

Weron, Rafał

**Working Paper**

## Computationally intensive Value at Risk calculations

Papers, No. 2004,32

**Provided in Cooperation with:**

CASE - Center for Applied Statistics and Economics, Humboldt University Berlin

*Suggested Citation:* Weron, Rafał (2004) : Computationally intensive Value at Risk calculations, Papers, No. 2004,32, Humboldt-Universität zu Berlin, Center for Applied Statistics and Economics (CASE), Berlin

This Version is available at:

<https://hdl.handle.net/10419/22205>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

---

# Computationally intensive Value at Risk calculations

Rafał Weron<sup>1</sup>

Hugo Steinhaus Center for Stochastic Methods, Wrocław University of Technology,  
50-370 Wrocław, Poland, [rweron@im.pwr.wroc.pl](mailto:rweron@im.pwr.wroc.pl)

## 1 Introduction

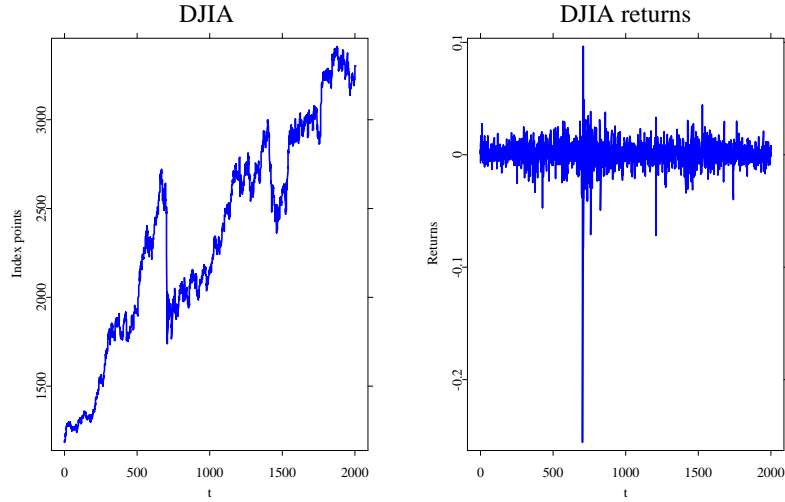
Market risks are the prospect of financial losses – or gains – due to unexpected changes in market prices and rates. Evaluating the exposure to such risks is nowadays of primary concern to risk managers in financial and non-financial institutions alike. Until late 1980s market risks were estimated through gap and duration analysis (interest rates), portfolio theory (securities), sensitivity analysis (derivatives) or "what-if" scenarios. However, all these methods either could be applied only to very specific assets or relied on subjective reasoning.

Since the early 1990s a commonly used market risk estimation methodology has been the Value at Risk (VaR). A VaR measure is the highest possible loss  $L$  incurred from holding the current portfolio over a certain period of time at a given confidence level (Dowd, 2002; Franke, Härdle and Stahl, 2000; Jorion, 2000):

$$\mathbb{P}(L > \text{VaR}) \leq 1 - c,$$

where  $c$  is the confidence level, typically 95%, 97.5% or 99%. By convention,  $L = -\Delta X(\tau)$ , where  $\Delta X(\tau)$  is the relative change (return) in portfolio value over the time horizon  $\tau$ . Hence, large values of  $L$  correspond to large losses (or large negative returns).

The VaR figure has two important characteristics: (i) it provides a common consistent measure of risk across different positions and risk factors and (ii) it takes into account the correlations or dependencies between different risk factors. Because of its intuitive appeal and simplicity, it is no surprise that in a few years Value at Risk has become the standard risk measure used around the world. However, VaR has a number deficiencies, among them the non-subadditivity – a sum of VaR's of two portfolios can be smaller than the VaR of the combined portfolio. To cope with these shortcomings, Artzner et al. (1999) proposed an alternative measure that satisfies the assumptions of a coherent, i.e. an adequate, risk measure. The Expected Shortfall (ES), also called Expected Tail Loss or Conditional VaR, is the expected value of the



**Fig. 1.** DJIA daily closing values,  $X_t$ , and daily returns,  $\log(X_{t+1}/X_t)$ , from the period January 2, 1985 – November 30, 1992. Note, that this period includes Black Monday, the worst stock market crash in Wall Street history. On October 19, 1987 DJIA lost 508 points or 25.6% of its value. Q: CSAfin01

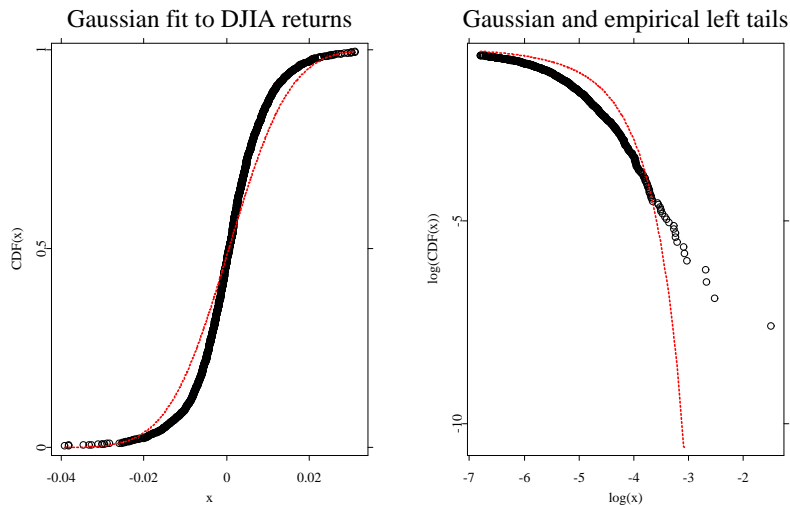
losses in excess of VaR:

$$\text{ES} = \mathbb{E}(L|L > \text{VaR}).$$

It is interesting to note, that – although new to the finance industry – Expected Shortfall has been familiar to insurance practitioners for a long time. It is very similar to the mean excess function which is used to characterize claim size distributions, see e.g. Cizek, Härdle and Weron (2004).

The essence of the VaR and ES computations is estimation of low quantiles in the portfolio return distributions. Hence, the performance of market risk measurement methods depends on the quality of distributional assumptions on the underlying risk factors. 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 and even the RiskMetrics variance-covariance approach to VaR – rest upon the assumption that asset returns follow a normal distribution. But is this assumption justified by empirical data?

No, it is not! It has been long known that asset returns are not normally distributed. Rather, the empirical observations exhibit excess kurtosis (fat tails). The Dow Jones Industrial Average (DJIA) index is a prominent example, see Figure 1 where the index itself and its returns (or log-returns) are depicted. In Figure 2 we plotted the empirical distribution of the DJIA index.



**Fig. 2.** Gaussian (dashed red) fit to the DJIA daily returns (black circles) empirical cumulative distribution function from the period January 2, 1985 – November 30, 1992. Top right panel is a magnification of the left tail fit on a double logarithmic scale clearly showing the discrepancy between the data and the normal distribution.

Q: CSAfin02

The contrast with the Gaussian law is striking. This heavy tailed or leptokurtic character of the distribution of price changes has been repeatedly observed in various markets and may be quantitatively measured by the kurtosis in excess of 3 (a value obtained for the normal distribution), see e.g. Bouchaud and Potters (2000), Carr et al. (2002), Gopikrishnan et al. (1998), Guillaume et al. (1997), and Rachev and Mittnik (2000).

This problem 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 ten day VaR at the  $c = 99\%$  confidence level multiplied by a safety factor  $s \in [3, 4]$ , with the exact value of  $s$  depending on the past performance of the model. It has been argued by Stahl (1997) and Danielsson, Hartmann and De Vries (1998) that the range of the safety factor comes from the heavy-tailed nature of the returns distribution. Indeed, if we assume that the asset returns distribution is symmetric and has finite variance  $\sigma^2$  then from Chebyshev's inequality (see e.g. Laha and Rohatgi, 1979) we obtain  $\mathbb{P}(L \geq \epsilon) \leq \sigma^2/2\epsilon^2$ , where  $L$  represents the random loss over the specified time horizon. So if we want to calculate the upper bound for a 99% VaR, setting  $\sigma^2/2\epsilon^2 = 1\%$  yields  $\epsilon = 7.07\sigma$ , which in turn implies that  $\text{VaR}_{99\%} \leq 7.07\sigma$ . However, if we assumed a Gaussian distribution of returns then we would have  $\text{VaR}_{99\%} \leq 2.33\sigma$ , which

is roughly three times lower than the bound obtained for a heavy-tailed, finite variance distribution.

Having said this much about the inadequacy of the Gaussian distribution for financial modeling and risk management we have no other choice but offer some heavy-tailed alternatives. We have to mention, though, that all distributional classes described in this chapter present computational challenge. Large parts of the text are thus devoted to numerical issues. In Section 2 we deal with the historically earliest alternative – the stable laws and briefly characterize their recent generalizations – the so-called truncated stable distributions. In Section 3 we study the class of generalized hyperbolic laws. Finally, in Section 4 we introduce the notion of copulas and discuss the relation between VaR, asset portfolios and heavy tails.

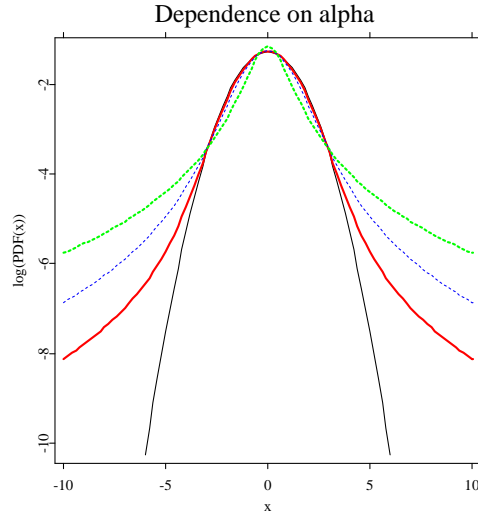
All theoretical results are illustrated by empirical examples which utilize the quantlets of the XploRe computing environment, see Härdle, Klinke and Müller (2000) and the XploRe website <http://www.xploRe-stat.de>. For reference, figure captions include names of the corresponding quantlets (Q). Currently, no other statistical computing environment offers a complete coverage of the issues discussed in this chapter. However, when available links to third-party libraries and specialized software are also provided.

## 2 Stable distributions

It is often argued that financial asset returns are the cumulative outcome of a vast number of pieces of information and individual decisions arriving almost continuously in time, see e.g. McCulloch (1996) and Rachev and Mittnik (2000). As such, since the pioneering work of Louis Bachelier in 1900, they have been modeled by the Gaussian distribution. The strongest statistical argument for it is based on the Central Limit Theorem, which states that the sum of a large number of independent, identically distributed variables from a finite-variance distribution will tend to be normally distributed. However, financial asset returns usually have heavier tails.

In response to the empirical evidence Mandelbrot (1963) and Fama (1965) proposed the stable distribution as an alternative model. There are at least two good reasons for modeling financial variables using stable distributions. Firstly, they are supported by the generalized Central Limit Theorem, which states that stable laws are the only possible limit distributions for properly normalized and centered sums of independent, identically distributed random variables, see e.g. Laha and Rohatgi (1979). Secondly, stable distributions are leptokurtic. Since they can accommodate the fat tails and asymmetry, they fit empirical distributions much better.

Stable laws – also called  $\alpha$ -stable, stable Paretian or Lévy stable – were introduced by Lévy (1925) during his investigations of the behavior of sums of independent random variables. A sum of two independent random variables



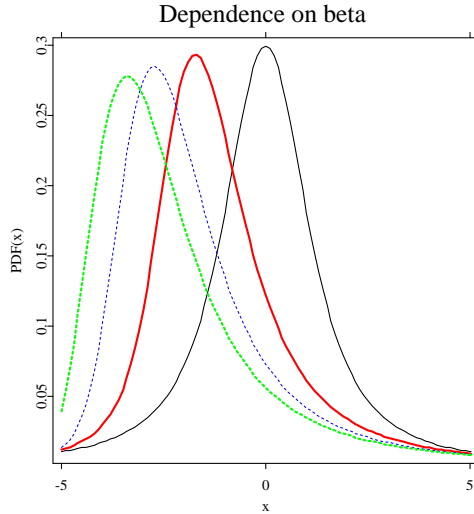
**Fig. 3.** A semilog plot of symmetric ( $\beta = \mu = 0$ )  $\alpha$ -stable probability density functions for  $\alpha = 2$  (thin black), 1.8 (red), 1.5 (thin, dashed blue) and 1 (dashed green). The Gaussian ( $\alpha = 2$ ) density forms a parabola and is the only  $\alpha$ -stable density with exponential tails. Q: STFstab01

having an  $\alpha$ -stable distribution with index  $\alpha$  is again  $\alpha$ -stable with the same index  $\alpha$ . This invariance property, however, does not hold for different  $\alpha$ 's.

The  $\alpha$ -stable distribution requires four parameters for complete description: an index of stability  $\alpha \in (0, 2]$  also called the tail index, tail exponent or characteristic exponent, a skewness parameter  $\beta \in [-1, 1]$ , a scale parameter  $\sigma > 0$  and a location parameter  $\mu \in \mathbb{R}$ . The tail exponent  $\alpha$  determines the rate at which the tails of the distribution taper off, see Figure 3. When  $\alpha = 2$ , a Gaussian distribution results. When  $\alpha < 2$ , the variance is infinite and the tails are asymptotically equivalent to a Pareto law, i.e. they exhibit a power-law behavior (Janicki and Weron, 1994a; Lévy, 1925; Samorodnitsky and Taqqu, 1994):

$$\mathbb{P}(X > x) \sim C_\alpha(1 + \beta)x^{-\alpha} \quad \text{and} \quad \mathbb{P}(X < -x) \sim C_\alpha(1 - \beta)x^{-\alpha}, \quad (1)$$

where  $C_\alpha = \frac{1}{\pi}\Gamma(\alpha)\sin\frac{\pi\alpha}{2}$ . When  $\alpha > 1$ , the mean of the distribution exists and is equal to  $\mu$ . In general, the  $p$ th moment of a stable random variable is finite if and only if  $p < \alpha$ . When the skewness parameter  $\beta$  is positive, the distribution is skewed to the right, i.e. the right tail is thicker, see Figure 4. When it is negative, it is skewed to the left. When  $\beta = 0$ , the distribution is symmetric about  $\mu$ . As  $\alpha$  approaches 2,  $\beta$  loses its effect and the distribution approaches the Gaussian distribution regardless of  $\beta$ . The last two parameters,



**Fig. 4.**  $\alpha$ -stable probability density functions for  $\alpha = 1.2$  and  $\beta = 0$  (thin black), 0.5 (red), 0.8 (thin, dashed blue) and 1 (dashed green). Q: STFstab02

$\sigma$  and  $\mu$ , are the usual scale and location parameters, i.e.  $\sigma$  determines the width and  $\mu$  the shift of the mode (the peak) of the distribution.

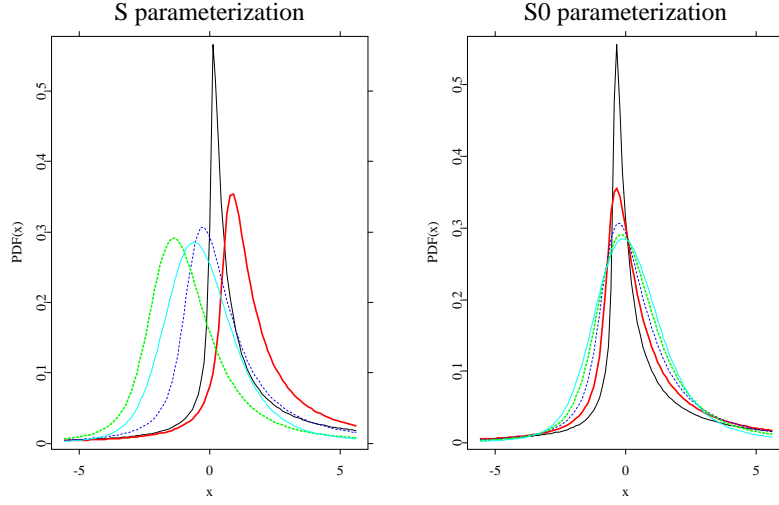
## 2.1 Characteristic function representation

From a practitioner's point of view the crucial drawback of the stable distribution is that, with the exception of three special cases, its probability density function (PDF) and cumulative distribution function (CDF) do not have closed form expressions. These exceptions include the well known Gaussian ( $\alpha = 2$ ) law, whose density function is given by:

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

and the lesser known Cauchy ( $\alpha = 1$ ,  $\beta = 0$ ) and Lévy ( $\alpha = 0.5$ ,  $\beta = 1$ ) laws.

Hence, the  $\alpha$ -stable distribution can be most conveniently described by its characteristic function  $\phi(t)$  – the inverse Fourier transform of the PDF. However, there are multiple parameterizations for  $\alpha$ -stable laws and much confusion has been caused by these different representations. The variety of formulas is caused by a combination of historical evolution and the numerous problems that have been analyzed using specialized forms of the stable distributions. The most popular parameterization of the characteristic function of  $X \sim S_\alpha(\sigma, \beta, \mu)$ , i.e. an  $\alpha$ -stable random variable with parameters  $\alpha$ ,  $\sigma$ ,  $\beta$  and  $\mu$ , is given by (Samorodnitsky and Taqqu, 1994; Weron, 1996):



**Fig. 5.** Comparison of  $S$  and  $S^0$  parameterizations:  $\alpha$ -stable probability density functions for  $\beta = 0.5$  and  $\alpha = 0.5$  (thin black),  $0.75$  (red),  $1$  (thin, dashed blue),  $1.25$  (dashed green) and  $1.5$  (thin cyan). Q: STFstab04

$$\log \phi(t) = \begin{cases} -\sigma^\alpha |t|^\alpha \{1 - i\beta \operatorname{sign}(t) \tan \frac{\pi\alpha}{2}\} + i\mu t, & \alpha \neq 1, \\ -\sigma |t| \{1 + i\beta \operatorname{sign}(t) \frac{2}{\pi} \log |t|\} + i\mu t, & \alpha = 1. \end{cases} \quad (3)$$

Note, that the traditional scale parameter  $\sigma$  of the Gaussian distribution is not the same as  $\sigma$  in the above representation. A comparison of formulas (2) and (3) yields the relation:  $\sigma_{\text{Gaussian}} = \sqrt{2}\sigma$ .

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

$$\log \phi_0(t) = \begin{cases} -\sigma^\alpha |t|^\alpha \{1 + i\beta \operatorname{sign}(t) \tan \frac{\pi\alpha}{2} [(\sigma|t|)^{1-\alpha} - 1]\} + i\mu_0 t, & \alpha \neq 1, \\ -\sigma |t| \{1 + i\beta \operatorname{sign}(t) \frac{2}{\pi} \log(\sigma|t|)\} + i\mu_0 t, & \alpha = 1. \end{cases} \quad (4)$$

The  $S_\alpha^0(\sigma, \beta, \mu_0)$  representation is a variant of Zolotarev's (1986) (M)-parameterization, with the characteristic function and hence the density and the distribution function jointly continuous in all four parameters, see Figure 5. In particular, percentiles and convergence to the power-law tail vary in a continuous way as  $\alpha$  and  $\beta$  vary. The location parameters of the two representations 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  $\alpha = 1$ .



## 2.2 Computation of stable density and distribution functions

The lack of closed form formulas for most stable densities and distribution functions has negative consequences. Numerical approximation or direct numerical integration have to be used, leading to a drastic increase in computational time and loss of accuracy. Of all the attempts to be found in the literature a few are worth mentioning. DuMouchel (1971) developed a procedure for approximating the stable distribution function using Bergström's (1952) series expansion. Depending on the particular range of  $\alpha$  and  $\beta$ , Holt and Crow (1973) combined four alternative approximations to compute the stable density function. Both algorithms are computationally intensive and time consuming, making maximum likelihood estimation a nontrivial task, even for modern computers. Recently, two other techniques have been proposed.

Mittnik, Doganoglu and Chenyao (1999) exploited the density function – characteristic function 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 a range of  $\alpha$ 's typically found for financial data ( $1.6 < \alpha < 1.9$ ). 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 probability density function calculations are concerned. When computing the cumulative distribution function 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 numerically integrating Zolotarev's (1986) integral formulas for the density or the distribution function. To save space we state only the formulas for the probability density function. Complete formulas can be also found in Borak, Härdle and Weron (2004).

Set  $\zeta = -\beta \tan \frac{\pi\alpha}{2}$ . Then the density  $f(x; \alpha, \beta)$  of a standard  $\alpha$ -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 yet another parametrization):

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

$$f(x; \alpha, \beta) = \frac{\alpha(x - \zeta)^{\frac{1}{\alpha-1}}}{\pi |\alpha - 1|} \int_{-\theta_0}^{\frac{\pi}{2}} V(\theta; \alpha, \beta) \exp \left\{ -(x - \zeta)^{\frac{\alpha}{\alpha-1}} V(\theta; \alpha, \beta) \right\} d\theta,$$

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

$$f(x; \alpha, \beta) = f(-x; \alpha, -\beta),$$

- when  $\alpha = 1$ :

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

where

$$\xi = \begin{cases} \frac{1}{\alpha} \arctan(-\zeta), & \alpha \neq 1, \\ \frac{\pi}{2}, & \alpha = 1, \end{cases} \quad (5)$$

and

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

XploRe offers the direct integration method through the `cdfstab` and `pdfstab` quantlets, see Borak, Härdle and Weron (2004) for a thorough exposition of quantlets related to stable distributions. On a PC equipped with a Pentium IV 1.6 GHz processor the calculation of the stable distribution or density function at 1000 points takes about 2 seconds. As default, the integrals found in the above formulas are numerically integrated using 2000 subintervals. These computational times can be slightly improved when using a numerically more efficient environment. For example, the program STABLE (downloadable from John Nolan's web page: <http://academic2.american.edu/~jpnolan/stable/stable.html>) needs about 1.6 seconds for performing corresponding calculations. It was written in Fortran and calls several external IMSL routines, see Nolan (1997) for details. Apart from speed, the STABLE program also exhibits higher relative accuracy (for default tolerance settings in both programs): about  $10^{-13}$  compared to  $10^{-5}$  for extreme tail events and  $10^{-10}$  for values used in typical financial applications (like approximating asset return distributions). Naturally, the accuracy of both programs can be increased at the cost of computational time.

It is interesting to note, that currently no other 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. John Nolan offers the STABLE program in library form through Robust Analysis Inc., see <http://www.robustanalysis.com>. This library (in C, Visual Basic or Fortran) provides interfaces to Matlab, S-plus (or its GNU version – R) and Mathematica. Diethelm Würtz has developed Rmetrics, an open source collection of software packages for S-plus/R, which may be useful for teaching computational finance, see <http://www.itp.phys.ethz.ch/econophysics/R/>. Stable PDF and CDF calculations are performed using the direct integration method, with the integrals being computed by R’s function `integrate`. Interestingly, for symmetric stable distributions Rmetrics utilizes McCulloch’s (1998) approximation, which was obtained by interpolating between the complements of the Cauchy and Gaussian CDFs in a transformed space. For  $\alpha$  in the range  $[0.92, 2.00]$  the absolute precision of the stable PDF and CDF approximation is  $6.6 \times 10^{-5}$  and  $2.2 \times 10^{-5}$ , respectively. The FFT based approach is utilized in Cognity, the first and probably the only commercial risk management platform that offers portfolio optimization based on the assumption of stably distributed returns, see <http://www.finanalytica.com>.

### 2.3 Simulation of $\alpha$ -stable variables

The complexity of the problem of simulating sequences of  $\alpha$ -stable random variables stems from the fact that there are no analytic expressions for the inverse  $F^{-1}(x)$  nor the cumulative distribution function  $F(x)$ . All standard approaches like the rejection or the inversion methods would require tedious computations. See Chapter ?? for a review of non-uniform random number generation techniques.

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) yielded 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:

$$X = (1 + \zeta^2)^{\frac{1}{2\alpha}} \frac{\sin\{\alpha(U + \xi)\}}{\{\cos(U)\}^{1/\alpha}} \left[ \frac{\cos\{U - \alpha(U + \xi)\}}{W} \right]^{\frac{1-\alpha}{\alpha}}, \quad (6)$$

- for  $\alpha = 1$  compute:

$$X = \frac{1}{\xi} \left\{ \left( \frac{\pi}{2} + \beta U \right) \tan U - \beta \log \left( \frac{\frac{\pi}{2} W \cos U}{\frac{\pi}{2} + \beta U} \right) \right\}, \quad (7)$$

where  $\xi$  is given by eqn. (5). This algorithm yields a random variable  $X \sim S_\alpha(1, \beta, 0)$ , in representation (3). For a detailed proof see Weron (1996).

Given the formulas for simulation of a standard  $\alpha$ -stable random variable, we can easily simulate a stable random variable for all admissible values of the parameters  $\alpha$ ,  $\sigma$ ,  $\beta$  and  $\mu$  using the following property: if  $X \sim S_\alpha(1, \beta, 0)$  then

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

is  $S_\alpha(\sigma, \beta, \mu)$ . It is interesting to note that for  $\alpha = 2$  (and  $\beta = 0$ ) the Chambers-Mallows-Stuck method reduces to the well known Box-Muller algorithm for generating Gaussian random variables, see e.g. Devroye (1986) and Janicki and Weron (1994b).

Many other approaches have been proposed in the literature, including application of Bergström (1952) and LePage (LePage, Woodroffe and Zinn, 1981) series expansions, see Mantegna (1994) and Janicki and Kokoszka (1992), respectively. However, this method is regarded as the fastest and the most accurate. In XploRe the algorithm is implemented in the `rndstab` quantlet. On a PC equipped with a Pentium IV 1.6 GHz processor one million variables are generated in about 7 seconds, compared to about 1 second for one million standard normal random variables obtained via the Box-Muller algorithm (`normal2`). Because of its unquestioned superiority and relative simplicity, the Chambers-Mallows-Stuck method 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.

## 2.4 Estimation of parameters

The estimation of stable law parameters is in general severely hampered by the lack of known closed-form density functions for all but a few members of the stable family. Numerical approximation or direct numerical integration are nontrivial and burdensome from a computational point of view. As a consequence, the maximum likelihood (ML) estimation algorithm based on such approximations is difficult to implement and time consuming for samples encountered in modern finance. However, there are numerical methods that have been found useful in practice and are discussed in this section.

All presented methods work quite well assuming that the sample under consideration is indeed  $\alpha$ -stable. Since there are no formal tests for assessing the  $\alpha$ -stability of a data set we suggest to first apply the "visual inspection" or tail exponent estimators, like the Hill (1975) and direct tail estimation methods, to see whether the empirical densities resemble those of  $\alpha$ -stable laws, see e.g. Borak, Härdle and Weron (2004) and Weron (2001).

Given a sample  $x_1, \dots, x_n$  from  $S_\alpha(\sigma, \beta, \mu)$ , in what follows, we provide estimates  $\hat{\alpha}$ ,  $\hat{\sigma}$ ,  $\hat{\beta}$  and  $\hat{\mu}$  of  $\alpha$ ,  $\sigma$ ,  $\beta$  and  $\mu$ , respectively. We start the discussion

with the simplest, fastest and ... least accurate quantile methods, then develop the slower, yet much more accurate sample characteristic function methods and, finally, conclude with the slowest but most accurate maximum likelihood approach.

### Sample quantiles methods

Fama and Roll (1971) provided very simple estimates for parameters of symmetric ( $\beta = 0, \mu = 0$ ) stable laws with  $\alpha > 1$ . They propose to estimate  $\sigma$  by:

$$\hat{\sigma} = \frac{\hat{x}_{0.72} - \hat{x}_{0.28}}{1.654}, \quad (8)$$

where  $x_f$  denotes the  $f$ -th population quantile, so that  $S_\alpha(\sigma, \beta, \mu)(x_f) = f$ . As McCulloch (1986) noticed, Fama and Roll based their estimator of  $\sigma$  on the fortuitous observation that  $(x_{0.72} - x_{0.28})/\sigma$  lies within 0.4% of 1.654 for all  $1 < \alpha \leq 2$ , when  $\beta = 0$ . This enabled them to estimate  $\sigma$  by (8) with less than 0.4% asymptotic bias without first knowing  $\alpha$ . However, when  $\beta \neq 0$ , the search for an invariant range such as the one they found becomes futile.

The characteristic exponent  $\alpha$ , on the other hand, can be estimated from the tail behavior of the distribution. Fama and Roll suggested to take  $\hat{\alpha}$  satisfying:

$$S_{\hat{\alpha}}\left(\frac{\hat{x}_f - \hat{x}_{1-f}}{2\hat{\sigma}}\right) = f. \quad (9)$$

They found that  $f = 0.95, 0.96, 0.97$  worked best for estimating  $\alpha$ . This method unnecessarily compounds the small asymptotic bias in the estimator of  $\sigma$  into the estimator of  $\alpha$ .

For  $1 < \alpha \leq 2$ , the stable distribution has finite mean. Hence, the sample mean is a consistent estimate of the location parameter  $\mu$ . However, a more robust estimate is the  $p\%$  truncated sample mean – the arithmetic mean of the middle  $p$  percent of the ranked observations. The 50% truncated mean is often suggested in the literature when the range of  $\alpha$  is unknown.

Fama and Roll's (1971) method is simple but suffers from a small asymptotic bias in  $\hat{\alpha}$  and  $\hat{\sigma}$  and restrictions on  $\alpha$  and  $\beta$ . McCulloch (1986) generalized and improved the quantile method. He analyzed stable law quantiles and provided consistent estimators of all four stable parameters, with the restriction  $\alpha \geq 0.6$ , while retaining the computational simplicity of Fama and Roll's method. After McCulloch define:

$$v_\alpha = \frac{x_{0.95} - x_{0.05}}{x_{0.75} - x_{0.25}} \quad \text{and} \quad v_\beta = \frac{x_{0.95} + x_{0.05} - 2x_{0.50}}{x_{0.95} - x_{0.05}}. \quad (10)$$

Statistics  $v_\alpha$  and  $v_\beta$  are functions of  $\alpha$  and  $\beta$  only, i.e. they are independent of both  $\sigma$  and  $\mu$ . This relationship may be inverted and the parameters  $\alpha$  and  $\beta$  may be viewed as functions of  $v_\alpha$  and  $v_\beta$ . Substituting  $v_\alpha$  and  $v_\beta$  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,  $\sigma$  and  $\mu$ , can be estimated in a similar way. However, due to the discontinuity of the characteristic function for  $\alpha = 1$  and  $\beta \neq 0$  in representation (3), this procedure is much more complicated. We refer the interested reader to the original work of McCulloch (1986). This estimation technique is implemented in XploRe in the `stabcull` quantlet.

### Sample characteristic function methods

Given an i.i.d. random sample  $x_1, \dots, x_n$  of size  $n$ , define the *sample characteristic function* 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 characteristic function  $\phi(t)$ .

Press (1972) proposed a simple estimation method, called the method of moments, based on transformations of the characteristic function. From (3) we have for all  $\alpha$ :

$$|\phi(t)| = \exp(-\sigma^\alpha |t|^\alpha). \quad (11)$$

Hence,  $-\log |\phi(t)| = \sigma^\alpha |t|^\alpha$ . Now, assuming  $\alpha \neq 1$ , choose two nonzero values of  $t$ , say  $t_1 \neq t_2$ . Then for  $k = 1, 2$  we have:

$$-\log |\phi(t_k)| = \sigma^\alpha |t_k|^\alpha. \quad (12)$$

Solving these two equations for  $\alpha$  and  $\sigma$ , and substituting  $\hat{\phi}(t)$  for  $\phi(t)$  yields:

$$\hat{\alpha} = \frac{\log \frac{\log |\hat{\phi}(t_1)|}{\log |\hat{\phi}(t_2)|}}{\log \left| \frac{t_1}{t_2} \right|}, \quad (13)$$

and

$$\log \hat{\sigma} = \frac{\log |t_1| \log(-\log |\hat{\phi}(t_2)|) - \log |t_2| \log(-\log |\hat{\phi}(t_1)|)}{\log \left| \frac{t_1}{t_2} \right|}. \quad (14)$$

In order to estimate  $\beta$  and  $\mu$  we have to choose two nonzero values of  $t$ , say  $t_3 \neq t_4$ , and apply a similar trick to  $\Im\{\log \phi(t)\}$ . The estimators are consistent since they are based upon estimators of  $\phi(t)$ ,  $\Im\{\phi(t)\}$  and  $\Re\{\phi(t)\}$ , which are known to be consistent. However, convergence to the population values depends on the choice of  $t_1, \dots, t_4$ . The optimal selection of these values is problematic and still is an open question. The XploRe implementation of the method of moments (the `stabmom` quantlet) uses  $t_1 = 0.2$ ,  $t_2 = 0.8$ ,  $t_3 = 0.1$ , and  $t_4 = 0.4$  as proposed by Koutrouvelis (1980) in his simulation study.

In the same paper Koutrouvelis 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  $\alpha$ . 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 characteristic function  $\phi(t)$ . First, from (3) we can easily derive:

$$\log(-\log|\phi(t)|^2) = \log(2\sigma^\alpha) + \alpha \log|t|. \quad (15)$$

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

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

and

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

The last two equations lead, apart from considerations of principal values, to:

$$\arctan\left(\frac{\Im\{\phi(t)\}}{\Re\{\phi(t)\}}\right) = \mu t + \beta \sigma^\alpha \tan \frac{\pi\alpha}{2} \mathbf{sign}(t) |t|^\alpha. \quad (16)$$

Equation (15) depends only on  $\alpha$  and  $\sigma$  and suggests that we estimate these parameters by regressing  $y = \log(-\log|\phi_n(t)|^2)$  on  $w = \log|t|$  in the model:

$$y_k = m + \alpha w_k + \epsilon_k, \quad (17)$$

where  $t_k$  is an appropriate set of real numbers,  $m = \log(2\sigma^\alpha)$ , and  $\epsilon_k$  denotes an error term. Koutrouvelis (1980) proposed to use  $t_k = \frac{\pi k}{25}$ ,  $k = 1, 2, \dots, K$ ; with  $K$  ranging between 9 and 134 for different values of  $\alpha$  and sample sizes.

Once  $\hat{\alpha}$  and  $\hat{\sigma}$  have been obtained and  $\alpha$  and  $\sigma$  have been fixed at these values, estimates of  $\beta$  and  $\mu$  can be obtained using (16). 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 the Fama-Roll estimator (8) and the 25% truncated mean for initial estimates of  $\sigma$  and  $\mu$ , 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 (4) of the characteristic function instead of the classical one (3) 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 method of Koutrouvelis, see Table 1. It has a significantly better performance near  $\alpha = 1$  and  $\beta \neq 0$  due

**Table 1.** Comparison of McCulloch’s quantile technique, the method of moments, the regression approach of Koutrouvelis and the method of Kogon and Williams for 100 simulated samples of two thousand  $S_{1.7}(0.005, 0.1, 0.001)$  random variables each. Parameter estimates are mean values over 100 samples. Values of the Mean Absolute Percentage Error ( $\text{MAPE}_\theta = \frac{1}{n} \sum_{i=1}^n |\hat{\theta} - \theta|/\theta$ ) are given in parentheses. In the last column average computational times for one sample of 2000 random variables are provided (on a PC equipped with a Pentium IV 1.6 GHz processor and running XploRe 4.6). Q: CSAfin03

| Method         | $\hat{\alpha}$     | $\hat{\sigma}$      | $\hat{\beta}$       | $\hat{\mu}$        | CPU time |
|----------------|--------------------|---------------------|---------------------|--------------------|----------|
| McCulloch      | 1.7005<br>(2.60%)  | 0.0050<br>(2.16%)   | 0.1045<br>(110.72%) | 0.0010<br>(22.01%) | 0.04s    |
| Moments        | 1.9895<br>(17.03%) | 0.0104<br>(107.64%) | 0.0712<br>(969.57%) | 0.0010<br>(33.56%) | 0.03s    |
| Koutrouvelis   | 1.6988<br>(1.66%)  | 0.0050<br>(1.69%)   | 0.0989<br>(108.21%) | 0.0010<br>(21.01%) | 0.69s    |
| Kogon-Williams | 1.6994<br>(1.95%)  | 0.0050<br>(1.77%)   | 0.0957<br>(110.59%) | 0.0010<br>(21.14%) | 0.18s    |

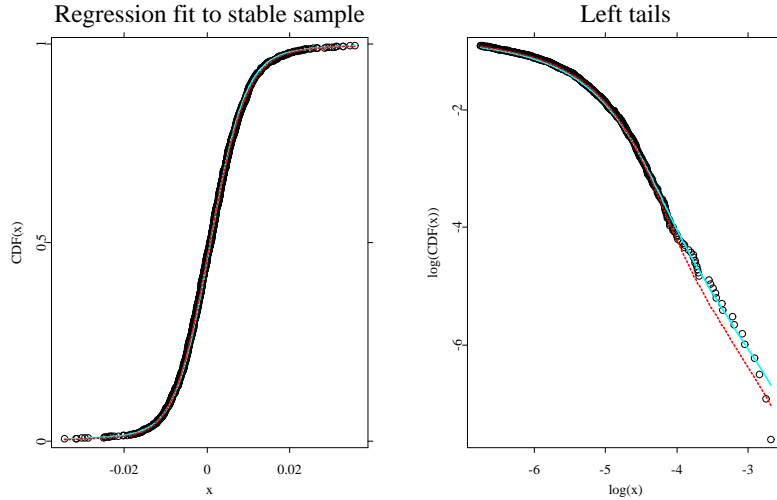
to the elimination of discontinuity of the characteristic function. However, it returns slightly worse results for other values of  $\alpha$ . In XploRe both regression algorithms are implemented in the `stabreg` quantlet. An optional parameter lets the user choose between the original Koutrouvelis code and the Kogon-Williams modification.

A typical performance of the described estimators is summarized in Table 1, see also Figure 6. McCulloch’s quantile technique, the method of moments, the regression approach of Koutrouvelis and the method of Kogon and Williams were applied to 100 simulated samples of two thousand  $S_{1.7}(0.005, 0.1, 0.001)$  random variables each. The method of moments yielded the worst estimates, clearly outside any admissible error range. McCulloch’s method came in next with acceptable results and computational time significantly lower than the regression approaches. On the other hand, both the Koutrouvelis and the Kogon-Williams implementations yielded good estimators with the latter performing considerably faster, but slightly less accurate. We have to say, though, that all methods had problems with estimating  $\beta$ . Like it or not, our search for the optimal estimation technique is not over yet. We have no other choice but turn to the last resort – the maximum likelihood method.

### Maximum likelihood method

The maximum likelihood (ML) estimation scheme for  $\alpha$ -stable distributions does not differ from that for other laws, at least as far as the theory is concerned. For a vector of observations  $x = (x_1, \dots, x_n)$ , the ML estimate of the





**Fig. 6.** Regression fit (dashed red), using Koutrouvelis' regression method, to 2000 simulated  $S_{1.7}(0.005, 0.1, 0.001)$  random variables (black circles). For comparison, the CDF of the original distribution is also plotted (cyan). Right panel is a magnification of the left tail fit on a double logarithmic scale. **Q: CSAfin04**

parameter vector  $\theta = (\alpha, \sigma, \beta, \mu)$  is obtained by maximizing the log-likelihood function:

$$L_{\theta}(x) = \sum_{i=1}^n \log \tilde{f}(x_i; \theta), \quad (18)$$

where  $\tilde{f}(\cdot; \theta)$  is the stable density function. The tilde denotes the fact that, in general, we do not know the explicit form of the stable PDF 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 maximum likelihood estimator is asymptotically normal with the variance specified by the Fischer information matrix, see 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. 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 (the fast Fourier transform for central values of  $x$  and series expansions for tails) to compute an approximate log-likelihood function. This function was then numerically maximized.

Applying Zolotarev's (1964) integral formulas, Brorsen and Yang (1990) formulated another approximate ML method, however, only for symmetric stable random variables. To avoid the discontinuity and nondifferentiability of the symmetric stable density function at  $\alpha = 1$ , the tail index  $\alpha$  was restricted to be greater than one.

Much better, in terms of accuracy and computational time, are more recent maximum likelihood 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 2.2.

As Ojeda (2001) observes, the ML estimates are almost always the most accurate, closely followed by the regression-type estimates, McCulloch's quantile method, and finally the method of moments. However, as we have already said in the introduction to this Section, maximum likelihood estimation techniques are certainly the slowest of all the discussed methods. For example, ML estimation for a sample of 2000 observations using a gradient search routine which utilizes the direct integration method needs 484 seconds or about 8 minutes! The calculations were performed on a PC equipped with a Pentium IV 1.6 GHz processor and running STABLE ver. 3.13 (see also Section 2.2 where the program was briefly described). For comparison, the STABLE implementation of the Kogon-Williams algorithm performs the same calculations in only 0.04 seconds (the XploRe quantlet `stabreg` needs roughly four times more time, see Table 1). 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 also 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  $(x, \alpha, \beta)$  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 – 0.57s for a sample of 2000 observations.

## 2.5 Are DJIA and DAX returns $\alpha$ -stable ?

In this paragraph we want to apply the discussed techniques to financial data. Due to limited space we chose only one estimation method – the regression approach of Koutrouvelis (1980), as it offers high accuracy at moderate computational time. We start the empirical analysis with the most prominent example – the Dow Jones Industrial Average (DJIA) index. The data set covers the period January 2, 1985 – November 30, 1992 and comprises 2000 returns. Recall, that this period includes the largest crash in Wall Street history – the Black Monday of October 19, 1987. Clearly the 1.66-stable law offers a much better fit to the DJIA returns than the Gaussian distribution, see Table 2. Its

**Table 2.**  $\alpha$ -stable and Gaussian fits to 2000 returns of the Dow Jones Industrial Average (DJIA) index from the period January 2, 1985 – November 30, 1992. Values of the Anderson-Darling and Kolmogorov goodness-of-fit statistics are also given. The former may be treated as a weighted Kolmogorov statistics which puts more weight to the differences in the tails of the distributions. Q: CSAfin05

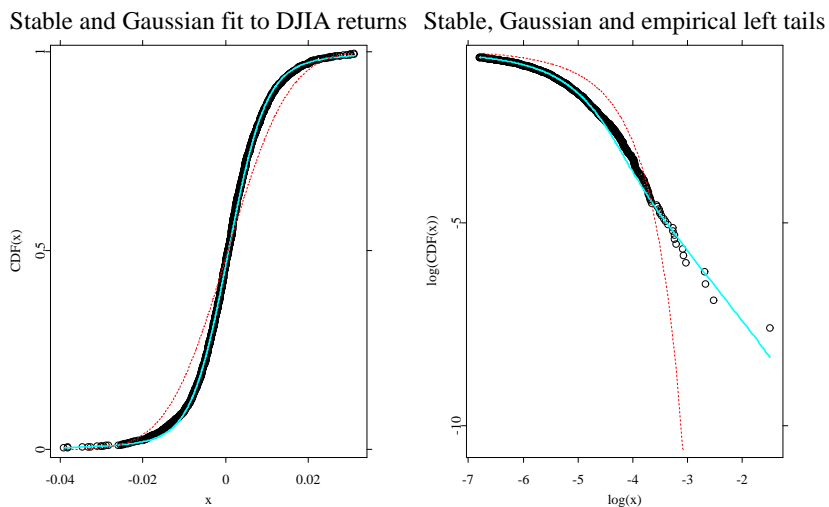
| Parameters           | $\alpha$         | $\sigma$   | $\beta$ | $\mu$  |
|----------------------|------------------|------------|---------|--------|
| $\alpha$ -stable fit | 1.6596           | 0.0053     | 0.0823  | 0.0009 |
| Gaussian fit         |                  | 0.0115     |         | 0.0006 |
| Test values          | Anderson-Darling | Kolmogorov |         |        |
| $\alpha$ -stable fit | 1.0044           | 0.8641     |         |        |
| Gaussian fit         | +INF             | 4.5121     |         |        |

superiority, especially in the tails of the distribution, is even better visible in Figure 7.

To make our statistical analysis more sound, we compare both fits through Anderson-Darling and Kolmogorov test statistics, see Cizek, Härdle and Weron (2004) and D’Agostino and Stephens (1986). The former may be treated as a weighted Kolmogorov statistics which puts more weight to the differences in the tails of the distributions. Although no asymptotic results are known for the stable laws, approximate critical values for these goodness-of-fit tests can be obtained via the bootstrap technique, see Borak, Härdle and Weron (2004) and Stute, Manteiga and Quindimil (1993). In this chapter, though, we will not perform hypothesis testing and just compare the test values. Naturally, the lower the values the better the fit. The stable law seems to be tailor-cut for the DJIA index returns. But does it fit other asset returns as well?

The second analyzed data set comprises 2000 returns of the Deutsche Aktienindex (DAX) index from the period January 2, 1995 – December 5, 2002. Also in this case the  $\alpha$ -stable law offers a much better fit than the Gaussian, see Table 3. However, the test statistics suggest that the fit is not as good as for the DJIA returns (observe that both data sets are of the same size and the test values in both cases can be compared). This can be also seen in Figure 8. The left tail seems to drop off at some point and the very tail is largely overestimated by the stable distribution. At the same time it is better approximated by the Gaussian law.

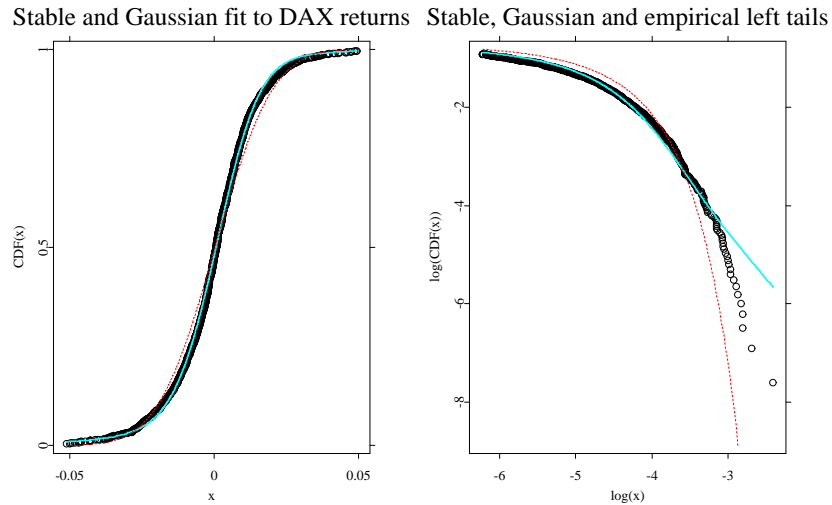
This example clearly shows that the  $\alpha$ -stable distribution is not a panacea. Although it gives a very good fit to a number of empirical data sets, there surely are distributions that recover the characteristics of other data sets better. We devote the rest of this chapter to such alternative heavy tailed distributions. We start with a modification of the stable law and in Section 3 concentrate on the class of generalized hyperbolic distributions.



**Fig. 7.** 1.66-stable (cyan) and Gaussian (dashed red) fits to the DJIA returns (black circles) empirical cumulative distribution function from the period January 2, 1985 – November 30, 1992. For better exposition of the fit in the central part of the distribution ten largest and ten smallest returns are not illustrated in the left panel. The largest loss (i.e. the smallest return) of  $-0.226$  or  $-22.6\%$  was sustained on Black Monday. The right panel is a magnification of the left tail fit on a double logarithmic scale clearly showing the superiority of the stable law. Q: CSAfin05

**Table 3.**  $\alpha$ -stable and Gaussian fits to 2000 returns of the Deutsche Aktienindex (DAX) index from the period January 2, 1995 – December 5, 2002. Q: CSAfin06

| Parameters           | $\alpha$         | $\sigma$ | $\beta$    | $\mu$   |
|----------------------|------------------|----------|------------|---------|
| $\alpha$ -stable fit | 1.7003           | 0.0088   | -0.3179    | -0.0002 |
| Gaussian fit         |                  | 0.0157   |            | 0.0004  |
| Test values          | Anderson-Darling |          | Kolmogorov |         |
| $\alpha$ -stable fit | 1.9149           |          | 1.1798     |         |
| Gaussian fit         | 16.4119          |          | 2.8197     |         |



**Fig. 8.** 1.7-stable (cyan) and Gaussian (dashed red) fits to the DAX returns (black circles) empirical cumulative distribution function from the period January 2, 1995 – December 5, 2002. For better exposition of the fit in the central part of the distribution ten largest and ten smallest returns are not illustrated in the left panel. The right panel is a magnification of the left tail fit on a double logarithmic scale. This time the stable law overfits the tails of the empirical distribution. **Q: CSAfin06**

## 2.6 Truncated stable distributions

Mandelbrot's (1963) pioneering work on applying  $\alpha$ -stable distributions to asset returns gained support in the first few years after its publication (Fama, 1965; Officer, 1972; Teichmoeller, 1971). Subsequent works, however, have questioned the stable distribution hypothesis (Akgiray and Booth, 1988; Blattberg and Gonedes, 1974). By the definition of the stability property, the sum of i.i.d. stable random variables is also stable. Thus, if short term asset returns are distributed according to a stable law, longer term returns should retain the same functional form. However, from the empirical data it is evident that as the time interval between price observations grows longer, the distribution of returns deviates from the short term heavy tailed distribution, and converges to the Gaussian law. This indicates that the returns probably are not  $\alpha$ -stable (but it could also mean that the returns are just not independent). Over the next few years, the stable distribution temporarily lost favor and alternative processes were suggested as mechanisms generating stock returns.

In mid 1990s the stable distribution hypothesis has made a dramatic comeback. Recently 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, see e.g. Cont, Potters and Bouchaud (1997) and Mantegna and Stanley (1995). For more extreme observations, however, the distribution they have found falls off approximately exponentially. To cope with such observations the truncated Lévy distributions (TLD) were introduced by Mantegna and Stanley (1994). The original definition postulated a sharp truncation of the  $\alpha$ -stable probability density function at some arbitrary point. However, later an exponential smoothing was proposed by Koponen (1995).

For  $\alpha \neq 1$  the characteristic function of a symmetric TLD random variable is given by:

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

where  $\alpha$  is the tail exponent,  $\sigma$  is the scale parameter and  $\lambda$  is the truncation coefficient. Clearly the TLD reduces to the symmetric  $\alpha$ -stable distribution ( $\beta = \mu = 0$ ) when  $\lambda = 0$ . The TLD distribution exhibits the following behavior: for small and intermediate returns it behaves like a stable distribution, but for extreme returns the truncation causes the distribution to converge to a Gaussian distribution. Thus the observation that the asset returns distribution is a TLD explains both the short-term  $\alpha$ -stable behavior and the long run convergence to the normal distribution.

Despite these interesting features the truncated Lévy distributions have not been applied extensively to date. The most probable reason for this being the complicated definition of the TLD law. Like for  $\alpha$ -stable distributions, only the characteristic function is known. No closed form formulas exist for the density or the distribution functions. Since no integral formulas, like Zolotarev's (1986) for the  $\alpha$ -stable laws, have been discovered as yet, statistical inference is, in general, limited to maximum likelihood utilizing the FFT technique for approximating the PDF. Moreover, compared to the stable distribution, the TLD introduces one more parameter (asymmetric TLD laws have also been considered in the literature, see e.g. Boyarchenko and Levendorskii (2000) and Koponen (1995)) making the estimation procedure even more complicated. Other parameter fitting techniques proposed in the literature comprise a combination of ad hoc approaches and moment matching, see e.g. Boyarchenko and Levendorskii (2000) and Matacz (2000). Better techniques have to be discovered before TLDs become a common tool in finance.

### 3 Hyperbolic distributions

In response to remarkable regularities discovered by geomorphologists in the 1940s, Barndorff-Nielsen (1977) introduced the hyperbolic law for modeling the grain size distribution of windblown sand. Excellent fits were also obtained for the log-size distribution of diamonds from a large mining area in South West Africa. Almost twenty years later the hyperbolic law was found to provide a very good model for the distributions of daily returns of stocks from

a number of leading German enterprises (Eberlein and Keller, 1995; Küchler et al., 1999), giving way to its today's use in stock price modeling (Bibby and Sørensen, 1997) and market risk measurement (Eberlein, Keller and Prause, 1998). The name of the distribution is derived from the fact that its log-density forms a hyperbola, see Figure 9. Recall that the log-density of the normal distribution is a parabola. Hence the hyperbolic distribution provides the possibility of modeling heavier tails.

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, see Barndorff-Nielsen (1977) and Barndorff-Nielsen and Blaesild (1981). More precisely, a random variable  $Z$  has the hyperbolic distribution if:

$$(Z|Y) \sim N(\mu + \beta Y, Y), \quad (19)$$

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

$$f_{\text{GIG}}(x) = \frac{(\psi/\chi)^{\lambda/2}}{2K_{\lambda}(\sqrt{\chi\psi})} x^{\lambda-1} e^{-\frac{1}{2}(\chi x^{-1} + \psi x)}, \quad x > 0, \quad (20)$$

where the normalizing constant:

$$K_{\lambda}(t) = \frac{1}{2} \int_0^{\infty} x^{\lambda-1} e^{-\frac{1}{2}t(x+x^{-1})} dx, \quad t > 0, \quad (21)$$

is the modified Bessel function of the third kind with index  $\lambda$ . Note that  $K_{\lambda}(t)$  is also known as the MacDonald function. 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  $\lambda$ , i.e.  $K_{\lambda}(t) = K_{-\lambda}(t)$ , and (ii) for  $\lambda = \pm\frac{1}{2}$  it can be written in a simpler form:

$$K_{\pm\frac{1}{2}}(t) = \sqrt{\frac{\pi}{2}} t^{-\frac{1}{2}} e^{-t}. \quad (22)$$

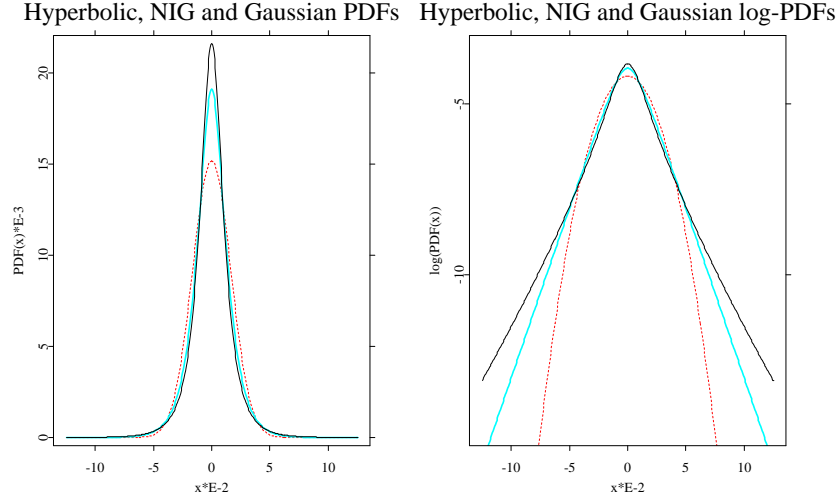
For other values of  $\lambda$  numerical approximations of the integral in eqn. (21) have to be used, see e.g. Campbell (1980), Press et al. (1992) or Temme (1975).

Relation (19) implies that a hyperbolic random variable  $Z \sim H(\psi, \beta, \chi, \mu)$  can be represented in the form:

$$Z \sim \mu + \beta Y + \sqrt{Y}N(0, 1),$$

with the characteristic function:

$$\phi_Z(u) = e^{iu\mu} \int_0^{\infty} e^{i\beta zu - \frac{1}{2}zu^2} dF_Y(z). \quad (23)$$



**Fig. 9.** Densities and log-densities of hyperbolic (cyan), NIG (thin black) and Gaussian (thin, dashed red) distributions having the same variance, see (30). 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. Q: CSAfin07

Here  $F_Y(z)$  denotes the distribution function of a generalized inverse Gaussian random variable  $Y$  with parameter  $\lambda = 1$ , see eqn. (20). Hence, the hyperbolic PDF is given by:

$$f_{\mathbf{H}}(x) = \frac{\sqrt{\psi/\chi}}{2\sqrt{\psi + \beta^2} K_1(\sqrt{\psi\chi})} e^{-\sqrt{\psi + \beta^2}\{\chi + (x - \mu)^2\} + \beta(x - \mu)}. \quad (24)$$

Sometimes another parameterization of the hyperbolic distribution with  $\delta = \sqrt{\chi}$  and  $\alpha = \sqrt{\psi + \beta^2}$  is used. Then the probability density function of the hyperbolic  $\mathbf{H}(\alpha, \beta, \delta, \mu)$  law can be written as:

$$f_{\mathbf{H}}(x) = \frac{\sqrt{\alpha^2 - \beta^2}}{2\alpha\delta K_1(\delta\sqrt{\alpha^2 - \beta^2})} e^{-\alpha\sqrt{\delta^2 + (x - \mu)^2} + \beta(x - \mu)}, \quad (25)$$

where  $\delta > 0$  is the scale parameter,  $\mu \in \mathbb{R}$  is the location parameter and  $0 \leq |\beta| < \alpha$ . The latter two parameters –  $\alpha$  and  $\beta$  – determine the shape, with  $\alpha$  being responsible for the steepness and  $\beta$  for the skewness. In XploRe the hyperbolic density and distribution functions are implemented in the `pdfhyp` and `cdfhyp` quantlets, respectively. The calculation of the PDF is straightforward, however, the CDF has to be numerically integrated from eqn. (25).

The hyperbolic law is a member of a more general class of generalized hyperbolic distributions, which also includes the normal-inverse Gaussian (NIG)



and variance-gamma distributions as special cases. The generalized hyperbolic law can be represented as a normal variance-mean mixture where the mixing distribution is the generalized inverse Gaussian (GIG) law with any  $\lambda \in \mathbb{R}$ . Hence, the generalized hyperbolic distribution is described by five parameters  $\theta = (\lambda, \alpha, \beta, \delta, \mu)$ . Its probability density function is given by:

$$f_{\text{GH}}(x) = \kappa \left\{ \delta^2 + (x - \mu)^2 \right\}^{\frac{1}{2}(\lambda - \frac{1}{2})} K_{\lambda - \frac{1}{2}} \left( \alpha \sqrt{\delta^2 + (x - \mu)^2} \right) e^{\beta(x - \mu)}, \quad (26)$$

where:

$$\kappa = \frac{(\alpha^2 - \beta^2)^{\frac{\lambda}{2}}}{\sqrt{2\pi} \alpha^{\lambda - \frac{1}{2}} \delta^\lambda K_\lambda(\delta \sqrt{\alpha^2 - \beta^2})}. \quad (27)$$

For  $|\beta + z| < \alpha$  its moment generating function takes the form:

$$M(z) = e^{\mu z} \left\{ \frac{\alpha^2 - \beta^2}{\alpha^2 - (\beta + z)^2} \right\}^{\frac{\lambda}{2}} \frac{K_\lambda \left( \delta \sqrt{\alpha^2 - (\beta + z)^2} \right)}{K_\lambda \left( \delta \sqrt{\alpha^2 - \beta^2} \right)}. \quad (28)$$

Note, that  $M(z)$  is smooth, i.e. infinitely many times differentiable, near 0 and hence every moment exists. If we set  $\zeta = \delta \sqrt{\alpha^2 - \beta^2} = \sqrt{\psi\chi}$  then the first two moments lead to the following formulas for the mean and variance of a generalized hyperbolic random variable:

$$\mathbb{E}X = \mu + \frac{\beta \delta^2}{\zeta} \frac{K_{\lambda+1}(\zeta)}{K_\lambda(\zeta)}, \quad (29)$$

$$\text{Var}X = \delta^2 \left[ \frac{K_{\lambda+1}(\zeta)}{\zeta K_\lambda(\zeta)} + \frac{\beta^2 \delta^2}{\zeta^2} \left\{ \frac{K_{\lambda+2}(\zeta)}{K_\lambda(\zeta)} - \left( \frac{K_{\lambda+1}(\zeta)}{\zeta K_\lambda(\zeta)} \right)^2 \right\} \right]. \quad (30)$$

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

$$f_{\text{NIG}}(x) = \frac{\alpha \delta}{\pi} e^{\delta \sqrt{\alpha^2 - \beta^2} + \beta(x - \mu)} \frac{K_1(\alpha \sqrt{\delta^2 + (x - \mu)^2})}{\sqrt{\delta^2 + (x - \mu)^2}}. \quad (31)$$

In XploRe the NIG density and distribution functions are implemented in the `pdfnig` and `cdfnig` quantlets, respectively. Like for the hyperbolic distribution the calculation of the PDF is straightforward, but the CDF has to be numerically integrated from eqn. (31).

At the "expense" of four parameters, the NIG distribution is able to model symmetric and asymmetric distributions with possibly long tails in both directions. Its tail behavior is often classified as "semi-heavy", i.e. the tails are lighter than those of non-Gaussian stable laws, but much heavier than Gaussian. Interestingly, if we let  $\alpha$  tend to zero the NIG distribution converges to

the Cauchy distribution (with location parameter  $\mu$  and scale parameter  $\delta$ ), which exhibits extremely heavy tails. The tail behavior of the NIG density is characterized by the following asymptotic relation:

$$f_{NIG}(x) \sim |x|^{-3/2} e^{(\mp\alpha+\beta)x} \quad \text{for } x \rightarrow \pm\infty. \quad (32)$$

In fact, this is a special case of a more general relation with the exponent of  $|x|$  being equal to  $\lambda - 1$  (instead of  $-3/2$ ), which is valid for all generalized hyperbolic laws, see Barndorff-Nielsen and Blaesild (1981). 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 an excellent fit of the NIG law to financial data, see e.g. Karlis (2002), Lillestøl (2001), Rydberg (1997) and 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 normal inverse Gaussian random variables with common parameters  $\alpha$  and  $\beta$  but having different scale and location parameters  $\delta_{1,2}$  and  $\mu_{1,2}$ , respectively, then  $X = X_1 + X_2$  is NIG( $\alpha, \beta, \delta_1 + \delta_2, \mu_1 + \mu_2$ ). This feature is especially useful in time scaling of risks, e.g. in deriving 10-day risks from daily risks. 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  $\delta$  is equal to 0. This is only possible for  $\lambda > 0$  and  $\alpha > |\beta|$ . The variance-gamma distributions (with  $\beta = 0$ ) were introduced to the financial literature by Madan and Seneta (1990).

### 3.1 Simulation of generalized hyperbolic variables

The most natural way of simulating generalized hyperbolic variables stems from the fact that they can be represented as normal variance-mean mixtures. Since the mixing distribution is the generalized inverse Gaussian law, the resulting algorithm reads as follows:

1. simulate a random variable  $Y \sim \text{GIG}(\lambda, \chi, \psi) = \text{GIG}(\lambda, \delta^2, \alpha^2 - \beta^2)$ ;
2. simulate a standard normal random variable  $N$ , e.g. using the Box-Muller algorithm, see Section 2.3;
3. return  $X = \mu + \beta Y + \sqrt{Y} N$ .

The algorithm is fast and efficient if we have a handy way of simulating generalized inverse Gaussian variates. Michael, Schucany and Haas (1976) proposed such a method for  $\lambda = -\frac{1}{2}$ , i.e. when sampling from the so-called inverse Gaussian (IG) distribution. The procedure utilizes a transformation that yields two roots. It starts with the observation that if we let  $\vartheta = \sqrt{\chi/\psi}$  then the  $\text{GIG}(-\frac{1}{2}, \chi, \psi) = \text{IG}(\chi, \psi)$  density, see eqn. (20), of  $Y$  can be written as:

$$f_Y(x) = \sqrt{\frac{\chi}{2\pi x^3}} e^{-\frac{1}{2}\chi(x-\vartheta)^2/x\vartheta^2}.$$

Hence:

$$V = \frac{\chi(Y - \vartheta)^2}{Y\vartheta^2}, \quad (33)$$

is distributed as a chi-square random variable with one degree of freedom, i.e.  $\chi^2(1)$ . 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. (33). Solving this equation for  $Y$  yields two roots:

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

According to Michael, Schucany and Haas (1976) the inverse Gaussian random variable  $Y$  can be simulated by choosing  $y_1$  with probability  $\vartheta/(\vartheta + y_1)$  and  $y_2$  with probability  $y_1/(\vartheta + y_1)$ , since  $Y$  has a representation as a mixture of the binomial distribution with two values ( $y_1$  and  $y_2$ ) and with  $\chi^2(1)$  as the mixing distribution. The `rndnig` quantlet of XploRe, as well as the `rnig` function of the Rmetrics collection of software packages for S-plus/R (see also Section 2.2 where Rmetrics was briefly described), utilize this routine.

In the general case, the GIG distribution – as well as the (generalized) hyperbolic law – can be simulated via the rejection algorithm. This method, in a version proposed by Atkinson (1982), is implemented in the `HyperbolicDist` package for S-plus/R developed by David Scott, see the R-project home page <http://cran.r-project.org/>. An adaptive version of the rejection method is used to obtain hyperbolic random numbers in the `rhyp` function of Rmetrics.

Ingenious ways of choosing the so-called hat functions (which provide an upper limit for the PDF of the sampled distribution) or squeeze functions (which provide upper and lower limits) are available. The proper choice of such functions can substantially increase the speed of computations, see also Chapter ?? where the rejection algorithm is discussed. Although the method works well for pseudo random numbers, it is not suited for the low discrepancy sequences (often called quasi random numbers, even though there is nothing random about them). The rejection method abandons some points from the sample, thereby destroying the structure of the sample. But often we want to use low discrepancy points since they promise faster convergence, at least for low dimensions.

A much more direct approach for sampling from a given distribution is the inversion algorithm, see Chapter ?? for an overview of this method. If we want to simulate generalized hyperbolic random numbers, all we have to do is to calculate (numerically) the inverse of the generalized hyperbolic CDF and apply it to a uniformly distributed sample. For example, the `rndhyp` quantlet of XploRe utilizes a binary search routine to obtain the inverse of the hyperbolic CDF. Another approach was proposed by Leobacher and Pillichshammer (2002) who suggested that the approximate inverse of the hyperbolic CDF be computed as the solution of a first-order differential equation.

### 3.2 Estimation of parameters

#### Maximum likelihood method

The parameter estimation of generalized hyperbolic distributions can be performed by the maximum likelihood 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  $\alpha$ -stable laws, but it still is considerable.

In general, the maximum likelihood estimation algorithm is as follows. For a vector of observations  $x = (x_1, \dots, x_n)$ , the ML estimate of the parameter vector  $\theta = (\lambda, \alpha, \beta, \delta, \mu)$  is obtained by maximizing the log-likelihood function:

$$L(x; \theta) = \log \kappa + \frac{\lambda - \frac{1}{2}}{2} \sum_{i=1}^n \log(\delta^2 + (x_i - \mu)^2) + \sum_{i=1}^n \log K_{\lambda - \frac{1}{2}}(\alpha \sqrt{\delta^2 + (x_i - \mu)^2}) + \sum_{i=1}^n \beta(x_i - \mu), \quad (34)$$

where  $\kappa$  is defined by (27). Obviously, for hyperbolic ( $\lambda = 1$ ) distributions the algorithm uses simpler expressions of the log-likelihood function due to relation (22).

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, see Blaesild and Sorensen (1992). In a large simulation study Prause (1999) utilized the bracketing method. The XploRe quantlets `mlehyp` and `mlebig` use yet another technique – the downhill simplex method of Nelder and Mead (1965), 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  $\kappa$ . For a data set with  $n$  independent observations we need to evaluate  $n$  and  $n + 1$  Bessel functions for NIG and generalized hyperbolic 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. For example, limited simulation studies performed in XploRe revealed a 25%, 55% and 85% reduction in CPU time for samples of size 500, 1000 and 2000, respectively.

We also have to say that the optimization is challenging. Some of the parameters are hard to separate since a flat-tailed generalized hyperbolic 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  $\alpha$  and  $\beta$  that can be used as starting values for the ML scheme. Starting from relation (32) for the tails of the NIG density they derived the following approximation:

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

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

where  $x_f$  is the  $f$ -th population quantile, see Section 2.4. After the choice of a suitable value for  $f$ , Venter and de Jongh (2002) used  $f = 5\%$ , the "tail estimates" of  $\alpha$  and  $\beta$  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 ( $\beta = \mu = 0$ ) generalized hyperbolic 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. Lillestöl (2001) made use of the Markov chain Monte Carlo technique (see Chapter ??), however, again the results obtained were not impressive. Karlis (2002) described an EM type algorithm (see Chapter ??) for maximum likelihood estimation of the normal inverse Gaussian 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. The EM scheme can be also generalized to the family of generalized hyperbolic distributions.

### 3.3 Are DJIA and DAX returns NIG distributed ?

It is always necessary to find a reasonable tradeoff between the introduction of additional parameters and the possible improvement of the fit. Barndorff-Nielsen and Blaesild (1981) mentioned the flatness of the likelihood function for the hyperbolic distribution. The variation in the likelihood function of the generalized hyperbolic distribution is even smaller for a wide range of parameters. Consequently, the generalized hyperbolic distribution applied as a model for financial data leads to overfitting (Prause, 1999). In the empirical analysis that follows we will concentrate on NIG distributions since they possess

**Table 4.** NIG and Gaussian fits to 2000 returns of the Dow Jones Industrial Average (DJIA) index from the period January 2, 1985 – November 30, 1992. Q: CSAfin08

| Parameters                | $\alpha$         | $\delta$ or $\sigma$ | $\beta$    | $\mu$  |
|---------------------------|------------------|----------------------|------------|--------|
| NIG fit ( $\delta$ )      | 68.0724          | 0.0074               | -1.8125    | 0.0008 |
| Gaussian fit ( $\sigma$ ) |                  | 0.0115               |            | 0.0006 |
| Test values               | Anderson-Darling |                      | Kolmogorov |        |
| NIG fit                   | 0.6724           |                      | 0.7216     |        |
| Gaussian fit              | +INF             |                      | 4.5121     |        |

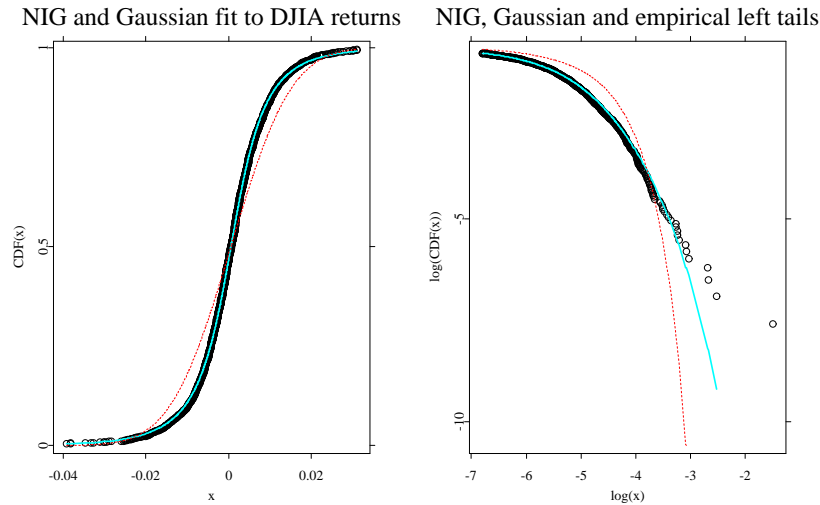
**Table 5.** NIG and Gaussian fits to 2000 returns of the Deutsche Aktienindex (DAX) index from the period January 2, 1995 – December 5, 2002. Q: CSAfin09

| Parameters                | $\alpha$         | $\delta$ or $\sigma$ | $\beta$    | $\mu$  |
|---------------------------|------------------|----------------------|------------|--------|
| NIG fit ( $\delta$ )      | 55.1771          | 0.0138               | -4.9035    | 0.0016 |
| Gaussian fit ( $\sigma$ ) |                  | 0.0157               |            | 0.0004 |
| Test values               | Anderson-Darling |                      | Kolmogorov |        |
| NIG fit                   | 0.3628           |                      | 0.7936     |        |
| Gaussian fit              | 16.4119          |                      | 2.8197     |        |

nice analytic properties and have been reported to fit financial data better than hyperbolic laws, see e.g. Karlis (2002), Lillestøl (2001) and Venter and de Jongh (2002).

Now, we can return to the empirical analysis. This time we want to check whether DJIA and/or DAX returns can be approximated by the NIG distribution. We fit the parameters using the maximum likelihood estimation technique as it seems to be the best available. As can be seen in Figure 10 the fitted NIG distribution misses the very extreme DJIA returns. However, it seems to give a better fit to the central part of the empirical distribution than the  $\alpha$ -stable law. This is confirmed by a lower value of the Kolmogorov statistics, compare Tables 2 and 4. Surprisingly, also the Anderson-Darling statistics returns a lower value, implying a better fit in the tails of the distribution as well.

The second analyzed data set comprises 2000 returns of the Deutsche Aktienindex (DAX) index. In this case the NIG law offers a much better fit than the  $\alpha$ -stable or the Gaussian, see Table 5 and compare with Table 3. This can be also seen in Figure 11. The drop off in the left tail of the empirical distribution is nicely caught by the NIG distribution.

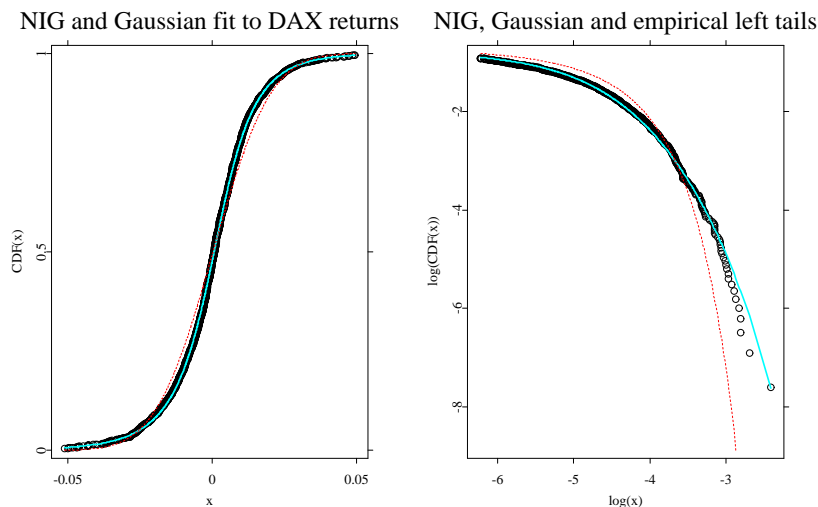


**Fig. 10.** NIG (cyan) and Gaussian (dashed red) fits to the DJIA returns (black circles) empirical cumulative distribution function from the period January 2, 1985 – November 30, 1992. For better exposition of the fit in the central part of the distribution ten largest and ten smallest returns are not illustrated in the left panel. The right panel is a magnification of the left tail fit on a double logarithmic scale. The NIG law slightly underfits the tails of the empirical distribution. Compare with Figure 6 where the stable law is shown to fit the DJIA returns very well. Q: CSAfin08

#### 4 Value at Risk, portfolios and heavy tails

The presented examples clearly show that we not only can, but must use heavy tailed alternatives to the Gaussian law in order to obtain acceptable estimates of market losses. But can we substitute the Gaussian distribution with other distributions in Value at Risk (Expected Shortfall) calculations? Recall, that the definition of VaR utilizes the quantiles of the portfolio returns distribution and not the returns distribution of individual assets in the portfolio. If all asset return distributions are assumed to be Gaussian then the portfolio distribution is multivariate normal and well known statistical tools can be applied. However, when asset returns are distributed according to a different law (or different laws!) then the multivariate distribution may be hard to tackle. In particular, linear correlation may no longer be a meaningful measure of dependence.

Luckily for us multivariate statistics offers the concept of copulas. In rough terms, a copula is a function  $C : [0, 1]^n \rightarrow [0, 1]$  with certain special properties, see Nelsen (1999) and Rank and Siegl (2002). What is important for VaR calculations is that it enables us to construct a multivariate distribution function



**Fig. 11.** NIG (cyan) and Gaussian (dashed red) fits to the DAX returns (black circles) empirical cumulative distribution function from the period January 2, 1995 – December 5, 2002. Right panel is a magnification of the left tail fit on a double logarithmic scale clearly showing the superiority of the NIG distribution. Compare with Figure 7 where the stable law is shown to overfit the DJIA returns. **Q: CSAfin09**

from the marginal distribution functions of individual asset returns in a way that takes their dependence structure into account. This dependence structure is no longer measured by correlation, but by other adequate functions like rank correlation, comonotonicity and, especially, tail dependence. For a treatment of VaR calculations, heavy tails and copulas consult also Bradley and Taqu (2003), Duffie and Pan (1997), Embrechts, McNeil and Straumann (2002), and Schmidt (2004).

## 5 Conclusions

In this chapter we have analyzed heavy tailed models of asset returns distributions. At first it seemed that the  $\alpha$ -stable distribution was a panacea. It had nice analytic properties (convolution, limiting behavior, excess kurtosis, etc.) and it fitted the DJIA returns almost perfectly. But later, as we tested the second sample (DAX returns), it turned out that the  $\alpha$ -stable distribution overfitted the tails. For a risk manager who likes to play safe this may not be a bad idea, as the stable law overestimates the risks and thus provides an upper limit of losses. But, in general, a distribution that fits the empirical data perfectly would be ideal. In the second part of the chapter we have seen



that there exist interesting heavy tailed alternatives to the stable distribution. The normal inverse Gaussian law gave a better fit to both data sets. Naturally, based on two examples we cannot conclude that the NIG distribution is the optimal one. However, we can say that there exist tractable heavy tailed alternatives to the Gaussian law that can be used in risk management, as well as in other branches of finance.

## References

- Akgiray, V. and Booth, G. G. (1988). The Stable-law Model of Stock Returns, *Journal of Business & Economic Statistics* **6**: 51–57.
- Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999). Coherent measures of risk, *Mathematical Finance* **9**: 203–228.
- 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.
- Basle Committee on Banking Supervision (1995). An internal model-based approach to market risk capital requirements, <http://www.bis.org>.
- Bergström, H. (1952). On some expansions of stable distributions, *Arkiv for Matematik* **II**: 375–378.
- Bibby, B. M. and Sørensen, M. (1997). A hyperbolic diffusion model for stock prices, *Finance & Stochastics* **1**: 25–41.
- 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.
- Blattberg, R. C. and Gonedes, N. J. (1974). A Comparison of the Stable and Student Distributions as Statistical Models of Stock Prices, *Journal of Business* **47**: 244–280.
- Borak, Sz., Härdle, W. and Weron, R. (2004). Stable Distributions in Finance, in P. Cizek, W. Härdle, R. Weron (eds.) *Statistical Tools for Finance and Insurance*, Springer.
- Bouchaud, J.-P. and Potters, M. (2000). *Theory of Financial Risk*, Cambridge University Press, Cambridge.
- Boyarchenko, S. I. and Levendorskii, S. Z. (2000). Option pricing for truncated Lévy processes, *International Journal of Theoretical and Applied Finance* **3**: 549–552.
- Bradley, B. O. and Taqqu, M. S. (2003). Financial Risk and Heavy Tails, in S. T. Rachev (ed.) *Handbook of Heavy-tailed Distributions in Finance*, North Holland.

- Brorsen, B. W. and Yang, S. R. (1990). Maximum Likelihood Estimates of Symmetric Stable Distribution Parameters, *Communications in Statistics – Simulations* **19**(4): 1459–1464.
- Campbell, J. B. (1980). *A FORTRAN IV subroutine for the modified Bessel functions of the third kind of real order and real argument*, Report NRC/ERB-925, National Research Council, Canada.
- 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.
- Cizek, P., Härdle, W. and Weron, R. (2004). *Statistical Tools for Finance and Insurance*, Springer. See also: <http://www.xplore-stat.de/ebooks/ebooks.html>.
- 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.
- 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.
- Devroye, L. (1986). *Non-Uniform Random Variate Generation*, Springer-Verlag, New York.
- Dowd, K. (2002). *Measuring Market Risk*, Wiley.
- Duffie, D. and Pan, J. (1997). An overview of value at risk, *Journal of Derivatives* **4**: 7–49.
- 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.
- Eberlein, E., Keller, U. and Prause, K. (1998). New insights into the smile, mispricing and Value at Risk: The hyperbolic model, *Journal of Business* **71**: 371–406.
- Embrechts, P., Kluppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*, Springer.
- Embrechts, P., McNeil, A. J. and Straumann, D. (2002). Correlation and Dependence in Risk Management: Properties and Pitfalls, in M. A. H. Dempster (ed.) *Risk Management: Value at Risk and Beyond*, Cambridge Univ. Press, Cambridge.
- Fama, E. F. (1965). The behavior of stock market prices, *Journal of Business* **38**: 34–105.
- Fama, E. F. and Roll, R. (1971). Parameter Estimates for Symmetric Stable Distributions, *Journal of the American Statistical Association* **66**: 331–338.
- Fofack, H. and Nolan, J. P. (1999). Tail Behavior, Modes and Other Characteristics of Stable Distributions, *Extremes* **2**: 39–58.
- Franke, J., Härdle, W. and Stahl, G. (2000). *Measuring Risk in Complex Stochastic Systems*, Springer. See also: <http://www.xplore-stat.de/ebooks/ebooks.html>.

- Gopikrishnan, P., Meyer, M., Amaral, L. A. N. and Stanley, H. E. (1998). Inverse cubic law for the distribution of stock price variations, *European Physical Journal B* **3**: 139–140.
- 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.
- Härdle, W., Klinke, S. and Müller, M. (2000). *XploRe Learning Guide*, Springer. See also: <http://www.xplore-stat.de/ebooks/ebooks.html>.
- Härdle, W., Kleinow, T. and Stahl, G. (2002). *Applied Quantitative Finance*, Springer. See also: <http://www.xplore-stat.de/ebooks/ebooks.html>.
- Hill, B. M. (1975). A Simple General Approach to Inference About the Tail of a Distribution, *Annals of Statistics* **3**: 1163–1174.
- Holt, D. R. and Crow, E. L. (1973). Tables and graphs of the stable probability density functions, *Journal of Research of the National Bureau of Standards B* **77B**: 143–198.
- Janicki, A. and Kokoszka, P. (1992). Computer investigation of the rate of convergence of LePage type series to alpha-stable random variables, *Statistica* **23**: 365–373.
- Janicki, A. and Weron, A. (1994a). Can one see  $\alpha$ -stable variables and processes, *Statistical Science* **9**: 109–126.
- Janicki, A. and Weron, A. (1994b). *Simulation and Chaotic Behavior of  $\alpha$ -Stable Stochastic Processes*, Marcel Dekker.
- Jorion, P. (2000). *Value at Risk: The New Benchmark for Managing Financial Risk*, McGraw-Hill.
- Karlis, D. (2002). An EM type algorithm for maximum likelihood estimation for the Normal Inverse Gaussian distribution, *Statistics and Probability Letters* **57**: 43–52.
- Khindarova, I., Rachev, S. and Schwartz, E. (2001). Stable Modeling of Value at Risk, *Mathematical and Computer Modelling* **34**: 1223–1259.
- 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.
- Küchler, U., Neumann, K., Sørensen, M. and Streller, A. (1999). Stock returns and hyperbolic distributions, *Mathematical and Computer Modelling* **29**: 1–15.
- Laha, R. G. and Rohatgi, V. K. (1979). *Probability Theory*, Wiley.
- Leobacher, G. and Pillichshammer, F. (2002). A Method for Approximate Inversion of the Hyperbolic CDF, *Computing* **69**: 291–303.
- LePage, R., Woodroffe, M. and Zinn, J. (1981). Convergence to a stable distribution via order statistics, *Annals of Probability* **9**: 624–632.
- Lévy, P. (1925). *Calcul des Probabilités*, Gauthier Villars.
- Lillestøl, J. (2001). *Bayesian Estimation of NIG-parameters by Markov chain Monte Carlo Methods*, Discussion paper 2001/3, Department of Finance and Management Science, The Norwegian School of Economics and Business Administration.

- 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. (1994). Fast, accurate algorithm for numerical simulation of Levy stable stochastic processes, *Physical Review E* **49**: 4677–4683.
- 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.
- Mantegna, R. N. and Stanley, H. E. (1995). Scaling behavior in the dynamics of an economic index, *Nature* **376**: 46–49.
- Matacz, A. (2000). Financial Modeling and Option Theory with the Truncated Lévy Process, *International Journal of Theoretical and Applied Finance* **3(1)**: 143–160.
- McCulloch, J. H. (1986). Simple Consistent Estimators of Stable Distribution Parameters, *Communications in Statistics – Simulations* **15**: 1109–1136.
- McCulloch, J. H. (1996). Financial Applications of Stable Distributions, in G. S. Maddala, C. R. Rao (eds.), *Handbook of Statistics, Vol. 14*, Elsevier, pp. 393–425.
- McCulloch, J. H. (1997). Measuring Tail Thickness to Estimate the Stable Index  $\alpha$ : A Critique, *Journal of Business & Economic Statistics* **15**: 74–81.
- McCulloch, J. H. (1998). Kogon, S. M. and Williams, D. B. (1998). Numerical Approximation of the Symmetric Stable Distribution and Density, in R. Adler, R. Feldman, M. Taqqu (eds.), *A Practical Guide to Heavy Tails*, Birkhauser, pp. 489–500.
- Michael, J. R., Schucany, W. R. and Haas, R. W. (1976). Generating Random Variates Using Transformations with Multiple Roots, *The American Statistician* **30**: 88–89.
- 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., Rachev, S. T., Doganoglu, T. and Chenyao, D. (1999). Maximum Likelihood Estimation of Stable Paretian Models, *Mathematical and Computer Modelling* **29**: 275–293.
- Nelder, J. A. and Mead, R. (1965). A Simplex Method for Function Minimization, *The Computer Journal* **7**: 308–313.
- Nelsen, R. B. (1999). *An Introduction to Copulas*, Springer, New York.
- Nolan, J. P. (1997). Numerical Calculation of Stable Densities and Distribution Functions, *Communications in Statistics – Stochastic Models* **13**: 759–774.
- Nolan, J. P. (1999). An Algorithm for Evaluating Stable Densities in Zolotarev’s (M) Parametrization, *Mathematical and Computer Modelling* **29**: 229–233.
- Nolan, J. P. (2001). Maximum Likelihood Estimation and Diagnostics for Stable Distributions, in O. E. Barndorff-Nielsen, T. Mikosch, S. Resnick (eds.), *Lévy Processes*, Birkhäuser, Boston.
- Officer, R. R. (1972). The Distribution of Stock Returns, *Journal of the American Statistical Association* **67**: 807–812.
- Ojeda, D. (2001). *Comparison of stable estimators*, Ph.D. Thesis, Department of Mathematics and Statistics, American University.

- 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.
- Press, W., Teukolsky, S., Vetterling, W. and Flannery, B. (1992). *Numerical Recipes in C*, Cambridge University Press. See also: <http://lib-www.lanl.gov/numerical/index.html>.
- Rachev, S. and Mittnik, S. (2000). *Stable Paretian Models in Finance*, Wiley.
- Rank, J. and Siegl, T. (2002). Applications of Copulas for the Calculation of Value-at-Risk, in Härdle, Kleinow and Stahl (2002).
- Rydberg, T. H. (1997). The Normal Inverse Gaussian Lévy Process: Simulation and Approximation, *Communications in Statistics – Simulations* **13(4)**: 887–910.
- Samorodnitsky, G. and Taqqu, M. S. (1994). *Stable Non-Gaussian Random Processes*, Chapman & Hall.
- Schmidt, R. (2004). Tail dependence, in P. Cizek, W. Härdle, R. Weron (eds.) *Statistical Tools for Finance and Insurance*, Springer.
- 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.
- Teichmoeller, J. (1971). A Note on the Distribution of Stock Price Changes, *Journal of the American Statistical Association* **66**: 282–284.
- Temme, N. M. (1975). On the numerical evaluation of the modified Bessel function of the third kind, *Journal of Computational Physics* **19**: 324–337.
- 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, *Correction to: On the Chambers-Mallows-Stuck Method for Simulating Skewed Stable Random Variables*, Research Report HSC/96/1, Wrocław University of Technology, 1996, <http://www.im.pwr.wroc.pl/~hugo/Publications.html>.
- 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.
- 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.

---

## Index

- Conditional VaR, 2
- Deutsche Aktienindex (DAX), 18, 29
- distribution
  - $\alpha$ -stable, 4, 5
    - $S$  parametrization, 6
    - $S^0$  parametrization, 7
  - characteristic function, 6
  - density function, 8
  - direct integration method, 8
  - distribution function, 8
  - Fama-Roll method, 12
  - maximum likelihood estimation, 15
  - method of moments, 13
  - regression method, 14
  - simulation, 10
  - STABLE program, 9, 17
- Cauchy, 6
- Gaussian, 6
- generalized hyperbolic
  - density function, 24
  - maximum likelihood estimation, 27
  - mean, 24
  - simulation, 25
  - variance, 24
- generalized inverse Gaussian (GIG),
  - 22
  - simulation, 25, 26
- hyperbolic, 21, 22
  - density function, 23
  - simulation, 26
- Lévy, 6
- Lévy stable, 4
- normal inverse Gaussian (NIG)
  - density function, 24
  - tail estimates, 28
  - tails, 25
  - stable Paretian, 4
  - truncated stable (TLD), 20
    - characteristic function, 21
- Dow Jones Industrial Average (DJIA),
  - 2, 3, 18, 29
- Expected Shortfall (ES), 2
- Expected Tail Loss (ETL), 2
- market risk, 1
- modified Bessel function, 22
- quantlets, 4
  - cdfhyp, 23
  - cdfnig, 24
  - cdfstab, 9
  - CSAfin01, 2
  - CSAfin02, 3
  - CSAfin03, 15
  - CSAfin04, 16
  - CSAfin05, 18, 19
  - CSAfin06, 19, 20
  - CSAfin07, 23
  - CSAfin08, 29, 30
  - CSAfin09, 29, 31
  - mlehyp, 27
  - mlebig, 27
  - normal2, 11
  - pdfhyp, 23
  - pdfnig, 24
  - pdfstab, 9

rndhyp, 26  
rndnig, 26  
rndstab, 11  
stabcull, 13  
stabmom, 13  
stabreg, 15, 17  
STFstab01, 5  
STFstab02, 6

STFstab04, 7

risk measure, 2

Value at Risk (VaR), 1, 4  
copulas, 30

XploRe, 4, 9, 24

