On the estimation of the second order parameter for heavy-tailed distributions

El Hadji Deme, Laurent Gardes, Stéphane Girard

On the estimation of the second order parameter for heavy-tailed distributions

El hadji DEME[1], Laurent GARDES[2] and Stéphane GIRARD[3,⋆]

[1] LERSTAD, Université Gaston Berger, Saint-Louis, Sénégal.

[2] Université de Strasbourg & CNRS, IRMA UMR 7501.

7 rue René Descartes, 67084 Strasbourg Cedex, France.

[3] Team Mistis, INRIA Rhône-Alpes & Laboratoire Jean Kuntzmann.

655 avenue de l'Europe, Montbonnot, 38334 Saint-Ismier Cedex, France.

⋆ Corresponding author, `Stephane.Girard@inria.fr`

Abstract

The extreme-value index γ is an important parameter in extreme-value theory since it controls the first order behavior of the distribution tail. In the literature, numerous estimators of this parameter have been proposed especially in the case of heavy-tailed distributions, which is the situation considered here. Most of these estimators depend on the k largest observations of the underlying sample. Their bias is controlled by the second order parameter ρ. In order to reduce the bias of γ's estimators or to select the best number k of observations to use, the knowledge of ρ is essential. In this paper, we propose a simple approach to estimate the second order parameter ρ leading to both existing and new estimators. We establish a general result that can be used to easily prove the asymptotic normality of a large number of estimators proposed in the literature or to compare different estimators within a given family. Some illustrations on simulations are also provided.

Keywords: Extreme-value theory; Heavy-tailed distribution; Extreme-value index; Second order parameter; Asymptotic properties.

AMS 2000 Subject Classifications 62G32 - 62G30 - 60G70 - 62E20

1 Introduction

Extreme-value theory establishes the asymptotic behavior of the largest observations in a sample. It provides methods for extending the empirical distribution function beyond the observed data. It is thus possible to estimate quantities related to the tail of a distribution such as small exceedance probabilities or extreme quantiles. We refer to [11, 25] for general accounts on extreme-value theory. More specifically, let $X_1, \ldots, X_n$ be a sequence of random variables (rv), independent and identically distributed from a cumulative distribution function (cdf) F. Extreme-value theory establishes that the asymptotic distribution of the maximum $X_{n,n} = \max\{X_1, \ldots, X_n\}$ properly rescaled is the extreme-value distribution with cdf

$$G_\gamma(x) = \exp\left(-(1 + \gamma x)_+\right)^{-1/\gamma}$$

where $y_+ = \max(y, 0)$. The parameter $\gamma \in \mathbb{R}$ is referred to as the extreme-value index. Here, we focus on the case where $\gamma > 0$. In such a situation, F is said to belong to the maximum domain of attraction of the Fréchet distribution. In this domain of attraction, a simple characterization of distributions is available: the quantile function $U(x) := F^{\leftarrow}(1 - 1/x)$ can be written as

$$U(x) = x^\gamma \ell(x),$$

where ℓ is a slowly varying function at infinity *i.e.* for all $\lambda > 0$,

$$\lim_{x\to\infty} \frac{\ell(\lambda x)}{\ell(x)} = 1. \tag{1}$$

The distribution F is said to be heavy tailed and the extreme-value parameter γ governs the heaviness of the tail. The estimation of γ is a central topic in the analysis of such distributions. Several estimators have thus been proposed in the statistical literature and their asymptotic distributions established under a second order condition: There exist a function $A(x) \to 0$ of constant sign for large values of x and a second order parameter $\rho < 0$ such that, for every $\lambda > 0$,

$$\lim_{x\to\infty} \frac{1}{A(x)} \log\left(\frac{\ell(\lambda x)}{\ell(x)}\right) = K_\rho(\lambda) := \int_1^\lambda u^{\rho-1} du. \tag{2}$$

Let us highlight that (2) implies that $|A|$ is regularly varying with index ρ, see [16]. Hence, as the second order parameter ρ decreases, the rate of convergence in (1) increases. Thus, the knowledge of ρ can be of high interest in real problems. For example, the second order parameter is of primordial importance in the adaptive choice of the best number of upper order statistics to be considered in the estimation of the extreme-value index [24]. The estimation of ρ can also be used to propose bias reduced estimators of the extreme value index (see for instance [4, 21, 23]) or of the Weibull tail-coefficient [9, 10], even though some bias reduction can be achieved with the canonical choice $\rho = -1$ as suggested in [12, 22]. For the above mentioned reasons, the estimation of the second order parameter ρ has received a lot of attention in the extreme-value literature, see for instance [3, 6, 13, 14, 17, 19, 26, 30, 31].

In this paper, we propose a simple and general approach to estimate ρ. Let $\mathbb{I} = {}^t(1,\dots,1) \in \mathbb{R}^d$. The two main ingredients of our approach are a random variable $T_n = T_n(X_1,\dots,X_n) \in \mathbb{R}^d$ verifying the following three assumptions:

(T1) There exists rvs ω_n, χ_n and a function $f : \mathbb{R}^- \to \mathbb{R}^d$ such that $\omega_n^{-1}(T_n - \chi_n \mathbb{I}) \xrightarrow{\mathbb{P}} f(\rho)$.

and a function $\psi : \mathbb{R}^d \to \mathbb{R}$ such that

(Ψ1) $\psi(x + \lambda \mathbb{I}) = \psi(x)$ for all $x \in \mathbb{R}^d$ and $\lambda \in \mathbb{R}$,

(Ψ2) $\psi(\lambda x) = \psi(x)$ for all $\lambda \in \mathbb{R} \setminus \{0\}$.

Note that **(T1)** imposes that T_n properly normalized converges in probability to some function of ρ, while **(Ψ1)** and **(Ψ2)** mean that ψ is both location and shift invariant. Starting from these three assumptions, we straightforwardly obtain that

$$\psi(\omega_n^{-1}(T_n - \chi_n \mathbb{I})) = \psi(T_n) \xrightarrow{\mathbb{P}} \psi(f(\rho)),$$

under a continuity condition on ψ. Denoting by $Z_n := \psi(T_n)$ and by $\varphi := \psi \circ f : \mathbb{R}^- \to \mathbb{R}$, we obtain $Z_n \xrightarrow{\mathbb{P}} \varphi(\rho)$. It is thus clear that, under an additional regularity assumption and assuming that both Z_n and φ are known, ρ can be consistently estimated thanks to $\varphi^{-1}(Z_n)$. This estimation principle is described more precisely in Section 2. The consistency and asymptotic normality of the proposed estimator is also established. Examples of T_n random variables are presented in Section 3. Some functions ψ are proposed in Section 4 and it is shown that the above mentioned estimators [6, 13, 14, 17, 19] can be read as particular cases of our approach. As a consequence, this remark permits to establish their asymptotic properties in a simple and unified way. We illustrate how several asymptotically Gaussian estimators can be derived from this framework. Finally, some estimators are compared in Section 5 both from the asymptotic and finite sample size performances points of view.

2 Main results

Recall that T_n is a $\mathbb{R}^d$- random vector verifying **(T1)** and ψ is a function $\mathbb{R}^d \to \mathbb{R}$ verifying **(Ψ1)** and **(Ψ2)**. We further assume that:

(Ψ3) There exist $J_0 \subseteq \mathbb{R}^-$ and an open interval $J \subset \mathbb{R}$ such that $\varphi = \psi \circ f$ is a bijection $J_0 \to J$.

Under this assumption, the following estimator of ρ may be considered:

$$\hat{\rho}_n = \left| \begin{array}{ll} \varphi^{-1}(Z_n) & \text{if } Z_n \in J \\ 0 & \text{otherwise.} \end{array} \right. \tag{3}$$

To derive the consistency of $\hat{\rho}_n$, an additional regularity assumption is introduced:

(Ψ4) ψ is continuous in a neighborhood of $f(\rho)$ and f is continuous in a neighborhood of ρ.

The proof of the next result is based on the heuristic consideration of Section 1 and is detailed in Section 6.

Theorem 1. *If* **(T1)** *and* **(Ψ1)**–**(Ψ4)** *hold then* $\hat{\rho}_n \xrightarrow{\mathbb{P}} \rho$ *as* $n \to \infty$.

The asymptotic normality of $\hat{\rho}_n$ can be established under a stronger version of **(Ψ4)**:

(Ψ5) ψ is continuously differentiable in a neighborhood of $f(\rho)$ and f is continuously differentiable in a neighborhood of ρ,

and the assumption that a normalized version of T_n is itself asymptotically Gaussian:

(T2) There exists two rvs ω_n, χ_n, a sequence $v_n \to \infty$, two functions f, $m : \mathbb{R}^- \to \mathbb{R}^d$ and a $d \times d$ matrix Σ such that $v_n(\omega_n^{-1}(T_n - \chi_n \mathbb{I}) - f(\rho)) \xrightarrow{d} \mathcal{N}_d(m(\rho), \gamma^2 \Sigma)$.

Theorem 2. *Suppose* **(T2)**, **(Ψ1)**–**(Ψ3)** *and* **(Ψ5)** *hold. If* $\rho \in J_0$ *and* $\varphi'(\rho) \neq 0$, *then*

$$v_n(\hat{\rho}_n - \rho) \xrightarrow{d} \mathcal{N}\left(\frac{m_\psi(\rho)}{\varphi'(\rho)}, \frac{\gamma^2 \sigma_\psi^2(\rho)}{(\varphi'(\rho))^2}\right),$$

with $\varphi'(\rho) = {}^t f'(\rho) \nabla \psi(f(\rho))$ *and where we have defined*

$$\begin{aligned} m_\psi(\rho) &:= {}^t m(\rho)\, \nabla \psi(f(\rho)), \\ \sigma_\psi^2(\rho) &:= {}^t \nabla \psi(f(\rho))\, \Sigma\, \nabla \psi(f(\rho)). \end{aligned}$$

3 Examples of T_n random variables

Let $X_{1,n} \leq \ldots \leq X_{n,n}$ be the sample of ascending order statistics and $k = k_n$ be an intermediate sequence *i.e.* such that $k \to \infty$ and $k/n \to 0$ as $n \to \infty$. Most extreme-value estimators are based either on the log-excesses $(\log X_{n-j+1,n} - \log X_{n-k,n})$ or on the rescaled log-spacings $j(\log X_{n-j+1,n} - \log X_{n-j,n})$ defined for $j = 1, \ldots, k$. In the following, two examples of T_n random variables are presented based on weighted means of the log-excesses and of the rescaled log-spacings. The first example is based on

$$R_k(\tau) = \frac{1}{k} \sum_{j=1}^{k} H_\tau\left(\frac{j}{k+1}\right) j(\log X_{n-j+1,n} - \log X_{n-j,n}), \tag{4}$$

where $H_\tau : [0,1] \to \mathbb{R}$ is a weight function indexed by a parameter $\tau \in (0, \infty)$. Without loss of generality, one can assume that H_τ integrates to one. This random variable is used for instance in [1] to estimate the extreme-value index γ, in [17, 26, 30] to estimate the second order parameter ρ and in [18] to estimate the third order parameter, see condition **(C2)** below. It is a particular case of the kernel statistic introduced in [7]. Let us also note that, in the case where $H_\tau(u) = 1$

for all $u \in [0,1]$, $R_k(\tau)$ reduces to the well-known Hill estimator [27]. The asymptotic properties of $R_k(\tau)$ require some technical condition (denoted by **(C1)**) on the weight function H_τ. It has been first introduced in [1] and it is recalled hereafter. Introducing the operator

$$\mu : h \in L_2([0,1]) \longrightarrow \mu(h) = \int_0^1 h(u)du \in \mathbb{R}$$

and $I_t(u) = u^{-t}$ for $t \leq 0$ and $u \in (0,1]$, the condition can be written as

(C1) $H_\tau \in L_2([0,1])$, $\mu(|H_\tau| I_{\rho+1+\varepsilon}) < \infty$ and

$$H_\tau(t) = \frac{1}{t}\int_0^t u(\nu)d\nu \text{ and}$$

for some $\varepsilon > 0$ and for some function u satisfying for all $j = 1, \ldots, k$

$$\left|(k+1)\int_{(j-1)/(k+1)}^{j/(k+1)} u(t)dt\right| \leq g\left(\frac{j}{k+1}\right),$$

where g is a positive continuous and integrable function defined on $(0,1)$. Furthermore, for $\eta \in \{0,1\}$, and $k \to \infty$:

$$\begin{aligned} \frac{1}{k}\sum_{j=1}^{k} H_\tau\left(\frac{j}{k+1}\right)\left(\frac{j}{k+1}\right)^{-\eta\rho} &= \mu(H_\tau I_{\eta\rho}) + o(k^{-1/2}), \\ \max_{j\in\{1,\ldots,k\}}\left|H_\tau\left(\frac{j}{k+1}\right)\right| &= o(k^{1/2}). \end{aligned}$$

It is then possible to define $T_n^{(R)}$ on the basis of $R_k(\tau)$, given in (4), as

$$T_n^{(R)} = \left(T_{n,i}^{(R)} = (R_k(\tau_i)/\gamma)^{\theta_i},\ i = 1,\ldots,d\right), \tag{5}$$

where θ_i, $i = 1, \ldots, d$ are positive parameters. In the next lemma, it is proved that $T_n^{(R)}$ satisfies condition **(T2)** under a third order condition, which is a refinement of (2):

(C2) There exist functions $A(x) \to 0$ and $B(x) \to 0$ both of constant sign for large values of x, a second order parameter $\rho < 0$ and a third order parameter $\beta < 0$ such that, for every $\lambda > 0$,

$$\lim_{x\to\infty} \frac{\left(\log \ell(\lambda x) - \log \ell(x)\right)/A(x) - K_\rho(\lambda)}{B(x)} = L_{(\rho,\beta)}(\lambda) := \int_1^\lambda s^{\rho-1}\int_1^s u^{\beta-1}duds,$$

and the functions $|A|$ and $|B|$ are regularly varying functions with index ρ and β respectively.

This condition is the cornerstone for establishing the asymptotic normality of estimators of ρ. Let us denote by $Y_{n-k,n}$ the $n-k$ largest order statistics from a n-sample of standard Pareto rv.

Lemma 1. *Suppose* **(C1)**, **(C2)** *hold and let $k = k_n$ be an intermediate sequence k such that*

$$k \to \infty,\ n/k \to \infty,\ k^{1/2}A(n/k) \to \infty,\ k^{1/2}A^2(n/k) \to \lambda_A \text{ and } k^{1/2}A(n/k)B(n/k) \to \lambda_B, \quad (6)$$

for $\lambda_A \in \mathbb{R}$ and $\lambda_B \in \mathbb{R}$. Then, the random vector $T_n^{(R)}$ satisfies **(T2)** *with $\omega_n^{(R)} = A(Y_{n-k,n})/\gamma$, $\chi_n^{(R)} = 1$, $v_n = k^{1/2}A(n/k)$,*

$$\begin{aligned} f^{(R)}(\rho) &= (\theta_i \mu(H_{\tau_i} I_\rho),\ i = 1, \ldots, d)\,, \\ m^{(R)}(\rho) &= \left(\lambda_A \frac{\theta_i(\theta_i - 1)}{2\gamma} \mu^2(H_{\tau_i} I_\rho) - \lambda_B \theta_i \mu(H_{\tau_i} I_\rho K_{-\beta});\ i = 1, \ldots, d\right), \end{aligned}$$

and, for $(i,j) \in \{1, \ldots, d\}^2$, $\Sigma_{i,j}^{(R)} = \theta_i \theta_j \mu(H_{\tau_i} H_{\tau_j})$.

The proof is a straightforward consequence of Theorem 2 and Appendix A.5 in [17].

The second example requires some additional notations. Let us consider the operator

$$\vartheta : (h_1, h_2) \in L_2([0,1]) \times L_2([0,1]) \longrightarrow \vartheta(h_1, h_2) = \int_0^1 \int_0^1 h_1(u) h_2(v)(u \wedge v - uv)dudv \in \mathbb{R}$$

and the two functions $\bar{I}_t(u) = (1-u)^{-t}$ and $J_t(u) = (-\log u)^{-t}$ defined for $t \le 0$ and $u \in (0,1]$. The random variables of interest are

$$S_k(\tau, \alpha) = \frac{1}{k} \sum_{j=1}^{k} G_{\tau,\alpha}\left(\frac{j}{k+1}\right) (\log X_{n-j+1,n} - \log X_{n-k,n})^\alpha\,, \quad (7)$$

where $G_{\tau,\alpha}$ is a positive function indexed by two positive parameters α and τ. Without loss of generality, it can be assumed that $\mu(G_{\tau,\alpha} J_{-\alpha}) = 1$. In [8, 20, 29] several estimators of γ based on $S_k(\tau, \alpha)$ are introduced in the particular case where G is constant. Most recently, in [6, 14, 26, 28], $S_k(\tau, \alpha)$ is used to estimate the parameters γ and ρ. The asymptotic distribution of these estimators is obtained under the following assumption on the function $G_{\tau,\alpha}$.

(C3) The function $G_{\tau,\alpha}$ is positive, non-increasing and integrable on $(0,1)$. Furthermore, there exists $\delta > 1/2$ such that $0 < \mu(G_{\tau,\alpha} I_\delta) < \infty$ and $0 < \mu(G_{\tau,\alpha} \bar{I}_\delta) < \infty$.

It is then possible to define $T_n^{(S)}$ on the basis of $S_k(\tau, \alpha)$, see (7), as

$$T_n^{(S)} = \left(T_{n,i}^{(S)} = (S_k(\tau_i, \alpha_i)/\gamma^{\alpha_i})^{\theta_i},\ i = 1, \ldots, d\right). \quad (8)$$

The following result is the analogous of Lemma 1 for the above random variables.

Lemma 2. *Suppose* **(C2)**, **(C3)** *hold. If the intermediate sequence k satisfy (6) then the random vector $T_n^{(S)}$ satisfies* **(T2)** *with $\omega_n^{(S)} = A(n/k)/\gamma$, $\chi_n^{(S)} = 1$, $v_n = k^{1/2}A(n/k)$,*

$$\begin{aligned} f^{(S)}(\rho) &= (-\theta_i \alpha_i \mu(G_{\tau_i,\alpha_i} J_{1-\alpha_i} K_{-\rho});\ i = 1, \ldots, d)\,, \\ m^{(S)}(\rho) &= \left(\lambda_A \frac{\theta_i \alpha_i(\alpha_i - 1)}{2\gamma} \mu(G_{\tau_i \alpha_i} J_{2-\alpha_i} K_{-\rho}^2) + \lambda_B \alpha_i \theta_i \mu(G_{\tau_i,\alpha_i} J_{1-\alpha_i} L_{(-\rho,-\beta)});\ i = 1, \ldots, d\right), \end{aligned}$$

and, for $(i,j) \in \{1, \ldots, d\}^2$, $\Sigma_{i,j}^{(S)} = \theta_i \theta_j \alpha_i \alpha_j \vartheta(G_{\tau_i,\alpha_i} J_{1-\alpha_i}, G_{\tau_j,\alpha_j} J_{1-\alpha_j})$.

The proof is a straightforward consequence of Proposition 3 and Lemma 1 in [6]. In the next section, we illustrate how the combination of $T_n^{(R)}$ or $T_n^{(S)}$ with some function ψ following (3) can lead to existing or new estimators of ρ.

4 Applications

In this section, we propose estimators of ρ based on the random variable $T_n^{(R)}$ (subsection 4.1) and $T_n^{(S)}$ (subsection 4.2). In both cases, $d = 8$ and the following function $\psi_\delta : \mathcal{D} \mapsto \mathbb{R} \setminus \{0\}$ is considered

$$\psi_\delta(x_1, \ldots, x_8) = \widetilde{\psi}_\delta(x_1 - x_2, x_3 - x_4, x_5 - x_6, x_7 - x_8), \tag{9}$$

where $\delta \geq 0$, $\mathcal{D} = \{(x_1, \ldots, x_8) \in \mathbb{R}^8;\ x_1 \neq x_2,\ x_3 \neq x_4,\ \text{and}\ (x_5 - x_6)(x_7 - x_8) > 0\}$, and $\widetilde{\psi}_\delta : \mathbb{R}^4 \mapsto \mathbb{R}$ is given by:

$$\widetilde{\psi}_\delta(y_1, \ldots, y_4) = \frac{y_1}{y_2}\left(\frac{y_4}{y_3}\right)^\delta.$$

Let us highlight that ψ_δ verifies the invariance properties **(Ψ1)** and **(Ψ2)**.

4.1 Estimators based on the random variable $R_k(\tau)$

Since $d = 8$, the random variable $T_n^{(R)}$ defined in (5) depends on 16 parameters: $\{(\theta_i, \tau_i) \in (0, \infty)^2,\ i = 1, \ldots, 8\}$. The following condition on these parameters is introduced. Let $\tilde{\theta} = (\tilde{\theta}_1, \ldots, \tilde{\theta}_4) \in (0, \infty)^4$ with $\tilde{\theta}_3 \neq \tilde{\theta}_4$.

(C4) $\{\theta_i = \tilde{\theta}_{\lceil i/2 \rceil},\ i = 1, \ldots, 8\}$ with $\delta = (\tilde{\theta}_1 - \tilde{\theta}_2)/(\tilde{\theta}_3 - \tilde{\theta}_4)$. Furthermore, $\tau_1 < \tau_2 \leq \tau_3 < \tau_4$, $\tau_5 < \tau_6 \leq \tau_7 < \tau_8$,

where $\lceil x \rceil = \inf\{n \in \mathbb{N} | x \leq n\}$. Under this condition, $T_n^{(R)}$ involves 12 free parameters. We also introduce the following notations: $Z_n^{(R)} = \psi_\delta(T_n^{(R)})$ and $\varphi_\delta^{(R)} = \psi_\delta \circ f^{(R)}$ where $f^{(R)}$ is given in Lemma 1. Note that, since $\delta = (\tilde{\theta}_1 - \tilde{\theta}_2)/(\tilde{\theta}_3 - \tilde{\theta}_4)$, it is easy to check that $Z_n^{(R)}$ does not depend on the unknown parameter γ. We now establish the asymptotic normality of the estimator $\hat{\rho}_n^{(R)}$ defined by (3) when $T_n^{(R)}$ and the function ψ_δ are used:

$$\hat{\rho}_n^{(R)} = \left| \begin{array}{ll} (\varphi_\delta^{(R)})^{-1}(Z_n^{(R)}) & \text{if } Z_n^{(R)} \in J \\ 0 & \text{otherwise.} \end{array} \right. \tag{10}$$

The following additional condition is required:

(C5) The function $\nu_\rho(\tau) = \mu(H_\tau I_\rho)$ is differentiable with, for all $\rho < 0$ and all $\tau \in \mathbb{R}$, $\nu'_\rho(\tau) > 0$.

Let us denote for $i \in \{1, \ldots, 4\}$,

$$m_A^{(R,i)} = \exp\left\{(\tilde{\theta}_i - 1)(\nu_\rho(\tau_{2i-1}) + \nu_\rho(\tau_{2i}))\right\},\ m_B^{(R,i)} = \exp\left\{\frac{\mu\left((H_{\tau_{2i-1}} - H_{\tau_{2i}})I_\rho K_{-\beta}\right)}{\nu_\rho(\tau_{2i-1}) - \nu_\rho(\tau_{2i})}\right\},$$

and for $u \in [0,1]$,

$$v^{(R,i)}(u) = \exp\left\{ \frac{H_{\tau_{2i-1}}(u) - H_{\tau_{2i}}(u)}{\nu_\rho(\tau_{2i-1}) - \nu_\rho(\tau_{2i})} \right\}.$$

For the sake of simplicity, we also introduce $m_A^{(R)} = (m_A^{(R,i)},\ i = 1,\dots,4)$, $m_B^{(R)} = (m_B^{(R,i)},\ i = 1,\dots,4)$ and $v^{(R)} = (v^{(R,i)},\ i = 1,\dots,4)$.

Corollary 1. *Suppose* **(C1)***,* **(C2)***,* **(C4)** *and* **(C5)** *hold. There exist two intervals J and J_0 such that for all $\rho \in J_0$ and for a sequence k satisfying (6),*

$$k^{1/2} A(n/k)(\hat{\rho}_n^{(R)} - \rho) \xrightarrow{d} \mathcal{N}\left(\frac{\lambda_A}{2\gamma} \mathcal{AB}_1^{(R)}(\delta,\rho) - \lambda_B \mathcal{AB}_2^{(R)}(\delta,\rho,\beta),\ \gamma^2 \mathcal{AV}^{(R)}(\delta,\rho) \right)$$

where

$$\begin{aligned}
\mathcal{AB}_1^{(R)}(\delta,\rho) &= \frac{\varphi_\delta^{(R)}(\rho)}{[\varphi_\delta^{(R)}]'(\rho)} \log \tilde{\psi}_\delta(m_A^{(R)}), \\
\mathcal{AB}_2^{(R)}(\delta,\rho,\beta) &= \frac{\varphi_\delta^{(R)}(\rho)}{[\varphi_\delta^{(R)}]'(\rho)} \log \tilde{\psi}_\delta(m_B^{(R)}), \\
\mathcal{AV}^{(R)}(\delta,\rho) &= \left(\frac{\varphi_\delta^{(R)}(\rho)}{[\varphi_\delta^{(R)}]'(\rho)} \right)^2 \mu\left(\log^2 \tilde{\psi}_\delta(v^{(R)}) \right).
\end{aligned}$$

Note that this result can be read as an extension of [17], Proposition 3, in two ways. First, we do not limit ourselves to the case $\delta = 1$. Second, we do not assume that the function $\varphi_\delta^{(R)}$ is a bijection, but it is shown to be a consequence of **(C4)**. Besides, the proof of Corollary 1 is very simple based on Theorem 2 and Lemma 1, see Section 6 for details.

As an example, the function $H_\tau : u \in [0,1] \mapsto \tau u^{\tau-1}$, $\tau \geq 1$ satisfies conditions **(C1)** and **(C5)** since $\nu_\rho(\tau) = \tau/(\tau-\rho)$. Letting $\tau_1 \leq \tau_5$, $\tau_2 = \tau_3$, $\tau_4 = \tau_8$ and $\tau_6 = \tau_7$ leads to a simple expression of $\varphi_\delta^{(R)}$:

$$\varphi_\delta^{(R)}(\rho) = \omega(\delta,\tilde{\theta}) \left(\frac{\tau_4 - \rho}{\tau_1 - \rho} \right) \left(\frac{\tau_5 - \rho}{\tau_4 - \rho} \right)^\delta \text{ where } \omega(\delta,\tilde{\theta}) = \left(\frac{\tilde{\theta}_1(\tau_1 - \tau_2)}{\tilde{\theta}_2(\tau_2 - \tau_4)} \right) \left(\frac{\tilde{\theta}_4(\tau_6 - \tau_4)}{\tilde{\theta}_3(\tau_5 - \tau_6)} \right)^\delta. \quad (11)$$

Moreover, one also has explicit forms for J_0 and J in two situations:

(i) If $0 \leq \delta \leq \delta_0 := (\tau_4 - \tau_1)/(\tau_4 - \tau_5)$ then $\varphi_\delta^{(R)}$ is increasing from $J_0 = \mathbb{R}^-$ to $J = \omega(\delta,\tilde{\theta}) \bullet (1, \widetilde{\psi}_\delta(\tau_4,\tau_1,\tau_4,\tau_5))$.

(ii) If $\delta \geq \delta_1 := \delta_0\tau_5/\tau_1$ then $\varphi_\delta^{(R)}$ is decreasing from $J_0 = \mathbb{R}^-$ to $J = \omega(\delta,\tilde{\theta}) \bullet (\widetilde{\psi}_\delta(\tau_4,\tau_1,\tau_4,\tau_5), 1)$.

Here, $\bullet$ denotes the scaling operator. The case $\delta \in (\delta_0, \delta_1)$ is not considered here, since one can show that, in this situation, $J_0 \subsetneq \mathbb{R}^-$ and thus the condition $\rho \in J_0$ of Corollary 1 is not necessarily satisfied. Let us now list some particular cases where the inverse function of $\varphi_\delta^{(R)}$ is explicit.

Example 1. *Let $\delta = 1$ i.e. $\tilde{\theta}_1 - \tilde{\theta}_2 = \tilde{\theta}_3 - \tilde{\theta}_4$. The rv $Z_n^{(R)}$ is denoted by $Z_{n,1}^{(R)}$. Since $\delta_0 > 1$, we are in situation (i) and*

$$\hat{\rho}_{n,1}^{(R)} = \frac{\tau_5 \omega(1,\tilde{\theta}) - \tau_1 Z_{n,1}^{(R)}}{\omega(1,\tilde{\theta}) - Z_{n,1}^{(R)}} \mathbb{1}\{Z_{n,1}^{(R)} \in \omega(1,\tilde{\theta}) \bullet (1, \widetilde{\psi}_1(\tau_4,\tau_1,\tau_4,\tau_5))\}.$$

Remark that this estimator coincides with the one proposed in [17], Lemma 1.

Example 2. *Let $\delta = 0$ i.e. $\tilde\theta_1 = \tilde\theta_2$. The rv $Z_n^{(R)}$ is thus denoted by $Z_{n,2}^{(R)}$. Again, we are in situation (i) and a new estimator of ρ is obtained*

$$\hat\rho_{n,2}^{(R)} = \frac{\tau_4\omega(0,\tilde\theta) - \tau_1 Z_{n,2}^{(R)}}{\omega(0,\tilde\theta) - Z_{n,2}^{(R)}} \mathbb{1}\{Z_{n,2}^{(R)} \in \omega(0,\tilde\theta) \bullet (1, \widetilde{\psi}_0(\tau_4,\tau_1,\tau_4,\tau_5))\}.$$

Example 3. *Let $\tau_1 = \tau_5$. In this case $\delta_0 = \delta_1 = 1$ and thus, we are in situation (i) if $\delta < 1$ and in situation (ii) otherwise. In this case, the rv $Z_n^{(R)}$ is denoted by $Z_{n,3}^{(R)}$. A new estimator of ρ is obtained:*

$$\hat\rho_{n,3}^{(R)} = \frac{\tau_4(Z_{n,3}^{(R)}/\omega(\delta,\tilde\theta))^{1/(\delta-1)} - \tau_1}{(Z_{n,3}^{(R)}/\omega(\delta,\tilde\theta))^{1/(\delta-1)} - 1} \mathbb{1}\{Z_{n,3}^{(R)} \in J\}.$$

4.2 Estimators based on the random variable $S_k(\tau,\alpha)$

The random variable $T_n^{(S)}$ defined in (8) depends on 24 parameters: $\{(\theta_i,\tau_i,\alpha_i) \in (0,\infty)^3,\ i = 1,\dots,8\}$. Let $(\zeta_1,\dots,\zeta_4) \in (0,\infty)^4$ with $\zeta_3 \neq \zeta_4$. In the following, we assume that

(C6) $\{\theta_i\alpha_i = \zeta_{\lceil i/2\rceil},\ i = 1,\dots,8\}$ with $\delta = (\zeta_1-\zeta_2)/(\zeta_3-\zeta_4)$. Furthermore, $(\tau_{2i-1},\alpha_{2i-1}) \neq (\tau_{2i},\alpha_{2i})$, for $i = 1,\dots,4$ and, for $i = 3,4$, $(\tau_{2i-1},\alpha_{2i-1}) < (\tau_{2i},\alpha_{2i})$,

where $(x,y) \neq (s,t)$ means that $x \neq s$ and/or $y \neq t$ and $(x,y) < (s,t)$ means that $x < s$ and $y \le t$ or $x = s$ and $y < t$. We introduce the notations: $Z_n^{(S)} = \psi_\delta(T_n^{(S)})$ and $\varphi_\delta^{(S)} = \psi_\delta \circ f^{(S)}$ where $f^{(S)}$ is given in Lemma 2. Under this condition, $T_n^{(S)}$ involves 20 free parameters. Besides, since $\delta = (\zeta_1-\zeta_2)/(\zeta_3-\zeta_4)$, it is easy to check that $Z_n^{(S)}$ does not depend on the unknown parameter γ. To establish the asymptotic distribution of the estimator $\hat\rho_n^{(S)}$, the following condition is required:

(C7) For all $\rho < 0$, the function $\nu_\rho(\tau,\alpha) = \mu(G_{\tau,\alpha}J_{1-\alpha}K_{-\rho})$ is differentiable with $\frac{\partial}{\partial\tau}\nu_\rho(\tau,\alpha) > 0$ and $\frac{\partial}{\partial\alpha}\nu_\rho(\tau,\alpha) > 0$ for all $\alpha > 0$ and all $\tau \in \mathbb{R}$.

For $i = 1,\dots,4$, we introduce the notations:

$$m_A^{(S,i)} = \exp\left\{\frac{(\alpha_{2i-1}-1)\mu(G_{\tau_{2i-1},\alpha_{2i-1}}J_{2-\alpha_{2i-1}}K_{-\rho}^2) - (\alpha_{2i}-1)\mu(G_{\tau_{2i},\alpha_{2i}}J_{2-\alpha_{2i}}K_{-\rho}^2)}{\nu_\rho(\tau_{2i},\alpha_{2i}) - \nu_\rho(\tau_{2i-1},\alpha_{2i-1})}\right\},$$

$$m_B^{(S,i)} = \exp\left\{\frac{\mu(G_{\tau_{2i-1},\alpha_{2i-1}}J_{1-\alpha_{2i-1}}L_{(-\rho,-\beta)}) - \mu(G_{\tau_{2i},\alpha_{2i}}J_{1-\alpha_{2i}}L_{(-\rho,-\beta)})}{\nu_\rho(\tau_{2i},\alpha_{2i}) - \nu_\rho(\tau_{2i-1},\alpha_{2i-1})}\right\},$$

and for $u \in [0,1]$,

$$v^{(S,i)}(u) = \frac{G_{\tau_{2i-1},\alpha_{2i-1}}(u)J_{1-\alpha_{2i-1}}(u) - G_{\tau_{2i},\alpha_{2i}}(u)J_{1-\alpha_{2i}}(u)}{\nu_\rho(\tau_{2i},\alpha_{2i}) - \nu_\rho(\tau_{2i-1},\alpha_{2i-1})}.$$

Let us also consider $m_A^{(S)} = (m_A^{(S,i)},\ i = 1,\dots,4)$ and $m_B^{(S)} = (m_B^{(S,i)},\ i = 1,\dots,4)$. The next result is a direct consequence of Theorem 2 and Lemma 2, see Section 6 for a short proof.

Corollary 2. *Suppose* **(C2)**, **(C3)**, **(C6)** *and* **(C7)** *hold. There exist two intervals J and J_0 such that for all $\rho \in J_0$ and for a sequence k satisfying (6),*

$$k^{1/2} A(n/k)(\hat{\rho}_n^{(S)} - \rho) \xrightarrow{d} \mathcal{N}\left(\frac{\lambda_A}{2\gamma}\mathcal{AB}_1^{(S)}(\delta,\rho) + \lambda_B \mathcal{AB}_2^{(S)}(\delta,\rho,\beta),\ \gamma^2 \mathcal{AV}^{(S)}(\delta,\rho)\right)$$

where

$$\begin{aligned}
\mathcal{AB}_1^{(S)}(\delta,\rho) &= \frac{\varphi_\delta^{(S)}(\rho)}{[\varphi_\delta^{(S)}]'(\rho)} \log \tilde{\psi}_\delta(m_A^{(S)}),\\
\mathcal{AB}_2^{(S)}(\delta,\rho,\beta) &= \frac{\varphi_\delta^{(S)}(\rho)}{[\varphi_\delta^{(S)}]'(\rho)} \log \tilde{\psi}_\delta(m_B^{(S)}),\\
\mathcal{AV}^{(S)}(\delta,\rho) &= \left(\frac{\varphi_\delta^{(S)}(\rho)}{[\varphi_\delta^{(S)}]'(\rho)}\right)^2 \vartheta\left(v^{(S,1)} - v^{(S,2)} - \delta(v^{(S,3)} - v^{(S,4)}), v^{(S,1)} - v^{(S,2)} - \delta(v^{(S,3)} - v^{(S,4)})\right).
\end{aligned}$$

Let us highlight that Proposition 5, Proposition 7 and Proposition 9 of [6] are particular cases of Corollary 2 for three different value of δ ($\delta = 2$, $\delta = 1$ and $\delta = 0$ respectively). The asymptotic normality of the estimators proposed in [19] and in [14] can also be easily established with Corollary 2.

As an example of function $G_{\tau,\alpha}$, one can consider the function defined on $[0,1]$ by:

$$G_{\tau,\alpha}(u) = \frac{g_{\tau-1}(u)}{\int_0^1 g_{\tau-1}(x) J_{-\alpha}(x)dx} \text{ for } \tau \geq 1 \text{ and } \alpha > 0,$$

where the function g_τ is given by

$$g_0(x) = 1, \quad g_{\tau-1}(x) = \frac{\tau}{\tau-1}(1 - x^{\tau-1}), \forall \tau > 1.$$

Clearly, the function $G_{\tau,\alpha}$ satisfies condition **(C7)** and, under **(C6)**, the expression of $\varphi_\rho^{(S)}$ is

$$\varphi_\delta^{(S)}(\rho) = \frac{\zeta_1}{\zeta_2}\left(\frac{\zeta_4}{\zeta_3}\right)^\delta \frac{\nu_\rho(\tau_1,\alpha_1) - \nu_\rho(\tau_2,\alpha_2)}{\nu_\rho(\tau_3,\alpha_3) - \nu_\rho(\tau_4,\alpha_4)} \left[\frac{\nu_\rho(\tau_7,\alpha_7) - \nu_\rho(\tau_8,\alpha_8)}{\nu_\rho(\tau_5,\alpha_5) - \nu_\rho(\tau_6,\alpha_6)}\right]^\delta$$

with

$$\nu_\rho(\tau,\alpha) = \frac{1 - (1-\rho)^{-\alpha} + (\tau-\rho)^{-\alpha} - \tau^{-\alpha}}{\alpha\rho(1 - \tau^{-\alpha-1})} \quad \text{if} \quad \tau \neq 1 \quad \text{and} \quad \nu_\rho(1,\alpha) = \frac{1}{\alpha\rho}\frac{(1-\rho)^\alpha - 1}{(1-\rho)^\alpha}.$$

Even if Corollary 2 ensures the existence of intervals J_0 and J, they are impossible to specify in the general case. In the following, we consider several sets of parameters where these intervals can be easily exhibited and for which the inverse function $\varphi_\delta^{(S)}$ admits an explicit form. To this end, it is assumed that $\tau_2 = \tau_3 = \tau_5 = \tau_6 = \tau_7 = \tau_8 = \alpha_7 = 1$, $\alpha_6 = 3$, $\alpha_8 = 2$ and the following notation is introduced:

$$\omega^*(\delta,\zeta) = \frac{\zeta_1}{\zeta_2}\left(\frac{3\zeta_4}{\zeta_3}\right)^\delta.$$

In all the examples below, $J_0 = \mathbb{R}^-$ and thus the condition $\rho \in J_0$ is always satisfied. The first three examples correspond to existing estimators of the second order parameter while the three last examples give rise to new estimators.

Example 4. *Let $\delta = 0$ (i.e. $\zeta_1 = \zeta_2$), $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1$, $\tau_1 = 2$ and $\tau_4 = 3$. Denoting by $Z_{n,4}^{(S)}$ the rv $Z_n^{(S)}$, the estimator of ρ is given by:*

$$\hat{\rho}_{n,4}^{(S)} = \frac{6(Z_{n,4}^{(S)} + 2)}{3Z_{n,4}^{(S)} + 4} 1\!\!1\{Z_{n,4}^{(S)} \in (-2, -4/3)\}.$$

Note that this estimator corresponds to the estimator $\hat{\rho}_{n,k}^{[2]}$ defined in [6], Section 5.2.

Example 5. *Let $\delta = 0$, $\alpha_1 = \alpha_3 = \alpha_4 = 1$ and $\tau_1 = \tau_4 = \alpha_2 = 2$. Denoting by $Z_{n,5}^{(S)}$ the rv $Z_n^{(S)}$, we find back the estimator $\hat{\rho}_{n,k}^{[3]}$ proposed in [6], Section 5.2:*

$$\hat{\rho}_{n,5}^{(S)} = \frac{2(Z_{n,5}^{(S)} - 2)}{2Z_{n,5}^{(S)} - 1} 1\!\!1\{Z_{n,5}^{(S)} \in (1/2, 2)\}.$$

Example 6. *Let $\alpha_1 = \zeta_1 = 4$, $\alpha_3 = \zeta_2 = \zeta_4 = 2$, $\zeta_3 = 3$ and $\alpha_2 = \alpha_4 = \alpha_5 = \tau_1 = \tau_4 = 1$. These choices entail $\delta = 2$. Denoting by $Z_{n,6}^{(S)}$ the rv $Z_n^{(S)}$, the estimator of ρ given by:*

$$\hat{\rho}_{n,6}^{(S)} = \frac{6Z_{n,6}^{(S)} - 4 + (3Z_{n,6}^{(S)} - 2)^{1/2}}{4Z_{n,6}^{(S)} - 3} 1\!\!1\{Z_{n,6}^{(S)} \in (2/3, 3/4)\}.$$

corresponds to the one proposed in [19], equation (12).

Example 7. *Consider the case $\delta = 1$ (i.e. $\zeta_1 - \zeta_2 = \zeta_3 - \zeta_4$), $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1$, $\tau_1 = \alpha_5 = 2$ and $\tau_4 = 3$. Denoting by $Z_{n,7}^{(S)}$ the rv $Z_n^{(S)}$, a new estimator of ρ is given by:*

$$\hat{\rho}_{n,7}^{(S)} = \frac{6Z_{n,7}^{(S)} + 4\omega^*(1,\zeta)}{3Z_{n,7}^{(S)} + 4\omega^*(1,\zeta)} 1\!\!1\{Z_{n,7}^{(S)} \in \omega^*(1,\zeta) \bullet (-4/3, -2/3)\}.$$

Example 8. *Let $\delta = 1$, $\alpha_1 = \alpha_3 = \alpha_4 = 1$ and $\tau_1 = \tau_4 = \alpha_2 = \alpha_5 = 2$. Denoting by $Z_{n,8}^{(S)}$ the rv $Z_n^{(S)}$, we obtain a new estimator of ρ:*

$$\hat{\rho}_{n,8}^{(S)} = \frac{6Z_{n,8}^{(S)} - 4\omega^*(1,\zeta)}{2Z_{n,8}^{(S)} - \omega^*(1,\zeta)} 1\!\!1\{Z_{n,8}^{(S)} \in \omega^*(1,\zeta) \bullet (1/2, 2/3)\}.$$

Example 9. *Let $\tau_1 = \tau_4 = \alpha_1 = 1$, $\alpha_2 = \alpha_3 = \alpha_5 = 2$ and $\alpha_4 = 3$. Denoting by $Z_{n,9}^{(S)}$ the rv $Z_n^{(S)}$, the estimator of ρ is given by:*

$$\hat{\rho}_{n,9}^{(S)} = \frac{3(Z_{n,9}^{(S)}/(3\omega^*(\delta,\zeta)))^{1/(\delta+1)} - 1}{(Z_{n,9}^{(S)}/(3\omega^*(\delta,\zeta)))^{1/(\delta+1)} - 1} 1\!\!1\{Z_{n,9}^{(S)} \in \omega^*(\delta,\zeta) \bullet (3^{-\delta}, 3)\}.$$

In the particular case where $\delta = 0$, this estimator corresponds to the one proposed in [13].

To summarize, we have illustrated how Theorem 2 may be used to prove the asymptotic normality of estimators built on $T_n^{(R)}$ or $T_n^{(S)}$: Corollary 1 and Corollary 2 cover a large number of estimators proposed in the literature. Five new estimators of ρ have been introduced: $\hat{\rho}_{n,2}^{(R)}$, $\hat{\rho}_{n,3}^{(R)}$, $\hat{\rho}_{n,7}^{(S)}$, $\hat{\rho}_{n,8}^{(S)}$ and $\hat{\rho}_{n,9}^{(S)}$. All of them are explicit and are asymptotically Gaussian. The comparison of their finite sample properties is a huge task since they may depend on their parameters (θ_i, τ_i, α_i) as well as on the simulated distribution. We conclude this study by proposing a method for selecting some "asymptotic optimal" parameters within a family of estimators. The performances and the limits of this technique are illustrated by comparing several estimators on simulated data.

5 Comparison of some estimators

Some estimators of ρ are now compared on a specific Pareto-type model, namely the Burr distribution with cdf $F(x) = 1 - (\zeta/(\zeta + x^\eta))^\lambda$, $x > 0$, $\zeta, \lambda, \eta > 0$, considered for instance in [2], equation (3). The associated extreme-value index is $\gamma = 1/(\lambda\eta)$ and this model satisfies the third order condition **(C2)** with $\rho = \beta = -1/\lambda$, $A(x) = \gamma x^\rho/(1 - x^\rho)$ and $B(x) = \rho x^\rho/(1 - x^\rho)$. We limit ourselves to the case $\zeta = 1$ and $\lambda = 1/\eta$ so that $\gamma = 1$.

5.1 Estimators based on the random variable $R_k(\tau)$

Let us first focus on the estimators of ρ based on the random variables $R_k(\tau_i)$ considered in Section 4.1 with kernel functions $H_{\tau_i}(u) = \tau_i u^{\tau_i - 1}$, for $i = 1, \ldots, 8$. The values of the parameters $\tau_1, ..., \tau_8$, $\tilde{\theta}_1$, $\tilde{\theta}_3$ and $\tilde{\theta}_4$ are taken as in [17, 30]: $\tau_1 = 1.25$, $\tau_2 = \tau_3 = 1.75$, $\tau_4 = \tau_8 = 2$, $\tau_5 = 1.5$, $\tau_6 = \tau_7 = 1.75$, $\tilde{\theta}_1 = 0.01$, $\tilde{\theta}_3 = 0.02$ and $\tilde{\theta}_4 = 0.04$. According to the authors, these values yield good results for distributions satisfying the third order condition **(C2)** with $\beta = \rho$. For these parameters, a simple expression of $\varphi_\delta^{(R)}$ is obtained, see (11), and we have $\delta_0 = 1.5$ and $\delta_1 = 1.8$. Recall that $\tilde{\theta}_2 = \tilde{\theta}_1 + \delta(\tilde{\theta}_4 - \tilde{\theta}_3)$ for $\delta \geq 0$. In the following, we propose to choose the remaining parameter δ using a method similar to the one proposed in [15]. It consists in minimizing with respect to δ an upper bound on the asymptotic mean-squared error. The method is described in Paragraph 5.1.1 and an example of application is presented in Paragraph 5.1.2.

5.1.1 Controlling the asymptotic mean-squared error

As in [17], we assume that $\rho = \beta$. Following Corollary 1, the asymptotic bias components of $\hat{\rho}_n^{(R)}$ are respectively proportional to $\mathcal{AB}_1^{(R)}(\delta, \rho)$ and $\mathcal{AB}_2^{(R)}(\delta, \rho, \rho)$ while its asymptotic variance is proportional to $\mathcal{AV}^{(R)}(\delta, \rho)$. The asymptotic mean-squared error of $\hat{\rho}_n^{(R)}$ can be defined as

$$\mathcal{AMSE}(\delta, \gamma, \rho) = \frac{1}{kA^2(n/k)} \left(\left(\frac{\lambda_A}{2\gamma} \mathcal{AB}_1^{(R)}(\delta, \rho) - \lambda_B \mathcal{AB}_2^{(R)}(\delta, \rho, \rho) \right)^2 + \gamma^2 \mathcal{AV}^{(R)}(\delta, \rho) \right). \quad (12)$$

One way to choose the parameter δ could be to minimize the above asymptotic mean-squared error. In practice, the parameters γ, ρ as well as the functions A and B are unknown and thus the asymptotic mean-squared error cannot be evaluated. To overcome this problem, it is possible to introduce an upper bound on $\mathcal{AMSE}(\delta, \gamma, \rho)$. Assuming that $\delta \in [0, \delta_0) \cup (\delta_1, \infty)$ and $\rho \in [\rho_{\min}, \rho_{\max}]$, it is easy to check that $|\mathcal{AB}_1^{(R)}(\delta, \rho)| \geq |\mathcal{AB}_1^{(R)}(\delta_1, \rho_{\max})|$, $|\mathcal{AB}_2^{(R)}(\delta, \rho, \rho)| \geq |\mathcal{AB}_2^{(R)}(\delta_0, \rho_{\min}, \rho_{\min})|$ and, numerically, we can see that $\mathcal{AV}^{(R)}(\delta,\rho) \geq \mathcal{AV}^{(R)}(1.32,-0.46)$. We thus have:

$$\mathcal{AMSE}(\delta,\gamma,\rho) \leq \frac{C\pi(\delta,\rho)}{kA^2(n/k)},$$

with $\pi(\delta,\rho) = (\mathcal{AB}_1^{(R)}(\delta,\rho)\mathcal{AB}_2^{(R)}(\delta,\rho,\rho))^2\mathcal{AV}^{(R)}(\delta,\rho)$ and where the constant C does not depend on δ and ρ. We thus consider for $\rho < 0$ the parameter δ minimizing the function $\pi(\delta,\rho)$. For instance, when ρ is in the neighborhood of 0, one can show that the optimal value is $\delta = \delta_0 = 1.5$.

5.1.2 Illustration on the Burr distribution

Three estimators are compared:

- the estimator $\hat{\rho}_{n,1}^{(R)}$ proposed in [17], and which corresponds to the case $\delta = 1$, see Example 1,
- the new explicit estimator $\hat{\rho}_{n,2}^{(R)}$ introduced in Example 2 which corresponds to the case $\delta = 0$,
- the new implicit estimator defined by $\hat{\rho}_{n,0}^{(R)} := \hat{\rho}_n^{(R)}$ with $\delta = \delta_0 = 1.5$, see equation (10).

First, the estimators are compared on the basis of their asymptotic mean-squared errors. Taking $\lambda_A = k^{1/2}A^2(n/k)$ and $\lambda_B = \rho\lambda_A/\gamma$, the asymptotic mean-squared errors are plotted on the left panel of Figure 1 as a function of $k \in \{1500,\dots,4999\}$ with $n = 5000$ and for $\rho \in \{-1,-0.25\}$. It appears that $\hat{\rho}_{n,0}^{(R)}$ yields the best results for $\rho = -1$. This is in accordance with the results from the previous paragraph: $\delta = 1.5$ is the "optimal" when ρ is close to 0. As a preliminary conclusion, the criterion $\pi(.)$ seems to be well-adapted for tuning the estimator parameters. At the opposite, when $\rho = -0.25$, the best estimator from the asymptotic mean-squared error point of view is $\hat{\rho}_{n,2}^{(R)}$.

Second, the estimators are compared on their finite sample size performances. For each estimator, and for each value of $k \in \{1500,\dots,4999\}$, the empirical mean-squared error is computed on 500 replications of the sample of size $n = 5000$. The results are displayed on the right panel of Figure 1. The conclusions are qualitatively the same: $\hat{\rho}_{n,0}^{(R)}$ yields the best results in the case $\rho \geq -1$ where as $\hat{\rho}_{n,2}^{(R)}$ yields the best results in the case $\rho < -1$. Let us note that, consequently, $\hat{\rho}_{n,1}^{(R)}$ is never the best estimator in the situation considered here. In practice, the case $\rho \geq -1$ is the more interesting one, since it corresponds to a strong bias. For this reason, it seems to us that $\hat{\rho}_{n,0}^{(R)}$ should be preferred.

5.2 Estimators based on the random variable $S_k(\tau,\alpha)$

Let us now consider the estimators of ρ based on the random variables $S_k(\tau_i,\alpha_i)$ for $i = 1,\dots,8$ considered in Section 4.2 in the case where $(\tau_1,\alpha_1) = (\tau_7,\alpha_7)$, $(\tau_2,\alpha_2) = (\tau_8,\alpha_8)$, $(\tau_3,\alpha_3) = (\tau_5,\alpha_5)$ and $(\tau_4,\alpha_4) = (\tau_6,\alpha_6)$. In Paragraph 5.2.1, we show that the asymptotic mean-squared error is independent of δ. In contrast, Paragraph 5.2.2 illustrates the finite sample behavior of the estimators when δ varies.

5.2.1 Comparison in terms of asymptotic mean-squared error

From Corollary 2, the asymptotic bias and variance components of $\hat{\rho}_n^{(S)}$ are respectively proportional to

$$
\begin{aligned}
\mathcal{AB}_1^{(S)}(\delta,\rho) &= \frac{g^{(S)}(\rho)}{(g^{(S)})'(\rho)}\left(\log m_A^{(S,1)} - \log m_A^{(S,2)}\right),\\
\mathcal{AB}_2^{(S)}(\delta,\rho) &= \frac{g^{(S)}(\rho)}{(g^{(S)})'(\rho)}\left(\log m_B^{(S,1)} - \log m_B^{(S,2)}\right),\\
\mathcal{AV}^{(S)}(\delta,\rho) &= \left(\frac{g^{(S)}(\rho)}{(g^{(S)})'(\rho)}\right)^2 \vartheta(v^{(S,1)} - v^{(S,2)}, v^{(S,1)} - v^{(S,2)}),
\end{aligned}
$$

where

$$
g^{(S)}(\rho) = \frac{\zeta_1}{\zeta_2}\frac{\mu(G_{\tau_1,\alpha_1}J_{1-\alpha_1}K_{-\rho}) - \mu(G_{\tau_2,\alpha_2}J_{1-\alpha_2}K_{-\rho})}{\mu(G_{\tau_3,\alpha_3}J_{1-\alpha_3}K_{-\rho}) - \mu(G_{\tau_4,\alpha_4}J_{1-\alpha_4}K_{-\rho})}.
$$

It thus appears that the asymptotic mean-squared error (defined similarly to (12)) does not depend on δ. From the asymptotic point of view, all the estimators $\hat{\rho}_n^{(S)}$ such that $(\tau_1,\alpha_1) = (\tau_7,\alpha_7)$, $(\tau_2,\alpha_2) = (\tau_8,\alpha_8)$, $(\tau_3,\alpha_3) = (\tau_5,\alpha_5)$ and $(\tau_4,\alpha_4) = (\tau_6,\alpha_6)$ are thus equivalent.

5.2.2 Comparison on the simulated Burr distribution

For the sake of simplicity, we fix $\alpha_1 = \alpha_7 = \theta_5 = \theta_7 = \tau_1 = \cdots = \tau_8 = 1$, $\alpha_2 = \alpha_3 = \alpha_5 = \alpha_8 = 2$, $\alpha_4 = \alpha_6 = 3$, $\theta_3 = \theta_8 = 1/2$, $\theta_4 = 1/3$, $\theta_6 = 2/3$, $\theta_1 = \delta + 1$ and $\theta_2 = (\delta+1)/2$ so that δ is the unique free parameter. The resulting estimator is $\hat{\rho}_{n,9}^{(S)}$, it coincides with the one proposed in [13] when $\delta = 0$. For each value of $k \in \{500,\dots,4999\}$, the empirical mean-squared error associated to $\hat{\rho}_{n,9}^{(S)}$ is computed on 500 replications of the sample of size $n = 5000$ for $\delta \in \{0,1,2\}$ and for $\rho \in \{-0.25,-1\}$. The results are displayed on Figure 2. It appears that $\delta = 0$ yields the best results for both values of ρ: the empirical mean-squared error is smaller than these associated to $\delta = 1$ or $\delta = 2$. This hierarchy cannot be observed on the asymptotic mean-squared error.

5.3 Tentative conclusion

The families of estimators of the second order parameter usually depend on a large set, say Θ, of parameters (12 parameters for estimators based on the random variables $R_k(\tau)$ and 20 parameters for $S_k(\tau,\alpha)$). The methodology proposed in Paragraph 5.1.1 permits to compute an upper bound $\pi(.)$ on the asymptotic mean-squared error $\mathcal{AMSE}$ associated to the estimators. This requires to show that the quantities $\mathcal{AB}_1$, $\mathcal{AB}_2$ and $\mathcal{AV}$ are lower bounded when Θ varies in some region R_Θ. Thus, it may be possible, for some well chosen region R_Θ, to find an "optimal" set of parameters minimizing $\pi(.)$. Unfortunately, the $\mathcal{AMSE}$ may not depend on all the parameters in Θ (see Paragraph 5.2.1) whereas the finite sample performances of the estimator does (see Paragraph 5.2.2). In such a case, the definition of a criterion for selecting an optimal Θ is an open question.

6 Proofs

Proof of Theorem 1. Clearly, **(Ψ1)** and **(Ψ2)** entail $Z_n = \psi(\omega_n^{-1}(T_n - \chi_n\mathbb{I}))$. Moreover, **(T1)** and **(Ψ4)** yield $Z_n \xrightarrow{\mathbb{P}} \psi(f(\rho)) = \varphi(\rho)$. For all $\varepsilon > 0$, we have

$$\begin{aligned}\mathbb{P}(|\hat{\rho}_n - \rho| > \varepsilon) &= \mathbb{P}(\{|\hat{\rho}_n - \rho| > \varepsilon\} \cap \{Z_n \in J\}) + \mathbb{P}(\{|\hat{\rho}_n - \rho| > \varepsilon\} \cap \{Z_n \notin J\}) \\ &\leq \mathbb{P}(\{|\hat{\rho}_n - \rho| > \varepsilon\} \cap \{Z_n \in J\}) + \mathbb{P}(\{Z_n \notin J\}) \\ &= \mathbb{P}(\{|\varphi^{-1}(Z_n) - \rho| > \varepsilon\} \cap \{Z_n \in J\}) + \mathbb{P}(\{Z_n \notin J\}).\end{aligned}$$

From **(Ψ3)** and **(Ψ4)**, φ^{-1} is also continuous in a neighborhood of $\varphi(\rho)$. Since $Z_n \xrightarrow{\mathbb{P}} \varphi(\rho)$, it follows that $\mathbb{P}(\{|\varphi^{-1}(Z_n) - \rho| > \varepsilon\} \cap \{Z_n \in J\}) \to 0$ as $n \to \infty$. Besides, $\rho \in J_0$ yields $\varphi(\rho) \in J$ and thus

$$\mathbb{P}(\{Z_n \notin J\}) \to 0 \text{ as } n \to \infty. \tag{13}$$

As a conclusion, $\mathbb{P}(|\hat{\rho}_n - \rho| > \varepsilon) \to 0$ as $n \to \infty$ and the result is proved. ■

Proof of Theorem 2. Recalling that $Z_n = \psi(\omega_n^{-1}(T_n - \chi_n\mathbb{I}))$, a first order Taylor expansion shows that there exists $\varepsilon \in (0, 1)$ such that

$$v_n(Z_n - \varphi(\rho)) = {}^t(v_n\xi_n)\, \nabla\psi(f(\rho) + \varepsilon\xi_n),$$

where we have defined $\xi_n = \omega_n^{-1}(T_n - \chi_n\mathbb{I}) - f(\rho)$. Therefore, $\xi_n \xrightarrow{\mathbb{P}} 0$ and **(Ψ5)** entail that $\nabla\psi(f(\rho) + \varepsilon\xi_n) \xrightarrow{\mathbb{P}} \nabla\psi(f(\rho))$. Thus, taking account of **(T2)**, we obtain that

$$v_n(Z_n - \varphi(\rho)) \xrightarrow{d} \mathcal{N}(m_\psi(\rho), \gamma^2\sigma_\psi^2(\rho)). \tag{14}$$

Now, $P_n(x) := \mathbb{P}(\{v_n(\hat{\rho}_n - \rho) \leq x\})$ can be rewritten as

$$\begin{aligned}P_n(x) &= \mathbb{P}(\{v_n(\hat{\rho}_n - \rho) \leq x\} \cap \{Z_n \in J\}) + \mathbb{P}(\{v_n(\hat{\rho}_n - \rho) \leq x\} \cap \{Z_n \notin J\}) \\ &= \mathbb{P}(\{v_n(\varphi^{-1}(Z_n) - \rho) \leq x\} \cap \{Z_n \in J\}) + \mathbb{P}(\{v_n(\hat{\rho}_n - \rho) \leq x\} \cap \{Z_n \notin J\}) \\ &=: P_{1,n}(x) + P_{2,n}(x).\end{aligned}$$

Let us first note that

$$0 \leq P_{2,n}(x) \leq \mathbb{P}(\{Z_n \notin J\}) \to 0 \text{ as } n \to \infty, \tag{15}$$

in view of (13) in the proof of Theorem 1. Focusing on $P_{1,n}(x)$, since φ is continuously differentiable in a neighborhood of ρ and $\varphi'(\rho) \neq 0$, it follows that φ is monotone in a neighborhood of ρ. Let us consider the case where φ is decreasing, the case φ increasing being similar. Writing $J = (a, b)$, it follows that

$$\begin{aligned}P_{1,n}(x) &= \mathbb{P}(\{a \vee \varphi(\rho + x/v_n) \leq Z_n \leq b\}) \\ &= \mathbb{P}(\{v_n(a \vee \varphi(\rho + x/v_n) - \varphi(\rho)) < v_n(Z_n - \varphi(\rho)) \leq v_n(b - \varphi(\rho))\}).\end{aligned}$$

Introducing G_n the cumulative distribution function of $v_n(Z_n - \varphi(\rho))$, we have

$$\begin{aligned}
1 - P_{1,n}(x) &= 1 - G_n(v_n(b - \varphi(\rho))) + G_n(v_n(a \vee \varphi(\rho + x/v_n) - \varphi(\rho))) \\
&= 1 - G_n(v_n(b - \varphi(\rho))) + G_n(v_n(a - \varphi(\rho))) \vee G_n(v_n(\varphi(\rho + x/v_n) - \varphi(\rho))) \\
&=: P_{1,1,n} + P_{1,2,n} \vee P_{1,3,n}(x).
\end{aligned}$$

Let G denote the cumulative distribution function of the $\mathcal{N}(m_\psi(\rho), \gamma^2\sigma_\psi^2(\rho))$ distribution. It is straightforward that

$$P_{1,1,n} \leq 1 - G(v_n(b - \varphi(\rho))) + \sup_{t \in \mathbb{R}} |G_n(t) - G(t)|.$$

Since $\rho \in J_0$, we have $\varphi(\rho) \in J = (a, b)$. In particular, $b > \varphi(\rho)$ yields $1 - G(v_n(b - \varphi(\rho))) \to 0$ as $n \to \infty$. Besides, (14) shows that $G_n(t) \to G(t)$ for all $t \in \mathbb{R}$ and thus $G_n(t) \to G(t)$ uniformly, see for instance [11], p.552. As a preliminary conclusion $P_{1,1,n} \to 0$ and, similarly, $P_{1,2,n} \to 0$ as $n \to \infty$. Finally,

$$|P_{1,3,n}(x) - G(x\varphi'(\rho))| \leq |G(v_n(\varphi(\rho + x/v_n) - \varphi(\rho)) - G(x\varphi'(\rho))| + \sup_{t \in \mathbb{R}} |G_n(t) - G(t)|$$

and, in view of **(Ψ5)**, $v_n(\varphi(\rho + x/v_n) - \varphi(\rho)) \to x\varphi'(\rho)$ as $n \to \infty$, which leads to $P_{1,3,n}(x) \to G(x\varphi'(\rho))$ as $n \to \infty$. We thus have shown that

$$P_{1,n}(x) \to 1 - G(x\varphi'(\rho)) = G(x|\varphi'(\rho)|) \text{ as } n \to \infty. \tag{16}$$

Collecting (15) and (16) yields

$$\mathbb{P}(\{v_n(\hat{\rho}_n - \rho) \leq x\}) \to G(x|\varphi'(\rho)|) \text{ as } n \to \infty$$

and concludes the proof. ■

Proof of Corollary 1. Clearly, ψ_δ given in (9) satisfies **(Ψ1)** and **(Ψ2)**. Moreover, Lemma 1 shows that **(T2)** holds. To apply Theorem 2 it only remains to prove that **(Ψ3)** and **(Ψ5)** are satisfied. First remark that under **(C4)** and **(C5)**, $\varphi_\delta^{(R)}(\rho)$ is well defined for all $\rho \leq 0$ since $f^{(R)}(\rho) \in \mathcal{D}$. Furthermore, from Lemma 1, we have for $i = 1, \ldots, 4$,

$$T_{n,2i-1}^{(R)} - T_{n,2i}^{(R)} = \frac{\tilde{\theta}_i A(Y_{n-k,n})}{\gamma}(\nu_\rho(\tau_{2i-1}) - \nu_\rho(\tau_{2i}))(1 + o_P(1)),$$

as n goes to infinity. Hence, conditions **(C4)** and **(C5)** imply that $T_n^{(R)} \in \mathcal{D}$. Finally, using Lerch's Theorem (see [5], page 345), condition **(C4)** implies that there exists $\rho_0 < 0$ such that the first derivative of $\varphi_\delta^{(R)}$ is non zero at ρ_0. Thus, the inverse function theorem insures the existence of intervals J_0 and J for which the function $\varphi_\delta^{(R)}$ is a continuously differentiable bijection from J_0 to J. In conclusion, conditions **(Ψ3)** and **(Ψ5)** are satisfied and Theorem 2 applies. ■

Proof of Corollary 2. The proof follows the same lines as the one of Corollary 1. It consists in remarking that, under **(C6)** and **(C7)**, one has $f^{(S)}(\rho) \in \mathcal{D}$ and $T_n^{(S)} \in \mathcal{D}$ since,

$$T_{n,2i-1}^{(S)} - T_{n,2i}^{(S)} = \frac{\zeta_i A(n/k)}{\gamma} \left(\nu_\rho(\tau_{2i}, \alpha_{2i}) - \nu_\rho(\tau_{2i-1}, \alpha_{2i-1})\right)(1 + o_P(1)),$$

in view of Lemma 2. ∎

References

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

[2] J. Beirlant, Y. Goegebeur, R. Verlaack and P. Vynckier. Burr regression and portfolio segmentation. *Insurance: Mathematics and Economics*, **23**, 231–250, 1998.

[3] J. Cai, L. de Haan and C. Zhou. Bias correction in extreme value statistics with index around zero. *Extremes*, to appear, DOI: 10.1007/s10687-012-0158-x, 2012.

[4] F. Cairo, M.I. Gomes and D. Pestana. Direct reduction of bias of the classical Hill estimator. *REVSTAT - Statistical Journal*, **3(2)**, 113–136, 2005.

[5] H.S. Carslaw and J.C. Jaeger. *Operational Methods in Applied Mathematics*, Oxford University Press, 1948.

[6] G. Ciuperca and C. Mercadier. Semi-parametric estimation for heavy tailed distributions. *Extremes*, **13**, 55–87, 2010.

[7] S. Csörgö, P. Deheuvels, and D.M. Mason. Kernel estimates of the tail index of a distribution. *The Annals of Statistics*, **13**, 1050–1077, 1985.

[8] A.L.M. Dekkers, J.H.J. Einmahl, and L. de Haan. A moment estimator for the index of an extreme-value distribution. *The Annals of Statistics*, **17**, 1833–1855, 1989.

[9] J. Diebolt, L. Gardes, S. Girard and A. Guillou. Bias-reduced estimators of the Weibull tail-coefficient. *Test*, **17(2)**, 311–331, 2008.

[10] J. Diebolt, L. Gardes, S. Girard and A. Guillou. Bias-reduced extreme quantile estimators of Weibull tail-distributions. *Journal of Statistical Planning and Inference*, **138(5)**, 1389–1401, 2008.

[11] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling extremal events*, Springer, 1997.

[12] A. Feuerverger and P. Hall. Estimating a tail exponent by modeling departure from a Pareto distribution. *The Annals of Statistics*, **27**, 760–781, 1999.

[13] M.I. Fraga Alves, M.I. Gomes, and L. de Haan. A new class of semi-parametric estimators of the second order parameter. *Portugaliae Mathematica*, **60**(2), 193–213, 2003.

[14] M.I. Fraga Alves, L. de Haan and T. Lin. Estimation of the parameter controlling the speed of convergence in extreme value theory. *Mathematical Methods of Statistics*, **12(2)**, 155–176, 2003.

[15] L. Gardes and S. Girard. Conditional extremes from heavy-tailed distributions: An application to the estimation of extreme rainfall return levels. *Extremes*, 13(2):177–204, 2010.

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

[17] Y. Goegebeur, J. Beirlant, and T. de Wet. Kernel estimators for the second order parameter in extreme value statistics. *Journal of Statistical Planning and Inference*, **140**, 2632–2652, 2010.

[18] Y. Goegebeur and T. de Wet. Estimation of the third-order parameter in extreme value statistics. *Test*, **21(2)**, 330–354, 2012.

[19] M.I. Gomes, L. de Haan and L. Peng. Semi-parametric estimation of the second order parameter in statistics of extreme. *Extremes*, **5**, 387–414, 2002.

[20] M.I. Gomes and M.J. Martins. Generalizations of the Hill estimator - asymptotic versus finite sample behaviour. *Journal of Statistical Planning and Inference*, **93**, 161–180, 2001.

[21] M.I. Gomes and M.J. Martins. "Asymptotically unbiased" estimators of the tail index based on external estimation of the second order parameter. *Extremes*, **5(1)**, 5–31, 2002.

[22] M.I. Gomes and M.J. Martins. Bias reduction and explicit semi-parametric estimation of the tail index. *Journal of Statistical Planning and Inference*, **124**, 361–378, 2004.

[23] M.I. Gomes, M.J. Martins and M. Neves. Improving second order reduced bias extreme value index estimator. *REVSTAT - Statistical Journal*, **5(2)**, 177–207, 2007.

[24] M.I. Gomes and O. Oliveira. The bootstrap methodology in statistics of extremes - Choice of the optimal sample fraction. *Extremes*, **4(4)**, 331–358, 2001.

[25] L. de Haan and A. Ferreira. *Extreme Value Theory: An Introduction*, Springer Series in Operations Research and Financial Engineering, Springer, 2006.

[26] P. Hall and A.H. Welsh. Adaptative estimates of parameters of regular variation. *The Annals of Statistics*, **13**, 331–341, 1985.

[27] B.M. Hill. A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, **3**, 1163–1174, 1975.

[28] L. Peng. Asymptotic unbiased estimators for the extreme value index. *Statistics and Probability Letters*, **38**, 107–115, 1998.

[29] J. Segers. Residual estimators. *Journal of Statistical Planning and Inference*, **98**, 15–27, 2001.

[30] T. de Wet, Y. Goegebeur, and M.R. Munch. Asymptotically unbiased estimation of the second order tail parameter. *Statistics and Probability Letters*, **82**, 565–573, 2012.

[31] J. Worms and R. Worms. Estimation of second order parameters using probability weighted moments. *ESAIM: Probability and Statistics*, **16**, 97–113, 2012.

Acknowledgements

The first author acknowledges support from AIRES-Sud (AIRES-Sud is a program from the French Ministry of Foreign and European Affairs, implemented by the "Institut de Recherche pour le Développement", IRD-DSF) and from the "Ministère de la Recherche Scientifique" of Sénégal.

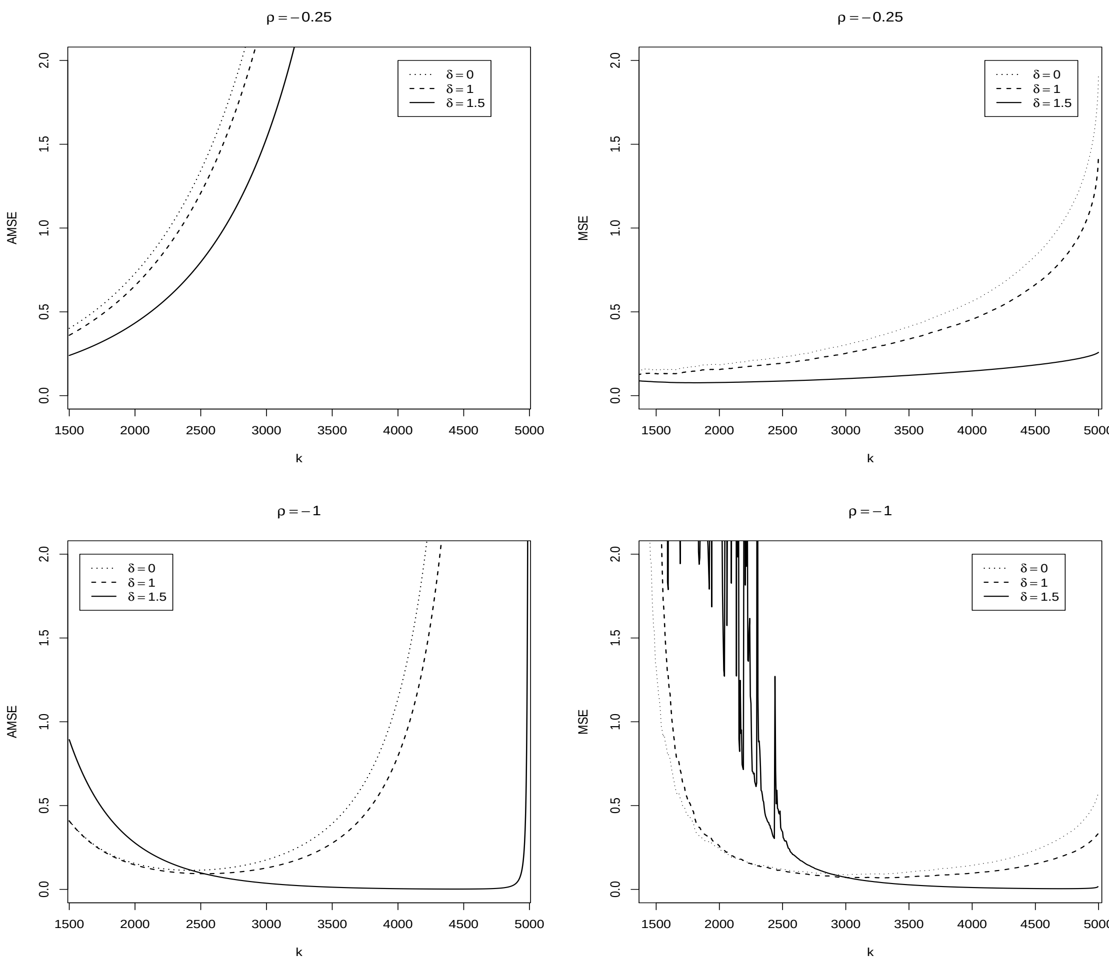

Figure 1: Asymptotic mean-squared errors (left) and empirical mean-squared errors (right) of $\hat{\rho}_{n,0}^{(R)}$, $\hat{\rho}_{n,1}^{(R)}$ and $\hat{\rho}_{n,2}^{(R)}$ as a function of k for a Burr distribution.

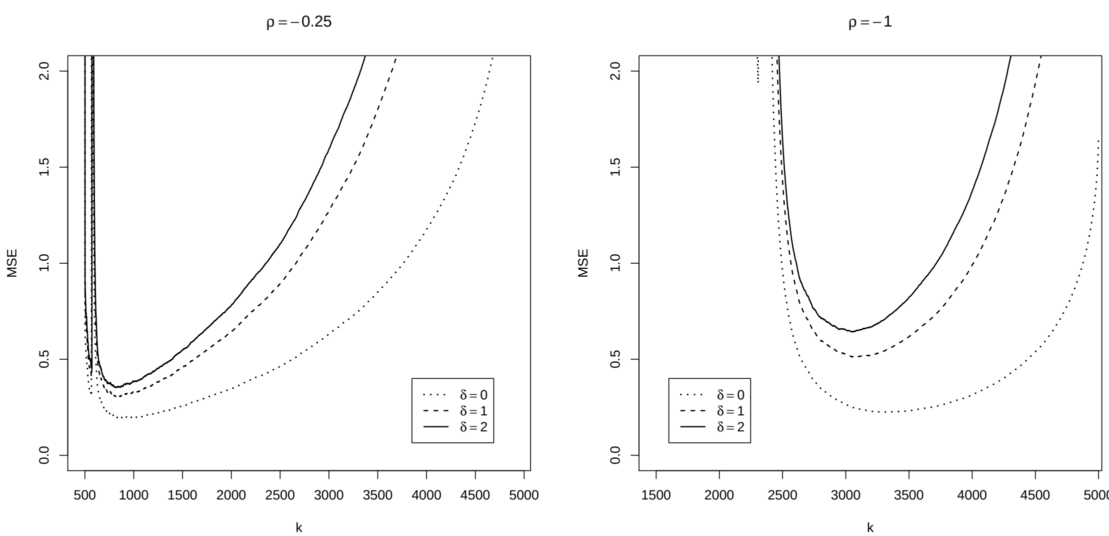

Figure 2: Empirical mean-squared errors of $\hat{\rho}_{n,9}^{(S)}$ as a function of k for a Burr distribution.