A Note on the Asymptotic Variance at Optimal Levels of a Bias-corrected Hill Estimator

Gomes, M. Ivette and Dinis Pestana
Faculdade de Ciências de Lisboa (DEIO) and CEAUL, Universidade de Lisboa
Campo Grande, 1749-016, Lisboa, Portugal
E-mails: ivette.gomes@fc.ul.pt and dinis.pestana@fc.ul.pt

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., if and only if \( F \in RV_{-1/\gamma} \), where the notation \( RV_\alpha \) stands for the class of regularly varying functions at infinity with index of regular variation equal to \( \alpha \), i.e., positive measurable functions \( g \) such that \( \lim_{t \to \infty} g(tx)/g(t) = x^\alpha \), for all \( x > 0 \). Equivalently, the quantile function \( U(t) = F^{-(1-1/t)} \), \( t \geq 1 \), with \( F^{-}(x) = \inf\{y : F(y) \geq x\} \), is of regular variation with index \( \gamma \), i.e., \( F \) is heavy-tailed \( \iff F \in RV_{-1/\gamma} \iff U \in RV_{\gamma} \), for some \( \gamma > 0 \). 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 \( F \in D_M(EV_{\gamma>0}) \). The parameter \( \gamma \) is the tail index, the primary parameters of extreme events.

The second order parameter, \( \rho (\leq 0) \), rules the rate of convergence in the first order condition, and is the parameter appearing in \( \lim_{n \to \infty} (\ln U(tx) - \ln U(t) - \gamma \ln x)/A(t) = (x^\rho - 1)/\rho \), which we assume to hold for every \( x > 0 \), and where \(|A(t)|\) must then be in \( RV_\rho \) (Geluk and de Haan, 1987). We shall 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 true 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 asymptotic bias of the second-order reduced-bias tail index estimators, we need further assuming a third order condition, ruling now the rate of convergence in the second order condition, and which guarantees that, for all \( x > 0 \),

\[
\lim_{t \to \infty} \left( \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} - \frac{x^\rho - 1}{\rho} \right)/B(t) = \frac{x^{\rho + \beta'} - 1}{\rho + \beta'},
\]

where \(|B(t)|\) must then be in \( RV_{\rho'} \). There appears then this extra third order parameter \( \rho' \leq 0 \), which we also assume to be negative. More restrictively, we shall here assume that we are in the class of models with a tail function \( 1 - F(x) = Cx^{-1/\gamma} \left( 1 + D_1 x^{\rho/\gamma} + D_2 x^{(\rho + \rho_1)/\gamma} + o \left( x^{(\rho + \rho_1)/\gamma} \right) \right) \), as \( x \to \infty \), with \( C > 0, D_1, D_2 \neq 0, \rho, \rho_1 < 0 \). Consequently, (1) holds and we may there choose

\[
A(t) = c t^\beta =: \gamma' t^\rho, \quad B(t) = c' t^{\beta'} =: \beta' t^{\rho'}, \quad \beta, \beta' \neq 0, \quad \rho \leq \rho' = \max(\rho, \rho_1) < 0.
\]

For intermediate \( k \), i.e., a sequence of integers \( k = k_n, k \in [1, n] \), such that \( k = k_n \to \infty \), and \( k_n = o(n) \) as \( n \to \infty \), we shall consider, as basic statistics, both the log-excesses over the random high level \( \{\ln X_{n-kn}\} \), i.e., \( V_k := \ln X_{n-i+1:n} - \ln X_{n-kn} \) and the scaled log-spacings, \( U_i := i \{\ln X_{n-i+1:n} - \ln X_{n-i:n}\}, 1 \leq i \leq k < n \), where \( X_{n:i} \) denotes, as usual, the \( i \)-th ascending order statistic (o.s.), \( 1 \leq i \leq n \), 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 \( k \), and under the first order framework, the log-excesses, \( V_{ik}, 1 \leq i \leq k \), 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, \( U_i, 1 \leq i \leq k \), are approximately i.i.d. and exponential with mean value \( \gamma \). Consequently the Hill estimator of \( \gamma \) (Hill, 1975), \( H(k) \equiv \)

\(^{1}\)Research partially supported by FCT/POCTI and POCI/FEDER
$H_n(k) = \frac{1}{k} \sum_{i=1}^{k} V_{ik} = \frac{1}{k} \sum_{i=1}^{k} U_i$, is consistent for the estimation of $\gamma$ under the first order framework and for intermediate $k$. Under the third order framework, the asymptotic distributional representation $H_n(k) \xrightarrow{d} \gamma + \frac{\gamma}{\sqrt{k}} Z_k^{(1)} + (A(n/k)/(1-\rho) + A(n/k) B(n/k))/(1-\rho - \rho')) (1 + o_p(1))$ holds true, where $Z_k^{(1)} = \sqrt{k} \left( \sum_{i=1}^{k} E_i/k - 1 \right)$, with $\{E_i\}$ i.i.d. standard exponential random variables (r.v.’s), is an asymptotically standard normal r.v.

The most simple second-order reduced-bias estimators in the literature are the bias-corrected Hill estimators in Caeiro et al. (2005), with the functional form,

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

dependent upon the Hill estimator $H(k)$ and $(\hat{\beta}, \hat{\rho})$, adequate consistent estimators of the second order parameters $\beta$ and $\rho$, respectively. Under the third order framework in (1),

$$\overline{H}_{\beta, \rho}(k) \xrightarrow{d} \gamma + \frac{\gamma}{\sqrt{k}} Z_k^{(1)} + A(n/k) \left( \frac{B(n/k)}{1-\rho - \rho'} - \frac{A(n/k)}{\gamma(1-\rho)^2} \right) + o_p \left( \frac{1}{\sqrt{k}} \right) \left( 1 + o_p(1) \right).$$

(4)

Consequently, even if $\sqrt{k} A(n/k) \to \infty$, with $\sqrt{k} A^2(n/k) \to \lambda_A$ and $\sqrt{k} A(n/k) B(n(k)) \to \lambda_B$, $\lambda_A$ and $\lambda_B$ finite, $\sqrt{k} \left( \overline{H}_{\beta, \rho}(k) - \gamma \right)$ is asymptotically normal with variance equal to $\gamma^2$. The asymptotic bias of $\overline{H}_{\beta, \rho}(k)$ is equal to $b_{\pi} = b_{\pi}(\gamma, \rho, \rho') \equiv ABIAS_{\pi} := \lambda_B/(1-\rho - \rho') - \lambda_A/(\gamma(1-\rho)^2)$. We may further state the following theorem.

**Theorem 1.1 (Caeiro et al., 2005)** Under the third order framework in (1) and for $(\hat{\beta}, \hat{\rho})$, any consistent estimator of the vector of second order parameters $(\beta, \rho)$, we may write

$$\overline{H}_{\beta, \rho}(k) - \overline{H}_{\beta, \rho}(k) \xrightarrow{L} A(n/k) \left\{ a(\hat{\beta} - \hat{\beta})/\beta + (\hat{\rho} - \rho) \left( a \ln(n/k) + b \right) \right\},$$

(5)

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

In section 2 of this paper, we shall briefly review the estimation of the two second order parameters $\beta$ and $\rho$. Next, in Section 3, we provide some information on the asymptotic behavior of $\sqrt{k} \left\{ \overline{H}_{\beta, \rho}(k) - \gamma \right\}$, whenever $\sqrt{k} A^2(n(k)) \to \lambda_A$ 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 estimators of $\beta$ and $\rho$ used before in papers like Caeiro et al. (2005), computed at their optimal levels $k_1$. These levels should be such that $\sqrt{k_1} A^2(n(k_1)) \to \lambda_A$, and $\sqrt{k_1} A(n(k_1)) B(n(k_1)) \to \lambda_B$, both finite, $\lambda_A$ or $\lambda_B \neq 0$, i.e., for the class of models under consideration, we have $k/k_1 \to q > 0$, whenever $n \to \infty$.

## 2 A brief review of the second order parameters’ estimators

### 2.1 The estimation of $\rho$

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

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

(6)
where \( M_n^{(j)}(k) := \frac{1}{k} \sum_{i=1}^{k} \left\{ \ln \frac{X_{n-i+1}^j}{X_{n-k}^j} \right\}^j, \ j \geq 1 \) \[ M_n^{(1)} = H, \text{ the Hill estimator}. \]

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.

**Proposition 2.1 (Fraga Alves et al., 2003)** Under the second order framework, \( \rho < 0 \), if \( k \) is intermediate, and if \( \sqrt{k} A(n/k) \to \infty \), as \( n \to \infty \), the statistic \( \hat{\rho}_n^{(\tau)}(k) \) in (6) converges in probability towards \( \rho \), as \( n \to \infty \), for any real \( \tau \neq 0 \). If (1) holds, we may further guarantee that there exist constants (\( u_\rho, v_\rho, \sigma_\rho \)) and an asymptotically standard normal r.v. \( W_k \), such that

\[
\hat{\rho}_n^{(\tau)}(k) - \rho \xrightarrow{d} \sigma_\rho \sqrt{W_k/(\sqrt{k} A(n/k) + (u_\rho A(n/k) + v_\rho B(n/k))(1 + o_p(1))}.
\]

Consequently, under the third order framework, if \( k \) is finite, \( \sqrt{k} A^2(n/k) \to \lambda_A \) and \( \sqrt{k} AB(n/k) \to \lambda_B \), both finite, \( \sqrt{k} A(n/k) \left( \hat{\rho}_n^{(\tau)}(k) - \rho \right) \) is asymptotically normal with a mean value \( \lambda_A u_\rho + \lambda_B v_\rho \) and variance

\[
\sigma_\rho^2 \equiv \sigma_\rho^2(\gamma) = \left( \gamma (1 - \rho^3)/\rho \right)^2 \left( 2 \rho^2 - 2\rho + 1 \right).
\]

Moreover, with \( Z_k^{(\alpha)} = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} E_n^0 / \Gamma(\alpha + 1) - \sqrt{k} \), for any \( \alpha \geq 1 \), \( \Gamma(t) \) denoting the complete gamma function, \( W_k = \left( (3 - \rho)Z_k^{(1)} - (3 - 2\rho)Z_k^{(2)} + (1 - \rho)Z_k^{(3)} \right) / \sqrt{2\rho^2 - 2\rho + 1} \).

### 2.2 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 \( U_i, 1 \leq i \leq k \). On the basis of a consistent estimator \( \hat{\rho} \) of the second order parameter \( \rho \), we shall consider the \( \beta \)-estimator, \( \hat{\beta}(k; \hat{\rho}) \), where, for any \( \rho < 0 \),

\[
\hat{\beta}(k; \rho) := \frac{\left( \frac{1}{k} \right)^{-\rho} \left( \frac{1}{k} \sum_{i=1}^{k} \left( \frac{1}{k} \right)^{-\rho} \right) \left( \frac{1}{k} \sum_{i=1}^{k} U_i \right) - \left( \frac{1}{k} \sum_{i=1}^{k} \left( \frac{1}{k} \right)^{-\rho} U_i \right)}{\left( \frac{1}{k} \sum_{i=1}^{k} \left( \frac{1}{k} \right)^{-\rho} \right) \left( \frac{1}{k} \sum_{i=1}^{k} \left( \frac{1}{k} \right)^{-\rho} U_i \right) - \left( \frac{1}{k} \sum_{i=1}^{k} \left( \frac{1}{k} \right)^{-2\rho} U_i \right)}.
\]

Gomes and Martins (2002) kept up to the second order framework. Here, we go into the third order framework in (1), assume that \( \beta \) and \( \rho \) are going to be estimated at the same level \( k \), and state the following result, proved in Gomes et al. (2005):

**Theorem 2.1 (Gomes et al., 2005)** Under the second order framework, with \( \rho < 0 \), if we consider \( \hat{\beta}(k; \hat{\rho}(k; \tau)) \), the rate of convergence of \( \hat{\beta}(k; \hat{\rho}(k; \tau)) \) is of the order of \( \{ \ln(n/k)/(\sqrt{k} A(n/k)) \} \), which must converge towards zero, so that \( \hat{\beta}(k; \hat{\rho}(k; \tau)) \) is consistent for the estimation of \( \beta \), and

\[
\sqrt{k} A(n/k) \left( \hat{\beta}(k; \hat{\rho}(k; \tau)) - \beta \right) / (\beta \ln(n/k)) \xrightarrow{L} -\sqrt{k} A(n/k) (\hat{\rho}(k; \tau) - \rho).
\]

If apart from \( \sqrt{k} A(n/k)/\ln(n/k) \to \infty \), we assume that, under the third order framework in (1), \( \sqrt{k} A^2(n/k) \to \lambda_A \) and \( \sqrt{k} AB(n/k) \to \lambda_B \), both finite, then, with \( \sigma_\rho \) given in (8), the statistic \( \sqrt{k} A(n/k) (\beta - \hat{\beta}(k; \hat{\rho}(k; \tau))) / (\beta \ln(n/k)) \) is asymptotically Normal \( \left( \lambda_A u_\rho + \lambda_B v_\rho, \sigma_\rho^2 \right) \).

### 3 The estimation of \( \gamma, \beta \) and \( \rho \) at levels of the same order

Let us assume now that we want to work with levels \( k \) that are optimal for the estimation of \( \gamma \) through the reduced-bias extreme value index estimators in this paper and for models such that (1) and (2) hold. Then those values of \( k \) should be 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., they should be levels of the same order of the ones leading to the optimal estimation of \( \beta \) and \( \rho \). We may then state the following:
Theorem 3.1 For models in (1), together with (2), let \( k \) and \( k_1 \) be sequences of intermediate integers such that \( k/k_1 \to q > 0 \), finite. Let us also 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. Let us consider \( \hat{H}(k;k_1) = H\beta(k_1,\dot{\rho}(k_1)(k) \), with \( \hat{H} \) the estimator (3). Then, \( \sqrt{k} (\hat{H}(k;k_1) - \gamma) \xrightarrow{d_{n \to \infty}} \text{Normal} (b^*, \sigma^2(q)) \), where

\[
b^* = \lambda_B \left( 1/(1-2\rho) - v/((1-\rho)^2) \right) - \lambda_A \left( 1/(\gamma(1-2\rho)) + u/((1-\rho)^2) \right),
\]

and

\[
\sigma^2(q) = \sigma^2(q,\gamma,\rho) = \gamma^2 \left( 1 + q^{-2}\rho \ln q + 1/(1-\rho)^2 - 2(\rho^2 - 2(1-\rho)/\rho^2) \right),
\]

i.e., we get the same rate of convergence, of the order of \( 1/\sqrt{k} \), for \( \hat{H}(k;k_1) \), but with an asymptotic variance increasing with \( q \).

Proof. From equations (5) and (9), \( \hat{H}(k;k_1) - \tilde{H}(\hat{\rho}) = (\hat{\rho} - \rho) A(n/k)(a \ln(k/k_1) + b) =: W_{k;k_1} \). If \( \sqrt{k} A^2(n/k_1) \to \lambda_{A1} \) and \( \sqrt{k} A(n/k_1)B(n/k_1) \to \lambda_{B1} \), both finite, then \( \hat{\rho} - \rho = O_p \left( 1/ (\sqrt{k} A(n/k_1)) \right) \) and

\[
\sqrt{k} W_{k;k_1} = O_p \left( \frac{\sqrt{k} A(n/k)}{\sqrt{k} A(n/k_1)} \left( a \ln \left( \frac{k}{k_1} \right) + b \right) \right) = O_p \left( \left( \frac{k}{k_1} \right)^{1/2} \rho \ln \left( \frac{k}{k_1} \right) \right).
\]

Let us think that \( k/k_1 \to q > 0 \). Since \( \sqrt{k} A(n/k_1) \sim \sqrt{k} A(n/k) q^{e^{-1/2} \left( 1 + o(1) \right)} \), \( A(n/k_1) \sim q^\rho A(n/k) \) and \( B(n/k_1) \sim q^\rho B(n/k) \),

\[
\sqrt{k} \left( \hat{H}(k;k_1) - \gamma \right) = \gamma Z_{k_1}^{(1)} + \sigma \rho \left( \frac{1}{2} q^{-\rho} \left( a \ln q + b \right) \right) W_{k_1}
\]

\[
+ \sqrt{k} A^2(n/k) \left( u + u_\rho q^\rho \left( a \ln q + b \right) \right) \left( 1 + o_p(1) \right)
\]

\[
+ \sqrt{k} A(n/k)B(n/k) \left( v + v_\rho q^\rho \left( a_v + b_v \right) \right) \left( 1 + o_p(1) \right).
\]

Since \( Cov \left( Z_{k_1}^{(1)}, W_{k_1} \right) = 0 \), the variance of \( \{ \gamma Z_{k_1}^{(1)} + \sigma \rho \left( \frac{1}{2} q^{-\rho} \left( a \ln q + b \right) \right) W_{k_1} \} \) is the value \( \sigma^2(q; \gamma, \rho) \) in (10).

The pattern of \( \sigma^2(q; \gamma, \rho) \), as a function of \( q \), is of the same type for all \( \gamma, \rho \): this variance converges towards \( \gamma^2 \), as \( q \to 0 \), next increases till a value slightly larger than \( \gamma^2 \), then decreases again till \( \gamma^2 \) at \( q = \exp \left( -1/(1-\rho) \right) < 1 \), and finally increases fast, taking the value \( \sigma^2 = \gamma^2 \left( 1 + ((1-\rho)/\rho^2) - 2(1-\rho)/\rho^2 \right) \); for \( q = 1 \). It is thus obvious that we should base the extreme value index estimation on a number \( k \) of top o.s., smaller than or equal to \( k_1 \), the number of o.s. used for the estimation of the second order parameters \( \beta \) and \( \rho \); from \( k_1 \) onwards, it will be convenient to use the same number of top o.s. for the estimation of the three parameters.

REFERENCES