

Hyperbolic Processes in Finance

Bo Martin Bibby

Institute of Mathematics and Physics
The Royal Veterinary and Agricultural University
Thorvaldsensvej 40
DK-1871 Frederiksberg C, Denmark

Michael Sørensen

Department of Statistics and Operations Research
Institute of Mathematical Sciences
University of Copenhagen
Universitetsparken 5
DK-2100 København Ø, Denmark

1 Introduction

Distributions that have tails heavier than the normal distribution are ubiquitous in finance. For purposes such as risk management and derivative pricing it is important to use relatively simple models that can capture the heavy tails and other relevant features of financial data. A class of distributions that is very often able to fit the distributions of financial data is the class of generalized hyperbolic distributions. This has been established in numerous investigations, see e.g. Eberlein & Keller (1995), Bibby & Sørensen (1997), Hurst (1997), Eberlein, Keller & Prause (1998), Rydberg (1999), Küchler et al. (1999), Jiang (2000), and Barndorff-Nielsen & Shephard (2001d). The class of generalized hyperbolic distributions includes the standard hyperbolic distributions, the normal inverse Gaussian distributions, the scaled t-distributions and the variance-gamma distributions. The use of scaled t-distributions in finance was studied by Praetz (1972) and Blattberg & Gonedes (1974), while Madan & Seneta (1990) introduced the variance-gamma distributions in the financial literature. The normal distribution appears as a limit of generalized hyperbolic distributions. The tail behaviour of the generalized hyperbolic distributions thus span a range from Gaussian tails via exponential tails to the power tails of the t-distributions.

In the Section 2 we present the generalized hyperbolic distributions and their most important properties. We also discuss the generalized inverse Gaussian distributions which play an important role in the theory of generalized hyperbolic distributions and processes. This class of distributions is also of interest in its own right as a model of positive quantities in finance. Its right hand tail behaviour spans a range from exponential decrease to a Pareto tail. In the following sections we present a number of stochastic process models for which marginal distributions or distributions of increments (or both) are generalized hyperbolic. The models are

increasingly complex. They are thus able to fit an increasing number of the stylized features of financial data. The well established features of financial data are for instance reviewed in Hurst (1997), Barndorff-Nielsen (1998), and Rydberg (2000). In Section 3 we discuss Lévy process models, while in Section 4 we discuss models defined by stochastic differential equations. These include classical diffusion models and Ornstein-Uhlenbeck models driven by Lévy processes as well as superpositions of such models. In the final Section 5 we present generalized hyperbolic stochastic volatility models.

2 Hyperbolic and related distributions

In this section we present the *generalized hyperbolic distributions* and describe their most important properties. We will also discuss the *generalized inverse Gaussian distributions* which play an important role in the theory of generalized hyperbolic distributions and processes. As mentioned earlier, this class of distributions is also of independent interest as a model of positive quantities in finance. We will present a few examples of how well these distributions fit financial data.

2.1 The generalized hyperbolic distribution

The generalized hyperbolic distributions were introduced by Barndorff-Nielsen (1977) and include, among others, the *hyperbolic distribution*, the *normal-inverse Gaussian (NIG) distribution*, the *scaled t -distributions* and the *variance-gamma distributions*. We shall discuss these sub-classes in more detail later. First we present the generalized hyperbolic distributions and their properties.

The generalized hyperbolic distribution has five parameters. If X follows a generalized hyperbolic distribution we write

$$X \sim H(\lambda, \alpha, \beta, \delta, \mu).$$

The probability density function of the generalized hyperbolic distribution is given by

$$\frac{(\gamma/\delta)^\lambda}{\sqrt{2\pi}K_\lambda(\delta\gamma)} \cdot \frac{K_{\lambda-\frac{1}{2}}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)}{\left(\sqrt{\delta^2 + (x-\mu)^2}/\alpha\right)^{\frac{1}{2}-\lambda}} \cdot e^{\beta(x-\mu)}, \quad x \in \mathbb{R}, \quad (2.1)$$

where $\gamma^2 = \alpha^2 - \beta^2$, and K_λ is the modified Bessel function of the third kind with index λ . Definitions and results concerning Bessel functions are collected in an Appendix.

The parameter domain for the generalized hyperbolic distribution is given by

$$\begin{aligned} \delta &\geq 0, & \alpha &> 0, & \alpha^2 &> \beta^2, & \text{if } \lambda > 0, \\ \delta &> 0, & \alpha &> 0, & \alpha^2 &> \beta^2, & \text{if } \lambda = 0, \\ \delta &> 0, & \alpha &\geq 0, & \alpha^2 &\geq \beta^2, & \text{if } \lambda < 0, \end{aligned}$$

In all cases $\mu \in \mathbb{R}$. If $\delta = 0$ or $\alpha^2 = \beta^2$ the generalized hyperbolic density in (2.1) is defined as the limit expression obtained by using (A.5). Note that if β is equal to zero, the distribution is symmetric.

The class of generalized hyperbolic distributions is closed under affine transformation. That is, if $X \sim H(\lambda, \alpha, \beta, \delta, \mu)$ and Y is defined as $Y = aX + b$, for some positive a , then we have that

$$Y \sim H(\lambda, \alpha/a, \beta/a, a\delta, a\mu + b). \quad (2.2)$$

From (2.2) we also see that the parameter λ is invariant under affine transformation of the generalized hyperbolic random variable.

From (A.3) it follows that the mode points for the generalized hyperbolic distribution are solutions to the equation

$$\frac{x - \mu}{\sqrt{\delta^2 + (x - \mu)^2}} \cdot \frac{K_{\lambda - \frac{3}{2}}(\alpha \sqrt{\delta^2 + (x - \mu)^2})}{K_{\lambda - \frac{1}{2}}(\alpha \sqrt{\delta^2 + (x - \mu)^2})} = \frac{\beta}{\alpha}. \quad (2.3)$$

If $\beta = 0$, it follows immediately that the distribution is unimodal with mode point μ . If $\lambda \geq \frac{3}{2}$ then the ratio of the modified Bessel functions in (2.3) increases monotonically from 0 to 1, and therefore the distribution is unimodal. See Blæsild (1978) for further discussion of features of the generalized hyperbolic density function.

The Laplace transform for the generalized hyperbolic distribution is given by

$$L(z) = e^{\mu z} \cdot \frac{\gamma^\lambda \cdot K_\lambda(\delta \gamma_z)}{\gamma_z^\lambda \cdot K_\lambda(\delta \gamma)}, \quad |\beta + z| < \alpha, \quad (2.4)$$

where $\gamma_z^2 = \alpha^2 - (\beta + z)^2$. From (A.3) we get that

$$EX = \mu + \frac{\delta \beta K_{\lambda+1}(\delta \gamma)}{\gamma K_\lambda(\delta \gamma)}, \quad (2.5)$$

and

$$\text{Var}X = \frac{\delta K_{\lambda+1}(\delta \gamma)}{\gamma K_\lambda(\delta \gamma)} + \frac{\beta^2 \delta^2}{\gamma^2} \left(\frac{K_{\lambda+2}(\delta \gamma)}{K_\lambda(\delta \gamma)} - \frac{K_{\lambda+1}^2(\delta \gamma)}{K_\lambda^2(\delta \gamma)} \right). \quad (2.6)$$

Expressions for the skewness and kurtosis involve modified Bessel functions in a rather complicated way and can be found in Barndorff-Nielsen & Blæsild (1980).

Sometimes it is useful to reparametrize the generalized hyperbolic density in terms of the parameters λ , τ , ζ , δ , and μ , where $\tau = \beta/\gamma$ and $\zeta = \delta\gamma$. Using this parametrization, the generalized hyperbolic density has the form,

$$\frac{\sqrt{\zeta}}{\sqrt{2\pi\delta}K_\lambda(\zeta)} \cdot \frac{K_{\lambda - \frac{1}{2}}\left(\zeta\sqrt{1 + \tau^2}\sqrt{1 + \left(\frac{x - \mu}{\delta}\right)^2}\right)}{\left(\sqrt{1 + \left(\frac{x - \mu}{\delta}\right)^2}/\sqrt{1 + \tau^2}\right)^{\frac{1}{2} - \lambda}} \cdot e^{\tau\zeta\frac{x - \mu}{\delta}}, \quad x \in \mathbb{R}. \quad (2.7)$$

The parameters τ , ζ , and λ are invariant under affine transformations of a random variable following the generalized hyperbolic distribution. More precisely, we have that the result equivalent to (2.2) is that $Y \sim H(\lambda, \tau, \zeta, a\delta, a\mu + b)$. From this result we see that δ is a scaling parameter and μ is a location parameter. In Figure 2.1 generalized hyperbolic densities are drawn for different values of λ , τ , and ζ . In all cases the mean value is 0 and the variance is 1.

The tail behaviour of the distributions is more easily seen in Figure 2.2, where the logarithm of the same densities are plotted.

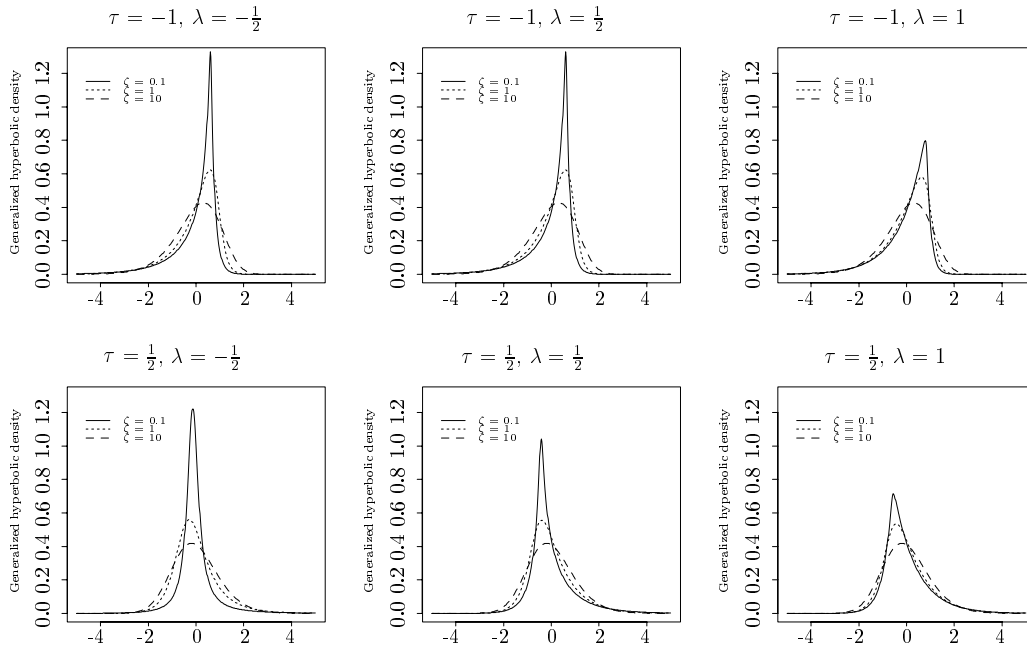


Figure 2.1: Generalized hyperbolic densities with mean 0 and variance 1 for different values of the parameters λ , τ , and ζ .

We shall now consider the important special cases of the generalized hyperbolic distribution mentioned earlier. The *hyperbolic distributions* is the subclass obtained when λ is equal to 1. With λ equal to 1 in (2.1), we get the following expression for the density of the hyperbolic distribution,

$$\frac{\gamma}{2\alpha\delta K_1(\delta\gamma)} \exp \left\{ -\alpha\sqrt{\delta^2 + (x - \mu)^2} + \beta(x - \mu) \right\}, \quad x \in \mathbb{R}. \quad (2.8)$$

From (2.8) we see that the log-density of the hyperbolic distribution is a hyperbola, which should be compared to the parabolic log-density of the normal distribution. The name of the hyperbolic distribution stems from this observation. In fact, the definition of the hyperbolic distributions was inspired by the empirical finding of the founding father of the physics of wind blown sand, Brigadier R. A. Bagnold, that the log-density of the distribution of the logarithm of the grain size of natural sand deposits looks more like a hyperbola than like a parabola, as had previously been assumed by geomorphologists, see Bagnold (1941).

For the hyperbolic distributions the equation (2.3), which determines the mode points of the generalized hyperbolic distribution, simplifies to

$$\frac{x - \mu}{\sqrt{\delta^2 + (x - \mu)^2}} = \frac{\beta}{\alpha},$$

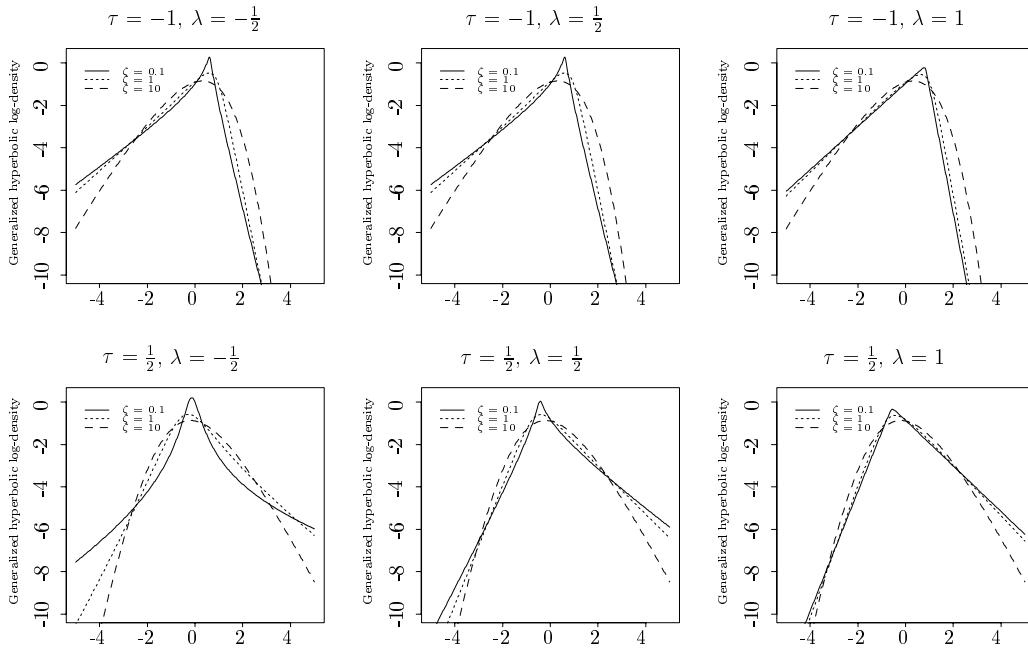


Figure 2.2: The logarithm of generalized hyperbolic densities with mean 0 and variance 1 for different values of the parameters λ , τ , and ζ .

giving that the distribution is unimodal with mode point

$$x = \mu + \frac{\delta\beta}{\gamma}.$$

Letting δ tend to zero and using (A.5), we get the asymmetric Laplace distribution as a special case of the hyperbolic distribution, that is,

$$\frac{\alpha^2 - \beta^2}{2\alpha} e^{\beta(x-\mu) - \alpha|x-\mu|}, \quad x \in \mathbb{R}.$$

The normal distribution can also be obtained as a limit case of the hyperbolic distribution. Letting $\alpha, \delta \rightarrow \infty$ in such a way that $\delta/\alpha \rightarrow \sigma^2$, we get, using (A.6), the normal density:

$$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad x \in \mathbb{R}.$$

According to Barndorff-Nielsen et al. (1985) we have that the skewness (γ_1) and kurtosis (γ_2) for large values of ζ and small values of β/α satisfy

$$(\gamma_1, \gamma_2) \sim (3\chi, 3\xi^2),$$

where

$$\chi = \frac{\beta/\alpha}{\sqrt{1+\zeta}}$$

and

$$\xi = \frac{1}{\sqrt{1+\zeta}}.$$

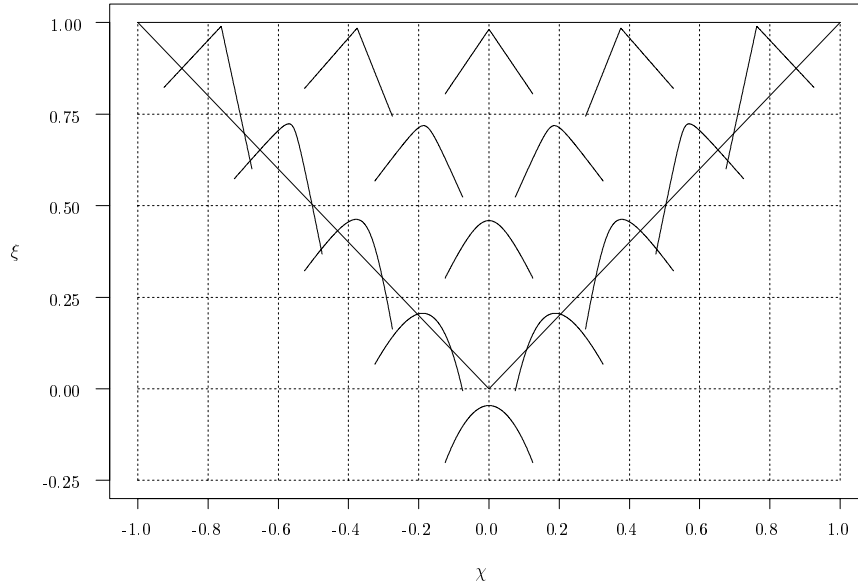


Figure 2.3: Hyperbolic log densities with mean 0 and variance 1 for different values of the parameters χ and ξ ($-0.8, -0.6, \dots, 0.8$ for χ and $0.0, 0.25, \dots, 1.0$ for ξ). The log densities are placed at the corresponding values of χ and ξ .

Based on this observation Barndorff-Nielsen et al. (1985) suggested that the parameters χ and ξ are natural measures of asymmetry and “kurtosis” for the hyperbolic distribution. Note that they are invariant under location-scale transformations. The parameters χ and ξ vary in the so-called **shape triangle** defined by

$$\{(\chi, \xi) \in \mathbb{R}^2 \mid 0 \leq |\chi| < \xi < 1\}. \quad (2.9)$$

Note that the normal and the (possibly skew) Laplace distributions are obtained as limit distributions when $\xi \rightarrow 1$ and $\xi \rightarrow 0$, respectively. In Figure 2.3 hyperbolic log density functions are drawn for different values of χ and ξ in the shape triangle.

In Figure 2.4 a histogram based on 2666 observations of the daily returns of IBM-stocks (increments on a logarithmic scale of the stock prices) in the period from the 1st of January 1990 to the 20th of March 2000 is given. Each point indicates the mid-point of the top of a column in the histogram. The best generalized hyperbolic, hyperbolic, and normal densities are superimposed on the histogram. The parameter values corresponding to the generalized hyperbolic density are $\alpha = 5.174$, $\beta = 0.0048$, $\delta = 0.0262$, $\mu = 0.0002$, and $\lambda = -1.933$. For the hyperbolic density the parameter values are $\alpha = 82.26$, $\beta = 3.725$, $\delta = 0.0060$, and $\mu = -0.0007$. In Figure 2.5 the logarithms of the same histogram points and the same densities are plotted.

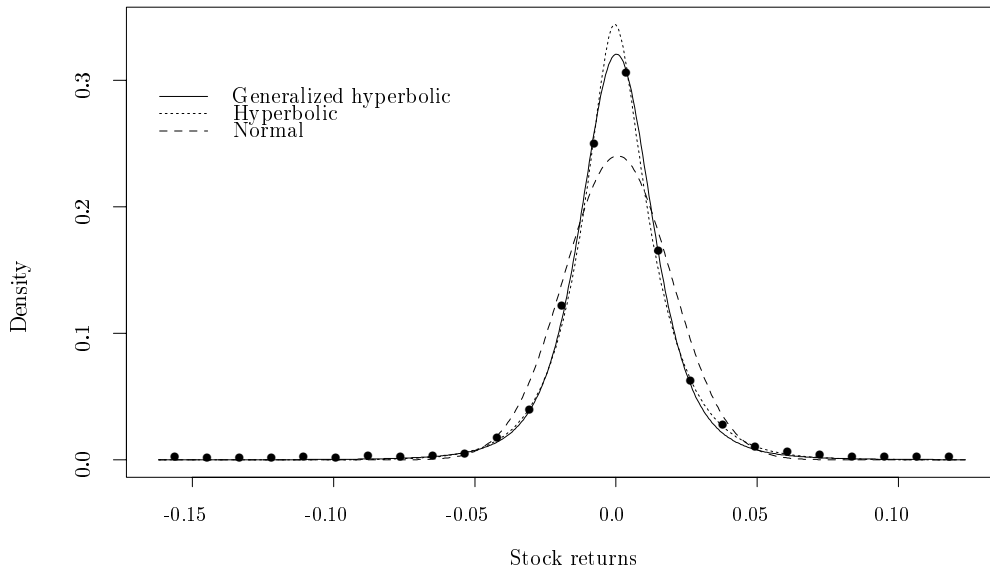


Figure 2.4: A histogram of 2666 daily IBM-stock returns. Superimposed are the best fitting generalized hyperbolic, hyperbolic, and normal densities. The parameter values corresponding to the generalized hyperbolic density are $\alpha = 5.174$, $\beta = 0.0048$, $\delta = 0.0262$, $\mu = 0.0002$, and $\lambda = -1.933$. The parameter values corresponding to the hyperbolic density are $\alpha = 82.26$, $\beta = 3.725$, $\delta = 0.0060$, and $\mu = -0.0007$.

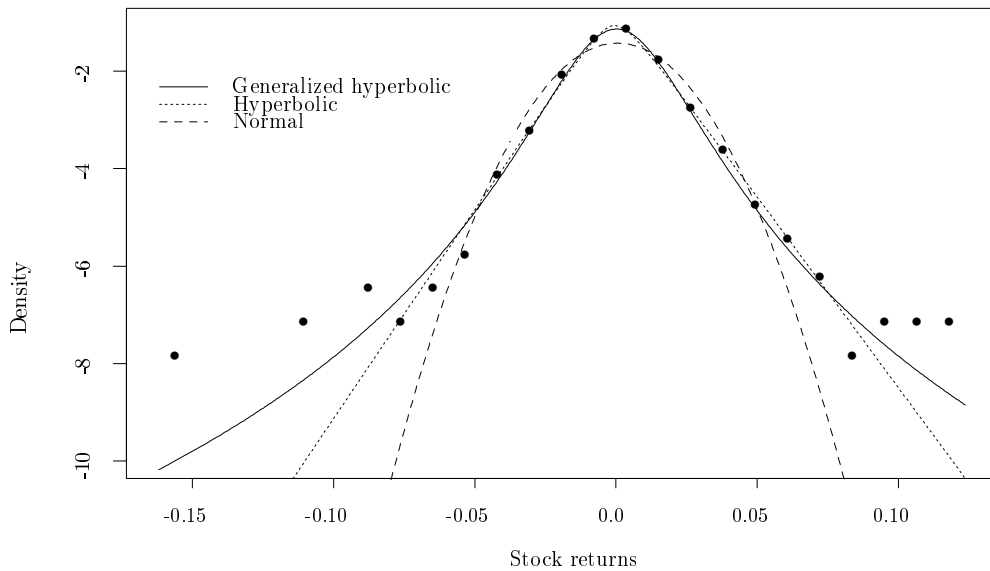


Figure 2.5: The logarithm of the histogram in Figure 2.4 of 2666 daily IBM-stock returns. Superimposed are the logarithms of the best fitting generalized hyperbolic, hyperbolic, and normal densities. The parameter values are as in Figure 2.4.

Log-histograms and log-densities are very useful when the interest is focussed on tail behaviour. From Figures 2.4 and 2.5 it is evident that a heavy-tailed distributions such as the generalized hyperbolic or hyperbolic distribution provides a good fit to the data, and certainly a much better fit than the normal distribution, in particular in the tails. A plot like Figure 2.5, which emphasizes differences in tail behaviour, reveals that the extreme tails of the histogram are a bit heavier than those of the fitted generalized hyperbolic distribution. There is no reason to be overly concerned about this minor discrepancy, because, first, it should be remembered that it is measured on a logarithmic scale, and secondly, the two log-histogram points in the extreme left tail are based on only 1 and 2 observations, respectively, while each of the two points in the extreme right tail corresponds to 2 observations.

The *normal-inverse Gaussian (NIG) distributions* is the subclass obtained for λ equal to $-\frac{1}{2}$. The density of the normal-inverse Gaussian distribution is given by

$$\frac{\alpha\delta}{\pi} e^{\delta\gamma} \cdot \frac{K_1\left(\alpha\sqrt{\delta^2 + (x - \mu)^2}\right)}{\sqrt{\delta^2 + (x - \mu)^2}} \cdot e^{\beta(x - \mu)}, \quad x \in \mathbb{R}. \quad (2.10)$$

If X follows the normal-inverse Gaussian distribution we write

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

If we let α tend to zero, it follows from (A.5) that the *NIG*-distribution converges to the Cauchy distribution with location parameter μ and scale parameter δ .

The Laplace transform of a *NIG*-distribution is especially simple:

$$L(z) = e^{\mu z + \delta(\gamma - \gamma z)}, \quad |\beta + z| < \alpha, \quad (2.11)$$

where $\gamma_z^2 = \alpha^2 - (\beta + z)^2$. Expressions for the mean and variance are also simple in the case of the *NIG*-distribution:

$$\begin{aligned} EX &= \mu + \frac{\delta\beta}{\gamma}, \\ \text{Var} X &= \frac{\delta\alpha^2}{\gamma^3}. \end{aligned}$$

The skewness is $3\delta\alpha^2\beta\gamma^{-5}$ and the kurtosis is $3\delta\alpha^2(\alpha^2 + 4\beta^2)\gamma^{-7}$, but also for the *NIG*-distributions it is informative to use the shape triangle, which can be defined in complete analogy with that for the hyperbolic distributions, see e.g. Rydberg (1997). In Figure 2.6 *NIG* log-density functions are drawn for different values of χ and ξ in the shape triangle defined in the same way as for the hyperbolic distribution.

Finally, but not least, the class of normal-inverse Gaussian distributions is closed under convolution, that is if X_1 and X_2 are independent so that $X_i \sim NIG(\alpha, \beta, \delta_i, \mu_i)$, $i = 1, 2$, then we have that

$$X_1 + X_2 \sim NIG(\alpha, \beta, \delta_1 + \delta_2, \mu_1 + \mu_2). \quad (2.12)$$

Only two subclasses of the generalized hyperbolic distributions are closed under convolution. The other class with this important property is the class of *variance-gamma (VG)* distributions, which is obtained when δ is equal to 0. This is only possible when $\lambda > 0$ and $\alpha > |\beta|$. The variance-gamma distributions (with $\beta = 0$) were introduced in the financial literature by Madan & Seneta (1990).

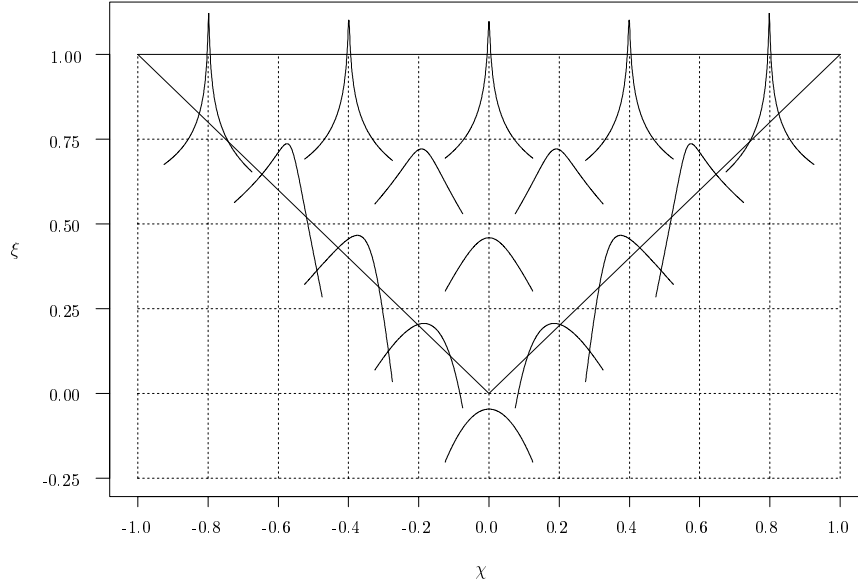


Figure 2.6: Normal-inverse Gaussian log densities with mean 0 and variance 1 for different values of the parameters χ and ξ ($-0.8, -0.6, \dots, 0.8$ for χ and $0.0, 0.25, \dots, 1.0$ for ξ). The log-densities are placed at the corresponding values of χ and ξ .

Another and perhaps more natural name for the full class is the *normal-gamma* (*NG*) distributions. The density function is given by

$$\frac{\gamma^{2\lambda}}{\sqrt{\pi}\Gamma(\lambda)(2\alpha)^{\lambda-\frac{1}{2}}}|x-\mu|^{\lambda-\frac{1}{2}}K_{\lambda-\frac{1}{2}}(\alpha|x-\mu|)e^{\beta(x-\mu)}, \quad x \in \mathbb{R}, \quad (2.13)$$

where Γ denotes the gamma-function. If X follows a variance-gamma distribution we write

$$X \sim VG(\lambda, \alpha, \beta, \mu).$$

The reader is reminded that the parameter domain is $\lambda > 0$, $\alpha > |\beta| \geq 0$ and $\mu \in \mathbb{R}$. The Laplace transform of a *VG*-distribution is simple:

$$L(z) = e^{\mu z} \left(\frac{\gamma}{\gamma_z} \right)^{2\lambda}, \quad |\beta + z| < \alpha, \quad (2.14)$$

where again $\gamma_z^2 = \alpha^2 - (\beta + z)^2$. From (2.14) (or from (2.5) and (2.6)) it easily follows that

$$\begin{aligned} EX &= \mu + \frac{2\beta\lambda}{\gamma^2}, \\ \text{Var}X &= \frac{2\lambda}{\gamma^2} \left(1 + 2 \left(\frac{\beta}{\gamma} \right)^2 \right). \end{aligned}$$

As mentioned above, the class of variance-gamma distributions is closed under convolution. If X_1 and X_2 are independent random variables such that $X_i \sim VG(\lambda_i, \alpha, \beta, \mu_i)$, $i = 1, 2$, then we have that

$$X_1 + X_2 \sim VG(\lambda_1 + \lambda_2, \alpha, \beta, \mu_1 + \mu_2). \quad (2.15)$$

This convolution property follows from (2.14).

By (A.6), the tails of a VG-distribution decrease as $|x|^{\lambda-1}e^{-\alpha|x|+\beta x}$ when $x \rightarrow \pm\infty$. The logarithm of the densities of variance-gamma distributions are plotted for different values of λ in Figure 2.7. In all cases $\beta = 0$, the mean is zero, and the variance is one. From this figure appears a disadvantage of the class of VG-distributions. The probability density is very peaked at the centre for $\lambda < 1$, while for $\lambda \geq 1$ the tail-behaviour does not fit the tails found in typical financial data like those in Figure 2.5 as well as other generalized hyperbolic distributions like for instance the *NIG*-distribution.

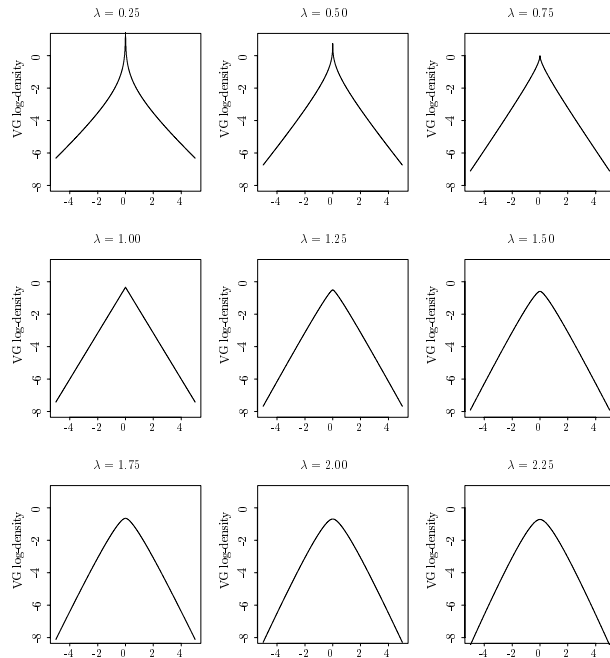


Figure 2.7: The logarithm of the densities of variance-gamma distributions with $\beta = 0$, mean 0, and variance 1 for different values of the parameter λ .

We will finally consider the subclass of the generalized hyperbolic distributions that is obtained when $\alpha = |\beta|$, or equivalently $\gamma = 0$. This is only possible when $\lambda < 0$ and $\delta > 0$. It is convenient to introduce the reparametrization $\nu = -2\lambda$. For $\gamma = 0$ we obtain the density function

$$\frac{\delta^\nu}{\sqrt{\pi}2^{(\nu-1)/2}\Gamma(\nu/2)} \cdot \frac{K_{(\nu+1)/2}(|\beta|\sqrt{\delta^2 + (x - \mu)^2})}{(\sqrt{\delta^2 + (x - \mu)^2}/|\beta|)^{(\nu+1)/2}} \cdot e^{\beta(x-\mu)}, \quad x \in \mathbb{R}, \quad (2.16)$$

where $\nu > 0$, $\delta > 0$, $\beta \in \mathbb{R}$ and $\mu \in \mathbb{R}$. A natural name for this distribution is the *asymmetric scaled t-distribution*, as will soon be clear. From (A.6) it follows that when β is positive, the left hand tail decreases as $|x|^{-(\nu/2+1)}e^{2\beta x}$, while the right hand tail decreases as $x^{-(\nu/2+1)}$. When β is negative, the behaviour of the two tails is interchanged. The expectation exists provided $\nu > 2$, and the variance exists when $\nu > 4$. More generally, the n th moment exists when $\nu > 2n$. The Laplace transform of the distribution given by (2.16) is

$$e^{\mu z} \frac{(-\delta z(z + 2\beta))^{\nu/2} K_{\nu/2}(-\delta z(z + 2\beta))}{\Gamma(\nu/2)2^{\nu/2-1}} \quad (2.17)$$

with domain $-2\beta < z \leq 0$ when $\beta > 0$ and $0 \leq z < -2\beta$ when $\beta < 0$. When $\beta = 0$, the domain is the

set $\{0\}$, and we obtain the density function

$$\frac{\Gamma((\nu + 1)/2)}{\delta\sqrt{\pi}\Gamma(\nu/2) \left(1 + \left(\frac{x-\mu}{\delta}\right)^2\right)^{(\nu+1)/2}}, \quad x \in \mathbb{R},$$

which is the well-known density of the scaled t-distribution with ν degrees of freedom.

2.2 The generalized inverse Gaussian distribution

The second class of distributions, that we consider in this section, is the class of *generalized inverse Gaussian (GIG)* distributions. The *GIG*-distributions are described by three parameters and defined on the positive half axis. The generalized inverse Gaussian density is of the form

$$\frac{(\gamma/\delta)^\lambda}{2K_\lambda(\delta\gamma)} \cdot x^{\lambda-1} \cdot \exp\left\{-\frac{1}{2}(\delta^2 x^{-1} + \gamma^2 x)\right\}, \quad x > 0. \quad (2.18)$$

The parameter domain is given by

$$\begin{aligned} \delta > 0, \quad \gamma \geq 0, & \quad \text{if } \lambda < 0, \\ \delta > 0, \quad \gamma > 0, & \quad \text{if } \lambda = 0, \\ \delta \geq 0, \quad \gamma > 0, & \quad \text{if } \lambda > 0. \end{aligned}$$

The class of generalized inverse Gaussian distributions was first proposed in 1946 by Étienne Halphen, who used it to model the distribution of the monthly flow of water in hydroelectric stations, see Seshadri (1997). The class was rediscovered by Sichel (1973) who used it to construct mixtures of Poisson distributions and by Barndorff-Nielsen (1977) who used it to construct the class of generalized hyperbolic distributions, but also realized its broad usefulness and initiated an in depth study of the class. We shall return to the relation to the generalized hyperbolic distributions later. The generalized inverse Gaussian distributions were briefly mentioned by Good (1953) as an intermediate between Pearson's curves of Type III and V. The class of generalized inverse Gaussian distributions was investigated extensively in Jørgensen (1982).

Using (A.5) we see that for $\lambda > 0$ and $\gamma > 0$ the gamma distribution emerges as limit distribution when δ tends to zero, that is we get the following density for positive λ and γ ,

$$\frac{(\frac{1}{2}\gamma^2)^\lambda}{\Gamma(\lambda)} \cdot x^{\lambda-1} \cdot e^{-\frac{1}{2}\gamma^2 x}, \quad x > 0.$$

Similarly, the inverse gamma distribution with density given by

$$\frac{(2/\delta^2)^\lambda}{\Gamma(-\lambda)} \cdot x^{\lambda-1} \cdot e^{-\frac{1}{2}\delta^2/x}, \quad x > 0.$$

is obtained when γ tends to zero for $\lambda < 0$ and $\delta > 0$. This distribution has a tail of the Pareto type. Finally, for $\lambda = -\frac{1}{2}$ we get the inverse Gaussian distribution with density function given by

$$\frac{\delta}{\sqrt{2\pi x^3}} \cdot e^{-\frac{\gamma(x-\delta/\gamma)^2}{2x}}, \quad x > 0.$$

The generalized inverse Gaussian distributions are unimodal with mode point given by

$$\begin{cases} \frac{\lambda-1+\sqrt{(\lambda-1)^2+\delta^2\gamma^2}}{\gamma^2} & \text{if } \gamma > 0, \\ \frac{\delta^2}{2(1-\lambda)} & \text{if } \gamma = 0. \end{cases}$$

If X has a generalized inverse Gaussian distribution, we write $X \sim GIG(\lambda, \delta, \gamma)$. In Figure 2.8 generalized inverse Gaussian densities are plotted for different values of λ and $\omega = \delta\gamma$. In all cases the variance is 1.

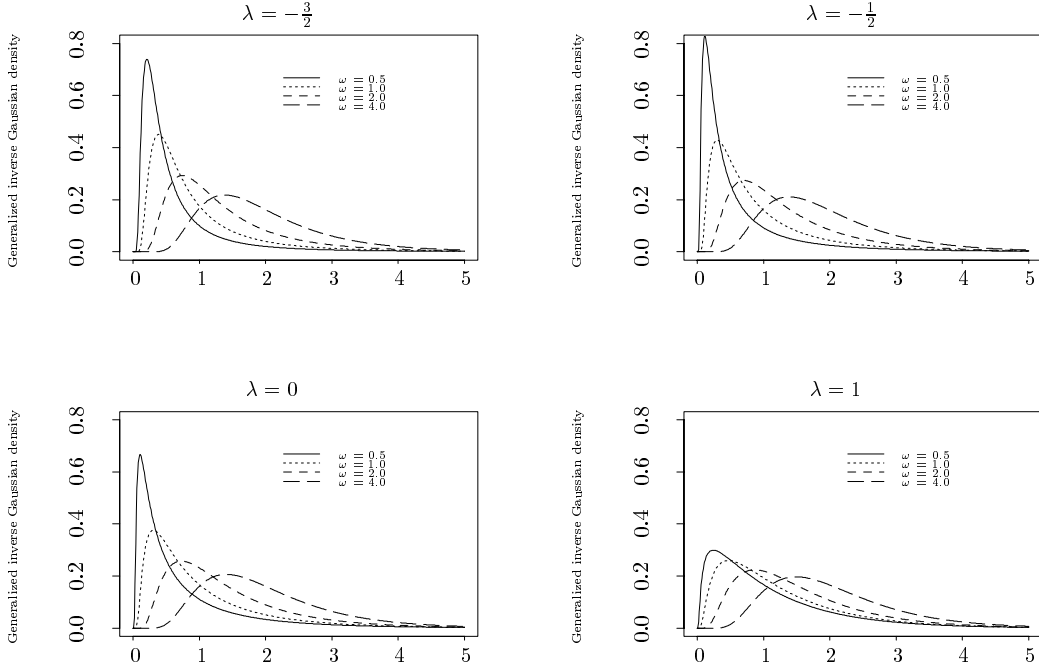


Figure 2.8: Generalized inverse Gaussian densities with variance 1 for different values of the parameters λ and $\omega = \delta\gamma$.

The Laplace transform of the $GIG(\lambda, \delta, \gamma)$ -distribution is

$$L(z) = \frac{K_\lambda\left(\omega\sqrt{1-\frac{2z}{\gamma^2}}\right)}{K_\lambda(\omega)\left(1-\frac{2z}{\gamma^2}\right)^{\lambda/2}} \quad (2.19)$$

for $\delta > 0$ and $\gamma > 0$. The domain of L is $z < \gamma^2/2$ when $\lambda \geq 0$ and $z \leq \gamma^2/2$ when $\lambda < 0$. In the cases $\delta = 0$ or $\gamma = 0$, the Laplace transform is obtained from (2.19) by (A.5). For $\delta = 0$,

$$L(z) = \left(1 - \frac{2z}{\gamma^2}\right)^{-\lambda}, \quad z < \gamma^2/2,$$

which is the well-known Laplace transform of the gamma-distribution. For $\gamma = 0$ we obtain

$$L(z) = \frac{2K_\lambda\left(\sqrt{-2\delta^2 z}\right)}{\Gamma(-\lambda)\left(-\frac{1}{2}\delta^2 z\right)^{\lambda/2}}, \quad z \leq 0.$$

For positive values of δ and γ the moments of X are given by

$$\text{EX}^j = (\delta/\gamma)^j \frac{K_{\lambda+j}(\omega)}{K_\lambda(\omega)}, \quad j = 1, 2, \dots, \quad (2.20)$$

When either δ or γ is zero, the moments of X are also known and are obtained as limits of (2.20). The variance of X is given by

$$\text{Var}X = (\delta/\gamma)^2 \left(\frac{K_{\lambda+2}(\omega)}{K_{\lambda}(\omega)} - \frac{K_{\lambda+1}^2(\omega)}{K_{\lambda}^2(\omega)} \right). \quad (2.21)$$

In Figure 2.9 a histogram of 307 monthly observations of interest rates in the period from June 1964 to December 1989 is given along with a fitted generalized inverse Gaussian density corresponding to the parameter values $\delta = 0.2693$, $\gamma = 11.23$, and $\lambda = -7.0707$. More precisely, the data are annualized monthly yields of U.S. one-month Treasury bills. The same data set was studied in Chan et al. (1992).

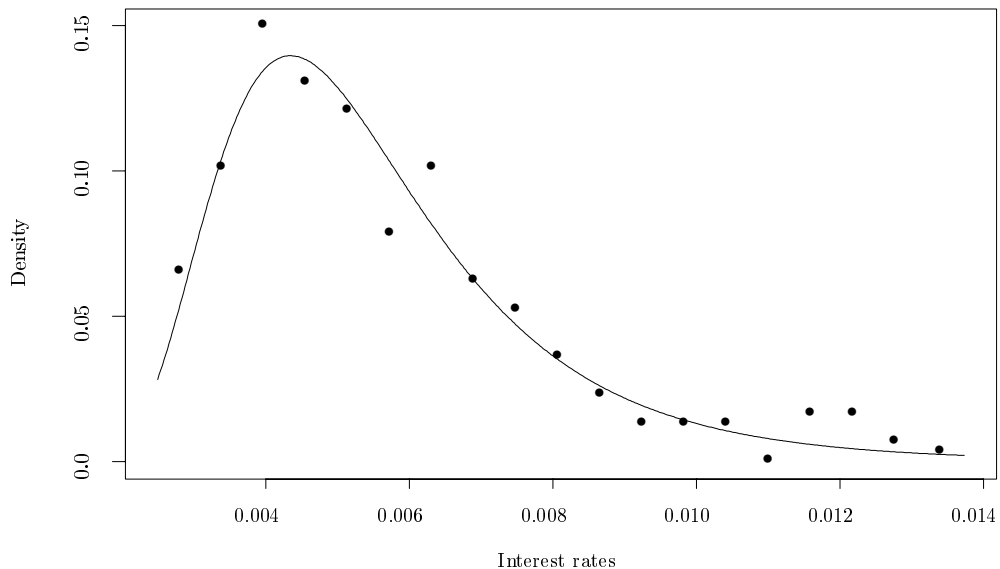


Figure 2.9: A histogram of 307 monthly interest rates. The generalized inverse Gaussian density with parameters $\delta = 0.2693$, $\gamma = 11.23$, and $\lambda = -7.0707$ is superimposed.

There is the following important relationship between the generalized hyperbolic distribution and the generalized inverse Gaussian distribution, which was, in fact, how the generalized hyperbolic distribution was originally derived in Barndorff-Nielsen (1977). The generalized hyperbolic distribution is a normal variance-mean mixture where the mixing distribution is generalized inverse Gaussian. What is meant by this is that if

$$X|W = w \sim N(\mu + \beta w, w),$$

and $W \sim GIG(\lambda, \delta, \gamma)$, then the marginal distribution of X will be generalized hyperbolic, $X \sim H(\lambda, \alpha, \beta, \delta, \mu)$, where $\alpha^2 = \beta^2 + \gamma^2$. This property provides a possible interpretation of non-Gaussian stochastic variation described by a generalized hyperbolic distribution.

As special cases we have that the normal-inverse Gaussian distribution appears when the mixing distribution is an inverse Gaussian distribution, and the variance-gamma distribution emerges as a normal variance-mean mixture where the mixing distribution is a gamma distribution. This explains the names of the distributions. The asymmetric scaled t -distribution is a normal variance-mean

mixture with an inverse gamma mixing distribution. As a special case we get the well-known result that the t -distribution is a normal variance mixture ($\beta = 0$) with an inverse gamma mixing distribution.

The mixing result implies that there is the following simple relationship between the Laplace transform, L_X , of the generalized hyperbolic distribution and that of the generalized inverse Gaussian distributions, L_W ,

$$L_X(z) = e^{\mu z} \cdot L_W(\beta z + \frac{1}{2}z^2).$$

Barndorff-Nielsen & Halgreen (1977) showed that generalized inverse Gaussian distributions are infinitely divisible. Using that the generalized hyperbolic distributions are normal variance-mean mixtures with generalized inverse Gaussian mixing distributions, they also proved that generalized hyperbolic distributions are infinitely divisible. Halgreen (1979) showed that generalized hyperbolic distributions and generalized inverse Gaussian distribution are even self-decomposable. In the following section, the properties of infinite divisibility and self-decomposability will turn out to be important because they allow the construction of certain hyperbolic stochastic process models.

2.3 Statistical inference

Inference for the parameters when dealing with independent and identically generalized hyperbolic or generalized inverse Gaussian distributed observations should be based on the likelihood function. The C-program HYP described in Blæsild & Sørensen (1992) can be used for maximum likelihood estimation in the situation where independent and identically (possibly multi-dimensional) hyperbolic distributed observations are considered. The program HYP also has the facility of basing the inference on the multinomial likelihood function obtained by only observing the number of observations in given intervals. More precisely, if I_1, \dots, I_k are disjoint intervals with union the entire real line and y_j denotes the number of observations in I_j , $j = 1, \dots, k$, then the multinomial log-likelihood function is given by

$$\ell(\alpha, \beta, \delta, \mu) = \sum_{j=1}^k y_j \log p_j, \tag{2.22}$$

where p_j is the probability that a hyperbolic distributed random variable takes a value in I_j , that is,

$$p_j = \int_{I_j} \frac{\gamma}{2\alpha\delta K_1(\delta\gamma)} \exp\{-\alpha\sqrt{\delta^2 + (x - \mu)^2} + \beta(x - \mu)\} dx, \quad j = 1, \dots, k. \tag{2.23}$$

Inference based on grouped observations from other distributions can of course be carried out in a similar way using (2.22) and the equivalent of (2.23). Küchler et al. (1999) note that if the observations are not independent then inference based on the multinomial likelihood function for grouped observations will be more robust to effects of the dependence than inference based on the original likelihood function for independent observations.

3 Lévy Processes

A homogeneous Lévy process X is a stochastic process with $X_0 = 0$ and with the property that its increments over non-overlapping time intervals are independent. Moreover, increments over any time interval of length t have the same distributions as X_t . The homogeneous Lévy processes are also called processes with independent, stationary increments or additive processes. The mathematical theory of Lévy processes can be found in Bertoin (1996) or Sato (1999). An example of a Lévy process that is

well-known from, for instance, the Black-Scholes-Merton option pricing theory is the Brownian motion (or Wiener process), where the increments are normally distributed.

For every generalized hyperbolic distribution there exists a homogeneous Lévy process X such that the probability distribution of the value of the process at a fixed time point t , X_t , is that particular generalized hyperbolic distribution. A thorough review of the theory of these generalized hyperbolic Lévy processes and their application in finance can be found in Eberlein (2001), see also Prause (1999) and Eberlein & Raible (2000). The distributions that can appear as the distribution of the instantaneous value of a homogeneous Lévy process are exactly those that have the property called infinite divisibility. As mentioned in Section 2 the generalized hyperbolic distributions are infinitely divisible. Usually, the distribution of the value X_s at a time point s different from t will not be generalized hyperbolic. However, in the case of the *NIG* and *VG* distributions, the convolution properties (2.12) and (2.15) imply that the value of the Lévy process will be *NIG*-distributed, respectively *VG*-distributed, at all time points. This makes the *NIG* and *VG* Lévy processes more natural generalized hyperbolic Lévy processes than the other generalized hyperbolic Lévy processes. Simulation of the *NIG* Lévy process was studied in Rydberg (1997).

A generalized hyperbolic Lévy processes can be written in the form

$$X_t = \lambda t + Z_t,$$

where Z_t is a pure jump martingale with infinitely many small jumps in every finite time interval, however small. The behaviour of Z_t is reflected in the so-called Lévy measure, see (3.4) and the discussion following this formula. The Lévy measure of the generalized hyperbolic distribution is

$$q(x) = \begin{cases} \frac{e^{\beta x}}{|x|} \left(\int_0^\infty \frac{\exp(-|x|\sqrt{2y+\alpha^2})}{\pi^2 y (J_\lambda^2(\delta\sqrt{2y}) + Y_\lambda^2(\delta\sqrt{2y}))} dy + \lambda e^{-\alpha|x|} \right) & \text{if } \lambda \geq 0 \\ \frac{e^{\beta x}}{|x|} \int_0^\infty \frac{\exp(-|x|\sqrt{2y+\alpha^2})}{\pi^2 y (J_{-\lambda}^2(\delta\sqrt{2y}) + Y_{-\lambda}^2(\delta\sqrt{2y}))} dy & \text{if } \lambda < 0. \end{cases} \quad (3.1)$$

Here J_λ and Y_λ denote Bessel functions of the first and second kind, respectively, see the Appendix. The Lévy measure was essentially found by Halgreen (1979), see also Prause (1999). For the *NIG*-distribution this expression simplifies to

$$q(x) = \pi^{-1} \delta \alpha |x|^{-1} K_1(\alpha|x|) e^{\beta x}, \quad (3.2)$$

where K_1 is a modified Bessel function of the third kind. The behaviour near zero is particularly important, so the following expansion for generalized hyperbolic distributions (Raible (2000)) is useful:

$$x^2 q(x) = \frac{\delta}{\pi} + \frac{\lambda + \frac{1}{2}}{2} |x| + \frac{\delta\beta}{\pi} x + o(|x|) \quad (3.3)$$

as $x \rightarrow 0$. We see that for every generalized hyperbolic distribution the Lévy measure has infinite mass in every neighbourhood of the origin. The process Z_t is given by

$$Z_t = \int_0^t \int_{\mathbb{R} \setminus \{0\}} x \left(\mu^X(du, dx) - q(x) du dx \right), \quad (3.4)$$

where the integer-valued random measure μ^X is defined by

$$\mu^X(dt, dx) = \sum_{s>0} 1_{\{\Delta X_s \neq 0\}} \varepsilon_{(s, \Delta X_s)}(dt, dx).$$

Here ε_a denotes the Dirac measure at a , and $\Delta X_s = X_s - X_{s-}$ is the jump of the process X at time s (for most time points $\Delta X_s = 0$). Integrals of the type (3.4) are treated in e.g. Jacod & Shiryaev (1987) or Protter (1990). The random measure μ^X is Poissonian with intensity measure $q(x)dxdt$. This implies that for any closed interval A that does not contain the origin, the number of jumps in the time interval $[0, t]$ with a size that belongs to A , i.e.

$$N_t^A = \mu^X([0, t], A),$$

is a Poisson process with intensity $\int_A q(x)dx$, which is a finite number. In particular, N_t^A is Poisson distributed with mean value $t \int_A q(x)dx$. As the boundary of the interval A tends to zero, the mean value goes to infinity, cf. (3.3). It is interesting to note that a generalized hyperbolic Lévy process has no continuous Brownian motion component and has infinitely many jumps on every time interval.

The generalized hyperbolic Lévy processes do, however, have a nice relation to the Brownian motion. Let B be a standard Brownian motion, and let $\tau(t)$ be a Lévy process for which the distribution of $\tau(1)$ is a generalized inverse Gaussian distribution. Then the process

$$X_t = \mu t + \beta \tau(t) + B_{\tau(t)} \quad (3.5)$$

is a generalized hyperbolic Lévy process. Because the increments of τ are generalized inverse Gaussian distributed and hence can only be positive, the process τ is increasing and can thus be interpreted as a time that increases with a randomly varying speed. A process τ with this property is called a subordinator, and the construction (3.5) is called *subordination*. The randomly increasing time τ has been interpreted as an *operational time* or a *business time* reflecting, for instance, the volume of trade at an exchange. Some times a lot is happening at the exchange and the business time increases rapidly. At other times the exchange is tranquil and the business time goes only slowly. That the distribution of X_1 is generalized hyperbolic follows because this distribution is a variance-mean mixture of normal distributions where the mixing distribution is the generalized inverse Gaussian distribution, see Subsection 2.2. The fact that a Lévy process τ exists such that $\tau(1)$ is generalized inverse Gaussian distributed follows because these distributions are infinitely divisible, as mentioned in Section 2. In the case of the *NIG*-distribution, the construction by subordination can be done in the following simple way (Barndorff-Nielsen (1998)). Let (U_t, V_t) be a two-dimensional standard Brownian motion starting at $(0, 0)$ and with drift vector (β, γ) , where $\gamma > 0$. Let $\tau(t)$ denote the first time the second component V attains the value $\delta t > 0$ with $\delta > 0$. Then $\{\tau(t) : t > 0\}$ is an inverse Gaussian Lévy process, and

$$X_t = \mu t + U_{\tau(t)}$$

is a *NIG*-Lévy process. Specifically, X_t is *NIG* $(\alpha, \beta, \delta t, \mu t)$ distributed, where $\alpha = \sqrt{\beta^2 + \gamma^2}$.

Construction of financial models by subordination was first proposed by Praetz (1972) who used a scaled t-distribution to model stock returns and obtained a good fit to weekly returns from the Sydney Stock Exchange. This is a particular example of a generalized hyperbolic distribution where the mixing distribution is an inverse gamma distribution, see Subsection 2.1. Praetz attributed the mixing of normals to the change in activity at the exchange. Clark (1973) and Epps & Epps (1976) found that there is a dependency between trading volume and the variance of returns, but did not suggest generalized hyperbolic models. These findings have been confirmed by Ané & Geman (2000). In Madan & Seneta (1990), Madan & Milne (1991) and Madan & Chang (1996) the so-called variance gamma model is introduced and studied as a model for share markets returns. This model is the generalized hyperbolic Lévy process with a gamma mixing distribution. For a discussion of the subordination approach in finance, see e.g. Hurst, Platen & Rachev (1997).

The use of generalized hyperbolic Lévy processes to model the prices of stocks and other assets and the corresponding theory of option pricing has been thoroughly investigated by Eberlein & Keller (1995),

Keller (1997), Eberlein, Keller & Prause (1998) and Eberlein & Prause (1998). Eberlein & Jacod (1997) proved that the set of equivalent martingale measures is large and that the corresponding price range is the entire non-arbitrage interval. A theory of the term structure of interest rates based on the hyperbolic Lévy process was developed in Eberlein & Raible (1999). A useful review can be found in Eberlein (2001).

For the processes discussed in this section, estimation based on observations at equidistant discrete time points is as easy as estimation for independent generalized hyperbolic distributions, because the increments of the process between the observation times are independent. Usually one would use a Lévy process for which the increments are generalized hyperbolic and then estimate the parameters, for instance by means of the computer program mentioned in Subsection 2.3. A simple check of the fit of the model to the data can be made as follows. If, for instance, the data are daily observations, then it should be checked that the distributions calculated from the estimated model of the increments over a number of suitably chosen longer time spans fit the corresponding increments calculated from the data. For the *NIG* and *VG* Lévy processes these distributions are simply given by the formulae (2.12) and (2.15). For an example of this procedure, see Eberlein (2001).

4 Stochastic differential equations

In this section we present various methods for constructing diffusion processes with generalized hyperbolic and generalized inverse Gaussian marginal distributions. A diffusion process is the solution of a stochastic differential equation driven by a Wiener process. Estimation of parameters based on discrete-time observations of a diffusion process is considered too. Furthermore, we consider Ornstein-Uhlenbeck type processes driven by Lévy processes and models given as sums of processes defined by stochastic differential equations.

4.1 Diffusion models

We consider a one-dimensional diffusion process $\{X_t\}$ and suppose that it is the unique weak solution to the stochastic differential equation

$$dX_t = b(X_t; \theta)dt + \sigma(X_t; \theta)dW_t, \quad (4.1)$$

where $\sigma(x; \theta)$ is positive for all x in the state space (l, r) ($-\infty \leq l < r \leq \infty$) and all θ in some p -dimensional parameter space Θ . We will focus on ergodic diffusions and denote the density of the corresponding invariant probability measure by μ_θ .

Diffusion processes with a specific marginal distribution are typically constructed by determining drift b and diffusion coefficient σ so that the invariant distribution is of the required type. This method will result in the appropriate marginal distribution for large values of t or for all t provided that the initial distribution is equal to the invariant distribution (i.e. $X_0 \sim \mu_\theta$). Under mild conditions we have the following relationship between the drift, diffusion coefficient and the density of the invariant distribution,

$$2b(x; \theta) - v'(x; \theta) = v(x; \theta) \frac{\mu_\theta'(x)}{\mu_\theta(x)}, \quad l < x < r, \quad \theta \in \Theta, \quad (4.2)$$

where v denotes the squared diffusion coefficient, $v(x; \theta) = \sigma^2(x; \theta)$.

Using (4.2), Bibby & Sørensen (2001b) discussed a method for constructing diffusion processes with

a prescribed marginal (invariant) distribution. Letting the drift be given by

$$b(x; \theta) = \frac{1}{2}v(x; \theta) \frac{d}{dx} \log[v(x; \theta)f(x)],$$

where f is a function that is integrable on the interval (l, r) , it was shown under some regularity conditions that the diffusion process given by (4.1) has invariant density μ_θ proportional to f , irrespective of the choice of the function v . Bibby & Sørensen (2001b) also considered the special case where

$$v(x; \theta) = \sigma^2 f(x)^{-\kappa}, \quad \sigma^2 > 0, \quad \kappa \in [0, 1],$$

in particular the situation where the invariant density was hyperbolic. This led to the following stochastic differential equation,

$$dX_t = \frac{1}{2}\sigma^2(1 - \kappa)f(X_t)^{-\kappa} \left[\beta - \frac{\alpha(X_t - \mu)}{\sqrt{\delta^2 + (X_t - \mu)^2}} \right] dt + \sigma f(X_t)^{-\frac{1}{2}\kappa} dW_t, \quad (4.3)$$

where f is proportional to the hyperbolic density function given by (2.8), that is

$$f(x) = \exp \left[-\alpha \sqrt{\delta^2 + (x - \mu)^2} + \beta(x - \mu) \right].$$

Note that the drift is towards the mode point of the hyperbolic distribution, $\mu + \beta\delta/\gamma$. The diffusion process given by (4.3) was successfully used to describe the logarithm of the price of VW-stocks after a linear trend had been subtracted.

In Bibby & Sørensen (1997) the special case where $\kappa = 1$ was considered in the situation of a hyperbolic invariant density. Note that this results in a diffusion process with no drift, that is the solution to the stochastic differential equation given by

$$dX_t = \sigma \exp \left\{ \frac{1}{2}\alpha \sqrt{\delta^2 + (X_t - \mu)^2} - \frac{1}{2}\beta(X_t - \mu) \right\} dW_t. \quad (4.4)$$

It turns out that this is an example of a local martingale which is not a martingale. Also the hyperbolic diffusion process given as the solution of (4.4) was fitted successfully to the logarithm of stock-prices (minus a linear trend) in Bibby & Sørensen (1997). The construction leading to the hyperbolic diffusion (4.3) can obviously be made similarly for any generalized hyperbolic distribution. In the special case $\kappa = 1$, this was done in Rydberg (1999), where the corresponding *NIG*-diffusion was fitted successfully to stock prices (minus a linear trend).

In Küchler et al. (1999) a hyperbolic diffusion process with constant diffusion coefficient was discussed. This corresponds to letting the function v be equal to a constant σ^2 , or to $\kappa = 0$ in (4.3), and gives the following stochastic differential equation,

$$dX_t = \frac{1}{2}\sigma^2 \left[\beta - \alpha \frac{X_t - \mu}{\sqrt{\delta^2 + (X_t - \mu)^2}} \right] dt + \sigma dW_t. \quad (4.5)$$

The hyperbolic diffusion process given by (4.5) was first proposed in Barndorff-Nielsen (1978).

For values of κ between the two extremes 0, corresponding to stationarity being obtained by pure reversion, and 1, where stationarity is obtained by pure diffusion, both these effects are present to varying degrees.

Sørensen (1997b) considers the construction of diffusion processes with generalized inverse Gaussian invariant distribution. If v is a positive function then the solution to the following stochastic differential equation,

$$dX_t = \left(v(X_t)v'(X_t) + \frac{1}{2}v(X_t)^2 \left[(\lambda - 1)X_t^{-1} - \frac{1}{2}\gamma^2 + \frac{1}{2}\delta^2 X_t^{-2} \right] \right) dt + v(X_t)dW_t. \quad (4.6)$$

will have a generalized inverse Gaussian invariant density given by (2.18) under suitable regularity conditions on v . The focus in Sørensen (1997b) is on the special case where $v(x) = \kappa x^\alpha$ for constants $\alpha \geq 0$ and $\kappa > 0$. With this choice of diffusion coefficient, the diffusion process is given by the solution to the stochastic differential equation given by

$$dX_t = \left(\beta_1 X_t^{2\alpha-1} - \beta_2 X_t^{2\alpha} + \beta_3 X_t^{2(\alpha-1)} \right) dt + \kappa X_t^\alpha dW_t, \quad (4.7)$$

where

$$\beta_1 = \frac{1}{2}\kappa^2(\lambda - 1) + \kappa^2\alpha, \quad \beta_2 = \frac{1}{4}(\kappa\gamma)^2, \quad \beta_3 = \frac{1}{4}(\kappa\delta)^2.$$

Note that if $\alpha = \frac{1}{2}$ and $\beta_3 = 0$, then the diffusion process is the solution to

$$dX_t = (\beta_1 - \beta_2 X_t) dt + \kappa \sqrt{X_t} dW_t, \quad (4.8)$$

that is the Cox-Ingersoll-Ross process (CIR-process) used in finance to describe short term interest rates, see Cox, Ingersoll, Jr. & Ross (1985).

A completely different way of constructing hyperbolic diffusion models was proposed in Jensen & Pedersen (1999). These authors consider processes given by $X_t = h(Y_t)$, where Y is a stationary Ornstein-Uhlenbeck process:

$$dY_t = -\alpha Y_t dt + \tau dW_t$$

with $\alpha > 0$ and $\tau > 0$. Suppose F is a cumulative distribution function of a given probability distribution, and let Φ denote the distribution function of the standard normal distribution. If $\tau^2 = 2\alpha$ and $h(y) = F^{-1}(\Phi(y))$, then the distribution of X_t will have the distribution function F . If, in particular, F is the distribution function of a generalized hyperbolic distribution, we obtain a hyperbolic diffusion process. Unfortunately, there is no explicit expression for the distribution function of a generalized hyperbolic distribution. An advantage of this approach is that there is an expression for the transition density involving the function h . Since the distribution function of a generalized hyperbolic distribution, and hence h , can be calculated numerically, it is relatively easy to calculate the likelihood function, which is usually not the case for diffusion models. A disadvantage is that the drift and diffusion coefficients of the diffusion process X are not explicit functions.

4.2 Statistical inference for diffusion processes

Inference for discretely observed diffusion processes is made difficult by the fact that the likelihood function is generally not tractable. In recent years many different methods have been proposed to overcome this obstacle. We will here briefly discuss the methods most commonly used in connection with financial data. For an excellent overview of a wide variety of procedures for estimating parameters based on discretely observed diffusions, see H. Sørensen (2000).

Approximate likelihood methods are considered by Pedersen (1995), Aït-Sahalia (1998), and Poulsen (1999). In Pedersen (1995) it is shown that the likelihood function can be calculated to any given precision using simulations and the Euler approximation in a clever way. Unfortunately, the method is very computer intensive. Honoré (1997) successfully applied the Pedersen method to the model for interest rates proposed by Chan et al. (1992). In Aït-Sahalia (1998) an analytical approximation to the likelihood function based on a truncated Hermite expansion is developed. Poulsen (1999) obtained an approximation to the likelihood function by numerically solving the Chapman-Kolmogorov forward equations. He used his method to fit the CKLS-model to interest rate data. Asymptotic results for the maximum likelihood estimator based on discrete time observations of a diffusion model were derived in Dacunha-Castelle & Florens-Zmirou (1986).

Inference for diffusion processes based on martingale estimating functions is considered in Bibby & Sørensen (1995), Bibby & Sørensen (1996), and Bibby & Sørensen (1997). For observations $X_{t_1}, X_{t_2}, \dots, X_{t_n}$ the martingale estimating functions introduced in Bibby & Sørensen (1995) and Bibby & Sørensen (1996) are of the form

$$G_n(\theta) = \sum_{i=1}^n g_i(X_{t_{i-1}}; \theta) [X_{t_i} - E_\theta(X_{t_i} | X_{t_{i-1}})] + \sum_{i=1}^n h_i(X_{t_{i-1}}; \theta) \left[(X_{t_i} - E_\theta(X_{t_i} | X_{t_{i-1}}))^2 - \text{Var}_\theta(X_{t_i} | X_{t_{i-1}}) \right]. \quad (4.9)$$

Note that in analogy with the unknown score function, the observations enter G_n in pairs, and G_n is a martingale with respect to the natural filtration. The conditional expectations in (4.9) can easily be calculated using simulations, and an estimator for the parameter θ is then obtained by solving the equation $G_n(\theta) = 0$. In Bibby & Sørensen (1995) the resulting estimator is shown to be consistent and asymptotically normal as the number of observations tends to infinity. An optimal choice of the functions g_i and h_i as well as simpler approximately optimal functions that are useful in practice are given in Bibby & Sørensen (1995) and Bibby & Sørensen (1996). As mentioned earlier the hyperbolic diffusion process given by (4.4) was fitted to the log-prices of stocks after a linear trend had been subtracted in Bibby & Sørensen (1997). The parameters in this hyperbolic diffusion model were estimated using the martingale estimating function

$$K_n(\theta) = \sum_{i=1}^n \frac{\dot{v}(X_{t_{i-1}}; \theta)}{(t_i - t_{i-1})v(X_{t_{i-1}}; \theta)^3} \left[(X_{t_i} - X_{t_{i-1}})^2 - E_\theta((X_{t_i} - X_{t_{i-1}})^2 | X_{t_{i-1}}) \right],$$

where v is the squared diffusion coefficient and a dot denotes differentiation with respect to the parameter θ . This is an approximately optimal modification of (4.9) taking into account that the diffusion has no drift.

Kessler & Sørensen (1999) considered martingale estimating functions based on eigenfunctions of the infinitesimal generator of the diffusion process. The advantage of such martingale estimating functions is that they are adapted to concrete models and are easy to calculate in cases where the eigenfunctions are explicitly known. Unfortunately this is not often the case.

It is usually easy to obtain an estimator from a simple estimating function of the form

$$F_n(\theta) = \sum_{i=1}^n f(X_{t_i}; \theta),$$

where the function f satisfies that $\int_l^r f(x, \theta) \mu_\theta(x) dx = 0$ with μ_θ denoting the density of the invariant probability measure. Such simple estimating functions were studied by Hansen & Scheinkman (1995), Kessler (2000), and Jacobsen (2001). The advantage of these estimating functions is that they are indeed simple and fast to work with because it is straightforward to explicitly find functions f with the property needed. The main disadvantages are that only parameters appearing in the invariant density can be estimated using simple estimating functions and that the estimators may be far from efficient because the dependence structure in the data is ignored. An improved version of the simple estimating function where each term in the sum depends on a pair of consecutive observations was considered by Hansen & Scheinkman (1995) and Jacobsen (2001). Optimality questions were treated in Kessler (2000) and Jacobsen (2001). For the improved version it is also not possible to estimate all parameters, see the discussion in Hansen & Scheinkman (1995). A review of estimating function inference for diffusion models was given in Sørensen (1997a).

Indirect inference procedures based on auxiliary models and extensive simulations were proposed by Gourieroux, Monfort & Renault (1993) and Gallant & Tauchen (1996). They have gained some

popularity in the finance literature under the name of the efficient method of moments. However, the quality of the estimators depend on the choice of the auxiliary model, which is not a straightforward matter.

Finally, Bayesian MCMC-methods have been applied to diffusion models by Eraker (1998) and Elerian, Chib & Shepard (2001). In these methods, the likelihood function is calculated in a way similar to that in Pedersen (1995).

4.3 Ornstein-Uhlenbeck processes

A stochastic process X is called a process of the Ornstein-Uhlenbeck type, if it satisfies a stochastic differential equation of the form

$$dX_t = -\lambda X_t dt + dZ_t, \quad (4.10)$$

where $\lambda > 0$ and where the driving process Z is a homogeneous Lévy process. It is not difficult to see that

$$X_t = e^{-\lambda t} X_0 + \int_0^t e^{-\lambda(t-s)} dZ_s. \quad (4.11)$$

If X is stationary and square integrable, the autocorrelation function of X is

$$\rho(u) = \exp(-\lambda u). \quad (4.12)$$

When the process Z is the standard Wiener process, the solution X is the usual Ornstein-Uhlenbeck process. Ornstein-Uhlenbeck type processes have been studied by Wolfe (1982), Sato & Yamazato (1982), Sato & Yamazato (1984) and Sato, Watanabe & Yamazato (1994); see also Jurek & Vervaat (1983), Jurek & Mason (1993), and Barndorff-Nielsen, Jensen & Sørensen (1998). A necessary and sufficient condition for (4.10) to have a stationary solution is that $E(\log(1 + |Z(1)|)) < \infty$.

For every generalized hyperbolic distribution there exists a stationary Ornstein-Uhlenbeck type process such that for all $t \geq 0$ the distribution of X_t is the given generalized hyperbolic distribution. The same is true for all generalized inverse Gaussian distributions. This is because these distributions have the property called self-decomposability, as discussed in Section 2. The Lévy process driving the *NIG* Ornstein-Uhlenbeck type process was studied by Barndorff-Nielsen (1998), while the process driving the symmetric variance-gamma Ornstein-Uhlenbeck type process, was found by Jiang (2000). For symmetric distributions, the driving Lévy process is, in the case of the *NIG* Ornstein-Uhlenbeck process, the sum of a *NIG* Lévy process and a compound Poisson process, while for the variance-gamma Ornstein-Uhlenbeck process, it is simply a compound Poisson process.

As for most ordinary diffusion processes, the likelihood function is usually not explicitly available for processes of the Ornstein-Uhlenbeck type. Since these processes are Markov processes, a simple and natural approach to statistical inference goes via estimating functions based on conditional moments defined in analogy with those discussed in Subsection 4.2.

4.4 Compound processes

Quite often, the exponentially decreasing autocorrelation function (4.12) is too simple to fit financial data. However, models with a much more flexible covariance structure are easily obtained by summing independent Ornstein-Uhlenbeck type processes, as was proposed by Barndorff-Nielsen, Jensen & Sørensen (1998). The process

$$X_t = X_t^{(1)} + \dots + X_t^{(m)}, \quad (4.13)$$

where the processes $X_t^{(i)}$, $i = 1, \dots, m$ are independent Ornstein-Uhlenbeck type processes given by

$$dX_t^{(i)} = -\lambda_i X_t^{(i)} dt + dZ_t^{(i)} \quad (4.14)$$

for independent Lévy processes $Z_t^{(i)}$, $i = 1, \dots, m$, has an autocorrelation function of the form

$$\rho(u) = \Phi_1 \exp(-\lambda_1 u) + \dots + \Phi_m \exp(-\lambda_m u), \quad (4.15)$$

where Φ_i is proportional to the variance of $X_t^{(i)}$, and $\Phi_1 + \dots + \Phi_m = 1$. Even for $m = 2$ a much better fit is often obtained. Examples can be found in Barndorff-Nielsen, Jensen & Sørensen (1998) and Barndorff-Nielsen & Shephard (2001d).

For every generalized hyperbolic distribution and for every generalized inverse Gaussian distribution there exists a stationary process X of the form (4.13) – (4.14) such that for all $t \geq 0$ the distribution of X_t is that particular distribution. Again this is because these distributions are self-decomposable, see Barndorff-Nielsen, Jensen & Sørensen (1998). More complex types of superpositions of Ornstein-Uhlenbeck type processes were investigated in Barndorff-Nielsen (2000).

The construction (4.13) can be made for diffusion models with linear drift and non-linear diffusion coefficient too, see Bibby & Sørensen (2001a). As an example, suppose we want a stationary stochastic process with autocorrelation function (4.15) for given values of $\lambda_1, \dots, \lambda_m$ and Φ_1, \dots, Φ_m , and such that the marginal distribution of X_t is a gamma distribution with shape parameter α and scale parameter β . This can be obtained by defining m independent processes as the stationary solutions to

$$dX_t^{(i)} = -\lambda_i (X_t^{(i)} - \Phi_i \alpha \beta) dt + \sqrt{2\beta \lambda_i X_t^{(i)}} dW_t^{(i)}, \quad (4.16)$$

$i = 1, \dots, m$. Each of the processes, $X_t^{(i)}$, is a CIR-process, (4.8), which is a particular example of the generalized inverse Gaussian diffusions given by (4.7). Since $X_t^{(i)}$ is gamma distributed with shape parameter $\alpha \Phi_i$ and scale parameter β , it follows that X_t defined by (4.13) has the required gamma distribution, and since the autocorrelation function of $X_t^{(i)}$ is $\exp(-\lambda_i u)$, the autocorrelation function of the sum X_t is given by (4.15). This construction will come in handy in Section 5, where processes of the type (4.13) will be used as models for stochastic volatility.

Empirical autocorrelations that might be interpreted as an indication of long range dependence, may often alternatively be approximated very well by autocorrelation functions of the type (4.15). However, if a model with genuine long range dependence is desirable, a *NIG*-process of this type can be constructed as follows.

Let $X^{(i)}$, $i = 1, 2, \dots$ be a sequence of independent *NIG* Ornstein-Uhlenbeck processes with *NIG*-parameters $(\alpha, \beta, 0, \delta_i)$, where

$$\delta_i \sim i^{-1-2(1-H)},$$

for some $H \in (0, 1)$, and with the same value of the drift parameter λ . Then Barndorff-Nielsen (1998) showed that the process

$$X(t) = \sum_{i=1}^{\infty} X^{(i)}(t/i), \quad (4.17)$$

which is stationary and well-defined as a mean-square limit, has as its marginal distribution the *NIG* distribution with parameters $(\alpha, \beta, 0, \delta)$, where $\delta = \sum_{i=1}^{\infty} \delta_i$. Moreover, the autocorrelation function $r(u)$ satisfies

$$r(u) \sim L(u) u^{-2(1-H)},$$

for some slowly varying function L . Thus if $\frac{1}{2} < H < 1$, the process exhibits long range dependence with exponent H . The construction of long range dependent processes by a sum of the type (4.17) is

similar to a construction proposed by Cox (1984). Almost the same construction was used in Barndorff-Nielsen, Jensen & Sørensen (1990). The construction (4.17) can also be applied to a sequence of independent stationary *NIG*-diffusions given as solutions of stochastic differential equations defined in analogy to (4.3).

Likelihood inference for the various compound processes considered here is complicated by the fact that the likelihood function is not explicitly available. A feasible alternative is provided by prediction-based estimating functions, see Sørensen (2000).

5 Stochastic volatility models

A generalization of the Black-Scholes model for the logarithm of an asset price

$$dX_t = (\mu + \beta\sigma^2)dt + \sigma dW_t, \quad (5.1)$$

that takes into account the empirical finding that the volatility σ^2 varies randomly over time is a stochastic volatility process:

$$dX_t = (\mu + \beta v_t)dt + \sqrt{v_t}dW_t. \quad (5.2)$$

Here the volatility v_t is a stochastic process that cannot be observed directly. If the data are observations at the time points Δi , $i = 0, 1, 2, \dots, n$, then the returns $Y_i = X_{i\Delta} - X_{(i-1)\Delta}$ can be written in the form

$$Y_i = \mu\Delta + \beta S_i + \sqrt{S_i}A_i, \quad (5.3)$$

where

$$S_i = \int_{(i-1)\Delta}^{i\Delta} v_t dt, \quad (5.4)$$

and where the A_i s are independent, standard normal distributed random variables. If the integrated volatility S_i is independent of A_i , and if it is generalized inverse Gaussian distributed, then the distribution of the return Y_i is generalized hyperbolic. This follows from the representation of the generalized hyperbolic distributions as variance-mean mixtures of normal distributions mentioned in Subsection 2.2. Unfortunately, no continuous time process v with the property that the integrated volatility (5.4) is exactly generalized inverse Gaussian distributed is presently known. Therefore we will instead consider models where the volatility process v is stationary with v_t generalized inverse Gaussian distributed. For small values of Δ , the distribution of S_i will then be close to a generalized inverse Gaussian distribution, and hence the distribution of Y_i will be close to a generalized hyperbolic distribution. Thus we obtain models that are not exactly hyperbolic, but which have marginal distribution with much the same tail properties when Δ is not too large. When Δ tends to infinity, the distribution of $\Delta^{-\frac{1}{2}}(Y_i - \mu\Delta - \beta S_i) = \sqrt{S_i/\Delta}A_i$ tends to a normal distribution with mean zero and variance equal to the mean volatility, $E(v_t)$, provided that the process v is ergodic. This is in accordance with the empirical finding that the distribution of returns over short periods have heavy tails and are well approximated by generalized hyperbolic distributions, whereas the distribution of returns over long periods is close to a normal distribution. Limit theorems relating, for small Δ , the distribution of Y_i to the generalized hyperbolic distribution obtained by assuming that S_i is exactly generalized inverse Gaussian distributed are given in Genon-Catalot, Jeantheau & Larédo (1998). A rather different type of discrete time stochastic volatility models with exactly generalized hyperbolic distributed returns was proposed in Barndorff-Nielsen (1997). It should be noted that stochastic volatility models can be interpreted as being obtained by *subordination*. Here the operational time or business time is the integral of the volatility process $\tau(t) = \int_0^t v_s ds$, which can be interpreted as discussed in Section 3.

A simple specification of the volatility process v is to assume that it is one of the stationary and ergodic generalized inverse Gaussian diffusions defined in Section 4 as the solution of (4.7). A particularly simple choice is to assume that v is the stationary CIR-model given by (4.8), for which v_t is gamma-distributed so that a variance-gamma stochastic volatility model is obtained. This model was proposed by Hull & White (1988) and was considered further by Heston (1993). Its advantage is that analytically it is relatively tractable. For instance, all moments and mixed moments can be found explicitly, see e.g. Sørensen (2000). A problem is that because of the linear drift, the autocorrelation function is an exponential function, whereas it is a well-established empirical fact that the autocorrelation function of the volatility process decreases more slowly than a single exponential function. Under relatively weak regularity conditions a diffusion model has an exponentially decreasing autocorrelation function. A sufficient condition is that it is ρ -mixing, for which simple conditions are given in Genon-Catalot, Jeantheau & Larédo (2000). For this reason, stochastic volatility models with a diffusion volatility process can usually not fit the autocorrelation of the volatility process well.

In applications where the autocorrelation of the volatility process is important, a solution is to use the construction in subsection 4.4, i.e. to define the volatility process as the sum

$$v_t = v_t^{(1)} + \dots + v_t^{(m)}, \quad (5.5)$$

where $v_t^{(1)}, \dots, v_t^{(m)}$ are independent CIR-processes, with $v_t^{(i)}$ defined like the process $X^{(i)}$ given by (4.16). Also in this case a variance-gamma model is obtained, which is exactly as analytically tractable as the variance-gamma model just discussed, but the autocorrelation structure of the volatility process (5.5) is given by (4.15) and is thus very flexible. This approach is studied for more general diffusion models in Bibby & Sørensen (2001a).

It has been found empirically that for equities a fall in the price is associated with an increase in the future volatility. This phenomenon is referred to as leverage, Black (1976) and Nelson (1991). Stochastic volatility models of the form (5.2), where the Wiener process driving the price process is independent of the volatility process, as we have so far assumed, cannot deal with leverage, because for such a model the future fluctuations of the volatility are independent of the present price. We can, however, easily generalize the model to allow for the leverage phenomenon. Again we let the volatility process v be given by (5.5), and denote the Wiener process driving the k th CIR-process $v_t^{(k)}$ by $B^{(k)}$. Then we define the log-price process by

$$dX_t = (\mu + \beta v_t)dt + \sqrt{v_t}d\tilde{W}_t, \quad (5.6)$$

where \tilde{W} is the standard Wiener process

$$\tilde{W}_t = \frac{W_t + \rho \bar{B}_t}{\sqrt{1 + \rho^2}}$$

with $\rho \in \mathbb{R}$ and

$$\bar{B}_t = (B_t^{(1)} + \dots + B_t^{(m)})/\sqrt{m}. \quad (5.7)$$

A lengthy calculation shows that for $\beta = 0$ the covariance between Y_i and Y_{i+j}^2 ($j \geq 1$) is

$$\frac{\rho}{\sqrt{1 + \rho^2}} \frac{1}{\sqrt{m}} \sum_{k=1}^m b_k e^{-\lambda_k \Delta j}.$$

Here

$$b_k = \sqrt{2\beta} e^{\lambda_k \Delta} \left(1 - e^{-\lambda_k \Delta}\right)^2 \lambda_k^{-3/2} \mathbb{E} \left(\sqrt{v_1^{(k)} v_1} \right) > 0,$$

where β is the shape parameter of the gamma distribution of the volatility, and λ_k is the speed of reversion of the k th volatility component. We see that the correlation between Y_i and Y_{i+j}^2 is negative

if $\rho < 0$, which is exactly what we wanted. For $\rho = 0$ there is no leverage effect as expected. Note that the effect decreases as j tends to infinity. The decrease is of the same type as that of the autocorrelation function (4.15), but with different weights. It is thus very flexible and can in particular be slow.

Barndorff-Nielsen & Shephard (2001c) and Barndorff-Nielsen & Shephard (2001d) proposed to model the volatility process v as an Ornstein-Uhlenbeck type process, i.e. a solution to the stochastic differential equation (4.10). Such a process can be chosen stationary with a generalized inverse Gaussian marginal distribution, as discussed in Subsection 4.3. Processes of this type have the advantage that the drift is linear and the coefficient in front of the driving Lévy process is constant, which, analogous to the situation for the classical Wiener-driven Ornstein-Uhlenbeck process, implies an unusual analytic tractability. For instance the integrated volatility, which is a key quantity in finance, has the simple structure

$$\int_s^t v_s ds = \lambda^{-1} ((Z_t - Z_s) - (v_t - v_s)),$$

where $s < t$, and where Z is the driving Lévy process. This relation implies, for instance, that stochastic volatility processes of this type can be simulated as accurately as the volatility process can be simulated. This is because the random variables S_i , given by (5.4), are simple functions of the processes Z and v . An efficient method of simulating Ornstein-Uhlenbeck type processes is based on results by Rosinski (1991) and Rosinski (2001), see the exposition in Barndorff-Nielsen & Shephard (2001c).

Barndorff-Nielsen & Shephard (2001b) have studied the distributional properties of integrated Ornstein-Uhlenbeck type processes in detail. For the Ornstein-Uhlenbeck type volatility process with inverse Gaussian marginal distributions they found that while the integrated volatility process is not distributed exactly as the inverse Gaussian distribution, its tails have the same behaviour as this distribution. This implies that the returns will have the expected *NIG* tail behaviour.

For an Ornstein-Uhlenbeck type volatility process v , the autocorrelations of the discrete time processes S_i , given by (5.4), and Y_i^2 , with Y_i being the returns given by (5.3), are

$$\text{cor}(S_i, S_{i+j}) = d \exp(-\lambda\Delta(j-1)), \quad (5.8)$$

and

$$\text{cor}(Y_i^2, Y_{i+j}^2) = c \exp(-\lambda\Delta(j-1)), \quad (5.9)$$

where

$$\begin{aligned} 1 &\geq d = \frac{[1 - \exp(-\lambda\Delta)]^2}{2[\exp(-\lambda\Delta) - 1 + \lambda\Delta]} \\ &\geq c = \frac{[1 - \exp(-\lambda\Delta)]^2}{6[\exp(-\lambda\Delta) - 1 + \lambda\Delta] + 2(\lambda\Delta)^2(\xi/\omega)^2} \geq 0, \end{aligned}$$

with ξ and ω denoting the mean and variance of the volatility v_t . Therefore, as discussed in Barndorff-Nielsen & Shephard (2001d), S and Y^2 are constrained ARMA(1,1) processes with common autoregressive parameter, and with the moving average root being stronger for S than for Y^2 . The ARMA structure implies that the return process Y is weak GARCH(1,1) in the sense of Drost & Nijman (1993). Note that the formulae (5.8) and (5.9) also hold for the stochastic volatility model discussed above, where the volatility process is a CIR-diffusion. Hence for this model, the processes S and Y^2 have the same ARMA structure.

Barndorff-Nielsen & Shephard (2001d) also proposed a model with a Lévy-driven Ornstein-Uhlenbeck volatility process that allows for the leverage phenomenon. The log-price is modelled by

$$dX_t = (\mu + \beta v_t)dt + \sqrt{v_t}dW_t + \rho d\bar{Z}_t, \quad (5.10)$$

where $\bar{Z}_t = Z_t - E(Z_t)$ is the centered version of the Lévy process Z that drives the volatility process. This model has properties similar to those of the model with leverage discussed above (when $m = 1$). It is not a generalized hyperbolic model in the sense of the other stochastic volatility models in this section because of the term $\rho d\bar{Z}_t$. It is not clear to what extent the model is approximately hyperbolic. A complication is that the log-price process is a diffusion with jumps rather than a classical diffusion process driven by a Wiener process.

As already mentioned in Subsection 4.3, the autocorrelation function of an Ornstein-Uhlenbeck type process decreases exponentially, which, as also mentioned earlier, is faster than what is typically found in financial data. Volatility processes of the form (5.5), where $v_t^{(1)}, \dots, v_t^{(m)}$ are independent, stationary Ornstein-Uhlenbeck processes such that the marginal distribution of v is a generalized inverse Gaussian distribution, has a much more flexible autocorrelation structure. That such a volatility process exists was discussed in Subsection 4.4. Stochastic volatility models of this type often provide a much better fit to financial data. An example of this is given in Barndorff-Nielsen & Shephard (2001d). Also models where the volatility process is a sum of independent Ornstein-Uhlenbeck processes are analytically tractable.

Statistical inference for stochastic volatility models cannot easily be based on the likelihood function as it is not explicitly available and quite hard to simulate. Harvey, Ruiz & Shephard (1994) proposed a pseudo-likelihood method based on a Gaussian approximation that allowed them to apply the Kalman filter. More recently, likelihood based methods for stochastic volatility models have been proposed by Kim, Shephard & Chib (1998), and simulation based Bayesian methods using Markov chain Monte Carlo have been developed by Elerian, Chib & Shepard (2001) and Eraker (1998). A new and quite simple way of obtaining an approximate likelihood function for stochastic volatility models, which seems very promising, has been proposed by H. Sørensen (2001). The method takes advantage of the fact that lag- k conditional densities are relatively easy to obtain by simulation for stochastic volatility models. Other methods are the indirect inference methods of Gouriéroux, Monfort & Renault (1993), Gallant & Tauchen (1996), and Gallant & Long (1997). The prediction-based estimating functions of Sørensen (2000) can be applied to all models discussed in this section, while the estimators proposed by Genon-Catalot, Jeantheau & Larédo (1999) based on limit results (where the time between observations goes to zero) in Genon-Catalot, Jeantheau & Larédo (1998) are developed for volatility processes of the diffusion type. Recently methods based on realized volatility have been proposed, see Gloter (1999) and Barndorff-Nielsen & Shephard (2001a). Surveys that discuss the literature on stochastic volatility models up to 1995 can be found in Ghysels, Harvey & Renault (1996) and Shephard (1996).

Acknowledgement:

We are grateful to Ole E. Barndorff-Nielsen for his several useful comments on an earlier version of this paper. The data were put at our disposal by the Centre for Analytical Finance, which is funded by the Danish Social Science Research Council.

Appendix

In this appendix a few relevant definitions and results concerning Bessel functions are collected.

The *modified Bessel function of the third kind* with index $\lambda \in \mathbb{R}$ can be defined by the following integral representation,

$$K_\lambda(x) = \frac{1}{2} \int_0^\infty u^{\lambda-1} e^{-\frac{1}{2}x(u+u^{-1})} du, \quad x > 0.$$

The modified Bessel function has the following properties:

$$K_{-\lambda}(x) = K_{\lambda}(x), \quad (\text{A.1})$$

$$K_{\lambda+1}(x) = \frac{2\lambda}{x}K_{\lambda}(x) + K_{\lambda-1}(x), \quad (\text{A.2})$$

$$K'_{\lambda}(x) = -\frac{\lambda}{x}K_{\lambda}(x) - K_{\lambda-1}(x). \quad (\text{A.3})$$

For $\lambda = n + \frac{1}{2}$, $n = 0, 1, 2, \dots$, we have that

$$K_{n+\frac{1}{2}}(x) = \sqrt{\frac{\pi}{2x}}e^{-x} \left\{ 1 + \sum_{i=1}^n \frac{(n+i)!}{(n-i)!i!} (2x)^{-i} \right\}. \quad (\text{A.4})$$

For small values of the argument it holds that

$$K_{\lambda}(x) \sim \Gamma(\lambda)2^{\lambda-1}x^{-\lambda}, \quad x \downarrow 0, \quad \text{if } \lambda > 0. \quad (\text{A.5})$$

Similarly, we have for large values of the argument that

$$K_{\lambda}(x) = \sqrt{\frac{\pi}{2x}} \cdot e^{-x} \cdot \left\{ 1 + \frac{4\lambda^2 - 1}{8x} + \frac{(4\lambda^2 - 1)(4\lambda^2 - 9)}{2!(8x)^2} + \frac{(4\lambda^2 - 1)(4\lambda^2 - 9)(4\lambda^2 - 25)}{3!(8x)^3} + \dots \right\}. \quad (\text{A.6})$$

The *Bessel function of the first kind* with index $\lambda \in \mathbb{R}$ can be defined for $x > 0$ by

$$J_{\lambda}(x) = \frac{1}{\pi} \int_0^{\pi} \cos(x \sin(u) - \lambda u) du - \frac{\sin(\lambda\pi)}{\pi} \int_0^{\infty} e^{-x \sinh(u) - \lambda u} du.$$

For $\lambda > -\frac{1}{2}$ we have

$$J_{\lambda}(x) = \frac{2(\frac{1}{2}x)^{\lambda}}{\sqrt{\pi}\Gamma(\lambda + \frac{1}{2})} \int_0^1 (1-u^2)^{\lambda-\frac{1}{2}} \cos(xu) du, \quad x \in \mathbb{R},$$

where Γ denotes the gamma function.

The *Bessel function of the second kind* with index $\lambda \in \mathbb{R}$ can be defined for $x > 0$ by

$$Y_{\lambda}(x) = \frac{1}{\pi} \int_0^{\pi} \sin(x \sin(u) - \lambda u) du - \frac{1}{\pi} \int_0^{\infty} [e^{\lambda u} + e^{-\lambda u} \cos(\lambda\pi)] e^{-x \sinh(u)} du.$$

The function $Y_{\lambda}(x)$ is often denoted $N_{\lambda}(x)$ and is some times called Weber's function. The relationship between $J_{\lambda}(x)$ and $Y_{\lambda}(x)$ is

$$Y_{\lambda}(x) = \frac{J_{\lambda}(x) \cos(\lambda\pi) - J_{-\lambda}(x)}{\sin(\lambda\pi)}.$$

In connection with the *NIG*-distribution, it is useful to know that

$$J_{\frac{1}{2}}(x) = \sqrt{\frac{2}{\pi x}} \sin(x) \quad \text{and} \quad Y_{\frac{1}{2}}(x) = -\sqrt{\frac{2}{\pi x}} \cos(x).$$

References

- Aït-Sahalia, Y. (1998). “Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-form Approach”. Revised version of Working Paper 467, Graduate School of Business, University of Chicago. Forthcoming in *Econometrica*.
- Ané, T. & Geman, H. (2000). “Order flow, transaction clock and normality of asset returns”. *J. Finance*, 55:2259–2284.
- Bagnold, R. A. (1941). *The physics of blown sand and desert dunes*. Methuen, London.
- Barndorff-Nielsen, O. E. (1977). “Exponentially decreasing distributions for the logarithm of particle size”. *Proc. Roy. Soc. London*, A(353):401–419.
- Barndorff-Nielsen, O. E. (1978). “Hyperbolic Distributions and Distributions on Hyperbolae”. *Scand. J. Statist.*, 5:151–157.
- Barndorff-Nielsen, O. E. (1997). “Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling”. *Scand. J. Statist.*, 24:1–13.
- Barndorff-Nielsen, O. E. (1998). “Processes of Normal Inverse Gaussian Type”. *Finance and Stochastics*, 2:41–68.
- Barndorff-Nielsen, O. E. (2000). “Superposition of Ornstein-Uhlenbeck Type Processes”. *Theor. Probab. Applic.*, 45.
- Barndorff-Nielsen, O. E.; Blæsild, P.; Jensen, J. L. & Sørensen, M. (1985). “The Fascination of Sand”. In Atkinson, A. C. & Fienberg, S. E., editors, *A Celebration of Statistics*, pages 57–87. Springer, New York.
- Barndorff-Nielsen, O. E. & Blæsild, P. (1980). “Hyperbolic Distributions and Ramifications: Contributions to Theory and Application”. Research Report 68, Department of Theoretical Statistics, Institute of Mathematics, University of Aarhus.
- Barndorff-Nielsen, O. E. & Halgreen, C. (1977). “Infinite divisibility of the hyperbolic and generalized inverse Gaussian distributions”. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 38:309–312.
- Barndorff-Nielsen, O. E.; Jensen, J. L. & Sørensen, M. (1990). “Parametric Modelling of Turbulence”. *Phil. Trans. R. Soc. Lond. A*, 332:439–455.
- Barndorff-Nielsen, O. E.; Jensen, J. L. & Sørensen, M. (1998). “Some Stationary Processes in Discrete and Continuous Time”. *Advances in Applied Probability*, 30:989–1007.
- Barndorff-Nielsen, O. E. & Shephard, N. (2001a). “Econometric analysis of realised volatility and its use in estimating stochastic volatility models”. Working paper, Centre for Analytical Finance, University of Aarhus.
- Barndorff-Nielsen, O. E. & Shephard, N. (2001b). “Integrated OU processes and non-Gaussian OU-based stochastic volatility models”. Research report, MaPhySto, University of Aarhus.
- Barndorff-Nielsen, O. E. & Shephard, N. (2001c). “Modelling by Lévy Processes for Financial Econometrics”. In Barndorff-Nielsen, O. E.; Mikosch, T. & Resnick, S., editors, *Lévy Processes: Theory and Applications*. Birkhäuser, Boston.

- Barndorff-Nielsen, O. E. & Shephard, N. (2001d). “Non-Gaussian Ornstein-Uhlenbeck-Based Models and some of their Uses in Financial Econometrics (with discussion)”. *Journal of the Royal Statistical Society B*, 63:167–241.
- Bertoin, J. (1996). *Lévy Processes*. Cambridge University Press.
- Bibby, B. M. & Sørensen, M. (1995). “Martingale Estimation Functions for Discretely Observed Diffusion Processes”. *Bernoulli*, 1:17–39.
- Bibby, B. M. & Sørensen, M. (1996). “On Estimation for Discretely Observed Diffusions: A Review”. *Theory of Stochastic Processes*, 2:49–56.
- Bibby, B. M. & Sørensen, M. (1997). “A Hyperbolic Diffusion Model for Stock Prices”. *Finance and Stochastics*, 1:25–41.
- Bibby, B. M. & Sørensen, M. (2001a). “Flexible Stochastic Volatility Models of the Diffusion type”. In preparation.
- Bibby, B. M. & Sørensen, M. (2001b). “Simplified Estimating Functions for Diffusion Models with a High-dimensional Parameter”. *Scand. J. Statist.*, 28(1):99–112.
- Black, F. (1976). “Studies of Stock Price Volatility Changes”. *Proc. Bus. Econ. Statist. Sect. Am. Statist. Ass.*, pages 177–181.
- Blæsild, P. & Sørensen, M. K. (1992). “HYP - A Computer program for Analyzing Data by Means of the Hyperbolic Distribution”. Research Report No. 248, Department of Theoretical Statistics, Institute of Mathematics, University of Aarhus.
- Blæsild, P. (1978). “The Shape of the Generalized Inverse Gaussian and Hyperbolic Distributions”. Research Report 37, Department of Theoretical Statistics, Institute of Mathematics, University of Aarhus.
- Blattberg, R. C. & Gonedes, N. (1974). “A Comparison of the Stable and Student Distributions as Models for Stock Prices”. *Journal of Business*, 47:244–280.
- Chan, K. C.; Karolyi, G. A.; Longstaff, F. A. & Sanders, A. B. (1992). “An Empirical Comparison of Alternative Models of the Short-term Interest Rate”. *Journal of Finance*, 47:1209–1227.
- Clark, P. K. (1973). “A Subordinated Stochastic Process with Finite Variance for Speculative Prices”. *Econometrica*, 41.
- Cox, D. R. (1984). “Long-range Dependence: A Review”. In David, H. A. & David, H. T., editors, *Statistics: An Appraisal*. Iowa State University Press.
- Cox, J. C.; Ingersoll, Jr., J. E. & Ross, S. A. (1985). “A Theory of the Term Structure of Interest Rates”. *Econometrica*, 53(2):385–407.
- Dacunha-Castelle, D. & Florens-Zmirou, D. (1986). “Estimation of the Coefficients of a Diffusion from Discrete Observations”. *Stochastics*, 19:263–284.
- Drost, F. C. & Nijman, T. E. (1993). “Temporal Aggregation of GARCH Processes”. *Econometrica*, 61:909–927.
- Eberlein, E. (2001). “Application of Generalized Hyperbolic Lévy Motions to Finance”. In Barndorff-Nielsen, O. E.; Mikosch, T. & Resnick, S., editors, *Lévy Processes – Theory and Applications*, pages 319–337. Birkhäuser, Boston.

- Eberlein, E. & Jacod, J. (1997). “On the Range of Option Prices”. *Finance and Stochastics*, 1:131–140.
- Eberlein, E. & Keller, U. (1995). “Hyperbolic Distributions in Finance”. *Bernoulli*, 1:281–299.
- Eberlein, E.; Keller, U. & Prause, K. (1998). “New Insights into Smile, Mispricing and Value at Risk: The Hyperbolic Model”. *Journal of Business*, 71:371–406.
- Eberlein, E. & Prause, K. (1998). “The Generalized Hyperbolic Model: Financial Derivatives and Risk Measures”. FDM-Preprint 56, Universität Freiburg.
- Eberlein, E. & Raible, S. (1999). “Term Structure Models Driven by General Lévy Processes”. *Mathematical Finance*, 9:31–53.
- Eberlein, E. & Raible, S. (2000). “Some analytic facts on the generalized hyperbolic model”. In *Proceedings of the Third European Meeting of Mathematicians*. Birkhäuser: Progress), Boston.
- Elerian, O.; Chib, S. & Shepard, N. (2001). “Likelihood Inference for Discretely Observed Non-linear Diffusions”. *Econometrica*, 69. Forthcoming.
- Epps, T. W. & Epps, M. L. (1976). “The Stochastic Dependence of Security Price Changes and Transaction Volumes: Implications for the Mixture-of-Distributions Hypothesis”. *Econometrica*, 44:305–321.
- Eraker, B. (1998). “MCMC Analysis of Diffusion Models with Application to Finance”. Discussion Paper 1998-5, Department of Finance and Management Science, Norwegian School of Economics and Business Administration.
- Gallant, A. R. & Long, J. R. (1997). “Estimating Stochastic Differential Equations Efficiently by Minimum Chi-square”. *Biometrika*, 84:125–141.
- Gallant, A. R. & Tauchen, G. (1996). “Which Moments to Match?”. *Econometric Theory*, 12:657–681.
- Genon-Catalot, V.; Jeantheau, T. & Larédo, C. (1998). “Limit Theorems for Discretely Observed Stochastic Volatility Models”. *Bernoulli*, 4:283–303.
- Genon-Catalot, V.; Jeantheau, T. & Larédo, C. (1999). “Parameter Estimation for Discretely Observed Stochastic Volatility Models”. *Bernoulli*, 5:855–872.
- Genon-Catalot, V.; Jeantheau, T. & Larédo, C. (2000). “Stochastic Volatility Models as Hidden Markov Models and Statistical Applications”. *Bernoulli*, 6:1051–1079.
- Ghysels, E.; Harvey, A. C. & Renault, E. (1996). “Stochastic Volatility”. In Rao, C. R. & Maddala, G. S., editors, *Statistical Methods in Finance*, pages 119–191. North-Holland, Amsterdam.
- Gloter, A. (1999). “Parameter estimation for a hidden diffusion”. Preprint 20/99, University of Marne-la-Vallée.
- Good, I. J. (1953). “The Population Frequences of Species and the Estimation of Population Parameters”. *Biometrika*, 40:237–264.
- Gourieroux, C.; Monfort, A. & Renault, E. (1993). “Indirect Inference”. *Journal of Applied Econometrics*, 8:S85–S118.
- Halgreen, C. (1979). “Self-decomposability of generalized inverse Gaussian and hyperbolic distributions”. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 47:13–17.
- Hansen, L. P. & Scheinkman, J. A. (1995). “Back to the Future: Generating Moment Implications for Continuous-time Markov Pocesess”. *Econometrica*, 63:767–804.

- Harvey, A. C.; Ruiz, E. & Shephard, N. (1994). “Multivariate Stochastic Variance Models”. *Review of Economic Studies*, 61:247–264.
- Heston, S. L. (1993). “A Closed-form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options”. *Review of Financial Studies*, 6:327–343.
- Honoré, P. (1997). “Maximum Likelihood Estimation of Non-linear Continuous-time Term-structure Models”. Working Paper 1997-7, Department of Finance, Aarhus School of Business.
- Hull, J. & White, A. (1988). “An Analysis of the Bias in Option Pricing Caused by a Stochastic Volatility”. *Advances in Futures and Options Research*, 3:29–61.
- Hurst, S. R. (1997). *On the Stochastic Dynamics of Stock Market Volatility*. PhD thesis, Australian National University.
- Hurst, S. R.; Platen, E. & Rachev, S. T. (1997). “Subordinated Market Index Models: A Comparison”. *Financial Engineering and the Japanese Markets*, 4:97–124.
- Jacobsen, M. (2001). “Discretely Observed Diffusions; Classes of Estimating Functions and Small Δ -optimality”. *Scand. J. Statist.*, 28(1):123–150.
- Jacod, J. & Shiryaev, A. N. (1987). *Limit Theorems for Stochastic Processes*. Springer, New York.
- Jensen, J. L. & Pedersen, J. (1999). “Ornstein-Uhlenbeck type processes with non-normal distribution”. *J. Appl. Prob.*, 36:389–402.
- Jiang, W. (2000). *Some Simulation-based Models towards Mathematical Finance*. PhD thesis, University of Aarhus.
- Jørgensen, B. (1982). *Statistical Properties of the Generalized Inverse Gaussian Distribution*. Springer-Verlag, New York. Lecture Notes in Statistics 9.
- Jurek, Z. J. & Mason, J. D. (1993). *Operator-Limit Distributions in Probability Theory*. Wiley, New York.
- Jurek, Z. J. & Vervaat, W. (1983). “An Integral Representation for Self-decomposable Banach Space Valued Random Variables”. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 62:247–262.
- Keller, U. (1997). *Realistic Modelling of Financial Derivatives*. PhD thesis, Universität Freiburg.
- Kessler, M. (2000). “Simple and Explicit Estimating Functions for a Discretely Observed Diffusion Process”. *Scand. J. Statist.*, 27:65–82.
- Kessler, M. & Sørensen, M. (1999). “Estimating Equations Based on Eigenfunctions for a Discretely Observed Diffusion Process”. *Bernoulli*, 5:299–314.
- Kim, S.; Shephard, N. & Chib, S. (1998). “Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models”. *Review of Economic Studies*, 65:361–393.
- Küchler, U.; Neumann, K.; Sørensen, M. & Streller, A. (1999). “Stock Returns and Hyperbolic Distributions”. *Mathematical and Computer Modelling*, 29:1–15.
- Madan, D. B. & Chang, E. C. (1996). “Volatility Smiles, Skewness Premia and Risk Metrics: Application of a Four Parameter Closed Form Generalization of Geometric Brownian Motion to the Pricing of Options”. Presented at a Conference on Mathematical Finance, University of Aarhus, Denmark.

- Madan, D. B. & Milne, F. (1991). “Option Pricing with V.G. Martingale Components”. *Mathematical Finance*, 1:39–55.
- Madan, D. B. & Seneta, E. (1990). “The Variance Gamma (V.G.) Model for Share Market Returns”. *Journal of Business*, 63:511–524.
- Nelson, D. B. (1991). “Conditional Heteroskedasticity in Asset Pricing: A New Approach”. *Econometrica*, 59:347–370.
- Pedersen, A. R. (1995). “A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete Observations”. *Scand. J. Statist.*, 22:55–71.
- Poulsen, R. (1999). “Approximate Maximum Likelihood Estimation of Discretely Observed Diffusion Processes”. Working Paper 29, Centre for Analytical Finance, Aarhus.
- Praetz, P. D. (1972). “The Distribution of Share Prices”. *Journal of Business*, 45:49–55.
- Prause, K. (1999). *The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measure*. PhD thesis, Universität Freiburg.
- Protter, P. (1990). *Stochastic Integration and Differential Equations – A New Approach*. Springer, New York.
- Raible, S. (2000). *Lévy Processes in Finance: Theory, Numerics, and Empirical Facts*. PhD thesis, Universität Freiburg.
- Rosinski, J. (1991). “On a Class of Infinitely Divisible Processes Represented as Mixtures of Gaussian Processes”. In Cambanis, S.; Samorodnitsky, G. & Taqqu, M. S., editors, *Stable Processes and Related Topics*, pages 27–41. Birkhäuser, Boston.
- Rosinski, J. (2001). “Series Representations of Lévy Processes from the Perspective of Point Processes”. In Barndorff-Nielsen, O. E.; Mikosch, T. & Resnick, S., editors, *Lévy Processes – Theory and Applications*. Birkhäuser, Boston.
- Rydberg, T. H. (1997). “The Normal Inverse Gaussian Lévy Process: Simulation and Approximation”. *Communications in Statistics: Stochastic Models*, 13:887–910.
- Rydberg, T. H. (1999). “Generalized Hyperbolic Diffusion Processes with Applications in Finance”. *Mathematical Finance*, 9:183–201.
- Rydberg, T. H. (2000). “Realistic Statistical Modelling of Financial Data”. *Int. Statist. Rev.*, 68:233–258.
- Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.
- Sato, K.; Watanabe, T. & Yamazato, M. (1994). “Recurrence Conditions for Multivariate Processes of Ornstein-Uhlenbeck Type”. *J. Math. Soc. Japan*, 46:245–265.
- Sato, K. & Yamazato, M. (1982). “Stationary Processes of the Ornstein-Uhlenbeck Type”. In Ito, K. & Prohorov, J. V., editors, *Probability Theory and Mathematical Statistics*. Springer-Verlag, Berlin. Lecture Notes in Math 1021.
- Sato, K. & Yamazato, M. (1984). “Operator-selfdecomposable Distributions as Limit Distributions of Processes of the Ornstein-Uhlenbeck Type”. *Stoch. Proc. Appl.*, 17:73–100.
- Seshadri, V. (1997). “Halphen’s laws”. In Kotz, S.; Read, C. B. & Banks, D. L., editors, *Encyclopedia of Statistical Sciences, Update Volume 1*, pages 302–306. Wiley, New York.

- Shephard, N. (1996). “Statistical Aspects of ARCH and Stochastic Volatility”. In Cox, D. R.; Hinkley, D. V. & Barndorff-Nielsen, O. E., editors, *Time Series Models in Econometrics, Finance and Other Fields*, pages 1–67. Chapman & Hall, London.
- Sichel, H. S. (1973). “Statistical evaluation of diamondiferous deposits”. *J. S. Afr. Inst. Min. Metall.*, 76:235–243.
- Sørensen, H. (2000). *Inference for Diffusion Processes and Stochastic Volatility Models*. PhD thesis, Department of Statistics and Operations Research, University of Copenhagen.
- Sørensen, H. (2001). “Simulated Likelihood Approximations for Stochastic Volatility Models”. Preprint 1, Department of Theoretical Statistics, University of Copenhagen.
- Sørensen, M. (1997a). “Estimating Functions for Discretely Observed Diffusions: A Review”. In Basawa, I. V.; Godambe, V. P. & Taylor, R. L., editors, *Selected Proceedings of the Symposium on Estimating Functions*. Hayward: Institute of Mathematical Statistics. IMS Lecture Notes – Monograph Series, Vol. 32.
- Sørensen, M. (1997b). “Exponential Family Inference for Diffusion Models”. Research Report No. 383, Department of Theoretical Statistics, Institute of Mathematics, University of Aarhus.
- Sørensen, M. (2000). “Prediction-based Estimating Functions”. *Econometrics Journal*, 3:123–147.
- Wolfe, S. J. (1982). “On a Continuous Analogue of the Stochastic Difference Equation $X_n = \rho X_{n-1} + B_n$ ”. *Stoch. Proc. Appl.*, 12:301–312.