ESAIM: PS 16 (2012) 97–113
DOI: 10.1051/ps/2010017

ESAIM: Probability and Statistics
www.esaim-ps.org

ESTIMATION OF SECOND ORDER PARAMETERS USING PROBABILITY WEIGHTED MOMENTS

Julien Worms[1] and Rym Worms[2]

Abstract. The P.O.T. method (Peaks Over Threshold) consists in using the generalized Pareto distribution (GPD) as an approximation for the distribution of excesses over a high threshold. In this work, we use a refinement of this approximation in order to estimate second order parameters of the model using the method of probability-weighted moments (PWM): in particular, this leads to the introduction of a new estimator for the second order parameter ρ, which will be compared to other recent estimators through some simulations. Asymptotic normality results are also proved. Our new estimator of ρ looks especially competitive when $|\rho|$ is small.

Mathematics Subject Classification. 62G32, 60G70.

Received February 27, 2009. Revised April 16, 2010.

1. Introduction

In statistical extreme value theory, one is often interested by the estimation of the far tail of a distribution. The quality of this estimation especially depends on knowledge about the so-called tail index $\gamma = \gamma(F)$ of the underlying model F, which is the shape parameter of the generalized Pareto distribution (GPD) with distribution function (d.f.)

$$G_{\gamma,\sigma}(x) = \begin{cases} 1-\Big(1 + \frac{\gamma x}{\sigma}\Big)^{-\frac{1}{\gamma}}, & \text{for } \gamma \neq 0 \\ 1 - \exp\Big(-\frac{x}{\sigma}\Big), & \text{for } \gamma = 0. \end{cases}$$

The GPD appears as the limiting d.f. of excesses over a high threshold u defined for $x \geq 0$ by

$$F_u(x) := \mathbb{P}(X - u \leq x \,|\, X > u), \quad \text{where } X \text{ has d.f. } F.$$

Keywords and phrases. Extreme values, domain of attraction, excesses, generalized Pareto distribution, probability-weighted moments, second order parameter, third order condition.

[1] Université de Versailles Saint-Quentin, Laboratoire de Mathématiques de Versailles (CNRS UMR 8100), UFR de Sciences, Bât. Fermat, 45 Av. des Etats-Unis, 78035 Versailles Cedex, France. worms@math.uvsq.fr

[2] Université Paris Est Créteil, Laboratoire d'Analyse et de Mathématiques Appliquées (CNRS UMR 8050), 61 Av. du G^l de Gaulle, 94010 Créteil Cedex, France. rym.worms@univ-paris12.fr

Article published by EDP Sciences

© EDP Sciences, SMAI 2012

https://doi.org/10.1051/ps/2010017 Published online by Cambridge University Press

It was established in Pickands' and Balkema and de Haan's results (see [17] and [1]) that F is in the domain of attraction of an extreme value distribution with shape parameter γ if and only if

$$\lim_{u\to s_+(F)} \sup_{0<x<s_+(F)-u} \left| F_u(x) - G_{\gamma,\sigma(u)}(x) \right| = 0$$

for some positive scaling function $\sigma(u)$ depending on u, where $s_+(F) = \sup\{x : F(x) < 1\}$. Since the far tail of the unknown underlying distribution F is closely tied to the d.f. of excesses over a high threshold, accurate modelisation of the distribution of excesses is an important topic.

In what follows, we suppose that F is twice differentiable and that its inverse F^{-1} exists. Let V and A be the two functions defined by

$$V(t) = \overline{F}^{-1}(\mathrm{e}^{-t}) \quad \text{and} \quad A(t) = \frac{V''(\ln t)}{V'(\ln t)} - \gamma.$$

We suppose the following first and second order conditions hold (RV_ρ below stands for the set of regularly varying functions with coefficient of variation ρ):

$$\lim_{t\to+\infty} A(t) = 0 \tag{1.1}$$

$$A \text{ is of constant sign at } \infty \text{ and there exists } \rho \le 0 \text{ such that } |A| \in RV_\rho. \tag{1.2}$$

Under these assumptions, it is proved in [18] that if (u_n) is a sequence of thresholds such that $u_n \to s_+(F)$ as $n \to \infty$, then we have the following development

$$\overline{F}_{u_n}(\sigma_n y) - \overline{G}_\gamma(y) = a_n D_{\gamma,\rho}(y) + o(a_n), \text{ as } n \to +\infty, \tag{1.3}$$

for all y, where $\overline{G}_\gamma(y) := 1 - G_{\gamma,1}(y)$,

$$\sigma_n := \sigma(u_n) = V'\left(V^{-1}(u_n)\right), \quad a_n := A\left(\mathrm{e}^{V^{-1}(u_n)}\right),$$

$$D_{\gamma,\rho}(y) := \begin{cases} C_{0,\rho}(y), & \text{if } \gamma = 0, \\ C_{\gamma,\rho}\left(\frac{1}{\gamma}\ln(1+\gamma y)\right) & \text{if } \gamma \neq 0, \end{cases}$$

and

$$C_{\gamma,\rho}(y) := \mathrm{e}^{-(1+\gamma)y} I_{\gamma,\rho}(y) \quad \text{and} \quad I_{\gamma,\rho}(y) := \int_0^y \mathrm{e}^{\gamma u} \int_0^u \mathrm{e}^{\rho s} \mathrm{d}s\mathrm{d}u.$$

The idea of the present work is that, according to the result (1.3), $\overline{G}_{\gamma,\sigma(u)}(x) + a_n D_{\gamma,\rho}(x/\sigma(u))$ is a better approximation of $\overline{F}_u(x)$ than $\overline{G}_{\gamma,\sigma(u)}(x)$ alone: this is the starting point of our method for the estimation of the second order parameters a_n and ρ.

The estimation of ρ is of great importance (for instance for the determination of the optimal sample fraction needed in the estimation of the tail index or of high quantiles) and has been studied by several authors during the last 15 years. Many of the existing estimators of ρ are based on functionals of the moment statistics $M_n^{(j)}(k_n) = {k_n}^{-1} \sum_{i=1}^{k_n} \left(\ln X_{n-i+1:n} - \ln X_{n-k_n:n}\right)^j$, where $X_{i:n}$ denotes the ith ascending order statistic associated to a sample $(X_1, \ldots, X_n)$ of d.f. F, and k_n is the number of excesses retained for the estimation (where $k_n \to \infty$ but slower than n). We can cite those introduced in [3,7–9,13,14,16].

The estimation of a_n can also be very useful. For instance, if we consider the estimation of the tail index γ by the PWM estimator, it was proved in [4] that the main component of the bias of this estimator is of order a_n. An estimation of the latter parameter could thus be used to reduce this bias. Moreover, it was proved in [21]

https://doi.org/10.1051/ps/2010017 Published online by Cambridge University Press

that, in the case $\rho = 0$, the GPD $G_{\gamma+a_n,\sigma_n}$ is a better approximation of the distribution of the excesses F_u than G_{γ,σ_n}; this is called the penultimate approximation, and the estimation of a_n is important in this framework.

In this work, we use the probability-weighted moments (PWM) techniques introduced by Hosking and Wallis in [15] to estimate the second order parameters ρ and a_n, as well as the scale parameter σ_n. The proposed estimators are based on an "external" estimation of γ: a similar procedure was undertaken in [8].

Under conditions (1.1) and (1.2), it is known (see [18]) that

$$\forall x \in \mathbb{R},\ \lim_{t\to+\infty} \frac{\frac{V(t+x)-V(t)}{V'(t)} - \int_0^x \mathrm{e}^{\gamma s}\,\mathrm{d}s}{A(\mathrm{e}^t)} = I_{\gamma,\rho}(x). \tag{1.4}$$

In order to achieve asymptotic normality results, we will need the following third order condition which specifies the rate of convergence in (1.4):

$$\forall x \in \mathbb{R},\ \lim_{t\to+\infty} \frac{1}{B(\mathrm{e}^t)} \left(\frac{\frac{V(t+x)-V(t)}{V'(t)} - \int_0^x \mathrm{e}^{\gamma s}\,\mathrm{d}s}{A(\mathrm{e}^t)} - I_{\gamma,\rho}(x) \right) = R_{\gamma,\rho,\beta}(\mathrm{e}^x), \tag{1.5}$$

where

$$R_{\gamma,\rho,\beta}(\mathrm{e}^x) := \int_0^x \mathrm{e}^{\gamma s} \int_0^s \mathrm{e}^{\rho z} \int_0^z \mathrm{e}^{\beta y} \mathrm{d}y\mathrm{d}z\mathrm{d}s$$

and the function B tends to 0 and is of constant sign at ∞ and $|B| \in RV_\beta$, for some $\beta \leq 0$. This condition has been introduced in [8] and studied in more details in [10].

Remark 1.1. We can choose, in our regular case, $B(t) := \frac{tA'(t)}{A(t)} - \rho$ (F should then be three times differentiable).

In Section 2, we introduce the new model based on (1.3) and the associated probability-weighted moments and establish the asymptotic normality of their estimators. In Section 3, we present our estimators for ρ, a_n and σ_n and establish their asymptotic normality, first when γ is supposed to be known and then for the unknown γ case. Section 4 contains some simulations illustrating the behaviour of our new estimator of ρ, by comparison to two other recent estimators.

2. Estimators for the probability-weighted moments

2.1. Definition of the probability-weighted moments

In [15], Hosking and Wallis introduced the PWM method in order to define estimators of γ and σ_n based on a sample with d.f. supposed to be an exact GPD. These estimators were obtained through a substitution method based on the following quantities, the probability-weighted moments

$$\nu_j = \mathbb{E}(X\overline{G}^j_{\gamma,\sigma_n}(X))$$

where $j \in \{0, 1\}$ and X has d.f. G_{γ,σ_n}. The results were generalized in [4] to the case where the sample was only supposed to be in the domain of attraction of a GPD.

In this work, more parameters are considered, and we note $\theta_n = (\gamma, \sigma_n, a_n, \rho)$. According to the asymptotic result (1.3), we define our extended model by the distribution function

$$B_{\theta_n}(x) = G_{\gamma,\sigma_n}(x) - a_n D_{\gamma,\rho}\left(\frac{x}{\sigma_n}\right),\ \text{for all } x,$$

and consider the corresponding first three PWM as follows, where X has d.f. B_{θ_n}

$$\tilde{v}_j = \mathbb{E}(X\overline{B}^j_{\theta_n}(X)),\ \text{for } j \in \{0, 1, 2\}.$$

https://doi.org/10.1051/ps/2010017 Published online by Cambridge University Press

It is easy to see that

$$\tilde{v}_j = \int_0^{+\infty} \frac{(1 - B_{\theta_n}(x))^{j+1}}{j+1}\, \mathrm{d}x.$$

Note that for all the PWM and their estimators, the subscript n is ommited in order to simplify the notations. The following lemmas provide expressions of these PWM as functions of the parameters.

Lemma 2.1. *For $j \in \{0,1,2\}$, $\rho \leq 0$ and $-1 < \gamma < 1$,*

$$\nu_j = \frac{\sigma_n}{(j+1)(j+1-\gamma)}.$$

Lemma 2.2.

$$\tilde{v}_0 = \nu_0 + a_n \int_0^{+\infty} D_{\gamma,\rho}\left(\frac{x}{\sigma_n}\right) \mathrm{d}x := v_0,$$

and, for $j \in \{1,2\}$,

$$\tilde{v}_j = \nu_j + a_n \int_0^{+\infty} \overline{G}^j_{\gamma,\sigma_n}(x)\; D_{\gamma,\rho}\left(\frac{x}{\sigma_n}\right) \mathrm{d}x + o(a_n) := v_j + o(a_n).$$

Lemma 2.3. *For $j \in \{0,1,2\}$, $\rho \leq 0$ and $-1 < \gamma < 1$, we have $\tilde{v}_j = v_j + o(a_n)$ where*

$$v_j := \frac{\sigma_n}{(j+1)(j+1-\gamma)} + \frac{a_n \sigma_n}{u_j} \quad \textit{and} \quad u_j := (j+1)(j+1-\gamma)(j+1-\gamma-\rho).$$

In the sequel, we will use the quantities v_0, v_1, v_2 (rather than $\tilde{v}_0, \tilde{v}_1, \tilde{v}_2$) in order to estimate ρ, a_n, σ_n, by a classical substitution method, relying on Lemma 2.3 above which gives the relations between the two triplets of parameters. The proof of lemmas 2.2 and 2.3 are given in Appendices A.1 and A.2 respectively. That of Lemma 2.1 can be found in [4] for $j = 0, 1$, the case $j = 2$ being similar.

2.2. Asymptotic behaviour of the estimators of the probability-weighted moments

Let $(X_1, \ldots, X_n)$ be n i.i.d. random variables with distribution function F, and $X_{1:n}, \ldots, X_{n:n}$ denote the corresponding order statistics. For a given thresold u_n, we introduce $Y_{1,N_n}, \ldots, Y_{N_n,N_n}$ the N_n excesses over u_n, in ascending order, *i.e.*

$$Y_{j,N_n} = X_{n-N_n+j\,:\,n} - u_n \quad \text{where} \quad N_n = \textstyle\sum_{i=1}^n \mathbb{I}_{X_i > u_n}.$$

According to (1.3), the distribution B_{θ_n} is then likely to be a good approximation for the distribution F_{u_n} of $Y_{1,N_n}, \ldots, Y_{N_n,N_n}$. This method is of the peak-over-threshold (POT) type.

Remark 2.4. Note that N_n is binomial distributed with mean $n(1 - F(u_n))$ which will be chosen as going to infinity: consequently, $N_n \to \infty$ and $N_n/(n(1 - F(u_n)) \to 1$ in probability as $n \to \infty$.

Definition 2.5. For $j \in \{0,1,2\}$, define the estimator of v_j by

$$\hat{v}_j := \int_0^{+\infty} \frac{(1 - \mathbb{A}_{n,u_n}(x))^{j+1}}{j+1}\, \mathrm{d}x,$$

where,

$$\mathbb{A}_{n,u_n}(x) := \frac{1}{N_n} \sum_{i=1}^{N_n} \mathbb{I}_{\{Y_i \leq x\}}.$$

https://doi.org/10.1051/ps/2010017 Published online by Cambridge University Press

It follows that, conditionnally on $N_n = k_n$,

$$\hat{v}_j = \frac{1}{j+1} \sum_{i=1}^{k_n} \left(\left(1 - \frac{i-1}{k_n}\right)^{j+1} - \left(1 - \frac{i}{k_n}\right)^{j+1} \right) Y_{i,k_n}.$$

Let $b_n = B(\mathrm{e}^{V^{-1}(u_n)})$, where function B appears in the third order condition (1.5).

Theorem 2.6. *Under assumptions (1.1), (1.2) and (1.5), with $-1 < \gamma < 1/2$, and if*

$$\lim_{n\to\infty} \sqrt{n\,(1-F(u_n))}\; a_n b_n = \lambda_1, \quad \lambda_1 \in \mathbb{R}, \tag{2.1}$$

$$\lim_{n\to\infty} \sqrt{n\,(1-F(u_n))}\; a_n^2 = \lambda_2, \quad \lambda_2 \geq 0, \tag{2.2}$$

$$\lim_{n\to\infty} \sqrt{n\,(1-F(u_n))}\; a_n = \infty, \tag{2.3}$$

we have, for almost all sequences $k_n \to +\infty$, conditionally on $N_n = k_n$,

$$\sqrt{k_n} \begin{pmatrix} \dfrac{\widehat{v}_0}{\sigma_n} - \dfrac{v_0}{\sigma_n} \\ \dfrac{\widehat{v}_1}{\sigma_n} - \dfrac{v_1}{\sigma_n} \\ \dfrac{\widehat{v}_2}{\sigma_n} - \dfrac{v_2}{\sigma_n} \end{pmatrix} \xrightarrow{d} \mathcal{N}_3\left(\lambda_1 C, \Gamma\right),$$

where

$$\Gamma = \begin{pmatrix} ((1-2\gamma)(1-\gamma)^2)^{-1} & ((2-\gamma)(1-\gamma)(2-2\gamma))^{-1} & ((3-\gamma)(1-\gamma)(3-2\gamma))^{-1} \\ ((2-\gamma)(1-\gamma)(2-2\gamma))^{-1} & ((3-2\gamma)(2-\gamma)^2)^{-1} & ((2-\gamma)(3-\gamma)(4-2\gamma))^{-1} \\ ((3-\gamma)(1-\gamma)(3-2\gamma))^{-1} & ((2-\gamma)(3-\gamma)(4-2\gamma))^{-1} & ((5-2\gamma)(3-\gamma)^2)^{-1} \end{pmatrix} \tag{2.4}$$

and

$$C = \begin{pmatrix} c^0_{\gamma,\rho,\beta} \\ c^1_{\gamma,\rho,\beta} \\ c^2_{\gamma,\rho,\beta} \end{pmatrix} \quad \text{where } c^j_{\gamma,\rho,\beta} = ((j+1)(j+1-\gamma)(j+1-\gamma-\rho)(j+1-\gamma-\rho-\beta))^{-1}.$$

Proof of Theorem 2.6. Note that

$$\mathbb{A}_{n,u_n} \overset{d}{=} F_{u_n} + \frac{1}{\sqrt{k_n}}\, \alpha_{k_n} \circ F_{u_n},$$

where α_{k_n} is the uniform empirical process based on k_n i.i.d. random variables uniformly distributed on $[0,1]$. We have, for $j \in \{0,1,2\}$,

$$\begin{aligned} \frac{\widehat{v}_j}{\sigma_n} - \frac{v_j}{\sigma_n} &= \frac{\widehat{v}_j}{\sigma_n} - \frac{\nu_j}{\sigma_n} - \frac{a_n}{u_j} \\ &= T^1_{j,k_n} - \frac{1}{\sqrt{k_n}}\, T^2_{j,k_n} + \frac{1}{\sqrt{k_n}}\, T^3_{j,k_n} - \frac{a_n}{u_j}, \end{aligned}$$

https://doi.org/10.1051/ps/2010017 Published online by Cambridge University Press

where,

$$
\begin{aligned}
T^1_{j,k_n} &= \frac{1}{j+1}\int_0^{+\infty}\left[(\overline{F}_{u_n}(\sigma_n y))^{j+1}-(\overline{G}_\gamma(y))^{j+1}\right]\,\mathrm{d}y\\
T^2_{j,k_n} &= \int_0^{+\infty}\left[\alpha_{k_n}\circ F_{u_n}(\sigma_n y)\right](F_{u_n}(\sigma_n y))^j\,\mathrm{d}y\\
T^3_{j,k_n} &= \begin{cases}\displaystyle\int_0^{+\infty}\int_0^1 (1-t)\tfrac{1}{\sqrt{k_n}}[\alpha_{k_n}\circ F_{u_n}]^2(\sigma_n y).j\left[\overline{F}_{u_n}(\sigma_n y)-\tfrac{t}{\sqrt{k_n}}\alpha_{k_n}\circ F_{u_n}(\sigma_n y)\right]^{j-1}\,\mathrm{d}t\,\mathrm{d}y & \text{if } j\in\{1,2\}\\ 0 & \text{if } j=0.\end{cases}
\end{aligned}
$$

This is indeed straightforward for $j=0$, whereas for $j\in\{1,2\}$ we use a Taylor expansion, as in the proof of Theorem 1 in [5] (p. 850), with power functions instead of their general weight functions, which have to be null at zero[3].

The following lemma concerns the terms T^1_{j,k_n} and will be proved in Appendix A.3.

Lemma 2.7. *Under the assumptions of Theorem 2.6, for $j\in\{0,1,2\}$,*

$$T^1_{j,k_n}-\frac{a_n}{u_j}=c^j_{\gamma,\rho,\beta}a_nb_n+o(a_nb_n).$$

T^2_{0,k_n} has been studied in [4]. The other terms T^2_{j,k_n} and T^3_{j,k_n}, for $j\in\{1,2\}$, have been treated in [5] (see pp. 851–853), in a more general framework. The results are stated in the following lemma and the proofs remain valid under our slightly different assumptions (where the role of the condition $\sqrt{k_n}a_n\to\lambda$ is replaced here by $\sqrt{k_n}a_n^2\to\lambda$).

Lemma 2.8. *Under the assumptions[4] of Theorem 2.6, for $j\in\{0,1,2\}$, as $n\to\infty$,*

$$T^3_{j,k_n}\xrightarrow{\mathbb{P}}0\quad\text{and}\quad T^2_{j,k_n}\xrightarrow{d}\int_0^1 t^{-\gamma-1}t^j\mathbb{B}(t)\,\mathrm{d}t,$$

where $\mathbb{B}$ is a Brownian Bridge on $[0,1]$. Moreover, the vector of coordinates $\int_0^1 t^{-\gamma-1}t^j\mathbb{B}(t)\,\mathrm{d}t$ with $j\in\{0,1,2\}$ has a multivariate normal distribution with mean 0 and covariance matrix Γ defined by (2.4).

We deduce from these lemmas that

$$\sqrt{k_n}\left(\frac{\widehat{v}_j}{\sigma_n}-\frac{v_j}{\sigma_n}\right)\overset{d}{=}c^j_{\gamma,\rho,\beta}\sqrt{k_n}\;a_nb_n+Z^j_n+o_{\mathbb{P}}(1),$$

where, using [19] (p. 18), the vector of coordinates Z^j_n (for $j\in\{0,1,2\}$) converges in distribution to $N_3(0,\Gamma)$. The statement of Theorem 2.6 follows by the assumption $\sqrt{k_n}\;a_nb_n\to\lambda_1$.

Remark 2.9. The third order condition is not used to prove Lemma 2.8. This implies that the consistence of the vector of coordinates $\hat{v}_j/\sigma_n$ could be obtained under weaker assumptions.

3. Asymptotic normality of the PWM estimators of the parameters

3.1. Asymptotic normality for known $\vec{\gamma}$

From now on we will use the following notations:

$$V_n=(v_0,v_1,v_2)^t\ ,\ \widehat{V}_n=(\widehat{v}_0,\widehat{v}_1,\widehat{v}_2)^t.$$

[3] This fact exludes the case $j=0$, where the weight function is identically equal to 1, from their study.

[4] The restriction $\gamma<\frac{1}{2}$ comes from the study of T^2_{0,k_n}.

https://doi.org/10.1051/ps/2010017 Published online by Cambridge University Press

The expressions of the probability weighted-moments as functions of the parameters ρ, a_n, σ_n are stated in Lemma 2.3. Elementary calculus leads to the following equations (recall that $u_j = (j+1)(j+1-\gamma)(j+1-\gamma-\rho)$ for $j \in \{0,1,2\}$):

$$\rho = \phi_{1,\gamma}(V_n)\ ,\ a_n = \phi_{2,\gamma}(V_n)\ ,\ \sigma_n = \phi_{3,\gamma}(V_n)$$

where

$$\begin{aligned}
\phi_{1,\gamma}\colon (x,y,z) &\mapsto \frac{(1-\gamma)^2 x - 4(2-\gamma)^2 y + 3(3-\gamma)^2 z}{(1-\gamma)x - 4(2-\gamma)y + 3(3-\gamma)z} \\
\phi_{2,\gamma}\colon (x,y,z) &\mapsto \frac{2((1-\gamma)x-2(2-\gamma)y)((1-\gamma)x-3(3-\gamma)z)(2(2-\gamma)y-3(3-\gamma)z)}{((1-\gamma)x-4(2-\gamma)y+3(3-\gamma)z)\,(6(3-\gamma)z[(1-\gamma)x-2(2-\gamma)y]-2(2-\gamma)y[(1-\gamma)x-3(3-\gamma)z])} \\
\phi_{3,\gamma}\colon (x,y,z) &\mapsto \frac{6(3-\gamma)z[(1-\gamma)x-2(2-\gamma)y]-2(2-\gamma)y[(1-\gamma)x-3(3-\gamma)z]}{(1-\gamma)x-4(2-\gamma)y+3(3-\gamma)z}.
\end{aligned}$$

First assuming that the first order parameter γ is known (the case γ unknown will be handled in the next section), we can then define our estimators of the parameters ρ, a_n, and σ_n as:

$$\begin{pmatrix} \hat\rho_\gamma \\ \hat a_{n,\gamma} \\ \hat\sigma_{n,\gamma} \end{pmatrix} := \begin{pmatrix} \phi_{1,\gamma}(\hat V_n) \\ \phi_{2,\gamma}(\hat V_n) \\ \phi_{3,\gamma}(\hat V_n) \end{pmatrix} \quad \text{hence} \quad \begin{pmatrix} \hat\rho_\gamma \\ \hat a_{n,\gamma} \\ \hat\sigma_{n,\gamma}/\sigma_n \end{pmatrix} = \begin{pmatrix} \phi_{1,\gamma}(\hat V_n/\sigma_n) \\ \phi_{2,\gamma}(\hat V_n/\sigma_n) \\ \phi_{3,\gamma}(\hat V_n/\sigma_n) \end{pmatrix}.$$

Proving the asymptotic normality of these estimators by the delta-method (see [20] for instance) would be straightforward if the functions $\phi_{j,\gamma}$ were well-defined at the limit

$$v_\gamma := \lim_{n\to\infty} \frac{V_n}{\sigma_n} = \left((1-\gamma)^{-1}, (2(2-\gamma))^{-1}, (3(3-\gamma))^{-1}\right)^t.$$

However this is not the case here, and the proof needs more care than it seems at first glance.

Proposition 3.1. *Suppose that $-1 < \gamma < 1/2$ is known, and assumptions (1.1), (1.2), (1.5), (2.1)–(2.3) hold. Then for almost all sequences $k_n \to +\infty$, we have, conditionally on $N_n = k_n$:*

$$\sqrt{k_n} a_n \begin{pmatrix} \hat\rho_\gamma - \rho \\ \hat a_{n,\gamma}/a_n - 1 \\ a_n^{-1}\,(\hat\sigma_{n,\gamma}/\sigma_n - 1) \end{pmatrix} \xrightarrow{d} \mathcal{N}_3(\lambda_1 \nabla H_\gamma C, \nabla H_\gamma \Gamma H_\gamma \nabla^t) \tag{3.1}$$

where ∇ and H_γ are 3×3 matrices defined in the proof of this proposition, and λ_1 in (2.1).

Proof of Proposition 3.1. Let H_γ denote the matrix

$$H_\gamma := \begin{pmatrix} 1-\gamma & 0 & 0 \\ 0 & 2(2-\gamma) & 0 \\ 0 & 0 & 3(3-\gamma) \end{pmatrix}$$

and let us define the following functions

$$\begin{aligned}
\psi_{1,\gamma}\ :\ (x,y,z) &\mapsto \frac{(1-\gamma)x - 2(2-\gamma)y + (3-\gamma)z}{x-2y+z} \\
\psi_2\ :\ (x,y,z) &\mapsto \frac{2(x-y)(x-z)(y-z)}{(x-2y+z)(2z(x-y)-y(x-z))} \\
\psi_3\ :\ (x,y,z) &\mapsto \frac{2z(x-y)-y(x-z)}{x-2y+z}.
\end{aligned}$$

https://doi.org/10.1051/ps/2010017 Published online by Cambridge University Press

If U denotes the subset of $\mathbb{R}^3$ on which ψ_2 is defined, we have $\phi_{1,\gamma}(u) = \psi_{1,\gamma}(H_\gamma u)$, $\phi_{2,\gamma}(u) = \psi_2(H_\gamma u)$ and $\phi_{3,\gamma}(u) = \psi_3(H_\gamma u)$, for every $u = (x, y, z)^t \in U$. In addition note that if $e := (1, 1, 1)^t$, then for every $u \in U$ we have

$$\psi_{1,\gamma}(u + e) = \psi_{1,\gamma}(u) \ , \ \psi_2(u + e) = \tilde{\psi}_2(u)(1 + d_2(u)/d_1(u))^{-1} \ , \ \psi_3(u + e) = \psi_3(u) + 1 \ , \tag{3.2}$$

where $d_1(x, y, z) = x - 2y + z$, $d_2(x, y, z) = 2z(x - y) - y(x - z)$ and

$$\tilde{\psi}_2 : (x, y, z) \ \mapsto \ \frac{2(x - y)(x - z)(y - z)}{(x - 2y + z)^2}.$$

The proof of the proposition relies on the introduction of the following modified probability-weighted moments

$$\widehat{V}'_{n,\gamma} = \frac{\widehat{V}_n/\sigma_n - v_\gamma}{a_n} \quad \text{and} \quad V' = \frac{V_n/\sigma_n - v_\gamma}{a_n} = \left(\frac{1}{u_0}, \frac{1}{u_1}, \frac{1}{u_2}\right)^t$$

(where the u_j are defined in the statement of Lem. 2.3). Defining now $\widehat{V}''_{n,\gamma} := H_\gamma \widehat{V}'_{n,\gamma}$ and

$$V'' := H_\gamma V' = ((1 - \gamma - \rho)^{-1}, (2 - \gamma - \rho)^{-1}, (3 - \gamma - \rho)^{-1})^t \tag{3.3}$$

and noticing that $H_\gamma v_\gamma = e$, $d_2(V'') = 0$ and $d_1(V'') \neq 0$, it is now easy to prove the following identities using (3.2):

$$\sqrt{k_n} a_n(\hat{\rho}_\gamma - \rho) \ = \ \sqrt{k_n} a_n(\phi_{1,\gamma}(\widehat{V}_n/\sigma_n) - \phi_{1,\gamma}(V_n/\sigma_n)) \ = \ \sqrt{k_n} a_n(\psi_{1,\gamma}(\widehat{V}''_{n,\gamma}) - \psi_{1,\gamma}(V'')) \tag{3.4}$$

$$\sqrt{k_n} a_n \left(\hat{a}_{n,\gamma}/a_n - 1\right) \ = \ \sqrt{k_n}(\phi_{2,\gamma}(\widehat{V}_n/\sigma_n) - \phi_{2,\gamma}(V_n/\sigma_n)) \ = \ \sqrt{k_n} a_n(\tilde{\psi}_2(\widehat{V}''_{n,\gamma}) - \tilde{\psi}_2(V'')) \ + \ R_n \tag{3.5}$$

$$\sqrt{k_n} \left(\hat{\sigma}_{n,\gamma}/\sigma_n - 1\right) \ = \ \sqrt{k_n}(\phi_{3,\gamma}(\widehat{V}_n/\sigma_n) - \phi_{3,\gamma}(V_n/\sigma_n)) \ = \ \sqrt{k_n} a_n(\psi_3(\widehat{V}''_{n,\gamma}) - \psi_3(V'')) \tag{3.6}$$

where

$$R_n = \sqrt{k_n} a_n \tilde{\psi}_2(\widehat{V}''_{n,\gamma}) \left((1 + a_n d_2(\widehat{V}''_{n,\gamma})/d_1(\widehat{V}''_{n,\gamma}))^{-1} - 1\right). \tag{3.7}$$

The point is that the functions $\psi_{1,\gamma}$, $\tilde{\psi}_2$ and ψ_3 and their derivatives are well-defined at V'' defined by (3.3) (it was not the case for the functions $\phi_{j,\gamma}$ at the limit $v_\gamma = \lim V_n/\sigma_n$). We have indeed

$$\nabla\psi_{1,\gamma}(V'') = \begin{pmatrix} p(1 - \gamma - \rho)/2 \\ -p(2 - \gamma - \rho) \\ p(3 - \gamma - \rho)/2 \end{pmatrix}, \qquad \nabla\psi_3(V'') = \begin{pmatrix} (1 - \gamma - \rho)^2/2 \\ -(2 - \gamma - \rho)^2 \\ (3 - \gamma - \rho)^2/2 \end{pmatrix},$$

$$\nabla\tilde{\psi}_2(V'') = \begin{pmatrix} \frac{1}{2}(1 - \gamma - \rho)(-5 + 7\gamma + 7\rho - 2(\gamma + \rho)^2) \\ 2(2 - \gamma - \rho)(4 - 4\gamma - 4\rho + (\gamma + \rho)^2) \\ \frac{1}{2}(3 - \gamma - \rho)(-9 + 9\gamma + 9\rho - 2(\gamma + \rho)^2) \end{pmatrix}$$

where $p = (1 - \gamma - \rho)(2 - \gamma - \rho)(3 - \gamma - \rho)$.

Defining ∇ as the Jacobian matrix $J_{\tilde{\psi}_\gamma}(V'')$ where

$$\tilde{\psi}_\gamma \ : \ (x, y, z) \mapsto \left(\psi_{1\gamma}(x, y, z), \tilde{\psi}_2(x, y, z), \psi_3(x, y, z)\right)^t,$$

relations (3.4) to (3.6) become

$$\sqrt{k_n} a_n \begin{pmatrix} \hat{\rho}_\gamma - \rho \\ \hat{a}_{n,\gamma}/a_n - 1 \\ a_n^{-1}\left(\hat{\sigma}_{n,\gamma}/\sigma_n - 1\right) \end{pmatrix} \ = \ \sqrt{k_n} a_n(\tilde{\psi}_\gamma(\widehat{V}''_{n,\gamma}) - \tilde{\psi}_\gamma(V'')) \ + \ \begin{pmatrix} 0 \\ R_n \\ 0 \end{pmatrix}$$

https://doi.org/10.1051/ps/2010017 Published online by Cambridge University Press

and thus the delta-method can be called upon to obtain relation (3.1) by combining Theorem 2.6 with the following equality

$$\sqrt{k_n}a_n(\widehat{V}''_{n,\gamma} - V'') \;=\; H_\gamma(\sqrt{k_n}(\widehat{V}_n/\sigma_n - V_n/\sigma_n))$$

provided that R_n (defined in (3.7)) converges to 0 in probability. This is indeed the case, since $\widehat{V}''_{n,\gamma} \to V''$ in probability as $n\to\infty$, and consequently

$$R_n = -\sqrt{k_n}a_n^2\tilde{\psi}_2(\widehat{V}''_{n,\gamma})\frac{d_2(\widehat{V}''_{n,\gamma})}{d_1(\widehat{V}''_{n,\gamma})}\left(1 + a_n d_2(\widehat{V}''_{n,\gamma})/d_1(\widehat{V}''_{n,\gamma})\right)^{-1}$$

vanishes to 0 as $n\to\infty$ (in probability) because $\tilde{\psi}_2(V'') = 1$, $d_1(V'') = 2/p$, $d_2(V'') = 0$, and using assumption (2.2) (which ensures that $\sqrt{k_n}a_n^2$ has a real limit as $n\to\infty$).

3.2. **Asymptotic normality for unknown $\vec{\gamma}$**

We can now define our final estimators of the parameters ρ, a_n, and σ_n, by plugging-in an external estimator of γ. We set

$$\begin{pmatrix} \hat{\rho} \\ \hat{a}_n \\ \hat{\sigma}_n \end{pmatrix} = \begin{pmatrix} \phi_{1,\hat{\gamma}}(\hat{V}_n) \\ \phi_{2,\hat{\gamma}}(\hat{V}_n) \\ \phi_{3,\hat{\gamma}}(\hat{V}_n) \end{pmatrix}$$

where $\hat{\gamma} = \hat{\gamma}_n$ defines a sequence of estimators of γ based on the $\tilde{N}_n$ upper excesses associated to a threshold $\tilde{u}_n$ such that $\tilde{u}_n \to s_+(F)$. Let $\tilde{a}_n = A(\mathrm{e}^{V^{-1}(\tilde{u}_n)})$ and $\tilde{\lambda}, c, d$ denote some real constants.

Theorem 3.2. *Let the assumptions of Proposition 3.1 hold with $\rho < 0$ and suppose that for some real constant $\tilde{\lambda} \neq 0$,*

$$\sqrt{n(1 - F(\tilde{u}_n))}\,\tilde{a}_n \to \tilde{\lambda} \quad \textit{as} \quad n\to\infty. \tag{3.8}$$

If conditionally on $\tilde{N}_n = \tilde{k}_n$

$$\tilde{k}_n^{1/2}(\hat{\gamma} - \gamma) \xrightarrow{d} \mathcal{N}(\tilde{\lambda}c, d) \quad \textit{as} \quad n\to\infty, \tag{3.9}$$

then for almost all sequences $k_n\to\infty$ and $\tilde{k}_n\to\infty$ such that $\tilde{k}_n = o(k_n)$, we have, conditionally on $N_n = k_n$ and $\tilde{N}_n = \tilde{k}_n$,

$$(\tilde{k}_n)^{1/2}a_n\begin{pmatrix} \hat{\rho} - \rho \\ \hat{a}_n/a_n - 1 \\ a_n^{-1}(\hat{\sigma}_n/\sigma_n - 1) \end{pmatrix} \;=\; (\tilde{k}_n)^{1/2}(\hat{\gamma} - \gamma)\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \;+\; o_{\mathbb{P}}(1) \tag{3.10}$$

for some constants c_1, c_2, c_3 depending on γ and ρ (which expressions are given in the proof of the theorem).

Remark 3.3. The condition $\tilde{k}_n = o(k_n)$ means that we take less excesses for the estimation of the first-order parameter γ than for the estimation of the second-order parameter ρ.

Remark 3.4. Proposition 3.1 is valid in the whole scope $\rho \leq 0$, whereas Theorem 3.2 excludes the case $\rho = 0$. However, according to (3.10) and the expression of c_1, it is worth noting that the asymptotic mean square error (AMSE) of $\hat{\rho}$ tends to 0 when $\rho \to 0$, while this is not the case for many other estimators of ρ studied in the litterature, for which the AMSE goes to infinity when $\rho \to 0$ (in [9] or [3] for instance).

This advantage is however weakened by the fact that the rate of convergence in Theorem 2 is not $\sqrt{k_n}a_n$, which is the rate attained by other estimators of ρ in the litterature: therefore our estimator is asymptotically not as efficient as one would expect. For finite n though, the performance of $\hat{\rho}$ appears to be quite comptetitive in a number of models, particularly in situations where $|\rho|$ is small, as it will be seen in the simulations below (Sect. 4).

Remark 3.5. In our simulations, we used for $\hat{\gamma}$ the PWM estimator defined by Hosking & Wallis in [15] (and studied in [4]).

https://doi.org/10.1051/ps/2010017 Published online by Cambridge University Press

Proof of Theorem 3.2. We keep using the notations previously introduced in the proof of Proposition 3.1 especially:

$$\widehat{V}''_{n,\gamma} = H_\gamma \frac{\widehat{V}_n/\sigma_n - v_\gamma}{a_n} = H_\gamma \widehat{V}'_{n,\gamma} \quad \text{where} \quad v_\gamma = \left((1-\gamma)^{-1}, (2(2-\gamma))^{-1}, (3(3-\gamma))^{-1}\right).$$

We first study the deviation

$$\begin{aligned} \hat{\rho} - \hat{\rho}_\gamma &= \psi_{1,\hat{\gamma}}(\widehat{V}''_{n,\hat{\gamma}}) - \psi_{1,\gamma}(\widehat{V}''_{n,\gamma}) \\ &= \psi_{1,\hat{\gamma}}(\widehat{V}''_{n,\hat{\gamma}}) - \psi_{1,\gamma}(\widehat{V}''_{n,\hat{\gamma}}) + \psi_{1,\gamma}(\widehat{V}''_{n,\hat{\gamma}}) - \psi_{1,\gamma}(\widehat{V}''_{n,\gamma}) \\ &= \gamma - \hat{\gamma} + \psi_{1,\gamma}(\widehat{V}''_{n,\hat{\gamma}}) - \psi_{1,\gamma}(\widehat{V}''_{n,\gamma}) \end{aligned} \tag{3.11}$$

where we used the fact that $\psi_{1,\gamma}(x,y,z) = -\gamma + (x-4y+3z)/(x-2y+z)$. We thus have to concentrate on the second term. If we note J the 3×3 diagonal matrix with diagonal coefficients 1, 2 and 3, after some calculations we obtain the following essential development (remind that $a_n \widehat{V}'_{n,\gamma} = \hat{V}_n/\sigma_n - v_\gamma = o_{\mathbb{P}}(1)$)

$$\begin{aligned} \widehat{V}''_{n,\hat{\gamma}} - \widehat{V}''_{n,\gamma} &= (H_{\hat{\gamma}} - H_\gamma)\widehat{V}'_{n,\gamma} + H_{\hat{\gamma}}(\widehat{V}'_{n,\hat{\gamma}} - \widehat{V}'_{n,\gamma}) \\ &= (\gamma - \hat{\gamma}) J\, \widehat{V}'_{n,\gamma} + H_{\hat{\gamma}}(v_\gamma - v_{\hat{\gamma}})/a_n \\ &= (\gamma - \hat{\gamma}) J\, \widehat{V}'_{n,\gamma} + \frac{\gamma - \hat{\gamma}}{a_n} J\, v_\gamma \\ &= \frac{\gamma - \hat{\gamma}}{a_n}(J\, v_\gamma + o_{\mathbb{P}}(1)). \end{aligned} \tag{3.12}$$

Note that, according to assumptions (3.9), (3.8) (with $\tilde{\lambda} \neq 0$) and the second order condition (1.2) (which ensures that $\frac{\tilde{a}_n}{a_n} \to 0$, when $\rho < 0$), we have

$$\frac{\gamma - \hat{\gamma}}{a_n} = \frac{\tilde{a}_n}{a_n} \frac{\tilde{k}_n^{1/2}(\gamma - \hat{\gamma})}{\tilde{k}_n^{1/2}\tilde{a}_n} \to 0 \qquad \text{as } n \to \infty, \text{ in probability.} \tag{3.13}$$

Since $\widehat{V}'_{n,\gamma} \to V'$ as $n \to \infty$, $\widehat{V}''_{n,\gamma} = H_\gamma \widehat{V}'_{n,\gamma}$ converges in probability to V'' (defined in (3.3)), and consequently relations (3.12) and (3.13) imply that

$$\lim_{n\to\infty} \widehat{V}''_{n,\hat{\gamma}} = \lim_{n\to\infty} \widehat{V}''_{n,\gamma} = V'' \quad \text{in probability.} \tag{3.14}$$

Therefore, combining (3.11)–(3.14), there exists some sequence W_n converging to V'' such that

$$\begin{aligned} \hat{\rho} - \rho &= (\gamma - \hat{\gamma}) + a_n^{-1}(\gamma - \hat{\gamma}) < \nabla\psi_{1,\gamma}(W_n)\,,\, J\, v_\gamma + o_{\mathbb{P}}(1) > + (\hat{\rho}_\gamma - \rho) \\ &= a_n^{-1}(\hat{\gamma} - \gamma)(c_1 + o_{\mathbb{P}}(1)) + (\hat{\rho}_\gamma - \rho) \end{aligned} \tag{3.15}$$

where

$$c_1 := -(\nabla\psi_{1,\gamma}(V''))^t J\, v_\gamma = -\frac{p}{2}\left(\frac{1-\gamma-\rho}{1-\gamma} - 2\frac{2-\gamma-\rho}{2-\gamma} + \frac{3-\gamma-\rho}{3-\gamma}\right) = \frac{p\rho}{(1-\gamma)(2-\gamma)(3-\gamma)}.$$

Note that the first term of the right-hand side of (3.15) makes it impossible to have $\sqrt{k_n}a_n$ as the rate for the asymptotic normality of $\hat{\rho} - \rho$ because, in view of (3.13) and assumptions (2.2) and (2.3), $\sqrt{k_n}a_n(\hat{\gamma} - \gamma) \to 0$ in probability but $\sqrt{k_n}(\hat{\gamma} - \gamma)$ does not.

https://doi.org/10.1051/ps/2010017 Published online by Cambridge University Press

On the other hand, we have (see end of the proof for details on $\hat{a}_n - \hat{a}_{n,\gamma}$)

$$\hat{a}_n - \hat{a}_{n,\gamma} \;=\; a_n(\tilde{\psi}_2(\widehat{V}''_{n,\hat{\gamma}}) - \tilde{\psi}_2(\widehat{V}''_{n,\gamma})) + a_n^2 o_{\mathbb{P}}(1) \quad \text{and} \quad \frac{\hat{\sigma}_n}{\sigma_n} - \frac{\hat{\sigma}_{n,\gamma}}{\sigma_n} \;=\; a_n(\psi_3(\widehat{V}''_{n,\hat{\gamma}}) - \psi_3(\widehat{V}''_{n,\gamma})). \tag{3.16}$$

Consequently, proceeding as for $\hat{\rho} - \rho$ and defining $c_2 := -(\nabla\tilde{\psi}_2(V''))^t J\, v_\gamma$ and $c_3 := -(\nabla\psi_3(V''))^t J\, v_\gamma$, it comes

$$\begin{aligned}(\tilde{k}_n)^{1/2} a_n \begin{pmatrix} \hat{\rho} - \rho \\ \hat{a}_n/a_n - 1 \\ a_n^{-1}\,(\hat{\sigma}_n/\sigma_n - 1) \end{pmatrix} \;=\; & (\tilde{k}_n)^{1/2}(\hat{\gamma} - \gamma) \begin{pmatrix} c_1 + o_{\mathbb{P}}(1) \\ c_2 + o_{\mathbb{P}}(1) \\ c_3 + o_{\mathbb{P}}(1) \end{pmatrix} + \begin{pmatrix} 0 \\ o_{\mathbb{P}}((\tilde{k}_n)^{1/2} a_n^2) \\ 0 \end{pmatrix} \\ & + (\tilde{k}_n/k_n)^{1/2} \sqrt{k_n} a_n \begin{pmatrix} \hat{\rho}_\gamma - \rho \\ \hat{a}_{n,\gamma}/a_n - 1 \\ a_n^{-1}\,(\hat{\sigma}_{n,\gamma}/\sigma_n - 1) \end{pmatrix}\end{aligned} \tag{3.17}$$

which, according to assumptions (3.8), (3.9) and Proposition 3.1, yields the result of Theorem 3.2.

We finish the proof by providing some details regarding (3.16). Setting $e = (1,1,1)^t$ and $h(u) = -\tilde{\psi}_2(u)(1 + d_2(u)/d_1(u))^{-1}(d_2(u)/d_1(u))$, using (3.2) yields

$$\begin{aligned}\hat{a}_n - \hat{a}_{n,\gamma} &= \phi_{2,\hat{\gamma}}(\widehat{V}_n/\sigma_n - v_{\hat{\gamma}} + v_{\hat{\gamma}}) - \phi_{2,\gamma}(\widehat{V}_n/\sigma_n - v_\gamma + v_\gamma) \\ &= \psi_2(a_n \widehat{V}''_{n,\hat{\gamma}} + e) - \psi_2(a_n \widehat{V}''_{n,\gamma} + e) \\ &= \tilde{\psi}_2(a_n \widehat{V}''_{n,\hat{\gamma}}) - \tilde{\psi}_2(a_n \widehat{V}''_{n,\gamma}) + h(a_n \widehat{V}''_{n,\hat{\gamma}}) - h(a_n \widehat{V}''_{n,\gamma}) \\ &= a_n(\tilde{\psi}_2(\widehat{V}''_{n,\hat{\gamma}}) - \tilde{\psi}_2(\widehat{V}''_{n,\gamma})) + a_n^2 o_{\mathbb{P}}(1)\end{aligned}$$

where we used (3.14) and the following facts: $d_2(V'') = 0$, $\tilde{\psi}_2(V'') = 1$, $\tilde{\psi}_2(\alpha u) = \alpha\tilde{\psi}_2(u)$ and $(d_2/d_1)(\alpha u) = \alpha(d_2/d_1)(u)$ for any $\alpha \neq 0$ and $u \in U$.

4. Simulation results

In this section, we shall present graphics obtained, concerning bias and mean square errors of our estimator of ρ, compared with two others, for two classes of underlying distributions and for values of γ in the scope of our theorem.

For the three estimators considered, the P.O.T. method we use consists in choosing a number k_n of excesses for the estimation of ρ, as well as a second number $\tilde{k}_n$ for the preliminary estimation of γ (only when necessary, since one of the estimators studied below does not rely on such an initial estimation of γ). Regarding our estimator, we take, for $\hat{\gamma} = \hat{\gamma}(\tilde{k}_n)$, the Hosking and Wallis' estimator defined in [15] and studied in [4] and [5].

We compare our estimator, denoted in this section by $\hat{\rho}_{PWM}$, with two others: the one presented in Fraga Alves *et al.* [9], which will be noted $\hat{\rho}_{FGH}$ and the one presented in Fraga Alves *et al.* [8], which will be noted $\hat{\rho}_{FHL}$. They are defined as

$$\hat{\rho}_{FGH} = -\left|\frac{3(T_n^{(\tau)}(k_n) - 1)}{T_n^{(\tau)}(k_n) - 3}\right|$$

(see [9] for the definition of $T_n^{(\tau)}(k_n)$), with the tuning parameter τ equal to 0 whenever one expects ρ to be in the range $[-1, 0)$ and equal to 1 otherwise (as suggested for instance in [2]), and

$$\hat{\rho}_{FHL} = 3 - 8\hat{\gamma}_-(\tilde{k}_n) + \frac{6 - 12\hat{\gamma}_-(\tilde{k}_n)}{T_n(k_n) - 3}$$

https://doi.org/10.1051/ps/2010017 Published online by Cambridge University Press

(see [8] for the definition of $T_n(k_n)$ and $\hat{\gamma}_-(\tilde{k}_n)$, the latter being an estimator of $\gamma_- = \min(0,\gamma)$). Recall that k_n is the number of excesses used for the estimation of ρ and the calculation of $\tilde{T}_n(k_n)$, and $\tilde{k}_n$ the one used for the estimation of γ_-. The estimator $\hat{\rho}_{FGH}$ does not depend on an initial estimation of the parameter γ, though.

The models presented in our simulations are the following:

- The $Burr(\lambda,\tau)$ distribution (for which $\gamma = 1/\lambda\tau$ and $\rho = -1/\lambda$) defined by

$$\overline{F}(x) = (1+x^\tau)^{-\lambda},\ x > 0.$$

- The model, for which $\gamma > 0$ and $\rho = 0$ (see [11]), defined by

$$U(t) = \overline{F}^{-1}(1/t) = t^\gamma(1+\ln t),\ t > 0. \tag{4.1}$$

We consider 1000 samples of size $n = 1000$ and present the bias and the mean square error of the three estimators of ρ considered above, as function of the proportion $p_n = k_n/n$ of the excesses used for the estimation of ρ. The sample fraction $\tilde{p}_n = \tilde{k}_n/n$ used in the calculation of $\hat{\rho}_{PWM}$ and $\hat{\rho}_{FHL}$ was chosen as 0.1 for both models. Note however that the simulations undertaken showed that $\hat{\rho}_{FHL}$ was less sensitive to the choice of $\tilde{k}_n$ than $\hat{\rho}_{PWM}$.

Our simulations (see Fig. 1) confirm that in order to estimate the correct value of ρ, one should generally use even more than half of the order statistics of the sample. This is coherent with our theoretical result which states that the number of order statistics to use for the estimation of ρ must be of larger order than the order needed for the estimation of the tail index γ.

The flat pattern of the RMSE of $\hat{\rho}_{PWM}$ for a reasonably wide region of sample fractions makes the exact determination of the optimal choice of the sample fraction p_n to use less relevant, from a practical point of view.

The figures presented here show that our estimator can be competitive especially when $|\rho|$ is small, despite its asymptotic non-efficiency. The same conclusions have been drawn for sample sizes $n = 500$ and $n = 5000$ for the distributions presented here. Note that each of the particular distributions presented here satisfies the restriction $-1 < \gamma < 1/2$ imposed in Theorem 3.2, but removing this restriction does not necessarily lead to a degradation of the performance of $\hat{\rho}_{PWM}$.

Appendix A

A.1. Proof of Lemma 2.2

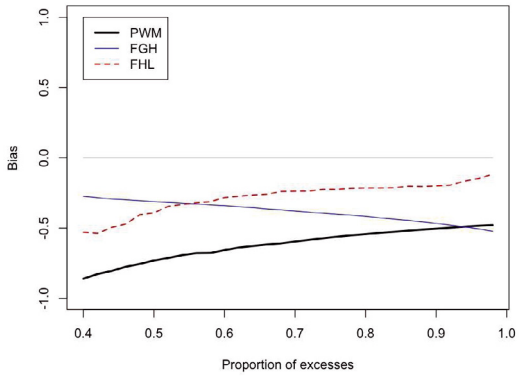
Since $B_{\theta_n}(x) = G_{\gamma,\sigma_n}(x) - a_n D_{\gamma,\rho}(\frac{x}{\sigma_n})$, we have

$$\begin{aligned}\tilde{v}_0 &= \int_0^{+\infty} \overline{B}_{\theta_n}(x)\ \mathrm{d}x \\ &= \int_0^{+\infty} \left(\overline{G}_{\gamma,\sigma_n}(x) + a_n D_{\gamma,\rho}(x/\sigma_n)\right) \mathrm{d}x \\ &= \nu_0 + a_n \int_0^{+\infty} D_{\gamma,\rho}(x/\sigma_n)\ \mathrm{d}x\end{aligned}$$

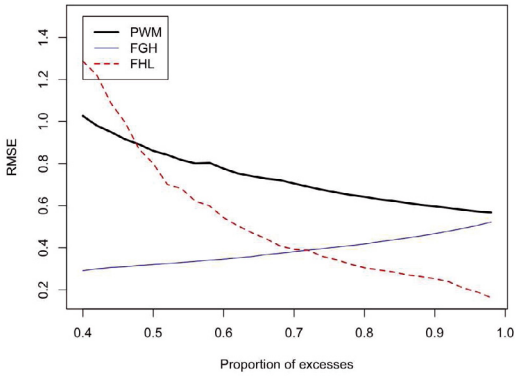
and, for $j \in \{1,2\}$,

$$\begin{aligned}\tilde{v}_j &= \int_0^{+\infty} \frac{\overline{B}^{j+1}_{\theta_n}(x)}{j+1}\ \mathrm{d}x \\ &= \tfrac{1}{j+1}\int_0^{+\infty} \left(\overline{G}_{\gamma,\sigma_n}(x) + a_n D_{\gamma,\rho}(x/\sigma_n)\right)^{j+1} \mathrm{d}x \\ &= \tfrac{1}{j+1}\int_0^{+\infty} \overline{G}^{j+1}_{\gamma,\sigma_n}(x)\ \mathrm{d}x + a_n \int_0^{s_+(B_{\theta_n})} \overline{G}^{j}_{\gamma,\sigma_n}(x)\ D_{\gamma,\rho}(x/\sigma_n)\ \mathrm{d}x + o(a_n) \\ &= \nu_j + a_n \int_0^{+\infty} \overline{G}^{j}_{\gamma,\sigma_n}(x)\ D_{\gamma,\rho}(x/\sigma_n)\ \mathrm{d}x + o(a_n).\end{aligned}$$

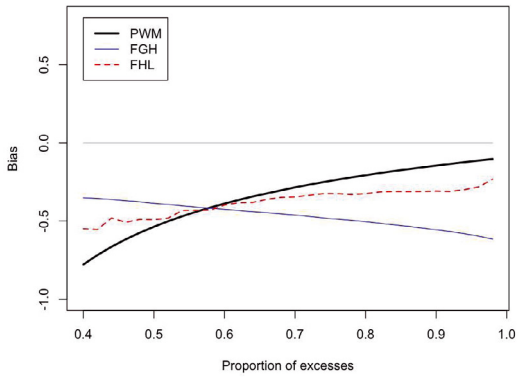
https://doi.org/10.1051/ps/2010017 Published online by Cambridge University Press



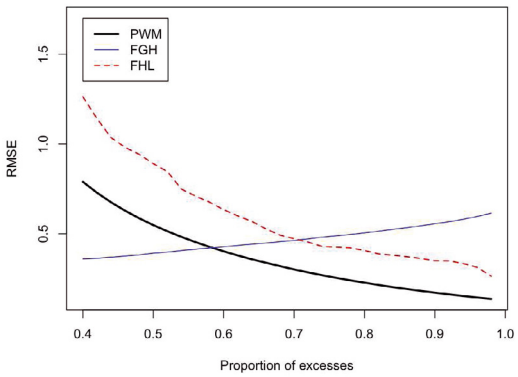
(a) Bias for Burr(5, $\frac{1}{2}$) model, $n = 1000$



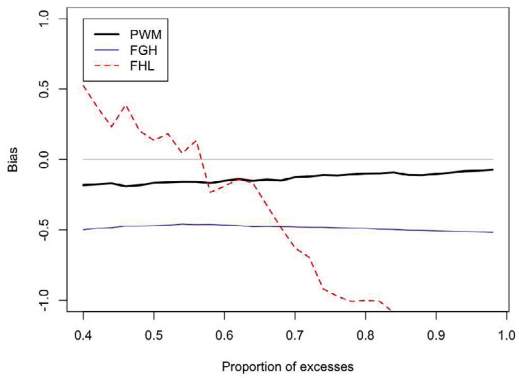
(b) RMSE for Burr(5, $\frac{1}{2}$) model, $n = 1000$



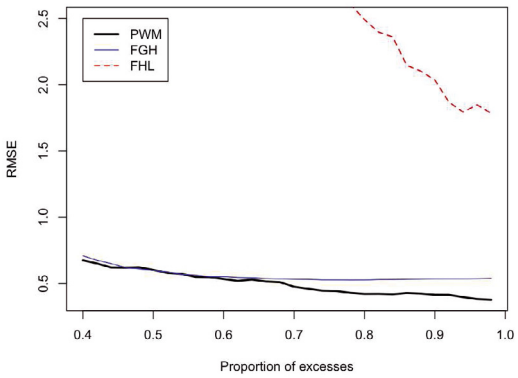
(c) Bias for Burr(10, $\frac{1}{4}$) model, $n = 1000$



(d) RMSE for Burr(10, $\frac{1}{4}$) model, $n = 1000$



(e) Bias for model (4.1) with $\gamma = 2/5$, $n = 1000$



(f) RMSE for model (4.1) with $\gamma = 2/5$, $n = 1000$

FIGURE 1. Bias and RMSE of three estimators of ρ for some models (Burr$(5, 1/2)$, Burr$(10, 1/4)$ and model (4.1)), for which $\gamma = 2/5$ and, respectively, $\rho = -1/5$, $\rho = -1/10$ and $\rho = 0$, as a function of the sample fraction p_n. The dashed line is for $\hat{\rho}_{FHL}$, the thin solid line for $\hat{\rho}_{FGH}$, and the thick solid line for $\hat{\rho}_{PWM}$.

https://doi.org/10.1051/ps/2010017 Published online by Cambridge University Press

4.1. **Proof of Lemma 2.3**

In the case $\gamma \neq 0$, $\rho < 0$ and $\gamma + \rho \neq 0$,

$$D_{\gamma,\rho}(x) = \frac{(1+\gamma x)^{-1/\gamma}}{\gamma\rho(\gamma+\rho)} \left(\gamma(1+\gamma x)^{\rho/\gamma} + \rho(1+\gamma x)^{-1} - \gamma - \rho\right).$$

At the end of this subsection, we will give expressions of $D_{\gamma,\rho}$ in the other cases.

We give the sketch of the proof in case $\gamma > 0$, $\rho < 0$ and $\gamma + \rho \neq 0$. All the other cases are similar. We will see below how the restriction $\gamma < 1$ appears (similarly, in case $\gamma < 0$, appears the restriction $\gamma > -1$). In the sequel, we will note σ instead of $\sigma(u_n)$.

(i) *Calculation of* v_0: we have,

$$\begin{aligned}\int_0^{+\infty} D_{\gamma,\rho}(x/\sigma)\ \mathrm{d}x &= \frac{1}{\rho(\gamma+\rho)} \int_0^{+\infty} (1+\gamma x/\sigma)^{(\rho-1)/\gamma}\ \mathrm{d}x + \frac{1}{\gamma(\gamma+\rho)} \int_0^{+\infty} (1+\gamma x/\sigma)^{-1/\gamma-1}\ \mathrm{d}x \\ &\quad - \frac{1}{\gamma\rho} \int_0^{+\infty} (1+\gamma x/\sigma)^{-1/\gamma}\ \mathrm{d}x \\ &= \frac{1}{\rho(\gamma+\rho)}\ I_1 + \frac{1}{\gamma(\gamma+\rho)}\ I_2 - \frac{1}{\gamma\rho}\ I_3,\end{aligned}$$

where, for $\gamma < 1$,

$$I_1 = \frac{\sigma}{1-\gamma-\rho},\ I_2 = \sigma,\ I_3 = \frac{\sigma}{1-\gamma}.$$

Therefore $\int_0^{+\infty} D_{\gamma,\rho}(x/\sigma)\ \mathrm{d}x = \frac{\sigma}{(1-\gamma)(1-\gamma-\rho)}$ and the result for v_0 follows.

(ii) *Calculation of* v_1: we have,

$$\begin{aligned}\int_0^{+\infty} \overline{G}_{\gamma,\sigma}(x)\ D_{\gamma,\rho}(x/\sigma)\ \mathrm{d}x &= \frac{1}{\rho(\gamma+\rho)} \int_0^{+\infty} (1+\gamma x/\sigma)^{(\rho-2)/\gamma}\ \mathrm{d}x \\ &\quad + \frac{1}{\gamma(\gamma+\rho)} \int_0^{+\infty} (1+\gamma x/\sigma)^{-2/\gamma-1}\ \mathrm{d}x - \frac{1}{\gamma\rho} \int_0^{+\infty} (1+\gamma x/\sigma)^{-2/\gamma}\ \mathrm{d}x \\ &= \frac{1}{\rho(\gamma+\rho)}\ I_1 + \frac{1}{\gamma(\gamma+\rho)}\ I_2 - \frac{1}{\gamma\rho}\ I_3,\end{aligned}$$

where,

$$I_1 = \frac{\sigma}{2-\gamma-\rho},\ I_2 = \sigma/2,\ I_3 = \frac{\sigma}{2-\gamma}.$$

Therefore $\int_0^{+\infty} \overline{G}_{\gamma,\sigma}(x)\ D_{\gamma,\rho}(x/\sigma)\ \mathrm{d}x = \frac{\sigma}{2(2-\gamma)(2-\gamma-\rho)}$ and the expression of v_1 follows.

(iii) *Calculation of* v_2: we have,

$$\begin{aligned}\int_0^{+\infty} \overline{G}^2_{\gamma,\sigma}(x)\ D_{\gamma,\rho}(x/\sigma)\ \mathrm{d}x &= \frac{1}{\rho(\gamma+\rho)} \int_0^{+\infty} (1+\gamma x/\sigma)^{(\rho-3)/\gamma}\ \mathrm{d}x \\ &\quad + \frac{1}{\gamma(\gamma+\rho)} \int_0^{+\infty} (1+\gamma x/\sigma)^{-3/\gamma-1}\ \mathrm{d}x \qquad - \frac{1}{\gamma\rho} \int_0^{+\infty} (1+\gamma x/\sigma)^{-3/\gamma}\ \mathrm{d}x \\ &= \frac{1}{\rho(\gamma+\rho)}\ I_1 + \frac{1}{\gamma(\gamma+\rho)}\ I_2 - \frac{1}{\gamma\rho}\ I_3,\end{aligned}$$

https://doi.org/10.1051/ps/2010017 Published online by Cambridge University Press

where

$$I_1 = \frac{\sigma}{3-\gamma-\rho},\ I_2 = \sigma/3,\ I_3 = \frac{\sigma}{3-\gamma}\cdot$$

Therefore $\int_0^{+\infty} \overline{G}^2_{\gamma,\sigma}(x)\ D_{\gamma,\rho}(x/\sigma)\ \mathrm{d}x = \frac{\sigma}{3(3-\gamma)(3-\gamma-\rho)}$ which yields the result for v_2.

Here are the expressions of $D_{\gamma,\rho}$ in the different other cases:

$$\begin{aligned}
D_{\gamma,\rho}(x) &= \frac{(1+\gamma x)^{-1/\gamma}}{\gamma\rho}\left((1+\gamma x)^{-1}(1+\ln(1+\gamma x))-1\right) && \text{if } \gamma\neq 0,\ \rho<0,\ \gamma+\rho=0\\
&= \frac{(1+\gamma x)^{-1/\gamma}}{\gamma^2}\left(\ln(1+\gamma x)-1+(1+\gamma x)^{-1}\right) && \text{if } \gamma\neq 0,\ \rho=0\\
&= \frac{\mathrm{e}^{-x}}{\rho}\left(\frac{1}{\rho}(\mathrm{e}^{\rho x}-1)-x\right) && \text{if } \gamma=0,\ \rho<0\\
&= \frac{x^2}{2}\mathrm{e}^{-x} && \text{if } \gamma=0,\ \rho=0.
\end{aligned}$$

4.2. **Proof of Lemma 2.7**

$$\begin{aligned}
T^1_{j,k_n} &= \frac{1}{j+1}\int_0^{+\infty}\left[(\overline{F}_{u_n}(\sigma_n y))^{j+1}-(\overline{G}_\gamma(y))^{j+1}\right]\ \mathrm{d}y\\
&= \frac{1}{j+1}\int_0^{+\infty}\left((\overline{F}_{u_n}(\sigma_n y))^{j+1}\right)-\frac{1}{j+1}\int_0^{+\infty}\left((\overline{G}_\gamma(y))^{j+1}\right)\ \mathrm{d}y\\
&=: I_1 - I_2.
\end{aligned}$$

Defining g_n and g by

$$g_n(x) = V^{-1}(u_n+\sigma_n x)-V^{-1}(u_n)\ \text{ and }\ g(x)=\frac{1}{\gamma}\ln(1+\gamma x),$$

we have,

$$\overline{F}_{u_n}(\sigma_n y)=\mathrm{e}^{-g_n(y)}\ \text{ and }\ \overline{G}_\gamma(y)=\mathrm{e}^{-g(y)}.$$

Setting $W(x)=\frac{x^{j+1}}{j+1}$, integration by parts yields

$$I_1 = \int_0^{+\infty} W(\mathrm{e}^{-g_n(y)})\ \mathrm{d}y = \int_0^{+\infty} yW'(\mathrm{e}^{-g_n(y)})g_n'(y)\mathrm{e}^{-g_n(y)}\ \mathrm{d}y = \int_0^{+\infty} g_n^{-1}(s)\mathrm{e}^{-(j+1)s}\ \mathrm{d}s,$$

(where we used the fact that $yW(\overline{F}_{u_n}(\sigma_n y))\to 0$ as $y\to+\infty$) and

$$I_2 = \int_0^{+\infty} W(\mathrm{e}^{-g(y)})\ \mathrm{d}y = \int_0^{+\infty} yW'(\mathrm{e}^{-g(y)})g'(y)\mathrm{e}^{-g(y)}\ \mathrm{d}y = \int_0^{+\infty} g^{-1}(s)\mathrm{e}^{-(j+1)s}\ \mathrm{d}s.$$

It follows that, for $j\in\{0,1,2\}$,

$$T^1_{j,k_n} = \int_0^{+\infty}\left(g_n^{-1}(s)-g^{-1}(s)\right)\mathrm{e}^{-(j+1)s}\ \mathrm{d}s = \int_0^{+\infty} p_{u_n}(s)\mathrm{e}^{-(j+1)s}\ \mathrm{d}s,$$

where $p_{u_n}(s)=\frac{V(s+V^{-1}(u_n))-u_n}{V'(V^{-1}(u_n))}-\int_0^s \mathrm{e}^{\gamma u}\,\mathrm{d}u$. Moreover, it is easy to see that

$$\int_0^{+\infty}\mathrm{e}^{-(j+1)s}I_{\gamma,\rho}(s)\ \mathrm{d}s=\frac{1}{u_j},$$

https://doi.org/10.1051/ps/2010017 Published online by Cambridge University Press

and therefore

$$\frac{T^1_{j,k_n} - \frac{a_n}{u_j}}{a_n b_n} = \int_0^{+\infty} \mathrm{e}^{-(j+1)s} \left[\frac{\frac{p_{u_n}(s)}{a_n} - I_{\gamma,\rho}(s)}{b_n} \right] \mathrm{d}s.$$

We deduce from the third order condition (1.5) that

$$\Gamma_n(s) := \frac{\frac{p_{u_n}(s)}{a_n} - I_{\gamma,\rho}(s)}{b_n} \overset{n\to\infty}{\longrightarrow} R_{\gamma,\rho,\beta}(\mathrm{e}^s).$$

In order to conclude the proof, it remains to use the dominated convergence theorem to show that $\frac{1}{a_n b_n}(T^1_{j,k_n} - \frac{a_n}{u_j})$ converges to $\int_0^{+\infty} \mathrm{e}^{-(j+1)s}\ R_{\gamma,\rho,\beta}(\mathrm{e}^s)\ \mathrm{d}s$. Under the third order condition (1.5), we can use the following bound which is proved in [10] (see Eq. (2.9) in Thm. 2.1 of [10]):

$$\forall \epsilon > 0,\ \exists n_0,\ \forall n \geq n_0,\ \forall s \geq V^{-1}(u_{n_0}) - V^{-1}(u_n),\ |\Gamma_n(s) - R_{\gamma,\rho,\beta}(\mathrm{e}^s)| \leq \epsilon \mathrm{e}^{\epsilon s} \mathrm{e}^{(\gamma+\rho+\beta)s}.$$

Therefore,

$$\left| \int_0^{+\infty} \mathrm{e}^{-(j+1)s} \left(\Gamma_n(s) - R_{\gamma,\rho,\beta}(\mathrm{e}^s)\right)\ \mathrm{d}s \right| \leq \epsilon \int_0^{+\infty} \mathrm{e}^{(j+1-\gamma-\rho-\beta-\epsilon)s}\ \mathrm{d}s,$$

with the right hand side of the inequality being bounded, since ϵ can be set sufficiently small to insure that $j + 1 - \gamma - \rho - \beta - \epsilon > 0$. Finally, elementary calculus leads to

$$\int_0^{+\infty} \mathrm{e}^{-(j+1)s}\ R_{\gamma,\rho,\beta}(\mathrm{e}^s)\ \mathrm{d}s = \frac{1}{(j+1)(j+1-\gamma)(j+1-\gamma-\rho)(j+1-\gamma-\rho-\beta)} = c^j_{\gamma,\rho,\beta}$$

which concludes the proof of Lemma 2.7.

References

[1] A. Balkema and L. de Haan, Residual life time at a great age. *Ann. Probab.* **2** (1974) 792–801.
[2] F. Caeiro, M.I. Gomes and D. Pestana, A note on the asymptotic variance at optimal levels of a bias-corrected Hill estimator. *Stat. Probab. Lett.* **79** (2009) 295–303.
[3] G. Ciuperca and C. Mercadier, Semi-parametric estimation for heavy tailed distributions. *Extremes* **13** (2010) 55–87.
[4] J. Diebolt, A. Guillou and R. Worms, Asymptotic behaviour of the probability-weighted moments and penultimate approximation. *ESAIM: PS* **7** (2003) 217–236.
[5] J. Diebolt, A. Guillou and I. Rached, Approximation of the distribution of excesses through a generalized probability-weighted moments method. *J. Statist. Plann. Inference* **137** (2007) 841–857.
[6] J. Diebolt, A. Guillou and I. Rached, Approximation of the distribution of excesses through a generalized probability-weighted moments method. *J. Statist. Plann. Inference* **137** (2007) 841–857.
[7] H. Drees and E. Kaufmann, Selecting the optimal sample fraction in univariate extreme value estimation. *Stoc. Proc. Appl.* **75** (1998) 149–172.
[8] M.I. Fraga Alves, L. de Haan and T. Lin, Estimation of the parameter controlling the speed of convergence in extreme value theory. *Math. Methods Stat.* **12** (2003) 155–176.
[9] 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** (2003) 193–213.
[10] M.I. Fraga Alves, L. de Haan and T. Lin, Third order extended regular variation. *Publ. Inst. Math.* **80** (2006) 109–120.
[11] M.I. Fraga Alves, M.I. Gomes, L. de Haan and C. Neves, A note on second order conditions in extreme value theory: linking general and heavy tail conditions. *REVSTAT Stat. J.* **5** (2007) 285–304.
[12] M.I. Gomes and J. Martins, "Asymptotically unbiased" estimators of the tail index based on external estimation of the second order parameter. *Extremes* **5** (2002) 5–31.
[13] M.I. Gomes, L. de Haan and L. Peng, Semi-parametric estimation of the second order parameter in statistics of extremes. *Extremes* **5** (2002) 387–414.
[14] P. Hall and A.H. Welsh, Adaptive estimates of parameters of regular variation. *Ann. Stat.* **13** (1985) 331–341.

https://doi.org/10.1051/ps/2010017 Published online by Cambridge University Press

[15] J. Hosking and J. Wallis, Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics* **29** (1987) 339–349.
[16] L. Peng, Asymptotically unbiased estimator for the extreme value index. *Statist. Prob. Lett.* **38** (1998) 107–115.
[17] J. Pickands III, Statistical inference using extreme order statistics. *Ann. Statist.* **3** (1975) 119–131.
[18] J.P. Raoult and R. Worms, Rate of convergence for the generalized Pareto approximation of the excesses. *Adv. Applied Prob.* **35** (2003) 1007–1027.
[19] R.J. Serfling, *Approximation Theorems of Mathematical Statistics*. Wiley & Son (1980).
[20] A.W. van der Vaart, Asymptotic Statistics. Cambridge Series in *Statistical and Probabilistic Mathematics* (2000).
[21] R. Worms, Penultimate approximation for the distribution of the excesses. *ESAIM: PS* **6** (2002) 21–31.

https://doi.org/10.1051/ps/2010017 Published online by Cambridge University Press