2, ..., n.$$

Under the Lévy assumption

$$f(y_1, ..., y_n) = \prod_{i=1}^{n} f_{Y_1}(y_i),$$

where f_{Y_1} is the density of Y_1.

Typically f_{Y_1} is indexed by some parameters which are written as θ. Inference on θ is carried out using the likelihood

$$\log f(y_1, ..., y_n; \theta) = \sum_{i=1}^{n} \log f_{Y_1}(y_i; \theta).$$

For concreteness we will now focus on the case where Y is a GH Lévy process. Then $f_{Y_1}(y; \theta)$ is

$$\frac{\gamma^{2\nu}\alpha^{1-2\nu}}{\overline{K}_{\nu}(\delta\gamma)\sqrt{2\pi}}\overline{K}_{\nu-\frac{1}{2}}\left\{\alpha\sqrt{\delta^2 + (y-\mu)^2}\right\}\exp\left\{\beta(y-\mu)\right\}.$$

The parameters of the GH distribution are $\theta = (\nu, \mu, \beta, \delta, \gamma)'$, where $\alpha = \sqrt{\beta^2 + \gamma^2}$. The maximum likelihood (ML) estimator of θ is given by

$$\widehat{\theta} = \underset{\theta}{\arg} \max \log f(y_1, ..., y_n; \theta),$$

which has to be determined by numerical optimisation, either directly or via the EM algorithm. The latter is particularly effective in the multivariate case and is used in Section 9.2.2 while the associated theory is detailed in Barndorff-Nielsen and Shephard (2012c). In the case where Y is univariate the optimisation is carried out using the Broyden, Fletcher, Goldfard and Shanno (BFGS) quasi-Newton algorithm made available in the matrix programming language `Ox` by Doornik (2001) and takes a handful of seconds for sample sizes of around $3,000$.

Confidence intervals for the parameters can be constructed via the asymptotic distribution of $\widehat{\theta}$ based on the Lévy assumption of *i.i.d.* increments. As already mentioned, the independence assumption is clearly unrealistic. In a later subsection we will discuss the impact of this misspecification on confidence intervals. For now we stand by the Lévy assumption.

The asymptotic theory for ML estimators means that

$$\sqrt{T}\left(\widehat{\theta} - \theta\right) \xrightarrow{\mathcal{L}} N(0, \mathcal{I}^{-1}), \quad \text{as} \quad n \to \infty, \tag{39}$$

where $\mathcal{I}$ is the expected information per observation which is

$$\mathcal{I} = -\mathrm{E}\left(\frac{\partial S_i}{\partial \theta'}\right) = \mathrm{Cov}\left(S_i\right), \quad \text{where} \quad S_i = \frac{\partial \log f(y_i; \theta)}{\partial \theta}. \tag{40}$$

For inference expected information is usually replaced by averaged observed quantities

$$\mathcal{I}_S = -\frac{1}{n}\sum_{i=1}^{n}\frac{\partial S_i}{\partial \theta'} \quad \text{or} \quad \mathcal{I}_O = \frac{1}{n}\sum_{i=1}^{n} S_i S_i'. \tag{41}$$

The terms S_i and $\partial S_i / \partial \theta'$ are found by numerical differentiation. This allow us to construct asymptotically valid t statistics for elements of $\widehat{\theta} - \theta$. In particular a 95% asymptotic confidence interval for ν can be found as

$$\widehat{\nu} \pm 1.96\sqrt{\frac{1}{n}\left(\mathcal{I}^{-1}\right)_{\nu\nu}}, \tag{42}$$

	MLE of GH parameters					Likelihoods		
	μ	β	γ	δ	ν	GH	$\beta=0$	$N(\mu,\sigma^2)$
ML estimates	-0.0133	0.179	0.419	1.95	-1.62	-626.74	-627.97	-819.32
Outer product (I_O)	(.0088)	(.113)	(.053)	(.43)	(.71)			
Second derivative (I_S)	(.0087)	(.115)	(.049)	(.39)	(.65)			
Robust: m=250	(.013)	(.14)	(.071)	(.39)	(.74)			
Robust: m=500	(.016)	(.17)	(.076)	(.39)	(.76)			

Table 2: *ML estimates of GH for the Canadian daily exchange rate. Brackets are the asymptotic standard errors computed using different estimates of the expected information:* $\mathcal{I}_O$ and $\mathcal{I}_S$. *GH column denotes the likelihood for the unrestricted model.* $\beta = 0$ *imposes symmetry. Robust, denotes robust standard errors computed using m lags, which will be explained in a moment.*

where $\left(\mathcal{I}^{-1}\right)_{\nu\nu}$ denotes the diagonal element of $\mathcal{I}^{-1}$ corresponding to ν.

To illustrate the above methods we go back to the daily exchange rate series for the Canadian Dollar rate against the US Dollar. The results are given in Table 2. The ML estimate of ν is quite negative, while μ and β are close to zero. The asymptotic standard errors for β and ν are quite large and suggest both μ and β are not significantly different from zero. Interestingly the standard errors based on $\mathcal{I}_O$ and $\mathcal{I}_S$ are very similar indeed.

The Table also gives the likelihood when β is constrained to be zero. The likelihood drops by around one, which again suggests β can be set to zero. Finally the Table shows the *GH* model improves upon the Gaussian likelihood fit by around 193, which is a very large improvement.

6.1.2 Confidence intervals via profile likelihoods

An alternative way of quantifying uncertainty is based on the likelihood ratio statistic. Again suppose our focus is on ν. Define $\omega = (\mu, \beta, \delta, \gamma)'$, so that $\theta = (\nu, \omega')'$, and

$$\widehat{\omega_\nu} = \arg\max_{\omega} \log f(y_1, ..., y_n; \nu, \omega).$$

$\widehat{\omega_\nu}$ is a constrained ML estimator of ω, imposing on θ an a priori fixed value of ν. Likelihood theory tells us that if we constrain ν correctly then the *likelihood ratio statistic*

$$2\left\{\log f(y_1, ..., y_n; \widehat{\theta}) - \log f(y_1, ..., y_n; \nu, \widehat{\omega_\nu})\right\} \xrightarrow{\mathcal{L}} \chi_1^2, \quad \text{as} \quad n \to \infty,$$

then the ratio should typically take on unusually large values. We will be plotting the *profile likelihood*

$$\log f(y_1, ..., y_n; \nu, \widehat{\omega_\nu}) - \log f(y_1, ..., y_n; \widehat{\theta}) \quad \text{against} \quad \nu$$

to indicate plausible values of ν. Of course, while the χ_1^2 distribution is only valid if the model is accurate, which certain features are demonstrably not, it is still interesting to plot out the profile.

The top left graph in Figure 10 draws the profile likelihood function for ν for the Canadian Dollar example. This gives a similar result to the t statistics given in Table 2 with ranges of

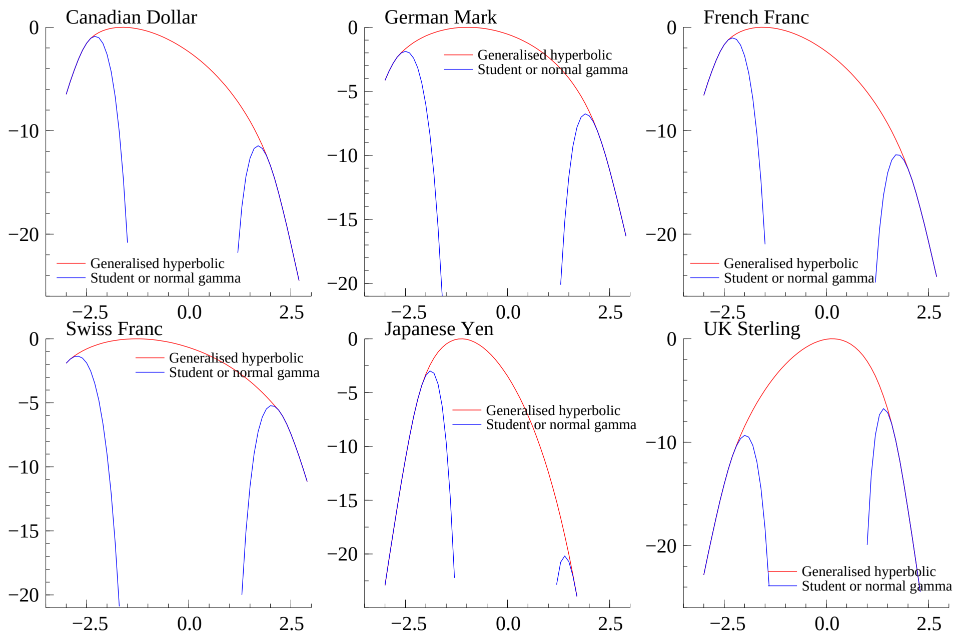

Figure 10: *Daily exchange rate data. Profile likelihood (truncated at -25) for the ν parameter of the GH. Also profile likelihood for the skewed Student ($\nu < 0$) and the normal gamma ($\nu > 0$). NIG case corresponds to the generalised hyperbolic curve at the point $\nu = -0.5$.*

approximately -2.5 to -0.2 supported by the data. The Figure also shows the profile likelihoods for the normal gamma and skewed Student's t, special cases of the *GH* model. Recall in the normal gamma model δ is set to zero, while in the skewed Student's t case $\gamma = 0$. Of course the likelihoods for these models cannot exceed that of the *GH* model, but this plot shows how far these models fall behind the *GH* model. We can see that for very negative ν the likelihood for the *GH* model is the same type of that as the skewed Student's t model, for the ML of γ turns out to be zero. The same effect can be seen for large values of ν for then the ML of δ is zero. The Figure shows that the skewed Student's t model performs quite well, but the normal gamma process has some very significant difficulties.

6.2 Model misspecification: robust standard errors

Lévy processes allow us to flexibly model the distribution of i.i.d. increments, however in financial economics returns exhibit volatility clustering. Later models will be developed which deal with this feature, but for now our Lévy models are misspecified. Even though our models are incorrect, estimation by ML methods makes sense. We will now be modelling the marginal distribution of the increments.

The theory of estimating equations implies

$$\sqrt{n}\left(\widehat{\theta} - \theta\right) \xrightarrow{\mathcal{L}} N\left(0, \mathcal{I}^{-1}\mathcal{J}\mathcal{I}^{-1}\right), \quad \text{as} \quad n \to \infty, \tag{43}$$

where

$$\mathcal{J} = \lim_{n\to\infty} \frac{1}{n}\mathrm{Cov}\left(\sum_{i=1}^{n} S_i\right), \quad \text{and} \quad \mathcal{I} = -\frac{1}{n}\mathrm{E}\left(\sum_{i=1}^{n} \frac{\partial S_i}{\partial \theta'}\right).$$

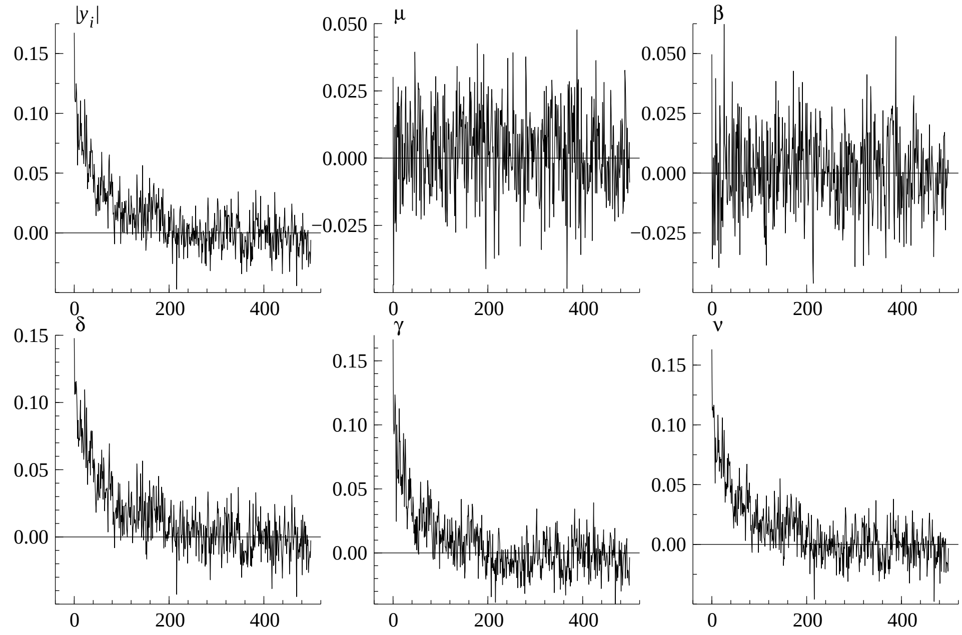

Figure 11: *Correlograms for the Canadian data drawn against lag length. Top left is of* $|y_i|$*, other graphs are for the five elements of* S_i*. Code:* `em_gh.ox`.

In order to construct robust standard errors for $\widehat{\theta}$ we estimate the elements of the *sandwich* $\mathcal{I}^{-1}\mathcal{J}\mathcal{I}^{-1}$. The empirical average $\mathcal{I}_S$ will consistently estimate $\mathcal{I}$ so long as the process is ergodic. However, unless the returns, and so the scores S_i, are i.i.d. $\mathcal{I}_O$ will not correctly estimate $\mathcal{J}$. Figure 11 shows the correlograms in the Canadian Dollar case (evaluated at $\widehat{\theta}$) of the elements of the score and $|y_i|$. It shows the scores for ν, δ and γ have correlograms which are close to that of $|y_i|$. The scores for μ and β are much less dependent.

There is a large literature on estimating $\mathcal{J}$ in the presence of autocorrelation. $\mathcal{J}$ is just the zero frequency of the spectral matrix of the vector S_i process, i.e.

$$\mathcal{J} = \mathrm{Cov}\left(S_i\right) + \sum_{s=1}^{\infty}\left\{\mathrm{Cov}\left(S_i, S_{i-s}\right) + \mathrm{Cov}(S_{i-s}, S_i)\right\},$$

which can be estimated by

$$\mathcal{J}_O = \frac{1}{n}\sum_{i=s+1}^{n} S_i S_i' + \sum_{s=1}^{m} K(j;m)\left\{\frac{1}{n}\sum_{i=s+1}^{n} S_i S_{i-s}' + \frac{1}{n}\sum_{i=s+1}^{n} S_{i-s} S_i'\right\},$$

where $K(j;m)$ denotes a non-negative Bartlett smoothing window

$$K(j;m) = \begin{cases} 1 - \left|\frac{j}{m+1}\right|, & \left|\frac{j}{m+1}\right| \leq 1, \\ 0, & \left|\frac{j}{m+1}\right| > 1, \end{cases}$$

while m is called the lag truncation parameter.

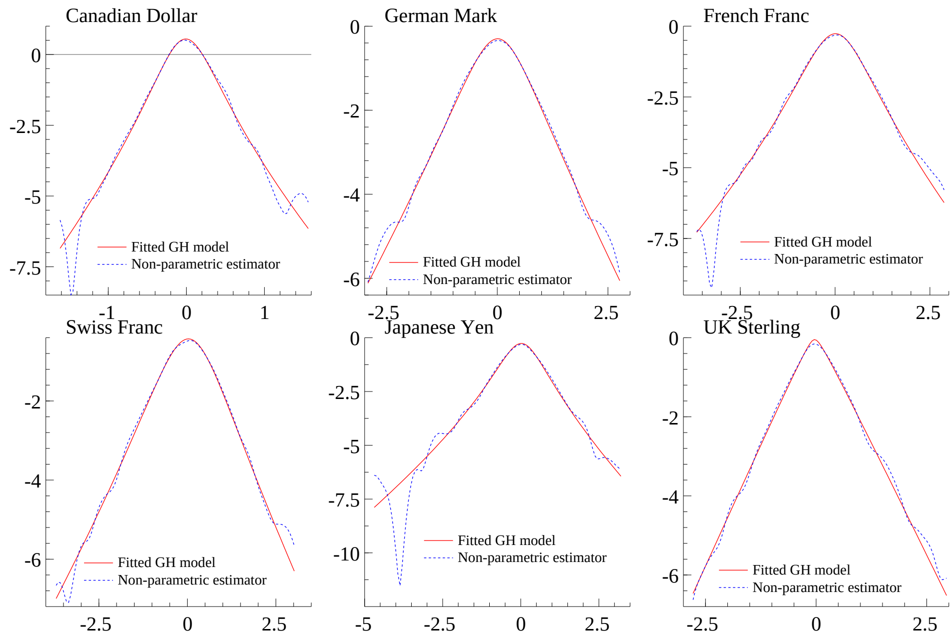

Figure 12: *Log of the estimates of the unconditional density of the returns for six exchange rates against the U.S. Dollar. Also plotted is the log-density for fitted GH model.*

Table 2, considered earlier, reports the results from using the above methods to compute the robust standard errors for the Canadian Dollar dataset. The results do not vary much with m.

The corresponding fitted generalised hyperbolic log-density for the Canadian Dollar is given in the upper left hand graph in Figure 12. This shows sub-log-linear tails in the marginal distribution. The fit of the model is very close to the drawn non-parametric estimate of the log-density. The non-parametric estimator is the log of the Gaussian kernel estimator coded in Applied Statistics Algorithm AS 176 by Bernard Silverman, which is available at `StatLib` and `NAG` and in many statistical software environments. The bandwidth is chosen to be $1.06\widehat{\sigma}n^{-1/5}$, where $\widehat{\sigma}^2 = \sum y_i^2/n$ (this is an optimal choice against a mean square error loss for Gaussian data).

6.3 Further empirical results

6.3.1 Six daily exchange rate movements

Table 3 gives the estimates of the parameters, together with their standard and robust standard errors, for our daily exchange rate return data sets. The corresponding fitted log-density for all six series was given in Figure 12. This shows sub-log-linear tails in the marginal distributions for all the fitted distributions except for and German Mark and Sterling, which has approximately log-linear tails. And, as already notes, the fit of the model is very close to the drawn non-parametric estimate of the log-density.

Partly repeating some of the discussion above, there are a number of common features across these results. First all the non-Gaussian models provide dramatic improvements over the fit of the normal likelihood. The ν parameters seems to take values between -2 and 0.5, while neither δ nor γ are close to zero. To reinforce this, Figure 10 showed the profile likelihood function for each of the datasets. Also drawn are the corresponding profile likelihoods for the skewed Student's t and normal gamma models. The results indicates that the normal gamma model is not really supported by the data. The skewed Student's t model fits better — primarily as it has fatter tails. Typically when $\gamma = 0$, ν is around -2, which corresponds to 4 degrees of freedom for the Student's t distribution. The fit of the distribution is very sensitive to this value. The skewed Student's t is dominated by GH models with $\gamma > 0$. The likelihood function is typically flat for GH models with ν between -2 and 2. Overall, however, the values between -2.0 and 0 seem best. Finally, the special cases of imposing $\beta = 0$ seems not to harm the fit a great deal for exchange rate data, although there is slight statistical significance in the negative skewness in the UK Sterling, Swiss Franc and Japanese Yen series.

Rate	MLE of GH parameters					Likelihoods				
	μ	β	δ	γ	ν	GH	$\beta = 0$	$\delta = 0$	$\gamma = 0$	N
Canada	-0.013	0.179	0.419	1.95	-1.62	-626.74	-627.97	-638.19	-627.61	-819.32
	(.016)	(.172)	(.076)	(.399)	(.769)					
DM	0.024	-0.064	0.873	1.40	-0.979	-3,903.1	-3,903.8	-3,909.8	-3,905.0	-4,052.2
	(.033)	(.063)	(.129)	(.265)	(1.10)					
FF	0.030	-0.074	0.930	0.923	-1.55	-3,797.1	-3,798.1	-3,809.3	-3,798.1	-3,988.2
	(.029)	(.060)	(.083)	(.208)	(.614)					
SF	0.073	-0.143	1.08	1.25	-1.27	-4,296.9	-4,300.6	-4,302.1	-4,298.3	-4,428.2
	(.037)	(.065)	(.094)	(.214)	(.834)					
JY	0.044	-0.120	0.800	0.841	-1.12	-4,022.6	-4,027.1	-4,042.8	-4,025.5	-4,310.3
	(.026)	(.041)	(.069)	(.144)	(.372)					
Pound	0.034	-0.089	0.430	1.89	0.145	-3485.7	-3487.5	-3492.5	-3495.1	-3704.8
	(.018)	(.039)	(.090)	(.097)	(.439)					

Table 3: *Fit of GH for daily exchange rates. GH denotes unrestricted model.* $\beta = 0$ *imposes symmetry,* $\delta = 0$ *skewed* $N\Gamma$*model,* $\gamma = 0$ *skewed Student. Robust S.E.s (*$m = 500$*) are in brackets.*

6.3.2 Daily equity indexes

Table 4 gives the estimates of our parameters for the daily equity return data. The corresponding profile likelihoods are given in Figure 13. These results are more mixed, with values of ν between -1 and 1 being roughly necessary. Overall again the normal inverse Gaussian usually does pretty well, never fitting really poorly. One conclusion from these fitted models is that there seems very little asymmetry in this data. This is perhaps surprising as this is always an important possibility for equity data. The improvement over the Gaussian fit is picked up very well in the discrepancy between the Gaussian and the GH likelihood fits. This holds across all the assets, but is less severe for FTSE — which is not surprising given its normal gamma like behaviour.

Index	MLE of GH parameters					Likelihoods				
	μ	β	δ	γ	ν	GH	$\beta=0$	$\delta=0$	$\gamma=0$	N
DAX 30	0.234	-0.113	0.983	0.821	-0.206	-2,735.5	-2,740.6	-2,739.1	-2,740.3	-2,849.4
	(.091)	(.037)	(.126)	(.133)	(.836)					
FTSE 100	0.043	-0.004	0.734	1.99	1.72	-2,488.7	-2,488.7	-2,488.8	-2,490.4	-2,515.4
	(.041)	(.037)	(.328)	(.190)	(1.95)					
S&P 500	0.124	-0.059	1.05	0.655	-1.03	-2,314.6	-2,315.6	-2,322.3	-2,315.8	-2,444.9
	(.027)	(.019)	(.191)	(.236)	(.872)					
Nikkei 500	0.009	-0.007	1.08	0.780	-0.654	-2,531.0	-2,531.0	-2,536.1	-2,532.4	-2,638.1
	(.038)	(.029)	(.293)	(.322)	(1.75)					

Table 4: *Fit of GH for daily equities. GH denotes unrestricted model.* $\beta = 0$ *imposes symmetry,* $\delta = 0$ skewed $N\Gamma$ *model,* $\gamma = 0$ *skewed Student. Robust S.E.s* $(m = 500)$ *are in brackets.*

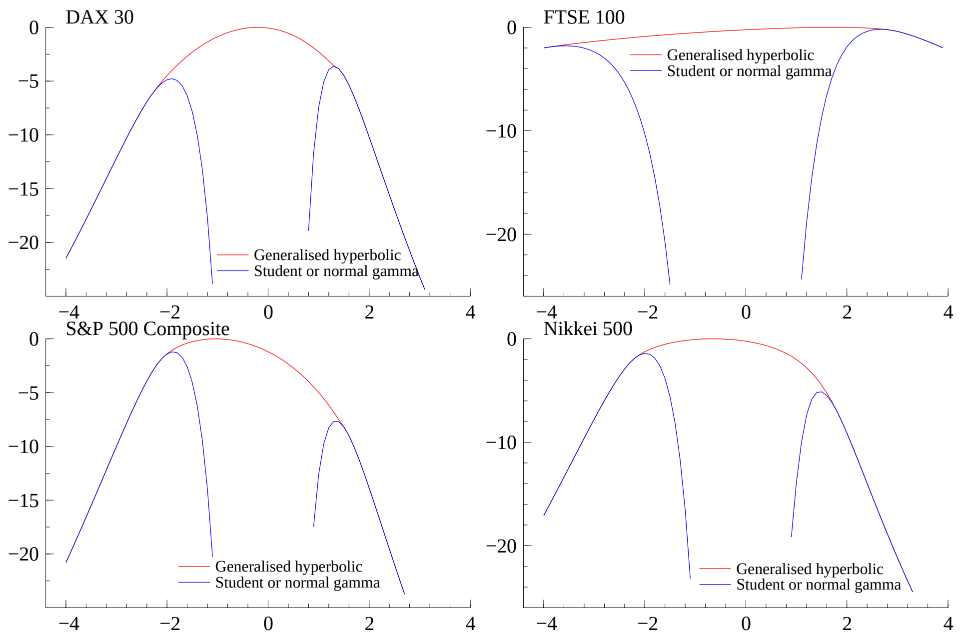

Figure 13: *Daily index return data. Profile likelihood for* ν *in GH model. Also profile for the skewed Student* $(\nu < 0)$ *and the skewed* $N\Gamma$ $(\nu > 0)$*.*

7 Quadratic variation

7.1 Basics

A central concept in financial econometrics, derivative pricing and stochastic analysis is the Quadratic Variation (QV) process. This has two steps. First, time is split into small intervals

$$t_0^r = 0 < t_1^r < ... < t_{m_r}^r = t.$$

Then the QV process is defined as

$$[Y]_t = \text{p}-\lim_{r\to\infty} \sum \left(Y_{t_{i+1}^r} - Y_{t_i^r} \right)^2, \tag{44}$$

where $\sup_i \left(t_{i+1}^r - t_i^r\right) \to 0$ for $r \to \infty$.

It is sometimes helpful to work with an alternative, and equivalent[2] (to (44)), definition of QV which is written in terms of a stochastic integral. It is that

$$[Y]_t = Y_t^2 - 2\int_0^t Y_{u-}\mathrm{d}Y_u. \tag{45}$$

This is discussed in some detail in our primer on stochastic analysis.

Example 8 *Let N be a Poisson process and let us check the consistency of the formulae (45) and (44). Suppose $N_t = n$. It is immediate from (44) that*

$$[N]_t = n$$

while, on the other hand,

$$N_t^2 - 2\int_0^t N_{u-}\mathrm{d}N_u = n^2 - 2\sum_{i=1}^{n}(i-1) = n^2 - 2\binom{n}{2} = n.$$

In general the QV process of a Lévy process is a subordinator, for the increments are non-negative, independent and stationary. If Y has a Lévy density u, recalling this means its Lévy measure v can be written as $v(\mathrm{d}y) = u(y)\mathrm{d}y$, then for $y > 0$ the Lévy density of the QV process is

$$\frac{u(\sqrt{y})}{2\sqrt{y}} + \frac{u(-\sqrt{y})}{2\sqrt{y}}.$$

This follows from the Lévy-Ito representation of Y.

Example 9 *Suppose $Y = \sigma W \circ T$ where W is Brownian motion and T is a Poisson process independent of W with intensity λ, then $Y_t = \sum_{j=1}^{T_t} C_j$, where $C_j \overset{i.i.d.}{\sim} N(0,\sigma^2)$, is a compound process (37) while*

$$[Y]_t = \sum_{j=1}^{T_t} C_j^2.$$

Hence $[Y]$ is a compound Poisson process subordinator with Lévy density

$$\lambda\sigma^{-1} f(\sigma^{-1}y \ddagger \chi_1^2),$$

where $f(y \ddagger \chi_1^2)$ denotes the probability density function of the χ_1^2 distribution with one degree of freedom.

[2] That they are exactly equivalent follows from Ito's lemma. Many probabilists use this second form as the original definition of QV, rather than via the limit of sums of squares. However, their formal equivalence implies we can use either definition provided we are in the semimartingale framework.

7.2 Brownian motion and quadratic variation

Suppose Y is a scaled Brownian motion with drift, such that $Y_1 \sim N(\mu, \sigma^2)$. Then

$$Y_{t_{i+1}^r} - Y_{t_i^r} \sim N(\left(t_{i+1}^r - t_i^r\right)\mu, \left(t_{i+1}^r - t_i^r\right)\sigma^2).$$

For small values of $\left(t_{i+1}^r - t_i^r\right)$ the variation in the series dominates — the standard deviation and drift are $O\left(\sqrt{t_{i+1}^r - t_i^r}\right)$ and $O\left(t_{i+1}^r - t_i^r\right)$, respectively. As a result

$$[Y]_t = t\sigma^2,$$

whatever the value of μ. The QV is non-stochastic and is the only non-trivial example of a Lévy process where the QV degenerates to a deterministic function of time.

From a statistical viewpoint it means we can theoretically estimate σ^2 without error using a tiny path of Brownian motion even in the presence of drift. Of course in practice this is a highly misleading argument for the continuous time model is unlikely to be perfectly specified at very short time horizons.

7.3 Realised QV process

The *realised QV process* is defined, for $\delta > 0$ and any stochastic process Y, by

$$[Y_\delta]_t = \sum_{j=1}^{\lfloor t/\delta \rfloor} \left(Y_{\delta j} - Y_{\delta(j-1)}\right)^2,$$

where $\lfloor t \rfloor$ denotes the largest integer less than or equal to t. We can see that if $Y \in \mathcal{SM}$ then

$$p - \lim_{\delta \downarrow 0} [Y_\delta]_t = [Y]_t,$$

that is the realised QV is a consistent estimator of QV. However, if Y is a Lévy process $[Y_\delta]$ is not a Lévy process — rather it jumps upwards at specified points in time and so is a discrete time random walk.

Example 10 *A numerical example of the realised QV process is given in Figure 14, which computes it for a NIG Lévy process. In this picture we have taken $\delta = 1$ and $\delta = 1/10$, so taking 1 and 10 squared observations per unit of time, respectively. Also given is the corresponding limit, the QV. We see that as δ gets small so $[Y_\delta]$ becomes a good approximation of $[Y]$.*

In applied economics it is often inappropriate to study returns over tiny time intervals for our models tend to be highly misspecified at that scale due to market frictions. In particular the idea of a unique price is a fiction, for the transaction price tends to depend upon, for example, the volume of the deal, the reputation of the buyer and seller, prevailing liquidity (and so time of day)

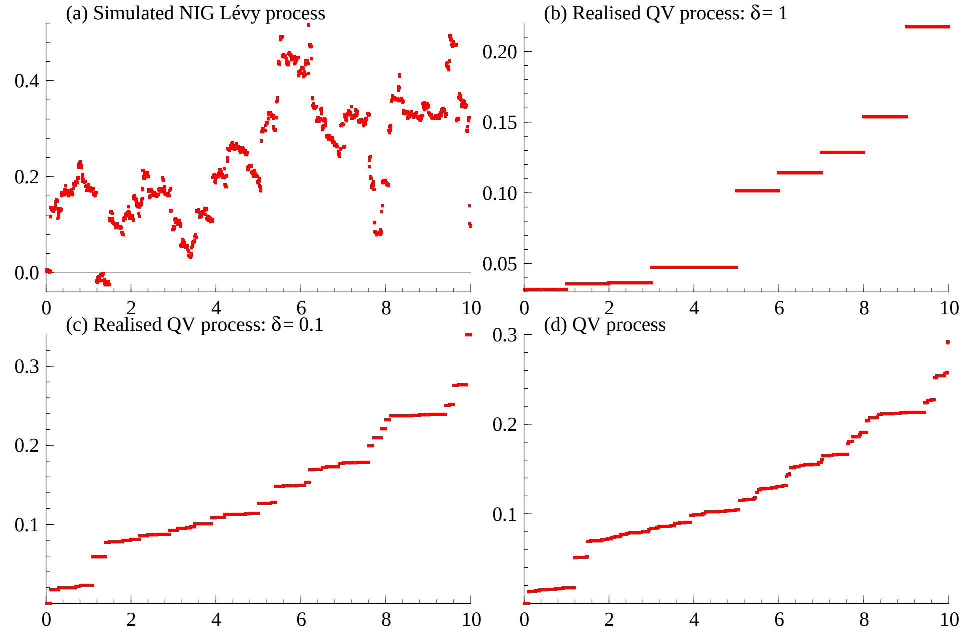

Figure 14: *Figure (a) Sample path of NIG(0.2,0,0,10) Lévy process. (b) Corresponding realised QV process taking $\delta = 1$. (c) Same but with $\delta = 1/10$. (d) QV of the process. Code:* `levy_code.ox`.

and the initiator (i.e. was it the buyer or the seller). These issues will be discussed at more length in later Chapters. To avoid the worst effects of misspecification, realised quadratic variation $[Y_\delta]$ with δ not too small is used.

In order to understand the connection between the Lévy, realised QV and QV processes it is helpful to think about the following calculation

$$\mathrm{E}\begin{pmatrix} Y_t \\ [Y_\delta]_t \\ [Y]_t \end{pmatrix} = t\begin{pmatrix} \kappa_1 \\ \kappa_2 \\ \kappa_2 \end{pmatrix}, \quad \mathrm{Cov}\begin{pmatrix} Y_t \\ [Y_\delta]_t \\ [Y]_t \end{pmatrix} = t\begin{pmatrix} \kappa_2 & \kappa_3 & \kappa_3 \\ \kappa_3 & \kappa_4 + 2\kappa_2^2\delta & \kappa_4 \\ \kappa_3 & \kappa_4 & \kappa_4 \end{pmatrix},$$

where κ_r denotes the r-th cumulant of Y_1. The only one of these results which is not straightforward is

$$\begin{aligned} \mathrm{Var}([Y_\delta]_t) &= \frac{t}{\delta}\left\{\mu_4(Y_\delta) - \mu_2(Y_\delta)^2\right\} \\ &= \frac{t}{\delta}\left\{\kappa_4(Y_\delta) + 2\kappa_2^2(Y_\delta)\right\} \\ &= t\left(\kappa_4 + 2\delta\kappa_2^2\right), \end{aligned}$$

where, generically, $\kappa_r(X)$ is the r-th cumulant of X. Finally we notice the implication that $[Y_\delta] - [Y]$ has a zero mean, while

$$\mathrm{Var}\left([Y]_t - [Y_\delta]_t\right) = 2\kappa_2^2 t\delta.$$

Hence we could use $[Y_\delta]$ as an estimator of $[Y]$.

Example 11 *Suppose Y is standard Brownian motion, then $\kappa_3 = \kappa_4 = 0$, which means that*

$$\mathrm{E}\begin{pmatrix} Y_t \\ [Y_\delta]_t \\ [Y]_t \end{pmatrix} = t\begin{pmatrix} \kappa_1 \\ \kappa_2 \\ \kappa_2 \end{pmatrix}, \quad \mathrm{Cov}\begin{pmatrix} Y_t \\ [Y_\delta]_t \\ [Y]_t \end{pmatrix} = t\begin{pmatrix} \kappa_2 & 0 & 0 \\ 0 & 2\kappa_2^2\delta & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

which makes sense, for $[Y]_t = t$.

Example 12 *Suppose Y is the homogeneous Poisson process, then all the cumulants are equal to κ_1. Thus, writing $\iota = (1\ 1\ ...\ 1)'$,*

$$\mathrm{E}\begin{pmatrix} Y_t \\ [Y_\delta]_t \\ [Y]_t \end{pmatrix} = t\kappa_1\iota, \quad \mathrm{Cov}\begin{pmatrix} Y_t \\ \{Y_\delta\}_t \\ [Y]_t \end{pmatrix} = t\kappa_1\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1+2\kappa_1\delta & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Notice the covariance is singular, for $Y = [Y]$ as was noted in Example 8.

Example 13 *Suppose $Y = \sigma W \circ T$. Then*

$$\mathrm{C}\{\zeta \ddagger Y_1\} = \overline{\mathrm{K}}\left\{\frac{1}{2}\zeta^2\sigma^2 \ddagger T_1\right\}.$$

This implies

$$\kappa_1 = 0, \quad \kappa_2 = \sigma^2\kappa_1(T), \quad \kappa_3 = 0, \quad \kappa_4 = 3\sigma^2\kappa_2(T).$$

Hence in this case

$$\mathrm{Var}\left([Y]_t - [Y_\delta]_t\right) = 2\sigma^4\kappa_1^2(T)t\delta,$$

which only depends upon the mean of the subordinator, not its variance.

8 Lévy processes and stochastic analysis

8.1 Ito's formula for Lévy processes

Any Lévy process Y is a semimartingale and so Ito's formula for semimartingales immediately applies such that, for any real function f which is twice continuously differentiable, then $f(Y_t) \in \mathcal{SM}$ and

$$\mathrm{d}f(Y_t) = f'(Y_{t-})\mathrm{d}Y_t + \frac{1}{2}f''(Y_{t-})\mathrm{d}[Y^c]_t + f(Y_t) - f(Y_{t-}) - f'(Y_{t-})\Delta Y_t, \tag{46}$$

where Y^c is the continuous part of the Lévy process. For Lévy processes $[Y^c]$ is exactly zero unless Y has a Brownian component, in which case $[Y^c]$ is proportional to t.

Example 14 *Suppose Y is a pure jump Lévy process and $V = \exp(Y)$, then*

$$\mathrm{d}V_t = V_{t-}\mathrm{d}Y_t + \Delta V_t - V_{t-}\Delta Y_t.$$

8.2 Stochastic exponential of a Lévy process

The stochastic exponential $\mathcal{E}(Y)$ of Y is defined as

$$\mathcal{E}(Y)_t = e^{Y_t - \frac{1}{2}[Y]_t} \prod_{0<u\leq t} (1 + \Delta Y_u) e^{-\Delta Y_u + \frac{1}{2}\Delta Y_u^2} \tag{47}$$

and $V = \mathcal{E}(Y)$ is the unique solution to the SDE

$$\mathrm{d}V_t = V_{t-}\mathrm{d}Y_t \tag{48}$$

with initial condition $V_0 = 1$. The process $\mathcal{E}(Y)_t$ is a semimartingale. Some of the above is familiar from the result on exponentiating a diffusion, where the result is $\mathcal{E}(Y)_t = \exp\left(Y_t - \frac{1}{2}[Y]_t\right)$.

8.3 Stochastic logarithm

The stochastic logarithm $\mathcal{L}(Y)$ of $Y \in \mathcal{SM}$, where Y is assumed to live on $\mathbb{R}_{>0}$, is defined as

$$\mathcal{L}(Y)_t = \int_0^t \frac{1}{Y_{u-}} \mathrm{d}Y_u. \tag{49}$$

Clearly, with $V = \mathcal{E}(Y)$,

$$\mathcal{L}(\mathcal{E}(Y))_t = \int_0^t \frac{1}{V_{u-}} \mathrm{d}V_u = \int_0^t \frac{1}{V_{u-}} V_{u-}\mathrm{d}Y_u = Y_t,$$

where we have used formula (48).

9 Multivariate Lévy processes

9.1 Elemental constructions

An important question is how to generate multivariate Lévy processes, that is processes with independent and stationary multivariate increments. Here we discuss just two approaches: linear transformation and time deformation, of independent Lévy process U and V.

If Θ is some deterministic matrix, then the linear combination of the original Lévy processes

$$Y = \Theta \begin{pmatrix} U \\ V \end{pmatrix},$$

is a bivariate Lévy process. The elements of Y are marginally Lévy processes. This type of argument generalises to any dimension.

We saw in Section 5 that subordination can be used to generate compelling Lévy processes. Here we use this idea to put

$$Y = \begin{pmatrix} U \\ V \end{pmatrix} \circ T,$$

where T is an independent, common subordinator. This means, for $Y_t^1 = U_{T_t}$ and $Y_t^2 = V_{T_t}$, and Y is a multivariate Lévy process. A concrete example of this is where (U, V) are independent standard Brownian motions, then

$$Y_t|T_t \sim N(0, T_t I),$$

which implies the elements of Y_t are uncorrelated but dependent. In particular

$$\text{Cov}\begin{pmatrix} Y_t^1 \\ Y_t^2 \end{pmatrix} = \text{E}(T_t)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

while

$$\text{E}\left\{\left(Y_t^1\right)^2\left(Y_t^2\right)^2\right\} = \text{E}\left(T_t^2\right) \neq \text{E}\left(T_t\right)^2 = \text{E}\left\{\left(Y_t^1\right)^2\right\}\text{E}\left\{\left(Y_t^2\right)^2\right\}.$$

More generally we could consider a bivariate Lévy process $T = \left(T^{(1)}, T^{(2)}\right)$, with both $T^{(1)}$ and $T^{(2)}$ being subordinators, setting $Y = \left(U \circ T^{(1)}, V \circ T^{(2)}\right)'$.

9.2 Multivariate generalised hyperbolic Lévy process

9.2.1 Background

Suppose we take V_t as a $d \times 1$ vector of correlated Brownian motions generated by

$$V_t = t\Sigma\beta + \Sigma^{1/2} W_t,$$

where W is a $d \times 1$ vector of independent, standard Brownian motions and Σ is a positive definite $d \times d$ matrix. Further we take T to be an independent subordinator and define the deformed series $Y_t = \mu t + V_{T_t}$. Then Y is a Lévy process with

$$Y_t|T_t \sim N(\mu t + T_t\Sigma\beta, T_t\Sigma).$$

Suppose we choose to make T a $GIG(\nu, \delta, \gamma)$ Lévy process, then we say that Y is a multivariate generalised hyperbolic Lévy process, following our earlier work on the univariate process discussed in Section 4.2.4. In particular the increments of such a process are independent and stationary while the density of Y_1 is known to follow a multivariate $GH(\nu, \alpha, \beta, \mu, \delta, \Sigma)$ density

$$f_{Y_1}(y) = \frac{(2\pi)^{-d/2}\gamma^{2\nu}\alpha^{d-2\nu}}{|\Sigma|\, K_\nu(\delta\gamma)} K_{\nu-d/2}\left\{\alpha\sqrt{\delta^2 + q}\right\}\exp\left\{\beta'(y-\mu)\right\}, \tag{50}$$

where

$$q = (y-\mu)'\,\Sigma^{-1}(y-\mu) \qquad \text{and} \qquad \alpha = \sqrt{\gamma^2 + \beta'\Sigma^{-1}\beta}.$$

Here Σ allows us to model the correlation between the processes, while ν, δ, and γ controls the tails of the density. The whole vector β freely parameterises the skewness of the returns. In order to

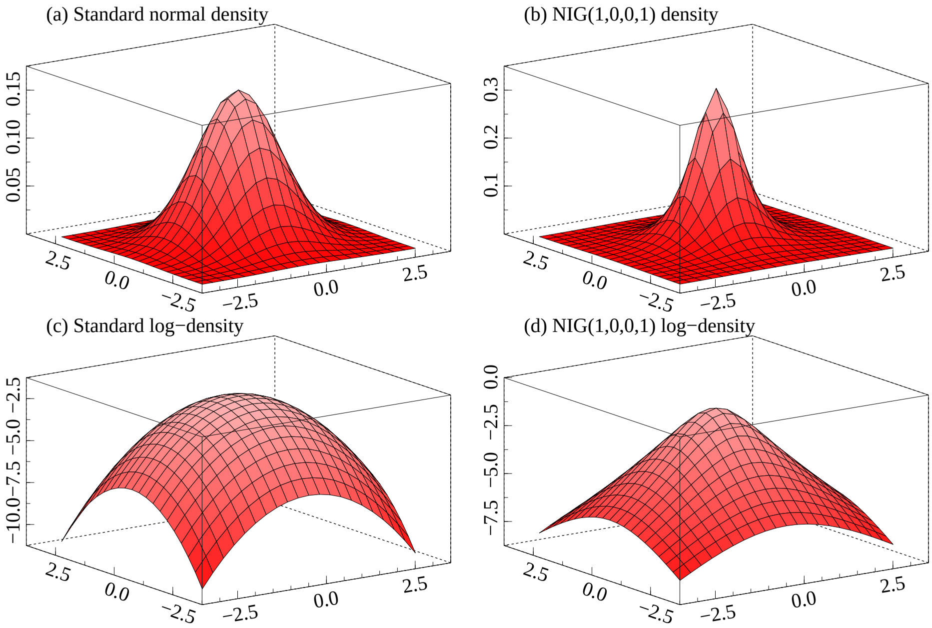

Figure 15: *Densities and log-densities of N(0,I) and NIG(1,0,0,1,I) variables. Densities: (a) N(0,I). (b) NIG. Log-densities: (c) N(0,I). (d) NIG. Code:* `levy_graphs.ox`.

enforce identification on this model it is typical to assume that $\det(\Sigma) = 1$, although Mencia and Sentana (2004) have normalised the model by setting $\delta = 1$. The multivariate GH density has many interesting special cases such as the multivariate skewed Student's t, normal gamma, normal inverse Gaussian, hyperbolic and Laplace. Of course, as the multivariate GH is a normal variance-mean mixture, then linear combinations of Y are also GH, while if we write $Y = (X', Z')'$ then $Z|X$, $X|Z$, X and Z are all GH. Hence many of the important attractive features of the multivariate Gaussian distribution carry over to the multivariate generalised hyperbolic distribution.

A simple example of the above multivariate distributions of Y_1 is given in Figure 15 which draws the density and log-density for the bivariate standard normal and the corresponding $NIG(1, 0, 0, 1, I)$ variables (chosen so that the marginal variances of the variables are 1). Again the log-densities show that the tails of the NIG variables are much thicker — looking roughly linear in all tails. This has a very big impact on the chance of observing two observations in the tails of the distribution.

9.2.2 Empirical fit of multivariate GH processes

Experience suggests that it is computationally convenient to compute the ML estimator of multivariate GH models using the EM algorithm. This approach, which is detailed in Barndorff-Nielsen and Shephard (2012c), becomes particularly attractive when d is large for the EM algorithm quickly converges to the ML estimator as the degree of missing data, which in this context is the unknown scale σ, lessens as d increases.

We start by fitting some bivariate models. The first example of this is a fit of the German DM and French Franc against the US Dollar which is reported in the first two lines of Table 5 and Figure 16. The Table shows the expected dramatic improvement in fit associated with these

DM	MLE of GH parameters							Likelihoods				
+	μ	β	δ	γ	ν	Σ		GH	$\beta=0$	$\delta=0$	$\gamma=0$	N
	.0221	-.0310	.365	1.03	-1.06	3.83	3.59	-2,865	-2,867	-2,933	-2,869	-3,913
FF	.0241	-.0143				3.59	3.62	*4,835*	*4,834*	*4,785*	*4,834*	*4,127*
	.0130	-.0434	.655	1.78	-1.46	2.31	.049	-4,518	-4,524	-4,535	-4,520	-4,868
Can	-.0116	.167				.049	.432	*11.8*	*7.4*	*12.6*	*12.2*	*3.2*
	.0429	.0283	.660	1.36	-1.62	2.24	2.27	-4,592	-4,605	-4,615	-4,594	-5,023
SF	.0548	-.129				2.27	2.76	*3,607*	*3,599*	*3,596*	*3,609*	*3,179*
	.0185	.0424	.737	1.37	-.848	1.15	0.60	-7,258	-7,264	-7,281	-7,279	-7,730
JY	.0578	-.175				0.60	1.18	*667*	*666*	*671*	*651*	*632*
	.0315	-.303	.543	1.95	-.391	1.57	-0.98	-5,975	-6,026	-5,993	-5,986	-6,457
BP	.0449	-.361				-0.98	1.24	*1,413*	*1,365*	*1,408*	*1,414*	*1,298*

Table 5: *Bivariate GH models. Fits the pair of the DM plus another currency against US Dollar. In italics are improvement in the log-likelihood compared to fit of the univariate models.* $\beta = 0$ imposes symmetry, $\delta = 0$ gives skewed $N\Gamma$ model, $\gamma = 0$ gives skewed Student distribution. *Code:* `em_gh_mult.ox`.

multivariate models, for the DM and FF are highly related currencies. This is shown up by the estimated Σ matrix. Again ν is estimated to be negative, while the fit of the GH model is very close to the bivariate skewed Student's t in this case. The normal gamma model is quite a lot poorer in this multivariate setting. The result in the italics gives the likelihood for the multivariate model minus the sum of the likelihoods for the DM and FF univariate models. So the number for the normal case shows an improvement in the likelihood of $4,127$. Although this is very substantial, the improvements for the other models are much higher. Hence the gains in using GH models is even higher in the multivariate case than one might have expected from the univariate analysis.

Figure 16 shows the fitted bivariate normal and GH densities for the DM and FF returns. The graphs have been drawn to show the densities in places where the log-density does not drop 12 from the mode. This gives an impression of the plausible scatter of points from this variable. The bivariate normal density is tightly packed, while the GH model gives a wider range of possible points while the tails of the log-density appear linear or slower in each direction.

Table 5 gives the results for all the bivariate exchange rate relationships which involve the Dollar. Broadly similar conclusions follow from the above, except the degree of dependence between the currencies is smaller in these other cases. Interestingly the UK Sterling is negatively related to the DM returns. Throughout the table the estimated values of ν ranges between about -0.5 and -1.5. This is an important common theme, again suggesting evidence against the use of normal gamma models.

Table 6 gives the GH fit to all six exchange rate return series. This high dimensional model has

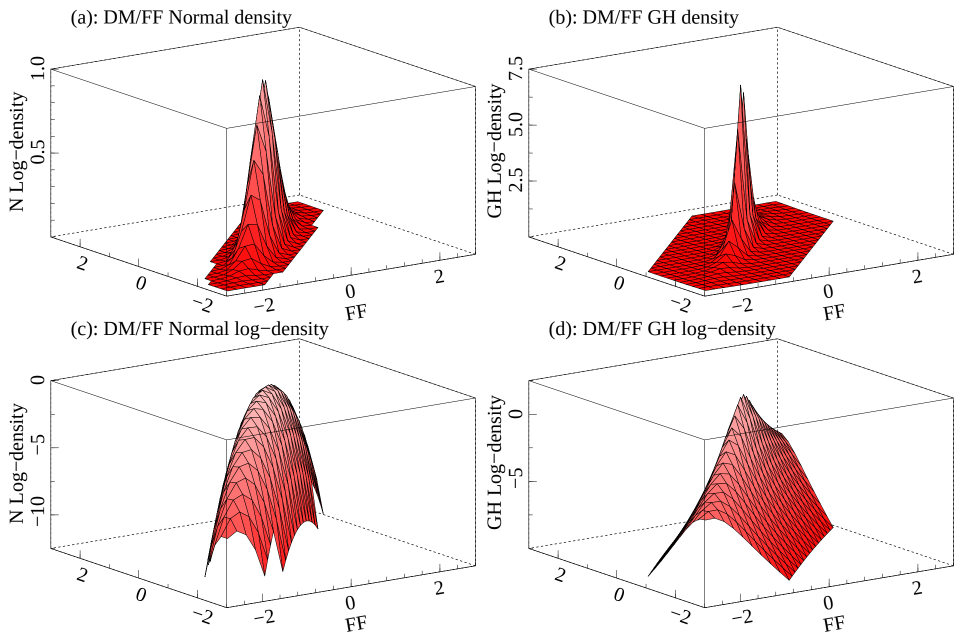

Figure 16: *Fit of bivariate Gaussian and GH models for the DM and FF against the Dollar. (a) ML fit of bivariate Gaussian density, (c) log-density. (b) ML fit of bivariate GH density, (d) log-density. Code:* `em_gh_mult.ox`.

a low value of ν, which suggests the fit is very close to a skewed multivariate Student's t distribution. The skewness parameters are important. The NIG fits worse than the skewed Student's t but is much better than the normal gamma model. All these models are enormous improvements over the multivariate normal fit to the data.

Table 7 gives the corresponding result for the four dimensional equity return data. Here the elements of β are all estimated to be negative, indicating common negative skewness. That is, the large negative movements have a tendency to occur in all the markets at the same time. In this case the non-symmetries are important, while the normal gamma is again considerably worse than the Student's t or the NIG distributions.

$\delta = .638, \quad \gamma = .586, \quad \nu = -2.22,$

$$\Sigma = \begin{pmatrix} .628 & .0692 & .0754 & .0760 & .0199 & -.115 \\ .0692 & 3.09 & 2.86 & 3.13 & 1.66 & -1.95 \\ .0754 & 2.86 & 2.91 & 2.99 & 1.57 & -1.86 \\ .0760 & 3.13 & 2.99 & 3.82 & 1.85 & -2.08 \\ .0199 & 1.66 & 1.57 & 1.85 & 3.45 & -1.19 \\ -.1157 & -1.95 & -1.86 & -2.08 & -1.19 & 2.54 \end{pmatrix}, \mu = \begin{pmatrix} -.0115 \\ .0307 \\ .0346 \\ .0494 \\ .0605 \\ .0239 \end{pmatrix}, \beta = \begin{pmatrix} .120 \\ .0875 \\ -.0647 \\ -.182 \\ -.148 \\ -.254 \end{pmatrix}.$$

Likelihoods					
GH	$\beta = 0$	$\delta = 0$	$\gamma = 0$	NIG	N
-9,671	-9,726	-9,791	-9,672	-9,697	-11,635

Table 6: *Multivariate GH model for CD, DM, FF, SF, JY and Sterling.* $\beta = 0$ *imposes symmetry.* $\delta = 0$ *gives skewed* $N\Gamma$ *model,* $\gamma = 0$ *gives skewed Student t distribution. Code:* `em_gh_mult.ox`.

$\delta = 1.47$, $\gamma = .660$, $\nu = -1.72$,					
$\Sigma = \begin{pmatrix} 1.54 & .778 & .654 & .145 \\ .778 & 1.18 & .510 & .181 \\ .654 & .510 & .977 & .128 \\ .145 & .181 & .128 & 1.26 \end{pmatrix}, \mu = \begin{pmatrix} .201 \\ .132 \\ .161 \\ .0872 \end{pmatrix}, \beta = \begin{pmatrix} -.0668 \\ -.0105 \\ -.0438 \\ -.0580 \end{pmatrix}.$					
Likelihoods					
GH	$\beta = 0$	$\delta = 0$	$\gamma = 0$	*NIG*	*N*
-9,268	-9,285	-9,293	-9,270	-9,271	-9,694

Table 7: *Multivariate GH model for DAX 30, FTSE 100, S&P500, Nikkei 100. $\beta = 0$ imposes symmetry, $\delta = 0$ implies the skewed $N\Gamma$ model, $\gamma = 0$ is the skewed Student t. Code:* `em_gh_mult.ox`.

9.3 Stochastic discount factors

In financial economics we typically price contingent payoffs $g(Y_T)$ as

$$C_t = \mathrm{E}\left(\frac{\widetilde{M}_T^*}{\widetilde{M}_t^*} g(Y_T)|\mathcal{F}_t\right), \quad \widetilde{M}_t^* = \exp\left(\widetilde{M}_t\right),$$

the expected discounted value of the claim where $T > t$, $\widetilde{M}^*$ is called the stochastic discount factor (SDF) and $\widetilde{M}$ is the log-SDF. For this setup to rule out trivial arbitrages we require that

$$\exp\left(\widetilde{M}_t Y_t\right) \quad \text{and} \quad \exp(\widetilde{M}_t)\exp(tr)$$

are local martingales, in order to avoid arbitrage, and where r is a riskless interest rate. A book length exposition of this approach to asset pricing is given by Cochrane (2001).

Suppose that $\widetilde{M}$ and Y together constitute a bivariate Lévy process. What constraints are imposed on the Lévy process in this setup? For $\exp(\widetilde{M}_t)\exp(tr)$ to be a martingale we need that

$$\mathrm{K}\left\{1 \ddagger \widetilde{M}_1 + Y_1\right\} = 0 \quad \text{and} \quad \mathrm{K}\left\{1 \ddagger \widetilde{M}_1\right\} = -r. \tag{51}$$

Example 15 *Suppose*

$$\begin{pmatrix} \widetilde{M} \\ Y \end{pmatrix}_t = \begin{pmatrix} \mu_{\widetilde{M}} \\ \mu_Y \end{pmatrix} t + \begin{pmatrix} \sigma_{\widetilde{M}} B \\ \sigma_Y W \end{pmatrix}_t,$$

where B, W are independent Brownian motions. Then (51) imply

$$\mu_Y = r - \frac{1}{2}\sigma_Y^2 \quad \textit{and} \quad \mu_{\widetilde{M}} = -r - \frac{1}{2}\sigma_{\widetilde{M}}^2.$$

Hence C_t does not depend upon the value of $\sigma_{\widetilde{M}}^2$. More generally, if $\widetilde{M} \perp\!\!\!\perp Y$ and $\widetilde{M}, Y$ are Lévy processes then we require

$$\mathrm{K}\left\{1 \ddagger Y_1\right\} + \mathrm{K}\left\{1 \ddagger \widetilde{M}\right\} = r \quad \textit{and} \quad \mathrm{K}\left\{1 \ddagger \widetilde{M}\right\} = -r.$$

10 From Lévy processes to semimartingales*

Lévy processes are determined by their characteristic triplets. For any Lévy process Y with characteristic triplet (a, b, v) the law of Y_t has characteristic triplet (ta, tb, tv). Suppose we think of extending the concept of Lévy processes by having a triplet of predictable processes (a_t, b_t, v_t) instead of (a, b, v). It turns out that the class of semimartingales can be seen as the natural answer to this quest — which in turn places them at the centre of modern continuous time asset pricing for arbitrage freeness and semimartingales are synonymous.

We will indicate the character of this, leaving the rather formidable technicalities aside.

Subject to minor regularity assumptions, if Y is a univariate semimartingale then there is a unique triplet $\partial Y = (a, b, v)$ of predictable processes such that Y has representation

$$\begin{aligned} Y_t &= Y_0 + \int_0^t a_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s + \int_0^t \int_{\mathbb{R}^d} \bar{c}(x)\,\mu(\mathrm{d}s\mathrm{d}x) \\ &\quad + \int_0^t \int_{\mathbb{R}^d} c(x)\left(\mu(\mathrm{d}s\mathrm{d}x) - \nu_s(\mathrm{d}x)\,\mathrm{d}s\right) \end{aligned} \tag{52}$$

where W is Brownian motion, $b = \sigma^2$, c is a truncation function, $\bar{c} = 1 - c$, μ is the random measure given by

$$\mu((0, t] \times A) = \sum_{0<s\leq t} 1_A(\Delta Y_s)$$

(where A is an arbitrary Borel set in $\mathbb{R}$) and $v_t(\mathrm{d}x)\,\mathrm{d}t$ is the compensator of $\mu(\mathrm{d}t\mathrm{d}x)$, i.e.

$$v_t(\mathrm{d}x)\,\mathrm{d}t = \mathrm{E}\{\mu(\mathrm{d}t\mathrm{d}x)\}.$$

Example 16 *If there are no jumps then*

$$Y_t = Y_0 + \int_0^t a_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s, \tag{53}$$

which is often called a Brownian semimartingale. The characteristic triplet of Y is in this case $\left(a_{t-}, \sigma_{t-}^2, 0\right)$. The term $\int_0^t \sigma_s \mathrm{d}B_s$ is a stochastic volatility process and such processes are the subject of the next Chapter.

Example 17 *Suppose N is a Poisson process jumping with intensity λ and let*

$$Y_t = \sum_{0\leq s\leq t} C_s 1_{\{1\}}(\Delta N_s) \tag{54}$$

where $\{C_t\}_{t\geq 0}$ is a family of independent but not necessarily identically distributed random variables. Then Y is semimartingale of the form (52) with triplet $(0, 0, \lambda\pi_t)$ where π_t is the law of C_t and is assumed to depend continuously on t. In this case,

$$Y_t = \mu((0, t] \times \{1\}).$$

If, more generally, the process N is a Poisson process with deterministic intensity λ_t we are still in the semimartingale setting but the structure is slightly more complicated than in (52) and a more general triplet concept is needed.

Example 18 *Adding (53) and (54), assuming they are independent, gives a semimartingale as in (52) and with triplet $\left(a_{t-}, \sigma_{t-}^2, \lambda\pi_t\right)$.*

The process (54) is an example of an *additive process*, that is a process with independent increments. Note that all such processes have deterministic triplet (a_t, b_t, ν_t).

11 Conclusion

This Chapter has provided an informal introduction to Lévy processes. Starting with subordinators, where the Lévy-Khintchine representation is more accessible, we have typically built models using the idea of time deformation. Such models are then imbedded within the familiar normal variance-mean mixture class of processes. We have emphasised the role of QV and have discussed the development of various multivariate Lévy processes. Barndorff-Nielsen and Shephard (2012a) provides a more detailed mathematical treatment of the material we have developed in this Chapter, as well as other topics involving Lévy processes, that we develop later. Background information of semimartingale theory is provided in Barndorff-Nielsen and Shephard (2012b).

12 Exercises

Exercise 8 *Suppose that Y and Y' are independent gamma processes indexed by ν, α and ν, α'; show that $Y - Y'$ is a skewed normal gamma process. Hint, recall that*

$$\overline{\mathrm{K}}\left\{\theta \ddagger Y_1\right\} = \nu \log\left(1 + \frac{\theta}{\alpha}\right).$$

Exercise 9 *Suppose that Y and Y' are independent IG processes indexed by δ, γ and δ, γ'; show that $Y - Y'$ is not a NIG process.*

Exercise 10 *Suppose that Y is a compound Poisson process*

$$Y_t = \sum_{j=1}^{N_t} C_j.$$

Prove directly that

$$[Y]_t = Y_t^2 - 2\int_0^t Y_{u-}\mathrm{d}Y_u = \sum_{j=1}^{N_t} C_j^2.$$

Exercise 11 *Let X be a random variable such that*

$$Y = sign(X) |X|^{1/2}$$

follows the standard normal distribution. Compare the law of X to that of the symmetric NRIG distribution, as regards the behaviour of the probability densities at 0 and $\pm\infty$.

Exercise 12 *Let X and Y be semimartingales and suppose that they are related by*

$$\mathcal{E}(Y)_t = e^{X_t}.$$

Show that Y is a Lévy process if and only if X is a Lévy process.

Exercise 13 *Show that the Cauchy motion (that is the Lévy process which at time one follows the Cauchy law) can be represented as the sum of an NIG motion and an independent compound Poisson process. Give an extension of this result.*

Exercise 14 *Show that if Y is a Lévy process of type G and if its subordinator T has Blumenthal-Getoor index α then the Blumenthal-Getoor index of Y is 2α.*

13 Bibliographic notes

13.1 Lévy processes

Lévy processes were introduced by Lévy (1937) who pioneered the theory of infinite divisibility. Modern accounts of the probability theory of Lévy processes are given in Bertoin (1996), Sato (1999) and Applebaum (2004). See also Ito (2004), Rogers and Williams (1994, pp. 73–84) and Bertoin (1999). In his notes after each chapter Sato (1999) gives a detailed discussion of the historical development of the subject. A reasonably accessible overview of the theory and uses of Lévy processes is given in Barndorff-Nielsen, Mikosch, and Resnick (2001). A compact account in the context of finance is presented by Shiryaev (1999, pp. 200-206), while a more extensive discussion of the uses of Lévy processes in finance is given in Cont and Tankov (2004) and Boyarchenko and Levendorskii (2002b).

The simulation of Lévy processes has to be carried out with some care. There are extensive results available. Some of the most useful are the infinite series representation developed by Rosinski (2001), Rosinski (2002) and Rosinski (2007). The special case of gamma process simulation is discussed by Wolpert and Ickstadt (1999), while some more general discussion is given in Walker and Damien (2000). We should also note the important contribution of Asmussen and Rosinski (2001). Cont and Tankov (2004, Ch. 6) is an excellent source on these numerical issues.

The connections between Lévy processes and semimartingales are well discussed in some technical depth by for example Kallsen (2006). In turn this relies strongly on results in Jacod and Shiryaev (2003), see also Jacod (1979).

13.2 Flexible distributions

Most of modern financial economics is built out of Brownian motion and the corresponding Itô calculus. In this Chapter we have discussed many familiar alternative Lévy processes like the Poisson, normal gamma, Student's t, Laplace and normal inverse Gaussian laws. All these processes have been used as empirical models for log-prices. Throughout we have emphasised the normal variance-mean mixture distributions

$$Y = \mu + \beta\sigma^2 + \sigma U,$$

where U is a standard normal variable which is independent from a random σ^2. When $\beta = 0$ these variables are called normal variance mixtures and are very familiar in econometrics. The extension to $\beta \neq 0$ is important in financial economics as this allows us to model skewness. This approach to building non-Gaussian densities is attractive for it leads naturally into Lévy processes which have a subordination interpretation. A general discussion of these types of mixtures in statistics is given in Barndorff-Nielsen, Kent, and Sørensen (1982). When the distribution of σ^2 is infinitely divisible a natural name for this distribution is type G. On the other hand, Steutel and Van Harn (2004) call this type of distribution B(V), while Chaumont and Yor (2003) use the label "Gaussian transforms".

In early work Praetz (1972) and Blattberg and Gonedes (1974) suggested modelling the increments to log-prices using a Student's t distribution. This model was not set in continuous time, but we have seen above that it is possible to construct a Lévy process to justify this type of modelling. Further the model can be extended to allow for asymmetry. More recently Granger and Ding (1995) have advocated the use of Laplace distributions to model discrete time returns, while the non-linear Brownian motion based Cox, Ingersoll, and Ross (1985) processes have gamma marginals and so normal gamma distributions are often implicitly used in econometrics. It turns out that fitted values of the normal gamma distribution are typically thinner tailed, in fact sub-log-linear, than the corresponding Student, normal inverse Gaussian or the Laplace.

In this Chapter we have placed a great deal of emphasis on generalised hyperbolic and generalised inverse Gaussian distributions. We have carried this out for they support Lévy processes, are empirically flexible, encompass many of the familiar models econometricians are accustomed to and are mathematically tractable. However, their generality and some of the special cases are not so familiar.

The hyperbolic distribution and its extension to the generalised hyperbolic distribution was introduced in Barndorff-Nielsen (1977) in order to describe the size distribution of sand grains in samples of windblown sands. This was motivated by empirical observations due to R. A. Bagnold who noted that in double logarithmic plotting (that is both the horizontal and vertical axes are plotted on the logarithmic scale) the histograms looked strikingly as following hyperbolae, the slopes of the asymptotes being related to the physical conditions under which the sand was deposited; see Bagnold (1941) (note the similarity to the Granger and Ding (1995) empirics). Subsequently, it was discovered that the hyperbolic shape, or shapes very close to that, occur in a very wide range of empirical studies, for instance in other areas of geology, in turbulence, in paleomagnetism, in relativity theory and in biology. For a survey of developments up till the mid-1980ies, see Barndorff-Nielsen, Blæsild, Jensen, and Sørensen (1985). The generalised inverse Gaussian distribution is due to Étienne Halphen in 1946 (see the review article by Seshadri (1997)), while it was briefly discussed by Good (1953). A detailed discussion of this distribution was given by Jørgensen (1982).

Following a suggestion by Barndorff-Nielsen, Ernst Eberlein and coworkers began an investigation of the applicability of the generalised hyperbolic laws in finance and this has developed into a major project. For their work on this, see Eberlein and Keller (1995), Eberlein, Keller, and Prause (1998), Eberlein (2001), Eberlein and Özkan (2003b) and Prause (1999). Bauer (2000) discusses the use of these models in the context of value at risk.

When deviations from the hyperbolic shape occurred they typically showed somewhat heavier tails than the hyperbolic. This led Barndorff-Nielsen to consider more closely another of the generalised hyperbolic laws, the normal inverse Gaussian, which had until then received no attention, but turned out not only to fit a much wider range of data but also to possess various nice mathematical properties not shared by the hyperbolic (Barndorff-Nielsen (1997), Barndorff-Nielsen (1998b), Barndorff-Nielsen (1998a)). Sørensen (2006) presents a model for the development of the size distribution of sand under transport by wind, leading to the log NIG distribution.

The class of tempered stable distributions was introduced by Tweedie (1984). Hougaard (1986) discussed their applicability in survival analysis. See also Jørgensen (1987) and Brix (1999). The normal variance-mean mixtures with TS mixing was introduced by Barndorff-Nielsen and Shephard (2001b), who also extended this concept to the normal modified stable distributions. Barndorff-Nielsen and Shephard (2001a) used some of these distributions in their work on stochastic volatility.

Multivariate Student's t distributions have been used since their introduction by Mardia (1970, p. 92) and Zellner (1971, pp. 383-389). The multivariate GH distribution was first defined in Barndorff-Nielsen (1977) and was discussed in considerable detail and with a biological application by Blæsild (1981). A rather extensive discussion of the use of multivariate GH distributions in financial econometrics is given by Mencia and Sentana (2004). Alternative skewed multivariate Student's t distributions are provided by Fernández and Steel (1998), Bauwens and Laurent (2005) and Bera and Premaratne (2002). Jones and Faddy (2003) provides a discussion of much of the literature. Applications of the multivariate NIG distribution have been discussed by Aas, Haff, and Dimakos (2006) and Øigard, Hanssen, Hansen, and Godtliebsen (2005).

In a series of papers Kou (2002), Kou and Wang (2003), Kou and Wang (2004), Kou, Petrella, and Wang (2005), Heyde and Kou (2004) and Glasserman and Kou (2003).

The Skellam distribution was introduced by Irwin (1937). Related mathematical finance work is carried out by Kirch and Runggaldier (2004). Sichel (1973) and Sichel (1975) studied Poisson-IG distributions. The Skellam process was introduced by Barndorff-Nielsen, Pollard, and Shephard (2012) together with various integer valued extensions. They used it in the context of high frequency financial data.

Bondesson (1992) is also a key reference for infinite divisibility. It also includes a detailed account of Thorin's pioneering work (e.g. Thorin (1977) and Thorin (1978)) on the log-normal distribution and some of the important work that followed.

A simple generic method to sample from the GIG distribution has been derived by Dagpunar (1988, pp. 133-5) (see also Atkinson (1982)). Tempered stable variables are simulated by Devroye (2009) (see also Zhang (2011)).

13.3 Lévy processes in finance

The use of normal gamma based Lévy processes in finance was pioneered by Madan and Seneta (1990) and Madan, Carr, and Chang (1998) who paid particular attention to their use in option pricing. Recent extensions of this work include Carr, Geman, Madan, and Yor (2002).

The thicker tailed hyperbolic distribution and Lévy process was studied extensively in the context of finance by Eberlein and Keller (1995), who also discussed the corresponding option pricing theory and practice in Eberlein, Keller, and Prause (1998) and Eberlein (2001). This work is possible because the generalised inverse Gaussian distribution were shown to be infinitely divisible by Barndorff-Nielsen and Halgreen (1977). In fact it has the stronger property of being selfdecomposable and the same holds for the GH distribution, see Halgreen (1979).

The even thicker tailed normal inverse Gaussian process is studied by Barndorff-Nielsen (1995), Barndorff-Nielsen (1997), while Rydberg (1997b) and Rydberg (1997a) discusses both fitting the process to financial data and simulating from such processes. Prause (1999) and Raible (2000) have recently written first rate Ph.D. theses on generalised hyperbolic Lévy processes under the supervision of Ernst Eberlein. Both of these theses have a wealth of information on this topic. Bingham and Kiesel (2000) looks at the use of hyperbolic processes in finance, while Bibby and

Sørensen (2003) reviews the area of generalised hyperbolic processes in finance.

The idea of time deformation or subordination is due to Bochner (1949) and Bochner (1955), while it was introduced into economics by Clark (1973) who suggested the use of volume statistics as a subordinator, placing particular weight on studying the implications of using a lognormal subordinator. At that stage it was not known that this was a valid mathematical construction for it was not until Thorin (1977) that the lognormal was shown to be infinitely divisible. See also Bondesson (2002) for up to date treatment of lognormal Lévy processes. Epps and Epps (1976) and Tauchen and Pitts (1983) further studied the relationship between volume and the variance of the increments to prices. Recent discussions of this includes Ané and Geman (2000). Stock (1988) used the concept of subordination in a wider economic context outside finance. The *Student-OU* process, and other stationary processes with Student marginals, are discussed in Heyde and Leonenko (2005).

Mandelbrot (1963) and Mandelbrot and Taylor (1967) introduced the concepts of self-similarity and stable Lévy processes into financial economics. Almost immediately the main stream academic profession rejected these models, after some initial support from Fama (1965), due to their lack of empirical fit, as most research papers suggested the existence of at least two moments for returns. An elegant discussion of the move away from these models and its importance is given in Campbell, Lo, and MacKinlay (1997, pp. 17-19). However, there still remains a small group of researchers who push in this area. Recent work is discussed by Rachev and Mittnik (2000). Taleb (2007) has popularised some direct uses of stable models, although without any detailed empirical work to back them up in the financial context nor discussion of heavy tailed alternatives which are both tractable mathematically and empirically more appealing. Put simply stable processes are a poor way to produce financial Black Swans.

Truncated Lévy flights were introduced by Mantegna and Stanley (1994), while it has been pioneered in finance in Mantegna and Stanley (1996) and Mantegna and Stanley (2000). The extended Koponen class has been considered by Novikov (1994), Koponen (1995), Mantegna and Stanley (2000), Boyarchenko and Levendorskii (2002b), Boyarchenko and Levendorskii (2000), Boyarchenko and Levendorskii (2002c), Boyarchenko and Levendorskii (2002a), Barndorff-Nielsen and Levendorskii (2001), Carr, Geman, Madan, and Yor (2002), and Rosinski (2001). Carr, Geman, Madan, and Yor (2002) called these models CGMY processes after their own initials. We have not followed that nomenclature. Meixner distributions were introduced by Schoutens and Teugels (1998) and have been studied in the context of Lévy based models for finance by Schoutens and Teugels (2001) and Grigelionis (1999). Ben-Hamou (2000) has studied estimating the parameters of the Lévy process from option prices.

The use of Lévy processes for term structure and credit risk has been an active recent area. Work includes, for example, Eberlein and Raible (1999), Özkan (2002), Eberlein, Jacod, and Raible (2005), Barndorff-Nielsen, Christiansen, and Nicolato (2001) and Eberlein and Özkan (2003a), Schoutens and Cariboni (2009) and Eberlein and Kluge (2007). See also Bocker and Klüppelberg (2007) for an application to operational risk. This highly stimulating literature is beyond the scope of this book and so will not be discussed further.

Kijima (2002), Cont and Tankov (2004) and Schoutens (2003) are general books which discuss the use of Lévy processes and finance.

13.4 Empirical fit of Lévy processes

There is a large literature on studying the fit of various parametric models to the marginal distribution of returns of speculative assets. Most of these papers are not based on a background of a Lévy process and so risked fitting an incoherent (from a continuous time viewpoint) model. An example of this is Praetz (1972) in his work on the Student t distribution. Notable exceptions are Mandelbrot (1963), where he used stable distributions and related this to stable processes. See also Eberlein and Keller (1995), Eberlein (2001) and Eberlein and Özkan (2003b).

The likelihood methods we used to fit the models are entirely standard. We have used profile likelihoods. A discussion of this literature is given in Barndorff-Nielsen and Cox (1994, Section 3.4). The use of profile likelihoods for ν in the generalised hyperbolic is new as was the use of the EM algorithm in this context. As well as our own work on the EM algorithm for GH distributions, independent and concurrent work on the use of the EM algorithm for this problem was carried out by Protassov (2004). An elegant discussion of the EM algorithm is given in Tanner (1996). The theory of robust standard errors for maximum likelihood estimation is standard in econometrics and statistics. Leading references are White (1982) and White (1994).

Barndorff-Nielsen and Prause (2001) showed that the Olsen scaling law is explainable by the NIG Lévy process.

References

Aas, K., I. H. Haff, and X. K. Dimakos (2006). Risk estimation using the multivariate inverse Gaussian distribution. *Journal of Risk 8*.

Ané, T. and H. Geman (2000). Order flow, transaction clock and normality of asset returns. *Journal of Finance 55*, 2259–2284.

Applebaum, D. (2004). *Lévy Processes and Stochastic Calculus*. Cambridge: Cambridge University Press.

Asmussen, S. and J. Rosinski (2001). Approximation of small jumps of Lévy processes with a view towards simulation. *Journal of Applied Probability 38*, 482–493.

Atkinson, A. C. (1982). The simulation of generalised inverse Gaussian and hyperbolic random variables. *SIAM Journal of Scientific and Statistical Computing 3*, 502–17.

Bagnold, R. A. (1941). *The Physics of Blown Sand and Desert Dunes*. London: Chapman and Hall.

Barndorff-Nielsen, O. E. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences 353*, 401–419.

Barndorff-Nielsen, O. E. (1995). Normal inverse Gaussian processes and the modelling of stock returns. Research Report 300, Dept. Theor. Statistics, Aarhus University.

Barndorff-Nielsen, O. E. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics 24*, 1–14.

Barndorff-Nielsen, O. E. (1998a). Probability and statistics: selfdecomposability, finance and turbulence. In L. Accardi and C. C. Heyde (Eds.), *Probability Towards 2000. Proceedings of a Symposium held 2-5 October 1995 at Columbia University*, pp. 47–57. New York: Springer-Verlag.

Barndorff-Nielsen, O. E. (1998b). Processes of normal inverse Gaussian type. *Finance and Stochastics 2*, 41–68.

Barndorff-Nielsen, O. E., P. Blæsild, J. L. Jensen, and M. Sørensen (1985). The fascination of sand. In A. C. Atkinson and S. E. Fienberg (Eds.), *A Celebration of Statistics*, pp. 57–87. New York: Springer-Verlag.

Barndorff-Nielsen, O. E., B. J. Christiansen, and E. Nicolato (2001). General stochastic volatility and explicit option prices in Heath-Jarrow-Morton term structure analysis. Unpublished paper, Department of Mathematical Sciences, Aarhus University.

Barndorff-Nielsen, O. E. and D. R. Cox (1994). *Inference and Asymptotics*. London: Chapman & Hall.

Barndorff-Nielsen, O. E. and C. Halgreen (1977). Infinite divisibility of the hyperbolic and generalised inverse Gaussian distributions. *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete 38*, 309–311.

Barndorff-Nielsen, O. E., J. T. Kent, and M. Sørensen (1982). Normal variance-mean mixtures and z distributions. *International Statistical Review 50*, 145–159.

Barndorff-Nielsen, O. E. and S. Levendorskii (2001). Feller processes of normal inverse Gaussian type. *Quantitative Finance 1*, 318–331.

Barndorff-Nielsen, O. E., T. Mikosch, and S. Resnick (2001). *Lévy Processes – Theory and Applications.* Boston: Birkhäuser.

Barndorff-Nielsen, O. E., D. G. Pollard, and N. Shephard (2012). Integer-valued Lévy processes and low latency financial econometrics. *Quantitative Finance 12*, 587–605.

Barndorff-Nielsen, O. E. and K. Prause (2001). Apparent scaling. *Finance and Stochastics 5*, 103–113.

Barndorff-Nielsen, O. E. and N. Shephard (2001a). Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics (with discussion). *Journal of the Royal Statistical Society, Series B 63*, 167–241.

Barndorff-Nielsen, O. E. and N. Shephard (2001b). Normal modified stable processes. *Theory of Probability and Mathematical Statistics 65*, 1–19.

Barndorff-Nielsen, O. E. and N. Shephard (2004). Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial Econometrics 2*, 1–48.

Barndorff-Nielsen, O. E. and N. Shephard (2012a). Mathematics of Lévy based processes. Unpublished chapter of the authors' book on Lévy Driven Volatility Models.

Barndorff-Nielsen, O. E. and N. Shephard (2012b). Primer on stochastic analysis. Unpublished chapter of the authors' book on Lévy Driven Volatility Models.

Barndorff-Nielsen, O. E. and N. Shephard (2012c). Toolboxes for financial processes: distributions. Unpublished chapter of the authors' book on Lévy Driven Volatility Models.

Bauer, C. (2000). Value at risk using hyperbolic distributions. *Journal of Economics and Business 52*, 455–467.

Bauwens, L. and S. Laurent (2005). A new class of multivariate skew densities, with application to generalised autoregressive conditional heteroscedasticity models. *Journal of Business and Economic Statistics 23*, 346–354.

Ben-Hamou, E. (2000). Option pricing with Lévy process. Financial Markets Group Working Paper.

Bera, A. and G. Premaratne (2002). Modelling asymmetry and excess kurtosis in stock return data. Unpublished paper, Economics Department, University of Illinois.

Bertoin, J. (1996). *Lévy Processes.* Cambridge: Cambridge University Press.

Bertoin, J. (1999). *Subordinators: Examples and Applications.* Berlin: Springer. Presented at Ecole d'Ete de St-Flour, 1997. Lecture Notes in Mathematics.

Bibby, B. M. and M. Sørensen (2003). Hyperbolic processes in finance. In S. Rachev (Ed.), *Handbook of Heavy Tailed Distributions in Finance*, pp. 211–248. Elsevier Science.

Bingham, N. and R. Kiesel (2000). Modelling asset returns with hyperbolic distributions. In J. Knight and S. E. Satchell (Eds.), *Asset Return Distributions*, pp. 1–20. Butterworth-Heinemann.

Blæsild, P. (1981). The two-dimensional hyperbolic distribution and related distributions, with an application to the Johannsen bean data. *Biometrika 68*, 251–263.

Blattberg, R. C. and N. J. Gonedes (1974). A comparison of the stable and student distributions as models for stock prices. *Journal of Business 47*, 244–280.

Bochner, S. (1949). Diffusion equation and stochastic processes. *Proceedings of the National Academy of Science of the United States of America 85*, 369–370.

Bochner, S. (1955). *Harmonic Analysis and the Theory of Probability.* Berkeley: University of California Press.

Bocker, K. and C. Klüppelberg (2007). Modelling and measuring multivariate operational risk with Lévy copulas. *Journal of Operational Risk 3*, 3–27.

Bondesson, L. (1992). *Generalized Gamma Convolutions and Related Classes of Distributions and Densities.* Heidelberg: Springer-Verlag.

Bondesson, L. (2002). On the Lévy measure of the lognormal and the logcauchy distributions. *Methodology and Computing in Applied Probability 4*, 243–256.

Boyarchenko, S. I. and S. Z. Levendorskii (2000). Option pricing for truncated Lévy processes. *International Journal of Theoretical and Applied Finance 3*, 549–552.

Boyarchenko, S. I. and S. Z. Levendorskii (2002a). Barrier options and touch-and-out options under regular Lévy processes of exponential type. *Annals of Applied Probability 12*, 1261–1298.

Boyarchenko, S. I. and S. Z. Levendorskii (2002b). *Non-Gaussian Merton-Black-Scholes Theory.* Singapore: World Scientific.

Boyarchenko, S. I. and S. Z. Levendorskii (2002c). Perpetual American options under Lévy processes. *SIAM Journal on Control and Optimization 40*, 1663–1696.

Brix, A. (1999). Generalized gamma measures and shot-noise Cox processes. *Advances in Applied Probability 31*, 929–953.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997). *The Econometrics of Financial Markets.* Princeton, New Jersey: Princeton University Press.

Carr, P., H. Geman, D. B. Madan, and M. Yor (2002). The fine structure of asset returns: an empirical investigation. *Journal of Business 75*, 305–332.

Chaumont, L. and M. Yor (2003). *Exercises in Probability.* Cambridge: Cambridge University Press.

Clark, P. K. (1973). A subordinated stochastic process model with fixed variance for speculative prices. *Econometrica 41*, 135–156.

Cochrane, J. H. (2001). *Asset Pricing.* Princeton: Princeton University Press.

Cont, R. and P. Tankov (2004). *Financial Modelling with Jump Processes.* London: Chapman and Hall.

Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985). A theory of the term structure of interest rates. *Econometrica 53*, 385–407.

Dagpunar, J. (1988). *Principles of Random Number Generation.* Oxford: Oxford University Press.

Devroye, L. (2009). Random variable generation for exponentially and polynomially tilted stable distributions. *ACM Transactions on Modeling and Computer Simulation 19*, 18:1–18:20.

Doornik, J. A. (2001). *Ox: Object Oriented Matrix Programming, 5.0.* London: Timberlake Consultants Press.

Eberlein, E. (2001). Application of generalized hyperbolic Lévy motion to finance. In O. E. Barndorff-Nielsen, T. Mikosch, and S. Resnick (Eds.), *Lévy Processes – Theory and Applications*, pp. 319–337. Boston: Birkhäuser.

Eberlein, E., J. Jacod, and S. Raible (2005). Lévy term structure models: no-arbitrage and completeness. *Finance and Stochastics 9*, 67–88.

Eberlein, E. and U. Keller (1995). Hyperbolic distributions in finance. *Bernoulli 1*, 281–299.

Eberlein, E., U. Keller, and K. Prause (1998). New insights into smile, mispricing and value at risk: the hyperbolic model. *Journal of Business 71*, 371–405.

Eberlein, E. and W. Kluge (2007). Calibration of Lévy term structure models. In M. Fu, R. A. Jarrow, J.-Y. Yen, and R. J. Elliott (Eds.), *In Advances in Mathematical Finance In Honor of D. B. Madan*, pp. 147–172. Birkhauser.

Eberlein, E. and F. Özkan (2003a). The defaultable Lévy term structure: ratings and restructing. *Mathematical Finance 13*, 277–300.

Eberlein, E. and F. Özkan (2003b). Time consistency of Lévy models. *Quantitative Finance 3*, 40–50.

Eberlein, E. and S. Raible (1999). Term structure models driven by general Lévy processes. *Mathematical Finance 9*, 31–53.

Epps, T. W. and M. L. Epps (1976). The stochastic dependence of security price changes and transaction volumes: implications for the mixture-of-distributions hypothesis. *Econometrica 44*, 305–321.

Fama, E. F. (1965). The behaviour of stock market prices. *Journal of Business 38*, 34–105.

Feller, W. (1971a). *An Introduction to Probability Theory and Its Applications* (2 ed.), Volume 1. New York: John Wiley.

Feller, W. (1971b). *An Introduction to Probability Theory and Its Applications* (2 ed.), Volume 2. New York: John Wiley.

Fernández, C. and M. F. J. Steel (1998). On Bayesian modelling of fat tails and skewness. *Journal of the American Statistical Association 93*, 359–371.

Glasserman, P. and S. G. Kou (2003). The term structure of simple foward rates with jump risk. *Mathematical Finance 13*, 383–410.

Good, I. J. (1953). The population frequencies of species and the estimation of population parameters. *Biometrika 40*, 237–264.

Granger, C. W. J. and Z. Ding (1995). Some properties of absolute returns, an alternative measure of risk. *Annals d'Economie et de Statistique 40*, 67–91.

Grigelionis, B. (1999). Processes of Meixner type. *Lithuanian Mathematics Journal 39*, 33–41.

Halgreen, C. (1979). Self-decomposability of the generalized inverse Gaussian and hyperbolic distributions. *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete 47*, 13–17.

Heyde, C. C. and S. G. Kou (2004). On the controversy over tailweight of distributions. *Operations Research Letters 32*, 399–408.

Heyde, C. C. and N. N. Leonenko (2005). Student processes. *Advances in Applied Probability 37*, 342–365.

Hougaard, P. (1986). Survival models for heterogeneous populations derived from stable distributions. *Biometrika 73*, 387–396.

Irwin, J. O. (1937). The frequency distribution of the difference between two independent variates following the same Poisson distribution. *Journal of the Royal Statistical Society, Series A 100*, 415–416.

Ito, K. (2004). *Stochastic Processes: Lectures given at Aarhus University*. Springer.

Jacod, J. (1979). *Calul Stochastique et Problems de Martingales*. Number 714 in Lecture Notes in Mathematics. Berlin: Springer.

Jacod, J. and P. Protter (2012). *Discretization of Processes*. Heidelberg: Springer.

Jacod, J. and A. N. Shiryaev (2003). *Limit Theorems for Stochastic Processes* (2 ed.). Springer: Berlin.

Jørgensen, B. (1982). *Statistical Properties of the Generalised Inverse Gaussian Distribution*. New York: Springer-Verlag.

Jørgensen, B. (1987). Exponential dispersion models (with discussion). *Journal of the Royal Statistical Society, Series B 49*, 127–162.

Jones, M. C. and M. J. Faddy (2003). A skew extension of the t-distribution, with applications. *Journal of the Royal Statistical Society, Series B 65*, 159–174.

Kallsen, J. (2006). A didactic note on affine stochastic volatility models. In Y. Kabanov, R. Lipster, and J. Stoyanov (Eds.), *From Stochastic Calculus to Mathematical Finance. The Shiryaev Festschrift*, pp. 343–368. Heidelberg: Springer.

Kijima, M. (2002). *Stochastic Processes with Applications to Finance*. London: Chapman and Hall.

Kirch, M. and W. Runggaldier (2004). Efficient hedging when asset prices follow a geometric Poisson process with unknown intensities. *SIAM Journal on Control and Optimization 43*, 1174–1195.

Koponen, I. (1995). Analytic approach to the problem of convergence of truncated Lévy flights towards the Gaussian stochastic process. *Physics Review E 52*, 1197–1199.

Kou, S. G. (2002). A jump-diffusion model for option pricing. *Management Science 48*, 1086–1101.

Kou, S. G., G. Petrella, and H. Wang (2005). Pricing path-dependent options with jump risk via laplace transforms. *Kyoto Economic Review 74*, 1–23.

Kou, S. G. and H. Wang (2003). First passage time of a jump diffusion process. *Advances in Applied Probability 35*, 504–531.

Kou, S. G. and H. Wang (2004). Option pricing under a double exponential jump-diffusion model. *Management Science 50*, 1178–1192.

Lévy, P. (1937). *Théories de L'Addition Aléatories*. Paris: Gauthier-Villars.

Madan, D. B., P. Carr, and E. Chang (1998). The variance gamma process and option pricing. *European Finance Review 2*, 79–105.

Madan, D. B. and E. Seneta (1990). The VG model for share market returns. *Journal of Business 63*, 511–524.

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business 36*, 394–419.

Mandelbrot, B. and H. Taylor (1967). On the distribution of stock price differences. *Operations Research 15*, 1057–1062.

Mantegna, R. and H. E. Stanley (2000). *Introduction to Econophysics: Correlation and Complexity in Finance*. Cambridge: Cambridge University Press.

Mantegna, R. N. and H. E. Stanley (1994). Stochastic process with ultraslow convergence to a Gaussian: the truncated Lévy flight. *Physics Reviews Letters 73*, 2946–2949.

Mantegna, R. N. and H. E. Stanley (1996). Turbulence and financial markets. *Nature 383*, 587–588.

Mardia, K. V. (1970). *Families of Bivariate Distributions*. Darien, Connecticut: Hafner Publishing Co.

Mencia, F. J. and E. Sentana (2004). Estimating and testing of dynamic models with generalised hyperbolic innovations. Unpublished paper: CEMFI, Madrid.

Novikov, E. A. (1994). Infinitely divisible distributions in turbulence. *Physics Reviews E 50*, R3303–R3305.

Øigard, T. A., A. Hanssen, R. E. Hansen, and F. Godtliebsen (2005). EM algorithm and modeling of heavy-tailed processes with the multivariate normal inverse Gaussian distribution. *Signal Processing 85*, 1655–1673.

Özkan, F. (2002). Lévy processes in credit risk and market models. Unpublished Ph.D. thesis, Freiburg University.

Praetz, P. D. (1972). The distribution of share prices. *Journal of Business 45*, 45–55.

Prause, K. (1999). The generalised hyperbolic models: estimation, financial derivatives and risk measurement. PhD Thesis. Mathematics Faculty, Freiburg University.

Protassov, R. S. (2004). EM-based maximum likelihood parameter estimation for multivariate generalized hyperbolic distributions with fixed λ. *Statistics and Computing 14*, 67–77.

Rachev, S. and S. Mittnik (2000). *Stable Paretian Models in Finance*. New York: Wiley.

Raible, S. (2000). Lévy processes in finance: Theory, numerics and empirical facts. PhD Thesis. Mathematics Faculty, Freiburg University.

Rogers, L. C. G. and D. Williams (1994). *Diffusions, Markov Processes and Martingales. Volume 1, Foundations* (2 ed.). Chichester: Wiley.

Rosinski, J. (2001). Series representations of Lévy processes from the perspective of point processes. In O. E. Barndorff-Nielsen, T. Mikosch, and S. Resnick (Eds.), *Lévy Processes – Theory and Applications*, pp. 401–415. Boston: Birkhäuser.

Rosinski, J. (2002). Tempered stable processes. In O. E. Barndorff-Nielsen (Ed.), *Miniproceedings from 2nd MaPhySto Conference on Lévy Processes: Theory and Applications*, pp. 215–220. Aarhus: http://www.maphysto.dk/publications/MPS-misc/2002/22.pdf.

Rosinski, J. (2007). Tempering stable processes. *Stochastic Processes and Their Applications 117*, 677–707.

Rydberg, T. H. (1997a). Existence of unique equivalent martingale measures in a Markovian setting. *Finance and Stochastics 1*, 251–257.

Rydberg, T. H. (1997b). The normal inverse Gaussian Lévy process: simulation and approximation. *Communications in Statistics: Stochastic Models 13*, 887–910.

Samorodnitsky, G. and M. S. Taqqu (1994). *Stable Non-Gaussian Random Processes*. New York: Chapman and Hall.

Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge: Cambridge University Press.

Schoutens, W. (2003). *Lévy Processes in Finance: Pricing Financial Derivatives*. New York: Wiley.

Schoutens, W. and J. Cariboni (2009). *Lévy Processes in Credit Risk*. New York: Wiley.

Schoutens, W. and J. L. Teugels (1998). Lévy processes, polynomials and martingales. *Statistics and Stochastic Models 14*, 335–349.

Schoutens, W. and J. L. Teugels (2001). Meixner processes in finance. Unpublished paper: Eurandom, K.U. Leuven.

Seshadri, S. (1997). Halphen's laws. In S. Kotz (Ed.), *Encyclopedia of Statistical Sciences*, pp. 302–306. New York: Wiley.

Shiryaev, A. N. (1999). *Essentials of Stochastic Finance: Facts, Models and Theory*. Singapore: World Scientific.

Sichel, H. S. (1973). Statistical valuation of diamondiferous deposits. *Journal of South African Institute of Mining and Metallurgy 73*, 235–243.

Sichel, H. S. (1975). On a distribution representing sentence-length in written prose. *Journal of the Royal Statistical Society, Series A 137*, 25–34.

Sørensen, M. (2006). On the size distribution of sand. Unpublished paper, Department of Department of Applied Mathematics and Statistics, University of Copenhagen.

Steutel, F. W. and K. Van Harn (2004). *Infinitely Divisibility of Probability Distributions on the Real Line*. Marcel Dekker.

Stock, J. H. (1988). Estimating continuous-time processes subject to time deformation: an application to postwar U.S. GNP. *Journal of the American Statistical Association 83*, 77–85.

Taleb, N. N. (2007). *The Black Swan*. Penguin.

Tanner, M. A. (1996). *Tools for Statistical Inference: Methods for Exploration of Posterior Distributions and Likelihood Functions* (3 ed.). New York: Springer-Verlag.

Tauchen, G. and M. Pitts (1983). The price variability-volume relationship on speculative markets. *Econometrica 51*, 485–505.

Thorin, O. (1977). On the infinite divisibility of the lognormal distribution. *Scandinavian Actuarial Journal 47*, 121–148.

Thorin, O. (1978). An extension of the notion of a generalized γ-convolution. *Scandinavian Actuarial Journal 48*, 141–149.

Tweedie, M. (1984). An index which distinguishes between some important exponential families. In J. Ghosh and J. Roy (Eds.), *Statistics: Applications and New Directions: Proc. Indian Statistical Institute Golden Jubilee International Conference*, pp. 579–604.

Walker, S. and P. Damien (2000). Representations of Lévy processes without Gaussian components. *Biometrika 87*, 477–483.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica 50*, 1–25.

White, H. (1994). *Estimation, Inference and Specification Analysis*. Cambridge: Cambridge University Press.

Wolpert, R. L. and K. Ickstadt (1999). Simulation of Lévy random fields. In D. D. Dey, M. Muller, and D. Sinha (Eds.), *Practical Nonparametric and Semiparametric Bayesian Statistics*, pp. 227–242. New York: Springer-Verlag.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*. New York: John Wiley and Sons.

Zhang, S. (2011). Exact simulation of tempered stable ornstein-uhlenbeck processes. *Journal of Statistical Computation and Simulation 81*, 1533–1544.