# A note on the asymptotic variance at optimal levels of a bias-corrected Hill estimator[†]

M. Ivette Gomes
CEAUL, DEIO, Faculty of Science
University of Lisbon, Portugal

Dinis Pestana
CEAUL, DEIO, Faculty of Science
University of Lisbon, Portugal

Frederico Caeiro
Department of Mathematics and CMA
New University of Lisbon, Portugal

December 8, 2007

For heavy tails, with a positive tail index $\gamma$, classical tail index estimators, like the Hill estimator, are known to be quite sensitive to the number of top order statistics $k$ used in the estimation. Second-order reduced-bias estimators, which are in a certain sense bias-corrected Hill estimators, show much less sensitivity to changes in $k$. Recently, in the minimum-variance reduced-bias (MVRB) tail index estimators, the estimation of the second order parameters in the bias is performed at a level $k_1$ of a larger order than that of the level $k$ at which we compute the tail index estimators, enabling us to keep the asymptotic variance of the new estimators equal to the asymptotic variance of the Hill estimator, for all values of $k$ which enable us to guarantee the asymptotic normality of the Hill statistics. These values of $k$, as well as larger values of $k$, will also enable us to guarantee the asymptotic normality of the reduced-bias estimators. However, to reach the minimal mean squared error of these new MVRB estimators, we need to work with levels $k$ and $k_1$ of the same order. In this note we derive the way the asymptotic variance varies as a function of $q$, the finite limiting value of $k/k_1$, as the sample size $n$ increases towards infinity.

**Keywords and phrases**: Statistics of extremes; reduced-bias semi-parametric estimation; tail

[†]Research partially supported by FCT / POCTI and POCI / FEDER.

index; asymptotic theory.

## 1 Introduction

In *statistics of extremes*, a model $F$ is said to be *heavy-tailed* whenever the *tail function*, $\overline{F} := 1 - F$, is a regularly varying function with a negative index of regular variation equal to $-1/\gamma$, $\gamma > 0$, i.e., whenever $\overline{F} \in RV_{-1/\gamma}$, where for any real $a$ the notation $RV_a$ stands for the class of *regularly varying* functions at infinity with an *index of regular variation* equal to $a$, i.e., positive measurable functions $g$ such that $\lim_{t\to\infty} g(tx)/g(t) = x^a$, for all $x > 0$. Equivalently, the quantile function $U(t) = F^{\leftarrow}(1 - 1/t),\ t \geq 1$, with $F^{\leftarrow}(x) = \inf\{y : F(y) \geq x\}$, is of regular variation with index $\gamma$, i.e.,

$$F \text{ is heavy-tailed} \quad \Longleftrightarrow \quad \overline{F} \in RV_{-1/\gamma} \quad \Longleftrightarrow \quad U \in RV_\gamma \tag{1}$$

for some $\gamma > 0$ (Gnedenko, 1943; de Haan, 1970). Then, we are in the domain of attraction for maxima of an *extreme value* distribution function (d.f.),

$$EV_\gamma(x) = \exp(-(1 + \gamma x)^{-1/\gamma}),\ x \geq -1/\gamma,$$

and we write $F \in \mathcal{D}_{\mathcal{M}}(EV_{\gamma>0})$. The parameter $\gamma$ is the *tail index*, the primary parameter of extreme events.

The *second order parameter*, $\rho$ ($\leq 0$), rules the rate of convergence in the first order condition in (1), and it is the parameter appearing in

$$\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} = \frac{x^\rho - 1}{\rho}, \tag{2}$$

which holds for every $x > 0$, and where $|A|$ must then be in $RV_\rho$ (Geluk and de Haan, 1987). We shall moreover assume that $\rho < 0$. This condition has been widely accepted as an appropriate condition to specify the tail of a Pareto-type distribution in a semi-parametric way, and it holds for most common Pareto-type models, like the Fréchet, the Generalized Pareto, the Burr and the Student's $t$.

In order to obtain information on the order of the asymptotic bias of any second-order reduced-bias tail index estimator, we need further assuming a third order condition, ruling now the rate of convergence in the second order condition in (2), and which guarantees that, for all $x > 0$,

$$\lim_{t\to\infty} \Big(\frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} - \frac{x^\rho - 1}{\rho}\Big)/B(t) = \frac{x^{\rho+\rho'} - 1}{\rho + \rho'}, \tag{3}$$

where $|B|$ must then be in $RV_{\rho'}$. There appears then this extra third order parameter $\rho' \le 0$, which we also assume to be negative. Such a condition has already been used in Gomes, de Haan and Peng (2002) and Fraga Alves, Gomes and de Haan (2003), for the full derivation of the asymptotic behaviour of the $\rho$-estimators there introduced, and in Gomes, Caeiro and Figueiredo (2004a), for the study of specific reduced-bias tail index estimators. More restrictively, and for some details in the paper, we shall assume that we can choose in (3),

$$A(t) = c\ t^\rho =: \gamma\ \beta\ t^\rho,\ \ B(t) = c'\ t^{\rho'} =: \beta'\ t^{\rho'}, \quad \beta,\ \beta' \ne 0 \quad \rho,\ \rho' < 0. \tag{4}$$

For *intermediate* $k$, i.e., a sequence of integers $k = k_n$, $k \in [1, n)$, such that

$$k = k_n \to \infty \quad \text{and} \quad k_n = o(n), \quad \text{as} \quad n \to \infty, \tag{5}$$

we shall consider, as basic statistics, both the log-excesses over the random high level $\{\ln X_{n-k:n}\}$, i.e.,

$$V_{ik} := \ln X_{n-i+1:n} - \ln X_{n-k:n}, \quad 1 \le i \le k < n, \tag{6}$$

and the scaled log-spacings,

$$W_i := i \left\{\ln X_{n-i+1:n} - \ln X_{n-i:n}\right\}, \quad 1 \le i \le k < n. \tag{7}$$

where $X_{i:n}$ denotes, as usual, the $i$-th ascending order statistic (o.s.), $1 \le i \le n$, associated to an independent, identically distributed (i.i.d.) random sample $(X_1, X_2, \cdots, X_n)$. It is well known that for intermediate $k$, i.e. if (5) holds, and under the first order framework in (1), the log-excesses, $V_{ik}$, $1 \le i \le k$, in (6), are approximately the $k$ o.s.'s from an exponential sample of size $k$ and mean value $\gamma$. Also, under the same conditions, the scaled log-spacings, $W_i$, $1 \le i \le k$, in (7), are approximately i.i.d. and exponential with mean value $\gamma$. Consequently the Hill estimator of $\gamma$ (Hill, 1975),

$$H(k) \equiv H_n(k) = \frac{1}{k}\sum_{i=1}^{k} V_{ik} = \frac{1}{k}\sum_{i=1}^{k} W_i, \tag{8}$$

is consistent for the estimation of $\gamma$ under the first order framework and for intermediate $k$. Note that the Hill estimator in (8) is the maximum likelihood estimator of the tail index $\gamma$, under a strict Pareto model, with d.f. $F_{_P}(x) = 1 - x^{-1/\gamma},\ x \geq 1$. Under the second order framework in (2) and for intermediate $k$, i.e., if (5) holds, the asymptotic distributional representation

$$H_n(k) \stackrel{d}{=} \gamma + \frac{\gamma}{\sqrt{k}}\ Z_k^{(1)} + \frac{A(n/k)}{1-\rho} + o_p(A(n/k)) \tag{9}$$

holds, where $Z_k^{(1)} = \sqrt{k}\big(\sum_{i=1}^k E_i/k - 1\big)$, with $\{E_i\}$ i.i.d. standard exponential random variables (r.v.'s), is an asymptotically standard normal r.v. (de Haan and Peng, 1998).

The most simple minimum-variance reduced-bias (MVRB) estimators in the literature are the bias-corrected Hill estimators in Caeiro, Gomes and Pestana (2005), with the functional form,

$$\overline{H}_{\hat\beta,\hat\rho}(k) := H(k)\Big(1 - \hat\beta\left(\frac{n}{k}\right)^{\hat\rho}/(1-\hat\rho)\Big), \tag{10}$$

dependent upon the Hill estimator $H(k)$ and $(\hat\beta, \hat\rho)$, adequate consistent estimators of the second order parameters $\beta$ and $\rho$, respectively. If $(\hat\beta, \hat\rho)$ is a consistent estimator of $(\beta, \rho)$ and $\hat\rho - \rho = o_p(1/\ln n)$, $\sqrt{k}(\overline{H}_{\hat\beta,\hat\rho}(k) - \gamma)$ is asymptotically normal with mean value equal to zero and variance $\gamma^2$ at least for intermediate values $k$ such that $\sqrt{k}A(n/k) \to \lambda$, finite (Caeiro *et al.*, 2005). Further information on the order of the asymptotic bias of the reduced-bias estimator in (10), for a slightly more restrictive class of models than the one in (3), is provided in Caeiro and Gomes (2007).

In Section 2 of this paper, we shall provide details on conditions under which we are able to keep the asymptotic variance of the MVRB-estimators in (10) equal to $\gamma^2$, together with their behavior under a third order framework. In Section 3, we shall briefly review the estimation of the second order parameters $\beta$ and $\rho$. Next, in Section 4, we provide some information on the asymptotic behavior of $\sqrt{k}\{\overline{H}_{\hat\beta,\hat\rho}(k) - \gamma\}$, whenever $\sqrt{k}A(n/k) \to \infty$, $\sqrt{k}\ A^2(n/k) \to \lambda_{_A}$ and $\sqrt{k}\ A(n/k)B(n/k) \to \lambda_{_B}$, both finite, $\lambda_{_A}$ or $\lambda_{_B} \neq 0$, and when we consider the estimator of $\rho$ used before in papers like Caeiro *et al.* (2005), computed at any "optimal" level $k_1$, in the sense of a level such that $\sqrt{k_1}\ A^2(n/k_1) \to \lambda_{_{A_1}}$ and $\sqrt{k_1}\ A(n/k_1)B(n/k_1) \to \lambda_{_{B_1}}$, both finite, $\lambda_{_{A_1}}$ or $\lambda_{_{B_1}} \neq 0$. For the class of models under consideration, i.e., models for which (3) holds, with $A$ and $B$ chosen as in (4) for arbitrary $\rho,\ \rho' < 0$, we thus have $k/k_1 \to q > 0$, whenever $n \to \infty$.

## 2 Asymptotic behaviour of the MVRB-estimators under a third order framework

In a trial to keep the asymptotic variance of the reduced-bias estimators in (10) equal to $\gamma^2$, we may state the following theorem, a generalization of Theorem 3.1 in Caeiro *et al.* (2005), where $U_n \overset{p}{\sim} V_n$ denotes that $U_n/V_n$ converges in probability towards one, as $n \to \infty$.

**Theorem 1.** *Under the third order framework in* (3)*, with* $A$ *and* $B$ *chosen as in* (4)*, if* (5) *holds, and with* $Z_k^{(1)}$ *the asymptotically standard normal r.v. in* (9)*,*

$$\overline{H}_{\beta,\rho}(k) \overset{d}{=} \gamma + \frac{\gamma}{\sqrt{k}}\, Z_k^{(1)} - A(n/k)\Big(\frac{A(n/k)}{\gamma(1-\rho)^2} - \frac{B(n/k)}{1-\rho-\rho'} + O_p\Big(\frac{1}{\sqrt{k}}\Big)\Big)(1+o_p(1)), \quad (11)$$

*with* $\overline{H}_{\hat\beta,\hat\rho}(k)$ *given in* (10)*. Consequently, even if*

$$\sqrt{k}\ A(n/k) \to \infty,\ \text{with}\ \sqrt{k}\ A^2(n/k) \to \lambda_A\ \text{and}\ \sqrt{k}\ A(n/k)\ B(n/k) \to \lambda_B, \quad (12)$$

$\lambda_A$ *and* $\lambda_B$ *finite,*

$$\sqrt{k}\ \left(\overline{H}_{\beta,\rho}(k) - \gamma\right) \underset{n\to\infty}{\overset{d}{\longrightarrow}} Normal\left(b_{\overline{H}}, \gamma^2\right)$$

*where the asymptotic bias is*

$$b_{\overline{H}} \equiv b_{\overline{H}}(\gamma,\rho,\rho') \equiv ABIAS_{\overline{H}} = -\lambda_A/(\gamma(1-\rho)^2) + \lambda_B/(1-\rho-\rho') =: \lambda_A u_{\overline{H}} + \lambda_B v_{\overline{H}}. \quad (13)$$

*Let* $(\hat\beta,\hat\rho)$ *be any consistent estimator of the vector of second order parameters* $(\beta,\rho)$*, such that*

$$\hat\rho - \rho = o_p(1/\ln n),\ \text{as}\ n \to \infty. \quad (14)$$

*Then, with* $a_{\overline{H}} = -1/(1-\rho)$*,*

$$\overline{H}_{\hat\beta,\hat\rho}(k) - \overline{H}_{\beta,\rho}(k) \overset{p}{\sim} a_{\overline{H}} A(n/k)\Big\{\big(\hat\beta-\beta\big)/\beta + (\hat\rho-\rho)\left[\ln(n/k) - a_{\overline{H}}\right]\Big\}. \quad (15)$$

*Consequently,* $\sqrt{k}\{\overline{H}_{\hat\beta,\hat\rho}(k) - \gamma\}$ *are asymptotically normal with null mean value and variance* $\sigma_0^2 = \gamma^2$*, not only when* $\sqrt{k}\ A(n/k) \to 0$*, but also whenever* $\sqrt{k}\ A(n/k) \to \lambda$*, finite. This same result still holds for levels* $k$ *such that* $\sqrt{k}\ A(n/k) \to \infty$*, provided that* $\sqrt{k}\ A^2(n/k) \to 0$*,* $\sqrt{k}\ A(n/k)B(n/k) \to 0$*, and* $\hat\beta - \beta$ *as well as* $(\hat\rho-\rho)\ln n$ *are* $o_p\big(1/\sqrt{k}\ A(n/k)\big)$*.*

*Proof.* Under the conditions in the theorem, the asymptotic distributional representation

$$H_n(k) \overset{d}{=} \gamma + \frac{\gamma}{\sqrt{k}}\, Z_k^{(1)} + \left(\frac{A(n/k)}{1-\rho} + \frac{A(n/k)\ B(n/k)}{1-\rho-\rho'} + O_p\Big(\frac{A(n/k)}{\sqrt{k}}\Big)\right)(1+o_p(1))$$

holds. On the other hand, as $\overline{H}_{\beta,\rho}(k) = H_n(k)(1 - A(n/k)/(\gamma(1-\rho)))$, we get (11), and consequently, the stated asymptotic normality of $\sqrt{k}\,\left(\overline{H}_{\beta,\rho}(k) - \gamma\right)$.

Next, and directly from the expression of $\overline{H}_{\beta,\rho}(k)$, we get

$$\frac{\partial \overline{H}_{\beta,\rho}}{\partial \beta} \overset{p}{\sim} -\frac{A(n/k)}{1-\rho}, \qquad \frac{\partial \overline{H}_{\beta,\rho}}{\partial \rho} \overset{p}{\sim} -\frac{A(n/k)}{1-\rho}\Big(\ln(n/k) + \frac{1}{1-\rho}\Big).$$

The use of Cramer's delta-method, together with the validity of (14), enables us to write

$$\overline{H}_{\hat\beta,\hat\rho}(k) = \overline{H}_{\beta,\rho}(k) - \frac{A(n/k)}{1-\rho}\Big\{\frac{\hat\beta-\beta}{\beta} + (\hat\rho-\rho)\Big(\ln(n/k) + \frac{1}{1-\rho}\Big)\Big\}(1+o_p(1)).$$

Consequently, (15) follows, as well as the remaining of the theorem, provided that we pay attention to the validity of (11). □

**Remark 1.** *Note that the optimal level, in the sense of minimal asymptotic mean square error, for the estimation of the tail index* $\gamma$ *through the reduced-bias estimator* $\overline{H}_{\beta,\rho}(k)$ (*assuming thus that* $\beta$ *and* $\rho$ *are known*) *is such that* (12) *holds. An important question to answer is how far can this same result be true for the reduced-bias estimator* $\overline{H}_{\hat\beta,\hat\rho}(k)$ *in* (10).

# 3 A brief review of the second order parameters' estimators

## 3.1 The estimation of $\rho$

We shall base the estimation of $\rho$ on estimators of the type of the ones in Fraga Alves *et al.* (2003). Such a class of estimators has been first parameterised in a tuning parameter $\tau \ge 0$, but $\tau$ can be more generally considered as a real number (Caeiro and Gomes, 2006), and can be defined as,

$$\widehat{\rho}_\tau(k) \equiv \widehat{\rho}_n(k;\tau) := -\left|\frac{3(T_n^{(\tau)}(k)-1)}{T_n^{(\tau)}(k)-3}\right|, \quad T_n^{(\tau)}(k) := \frac{\left(M_n^{(1)}(k)\right)^{\tau} - \left(M_n^{(2)}(k)/2\right)^{\tau/2}}{\left(M_n^{(2)}(k)/2\right)^{\tau/2} - \left(M_n^{(3)}(k)/6\right)^{\tau/3}}, \quad (16)$$

for $\tau \neq 0$ and with the usual continuation for $\tau = 0$, where, with $V_{ik}$ given in (6),

$$M_n^{(j)}(k) := \frac{1}{k}\sum_{i=1}^{k} V_{ik}^j, \; j \ge 1 \qquad \Big[M_n^{(1)} \equiv H, \text{ the Hill estimator in (8)}\Big].$$

We shall here summarize a result proved in Fraga Alves *et al.* (2003), making now explicit the random behaviour of the term leading to the asymptotic variance, needed later on, when dealing with the estimation of the three parameters, $\gamma$, $\beta$ and $\rho$, at levels of the same order.

**Theorem 2** (Fraga Alves *et al.*, 2003)**.** *Under the second order framework in* (2)*, if $k$ is intermediate, i.e.,* (5) *holds, and if $\sqrt{k}\ A(n/k) \to \infty$, as $n \to \infty$, the statistics $\widehat{\rho}_n(k;\tau)$ in* (16) *converge in probability towards $\rho$, as $n \to \infty$, for any real $\tau$. If* (3) *holds, with $\rho < 0$, we may further guarantee that there exist constants $(u_{\rho,\tau}, v_{\rho,\rho'}, \sigma_\rho)$ and an asymptotically standard normal r.v. $W_k^R$, such that*

$$\widehat{\rho}_n(k;\tau) - \rho \ \stackrel{d}{=}\ \frac{\sigma_\rho\ W_k^R}{\sqrt{k}\ A(n/k)} + \big(u_{\rho,\tau} A(n/k) + v_{\rho,\rho'} B(n/k)\big)(1 + o_p(1)). \tag{17}$$

*Moreover, with $Z_k^{(\alpha)} = \frac{1}{\sqrt{k}}\ \sum_{i=1}^{k} E_i^\alpha/\Gamma(\alpha+1) - \sqrt{k}$, for any $\alpha \ge 1$, $\Gamma(t)$ denoting the complete Gamma function, we may write*

$$W_k^R = \Big((3-\rho)Z_k^{(1)} - (3-2\rho)Z_k^{(2)} + (1-\rho)Z_k^{(3)}\Big)\,/\sqrt{2\rho^2 - 2\rho + 1}. \tag{18}$$

*Consequently, if* (12) *holds, $\sqrt{k}\ A(n/k)\,(\widehat{\rho}_n(k;\tau) - \rho)$ is asymptotically normal with a mean value $\lambda_{_A} u_{\rho,\tau} + \lambda_{_B} v_{\rho,\rho'}$ and variance*

$$\sigma_\rho^2 \equiv \sigma_\rho^2(\gamma) = \big(\gamma(1-\rho)^3/\rho\big)^2\,\big(2\rho^2 - 2\rho + 1\big)\,. \tag{19}$$

Let us next assume that $k_1$ is "optimal" for the estimation of $\rho$, in the sense of a value $k_1$ that enables us to guarantee the asymptotic normality of the $\rho$-estimator with a non-null asymptotic bias. That level $k_1$ is then such that $\sqrt{k_1}\ A(n/k_1)B(n/k_1) \ \to\ \lambda_{_{B_1}}$, finite, and $\sqrt{k_1}\ A^2(n/k_1) \ \to\ \lambda_{_{A_1}}$, also finite, with at least one of them non-null, let us say $\lambda_{_{B_1}}$. We then get $k_1 = O\big(n^{-2(\rho+\rho')/(1-2(\rho+\rho'))}\big)$. Denoting $\hat\rho = \hat\rho(k_1;\tau)$ for any $\tau$, i.e., any of the $\rho$-estimators in this section computed at such a level $k_1$, $\{\hat\rho - \rho\}$ is, in probability, of the order of $1/\big(\sqrt{k_1}A(n/k_1)\big) = O\big(n^{\rho'/(1-2(\rho+\rho'))}\big) = o(1/\ln n)$, i.e., (14) holds, and this is the crucial condition on $\hat\rho$, needed in Theorem 1, if we want to keep equal to $\gamma^2$ the asymptotic variance of the reduced-bias estimator in (10). At the current state-of-the-art, such a $k_1$ has only a theoretical interest. From a practical point of view additional research is needed, in order to have adaptive ways of selecting this optimal threshold $k_1$. For non-optimal, but practical ways of choosing $k_1$, see Gomes, de Haan and Henriques Rodrigues (2004b) and Gomes, Martins and Neves (2007).

### 3.2 Estimation of $\beta$ based on the scaled log-spacings

We have here considered the $\beta$-estimator obtained in Gomes and Martins (2002) and based on the scaled log-spacings $W_i$, $1 \le i \le k$, in (7). On the basis of any consistent estimator $\hat\rho$ of the second order parameter $\rho$, we shall consider the $\beta$-estimator, $\hat\beta(k;\hat\rho)$, where, for any $\rho < 0$,

$$\hat\beta(k;\rho) := \frac{\left(\frac{k}{n}\right)^{\rho}\left\{\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\rho}\right)\left(\frac{1}{k}\sum_{i=1}^{k}W_i\right)-\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\rho}\ W_i\right)\right\}}{\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\rho}\right)\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\rho}\ W_i\right)-\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-2\rho}\ W_i\right)}. \tag{20}$$

Gomes and Martins (2002) kept up to the second order framework and studied the behaviour of $\hat\beta(k;\rho)$ in (20). Here, we go into the third order framework in (3), and assume, just as in Gomes *et al.* (2004b), that $\beta$ and $\rho$ are going to be estimated at the same level $k$. We state, without proof:

**Theorem 3.** *Under the second order framework in* (2)*, with* $\rho < 0$ *and* $\hat\rho(k;\tau) \equiv \widehat{\rho_\tau}(k)$ *in* (16)*, the rate of convergence of* $\hat\beta(k;\hat\rho(k;\tau))$ *is of the order of* $\left\{\ln(n/k)/(\sqrt{k}\ A(n/k))\right\}$*, which must converge towards zero, so that* $\hat\beta(k;\hat\rho(k;\tau))$ *is consistent for the estimation of* $\beta$*. If* (14) *holds for* $\hat\rho(k;\tau)$*,*

$$\hat\beta(k;\hat\rho(k;\tau)) - \beta \ \overset{p}{\sim}\ -\beta\ln(n/k)\left(\hat\rho(k;\tau)-\rho\right). \tag{21}$$

*If apart from* $\sqrt{k}\ A(n/k)/\ln(n/k) \to \infty$*, we assume that, under the third order framework in* (3)*,* (12) *holds, then, with* $\sigma_\rho$ *given explicitly in* (19)*,* $\sqrt{k}\ A(n/k)\big(\beta - \hat\beta(k;\hat\rho(k;\tau))\big)/(\beta\ln(n/k))$ *is asymptotically* $Normal\big(\lambda_{_A} u_{\rho,\tau} + \lambda_{_B} v_{\rho,\rho'},\ \sigma_\rho^2\big)$*, with* $(u_{\rho,\tau}, v_{\rho,\rho'}, \sigma_\rho)$ *implicitly given in* (17)*.*

## 4 The estimation of $\gamma$, $\beta$ and $\rho$ at levels of the same order

On the basis of Theorems 1, 2 and 3, we state the following result.

**Proposition 1.** *For the class of models in* (3)*, for which* $A$ *and* $B$ *can be chosen as in* (4)*, let us consider a value* $k_1$ *"optimal" for the estimation of* $\rho$*, in the sense that* $\sqrt{k_1}\ A^2(n/k_1) \to \lambda_{_{A_1}}$*, finite, and* $\sqrt{k_1}\ A(n/k_1)B(n/k_1) \to \lambda_{_{B_1}}$*, also finite, with at least one of them non-null. Let us then denote* $(\hat\rho,\hat\beta) = \big(\hat\rho(k_1,\tau), \hat\beta(k_1,\hat\rho)\big)$*,* $\hat\rho(k;\tau)$ *and* $\hat\beta(k,\rho)$ *given in expressions* (16) *and* (20) *respectively. We can then write*

$$\hat{R}_n(k;k_1) := \sqrt{k}\left\{\overline{H}_{\hat\beta,\hat\rho}(k) - \overline{H}_{\beta,\rho}(k)\right\} \overset{p}{\sim} \sqrt{k}A(n/k)\,(\hat\rho-\rho)\,\frac{\ln(k/k_1)}{1-\rho}$$

$$= O_p\Big(\Big(\frac{k}{k_1}\Big)^{\frac{1}{2}-\rho}\ln\Big(\frac{k}{k_1}\Big)\Big). \qquad (22)$$

*Consequently, if* $k = o(k_1)$, $\hat{R}_n(k;k_1)$ *converges in probability towards zero, as* $n\to\infty$, *and* $\sqrt{k}\big(\overline{H}_{\hat\beta,\hat\rho}(k)-\gamma\big)$ *is asymptotically normal with null mean value and variance equal to* $\gamma^2$.

*Proof.* From equations (15) and (21),

$$\hat{R}_n(k;k_1) \overset{p}{\sim} a_{\overline{H}}\,(\widehat\rho-\rho)\,\sqrt{k}A\,(n/k)\,\big(\ln(k/k_1) - a_{\overline{H}}\big)\,.$$

If $\sqrt{k_1}\;A^2(n/k_1) \to \lambda_{A_1}$ and $\sqrt{k_1}\;A(n/k_1)B(n/k_1) \to \lambda_{B_1}$, both finite, then $\widehat\rho-\rho = O_p\left(1/\left(\sqrt{k_1}\;A(n/k_1)\right)\right)$ and (22) follows together with the remaining of the proposition. □

**Remark 2.** *Note that if* $k = o(k_1)$, *the values* $k$ *such that* $\sqrt{k}A(n/k)\to\infty$, *in Theorem 1, should be such that both* $\sqrt{k}A^2(n/k)$ *and* $\sqrt{k}A(n/k)B(n/k)$ *converge towards zero, as* $n\to\infty$, *as already mentioned.*

Let us assume now that we want to work with levels $k$ such that $\sqrt{k}\;A^2(n/k)\to\lambda_A$ and $\sqrt{k}\;A(n/k)B(n/k)\to\lambda_B$, both finite, as $n\to\infty$, $\lambda_A$ or $\lambda_B\neq 0$, i.e., levels possibly of the same order of the ones leading to the "optimal" estimation of $\beta$ and $\rho$. We may then state the following:

**Theorem 4.** *For models in* (3), *jointly with* (4), *let us assume that* $\sqrt{k}\;A(n/k)\to\infty$, *with* $\sqrt{k}\;A^2(n/k)\to\lambda_A$ *and* $\sqrt{k}\;A(n/k)B(n/k)\to\lambda_B$, *both finite. As in Theorem 1, let* $k_1$ *be "optimal" for the estimation of* $\rho$, *in the sense that* $\sqrt{k_1}\;A(n/k_1)B(n/k_1)\to\lambda_{B_1}$, *finite, and* $\sqrt{k_1}\;A^2(n/k_1)\to\lambda_{A_1}$, *also finite, with at least one of them non-null. More specifically, let* $k$ *and* $k_1$ *be sequences of intermediate integers such that* $k/k_1\to q>0$, *finite, and let us consider, for any real* $\tau$,

$$\widetilde{H}_\tau(k;k_1) = \overline{H}_{\hat\beta(k_1;\hat\rho(k_1;\tau)),\hat\rho(k_1;\tau)}(k), \qquad (23)$$

*with* $\overline{H}$ *the estimator* (10). *Then,*

$$\sqrt{k}\left(\widetilde{H}_\tau(k;k_1)-\gamma\right)\underset{n\to\infty}{\overset{d}{\to}} Normal\left(b^*,\ \sigma^2(q)\right),$$

*where, with* $(u_{\overline{H}}, v_{\overline{H}})$ *defined in* (13), $(u_{\rho,\tau}, v_{\rho,\rho'})$ *the constants in* (17) *and* $a_{\overline{H}} = -1/(1-\rho)$,

$$b^* = \lambda_{_A} \left(u_{\overline{H}} + u_{\rho,\tau}\ q^\rho a_{\overline{H}} \left(\ln q - a_{\overline{H}}\right)\right) + \lambda_{_B} \left(v_{\overline{H}} + v_{\rho,\rho'}\ q^{\rho'} a_{\overline{H}} \left(\ln q - a_{\overline{H}}\right)\right),$$

*and*

$$\sigma^2(q) = \sigma^2(q;\gamma,\rho) = \gamma^2\Big(1 + q^{1-2\rho}\big(\ln q + 1/(1-\rho)\big)^2(1-\rho)^4(2\rho^2 - 2\rho + 1)/\rho^2\Big), \tag{24}$$

*i.e., we get the same rate of convergence, of the order of* $1/\sqrt{k}$, *for* $\widetilde{H}_\tau(k;k_1)$, *but with an asymptotic variance dependent upon* $q$. *Such an asymptotic variance is equal to* $\gamma^2$ *for* $q = 0$ *and* $q = q_0 = \exp(-1/(1-\rho))$, *and increases with* $q$, *after the value* $q_0$.

*Proof.* Let us think that $k/k_1 \to q > 0$. Since $\sqrt{k_1}\ A(n/k_1) \sim \sqrt{k}\ A(n/k)\ q^{\rho-\frac{1}{2}}$, $A(n/k_1) \sim q^\rho A(n/k)$ and $B(n/k_1) \sim q^{\rho'} B(n/k)$, we can write

$$\begin{aligned}\sqrt{k}\left(\widetilde{H}_\tau(k;k_1) - \gamma\right) &= \gamma\ Z_k^{(1)} + \sqrt{k}A(n/k)\left(u_{\overline{H}}A(n/k) + v_{\overline{H}}\ B(n/k)\right)(1 + o_p(1)) \\ &+ \left(\sigma_\rho W_{k_1}^R + \sqrt{k}\ A(n/k) q^{\rho-\frac{1}{2}}\left(u_{\rho,\tau}A(n/k)\ q^\rho + v_{\rho,\rho'}B(n/k)\ q^{\rho'}\right)(1 + o_p(1))\right) \\ &\qquad \times q^{\frac{1}{2}-\rho}\ a_{\overline{H}}\left(\ln q - a_{\overline{H}} + o(1)\right)(1 + o(1)),\end{aligned}$$

with $a_{\overline{H}} = -1/(1-\rho)$, $(u_{\overline{H}}, v_{\overline{H}})$ defined in (13) and $(u_{\rho,\tau}, v_{\rho,\rho'}, \sigma_\rho, W_k^R)$ given in (17), $W_k^R$ and $\sigma_\rho$ made explicit in (18) and (19), respectively. Then

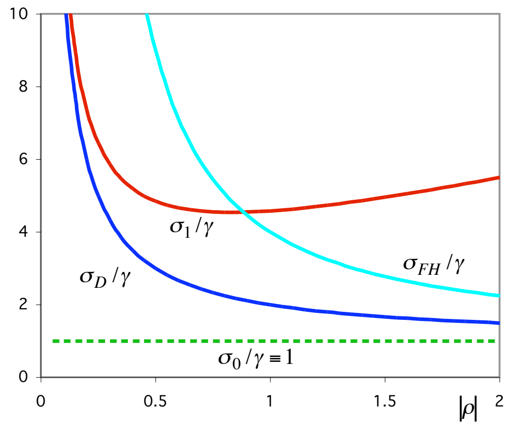
$$\begin{aligned}\sqrt{k}\left(\widetilde{H}_\tau(k;k_1) - \gamma\right) &= \gamma\ Z_k^{(1)} + \sigma_\rho\ q^{\frac{1}{2}-\rho}\ a_{\overline{H}}\left(\ln q - a_{\overline{H}}\right)\ W_{k_1}^R \\ &+ \sqrt{k}\ A^2(n/k)\left(u_{\overline{H}} + u_{\rho,\tau}\ q^\rho a_{\overline{H}}\left(\ln q - a_{\overline{H}}\right)\right)(1 + o_p(1)) \\ &+ \sqrt{k}\ A(n/k)B(n/k)\left(v_{\overline{H}} + v_{\rho,\rho'}\ q^{\rho'} a_{\overline{H}}\left(\ln q - a_{\overline{H}}\right)\right)(1 + o_p(1)).\end{aligned}$$

As $Cov\big(Z_k^{(1)},\ W_{k_1}^R\big) = 0$, the variance of $\left\{\gamma\ Z_k^{(1)} + \sigma_\rho\ q^{\frac{1}{2}-\rho}\ a_{\overline{H}}\left(\ln q - a_{\overline{H}}\right)\ W_{k_1}^R\right\}$ is the value $\sigma^2(q;\ \gamma,\rho)$ in (24), and the remaining of the proof follows straightforwardly. □

**Remark 3.** *If we compare Theorems 1 and 4, we see that the estimation of* $\gamma$, $\beta$ *and* $\rho$ *at the same level* $k$ *induces an increase in the asymptotic variance of the final* $\gamma$*-estimator. The use of* $q = 1$ *in* (24) *leads us to the asymptotic variance* $\sigma_1^2 = \sigma^2(1) = \gamma^2\{1 + ((1-\rho)/\rho)^2 - 2(1-\rho)^3/\rho\}$, *greater than* $\sigma_0^2 = \gamma^2$, *the value associated to* $q = 0$ *in* (24). *As noticed before in Gomes and Martins* (2002), *the asymptotic variance of the estimator in Feuerverger and Hall* (1999) (*where* *also the three parameters are computed at the same level $k$) is given by $\sigma^2_{FH} := \gamma^2\left((1-\rho)/\rho\right)^4$, the asymptotic variance also achieved by Peng and Qi* (2004)*, for an approximate second-order reduced-bias maximum likelihood tail index estimator. Moreover, it is also known that if we estimate $\rho$ at a level $k_1$, but estimate both $\gamma$ and $\beta$ at the same level $k$, we already induce an extra increase in the asymptotic variance of the final $\gamma$-estimator, which is then equal to $\sigma^2_D = \gamma^2((1-\rho)/\rho)^2$, the minimal asymptotic variance of any "asymptotically unbiased" estimator in Drees' class of functionals* (*Drees,* 1998)*. We have*

$$\sigma_0 < \sigma_D < \sigma_1 < \sigma_{FH} \quad \textit{if} \quad |\rho| < 0.8832 \quad \textit{and} \quad \sigma_0 < \sigma_D < \sigma_{FH} < \sigma_1 \quad \textit{if} \quad |\rho| > 0.8832.$$

*In Figure* 1 *we provide both a picture and some values of $\sigma_0/\gamma \equiv 1$, $\sigma_D/\gamma$, $\sigma_1/\gamma$ and $\sigma_{FH}/\gamma$, as functions of $|\rho|$.*




| $\lvert\rho\rvert$ | $\sigma_0/\gamma$ | $\sigma_D/\gamma$ | $\sigma_1/\gamma$ | $\sigma_{FH}/\gamma$ |
|---|---|---|---|---|
| 0.1 | 1.00 | 11.00 | 12.19 | 121.00 |
| 0.2 | 1.00 | 6.00 | 7.37 | 36.00 |
| 0.3 | 1.00 | 4.33 | 5.87 | 18.78 |
| 0.4 | 1.00 | 3.50 | 5.19 | 12.25 |
| 0.5 | 1.00 | 3.00 | 4.85 | 9.00 |
| 1.0 | 1.00 | 2.00 | 4.58 | 4.00 |
| 1.5 | 1.00 | 1.67 | 4.96 | 2.78 |
| 2.0 | 1.00 | 1.50 | 5.50 | 2.25 |
| 2.5 | 1.00 | 1.40 | 6.10 | 1.96 |
| 3.0 | 1.00 | 1.33 | 6.74 | 1.78 |

Figure 1: Asymptotic standard deviations, $\sigma_0$, $\sigma_D$ and $\sigma_1$, together with $\sigma_{FH}$, for $\gamma = 1$

*It is obvious from Figure* 1 *that, whenever possible, it seems convenient to estimate both $\beta$ and $\rho$ externally, at a $k_1$-value higher than the one used for the estimation of the tail index $\gamma$, if we want to work with a tail index estimator potentially better than the Hill estimator for all $k$.*

The pattern of $\sigma^2(q;\ \gamma,\rho)$ in (24), as a function of $q$, is of the same type for all $(\gamma,\ \rho)$, and is pictured in Figure 2: this variance converges towards $\sigma^2_0 = \gamma^2$, as $q \to 0$, next increases till a value slightly larger than $\gamma^2$, then decreases again till $\gamma^2$ at $q_0 = \exp\left(-1/(1-\rho)\right)$, $\exp(-1) < q_0 < 1$, and finally increases fast, taking the value $\sigma^2_1 = \gamma^2\big(1 + ((1-\rho)/\rho)^2 - 2(1-\rho)^3/\rho\big)$, for $q = 1$.

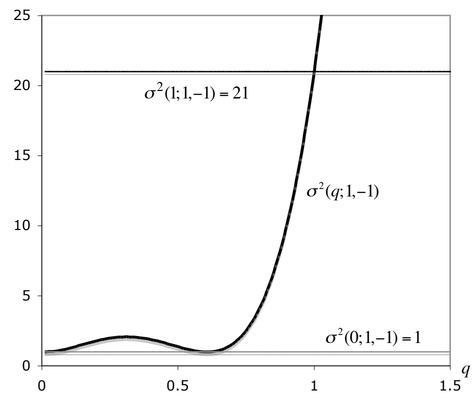



Figure 2: Pattern of $\sigma^2(q;\gamma,\rho)$, as a function of $q$, for $\gamma = 1$ and $\rho = -1$.

As the variance of the estimator $\widetilde{H}_\tau(k;k_1)$ in (23), with $k/k_1 \to q > 0$, is well approximated by $\sigma^2(q;\gamma,\rho)/k$, we provide in Figure 3, the pattern of $\sigma^2(q;\gamma,\rho)/q$, which indicates the behaviour of the variance of our estimator as a function of $q$.

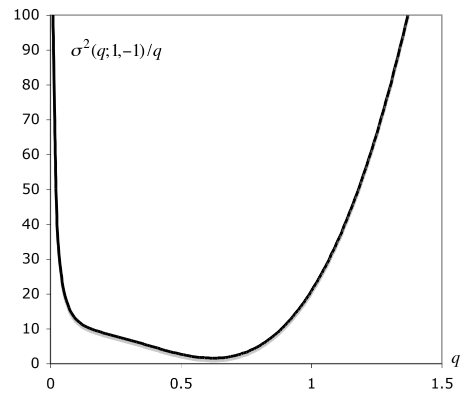



Figure 3: Pattern of $\sigma^2(q;\gamma,\rho)/q$, as a function of $q$, for $\gamma = 1$ and $\rho = -1$, an indicator of the variance of the estimator in equation (23).

It is thus obvious that we should base the tail index estimation on a number $k$ of top o.s., smaller than $k_1$, the number of o.s. that should be optimally used for the estimation of the second order parameters $\beta$ and $\rho$. The optimal rate $k/k_1$ depends obviously in $\rho$, but it is not a long way from 0.5 for the most common models. Indeed, for $\rho = -0.25, -0.5, -1, -1.5, -2$ we get $q_{min} := \arg\min_q \sigma^2(q;\gamma,\rho)/q = 0.46, 0.53, 0.62, 0.68, 0.72$, respectively, a value quite close to $q_0 = \exp(-1/(1-\rho))$.

## References


[1] Caeiro, F. and Gomes, M.I. (2006). A new class of estimators of a "scale" second order parameter. *Extremes* **9**, 193-211, 2007.

[2] Caeiro, C. and Gomes, M.I. (2007). *Minimum-variance reduced-bias tail index and quantile estimation*, UNL Preprint (submitted).

[3] Caeiro, F., Gomes, M.I. and Pestana, D. D. (2005). Direct reduction of bias of the classical Hill estimator. *Revstat* **3**:2, 113-136.

[4] Drees, H. (1998). A general class of estimators of the extreme value index. *J. Statist. Planning and Inference* **98**, 95-112.

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

[6] Fraga Alves, M.I., Gomes, M.I. and de Haan, L. (2003). A new class of semi-parametric estimators of the second order parameter. *Portugaliae Mathematica* **60**, 193-214.

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

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

[9] Gomes, M.I., Caeiro, F., and Figueiredo, F. (2004a). Bias reduction of an extreme value index estimator trough an external estimation of the second order parameter. *Statistics* **38**(6), 497-510.

[10] Gomes, M.I., de Haan, L. and Henriques Rodrigues, L. (2004b). Tail index estimation for heavy-tailed models: accommodation of bias in the weighted log-excesses. *J. Royal Statistical Society*, DOI: 10.1111/j.1467-9869.2007.00620.x, 2007.

[11] Gomes, M.I., Haan, L. de, and Peng, L. (2002). Semi-parametric estimation of the second order parameter — asymptotic and finite sample behaviour. *Extremes* **5**:4, 387-414.

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

[13] Gomes, M.I., Martins, M.J., and Neves, M. (2007). Improving second order reduced bias extreme value index estimation. *Revstat* **5**:2, 177–207.

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

[15] Haan, L. de, and Peng, L. (1998). Comparison of tail index estimators. *Statistica Neerlandica* **52**, 60-70.

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

[17] Peng, L. and Qi, Y. (2004). Estimating the first and second order parameters of a heavy tailed distribution. *Australian and New Zealand J. of Statist.* **46**, 305-312.