PROCEEDINGS

ASMDA 2021

International Conference

and

Demographics 2021 Workshop

http://www.asmda.es/asmda2021.html

Virtual 1 – 4 June 2021

Christos H Skiadas

Editor

A Generalized Mean under a Non-regular Framework and Extreme Value Index Estimation

M. Ivette Gomes[1], Lígia Henriques-Rodrigues[2], and Dinis Pestana[3]

[1] Centro de Estatística e Aplicações (CEAUL), Universidade de Lisboa, Lisboa, Portugal (E-mail: ivette.gomes@fc.ul.pt)
[2] Centro de Investigação em Matemática e Aplicações (CIMA), Universidade de Évora, Évora, Portugal (E-mail: ligiahr@uevora.pt)
[3] Centro de Estatística e Aplicações (CEAUL), Universidade de Lisboa, Lisboa, Portugal (E-mail: dinis.pestana@fc.ul.pt)

Abstract. The Hill estimator, one of the most popular *extreme value index* (EVI) estimators under a heavy right-tail framework, i.e. for a positive EVI, here denoted by ξ, is an average of the log-excesses. Consequently, it can be regarded as the logarithm of the geometric mean or mean of order $p = 0$ of an adequate set of systematic statistics. We can thus more generally consider any real p, the mean of order p (MO_p) of those same statistics and the associated MO_p EVI-estimators, also called harmonic moment EVI-estimators. The normal asymptotic behaviour of these estimators has been obtained for $p < 1/(2\xi)$, with consistency achieved for $p < 1/\xi$. The non-regular framework, i.e. the case $p \geq 1/(2\xi)$, will be now considered. Consistency is no longer achieved for $p > 1/\xi$, but an almost degenerate behavior appears for $p = 1/\xi$. Results are illustrated on the basis of large-scale simulation studies. An algorithm providing an almost degenerate MO_p EVI-estimation is suggested.
Keywords: Generalized Means, Non-regular Frameworks, Statistics of Extremes.

1 Introduction

Given $X_1, \ldots, X_n$, a sample of size n of *independent, identically distributed* (IID), or possibly stationary weakly dependent *random variables* (RVs), with a *cumulative distribution function* (CDF) F, let us consider the notation $X_{1:n} \leq \cdots \leq X_{n:n}$ for the associated ascending *order statistics* (OSs). Let us further assume that there exist real constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that the linearly normalized maximum, $(X_{n:n} - b_n)/a_n$, converges in distribution to a non-degenerate RV. Then, with $\lambda \in \mathbb{R}$, a location parameter, and $\delta > 0$, a scale parameter, the limit distribution is necessarily the general *extreme value* (EV) CDF (von Misès[21], Jenkinson[20]), given by

$$\mathrm{EV}_\xi\big(\tfrac{x-\lambda}{\delta}\big) = \begin{cases} \exp\Big(-\Big(1+\frac{\xi(x-\lambda)}{\delta}\Big)^{-1/\xi}\Big), \ 1+\frac{\xi(x-\lambda)}{\delta} > 0, & \text{if } \xi \neq 0, \\ \exp\big(-\exp\big(-\frac{x-\lambda}{\delta}\big)\big), \ x \in \mathbb{R}, & \text{if } \xi = 0. \end{cases} \quad (1)$$

The CDF F is said to belong to the *max-domain of attraction* (MDA) of EV_ξ, in (1), the unique max-stable laws, under the aforementioned framework, and, as usual, the notation $F \in \mathcal{D}_\mathcal{M}(\mathrm{EV}_\xi)$ is used. The parameter ξ is the *extreme value index* (EVI), the primary parameter of extreme events.

The EVI measures the heaviness of the *right-tail* function $\overline{F}(x) := 1 - F(x)$, and the heavier the right-tail, the larger ξ is. Let us further use the notation $\mathcal{R}_a$ for the class of regularly varying functions at infinity, with an index of regular variation equal to $a \in \mathbb{R}$ (see Seneta[25] and Bingham *et al.*[2], among others, for details on regular variation theory). In this paper we work with Pareto-type underlying models, with a positive EVI, or equivalently, CDFs such that $\overline{F}(x) = x^{-1/\xi} L(x)$, $\xi > 0$, with $L \in \mathcal{R}_0$, a slowly varying function at infinity, i.e. a regularly varying function with an index of regular variation equal to zero. These heavy-tailed models are quite common in many areas of application, like bibliometrics, biostatistics, computer science, finance, insurance, statistical quality control and telecommunications, among others.

For Pareto-type models, the classical EVI-estimators are the Hill (H) EVI–estimators (Hill[19]), which are the averages of the log-excesses,

$$V_{ik} := \ln X_{n-i+1:n} - \ln X_{n-k:n} = \ln \frac{X_{n-i+1:n}}{X_{n-k:n}} =: \ln U_{ik}, \quad 1 \le i \le k < n. \quad (2)$$

We can thus write

$$\mathrm{H}(k) := \frac{1}{k} \sum_{i=1}^{k} V_{ik} = \sum_{i=1}^{k} \ln U_{ik}^{1/k} = \ln \left(\prod_{i=1}^{k} U_{ik} \right)^{1/k}, \quad 1 \le k < n, \quad (3)$$

i.e. the Hill estimator is the logarithm of the *geometric mean* (or *mean-of-order-*0) of the statistics U_{ik}, defined in (2). More generally, first Brilhante *et al.*[3], for $p \in \mathbb{R}_0^+$, and more recently Caeiro *et al.*[5], for $p \in \mathbb{R}$, considered as basic statistics the *mean-of-order-p* (MO_p) of U_{ik}, i.e. the class of statistics

$$T_p(k) = \begin{cases} \left(\frac{1}{k} \sum\limits_{i=1}^{k} U_{ik}^p \right)^{1/p}, & \text{if } p \neq 0, \\ \left(\prod\limits_{i=1}^{k} U_{ik} \right)^{1/k}, & \text{if } p = 0, \end{cases}$$

and the class of functionals,

$$\mathrm{H}_p(k) \equiv \mathrm{MO}_p(k) := \begin{cases} \left(1 - T_p^{-p}(k)\right)/p, & \text{if } p \neq 0, \\ \ln T_0(k) = \mathrm{H}(k), & \text{if } p = 0, \end{cases} \quad (4)$$

with $H_0(k) \equiv H(k)$, given in (3). This class of MO_p functionals depends now on this *tuning* parameter $p \in \mathbb{R}$, which makes it highly flexible, and has been studied asymptotically and for finite samples in the aforementioned articles, and also, independently, in Paulauskas and Vaičiulis[22] and Beran *et al.*[1]. Consistency was shown for $p < 1/\xi$, with an asymptotic normal behaviour holding for $p < 1/(2\xi)$. See also Paulauskas and Vaičiulis[23], among others.

In Section 2 of this paper, apart from the introduction of a few technical details in the field of *extreme value theory* (EVT), we provide a few details on sum-stable laws and the asymptotic behaviour of the MO_p EVI-estimators under regular and non-regular frameworks is put forward. In Section 3 we provide a short illustration of the finite-sample behaviour of the MO_p class of functionals under regular and non-regular frameworks. In Section 4, a method for the adaptive choice of the tuning parameters k and p, essentially on the basis of sample path stability, is provided. The behaviour of the new adaptive EVI-estimators is illustrated through an application to simulated random samples, and in Section 5 some concluding remarks are presented.

2 Preliminary results in the area of EVT for heavy tails and asymptotic behaviour of MO_p functionals

In the area of statistics of univariate extremes and whenever working with large values, i.e. with the right-tail of the model F underlying the data, a model F is often said to be *heavy-tailed* whenever the right-tail function $\overline{F} \in \mathcal{R}_{-1/\xi}$, $\xi > 0$. Then, $F \in \mathcal{D}_{\mathcal{M}}\left(\mathrm{EV}_\xi\right)_{\xi>0} =: \mathcal{D}_{\mathcal{M}}^{+}$ and reciprocally, if $F \in \mathcal{D}_{\mathcal{M}}^{+}$ we necessarily have $\overline{F} \in \mathcal{R}_{-1/\xi}$ (Gnedenko[8]).

2.1 A brief review of first and second-order conditions

If $F \in \mathcal{D}_{\mathcal{M}}^{+}$, and with the notation $F^{\leftarrow}(t) := \inf\{x : F(x) \geq t\}$ for the generalised inverse function of F, the reciprocal tail quantile function $U(t) := F^{\leftarrow}(1 - 1/t)$ is of regular variation with index ξ (de Haan[15]), i.e. we usually work with any of the first-order conditions:

$$F \in \mathcal{D}_{\mathcal{M}}^{+} \quad \Longleftrightarrow \quad \overline{F} \in \mathcal{R}_{-1/\xi} \quad \Longleftrightarrow \quad U \in \mathcal{R}_{\xi}. \tag{5}$$

The *second-order parameter* ρ (≤ 0) rules the rate of convergence in the first-order condition, in (5), and can be defined as the non-positive parameter appearing in the limiting relation

$$\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \xi \ln x}{A(t)} = \psi_\rho(x) := \begin{cases} \frac{x^\rho - 1}{\rho}, & \text{if } \rho < 0, \\ \ln x, & \text{if } \rho = 0, \end{cases} \tag{6}$$

which is assumed to hold for every $x > 0$, and where $|A| \in \mathcal{R}_\rho$ (Geluk and de Haan[7]). This condition has been widely accepted as an appropriate condition to specify the right-tail of a Pareto-type distribution in a semi-parametric way and enables easily the derivation of the non-degenerate asymptotic bias of EVI-estimators, under a semi-parametric framework.

2.2 Asymptotic behaviour of the Hill EVI-estimators

To have consistency of the Hill estimators, in all $\mathcal{D}_{\mathcal{M}}^{+}$, we need to work with intermediate values of k, i.e. a sequence of integers $k = k_n$, $1 \leq k < n$, such that

$$k = k_n \to \infty \quad \text{and} \quad k_n = o(n), \text{ as } n \to \infty. \tag{7}$$

Under the aforementioned second-order framework, in (6), the asymptotic distributional representation

$$\mathrm{H}(k) - \xi \overset{d}{=} \frac{\xi}{\sqrt{k}} Z_k + \frac{A(n/k)}{1-\rho}\,(1 + o_p(1)) \tag{8}$$

holds (de Haan and Peng[17]), where, with $\{E_i\}$ a sequence of IID standard exponential RVs, $Z_k = \sqrt{k}\Big(\sum_{i=1}^{k} E_i/k - 1\Big)$ is asymptotically standard normal.

Remark 1. For the Hill estimator, and whenever needed, we often consider the most common estimate of $k_{0|0} \equiv k_{0|H}(n) := \arg\min_k \mathrm{MSE}(\mathrm{H}_0(k))$ (Hall[18]), given by

$$\hat{k}_{0|0} = \min\Big(n-1, \Big[\big((1-\hat\rho)^2 n^{-2\hat\rho}/\big(-2\ \hat\rho\ \hat\beta^2\big)\big)^{1/(1-2\hat\rho)}\Big] + 1\Big), \tag{9}$$

with $(\hat\beta, \hat\rho)$ adequate estimates of the vector (β, ρ) of second-order parameters. References to the estimation of second-order parameters can be found in [5], among others. But the estimate $\widehat{\xi}_0 := \mathrm{H}(\hat{k}_{0|0})$ has usually a high positive bias, and alternatives are often needed.

2.3 Asymptotic behaviour of MO$_p$ EVI-estimators under a regular framework

More generally that the distributional representation in (8), we refer the main theorem in Brilhante *et al.*[3] (see also Gomes and Caeiro[11] and Caeiro *et al.*[5]):

Theorem 1 ([3], [11], [5]). *Under the validity of the first-order condition, in* (5)*, and for intermediate sequences $k = k_n$, i.e. if* (7) *holds, the class of estimators $\mathrm{H}_p(k)$, in* (4)*, is consistent for the estimation of ξ, provided that $p < 1/\xi$.*

If we moreover assume the validity of the second-order condition in (6)*, the asymptotic distributional representation*

$$\mathrm{H}_p(k) - \xi \ \stackrel{d}{=}\ \frac{\sigma_p(\xi)\ Z_k^{(p)}}{\sqrt{k}} + b_p(\xi|\rho)\ A(n/k) + o_p(A(n/k))$$

holds for all $p < 1/(2\xi)$ and $\rho \le 0$, with $Z_k^{(p)}$ asymptotically standard normal.

Remark 2. For details on the explicit expression of $\sigma_p(\xi)$ and $b_p(\xi, \rho)$, see any of the aforementioned articles.

2.4 A brief reference to additive stable laws

Given the sequence $\{X_n\}_{\ge 1}$ of IID RVs, let us assume that for each $n \in \mathbb{N}$ there exist normalizing constants $A_n > 0$, $B_n \in \mathbb{R}$, such that the sum linearly normalized converges in distribution to a non-degenerate RV, as $n \to \infty$. Then, such an RV is necessarily an additive or sum stable law, denoted by $S_{\alpha,\beta} = S_{\alpha,\beta}(\lambda, \delta)$, i.e., as $n \to \infty$,

$$\frac{\sum_{i=1}^{n} X_i - B_n}{A_n} \ \stackrel{d}{\longrightarrow}\ S_{\alpha,\beta}(\lambda, \delta). \tag{10}$$

Among other books in the topic, see Gnedenko and Kolmogorov[10]. The scale parameter $\alpha \in (0,2]$ is the so-called *characteristic exponent* (CE), also related to the tail weight of F and strongly linked to the EVI. The practical use of additive stable laws has been seriously hampered by the fact that the unique so far explicitly known additive stable probability density functions are the ones corresponding to $\alpha = 2$, the normal case, $(\alpha, \beta) = (1, 0)$, the Cauchy case, and $(\alpha, \beta) = (1/2, 1)$, the Lévy case.

Regarding the common CDF F, the *generalized central limit theorem* (GCLT) (Samorodnitsky and Taqqu[24]) states that, as $n \to \infty$, the result in (10) holds if and only if

$$1 - F(x) + F(-x) \in \mathcal{R}_{-\alpha} \quad \text{and} \quad \frac{F(-x)}{1 - F(x) + F(-x)} \underset{x\to\infty}{\longrightarrow} \frac{1-\beta}{2}. \qquad (11)$$

The particular Pareto case. Let us now consider a unit Pareto RV, Y, with CDF $F_1(y) = 1 - 1/y,\ y \geq 1$, and for $\xi > 0$, Y^ξ, an RV with CDF $F_\xi(y) = 1 - y^{-1/\xi},\ y \geq 1$, with a right-tail function

$$\overline{F}(y) = 1 - F(y) = y^{-1/\xi}, \quad y \geq 1.$$

On the basis of (11), since $\overline{F}_\xi \in \mathcal{R}_{-1/\xi}$ and $\beta = 1$, Y^ξ belongs to the additive domain of attraction of a stable law with $\beta = 1$ and a CE given by $\alpha(\xi) = \min\{2, 1/\xi\}$. Let us use the notation $S_\alpha = S_{\alpha,1}$. Then, and asymptotically,

$$\frac{1}{k}\sum_{i=1}^{k} Y_i^\xi = \begin{cases} \frac{1}{1-\xi} + \frac{\xi}{1-\xi}\sqrt{\frac{1}{k(1-2\xi)}}\ S_{2,0}(1+o_{\mathbb{P}}(1)), & \text{if } \xi < \frac{1}{2}, \\ \frac{1}{1-\xi} + \sqrt{\frac{\ln k}{k}}\ S_{2,0}(1+o_{\mathbb{P}}(1)), & \text{if } \xi = \frac{1}{2}, \\ \frac{1}{1-\xi} + k^{\xi-1}\left\{\frac{\xi\Gamma(2-1/\xi)|\cos(\pi/(2\xi))|}{1-\xi}\right\}^{\xi} S_{1/\xi}(1+o_{\mathbb{P}}(1)), & \text{if } \frac{1}{2} < \xi < 1, \\ \ln k + 1 - \gamma - \ln(2/\pi) + \frac{\pi}{2} S_1(1+o_{\mathbb{P}}(1)), & \text{if } \xi = 1, \\ k^{\xi-1}\{\Gamma(1-1/\xi)\cos(\pi/(2\xi))\}^{\xi} S_{1/\xi}(1+o_{\mathbb{P}}(1)), & \text{if } \xi > 1. \end{cases} \qquad (12)$$

2.5 Asymptotic behaviour of EVI-estimators under a non-regular framework

We next state the main theorem in Gomes *et al.*[13], a generalization of **Theorem 1** to non-regular cases.

Theorem 2 ([13]). *Under the validity of the first-order condition, in* (5)*, and for intermediate sequences* $k = k_n$*, i.e. if* (7) *holds, the class of functionals* $\mathrm{H}_p(k)$*, in* (4)*, is consistent for the estimation of* ξ*, provided that* $p \leq 1/\xi$ *and converges to* $1/p$ $(< \xi)$ *if* $p > 1/\xi$*.*

In the region of values of p *out of the scope of* **Theorem 1***, i.e.* $1/(2\xi) \leq p \leq 1/\xi$*, if we assume the validity of the second-order condition in* (6)*, the following asymptotic distributional representations hold.*

(ii) *If* $p = 1/(2\xi)$, *with* $S_{2,0} \equiv \mathcal{N}(0,1)$, *an asymptotically centered normal,*

$$\mathrm{H}_p(k) - \xi \ \stackrel{d}{=} \ \frac{S_{2,0}}{4p\sqrt{k/\ln k}} + \frac{A(n/k)}{1-2\rho} + o_p(A(n/k)).$$

(i) *If* $1/(2\xi) < p < 1/\xi$, *with* $S_\alpha \equiv S_{\alpha,1}$ *asymptotically stable with a* CE $\alpha = 1/(p\xi)$, *we get the validity of the asymptotic distributional representation,*

$$\mathrm{H}_p(k) - \xi \ \stackrel{d}{=} \ \frac{\sigma_p^*(\xi)\ S_\alpha}{k^{1-p\xi}} + b_p(\xi|\rho)\ A(n/k) + o_p(A(n/k)).$$

(iii) *If* $p = 1/\xi$, *with* S_1 *asymptotically sum-stable with a* CE, $\alpha = 1$, *and with* $\omega = 1 - \gamma - \ln(2/\pi)$, γ *denoting Euler's constant,*

$$\mathrm{H}_p(k) - \xi \ \stackrel{d}{=} \ -\frac{\xi}{\ln k} + \xi\left(\frac{w + \pi/2\ S_1}{\ln^2 k} - \frac{pA(n/k)}{\rho \ln k}\right)(1 + o_{\mathbb{P}}(1)).$$

We now further state the following:

Theorem 3. *Under the first order condition in* (5), *and for any* $k = k_n \to \infty$, *not necessarily intermediate,*

$$\mathrm{H}_p(k) - \xi \ \stackrel{d}{=} \ -\frac{\xi}{\ln k}(1 + o_{\mathbb{P}}(1)), \quad \textit{if} \quad p = 1/\xi, \tag{13}$$

and

$$\mathrm{H}_p(k) - 1/p \ \stackrel{d}{=} \ O_{\mathbb{P}}\left(1/k^{p\xi-1}\right), \quad \textit{if} \quad p > 1/\xi. \tag{14}$$

Proof. For any $1 \le k < n$, the distributional identity,

$$\mathrm{H}_p(k) \stackrel{d}{=} \ \frac{1}{p}\left(1 - \left(\frac{1}{k}\sum_{i=1}^{k} Y_i^{\xi p}(1 + o_{\mathbb{P}}(1))\right)^{-1}\right)$$

holds. Just as mentioned in [5], the term $o_{\mathbb{P}}(1)$ above is uniform in i, $1 \le i \le k$. This comes from the results in [6] (see Theorem B.2.18 in [16]), jointly with the fact that for uniform order statistics $U_{i:n}, 1 \le i \le n$, we have that $1/U_{i:n}$ can be uniformly bounded in probability by $C[i/(n+1)]^{-1}$ (for some constant C). From (12), (13) and (14) follow.

3 Finite-sample behaviour of MO$_p$ functionals

The Monte-Carlo simulations, already performed in Gomes *et al.*[13,14], were extended to other models, like the Burr$_{\xi,\rho}$ models, with CDF $F(x) = 1 - \left(1 + x^{-\rho/\xi}\right)^{1/\rho}$, $x \ge 0$, with $\xi > 0$ and $\rho < 0$, the ones presented in this paper. Multi-sample Monte-Carlo simulations of size 5000×20 (20 replicates of 5000 runs each) were performed. Details on multi-sample simulation can be seen in Gomes and Oliveira[12].

Apart from the simulated mean values, at optimal levels, in the sense of minimal *mean square error* (MSE), presented in Table 1, we also present in Table 2 an indicator of the *relative efficiency* (REFF), given by

$$\mathrm{REFF}_{p|0} = \frac{\mathrm{RMSE}(\mathrm{H}_{00})}{\mathrm{RMSE}(\mathrm{H}_{p0})} := \sqrt{\frac{\mathrm{MSE}(\mathrm{H}_{00})}{\mathrm{MSE}(\mathrm{H}_{p0})}} =: \frac{\mathrm{RMSE}_{00}}{\mathrm{RMSE}_{p0}},$$

where H_{p0} is the $\mathrm{H}_p(k)$ EVI-estimator computed at the simulated value of $k_{0|p} := \arg\min_k \mathrm{MSE}(\mathrm{H}_p(k))$. For each model, the smallest bias, in Table 1, and the highest REFF, in Table 2, are written in **bold**.

n	100	200	2000	5000
$\mathrm{BURR}_{\xi,\rho}$ parent, $\xi = 1$, $\rho = -0.1$				
H_0	3.256 ± 0.0088	2.814 ± 0.0069	2.100 ± 0.0137	1.947 ± 0.0101
$\mathrm{H}_{0.1}$	2.588 ± 0.0057	2.328 ± 0.0049	1.812 ± 0.0095	1.717 ± 0.0141
$\mathrm{H}_{0.4}$	1.949 ± 0.0030	1.824 ± 0.0028	1.523 ± 0.0031	1.435 ± 0.0028
$\mathrm{H}_{0.5}$	1.698 ± 0.0021	1.612 ± 0.0022	1.389 ± 0.0024	1.320 ± 0.0023
$\mathrm{H}_{0.9}$	1.084 ± 0.0019	1.080 ± 0.0014	1.063 ± 0.0011	1.057 ± 0.0009
H_1	$\mathbf{1.000} \pm 0.0000$	$\mathbf{1.000} \pm 0.0000$	$\mathbf{1.000} \pm 0.0000$	$\mathbf{1.000} \pm 0.0000$
$\mathrm{H}_{1.1}$	0.909 ± 0.0000	0.909 ± 0.0000	0.909 ± 0.0000	0.909 ± 0.0000
$\mathrm{H}_{1.2}$	0.833 ± 0.0000	0.833 ± 0.0000	0.833 ± 0.0000	1.000 ± 0.0000
$\mathrm{BURR}_{\xi,\rho}$ parent, $\xi = 1$, $\rho = -0.5$				
H_0	1.295 ± 0.0086	1.242 ± 0.0048	1.131 ± 0.0027	1.102 ± 0.0022
$\mathrm{H}_{0.1}$	1.233 ± 0.0065	1.198 ± 0.0046	1.115 ± 0.0026	1.092 ± 0.0020
$\mathrm{H}_{0.4}$	1.156 ± 0.0047	1.142 ± 0.0035	1.098 ± 0.0018	1.083 ± 0.0018
$\mathrm{H}_{0.5}$	1.119 ± 0.0041	1.113 ± 0.0039	1.083 ± 0.0017	1.072 ± 0.0010
$\mathrm{H}_{0.9}$	1.028 ± 0.0008	1.023 ± 0.0006	1.014 ± 0.0002	1.012 ± 0.002
H_1	$\mathbf{1.000} \pm 0.0000$	$\mathbf{1.000} \pm 0.0000$	$\mathbf{1.000} \pm 0.0000$	$\mathbf{1.000} \pm 0.0000$
$\mathrm{H}_{1.1}$	0.909 ± 0.0000	0.909 ± 0.0000	0.909 ± 0.0000	0.909 ± 0.0000
$\mathrm{H}_{1.2}$	0.833 ± 0.0000	0.833 ± 0.0000	0.833 ± 0.0000	1.000 ± 0.0000
$\mathrm{BURR}_{\xi,\rho}$ parent, $\xi = 1$, $\rho = -1$				
H_0	1.138 ± 0.0042	1.110 ± 0.0033	1.050 ± 0.0010	1.037 ± 0.0009
$\mathrm{H}_{0.1}$	1.111 ± 0.0027	1.093 ± 0.0030	1.045 ± 0.0010	1.033 ± 0.0006
$\mathrm{H}_{0.4}$	1.092 ± 0.0026	1.080 ± 0.0020	1.046 ± 0.0007	1.035 ± 0.0007
$\mathrm{H}_{0.5}$	1.075 ± 0.0024	1.065 ± 0.0018	1.042 ± 0.0008	1.034 ± 0.0006
$\mathrm{H}_{0.9}$	1.018 ± 0.0005	1.015 ± 0.0001	1.010 ± 0.0001	1.008 ± 0.0001
H_1	$\mathbf{0.998} \pm 0.0001$	$\mathbf{0.999} \pm 0.0000$	$\mathbf{1.000} \pm 0.0000$	$\mathbf{1.000} \pm 0.0000$
$\mathrm{H}_{1.1}$	0.908 ± 0.0001	0.909 ± 0.0000	0.909 ± 0.0000	0.909 ± 0.0000
$\mathrm{H}_{1.2}$	0.833 ± 0.0000	0.833 ± 0.0000	0.833 ± 0.0000	1.000 ± 0.0000

Table 1. Simulated mean values, at optimal levels, of $\mathrm{H}_p(k)/\xi$, $p\xi = 0, 0.1, 0.4, 0.5, 0.9, 1.0, 1.1, 1.2$, for $\mathrm{BURR}_{\xi,\rho}$ underlying parents, $\xi = 1$, $\rho = -0.1, -0.5, -1$, together with 95% confidence intervals

n	100	200	2000	5000
BURR$_{\xi,\rho}$ parent, $\xi = 1,\ \rho = -0.1$				
RMSE$_{00}$	2.592 ± 0.2304	2.133 ± 0.2287	1.326 ± 0.2224	1.134 ± 0.2197
H$_{0.1}$	1.188 ± 0.0013	1.163 ± 0.0009	1.113 ± 0.0030	1.092 ± 0.0022
H$_{0.2}$	1.475 ± 0.0027	1.408 ± 0.0016	1.291 ± 0.0042	1.249 ± 0.0037
H$_{0.4}$	2.535 ± 0.0065	2.323 ± 0.0039	1.960 ± 0.0080	1.855 ± 0.0057
H$_{0.5}$	3.467 ± 0.0097	3.133 ± 0.0062	2.555 ± 0.0112	2.390 ± 0.0077
H$_{0.9}$	26.030 ± 0.1132	22.231 ± 0.0936	15.580 ± 0.0810	14.011 ± 0.0708
H$_1$	∞	∞	∞	∞
H$_{1.1}$	28.514 ± 0.1025	23.464 ± 0.0777	14.588 ± 0.0757	12.469 ± 0.0453
H$_{1.2}$	15.553 ± 0.0559	12.798 ± 0.0424	7.957 ± 0.0413	6.801 ± 0.0247
BURR$_{\xi,\rho}$ parent, $\xi = 1,\ \rho = -0.5$				
RMSE$_{00}$	0.478 ± 0.1981	0.381 ± 0.1917	0.193 ± 0.1679	0.150 ± 0.1579
H$_{0.1}$	1.074 ± 0.0011	1.059 ± 0.0008	1.034 ± 0.0008	1.029 ± 0.0006
H$_{0.4}$	1.484 ± 0.0047	1.375 ± 0.0047	1.159 ± 0.0043	1.110 ± 0.0055
H$_{0.5}$	1.761 ± 0.0058	1.603 ± 0.0063	1.260 ± 0.0051	1.175 ± 0.0073
H$_{0.9}$	8.404 ± 0.0343	7.448 ± 0.0411	4.998 ± 0.0141	4.277 ± 0.0230
H$_1$	∞	∞	∞	∞
H$_{1.1}$	5.253 ± 0.0225	4.196 ± 0.0231	2.128 ± 0.0090	1.654 ± 0.0083
H$_{1.2}$	2.865 ± 0.0123	$2.289 \pm 0,0126$	1.160 ± 0.0049	0.9020 ± 0.0045
BURR$_{\xi,\rho}$ parent, $\xi = 1,\ \rho = -1$				
RMSE$_{00}$	0.266 ± 0.1740	0.205 ± 0.1656	0.090 ± 0.1356	0.066 ± 0.1231
H$_{0.1}$	1.047 ± 0.0013	1.038 ± 0.0011	1.023 ± 0.0009	1.020 ± 0.0009
H$_{0.4}$	1.257 ± 0.0048	1.178 ± 0.0056	1.025 ± 0.0071	0.985 ± 0.0074
H$_{0.5}$	1.415 ± 0.0063	1.288 ± 0.0068	1.018 ± 0.0081	0.943 ± 0.0090
H$_{0.9}$	5.743 ± 0.0331	4.924 ± 0.0231	2.849 ± 0.0170	2.268 ± 0.0134
H$_1$	$\mathbf{88.50} \pm 0.782$	$\mathbf{151.87} \pm 10.13$	∞	∞
H$_{1.1}$	2.894 ± 0.0140	2.2433 ± 0.0112	0.994 ± 0.0039	0.724 ± 0.0029
H$_{1.2}$	1.590 ± 0.0077	1.228 ± 0.0062	0.542 ± 0.0021	0.395 ± 0.0016

Table 2. Simulated RMSE of H$_0$ $/\xi$ (first row of any model) and REFF-indicators of H$_p$, $p\xi = 0.1, 0.4, 0.5, 0.9, 1.0, 1.1, 1.2$, for BURR$_{\xi,\rho}$ underlying parents, $\xi = 1$, $\rho = -0.1, -0.5, -1$, together with 95% confidence intervals

A visualization of both tables is provided in Figure 1.

On the basis of the first replica of size 5000, we finally present, in Figure 2, the simulated mean values and *root mean squared errors* (RMSE's), as a function of k, and also for a BURR$_{\xi,\rho}$ model with $\xi = -\rho = 1$.

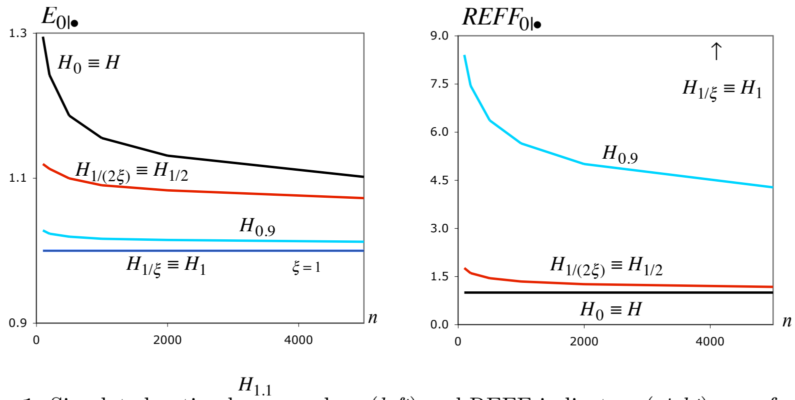

Fig. 1. Simulated optimal mean values (*left*) and REFF-indicators (*right*), as a function of n for a $\text{BURR}_{\xi,\rho}$ parent with $\xi = 1$ and $\rho = -0.5$.

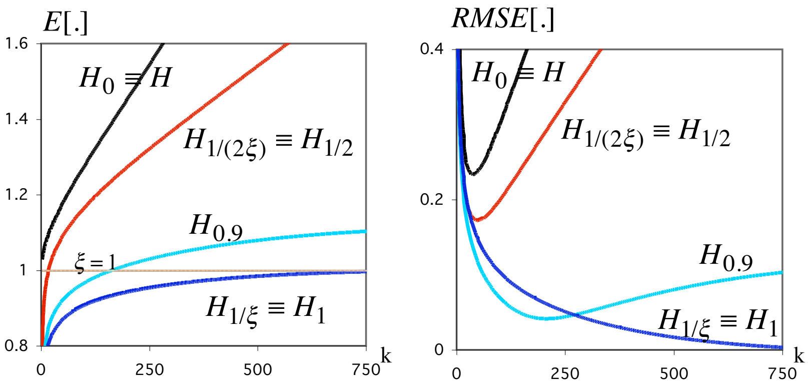

Fig. 2. Mean values (*left*) and RMSE's (*right*) as a function of k, for a $\text{BURR}_{\xi,\rho}$ model with $\xi = -\rho = 1$.

4 A non-regular adaptive choice of p and k

It seems sensible to take $k = n - 1$ and $p = 1/\xi$, but ξ is not known. And if $p\xi > 1$, $H_p(k)$ converges to $1/p < \xi$. Let us thus consider the following procedure:

Algorithm

- Compute an initial EVI-estimate, $\widehat{\xi}_0 = \mathrm{H}(\hat{k}_{0|0})$, with $\hat{k}_{0|0}$ given in (9);
- For p=0.01(0.01)$\cdots$, compute $p_{min} := \arg\min_p (H_p(n-1) - \widehat{\xi}_0)^2$;
- Consider the estimate $\hat{\xi} = 1/p_{min}$.

To ilustrate the performance of the Algorithm, sample sizes from n = 100 until n = 5000 were simulated from a set of underlying models that include the $\mathrm{Burr}_{\xi,\rho}$ models with $\xi = \{0.5, 1\}$ and $\rho = \{-0.5, -1\}$, the Student-t_ν, with $\nu = 2$ degrees-of-freedom ($\xi = 1/\nu = 0.5$), the Cauchy model, $\xi = 1$, the Generalised Pareto model, GP_ξ, with $\xi = \{0.5, 1\}$ and the Fréchet model with $\xi = 0.5$. The results are summarized in Table 3.

n	100	200	1000	2000	5000
$\mathrm{Burr}_{\xi,\rho}$ parent, $\xi = 0.5$, $\rho = -0.5$					
$\lvert\hat{\xi} - \xi\rvert$	0.0098	0.0074	0.0291	0.0051	0.0098
$\mathrm{Burr}_{\xi,\rho}$ parent, $\xi = 0.5$, $\rho = -1$					
$\lvert\hat{\xi} - \xi\rvert$	0.0435	0.0348	0.0076	0.0098	0.0051
$\mathrm{Burr}_{\xi,\rho}$ parent, $\xi = 1$, $\rho = -0.5$					
$\lvert\hat{\xi} - \xi\rvert$	0.1364	0.1111	0.0204	0.0989	0.0417
$\mathrm{Burr}_{\xi,\rho}$ parent, $\xi = 1$, $\rho = -1$					
$\lvert\hat{\xi} - \xi\rvert$	0.0638	0.0204	0.01010	0.0291	0.0291
Student t_2 parent, $\xi = 0.5$					
$\lvert\hat{\xi} - \xi\rvert$	0.0076	0.0102	0.0025	0.0098	0.0051
Cauchy parent, $\xi = 1$, $\rho = -2$					
$\lvert\hat{\xi} - \xi\rvert$	0.0417	0.0638	0.0196	0.0101	0.0099
GP_ξ parent, $\xi = 0.5$					
$\lvert\hat{\xi} - \xi\rvert$	0.0263	0.0128	0.0128	0.0025	0.0025
GP_ξ parent, $\xi = 1$					
$\lvert\hat{\xi} - \xi\rvert$	0.0870	0.0870	0.0309	0.0000	0.0204
$\mathrm{Fréchet}_\xi$ parent, $\xi = 0.5$					
$\lvert\hat{\xi} - \xi\rvert$	0.0025	0.0098	0.0122	0.0283	0.0902

Table 3. Absolute bias of the adaptive MO_p EVI-estimates provided by the Algorithm for several selected models

The results obtained for the simulated samples show that the Algorithm performs well for most of the models. However, the performance of the Algorithm is strongly dependent of the initial EVI-estimate and if the bias of the initial EVI estimate is small then the bias of the MO_p EVI-estimate will also be small. Other initial EVI-estimates, like the reduced bias EVI-estimates in Caeiro *et al.*[4] could be considered. These results claim for a simulation study of the Algorithm comparatively to other adequate adaptive algorithms, a topic out of the scope of this paper.

5 A few concluding remarks

It has been clear for a long time that the H EVI-estimators lead often to a high over-estimation of the EVI, even at optimal levels, in the sense of minimal MSE. The use of the extra uning parameter $p \in \mathbb{R}$ and the MO_p methodology can thus provide a much more adequate EVI-estimation, with asymptotic normality achieved for $p \leq 1/(2\xi)$. But we can now go up to $p = 1/\xi$, getting then a sum-stable behavior, with an index of sum-stability $\alpha = 1/(p\xi)$. And for $p = 1/\xi$, we get, for $\mathrm{H}_p(k) - \xi$, a deterministic dominant component, of the order of $1/\ln k$. The challenge is not to go beyond $p = 1/\xi$. But the algorithm discussed in Section 4 is able to perform such a goal in a great variety of situations. Further notice that for the adaptive choice of (p, k) a double-bootstrap algorithm, of the type of the one in [3], can be used, with some minor modifications. However, such an algorithm relies too much on the finite variance of the normal asymptotic RV associated with the asymptotic behaviour of $\mathrm{H}_p(k)$, needs to be slightly modified, being still under study. And the simple heuristic algorithm presented above seems to work adequately in a great variety of situations.

References

1. J. Beran, D. Schell and M. Stehlik. The harmonic moment tail index estimator: asymptotic distribution and robustness. *Ann. Inst. Statist. Math.*, **66**, 193–220, 2014.
2. N. Bingham, C.M. Goldie and J.L. Teugels. *Regular Variation*. Cambridge Univ. Press, Cambridge, 1987.
3. M.F. Brilhante, M.I. Gomes and D.D. Pestana. A Simple Generalization of the Hill Estimator. *Computational Statistics and Data Analysis* **57**, 1, 518–535, 2013.
4. F. Caeiro, M.I. Gomes, and D.D. Pestana. Direct reduction of bias of the classical Hill estimator. *Revstat—Statist. J.* **3**:2, 113–136, 2005.
5. F. Caeiro, M.I. Gomes, J. Beirlant and de T. Wet. Mean-of-order-p reduced-bias extreme value index estimation under a third-order framework. *Extremes* **19**, 4, 561–589, 2016.
6. H. Drees. A general class of estimators of the extreme value index. *J. Statist. Planning and Inference* **98**, 95–112, 1998.
7. J. Geluk and L. de Haan. *Regular Variation, Extensions and Tauberian Theorems.* CWI Tract 40, Center for Mathematics and Computer Science, Amsterdam, Netherlands, 1987.
8. B.V. Gnedenko. Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics* **44**, 6, 423–453, 1943.
9. B.V. Gnedenko. *The Theory of Probability*, 4th edition, New York: Chelsea, 1968.

10. B.V. Gnedenko and A.N. Kolmogorov. *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley, Reading, MA, 1954.
11. M.I. Gomes and F. Caeiro. Efficiency of partially reduced-bias mean-of-order-p versus minimum-variance reduced-bias extreme value index estimation. In M. Gilli *et al.* (eds.), Proceedings of COMPSTAT 2014, The International Statistical Institute/International Association for Statistical Computing, 289-298, 2014.
12. M.I. Gomes and O. Oliveira. The bootstrap methodology in Statistics of Extremes: choice of the optimal sample fraction. *Extremes* **4**:4, 331–358, 2001.
13. M.I. Gomes, L. Henriques-Rodrigues and D. Pestana. *Non-regular Frameworks and the Mean-of-order p Extreme Value Index Estimation*, `https://doi.org/10.13140/RG.2.2.28347.64804`, 2020.
14. M.I. Gomes, L. Henriques-Rodrigues and D. Pestana. Estimação de um índice de valores extremos positivo através de médias generalizadas e em ambiente de não–regularidade. In P Milheiro *et al.* (eds.), *Estatística: Desafios Transversais às Ciências com Dados—Atas do XXIV Congresso da Sociedade Portuguesa de Estatística*. Edições SPE, 213–226, 2021.
15. L. de Haan. Slow variation and characterization of domains of attraction, In J. Tiago de Oliveira, ed., *Statistical Extremes and Applications*. D. Reidel, Dordrecht, 31–48, 1984.
16. L. de Haan and A. Ferreira. *Extreme Value Theory: an Introduction*. Springer Science+Business Media, LLC, New York, 2006.
17. L. de Haan and L. Peng. Comparison of extreme value index estimators. *Statistica Neerlandica* **52**, 60–70, 1998.
18. P. Hall. On some simple estimates of an exponent of regular variation. *J. Royal Statistical Society B* **44**, 37–42, 1982.
19. B.M. Hill. A simple general approach to inference about the tail of a distribution, *Ann. Statist.* **3**, 1163–1174, 1975.
20. A.F. Jenkinson. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quart. J. Royal Meteorol. Soc.*, **81**, 158–171, 1955.
21. R. von Misès. La distribution de la plus grande de n valeurs. In *Selected Papers*, Vol II, American Mathematical Society, 271–294, 1954.
22. V. Paulauskas and M. Vaičiulis. On the improvement of Hill and some others estimators. *Lith. Math. J.* **53**, 336–355, 2013.
23. V. Paulauskas and M. Vaičiulis. A class of new tail index estimators. *Ann. Instit. Statist. Math.* **69**, 661–487, 2017.
24. G. Samorodnitsky and M. Taqqu. *Stable Non-Gaussian Random Processes–Stochastic Models with Infinite Variance*. Chapman & Hall, USA, 1994.
25. E. Seneta. *Regularly Varying Functions*. Lecture Notes in Math. 508, Springer-Verlag, Berlin, 1976.