

Stable Distributions: a survey on simulation and  
calibration methodologies

# Contents

0.1	Introduction . . . . .	1
<b>1</b>	<b>An overview on calibration methods and models for heavy tail series</b>	<b>2</b>
1.1	Empirical Characteristic Function(ECF) based methods . . . . .	2
1.2	Quantile based methods . . . . .	4
<b>2</b>	<b>Some calibration methods for Stable Distributions</b>	<b>7</b>
2.1	Introduction to Stable Distribution . . . . .	7
2.1.1	Why stable distributions? . . . . .	7
2.1.2	Definition and properties of Stable Distributions . . . . .	7
2.2	Weron's exact simulation method . . . . .	10
2.2.1	Description . . . . .	10
2.2.2	Theoretical Basement . . . . .	10
2.2.3	Algorithm for generate stable numbers . . . . .	10
2.2.4	Numerical Implementation and results . . . . .	11
2.2.5	Advantages and Drawbacks . . . . .	12
2.3	Tail estimators . . . . .	13
2.3.1	Tail Regression . . . . .	13
2.3.2	Hill estimator and their variants . . . . .	16
2.4	Quantile based calibration method . . . . .	18
2.4.1	McCulloch Method . . . . .	18
2.4.2	<i>L-Moments Method</i> . . . . .	21
2.5	Maximum Likelihood Methods . . . . .	24
2.5.1	Zolotarev's Integral Representation Method . . . . .	24
2.5.2	Monte Carlo Markov Chain Simulated Annealing Method . . . . .	28
2.6	Sample characteristic function methods . . . . .	30
2.6.1	Minimum Distance Method . . . . .	30
2.6.2	Regression Method . . . . .	32
2.6.3	Method of Moments . . . . .	34
2.7	Appendix . . . . .	38

## 0.1 Introduction

It is well known that the Gaussian assumption on market data is not supported by empirical evidence. Particularly the presence of skewness and a kurtosis larger than in the normal law, can affect in a dramatically way the Risk Managing Analysis, specially, the VaR calculation through quantile estimators. Therefore the necessity of searching for other distribution classes explaining asymmetry and heavy tail phenomena. In this sense stable, generalized hyperbolic and gaussian mixing distribution have been used with considerable success.

On the other hand, heavy tails affect standard calibration and model testing procedures, due to the frequent presence of outliers, calling for more robust methods .

The first part of the report is a review on some methods for calibration models with robust properties.

In particular, minimum distance methods using characteristic functions and quantile based methods are studied.

The second part focuses on Stable Distributions, for which a simulation method has been studied and implemented . Also, eight calibration methods have been compared from simulated and financial data and a resulting methodology is established.

Simulation and calibration methods were implemented in C++ for PC. An oriented-object system was created, allowing to perform simulations, calibrations and model test. Some graphic facilities are implemented.

# Chapter 1

## An overview on calibration methods and models for heavy tail series

### 1.1 Empirical Characteristic Function(ECF) based methods

In recent years, alternative methods to the maximum likelihood approach have been used by practitioners in the finance field.

Despite its generality and well known asymptotic properties as consistency, normality and efficiency, likelihood function can be no tractable in many situations due to its boundless over the parametric space, instabilities, numerical inaccuracies in the evaluation of the likelihood function or the existence of many local maximums.

The problem is found very often in models describing financial data, for example the mixture of Normal Distributions, Switching Models, Variance Gamma Distributions, Stable Distributions and Stable ARMA Models, Diffusion-Jumps and Stochastic Volatility continuous time models among others.

Although these difficulties are partially solved inside the likelihood framework, using alternative methods like *Simulated Annealing Monte Carlo Markov Chain* to obtain global maximum (see 2.5.2 below or [2],[14]), or by establishing more robust procedures like using a multinomial likelihood function instead, see [3], still in many situations m.l. method leads to non satisfactory situations .

Moreover, for some class of distribution families (see section 2.1 below), like Stable Distributions, used in the modeling of heavy tails phenomena, there is not a closed form for the density function.

On the contrary, the characteristic function(CF) is always bounded and is available in a simpler form than the density in some important cases.

The main idea consists in minimize the distance between the CF and the empirical characteristic function(ECF) in an appropriate norm. The motivation of the method lies in the one to one correspondence between distribution function and their corresponding Fourier-Stieltjes transform.

The ECF procedure for i.i.d. random variables draw back to [32] and [20]. For more recent developments we have [38],[4].

While the minimization procedure imply a lot of calculations, some simpler variants, exploiting particular relations derived from CF expression in Stable Distributions, have been used, see section 2.6 below.

In [39] it can be found a comparison of several methods (see also tables 5-? in this report).

The general method can be described in the following way:

Denote by  $\phi$  the characteristic function and by

$$\hat{\phi}(t) = \frac{1}{n} \sum_{j=1}^n \exp itX_j$$

the empirical characteristic function.

Since  $|\hat{\phi}(t)|$  is bounded all moments of  $\hat{\phi}(t)$  are finite for any fixed  $t$ . By the Law of Large Numbers  $\hat{\phi}(t)$  is a consistent estimator of  $\phi(t)$ .

The method finds

$$\hat{\theta} = \operatorname{argmin}_{\theta \in \Theta} \|\phi(t) - \hat{\phi}(t)\|$$

where  $\theta$  is a point in the parametric space  $\Theta$  and  $\|\cdot\|$  is a norm, usually the  $L^\infty$  or a  $L^r$  weighted norms.

This last is more useful for implementation and it can be written as:

$$h(\theta) = \int_{-\infty}^{\infty} |\phi(t) - \hat{\phi}(t)|^r W(t) dt \quad (1.1)$$

where  $W(t)$  is some weight function.

More generally:

$$h(\theta) = \int_{-\infty}^{\infty} |\phi(t) - \hat{\phi}(t)|^r dG(t). \quad (1.2)$$

Here  $G(t)$  is some distribution function.

The maximization procedure is performed on a finite interval, containing then sample values.

When  $G(t)$  is an step function and  $r = 2$  the right side of (1.2) becomes:

$$\sum_{i=1}^p |\phi(t_i) - \hat{\phi}(t_i)|^2 g(t_i)$$

and the procedure is similar to the weighted least square.

Optimal selection of discrete points  $t_1, t_2, \dots, t_p$  is discussed in [36],[22] and [4].

Another possibility is to choose a function  $G(t)$  with density  $g(t)$  with respect to the Lebesgue measure, typically the exponential law. This selection has an advantage related to numerical quadrature optimization procedures, also it gives more importance to the CF in a neighborhood of zero.

Another advantage of ECF methods is that they can be extended to the non i.i.d case, particularly to dynamic models with heteroscedastic volatility, by considering a multivariate

or conditional CF instead. Asymptotic consistency and normality are conserved([11],[19])in this case.

Also validation test based on ECF have been proposed in [1] for multivariate data while attempts to extend Kolmogov-Smirnov and Cramer-Von Mises test reveal an unbearable computational burden for dimension greater than two.

## 1.2 Quantile based methods

Quantile methods to measure the risk in finance context are usually linked to Value-at-Risk measures(VaR).

Beyond the Gaussian assumption, which reveals non adequacy to explain empirical data, several methods have been proposed to estimate the quantiles.

Some methods point out to fit a distribution law to empirical data, then to estimate quantiles under an appropriate inverse transformation. Extreme Value Theory can be useful in this direction([9], [37]).

A non parametric approach is used, without any assumption about distribution law, by estimating quantiles from the sample order statistics. While consistent, these estimators reveal lack of accuracy concerning VaR calculation. The problem have been studied in [25] and [24].

In order to overcome this difficulty a new approach has been proposed consisting in a weighted sum or sample quantiles instead of a single one, making the estimation more robust.

In this case the quantile at level  $\alpha$  is estimated by:

$$Q_\alpha = \sum_{i=1}^r w_{i,k} X_{n:k}$$

and

$$\sum_{i=1}^r w_{i,k} = 1$$

where  $X_{n:k}$  represents the k-th sample order statistic.

Weights are chosen according to some criteria, for example kernels are used as weights in order to smooth the sample distribution within a given windows([35]).

In this case consistency and its rate of convergence, as well as asymptotic normality, are nice properties of this estimators.

Another variant for the weights leads to the the Harrell-Davis estimator,[15], which is nothing but the bootstrap estimator of the mean value for the sample quantile.

On the other hand, robust quantile estimators can be used to estimate parameters of a given distribution law. One of this methods is the *L-moments* consisting in matching sample weighted quantiles to the theoretical ones.

In this sense *L-moments* is an alternative way to describe the shape of a probability distribution. They are based on the probability weighted moment, which is an attempt to make method of moments more robust by smoothing the tail distribution.

According to Greenwood([13])the theoretical L-moments are given, for a random variable X

with CDF  $F$ , by the weighted quantiles:

$$M_{p,r,s} = E\{X^p[F(X)]^r[1 - F(X)]^s\}$$

Of particular interest result the cases  $\alpha_r = M_{1,0,r}$  and  $\beta_r = M_{1,r,0}$  which can be written as

$$\alpha_r = \int_0^1 Q(u)(1-u)^r du \quad (1.3)$$

$$\beta_r = \int_0^1 Q(u)u^r du \quad (1.4)$$

where  $Q$  is the quantile function.

According to the last expression, *L-moments* can be view as an the conventional moments weighted by polynomials  $u$  or  $1 - u$ .

*L-moments* themselves are difficult to interpret, however, certain linear combination of them can be interpreted in terms of location parameters,  $\lambda_1 = \alpha_0$ , scale parameter,  $\lambda_2 = \alpha_0 - 2\alpha_1$  or  $2\beta_1 - \beta_0$ , skewness,  $\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0$  and kurtosis,  $\lambda_4 = \alpha_0 - 12\alpha_1 + 30\alpha_2 - 20\alpha_3$ , see [13] for details.

An intuitive justification for these parameters can be obtained considering the order statistics  $X_{1:n}, X_{2:n}, \dots, X_{n:n}$  and noting that:

$$\begin{aligned} \lambda_1 &= E(X_{1:1}) \\ \lambda_2 &= E(X_{2:2} - X_{1:2}) \\ \lambda_3 &= E(X_{3:3} - 2X_{2:3} + X_{1:3}) \\ \lambda_4 &= E((X_{4:4} - 3X_{3:4} + 3X_{2:4}) - X_{1:4}) \end{aligned}$$

Natural estimators for the four parameters can be obtained from sample data in the following way:

An unbiased estimator for  $\beta_r$  is;

$$b_r = n^{-1} \binom{n-1}{r} \sum_{j=r+1}^n \binom{j-1}{r} X_{j:n} \quad (1.5)$$

which leads to the sample L-moments estimators

$$l_r = \sum_{k=0}^{r-1} p_{r-1,k}^* b_k = n^{-1} \sum_{j=1}^n w_{j:n}^{(r)} X_{j:n} \quad (1.6)$$

for  $r = 1, 2, 3, 4$ . The expressions for  $p_{r,k}^*$  are derived from Legendre polynomials, then the weights can be calculated easily.

Ordinary moments put a greater weight in the tails of the distribution, therefore they are more affected by heavy tail phenomena by opposite to L-moments where the smaller weights on the tail distribution make them more robust.

Moreover, *L-moments* exist when the mean distribution is finite.

For some distributions arising in extreme value theory or with heavy tail behavior, *L-moments* have been calculated as function of their parameters, for ex. generalized Pareto and Gumbel distributions.

In [18], estimations for historical series of exchanges and interest rates are compared, see also [27].

Within a non-parametric approach [16] propose the use of *L-moments* to recover the CDF itself using maximum entropy techniques.

For stable distributions, the exact relationship between parameters and theoretical *L-moments* can not be explicitly calculate, nevertheless it is possible to follow a numerical approach to solve the corresponding equations as result of equate sample and theoretical *L-moments*. This procedure involves four approximations hence numerical calculations should be follow with extreme care, see section 2.4.2 below for details.

Also for stable distributions, another different procedure for calibration is proposed by McCulloch, based on quantiles. Functions of the quantiles in the tails are found to be independent on the scale and position parameters, then a numerical function is constructed to estimate the tail and skewness parameters. Once these two parameters are established, another two other functions on quantiles allow to estimate scale and position parameters.

The procedure gives reasonable good results and numerical calculations are fast once a table with these functions are pre-calculated(see section 2.4 below for details).

# Chapter 2

## Some calibration methods for Stable Distributions

### 2.1 Introduction to Stable Distribution

#### 2.1.1 Why stable distributions?

Stable distributions are one of the several possible attempts to explain empirical evidence in the financial markets that cannot be described assuming a Gaussian distribution of returns, as the heavy tail phenomena already mentioned.

In the 60's by Mandelbrot([23]) and Fama([10]) proposed the  $\alpha$ -stable distribution the basis of their attractive properties.

The family of stable distributions not only describes heavy tails and asymmetric behavior but also, the dependency on four parameters, instead for two parameters in the Gaussian case, makes them more flexible to adapt empirical data for calibration and model testing.

Another nice property is that they have domain of attraction, i.e. limits of sums of i.i.d stable random variables are also stables, making them robust models, i.e. an slight modification in the data will produce a distribution law that still lies in the family.

Moreover, it is possible to consider stable processes to describe the random changes in time. Stable distributions have been extended also to a multivariate framework, allowing to express a more complex kind of dependency among variables than covariances in the Gaussian context.

Also, option pricing theory can be set by assuming an stable process.

Several methods for calibration and simulation of stable laws are now available and they have been tested and compared successfully.

#### 2.1.2 Definition and properties of Stable Distributions

In this section we introduce the stable distribution, some of its properties, different parameterizations and the simulation technique.

The following conditions are equivalent definitions of a stable random variable  $X$ :

- a) Let  $a, b$  two real positive numbers and  $X_1$  and  $X_2$  be independent random variables equally distributed to  $X$ . There exist  $c \in \mathbf{R}^+$  and  $d \in \mathbf{R}$  such that:

$$aX_1 + bX_2 = cX + d \text{ in distribution}$$

- b) Let  $n$  be a positive integer,  $n \geq 2$ , and  $X_1, X_2, \dots, X_n$  be independent copies of  $X$ . There exist  $c_n \in \mathbf{R}^+$  and  $d_n \in \mathbf{R}$  such that

$$X_1 + X_2 + \dots + X_n = n^{\frac{1}{\alpha}}X + \mu(n - n^{\frac{1}{\alpha}}) \text{ if } \alpha \neq 1$$

and

$$X_1 + X_2 + \dots + X_n = nX + \frac{2}{\pi}\sigma\beta n \log n \text{ if } \alpha = 1$$

- c)  $X$  has a domain of attraction, i.e., there are a sequence of i.i.d. random variables  $\{Y_i\}_{i \in \mathbf{N}}$ , a real positive sequence  $\{a_i\}_{i \in \mathbf{N}}$  and a real sequence  $\{b_i\}_{i \in \mathbf{N}}$  such that

$$\frac{1}{a_n} \sum_{i=1}^n Y_i - b_n \rightarrow X \text{ in distribution}$$

- d) The characteristic function of  $X$  admits the following form:

$$\phi_X t = \begin{cases} \exp(-\sigma^\alpha |t|^\alpha (1 - i\beta \operatorname{sgn}(t) \tan \pi\alpha/2) + i\mu t) & \text{for } \alpha \neq 1 \\ \exp(-\sigma |t| (1 + i\beta \frac{\pi}{2} \operatorname{sgn}(t) \log |t|) + i\mu t) & \text{for } \alpha = 1 \end{cases} \quad (2.1)$$

where the parameters satisfy the constraints  $\alpha \in (0, 2], \sigma \in \mathbf{R}_0^+, \beta \in [-1, 1], \mu \in \mathbf{R}$ .

A random variable  $X$  with stable distribution of parameters  $\alpha, \beta, \sigma, \mu$  is noted as  $X \sim S(\alpha, \beta, \sigma, \mu)$ .

**Theorem 2.1.1** *The following properties hold for an stable distribution:*

- i) Let  $X_1$  and  $X_2$  independent stable random variables, with  $X_i \sim S(\alpha_i, \beta_i, \sigma_i, \mu_i)$ , for  $i = 1, 2$  then  $X_1 + X_2 \sim S(\alpha, \beta, \sigma, \mu)$  with

$$\sigma = (\sigma_1^\alpha + \sigma_2^\alpha)^{\frac{1}{\alpha}}$$

$$\beta = \frac{\beta_1 \sigma_1^\alpha + \beta_2 \sigma_2^\alpha}{\sigma_1^\alpha + \sigma_2^\alpha}$$

and

$$\mu = \mu_1 + \mu_2$$

ii)  $X \sim S(\alpha, \beta, \sigma, \mu)$  and  $a \in \mathbf{R}$  then  $X + a \sim S(\alpha, \beta, \sigma, \mu + a)$ ,

iii)  $X \sim S(\alpha, \beta, \sigma, \mu)$  and  $a \in \mathbf{R}$ .

Also:

$$aX \sim \begin{cases} S(\alpha, \operatorname{sgn}(\alpha)\beta, |\alpha|\sigma, a\mu) & \text{for } \alpha \neq 1 \\ S(1, \operatorname{sgn}(\alpha)\beta, |\alpha|\sigma, a\mu - \frac{2}{\pi}a(\log|a|)\sigma\beta) & \text{for } \alpha = 1 \end{cases} \quad (2.2)$$

iv)  $\alpha \in (0, 2)$  and  $X \sim S(\alpha, \beta, \sigma, 0)$  then  $-X \sim S(\alpha, -\beta, \sigma, 0)$ .

v) For  $\alpha \in (0, 2)$  and  $X \sim S(\alpha, \beta, \sigma, 0)$  then

$$\lim_{x \rightarrow \infty} x^\alpha P(X > x) = C_\alpha(1 + \beta)\sigma^\alpha \quad (2.3)$$

where  $C_\alpha$  is a constant depending on  $\alpha$ .

vi) For  $\alpha \in (0, 2)$  and  $X \sim S(\alpha, \beta, \sigma, 0)$  then  $E|X|^p$  exists for  $0 < p < \alpha$  and  $E|X|^p = \infty$  for  $p \geq \alpha$ .

vii) Let  $(X_t)_{t \in \mathbf{N}}$  an stochastic process with symmetrical stable finite-dimensional distributions and tail index  $\alpha$  then for every  $c > 0$  and  $t_1 < t_2 < \dots < t_n$ :

$$(X_{ct_1}, X_{ct_2}, \dots, X_{ct_n}) = c^{\frac{1}{\alpha}}(X_{t_1}, X_{t_2}, \dots, X_{t_n}) \text{ in distribution} \quad (2.4)$$

Note that properties ii)-v) allow interpretation of  $\alpha$  as a tail parameter,  $\beta$  as a coefficient of skewness,  $\sigma$  as an scale parameter and  $\mu$  as a location parameter.

Properties v) and vi) express the heavy tail behavior and property vii) a self-similarity behavior.

Note also that Gaussian law is included in the family. The CF given in (2.1) is not continuous with respect to parameters at  $\alpha = 1$ , then the re-parametrization proposed by Zolotarev([40]) is often used:

$$\Phi_X(t) = \begin{cases} \exp(-\sigma_2^\alpha |t|^\alpha \exp(-i\beta_2 \operatorname{sgn}(t) \frac{\pi}{2} K(\alpha)) + i\mu t) & \text{for } \alpha \neq 1 \\ \exp(-\sigma_2 |t|(\frac{\pi}{2} + i\beta_2 \operatorname{sgn}(t) \ln |t|)) & \text{for } \alpha = 1 \end{cases} \quad (2.5)$$

where

$$K(\alpha) = \alpha - 1 + \operatorname{sgn}(1 - \alpha) = \begin{cases} \alpha & \text{for } \alpha \leq 1 \\ \alpha - 2 & \text{for } \alpha \geq 1 \end{cases}$$

$$\sigma_2 = \sigma(1 + \beta^2 \tan^2 \frac{\pi\alpha}{2})^{\frac{1}{2\alpha}}$$

$$\tan(\beta_2 \frac{\pi K(\alpha)}{2}) = \beta \tan \frac{\pi\alpha}{2} \text{ for } \alpha \neq 1.$$

For  $\alpha = 1$  we have:

$$\sigma_2 = \frac{2}{\pi}\sigma \text{ and } \beta_2 = \beta$$

The tail index and the location parameter coincide with the standard parametrization.

## 2.2 Weron's exact simulation method

### 2.2.1 Description

In order to test calibration methods and to generate future scenarios it is necessary to have a fast Monte Carlo generator of pseudo-random stable distributed numbers.

The main difficulty is the non existence of explicit density functions, with the exception of a few particular cases, thus classical methods, as for example the inverse transform method, don't apply.

In [5] a method was developed based on an integral representation of the stable density function following, see [40] representation.

More recently in [39] it was constructed an exact technique to this end by mean of a suitable non-linear transform of a pair of independently uniform and exponentially random variable.

### 2.2.2 Theoretical Basement

Let  $X \sim S(\alpha, \mu, \beta, \sigma)$  an stable random variable with characteristic function given by (2.5). The result is the following:

**Theorem 2.2.1** (Weron(1996)) *Let  $U$  a uniform random variable on  $(-\frac{\pi}{2}, \frac{\pi}{2})$ ,  $E$  an exponential random variable with mean 1 independent of  $U$ , and*

$$U_0 = -\frac{\pi}{2}\beta_2 \frac{K(\alpha)}{\alpha}$$

then the random variable:

$$X = \begin{cases} \frac{\sin\alpha(U-U_0)}{(\cos U)^{\frac{1}{\alpha}}} \left( \frac{\cos(U-\alpha(U-U_0))}{E} \right)^{\frac{1-\alpha}{\alpha}} & \text{for } \alpha \neq 1 \\ (\frac{\pi}{2} + \beta_2 U) \tan U - \beta_2 \ln\left(\frac{E \cos U}{\frac{\pi}{2} + \beta_2 U}\right) & \text{for } \alpha = 1 \end{cases} \quad (2.6)$$

### 2.2.3 Algorithm for generate stable numbers

From the theorem above we establish the following algorithm:

**Algorithm**

1. Generate  $U_1 \sim (-\frac{\pi}{2}, \frac{\pi}{2})$  and  $U_2 \sim U(0, 1)$ .
2.  $E = -\ln U_2$  (by transform inverse method  $E \sim \exp(1)$ ).
3. If  $\alpha \neq 1$  compute:
  - a)  $B(\alpha, \beta) = \frac{\arctan\left(\beta \tan \frac{\pi\alpha}{2}\right)}{\alpha}$

b)  $S(\alpha, \beta) = (1 + \beta^2 \tan^2 \frac{\pi\alpha}{2})^{\frac{1}{2\alpha}}$   
c)

$$X = S(\alpha, \beta) \frac{\sin \alpha (U + B(\alpha, \beta))}{(\cos U)^{\frac{1}{\alpha}}} \left( \frac{\cos (U - \alpha(U + B(\alpha, \beta)))}{E} \right)^{\frac{1-\alpha}{\alpha}}$$

4. If  $\alpha = 1$  compute:

$$X = \frac{2}{\pi} \left( \frac{\pi}{2} + \beta U \right) \tan U - \beta \ln \frac{\frac{\pi}{2} E \cos U}{\frac{\pi}{2} + \beta U}$$

5. Set

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

## 2.2.4 Numerical Implementation and results

The algorithm is linear time with speed  $O(N)$ .

Possible overflow problems, as consequence of division by zero in (2.6) can be avoided by truncating. It doesn't constitute a serious approximation problem due to the double arithmetic precision used in the implementation in C++.

Also for  $\alpha$  close to zero the same problem can be present. Again most applications consider  $\alpha$  greater then one.

In Figures 2.1, 2.5 and 2.3 histograms for simulated data for a given set of parameters are shown, they are compared with the respective theoretical functions obtained by inverting the CF with the same parameters.

In figure 2.4 the difference between respective distribution functions is showed together with the difference with Gaussian distribution.

Figure 2.1: Empirical simulated density function (solid line) and theoretical calculated density function (dashed line) for an stable distribution with  $\alpha = 1.5$ ,  $\beta = 0$ ,  $\sigma = 1$  and  $\mu = 0$

Figure 2.2: Empirical simulated distribution function(solid line) and theoretical calculated distribution function(dash line) for an stable distribution with  $\alpha = 1.5$ ,  $\beta = 0$ ,  $\sigma = 1$  and  $\mu = 0$

Figure 2.3: Empirical simulated quantile function(solid line)and theoretical calculated density function(dash line) for an stable distribution with  $\alpha = 1.5$ ,  $\beta = 0$ ,  $\sigma = 1$  and  $\mu = 0$

Note that the differences between both functions are small. For other set of parameters the results are similar.

In Tables 2.1 and 2.2 results of a Kolmogov-Smirnov test applied to simulated data are shown. There, 100 000 simulated groups of data of sizes 250, 500 and 1000 were tested. The fraction of rejected data is well below the significance level.

## 2.2.5 Advantages and Drawbacks

The advantage is a direct and simple to implement method as opposite to former recurrence methods.

Figure 2.4: Differences between empirical and inverted theoretical distribution functions (solid line) and differences between empirical calculated distribution function with Gaussian distribution function (dashed line) for a stable distribution with  $\alpha = 1.5$ ,  $\beta = 0$ ,  $\sigma = 1$  and  $\mu = 0$

## 2.3 Tail estimators

### 2.3.1 Tail Regression

#### Description

It is a tail based estimator method for index  $\alpha$ . From the known result about the decay of tails in stable distributions as Pareto tails with a power given by index  $\alpha$ , see 2.3, a linear regression is implemented in logarithm scale.

For  $x$  large enough and taking logarithms on both sides we have:

$$\log P(X > x) = C_\alpha(1 + \beta)\sigma^\alpha - \alpha \log x$$

then a standard linear simple regression is set between the data and corresponding frequencies. The slope of the straight line is  $-\alpha$ .

#### Theoretical Basement and Properties

Least square estimators for linear regression estimators, such as minimal mean square error among linear estimators, unbiased property, consistency and asymptotic normality are well known.

#### Algorithm

1. Order the sample data  $X_1, X_2, \dots, X_n$  and from a chosen one and for a given step select some points  $x_1, x_2, \dots, x_k$ .

Figure 2.5: Empirical simulated distribution function(solid line) and theoretical calculated distribution function(dashed line) for an stable distribution with  $\alpha = 1.5$ ,  $\beta = 0$ ,  $\sigma = 1$  and  $\mu = 0$

2. From a chosen one, compute the corresponding relative frequency of data above the chosen points, i.e.,  $y_i = \frac{\text{number of } X_i > x_i}{n} \simeq P(X > x_i)$
3. Take logarithms on both sides in the expression above and fit the linear regression model:

$$\log y_i = k_\alpha - \alpha \log x_i.$$

4. Set  $\hat{\alpha}$  as:

$$\hat{\alpha} = -\frac{\sum_{i=1}^k (\log y_i - \overline{\log y_i})}{\sum_{i=1}^k (\log x_i - \overline{\log x_i})^2}$$

## Implementation and Numerical Experimentation

The method is implemented in  $C_{++}$ . Tables 2.3-?? show a comparison between the true value of the parameter with the estimator obtained with the method using simulated data. Different sample sizes were analyzed. The selected points were taken from the percentile 85 %.

When  $\alpha$  is greater than 1.5 and the data achieve and asymmetric behavior the regression tail estimator overestimate the true value of the parameter. This fact was also reported for several author(see for example ([31])).

## Advantages and Drawbacks

It is very simple and easy to implement. It provides only an estimation of stability index  $\alpha$ . Expression (2.3) is fulfilled only for large values of  $x$ , so in practice is difficult to asses whether we are in the tail position, moreover it depends on which parameter  $\alpha$  we are estimating. On the other hand, if large tail points are taken as starting points less point are available for the tail estimates.

In this sense, empirical studies suggest to start from the 90% quantile.

In [6] it is reported overestimation of  $\alpha$  when it is applied to a simulated data from parameter  $\alpha \in (1, 2)$ .

In [12] it is shown that for  $\alpha$  close to 2, the tail of the stable distribution is approximate on a reasonable way by Pareto tail only from large tail points.

## 2.3.2 Hill estimator and their variants

### Description

As an alternative to regression tail estimator, it was proposed the Hill's estimator (see [17]) based in the differences between logarithms of the order statistics.

A key point is the choice of the window size  $k$ . It is always a trade off between the position in the extreme of the tail and the variance of the estimator.

Indeed, the window size needs to be small enough to capture the tail position and large enough to control the variance as can be seen from expression (2.8).

Several variants have been proposed like Pickands and modified Unconditional Pickands (see [26]) among others but solutions remain, in general, non satisfactory because of the critical dependence on the window  $k$ .

The information about this dependence contained in the Hill estimator as function of  $k$ , say  $\hat{\alpha}$ , is used in [29] to establish, from empirical evidence, a linear dependency, then a simple linear regression is adjusted, so intercept and slope are estimated using standard least square method then compared with previously calculated value from simulation studies. Finally by interpolation, surface response or again a polynomial regression also from empirical evidence the Hill estimator is obtained.

Most variants on Hill estimator are supported by empirical studies rather than theoretical analysis.

### Theoretical Basement and Properties

The expression for the Hill estimator is:

$$\hat{\alpha} = \frac{1}{\frac{1}{k} \sum_{j=1}^k \log X_{n+1-j:n} - \log X_{n-k:n}} \quad (2.7)$$

where  $X_{j:n}$  is the sample  $j$ -th order statistics. Its standard error is

$$MSE(\hat{\alpha}) = \frac{k\hat{\alpha}}{(k-1)(k-2)^{1/2}} \text{ for } k > 2 \quad (2.8)$$

### Algorithm for calculating the Hill estimator

1. Order the sample to obtain  $X_{1:n}, X_{2:n}, \dots, X_{n:n}$ .
2. Fix  $k$
3. Calculate  $\hat{\alpha}$
4. Calculate the confidence interval

$$\left( \frac{1}{Z_{1-\frac{\alpha}{2}} \hat{\alpha}^{-2} + \hat{\alpha}^{-1}}, \frac{1}{\hat{\alpha}^{-1} - Z_{1-\frac{\alpha}{2}} \hat{\alpha}^{-2}} \right)$$

### Algorithm for the selection of the window

1. From sample adjust a linear regression according to

$$\hat{\alpha}(k) = \frac{k}{1000}\hat{m} + \hat{b}$$

2. For a grid of  $(\alpha, n)$  adjust from simulated data the linear regression

$$\hat{\alpha}(k) = \frac{k}{1000}m(\hat{\alpha}, n) + b(\hat{\alpha}, n)$$

(in both steps  $\hat{\alpha}(k)$  are calculated by the first algorithm)

3. Construct throughout a table the function  $\phi : (b, m) \rightarrow (\alpha, n)$ .
4. a) Interpolate the value  $(\hat{b}, \hat{m})$  calculated in step 1 using the tabulated value obtained in step 2,  
or  
b) fit

$$\hat{\alpha}(k) = 0.810 - 0.3079\hat{b} + 2.0278\hat{b}^{0.5}$$

### Implementation and Numerical Experimentation

A program in  $C_{++}$  was implemented, giving the sample values, the confidence level and the value of  $k$  the program calculate the Hill estimator and the confidence interval. See 2.3-?? for results.

Also it fits a linear regression to the values  $(k, \hat{\alpha}(k))$  for a dense enough grid of points  $(n, \alpha)$  with simulated data,  $k$  values are such that  $0.2 \leq \frac{k}{n} \leq 0.8$ , another linear regression is fitted to sample data and corresponding intercept and slope are identified thus the right  $\alpha$  is identified either by interpolation or according to [29] method as explained in the algorithm.

### Advantages and Drawbacks

The simplicity is one the advantages.

Heavy dependency on the window size is the main problem. Nevertheless the difficulty can be overcome with a careful, rather sophisticated choice of  $k$  based on the study of this dependency for large simulated data.

Report of several authors point out overestimation for index  $\alpha$  close to 2.

## 2.4 Quantile based calibration method

### 2.4.1 McCulloch Method

#### Description

The method is based on quantile estimation. The main idea is to use differences in quantiles, properly normalized, in order to get rid dependence on location and scale parameters. Then, two functions on the stability index and the skewness are numerically calculate from the sample quantiles values and inverted to get the corresponding parameter estimates. An interpolation algorithm allows get more precise functional values. The idea goes back to McCulloch(1986), see [30].

Let  $X_1, X_2, \dots, X_n$  a sample of independent and identical distributed random variables identically distributed to  $X \sim S_\alpha(\sigma, \beta, \mu)$  and  $x_p$  the p-th population quantile. A quantile estimation is obtained from the p100% ordered sample observation whether this value is present in the sample. Otherwise the estimator is obtained by linear interpolation. For  $q(i) < p \leq q(i+1)$  we have:

$$\hat{x}_p = n(\hat{x}_{q(i+1)} - x_{q(i)})(p - x_{q(i)})$$

where  $0 < q(i) < \dots < q(i) = \frac{2i-1}{n} < \dots < 1$ .

Then the following index are established:

$$\nu_\alpha = \frac{x_{0.95} - x_{0.05}}{x_{0.75} - x_{0.25}}$$

$$\nu_\beta = \frac{x_{0.95} + x_{0.05} - 2x_{0.5}}{x_{0.95} - x_{0.05}}$$

These two indexes don't depend on  $\mu$  and  $\sigma$ . Moreover they are respectively decreasing and increasing function of  $\alpha$  and  $\beta$ .

From quantile estimates we get index estimates  $(\nu_\alpha, \nu_\beta)$ , then by inversion the corresponding estimates  $(\hat{\alpha}, \hat{\beta})$

In addition two new indexes are established as:

$$\nu_\sigma = \frac{x_{0.75} - x_{0.25}}{\sigma}$$

and

$$\nu_\zeta = \frac{\zeta - x_{0.5}}{\sigma}$$

where

$$\zeta = \begin{cases} \mu + \beta\sigma \tan \frac{\pi\alpha}{2} & \text{for } \alpha \neq 1 \\ \mu & \text{for } \alpha = 1 \end{cases}$$

The first one doesn't depend on  $\mu$ , the second one doesn't depend on  $\sigma$ .

In this way two new estimates can be obtained:

$$\hat{\sigma} = \frac{\hat{x}_{0.75} - \hat{x}_{0.25}}{\hat{\nu}_\sigma}$$

and

$$\hat{\zeta} = \hat{x}_{0.5} + \hat{\sigma}\hat{\nu}_{\zeta}$$

hence

$$\hat{\mu} = \hat{\zeta} - \hat{\beta}\hat{\sigma} \tan \frac{\pi\hat{\alpha}}{2}$$

### Theoretical Basement

$\hat{x}_p$  is a consistent estimate of  $x_p$  for large sample size, hence  $\hat{\nu}_{\alpha}, \hat{\nu}_{\beta}, \hat{\nu}_{\sigma}, \hat{\nu}_{\zeta}$  estimate consistently the corresponding index, their inverses estimate in a consistent way the parameters as indicated above.

Asymptotic covariances, and their corresponding standard deviations can be found in [10] and [30].

### Algorithm

The mentioned procedure can be summarized as:

1. Order the sample in increasing sense to obtain the order statistics  $X_{1:n}, X_{2:n}, \dots, X_{n:n}$ .
2. If  $\frac{2n+1}{p}$  is an integer number take  $i = \frac{2n+1}{p}$  and  $\hat{x}_p = x_{1:n}$ .
3. Otherwise take  $i = \lceil \frac{2n+1}{p} \rceil$  and

$$\hat{x}_p = n(x_{i+1:n} - x_{i:n})(p - x_{i:n})$$

4. Calculate

$$\nu_{\alpha} = \frac{\hat{x}_{0.95} - \hat{x}_{0.05}}{\hat{x}_{0.75} - \hat{x}_{0.25}}$$

$$\nu_{\beta} = \frac{\hat{x}_{0.95} + \hat{x}_{0.05} - 2\hat{x}_{0.5}}{\hat{x}_{0.95} - \hat{x}_{0.05}}$$

5. By two-dimensional interpolation calculate

$$\hat{\alpha} = \Psi^{-1}(\hat{\nu}_{\alpha}, \hat{\nu}_{\beta})$$

6. Tabulate  $\hat{\nu}_{\sigma}$  as a function  $\Phi(\hat{\alpha}, \hat{\beta})$  depending on  $\alpha$  and  $\beta$

7. Calculate

$$\hat{\sigma} = \frac{\hat{x}_{0.75} - \hat{x}_{0.25}}{\hat{\nu}_{\sigma}}$$

8. Calculate

$$\hat{\nu}_{\zeta} = \frac{\zeta - \hat{x}_{0.5}}{\hat{\sigma}}$$

- 9.

$$\hat{\zeta} = \hat{x}_{0.5} + \hat{\sigma}\hat{\nu}_{\zeta}$$

10.

$$\hat{\mu} = \hat{\zeta} - \hat{\beta}\hat{\sigma} \tan \frac{\pi\hat{\alpha}}{2}$$

### **Numerical implementation and results**

The key point is the procedure for calculate the inverse function on index to parameters. In principle DuMouchel's tables can be used [6] and then proceed by bi-linear interpolation.

For more precision a more sophisticated inversion method is implemented, consisting in finding the solution by moving through segments of the grid of points  $(\alpha, \beta)$ . Previously, a grid of  $200 \times 200$  points needs to be calculate.

### **Advantages and Drawbacks**

It performs well except for values of  $\alpha$  close to two and a combination of skewness parameter  $\beta$  large together with large location parameter  $\sigma$ . Precise tabulated values require a large amount of computation, though it is a pre-processing procedure.

## 2.4.2 *L-Moments Method*

### Description

*L-Moments Method* consists in matching sample weighted quantiles to the theoretical one. *L-moments* is an alternative system describing the shape of a probability distribution. They are based on the *probability weighted moment*, which is an attempt to make method of moments more robust.

Moments put a great weight in the tails of the distribution, therefore they are more affected by heavy tail phenomena, that not the case of *L-moments*. Moreover *L-moments* exist when the mean distribution is finite.

sampled *L-moments* are calculated according to 1.6.

The particularity for stable distributions is that the exact *L-moment* can not be calculate, nevertheless it is possible to follow a numerical approach to find the root of the corresponding equations as result of equate sample and theoretical *L-moments* .

### Theoretical Basement and Properties

*L-moments* are defined, for a random variable X with Cumulative Distribution Function  $F$  as:

$$M_{p,r,s} = E\{X^p[F(X)]^r[1 - F(X)]^s\}$$

Cases  $\alpha_r = M_{1,0,r}$  and  $\beta_r = M_{1,r,0}$  can be written as

$$\alpha_r = \int_0^1 Q(u)(1-u)^r du \quad (2.9)$$

$$\beta_r = \int_0^1 Q(u)u^r du \quad (2.10)$$

According to the last expression, *L-moment* can be view as an the conventional moments weighted by polynomials  $u$  or  $1-u$ .

*L-moments* itself are difficult to interpret, however, certain linear combination of them can be interpreted in terms of location parameters  $\lambda_1 = \alpha_0$ , scale parameter,  $\lambda_2 = \alpha_0 - 2\alpha_1$  or  $2\beta_1 - \beta_0$ , skewness,  $\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0$  and kurtosis,  $\lambda_4 = \alpha_0 - 12\alpha_1 + 30\alpha_2 - 20\alpha_3$ , see [13] for details. An intuitive justification for these parameters can be obtained considering the order statistics  $X_{1:n}, X_{2:n}, \dots, X_{n:n}$  and noting that:

$$\begin{aligned} \lambda_1 &= E(X_{1:1}) \\ \lambda_2 &= E(X_{2:2} - X_{1:2}) \\ \lambda_3 &= E(X_{3:3} - 2X_{2:3} + X_{1:3}) \\ \lambda_4 &= E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}) \end{aligned}$$

A natural estimator for the four parameters can be obtained from sampled data in the following way;

An unbiased estimator for  $\beta_r$  is;

$$b_r = n^{-1} \binom{n-1}{r} \sum_{j=r+1}^n \binom{j-1}{r} X_{j:n} \quad (2.11)$$

which lead to the sample *L-moments* estimators

$$\begin{aligned} l_r &= \sum_{k=0}^{r-1} p_{r-1,k}^* b_k \\ &= n^{-1} \sum_{j=1}^n j = 1^n w_{j:n}^{(r)} X_{j:n} \end{aligned}$$

for certain weights.

Where  $p_{r,k}^*$  are given by:

$$p_{r,k}^* = \frac{(-1)^{r-k} (r+k)!}{(k!)^2 (r-k)!} \quad (2.12)$$

Though exact distribution for the estimators are difficult to obtain, hence confidence intervals are not, in general available, they can be obtained from large-sample approximation using asymptotic theory. For most standard distributions, l-moments estimators of parameters and quantiles are asymptotically normal, then we can find standard error and confidence intervals (see [18]).

For stable distributions only approximate theoretical L-moments can be obtained, then, the difference with the general procedure is that equations matching theoretical and sample moments should to be solved numerically.

#### Algorithm for L-moment estimation

1. sample probability weighted moments calculation:

a) Order the sample giving  $X_{1:n}, X_{2:n}, \dots, X_{n:n}$ .

b) Calculate:

b1)  $b_0 = \frac{1}{n} \sum_{j=1}^n X_{j:n}$ ,

b2)  $b_1 = \frac{1}{n} \sum_{j=2}^n \frac{(j-1)}{(n-1)} X_{j:n}$ ,

b3)  $b_2 = \frac{1}{n} \sum_{j=3}^n \frac{(j-1)(j-2)}{(n-1)} X_{j:n}$ ,

b4)  $b_3 = \frac{1}{n} \sum_{j=4}^n \frac{(j-1)(j-2)(j-3)}{(n-1)} X_{j:n}$ ,

2. Sample L-moments calculation:

Calculate a)  $l_1 = b_0$ ,

b)  $l_2 = 2b_1 - b_0$ ,

c)  $l_3 = 6b_2 - 6b_1 + b_0$ ,

d)  $l_4 = 20b_3 - 30b_2 + 12b_1 - b_0$ .

3. Theoretical probability weighted moments calculation:

For a given set of parameters  $(\alpha, \beta, \mu, \sigma)$  calculate:

- 3a)the approximated density function by inversion in a suitable grid of points(see 2.5.1).  
 3b)The corresponding cumulative distribution function F(ref. Nolan),  
 3c)The corresponding quantile function Q(inverse of F).

4. Calculate numerically by the trapezoid method the expected values  $E(X_{r:k})$  as:

$$e_{r,k}(\alpha, \beta, \mu, \sigma) = \frac{k}{(r-1)!(k-r)!} \int_0^1 Q(u)u^{r-1}(1-u)^{k-r} du$$

for  $r = 1, 2, 3, 4$  and  $k = 1, 2, 3, 4$ .

5. Theoretical L-moments calculation:  
 Calculate:

$$\begin{aligned} \lambda_1 &= \lambda_1(\alpha, \beta, \mu, \sigma) = e_{1,1} \\ \lambda_2 &= \lambda_2(\alpha, \beta, \mu, \sigma) = \frac{1}{2}(e_{2,2} - e_{1,2}) \\ \lambda_3 &= \lambda_3(\alpha, \beta, \mu, \sigma) = \frac{1}{3}(e_{3,3} - 2e_{2,3} + e_{1,3}) \\ \lambda_4 &= \lambda_4(\alpha, \beta, \mu, \sigma) = \frac{1}{4}(e_{4,4} - 3e_{3,4} + 3e_{2,4} - e_{1,4}) \end{aligned}$$

6. Solve

$$\lambda_k(\alpha, \beta, \mu, \sigma) - l_k = 0 \text{ for } k = 1, 2, 3, 4 \quad (2.13)$$

### Implementation and Numerical Experimentation

Solving equations 2.13 by standard numerical methods involves several approximations at a high computational cost. To avoid numerical inaccuracies, at least partially, a table is calculated linking the parameters  $(\alpha, \beta)$  with the theoretical L-moments  $\lambda_k(\alpha, \beta, 0, 1)$  then a relationship is established in order to calculate  $\lambda_k(\alpha, \beta, \mu, \sigma)$ . As in McCulloch method the table is constructed previously so the solution of a concrete problem.

### Advantages and Drawbacks

## 2.5 Maximum Likelihood Methods

### 2.5.1 Zolotarev's Integral Representation Method

#### Description

Classical Maximum Likelihood Method has been implemented for stable distributions in early date(see [7],[8] and [?]).

The main difficult is again that the density is unknown.

One possibility is to approximate density by inverting the characteristic function via Fast Fourier Transform. Other related method relies on the Zolotarev integral representation, [28] which has the advantage of perform calculation over a finite interval.

Once the PDF is calculated on a grid, a quasi-newton method is implemented to maximize the likelihood:

$$l(\alpha, \beta, \theta, \zeta) = \sum_{i=1}^n \log f(X_i; \alpha, \beta, \theta, \zeta) \quad (2.14)$$

where the parametric space is defined in 2.1.

As most quasi-Newton methods, the idea is to construct an approximation of the inverse Hessian using the information gathered in at the descent process, driven by the gradient progress. The current approximation converge to the inverse of the Hessian like classic Newton method but without the necessity to calculate the later at every point which has a high computational cost. Two stopping criteria are used: a limit in the number of evaluations and the amount of the increment in the likelihood.

As starting point the McCulloch estimator is taken.

Also, confidence intervals are numerically calculated from approximated Fisher information matrix that is closed related to the Hessian calculated for the quasi-newton method, see [21].

#### Theoretical Basement and Properties

Consistency, asymptotic normality and efficiency are well known properties of the m.l.e.

Moreover, from general well known results for m.l.e., if  $\hat{\theta}$  is the m.l.e. of parameter  $\theta$  then for large samples  $\hat{\theta} \sim N(\theta, n^{-1}B)$  where B is the inverse of the Fishers information matrix with entries:

$$I_{ij} = \int_{-\infty}^{\infty} \frac{\partial f}{\partial \theta_i}(x) \frac{\partial f}{\partial \theta_j}(x) \frac{1}{f}(x) dx$$

where f is the stable pdf.

So the large sample confidence intervals for each of the parameters are:

$$\hat{\theta}_i \pm Z_{1-\frac{\alpha}{2}} \frac{\sigma_{\hat{\theta}_i}}{\sqrt{n}}$$

where  $\sigma_{\hat{\theta}_1}, \sigma_{\hat{\theta}_2}, \sigma_{\hat{\theta}_3}$  and  $\sigma_{\hat{\theta}_4}$  are the square rots of the diagonal entries of B.

It is known that inverting characteristic function is assured by inversion transform formula

and then, Theorems 2 and 3 in [28] guarantee inversion formulas in the Zolotarev's representation method.

Following this method the PDF can be calculated as:

$$f(x; \alpha, \beta) = c_2(x, \alpha, \beta) \int_{\theta_0}^{\frac{\pi}{2}} g(\theta; x, \alpha, \beta) \exp -g(\theta; x, \alpha, \beta) d\theta \quad (2.15)$$

with

$$c_2(x, \alpha, \beta) = \begin{cases} \frac{\alpha}{\pi|\alpha-1|(x-\zeta)} & \text{for } \alpha \neq 1 \\ \frac{1}{2|\beta|} & \text{for } \alpha = 1 \end{cases} \quad (2.16)$$

$$c_3(\alpha) = \begin{cases} \frac{\text{sign}(1-\alpha)}{\pi} & \text{for } \alpha \neq 1 \\ \frac{1}{\pi} & \text{for } \alpha = 1 \end{cases} \quad (2.17)$$

$$g(\theta; x, \alpha, \beta) = \begin{cases} (x - \zeta)^{\frac{\alpha}{\alpha-1}} V(\theta; \alpha, \beta) & \text{for } \alpha \neq 1 \\ \exp \frac{-\pi x}{2\beta} V(\theta; \alpha, \beta) & \text{for } \alpha = 1 \end{cases} \quad (2.18)$$

$$\zeta = \begin{cases} -\beta \tan \frac{\pi\alpha}{2} & \text{for } \alpha \neq 1 \\ \exp \frac{-\pi x}{2\beta} & \text{for } \alpha = 1 \end{cases} \quad (2.19)$$

$$\theta_0 = \begin{cases} \frac{1}{\alpha} \arctan(\beta \tan \frac{\pi\alpha}{2}) & \text{for } \alpha \neq 1 \\ \frac{-\pi x}{2} & \text{for } \alpha = 1 \end{cases} \quad (2.20)$$

and

$$V(\theta; \alpha, \beta) = \begin{cases} (\cos \alpha \theta_0)^{\frac{1}{\alpha-1}} \frac{\cos \theta}{\sin \alpha(\theta_0+\theta)} \frac{\cos(\alpha\theta_0+(\alpha-1)\theta)}{\cos \theta} & \text{for } \alpha \neq 1 \\ \frac{-\pi x}{2} & \text{for } \alpha = 1, \beta \neq 0 \end{cases} \quad (2.21)$$

The following quantities are involved in the method:

$$g_k = \nabla l(\theta_k)$$

$$q_k = g_{k+1} - g_k$$

$$p_k = \theta_{k+1} - \theta_k$$

where  $\theta_k$  is the k-th iteration of the parameter vector:

$$S_{k+1} = S_k \frac{p_k p'_k}{p_k q'_k} - \frac{S_k q_k q'_k S_k}{q'_k S_k q_k}$$

The Hessian is approximate using the rank one correction by the sequence  $S_k$  of 4x4 matrices, starting from the identity matrix.

Movement from one point to another is given by calculating the direction  $d_k = -S_k g_k$  and then maximizing  $l(\theta_k + \delta d_k)$  with respect to  $\delta$  to find the optimal point  $\delta_k$ . This determines  $\theta_{k+1} = \theta_k + \delta_k S_k$ .

The convergence is theoretically assured. Although it has some disadvantages concerning the positive definiteness condition for matrix  $S_{k+1}$  given by  $q'_k(p - k - S_k q_k) > 0$ . When this condition is not fulfilled is recommended to take  $S_{k+1} = S_k$ .

The Davidon-Fletcher-Powell method improves the result concerning the positive definiteness condition .

### Algorithm for inverting pdf following Zolotarev's integral representation.

Given a set of points  $(x, \alpha, \beta)$ :

1. Calculate  $\zeta$ .
2. Fix  $\theta_0$ .
3. Calculate  $c_2(x, \alpha, \beta)$ .
4. Calculate  $c_3(\alpha)$ .
5. Calculate  $\zeta$ .
6. Calculate  $V(\theta; \alpha, \beta)$ .
7. For a partition  $\theta_0, \theta_1, \dots, \theta_n$  on  $(-\theta_0, \frac{\pi}{2})$  evaluate:  
 $g(\theta; x_i, \alpha, \beta) \cdot \exp\{-g(\theta; x_i, \alpha, \beta)\}$ .
8. Calculate

$$f(x; \alpha, \beta) = c_2(x, \alpha, \beta) \sum_{i=0}^n (g(\theta_i; x, \alpha, \beta) \exp\{-g(\theta_i; x, \alpha, \beta)\}) (\theta_{i+1} - \theta_i) \quad (2.22)$$

$$+(\theta_{i+1} - \theta_i) g(\theta_{i+1}; x, \alpha, \beta) \exp\{-g(\theta_{i+1}; x, \alpha, \beta)\} - g(\theta_i; x, \alpha, \beta) \exp\{-g(\theta_i; x, \alpha, \beta)\}$$

### Maximization Algorithm with Davidon-Fletcher-Powell method

1. (initialization) Take  $\theta_0$  as the estimation in Mcculloch quantile method and  $S_k = I$ . From a parameter point  $\theta_k$  calculate  $\theta_{k+1}$  as:
2. Calculate the approximate gradient  $g_k$ .
3. Calculate the direction from  $d_k = -S_k g_k$ .
4. Maximize  $l(\theta + \delta d_k)$  with respect to  $\delta > 0$  determining the optimal point  $\delta_k$ .

5. Calculate  $\theta_{k+1} = \theta_k + \delta_k d_k$ .
6. Calculate  $p_k = \alpha_k d_k$ .
7. Calculate  $g_{k+1}$ .
8. Set  $q_k = g_{k+1} - g_k$  and calculate

$$S_{k+1} = S_k + \frac{p_k p_k'}{p_k q_k'} - \frac{S_k q_k q_k' S_k}{q_k' S_k q_k}$$

9. Update and return to step 1.

### Algorithm for large sample confidence interval calculation

1. Compute the PDF  $f$  according to the first algorithm evaluated at points  $(x, \hat{\theta})$ .
2. Calculate numerically the integral defining the entries of the Fisher's information.
3. Calculate the confidence interval.

### Implementation and Numerical Experimentation

Three programs in  $C++$  have been implemented, concerning the corresponding algorithms mentioned above.

The first one, for a given point  $(x, \theta)$  allows to calculate the PDF at the point. Results are shown in tables 2.3-???. We encountered some numerical inaccuracies in evaluating the likelihood function for large sample values.

### Advantages and Drawbacks

It is theoretically the most efficient calibration method concerning its variance for large samples but at a high computational cost. Two main numerical approximations are involved: the density itself and the maximum search, so carefully implementation is need.

## 2.5.2 Monte Carlo Markov Chain Simulated Annealing Method

### Description

It is an alternative to m.l.

The main idea is to construct a grid on the parametric space and find the maximum by moving through neighbor points in the grid.

A neighbor is chosen at random with equal probability and, if the likelihood evaluated at this point is greater than the previous one the system will move to it with a certain probability. This probability is chosen in a way that points where the system moves through form a reversible Markov Chain whose stationary probability law is the desired one. It can be done using the Hansting-Metropolis Algorithm.

It turns out that the limit probability law depends on a parameter called temperature.

In order to assure this probability law charges only optimum points the temperature is raised slowly to infinity.

### Theoretical Basement

Let  $l(\theta)$  be the log-likelihood and  $l^* = \max_{\theta \in G} l(\theta)$ , where  $G \subset \Theta$  is the set of points in the parametric space belonging to the grid.

Usually the cardinal of  $G$  is very large, then to try to find the maximum points by direct evaluation of all its points is an intractable problem.

Let  $m = \{x \in G / l(x) = l^*\}$  be the set of all optimal points.

For  $\lambda > 0$  a probability law on  $G$  is chosen as:

$$p_\lambda(x) = \frac{\exp \lambda(l(x) - l^*)}{|M| + \sum_{y \notin M} \exp \lambda(l(y) - l^*)}$$

where  $|M|$  is the cardinal of the set  $M$ .

Note that under usual regular conditions  $|M| = 1$  and also:

$$p_\lambda(x) \xrightarrow{\lambda \rightarrow +\infty} \frac{\delta_x(M)}{|M|}$$

where  $\delta_x(\cdot)$  is the Dirac measure concentrated at point  $x$ .

Hence, for  $\lambda$  large, the probability law  $\{p_\lambda(x), x \in G\}$  concentrates on points in  $M$ .

If at a given time the system is at point  $\theta$ , then a neighbor is chosen with equal probability, say  $\theta'$ , and the system will move to  $\theta'$  if  $l(\theta') \geq l(\theta)$ , otherwise the system changes to  $\theta'$  with probability:

$$\exp\{\lambda(l(\theta') - l(\theta))\}$$

or it remains in  $\theta$ .

At step  $n$ ,  $\lambda$  is taken as

$$\lambda_n = C \log(1 + n) \tag{2.23}$$

where  $C$  is a constant.

The sequence of chosen points  $(X_n)_{n \in \mathbb{N}}$  form an stationary Markov Chain with transition

probabilities:

$$p_{\theta,\theta'} = \frac{1}{8} \min\{1, \exp \lambda(l(\theta') - l(\theta))\}$$

Indeed, the chain is reversible, satisfying:

$$p_{\lambda}(\theta)p_{\theta,\theta'} = p_{\lambda}(\theta')p_{\theta',\theta}$$

then these satisfy the stationary conditions for an irreducible Markov Chain, hence the law  $p_{\lambda}(\theta)_{\theta \in G}$  is the limit distribution.

With a slight modification the argument is valid for  $\lambda$  variable in the way signaled in 2.23.

### Algorithm

1. *Initialization part*

a) Take any  $\theta_0 = (\alpha_0, \beta_0, \sigma_0, \mu_0)$  and  $\lambda_0$ .

b) Define  $h$  and take a grid  $G$  on  $\Theta$  with distance  $h$  between neighbors.

2. *Iterative part*

At step  $n$  the system is at point  $\theta_n$  then:

3. Generate a uniform discrete random number on  $[1,8]$  and select the target point  $\theta_{n+1}$  among neighbors (neighbors should be ordered previously).

4. Evaluate  $l(\theta_{n+1})$ .

5. If  $l(\theta_{n+1}) \geq l(\theta_n)$  then  $W = 0$  and move to  $\theta_{n+1}$ , else:

6. With probability:

$$\exp \lambda(l(\theta_{n+1}) - l(\theta_n))$$

move to  $\theta_{n+1}$ , fix  $W = 0$  else  $W = W + 1$  and repeat the process from  $\theta_n$ .

7. *Ending part*

If  $W > \varepsilon$  stop the process and set  $l^* = l(\theta_n)$  and  $\theta^* = \theta_n$ .

## Numerical implementation

The technique has been implemented in C++ language program, a grid of points was taken. As stopping criteria was used a time of remaining at the same stage.

The results are shown in tables 2.3-???. Despite the larger computing time the method doesn't seem to be superior to the others, in some cases is worse. It is due to numerical problems in evaluating the likelihood for large sample values.

## Advantages and Drawbacks

The main advantage is that global optimum points are obtained as opposite to conventional m.l method where we have only local optimum points.

The weakness is that a large amount of computation is required to reach the optimum, see the appendix to compare the results.

## 2.6 Sample characteristic function methods

### 2.6.1 Minimum Distance Method

#### Description

Several methods using the empirical characteristic function has been implemented.

This method consist in maximize the distance between theoretical characteristic function  $\phi$  and the sample characteristic function,  $\hat{\phi}(t)$  directly without passing throughout the PDF.

As distance it is usually taken the norm infinity or preferable the norm  $L^2$  between functions, weighted in a convenient way in order to make the norm convergent.

Distance calculation is accomplished by mean of an Hermitian quadrature formula and .The procedure has been accomplished numerically using a 20-points to calculate:

$$\int_{-\infty}^{\infty} |\phi(t) - \hat{\phi}(t)|^r W(t) dt = \sum_{i=1}^n w_k \| (\phi(t) - \hat{\phi}(t)) \| \quad (2.24)$$

where  $u'_k$ s are zero of the Hermite polynomials and  $w'_k$ s are the respective weights for implementation details. A gradient projection routine iterates to obtain optimal values of estimates from an starting point.

#### Theoretical Basement and Properties

Denote by:

$$\hat{\phi}(t) = \frac{1}{n} \sum_{j=1}^n \exp itX_j$$

the sample characteristic function.

Since  $|\hat{\phi}(t)|$  is bounded all moments of  $\hat{\phi}(t)$  are finite for any fixed  $t$ .

By Law of Large Numbers  $\hat{\phi}(t)$  is a consistent estimator of the cf  $\phi(t)$ .

Furthermore in the symmetric case the cf is real and

$$\hat{\phi}(t) = \frac{1}{n} \sum_{j=1}^n \cos(itX_j)$$

Let

$$D(\alpha, \beta, \sigma, \mu) = \| \phi(t) - \hat{\phi}(t) \|$$

the method finds

$$\hat{\theta} = \operatorname{argmin}_{\theta \in \Theta} D(\theta)$$

where  $\theta = (\alpha, \beta, \sigma, \mu)$  is a point in the parametric space and  $\| \cdot \|$  is a norm, usually the  $L^\infty$  or  $L^r$  weighted norms. The last is more suited for implementation and it can be expressed by:

$$h(\theta) = \int_{-\infty}^{\infty} |\phi(t) - \hat{\phi}(t)|^r W(t) dt$$

where  $W(t)$  is the weight making the integral to be convergent. Typically  $r = 2$  and  $W(t) = \exp(-t^2)$  (see [33, 34]).

The estimators are consistent and asymptotically normal in the symmetrical case although an efficiency study is not reported.

### Algorithm

1. From an starting parameter set  $\theta_0$  solve the Hermitian quadrature (2.24) for  $\theta_k$  to obtain  $D(\theta_k)$ .
2. Calculate the approximated gradient  $\nabla D(\theta_k)$  and search a direction  $d$  such that  $\nabla D(\theta_k)d > 0$ ,
3. Calculate  $\theta_{k+1} = \theta_k + d$  and repeat steps 1 and 2,
4. Stop if a given number of iteration is reached or  $D(\theta_{k+1}) - D(\theta_k) < R$  for a given  $d$ .

### Implementation and Numerical Results

The method was not implemented.

### Advantages and Drawbacks

The method has the advantage to m.l. that minimize directly CF avoiding the inversion procedure. Nevertheless direct minimization of the distance has a high computational cost, instabilities are reported for  $\sigma$  and  $\mu$  far from unity and zero respectively.

Following the same idea other methods reveal to be more efficient and simple.

## 2.6.2 Regression Method

### Description

In [20] a regression-type method using the CF is presented. From the general expression of the CF a linear expression is obtained between certain functionals of the CF and the parameters  $\alpha$  and  $\sigma$  only, then from sample cf it is fitted a linear regression to estimate  $\alpha$  which is precisely the slope of the straight line.

Again from the CF another linear expression is obtained involving this time the parameters  $\beta$  and  $\mu$  together with non-linear relationship on  $\alpha$  and  $\sigma$ , then once the last two are estimated it is proceed to estimate the first ones by fitting a linear regression.

The first adjustment can be repeated a number of times to achieve better precision in the estimation.

A variant is proposed, called *Iterative Regression Method* consisting in a recursive estimation of parameters, once an estimation set is obtained, the data is standardized by subtraction the location parameter and divided by the scale parameter. This variant aims to make use of the empirical evidence from simulation that the original method performs better for parameters  $\mu$  and  $\sigma$  close to zero and one respectively.

### Theoretical Basement and Properties

Least square estimator properties are well known. The first equation is obtained from general expression of the cf of stable law namely:

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

then a simple linear regression  $y_i = m + \alpha w_k + e_k$  is fitted with  $y_k = \log(-\log |\hat{\phi}(t_k)|^2)$  with  $t_k = \frac{\pi k}{25}, k = 1, 2, \dots, K$  and  $K$  ranging from 9 to 134 according to Koutrouvelis proposal. Here  $m = \log(2\sigma^\alpha)$  and  $w_k = \log|t_k|$ .

Once  $\hat{\alpha}$  and  $\hat{\sigma}$  have been obtained then estimates of  $\beta$  and  $\mu$  can be obtained by using:

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

derived also from the expression of the CF. Here arctan denotes the principal value of the arctan function.

Also, writing  $g_n(t) = \arctan\left(\frac{Im(\hat{\phi})(t)}{Re(\hat{\phi})(t)}\right)$  and  $z = g_n(u) + \pi k_n(u)$  the following regression model is obtained:

$$z_i = \mu u_i + \beta \sigma^\alpha \tan \frac{\pi\alpha}{2} \text{sign}(u_i)|u_i|^\alpha + \eta_i, \quad i = 1, 2, \dots, L.$$

where  $u_i$  are appropriate real numbers and  $\eta_i$  denotes the error terms.

### Algorithm

1. Fix an appropriate  $K$  and, for points  $t_k = \frac{\pi k}{25}, k = 1, 2, \dots, K$ , calculate the sample CF  $\hat{\phi}(t)$ .

2. Calculate  $y_k = \log(-\log |\hat{\phi}(t_k)|^2)$ .
3. Fit the linear regression  $y_i = m + \alpha w_k + e_k$ , finding the estimates  $\hat{\alpha}$  and  $\hat{m}$ .
4. Find the estimate  $\hat{\sigma}$  by solving  $m = \log(2\hat{\sigma}^{\hat{\alpha}})$ .
5. Calculate, for an appropriate L,  $z(u_i) = g_n(u_i) + \pi k_n(u_i)$  for  $u_i = \frac{\pi i}{50}$ .
6. Fit the regression

$$z_i = \mu u_i + \beta \hat{\sigma}^{\alpha} \tan \frac{\pi \alpha}{2} \text{sign}(u_i) |u_i|^{\alpha} + \eta_i, \quad i = 1, 2, \dots, L.$$

to obtain estimators  $\mu$  and  $\beta$ .

### Implementation and Numerical Experimentation

The second fit requires numerical inversion of matrices, for large values of L it is time consuming, even when accurate inversion algorithms are implemented. A recursive procedure can be implemented in order to fit K automatically (see [39]).

### Advantages and Drawbacks

It is simpler to implement than the minimum distance estimator or maximum likelihood, though the efficiency is lesser than the last one.

## 2.6.3 Method of Moments

### Description

In [33] and [34] is proposed another simple estimation method based on transformation of the CF.

this transformation gives a relationship between the logarithm of the CF,  $\alpha$ -power of its argument and  $\sigma$ . Solving this equation for two values, replacing cf by sample characteristic function at two points produces estimates for both parameters.

in a similar way another relationship is obtained for  $\beta$  and  $\mu$  the evaluation for two other points and solving the equation the corresponding estimators are obtained.

### Theoretical Basement and Properties

The estimators are consistent since they are based upon estimation of  $\phi(t)$ ,  $Im\phi(t)$  and  $Re\phi(t)$  which are consistent, as was discussed above.

The two relationship, obtained from 2.1 are, for  $\alpha \neq 1$ :

$$-\log|\phi(t)| = \sigma^\alpha |t|^\alpha$$

and

$$Im\phi(t) = \mu t + \sigma^\alpha |t|^\alpha \beta sign(t) \tan \frac{\alpha\pi}{2}.$$

Solving the first equation with points  $t_1$  and  $t_2$  gives:

$$\hat{\alpha} = \frac{\log \frac{\log|\phi(\hat{t}_1)|}{\log|\phi(\hat{t}_2)|}}{\log \frac{|t_1|}{|t_2|}} \quad (2.25)$$

and

$$\log \hat{\sigma} = \frac{\log(-\log|\phi(t_1)|) \hat{\log} |(t_1)| \log(-\log|\phi(\hat{t}_2)| \log(|(t_2)|))}{\log \frac{|t_1|}{|t_2|}} \quad (2.26)$$

Solving the second one for  $t_3$  and  $t_4$  gives:

$$\hat{\beta} = \frac{\frac{u(\hat{t}_4)}{t_4} - \frac{u(\hat{t}_3)}{t_3}}{[|t_4|^{\hat{\alpha}-1} - |t_3|^{\hat{\alpha}-1}] \hat{\sigma}^{\hat{\alpha}} \tan \frac{\hat{\alpha}\pi}{2}} \quad (2.27)$$

and

$$\hat{\mu} = \frac{|t_4|^{\hat{\alpha}-1} \frac{u(\hat{t}_4)}{t_4} - |t_3|^{\hat{\alpha}-1} \frac{u(\hat{t}_3)}{t_3}}{[|t_4|^{\hat{\alpha}-1} - |t_3|^{\hat{\alpha}-1}]} \quad (2.28)$$

For  $\alpha = 1$  similar simpler equations are obtained.

### Algorithm

1. Choose two points  $t_1$  and  $t_2$  and calculate  $\phi(\hat{t}_1)$  and  $\phi(\hat{t}_2)$ ,
2. Calculate  $\hat{\alpha}$  and  $\hat{\sigma}$  from expressions (2.25) and (2.26),
3. Choose two points  $t_3$  and  $t_4$  and calculate  $\hat{\beta}$  and  $\hat{\mu}$  from expressions (2.27) and (2.28).

### Implementation and Numerical Experimentation

A key point is the optimal choice of points  $t_1, t_2, \dots, t_4$ , for stable symmetric distributions positive experiences are reported for  $t_1 = 0.2$  and  $t_2 = 0.8$ .

Attempts to fix  $t_3$  and  $t_4$  in dependency of the estimations of  $\hat{\alpha}$  and  $\hat{\sigma}$  haven't arrived to a satisfactory formula.

See tables 2.3-?? in appendix for results which are not as entirely satisfactory as in other cases.

### Advantages and Drawbacks

Simplicity is an attractive feature. The method is reported to give significant deviations from the true parameters when the distribution exhibits a non-symmetric behavior ([39]).

# Bibliography

- [1] BERENDO, M., HODGES, S. D. Multivariate Distributional Test in Risk Management : Empirical Characteristic Function Approach *Working Paper*, [www.gloriamundi.org.com](http://www.gloriamundi.org.com), **2002**.
- [2] BESAG, J. Markov Chain Monte Carlo for Statistical Inference. *Working Paper*, Center for Statistics and the Social Sciences.
- [3] BIBBY B. M., SORENSEN M. Hyperbolic Processes in Finance *working paper*, **2002**.
- [4] CARRASCO, M., FLORENS, J. Generalization of GMM to a continuum moment condition. *Econometric Theory*, **16**: 767-834(2000).
- [5] J. M. CHAMBERS, C. L. MALLOWS, B. W. STUCKA Method for simulating Stable Random Variables. *Journal of the American Statistical Association* **71**: 340-344, 1976.
- [6] W. H. DUMOUCHEL *Stable distributions in statistical inference*. PhD thesis, University of Ann Arbor, Ann Arbor, MI, 1971.
- [7] W. H. DUMOUCHEL On asymptotical normality of the m.l.e. when sampling from stable , *Annals of Statistics*, **1**: 948-957, 1973.
- [8] DUMOUCHEL. Stable distribution in statistical Inference *PhD Thesis, Dept. of Statistics, Yale univeristy*.
- [9] EMBRECHTS, P., RESNIK S. I., SAMORODNITSKY, G. Extreme Value theory as risk management tool. *North american Actuarial*, **3**: 30-41, 1999.
- [10] E. FAMA AND R. ROLL Parameters estimates for Symmetric Stable Distributions. *Journal of the American Statistical Association*, **66**: 331-339, 1971.
- [11] FEUERVERGER, A. An efficient result for the empirical characteristic function in stationary time-series models. *The Canadian Journal of Statistics* . **18**: 155-161 (1990).
- [12] FOFAK AND J. NOLAN Tail behavior, Modes and other characteristics of stable distributions, 1998 (preprint).

- [13] GREENWOOD. Regional Frequency Analysis **1979**.
- [14] GUYON X.. *Random Fields on a Network*. Springer-Verlag, New York, 1995.
- [15] HARRELL, F.E., DAVIS, C.E.. A new distribution-free quantile estimator *biometrika*, **69(3)**:635-640.
- [16] HERNANDEZ, N., SECO, L. .Non-gaussian univariate distributions *working paper*.
- [17] B.M. HILL A simple general approach to inference about the tail of a stable distribution. *Annals of Statistics* **3(5)**, **1163-1174**, **1975**.
- [18] HOSKING, J.R.M.. L-moments: analysis and estimation of distributions using linear combination of order statistics *J.R. Statist. Soc., Ser. B*, **52(1)**:105-124(1990).
- [19] KNIGHT, J.L., YU, J. The empirical characteristic function in time series estimation. *Econometrics Theory*, *forthcoming*.
- [20] KOUTROVELIS, I.A. Regression-type estimation of the parameters of stable laws. *Journal of the American Statistical Association* **75**: 918-928(1980).
- [21] LUENBERGER, D.G.. Introduction to Linear and non-linear programming(1974).
- [22] MADAN, D.B., SENETA E. The variance gamma model from shared market returns *Journal of Business*, **63**: 511-524(1990).
- [23] MANDELBROT, B. The variation of certain speculative prices, *J. of Bus.*, **26**:394-419(1963).
- [24] MAUSSEL, H.. Calculating Quantile-based Risk Analytics with l-estimators *Algo Research Quarterly*, **33**: vol 4, no.4(2001).
- [25] MAUSSEL, H., ROSEN D.. Beyond VaR: From measuring Risk to managing risk *Algo Research Quarterly*, **1(2)**: 5-20(1998).
- [26] MITNIK, S., RACHEV, S.T. Tail estimation of the Stable Index  $\alpha$ . *Applied Mathematic Letters* **9(3)**:53-56(1996). 53-56,1996.
- [27] HERNANDEZ, N. PHD Thesis. Toronto. Canada(2002).
- [28] NOAL, J.P.. Numerical calculation of stable densities *Stochastic Models*, **4**.
- [29] MITNIK, PAOELLAA Simple Estimator for the Characteristic Exponent of the Stable Paretian Estimator. *Mathematical and Computer Modelling*:29,161-176,1999.
- [30] J.H. McCULLOCH Simple consistent estimators of stable distribution parameters. *Commun. Statistist.-Simula* **15**:1109-1136,1986.

- [31] J.H.McCULLOCH Measuring tail thickness in order to estimate the stable index  $\alpha$ : a critique. *Business and Economic Statistics* **15:74-81**, 1997.
- [32] PARZEN, E. On estimation of a probability density function and mode *Annals of Mathematics Statistics*, **33**:1065-1076(1962).
- [33] PRESS, S.J. *Applied multivariate Analysis*. Holt, Rinehart & Wiston Inc., New York.
- [34] PRESS, S.J. Estimation in Univariate and Multivariate Stable Distribution. *J. Amer. Statist. Assoc.* **67**:842-846.
- [35] SHEATER, S.J., MARRON, J.S. Kernel Quantile Estimator *Journal of American Statistical Association*, **85(410)**:410-416(1990).
- [36] SCHMIDT, P. An improved version of Quandt-Ramsey MGF estimator for mixtures of normal distributions and switching regressions. *Econometrica*, **50**:501-524(1982).
- [37] BRADLEY B.O., TAQQU M.S. Financial Risk and Heavy Tails *Heavy-tailed distributions in Finance*, Rachev S.T. editor, **2000**:
- [38] TRAN, K.C. Estimating mixtures of normal distribution via empirical characteristic function. *Economic Review*, **17**:167-83(1998).
- [39] R. WERON On the Chambers-Mallows-Stuck method for simulating skewed stable random variables. *Statistics and Probability Letters* **28**:165-171, 1996.
- [40] V.M. ZOLOTAREV On representation of stable laws by integrals *Selected Translation in Mathematical Statistics and Probability* **6**:84-88, 1986.

## 2.7 Appendix

Table 2.1: Kolmogorov-Smirnov test applied to simulated data. Significance Level: 0.05, No. of observations: 500, No. of samples: 100 000.

Parameters	fraction rejected
$\alpha = 1, \beta = 0, \sigma = 1, \mu = 0$	0.005049
$\alpha = 1, \beta = 0, \sigma = 3, \mu = 0$	0.004966
$\alpha = 1, \beta = 0, \sigma = 1, \mu = 3$	0.004483
$\alpha = 1, \beta = 0, \sigma = 3, \mu = 2$	0.004198
$\alpha = 1, \beta = 0, \sigma = 1, \mu = 10$	0.004887
$\alpha = 1, \beta = 0, \sigma = 10, \mu = 0$	0.004239
$\alpha = 1.5, \beta = 0, \sigma = 10, \mu = 0$	0.005179

Table 2.2: Kolmogorov-Smirnov test applied to simulated data. Significance Level: 0.01, No. of observations: 1000, No. of samples: 100 000.

<b>Parameters</b>	<b>fraction rejected</b>
$\alpha = 1, \beta = 0.5, \sigma = 1, \mu = 0$	0.0158
$\alpha = 1.5, \beta = 0.5, \sigma = 1, \mu = 0$	0.0133

Table 2.3: Mean estimates and their standard deviations from 25 groups of simulated stable data with selected parameters results and size n=250

<i>estimator</i>	$\alpha = 0.5$	$\beta = 0$	$\sigma = 1$	$\mu = 0$
CF-Regression	0.435915	-0.013588	1.14528	0.00845163
	0.0552721	0.169768	0.386165	0.121372
McCulloch	0.48017	0.0764	0.860747	-0.0504369
	0.049019	0.23643	0.274849	0.181776
MLE	0.495594	0.0558397	0.890342	-0.0213346
	0.0388056	0.202488	0.254329	0.127616
CF-Moment	0.499293	2.43702	1.00773	-2.01898
	0.108261	10.3082	0.162915	18.8282
Iterative CF-regression	0.454551	0.0128717	1.06439	-0.227646
	0.04446	0.18381	0.207676	0.74108
MCMC	0.52451	-0.0156	0.936816	0.00492654
	0.0280826	0.112947	0.155678	0.105463

Table 2.4: Mean estimates and their standard deviations from 25 groups of simulated stable data with selected parameters results and size n=250

<i>estimator</i>	$\alpha = 1$	$\beta = 0$	$\sigma = 1$	$\mu = 0$
CF-Regression	0.991621	-0.0349896	1.01025	-0.558531
	0.0858409	0.190262	0.0953532	3.86229
McCulloch	1.01369	-0.0712	0.982848	-1.00437
	0.118403	0.188156	0.109607	3.40732
MLE	0.998356	0.00224512	1.06575	-0.96617
	0.148866	0.300294	0.202728	3.37208
CF-Moment	1.02666	-3.26098	1.03054	290.468
	0.119731	15.9729	0.0936238	937.049
Iterative CF-regression	0.720715	-0.0487408	1.10693	4.82717
	0.267385	0.172512	0.142423	41.6437
MCMC	1.04635	-0.0222	0.992922	0.677912
	0.086774	0.159257	0.138151	2.59554

Table 2.5: Mean estimates and their standard deviations from 25 groups of simulated stable data with selected parameters results and size n=250

<i>estimator</i>	$\alpha = 1.5$	$\beta = 0$	$\sigma = 1$	$\mu = 0$
CF-Regression	1.52003 0.118926	-0.0730552 0.274792	0.981433 0.0635601	-0.0762012 0.211226
McCulloch	1.51484 0.138123	-0.0556 0.283638	0.961019 0.069596	-0.0890961 0.253612
MLE	1.48544 0.150724	-0.0409574 0.249337	1.02273 0.136428	-0.0745139 0.24507
CF-Moment	1.52411 0.128137	0.886024 57.046	0.976487 0.108296	-4.5581 29.9291
Iterative CF-regression	1.52361 0.123387	-0.106227 0.266413	0.982847 0.0664215	-0.0975127 0.203274
MCMC	1.47301 0.112791	0.0916 0.141014	1.29567 0.525525	0.0468806 0.207485

Table 2.6: Mean estimates and their standard deviations from 25 groups of simulated stable data with selected parameters results and size n=500

<i>Estimator</i>	$\alpha = 0.5$	$\beta = 0$	$\sigma = 1$	$\mu = 0$
CF-Regression	0.478823 (0.0479067)	0.0251645 (0.154217)	1.00838 (0.142711)	-0.00920822 (0.105985)
McCulloch	0.50747 (0.0313318)	-0.0096 (0.167667)	0.926555 (0.202289)	-0.0172259 (0.175157)
MLE	0.514286 (0.0337152)	-0.017812 (0.154116)	0.931027 (0.195897)	-0.00900449 (0.139462)
CF-Moment	0.498883 (0.0542758)	-2.47976 (9.35372)	0.994678 (0.088882)	4.50854 (15.1219)
Iterative CF-Regression	0.48404 (0.0448191)	0.0238453 (0.148163)	1.00044 (0.138822)	-0.0057291 (0.108968)
MCMC	0.52483 (0.0214436)	0.0206 (0.0638899)	0.944183 (0.084864)	-0.0193569 (0.0546881)

Table 2.7: Mean estimates and their standard deviations from 25 groups of simulated stable data with selected parameters results and size n=500

<i>estimator</i>	$\alpha = 0.5$	$\beta = 0.5$	$\sigma = 1$	$\mu = 0$
FC-Regression	0.471069	0.563546	1.02589	-0.0151985
	0.0404129	0.118261	0.124909	0.0727878
McCulloch	0.52151	0.5392	1.01807	-0.137257
	0.0511976	0.180899	0.331423	0.300505
MLE	0.523811	0.536335	1.02469	0.0707467
	0.0691871	0.169322	0.323351	0.293413
FC-Moment	0.514503	7.13127	0.994288	-8.87895
	0.0860338	0.462517	0.0884706	3.04283
Iterative FC-regression	0.475166	0.306662	1.01701	-1.10598
	0.0418552	0.373598	0.106244	1.50213
MCMC	0.52315	0.5006	0.953402	-0.0376057
	0.0448123	0.0729486	0.148418	0.175335

Table 2.8: Mean estimates and their standard deviations from 25 groups of simulated stable data with selected parameters results and size n=500

<i>estimator</i>	$\alpha = 1.5$	$\beta = 0$	$\sigma = 1$	$\mu = 0$
FC-Regression	1.50816	0.00271059	1.00943	-0.0262886
	0.0646233	0.170237	0.0543769	0.138454
McCulloch	1.51016	-0.0292	0.99899	-0.05783
	0.0718287	0.176563	0.0565833	0.139363
MLE	1.3814	-0.0169418	1.02438	-0.0462738
	0.0826959	0.144085	0.0557254	0.152045
FC-Moment	1.50839	-19.1431	1.02116	-14.5376
	0.0927253	41.3347	0.0925347	27.0163
Iterative FC-regression	1.51106	-0.00476165	1.01097	-0.0299716
	0.0598664	0.163959	0.0506602	0.132402
MCMC	1.56489	-0.054	0.992534	-0.02136
	0.0745197	0.17547	0.042264	0.12367

Table 2.9: Mean estimates and their standard deviations from 25 groups of simulated stable data with selected parameters results and size n=500

<i>estimator</i>	$\alpha = 0.5$	$\beta = 0$	$\sigma = 1$	$\mu = 10$
CF-Regression	0.503946 (0.0416664)	0.0724125 (0.104725)	0.98075 (0.114787)	-0.0467315 (0.122244)
McCulloch	0.50864 (0.0291795)	-0.0144 (0.17919)	0.94614 (0.126873)	9.99988 (0.207314)
MLE	0.525522 (0.033946)	-0.0260277 (0.184907)	0.970471 (0.116694)	9.99456 (0.170739)
CF-Moment	0.496124 (0.0580103)	2.49817 (0.170561)	0.991917 (0.100128)	-2.15681 (0.601381)
Iterative CF-regression	0.474928 (0.0449769)	0.00697228 (0.128444)	1.01771 (0.130482)	9.88486 (0.602022)
MCMC	0.52058 0.0272568	0.0346 0.121372	0.978745 0.140357	9.9793 0.0841136

Table 2.10: Mean estimates and their standard deviations from 25 groups of simulated stable data with selected parameters results and size n=500

<i>estimator</i>	$\alpha = 1$	$\beta = 0$	$\sigma = 1$	$\mu = 0$
CF-Regression	0.99959 0.0587057	-0.0267108 0.10601	1.00133 0.0564359	1.87928 12.1369
McCulloch	1.0133 0.0819135	-0.0616 0.13082	0.993773 0.0807037	-1.17195 8.09328
MLE	1.01474 0.079546	-0.0549247 0.134045	1.00263 0.0905477	-1.17332 8.09307
CF-Moment	0.978291 0.098298	-9.02552 15.3926	0.985756 0.0758809	-284.767 1518.56
Iterative CF-regression	0.697322 0.22667	0.095555 0.592179	1.11277 0.104518	-23.6447 151.714
MCMC	0.94752 0.294827	-0.0404 0.154049	0.980011 0.306989	-1.93803 6.78531

Table 2.11: Mean estimates and their standard deviations from 25 groups of simulated stable data with selected parameters results and size n=1000

<i>estimator</i>	$\alpha = 0.5$	$\beta = 0$	$\sigma = 1$	$\mu = 0$
FC-Regression	0.489199	0.0217444	1.0129	-0.00476361
	0.0319408	0.0799249	0.11366	0.061289
McCulloch	0.50591	-0.0056	0.959389	0.00698019
	0.0218525	0.1421	0.126558	0.143793
MLE	0.506131	-0.00185082	0.964635	0.000190246
	0.022164	0.144572	0.122239	0.133682
FC-Moment	0.492145	0.612828	0.985277	0.509109
	0.053143	10.1172	0.0872646	17.4422
Iterative FC-regression	0.48885	0.023468	1.00843	-0.0101519
	0.0262799	0.0919269	0.0974687	0.0657819
MCMC	0.52723	0.01	0.943632	-0.0111689
	0.0137999	0.0490535	0.0730361	0.0442244

Table 2.12: Mean estimates and their standard deviations from 25 groups of simulated stable data with selected parameters results and size n=1000

<i>estimator</i>	$\alpha = 1$	$\beta = 0$	$\sigma = 1$	$\mu = 0$
FC-Regression	1.00351	0.0255765	0.999121	0.219299
	0.0511211	0.0952672	0.0518115	2.94331
McCulloch	1.02461	0.0052	1.0022	0.520606
	0.0554037	0.10697	0.0526366	6.27534
MLE	1.03124	0.00984981	1.00577	0.520582
	0.0627518	0.118616	0.0539564	6.27512
CF-Moment	0.997683	1.6313	0.997054	-121.591
	0.0476241	19.0541	0.0633366	1293.47
Iterative CF-regression	0.81558	-0.149616	1.07341	20.1319
	0.212896	0.776966	0.0976458	439.152
MCMC	1.0454	-0.013	1.06686	0.243609
	0.0747189	0.128914	0.0807836	2.7496

Table 2.13: Mean estimates and their standard deviations from 25 groups of simulated stable data with selected parameters results and size n=1000

<i>estimator</i>	$\alpha = 1.5$	$\beta = 0$	$\sigma = 1$	$\mu = 0$
FC-Regression	1.49662	-0.0335114	1.00072	-0.0271406
	0.0546952	0.127774	0.0266588	0.100014
McCulloch	1.50977	-0.0548	0.99735	-0.0482162
	0.0812103	0.131692	0.0333893	0.107477
MLE	1.48649	-0.0533887	1.02008	-0.0420542
	0.0994638	0.1323	0.053164	0.103385
CF-Moment	1.49635	-8.62833	0.997025	-6.12921
	0.0594471	45.6172	0.0476839	29.6201
Iterative CF-regression	1.49599	-0.0283467	1.00023	-0.0252189
	0.0563014	0.114953	0.026774	0.0892421
MCMC	1.47291	0.0786	1.296	-0.0178058
	0.0701458	0.137814	0.35031	0.157702

Table 2.14: Mean estimates and their standard deviations from 25 groups of simulated stable data with selected parameters results and size n=1000

<i>estimator</i>	$\alpha = 1.5$	$\beta = 0.5$	$\sigma = 1$	$\mu = 0$
CF-Regression	1.51938	0.503099	1.00126	-0.0202522
	0.0476176	0.128898	0.0315752	0.104531
McCulloch	1.53278	0.5256	0.994742	-0.0316673
	0.068071	0.100957	0.0313396	0.12937
MLE	1.4835	0.48945	1.09102	-0.0787496
	0.0948437	0.0979627	0.128446	0.122833
CF-Moment	1.52284	-37.4717	1.00304	-24.122
	0.0673976	19.7351	0.0492423	13.7309
Iterative CF-regression	1.5186	0.508533	1.00046	-0.0148481
	0.0485777	0.131164	0.0327815	0.110699
MCMC	1.55586	0.5602	0.992077	-0.0251141
	0.0535754	0.101113	0.0309532	0.0909359

Table 2.15: Table: Mean estimates and their standard deviations from 25 groups of simulated stable data with selected parameters results and size n=10000

Estimator	$\alpha = 0.5$	$\beta = 0$	$\sigma = 1$	$\mu = 0$
FC-regression	0.502867	0.00734008	0.99637	-0.00632182
	0.0161219	0.033443	0.0362375	0.0226509
McCulloch	0.50825	0.0348	0.986627	-0.0438102
	0.00844375	0.115222	0.0421706	0.122733
MLE	0.5048	0.036639	0.987171	-0.0384125
	0.00929275	0.115121	0.0408825	0.105798
FC-Moment	0.503792	0.670577	1.00451	-0.339731
	0.0181633	9.25442	0.0184257	14.5267
Iterative FC-regression	0.503484	0.00679943	0.996457	-0.00574747
	0.0163332	0.0330291	0.0364275	0.0223504
MCMC	0.52776	0.005	0.970185	-0.00221444
	0.00671615	0.0139194	0.0297314	0.0144972

Table 2.16: Table: Mean estimates and their standard deviations from 25 groups of simulated stable data with selected parameters results and size n=10000

Estimator	$\alpha = 1$	$\beta = 0$	$\sigma = 1$	$\mu = 0$
FC-regression	0.999604	-0.00558898	0.995424	-3.44274
	0.0119959	0.029460	0.0156398	13.7337
McCulloch	1.02812	-0.0036	1.00308	-0.173522
	0.0186103	0.0363868	0.0218896	1.00963
MLE	1.05438	-0.00973607	1.02346	-0.187719
	0.0331323	0.062522	0.0759836	1.00149
FC-Moment	0.997997	-7.44202	0.994112	167.113
	0.0206089	14.999	0.0248925	2591.47
Iterative FC-regression	0.837399	0.125548	1.07153	292.004
	0.144111	0.743476	0.0792006	1473.93
MCMC	1.08714	-0.019	1.15779	0.114982
	0.0208029	0.0845454	0.0575419	1.66797

Table 2.17: Table: Mean estimates and their standard deviations from 25 groups of simulated stable data with selected parameters results and size n=10000

Estimator	$\alpha = 1.5$	$\beta = 0$	$\sigma = 1$	$\mu = 0$
FC-regression	1.50007	0.0116706	1.0023	0.00704792
	0.0143106	0.0360941	0.0099203	0.025625
McCulloch	1.51718	-6.93889e-019	1.00104	-0.0123177
	0.019926	0.0387298	0.013014	0.0272368
MLE	1.37079	0.000601176	1.02523	-0.00022673
	0.0161103	0.0348947	0.0181687	0.0439337
FC-Moment	1.50013	7.96978	1.00296	4.65848
	0.0182102	44.6799	0.0164352	29.7465
Iterative FC-regression	1.50008	0.0118904	1.0024	0.00721545
	0.0143319	0.0358416	0.00990972	0.0254119
MCMC	1.55424	-0.0092	0.996717	-0.00157491
	0.0180577	0.0354342	0.00890782	0.0233374