SEMIPARAMETRIC ESTIMATORS FOR HEAVY TAILED DISTRIBUTIONS

Valeria Caviezel*
Department MSIA
University of Bergamo,
Via dei Caniana, 2 - 24127-Bergamo- Italy
E-mail: valeria.caviezel@unibg.it

Sergio Ortobelli
Department MSIA
University of Bergamo,
Via dei Caniana, 2 - 24127-Bergamo- Italy
E-mail: sergio.ortobelli@unibg.it

Prof. Svetlozar (Zari) T. Rachev
Chair of Econometrics, Statistics and Mathematical Finance
School of Economics and Business Engineering
University of Karlsruhe and KIT,
Kollegium am Schloss, Bau II, 20.12, R210
Postfach 6980, D-76128, Karlsruhe, Germany and
Department of Statistics and Applied Probability
University of California, Santa Barbara
CA 93106-3110, USA and
Chief Scientist of FinAnalytica INC.
E-mail: rachev@statistik.uni-karlsruhe.de

Abstract: In this paper we describe and apply estimating function theory to evaluate the parameters of parametric distributions uniquely defined by their characteristic functions. We first implement an estimating function model based on the first four moments of a parametric function of the underlying random variables. For instance we propose two parametric functions of the underlying random variables so as to obtain its first moments more easily by the simple knowledge of the characteristic function. Thus we consider the estimates that present the minimal asymptotic variance with respect to the parameter of the function. Then we propose an empirical analysis based on simulated stable Paretoian distributions. Using simulated data of stable distributions we evaluate the forecasting power of the proposed methodology comparing it with the analogous maximum likelihood estimates. Moreover we show how to apply the same methodology to some well-known infinitely divisible distributions.

Key words: Estimating function, Stable distributions, infinitely divisible distributions, asymptotic variance.

*Corresponding author
1 Introduction

Several empirical investigations have shown that empirical distributions of many observed economic and financial data deviate from the ideal Gaussian law, since they often exhibit skewness and fat tails. Thus many alternative distributions have been proposed to approximate the underlying data. The classical example is the case of infinitely divisible distributions uniquely defined by their characteristic functions. Infinitely divisible distributions take into account the skewness and kurtosis of return series. In addition their associated Lévy processes can be used for valuing intertemporal financial products such as the derivatives. Thus several Lévy processes have been widely used in recent financial literature (see, among others, Rachev and Mittnik (2000) and the references therein). In particular, the stable Pareto distributions probably present the most attractive modeling properties, since they not only provide a better empirical fit, but they also posses heavy tails and result as limit distribution of sums of i.i.d. random variables.

In this paper, we deal with the problem of valuing the parameters of distributions uniquely defined by their characteristic function. Typically, if we know the characteristic function, we can determine the approximated distribution by inverting the characteristic function with the Fast Fourier Transform (FFT) and then we can value the maximum likelihood estimates (MLE) of the parameters. However, even if this methodology is widely used it does not permit an a priori valuation of the efficiency of the estimates. In this paper we propose a semiparametric methodology that gives optimal estimates of distribution parameters based on a valuation a priori of the asymptotic variance of the estimates. The semiparametric valuation is based on estimating function (EF) theory (see Godambe (1991) and the references therein). We suggest using this estimation either for the moments curve or for a bounded parametric function of the underlying distribution that admits finite the first four moments. Thus we propose to minimize the asymptotic variance of the estimates subject to some estimating equations. Finally we value the forecasting power of some of these semiparametric estimators applied to stable Pareto distributions. In particular, we compare the absolute difference between simulated data, and either EF estimates or MLE ones obtained for stable Pareto parameters. We also show how to use the same methodology for some particular infinitely divisible distributions.

The paper is organized as follows: Section 2 introduces semiparametric estimators based on estimating function theory, and in Section 3 we compare semiparametric and maximum likelihood methods. Finally, we briefly summarize the results.

2 Estimating Function parameter estimation

Suppose we have a sample $X = (X_1, X_2, ..., X_T)$ of i.i.d. observations whose distribution family $\mathcal{F}(\theta)$ is parametrized by $\theta = (\theta_1, \theta_2, ..., \theta_p) \in B \subseteq \mathbb{R}^p$. In
the theory of estimating functions the optimum has two components: "unbiasedness" of EF and "smallness of the variance" of the standardized EFs. An estimating function \( h_i(X, \theta) \) is called unbiased if \( E(h_i(X, \theta)) = 0 \) for all admissible \( \theta \). In particular, when we consider a single parameter \( \theta \) the score function \( \frac{\partial \log f(X, \theta)}{\partial \theta} \) represents a typical example of unbiased estimating function. We say that the unbiased estimating functions \( h_i(X_s, \theta) \) are mutually orthogonal, when \( E(h_i(X_s, \theta)h_j(X_s, \theta)) = 0 \) for every \( i \neq j; \ i, j = 1, ..., n \). Among all the linear combinations \( l_{\theta,k} = \sum_{s=1}^{T} \sum_{i=1}^{n} a_{k,i}(\theta)h_i(X_s, \theta), \ (k = 1, ..., p \) of unbiased, mutually orthogonal estimating functions \( h_i(X_s, \theta) \), the estimating functions \( l_{\theta,k} = \sum_{s=1}^{T} \sum_{i=1}^{n} a^{*}_{k,i}(\theta)h_i(X_s, \theta) \) with coefficients \( a^{*}_{k,i}(\theta) = E\left(\frac{\partial h_i}{\partial \theta_k}\right)/E(h^2_i) \), \( \forall i = 1, ..., n; \forall k = 1, ..., p \), are optimal (with minimum variance). According to estimating function theory the optimal EFs \( \hat{\theta}_{(T)} = \left[\hat{\theta}_{(T),1}, ..., \hat{\theta}_{(T),p}\right] \) obtained as consistent solution of the system: \( l_{\theta,k}^* = 0, \ k = 1, ..., p \); after orthogonalization, standardization and optimal combination is asymptotically Gaussian, i.e.,

\[
\sqrt{T}\left(\hat{\theta}_{(T)} - \theta\right) \rightarrow MVN(0, \text{Var}_{\text{EF}}(\theta)) \text{ for } T \rightarrow \infty,
\]

where \( \text{Var}_{\text{EF}}(\theta) = [v_{i,j}]_{i,j=1,...,p} \) and \( v_{i,j} = E\left(\frac{\partial \theta_i}{\partial \theta_j}\right) \); \( i, j = 1, ..., p \). These results can be extended to stationary series that are not necessarily independent (see Li and Turtle (2000)). Moreover, we get similar results using GMM (General Method of Moments), but in this case we need a recursive methodology to determine the parameters in the linear combination \( l_{\theta,j}^* \). Thus any algorithm based on GMM needs additional computational time.

Typical examples of optimal estimating functions are those proposed by Godambe and Thompson (1989) where they use two unbiased and mutually orthogonal estimating functions: \( h_1(X_s, \theta) = f(X_s) - m_f(\theta) \) and \( h_2(X_s, \theta) = \left( f(X_s) - m_f(\theta) \right)^2 - \sigma^2_f(\theta) - s_f(\theta)\sigma_f(\theta) \left( f(X_s) - m_f(\theta) \right) \), where \( f \) is a measurable real function, \( m_f(\theta) = E(f(X_s)), \ \sigma^2_f(\theta) = E\left(\left(f(X_s)\right)^2\right) - m^2_f(\theta), \) and \( s_f(\theta) = \frac{E\left(\left(f(X_s) - m_f(\theta)\right)^3\right)}{\sigma^3_f(\theta)}. \) Therefore, we get the EF optimal estimator \( \hat{\theta}_{(T)} \) of the vector of parameters \( \theta \) solving the equations (for \( k = 1, ..., p \))

\[
l_{\theta,k}^* = \sum_{s=1}^{T} \left( a_{k,1}h_1(X_s, \theta) + a_{k,2}h_2(X_s, \theta) \right) = 0,
\]

where

\[
a_{k,1} = \frac{E\left(\frac{\partial h_1(X_s, \theta)}{\partial \theta_k}\right)}{E(h^2_1(X_s, \theta))} = -\frac{\partial m_f(\theta)}{\partial \theta_k} \sigma^2_f(\theta),
\]

\[
a_{k,2} = \frac{E\left(\frac{\partial h_2(X_s, \theta)}{\partial \theta_k}\right)}{E(h^2_1(X_s, \theta))} = \frac{\partial \sigma^2_f(\theta)}{\partial \theta_k} + \frac{\partial m_f(\theta)}{\partial \theta_k}s_f(\theta)\sigma_f(\theta)\left(k_f(\theta) - 1 - \sigma^2_f(\theta)\right),
\]
and \( k_f(\theta) = \frac{E \left( (f(X_s) - m_f(\theta))^4 \right)}{\sigma^4_f(\theta)} \). We call this class of estimating functions GT (Godambe and Thompson) estimating functions. Note that we can also easily obtain an analytical formulation of the asymptotic variance of GT estimating functions when we know \( \frac{\partial m_f(\theta)}{\partial \theta_k} \) and \( \frac{\partial^2 \sigma_f^2(\theta)}{\partial \theta_k^2} \). As a matter of fact, from (1) we deduce that the asymptotic variance \( V_{EF}(\theta) \) is the inverse of \( V_{EF}(\theta) = [v_{ij}(\theta)]_{i,j=1,...,p} \), where \( v_{ij}(\theta) = E \left( \frac{\partial^2}{\partial \theta_i \partial \theta_j} f(\theta) \right) \). Thus, we get

\[
v_{ij}(\theta) = T \frac{\partial m_f(\theta)}{\partial \theta_i} \frac{\partial m_f(\theta)}{\partial \theta_j} +\]

\[+ T \left( -\frac{\partial^2 \sigma_f^2(\theta)}{\partial \theta_i \partial \theta_j} + \frac{\partial m_f(\theta)}{\partial \theta_i} \sigma_f(\theta) \right) \left( -\frac{\partial^2 \sigma_f^2(\theta)}{\partial \theta_j \partial \theta_j} + \frac{\partial m_f(\theta)}{\partial \theta_j} \sigma_f(\theta) \right) \]

(4)

It has been demonstrated (see Ortobelli and Topaloglou (2008)) that the solutions \( l_{\theta,k}^0 = 0 \) are given by the values \( \theta_k \), solutions of the simplified equations for \( k = 1, \ldots, p \):

\[
m_f(\theta) = \begin{cases} \frac{1}{T} \sum_{s=1}^T f(X_s) + c_k - b_k & \text{if } c_k > 0 \\ \frac{1}{T} \sum_{s=1}^T f(X_s) + c_k + b_k & \text{if } c_k \leq 0 \end{cases}
\]

(5)

where

\[
b_k = \left( \frac{2}{c_k^2} - \frac{1}{T} \sum_{s=1}^T \left( f(X_s) - \frac{1}{T} \sum_{s=1}^T f(X_s) \right)^2 + \sigma_f^2(\theta) \right)^{1/2}, \quad \text{and} \quad c_k = \frac{a_{m,1} - a_{m,2} s_f(\theta) \sigma_f(\theta)}{2 a_{k,2}}.
\]

In particular, when \( m_f(\theta) \) is itself a possible parameter, then a consistent estimating function of \( m_f(\theta) \) is given by (5) where \( c_k = \frac{a_{m,1} - a_{m,2} s_f(\theta) \sigma_f(\theta)}{2 a_{k,2}} \). According to formulas (2) (3), the above estimating equations (5) are uniquely determined by \( m_f(\theta) = \sigma_f^2(\theta) = s_f(\theta) = k_f(\theta) = \frac{\partial m_f(\theta)}{\partial \theta_k} \), and \( \frac{\partial^2 \sigma_f^2(\theta)}{\partial \theta_k \partial \theta_k} \) for a given \( k \). According to formulas (2) (3), the above estimating equations (5) are uniquely determined by \( m_f(\theta), \sigma_f^2(\theta), s_f(\theta), k_f(\theta), \frac{\partial m_f(\theta)}{\partial \theta_k} \), and \( \frac{\partial^2 \sigma_f^2(\theta)}{\partial \theta_k \partial \theta_k} \) that are respectively the mean, the variance, the skewness and the kurtosis of \( f(X) \), and the partial derivatives of \( m_f(\theta) \), and \( \sigma_f^2(\theta) \). Thus, if \( X \) admits finite the first four moments, we can use the function \( f(X) = X \) to get some estimators of the distributional parameters. Otherwise, we can use a bounded function \( f \).
Moreover, in several cases it could be useful to adopt a parametric differentiable function \( f : A \rightarrow \mathbb{R} \), where \( A = \mathbb{R} \times [a, b] \subseteq \mathbb{R}^2 \). As a matter of fact, in this case the previous estimators and functionals (i.e., \( m_f(\theta, q) \), \( \sigma_f^2(\theta, q) \), \( s_f(\theta, q) \) and \( k_f(\theta, q) \)) also depend on a parameter \( q \in [a, b] \) of the function \( f(X, q) \). This aspect can be used to get an optimal estimator with the minimum asymptotic variance as suggested in the following remark.

**Remark** Suppose we have a function \( f : A \rightarrow \mathbb{R} \) (with \( A \subseteq \mathbb{R}^2 \)) such that for any real random variable \( X \in \mathcal{F}(\theta) \) (with \( \theta \in B \subseteq \mathbb{R}^p \)) \( f(X,q) \) admits finite the first four moments for any admissible \( q \). Then we get an optimal GT estimator with minimum asymptotic variance solving the following optimization problem:

\[
\max_{(\theta, q) \in A} |V_{EF}(\theta, q)| \quad \text{subject to} \quad \begin{align*}
(\theta, q) &
\end{align*}
\]

where \( |V_{EF}(\theta, q)| \) is the determinant of the inverse asymptotic variance (1)

\[
V_{EF}(\theta, q) = [v_{i,j}(\theta, q)]_{i,j=1,...,p} \quad \text{and} \quad v_{i,j}(\theta, q) = E \left( \frac{\partial f}{\partial \theta_i} \frac{\partial f}{\partial \theta_j} \right). 
\]

In order to reduce the computational complexity of (6) we can use \( \sum_{i=1}^p v_{i,j}(\theta, q) \) as objective function instead of \( |V_{EF}(\theta, q)| \).

Clearly the idea of minimizing the asymptotic variance subject to the constraints of some equations can be applied to many other estimators (for example the maximum likelihood estimator). Using a GT estimator we do not necessarily need to know a closed form of the density (or the cumulative) distribution of the underlying random variables. As a matter of fact, GT estimators can be used even for those distributions uniquely defined by their characteristic function \( \phi_X(t) = E(\exp(itX)) \). Let us consider two possible examples.

**Moment estimator:** Suppose all parametric random variables belong to the space \( L' = \{ X/E(|X'|) < \infty \} \). By using the derivatives of the characteristic function \( \phi_X^{(k)}(0) \) we can determine all the existing integer moments of \( X \), since \( \phi_X^{(k)}(0) = i^kE(X^k) \). Then, as parametric functions we can use:

- \( f(X, q) = |X|^q \) for any \( q \in [0, r/4] \) (in this case \( m_f(\theta, q) = E(|X|^q) \) represents the moments curve that characterizes the distribution of \( |X| \));
- \( f(X, q) = X^q \) for any integer \( q \in [0, r/4] \).

**Trigonometric estimator:** Suppose \( f(X,t) = \sin(tX) \) for some given \( t \neq 0 \). Then if we know the characteristic function \( \phi_X(t) = E(\exp(itX)) \) we can easily determine the first four moments of \( f(X,t) \), since \( E(\cos(tX)) = \text{Re}(\phi_X(t)) \); \( E(\sin(tX)) = \text{Im}(\phi_X(t)) \). So, if \( f(X,t) = \sin(tX) \) for a \( t \neq 0 \) the first four moments of \( \sin(tX) \) are given by:

\[
E(\sin(tX)) = \text{Im}(\phi_X(t)); \quad E(\sin^2(tX)) = 0.5(1 - \text{Re}(\phi_X(2t)))
\]

\[
E(\sin^3(tX)) = 1 - (3 \text{Im}(\phi_X(t)) - \text{Im}(\phi_X(3t)));
\]

\[
E(\sin^4(tX)) = 1 - (3 - 4 \text{Re}(\phi_X(2t)) + \text{Re}(\phi_X(4t)));
\]

5
and together with these we have to know \( \frac{\partial m_X(\theta,t)}{\partial \theta_k} \) and \( \frac{\partial \sigma^2_X(\theta)}{\partial \theta_k} \). This method can be easily applied to estimate the parameters of all distributions defined by their characteristic function, when from the characteristic function we can easily distinguish the imaginary part from the real one. For example, all the infinitely divisible random variables \( X \) have characteristic function \( \phi_X(u) \) uniquely determined by the triplet \( [\gamma, \sigma^2, \nu] \) that identifies the so called Lévy-Khintchine characteristic exponent \( \psi_X(u) = \log \phi_X(u) \) given by:

\[
\psi_X(u) = i\gamma u - \frac{1}{2}\sigma^2 u^2 + \int_{-\infty}^{+\infty} (\exp(iux) - 1 - iux1_{\{|x|<1\}})\nu(dx) = \\
i \left( \gamma u + \int_{-\infty}^{+\infty} (\sin(ux) - ux1_{\{|x|<1\}})\nu(dx) \right) - \\
\frac{1}{2}\sigma^2 u^2 + \int_{-\infty}^{+\infty} (\cos(ux) - 1)\nu(dx)
\]

where \( \gamma \in \mathbb{R}, \sigma^2 > 0 \) and \( \nu \) is a measure on \( \mathbb{R}\backslash\{0\} \) with \( \int_{-\infty}^{+\infty} (1 \land x^2)\nu(dx) < \infty \). Therefore for any infinitely divisible random variables \( X \) we can easily identify the real and the imaginary one of the characteristic function given by:

\[
\text{Re} (\phi_X(u)) = \exp \left( -\frac{1}{2}\sigma^2 u^2 + \int_{-\infty}^{+\infty} (\cos(ux) - 1)\nu(dx) \right) \times \\
\times \cos \left( \gamma u + \int_{-\infty}^{+\infty} (\sin(ux) - ux1_{\{|x|<1\}})\nu(dx) \right),
\]

\[
\text{Im} (\phi_X(u)) = \exp \left( -\frac{1}{2}\sigma^2 u^2 + \int_{-\infty}^{+\infty} (\cos(ux) - 1)\nu(dx) \right) \times \\
\times \sin \left( \gamma u + \int_{-\infty}^{+\infty} (\sin(ux) - ux1_{\{|x|<1\}})\nu(dx) \right).
\]

In particular, the Lévy triplet \( [\gamma, \sigma^2, \nu] \) identifies the three main components of any Lévy process: the deterministic component \( (\gamma) \), the Brownian component \( (\sigma^2) \) and the pure jump component \( (\nu) \). For further details on theoretical aspects we refer to Sato (1999).

### 2.1 Examples of infinitely divisible distributions

In this subsection we consider some particular infinitely divisible distributions: stable Pareto distributions, tempered stable (TS) distributions, normal inverse Gaussian (NIG) distributions, and Carr, Geman, Madan, Yor (CGMY) distributions. For each of these distributions we describe:

- the characteristic function distinguishing the real and imaginary parts;
- the moments curve in the stable Pareto case and the mean, the variance, the skewness and the kurtosis for the other infinitely divisible distributions;
• In Table I we give the derivatives $\frac{\partial m_f(\theta)}{\partial \theta_k}$ and $\frac{\partial \sigma^2_f(\theta)}{\partial \theta_k}$ for the GT (EF) moment estimator;

• In Tables II and III we give the derivatives $\frac{\partial \text{Im}(\phi_X(t))}{\partial \theta_k}$ and $\frac{\partial \text{Re}(\phi_X(t))}{\partial \theta_k}$ for the GT (EF) trigonometric estimator.

Doing so, we are able to derive:

a) a GT trigonometric estimator when $f(X,t) = \sin(tX)$ for all these distributions;

b) a GT moment estimator when $f(X) = X$ for TS, NIG and CGMY distributions;

c) a GT parametric moment estimator when $f(X,p) = |X - \mu|^p$ for stable Paretian distributions.

Stable distributions (see, among others, Rachev and Mittnik (2000) and the references therein) A univariate stable distribution $X \overset{d}{=} S_\alpha(\sigma, \beta, \mu)$ is characterized by four parameters. These are: the index of stability $\alpha \in (0, 2]$, the scale parameter $\sigma \in \mathbb{R}^+$, the skewness parameter $\beta \in [-1, 1]$ and the shift parameter $\mu \in \mathbb{R}$. The stable distribution is Gaussian when $\alpha = 2$, and in this case, $\sigma$ is proportional to the standard deviation, $\beta$ can be taken to be zero and $\mu$ is the mean. An $\alpha$ stable non Gaussian distribution admits finite $p$-th fractional moment $E(|X|^p) < \infty$ for any $p \in (-1, \alpha)$. A stable distribution can be defined in different equivalent ways. We use the characteristic function extensively because few stable density functions are known in closed form. The characteristic function of $X \overset{d}{=} S_\alpha(\sigma, \beta, \mu)$ is given by:

$$
\phi_X(t) = \begin{cases} 
\exp \left(-|t\sigma|^\alpha \left(1 - i \text{sgn}(t) \beta \tan \frac{\pi \alpha}{2}\right) + it\mu\right) & \text{if } \alpha \in (0, 2]; \alpha \neq 1 \\
\exp \left(-|t\sigma| \left(1 + i \beta \frac{2}{\pi} \text{sgn}(t) \log |t|\right) + it\mu\right) & \text{if } \alpha = 1.
\end{cases}
$$

Thus, $\text{Re}(\phi_X(t)) = \exp(-|t\sigma|^\alpha) \cos(b)$; $\text{Im}(\phi_X(t)) = \exp(-|t\sigma|^\alpha) \sin(b)$, where

$$
b = \begin{cases} 
|t\sigma|^\alpha \text{sgn}(t) \beta \tan \frac{\pi \alpha}{2} + t\mu & \text{if } \alpha \in (0, 2]; \alpha \neq 1 \\
t\mu - |t\sigma| \beta \frac{2}{\pi} \text{sgn}(t) \log |t| & \text{if } \alpha = 1,
\end{cases}
$$

and we can apply formula (5) to approximate $\alpha, \sigma, \beta, \mu$ with the GT trigonometric estimator when we assume $f(X,t) = \sin(tX)$. Thus using formulas (2) (3) (7) and (4) we can solve the problem (6):

$$
\max_{\alpha, \sigma, \beta, \mu, t} |V_{EF}(\alpha, \sigma, \beta, \mu, t)|
$$

subject to (5); $t \neq 0$, 

7
Table I

This table summarizes the derivatives used in the moment estimators for TS, NIG, CGMY and Stable distributions.

### $X = TS(a, b)$

<table>
<thead>
<tr>
<th>$\frac{\partial E[X]}{\partial a}$</th>
<th>$\frac{\partial Var[X]}{\partial a}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2ab^{a-1} \left( 1 + \frac{1}{k} \ln b \right)$</td>
<td>$4a b^{a-1} \left( \frac{k}{k - 1} b - \ln b \right)$</td>
</tr>
</tbody>
</table>

### $X = NIG(\alpha, \beta, \delta)$

<table>
<thead>
<tr>
<th>$\frac{\partial E[X]}{\partial \alpha}$</th>
<th>$\frac{\partial E[X]}{\partial \beta}$</th>
<th>$\frac{\partial E[X]}{\partial \delta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>$\beta$</td>
<td>$\delta$</td>
</tr>
</tbody>
</table>

### $X = CGMY(\alpha, \beta, \gamma, \mu)$

<table>
<thead>
<tr>
<th>$\frac{\partial E[X]}{\partial \alpha}$</th>
<th>$\frac{\partial E[X]}{\partial \beta}$</th>
<th>$\frac{\partial E[X]}{\partial \gamma}$</th>
<th>$\frac{\partial E[X]}{\partial \mu}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>$\beta$</td>
<td>$\gamma$</td>
<td>$\mu$</td>
</tr>
</tbody>
</table>

### $X = S_a(\beta, \sigma, \mu)$

<table>
<thead>
<tr>
<th>$\frac{\partial [v]}{\partial \beta}$</th>
<th>$\frac{\partial [v]}{\partial \sigma}$</th>
<th>$\frac{\partial [v]}{\partial \mu}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{1}{a}$</td>
<td>$\frac{1}{\sqrt{\pi}}$</td>
<td>$\frac{1}{\sqrt{\pi}}$</td>
</tr>
</tbody>
</table>

where: $v = \cos \nu - \frac{\nu}{2}$, $\sigma = \tan \frac{\nu}{2}$, $\mu = \frac{\nu}{2} \arctan(\beta \nu)$, $\gamma = \frac{\nu}{2} \cos \left( \frac{\nu}{2} \right)$, $\delta = \frac{\nu}{2} \sin \left( \frac{\nu}{2} \right)$.
Table II
This table summarizes the derivatives used in the trigonometric estimators for TS, and CGMY distributions.

<table>
<thead>
<tr>
<th>CGMY ({C, G, M, Y})</th>
<th>where</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{Im} (\frac{\partial}{\partial y} (\cos(y) - \sin(y))))</td>
<td>(z = \text{atan}(x))</td>
</tr>
<tr>
<td>(\text{Re} (\frac{\partial}{\partial y} (\cos(y) - \sin(y))))</td>
<td>(\text{arc}(\cos(y) - \sin(y)))</td>
</tr>
<tr>
<td>(\text{Im} (\frac{\partial}{\partial y} (\cos(y) + \sin(y))))</td>
<td>(z = \exp(\text{arc}(\cos(y) - \sin(y))))</td>
</tr>
<tr>
<td>(\text{Re} (\frac{\partial}{\partial y} (\cos(y) + \sin(y))))</td>
<td>(\text{arc}(\cos(y) - \sin(y)))</td>
</tr>
</tbody>
</table>
where $\frac{\partial \text{Im}(\phi_X(t))}{\partial \sigma}$ and $\frac{\partial \text{Re}(\phi_X(t))}{\partial \sigma}$ are given in Table III. Moreover, we can propose a GT moment estimator based on the moments curve. As a matter of fact the absolute central moments for any $p \in (-1, \alpha)$, $\alpha \neq 1$ are given by:

$$E(\lvert X - \mu \rvert^p) = \sigma^p \left(1 + \beta^2 \tan^2 \frac{\alpha \pi}{2}\right)^{0.5p/\alpha} \cos \left(\frac{\beta}{\alpha} \arctan \left(\frac{\alpha \pi}{2}\right)\right) A(p),$$

where $A(p) = \frac{2^p \Gamma \left(\frac{\alpha - p}{\alpha}\right) \Gamma \left(\frac{p + 1}{2}\right)}{\Gamma \left(\frac{2 - p}{2}\right) \sqrt{\pi}}$ and $\Gamma(c) = \int_0^{+\infty} z^{c - 1} e^{-z} dz$ for $c > 0$.

while we refer to Abramowitz and Stegun’s definition (see Abramowitz and Stegun 1970) of the Gamma function $\Gamma(.)$ of a negative real (non integer) number. Thus, if we use $f(X) = \lvert X - \mu \rvert^p$ such that $p$ is small enough (i.e., $p \in (-0.25, \alpha/4)$) we can easily get the first four moments $E(f(X)^j) = E(\lvert X - \mu \rvert^{pj})$ $j = 1, \ldots, 4$, in order to obtain a GT moment estimator for the stable parameters. Thus we can apply formulas (2) (3) (using $\frac{\partial E(\lvert X - \mu \rvert^p)}{\partial \sigma}$, $\frac{\partial E(\lvert X - \mu \rvert^p)}{\partial \pi}$ and $\frac{\partial E(\lvert X - \mu \rvert^p)}{\partial \alpha}$ given in Table I) to determine estimating functions (5) for $\alpha, \sigma, \beta$.

Since we need four equations for four parameters and we do not know a closed form of $E(\lvert X - \mu \rvert^p \text{sgn}(X - \mu))$, we suggest using, as fourth estimating equation, the consistent estimating function of $m_f(\theta)$ (5) assuming that $m_f(\theta)$ is itself a possible parameter. Since $\sigma_f^2(\alpha, \sigma, \beta) = E(\lvert X - \mu \rvert^2) - m_f(\alpha, \sigma, \beta)^2$,

then $\frac{\partial \sigma_f^2(\alpha, \sigma, \beta)}{\partial m_f} = \frac{\partial \sigma_f^2(\alpha, \sigma, \beta)}{\partial m_f}$ and $\frac{\partial m_f(\alpha, \sigma, \beta)}{\partial m_f} = 1$, and the fourth estimating equation is

$$l_m = \sum_{k=1}^{\infty} (a_1 h_1(Xs, \theta) + a_2 h_2(Xs, \theta)) = 0 \quad \text{where} \quad a_1 = \frac{-1}{\sigma_f^2(\alpha, \sigma, \beta)} \left(\frac{\partial \sigma_f^2(\theta)}{\partial m_f} + s_f(\theta) \sigma_f(\theta)\right) \left(k_f(\theta) - 1 - s_f(\theta)\right).$$

Analogously we get the elements of the asymptotic variance (4). Doing so, we can maximize the inverse of the asymptotic variance determinant $\max_{\alpha,\sigma,\beta,\mu,\mu} |V_{EF}(\alpha, \sigma, \beta, \mu, p)|$ associated with the GT moment estimator of stable paretian distributions.

**Tempered Stable (TS) distribution** (see Tweedie (1984) and, for more general definitions, see Kim et al. (2008)) Tempered stable distributions depend on three parameters $\alpha > 0$; $b \geq 0$; $0 < k < 1$ and their characteristic function is given by:

$$\phi_{TS}(u; k, a, b) = \exp \left(ab - a \left(b^{1/k} - 2iu\right)^k \right).$$

Thus, $\text{Re}(\phi_X(t)) = c \cos(d)$; $\text{Im}(\phi_X(t)) = c \sin(d)$, where

$$c = \exp \left(ab - a \left(b^{2/k} + 4t^2\right)^{k/2} \cos \left(k \arctan \frac{-2t}{b^{1/k}}\right)\right);$$

$$d = -a \left(b^{2/k} + 4t^2\right)^{k/2} \sin \left(k \arctan \frac{-2t}{b^{1/k}}\right).$$

10
Thus, the mean, the variance, the skewness and the kurtosis of tempered stable distributions are given by:

\[ m_f(a, b, k) = 2abk^{(k-1)/k}; \quad \sigma_f^2(a, b, k) = 4ak(1-k)b^{(k-2)/k}; \]
\[ s_f(a, b, k) = \frac{(k-2)}{[abk(1-k)]^{1/2}}; \quad k_f(a, b, k) = 3 + \frac{4k - 6 - k(1-k)}{abk(1-k)}. \]

Therefore, considering the derivatives \( \frac{\partial m_f(t)}{\partial \theta}, \frac{\partial \sigma_f^2(t)}{\partial \theta}, \frac{\partial \text{Im}(\phi_X(t))}{\partial \theta}, \) and \( \frac{\partial \text{Re}(\phi_X(t))}{\partial \theta} \) given in Tables I and III, we can obtain GT moment and trigonometric estimators.

**Normal Inverse Gaussian (NIG) distribution** (see Rachev and Mittnik (2000) and the references therein) A NIG distribution depends on three parameters \( \alpha > 0; \delta > 0; \beta \in (-\alpha, \alpha) \) and its characteristic function is given by:

\[ \phi_{NIG}(u; \alpha, \beta, \delta) = \exp \left\{ -\delta \left( \sqrt{\alpha^2 - (\beta + iu)^2} - \sqrt{\alpha^2 - \beta^2} \right) \right\} \]

Thus, \( \text{Re}(\phi_X(u)) = c \cos(d); \text{Im}(\phi_X(u)) = c \sin(d), \) where

\[ c = \exp \left( \delta \sqrt{\alpha^2 - \beta^2} - \delta \left( (\alpha^2 - \beta^2 + u^2)^2 + 4\alpha^2\beta^2 \right)^{0.25} \cos(e) \right), \]
\[ d = -\delta \left( (\alpha^2 - \beta^2 + u^2)^2 + 4\alpha^2\beta^2 \right)^{0.25} \sin(e), \]
\[ e = 0.5 \arctan \frac{-2u\beta}{\alpha^2 - \beta^2 + u^2}. \]

Moreover, if we assume \( f(X) = X, \) the mean, the variance, the skewness and the kurtosis of Normal Inverse Gaussian distributions are given by:

\[ m_f(\alpha, \beta, \delta) = \delta \beta (\alpha^2 - \beta^2)^{-1/2}; \quad \sigma_f^2(\alpha, \beta, \delta) = \alpha^2 \delta (\alpha^2 - \beta^2)^{-3/2}; \]
\[ s_f(\alpha, \beta, \delta) = \frac{3\beta}{\alpha \sqrt{\delta \cdot \sqrt{\alpha^2 - \beta^2}}}; \quad k_f(\alpha, \beta, \delta) = 3 \left( 1 + \frac{\alpha^2 + 4\beta^2}{\delta \alpha^2 \sqrt{\alpha^2 - \beta^2}} \right). \]

The derivatives \( \frac{\partial m_f(t)}{\partial \theta}, \frac{\partial \sigma_f^2(t)}{\partial \theta}, \frac{\partial \text{Im}(\phi_X(t))}{\partial \theta}, \) and \( \frac{\partial \text{Re}(\phi_X(t))}{\partial \theta} \) are given in Tables I and III.

**The Carr, Geman, Madan, Yor (CGMY) distribution** (see Carr et al. (2002)) A CGMY distribution depends on four parameters \( C, G, M > 0; Y < 2 \) and its characteristic function is given by:

\[ \phi_{CGMY}(u; C, G, M, Y) = \exp \left\{ CT(-Y) \left[ (M - iu)^Y - M^Y + (G + iu)^Y - G^Y \right] \right\}. \]

Thus, \( \text{Re}(\phi_X(t)) = c \cos(d); \text{Im}(\phi_X(t)) = c \sin(d), \) where

\[ c = \exp \left\{ CT(-Y) \left[ -M^Y - G^Y + (M^2 + t^2)^Y \cos \left( \frac{Y \arctan \frac{-t}{M}}{M} \right) \right. \right. \]
\[ + \left. \left. (G^2 + t^2)^Y \cos \left( Y \arctan \frac{t}{G} \right) \right] \right\} \]
Table III
This table summarizes the derivatives used in the trigonometric estimators for NIG, and Stable distributions.

<table>
<thead>
<tr>
<th></th>
<th>( S_0(\beta, \sigma, \mu) )</th>
<th>( \text{where} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{\partial \Re}{\partial \alpha} )</td>
<td>( \frac{1}{2} \beta (\pi + 2q \ln</td>
<td>\sigma</td>
</tr>
<tr>
<td>( \frac{\partial \Re}{\partial \beta} )</td>
<td>(-pq \text{sgn} \sin m )</td>
<td>( q = \sin \frac{\mu}{2} )</td>
</tr>
<tr>
<td>( \frac{\partial \Re}{\partial \sigma} )</td>
<td>(- \pi t \sin m )</td>
<td>( m = \mu + \beta \sigma \epsilon^t \text{sgn}(t) )</td>
</tr>
<tr>
<td>( \frac{\partial \Re}{\partial \mu} )</td>
<td>(-\pi \mu \text{sgn}(\sigma) \cos m \sin m )</td>
<td></td>
</tr>
<tr>
<td>( \frac{\partial \Im}{\partial \alpha} )</td>
<td>( \frac{1}{2} \beta (\pi + 2q \ln</td>
<td>\sigma</td>
</tr>
<tr>
<td>( \frac{\partial \Im}{\partial \beta} )</td>
<td>( p \epsilon^t \text{sgn} \cos m )</td>
<td></td>
</tr>
<tr>
<td>( \frac{\partial \Im}{\partial \sigma} )</td>
<td>(-\pi \mu \text{sgn}(\sigma) \sin m - \beta q \epsilon^t \text{sgn}(m) )</td>
<td></td>
</tr>
<tr>
<td>( \frac{\partial \Im}{\partial \mu} )</td>
<td>(-\epsilon^t \text{sgn}(\sigma) \sin m )</td>
<td></td>
</tr>
</tbody>
</table>

\[ d = C \Gamma(-Y) \left( M^2 + t^2 \right)^{Y/2} \sin \left( Y \arctan \frac{-t}{M} \right) + C \Gamma(-Y) \left( G^2 + t^2 \right)^{Y/2} \sin \left( Y \arctan \frac{t}{G} \right). \]

Moreover, if we assume \( f(X) = X \), the mean, the variance, the skewness and the kurtosis of CGMY distributions are given by:

\[ m_f(C, G, M, Y) = C \left( M^{Y-1} - G^{Y-1} \right) \Gamma(1 - Y); \]
\[ \sigma_f^2(C, G, M, Y) = C \left( M^{Y-2} + G^{Y-2} \right) \Gamma(2 - Y); \]
The derivatives \( \frac{\partial m_f(\theta)}{\partial \theta_k}, \frac{\partial^2 m_f(\theta)}{\partial \theta_k^2}, \frac{\partial \text{Im}(\phi_X(t))}{\partial \theta_k}, \text{and} \frac{\partial \text{Re}(\phi_X(t))}{\partial \theta_k} \) are given in Tables I and II.

3 An empirical comparison between the GT moment estimator and the maximum likelihood estimator of stable Paretian distributions

In this section we compare the GT moment estimator and the MLE obtained by inverting the characteristic function of stable Paretian distributions (see Rachev and Mittnik (2000)). In particular, we test the above semi-parametric estimator for stable distributions using simulated data. Therefore, using the algorithm proposed by Chambers et al. (see Chambers et al. (1976)), we generate \( N \) \( (N=200, \ldots, 5000 \text{ with step 100}) \) stable distributions \( S_{\alpha}(\sigma, \beta, \mu) \) with parameters \( \alpha = 0.51, 0.76, 1.26, 1.51, 1.76; \beta = -1, -0.5, 0.5, 1; \sigma = 1; \mu = 1 \). Then we estimate the parameters on the simulated data (for each \( N \)) considering both estimating methods: the GT moment estimator, and MLE valued inverting the characteristic function with the FFT. As starting point for the GT moment and MLE estimators we use the parameters obtained estimating the series of \( N \) elements with the McCulloch quantile method (see, among others, Rachev and Mittnik (2000)). We obtain the results of GT moment minimizing the sum of the asymptotic variances subject to the usual constraints. This computation requires less time than the MLE approximation. Minimizing the determinant of the asymptotic variance matrix we get more robust results, but we need much more computational time to approximate the parameters.

We measure (on average) the absolute value of the percentage of the distance between the parameters of simulated series and the estimated ones for the different \( \alpha, \sigma, \beta, \) and \( \mu \) i.e., we compute the average (varying \( N \)) of \( |\Delta \theta_{GT}| = \frac{\theta_{GT} - \theta_{simulation}}{\theta_{simulation}} \), and similarly of \( |\Delta \theta_{MLE}| = \frac{\theta_{MLE} - \theta_{simulation}}{\theta_{simulation}} \), where \( \theta = \alpha, \sigma, \beta, \mu \). These results are given in Table IV where we remark in bold the best approximations. We observe that the sum of the all absolute errors is higher for the MLE. In addition, the empirical analysis shows that the GT moment estimators present a very good performance even in comparison to those obtained with the MLE method.

4 Concluding remarks

In the paper we discussed the application of the estimating function method to value the parameters of distributions defined only by their characteristic func-
This table summarizes the average of the absolute errors we have using either a GT moment estimator or a MLE estimator for different values of $\alpha$ ($\alpha = 0.51; 0.76; 1.26; 1.51; 1.76$) and $\beta$ ($\beta = -1; -0.5; 0.5; 1$) and $\sigma = 1$, $\mu = 1$.

<table>
<thead>
<tr>
<th>Moment</th>
<th>Beta</th>
<th>0.51</th>
<th>0.76</th>
<th>1.26</th>
<th>1.51</th>
<th>1.76</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha</td>
<td>$</td>
<td>\alpha\Delta\sigma</td>
<td>$, $</td>
<td>\alpha\Delta\mu</td>
<td>$, $</td>
<td>\beta\Delta\sigma</td>
</tr>
<tr>
<td>0.51</td>
<td>0.07</td>
<td>0.04</td>
<td>0.04</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>0.76</td>
<td>0.02</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>1.26</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>1.51</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>1.76</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MLE</th>
<th>Beta</th>
<th>0.51</th>
<th>0.76</th>
<th>1.26</th>
<th>1.51</th>
<th>1.76</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha</td>
<td>$</td>
<td>\alpha\Delta\sigma</td>
<td>$, $</td>
<td>\alpha\Delta\mu</td>
<td>$, $</td>
<td>\beta\Delta\sigma</td>
</tr>
<tr>
<td>0.51</td>
<td>0.07</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>0.76</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>1.26</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>1.51</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>1.76</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>

The proposed methodology showed good versatility, since it could be applied to any bounded function of the underlying random variable. In particular, we propose two EF estimators for the parameters of stable Pareto distributions and other infinitely divisible distributions. Finally, we have proposed an empirical comparison based on simulated data of stable Pareto distributions. The good results obtained with the EF moment estimator even with respect to the MLE method, suggest that probably we could make further improvements in parameter estimation using other bounded functions.

Acknowledgement: The authors thank for helpful comments seminar audiences at AMAT 2008 (Memphis, USA). For their helpful support in writing the tables, we thank Valentina Acerbis, Marco Cassader and Sebastiano Vitali. Rachev’s research was supported by grants from the Division of Mathematical, Life and Physical Science, College of Letters and Science, University of California, Santa Barbara, and the Deutschen Forschungsgemeinschaft. Sergio Ortobelli and Valeria Caviezel were partially supported by grants from ex-murst 60%, 2007 and 2008.

References


