HSC Research Report

HSC/10/01

Models for Heavy-tailed Asset Returns

Szymon Borak*
Adam Misiorek**
Rafał Weron***

* Deutsche Bank AG, London, U.K.
** Santander Consumer Bank S.A., Wrocław, Poland
*** Institute of Organization and Management, Wrocław
University of Technology, Poland

Models for Heavy-tailed Asset Returns[1]

Szymon Borak[2], Adam Misiorek[3], and Rafał Weron[4]

Abstract: Many of the concepts in theoretical and empirical finance developed over the past decades – including the classical portfolio theory, the Black-Scholes-Merton option pricing model or the RiskMetrics variance-covariance approach to VaR – rest upon the assumption that asset returns follow a normal distribution. But this assumption is not justified by empirical data! Rather, the empirical observations exhibit excess kurtosis, more colloquially known as fat tails or heavy tails. This chapter is intended as a guide to heavy-tailed models. We first describe the historically oldest heavy-tailed model – the stable laws. Next, we briefly characterize their recent lighter-tailed generalizations, the so-called truncated and tempered stable distributions. Then we study the class of generalized hyperbolic laws, which – like tempered stable distributions – can be classified somewhere between infinite variance stable laws and the Gaussian distribution. Finally, we provide numerical examples.

Keywords: Heavy-tailed distribution; Stable distribution; Tempered stable distribution; Generalized hyperbolic distribution; Asset return; Random number generation; Parameter estimation

JEL: C13, C15, C16, G32

---

[2] Deutsche Bank AG, London, U.K.
[3] Santander Consumer Bank S.A., Wrocław, Poland
[4] Institute of Organization and Management, Wrocław University of Technology, Poland

1 Models for heavy-tailed asset returns

Szymon Borak, Adam Misiorek, and Rafał Weron

1.1 Introduction

Many of the concepts in theoretical and empirical finance developed over the past decades – including the classical portfolio theory, the Black-Scholes-Merton option pricing model or the RiskMetrics variance-covariance approach to Value at Risk (VaR) – rest upon the assumption that asset returns follow a normal distribution. But this assumption is not justified by empirical data! Rather, the empirical observations exhibit excess kurtosis, more colloquially known as *fat tails* or *heavy tails* (Guillaume et al., 1997; Rachev and Mittnik, 2000). The contrast with the Gaussian law can be striking, as in Figure 1.1 where we illustrate this phenomenon using a ten-year history of the Dow Jones Industrial Average (DJIA) index.

In the context of VaR calculations, the problem of the underestimation of risk by the Gaussian distribution has been dealt with by the regulators in an *ad hoc* way. The Basle Committee on Banking Supervision (1995) suggested that for the purpose of determining minimum capital reserves financial institutions use a 10-day VaR at the 99% confidence level multiplied by a safety factor $s \in [3, 4]$. Stahl (1997) and Danielsson, Hartmann and De Vries (1998) argue convincingly that the range of s is a result of the heavy-tailed nature of asset returns. Namely, if we assume that the distribution is symmetric and has finite variance σ^2 then from Chebyshev's inequality we have $\mathbb{P}(Loss \geq \epsilon) \leq \frac{1}{2}\sigma^2\epsilon^2$. Setting the right hand side to 1% yields an upper bound for $\text{VaR}_{99\%} \leq 7.07\sigma$. On the other hand, if we assume that returns are normally distributed we arrive at $\text{VaR}_{99\%} \leq 2.33\sigma$, which is roughly three times lower than the bound obtained for a heavy-tailed, finite variance distribution.
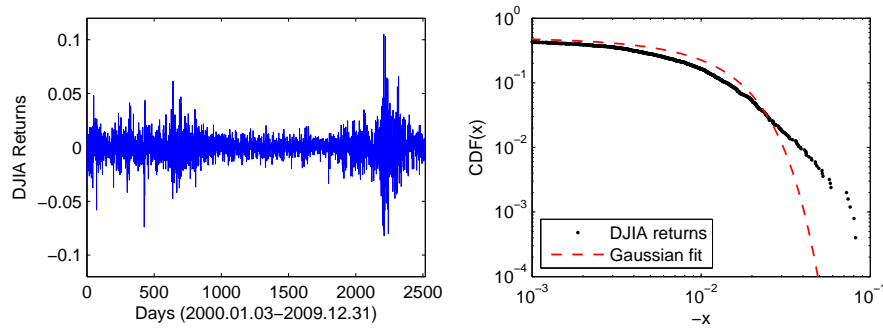
Figure 1.1: *Left panel*: Returns $\log(X_{t+1}/X_t)$ of the DJIA daily closing values X_t from the period January 3, 2000 – December 31, 2009. *Right panel*: Gaussian fit to the empirical cumulative distribution function (cdf) of the returns on a double logarithmic scale (only the left tail fit is displayed).

STFstab01

Being aware of the underestimation of risk by the Gaussian law we should consider using heavy-tailed alternatives. This chapter is intended as a guide to such models. In Section 1.2 we describe the historically oldest heavy-tailed model – the stable laws. Next, in Section 1.3 we briefly characterize their recent lighter-tailed generalizations, the so-called truncated and tempered stable distributions. In Section 1.4 we study the class of generalized hyperbolic laws, which – like tempered stable distributions – can be classified somewhere between infinite variance stable laws and the Gaussian distribution. Finally, in Section 1.5 we provide numerical examples.

1.2 Stable distributions

1.2.1 Definitions and basic properties

The theoretical rationale for modeling asset returns by the Gaussian distribution comes from the Central Limit Theorem (CLT), which states that the sum of a large number of independent, identically distributed (i.i.d.) variables – say, decisions of investors – from a finite-variance distribution will be (asymp-
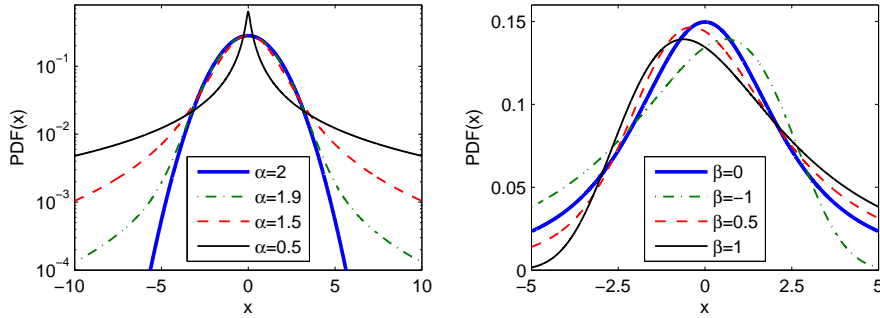
Figure 1.2: *Left panel*: A semi-logarithmic plot of symmetric ($\beta = \mu = 0$) stable densities for four values of α. Note, the distinct behavior of the Gaussian ($\alpha = 2$) distribution. *Right panel*: A plot of stable densities for $\alpha = 1.2$ and four values of β.

STFstab02

totically) normally distributed. Yet, this beautiful theoretical result has been notoriously contradicted by empirical findings. Possible reasons for the failure of the CLT in financial markets are (i) infinite-variance distributions of the variables, (ii) non-identical distributions of the variables, (iii) dependences between the variables or (iv) any combination of the three. If only the finite variance assumption is released we have a straightforward solution by virtue of the generalized CLT, which states that the limiting distribution of sums of such variables is stable (Nolan, 2010). This, together with the fact that stable distributions are leptokurtic and can accommodate fat tails and asymmetry, has led to their use as an alternative model for asset returns since the 1960s.

Stable laws – also called α-stable, stable Paretian or Lévy stable – were introduced by Paul Lévy in the 1920s. The name 'stable' reflects the fact that a sum of two independent random variables having a stable distribution with the same index α is again stable with index α. This invariance property holds also for Gaussian variables. In fact, the Gaussian distribution is stable with $\alpha = 2$.

For complete description the stable distribution requires four parameters. The index of stability $\alpha \in (0, 2]$, also called the tail index, tail exponent or characteristic exponent, determines the rate at which the tails of the distribution taper off, see the left panel in Figure 1.2. The skewness parameter $\beta \in [-1, 1]$ defines the asymmetry. When $\beta > 0$, the distribution is skewed to the right, i.e.

the right tail is thicker, see the right panel in Figure 1.2. When it is negative, it is skewed to the left. When $\beta = 0$, the distribution is symmetric about the mode (the peak) of the distribution. As α approaches 2, β loses its effect and the distribution approaches the Gaussian distribution regardless of β. The last two parameters, $\sigma > 0$ and $\mu \in \mathbb{R}$, are the usual scale and location parameters, respectively.

A far-reaching feature of the stable distribution is the fact that its probability density function (pdf) and cumulative distribution function (cdf) do not have closed form expressions, with the exception of three special cases. The best known of these is the Gaussian ($\alpha = 2$) law whose pdf is given by:

$$f_G(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}. \tag{1.1}$$

The other two are the lesser known Cauchy ($\alpha = 1$, $\beta = 0$) and Lévy ($\alpha = 0.5$, $\beta = 1$) laws. Consequently, the stable distribution can be most conveniently described by its characteristic function (cf) – the inverse Fourier transform of the pdf. The most popular parameterization of the characteristic function $\phi(t)$ of $X \sim S_\alpha(\sigma, \beta, \mu)$, i.e. a stable random variable with parameters α, σ, β and μ, is given by (Samorodnitsky and Taqqu, 1994; Weron, 1996):

$$\log \phi(t) = \begin{cases} -\sigma^\alpha |t|^\alpha \{1 - i\beta \mathrm{sign}(t) \tan \frac{\pi\alpha}{2}\} + i\mu t, & \alpha \neq 1, \\ \\ -\sigma|t|\{1 + i\beta \mathrm{sign}(t) \frac{2}{\pi} \log|t|\} + i\mu t, & \alpha = 1. \end{cases} \tag{1.2}$$

Note, that the traditional scale parameter σ of the Gaussian distribution is not the same as σ in the above representation. A comparison of formulas (1.1) and (1.2) yields the relation: $\sigma_{\mathrm{Gaussian}} = \sqrt{2}\sigma$.

For numerical purposes, it is often useful to use Nolan's (1997) parameterization:

$$\log \phi_0(t) = \begin{cases} -\sigma^\alpha |t|^\alpha \{1 + i\beta \mathrm{sign}(t) \tan \frac{\pi\alpha}{2}[(\sigma|t|)^{1-\alpha} - 1]\} + i\mu_0 t, & \alpha \neq 1, \\ \\ -\sigma|t|\{1 + i\beta \mathrm{sign}(t) \frac{2}{\pi} \log(\sigma|t|)\} + i\mu_0 t, & \alpha = 1, \end{cases} \tag{1.3}$$

which yields a cf (and hence the pdf and cdf) jointly continuous in all four parameters. The location parameters of the two representations (S and S^0) are related by $\mu = \mu_0 - \beta\sigma \tan \frac{\pi\alpha}{2}$ for $\alpha \neq 1$ and $\mu = \mu_0 - \beta\sigma \frac{2}{\pi} \log \sigma$ for $\alpha = 1$.

The 'fatness' of the tails of a stable distribution can be derived from the following property: the pth moment of a stable random variable is finite if and only if $p < \alpha$. Hence, when $\alpha > 1$ the mean of the distribution exists (and is equal to μ). On the other hand, when $\alpha < 2$ the variance is infinite and the tails exhibit a power-law behavior (i.e. they are asymptotically equivalent to a Pareto law). More precisely, using a CLT type argument it can be shown that (Janicki and Weron, 1994a; Samorodnitsky and Taqqu, 1994):

$$\begin{cases} \lim_{x \to \infty} x^\alpha \mathbb{P}(X > x) = C_\alpha(1 + \beta)\sigma^\alpha, \\ \lim_{x \to \infty} x^\alpha \mathbb{P}(X < -x) = C_\alpha(1 + \beta)\sigma^\alpha, \end{cases} \tag{1.4}$$

where $C_\alpha = \left(2 \int_0^\infty x^{-\alpha} \sin(x) dx\right)^{-1} = \frac{1}{\pi}\Gamma(\alpha)\sin\frac{\pi\alpha}{2}$. The convergence to the power-law tail varies for different α's and is slower for larger values of the tail index. Moreover, the tails of stable cdfs exhibit a crossover from an approximate power decay with exponent $\alpha > 2$ to the true tail with exponent α. This phenomenon is more visible for large α's (Weron, 2001).

1.2.2 Computation of stable density and distribution functions

The lack of closed form formulas for most stable densities and distribution functions has far-reaching consequences. Numerical approximation or direct numerical integration have to be used instead of analytical formulas, leading to a drastic increase in computational time and loss of accuracy. Despite a few early attempts in the 1970s, efficient and general techniques have not been developed until late 1990s.

Mittnik, Doganoglu and Chenyao (1999) exploited the pdf–cf relationship and applied the fast Fourier transform (FFT). However, for data points falling between the equally spaced FFT grid nodes an interpolation technique has to be used. The authors suggested that linear interpolation suffices in most practical applications, see also Rachev and Mittnik (2000). Taking a larger number of grid points increases accuracy, however, at the expense of higher computational burden. Setting the number of grid points to $N = 2^{13}$ and the grid spacing to $h = 0.01$ allows to achieve comparable accuracy to the direct integration method (see below), at least for typically used values of $\alpha > 1.6$.

As for the computational speed, the FFT based approach is faster for large samples, whereas the direct integration method favors small data sets since it can be computed at any arbitrarily chosen point. Mittnik, Doganoglu and Chenyao (1999) report that for $N = 2^{13}$ the FFT based method is faster

for samples exceeding 100 observations and slower for smaller data sets. We must stress, however, that the FFT based approach is not as universal as the direct integration method – it is efficient only for large alpha's and only as far as the pdf calculations are concerned. When computing the cdf the former method must numerically integrate the density, whereas the latter takes the same amount of time in both cases.

The direct integration method, proposed by Nolan (1997, 1999), consists of a numerical integration of Zolotarev's (1986) formulas for the density or the distribution function. Set $\zeta = -\beta \tan \frac{\pi\alpha}{2}$. Then the density $f(x; \alpha, \beta)$ of a standard stable random variable in representation S^0, i.e. $X \sim S_\alpha^0(1, \beta, 0)$, can be expressed as (note, that Zolotarev (1986, Section 2.2) used another parametrization):

- when $\alpha \neq 1$ and $x \neq \zeta$:

$$f(x; \alpha, \beta) = \frac{\alpha(x-\zeta)^{\frac{1}{\alpha-1}}}{\pi \mid \alpha - 1 \mid} \int_{-\xi}^{\frac{\pi}{2}} V(\theta; \alpha, \beta) \exp\left\{-(x-\zeta)^{\frac{\alpha}{\alpha-1}} V(\theta; \alpha, \beta)\right\} d\theta,$$

(1.5)

for $x > \zeta$ and $f(x; \alpha, \beta) = f(-x; \alpha, -\beta)$ for $x < \zeta$,

- when $\alpha \neq 1$ and $x = \zeta$:

$$f(x; \alpha, \beta) = \frac{\Gamma(1 + \frac{1}{\alpha}) \cos(\xi)}{\pi(1 + \zeta^2)^{\frac{1}{2\alpha}}},$$

- when $\alpha = 1$:

$$f(x; 1, \beta) = \begin{cases} \frac{1}{2|\beta|} e^{\frac{\pi x}{2\beta}} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} V(\theta; 1, \beta) \exp\left\{-e^{\frac{\pi x}{2\beta}} V(\theta; 1, \beta)\right\} d\theta, & \beta \neq 0, \\[2mm] \frac{1}{\pi(1+x^2)}, & \beta = 0, \end{cases}$$

where

$$\xi = \begin{cases} \frac{1}{\alpha} \arctan(-\zeta), & \alpha \neq 1, \\ \frac{\pi}{2}, & \alpha = 1, \end{cases}$$

(1.6)

and

$$V(\theta; \alpha, \beta) = \begin{cases} (\cos \alpha\xi)^{\frac{1}{\alpha-1}} \left(\frac{\cos\theta}{\sin\alpha(\xi+\theta)}\right)^{\frac{\alpha}{\alpha-1}} \frac{\cos\{\alpha\xi + (\alpha-1)\theta\}}{\cos\theta}, & \alpha \neq 1, \\ \frac{2}{\pi} \left(\frac{\frac{\pi}{2} + \beta\theta}{\cos\theta}\right) \exp\left\{\frac{1}{\beta}(\frac{\pi}{2} + \beta\theta) \tan\theta\right\}, & \alpha = 1, \beta \neq 0. \end{cases}$$

The distribution $F(x; \alpha, \beta)$ of a standard stable random variable in representation S^0 can be expressed as:

- when $\alpha \neq 1$ and $x \neq \zeta$:

$$F(x; \alpha, \beta) = c_1(\alpha, \beta) + \frac{\text{sign}(1 - \alpha)}{\pi} \int_{-\xi}^{\frac{\pi}{2}} \exp\left\{-(x - \zeta)^{\frac{\alpha}{\alpha - 1}} V(\theta; \alpha, \beta)\right\} d\theta,$$

  for $x > \zeta$ and $F(x; \alpha, \beta) = 1 - F(-x; \alpha, -\beta)$ for $x < \zeta$, where

$$c_1(\alpha, \beta) = \begin{cases} \frac{1}{\pi}\left(\frac{\pi}{2} - \xi\right), & \alpha < 1, \\ 1, & \alpha > 1, \end{cases}$$

- when $\alpha \neq 1$ and $x = \zeta$:

$$F(x; \alpha, \beta) = \frac{1}{\pi}\left(\frac{\pi}{2} - \xi\right),$$

- when $\alpha = 1$:

$$F(x; 1, \beta) = \begin{cases} \frac{1}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \exp\left\{-e^{-\frac{\pi x}{2\beta}} V(\theta; 1, \beta)\right\} d\theta, & \beta > 0, \\[2mm] \frac{1}{2} + \frac{1}{\pi} \arctan x, & \beta = 0, \\[2mm] 1 - F(x, 1, -\beta), & \beta < 0. \end{cases}$$

Formula (1.5) requires numerical integration of the function $g(\cdot) \exp\{-g(\cdot)\}$, where $g(\theta; x, \alpha, \beta) = (x - \zeta)^{\frac{\alpha}{\alpha - 1}} V(\theta; \alpha, \beta)$. The integrand is 0 at $-\xi$, increases monotonically to a maximum of $\frac{1}{e}$ at point θ^* for which $g(\theta^*; x, \alpha, \beta) = 1$, and then decreases monotonically to 0 at $\frac{\pi}{2}$ (Nolan, 1997). However, in some cases the integrand becomes very peaked and numerical algorithms can miss the spike and underestimate the integral. To avoid this problem we need to find the argument θ^* of the peak numerically and compute the integral as a sum of two integrals: one from $-\xi$ to θ^* and the other from θ^* to $\frac{\pi}{2}$.

To the best of our knowledge, currently no statistical computing environment offers the computation of stable density and distribution functions in its standard release. Users have to rely on third-party libraries or commercial products. A few are worth mentioning. The standalone program STABLE is probably the most efficient (downloadable from John Nolan's web page: academic2.american.edu/˜jpnolan/stable/stable.html). It was written in Fortran

and calls several external IMSL routines, see Nolan (1997) for details. Apart from speed, the STABLE program also exhibits high relative accuracy (ca. 10^{-13}; for default tolerance settings) for extreme tail events and 10^{-10} for values used in typical financial applications (like approximating asset return distributions). The STABLE program is also available in library form through Robust Analysis Inc. (www.robustanalysis.com). This library provides interfaces to Matlab, S-plus/R and Mathematica.

In the late 1990s Diethelm Würtz has initiated the development of Rmetrics, an open source collection of S-plus/R software packages for computational finance (www.rmetrics.org). In the *fBasics* package stable pdf and cdf calculations are performed using the direct integration method, with the integrals being computed by R's function *integrate*. On the other hand, the FFT based approach is utilized in Cognity, a commercial risk management platform that offers derivatives pricing and portfolio optimization based on the assumption of stably distributed returns (www.finanalytica.com). The FFT implementation is also available in Matlab (*stablepdf_fft.m*) from the Statistical Software Components repository (ideas.repec.org/c/boc/bocode/m429004.html).

1.2.3 Simulation of stable variables

Simulating sequences of stable random variables is not straightforward, since there are no analytic expressions for the inverse $F^{-1}(x)$ nor the cdf $F(x)$ itself. All standard approaches like the rejection or the inversion methods would require tedious computations. A much more elegant and efficient solution was proposed by Chambers, Mallows and Stuck (1976). They noticed that a certain integral formula derived by Zolotarev (1964) led to the following algorithm:

- generate a random variable U uniformly distributed on $(-\frac{\pi}{2}, \frac{\pi}{2})$ and an independent exponential random variable W with mean 1;

- for $\alpha \neq 1$ compute:

$$X = (1 + \zeta^2)^{\frac{1}{2\alpha}} \frac{\sin\{\alpha(U + \xi)\}}{\{\cos(U)\}^{1/\alpha}} \left[\frac{\cos\{U - \alpha(U + \xi)\}}{W}\right]^{\frac{1-\alpha}{\alpha}}, \qquad (1.7)$$

- for $\alpha = 1$ compute:

$$X = \frac{1}{\xi}\left\{\left(\frac{\pi}{2} + \beta U\right)\tan U - \beta \log\left(\frac{\frac{\pi}{2}W\cos U}{\frac{\pi}{2} + \beta U}\right)\right\}, \qquad (1.8)$$

where ξ is given by eqn. (1.6). This algorithm yields a random variable $X \sim S_\alpha(1, \beta, 0)$, in representation (1.2). For a detailed proof see Weron (1996).

Given the formulas for simulation of a standard stable random variable, we can easily simulate a stable random variable for all admissible values of the parameters α, σ, β and μ using the following property. If $X \sim S_\alpha(1, \beta, 0)$ then

$$Y = \begin{cases} \sigma X + \mu, & \alpha \neq 1, \\ \sigma X + \frac{2}{\pi}\beta\sigma \log \sigma + \mu, & \alpha = 1, \end{cases} \tag{1.9}$$

is $S_\alpha(\sigma, \beta, \mu)$. It is interesting to note that for $\alpha = 2$ (and $\beta = 0$) the Chambers-Mallows-Stuck (CMS) method reduces to the well known Box-Muller algorithm for generating Gaussian random variables.

Many other approaches have been proposed in the literature, including application of Bergström and LePage series expansions (Janicki and Weron, 1994b). However, the CMS method is regarded as the fastest and the most accurate. Because of its unquestioned superiority and relative simplicity, it is implemented in some statistical computing environments (e.g. the *rstable* function in S-plus/R) even if no other routines related to stable distributions are provided. It is also available in Matlab (function *stablernd.m*) from the SSC repository (ideas.repec.org/c/boc/bocode/m429003.html).

1.2.4 Estimation of parameters

The lack of known closed-form density functions also complicates statistical inference for stable distributions. For instance, maximum likelihood (ML) estimates have to be based on numerical approximations or direct numerical integration of the formulas presented in Section 1.2.2. Consequently, ML estimation is difficult to implement and time consuming for samples encountered in modern finance. However, there are also other numerical methods that have been found useful in practice and are discussed in this section.

Given a sample $x_1, ..., x_n$ of i.i.d. $S_\alpha(\sigma, \beta, \mu)$ observations, in what follows, we provide estimates $\hat{\alpha}$, $\hat{\sigma}$, $\hat{\beta}$ and $\hat{\mu}$ of all four stable law parameters. We start the discussion with the simplest, fastest and ... least accurate quantile methods, then develop the slower, yet much more accurate sample cf methods and, finally, conclude with the slowest but most accurate ML approach. All of the presented methods work quite well assuming that the sample under consideration is indeed stable.

However, testing for stability is not an easy task. Despite some more or less successful attempts (Brcich, Iskander and Zoubir, 2005; Paolella, 2001; Matsui and Takemura, 2008), there are no standard, widely-accepted tests for assessing stability. A possible remedy may be to use bootstrap (or Monte Carlo simulation) techniques, as discussed in Chapter **??** in the context of insurance loss distributions. Other proposed approaches involve using tail exponent estimators for testing if α is in the admissible range (Fan, 2006; Mittnik and Paolella, 1999) or simply 'visual inspection' to see whether the empirical densities resemble those of stable laws (Nolan, 2001; Weron, 2001).

Sample Quantile Methods. The origins of sample quantile methods for stable laws go back to Fama and Roll (1971), who provided very simple estimates for parameters of symmetric ($\beta = 0, \mu = 0$) stable laws with $\alpha > 1$. A decade later McCulloch (1986) generalized their method and provided consistent estimators of all four stable parameters (with the restriction $\alpha \geq 0.6$). After McCulloch define:

$$v_\alpha = \frac{x_{0.95} - x_{0.05}}{x_{0.75} - x_{0.25}} \quad \text{and} \quad v_\beta = \frac{x_{0.95} + x_{0.05} - 2x_{0.50}}{x_{0.95} - x_{0.05}}, \qquad (1.10)$$

where x_f denotes the f-th population quantile, so that $S_\alpha(\sigma, \beta, \mu)(x_f) = f$. Statistics v_α and v_β are functions of α and β only, i.e. they are independent of both σ and μ. This relationship may be inverted and the parameters α and β may be viewed as functions of v_α and v_β:

$$\alpha = \psi_1(v_\alpha, v_\beta) \quad \text{and} \quad \beta = \psi_2(v_\alpha, v_\beta). \qquad (1.11)$$

Substituting v_α and v_β by their sample values and applying linear interpolation between values found in tables given in McCulloch (1986) yields estimators $\hat{\alpha}$ and $\hat{\beta}$. Scale and location parameters, σ and μ, can be estimated in a similar way. However, due to the discontinuity of the cf for $\alpha = 1$ and $\beta \neq 0$ in representation (1.2), this procedure is much more complicated.

In a recent paper, Dominicy and Veredas (2010) further extended the quantile approach by introducing the method of simulated quantiles. It is a promising approach which can also handle multidimensional cases as, for instance, the joint estimation of N univariate stable distributions (but with the constraint of a common tail index).

Sample Characteristic Function Methods. Given an i.i.d. random sample $x_1, ..., x_n$ of size n, define the sample cf by: $\hat{\phi}(t) = \frac{1}{n} \sum_{j=1}^{n} \exp(itx_j)$. Since $|\hat{\phi}(t)|$ is bounded by unity all moments of $\hat{\phi}(t)$ are finite and, for any fixed t, it is the sample average of i.i.d. random variables $\exp(itx_j)$. Hence, by the law of large numbers, $\hat{\phi}(t)$ is a consistent estimator of the cf $\phi(t)$.

To the best of our knowledge, Press (1972) was the first to use the sample cf in the context of statistical inference for stable laws. He proposed a simple estimation method for all four parameters, called the method of moments, based on transformations of the cf. However, the convergence of this method to the population values depends on the choice of four estimation points, whose selection is problematic.

Koutrouvelis (1980) presented a much more accurate regression-type method which starts with an initial estimate of the parameters and proceeds iteratively until some prespecified convergence criterion is satisfied. Each iteration consists of two weighted regression runs. The number of points to be used in these regressions depends on the sample size and starting values of α. Typically no more than two or three iterations are needed. The speed of the convergence, however, depends on the initial estimates and the convergence criterion. The regression method is based on the following observations concerning the cf $\phi(t)$. First, from (1.2) we can easily derive:

$$\log(-\log|\phi(t)|^2) = \log(2\sigma^\alpha) + \alpha \log|t|. \tag{1.12}$$

The real and imaginary parts of $\phi(t)$ are for $\alpha \neq 1$ given by:

$$\Re\{\phi(t)\} = \exp(-|\sigma t|^\alpha) \cos \left[\mu t + |\sigma t|^\alpha \beta \text{sign}(t) \tan \frac{\pi\alpha}{2} \right], \tag{1.13}$$

$$\Im\{\phi(t)\} = \exp(-|\sigma t|^\alpha) \sin \left[\mu t + |\sigma t|^\alpha \beta \text{sign}(t) \tan \frac{\pi\alpha}{2} \right]. \tag{1.14}$$

Apart from considerations of principal values, equations (1.13)-(1.14) lead to:

$$\arctan \left(\frac{\Im\{\phi(t)\}}{\Re\{\phi(t)\}} \right) = \mu t + \beta\sigma^\alpha \tan \frac{\pi\alpha}{2} \text{sign}(t)|t|^\alpha. \tag{1.15}$$

Equation (1.12) depends only on α and σ and suggests that we can estimate these two parameters by regressing $y = \log(-\log|\phi_n(t)|^2)$ on $w = \log|t|$ in the model: $y_k = m + \alpha w_k + \epsilon_k$, where t_k is an appropriate set of real numbers, $m = \log(2\sigma^\alpha)$, and ϵ_k denotes an error term. Koutrouvelis (1980) proposed to use $t_k = \frac{\pi k}{25}, k = 1, 2, ..., K$; with K ranging between 9 and 134 for different values of α and sample sizes.

Once $\hat{\alpha}$ and $\hat{\sigma}$ have been obtained and α and σ have been fixed at these values, estimates of β and μ can be obtained using (1.15). Next, the regressions are repeated with $\hat{\alpha}$, $\hat{\sigma}$, $\hat{\beta}$ and $\hat{\mu}$ as the initial parameters. The iterations continue until a prespecified convergence criterion is satisfied. Koutrouvelis proposed to use Fama and Roll's (1971) formula and the 25% truncated mean for initial estimates of σ and μ, respectively.

Kogon and Williams (1998) eliminated this iteration procedure and simplified the regression method. For initial estimation they applied McCulloch's method, worked with the continuous representation (1.3) of the cf instead of the classical one (1.2) and used a fixed set of only 10 equally spaced frequency points t_k. In terms of computational speed their method compares favorably to the original regression method. It is over five times faster than the procedure of Koutrouvelis, but still about three times slower than the quantile method of McCulloch (Weron, 2011). It has a significantly better performance near $\alpha = 1$ and $\beta \neq 0$ due to the elimination of discontinuity of the cf. However, it returns slightly worse results for other values of α. Matlab implementations of McCulloch's quantile technique (*stabcull.m*) and the regression approach of Koutrouvelis (*stabreg.m*) are distributed with the MFE Toolbox accompanying the monograph of Weron (2006) and can be downloaded from www.ioz.pwr.wroc.pl/pracownicy/weron/MFE.htm.

Maximum Likelihood Method. For a vector of observations $x = (x_1, ..., x_n)$, the maximum likelihood (ML) estimate of the parameter vector $\theta = (\alpha, \sigma, \beta, \mu)$ is obtained by maximizing the log-likelihood function:

$$L_\theta(x) = \sum_{i=1}^{n} \log \tilde{f}(x_i; \theta), \tag{1.16}$$

where $\tilde{f}(\cdot; \theta)$ is the stable pdf. The tilde reflects the fact that, in general, we do not know the explicit form of the stable density and have to approximate it numerically. The ML methods proposed in the literature differ in the choice of the approximating algorithm. However, all of them have an appealing common feature – under certain regularity conditions the ML estimator is asymptotically normal with the variance specified by the Fischer information matrix (DuMouchel, 1973). The latter can be approximated either by using the Hessian matrix arising in maximization or, as in Nolan (2001), by numerical integration.

Because of computational complexity there are only a few documented attempts of estimating stable law parameters via maximum likelihood worth mentioning. DuMouchel (1971) developed an approximate ML method, which was based on

grouping the data set into bins and using a combination of means to compute the density (FFT for the central values of x and series expansions for the tails) to compute an approximate log-likelihood function. This function was then numerically maximized.

Much better, in terms of accuracy and computational time, are more recent ML estimation techniques. Mittnik et al. (1999) utilized the FFT approach for approximating the stable density function, whereas Nolan (2001) used the direct integration method. Both approaches are comparable in terms of efficiency. The differences in performance are the result of different approximation algorithms, see Section 1.2.2. Matsui and Takemura (2006) further improved Nolan's method for the boundary cases, i.e. in the tail and mode of the densities and in the neighborhood of the Cauchy and the Gaussian distributions, but only in the symmetric stable case.

As Ojeda (2001) observes, the ML estimates are almost always the most accurate, closely followed by the regression-type estimates and McCulloch's quantile method. However, ML estimation techniques are certainly the slowest of all the discussed methods. For instance, ML estimation for a sample of 2000 observations using a gradient search routine which utilizes the direct integration method is over 11 thousand (!) times slower than the Kogon-Williams algorithm (calculations performed on a PC running STABLE ver. 3.13; see Section 1.2.2 where the program was briefly described). Clearly, the higher accuracy does not justify the application of ML estimation in many real life problems, especially when calculations are to be performed on-line. For this reason the program STABLE offers an alternative – a fast quasi ML technique. It quickly approximates stable densities using a 3-dimensional spline interpolation based on pre-computed values of the standardized stable density on a grid of (x, α, β) values. At the cost of a large array of coefficients, the interpolation is highly accurate over most values of the parameter space and relatively fast – only ca. 13 times slower than the Kogon-Williams algorithm.

Alternative Methods. Besides the popular methods discussed so far other estimation algorithms have been proposed in the literature. A Bayesian Markov chain Monte Carlo (MCMC) approach was initiated by Buckle (1995). It was later modified by Lombardi (2007) who used an approximated version of the likelihood, instead of the twice slower Gibbs sampler, and by Peters, Sisson and Fan (2009) who proposed likelihood-free Bayesian inference for stable models.

In a recent paper Garcia, Renault and Veredas (2010) estimate the stable law parameters with (constrained) indirect inference, a method particularly suited to situations where the model of interest is difficult to estimate but relatively

easy to simulate. They use the skewed-t distribution as an auxiliary model, since it has the same number of parameters as the stable with each parameter playing a similar role.

1.3 Truncated and tempered stable distributions

Mandelbrot's (1963) seminal work on applying stable distributions in finance gained support in the first few years after its publication, but subsequent works have questioned the stable distribution hypothesis, in particular, the stability under summation (for a review see Rachev and Mittnik, 2000). Over the next few years, the stable law temporarily lost favor and alternative processes were suggested as mechanisms generating stock returns.

In the mid 1990s the stable distribution hypothesis has made a dramatic comeback, at first in the econophysics literature. Several authors have found a very good agreement of high-frequency returns with a stable distribution up to six standard deviations away from the mean (Cont, Potters and Bouchaud, 1997). For more extreme observations, however, the distribution they found fell off approximately exponentially. To cope with such observations the so called truncated Lévy distributions (TLD) were introduced by Mantegna and Stanley (1994). The original definition postulated a sharp truncation of the stable pdf at some arbitrary point. Later, however, exponential smoothing was proposed by Koponen (1995) leading to the following characteristic function:

$$\log \phi(t) = -\frac{\sigma^\alpha}{\cos \frac{\pi\alpha}{2}} \left[(t^2 + \lambda^2)^{\alpha/2} \cos \left\{ \alpha \arctan \frac{|t|}{\lambda} \right\} - \lambda^\alpha \right], \qquad (1.17)$$

where $\alpha \neq 1$ is the tail exponent, σ is the scale parameter and λ is the truncation coefficient (for simplicity β and μ are set to zero here). Clearly the symmetric (exponentially smoothed) TLD reduces to the symmetric stable distribution ($\beta = \mu = 0$) when $\lambda = 0$. For small and intermediate returns the TLD behaves like a stable distribution, but for extreme returns the truncation causes the distribution to converge to the Gaussian and, hence, all moments are finite. In particular, the variance and kurtosis are given by:

$$\mathrm{Var}(X) = \frac{\alpha(1-\alpha)}{\cos(\frac{\pi\alpha}{2})} \sigma^\alpha \lambda^{\alpha-2}, \qquad (1.18)$$

$$k(X) = \frac{\cos(\frac{\pi\alpha}{2})(\alpha-2)(\alpha-3)}{\alpha(1-\alpha)\sigma^\alpha \lambda^\alpha}. \qquad (1.19)$$
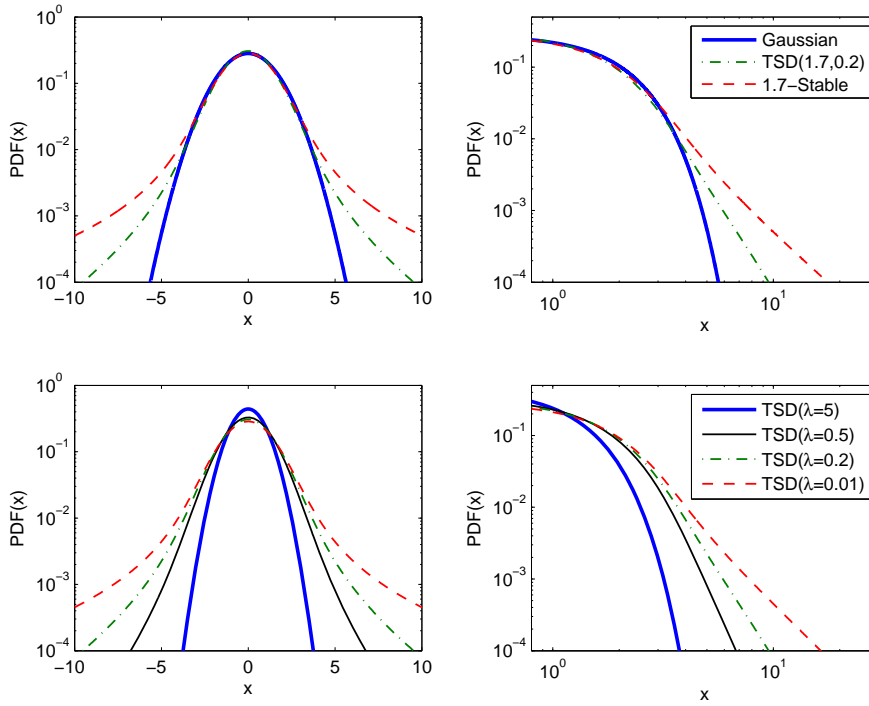
Figure 1.3: *Top panels*: Semilog and loglog plots of symmetric 1.7-stable, symmetric tempered stable (TSD) with $\alpha = 1.7$ and $\lambda = 0.2$, and Gaussian pdfs. *Bottom panels*: Semilog and loglog plots of symmetric TSD pdfs with $\alpha = 1.7$ and four truncation coefficients: $\lambda = 5, 0.5, 0.2, 0.01$. Note, that for large λ's the distribution approaches the Gaussian (though with a different scale) and for small λ's the stable law with the same shape parameter α.

STFstab03

The convergence to the Gaussian and stable laws can be seen in Figure 1.3, where we compare stable and exponentially smoothed TLDs (or TSDs, see below) for a typically reported for financial data value of the tail exponent ($\alpha = 1.7$). Thus the observation that the asset returns distribution is a TLD explains both the short-term stable behavior and the long run convergence to the normal distribution (for interesting insights on the CLT-type behavior of the TLD see a recent paper of Grabchak and Samorodnitsky, 2010).

The (exponentially smoothed) TLD was not recognized in finance until the introduction of the KoBoL (Boyarchenko and Levendorskii, 2000) and CGMY models (Carr et al., 2002). Around this time Rosinski coined the term under which the exponentially smoothed TLD is known today in the mathematics literature – tempered stable distribution (TSD; see Rosinski, 2007).

Despite the interesting statistical properties, the TSDs (TLDs) have not been applied extensively to date. The most probable reason for this being the complicated definition of the TSD. Like for stable distributions, only the characteristic function is known. No closed form formulas exist for the density or the distribution functions. No integral formulas, like Zolotarev's (1986) for the stable laws (see Section 1.2.2), have been discovered to date. Hence, statistical inference is, in general, limited to ML utilizing the FFT technique for approximating the pdf (Bianchi et al., 2010; Grabchak, 2008). Moreover, compared to the stable distribution, the TSD introduces one more parameter (the truncation λ) making the estimation procedure even more complicated. Other parameter fitting techniques proposed so far comprise a combination of *ad hoc* approaches and moment matching (Boyarchenko and Levendorskii, 2000; Matacz, 2000). Apart from a few special cases, also the simulation of TSD variables is cumbersome and numerically demanding (Bianchi et al., 2010; Kawai and Masuda, 2010; Poirot and Tankov, 2006).

1.4 Generalized hyperbolic distributions

1.4.1 Definitions and basic properties

The hyperbolic distribution saw its appearance in finance in the mid-1990s, when a number of authors reported that it provides a very good model for the empirical distributions of daily stock returns from a number of leading German enterprises (Eberlein and Keller, 1995; Küchler et al., 1999). Since then it has become a popular tool in stock price modeling and market risk measurement (Bibby and Sørensen, 2003; Chen, Härdle and Jeong, 2008; McNeil, Rüdiger and Embrechts, 2005). However, the origins of the hyperbolic law date back to the 1940s and the empirical findings in geophysics. A formal mathematical description was developed years later by Barndorff-Nielsen (1977).

The hyperbolic law provides the possibility of modeling heavier tails than the Gaussian, since its log-density forms a hyperbola while that of the Gaussian is a parabola (see Figure 1.4), but lighter than the stable. As we will see later in this
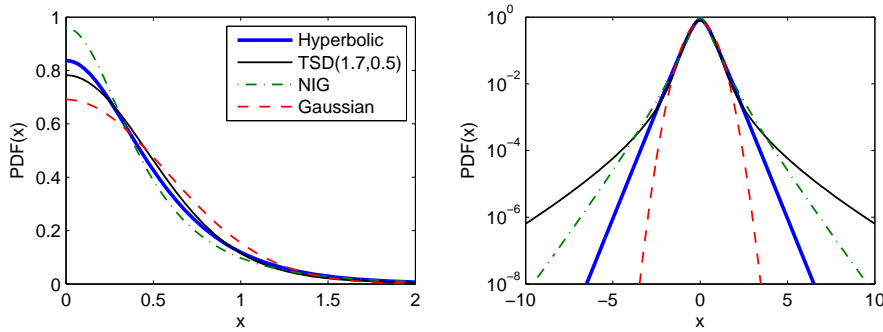
Figure 1.4: Densities and log-densities of symmetric hyperbolic, TSD, NIG, and Gaussian distributions having the same variance, see eqns. (1.18) and (1.33). The name of the hyperbolic distribution is derived from the fact that its log-density forms a hyperbola, which is clearly visible in the right panel.

STFstab04

Section, the hyperbolic law is a member of a larger, versatile class of generalized hyperbolic (GH) distributions, which also includes the normal-inverse Gaussian (NIG) and variance-gamma (VG) distributions as special cases. For a concise review of special and limiting cases of the GH distribution see Chapter 9 in Paolella (2007).

The Hyperbolic Distribution. The hyperbolic distribution is defined as a normal variance-mean mixture where the mixing distribution is the generalized inverse Gaussian (GIG) law with parameter $\lambda = 1$, i.e. it is conditionally Gaussian (Barndorff-Nielsen, 1977; Barndorff-Nielsen and Blaesild, 1981). More precisely, a random variable Z has the hyperbolic distribution if:

$$(Z|Y) \sim \mathrm{N}\left(\mu + \beta Y, Y\right), \tag{1.20}$$

where Y is a generalized inverse Gaussian $\mathrm{GIG}(\lambda = 1, \chi, \psi)$ random variable and $\mathrm{N}(m, s^2)$ denotes the Gaussian distribution with mean m and variance s^2. The GIG law is a positive domain distribution with the pdf given by:

$$f_{\mathrm{GIG}}(x) = \frac{(\psi/\chi)^{\lambda/2}}{2\mathrm{K}_\lambda(\sqrt{\chi\psi})} x^{\lambda-1} e^{-\frac{1}{2}(\chi x^{-1} + \psi x)}, \quad x > 0, \tag{1.21}$$

where the three parameters take values in one of the ranges: (i) $\chi > 0, \psi \geq 0$ if $\lambda < 0$, (ii) $\chi > 0, \psi > 0$ if $\lambda = 0$ or (iii) $\chi \geq 0, \psi = 0$ if $\lambda > 0$. The generalized inverse Gaussian law has a number of interesting properties that we will use later in this section. The distribution of the inverse of a GIG variable is again GIG but with a different λ, namely if:

$$Y \sim \text{GIG}(\lambda, \chi, \psi) \quad \text{then} \quad Y^{-1} \sim \text{GIG}(-\lambda, \chi, \psi). \tag{1.22}$$

A GIG variable can be also reparameterized by setting $a = \sqrt{\chi/\psi}$ and $b = \sqrt{\chi\psi}$, and defining $Y = a\tilde{Y}$, where:

$$\tilde{Y} \sim \text{GIG}(\lambda, b, b). \tag{1.23}$$

The normalizing constant $K_\lambda(t)$ in formula (1.21) is the modified Bessel function of the third kind with index λ, also known as the MacDonald function. It is defined as:

$$K_\lambda(t) = \frac{1}{2} \int_0^\infty x^{\lambda-1} e^{-\frac{1}{2}t(x+x^{-1})} dx, \qquad t > 0. \tag{1.24}$$

In the context of hyperbolic distributions, the Bessel functions are thoroughly discussed in Barndorff-Nielsen and Blaesild (1981). Here we recall only two properties that will be used later. Namely, (i) $K_\lambda(t)$ is symmetric with respect to λ, i.e. $K_\lambda(t) = K_{-\lambda}(t)$, and (ii) for $\lambda = \pm\frac{1}{2}$ it can be written in a simpler form:

$$K_{\pm\frac{1}{2}}(t) = \sqrt{\frac{\pi}{2}} t^{-\frac{1}{2}} e^{-t}. \tag{1.25}$$

Relation (1.20) implies that a hyperbolic random variable $Z \sim \text{H}(\psi, \beta, \chi, \mu)$ can be represented in the form: $Z \sim \mu + \beta Y + \sqrt{Y} N(0,1)$, with the cf:

$$\phi_Z(u) = e^{iu\mu} \int_0^\infty e^{i\beta zu - \frac{1}{2}zu^2} dF_Y(z). \tag{1.26}$$

Here $F_Y(z)$ denotes the distribution function of a GIG random variable Y with parameter $\lambda = 1$, see eqn. (1.21). Hence, the hyperbolic pdf is given by:

$$f_{\text{H}}(x; \psi, \beta, \chi, \mu) = \frac{\sqrt{\psi/\chi}}{2\sqrt{\psi + \beta^2} K_1(\sqrt{\psi\chi})} e^{-\sqrt{\{\psi+\beta^2\}\{\chi+(x-\mu)^2\}}+\beta(x-\mu)}, \tag{1.27}$$

or in an alternative parameterization (with $\delta = \sqrt{\chi}$ and $\alpha = \sqrt{\psi + \beta^2}$) by:

$$f_{\text{H}}(x; \alpha, \beta, \delta, \mu) = \frac{\sqrt{\alpha^2 - \beta^2}}{2\alpha\delta K_1(\delta\sqrt{\alpha^2 - \beta^2})} e^{-\alpha\sqrt{\delta^2+(x-\mu)^2}+\beta(x-\mu)}. \tag{1.28}$$

The latter is more popular and has the advantage of $\delta > 0$ being the traditional scale parameter. Out of the remaining three parameters, α and β determine the shape, with α being responsible for the steepness and $0 \leq |\beta| < \alpha$ for the skewness, and $\mu \in R$ is the location parameter. Finally, note that if we only have an efficient algorithm to compute K_1, the calculation of the pdf is straightforward. However, the cdf has to be numerically integrated from (1.27) or (1.28).

The General Class. The generalized hyperbolic (GH) law can be represented as a normal variance-mean mixture where the mixing distribution is the generalized inverse Gaussian law with any $\lambda \in \mathbb{R}$. Hence, the GH distribution is described by five parameters $\theta = (\lambda, \alpha, \beta, \delta, \mu)$, using parameterization (1.28), and its pdf is given by:

$$f_{\mathrm{GH}}(x; \theta) = \kappa \left\{ \delta^2 + (x - \mu)^2 \right\}^{\frac{1}{2}(\lambda - \frac{1}{2})} \mathrm{K}_{\lambda - \frac{1}{2}} \left(\alpha \sqrt{\delta^2 + (x - \mu)^2} \right) e^{\beta(x - \mu)}, \tag{1.29}$$

where:

$$\kappa = \frac{(\alpha^2 - \beta^2)^{\frac{\lambda}{2}}}{\sqrt{2\pi} \alpha^{\lambda - \frac{1}{2}} \delta^\lambda \mathrm{K}_\lambda(\delta \sqrt{\alpha^2 - \beta^2})}. \tag{1.30}$$

The tail behavior of the GH density is 'semi-heavy', i.e. the tails are lighter than those of non-Gaussian stable laws and TSDs with a relatively small truncation parameter (see Figure 1.4), but much heavier than Gaussian. Formally, the following asymptotic relation is satisfied (Barndorff-Nielsen and Blaesild, 1981):

$$f_{\mathrm{GH}}(x) \approx |x|^{\lambda - 1} e^{(\mp \alpha + \beta)x} \quad \text{for} \quad x \to \pm\infty, \tag{1.31}$$

which can be interpreted as exponential 'tempering' of the power-law tails (compare with the TSD described in Section 1.3). Consequently, all moments of the GH law exist. In particular, the mean and variance are given by:

$$\mathbb{E}(X) = \mu + \frac{\beta \delta^2}{\zeta} \frac{\mathrm{K}_{\lambda+1}(\zeta)}{\mathrm{K}_\lambda(\zeta)}, \tag{1.32}$$

$$\mathrm{Var}(X) = \delta^2 \left[\frac{\mathrm{K}_{\lambda+1}(\zeta)}{\zeta \mathrm{K}_\lambda(\zeta)} + \frac{\beta^2 \delta^2}{\zeta^2} \left\{ \frac{\mathrm{K}_{\lambda+2}(\zeta)}{\mathrm{K}_\lambda(\zeta)} - \left(\frac{\mathrm{K}_{\lambda+1}(\zeta)}{\zeta \mathrm{K}_\lambda(\zeta)} \right)^2 \right\} \right]. \tag{1.33}$$

The Normal-Inverse Gaussian Distribution. The normal-inverse Gaussian (NIG) laws were introduced by Barndorff-Nielsen (1995) as a subclass of the generalized hyperbolic laws obtained for $\lambda = -\frac{1}{2}$. The density of the NIG

distribution is given by:

$$f_{\mathrm{NIG}}(x) = \frac{\alpha\delta}{\pi} e^{\delta\sqrt{\alpha^2-\beta^2}+\beta(x-\mu)} \frac{\mathrm{K}_1(\alpha\sqrt{\delta^2+(x-\mu)^2})}{\sqrt{\delta^2+(x-\mu)^2}}. \qquad (1.34)$$

Like for the hyperbolic law the calculation of the pdf is straightforward, but the cdf has to be numerically integrated from eqn. (1.34).

At the expense of four parameters, the NIG distribution is able to model asymmetric distributions with 'semi-heavy' tails. However, if we let $\alpha \to 0$ the NIG distribution converges to the Cauchy distribution (with location parameter μ and scale parameter δ), which exhibits extremely heavy tails. Obviously, the NIG distribution may not be adequate to deal with cases of extremely heavy tails such as those of Pareto or non-Gaussian stable laws. However, empirical experience suggests excellent fits of the NIG law to financial data (Karlis, 2002; Karlis and Lillestöl, 2004; Venter and de Jongh, 2002).

Moreover, the class of normal-inverse Gaussian distributions possesses an appealing feature that the class of hyperbolic laws does not have. Namely, it is closed under convolution, i.e. a sum of two independent NIG random variables is again NIG (Barndorff-Nielsen, 1995). In particular, if X_1 and X_2 are independent NIG random variables with common parameters α and β but having different scale and location parameters $\delta_{1,2}$ and $\mu_{1,2}$, respectively, then $X = X_1 + X_2$ is NIG$(\alpha, \beta, \delta_1 + \delta_1, \mu_1 + \mu_2)$. Only two subclasses of the GH 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 for $\lambda > 0$ and $\alpha > |\beta|$. The VG distributions (with $\beta = 0$) were introduced to finance by Madan and Seneta (1990), long before the popularity of GH and NIG laws.

1.4.2 Simulation of generalized hyperbolic variables

The most natural way of simulating GH variables is derived from the property that they can be represented as normal variance-mean mixtures. Since the mixing distribution is the GIG law, the resulting algorithm reads as follows:

1. simulate a random variable $Y \sim \mathrm{GIG}(\lambda, \chi, \psi) = \mathrm{GIG}(\lambda, \delta^2, \alpha^2 - \beta^2)$;

2. simulate a standard normal random variable N;

3. return $X = \mu + \beta Y + \sqrt{Y}N$.

The algorithm is fast and efficient if we have a handy way of simulating GIG variates. For $\lambda = -\frac{1}{2}$, i.e. when sampling from the so-called inverse Gaussian (IG) distribution, there exists an efficient procedure that utilizes a transformation yielding two roots. It starts with the observation that if we let $\vartheta = \sqrt{\chi/\psi}$ then the IG(χ, ψ) density (= GIG$(-\frac{1}{2}, \chi, \psi)$; see eqn. (1.21)) of Y can be written as:

$$f_{\mathrm{IG}}(x) = \sqrt{\frac{\chi}{2\pi x^3}} \exp\left\{\frac{-\chi(x-\vartheta)^2}{2x\vartheta^2}\right\}. \tag{1.35}$$

Now, following Shuster (1968) we may write:

$$V = \frac{\chi(Y-\vartheta)^2}{Y\vartheta^2} \sim \chi^2_{(1)}, \tag{1.36}$$

i.e. V is distributed as a chi-square random variable with one degree of freedom. As such it can be simply generated by taking a square of a standard normal random number. Unfortunately, the value of Y is not uniquely determined by eqn. (1.36). Solving this equation for Y yields two roots:

$$y_1 = \vartheta + \frac{\vartheta}{2\chi}\left(\vartheta V - \sqrt{4\vartheta\chi V + \vartheta^2 V^2}\right) \quad \text{and} \quad y_2 = \frac{\vartheta^2}{y_1}.$$

The difficulty in generating observations with the desired distribution now lies in choosing between the two roots. Michael, Schucany and Haas (1976) showed that Y can be simulated by choosing y_1 with probability $\vartheta/(\vartheta + y_1)$. So for each random observation V from a $\chi^2_{(1)}$ distribution the smaller root y_1 has to be calculated. Then an auxiliary Bernoulli trial is performed with probability $p = \vartheta/(\vartheta + y_1)$. If the trial results in a 'success', y_1 is chosen; otherwise, the larger root y_2 is selected. This routine is, for instance, implemented in the *rnig* function of the Rmetrics collection of software packages for S-plus/R (see also Section 1.2.2 where Rmetrics was briefly described).

In the general case, the GIG distribution – as well as the (generalized) hyperbolic law – can be simulated via the rejection algorithm. An adaptive version of this technique is used to obtain hyperbolic random numbers in the *rhyp* function of Rmetrics. Rejection is also implemented in the HyperbolicDist package for S-plus/R developed by David Scott, see the R-project home page cran.r-project.org. The package utilizes a version of the algorithm proposed by Atkinson (1982), i.e. rejection coupled either with a two ('GIG algorithm' for any admissible value of λ) or a three ('GIGLT1 algorithm' for $0 \leq \lambda < 1$) part envelope (or majorizing function). However, finding the appropriate parameter values for these envelopes requires optimization and makes the technique burdensome.

This difficulty led to a search for a short algorithm which would give comparable efficiencies but without the drawback of extensive numerical optimizations. A solution, based on the 'ratio-of-uniforms' method, was provided by Dagpunar (1989). First, recalling properties (1.22) and (1.23), observe that we only need to find a method to simulate $\tilde{Y} \sim \text{GIG}(\lambda, b, b)$ variables and only for non-negative λ's. Next, define the relocated variable $\tilde{Y}_m = \tilde{Y} - m$, where the shift $m = \frac{1}{b}(\lambda - 1 + \sqrt{(\lambda - 1)^2 + b^2})$ is the mode of the density of $\tilde{Y}$. Then $\tilde{Y}$ can be generated by setting $\tilde{Y}_m = V/U$, where the pair (U, V) is uniformly distributed over the region $\{(u, v) : 0 \le u \le \sqrt{h(v/u)}\}$ with:

$$h(t) = (t + m)^{\lambda - 1} \exp\left(-\frac{b}{2}\frac{t + m + 1}{t + m}\right), \quad \text{for} \quad t \ge -m.$$

Since this region is irregularly shaped, it is more convenient to generate the pair (U, V) uniformly over a minimal enclosing rectangle $\{(u, v) : 0 \le u \le u_+, v_- \le v \le v_+\}$. Finally, the variate (V/U) is accepted if $U^2 \le h(V/U)$. The efficiency of the algorithm depends on the method of deriving and the actual choice of u_+ and $v_\pm$. Further, for $\lambda \le 1$ and $b \le 1$ there is no need for the shift at mode m. Such a version of the algorithm is implemented in UNU.RAN, a library of C functions for non-uniform random number generation developed at the Vienna University of Economics, see statistik.wu-wien.ac.at/unuran.

1.4.3 Estimation of parameters

Maximum Likelihood Method. The parameter estimation of GH distributions can be performed by the ML method, since there exist closed-form formulas (although, involving special functions) for the densities of these laws. The computational burden is not as heavy as for stable laws, but it still is considerable. In general, the ML estimation algorithm is as follows. For a vector of observations $x = (x_1, ..., x_n)$, the ML estimate of the parameter vector $\theta = (\lambda, \alpha, \beta, \delta, \mu)$ is obtained by maximizing the log-likelihood function:

$$
\begin{aligned}
L(x; \theta) \quad = \quad & \log \kappa + \frac{\lambda - \frac{1}{2}}{2} \sum_{i=1}^{n} \log(\delta^2 + (x_i - \mu)^2) + \\
& + \sum_{i=1}^{n} \log \mathrm{K}_{\lambda - \frac{1}{2}}(\alpha\sqrt{\delta^2 + (x_i - \mu)^2}) + \sum_{i=1}^{n} \beta(x_i - \mu), \ (1.37)
\end{aligned}
$$

where κ is defined by (1.30). Obviously, for hyperbolic ($\lambda = 1$) distributions the algorithm uses simpler expressions of the log-likelihood function due to relation (1.25).

The routines proposed in the literature differ in the choice of the optimization scheme. The first software product that allowed statistical inference with hyperbolic distributions, the HYP program, used a gradient search technique (Blaesild and Sorensen, 1992). In a large simulation study Prause (1999) utilized the bracketing method. Matlab functions *hypest.m* and *nigest.m* distributed with the MFE Toolbox (Weron, 2006) use the downhill simplex method, with slight modifications due to parameter restrictions.

The main factor for the speed of the estimation is the number of modified Bessel functions to compute. Note, that for $\lambda = 1$ (i.e. the hyperbolic distribution) this function appears only in the constant κ. For a data set with n independent observations we need to evaluate n and $n+1$ Bessel functions for NIG and GH distributions, respectively, whereas only one for the hyperbolic. This leads to a considerable reduction in the time necessary to calculate the likelihood function in the hyperbolic case. Prause (1999) reported a reduction of ca. 33%, however, the efficiency results are highly sample and implementation dependent.

We also have to say that the optimization is challenging. Some of the parameters are hard to separate since a flat-tailed GH distribution with a large scale parameter is hard to distinguish from a fat-tailed distribution with a small scale parameter, see Barndorff-Nielsen and Blaesild (1981) who observed such a behavior already for the hyperbolic law. The likelihood function with respect to these parameters then becomes very flat, and may have local minima. In the case of NIG distributions Venter and de Jongh (2002) proposed simple estimates of α and β that can be used as staring values for the ML scheme. Starting from relation (1.31) for the tails of the NIG density (i.e. with $\lambda = -1/2$) they derived the following approximation:

$$
\alpha - \beta \quad \sim \quad \frac{1}{2} \frac{x_{1-f} + \mathbb{E}(X|X > x_{1-f})}{\mathbb{E}(X^2|X > x_{1-f}) - x_{1-f}\mathbb{E}(X|X > x_{1-f})},
$$

$$
\alpha + \beta \quad \sim \quad -\frac{1}{2} \frac{x_f + \mathbb{E}(X|X < x_f)}{\mathbb{E}(X^2|X < x_f) - x_f\mathbb{E}(X|X < x_f)},
$$

where x_f is the f-th population quantile, see Section 1.2.4. After the choice of a suitable value for f, Venter and de Jongh used $f = 5\%$, the 'tail estimates' of α and β are obtained by replacing the quantiles and expectations by their sample values in the above relations.

Another method of providing the starting values for the ML scheme was suggested by Prause (1999). He estimated the parameters of a symmetric ($\beta = \mu = 0$) GH law with a reasonable kurtosis (i.e. with $\delta\alpha \approx 1.04$) that had the variance equal to that of the empirical distribution.

Other Methods. Besides the ML approach other estimation methods have been proposed in the literature. Prause (1999) tested different estimation techniques by replacing the log-likelihood function with other score functions, like the Anderson-Darling and Kolmogorov statistics or L^p-norms. But the results were disappointing. Karlis and Lillestöl (2004) made use of the MCMC technique, however, again the results obtained were not impressive. Karlis (2002) described an Expectation-Maximization (EM) type algorithm for ML estimation of the NIG distribution. The algorithm can be programmed in any statistical package supporting Bessel functions and it has all the properties of the standard EM algorithm, like sure, but slow, convergence, parameters in the admissible range, etc. Recently Fragiadakis, Karlis and Meintanis (2009) used this approach to construct goodness-of-fit tests for symmetric NIG distributions. The tests are based on a weighted integral incorporating the empirical cf of suitably standardized data. The EM scheme can be also generalized to multivariate GH distributions (but with fixed λ, see Protassov, 2004).

1.5 Empirical evidence

The empirical facts presented in Section 1.1 show that we should use heavy tailed alternatives to the Gaussian law in order to obtain acceptable estimates of market losses. In this section we apply the techniques discussed so far to two samples of financial data: the Dow Jones Industrial Average (DJIA) index and the Polish WIG20 index. Both are blue chip stock market indexes. DJIA is composed of 30 major U.S. companies that trade on NYSE and NASDAQ. WIG20 consists of 20 major Polish companies listed on the Warsaw Stock Exchange. We use daily closing index values from the period January 3, 2000 – December 31, 2009. Eliminating missing values (mostly U.S. and Polish holidays) we end up with 2494 (log-)returns for each index, see the top left panels in Figures 1.5 and 1.6.

Like most financial time series, also these index returns contain volatility clusters which imply that the probability of a specific incurred loss is not the same on each day. During days of higher volatility we should expect larger than usual losses and during calmer days – smaller than usual. To remove volatility clusters it is necessary to model the process that generates them. Following Barone-Adesi, Giannopoulos, and Vosper (1999) and Kuester, Mittnik, and Paolella (2006) we eliminate volatility clusters by filtering the returns r_t with

Table 1.1: Gaussian, hyperbolic, NIG, and stable fits to 2516 standardized (rescaled by the sample standard deviation) and σ_t-filtered returns of the DJIA from the period January 3, 2000 – December 31, 2009, see also Figure 1.5. The values of the Kolmogorov (K) and Anderson-Darling (AD) goodness-of-fit statistics suggest the hyperbolic and NIG distributions as the best models for filtered and standardized returns, respectively. The symbols ***, **, and * denote significance of the goodness-of-fit tests at the 10%, 5%, and 1% level, respectively.

Distribution	Parameters				Statistics	
	α	$\sigma(\delta)$	β	μ	K	AD
Returns (standardized)						
Gaussian		1.0000		-0.0026	4.083	38.61
Hyperbolic	1.5003	0.0700	-0.0794	0.0691	0.756**	0.671*
NIG	0.6988	0.6829	-0.0586	0.0548	**0.651*****	**0.414****
Stable	1.6150	0.4982	-0.1624	-0.0247	1.171***	1.792***
Filtered returns						
Gaussian		1.0000		-0.0021	1.857	4.328
Hyperbolic	2.1899	1.3793	-0.2852	0.2755	**0.713****	**0.658***
NIG	1.8928	1.8172	-0.2953	0.2849	0.743**	0.711*
Stable	1.9327	0.6669	-0.8026	-0.0101	1.074***	1.504***

a GARCH(1,1) process:

$$r_t = \sigma_t \epsilon_t, \quad \text{with} \quad \sigma_t = c_0 + c_1 r_{t-1}^2 + d_1 \sigma_{t-1}^2, \tag{1.38}$$

where c_0, c_1 and d_1 are constants and $\epsilon_t = r_t/\sigma_t$ are the *filtered returns*, see the top right panels in Figures 1.5 and 1.6. We could also insert a moving average term in the conditional mean to remove any serial dependency if needed.

To find the right model class for each dataset we fit four distributions: Gaussian, hyperbolic, NIG, and stable to standardized (rescaled by the sample standard deviation) and σ_t-filtered returns. The results are presented in Tables 1.1 and 1.2, see also Figures 1.5 and 1.6. We compare the fits using Kolmogorov (K) and Anderson-Darling (AD) test statistics (D'Agostino and Stephens, 1986). The latter may be treated as a weighted Kolmogorov statistics which puts more weight to the differences in the tails of the distributions. Naturally, the lower the values the better the fit. For both datasets, both statistics suggest the

Table 1.2: Gaussian, hyperbolic, NIG and stable fits to 2514 standardized
(rescaled by the sample standard deviation) and σ_t-filtered returns
of the Polish WIG20 index from the period January 3, 2000 – De-
cember 31, 2009, see also Figure 1.6. The values of the Kolmogorov
(K) and Anderson-Darling (AD) goodness-of-fit statistics suggest
the hyperbolic distribution as the best model, with the NIG law fol-
lowing closely by. The symbols ***, **, and * denote significance of
the goodness-of-fit tests at the 10%, 5%, and 1% level, respectively.

Distribution	Parameters				Statistics	
	α	$\sigma(\delta)$	β	μ	K	AD
Returns (standardized)						
Gaussian		1.0000		0.0057	2.337	11.19
Hyperbolic	1.5779	0.4884	-0.0069	0.0126	**0.448*****	**0.247*****
NIG	1.0972	1.1054	-0.0042	0.0100	0.557***	0.360***
Stable	1.7841	0.5977	0.1069	0.0222	0.926***	1.550***
Filtered returns						
Gaussian		1.0000		0.0126	1.598	3.681
Hyperbolic	2.0121	1.2288	0.0161	-0.0036	**0.639*****	**0.379****
NIG	1.6855	1.6984	0.0140	-0.0015	0.669***	0.435**
Stable	1.8891	0.6529	0.1053	0.0200	0.890***	1.062***

hyperbolic distribution as the best model, with the NIG law following closely by.
Although no asymptotic results are known for stable or generalized hyperbolic
laws, approximate critical values for these goodness-of-fit tests can be obtained
via the bootstrap (or simulation) technique, see Stute, Manteiga and Quindimil
(1993), Ross (2002), and Chapter **??** where Monte Carlo simulations are used
for this purpose in the context of insurance loss distributions. Like in the latter,
also here the significance is computed on the basis of 1000 simulated samples.
Clearly, the Gaussian law cannot be accepted for any dataset. On the other
hand, the stable law yields moderately high test values (indicating a poor fit),
but high p-values. This is likely due to the significant dispersion of stable law
parameter estimates for samples of this size. Finally, note that for the DJIA
filtered returns the left tail is much heavier than the right tail, with the latter
being pretty well modeled by the Gaussian law, see the bottom right panel in
Figure 1.5. This is also confirmed by strongly negative skewness parameters
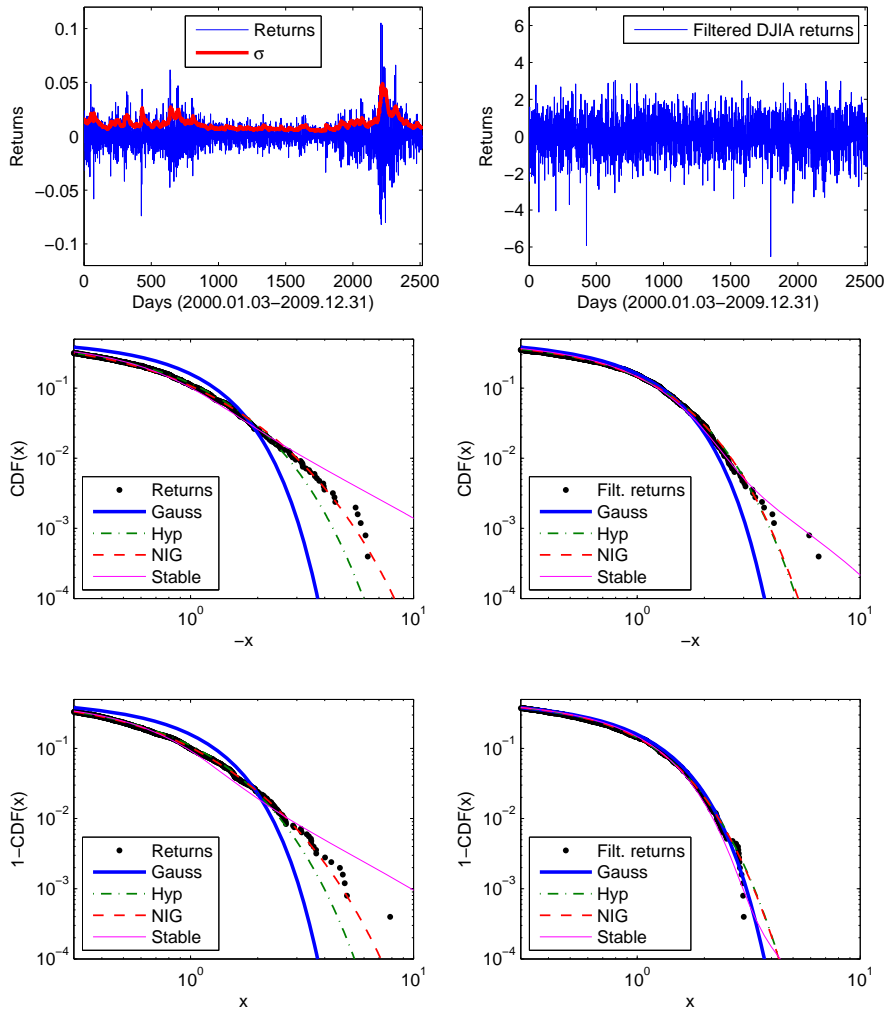(β). In contrast, the WIG20 filtered returns are roughly symmetric.

Figure 1.5: *Top left*: DJIA index (log-)returns and the GARCH(1,1)-based daily volatility estimate σ_t. *Top right*: σ_t-filtered DJIA returns. *Middle*: The left tails of the cdf of (standardized) returns (*left*) or filtered returns (*right*) and of four fitted distributions. *Bottom*: The corresponding right tails.
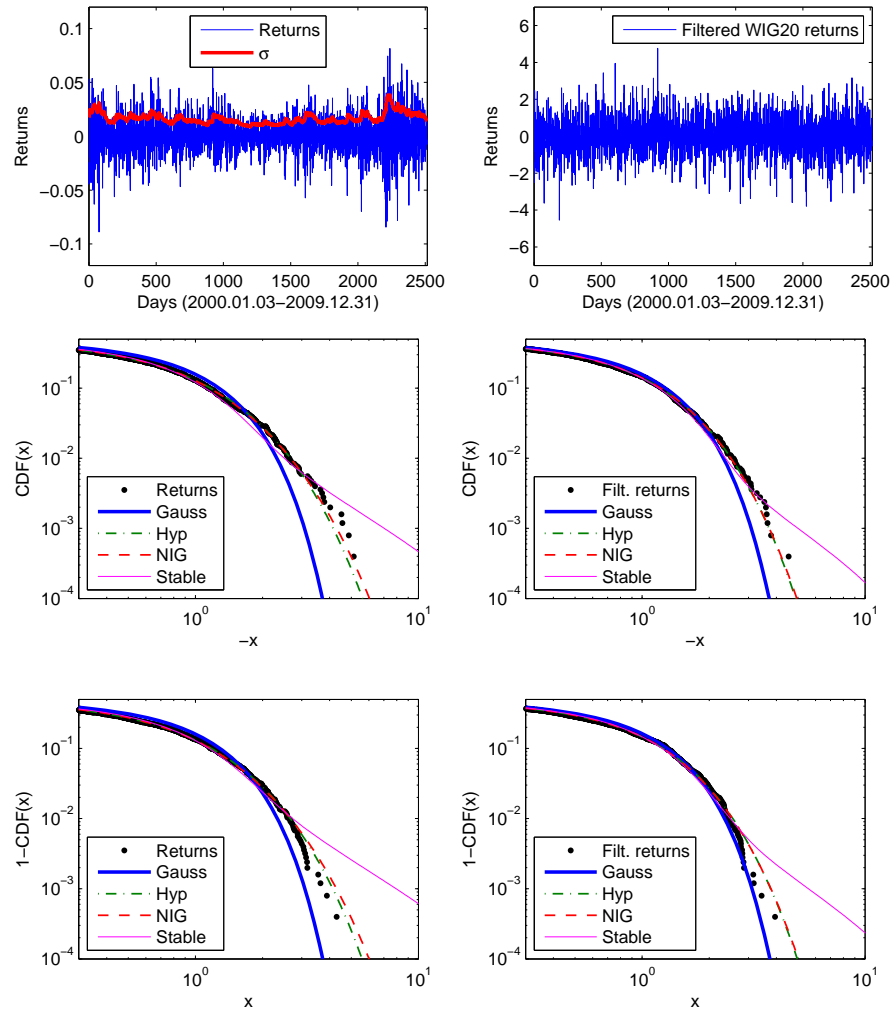
STFstab05

Figure 1.6: *Top left*: WIG20 index (log-)returns and the GARCH(1,1)-based daily volatility estimate σ_t. *Top right*: σ_t-filtered WIG20 returns. *Middle*: The left tails of the cdf of (standardized) returns (*left*) or filtered returns (*right*) and of four fitted distributions. *Bottom*: The corresponding right tails.

STFstab06

Bibliography

Atkinson, A. C. (1982). The simulation of generalized inverse Gaussian and hyperbolic random variables, *SIAM Journal of Scientific & Statistical Computing* **3**: 502–515.

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

Barndorff-Nielsen, O. E. (1995). *Normal\\Inverse Gaussian Processes and the Modelling of Stock Returns*, Research Report 300, Department of Theoretical Statistics, University of Aarhus.

Barndorff-Nielsen, O. E. and Blaesild, P. (1981). Hyperbolic distributions and ramifications: Contributions to theory and applications, *in* C. Taillie, G. Patil, B. Baldessari (eds.) *Statistical Distributions in Scientific Work*, Volume 4, Reidel, Dordrecht, pp. 19–44.

Barone-Adesi, G., Giannopoulos, K., and Vosper, L., (1999). VaR without correlations for portfolios of derivative securities, *Journal of Futures Markets* **19(5)**: 583–602.

Basle Committee on Banking Supervision (1995). An internal model-based approach to market risk capital requirements, http://www.bis.org.

Bianchi, M. L., Rachev, S. T., Kim, Y. S. and Fabozzi, F. J. (2010). Tempered stable distributions and processes in finance: Numerical analysis, *in* M. Corazza, P. Claudio, P. (Eds.) *Mathematical and Statistical Methods for Actuarial Sciences and Finance*, Springer.

Bibby, B. M. and Sørensen, M. (2003). Hyperbolic processes in finance, *in* S. T. Rachev (ed.) *Handbook of Heavy-tailed Distributions in Finance*, North Holland.

Blaesild, P. and Sorensen, M. (1992). *HYP – a Computer Program for Analyzing Data by Means of the Hyperbolic Distribution*, Research Report 248, Department of Theoretical Statistics, Aarhus University.

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

Brcich, R. F., Iskander, D. R. and Zoubir, A. M. (2005). The stability test for symmetric alpha stable distributions, *IEEE Transactions on Signal Processing* **53**: 977–986.

Buckle, D. J. (1995). Bayesian inference for stable distributions, *Journal of the American Statistical Association* **90**: 605–613.

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

Chambers, J. M., Mallows, C. L. and Stuck, B. W. (1976). A Method for Simulating Stable Random Variables, *Journal of the American Statistical Association* **71**: 340–344.

Chen, Y., Härdle, W. and Jeong, S.-O. (2008). Nonparametric risk management with generalized hyperbolic distributions, *Journal of the American Statistical Association* **103**: 910–923.

Cont, R., Potters, M. and Bouchaud, J.-P. (1997). Scaling in stock market data: Stable laws and beyond, *in* B. Dubrulle, F. Graner, D. Sornette (eds.) *Scale Invariance and Beyond*, Proceedings of the CNRS Workshop on Scale Invariance, Springer, Berlin.

D'Agostino, R. B. and Stephens, M. A. (1986). *Goodness-of-Fit Techniques*, Marcel Dekker, New York.

Dagpunar, J. S. (1989). An Easily Implemented Generalized Inverse Gaussian Generator, *Communications in Statistics – Simulations* **18**: 703–710.

Danielsson, J., Hartmann, P. and De Vries, C. G. (1998). The cost of conservatism: Extreme returns, value at risk and the Basle multiplication factor, *Risk* **11**: 101–103.

Dominicy, Y. and Veredas, D. (2010). The method of simulated quantiles, *ECARES working paper*, 2010–008.

DuMouchel, W. H. (1971). *Stable Distributions in Statistical Inference*, Ph.D. Thesis, Department of Statistics, Yale University.

DuMouchel, W. H. (1973). On the Asymptotic Normality of the Maximum–Likelihood Estimate when Sampling from a Stable Distribution, *Annals of Statistics* **1(5)**: 948–957.

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

Fama, E. F. and Roll, R. (1971). Parameter Estimates for Symmetric Stable Distributions, *Journal of the American Statistical Association* **66**: 331–338.

Fan, Z. (2006). Parameter estimation of stable distributions, *Communications in Statistics – Theory and Methods* **35(2)**: 245–255.

Fragiadakis, K., Karlis, D. and Meintanis, S. G. (2009). Tests of fit for normal inverse Gaussian distributions, *Statistical Methodology* **6**: 553–564.

Garcia, R., Renault, E. and Veredas, D. (2010). Estimation of stable distributions by indirect inference, *Journal of Econometrics*, Forthcoming.

Grabchak, M. (2010). Maximum likelihood estimation of parametric tempered stable distributions on the real line with applications to finance, *Ph.D. thesis*, Cornell University.

Grabchak, M. and Samorodnitsky, G. (2010). Do financial returns have finite or infinite variance? A paradox and an explanation, *Quantitative Finance*, DOI: 10.1080/14697680903540381.

Guillaume, D. M., Dacorogna, M. M., Dave, R. R., Müller, U. A., Olsen, R. B. and Pictet, O. V. (1997). From the birds eye to the microscope: A survey of new stylized facts of the intra-daily foreign exchange markets, *Finance & Stochastics* **1**: 95–129.

Janicki, A. and Weron, A. (1994a). Can one see α-stable variables and processes, *Statistical Science* **9**: 109–126.

Janicki, A. and Weron, A. (1994b). *Simulation and Chaotic Behavior of α-Stable Stochastic Processes*, Marcel Dekker.

Karlis, D. (2002). An EM type algorithm for maximum likelihood estimation for the Normal Inverse Gaussian distribution, *Statistics and Probability Letters* **57**: 43–52.

Karlis, D. and Lillestöl, J. (2004). Bayesian estimation of NIG models via Markov chain Monte Carlo methods, *Applied Stochastic Models in Business and Industry* **20(4)**: 323–338.

Kawai, R. and Masuda, H. (2010). On simulation of tempered stable random variates, *Preprint*, Kyushu University.

Kogon, S. M. and Williams, D. B. (1998). Characteristic function based estimation of stable parameters, *in* R. Adler, R. Feldman, M. Taqqu (eds.), *A Practical Guide to Heavy Tails*, Birkhauser, pp. 311–335.

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

Koutrouvelis, I. A. (1980). Regression–Type Estimation of the Parameters of Stable Laws, *Journal of the American Statistical Association* **75**: 918–928.

Kuester, K., Mittnik, S., and Paolella, M.S. (2006). Value-at-Risk prediction: A comparison of alternative strategies, *Journal of Financial Econometrics* **4(1)**: 53–89.

Küchler, U., Neumann, K., Sørensen, M. and Streller, A. (1999). Stock returns and hyperbolic distributions, *Mathematical and Computer Modelling* **29**: 1–15.

Lombardi, M. J. (2007). Bayesian inference for α-stable distributions: A random walk MCMC approach, *Computational Statistics and Data Analysis* **51(5)**: 2688–2700.

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

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

Mantegna, R. N. and Stanley, H. E. (1994). Stochastic processes with ultraslow convergence to a Gaussian: The truncated Lévy flight, *Physical Review Letters* **73**: 2946–2949.

Matacz, A. (2000). Financial Modeling and Option Theory with the Truncated Lévy Process, *International Journal of Theoretical and Applied Finance* **3(1)**: 143–160.

Matsui, M. and Takemura, A. (2006). Some improvements in numerical evaluation of symmetric stable density and its derivatives, *Communications in Statistics – Theory and Methods* **35(1)**: 149–172.

Matsui, M. and Takemura, A. (2008). Goodness-of-fit tests for symmetric stable distributions – empirical characteristic function approach, *TEST* **17(3)**: 546–566.

McCulloch, J. H. (1986). Simple consistent estimators of stable distribution

parameters, *Communications in Statistics – Simulations* **15**: 1109–1136.

McNeil, A. J., Rüdiger, F. and Embrechts, P. (2005). *Quantitative Risk Management*, Princeton University Press, Princeton, NJ.

Michael, J. R., Schucany, W. R. and Haas, R. W. (1976). Generating Random Variates Using Transformations with Multiple Roots, *The American Statistician* **30**: 88–90.

Mittnik, S., Doganoglu, T. and Chenyao, D. (1999). Computing the Probability Density Function of the Stable Paretian Distribution, *Mathematical and Computer Modelling* **29**: 235–240.

Mittnik, S. and Paolella, M. S. (1999). A simple estimator for the characteristic exponent of the stable Paretian distribution, *Mathematical and Computer Modelling* **29**: 161–176.

Mittnik, S., Rachev, S. T., Doganoglu, T. and Chenyao, D. (1999). Maximum Likelihood Estimation of Stable Paretian Models, *Mathematical and Computer Modelling* **29**: 275–293.

Nolan, J. P. (1997). Numerical Calculation of Stable Densities and Distribution Functions, *Communications in Statistics – Stochastic Models* **13**: 759–774.

Nolan, J. P. (2001). Maximum Likelihood Estimation and Diagnostics for Stable Distributions, *in* O. E. Barndorff-Nielsen, T. Mikosch, S. Resnick (eds.), *Lévy Processes*, Brikhäuser, Boston.

Nolan, J. P. (2010). *Stable Distributions – Models for Heavy Tailed Data*, Birkhäuser, Boston. *In progress*, Chapter 1 online at academic2.american.edu/∼jpnolan.

Ojeda, D. (2001). *Comparison of stable estimators*, Ph.D. Thesis, Department of Mathematics and Statistics, American University.

Paolella, M. S. (2001). Testing the stable Paretian assumption, *Mathematical and Computer Modelling* **34**: 1095–1112.

Paolella, M. S. (2007). *Intermediate Probability: A Computational Approach*, Wiley, Chichester.

Peters, G. W., Sisson, S. A. and Fan, Y. (2009). Likelihood-free Bayesian inference for α-stable models, *Preprint*: http://arxiv.org/abs/0912.4729.

Poirot, J. and Tankov, P. (2006). Monte Carlo option pricing for tempered stable (CGMY) processes, *Asia-Pacific Financial Markets* **13(4)**: 327-

344.

Prause, K. (1999). *The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measures*, Ph.D. Thesis, Freiburg University, http://www.freidok.uni-freiburg.de/volltexte/15.

Press, S. J. (1972). Estimation in Univariate and Multivariate Stable Distribution, *Journal of the American Statistical Association* **67**: 842–846.

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

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

Rosinski, J. (2007). Tempering stable processes, *Stochastic Processes and Their Applications* **117(6)**: 677–707.

Ross, S. (2002). *Simulation*, Academic Press, San Diego.

Samorodnitsky, G. and Taqqu, M. S. (1994). *Stable Non–Gaussian Random Processes*, Chapman & Hall.

Shuster, J. (1968). On the Inverse Gaussian Distribution Function, *Journal of the American Statistical Association* **63**: 1514–1516.

Stahl, G. (1997). Three cheers, *Risk* **10**: 67–69.

Stute, W., Manteiga, W. G. and Quindimil, M.P. (1993). Bootstrap Based Goodness-Of-Fit-Tests, *Metrika* **40**: 243–256.

Venter, J. H. and de Jongh, P. J. (2002). Risk estimation using the Normal Inverse Gaussian distribution, *The Journal of Risk* **4**: 1–23.

Weron, R. (1996). On the Chambers-Mallows-Stuck Method for Simulating Skewed Stable Random Variables, *Statistics and Probability Letters* **28**: 165–171. *See also* R. Weron (1996) *Correction to: On the Chambers-Mallows-Stuck Method for Simulating Skewed Stable Random Variables*, Working Paper, Available at MPRA: http://mpra.ub.uni-muenchen.de/20761/.

Weron, R. (2001). Levy–Stable Distributions Revisited: Tail Index > 2 Does Not Exclude the Levy–Stable Regime, *International Journal of Modern Physics C* **12**: 209–223.

Weron, R. (2006). *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*, Wiley, Chichester.

Weron, R. (2011). Computationally Intensive Value at Risk Calculations, *in* J. E. Gentle, W. Härdle, Y. Mori (eds.) *Handbook of Computational Statistics*, 2nd edition, Springer, Berlin, 911–950.

Zolotarev, V. M. (1964). On representation of stable laws by integrals, *Selected Translations in Mathematical Statistics and Probability* **4**: 84–88.

Zolotarev, V. M. (1986). *One–Dimensional Stable Distributions*, American Mathematical Society.

HSC Research Report Series 2010

For a complete list please visit http://ideas.repec.org/s/wuu/wpaper.html

01   *Models for Heavy-tailed Asset Returns* by Szymon Borak, Adam Misiorek and Rafał Weron