PORTUGALIAE MATHEMATICA
Vol. 60 Fasc. 2 – 2003
Nova Série

A NEW CLASS OF SEMI-PARAMETRIC ESTIMATORS OF THE SECOND ORDER PARAMETER *

M.I. Fraga Alves, M. Ivette Gomes and Laurens de Haan

Abstract: The main goal of this paper is to develop, under a semi-parametric context, asymptotically normal estimators of the *second order parameter* ρ, a parameter related to the rate of convergence of maximum values, linearly normalized, towards its limit. *Asymptotic normality* of such estimators is achieved under a *third order condition* on the tail $1-F$ of the underlying model F, and for suitably large intermediate ranks. The class of estimators introduced is dependent on some *control* or *tuning* parameters and has the advantage of providing estimators with stable sample paths, as functions of the number k of top order statistics to be considered, for large values of k; such a behaviour makes obviously less important the choice of an optimal k. The practical validation of asymptotic results for small finite samples is done by means of simulation techniques in Fréchet and Burr models.

1 – Introduction

In *Statistical Extreme Value Theory* we are essentially interested in the estimation of *parameters of rare events* like *high quantiles* and *return periods of high levels*. Those parameters depend on the *tail index* $\gamma = \gamma(F)$, of the underlying model $F(.)$, which is the shape parameter in the *Extreme Value* (EV) distribution function (d.f.),

$$(1.1)\qquad G(x) \equiv G_\gamma(x) := \begin{cases} \exp\Big\{-(1+\gamma\,x)^{-1/\gamma}\Big\}, & 1+\gamma\,x>0 \quad \text{if } \gamma\neq 0\ , \\ \exp\Big(-\exp(-x)\Big), & x\in\mathbb{R} \qquad\quad \text{if } \gamma=0\ . \end{cases}$$

Received: June 26, 2001; *Revised*: February 15, 2002.
AMS Subject Classification: Primary 60G70, 62G32; Secondary 62G05, 62E20, 65C05.
Keywords and Phrases: extreme value theory; tail inference; semi-parametric estimation; asymptotic properties.
* Research partially supported by FCT/POCTI/FEDER.

This d.f. appears as the non-degenerate limiting d.f. of the sequence of maximum values, $\{X_{n:n} = \max(X_1, X_2, \ldots, X_n)\}_{n\geq 1}$, linearly normalized, with $\{X_i\}_{i\geq 1}$ a sequence of independent, identically distributed (i.i.d.), or possibly weakly dependent random variables (r.v.'s) (Galambos [9]; Leadbetter and Nandagopalan [21]). Whenever there is such a non-degenerate limit we say that F is in the *domain of attraction* of G_γ, and write $F \in D(G_\gamma)$. Putting $U(t) := F^{\leftarrow}(1-1/t)$ for $t > 1$, where $F^{\leftarrow}(t) = \inf\{x\colon F(x) \geq t\}$ denotes the generalized inverse function of F, we have, for heavy tails, i.e., for $\gamma > 0$,

$$F \in D(G_\gamma) \quad \text{iff} \;\; 1 - F \in RV_{-1/\gamma} \quad \text{iff} \;\; U \in RV_\gamma \;, \tag{1.2}$$

where RV_β stands for the class of *regularly varying* functions at infinity with *index of regular variation* equal to β, i.e., functions $g(.)$ with infinite right endpoint, and such that $\lim_{t\to\infty} g(t\,x)/g(t) = x^\beta$, for all $x > 0$. The conditions in (1.2) characterize completely the first order behaviour of $F(\cdot)$ (Gnedenko [10]; de Haan [17]).

The second order theory has been worked out in full generality by de Haan and Stadtmüller [18]. Indeed, for a large class of heavy tail models there exists a function $A(t) \to 0$ of constant sign for large values of t, such that

$$\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} \;=\; \frac{x^\rho - 1}{\rho} \tag{1.3}$$

for every $x > 0$, where ρ (≤ 0) is a *second order parameter*, which also needs to be properly estimated from the original sample. The limit function in (1.3) must be of the stated form, and $|A(t)| \in RV_\rho$ (Geluk and de Haan [11]). Notice that as $|\rho|$ increases, the rate of convergence in the first order approximation increases as well, and this is important for approximations in real problems.

Here, for part of our results, we shall assume the validity of a third order framework, i.e., we shall assume there is a function $B(t) \to 0$, also of constant sign for large values of t, and

$$\lim_{t\to\infty} \frac{\frac{\ln U(t\,x) - \ln U(t) - \gamma \ln x}{A(t)} - \frac{x^\rho - 1}{\rho}}{B(t)} \;=\; \frac{1}{\beta}\left\{\frac{x^{\rho+\beta} - 1}{\rho+\beta} - \frac{x^\rho - 1}{\rho}\right\}. \tag{1.4}$$

Then $|B(t)| \in RV_\beta$, $\beta \leq 0$.

Under the validity of (1.3) and (1.4) we have, for every $x > 0$, and as $t \to \infty$,

$$\ln U(t\,x) - \ln U(t) \;=\; \gamma \ln x + A(t)\,\frac{x^\rho - 1}{\rho} + A(t)\,B(t)\,H(x;\rho,\beta)\,\Big(1 + o(1)\Big) \;,$$

where

$$H(x;\rho,\beta) \;=\; \frac{1}{\beta}\left\{\frac{x^{\rho+\beta}-1}{\rho+\beta}-\frac{x^{\rho}-1}{\rho}\right\}.$$

For heavy tails, estimation of the tail index γ may be based on the statistics

$$M_n^{(\alpha)}(k) := \frac{1}{k}\sum_{i=1}^{k}\Big[\ln X_{n-i+1:n}-\ln X_{n-k:n}\Big]^{\alpha}, \qquad \alpha\in\mathbb{R}^+ \;, \tag{1.5}$$

where $X_{i:n}$, $1\le i\le n$, is the sample of ascending order statistics (o.s.) associated to our original sample $(X_1,X_2,\ldots,X_n)$. These statistics were introduced and studied under a second order framework by Dekkers *et al.* [5]. For more details on these statistics, and the way they may be used to build alternatives to the Hill estimator given by (1.5) and $\alpha=1$ (Hill [20]), see Gomes and Martins [12].

In this paper we are interested in the estimation of the second order parameter ρ in (1.3). The second order parameter ρ is of primordial importance in the adaptive choice of the best threshold to be considered in the estimation of the tail index γ, like may be seen in the papers by Hall and Welsh [19], Beirlant et al. ([1], [2]), Drees and Kaufmann [7], Danielsson et al. [4], Draisma et al. [6], Guillou and Hall [16], among others. Also, most of the recent research devised to improve the classical estimators of the tail index, try to reduce the main component of their asymptotic bias, which also depends strongly on ρ. So, an *a priori* estimation of ρ is needed, or at least desirable, for the adequate reduction of bias. Some of the papers in this area are the ones by Beirlant et al. [3], Feuerverger and Hall [8], Gomes and Martins ([12], [13]) and Gomes et al. [15].

All over the paper, and in order to simplify the proof of theoretical results, we shall only assume the situation $\rho,\beta<0$. We shall also assume everywhere that k is an intermediate rank, i.e.

$$k=k_n\to\infty,\quad k/n\to 0, \qquad \text{as}\;\; n\to\infty \;. \tag{1.6}$$

The starting point to obtain the class of estimators we are going to consider, is a well-known expansion of $M_n^{(\alpha)}(k)$ for any real $\alpha>0$, valid for intermediate k,

$$M_n^{(\alpha)}(k) \;=\; \gamma^{\alpha}\mu_\alpha^{(1)}+\gamma^{\alpha}\sigma_\alpha^{(1)}\frac{1}{\sqrt{k}}\,P_n^{(\alpha)}+\alpha\,\gamma^{\alpha-1}\mu_\alpha^{(2)}(\rho)\,A(Y_{n-k:n})+o_p(A(n/k)) \;,$$

where $P_n^{(\alpha)}$ is asymptotically a standard normal r.v. (cf. section 2 below where the notation is explained). The reasoning is then similar to the one in Gomes *et al.* [14]: first, for sequences $k=k(n)\to\infty$ with $\sqrt{k}\,A(n/k)=O(1)$, as $n\to\infty$, this gives an asymptotically normal estimator of a simple function of γ; but by taking sequences $k(n)$ of greater order than the previous ones, i.e. such that

$$\sqrt{k}\, A(n/k) \to \infty, \quad \text{as} \;\; n \to \infty \;, \tag{1.7}$$

we can emphasize other parts of this equation as follows.

First we get rid of the first term on the right by composing a linear combination of powers of two different $M_n^{(\alpha)}(k)$'s (i.e., for two different α), suitably normalized. We have here considered, for positive real τ and $\theta_1 \neq 1$,

$$\frac{\left(\frac{M_n^{(\alpha)}(k)}{\mu_\alpha^{(1)}}\right)^{\tau} - \left(\frac{M_n^{(\alpha\theta_1)}(k)}{\mu_{\alpha\theta_1}^{(1)}}\right)^{\tau/\theta_1}}{A(Y_{n-k:n})} \longrightarrow \alpha\,\tau\,\gamma^{\alpha\tau-1}\left(\frac{\mu_\alpha^{(2)}(\rho)}{\mu_\alpha^{(1)}} - \frac{\mu_{\alpha\theta_1}^{(2)}(\rho)}{\mu_{\alpha\theta_1}^{(1)}}\right) ,$$

which is a function of both parameters of the model, γ and ρ. We then get rid of the unknown $A(Y_{n-k:n})$ and of γ, by composing, for positive real values $\theta_1 \neq \theta_2$, both different from 1, a quotient of the type

$$T_n^{(\alpha,\theta_1,\theta_2,\tau)}(k) := \frac{\left(\frac{M_n^{(\alpha)}(k)}{\mu_\alpha^{(1)}}\right)^{\tau} - \left(\frac{M_n^{(\alpha\theta_1)}(k)}{\mu_{\alpha\theta_1}^{(1)}}\right)^{\tau/\theta_1}}{\left(\frac{M_n^{(\alpha\theta_1)}(k)}{\mu_{\alpha\theta_1}^{(1)}}\right)^{\tau/\theta_1} - \left(\frac{M_n^{(\alpha\theta_2)}(k)}{\mu_{\alpha\theta_2}^{(1)}}\right)^{\tau/\theta_2}} , \tag{1.8}$$

which, under conditions (1.6) and (1.7), converges in probability towards

$$t_{\alpha,\theta_1,\theta_2}(\rho) := \frac{\frac{\mu_\alpha^{(2)}(\rho)}{\mu_\alpha^{(1)}} - \frac{\mu_{\alpha\theta_1}^{(2)}(\rho)}{\mu_{\alpha\theta_1}^{(1)}}}{\frac{\mu_{\alpha\theta_1}^{(2)}(\rho)}{\mu_{\alpha\theta_1}^{(1)}} - \frac{\mu_{\alpha\theta_2}^{(2)}(\rho)}{\mu_{\alpha\theta_2}^{(1)}}} ,$$

and where the admissible values of the *tuning* parameters are $\alpha, \theta_1, \theta_2, \tau \in \mathbb{R}^+$, $\theta_1 \neq 1$, and $\theta_1 \neq \theta_2$.

We thus obtain a consistent estimator of a function of ρ which leads to a consistent estimator of ρ, as developed in section 2. In section 3 a somewhat more refined analysis again on the lines of Gomes *et al.* [14], using third order regular variation, brings the terms $\gamma^\alpha \sigma_\alpha^{(1)} \frac{1}{\sqrt{k}} P_n^{(\alpha)}$ back into play, and this will lead to a proof of the asymptotic normality of our estimators. We shall pay particular attention to the statistic obtained for $\alpha = 1$, $\theta_1 = 2$ and $\theta_2 = 3$, which involves only the first three moment statistics $M_n^{(i)}(k)$, $i = 1, 2, 3$, also handled in Draisma *et al.* [6], under a different general framework and for the estimation of

$\gamma \in \mathbb{R}$. We shall also advance with some indication on a possible way to choose the control parameters, in order to get estimators with stable sample paths and flat Mean Square Error (*MSE*) patterns, for large values of k, the number of top order statistics used in their construction. Finally, the practical validation of asymptotic results for small finite samples is done in section 4, by means of simulation techniques in Fréchet and Burr models.

2 – A class of semi-parametric estimators of the second order parameter

Let W denote an exponential r.v., with d.f. $F_W(x) = 1-\exp(-x)$, $x>0$, and, with the same notation as in Gomes *et al.* [14], let us put

$$\mu_\alpha^{(1)} := E[W^\alpha] = \Gamma(\alpha+1) \ , \tag{2.1}$$

$$\sigma_\alpha^{(1)} := \sqrt{\mathrm{Var}[W^\alpha]} = \sqrt{\Gamma(2\,\alpha+1)-\Gamma^2(\alpha+1)} \ , \tag{2.2}$$

$$\mu_\alpha^{(2)}(\rho) := E\left[W^{\alpha-1}\Big(\frac{e^{\rho W}-1}{\rho}\Big)\right] = \frac{\Gamma(\alpha)}{\rho}\,\frac{1-(1-\rho)^\alpha}{(1-\rho)^\alpha} \ , \tag{2.3}$$

$$\sigma_\alpha^{(2)}(\rho) := \sqrt{\mathrm{Var}\left[W^{\alpha-1}\Big(\frac{e^{\rho W}-1}{\rho}\Big)\right]} = \sqrt{\mu_{2\alpha}^{(3)}(\rho)-\Big(\mu_\alpha^{(2)}(\rho)\Big)^2} \ , \tag{2.4}$$

with

$$\begin{aligned}\mu_\alpha^{(3)}(\rho) &:= E\left[W^{\alpha-2}\Big(\frac{e^{\rho W}-1}{\rho}\Big)^2\right]\\ &= \begin{cases}\dfrac{1}{\rho^2}\ln\dfrac{(1-\rho)^2}{1-2\,\rho} & \text{if } \alpha=1,\\[2ex] \dfrac{\Gamma(\alpha)}{\rho^2(\alpha-1)}\left\{\dfrac{1}{(1-2\rho)^{\alpha-1}}-\dfrac{2}{(1-\rho)^{\alpha-1}}+1\right\} & \text{if } \alpha\neq 1 \ ,\end{cases}\end{aligned} \tag{2.5}$$

and

$$\begin{aligned}\mu_\alpha^{(4)}(\rho,\beta) &:= E\left[\frac{1}{\beta}\,W^{\alpha-1}\left(\frac{e^{(\rho+\beta)W-1}}{\rho+\beta}-\frac{e^{\rho W}-1}{\rho}\right)\right]\\ &= \frac{1}{\beta}\Big(\mu_\alpha^{(2)}(\rho+\beta)-\mu_\alpha^{(2)}(\rho)\Big) \ .\end{aligned} \tag{2.6}$$

Then, under the third order condition in (1.4), assuming that (1.6) holds, and using the same arguments as in Dekkers *et al.* [5], in lemma 2 of Draisma *et al.* [6] and more recently in Gomes *et al.* [14], we may write the distributional representation

$$\begin{aligned}
\left(\frac{M_n^{(\alpha\theta)}(k)}{\mu_{\alpha\theta}^{(1)}\gamma^{\alpha\theta}}\right)^{\tau/\theta} &= 1+\frac{\tau}{\theta}\,\overline{\sigma}_{\alpha\theta}^{(1)}\frac{1}{\sqrt{k}}\,P_n^{(\alpha\theta)} \\
&\quad+\frac{\alpha\tau}{\gamma}\,\overline{\mu}_{\alpha\theta}^{(2)}(\rho)\,A(n/k)+\frac{\alpha\tau}{\gamma}\,\overline{\sigma}_{\alpha\theta}^{(2)}(\rho)\,\frac{A(n/k)}{\sqrt{k}}\,\overline{P}_n^{(\alpha\theta)} \\
&\quad+\frac{\alpha\tau}{2\gamma^2}\,A^2(n/k)\Big((\alpha\theta-1)\,\overline{\mu}_{\alpha\theta}^{(3)}(\rho)+\alpha(\tau-\theta)\,(\overline{\mu}_{\alpha\theta}^{(2)}(\rho))^2\Big)\,(1+o_p(1)) \\
&\quad+\frac{\alpha\tau}{\gamma}\,\overline{\mu}_{\alpha\theta}^{(4)}(\rho,\beta)\,A(n/k)\,B(n/k)\,(1+o_p(1))\ ,
\end{aligned}\tag{2.7}$$

where $P_n^{(\alpha\theta)}$ and $\overline{P}_n^{(\alpha\theta)}$ are asymptotically standard Normal r.v.'s, and

$$\begin{aligned}
\overline{\mu}_\alpha^{(j)}(\rho) &= \frac{\mu_\alpha^{(j)}(\rho)}{\mu_\alpha^{(1)}},\quad j=2,3\,, \qquad \overline{\mu}_\alpha^{(4)}(\rho,\beta)=\frac{\mu_\alpha^{(4)}(\rho,\beta)}{\mu_\alpha^{(1)}}\ ,\\
\overline{\sigma}_\alpha^{(1)} &= \frac{\sigma_\alpha^{(1)}}{\mu_\alpha^{(1)}}\,, \qquad \overline{\sigma}_\alpha^{(2)}(\rho)=\frac{\sigma_\alpha^{(2)}(\rho)}{\mu_\alpha^{(1)}}\ .
\end{aligned}$$

If we now take the difference between two such expressions, we get a r.v. converging towards 0:

$$\begin{aligned}
D_n^{(\alpha,\theta_1,\theta_2,\tau)}(k) &:= \left(\frac{M_n^{(\alpha\theta_1)}(k)}{\mu_{\alpha\theta_1}^{(1)}\gamma^{\alpha\theta_1}}\right)^{\tau/\theta_1}-\left(\frac{M_n^{(\alpha\theta_2)}(k)}{\mu_{\alpha\theta_2}^{(1)}\gamma^{\alpha\theta_2}}\right)^{\tau/\theta_2} \\
&= \frac{\tau}{\sqrt{k}}\left(\frac{\overline{\sigma}_{\alpha\theta_1}^{(1)}}{\theta_1}\,P_n^{(\alpha\theta_1)}-\frac{\overline{\sigma}_{\alpha\theta_2}^{(1)}}{\theta_2}\,P_n^{(\alpha\theta_2)}\right) \\
&\quad+\frac{\alpha\tau}{\gamma}\,\Big(\overline{\mu}_{\alpha\theta_1}^{(2)}(\rho)-\overline{\mu}_{\alpha\theta_2}^{(2)}(\rho)\Big)\,A(n/k) \\
&\quad+\frac{\alpha\tau}{\gamma}\,\Big(\overline{\sigma}_{\alpha\theta_1}^{(2)}(\rho)\,\overline{P}_n^{(\alpha\theta_1)}-\overline{\sigma}_{\alpha\theta_2}^{(2)}(\rho)\,\overline{P}_n^{(\alpha\theta_2)}\Big)\,\frac{A(n/k)}{\sqrt{k}} \\
&\quad+\frac{\alpha\tau}{2\,\gamma^2}\Big((\alpha\theta_1-1)\,\overline{\mu}_{\alpha\theta_1}^{(3)}(\rho)+\alpha(\tau-\theta_1)\,(\overline{\mu}_{\alpha\theta_1}^{(2)}(\rho))^2 \\
&\quad-(\alpha\theta_2-1)\,\overline{\mu}_{\alpha\theta_2}^{(3)}(\rho)-\alpha(\tau-\theta_2)\,(\overline{\mu}_{\alpha\theta_2}^{(2)}(\rho))^2\Big)\,A^2(n/k)\,(1+o_p(1)) \\
&\quad+\frac{\alpha\tau}{\gamma}\,\Big(\overline{\mu}_{\alpha\theta_1}^{(4)}(\rho,\beta)-\overline{\mu}_{\alpha\theta_2}^{(4)}(\rho,\beta)\Big)\,A(n/k)\,B(n/k)\,(1+o_p(1))\ .
\end{aligned}\tag{2.8}$$

If we assume that (1.7) holds, the second term in the right hand side of (2.8) is the dominant one, and

$$
\begin{aligned}
\frac{D_n^{(\alpha,\theta_1,\theta_2,\tau)}(k)}{A(n/k)} = \; & \frac{\alpha\tau}{\gamma}\left(\overline{\mu}^{(2)}_{\alpha\theta_1}(\rho) - \overline{\mu}^{(2)}_{\alpha\theta_2}(\rho)\right) \\
& + \frac{\tau}{\sqrt{k}\,A(n/k)}\left(\frac{\overline{\sigma}^{(1)}_{\alpha\theta_1}}{\theta_1}\,P_n^{(\alpha\theta_1)} - \frac{\overline{\sigma}^{(1)}_{\alpha\theta_2}}{\theta_2}\,P_n^{(\alpha\theta_2)}\right) \\
& + \frac{\alpha\tau}{2\gamma^2}\Big((\alpha\theta_1-1)\,\overline{\mu}^{(3)}_{\alpha\theta_1}(\rho) + \alpha(\tau-\theta_1)\,(\overline{\mu}^{(2)}_{\alpha\theta_1}(\rho))^2 \\
& - (\alpha\theta_2-1)\,\overline{\mu}^{(3)}_{\alpha\theta_2}(\rho) - \alpha(\tau-\theta_2)\,(\overline{\mu}^{(2)}_{\alpha\theta_2}(\rho))^2\Big)\,A(n/k)\,(1+o_p(1)) \\
& + \frac{\alpha\tau}{\gamma}\left(\overline{\mu}^{(4)}_{\alpha\theta_1}(\rho,\beta) - \overline{\mu}^{(4)}_{\alpha\theta_2}(\rho,\beta)\right)B(n/k)\,(1+o_p(1)) \;.
\end{aligned}
\tag{2.9}
$$

Consequently, for $\theta_1 \neq \theta_2$, the statistic in (1.8), which may be written as

$$
T_n^{(\alpha,\theta_1,\theta_2,\tau)}(k) = \frac{D_n^{(\alpha,1,\theta_1,\tau)}(k)}{D_n^{(\alpha,\theta_1,\theta_2,\tau)}(k)} \;, \tag{2.10}
$$

converges in probability, as $n \to \infty$, towards

$$
t_{\alpha,\theta_1,\theta_2}(\rho) := \frac{\overline{\mu}^{(2)}_{\alpha}(\rho) - \overline{\mu}^{(2)}_{\alpha\theta_1}(\rho)}{\overline{\mu}^{(2)}_{\alpha\theta_1}(\rho) - \overline{\mu}^{(2)}_{\alpha\theta_2}(\rho)} = \frac{d_{\alpha,1,\theta_1}(\rho)}{d_{\alpha,\theta_1,\theta_2}(\rho)} \;, \tag{2.11}
$$

independently of τ, where

$$
d_{\alpha,\theta_1,\theta_2}(\rho) := \overline{\mu}^{(2)}_{\alpha\theta_1}(\rho) - \overline{\mu}^{(2)}_{\alpha\theta_2}(\rho) \;. \tag{2.12}
$$

Straightforward computations lead us to the expression

$$
t_{\alpha,\theta_1,\theta_2}(\rho) = \theta_2\,\frac{(\theta_1-1)\,(1-\rho)^{\alpha\theta_2} - \theta_1(1-\rho)^{\alpha(\theta_2-1)} + (1-\rho)^{\alpha(\theta_2-\theta_1)}}{(\theta_2-\theta_1)\,(1-\rho)^{\alpha\theta_2} - \theta_2(1-\rho)^{\alpha(\theta_2-\theta_1)} + \theta_1} \;. \tag{2.13}
$$

We have

$$
\lim_{\rho\to 0^-} t_{\alpha,\theta_1,\theta_2}(\rho) = \frac{\theta_1-1}{\theta_2-\theta_1} \;; \qquad \lim_{\rho\to -\infty} t_{\alpha,\theta_1,\theta_2}(\rho) = \frac{\theta_2(\theta_1-1)}{\theta_2-\theta_1} \;, \tag{2.14}
$$

and for negative values of ρ and $\alpha>0$, $t_{\alpha;\theta_1,\theta_2}(\rho)$ is always a decreasing (increasing) function of ρ, provided that $1<\theta_1<\theta_2$ $(\theta_1>\theta_2>1)$.

We have thus got a consistent estimator of a function of ρ, which needs to be inverted, i.e., the estimator of ρ to be studied in the following section is

$$(2.15)\qquad \widehat{\rho}_{n|T}^{(\alpha,\theta 1,\theta_2,\tau)}(k) := t^{\leftarrow}_{\alpha,\theta_1,\theta_2}\big(T_n^{(\alpha,\theta_1,\theta_2,\tau)}(k)\big), \quad \text{provided that}\quad \frac{\theta_1-1}{|\theta_2-\theta_1|} \le \Big|T_n^{(\alpha,\theta_1,\theta_2,\tau)}(k)\Big| < \frac{\theta_2(\theta_1-1)}{|\theta_2-\theta_1|}\ ,$$

in order to get the right sign for the ρ-estimator.

The easiest situation is the one associated to values (θ_1,θ_2) such that $\theta_2-\theta_1=1$ and $\theta_2-1=2$ (look at expression (2.11)), i.e. to the values $(\theta_1=2,\ \theta_2=3)$, for which we get

$$(2.16)\qquad t_\alpha(\rho) \;=\; t_{\alpha,2,3}(\rho) \;=\; 3(1-\rho)^\alpha\,\frac{(1-\rho)^{2\alpha}-2(1-\rho)^\alpha+1}{(1-\rho)^{3\alpha}-3(1-\rho)^\alpha+2} \;=\; \frac{3(1-\rho)^\alpha}{(1-\rho)^\alpha+2}\ ,$$

which, for any $\alpha>0$, must be between 1 and 3 to provide, by inversion, negative values of ρ. We then get an explicit analytic expression for the estimator of ρ. More specifically, we get

$$(2.17)\qquad \widehat{\rho}_{n|T}^{(\alpha,2,3,\tau)}(k) := 1-\left(\frac{2\,T_n^{(\alpha,2,3,\tau)}(k)}{3-T_n^{(\alpha,2,3,\tau)}(k)}\right)^{1/\alpha}\ , \quad \text{provided that}\quad 1\le T_n^{(\alpha,2,3,\tau)}(k)<3\ .$$

For the particular case $\alpha=1$, we have

$$(2.18)\qquad \widehat{\rho}_{n|T}^{(1,2,3,\tau)}(k) := \frac{3\Big(T_n^{(1,2,3,\tau)}(k)-1\Big)}{T_n^{(1,2,3,\tau)}(k)-3}\ , \quad \text{provided that}\quad 1\le T_n^{(1,2,3,\tau)}(k)<3\ .$$

We have thus proved the following

Theorem 2.1. *Suppose that the second order condition (1.3) holds, with $\rho<0$. For sequences of integers $k=k(n)$ satisfying $k(n)=o(n)$ and $\sqrt{k}\,A(n/k)\to\infty$, as $n\to\infty$, we have*

$$\lim_{n\to\infty}\widehat{\rho}_{n|T}^{(\alpha,\theta_1,\theta_2,\tau)}(k)=\rho$$

in probability for any $\alpha,\tau>0\in\mathbb{R}^+$, and $\theta_1,\theta_2\in\mathbb{R}^+\backslash\{1\}$, $\theta_1\neq\theta_2$, with $\widehat{\rho}_{n|T}^{(\alpha,\theta_1,\theta_2,\tau)}(k)$ defined in (2.15), and with an explicit analytic expression given by (2.17) for $(\theta_1,\theta_2)=(2,3)$.

3 – The asymptotic normality of the estimators of the second order parameter

From (2.9), and under the validity of (1.7),

$$\begin{aligned}\frac{D_n^{(\alpha,\theta_1,\theta_2,\tau)}(k)}{\alpha\,\tau\,\gamma^{-1}A(n/k)} \;&=\; d_{\alpha,\theta_1,\theta_2}(\rho)+\frac{\gamma}{\alpha\,\sqrt{k}\,A(n/k)}\,W_n^{(\alpha,\theta_1,\theta_2)}\\ &\quad+\Big\{u_{\alpha,\theta_1,\theta_2,\tau}(\rho)\,A(n/k)+v_{\alpha,\theta_1,\theta_2}(\rho,\beta)\,B(n/k)\Big\}\,(1+o_p(1))\;, \end{aligned}\tag{3.1}$$

where

$$W_n^{(\alpha,\theta_1,\theta_2)} \;:=\; \frac{\overline{\sigma}^{(1)}_{\alpha\theta_1}}{\theta_1}\,P_n^{(\alpha\theta_1)}-\frac{\overline{\sigma}^{(1)}_{\alpha\theta_2}}{\theta_2}\,P_n^{(\alpha\theta_2)}\;,\tag{3.2}$$

$$\begin{aligned}u_{\alpha,\theta_1,\theta_2,\tau}(\rho) \;:=\; \frac{1}{2\,\gamma}\Big\{&(\alpha\theta_1-1)\,\overline{\mu}^{(3)}_{\alpha\theta_1}(\rho)+\alpha(\tau-\theta_1)\,(\overline{\mu}^{(2)}_{\alpha\theta_1}(\rho))^2\\ &-(\alpha\theta_2-1)\,\overline{\mu}^{(3)}_{\alpha\theta_2}(\rho)-\alpha(\tau-\theta_2)\,(\overline{\mu}^{(2)}_{\alpha\theta_2}(\rho))^2\Big\}\end{aligned}\tag{3.3}$$

and

$$v_{\alpha,\theta_1,\theta_2}(\rho,\beta) \;:=\; \overline{\mu}^{(4)}_{\alpha\theta_1}(\rho,\beta)-\overline{\mu}^{(4)}_{\alpha\theta_2}(\rho,\beta)\;.\tag{3.4}$$

Then, since $T_n^{(\alpha,\theta_1,\theta_2,\tau)}(k)=D_n^{(\alpha,1,\theta_1,\tau)}(k)/D_n^{(\alpha,\theta_1,\theta_2,\tau)}(k)$, we have whenever

$$k=k_n\to\infty\,,\qquad k/n\to 0\,,\qquad \sqrt{k}\,A(n/k)\to\infty,\quad \text{as }\; n\to\infty\;,\tag{3.5}$$

$$\begin{aligned}T_n^{(\alpha,\theta_1,\theta_2,\tau)}(k) \;&=\; t_{\alpha,\theta_1,\theta_2}(\rho)\\ &\quad+\frac{\gamma}{\alpha\,\sqrt{k}\,A(n/k)}\,\frac{1}{d_{\alpha,\theta_1,\theta_2}(\rho)}\,\Big\{W_n^{(\alpha,1,\theta_1)}-t_{\alpha,\theta_1,\theta_2}(\rho)W_n^{(\alpha,\theta_1,\theta_2)}\Big\}\\ &\quad+\Bigg\{\frac{u_{\alpha,1,\theta_1,\tau}(\rho)-t_{\alpha,\theta_1,\theta_2}(\rho)\,u_{\alpha,\theta_1,\theta_2,\tau}(\rho)}{d_{\alpha,\theta_1,\theta_2}(\rho)}\,A(n/k)\\ &\quad+\frac{v_{\alpha,1,\theta_1}(\rho,\beta)-t_{\alpha,\theta_1,\theta_2}(\rho)\,v_{\alpha,\theta_1,\theta_2}(\rho,\beta)}{d_{\alpha,\theta_1,\theta_2}(\rho)}\,B(n/k)\Bigg\}\,(1+o_p(1))\;.\end{aligned}\tag{3.6}$$

From the asymptotic covariance between $\overline{\sigma}^{(1)}_{\alpha\theta_1}P_n^{(\alpha\theta_1)}$ and $\overline{\sigma}^{(1)}_{\alpha\theta_2}P_n^{(\alpha\theta_2)}$ (see Gomes and Martins [12]), given by

$$\frac{\alpha(\theta_1+\theta_2)\;\Gamma\Big(\alpha(\theta_1+\theta_2)\Big)}{\alpha^2\,\theta_1\,\theta_2\,\Gamma(\alpha\theta_1)\,\Gamma(\alpha\theta_2)}-1\;,$$

we easily derive the asymptotic covariance between $W_n^{(\alpha,1,\theta_1)}$ and $W_n^{(\alpha,\theta_1,\theta_2)}$, given by

$$\begin{aligned}\sigma_{W|\alpha,1,\theta_1,\theta_2} \;=\; &\frac{1}{\alpha}\left(\frac{(\theta_1+1)\,\Gamma\Big(\alpha(\theta_1+1)\Big)}{\theta_1^2\,\Gamma(\alpha)\,\Gamma(\alpha\theta_1)} - \frac{(\theta_2+1)\,\Gamma\Big(\alpha(\theta_2+1)\Big)}{\theta_2^2\,\Gamma(\alpha)\,\Gamma(\alpha\theta_2)} - \frac{2\,\Gamma(2\,\alpha\theta_1)}{\theta_1^3\,\Gamma^2(\alpha\theta_1)}\right.\\ &\left.+\frac{(\theta_1+\theta_2)\,\Gamma\Big(\alpha(\theta_1+\theta_2)\Big)}{\theta_1^2\,\theta_2^2\,\Gamma(\alpha\theta_1)\,\Gamma(\alpha\theta_2)}\right) - \left(1-\frac{1}{\theta_1}\right)\left(\frac{1}{\theta_1}-\frac{1}{\theta_2}\right) . \end{aligned} \tag{3.7}$$

The asymptotic variance of $W_n^{(\alpha,\theta_1,\theta_2)}$ is

$$\begin{aligned}\sigma^2_{W|\alpha,\theta_1,\theta_2} \;=\; &\frac{2}{\alpha}\left(\frac{\Gamma(2\,\alpha\theta_1)}{\theta_1^3\,\Gamma^2(\alpha\theta_1)} + \frac{\Gamma(2\,\alpha\theta_2)}{\theta_2^3\,\Gamma^2(\alpha\theta_2)} - \frac{(\theta_1+\theta_2)\,\Gamma\Big(\alpha(\theta_1+\theta_2)\Big)}{\theta_1^2\,\theta_2^2\,\Gamma(\alpha\theta_1)\,\Gamma(\alpha\theta_2)}\right)\\ &-\left(\frac{1}{\theta_1}-\frac{1}{\theta_2}\right)^2 . \end{aligned} \tag{3.8}$$

Consequently, if apart from the previous conditions in (3.5), we also have

$$\lim_{n\to\infty} \sqrt{k}\,A^2(n/k) = 0 \quad \text{and} \quad \lim_{n\to\infty} \sqrt{k}\,A(n/k)\,B(n/k) = 0 \;, \tag{3.9}$$

there is a null asymptotic bias, and

$$\sqrt{k}\,A(n/k)\left(T_n^{(\alpha,\theta_1,\theta_2,\tau)}(k) - t_{\alpha,\theta_1,\theta_2}\right) \xrightarrow{d} Z_\alpha \;, \tag{3.10}$$

where Z_α is a Normal r.v. with null mean and variance given by

$$\sigma^2_{T|\alpha,\theta_1,\theta_2} = \frac{\gamma^2\left(\sigma^2_{W|\alpha,1,\theta_1} + t^2_{\alpha,\theta_1,\theta_2}(\rho)\,\sigma^2_{W|\alpha,\theta_1,\theta_2} - 2\,t_{\alpha,\theta_1,\theta_2}(\rho)\,\sigma_{W|\alpha,1,\theta_1,\theta_2}\right)}{\alpha^2\,d^2_{\alpha,\theta_1,\theta_2}(\rho)} \,, \tag{3.11}$$

with $t_{\alpha,\theta_1,\theta_2}$, $d_{\alpha,\theta_1,\theta_2}(\rho)$, $\sigma_{W|\alpha,1,\theta_1,\theta_2}$ and $\sigma^2_{W|\alpha,\theta_1,\theta_2}$ given in (2.11), (2.12), (3.7) and (3.8), respectively.

In the more general case

$$\lim_{n\to\infty} \sqrt{k}\,A^2(n/k) = \lambda_1 \quad \text{and} \quad \lim_{n\to\infty} \sqrt{k}\,A(n/k)\,B(n/k) = \lambda_2 \;, \tag{3.12}$$

we have to take into account a non-null asymptotic bias, i.e.

$$\sqrt{k}\,A(n/k)\left\{T_n^{(\alpha,\theta_1,\theta_2,\tau)}(k) - t_{\alpha,\theta_1,\theta_2}\right\},$$

is asymptotically Normal with mean value equal to

$$\mu_{T|\alpha,\theta_1,\theta_2,\tau} = \lambda_1\,u_{T|\alpha,\theta_1,\theta_2,\tau} + \lambda_2\,v_{T|\alpha,\theta_1,\theta_2} \;, \tag{3.13}$$

where

$$u_{T|\alpha,\theta_1,\theta_2,\tau} \equiv u_{T|\alpha,\theta_1,\theta_2,\tau}(\rho) = \frac{u_{\alpha,1,\theta_1,\tau}(\rho) - t_{\alpha,\theta_1,\theta_2}(\rho)\, u_{\alpha,\theta_1,\theta_2,\tau}(\rho)}{d_{\alpha,\theta_1,\theta_2}(\rho)}\ , \tag{3.14}$$

$$v_{T|\alpha,\theta_1,\theta_2} \equiv v_{T|\alpha,\theta_1,\theta_2}(\rho,\beta) = \frac{v_{\alpha,1,\theta_1}(\rho,\beta) - t_{\alpha,\theta_1,\theta_2}(\rho)\, v_{\alpha,\theta_1,\theta_2}(\rho,\beta)}{d_{\alpha,\theta_1,\theta_2}(\rho)} \tag{3.15}$$

and variance given by (3.11), as well as before.

Figure 1 illustrates, for $\theta_1 = 2$ and $\theta_2 = 3$, the behaviour of $\sigma_{T|\alpha,\theta_1,\theta_2}/\gamma$, $\gamma\, u_{T|\alpha,\theta_1,\theta_2,\tau}(\rho)$ and $v_{T|\alpha,\theta_1,\theta_2}(\rho,\beta)$ as functions of α, for $\tau = -\rho = -\beta = 1$.

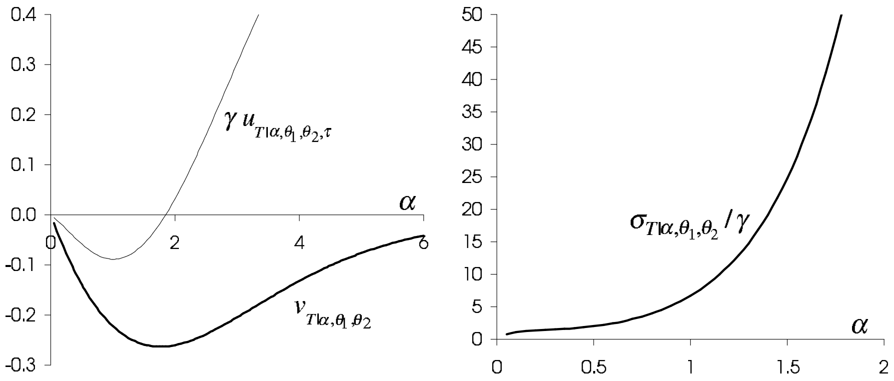

Fig. 1: $\sigma_{T|\alpha,\theta_1,\theta_2}/\gamma$, $\gamma u_{T|\alpha,\theta_1,\theta_2,\tau}(\rho)$ and $v_{T|\alpha,\theta_1,\theta_2}(\rho,\beta)$ as functions of α, for $\theta_1 = 2$, $\theta_2 = 3$, $\tau = 1$ and assuming $\rho = \beta = -1$.

Then, it follows that for the ρ-estimator, $\widehat{\rho}_{n|T}^{(\alpha,\theta_1,\theta_2,\tau)}(k)$, defined in (2.15), we have that, under (3.12),

$$\sqrt{k}\, A(n/k) \left\{\widehat{\rho}_{n|T}^{(\alpha,\theta_1,\theta_2,\tau)}(k) - \rho\right\},$$

is asymptotically Normal with mean value equal to

$$\mu_{\rho|T}^{(\alpha,\theta_1,\theta_2,\tau)} = \mu_{T|\alpha,\theta_1,\theta_2,\tau}/t'_{\alpha,\theta_1,\theta_2}(\rho) =: \lambda_1\, u_{\rho|T}^{(\alpha,\theta_1,\theta_2,\tau)} + \lambda_2\, v_{\rho|T}^{(\alpha,\theta_1,\theta_2)}\ , \tag{3.16}$$

and with variance given by

$$\sigma^2_{\rho|T,\alpha,\theta_1,\theta_2} = \left(\frac{\sigma_{T|\alpha,\theta_1,\theta_2}}{t'_{\alpha,\theta_1,\theta_2}(\rho)}\right)^2, \tag{3.17}$$

where $t'_{\alpha,\theta_1,\theta_2}(\rho)$ is such that

$$\begin{aligned} t'_{\alpha,\theta_1,\theta_2}(\rho)\,(1-\rho)\,\Big((\theta_2-\theta_1)\,(1-\rho)^{\alpha\theta_2} - \theta_2(1-\rho)^{\alpha(\theta_2-\theta_1)} + \theta_1\Big)^2 \;&= \\ =\; \alpha\,\theta_1\theta_2\,\Big\{\theta_1(\theta_2-1)\,(1-\rho)^{\alpha(\theta_2-1)}\,&\Big(1+(1-\rho)^{\alpha(\theta_2-\theta_1+1)}\Big) \\ -\,(\theta_2-\theta_1)\,(1-\rho)^{\alpha(\theta_2-\theta_1)}\,&\Big(1+(1-\rho)^{\alpha(\theta_2+\theta_1-1)}\Big) \\ -\,\theta_2(\theta_1-1)\,(1-\rho)^{\alpha\theta_2}\,&\Big(1+(1-\rho)^{\alpha(\theta_2-\theta_1-1)}\Big)\Big\}\;. \end{aligned}$$

For the particular, but interesting case $\alpha=1$, $\theta_1=2$, $\theta_2=3$ and under the same conditions as before, we have that, with $\widehat{\rho}^{(1,2,3,\tau)}_{n|T}(k)$ given in (2.18),

$$\sqrt{k}\,A(n/k)\,\Big\{\widehat{\rho}^{(1,2,3,\tau)}_{n|T}(k)-\rho\Big\}$$

is asymptotically Normal, with variance given by

$$\sigma^2_{\rho|T,1,2,3} \;=\; \left(\frac{\gamma(1-\rho)^3}{\rho}\right)^2\Big(2\,\rho^2-2\,\rho+1\Big)\,. \tag{3.18}$$

The asymptotic bias is either null or given by (3.16) according as (3.9) or (3.12) hold, respectively.

In Figure 2 we present asymptotic characteristics of $\widehat{\rho}^{(\alpha,\theta_1,\theta_2,\tau)}_{n|T}(k)$ for the same particular values of the control parameters, namely $\theta_1=2$, $\theta_2=3$, and $\tau=-\rho=-\beta=1$.

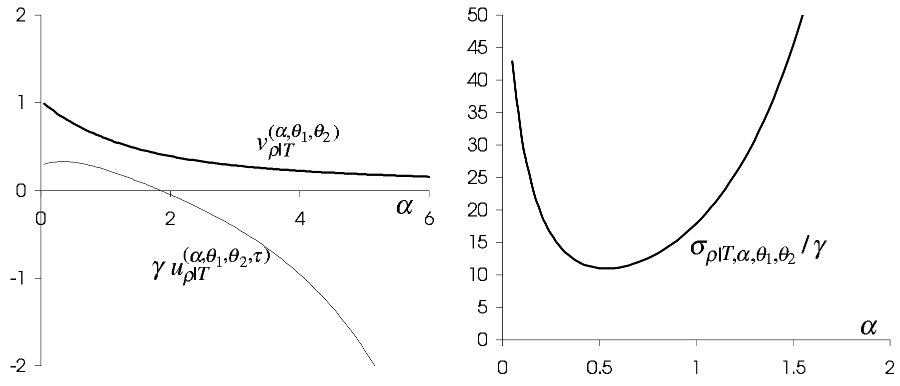

Fig. 2: $\sigma_{\rho|T(\alpha,\theta_1,\theta_2)}/\gamma$, $\gamma u^{(\alpha,\theta_1,\theta_2,\tau)}_{\rho|T}(\rho)$, and $v^{(\alpha,\theta_1,\theta_2)}_{\rho|T}(\rho,\beta)$ as functions of α, for $\theta_1=2$, $\theta_2=3$, $\tau=1$ and assuming $\rho=\beta=-1$.

We have thus proved the following

Theorem 3.1. *Suppose that third order condition (1.4) holds with $\rho < 0$. For intermediate sequences of integers $k = k(n)$ satisfying*

$$\begin{aligned} &\lim_{n\to\infty} \sqrt{k}\, A(n/k) = \infty\;, \\ (3.19)\qquad &\lim_{n\to\infty} \sqrt{k}\, A^2(n/k) = \lambda_1, \quad \textit{finite}\;, \\ &\lim_{n\to\infty} \sqrt{k}\, A(n/k) B(n/k) = \lambda_2, \quad \textit{finite}\;, \end{aligned}$$

we have that for every positive real numbers $\theta_1 \neq \theta_2$, both different from 1, and $\alpha, \tau > 0$

$$\sqrt{k}\, A(n/k) \left\{ \widehat{\rho}_{n|T}^{(\alpha,\theta_1,\theta_2,\tau)}(k) - \rho \right\} \tag{3.20}$$

is asymptotically normal with mean given in (3.16) and with variance given in (3.17). Note that the variance does not depend on λ_1 or λ_2.

Remarks:

1. We again enhance the fact that, for any $\tau > 0$, the statistic $T_n^{(\alpha,\theta_1,\theta_2,\tau)}(k)$ in (1.8) converges in probability to the same limit. This leads to an adequate control management about the parameter τ, which can be useful in the study of the exact distributional patterns of this class of estimators.

2. If we let $\tau \to 0$, we get the statistic

$$T_n^{(\alpha,\theta_1,\theta_2,0)}(k) := \frac{\ln\!\left(\frac{M_n^{\alpha}(k)}{\mu_\alpha^{(1)}}\right) - \ln\!\left(\frac{M_n^{\alpha\theta_1}(k)}{\mu_{\alpha\theta_1}^{(1)}}\right)\Big/\theta_1}{\ln\!\left(\frac{M_n^{\alpha\theta_1}(k)}{\mu_{\alpha\theta_1}^{(1)}}\right)\Big/\theta_1 - \ln\!\left(\frac{M_n^{\alpha\theta_2}(k)}{\mu_{\alpha\theta_2}^{(1)}}\right)\Big/\theta_2}\;, \tag{3.21}$$

and Theorems 2.1 and 3.1 hold true, with τ replaced by 0 everywhere. □

4 – An illustration of distributional and sample path properties of the estimators

We shall present in Figures 3 and 4 the simulated mean values and root mean square errors of the estimators $\widehat{\rho}_{n|T}^{(1,2,3,\tau)}(k)$ in (2.18), $\tau = 0, 0.5, 1, 2$, for a sample of size $n = 5000$ from a Fréchet model, $F(x) = \exp(-x^{-1/\gamma})$, $x \geq 0$, with $\gamma = 1$ ($\rho = -1$) and a Burr model, $F(x) = 1 - (1 + x^{-\rho/\gamma})^{1/\rho}$, $x \geq 0$, also with $\rho = -1$ and $\gamma = 1$, respectively. Simulations have been carried out with 5000 runs.

Notice that we consider, in both pictures, values of $k \geq 4000$. For smaller values of k we get high volatility of the estimators characteristics, and admissibility probabilities, associated to (2.18), slightly smaller than one. Those probabilities are equal to one whenever $k \geq 4216$ in the Fréchet model, and $k \geq 2986$ in the Burr model.

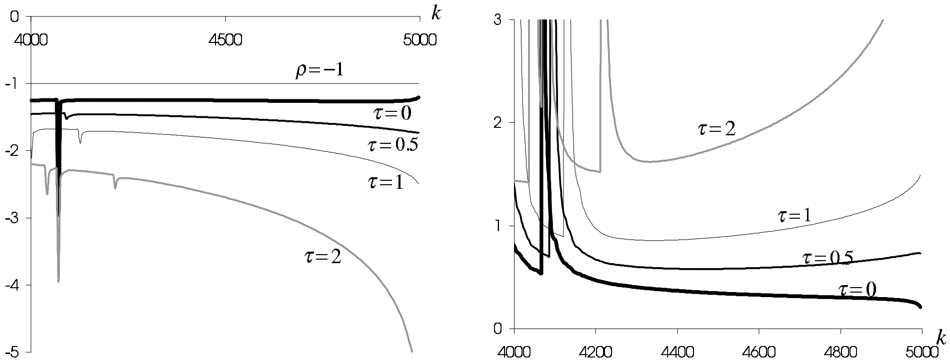

Fig. 3: Simulated mean values (*left*) and mean square errors (*right*) of $\widehat{\rho}_{n|T}^{(1,2,3,\tau)}(k)$ in (2.18), for $\tau = 0, 0.5, 1, 2$, and for a sample of size $n = 5000$ from a Fréchet(1) model ($\rho = -1$).

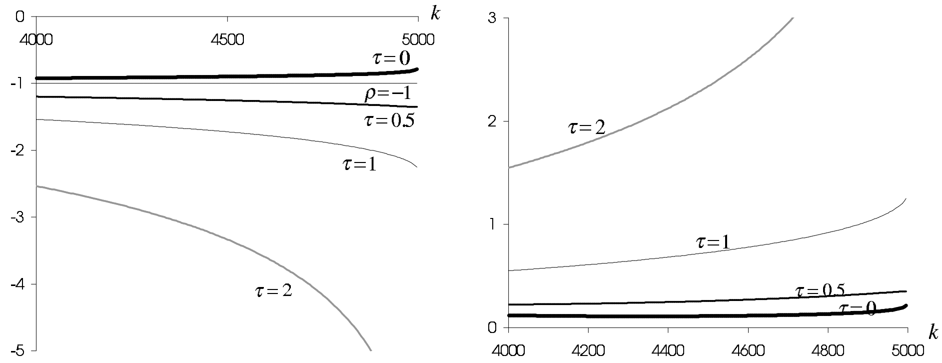

Fig. 4: Simulated mean values (*left*) and mean square errors (*right*) of $\widehat{\rho}_{n|T}^{(1,2,3,\tau)}(k)$ in (2.18), for $\tau = 0, 0.5, 1, 2$, and for a sample of size $n = 5000$ from a Burr model with $\rho = -1$ and $\gamma = 1$.

In Figure 5 we picture, for $k \geq 2000$ the simulated mean values and root mean square errors of the same estimators, for the same sample size but for a Burr model with $\gamma = 1$ and $\rho = -0.5$. Here we get admissibility probabilities equal to one for $k \geq 1681$.

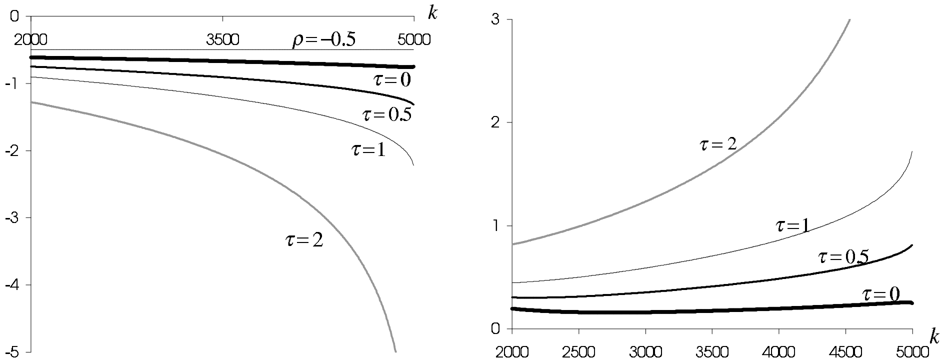

Fig. 5 : Simulated mean values (*left*) and mean square errors (*right*) of $\widehat{\rho}^{(1,2,3,\tau)}_{n|T}(k)$ in (2.18), for $\tau = 0, 0.5, 1, 2$, and for a sample of size $n = 5000$ from a Burr model with $\rho = -0.5$ and $\gamma = 1$.

The previous pictures seem to suggest the choice $\tau = 0$ for the *tuning* parameter τ. Notice however that such a choice is *not always* the best one, as may be seen from Figure 6, which is equivalent to Figure 5, but for a Burr model with $\gamma = 1$ and $\rho = -2$. This graph is represented for $k \geq 4000$, since the admissibility probabilities of the estimators under play are all equal to one provided that $k \geq 4301$. However, since values of ρ with such magnitude are not common in practice, the choice $\tau = 0$ seems to be a sensible one.

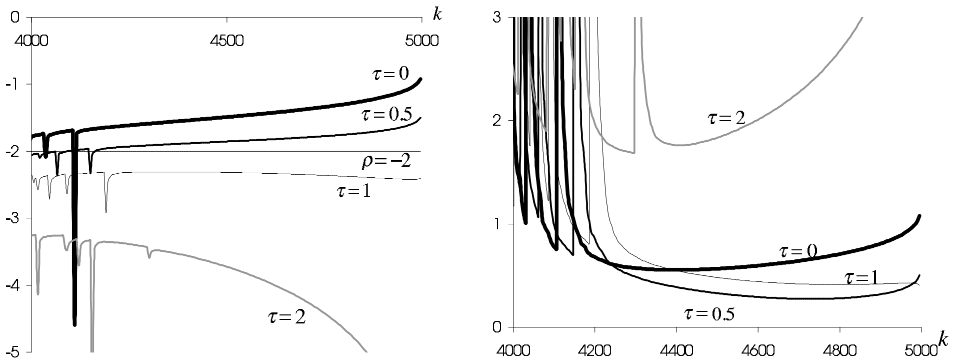

Fig. 6 : Simulated mean values (*left*) and mean square errors (*right*) of $\widehat{\rho}^{(1,2,3,\tau)}_{n|T}(k)$ in (2.18), for $\tau = 0, 0.5, 1, 2$, and for a sample of size $n = 5000$ from a Burr model with $\rho = -2$ and $\gamma = 1$.

We anyway advise the plot of sample paths of $\widehat{\rho}_{n|T}^{(1,2,3,\tau)}(k)$, for a few values of τ, like for instance the ones mentioned before, $\tau = 0, 0.5, 1$ and 2, and the choice of the value of τ which provides the highest stability in the region of large k values for which we get admissible estimates of ρ.

Finally, in Figure 7, we picture, for the values of k which provide admissible estimates of ρ, a sample path of the same estimators, for the same sample size $n = 5000$ and for two generated samples, one from a Fréchet model with $\gamma = 1$ ($\rho = -1$) and another from a Burr model with $\rho = -0.5$ and $\gamma = 1$.

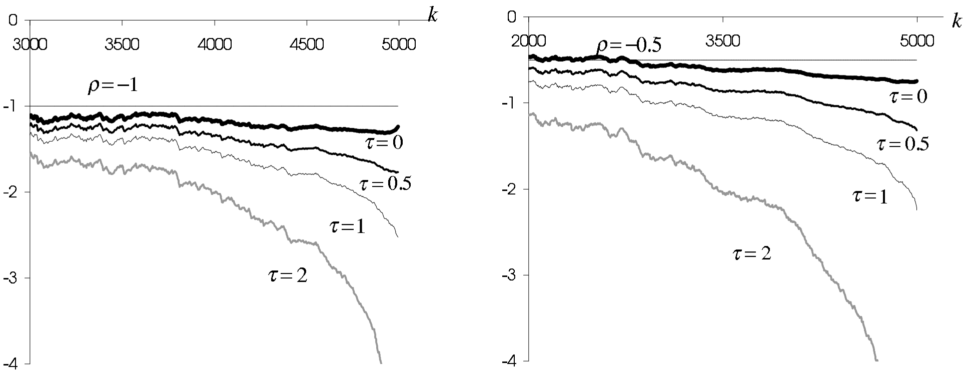

Fig. 7: Sample path of the estimators $\widehat{\rho}_{n|T}^{(1,2,3,\tau)}(k)$ in (2.18), $\tau = 0, 0.5, 1, 2$, for one sample of size $n = 5000$ from a Fréchet model with $\gamma = 1$ ($\rho = -1$) (*left*) and another sample from a Burr model with $\rho = -0.5$ and $\gamma = 1$ (*right*).

We have also carried out a large scale simulation, based on a multi-sample simulation of size 5000×10 (10 replicates with 5000 runs each), for the estimators associated to the control parameters $(\alpha, \theta_1, \theta_2) = (1, 2, 3)$ (which provide the explicit expression in (2.18)), and for $\tau = 0, 0.5, 1, 2$ and 6. The estimators $\widehat{\rho}_{n|T}^{(1,2,3,\tau)}(k)$, $\tau = 0, 0.5, 1, 2$ and 6, will be here denoted by $\widehat{\rho}_n^{(j)}(k)$, $1 \le j \le 5$, respectively.

In Tables 1 and 2 we present for a Fréchet parent with $\gamma = 1$ (for which $\rho = -1$), the simulated distributional properties of the five estimators computed at the optimal level, i.e. of $\widehat{\rho}_{n,0}^{(j)} := \widehat{\rho}_n^{(j)}(\widehat{k}_0^{(j)}(n))$, $\widehat{k}_0^{(j)}(n) := \arg\min_k MSE[\widehat{\rho}_n(j)]$, $1 \le j \le 5$. The standard error associated to each simulated characteristic is placed close to it and between parenthesis.

Table 1: Simulated mean values of $\widehat{\rho}^{(j)}_{n,0}$, $1 \le j \le 5$, for a Fréchet parent.

n	$E[\widehat{\rho}^{(1)}_{n,0}]$	$E[\widehat{\rho}^{(2)}_{n,0}]$	$E[\widehat{\rho}^{(3)}_{n,0}]$	$E[\widehat{\rho}^{(4)}_{n,0}]$	$E[\widehat{\rho}^{(5)}_{n,0}]$
100	-1.4287 (.0198)	-1.7371 (.0453)	-2.1752 (.0933)	-1.9417 (.0454)	-7.8318 (.1513)
200	-1.3118 (.0083)	-1.7966 (.0224)	-2.4823 (.0153)	-2.1936 (.2305)	-4.4530 (.0765)
500	-1.2237 (.0021)	-1.7372 (.0009)	-2.4216 (.0088)	-4.1152 (.1625)	-2.8365 (.0496)
1000	-1.1989 (.0006)	-1.7268 (.0007)	-2.1727 (.0199)	-3.4344 (.0673)	-2.5314 (.0673)
2000	-1.1812 (.0009)	-1.6226 (.0115)	-1.8841 (.0190)	-2.4138 (.1431)	-1.8778 (.0691)
5000	-1.1607 (.0006)	-1.4911 (.0047)	-1.6978 (.0102)	-2.1341 (.0252)	-1.6361 (.0575)
10000	-1.1467 (.0003)	-1.4184 (.0042)	-1.5763 (.0064)	-1.9418 (.0149)	-1.5647 (.0426)
20000	-1.1331 (.0002)	-1.3560 (.0038)	-1.4815 (.0057)	-1.7397 (.0131)	-2.9302 (.1668)

Table 2: Simulated Mean Square Errors of $\widehat{\rho}^{(j)}_{n,0}$, $1 \le j \le 5$, for a Fréchet parent.

n	$MSE[\widehat{\rho}^{(1)}_{n,0}]$	$MSE[\widehat{\rho}^{(2)}_{n,0}]$	$MSE[\widehat{\rho}^{(3)}_{n,0}]$	$MSE[\widehat{\rho}^{(4)}_{n,0}]$	$MSE[\widehat{\rho}^{(5)}_{n,0}]$
100	17.9556 (2.2790)	19.1709 (2.5617)	26.1292 (2.4915)	37.2158 (2.1936)	1256.9121 (98.8912)
200	2.2329 (.2858)	7.0139 (1.4370)	12.1975 (1.8059)	34.7150 (1.8443)	243.8227 (21.1588)
500	0.1247 (.0054)	0.6818 (.0142)	2.3807 (.0288)	17.1601 (.6100)	69.8044 (3.4082)
1000	0.0594 (.0004)	0.5597 (.0012)	1.7576 (.0233)	8.3825 (.1762)	71.8353 (5.1633)
2000	0.0409 (.0004)	0.4844 (.0110)	1.1234 (.0373)	3.7547 (.1133)	12.4722 (1.0556)
5000	0.0287 (.0002)	0.3283 (.0048)	0.6721 (.0097)	1.9210 (.0355)	9.5079 (.7900)
10000	0.0230 (.0001)	0.2414 (.0028)	0.4678 (.0075)	1.2322 (.0291)	8.8681 (.4593)
20000	0.0186 (.0001)	0.1803 (.0025)	0.3315 (.0054)	0.7822 (.0175)	7.5882 (.2331)

In the Tables 3 and 4 we present the distributional behaviour of the above mentioned estimators for a Burr model with $\gamma = 1$, and for values $\rho = -2, -1, -0.5, -0.25$.

Some final remarks:

1. The choice of the tuning parameters $(\theta_1,\ \theta_2)$ seems to be uncontroversial: the pair $(\theta_1,\ \theta_2) = (2, 3)$ seems to be the most convenient. The tuning parameter α can be any real positive number, but the value $\alpha = 1$ is the easiest choice, mainly due to the fact that the computation of an estimate for a given (perhaps large) data set is much less time-consuming whenever we work with $M_n^{(\alpha)}$, for positive integer α. The choice of τ is more open, and depends obviously on the model. This gives a higher flexibility to the choice of the adequate estimator of ρ, within the class of estimators herewith studied.

2. Indeed, the most interesting feature of this class of estimators is the fact that the consideration of the sample paths $\widehat{\rho}_n^{(1,2,3,\tau)}(k)$, as a function of k, for large k, and for a few values of τ, like $\tau = 0, 0.5, 1$ and 2, enables us to identify easily the most stable sample path, and to get an estimate of ρ.

Table 3: Simulated Mean Values of $\widehat{\rho}^{(j)}_{n,0}$, $1 \le j \le 5$, for a Burr parent.

n	$E[\widehat{\rho}^{(1)}_{n,0}]$	$E[\widehat{\rho}^{(2)}_{n,0}]$	$E[\widehat{\rho}^{(3)}_{n,0}]$	$E[\widehat{\rho}^{(4)}_{n,0}]$	$E[\widehat{\rho}^{(5)}_{n,0}]$
$\rho = -0.25$					
100	-0.5900 (.0018)	-0.8222 (.0031)	-1.0847 (.0078)	-1.7716 (.0255)	-21.4311 (2.7554)
200	-0.5648 (.0012)	-0.7612 (.0027)	-0.9793 (.0088)	-1.5626 (.0156)	-14.2221 (.1938)
500	-0.5330 (.0014)	-0.6929 (.0027)	-0.8635 (.0051)	-1.3154 (.0116)	-8.0533 (.1004)
1000	-0.5104 (.0014)	-0.6474 (.0026)	-0.8007 (.0053)	-1.1684 (.0114)	-4.9106 (.1604)
2000	-0.4845 (.0018)	-0.6065 (.0029)	-0.7313 (.0054)	-1.0365 (.0085)	-3.7967 (.0930)
5000	-0.4615 (.0015)	-0.5598 (.0024)	-0.6666 (.0031)	-0.8994 (.0061)	-2.7274 (.0537)
10000	-0.4426 (.0010)	-0.5294 (.0019)	-0.6213 (.0026)	-0.8279 (.0058)	-2.2065 (.0315)
20000	-0.4254 (.0013)	-0.5019 (.0015)	-0.5838 (.0012)	-0.7547 (.0053)	-1.7996 (.0162)
50000	-0.4044 (.0008)	-0.4712 (.0006)	-0.5381 (.0025)	-0.6837 (.0028)	-1.4920 (.0128)
$\rho = -.5$					
100	-0.7288 (.0008)	-1.0197 (.0040)	-1.3469 (.0149)	-2.2467 (.0326)	-15.2665 (.2814)
200	-0.7132 (.0009)	-0.9628 (.0042)	-1.2580 (.0086)	-2.0360 (.0242)	-10.7777 (.2224)
500	-0.6913 (.0009)	-0.9018 (.0025)	-1.1282 (.0074)	-1.7361 (.0144)	-5.2840 (.0763)
1000	-0.6742 (.0009)	-0.8517 (.0037)	-1.0399 (.0066)	-1.5447 (.0207)	-4.1265 (.2068)
2000	-0.6595 (.0016)	-0.8145 (.0021)	-0.9806 (.0060)	-1.3444 (.0120)	-3.8262 (.3132)
5000	-0.6405 (.0007)	-0.7635 (.0027)	-0.8945 (.0038)	-1.1838 (.0056)	-3.3592 (.0650)
10000	-0.6268 (.0006)	-0.7384 (.0028)	-0.8467 (.0045)	-1.0907 (.0064)	-2.5872 (.0545)
20000	-0.6139 (.0005)	-0.7096 (.0012)	-0.8037 (.0030)	-1.0051 (.0054)	-2.1792 (.0219)
50000	-0.5986 (.0005)	-0.6770 (.0011)	-0.7528 (.0012)	-0.9073 (.0027)	-1.7432 (.0176)
$\rho = -1$					
100	-0.8203 (.0003)	-1.3360 (.0011)	-1.9292 (.0161)	-3.6100 (.0959)	-11.1357 (.2070)
200	-0.8460 (.0007)	-1.3202 (.0018)	-1.8284 (.0072)	-3.3258 (.0547)	-6.7476 (.1156)
500	-0.8673 (.0005)	-1.2840 (.0013)	-1.6813 (.0074)	-2.7863 (.0344)	-3.4806 (.0529)
1000	-0.8812 (.0004)	-1.2545 (.0018)	-1.5912 (.0074)	-2.4848 (.0292)	-2.8536 (.0796)
2000	-0.8931 (.0005)	-1.2274 (.0022)	-1.5066 (.0041)	-2.1151 (.0185)	-2.3803 (.0336)
5000	-0.9078 (.0004)	-1.1959 (.0020)	-1.4055 (.0054)	-1.8632 (.0134)	-3.7971 (.5144)
10000	-0.9183 (.0003)	-1.1716 (.0011)	-1.3506 (.0044)	-1.7248 (.0068)	-4.2253 (.0636)
20000	-0.9275 (.0003)	-1.1494 (.0007)	-1.2949 (.0022)	-1.5900 (.0068)	-3.3916 (.0742)
50000	-0.9377 (.0003)	-1.1244 (.0006)	-1.2410 (.0011)	-1.4553 (.0034)	-2.5589 (.0286)
$\rho = -2$					
100	-1.1666 (.0287)	-1.7567 (.0280)	-2.5088 (.0126)	-2.9460 (.3943)	-7.4078 (.2161)
200	-1.0819 (.0082)	-1.5870 (.0072)	-2.4282 (.0027)	-6.0270 (.3222)	-4.0766 (.0774)
500	-1.2270 (.0074)	-1.6332 (.0025)	-2.4098 (.0005)	-5.3965 (.0564)	-2.7471 (.0392)
1000	-1.3423 (.0112)	-1.6906 (.0028)	-2.4025 (.0003)	-4.7159 (.0490)	-2.4740 (.0511)
2000	-1.4403 (.0095)	-1.7355 (.0021)	-2.3945 (.0002)	-4.0446 (.0422)	-2.338 (.0476)4
5000	-1.5421(.0079)	-1.7831 (.0027)	-2.3833 (.0001)	-3.4533 (.0303)	-2.3393 (.0717)
10000	-1.6000 (.0051)	-1.8136 (.0008)	-2.3752 (.0001)	-3.2113 (.0117)	-2.2954 (.0409)
20000	-1.6603 (.0016)	-1.8411 (.0006)	-2.2938 (.0014)	-2.9380 (.0096)	-2.3372 (.0350)
50000	-1.7231 (.0007)	-1.8705 (.0005)	-2.2427 (.0009)	-2.6915 (.0039)	-5.4448 (.0711)

In Gomes and Martins [13], the choice of the level $\widehat{k}_1 := \min(n-1, [2n/\ln\ln n])$ led to quite nice results, not a long way from the ones got for the optimal $\widehat{k}_0$, which is presently still ideal, in the sense that we do not have yet a practical adequate way of estimating the optimal sample fraction to be taken in the estimation of ρ, made through the semi-parametric estimators we have been discussing.

Table 4: Simulated Mean Square Errors of $\widehat{\rho}_{n,0}^{(j)}$, $1 \leq j \leq 5$, for a Burr parent.

n	$MSE[\widehat{\rho}_{n,0}^{(1)}]$	$MSE[\widehat{\rho}_{n,0}^{(2)}]$	$MSE[\widehat{\rho}_{n,0}^{(3)}]$	$MSE[\widehat{\rho}_{n,0}^{(4)}]$	$MSE[\widehat{\rho}_{n,0}^{(5)}]$
			$\rho = -0.25$		
100	0.1481 (.0005)	0.4060 (.0025)	0.8942 (.0076)	3.3215 (.0387)	9374.8958 (654.8213)
200	0.1260 (.0005)	0.3259 (.0018)	0.6848 (.0056)	2.3115 (.0287)	4913.0523(386.7304)
500	0.1013(.0005)	0.2461 (.0013)	0.4843 (.0029)	1.4618 (.0167)	1002.7483 (67.6996)
1000	0.0856 (.0005)	0.1986 (.0016)	0.3796 (.0040)	1.0545 (.0134)	121.6084 (6.6267)
2000	0.0715 (.0005)	0.1611 (.0013)	0.2980 (.0032)	0.7748 (.0072)	23.6176 (.6346)
5000	0.0570 (.0004)	0.1231 (.0008)	0.2192 (.0021)	0.5358 (.0038)	8.5425 (.1823)
10000	0.0473 (.0003)	0.1000 (.0007)	0.1746 (.0010)	0.4148 (.0037)	5.0895 (.0929)
20000	0.0393 (.0002)	0.0812 (.0006)	0.1400 (.0009)	0.3236 (.0027)	3.1931 (.0423)
50000	0.0309 (.0002)	0.0623 (.0004)	0.1052 (.0007)	0.2328 (.0022)	1.9021 (.0303)
			$\rho = -.5$		
100	0.0626 (.0002)	0.3349 (.0023)	0.9437 (.0095)	4.6875 (.1036)	6276.5771 (618.7948)
200	0.0561 (.0002)	0.2718 (.0019)	0.7272 (.0083)	3.1831 (.0518)	2456.0292 (178.0512)
500	0.0465 (.0002)	0.2034 (.0016)	0.5079 (.0042)	1.9673 (.0257)	308.1235 (21.5135)
1000	0.0396 (.0002)	0.1612 (.0019)	0.3863 (.0042)	1.3773 (.0278)	80.6911 (6.8881)
2000	0.0332 (.0003)	0.1278 (.0010)	0.2951 (.0034)	0.9443 (.0127)	35.0274 (1.4872)
5000	0.0262 (.0002)	0.0951 (.0009)	0.2063 (.0022)	0.6051 (.0056)	12.3445 (.5075)
10000	0.0215 (.0001)	0.0754 (.0008)	0.1599 (.0021)	0.4464 (.0052)	6.3303 (.1515)
20000	0.0175 (.0001)	0.0586 (.0004)	0.1212 (.0012)	0.3247 (.0047)	3.6238 (.0577)
50000	0.0133 (.0001)	0.0427 (.0003)	0.0856 (.0006)	0.2165 (.0019)	1.9328 (.0301)
			$\rho = -1$		
100	0.0414 (.0010)	0.1283 (.0015)	1.0560 (.0139)	11.6701 (.4114)	2906.6778 (194.3627)
200	0.0312 (.0001)	0.1178 (.0004)	0.8249 (.0083)	7.5619 (.1788)	718.3368 (45.3380)
500	0.0237 (.0001)	0.0992 (.0005)	0.5891 (.0058)	4.2446 (.0733)	118.5709 (7.1232)
1000	0.0194 (.0001)	0.0834 (.0007)	0.4447 (.0071)	2.7866 (.0821)	77.6078 (4.1954)
2000	0.0159 (.0001)	0.0685 (.0006)	0.3320 (.0039)	1.7492 (.0359)	52.1321 (4.3634)
5000	0.0120 (.0000)	0.0522 (.0006)	0.2276 (.0026)	1.0084 (.0136)	46.4450 (3.0347)
10000	0.0098 (.0000)	0.0411 (.0002)	0.1699 (.0017)	0.6957 (.0105)	16.7723 (.5231)
20000	0.0078 (.0000)	0.0320 (.0002)	0.1240 (.0010)	0.4733 (.0052)	7.8969 (.3094)
50000	0.0058 (.0000)	0.0231 (.0001)	0.0842 (.0004)	0.2913 (.0022)	3.3347 (.0527)
			$\rho = -2$		
100	3.3284 (.6156)	6.5954 (1.4812)	9.2325 (1.6795)	46.6354 (4.1232)	988.7566 (91.1488)
200	0.9969 (.0134)	0.3587 (.0287)	0.5581 (.0911)	32.1181 (2.9150)	208.8464 (16.7336)
500	0.7151 (.0065)	0.1823 (.0011)	0.1878 (.0010)	15.2818 (.2952)	75.3674 (5.2986)
1000	0.5595 (.0053)	0.1388 (.0010)	0.1675 (.0003)	9.6053 (.1834)	63.0611 (4.3781)
2000	0.4387 (.0026)	0.1071 (.0006)	0.1575 (.0002)	5.7564 (.1165)	51.8075 (2.4815)
5000	0.3146 (.0033)	0.0759 (.0006)	0.1475 (.0001)	3.0080 (.0522)	52.0707 (4.3049)
10000	0.2442 (.0019)	0.0575 (.0003)	0.1410 (.0000)	1.9853 (.0399)	50.2123 (3.0299)
20000	0.1851 (.0008)	0.0434 (.0001)	0.1178 (.0008)	1.2528 (.0188)	47.4506 (2.2278)
50000	0.1296 (.0003)	0.0297 (.0001)	0.0855 (.0003)	0.7138 (.0045)	18.8070 (.5143)

3. So, the simulated behaviour of the estimators at the optimal level, herewith presented, has not yet any application in practice. These results merely exhibit the optimal MSE we may obtain, should we be able to choose k in an optimal way.

REFERENCES

[1] Beirlant, J.; Vynckier, P. and Teugels, J.L. – Excess function and estimation of the extreme-value index, *Bernoulli,* 2 (1996), 293–318.

[2] Beirlant, J.; Vynckier, P. and Teugels, J.L. – Tail index estimation, Pareto quantile plots, and regression diagnostics, *J. Amer. Statist. Assoc.,* 91 (1996), 1659–1667.

[3] Beirlant, J.; Dierckx, G.; Goegebeur, Y. and Matthys, G. – Tail index estimation and an exponential regression model, *Extremes,* 2:2 (1999), 177–200.

[4] Danielsson, J.; Haan, L. de; Peng, L. and Vries, C.G. de – Using a bootstrap method to choose the sample fraction in the tail index estimation, *J. Multivariate Analysis,* 76 (1998), 226–248.

[5] Dekkers, A.L.M.; Einmahl, J.H.J. and Haan, L. de – A moment estimator for the index of an extreme-value distribution, *Ann. Statist.,* 17 (1989), 1833–1855.

[6] Draisma, G.; Haan, L. de; Peng, L. and Pereira, T.T. – A bootstrap-based method to achieve optimality in estimating the extreme value index, *Extremes,* 2:4 (1999), 367–404.

[7] Drees, H. and Kaufman, E. – Selecting the optimal sample fraction in univariate extreme value estimation, *Stoch. Proc. and Appl.,* 75 (1998), 149–172.

[8] Feuerverger, A. and Hall, P. – Estimating a tail exponent by modelling departure from a Pareto distribution, *Ann. Statist.,* 27 (1999), 760–781.

[9] Galambos, J. – *The Asymptotic Theory of Extreme Order Statistics* (2nd edition), Krieger, 1987.

[10] Gnedenko, B.V. – Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math.,* 44 (1943), 423–453.

[11] Geluk, J. and Haan, L. de – *Regular Variation, Extensions and Tauberian Theorems*, CWI Tract 40, Center for Mathematics and Computer Science, Amsterdam, Netherlands, 1987.

[12] Gomes, M.I. and Martins, M.J. – Generalizations of the Hill estimator – asymptotic versus finite sample behaviour, *J. Statist. Planning and Inference,* 93 (2001), 161–180.

[13] Gomes, M.I. and Martins, M.J. – *Generalized Jackknife estimators of the tail index based on the estimation of the second order parameter*, Notas e Comunicações CEAUL 9/2001, Submitted.

[14] Gomes, M.I.; de Haan, L. and Peng, L. – *Semi-parametric estimation of the second order parameter – asymptotic and finite sample behaviour*, Notas e Comunicações CEAUL 8/2000, To appear in *Extremes*.

[15] Gomes, M.I.; Martins, M.J. and Neves, M. – Alternatives to a semi-parametric estimator of parameters of rare events — the Jackknife methodology, *Extremes,* 3:3 (2000), 207–229.

[16] Guillou, A. and Hall, P. – A diagnostic for selecting the threshold in extreme value analysis, *J.R. Statist. Soc.,* B63 (2001), 293–305.

[17] Haan, L. de – *On Regular Variation and its Application to the Weak Convergence of Sample Extremes*, Mathematical Centre Tract 32, Amsterdam, 1970.

[18] Haan, L. de and Stadtmüller, U. – Generalized regular variation of second order, *J. Austral. Math. Soc. (A),* 61 (1996), 381–395.

[**19**] Hall, P. and Welsh, A.H. – Adaptive estimates of parameters of regular variation, *Ann. Statist.,* 13 (1985), 331–341.

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

[**21**] Leadbetter, M.R. and Nandagopalan, S. – *On exceedance point processes for stationary sequences under mild oscillation restrictions.* In "Extreme Value Theory" (J. Hüsler and R.-D. Reiss, Eds.), Springer-Verlag, 1989, pp. 69–80.

M.I. Fraga Alves and M. Ivette Gomes,
D.E.I.O. and C.E.A.U.L., University of Lisbon

and

Laurens de Haan,
Erasmus University Rotterdam