Estimation of the third order parameter in extreme value statistics

Goegebeur, Yuri
University of Southern Denmark, Department of Mathematics and Computer Science
Campusvej 55
5230 Odense M, Denmark
E-mail: yuri.goegebeur@stat.sdu.dk

de Wet, Tertius
Stellenbosch University, Department of Statistics and Actuarial Science
Private Bag X1
7602 Matieland, South Africa
E-mail: tdewet@sun.ac.za

Introduction

A distribution is said to be of Pareto-type if for some $\gamma > 0$ its survival function is of the form:

(1) $$1 - F(x) = x^{-1/\gamma}\ell_F(x), \qquad x > 0,$$

where ℓ_F denotes a slowly varying function at infinity, i.e.

(2) $$\frac{\ell_F(\lambda x)}{\ell_F(x)} \to 1 \text{ as } x \to \infty \text{ for all } \lambda > 0.$$

The Pareto-type model can also be stated in an equivalent way in terms of the tail quantile function U, where $U(x) := \inf\{y : F(y) \geq 1 - 1/x\}$, $x > 1$, as follows

(3) $$U(x) = x^{\gamma}\ell_U(x),$$

with ℓ_U again a slowly varying function at infinity (Gnedenko, 1943). The estimation of the parameter γ has received a lot of attention in the extreme value literature; we refer to Beirlant *et al.* (2004) and de Haan and Ferreira (2006) for good accounts of such estimation procedures. Consistency of estimators for γ can typically be achieved under (1), which is a first order condition, whereas establishing asymptotic normality requires more structure on the tail of F. This extra structure is typically formulated in terms of a second order condition on the tail behavior, the so-called slow variation with remainder condition; see de Haan and Ferreira (2006). Let $D_\rho(x) := \int_1^x u^{\rho-1}du = (x^\rho - 1)/\rho$, $x > 0$, $\rho < 0$.

Second order condition $(\mathcal{R}_2)$ *There exists a positive real parameter γ, a negative real parameter ρ and a function b with $b(t) \to 0$ for $t \to \infty$, of constant sign for large values of t, such that*

$$\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{b(t)} = D_\rho(x), \quad \forall x > 0.$$

Let us remark here that condition $(\mathcal{R}_2)$ implies that $|b|$ is regularly varying of index ρ (Geluk and de Haan, 1987), and hence the parameter ρ determines the rate of convergence of $\ln U(tx) - \ln U(t)$ to its limit, being $\gamma \ln x$, as $t \to \infty$.

Although the estimation of the second order parameter ρ is challenging, both from a theoretical and an applied perspective, several estimators for ρ that work well in practice have recently been introduced, and their asymptotic properties studied. Consistency of these estimators can be established under

$(\mathcal{R}_2)$, whereas asymptotic normality requires again a further condition on the tail behavior, the third order condition. Denote

$$\begin{aligned} H_{\rho,\beta}(x) &:= \int_1^x y^{\rho-1} \int_1^y u^{\beta-1} du\, dy \\ &= \frac{1}{\beta}\left(\frac{x^{\rho+\beta}-1}{\rho+\beta} - \frac{x^\rho - 1}{\rho}\right), \qquad x>0, \rho, \beta < 0. \end{aligned}$$

Third order condition $(\mathcal{R}_3)$ *There exists a positive real parameter* γ*, negative real parameters* ρ *and* β*, functions* b *and* $\tilde{b}$ *with* $b(t) \to 0$ *and* $\tilde{b}(t) \to 0$ *for* $t \to \infty$*, both of constant sign for large values of* t*, such that*

$$\lim_{t\to\infty} \frac{\frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{b(t)} - D_\rho(x)}{\tilde{b}(t)} = H_{\rho,\beta}(x), \;\; \forall x > 0.$$

Note that under condition $(\mathcal{R}_3)$ it can be proven that $|\tilde{b}|$ is regularly varying of index β. In this paper we develop a class of estimators for the third order parameter β.

Construction of the estimator

Consider $X_1, \ldots, X_n$ independent and identically distributed (i.i.d.) random variables according to a distribution function F satisfying (1), with associated order statistics $X_{1,n} \le \cdots \le X_{n,n}$. Let $T_{n,k}(K)$ denote a kernel statistic with kernel function K:

$$(4) \quad T_{n,k}(K) := \frac{1}{k}\sum_{j=1}^k K\left(\frac{j}{k+1}\right) Z_j$$

where $Z_j := j(\ln X_{n-j+1,n} - \ln X_{n-j,n})$, the so-called scaled log-spacings of successive order statistics introduced in Beirlant *et al.* (1999) and Feuerverger and Hall (1999). Further, for $\beta, \rho < 0$, let

$$\begin{aligned} \mu(K) &:= \int_0^1 K(u)du, \qquad \sigma^2(K) := \int_0^1 K^2(u)du, \\ I_1(K,\rho) &:= \int_0^1 K(u)u^{-\rho}du, \qquad I_2(K,\rho,\beta) := \int_0^1 K(u)u^{-\rho}\frac{u^{-\beta}-1}{\beta}du. \end{aligned}$$

In this paper we will consider statistics of type (4) with a log-weight function

$$\mathbb{L}_\alpha(u) := \frac{(-\ln u)^\alpha}{\Gamma(\alpha+1)}, \quad u \in (0,1),$$

with $\alpha > 0$ and where Γ denotes the Gamma function, i.e. $\Gamma(\omega) = \int_0^\infty e^{-x}x^{\omega-1}dx$ $(\omega > 0)$, as basic building blocks for the class of estimators for β. Therefore, we start with studying their asymptotic behavior. The following proposition gives the asymptotic distributional representation of $T_{n,k}(\mathbb{L}_\alpha)$ under $(\mathcal{R}_3)$. Let $Y_{1,n} \le \cdots \le Y_{n,n}$ denote the order statistics of a random sample of size n from the unit Pareto distribution, with distribution function $F(y) = 1 - 1/y$, $y > 1$, and

$$N_k(\mathbb{L}_\alpha) := \sqrt{k}\, \frac{\frac{1}{k}\sum_{j=1}^k \mathbb{L}_\alpha\left(\frac{j}{k+1}\right)(F_j - 1)}{\sigma(\mathbb{L}_\alpha)}$$

with $F_1, \ldots, F_k$ i.i.d. Exp(1) random variables.

Proposition 1 *Let* $X_1, \ldots, X_n$ *be i.i.d. random variables according to a distribution satisfying* $(\mathcal{R}_3)$*. If* $k \to \infty$ *such that* $k/n \to 0$ *we have that*

$$\begin{aligned} T_{n,k}(\mathbb{L}_\alpha) &\stackrel{\mathcal{D}}{=} \gamma + \gamma\sigma(\mathbb{L}_\alpha)\frac{N_k(\mathbb{L}_\alpha)}{\sqrt{k}} + b(Y_{n-k,n})I_1(\mathbb{L}_\alpha, \rho) \\ &\quad + b(Y_{n-k,n})\tilde{b}(Y_{n-k,n})I_2(\mathbb{L}_\alpha,\rho,\beta)(1+o_{\mathbb{P}}(1)) + b(Y_{n-k,n})O_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right). \end{aligned}$$

This proposition follows immediately from the fact that $T_{n,k}(\mathbb{L}_\alpha)$ is an element of the class of kernel statistics studied in Theorem 2 of Goegebeur *et al.* (2010). Note that

$$\begin{aligned} I_1(\mathbb{L}_\alpha,\rho) &= \frac{1}{(1-\rho)^{\alpha+1}}, \qquad I_2(\mathbb{L}_\alpha,\rho,\beta) = \frac{1}{\beta}\left[\frac{1}{(1-\rho-\beta)^{\alpha+1}} - \frac{1}{(1-\rho)^{\alpha+1}}\right], \\ \sigma^2(\mathbb{L}_\alpha) &= \frac{\Gamma(2\alpha+1)}{\Gamma^2(\alpha+1)}. \end{aligned}$$

The basic idea is now to use Proposition 1 to construct differences and ratios of statistics $T_{n,k}(\mathbb{L}_\alpha)$, chosen in such a way that the term involving $I_2(\mathbb{L}_\alpha,\rho,\beta)$ becomes the dominant one.

For arbitrary positive real parameters α and $\tilde{\alpha}$, we denote

$$I_1(\alpha,\tilde{\alpha},\rho) := I_1(\mathbb{L}_\alpha,\rho) - I_1(\mathbb{L}_{\tilde{\alpha}},\rho), \qquad I_2(\alpha,\tilde{\alpha},\rho,\beta) := I_2(\mathbb{L}_\alpha,\rho,\beta) - I_2(\mathbb{L}_{\tilde{\alpha}},\rho,\beta).$$

From Proposition 1 we have that if $\sqrt{k}b(n/k) \to \infty$

$$\Psi_{n,k}(\alpha,\tilde{\alpha},\breve{\alpha}) := \frac{T_{n,k}(\mathbb{L}_\alpha) - T_{n,k}(\mathbb{L}_{\tilde{\alpha}})}{T_{n,k}(\mathbb{L}_{\breve{\alpha}}) - T_{n,k}(\mathbb{L}_{\tilde{\alpha}})} \xrightarrow{\mathbb{P}} \psi(\alpha,\tilde{\alpha},\breve{\alpha},\rho) := \frac{I_1(\alpha,\tilde{\alpha},\rho)}{I_1(\breve{\alpha},\tilde{\alpha},\rho)} = \frac{(1-\rho)^{\tilde{\alpha}-\alpha}-1}{(1-\rho)^{\tilde{\alpha}-\breve{\alpha}}-1},$$

where $\breve{\alpha}$ is again an arbitrary positive real parameter. Now, in order to make the term involving $I_2(\alpha,\rho,\beta)$ the dominant one, we construct a ratio of differences of statistics $\Psi_{n,k}(\alpha,\tilde{\alpha},\breve{\alpha})$, with appropriately chosen parameters α, $\tilde{\alpha}$ and $\breve{\alpha}$. Setting $0<\tau_1<\tau_2<\tau_3$ and $0<\delta_1,\delta_2<\tau_1$, consider the statistic

$$\Lambda_{n,k}(\boldsymbol{\tau},\boldsymbol{\delta}) := \frac{\Psi_{n,k}(\tau_1-\delta_1,\tau_1,\tau_1-\delta_2) - \Psi_{n,k}(\tau_2-\delta_1,\tau_2,\tau_2-\delta_2)}{\Psi_{n,k}(\tau_2-\delta_1,\tau_2,\tau_2-\delta_2) - \Psi_{n,k}(\tau_3-\delta_1,\tau_3,\tau_3-\delta_2)}$$

where $\boldsymbol{\tau}' := (\tau_1,\tau_2,\tau_3)$ and $\boldsymbol{\delta}' := (\delta_1,\delta_2)$, and the function

$$\begin{aligned} \Lambda(\boldsymbol{\tau},\rho,\beta) &:= \frac{c_1(\tau_1-\delta_1,\tau_1,\tau_1-\delta_2,\rho,\beta) - c_1(\tau_2-\delta_1,\tau_2,\tau_2-\delta_2,\rho,\beta)}{c_1(\tau_2-\delta_1,\tau_2,\tau_2-\delta_2,\rho,\beta) - c_1(\tau_3-\delta_1,\tau_3,\tau_3-\delta_2,\rho,\beta)} \\ &= (1-\rho-\beta)^{\tau_3-\tau_1}\frac{(1-\rho-\beta)^{\tau_2}(1-\rho)^{\tau_1} - (1-\rho-\beta)^{\tau_1}(1-\rho)^{\tau_2}}{(1-\rho-\beta)^{\tau_3}(1-\rho)^{\tau_2} - (1-\rho-\beta)^{\tau_2}(1-\rho)^{\tau_3}}, \end{aligned}$$

with

$$(5) \quad c_1(\alpha,\tilde{\alpha},\breve{\alpha},\rho,\beta) := \frac{I_2(\alpha,\tilde{\alpha},\rho,\beta) - \psi(\alpha,\tilde{\alpha},\breve{\alpha},\rho)I_2(\breve{\alpha},\tilde{\alpha},\rho,\beta)}{I_1(\breve{\alpha},\tilde{\alpha},\rho)}.$$

Lemma 1 *The function $\beta \mapsto \Lambda(\boldsymbol{\tau},\rho,\beta)$ is decreasing for $\beta \in (-\infty,0)$ with*

$$\lim_{\beta\to 0-}\Lambda(\boldsymbol{\tau},\rho,\beta) = -\frac{\tau_2-\tau_1}{\tau_2-\tau_3} \qquad \textit{and} \qquad \lim_{\beta\to-\infty}\Lambda(\boldsymbol{\tau},\rho,\beta) = +\infty.$$

Now, consider the estimating equation $\Lambda_{n,k}(\boldsymbol{\tau},\boldsymbol{\delta}) = \Lambda(\boldsymbol{\tau},\rho,\beta)$, from which the estimator for β can be obtained by inversion:

$$\hat{\beta}_{n,k}(\boldsymbol{\tau},\boldsymbol{\delta},\rho) := \Lambda^{\leftarrow}(\boldsymbol{\tau},\rho,\Lambda_{n,k}(\boldsymbol{\tau},\boldsymbol{\delta})),$$

provided $\Lambda_{n,k}(\boldsymbol{\tau},\boldsymbol{\delta}) > -(\tau_2-\tau_1)/(\tau_2-\tau_3)$.

Asymptotic properties

Theorem 1 *Let $X_1,\ldots,X_n$ be i.i.d. random variables according to a distribution satisfying $(\mathcal{R}_3)$. Then, if $k\to\infty$ such that $k/n\to 0$ and $\sqrt{k}b(n/k)\tilde{b}(n/k)\to\infty$ we have that $\Lambda_{n,k}(\boldsymbol{\tau},\boldsymbol{\delta}) \xrightarrow{\mathbb{P}} \Lambda(\boldsymbol{\tau},\rho,\beta)$ and that $\hat{\beta}_{n,k}(\boldsymbol{\tau},\boldsymbol{\delta},\rho) \xrightarrow{\mathbb{P}} \beta$. Further, if $\hat{\rho}_{n,\breve{k}}$ is a consistent estimator for ρ, then also $\hat{\beta}_{n,k}(\boldsymbol{\tau},\boldsymbol{\delta},\hat{\rho}_{n,\breve{k}}) \xrightarrow{\mathbb{P}} \beta$.*

The recent extreme value literature contains several consistent estimators for the second order parameter ρ that work very well in practice; see e.g. Gomes *et al.* (2002), Fraga Alves *et al.* (2003), Ciuperca and Mercadier (2010), and Goegebeur *et al.* (2010). In these papers it is shown that consistency of the estimators for ρ can be obtained when the sequence $\breve{k}$ satisfies $\breve{k} \to \infty$ with $\breve{k}/n \to 0$ and $\sqrt{\breve{k}}b(n/\breve{k}) \to \infty$. Moreover, if additionally $\sqrt{\breve{k}}b(n/\breve{k})\tilde{b}(n/\breve{k}) \to \lambda_1$ and $\sqrt{\breve{k}}b^2(n/\breve{k}) \to \lambda_2$ we can obtain asymptotic normality for $\hat{\rho}_{n,\breve{k}}$, when appropriately normalized. In the framework of Theorem 1, $\breve{k}$ can be taken equal to k, the number of extreme values used in the estimation of β. Indeed, $\sqrt{k}b(n/k)\tilde{b}(n/k) \to \infty$ implies that $\sqrt{k}b(n/k) \to \infty$, and hence $\hat{\rho}_{n,k}$ will be consistent for ρ, though for such a sequence we may no longer guarantee asymptotic normality of the normalized $\hat{\rho}_{n,k}$.

We now establish that our class of estimators for the third order parameter, when appropriately normalized, converges in distribution to a normal random variable.

Denote

$$\begin{aligned} R_{\rho,\beta,\eta}(x) &:= \int_1^x y^{\rho-1} \int_1^y u^{\beta-1} \int_1^u s^{\eta-1} ds\, du\, dy \\ &= \frac{1}{\eta}\left[\frac{1}{\beta+\eta}\left(\frac{x^{\rho+\beta+\eta}-1}{\rho+\beta+\eta} - \frac{x^\rho - 1}{\rho}\right) - \frac{1}{\beta}\left(\frac{x^{\rho+\beta}-1}{\rho+\beta} - \frac{x^\rho-1}{\rho}\right)\right], \quad x>0, \rho,\beta,\eta<0. \end{aligned}$$

Fourth order condition $(\mathcal{R}_4)$ *There exists a positive real parameter γ, negative real parameters ρ, β and η, functions b, $\tilde{b}$ and $\breve{b}$ with $b(t) \to 0$, $\tilde{b}(t) \to 0$ and $\breve{b}(t) \to 0$ for $t \to \infty$, all of constant sign for large values of t, such that*

$$\lim_{t\to\infty} \frac{1}{\breve{b}(t)}\left[\frac{1}{\tilde{b}(t)}\left(\frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{b(t)} - D_\rho(x)\right) - H_{\rho,\beta}(x)\right] = R_{\rho,\beta,\eta}(x), \quad \forall x > 0.$$

It is shown in Fraga Alves *et al.* (2006) that $|\breve{b}|$ is regularly varying with index of regular variation η.

We start with the expansion of the statistic $T_{n,k}(K)$ under $(\mathcal{R}_4)$. This result is more general than what is needed in the present paper, but is of interest on its own as such an expansion for a general kernel statistic is to date not available.

We impose the following conditions on the kernel function K.

Condition $(\mathcal{K})$ *Let K be a function defined on $(0,1)$ such that*
(i) $K(t) = \frac{1}{t}\int_0^t u(v)dv$ *for some function u satisfying* $\left|(k+1)\int_{(j-1)/(k+1)}^{j/(k+1)} u(t)dt\right| \le f(\frac{j}{k+1})$ *for some positive continuous and integrable function f defined on $(0,1)$,*
(ii) $\frac{1}{k}\sum_{j=1}^k K\left(\frac{j}{k+1}\right) = \mu(K) + o(1/\sqrt{k})$ *for* $k \to \infty$,
(iii) $\frac{1}{k}\sum_{j=1}^k K\left(\frac{j}{k+1}\right)\left(\frac{j}{k+1}\right)^{-\rho} = I_1(K,\rho) + o(1/\sqrt{k})$ *for* $k \to \infty$,
(iv) $\frac{1}{k}\sum_{j=1}^k K\left(\frac{j}{k+1}\right)\left(\frac{j}{k+1}\right)^{-\rho-\beta} = I_1(K,\rho+\beta) + o(1/\sqrt{k})$ *for* $k \to \infty$,
(v) $\sigma^2(K) < \infty$,
(vi) $\max_{i\in\{1,\ldots,k\}}\left|K\left(\frac{i}{k+1}\right)\right| = o(\sqrt{k})$ *for* $k \to \infty$,
(vii) $\int_0^1 |K(u)|\, u^{|\rho|-1-\varepsilon} du < \infty$ *for some* $\varepsilon > 0$.

Denote

$$I_3(K,\rho,\beta,\eta) := \frac{1}{\eta}\int_0^1 K(u)u^{-\rho}\left(\frac{u^{-\beta-\eta}-1}{\beta+\eta} - \frac{u^{-\beta}-1}{\beta}\right)du.$$

Theorem 2 *Let $X_1, \ldots, X_n$ be i.i.d. random variables according to a distribution satisfying $(\mathcal{R}_4)$. If further $(\mathcal{K})$ holds, then for $k \to \infty$ such that $k/n \to 0$ we have*

$$\begin{aligned} T_{n,k}(K) &\stackrel{\mathcal{D}}{=} \gamma\mu(K) + \gamma\sigma(K)\frac{N_k(K)}{\sqrt{k}}(1+o_{\mathbb{P}}(1)) + b(Y_{n-k,n})I_1(K,\rho) + b(Y_{n-k,n})O_{\mathbb{P}}(1/\sqrt{k}) \\ &+ b(Y_{n-k,n})\tilde{b}(Y_{n-k,n})I_2(K,\rho,\beta) + b(Y_{n-k,n})\tilde{b}(Y_{n-k,n})O_{\mathbb{P}}(1/\sqrt{k}) \\ &+ b(Y_{n-k,n})\tilde{b}(Y_{n-k,n})\breve{b}(Y_{n-k,n})I_3(K,\rho,\beta,\eta)(1+o_{\mathbb{P}}(1)), \end{aligned}$$

where $N_k(K)$ is an asymptotic standard normally distributed random variable.

For the log kernel function $\mathbb{L}_\alpha$ we have

$$\begin{aligned} I_3(\mathbb{L}_\alpha,\rho,\beta,\eta) &= \frac{1}{\eta}\left\{\frac{1}{\beta+\eta}\left[\frac{1}{(1-\rho-\beta-\eta)^{\alpha+1}} - \frac{1}{(1-\rho)^{\alpha+1}}\right]\right. \\ &\left. - \frac{1}{\beta}\left[\frac{1}{(1-\rho-\beta)^{\alpha+1}} - \frac{1}{(1-\rho)^{\alpha+1}}\right]\right\}. \end{aligned}$$

We now establish the asymptotic normality of our estimator. In this we will use the following notation:

$$\begin{aligned} I_3(\alpha,\tilde{\alpha},\rho,\beta,\eta) &:= I_3(\mathbb{L}_\alpha,\rho,\beta,\eta) - I_3(\mathbb{L}_{\tilde{\alpha}},\rho,\beta,\eta), \\ c_2(\alpha,\tilde{\alpha},\breve{\alpha},\rho,\beta,\eta) &:= \frac{I_3(\alpha,\tilde{\alpha},\rho,\beta,\eta) - \psi(\alpha,\tilde{\alpha},\breve{\alpha},\rho)I_3(\breve{\alpha},\tilde{\alpha},\rho,\beta,\eta)}{I_1(\breve{\alpha},\tilde{\alpha},\rho)}, \\ c_3(\alpha,\tilde{\alpha},\breve{\alpha},\rho,\beta) &:= \frac{I_2(\alpha,\tilde{\alpha},\rho,\beta)I_2(\breve{\alpha},\tilde{\alpha},\rho,\beta) - \psi(\alpha,\tilde{\alpha},\breve{\alpha},\rho)I_2^2(\breve{\alpha},\tilde{\alpha},\rho,\beta)}{I_1^2(\breve{\alpha},\tilde{\alpha},\rho)}, \\ N_k(\alpha,\tilde{\alpha},\gamma) &:= \gamma[\sigma(\mathbb{L}_\alpha)N_k(\mathbb{L}_\alpha) - \sigma(\mathbb{L}_{\tilde{\alpha}})N_k(\mathbb{L}_{\tilde{\alpha}})], \\ N_k(\alpha,\tilde{\alpha},\breve{\alpha},\gamma,\rho) &:= \frac{N_k(\alpha,\tilde{\alpha},\gamma) - \psi(\alpha,\tilde{\alpha},\breve{\alpha},\rho)N_k(\breve{\alpha},\tilde{\alpha},\gamma)}{I_1(\breve{\alpha},\tilde{\alpha},\rho)}, \\ V_k(\tau_1,\tau_2,\boldsymbol{\delta},\gamma,\rho) &:= N_k(\tau_1-\delta_1,\tau_1,\tau_1-\delta_2,\gamma,\rho) - N_k(\tau_2-\delta_1,\tau_2,\tau_2-\delta_2,\gamma,\rho), \\ d_1(\tau_1,\tau_2,\boldsymbol{\delta},\rho,\beta) &:= c_1(\tau_1-\delta_1,\tau_1,\tau_1-\delta_2,\rho,\beta) - c_1(\tau_2-\delta_1,\tau_2,\tau_2-\delta_2,\rho,\beta), \\ d_2(\tau_1,\tau_2,\boldsymbol{\delta},\rho,\beta,\eta) &:= c_2(\tau_1-\delta_1,\tau_1,\tau_1-\delta_2,\rho,\beta,\eta) - c_2(\tau_2-\delta_1,\tau_2,\tau_2-\delta_2,\rho,\beta,\eta), \\ d_3(\tau_1,\tau_2,\boldsymbol{\delta},\rho,\beta) &:= c_3(\tau_1-\delta_1,\tau_1,\tau_1-\delta_2,\rho,\beta) - c_3(\tau_2-\delta_1,\tau_2,\tau_2-\delta_2,\rho,\beta), \\ V_k(\boldsymbol{\tau},\boldsymbol{\delta},\gamma,\rho,\beta) &:= \frac{V_k(\tau_1,\tau_2,\boldsymbol{\delta},\gamma,\rho) - \Lambda(\boldsymbol{\tau},\rho,\beta)V_k(\tau_2,\tau_3,\boldsymbol{\delta},\gamma,\rho)}{d_1(\tau_2,\tau_3,\boldsymbol{\delta},\rho,\beta)} \\ d_2(\boldsymbol{\tau},\boldsymbol{\delta},\rho,\beta,\eta) &:= \frac{d_2(\tau_1,\tau_2,\boldsymbol{\delta},\rho,\beta,\eta) - \Lambda(\boldsymbol{\tau},\rho,\beta)d_2(\tau_2,\tau_3,\boldsymbol{\delta},\rho,\beta,\eta)}{d_1(\tau_2,\tau_3,\boldsymbol{\delta},\rho,\beta)}, \\ d_3(\boldsymbol{\tau},\boldsymbol{\delta},\rho,\beta) &:= \frac{d_3(\tau_1,\tau_2,\boldsymbol{\delta},\rho,\beta) - \Lambda(\boldsymbol{\tau},\rho,\beta)d_3(\tau_2,\tau_3,\boldsymbol{\delta},\rho,\beta)}{d_1(\tau_2,\tau_3,\boldsymbol{\delta},\rho,\beta)}, \\ v^2(\boldsymbol{\tau},\boldsymbol{\delta},\gamma,\rho,\beta) &:= \mathrm{Avar}(V_k(\boldsymbol{\tau},\boldsymbol{\delta},\gamma,\rho,\beta)). \end{aligned}$$

Theorem 3 *Let $X_1, \ldots, X_n$ be i.i.d. random variables according to a distribution satisfying $(\mathcal{R}_4)$. Then, if $k \to \infty$ such that $k/n \to 0$, $\sqrt{k}b(n/k)\tilde{b}(n/k) \to \infty$, $\sqrt{k}b(n/k)\tilde{b}(n/k)\breve{b}(n/k) \to \lambda_1$ and $\sqrt{k}b(n/k)\tilde{b}^2(n/k) \to \lambda_2$ we have that*

$$\sqrt{k}b(n/k)\tilde{b}(n/k)[\Lambda_{n,k}(\boldsymbol{\tau},\boldsymbol{\delta}) - \Lambda(\boldsymbol{\tau},\rho,\beta)] \stackrel{\mathcal{D}}{\to} N(\lambda_1 d_2(\boldsymbol{\tau},\boldsymbol{\delta},\rho,\beta,\eta) - \lambda_2 d_3(\boldsymbol{\tau},\boldsymbol{\delta},\rho,\beta), v^2(\boldsymbol{\tau},\boldsymbol{\delta},\gamma,\rho,\beta)).$$

Let $\Lambda'(\boldsymbol{\tau},\rho,\beta) := d\Lambda(\boldsymbol{\tau},\rho,\beta)/d\beta$ and set

$$\begin{aligned} \tilde{d}_2(\boldsymbol{\tau},\boldsymbol{\delta},\rho,\beta,\eta) &:= \frac{d_2(\boldsymbol{\tau},\boldsymbol{\delta},\rho,\beta,\eta)}{\Lambda'(\boldsymbol{\tau},\rho,\beta)}, \\ \tilde{d}_3(\boldsymbol{\tau},\boldsymbol{\delta},\rho,\beta) &:= \frac{d_3(\boldsymbol{\tau},\boldsymbol{\delta},\rho,\beta)}{\Lambda'(\boldsymbol{\tau},\rho,\beta)}, \\ \tilde{v}^2(\boldsymbol{\tau},\boldsymbol{\delta},\gamma,\rho,\beta) &:= \frac{v^2(\boldsymbol{\tau},\boldsymbol{\delta},\gamma,\rho,\beta)}{[\Lambda'(\boldsymbol{\tau},\rho,\beta)]^2}. \end{aligned}$$

A straightforward application of the delta method establishes the asymptotic normality of the estimators $\hat{\beta}_{n,k}(\boldsymbol{\tau}, \boldsymbol{\delta}, \rho)$ and $\hat{\beta}_{n,k}(\boldsymbol{\tau}, \boldsymbol{\delta}, \hat{\rho}_{n,\breve{k}})$ for β. The result is stated formally in the next theorem.

Theorem 4 *Let $X_1, \dots, X_n$ be i.i.d. random variables according to a distribution satisfying $(\mathcal{R}_4)$. Then, if $k \to \infty$ such that $k/n \to 0$, $\sqrt{k}b(n/k)\tilde{b}(n/k) \to \infty$, $\sqrt{k}b(n/k)\tilde{b}(n/k)\breve{b}(n/k) \to \lambda_1$ and $\sqrt{k}b(n/k)\tilde{b}^2(n/k) \to \lambda_2$ we have that*

$$\sqrt{k}b(n/k)\tilde{b}(n/k)[\hat{\beta}_{n,k}(\boldsymbol{\tau}, \boldsymbol{\delta}, \rho) - \beta] \stackrel{\mathcal{D}}{\to} N(\lambda_1 \tilde{d}_2(\boldsymbol{\tau}, \boldsymbol{\delta}, \rho, \beta, \eta) - \lambda_2 \tilde{d}_3(\boldsymbol{\tau}, \boldsymbol{\delta}, \rho, \beta), \tilde{v}^2(\boldsymbol{\tau}, \boldsymbol{\delta}, \gamma, \rho, \beta)).$$

The result continues to hold when ρ in $\hat{\beta}_{n,k}(\boldsymbol{\tau}, \boldsymbol{\delta}, \rho)$ is replaced by an external estimator $\hat{\rho}_{n,\breve{k}}$ which is such that $\hat{\rho}_{n,\breve{k}} - \rho = O_{\mathbb{P}}(1/(\sqrt{\breve{k}}b(n/\breve{k})))$, when $\sqrt{\breve{k}}b(n/\breve{k}) \to \infty$, provided

$$(6) \qquad \frac{\sqrt{k}b(n/k)\tilde{b}(n/k)}{\sqrt{\breve{k}}b(n/\breve{k})} \to 0.$$

Many important members of the class of Pareto-type models like the Fréchet, Burr and student t, satisfy condition $(\mathcal{R}_4)$ with $b(x) \sim a_1 x^{\rho}$, $\tilde{b}(x) \sim a_2 x^{\rho}$ and $\breve{b}(x) = a_3 x^{\rho}$ as $x \to \infty$, for some constants a_1, a_2 and a_3. For such models one easily obtains from Theorem 4 that the asymptotic mean squared error (AMSE) optimal k is of the order $n^{-6\rho/(1-6\rho)}$, which is larger than the orders $n^{-2\rho/(1-2\rho)}$ and $n^{-4\rho/(1-4\rho)}$ one typically achieves for the estimation of first and second order tail parameters, respectively. When using this AMSE optimal k sequence for the estimation of β, the rate of convergence in Theorem 4 is of the order $n^{\rho/(1-6\rho)}$. Also, in this regard, condition (6) can be easily verified to be satisfied when using the AMSE optimal sequences for k and $\breve{k}$.

REFERENCES (RÉFERENCES)

- Beirlant, J., Dierckx, G., Goegebeur, Y., Matthys, G., 1999. Tail index estimation and an exponential regression model. Extremes 2, 177–200.
- Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., 2004. Statistics of Extremes – Theory and Applications. Wiley Series in Probability and Statistics.
- Ciuperca, G., Mercadier, C., 2010. Semi-parametric estimation for heavy tailed distributions. Extremes 13, 55–87.
- de Haan, L., Ferreira, A., 2006. *Extreme Value Theory: An Introduction.* Springer.
- Feuerverger, A., Hall, P., 1999. Estimating a tail exponent by modelling departure from a Pareto distribution. The Annals of Statistics 27, 760–781.
- Fraga Alves, I., de Haan, L., Lin, T., 2006. Third order extended regular variation. Publications de l'Institut Mathématique 80, 109–120.
- Fraga Alves, M.I., Gomes, M.I., de Haan, L., 2003. A new class of semi-parametric estimators of the second order parameter. Portugaliae Mathematica 60, 193–213.
- Geluk, J.L., de Haan, L., 1987. *Regular Variation, Extensions and Tauberian Theorems.* CWI Tract 40, Center for Mathematics and Computer Science, Amsterdam.
- Gnedenko, B.V., 1943. Sur la distribution limite du terme maximum d'une série aléatoire. Annals of Mathematics 44, 423–453.
- Goegebeur, Y., Beirlant, J., de Wet, T., 2010. Kernel estimators for the second order parameter in extreme value statistics. Journal of Statistical Planning and Inference 140, 2632–2652.
- Gomes, M.I., de Haan, L., Peng, L., 2002. Semi-parametric estimation of the second order parameter in statistics of extremes. Extremes 5, 387–414.