Parameter estimation for exponentially tempered power law distributions*

Mark M. Meerschaert†, Parthanil Roy, and Qin Shao‡
Department of Statistics and Probability
Michigan State University
East Lansing MI 48824

February 19, 2010

Abstract

Tail estimates are developed for power law probability distributions with exponential tempering, using a conditional maximum likelihood approach based on the upper order statistics. Tempered power law distributions are intermediate between heavy power-law tails and Laplace or exponential tails, and are sometimes called "semi-heavy" tailed distributions. The estimation method is demonstrated on simulated data from a tempered stable distribution, and for several data sets from geophysics and finance that show a power law probability tail with some tempering.

1 Introduction

Probability distributions with heavy, power law tails are important in many areas of application, including physics [20, 21, 31], finance [9, 13, 23, 26, 25], and hydrology [6, 7, 27, 28]. Stable Lévy motion with power law tails is useful to model anomalous diffusion, where long particle jumps lead to anomalous superdiffusion [16, 17].

* *Key words and phrases:* Tail estimation, heavy tails, exponential tempering, tempered stable.
† Partially supported by NSF grants EAR-0823965 and DMS-0803360
‡ On leave from the University of Toledo.

Often the power law behavior does not extend indefinitely, due to some truncation or tapering effects. Truncated Lévy flights were proposed by Mantegna and Stanley [14, 15] as a modification of the α-stable Lévy motion, to avoid infinite moments. In that model, the largest jumps are simply discarded. Tempered stable Lévy motion takes a different approach, exponentially tapering the probability of large jumps, so that all moments exist [24]. Tempered stable laws were recently applied in geophysics [19]. The problem of parameter estimation for tempered stable laws remains open. Normal inverse Gaussian distributions have the same asymptotic tail behavior, which Barndorff-Neilson calls "semi-heavy tails" [4]. These distributions are important in finance [3] and turbulence [5]. Parameter estimation for the normal inverse Gaussian distribution is a difficult problem [22]. Laplace distributions have also found many applications in engineering, finance, biology, and environmental science. The book [12] contains a comprehensive introduction to the theory and application of Laplace distributions and processes, as well as a number of applications. See [18] for some additional applications to geophysics.

In practical applications, it is often apparent that data tails are too heavy to admit a Gaussian model. Fitting alternative models with a heavier tail requires a judgment about whether the tails are exponential (Laplace, gamma, Weibull, etc.), power-law (Pareto, stable, geometric stable, etc.), or something in between. This judgment starts with an examination of the empirical tail distribution. If the tail appears to follow a pure exponential, then a Laplace or related model may be appropriate. If it follows a pure power-law, then a stable or related model may suffice. For cases in between, where the tail gradually transitions from power-law to exponential, the methods of this paper can be useful. Some further discussion, along with a test for determining the extent of truncation/tempering, appears in [8].

This paper treats exponentially tempered Pareto distributions $P(X > x) = \gamma x^{-\alpha} e^{-\beta x}$ where γ is a scale parameter, α controls the power law tail, and β governs the exponential truncation. In practical applications, the truncation parameter is relatively small, so that the data seems to follow the power law distribution until the largest values are exponentially cooled. A log-log plot of the data versus rank is linear until the tempering causes the plot to curve downward. Such plots are often observed in real data applications. It is also common that the power law behavior emerges only for large data values, so that the tail of the data is fit to this model. Hence we will consider parameter estimates based on the largest order statistics. The main technical tool is the Rényi representation for the order statistics, and the mathematical details are similar to Hill's estimator [10, 11] for the traditional Pareto distribution. A related paper [1] considered parameter estimation for the truncated Pareto distribution, relevant to the original model of Mantegna and Stanley.

2 Estimation

Suppose $X_1, X_2, \ldots, X_n$ is a random sample from the tempered Pareto distribution with the survival function

$$\bar{F}_X(x;\boldsymbol{\theta}) = P\{X_1 > x\} = \gamma x^{-\alpha} e^{-\beta x}, \quad x \geq x_0, \tag{2.1}$$

where $\boldsymbol{\theta} := \{\alpha, \beta, \gamma\}$ are unknown parameters and $x_0 > 0$ satisfies $\gamma = x_0^{\alpha} e^{\beta x_0}$. Clearly the corresponding density function is given by

$$f_X(x;\boldsymbol{\theta}) = \gamma x^{-\alpha-1} e^{-\beta x}(\alpha + x\beta), \quad x \geq x_0. \tag{2.2}$$

Let $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ be the order statistics of the sample, $z := 1/x$, $z_0 := 1/x_0$ and $Z_i := 1/X_i$ for $i = 1, 2, \ldots, n$. Then $Z_{(k)} := 1/X_{(k)} > Z_{(k+1)}$ for all $k = 1, \ldots, n-1$ and we have

$$P\{Z_1 \leq z\} = F_Z(z;\boldsymbol{\theta}) = \gamma z^{\alpha} e^{-\beta/z}, \quad z \leq z_0,$$

which implies that for $z \leq z_0$,

$$\frac{dF_Z(z;\boldsymbol{\theta})}{dz} = \gamma z^{\alpha} e^{-\beta/z}\left(\frac{\alpha}{z} + \frac{\beta}{z^2}\right).$$

By the formula (2.5) in [11], it follows that that the conditional log-likelihood of $\{Z_{(n-k+1)}, \ldots, Z_{(n)}\}$ given $Z_{(n-k+1)} < d_z \leq Z_{(n-k)}$ is proportional to the following

$$\begin{aligned}
&\log[1 - F_Z(d_z,\boldsymbol{\theta})]^{n-k} + \sum_{i=1}^{k} \log \frac{dF_Z(z_{(n-i+1)})}{dz_{(n-i+1)}} \\
\propto\ & (n-k)\log\left[1 - \gamma d_z^{\alpha} e^{-\beta/d_z}\right] + k\log\gamma + \alpha\sum_{i=1}^{k}\log z_{(n-i+1)} - \beta\sum_{i=1}^{k} z_{(n-i+1)}^{-1} \\
&+ \sum_{i=1}^{k}\log\left(\frac{\alpha}{z_{(n-i+1)}} + \frac{\beta}{z_{(n-i+1)}^2}\right).
\end{aligned}$$

Let $\boldsymbol{x_k} = \{x_{(n-k+1)} \ldots, x_{(n)}\} := \left\{z_{(n-k+1)}^{-1}, \ldots, z_{(n)}^{-1}\right\}$ and $d_x = 1/d_z$. Using the change of variable formula, the conditional log-likelihood of $\boldsymbol{X_k} = \{X_{(n-k+1)}, \ldots, X_{(n)}\}$ given

$X_{(n-k+1)} > d_x \geq X_{(n-k)}$ is of the form

$$\begin{aligned}\log L_c(\boldsymbol{\theta};\boldsymbol{x_k}) \quad \propto \quad & (n-k)\log\left[1-\gamma d_x^{-\alpha}e^{-\beta d_x}\right] + k\log\gamma - (\alpha+2)\sum_{i=1}^{k}\log x_{(n-i+1)} \\ & -\beta\sum_{i=1}^{k}x_{(n-i+1)} + \sum_{i=1}^{k}\log\left(\alpha x_{(n-i+1)} + \beta x_{(n-i+1)}^2\right). \end{aligned} \tag{2.3}$$

The following result gives the normal equations of the conditional likelihood problem with the notation introduced above.

Proposition 2.1. *(a) The conditional MLE $\hat{\boldsymbol{\theta}} = \{\hat{\alpha},\hat{\beta},\hat{\gamma}\}$ of $\boldsymbol{\theta} = \{\alpha,\beta,\gamma\}$ given $X_{(n-k+1)} > d_x \geq X_{(n-k)}$ satisfies the normal equations*

$$\sum_{i=1}^{k}(\log d_x - \log X_{(n-i+1)}) + \sum_{i=1}^{k}\frac{1}{\hat{\alpha}+\hat{\beta}X_{(n-i+1)}} = 0, \tag{2.4}$$

$$\sum_{i=1}^{k}(d_x - X_{(n-i+1)}) + \sum_{i=1}^{k}\frac{X_{(n-i+1)}}{\hat{\alpha}+\hat{\beta}X_{(n-i+1)}} = 0, \tag{2.5}$$

$$\hat{\gamma} = \frac{k}{n}d_x^{\hat{\alpha}}e^{\hat{\beta}d_x}. \tag{2.6}$$

(b) If the above system of normal equations has a solution $\hat{\boldsymbol{\theta}}$ with $\hat{\alpha} > 0$ and $\hat{\beta} > 0$, then it is the unique conditional MLE.

Proof. (a) Defining $\lambda = \gamma d_x^{-\alpha}e^{-\beta d_x}$, the conditional log-likelihood in (2.3) is simplified to

$$\begin{aligned}\log L_c(\boldsymbol{\theta};\boldsymbol{x_k}) \quad \propto \quad & (n-k)\log(1-\lambda) + k\log\gamma - (\alpha+2)\sum_{i=1}^{k}\log x_{(n-i+1)} \\ & -\beta\sum_{i=1}^{k}x_{(n-i+1)} + \sum_{i=1}^{k}\log\left(\alpha x_{(n-i+1)} + \beta x_{(n-i+1)}^2\right). \end{aligned} \tag{2.7}$$

The estimates $\hat{\boldsymbol{\theta}}$ satisfies the following normal equations obtained by $\dfrac{\partial \log L_c(\boldsymbol{\theta};\boldsymbol{x_k})}{\partial\boldsymbol{\theta}}$

$$\frac{(n-k)\lambda\log d_x}{1-\lambda} - \sum_{i=1}^{k}\log x_{(n-i+1)} + \sum_{i=1}^{k}\frac{1}{\alpha+\beta x_{(n-i+1)}} = 0, \tag{2.8}$$

$$\frac{(n-k)\lambda d_x}{1-\lambda} - \sum_{i=1}^{k}x_{(n-i+1)} + \sum_{i=1}^{k}\frac{x_{(n-i+1)}}{\alpha+\beta x_{(n-i+1)}} = 0, \tag{2.9}$$

$$\frac{k-n\lambda}{\gamma(1-\lambda)} = 0. \tag{2.10}$$

From (2.10), we have $\lambda = k/n$ from which (2.6) follows. Thus (2.8) and (2.9) are simplified to (2.4) and (2.5), respectively. This proves (a).

(b) We plug $\lambda = k/n$ in (2.7) and obtain

$$\begin{aligned}\log L_c^*(\alpha, \beta; \boldsymbol{x_k}) \quad &\propto \quad (n-k)\log\left(1-\frac{k}{n}\right) - (\alpha+2)\sum_{i=1}^{k}\log x_{(n-i+1)} - \beta\sum_{i=1}^{k} x_{(n-i+1)} \\ &\quad + \sum_{i=1}^{k}\log\left(\alpha x_{(n-i+1)} + \beta x_{(n-i+1)}^2\right) + k\log\frac{k}{n} + k\alpha\log d_x + k\beta d_x.\end{aligned}$$

Taking the second partial derivatives of $L_c^*(\alpha, \beta; \boldsymbol{x_k})$ yields:

$$\begin{aligned}\frac{\partial^2 \log L_c^*(\alpha, \beta; \boldsymbol{x_k})}{\partial \alpha^2} \quad &\propto \quad -\sum_{i=1}^{k}\frac{1}{(\alpha+\beta x_{(n-i+1)})^2}, \\ \frac{\partial^2 \log L_c^*(\alpha, \beta; \boldsymbol{x_k})}{\partial \beta^2} \quad &\propto \quad -\sum_{i=1}^{k}\frac{x_{(n-i+1)}^2}{(\alpha+\beta x_{(n-i+1)})^2}, \\ \frac{\partial^2 \log L_c^*(\alpha, \beta; \boldsymbol{x_k})}{\partial \alpha \partial \beta} \quad &\propto \quad -\sum_{i=1}^{k}\frac{x_{(n-i+1)}}{(\alpha+\beta x_{(n-i+1)})^2}.\end{aligned}$$

Observe that $\dfrac{\partial^2 \log L_c^*(\alpha, \beta; \boldsymbol{x_k})}{\partial \alpha^2} < 0$ and by Cauchy-Schwarz inequality,

$$\left(\frac{\partial^2 \log L_c^*(\alpha, \beta; \boldsymbol{x_k})}{\partial \alpha \partial \beta}\right)^2 < \left(\frac{\partial^2 \log L_c^*(\alpha, \beta; \boldsymbol{x_k})}{\partial \alpha^2}\right)\left(\frac{\partial^2 \log L_c^*(\alpha, \beta; \boldsymbol{x_k})}{\partial \beta^2}\right),$$

for all α, β and $\boldsymbol{x_k}$. Hence, it follows that $L_c(\boldsymbol{\theta}; \boldsymbol{x_k})$ has at most one local maximum, which, if exists, has to be the unique global maximum. This completes the proof of (b). □

Next we consider the important question of whether the system of normal equations (2.4) and (2.5) has a positive solution. In order to answer this question, we introduce some notation so that we can eliminate the secondary parameter β and focus on the tail parameter α, which is the main parameter of interest. We start by defining

$$T_1 : \quad = \quad \sum_{i=1}^{k}(\log X_{(n-i+1)} - \log d_x) = \sum_{i=1}^{k}\log\left(X_{(n-i+1)}/d_x\right)$$

and

$$T_2 : \;=\; \sum_{i=1}^{k}(X_{(n-i+1)} - d_x).$$

Observe that both T_1 and T_2 are positive. Also, for $n \geq 1$ and $1 \leq k \leq n$, define

$$G_{n,k}(u; \boldsymbol{x_k}) := \sum_{i=1}^{k} \frac{x_{(n-i+1)}}{kx_{(n-i+1)} + u(T_2 - T_1 x_{(n-i+1)})} - 1, \quad u \in [0, k/T_1] \tag{2.11}$$

and note that $G_{n,k}(0; \boldsymbol{x_k}) = 0$. With these notation we have the following result which gives the normal equation for $\hat{\alpha}$.

Proposition 2.2. *(a) For any order statistics $\boldsymbol{x_k}$ of a given sample with size $n \geq 1$, $G_{n,k}(u; \boldsymbol{x_k})$ is a well-defined continuous function of u at every point of the closed interval $[0, k/T_1]$.*
(b) $(\hat{\alpha}, \hat{\beta}) \in (0, \infty) \times (0, \infty)$ satisfies the normal equations (2.4) and (2.5) if and only if $\hat{\alpha} \in (0, k/T_1)$ satisfies

$$G_{n,k}(\hat{\alpha}; \boldsymbol{X_k}) = 0 \tag{2.12}$$

and $\hat{\beta} = (k - \hat{\alpha}T_1)/T_2$.
(c) There is at most one $\hat{\alpha} \in (0, k/T_1)$ satisfying (2.12).

Proof. (a) We just need to verify that $G_{n,k}$ has no pole in $[0, k/T_1]$. This is obvious because for each $i = 1, 2, \ldots, k$, and $u \in [0, k/T_1]$,

$$kx_{(n-i+1)} + u(T_2 - T_1 x_{(n-i+1)}) > (k - uT_1)\, x_{(n-i+1)} > 0.$$

(b) If $(\hat{\alpha}, \hat{\beta}) \in (0, \infty) \times (0, \infty)$ satisfies (2.4) and (2.5), then multiplying (2.4) by $\hat{\alpha}$ and (2.5) by $\hat{\beta}$ and adding we have $\hat{\alpha}T_1 + \hat{\beta}T_2 = k$ which gives $\hat{\beta} = (k - \hat{\alpha}T_1)/T_2$. Using this we can eliminate $\hat{\beta}$ from (2.5) to get (2.12). The positiveness of $\hat{\beta}$ implies $\hat{\alpha} \in (0, k/T_1)$.

To prove the converse observe that (2.12) and $\hat{\beta} = (k - \hat{\alpha}T_1)/T_2$ yields $\hat{\beta} > 0$ and (2.5). Multiplying both sides of (2.5) by $\hat{\beta}$, we have

$$\sum_{i=1}^{k} \frac{\hat{\beta} X_{(n-i+1)}}{\hat{\alpha} + \hat{\beta} X_{(n-i+1)}} = \hat{\beta}T_2 = k - \hat{\alpha}T_1,$$

which yields

$$\sum_{i=1}^{k} \frac{\hat{\alpha}}{\hat{\alpha} + \hat{\beta} X_{(n-i+1)}} = \hat{\alpha}T_1$$

from which (2.4) follows since $\hat{\alpha} > 0$.
(c) This part follows from part (b) of Proposition 2.1. □

The above result is important for two reasons - firstly it gives normal equations for the tail parameter (as well as for the parameter β) and secondly it enables us to establish the existence (with high probability for a large sample), consistency, and asymptotic normality of the unconditional MLE, the parameter estimates based on the entire data set.

Remark 2.3. Putting $k = n$ and $d_x = \hat{x}_0 := X_{(1)}$ in (2.4) - (2.6) we obtain the normal equations for the unconditional MLE of $\boldsymbol{\theta}$ as follows:

$$\sum_{i=1}^{n} \frac{1}{\hat{\alpha} + \hat{\beta} X_i} = \sum_{i=1}^{n} \log \frac{X_i}{\hat{x}_0}, \tag{2.13}$$

$$\sum_{i=1}^{n} \frac{X_i}{\hat{\alpha} + \hat{\beta} X_i} = \sum_{i=1}^{n} (X_i - \hat{x}_0), \tag{2.14}$$

$$\hat{\gamma} = \hat{x}_0^{\hat{\alpha}} e^{\hat{\beta}\hat{x}_0}. \tag{2.15}$$

Theorem 2.4. *(a) The probability that the normal equations* (2.13)-(2.15) *of the unconditional MLE for $\boldsymbol{\theta}$ has a unique solution $\hat{\boldsymbol{\theta}}_{\boldsymbol{n}} := \{\hat{\alpha}_n, \hat{\beta}_n, \hat{\gamma}_n\}$ converges to* 1 *as $n \to \infty$ and the unconditional MLE $\hat{\boldsymbol{\theta}}_{\boldsymbol{n}}$ is consistent for $\boldsymbol{\theta}$.*
(b) $\hat{\alpha}_n$ and $\hat{\beta}_n$ are asymptotically jointly normal with asymptotic means α and β respectively and asymptotic variance-covariance matrix $\frac{1}{n}\boldsymbol{W}^{-1}$, where

$$\boldsymbol{W} := \begin{pmatrix} E\left((\alpha + \beta X_1)^{-2}\right) & E\left(X_1(\alpha + \beta X_1)^{-2}\right) \\ E\left(X_1(\alpha + \beta X_1)^{-2}\right) & E\left(X_1^2(\alpha + \beta X_1)^{-2}\right) \end{pmatrix}, \tag{2.16}$$

which is invertible by Cauchy-Schwarz inequality.

In order to prove this theorem, we will need the following lemmas.

Lemma 2.5. *(a)* $E\left(\dfrac{X_1}{\alpha + \beta X_1}\right) = E(X_1 - x_0)$.
(b) $E\left(\dfrac{1}{\alpha + \beta X_1}\right) = E\left(\log \dfrac{X_1}{x_0}\right)$.

Proof. (a) Using the forms of the density and the survival functions of X_1, we have

$$\begin{aligned} E\left(\frac{X_1}{\alpha + \beta X_1}\right) &= \int_{x_0}^{\infty} \frac{x}{\alpha + \beta x} \gamma x^{-\alpha - 1} e^{-\beta x} (\alpha + \beta x) dx \\ &= \int_{x_0}^{\infty} \gamma x^{-\alpha} e^{-\beta x} dx \\ &= \int_{0}^{\infty} P(X_1 > x) dx - x_0 = E(X_1 - x_0). \end{aligned}$$

(b) To prove this observe that

$$\begin{aligned}
E\left(\log\frac{X_1}{x_0}\right) &= \int_0^\infty P\left(\log\frac{X_1}{x_0} > t\right)dt \\
&= \int_0^\infty P\left(X_1 > x_0e^t\right)dt \\
&= \int_0^\infty \gamma(x_0e^t)^{-\alpha}\exp\left(-\beta x_0e^t\right)dt,
\end{aligned}$$

which by a change of variable $x = x_0e^t$ becomes

$$\begin{aligned}
&= \int_{x_0}^\infty \gamma x^{-\alpha-1}e^{-\beta x}dx \\
&= \int_{x_0}^\infty \frac{1}{\alpha+\beta x}\gamma x^{-\alpha-1}e^{-\beta x}(\alpha+\beta x)dx \\
&= E\left(\frac{1}{\alpha+\beta X_1}\right)
\end{aligned}$$

and this completes the proof. □

Lemma 2.6. *For $\hat{x}_0 = X_{(1)}$ the following hold.*
(a) $\sqrt{n}(\hat{x}_0 - x_0) \xrightarrow{p} 0$.
(b) $\sqrt{n}\left(\log\hat{x}_0 - \log x_0\right) \xrightarrow{p} 0$.

Proof. (a) By the Markov inequality, it is enough to show that $E\left(\sqrt{n}(\hat{x}_0 - x_0)\right) \to 0$ as $n \to \infty$. Observe that

$$\begin{aligned}
0 \le E\left(\sqrt{n}(\hat{x}_0 - x_0)\right) &= \int_0^\infty P\left(\sqrt{n}(\hat{x}_0 - x_0) > t\right)dt \\
&= \int_0^\infty \left(P\left(X_1 > x_0 + \frac{t}{\sqrt{n}}\right)\right)^n dt \\
&= \int_0^\infty \left(1 + \frac{t}{x_0\sqrt{n}}\right)^{-\alpha n} e^{-\beta t\sqrt{n}}dt \\
&\le \int_0^\infty e^{-\beta t\sqrt{n}}dt = \frac{1}{\beta\sqrt{n}} \to 0
\end{aligned}$$

as $n \to \infty$ and this finishes the proof of part (a).
(b) Part (b) follows from part (a) using the inequality $|\log y - \log x_0| \le C|y - x_0|$ for all y in a neighborhood of x_0 and for some $C > 0$. □

Proof of Theorem 2.4. (a) In this case $k = n$ (see Remark 2.3). For simplicity of notation we use G_n to denote $G_{n,n}$, i.e.,

$$G_n(u; \boldsymbol{X_n}) = \frac{1}{n} \sum_{i=1}^{n} \frac{X_i}{X_i + u\left(\frac{T_2}{n} - \frac{T_1}{n} X_i\right)} - 1, \quad u \in [0, n/T_1].$$

We will eventually show that for all $\epsilon > 0$,

$$\lim_{n\to\infty} P\Big[G_n(u; \boldsymbol{X_n}) = 0 \text{ has a solution in } (\alpha - \epsilon, \alpha + \epsilon)\Big] = 1. \tag{2.17}$$

To prove (2.17), we start by introducing some notation. Define

$$\widetilde{G}_n(u; \boldsymbol{X_n}) := \frac{1}{n} \sum_{i=1}^{n} \frac{X_i}{X_i + u\,(B - AX_i)} - 1,$$

where, in view of Lemma 2.5,

$$\begin{aligned} A &:= E\left(\frac{1}{\alpha + \beta X_1}\right) = E\left(\log \frac{X_1}{x_0}\right) > 0, \text{ and} \\ B &:= E\left(\frac{X_1}{\alpha + \beta X_1}\right) = E(X_1 - x_0) > 0. \end{aligned} \tag{2.18}$$

Define

$$G(u) := E\left(\frac{X_1}{X_1 + u(B - AX_1)}\right) - 1. \tag{2.19}$$

Since

$$\alpha A + \beta B = 1, \tag{2.20}$$

it follows that $1 - uA > 0$ for all u in a small enough neighborhood $N(\alpha)$ of α, which yields that $\widetilde{G}_n(u; \boldsymbol{X_n})$ has no pole and hence is well-defined on $N(\alpha)$. Similarly $G(u)$ is also well-defined on $N(\alpha)$ because for all $u \in N(\alpha)$, $-1 < G(u) < E\,(X_1/uB) - 1 < \infty$.

From (2.20) and (2.18) we obtain

$$G(\alpha) = 0. \tag{2.21}$$

By a dominated convergence argument, we differentiate under the integral sign in (2.19) and obtain

$$G'(\alpha) = E\left[\frac{(AX_1 - B)X_1}{(X_1 + \alpha(B - AX_1))^2}\right] = E(YZ)$$

where

$$Y = \frac{AX_1 - B}{X_1 + \alpha(B - AX_1)} \text{ and } Z = \frac{X_1}{X_1 + \alpha(B - AX_1)}.$$

Since $Z = 1 + \alpha Y$, we have $\mathrm{Cov}(Y, Z) = \alpha \mathrm{Var(Y)} > 0$ and $E(Y) = \alpha^{-1}(E(Z) - 1) = 0$ by (2.21). This shows

$$G'(\alpha) = E(YZ) > E(Y)E(Z) = 0. \tag{2.22}$$

By (2.20), we can find a small enough neighborhood $N^*(\alpha) \subseteq N(\alpha)$ of α and $\delta > 0$ such that $1 - ux > 0$ whenever $u \in N^*(\alpha)$ and $|x - A| < \delta$. We will show that $G_n(u; \boldsymbol{X_n}) \xrightarrow{p} G(u)$ as $n \to \infty$ for all $u \in N^*(\alpha)$ in the following. Since, by the weak law of large numbers, $\widetilde{G}_n(u; \boldsymbol{X_n}) \xrightarrow{p} G(u)$ as $n \to \infty$ for all $u \in N(\alpha)$, it is enough to show that

$$G_n(u; \boldsymbol{X_n}) - \widetilde{G}_n(u; \boldsymbol{X_n}) = o_p(1) \tag{2.23}$$

for all $u \in N^*(\alpha) \subseteq N(\alpha)$. Since $T_1/n \xrightarrow{p} A$ and $T_2/n \xrightarrow{p} B$ as $n \to \infty$, we obtain using bivariate mean value theorem that

$$G_n(u; \boldsymbol{X_n}) - \widetilde{G}_n(u; \boldsymbol{X_n}) = \left(\frac{T_1}{n} - A\right) R_n + \left(B - \frac{T_2}{n}\right) S_n ,$$

where

$$R_n = \frac{1}{n} \sum_{i=1}^{n} \frac{uX_i^2}{(X_i + u(\xi_n - \eta_n X_i))^2}$$

and

$$S_n = \frac{1}{n} \sum_{i=1}^{n} \frac{uX_i}{(X_i + u(\xi_n - \eta_n X_i))^2}$$

with $\eta_n \xrightarrow{p} A$ and $\xi_n \xrightarrow{p} B$ as $n \to \infty$. Then in order to show (2.23), it is enough to establish that both R_n and S_n are tight for all $u \in N^*(\alpha)$. Let $\Omega_n := \{|\eta_n - A| < \delta, \xi_n > B/2\}$, where δ is as above. Then on the event Ω_n, we have for all $u \in N^*(\alpha)$,

$$R_n \le \frac{4}{uB^2} \frac{1}{n} \sum_{i=1}^{n} X_i^2,$$

which is tight because $n^{-1} \sum_{i=1}^{n} X_i^2 \xrightarrow{p} E(X_1^2) < \infty$. Similarly on Ω_n, for all $u \in N^*(\alpha)$,

$$|S_n| \le \frac{4}{uB^2} \frac{1}{n} \sum_{i=1}^{n} |X_i|,$$

which is also tight. Since $P(\Omega_n) \to 1$ as $n \to \infty$, the required tightnesses follow establishing (2.23) and hence $G_n(u; \boldsymbol{X_n}) \xrightarrow{p} G(u)$ as $n \to \infty$ for all $u \in N^*(\alpha)$.

We are now all set to complete the proof. By (2.21) and (2.22), we can find $\epsilon_0 > 0$ small enough such that $(\alpha - \epsilon_0, \alpha + \epsilon_0) \subseteq N^*(\alpha)$, $G(u) > 0$ for all $u \in (\alpha, \alpha + \epsilon_0)$ and $G(u) < 0$ for all $u \in (\alpha - \epsilon_0, \alpha)$. Since $G_n(u; \boldsymbol{X_n}) \xrightarrow{p} G(u)$ as $n \to \infty$ for all $u \in (\alpha - \epsilon_0, \alpha + \epsilon_0)$, (2.17) follows for all $\epsilon \in (0, \epsilon_0)$ and hence for all $\epsilon > 0$. The existence of $\hat{\boldsymbol{\theta}}_{\boldsymbol{n}}$ (with probability tending to 1 as $n \to \infty$) follows from (2.17) by continuity of $u \mapsto G_n(u; \boldsymbol{X_n})$, part (b) of Proposition 2.2 and the observation that $nT_1^{-1} \xrightarrow{p} A^{-1} > \alpha + \epsilon$ for small enough $\epsilon > 0$. The uniqueness is obvious from part (c) of Proposition 2.2. The consistency of $\hat{\alpha}_n$ follows by choosing $\epsilon > 0$ arbitrarily small in (2.17). The consistency of $\hat{\beta}_n$ is straightforward from the equations $\hat{\beta}_n = (n - \hat{\alpha}_n T_1)/T_2$ and (2.20) and finally (2.15) yields the consistency of $\hat{\gamma}_n$.

(b) In order to establish the asymptotic normality of $\hat{\alpha}_n$ and $\hat{\beta}_n$, we start with the following notation. Define, for $s, t, x > 0$,

$$H_n(s,t,x) := \frac{1}{n}\sum_{i=1}^{n} \frac{1}{s + tX_i} - \frac{1}{n}\sum_{i=1}^{n} \log \frac{X_i}{x},$$
$$K_n(s,t,x) := \frac{1}{n}\sum_{i=1}^{n} \frac{X_i}{s + tX_i} - \frac{1}{n}\sum_{i=1}^{n} (X_i - x).$$

By Remark 2.3, $\hat{\alpha}_n$ and $\hat{\beta}_n$ satisfies $H_n(\hat{\alpha}_n, \hat{\beta}_n, \hat{x}_0) = K_n(\hat{\alpha}_n, \hat{\beta}_n, \hat{x}_0) = 0$. Let $\tilde{\alpha}_n$ and $\tilde{\beta}_n$ be the unconditional MLE of α and β respectively when x_0 is known. Then $\tilde{\alpha}_n$ and $\tilde{\beta}_n$ satisfies $H_n(\tilde{\alpha}_n, \tilde{\beta}_n, x_0) = K_n(\tilde{\alpha}_n, \tilde{\beta}_n, x_0) = 0$. When x_0 is known, the support of the density (2.2) does not depend on the parameter values and the corresponding information matrix is given by $\boldsymbol{W}$ as in (2.16). Hence using the asymptotic properties of the MLE (see, for example, [30, p. 564]), it can easily be deduced that $\sqrt{n}\big(\tilde{\alpha}_n - \alpha, \tilde{\beta}_n - \beta\big)$ converges in distribution to a bivariate normal distribution with mean vector $\mathbf{0}$ and variance-covariance matrix $\boldsymbol{W}$. To complete the proof it is enough to show that $\sqrt{n}(\tilde{\alpha}_n - \hat{\alpha}_n) \xrightarrow{p} 0$ and $\sqrt{n}(\tilde{\beta}_n - \hat{\beta}_n) \xrightarrow{p} 0$.

To this end, define

$$\widetilde{H}_n(s,t) := \frac{1}{n}\sum_{i=1}^{n} \frac{1}{s + tX_i} - E\left(\log \frac{X_1}{x_0}\right),$$
$$\widetilde{K}_n(s,t) := \frac{1}{n}\sum_{i=1}^{n} \frac{X_i}{s + tX_i} - E\left(X_1 - x_0\right)$$

for $s,t>0$. Then we have

$$\begin{aligned}
U_n &:= \sqrt{n}\left(\widetilde{H}_n(\hat{\alpha}_n,\hat{\beta}_n)-\widetilde{H}_n(\tilde{\alpha}_n,\tilde{\beta}_n)\right)\\
&= \sqrt{n}\left(\widetilde{H}_n(\hat{\alpha}_n,\hat{\beta}_n)-H_n(\hat{\alpha}_n,\hat{\beta}_n,\hat{x}_0)\right)+\sqrt{n}\left(H_n(\tilde{\alpha}_n,\tilde{\beta}_n,x_0)-\widetilde{H}_n(\tilde{\alpha}_n,\tilde{\beta}_n)\right)\\
&= \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\log\frac{X_i}{\hat{x}_0}-E\left(\log\frac{X_1}{x_0}\right)\right)+\sqrt{n}\left(E\left(\log\frac{X_1}{x_0}\right)-\frac{1}{n}\sum_{i=1}^{n}\log\frac{X_i}{x_0}\right)\\
&= \sqrt{n}\left(\log x_0-\log\hat{x}_0\right)\xrightarrow{p}0 \qquad (2.24)
\end{aligned}$$

by Lemma 2.6. Similarly we can show

$$V_n := \sqrt{n}\left(\widetilde{K}_n(\hat{\alpha}_n,\hat{\beta}_n)-\widetilde{K}_n(\tilde{\alpha}_n,\tilde{\beta}_n)\right)\xrightarrow{p}0. \qquad (2.25)$$

Using bivariate mean value theorem we have

$$\begin{aligned}
U_n &= \sqrt{n}(\tilde{\alpha}_n-\hat{\alpha}_n)\frac{\partial\widetilde{H}_n}{\partial s}(\rho_n,\zeta_n)+\sqrt{n}(\tilde{\beta}_n-\hat{\beta}_n)\frac{\partial\widetilde{H}_n}{\partial t}(\rho_n,\zeta_n),\\
&=: \sqrt{n}(\tilde{\alpha}_n-\hat{\alpha}_n)W_{11}^{(n)}+\sqrt{n}(\tilde{\beta}_n-\hat{\beta}_n)W_{12}^{(n)}, \qquad (2.26)
\end{aligned}$$

where $\rho_n\xrightarrow{p}\alpha$ and $\zeta_n\xrightarrow{p}\beta$ as $n\to\infty$, and

$$\begin{aligned}
V_n &= \sqrt{n}(\tilde{\alpha}_n-\hat{\alpha}_n)\frac{\partial\widetilde{K}_n}{\partial s}(\tilde{\rho}_n,\tilde{\zeta}_n)+\sqrt{n}(\tilde{\beta}_n-\hat{\beta}_n)\frac{\partial\widetilde{K}_n}{\partial t}(\tilde{\rho}_n,\tilde{\zeta}_n)\\
&=: \sqrt{n}(\tilde{\alpha}_n-\hat{\alpha}_n)W_{21}^{(n)}+\sqrt{n}(\tilde{\beta}_n-\hat{\beta}_n)W_{22}^{(n)}, \qquad (2.27)
\end{aligned}$$

where $\tilde{\rho}_n\xrightarrow{p}\alpha$ and $\tilde{\zeta}_n\xrightarrow{p}\beta$ as $n\to\infty$. Defining $\boldsymbol{W^{(n)}} := (W_{ij}^{(n)})_{1\le i,j\le 2}$ we get from the equations (2.26) and (2.27) that

$$\sqrt{n}(\tilde{\alpha}_n-\hat{\alpha}_n)=\frac{W_{22}^{(n)}U_n-W_{12}^{(n)}V_n}{\det\boldsymbol{W^{(n)}}} \qquad (2.28)$$

and using an argument similar to the proof of (2.23), we get that $\boldsymbol{W^{(n)}}\xrightarrow{p}-\boldsymbol{W}$, where $\boldsymbol{W}$ is as in (2.16). This, in particular, implies $\det\boldsymbol{W^{(n)}}\xrightarrow{p}\det\boldsymbol{W}>0$ by Cauchy-Schwarz inequality. Hence using (2.24), (2.25) and (2.28), it follows that $\sqrt{n}(\tilde{\alpha}_n-\hat{\alpha}_n)\xrightarrow{p}0$. By a similar argument we can also show that $\sqrt{n}(\tilde{\beta}_n-\hat{\beta}_n)\xrightarrow{p}0$ and this completes the proof of Theorem 2.4. □

3 Applications

Extensive simulation trials were conducted to validate the estimator developed in the previous section. For simulated data from the exponentially tempered Pareto

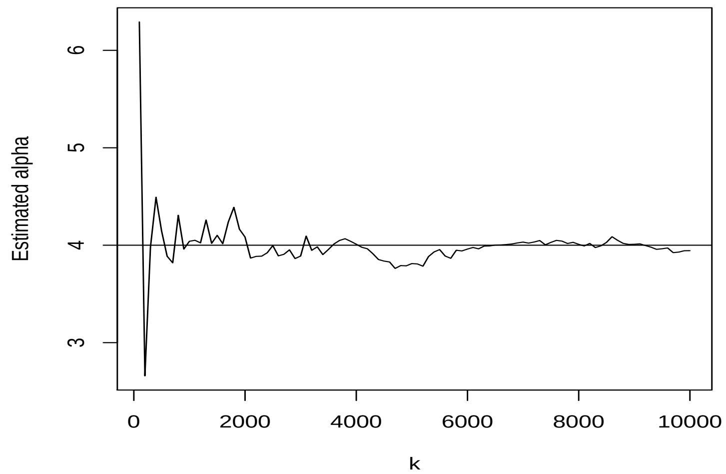

Figure 1: Behavior of α estimate as a function of the number k of upper order statistics used.

distribution (2.1) the parameter estimates were generally close to the assumed values, so long as the data range was sufficient to capture both the power law behavior and the subsequent tempering at the highest values. If the tempering parameter β is very large, so that the term $x^{-\alpha}$ in (2.1) hardly varies over the data range, then estimates of α are unreliable. If β is so small that the term $e^{-\beta x}$ in (2.1) hardly varies over the data range, then estimates of β are widely variable. In either case, a simpler model (exponential or power law) is indicated. Naturally the tempered Pareto model (2.1) is only appropriate when both terms are significant. This can be seen in a log-log plot of data versus ranks, where a straight line eventually falls away due to tempering. If the data follows a straight line over the entire tail, then a simpler Pareto model is indicated. If the data follows a straight line on a semi-log plot, then an exponential model is appropriate. Several illustrative examples follow.

Figure 1 illustrates the behavior of the tail estimate $\hat{\alpha}$ from Proposition 2.2 as a function of the number k of upper order statistics used. The simulated data comes from the tempered Pareto distribution (2.1) with lower limit $x_0 = 1$, tail parameter $\alpha = 4$, and tempering parameter $\beta = 0.5$. Simulation was performed using a standard rejection method. It is apparent that, once the number k of upper order statistics used reaches a few percent of the total sample size of $n = 10,000$, the α estimate settles down to a reasonable value. In practice, any sufficiently large value of k will

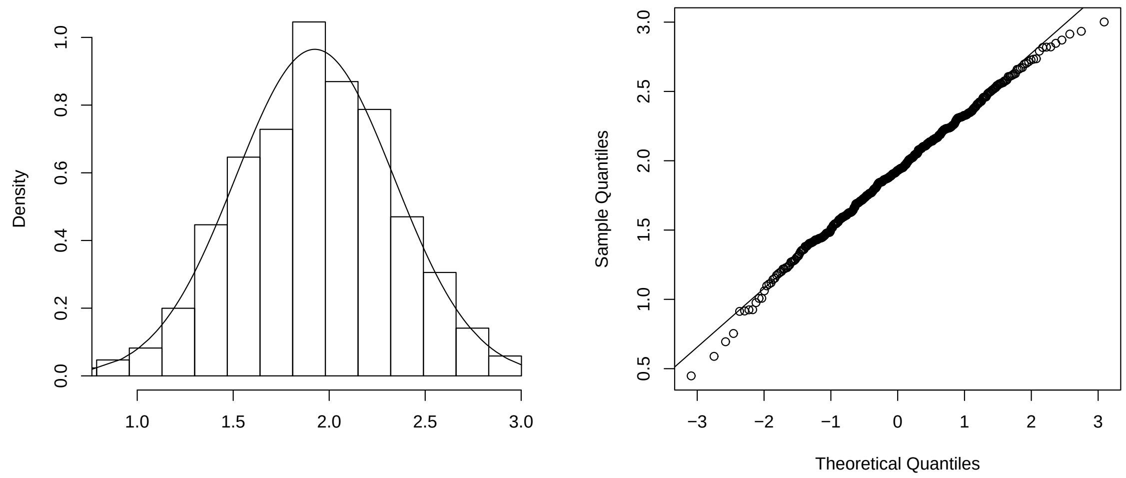

Figure 2: Evidence of normal sampling distribution for α estimates.

give a reasonable fit.

Figure 2 shows a histogram and normal quantile-quantile plot for α estimates obtained from 500 replications of the same tempered Pareto simulation. In this case we set $\alpha = 2$, $\beta = 0.5$ and $x_0 = 1$. The sample size is $n = 1,000$ and we use the $k = 500$ largest observations to estimate the distribution parameters. The corresponding plots are similar for various values of the parameters. We conclude that the sampling distribution of the parameters is reasonably well approximated by a normal distribution. Note that the asymptotic normality of the parameter estimates based on the entire data set was established in Theorem 2.4. The asymptotic theory for the general case $k < n$ is much more difficult.

Tempered stable laws [24] have power law tails modified by exponential tempering. Therefore, the exponentially tempered Pareto model (2.1) gives a simple way to approximate the tail behavior, and estimate the parameters. The simple and efficient exponential rejection method of [2] was used to simulate tempered stable random variates. Figure 3 shows the upper tail of simulated data following a tempered stable distribution. The largest $k = 100$ of the $n = 1000$ order statistics are plotted. The underlying stable distribution has tail parameter $\alpha = 1.5$, skewness 1, mean 0, and scale $\sigma = 4$ in the usual parameterization [29], and the truncation parameter is $\beta = 0.01$. Figure 3 is a log-log plot of the sorted data $X_{(i)}$ versus rank $(n - i)/n$ exhibiting the power-law tail as a straight line that eventually falls off due to temper-

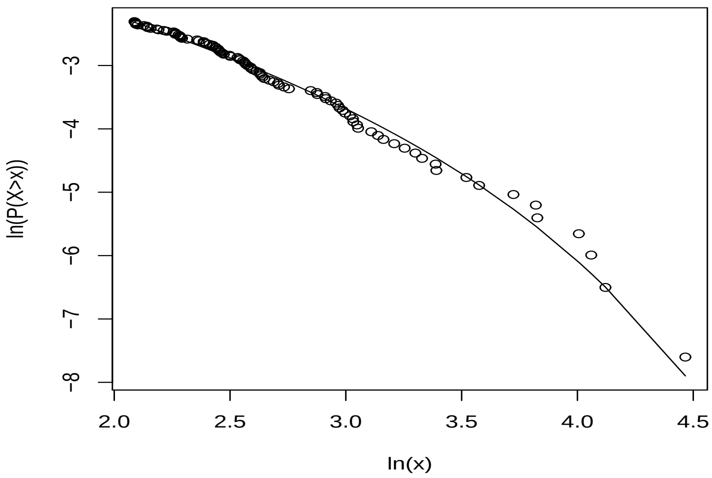

Figure 3: Tempered Pareto fit to the upper tail of simulated tempered stable data.

ing. It is apparent that the tempered Pareto model gives a reasonable fit to the more complicated tempered stable distribution, which has no closed form. Similar results were obtained for other values of the parameters. For smaller values of β the data plot more closely resembles a straight line (power law tail).

Next we apply the conditional MLE developed in this paper to several real data sets. First we consider a data set from hydrology. Hydraulic conductivity K measures the ability of water to pass through a porous medium. This is a function of porosity (percent of the material consisting of pore space) as well as connectivity. K data was collected in boreholes at the MAcroDispersion Experiment (MADE) site on a Air Force base near Columbus MS. The data set has been analyzed by several researchers, see for example [7]. Figure 4 shows a log-log plot of the largest 10% of the data, with the best-fitting tempered Pareto model (2.1), where the parameters were fit using Proposition 2.2. Absolute values of K were modeled in order to combine the heavy tails at both extremes. The largest $k = 262$ (approximately 10%) values of the data were used. It is apparent that the tempered Pareto model gives a good fit to the data. Since the data deviates from a straight line on this log-log plot, a simple Pareto model would be inadequate.

Figure 5 shows the constraint function $G_{n,k}(u; \boldsymbol{x_k})$ from Proposition 2.2 for the same data set, as a function of u. The vertical line on the graph is the upper bound of $u = k/T_1$. The constraint function has roots at $u = 0$ (by definition) and at

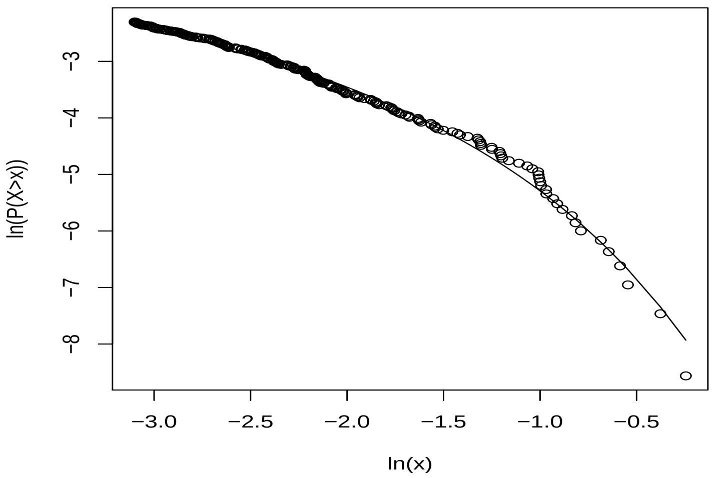

Figure 4: Tempered Pareto model for hydraulic conductivity data.

$u = \hat{\alpha} = 0.6171$ which is the estimate of the tail parameter. In view of Proposition 2.2 this is the unique solution to the normal equations. The remaining parameter estimates are $\hat{\beta} = 5.2397$ and $\hat{\gamma} = 0.0187$. This is a relatively heavy tail with a strong truncation.

Figure 6 fits a tempered Pareto model to absolute log returns in the daily price of stock for Amazon, Inc. The ticker symbol is AMZN. The data ranges from 1 January 1998 to 30 June 2003 ($n = 1378$). Based on the upper 10% of the data $k = 138$, the best fitting parameter values (conditional MLE from Proposition 2.2) are $\hat{\alpha} = 0.578$, $\hat{\beta} = 0.281$, and $\hat{\gamma} = 0.567$. The data shows a classic power law shape, linear on this log-log plot, but eventually falls off at the largest values. This indicates an opportunity to improve prediction using a tempered model.

Figure 7 shows the tempered Pareto fit to daily precipitation data at Tombstone AZ between 1 July 1893 and 31 December 2001. The fit is based on the largest $k = 2,608$ observations, which constitutes the upper half of the nonzero data. The fitted parameters were $\hat{\alpha} = 0.212$, $\hat{\beta} = 0.00964$, and $\hat{\gamma} = 1.56$. The data tail is clearly lighter than a pure power law model (straight line) but is fit well by the tempered model. A semi-log plot (not shown) was examined to rule out a simpler exponential fit.

We conclude this section with some practical advice for tail modeling. If a data set exhibits tails that are heavier than Gaussian (e.g., if there are numerous outliers),

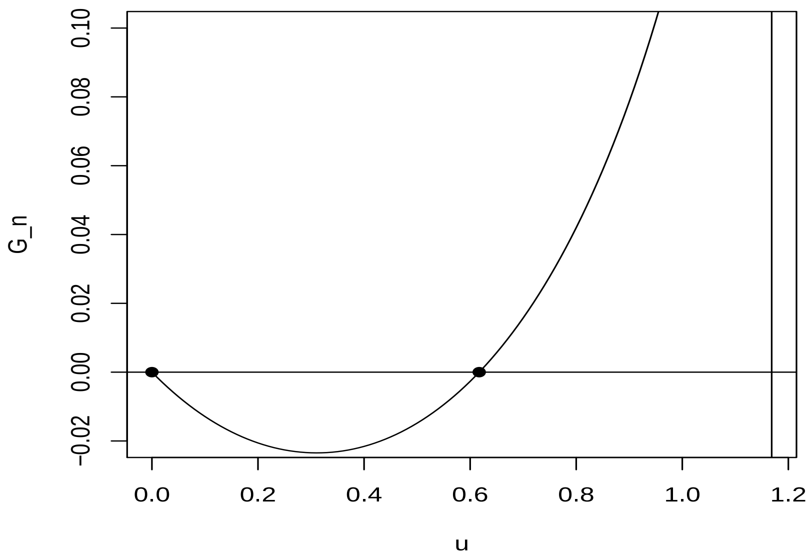

Figure 5: Constraint function $G_{n,k}(u; \boldsymbol{x_k})$ for the hydraulic conductivity data, showing unique positive root.

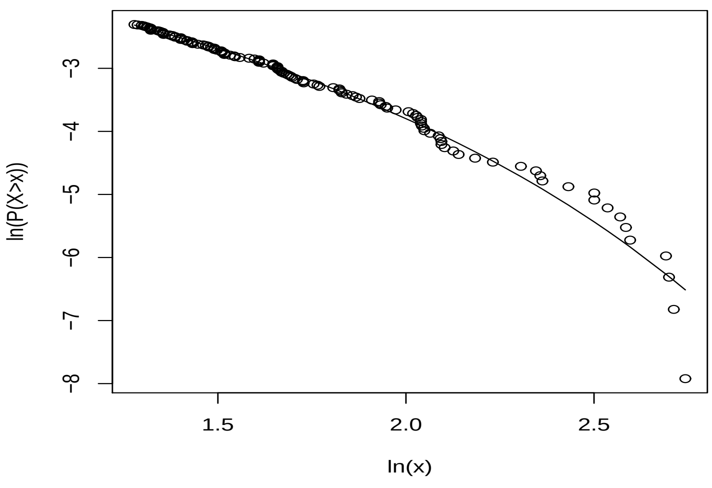

Figure 6: Tempered Pareto model for AMZN stock daily price returns.

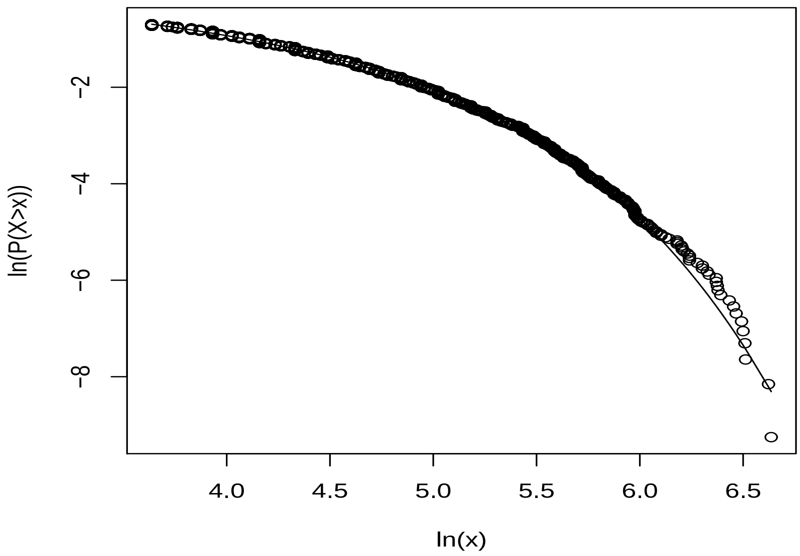

Figure 7: Tempered Pareto model for daily precipitation data.

then it makes sense to consider alternative models. Plotting order statistics of the data $X_{(i)}$ versus ranks $(n-i)/n$ gives a simple method for initial model selection. For signed data, it is often advisable to begin by examining the absolute values. If the points corresponding to the largest data values follow a straight line on a log-log plot, this indicates a power-law probability tail. If a semi-log plot of the same data appears to follow a straight line, this suggests an exponential tail (e.g., a Laplace model, if the data distribution appears symmetric). Upward or downward curvature on those plots indicates a heavier or lighter tail than the pure power-law or exponential. For example, the downward curvature in Figure 7 indicates that the data tail is lighter than a power-law. A semi-log plot of the same data (not shown) produced an upward curving shape, indicating that the data tail is heavier than an exponential. Then it makes sense to consider an intermediate model with "semi-heavy" tails, for which the methods of this paper may be useful. Some additional discussion of these intermediate models may be found in [8].

4 Conclusions

Tempered Pareto distributions are useful to model heavy tailed data, in cases where a pure power law places too much probability mass on the extreme tail. The simple form of this probability law facilitates the development of a maximum likelihood estimator (MLE) for the distribution parameters, accomplished in this paper. Those estimates are proven to be consistent and asymptotically normal. In some practical applications, including the tempered stable model, it is only the upper tail of the data that follows a tempered power law. For that reason, we also develop a conditional MLE in this paper, based on the upper tail of the data. The conditional MLE is easily computable. A detailed simulation study was performed to validate the performance of the MLE. In cases where the tempered Pareto model would be appropriate, the conditional MLE is reasonably accurate, and its sampling distribution appears to be well approximated by a normal law. Tempered stable laws are useful models in geophysics, but parameter estimation for this model is an open problem. Simulation demonstrates that the tempered Pareto model is a reasonable approximation, for which efficient parameter estimation can be accomplished, via the methods of this paper. Finally, data sets from hydrology, finance, and atmospheric science are examined. In each case, the methods of this paper are used to fit a reasonable and predictive tempered Pareto model.

References

[1] I.B. Aban, M.M. Meerschaert and A.K. Panorska (2006) Parameter estimation methods for the truncated Pareto distribution. *J. Amer. Statist. Assoc.: Theory and Methods* **101**, 270–277.

[2] B. Baeumer and M.M. Meerschaert (2010) Tempered stable Lévy motion and transient super-diffusion. *J. Comput. Appl. Math.*, **233**, 2438–2448.

[3] O.E. Barndorff-Nielsen (1997) Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling. *Scand. J. Statist.* **41**(1), 1–13.

[4] O.E. Barndorff-Nielsen (1998) Processes of normal inverse Gaussian type. *Finance Stoch.* **2**, 41–68.

[5] O.E. Barndorff-Nielsen and N.N. Leonenko (2005) Burgers' turbulence problem with linear or quadratic external potential. *J. Appl. Probab.* **42**, 550–565.

[6] D.A. Benson, S.W. Wheatcraft and M.M. Meerschaert (2000) Application of a fractional advection-dispersion equation. *Water Resour. Res.* **36**, 1403–1412.

[7] D.A. Benson, R. Schumer, M.M. Meerschaert and S.W. Wheatcraft (2001) Fractional dispersion, Lévy motions, and the MADE tracer tests. *Transport in Porous Media* **42**, 211–240.

[8] A. Chakrabarty and G. Samorodnitsky (2009) Understanding heavy tails in a bounded world or, is a truncated heavy tail heavy or not? Preprint available at `people.orie.cornell.edu/~gennady/newtechreports.html`

[9] R. Gorenflo, F. Mainardi, E. Scalas and M. Raberto (2001) Fractional calculus and continuous-time finance. III. The diffusion limit. Mathematical finance (Konstanz, 2000), 171–180, *Trends Math.*, Birkhuser, Basel.

[10] P. Hall (1982) On some simple estimates of an exponent of regular variation. *J. Royal Statist. Soc. B* **44**(1), 37–42.

[11] B. Hill (1975) A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3**, 1163–1174.

[12] S. Kotz, T.J. Kozubowski and K. Podgórski (2001) *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance.* Birkhaüser, Boston.

[13] F. Mainardi, M. Raberto, R. Gorenflo and E. Scalas (2000) Fractional Calculus and continuous-time finance II: the waiting-time distribution. *Physica A* **287**, 468–481.

[14] R.N. Mantegna and H.E. Stanley (1994) Stochastic process with ultraslow convergence to a Gaussian: The truncated Lévy flight. *Phys. Rev. Lett.* **73**(22), 2946–2949.

[15] R.N. Mantegna and H.E. Stanley (1995) Scaling behavior in the dyamics of an economic index. *Nature* **376**(6535), 46–49.

[16] M.M. Meerschaert, D.A. Benson and B. Baeumer (1999) Multidimensional advection and fractional dispersion. *Phys. Rev. E* **59**, 5026–5028.

[17] M.M. Meerschaert and H.P. Scheffler (2004) Limit theorems for continuous time random walks with infinite mean waiting times. *J. Applied Probab.* **41**, No. 3, 623–638.

[18] M.M. Meerschaert, T.J. Kozubowski, F.J. Molz and S. Lu (2004) Fractional Laplace model for hydraulic conductivity, *Geophys. Res. Lett.* **31**, L08501.

[19] M.M. Meerschaert, Y. Zhang and B. Baeumer (2008) Tempered anomalous diffusion in heterogeneous systems. *Geophys. Res. Lett.* **35**, L17403.

[20] R. Metzler and J. Klafter (2000) The random walk's guide to anomalous diffusion: A fractional dynamics approach. *Phys. Rep.* **339**, 1–77.

[21] R. Metzler and J. Klafter (2004) The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics. *J. Physics A* **37**, R161–R208.

[22] R.S. Protassov (2004) EM-based maximum likelihood parameter estimation for multivariate generalized hyperbolic distributions with fixed λ. *Statistics and Computing*, **14**(1), 67–77.

[23] M. Raberto, E. Scalas and F. Mainardi (2002) Waiting-times and returns in high-frequency financial data: an empirical study. *Physica A* **314**, 749–755.

[24] J. Rosiński (2007) Tempering stable processes. *Stoch. Proc. Appl.* **117**, 677–707.

[25] L. Sabatelli, S. Keating, J. Dudley and P. Richmond (2002) Waiting time distributions in financial markets. *Eur. Phys. J. B* **27**, 273–275.

[26] E. Scalas, R. Gorenflo and F. Mainardi (2000) Fractional calculus and continuous time finance. *Phys. A* **284**, 376–384.

[27] R. Schumer, D.A. Benson, M.M. Meerschaert and S.W. Wheatcraft (2001) Eulerian derivation of the fractional advection-dispersion equation. *J. Contaminant Hydrol.* **48**, 69–88.

[28] R. Schumer, D.A. Benson, M.M. Meerschaert and B. Baeumer (2003) Multiscaling fractional advection-dispersion equations and their solutions. *Water Resour. Res.* **39**, 1022–1032.

[29] G. Samorodnitsky and M. Taqqu (1994) *Stable non-Gaussian Random Processes.* Chapman and Hall, New York.

[30] G.B. Shorack (2000) *Probability for Statisticians.* Springer Texts in Statistics, Springer-Verlag, New York.

[31] V.V. Uchaikin and V.M. Zolotarev (1999) *Chance and stability. Stable distributions and their applications.* VSP, Utrecht.