Journal of Statistical and Econometric Methods, vol.2, no.2, 2013, 85-106
ISSN: 1792-6602 (print), 1792-6939 (online)
Scienpress Ltd, 2013

Joint robust parameter estimation for symmetric stable distributions

Csilla Csendes[1]

Abstract

In this paper we present a robust parameter estimation method for jointly estimating shape parameter α, scale parameter γ and location parameter δ of a symmetric stable distribution. The proposed estimation method is based on Probability Integral Transformation (PIT) and robust M-estimators. The procedure is ready to use as besides the theoretical description we provide the numerical algorithm and all constants and approximations of functions necessary to compute the estimators. Robust characteristics and the asymptotic behaviour of estimator $\hat{\alpha}$ is investigated. A simulation sequence was carried out to examine statistical properties of the estimator $\hat{\alpha}$. We perform an application for a data set of returns of some assets listed in Budapest Stock Exchange.

Mathematics Subject Classification: 60E07, 65C60, 62F35
Keywords: symmetric stable distributions, M-estimator, robust statistics

1 Introduction

Stable distributions are often mentioned as limit distributions of normalized

[1] Corvinus University of Budapest, Hungary

Article Info: *Received* : March 1, 2013. *Revised* : April 16, 2013
Published online : June 1, 2013

sum of independent, identically distributed (iid) random variables:

Definition 1.1. *Let $X, X_1, X_2, ...$ be iid random variables. The distribution of X is stable, if it is not concentrated at one point and if for each n there exist constants $a_n > 0$ and b_n such that*

$$\frac{X_1 + X_2 + ... + X_n}{a_n} - b_n \tag{1}$$

has the same distribution as X.

Definition 1.2. *If there exist $a_n > 0$ and b_n such that (1) converges in distribution to a non-degenerate random variable X, then X_1 is in the* domain of attraction *of X.*

Let's assume, that variables $X, X_1, X_2, ...$ have finite variance then we have the central limit theorem (CLT) and the normal law as a limit distribution. In that way stable distributions have an important role in probability theory as this is the only possible distribution family which can be the solution for the *domain of attraction problem* derived from generalized CLT, i.e., determining what random variables are attracted to what limiting random variables regardless the finite variance assumption.

Studying the stable distribution family is an active field of research both in theory and practice. For an extensive bibliography with the latest papers of stable laws we refer to Nolan's webpage [13]. Practitioners came upon stable distributions again and again in various applications at the fields of biology, physics, economics, information technology, etc. Characteristics of the observed data usually coming from summing huge number of small amounts with high or infinite variance often coincides with a stable distribution.

A wide theoretical research work was made on characterization of stable laws, some recent monographs are Zolotarev [22], Uchaikin and Zolotarev [20] and Samorodnitsky and Taqqu [19]. Interestingly, for the probability density function (pdf) of stable random variables no closed formula is known, except for the normal ($\alpha = 2$, $\beta = 0$), Levy ($\alpha = 0.5$, $\beta = 1$) and Cauchy distribution ($\alpha = 1$, $\beta = 0$). The characteristic function (chf) can be used to describe a stable distribution which depends on four parameters: characteristic exponent or index of stability $\alpha \in (0, 2]$, skewness or symmetry $\beta \in [-1, 1]$, scale $\gamma > 0$, location $\delta \in \mathbb{R}$.

The characteristic function of stable random variable Z is

$$\phi(u|\alpha,\beta,\gamma,\delta) = \mathrm{E}\exp(iuZ) = \exp(-\gamma^{\alpha}[|u|^{\alpha} + i\beta\eta(u,\alpha)] + iu\delta), \qquad (2)$$

$$\eta(u,\alpha) = \begin{cases} -(\mathrm{sign}(u))\tan(\pi\alpha/2)|u|^{\alpha}, & \text{if } \alpha \neq 1, \\ (2/\pi)u\ln|u|, & \text{if } \alpha = 1. \end{cases}$$

Besides the absence of closed pdf, another obstacle in applications is that the variance of a stable variable is infinite. Furthermore, all moments $E|X|^p$ with $p \geq \alpha$ do not exist, if $\alpha < 2$. These properties make stable distributions difficult to use in practice and force researchers to find alternative statistical methods. In [1] various practical techniques are collected primarily for data analysis of heavy-tailed data sets.

About proposed estimators of stable law parameters Weron [21] gives a well established survey providing comparison between methods by performance, reliability and computational issues. The estimation of parameters has to be based on special characteristics of stable laws since the classical maximum likelihood (ML) estimation which involves pdf cannot be computed directly. The estimators can be classified basically to procedures concerning the tail behaviour, the sample quantiles and the empirical characteristic function.

Nolan [14] presented STABLE program that includes routines for evaluation of ML estimators for all parameters numerically. For numerical calculations of pdf two approach were proposed, a method based on Fast Fourier Transform and direct integration, see Nolan [15]. The most problematic part of using ML is the very high computational demand. ML estimation has usually the best statistical properties among estimation methods, but it is certainly the slowest one. As Weron [21] points out good performance may not worth such computational effort and for real time (online) calculations this method is unusable. Advantages of ML estimation are the attractive feature of consistency and asymptotic normality.

A remarkable class of estimators are based on empirical percentiles. Among the first contributors of studying parameter estimation for stable distributions Fama and Roll [3] proposed estimators for α and γ at the symmetric case ($\beta = 0, \delta = 0$). McCulloch [12] improved their method and gave estimators for all four parameters with restriction $\alpha > 0.6$. Estimation is fast and simple to compute, but gives inaccurate results. The method can be used as an initial estimation in iterative procedures. Disadvantages of these estimators besides inaccuracy are that they include some asymptotic bias and there are restrictions for α and β.

Literature of empirical characteristic function based methods is also significant. Press [17] proposed a branch of methods based on transformations of the stable chf. His estimators are consistent but no information about efficiency is given. Moreover, properly determining auxiliary values for the algorithm is still an open question. Koutrouvelis(1980) [10] presented a regression-type estimator of the four parameters. The method concerns a linear regression model in α where the number of base points $t_k, k = 1, ..., K$ depend on the sample size and parameter α. Computational time is significant while the method requires numerical inversion of $n \times n$ matrices in a recursive algorithm. Kogon and Williams [11] improved the regression method by simplifying calculation and achieved better performance. Their method is most likely recommended according to the simulation results presented in [21].

Tail index estimators are used to determine α. This parameter is generally in focus at model definitions, because shape parameter α is the primary parameter that describes heaviness of the tails and the leptokurtic behaviour. The well-known Hill estimator, [7], and its vast number of modifications use the fact that asymptotically the tails are Pareto distributed. The drawback of these procedures is that nobody knows where the power decay starts that effect the distribution. In fact, there is no simple formula for determining which outliers should be investigated. It was shown, that power decay also depends on the parametrization, and it is a difficult function of parameter α and β, Nolan [14]. Results with acceptable estimation errors may require large samples. Borak *et. al* [2] show an example where a generated α-stable sample with $\alpha = 1.9$ resulted tail index estimation $\hat{\alpha} = 3.7$ which is an incisive critic of these procedures. These methods are very simple but should not be used solely while they can result parameter estimates far beyond the valid parameter space $0 < \alpha \leq 2$, hence indicating rejection of a stable model fit.

We mention some very recent paper which represents new methods based on Bayesian inference. Garcia *et. al* [5] uses indirect inference which involves the use of an auxiliary skewed-t distribution model, where parameters of the skewed t-distribution are associated to stable parameters and calculation is through the pseudo-likelihood of the auxiliary model. The method by Oral and Erdemir [16] uses Metropolis random walk chain and direct numerical integration.

Indeed, parameter estimation is a well studied field of theory of stable laws, but the lack of fast, easy to compute and reliable methods especially for practical applications motivated us to turn to robust statistics. In Section 2 we introduce a new estimation procedure based on the concept of robust statistics introduced by Huber[8] and Hampel *et. al* [6] called M-estimators. Our method provides joint (simultaneous) estimation of all three unknown parameters of a symmetric ($\beta = 0$) stable distribution. The motivation is to take advantage of down-weighting of outliers and known robustness properties of M-estimators.

The rest of our paper is arranged as follows: in Section 3 we give numerical approximations of the used functions in order to utilize application of the method. Section 4 is devoted to some simulation results. We investigate asymptotic normality of the estimator $\hat{\alpha}$ and present tables about performance properties. Finally, in the last section we perform an application of presented estimators for a data set of logarithmic returns of some stocks listed in Budapest Stock Exchange (BSE).

2 Method and Main Results

2.1 Estimators of location and scale

M-estimators (maximum likelihood type estimators) are generalizations of ML estimators while instead of minimizing $\sum_{i=1}^{n} -\log f(x_i)$, the M- estimator minimizes $\sum_{i=1}^{n} \rho(x_i)$, where ρ is an appropriate function. The minimization can be done by differentiation of ρ with $\psi(x) = d\rho(x)/dx$ and then solving $\sum_{i=1}^{n} \psi(x_i) = 0$. Choices of ρ and ψ yields different estimators. M-estimators were originally proposed by Huber [9].

Now let us consider the multiparameter problem defined by Huber [8] (Chapter 6) called simultaneous M-estimate of location and scale parameters $T$and S, respectively. The estimation is formulated through the implicit functions:

$$\sum \psi\Big(\frac{x_i - T_n}{S_n}\Big) = 0, \tag{4}$$

$$\sum \chi\Big(\frac{x_i - T_n}{S_n}\Big) = 0, \tag{5}$$

where x_i denotes the sample elements, T_n and S_n are estimates of T and S, ψ and χ are smooth functions, which control the weighting. This simultaneous version of M- estimation can be applied to any location-scale family of densities $\frac{1}{S} f\left(\frac{x-T}{S}\right)$ where T and S is a location and a scale parameter of the family. We define a new multiparameter M-estimator with ψ an χ functions of equations (4) and (5) for a symmetric stable variable $X \sim S(\alpha, 0, \gamma, \delta)$. Our notations follow Fegyverneki [4] and Huber [8].

Let $F(x) = F_0((x - T)/S)$, F and F_0 has the same type, F_0 is a special representative of the type. For example, with normal underlying it can be considered as the standard normal cdf Φ. Location T and scale parameter S are defined according to F_0.

We use Probability Integral Transformation (PIT) to construct a new uniform variable using the inverse distribution function method. If cdf F of a continuous random variable ξ can be inverted, then $F(\xi)$ is uniformly distributed on $[0, 1]$. From the method of moments applied to the transformed uniform variable we have:

$$E_F\Big(F_0\Big(\frac{\xi - T}{S}\Big)\Big) = \frac{1}{2}, \tag{6}$$

$$D^2_F\Big(F_0\Big(\frac{\xi - T}{S}\Big)\Big) = \frac{1}{12}. \tag{7}$$

From these moments Huber's functions $\psi(x)$ and $\chi(x)$ of equation (4) and (5) are now defined as

$$\psi(x) = F(x) - \frac{1}{2}, \tag{8}$$

$$\chi(x) = \Big(F(x) - \frac{1}{2}\Big)^2 - \frac{1}{12}, \tag{9}$$

and the equation system from (6) and (7) is

$$\sum \psi\Big(\frac{x_i - T_n}{S_n}\Big) = 0, \tag{10}$$

$$\sum \psi^2\Big(\frac{x_i - T_n}{S_n}\Big) = (n-1)\mathcal{B}, \tag{11}$$

where $\mathcal{B}$ denotes a constant. It is equal to $\frac{1}{12}$ if the sample has the same type as F_0 and it is

$$\mathcal{B} = D^2_{F_\xi}(\psi(\xi)), \tag{12}$$

otherwise.

For solving equation system (10) and (11) if $\mathcal{B}$ is known, i.e. the estimation procedure is used for a completely characterised underlying distribution function F_ξ of the sample, e.g. $\Phi(x)$, an iterative algorithm called *ping-pong method* was constructed, where only the scale and location parameters are the unknowns, see Fegyverneki [4]. The numerical algorithm is based on the modified Newton method, and convergence of approximations follows from Banach's fixpoint theorem, as the procedure is a contraction. In each step equations (13) and (14) are evaluated in turns until reaching a predefined precision.
Step for location parameter:

$$T_n^{(m+1)} = T_n^{(m)} + \frac{1}{n} S_n^{(m)} \sum_{i=1}^{n} \psi\Big(\frac{x_i - T_n^{(m)}}{S_n^{(m)}}\Big), \tag{13}$$

Step for scale parameter:

$$[S_n^{(m+1)}]^2 = \frac{1}{(n-1)\mathcal{B}} \sum_{i=1}^{n} \psi^2\Big(\frac{x_i - T_n^{(m+1)}}{S_n^{(m)}}\Big)[S_n^{(m)}]^2, \tag{14}$$

where the initial approximations are

$$T_n^{(0)} = med\{x_i\}, \tag{15}$$

$$S_n^{(0)} = C \cdot MAD, \tag{16}$$

where $med\{x_i\}$ is the median, MAD denotes median absolute deviation $MAD = med\{|x_i - med\{x_i\}|\}$, $S_n^{(m)}$ and $T_n^{(m)}$ are current estimations of S and T in step m of the iteration, respectively. Constant $C = F_0^{-1}(3/4)$ is used to have the initial estimation unbiased (F_0 is symmetric). Huber [8] shows that the median and median absolute deviation are consistent estimators of location and scale parameters, if constant C is properly chosen.

2.2 Estimator of characteristic exponent α

If we want to apply the above described method for the stable parameters, $\mathcal{B}$ involves the shape parameter α through cdf F_{ξ_α} in equation (12).

Our estimation procedure avoids direct use of the unknown α-stable cdf by a modification in $\mathcal{B}$. The modification is done by fixing in (12) the distribution function F at $\psi(x) = F(x) - 1/2$ to the known cdf's of stable family members Cauchy and normal cdf, respectively.

Definition 2.1. (Normal distribution, $\alpha = 2$)

$$f(\gamma, \delta; x) = \frac{1}{\sqrt{4\pi\gamma^2}} \exp\Big(-\frac{(x-\delta)^2}{4\,\gamma^2}\Big),$$

Definition 2.2. (Cauchy distribution, $\alpha = 1$)

$$f(\gamma, \delta; x) = \frac{1}{\pi}\frac{\gamma}{\gamma^2 + (x-\delta)^2}.$$

Denote $F_{\xi_\alpha} = F_\alpha$. Substitution of normal cdf $\Phi(x)$ and cdf of Cauchy distributions into (12) we have

$$\int_{-\infty}^{\infty} \Big(\frac{1}{\pi}\arctan x\Big)^2 dF_\alpha = \int_{-\infty}^{\infty} \Big(\frac{1}{\pi}\arctan x\Big)^2 f_\alpha(x)dx = \mathcal{B}_1(\alpha), \tag{17}$$

when F_0 is considered as standard Cauchy, and

$$\int_{-\infty}^{\infty} \Big(\Phi(x) - \frac{1}{2}\Big)^2 dF_\alpha = \int_{-\infty}^{\infty} \Big(\Phi(x) - \frac{1}{2}\Big)^2 f_\alpha(x)dx = \mathcal{B}_2(\alpha), \tag{18}$$

when F_0 is considered as standard normal.

Function $\mathcal{B}$ has a key role in the joint estimation of the parameters, while it effects the scale parameter approximation in ping-pong method. This modification adds some sort of bias in each case to the scale parameter estimation, because not the sample's correct α was used for $F_{0,\alpha}$. This bias is intentional and will be used to estimate the shape parameter. Although calculation of $\mathcal{B}$ also requires the pdf f_α and through f_α it depends on the shape parameter, both $\mathcal{B}_1(\alpha)$ and $\mathcal{B}_2(\alpha)$ functions were pre-calculated for different α's and approximated by a rational fraction function between these points, see Section 3.

For the joint estimation of the three parameters we construct another iteration around ping-pong method in an external level, in which $\hat{\alpha}$ changes systematically. Initially, both $\hat{\alpha} = 1$ and $\hat{\alpha} = 2$ are chosen, and with ping-pong method according to both functions $\mathcal{B}_1(\alpha)$ and $\mathcal{B}_2(\alpha)$ scale parameter estimations are calculated. Denote the scale parameter estimation calculated with $\mathcal{B}_1(\alpha)$ in equation (14) as $S_1(\alpha)$ and with $\mathcal{B}_2(\alpha)$ as $S_2(\alpha)$. Note, that S_1 and S_2 are dependent on α only through the $\mathcal{B}$ functions.

With initial $\hat{\alpha} = 1$ and $\hat{\alpha} = 2$ we have four scale estimation $S_1(1)$, $S_2(1)$, $S_1(2)$ and $S_2(2)$. The scale estimations involve the bias because of the modification in $\mathcal{B}$, that means for both $\alpha = 1$ and $\alpha = 2$ the two scale estimation differs, but at $\alpha = 1$ and $\alpha = 2$ the opposite one is greater. If we would choose some estimation for α, say a in interval $(1, 2)$ and calculate scale parameter $S_1(a)$ and $S_2(a)$ the difference between scale estimations would be smaller, and by getting closer to sample's α, the difference would tend to zero.

Hence, if we treat S_1 and S_2 as functions of α, then $S_1(\alpha)$ and $S_2(\alpha)$ form two concave, monotonically increasing functions on interval $\alpha \in (1, 2)$. The two scale parameter curve have only one intersection point where $S_1(a) = S_2(a)$, where the bias from choosing $F_{0,\alpha}$ *incorrectly* appears. Estimation of shape parameter is then $\hat{\alpha} = a$.

Intersection point of curves $S_1(\alpha)$ and $S_2(\alpha)$ is determined by cut-and-try method, with arbitrary precision. Estimators of location δ and scale γ are due to the last iteration of cut-and-try method. We provide the algorithm of the external iteration. It's logic is familiar to bisection search.

Algorithm

1. Set ϵ accuracy.

2. Set $a_0 = a_L = 1$ and $a_1 = a_U = 2$ and calculate scale estimations $S_1(a_L)$, $S_2(a_L)$, $S_1(a_U)$, $S_2(a_U)$.

3. The initial condition is $S_1(a_L) < S_2(a_L)$ and $S_2(a_U) < S_1(a_U)$ basically true, otherwise there is no intersection point and the method failed (return -1).

4. While $|a_{i-1} - a_i| > \epsilon$

    Set $a_i = (a_U + a_L)/2$ and calculate $S_1(a_i)$ and $S_2(a_i)$.

    If $S_1(a_i) < S_2(a_i)$, then set $a_L = a_i$, else set $a_U = a_i$.

5. return a_i as $\hat{\alpha}$

Computation time of cut-and-try method is logarithmic ($\log_2$), as in each iteration we consider only the half of the interval of the previous iteration. However, we calculate approximations S_1 and S_2 only at discrete points found by cut-and-try algorithm, $S_1(\alpha)$ and $S_2(\alpha)$ were perceived as continuous functions of α. In Figure 1. some scale parameter curves are plotted in case of various sample sizes and α parameters. One can see, how the procedure become more accurate, when the sample size is increased.

Our method has some limitations. It considers only symmetric α-stable distributions, where $1 \leq \alpha \leq 2$. Usage of Cauchy and normal distribution highly lean on symmetry. Our method is probably extendible to the case where $\alpha < 1$, at least approximations can be computed numerically. However, in this case the expectation does not exist, that's why for the practical point of view this case has much more lower relevance.[2]

Fegyverneki [4] shows results about the convergence of M-estimators of location and scale parameter. Let $\xi = \sigma\eta + \mu$ where the distribution of variable η is $F_0(x)$. Let T_n and S_n the estimator of location $\mu \in \mathbb{R}$ and scale parameter $\sigma > 0$, respectively. Given the sample $\xi_1, \xi_2, ..., \xi_n$ and cdf. $F_0(x)$, the distribution of ξ_i is $F_0\big((x-\mu)/\sigma\big)$.

Theorem 2.3. *(Fegyverneki [4], Theorem 2.) If $F_0(x)$ is differentiable, strictly monotone increasing and $F_0(0) = 0.5$ then the two-dimensional joint distribution of (T_n, S_n) converges to a normal one,*

$$\sqrt{n}((T_n, S_n) - \mu, \sigma) \to N(0, \Sigma).$$

Covariance matrix Σ is also given in Fegyverneki [4]. Note that convergence to normality alters according to type F_0 because covariance matrices are depend on the underlying distribution.

[2] Generally the expectation play a significant role in real data models, for example in portfolio selection problem the expectation coincides the return or log-return of the investment and if $\alpha < 1$, then the problem has no solution in the sense that expected return can not be infinite.

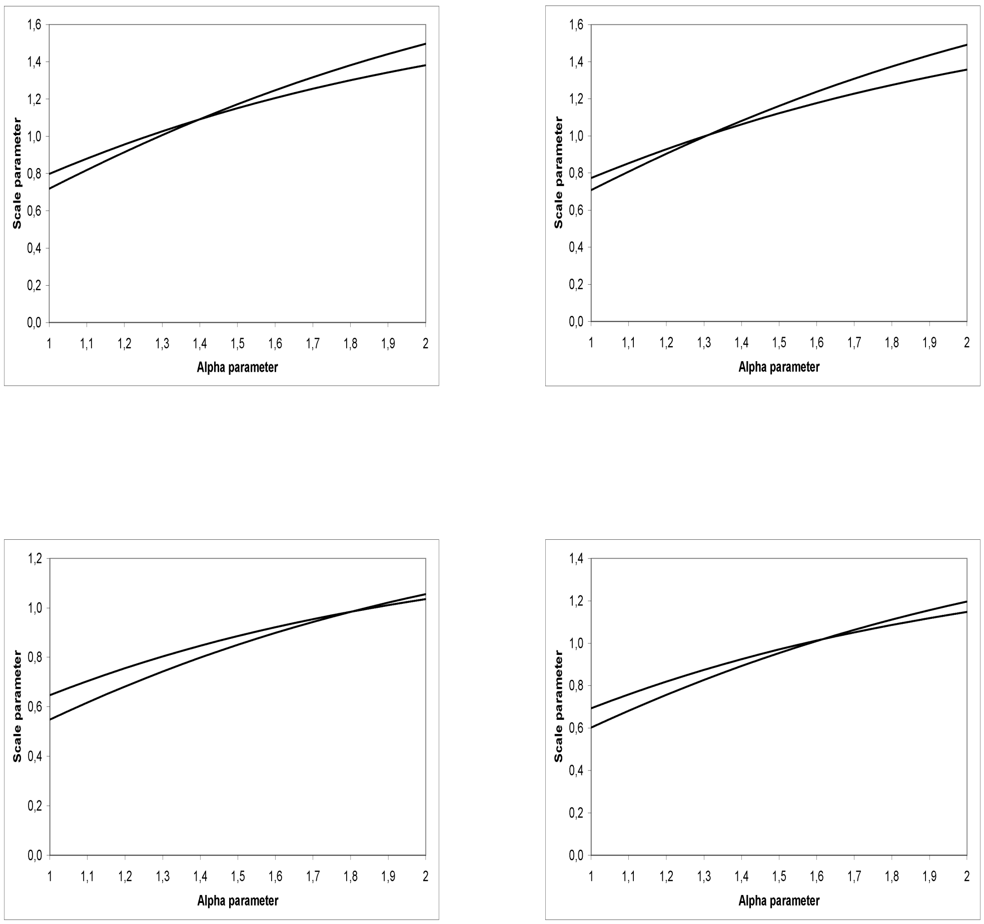

Figure 1: $S_1(\alpha)$ and $S_2(\alpha)$ scale parameter curves computed with ping-pong method with modified $\mathcal{B}$ functions, $\mathcal{B}_1(\alpha)$ and $\mathcal{B}_2(\alpha)$ (I. sample size 1000 elements, $\alpha = 1.3$(top-left); II. sample size 3000 elements, $\alpha = 1.3$ (top-right); III. sample size 3000 elements, $\alpha = 1.8$ (bottom-left); IV. sample size 3000 elements, $\alpha = 1.6$ (bottom-right)

3 Numerical Approximations of Functions $\mathcal{B}$

In this section we present numerical approximations of functions $\mathcal{B}$ which is necessary to apply the proposed PIT estimation procedure. While in equation (17) and (18) expressions involve the α-stable density, we are faced to the problem of unknown α-stable pdf. Obviously, $\mathcal{B}_1(\alpha)$ and $\mathcal{B}_2(\alpha)$ have to be approximated numerically. To avoid high computational demand of numerical integration that could extremely slow down the algorithm, a polynomial approximation for $\mathcal{B}$ s was determined.

The rational fractional function had to be determined only once, hence, no further time consuming computation is needed. Evaluation of the algorithm requires no integration, that's why faster than the proposed methods which calculates the density during the procedure.

Recall, that function $\mathcal{B}$ is the expectation of $(\psi(x))^2$, hence can be approximated with the mean because of the law of large numbers:

$$\mathcal{B}_1(\alpha) \approx \frac{1}{n}\sum_{i=1}^{n}(\Phi(x_i) - 0.5)^2, \tag{19}$$

$$\mathcal{B}_2(\alpha) \approx \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\pi}\arctan x_i\right)^2. \tag{20}$$

Table 1: Values for $\mathcal{B}_1$ and $\mathcal{B}_2$ depend on α

α	$\mathcal{B}_1(\alpha)$	$\mathcal{B}_2(\alpha)$
1	0.0833333333333333	0.126807877965645
1.1	0.0758844534723818	0.118966259082521
1.2	0.0697612892957584	0.112284032323310
1.3	0.0646999988570841	0.106570898511029
1.4	0.0604648399039825	0.101682622480835
1.5	0.0569093151515006	0.097443890906153
1.6	0.0538933607717261	0.093798682214659
1.7	0.0513226066667932	0.090637100836610
1.8	0.0491126022363082	0.087875629036068
1.9	0.0472087085432832	0.085445768785679
2	0.0455654051822800	0.083333333333333

In order to approximate (19) and (20) samples with parameters $\alpha = 1, \alpha = 1.1, \alpha = 1.2, ..., \alpha = 2$ of 20 million elements were generated. Table 1. shows calculated values of $\mathcal{B}_1(\alpha)$ and $\mathcal{B}_2(\alpha)$ in $\alpha = 1, \alpha = 1.1, \alpha = 1.2, ..., \alpha = 2$ base points. Afterwards, a rational fraction function approximation for $\mathcal{B}_1(\alpha)$ and

Table 2: Coefficients for the rational fraction function

coef.	$\mathcal{B}_1$	$\mathcal{B}_2$
b_3	-3.83008202167381	-5.44424585925350
b_2	4.78393407388667	8.41128641608921
b_1	-2.07519730244991	-0.91519048820337
b_0	0.18011293964047	-4.11676503125739
a_5	0.00557315701358	0.02536047583564
a_4	-0.02655295929925	-0.20581738159343
a_3	0.12169846973572	0.84848196461615
a_2	-0.31598423581221	-2.18605774383017
a_1	0.34060543137722	3.06692780009580
a_0	-0.12044269481921	-1.68393473522529

$\mathcal{B}_2(\alpha)$ were constructed, based on the base points of Table 1. The approximation formula is

$$\mathcal{B}_i(x) = \frac{a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0}{x^4 + b_3x^3 + b_2x^2 + b_1x + b_0}, \tag{21}$$

where $i = 1, 2$. Coefficients $a_5, ..., a_0, b_3, ..., b_0$ were calculated from the linear equations system build up from data in Table 1. The equations system is overdetermined, while we have twelve equations and eleven unknown variables. We solved the system of eleven equations by eliminating one equation. Table 2 contains the coefficients for approximations of $a_5, ..., a_0, b_3, ..., b_0$ from the system with the smallest least squares differences from the twelve base points.

4 Simulation Study

A Monte-Carlo simulation sequence was made to examine statistical properties of estimators $\hat{\alpha}$, $\hat{\gamma}$ and $\hat{\delta}$. The examined α values were $1.3, 1.5, 1.7$, in each case samples were generated with $n = 50$, 100, 400, 2500 elements, and Monte-Carlo replications were $r = 100$, $r = 400$ and $r = 2500$, respectively.

Random number generation of standard ($\gamma = 1$, $\delta = 0$) α-stable samples was computed according to the formula by Zolotarev [22]. An α - stable variable Z is generated as

$$Z(\alpha,0)=\frac{sin(\alpha\xi)}{(cos\xi)^{\frac{1}{\alpha}}}\left(\frac{cos((1-\alpha)\xi)}{\eta}\right)^{\frac{1-\alpha}{\alpha}}, \tag{22}$$

where η is a standard exponential variable, ξ is uniform on $(-\pi/2,\pi/2)$. For standardized variables $Z(\alpha,0)/\alpha^{\frac{1}{\alpha}}$ can be used.

Our simulations showed that if parameter α is near the endpoints of interval $[1,2]$ and the sample is small i.e. less than 100 elements, it is possible, that there is no intersection point because of random effect and inaccuracy of estimation. Presence of this problem in the simulation could be reduced by increasing the number of elements of the sample. In practice if the method fails we suggest to fit a normal or Cauchy model to the data set. Table 3. shows numbers of valid estimation results in our simulation study.

Table 3: Number of valid estimators (r: Monte-Carlo repl., n: sample size)

	r=100			r=400			r=2500		
α	1.3	1.5	1.7	1.3	1.5	1.7	1.3	1.5	1.7
n=50	94	99	91	382	388	357	2389	2429	2244
n=100	97	100	99	395	399	387	2479	2493	2443
n=400	100	100	100	400	400	400	2499	2500	2500
n=2500	100	100	100	400	400	400	2500	2500	2500

We examined asymptotic normality of the estimators $\hat{\gamma}$, $\hat{\delta}$, $\hat{\alpha}$ by χ^2 goodness-of-fit test. We summarize resulted p-values in Table 4. for only $\hat{\alpha}$, while convergence to normality of the scale and location estimator is proven theoretically. In fact, we are interested in the normality of the estimator $\hat{\alpha}$. Tests were based on originally 10 bins and calculation was done by function '*chi2gof*' of MATLAB Software package. We can accept normality if the chosen Type I. error is smaller than the p-value, generally 0.05. Normality of the estimator is acceptable in most of the simulation cases, except with small samples ($n=50, n=100$) replicated 2500 times. Rejection of normality is maybe because of the asymmetry of the samples as they are trimmed near 1 ($\alpha=1.3$ case) or 2 ($\alpha=1.7$ case).

We also present results on the performance of our new estimators. The differences $(\overline{\hat{\alpha}}-\alpha)$, $(\overline{\hat{\gamma}}-\gamma)$, $(\overline{\hat{\delta}}-\delta)$, the standard deviations, and correlation of each estimator pairs are calculated. The estimation procedure uses the scale estimator in the iteration to estimate α, hence it is expectable, that among $\hat{\alpha}$ and $\hat{\gamma}$ some relation would be found. Correlation coefficients are around $0.4-0.6$. In Table 5. and Table 6. we give results for the case where number of Monte-Carlo replications was $r=2500$. Table 7. contains the asymptotic ($r=2500$) mean squared errors (MSE) for both parameters.

Table 4: p-values to testing for normality with χ^2 test of par. estimate $\hat{\alpha}$

		n=50	n=100	n=400	n=2500
$\alpha=1.3$	r=100	0.4851	0.3224	0.3312	0.0350
	r=400	0.0001	0.0180	0.4774	0.0076
	r=2500	0.0000	0.0000	0.2303	0.0000
$\alpha=1.5$	r=100	0.0969	0.7050	0.7454	0.2131
	r=400	0.3654	0.1979	0.0694	0.7535
	r=2500	0.0005	0.0236	0.0666	0.9928
$\alpha=1.7$	r=100	0.7340	0.9458	0.2714	0.6293
	r=400	0.0000	0.2394	0.5478	0.1210
	r=2500	0.0000	0.0002	0.3554	0.4432

5 Application

We present an application of our PIT estimation method to a financial data set of daily closing prices of some assets. Similar applications are widely accepted in literature, as the characteristics of price change data was proved many times to be well fitted to symmetric α-stable distributions, Rachev and Mittnik[18].

Assets investigated are the dominant papers at Budapest Stock Exchange (BSE): BUX (official index), OTP, Richter, Egis, Magyar Telekom (MTelekom), MOL. Samples contain around 2700 observations except OTP, that paper is investigated from March 2002, the other papers from January, 2001. If a daily closing price was missing from the data set, the price change was calculated from the price of the following day.

Table 5: Accuracy for estimates $\hat{\alpha}$, $\hat{\gamma}$, $\hat{\delta}$, Monte-Carlo replication $r = 2500$

α	n	$\bar{\hat{\alpha}} - \alpha$	$\bar{\hat{\gamma}} - \gamma$	$\bar{\hat{\delta}} - \delta$
1.3	50	0.06246	0.05074	-0.00018
	100	0.02878	0.02032	-0.00026
	400	0.00664	-0.00081	0.00013
	2500	0.00192	-0.00699	0.00033
1.5	50	0.03528	0.02199	-0.00226
	100	0.02435	0.01191	-0.00067
	400	0.00840	0.00362	0.00104
	2500	0.00133	0.00041	0.00011
1.7	50	-0,00452	0.00542	-0.00944
	100	0.01020	0.00968	0.00227
	400	0.00524	0.00152	-0.00040
	2500	0.00012	0.00098	-0.00044

Asset price changes are examined in the logarithmic model, called continuously - compounded rate of returns:

$$r_{log} = \ln \frac{P_{i+1}}{P_i}, \tag{23}$$

where P_i denotes the price of the asset in time i. Note, that another possibility is the use of the so called discrete returns $r_{discrete} = (P_{i+1} - P_i)/P_i$, but for small returns, logarithmic and discrete returns are close, if $x \approx 0$ then $\log(x+1) \approx x$.

In Table 8. one can see estimations of parameters $\hat{\alpha}, \hat{\gamma}, \hat{\delta}$ assuming that price changes follow a symmetric stable distribution. Confidence intervals for estimation is with $\sigma(\hat{\alpha}) = 0.03$ according to Table 6. and $n \approx 2500$ is $\hat{\alpha} \pm 1.96 * (0.03)/\sqrt{2500} = \hat{\alpha} \pm 0.0113$. Based on calculated $\hat{\alpha}$ parameters we can classify the assets into a riskier (BUX, Egis, OTP) and a less riskier group (MOL, MTelekom, Richter), as a lower α value means a more volatile and hence risky investment.

In Figure 2, histograms of logarithmic returns computed from daily closing prices of the assets with normal curves are displayed. It is seen, that the data is not really fit the normal curve. We tested both the normal nullhypothesis and the α-stable nullhypothesis with χ^2 and Kolmogorov-Smirnov (KS) type goodness-of fit tests using our estimated parameter values. Data sets of returns

Table 6: Standard deviation and correlation coefficients for estimates $\hat{\alpha}$, $\hat{\gamma}$, $\hat{\delta}$, Monte-Carlo replication $r = 2500$

α	n	$\sigma(\hat{\alpha})$	$\sigma(\hat{\gamma})$	$\sigma(\hat{\delta})$	$r_{\alpha,\gamma}$	$r_{\alpha,\delta}$	$r_{\gamma,\delta}$
1.3	50	0.20375	0.21203	0.19909	0.56971	-0.01833	0.02058
	100	0.15317	0.15684	0.13783	0.62494	0.01291	0.02614
	400	0.07359	0.07872	0.06765	0.60536	-0.02671	-0.00753
	2500	0.02979	0.03183	0.02716	0.36814	-0.00482	-0.00743
1.5	50	0.21196	0.18609	0.17820	0.54814	-0.00804	-0.00410
	100	0.15722	0.13100	0.12392	0.59489	-0.01344	-0.01445
	400	0.07691	0.06304	0.06139	0.56732	-0.01081	-0.02274
	2500	0.03081	0.02546	0.02456	0.58733	0.00812	0.00336
1.7	50	0.18222	0.15643	0.15853	0.48490	-0.02716	-0.02113
	100	0.14298	0.10729	0.11421	0.50638	0.01945	-0.02351
	400	0.07412	0.05346	0.05642	0.51379	-0.00698	0.00765
	2500	0.02951	0.02148	0.02215	0.51498	0.00720	0.03302

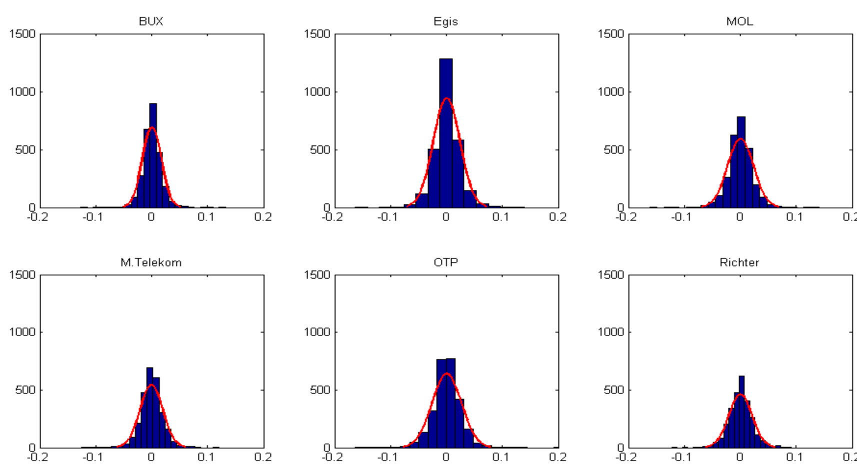

Figure 2: Histograms of logarithmic returns of some stocks listed in BSE (2001-2010) computed from daily closing prices (with fitted normal curve)

are standardized with the mean and standard deviation, and location and scale parameter, respectively, and was tested against standard cdf's. Because precise critical values for the stable underlying hypothesis were not available for us,

Table 7: Asymptotic MSE values of PIT estimators of α, γ, δ (Monte Carlo repetition=2500)

α	n	$MSE(\hat{\alpha})$	$MSE(\hat{\gamma})$	$MSE(\hat{\delta})$
$\alpha = 1.3$	50	0.045396	0.047511	0.039618
	100	0.024279	0.025002	0.018990
	400	0.005457	0.006195	0.004574
	2500	0.000891	0.001062	0.000737
$\alpha = 1.5$	50	0.046154	0.035099	0.031747
	100	0.025300	0.017296	0.015351
	400	0.005984	0.003986	0.003768
	2500	0.000950	0.000648	0.000603
$\alpha = 1.7$	50	0.033211	0.024488	0.025209
	100	0.020538	0.011600	0.013044
	400	0.005520	0.002859	0.003182
	2500	0.000870	0.000462	0.000490

we used the commonly known tables for KS test.

The χ^2 test resulted 0.0000 p-value for all asset in case of testing for normality as we expected. For the tested α-stable goodness-of-fit the p-values are overall greater, increased in compare to the normal case, but they not always exceed 0.05. Hence, the α-stable fit is not perfect, but the well-known drawbacks of the χ^2 test as by grouping it can mask over tail behaviour and number of the intervals and the edges of grouping can have an effect on the statistic may indicate this phenomena.

On the other hand, the KS test gave reasonable results. The absolute differences between empirical cdf and hypothetised cdf's according to KS test are smaller in the case of stable nullhypothesis than in the case of normal nullhypothesis for every investigated asset. Apart from a few paper, there is no reason to reject the stable model fit. Table 9. and Table 10. show results of performed χ^2 and KS test for the distribution fitting of logarithmic returns. We also note, that empirical studies of such financial time series data are based on either on a time series approach either an independent, identically distributed model. Here, we did not consider elements of time series theory

Table 8: Parameters estimated with PIT method assuming symmetric stable laws of logarithmic returns from BSE data set

Asset	$\hat{\alpha}$	$\hat{\gamma}$	$\hat{\delta}$
BUX	1.7525	0.0132	0.0003
Egis	1.6890	0.0171	0.0002
MOL	1.7265	0.0176	0.0004
M. Telekom	1.7518	0.0162	-0.0004
OTP	1.7062	0.0203	0.0007
Richter	1.7977	0.0175	0.0003

i.e. autocorrelation or partial autocorrelation functions of the data.

Table 9: Resulted p-values and test statistic values with χ^2 goodness-of-fit test to estimated stable parameters of BSE data set

	normal fit		stable fit	
Asset	p -value	test statistic	p -value	test statistic
BUX	0.0000	23.2047	0.3441	10.0796
Egis	0.0000	69.7031	0.0001	35.1439
Mol	0.0000	130.4229	0.0167	20.2015
M. Telekom	0.0000	98.4660	0.1516	13.2507
OTP	0.0000	109.4452	0.7635	5.7621
Richter	0.0000	118.9237	0.0120	21.1394

Table 10: Resulted p-values and test statistic values with Kolmogorov-Smirnov goodness-of-fit test to estimated stable parameters of BSE data set

	normal fit		stable fit		
Asset	p -value	teststat.	p -value	teststat.	critical value
BUX	0.0000	0.0515	0.6246	0.0144	0.0260
Egis	0.0000	0.0687	0.0676	0.0249	0.0260
Mol	0.0000	0.0576	0.3997	0.0171	0.0261
M. Telekom	0.0000	0.0531	0.0589	0.0254	0.0260
OTP	0.0000	0.0655	0.6236	0.0161	0.0291
Richter	0.0004	0.0395	0.0481	0.0262	0.0260

6 Summary

In this paper we presented a parameter estimation method which is used to determine α, γ, δ parameters of a symmetric stable distribution jointly. We used Probability Integral Transformation to derive our procedure, which is a variant of M-estimates. Characteristics of the estimation procedure is investigated in a Monte-Carlo simulation sequence. Our method is similar in performance to other proposed methods in literature. The method is applied to a financial data set and it was shown that it can be useful for stable portfolio optimization. Further work is to investigate asymptotic behaviour of parameter estimation more theoretically.

Acknowledgements. The described work was carried out as part of the TÁMOP-4.2.2/$B-10/1-2010-0008$ project in the framework of the New Hungarian Development Plan. The realization of this project is supported by the European Union, co-financed by the European Social Fund.

References

[1] J.R. Adler, R.E. Feldman and M.S. Taqqu (Editors), *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, Birkhauser, Boston, 1998.

[2] Sz. Borak, W. Härdle and R. Weron, Stable Distributions, *SFB 649 Discussion Papers*, SFB649DP2005-008, Humboldt University, Berlin, Germany, (2005).

[3] E.F. Fama and R. Roll, Parameter Estimates of Symmetric Stable Distributions *J. Amer. Statis. Assoc.*, **66**, (1971), 331-338.

[4] S. Fegyverneki, Robust Estimators and Probability Integral Transformations, *Math. Comput. Modelling*, **38**, (2003), 803-814.

[5] R. Garcia, E. Renault and D. Veredas, Estimation of Stable Distributions by Indirect Inference, *Journal of Econometrics*, **161**, (2011), 325-337.

[6] F.R. Hampel, E.M. Ronchetti, P.J. Rousseeuw and W.A. Stahel, *Robust Statistics - The Approach Based on Influence Functions*, John Wiley & Sons, New York, 1986.

[7] B.M. Hill, A Simple General Approach to Inference about the Tail of a Distribution, *Annals of Statistics*, **3**, (1975), 1163-1174.

[8] P.J. Huber, *Robust Statistics*, John Wiley & Sons, New York, 1981.

[9] P.J. Huber, Robust Estimation of a Location Parameter, *Ann. Math. Statist.*, **35**, (1964), 73-101.

[10] I.A. Koutrouvelis, Regression-type Estimation of the Parameters of Stable Laws, *J. Amer. Statis. Assoc.*, **75**, (1980), 918-928.

[11] S.M. Kogon and D.B. Williams, *Characteristic Function Based Estimation of Stable Parameters*, in R. Adler, R. Feldman, M. Taqqu (eds.), *A Practical Guide to Heavy Tails*, Birkhauser, 311-335, 1998.

[12] J.H. McCulloch, Simple Consistent Estimators of Stable Distribution Parameters, *Commun. Statist. - Simula.*, **15**(4), (1986), 1109-1136.

[13] J.P. Nolan's web page: http://academic2.american.edu/ jpnolan/stable/StableBibliography.pdf

[14] J.P. Nolan, Maximum Likelihood Estimation of Stable Parameters, in O.E. Barndorff-Nielsen, T. Mikosch and S.I. Resnick (Eds.), *Levy Processes: Theory and Applications*, Birkhauser, Boston, 379-400, 2001.

[15] J.P. Nolan, Numerical Calculation of Stable Densities and Distribution Functions, *Commun. in Statist. - Stochastic Models*, **13**, (1997), 759-774.

[16] E. Oral and C. Erdemir, A Bayesian Estimation of Stable Distributions, *Journal of Statistical and Econometric Methods*, **1**(3), (2012), 39–52.

[17] S.J. Press, *Applied Multivariate Analysis*, Holt, Rinehart and Winston, New York, 1972.

[18] S.T. Rachev and S. Mittnik, *Stable Paretian Models in Finance*, Wiley, New York, 2000.

[19] G. Samorodnitsky and M. Taqqu, *Stable Non-Gaussian Random Processes*, Chapman and Hall, New York, 1994.

[20] V.V. Uchaikin and V.M. Zolotarev, *Chance and Stability - Stable Distributions and their Applications*, VSP, Utrecht, 1999.

[21] R. Weron, Performance of the Estimators of Stable Law Parameters, *Research Report HSC/95/1*, Hugo Steinhaus Center for Stochastic Methods, (1995).

[22] V.M. Zolotarev, *One Dimensional Stable Distributions*, American Mathematical Society, Providence, 1986.