Bootstrap methods in statistics of extremes*

M. Ivette Gomes

Universidade de Lisboa, FCUL, DEIO and CEAUL

Frederico Caeiro

Universidade Nova de Lisboa, FCT and CMA

Lígia Henriques-Rodrigues

Instituto Politécnico de Tomar and CEAUL

B.G. Manjunath

Universidade de Lisboa, CEAUL

Abstract

In this chapter we provide an overview of the bootstrap methodology together with its possible use in the reliable estimation of any parameter of extreme or even rare events. For an asymptotically consistent choice of the *thresholds* to use in the estimation of the *extreme value index* (EVI), ξ, we suggest and discuss a double-bootstrap algorithm for the adaptive estimation of a positive EVI, the primary parameter in statistics of univariate extremes. Apart from the classical Hill and *peaks over random threshold* (PORT)-Hill EVI-estimators, we consider a class of minimum-variance reduced-bias (MVRB) EVI-estimators and associated PORT-MVRB EVI-estimators. The algorithm is described for the EVI-estimation, but it can work similarly for the estimation of other parameters of extreme events, like a *high quantile*, the *probability of exceedance* or the *return period of a high level*.

AMS 2000 subject classification. Primary 62G32, 62E20; Secondary 65C05.

Keywords and phrases. *Bootstrap methodology; semi-parametric estimation; statistics of extremes.*

1 Introduction and outline of the chapter

Let $(X_1, \ldots, X_n)$ be a random sample from an underlying *cumulative distribution function* (CDF) F. If we assume that F is known, we can easily estimate the sampling distribution

*Research partially supported by National Funds through **FCT** — Fundação para a Ciência e a Tecnologia, projects PEst-OE/MAT/UI0006/2014 (CEAUL), PEst-OE/MAT/UI0297/2014 (CMA/UNL), and Post-Doc Grants SFRH/BPD /77319/2011 and SFRH/BPD/72184/2010.

of any estimator $\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$ of an unknown parameter θ through the use of a Monte-Carlo simulation, described in the following algorithm:

S1. For $r = 1, \ldots, R$,
  S1.1 generate random samples $x_{1r}, \ldots, x_{nr} \frown F$,
  S1.2 and compute $\hat{\theta}_r = \hat{\theta}(x_{1r}, \ldots, x_{nr})$.

S2. On the basis of the output $(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_R)$, after the R iterations in Step S1., use such a sample to estimate the sampling distribution of $\hat{\theta}$, through either the associated *empirical distribution function* (EDF) or any kernel estimate, among others.

If R goes to infinity, we should then get a perfect match to the theoretical calculation, if available, i.e. the Monte-Carlo error should disappear. But, in practice, R is finite, and we thus have to cope with some quantifiable error. Moreover, F is usually unknown. How to proceed? The use of the bootstrap methodology is a possible way.

Bootstrapping (Efron, 1979) is essentially a computer-based and computer-intensive method for assigning measures of accuracy to sample estimates (see Efron and Tibshirani, 1994, and Davison and Hinkley, 1997, among others). Concomitantly, this technique also allows estimation of the sampling distribution of almost any statistic using only very simple resampling methods, based on the observed value of the EDF, given by

$$F_n^*(x) := \frac{1}{n} \sum_{i=1}^{n} I_{\{X_i \le x\}}. \tag{1.1}$$

We can replace in the previously sketched algorithm F by F_n^*, the EDF associated with the original observed data, $x_1, \ldots, x_n$, which puts mass $1/n$ on each of the $x_j, 1 \le j \le n$, generating with replacement $x_{1r}^*, \ldots, x_{nr}^* \frown F_n^*$, in Step S1.1 of the algorithm above, computing $\hat{\theta}_r^* = \hat{\theta}(x_{1r}^*, \ldots, x_{nr}^*)$, $1 \le r \le R$, in Step S1.2, and using next such a sample in Step S2.

The main goal of this chapter is to enhance the role of the *bootstrap* methodology in the field of *statistics of univariate extremes* (SUE). In SUE, the bootstrap has been commonly used in the choice of the number k of top *order statistics* (OSs) or of the *optimal sample fraction* (OSF), k/n, to be taken in the semi-parametric estimation of a parameter of extreme or even rare events. For an asymptotically consistent choice of the *thresholds* to use in the adaptive estimation of a positive *extreme value index* (EVI), ξ, the primary parameter in SUE, we suggest and discuss a double-bootstrap algorithm. In such algorithm, apart from the classical Hill (Hill, 1975) and *peaks over random threshold* (PORT)-Hill EVI-estimators (Araújo Santos *et al.*, 2006), we consider a class of minimum-variance reduced-bias (MVRB), the simplest one in Caeiro *et al.* (2005), and associated PORT-MVRB (Gomes *et al.*, 2011a, 2013) EVI-estimators.

After providing, in Section 2, a few technical details in the area of *extreme value theory* (EVT), related to the EVI-estimators under consideration in this chapter, we shall briefly discuss, in Section 3, the main ideas behind the bootstrap methodology and OSF-estimation. In the lines of Gomes *et al.* (2011b, 2012, 2014), we propose an algorithm for the adaptive consistent estimation of a positive EVI, through the use of resampling computer-intensive methods. The **Algorithm** is described for the Hill EVI-estimator and associated PORT-Hill, MVRB and PORT-MVRB EVI-estimators, but it can work similarly for the estimation of other parameters of extreme events, like a high quantile, the probability of exceedance or the return period of a high level. Section 4 is entirely dedicated to the application of the **Algorithm** to three simulated samples. Finally, in Section 5, we draw some overall conclusions.

2 A few details on EVT

The key result obtained by Fisher and Tippet (1928) on the possible limiting laws of the sample maxima, formalized by Gnedenko (1943), and used by Gumbel (1958) for applications of EVT in engineering subjects, are some of the key tools that led to the way statistical EVT has been exploding in the last decades. In this chapter, we focus on the behaviour of extreme values of a data set, dealing with maximum values and other top OSs in a univariate framework, working thus in the field of SUE.

Let us assume that we have access to a random sample, $(X_1, \ldots, X_n)$ of *independent, identically distributed*, or possibly stationary and weakly dependent, *random variables* (RVs) from an underlying model F, and let us denote by $(X_{1:n} \leq \cdots \leq X_{n:n})$ the sample of associated ascending OSs. As usual, let us further assume that it is possible to linearly normalize the sequence of maximum values, $\{X_{n:n}\}_{n\geq 1}$, so that we get a non-degenerate limit. Then (Gnedenko, 1943), that limiting RV has a CDF of the type of the *extreme value* (EV) CDF, given by

$$\mathrm{EV}_\xi(x) = \begin{cases} \exp\left(-(1+\xi x)^{-1/\xi}\right), \ 1+\xi x \geq 0, & \text{if} \quad \xi \neq 0, \\ \exp(-\exp(-x)), \ x \in \mathbb{R}, & \text{if} \quad \xi = 0, \end{cases} \tag{2.1}$$

and ξ is the so-called EVI, the primary parameter in SUE. We then say that F is in the max-domain of attraction of EV_ξ, in (2.1), and use the notation $F \in \mathcal{D}_{\mathcal{M}}(\mathrm{EV}_\xi)$. The EVI measures essentially the weight of the *right tail-function* (RTF), $\overline{F} := 1 - F$. If $\xi < 0$, the right tail is short and light, since F has compulsory a finite right endpoint, i.e. $x^F := \sup\{x : F(x) < 1\}$ is finite. If $\xi > 0$, the right tail is heavy, of a negative polynomial type, and F has an infinite right endpoint. If $\xi = 0$, the right tail is of an exponential type, and the right endpoint can then be either finite or infinite.

Slightly more restrictively than the full max-domain of attraction of the EV CDF, we now consider a positive EVI, i.e. we work with heavy-tailed models F in $\mathcal{D}_{\mathcal{M}}\left(\mathrm{EV}_{\xi}\right)_{\xi>0} =: \mathcal{D}_{\mathcal{M}}^{+}$. Heavy-tailed models appear often in practice in fields like bibliometrics, biostatistics, finance, insurance and telecommunications. Power laws, such as the Pareto distribution and the Zipf's law, have been observed a few decades ago in some important phenomena in economics and biology, and have seriously attracted scientists in recent years. As usual, we shall further use the notations $F^{\leftarrow}(y) := \inf\{x : F(x) \geq y\}$ for the *generalized inverse* function of F, and $\mathcal{R}_{\mathrm{a}}$ for the class of *regularly varying* functions at infinity with an index of regular variation a, i.e. positive Borel measurable functions $g(\cdot)$ such that $g(tx)/g(t) \to x^{\mathrm{a}}$, as $t \to \infty$, for all $x > 0$ (see Bingham *et al.*, 1987, for details on regular variation). Let us further use the notation $U(t) := F^{\leftarrow}\big(1 - 1/t\big)$, for the *tail quantile function* (TQF). For heavy-tailed models we have the validity of the following *first-order conditions*,

$$F \in \mathcal{D}_{\mathcal{M}}^{+} \quad \Longleftrightarrow \quad \overline{F} \in \mathcal{R}_{-1/\xi} \quad \Longleftrightarrow \quad U \in \mathcal{R}_{\xi}. \tag{2.2}$$

The first necessary and sufficient condition above, related to the RTF behaviour, was proved by Gnedenko (1943) and the second one, related to the TQF behaviour, was proved in de Haan (1984).

For these heavy-tailed models, and given a sample $\underline{\mathbf{X}}_n = (X_1, \ldots, X_n)$, the classical EVI-estimators are Hill estimators (Hill, 1975), with the functional expression

$$H_{k,n} \equiv H(k; \underline{\mathbf{X}}_n) := \frac{1}{k} \sum_{i=1}^{k} \ln \frac{X_{n-i+1:n}}{X_{n-k:n}}, \quad 1 \leq i \leq k < n. \tag{2.3}$$

They are thus the average of the k log-excesses, $V_{ik} := \ln X_{n-i+1:n} - \ln X_{n-k:n}$, $1 \leq i \leq k$, above the random level or threshold $X_{n-k:n}$. Such a random threshold $X_{n-k:n}$ compulsory needs to be an *intermediate* OS, i.e. we need to have

$$k = k_n \to \infty \quad \text{and} \quad k/n \to 0, \text{ as } n \to \infty, \tag{2.4}$$

if we want to have consistent EVI-estimation in the whole $\mathcal{D}_{\mathcal{M}}^{+}$. Indeed, under any of the first-order frameworks in (2.2), the *log-excesses*, $V_{ik}, 1 \leq i \leq k$, are approximately the k OSs of an exponential sample of size k, with mean value ξ. Hence, the reason for the EVI-estimators in (2.3).

Under adequate second-order conditions, that rule the rate of convergence in any of the first-order conditions in (2.2), Hill estimators, $H_{k,n}$, have usually a high asymptotic bias, and recently, several authors have considered different ways of reducing bias in the area of SUE (see the overviews in Gomes *et al.*, 2007b, Chapter 6 of Reiss and Thomas, 2007; Gomes *et al.*, 2008a; Beirlant *et al.*, 2012). A simple class of MVRB EVI-estimators is the class studied in Caeiro *et al.* (2005), to be introduced in Section 2.2. These MVRB EVI-estimators depend on the adequate estimation of second-order parameters, and the kind of second-order parameters' estimation which enables the building of MVRB EVI-estimators, i.e. EVI-estimators which outperform the Hill estimator for all k, is sketched in Sections 2.2.1 and 2.2.2.

But both the Hill and the MVRB EVI-estimators are invariant to changes in scale but not invariant to changes in location. And particularly the Hill EVI-estimators can suffer drastic changes when we induce an arbitrary shift in the data, as can be seen in Figure 1.

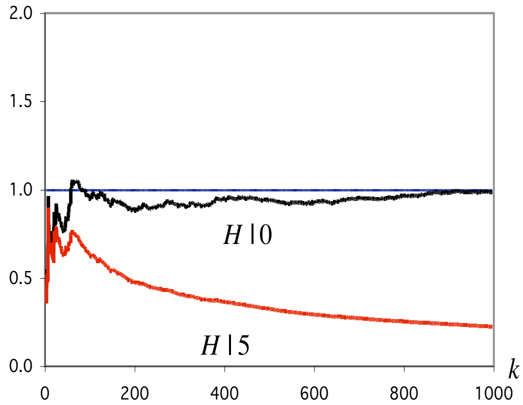

Figure 1: Hill-plots associated to a unit Pareto sample of size $n = 1000$, from model $F_0(x) = 1 - x^{-1}$, $x \geq 1$, $H|0$, and to the shifted sample from $F_5(x) = 1 - (x-5)^{-1}$, $x \geq 6$, $H|5$.

Indeed, even if a Hill-plot (a function of $H_{k,n}$ versus k) looks stable, as happens in Figure 1, with the $H|0$ sample path, where data, $(x_1, \ldots, x_n)$, $n = 1000$, come from a unit Pareto CDF, $F(x) = F_0(x) = 1 - x^{-\alpha}$, $x \geq 1$, for $\alpha = 1$ ($\xi = 1/\alpha = 1$), we easily come to the so-called 'Hill horror-plots', a terminology used in Resnick (1997), when we induce a shift to the data. This can be seen also in Figure 1 (look now at $H|5$), where we present the Hill-plot associated to the shifted sample $(x_1 + 5, \ldots, x_n + 5)$, from the CDF $F_s(x) = 1 - (x-s)^{-1}$, $x \geq 1 + s$, now for $s = 5$. This led Araújo Santos *et al.* (2006) to introduce the so-called PORT methodology, to be sketched in Section 2.3. The asymptotic behaviour of the EVI-estimators under consideration is discussed in Section 2.4.

2.1 Second-order reduced-bias (SORB) EVI-estimation

For consistent semi-parametric EVI-estimation, in the whole $\mathcal{D}^+_{\mathcal{M}}$, we have already noticed that we merely need to work with adequate functionals, dependent on an *intermediate tuning* parameter k, the number of top OSs involved in the estimation, i.e. (2.4) should hold. To obtain full information on the non-degenerate asymptotic behaviour of semi-parametric

EVI-estimators, we often need further assuming a *second-order condition*, ruling the rate of convergence in any of the *first-order conditions*, in (2.2). It is often assumed that there exists a function $A(\cdot)$, such that

$$\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \xi \ln x}{A(t)} = \psi_\rho(x) := \begin{cases} (x^\rho - 1)/\rho, & \text{if} \quad \rho \neq 0, \\ \ln x, & \text{if} \quad \rho = 0. \end{cases} \tag{2.5}$$

Then, we have $|A| \in \mathcal{R}_\rho$. Moreover, if the limit in the left hand-side of (2.5) exists, we can choose $A(\cdot)$ so that such a limit is compulsory equal to the above defined $\psi_\rho(\cdot)$ function (Geluk and de Haan, 1987).

Whenever dealing with SORB estimators of parameters of extreme events, and essentially due to technical reasons, it is common to slightly restrict the domain of attraction, $\mathcal{D}_{\mathcal{M}}^+$, and to consider a Pareto-type class of models, assuming that, with $C, \xi > 0$, $\rho < 0$, $\beta \neq 0$, and as $t \to \infty$,

$$U(t) = Ct^\xi \big(1 + A(t)/\rho + o(t^\rho)\big), \quad A(t) := \xi\beta t^\rho. \tag{2.6}$$

The class in (2.6) is however a wide class of models, that contains most of the heavy-tailed parents useful in applications, like the *Fréchet*, the *generalized Pareto* and the *Student*-t_ν, with ν degrees of freedom. Note that the validity of (2.5) with $\rho < 0$ is equivalent to (2.6). To obtain information on the bias of MVRB EVI-estimators it is even common to slightly restrict the class of models in (2.6), further assuming the following third-order condition,

$$U(t) = Ct^\xi \big(1 + A(t)/\rho + \beta' t^{2\rho} + o(t^{2\rho})\big), \tag{2.7}$$

as $t \to \infty$, with $\beta' \neq 0$. All the above mentioned models still belong to this class. Slightly more generally, we could have assumed a *general third-order condition*, ruling now the rate of convergence in the second-order condition in (2.5), which guarantees that, for all $x > 0$,

$$\lim_{t\to\infty} \frac{\frac{\ln U(tx) - \ln U(t) - \xi \ln x}{A(t)} - \psi_\rho(x)}{B(t)} = \psi_{\rho+\rho'}(x), \tag{2.8}$$

where $|B|$ must then be in $\mathcal{R}_{\rho'}$. Equation (2.7) is equivalent to equation (2.8) with $\rho = \rho' < 0$. Further details on the topic can be found in de Haan and Ferreira (2006).

Provided that (2.4) and (2.5) hold, Hill EVI-estimators, $H_{k,n}$, have usually a high asymptotic bias. The adequate accommodation of this bias has recently been extensively addressed. Among the pioneering papers, we mention Peng (1998), Beirlant *et al.* (1999), Feuerverger and Hall (1999) and Gomes *et al.* (2000). In these papers, authors are led to SORB EVI-estimators, with asymptotic variances larger than or equal to $(\xi\ (1-\rho)/\rho)^2$, where $\rho(< 0)$ is the aforementioned 'shape' second-order parameter in (2.5). Recently, as sketched in Section 2.2, Caeiro *et al.* (2005) and Gomes *et al.* (2007a; 2008c) have been able to *reduce the bias without increasing the asymptotic variance*, kept at ξ^2, just as happens with the Hill EVI-estimator.

2.2 MVRB EVI-estimation

To reduce bias, keeping the asymptotic variance at the same level, we merely need to use an adequate 'external' and a bit more than consistent estimation of the pair of second-order parameters, $(\beta, \rho) \in (\mathbb{R}, \mathbb{R}^-)$, in (2.6). The MVRB EVI-estimators outperform the classical Hill EVI-estimators for all k, and among them, we now consider the simplest class in Caeiro *et al.* (2005), used for Value-at-Risk (VaR) estimation in Gomes and Pestana (2007b). Such a class, denoted by $\overline{H} \equiv \overline{H}_{k,n}$, has the functional form

$$\overline{H}_{k,n} \equiv \overline{H}_{\hat{\beta},\hat{\rho}}(k; \underline{\mathbf{X}}_n) := H_{k,n}\big(1 - \hat{\beta}(n/k)^{\hat{\rho}}/(1 - \hat{\rho})\big), \tag{2.9}$$

where $(\hat{\beta}, \hat{\rho})$ is an adequate consistent estimator of (β, ρ), with $\hat{\beta}$ and $\hat{\rho}$ based on a number of top OSs k_1 usually of a higher order than the number of top OSs k used in the EVI-estimation, as explained in Sections 2.2.1 and 2.2.2. For different algorithms for the estimation of (β, ρ), see Gomes and Pestana (2007a,b).

2.2.1 Estimation of the 'shape' second-order parameter

We consider the most commonly used ρ-estimators, the ones studied in Fraga Alves *et al.* (2003), briefly introduced in the sequel. Given the sample $\underline{\mathbf{X}}_n$, the ρ-estimators in Fraga Alves *et al.* (2003) are dependent on the statistics

$$V_\tau(k; \underline{\mathbf{X}}_n) := \begin{cases} \dfrac{\big(M_n^{(1)}(k;\underline{\mathbf{X}}_n)\big)^{\tau} - \big(M_n^{(2)}(k;\underline{\mathbf{X}}_n)/2\big)^{\tau/2}}{\big(M_n^{(2)}(k;\underline{\mathbf{X}}_n)/2\big)^{\tau/2} - \big(M_n^{(3)}(k;\underline{\mathbf{X}}_n)/6\big)^{\tau/3}}, & \text{if } \tau \neq 0, \\[2ex] \dfrac{\ln M_n^{(1)}(k;\underline{\mathbf{X}}_n) - \ln\big(M_n^{(2)}(k;\underline{\mathbf{X}}_n)/2\big)/2}{\ln\big(M_n^{(2)}(k;\underline{\mathbf{X}}_n)/2\big)/2 - \ln\big(M_n^{(3)}(k;\underline{\mathbf{X}}_n)/6\big)/3}, & \text{if } \tau = 0, \end{cases} \tag{2.10}$$

defined for any *tuning parameter* $\tau \in \mathbb{R}$, and where

$$M_n^{(j)}(k; \underline{\mathbf{X}}_n) := \frac{1}{k} \sum_{i=1}^{k} \{\ln X_{n-i+1:n} - \ln X_{n-k:n}\}^j, \quad j = 1, 2, 3.$$

Under mild restrictions on k, i.e., if (2.4) holds and $\sqrt{k}\, A(n/k) \to \infty$, with $A(\cdot)$ the function in (2.6), the statistics in (2.10) converge towards $3(1-\rho)/(3-\rho)$, independently of the *tuning* parameter τ, and we can consequently consider the class of admissible ρ-estimators,

$$\hat{\rho}_\tau(k) \equiv \hat{\rho}_\tau(k; \underline{\mathbf{X}}_n) := -\left|\frac{3(V_\tau(k; \underline{\mathbf{X}}_n) - 1)}{V_\tau(k; \underline{\mathbf{X}}_n) - 3}\right|. \tag{2.11}$$

Under adequate general conditions, and for an appropriate tuning parameter τ, the ρ-estimators in (2.11) show highly stable sample paths as functions of k, the number of top OSs used, for a range of large k-values. Again, it is sensible to advise practitioners not to choose blindly the value of τ in (2.11). Sample paths of $\hat\rho_\tau(k)$, as functions of k, for a few values of τ, should be drawn, in order to elect the value of τ which provides higher stability for large k, by means of any stability criterion. For the most common stability criterion, see Gomes and Pestana, 2007b, and Remark 3.1. The value $\tau = 0$, considered in the description of the **Algorithm** in Section 3.2, has revealed to be the most adequate choice whenever we are in the region $|\rho| \le 1$, a common region in applications and the region where bias reduction is indeed needed. Distributional properties of the estimators in (2.11) can be found in Fraga Alves *et al.* (2003). Interesting alternative classes of ρ-estimators have recently been introduced in Goegebeur *et al.* (2008, 2010), Ciuperca and Mercadier (2010) and Caeiro and Gomes (2012a,b).

2.2.2 Estimation of the 'scale' second-order parameter

For the estimation of the scale second-order parameter β, on the basis of

$$d_\alpha(k) := \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\alpha} \quad \text{and} \quad D_\alpha(k) := \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\alpha} i \ln\frac{X_{n-i+1:n}}{X_{n-i:n}}, \ \alpha \in \mathbb{R},$$

we shall consider the estimator in Gomes and Martins (2002),

$$\hat\beta_{\hat\rho}(k) \equiv \hat\beta_{\hat\rho}(k; \underline{\mathbf{X}}_n) := \left(\frac{k}{n}\right)^{\hat\rho} \frac{d_{\hat\rho}(k)\ D_0(k) - D_{\hat\rho}(k)}{d_{\hat\rho}(k)\ D_{\hat\rho}(k) - D_{2\hat\rho}(k)}, \tag{2.12}$$

dependent on an adequate ρ-estimator, $\hat\rho$. It has been advised the computation of these second-order parameters' estimators at a k-value given by

$$k_1 = \lfloor n^{1-\epsilon} \rfloor, \quad \epsilon = 0.001. \tag{2.13}$$

The estimator $\hat\rho$, to be plugged in (2.12), is thus $\hat\rho := \hat\rho_\tau(k_1; \underline{\mathbf{X}}_n)$, with $\hat\rho_\tau(k; \underline{\mathbf{X}}_n)$ and k_1 given in (2.11) and (2.13), respectively.

Remark 2.1. *Note that only the external estimation of both β and ρ at a level k_1 adequately chosen, and the* EVI*-estimation at a level $k = o(k_1)$, or at a specific value $k = O(k_1)$, can lead to a* MVRB EVI*-estimator, with an asymptotic variance ξ^2. Such a choice of (k, k_1) is theoretically possible, as shown in* Gomes *et al.* (2008d) *and* Caeiro *et al.* (2009), *but under conditions difficult to guarantee in practice. As a compromise between theoretical and practical results, we have so far advised any choice $k_1 = \lfloor n^{1-\epsilon} \rfloor$, with ϵ small (see* Caeiro *et al.*, 2005, 2009, *and* Gomes *et al.*, 2007a,b, 2008c, *among others). With the choice of k_1 in* (2.13), *we have obviously the validity of condition* (2.4), *and whenever $\sqrt{k_1}\ A(n/k_1) \to \infty$, as $n \to \infty$ (an almost irrelevant restriction, from a practical point of view), we get $\hat\rho - \rho := \hat\rho_\tau(k_1) - \rho = o_p(1/\ln n)$, a condition needed, in order not to have any increase in the asymptotic variance of the bias-corrected* Hill EVI*-estimator in equation* (2.9), *comparatively with the one of the* Hill EVI*-estimator, in* (2.3).

Remark 2.2. *Further note that the estimation of ξ, β and ρ at the same value k would lead to a high increase in the asymptotic variance of the* SORB EVI-*estimators* $\overline{H}_{k,n;\hat\beta,\hat\rho}$ *in* (2.9)*, which would become* $\xi^2\left((1-\rho)/\rho\right)^4$ (*see* Feuerverger and Hall, 1999; Beirlant *et al.*, 1999; Peng and Qi, 2004*, also among others*)*. The external estimation of ρ at k_1, but the estimation of ξ and β at the same $k = o(k_1)$, enables a slight decreasing of the asymptotic variance to* $\xi^2\left((1-\rho)/\rho\right)^2$*, still greater than* ξ^2 (*see* Gomes and Martins, 2002*, again among others*)*.*

Details on the distributional behaviour of the estimator in (2.12) can be found in Gomes and Martins (2002) and more recently in Gomes *et al.* (2008c) and Caeiro *et al.* (2009). Again, consistency is achieved for models in (2.6), and k values such that (2.4) holds and $\sqrt{k}\ A(n/k) \to \infty$, as $n \to \infty$. Alternative estimators of β can be found in Caeiro and Gomes (2006) and Gomes *et al.* (2010). Due to the fact that $\hat\beta = \hat\beta_{\hat\rho}(k_1)$ and $\hat\rho = \hat\rho_\tau(k_1)$, with $\hat\rho_\tau(k)$, $\hat\beta_{\hat\rho}(k)$ and k_1 given in (2.11), (2.12) and (2.13), respectively, depend on $\tau \in \mathbb{R}$, we often use the notation $\overline{H}_\tau = \overline{H}$. But when we work with $\tau = 0$ only, as happens in Section 3.2, we shall not use the subscript $\tau = 0$. Note however that the **Algorithm** in Section 3.2 can also be used for another fixed choice of τ, as well as for a data-driven choice of τ provided by any of the algorithms in Gomes and Pestana (2007a,b), among others.

2.3 PORT EVI-estimation

The estimators in (2.3) and (2.9) are scale invariant but not location invariant. In order to achieve location invariance for a class of modified Hill EVI-estimators and adequate properties for VaR-estimators, Araújo Santos *et al.* (2006) introduced the so-called PORT methodology. The estimators are then functionals of a sample of excesses over a random level $X_{n_q:n}$, $n_q := \lfloor nq \rfloor + 1$, i.e. functionals of the sample,

$$\underline{\mathbf{X}}_n^{(q)} := \left(X_{n:n} - X_{n_q:n}, \ldots, X_{n_q+1:n} - X_{n_q:n}\right). \tag{2.14}$$

Generally, we can have $0 < q < 1$, for any $F \in \mathcal{D}_{\mathcal{M}}^+$ (*the random level is an empirical quantile*). If the underlying model F has a finite *left endpoint*, $x_{_F} := \inf\{x : F(x) \geq 0\}$, we can also use $q = 0$ (*the random level can then be the minimum*).

If we think, for instance, on Hill EVI-estimators, in (2.3), the new classes of PORT-Hill EVI-estimators, theoretically studied in Araújo Santos *et al.* (2006), and for finite samples in Gomes *et al.* (2008b), are given by

$$H_{k,n}^{(q)} := H(k; \underline{\mathbf{X}}_n^{(q)}) = \frac{1}{k}\ \sum_{i=1}^{k} \left\{ \ln\ \frac{X_{n-i+1:n} - X_{n_q:n}}{X_{n-k:n} - X_{n_q:n}} \right\}, \quad 0 \leq q < 1. \tag{2.15}$$

Similarly, if we think on the MVRB EVI-estimators, in (2.9), the new classes of PORT-MVRB EVI-estimators, studied for finite samples in Gomes *et al.* (2011a; 2013), are given by

$$\overline{H}^{(q)}_{k,n} := \overline{H}_{\hat\beta^{(q)},\hat\rho^{(q)}}(k;\underline{\mathbf{X}}^{(q)}_n) = H^{(q)}_{k,n}\Big(1-\hat\beta^{(q)}(n^{(q)}/k)^{\hat\rho^{(q)}}/(1-\hat\rho^{(q)})\Big),\quad 0\le q<1, \tag{2.16}$$

with $H^{(q)}_{k,n}$ in (2.15), $n^{(q)} = n - n_q$ and $\big(\hat\beta^{(q)} := \hat\beta(k_1,\underline{\mathbf{X}}^{(q)}_n), \hat\rho^{(q)} := \hat\rho(k_1,\underline{\mathbf{X}}^{(q)}_n)\big)$ any adequate estimator of (β_q,ρ_q), the vector of second-order parameters associated with the shifted model, based on the sample $\underline{\mathbf{X}}^{(q)}_n$, in (2.14).

These PORT EVI -estimators are thus dependent on a *tuning parameter* q, $0\le q<1$, that makes them highly flexible. Moreover, they are invariant to changes in both location and scale. Just as in Gomes *et al.* (2013; 2014), we shall further include in the algorithm the value $q=1$, so that with H, $\overline{H}$, $H^{(q)}$ and $\overline{H}^{(q)}$, given in (2.3), (2.9), (2.15) and (2.16), respectively, we can consider that $H = H^{(q)}$ and $\overline{H} = \overline{H}^{(q)}$ for $q=1$ (with the notations $n_1 = n+1 \equiv 0$, $X_{n+1:n} = X_{0:n}\equiv 0$, so that $\underline{\mathbf{X}}^{(1)}_n = \underline{\mathbf{X}}_n$, $n^{(1)} = n-n_1 = n$, $\hat\beta^{(1)} = \hat\beta$, $\hat\rho^{(1)} = \hat\rho$).

Remark 2.3. *The* PORT VaR_p*-estimators at a level p, $0<p<1$, introduced in* Araújo Santos *et al.* (2006)*, are also semi-parametric in nature, and connected to the* Weissman-Hill *estimator*

$$\widehat{\chi}_p(k) \equiv \widehat{\chi}_p(k;\underline{X}_n) := X_{n-k:n}\big(k/(np)\big)^{H_{k,n}}, \tag{2.17}$$

studied in Weissman (1978)*. However, they satisfy the empirical counterpart of the theoretical linear property of a quantile $\chi_p(X) := F^{\leftarrow}(1-p)$, given by $\chi_p(\lambda+\delta X) = \lambda+\delta\chi_p(X)$ for any real λ and positive real δ. Up to an adequate translation, they have the same functional expression of the* VaR_p*-estimators in* (2.17)*, but applied to the sample of excesses in* (2.14)*. More precisely, they are given by*

$$\widehat{\chi}^{(q)}_p(k) := (X_{n-k:n} - X_{n_q:n})\big(k/(np)\big)^{H^{(q)}_{k,n}} + X_{n_q:n}, \tag{2.18}$$

with $H^{(q)}_{k,n}$ the PORT-Hill EVI*-estimator in* (2.15)*. An expression similar to* (2.18) *can be written for a* PORT-MVRB VaR_p*-estimator, provided that we replace in* (2.18)*, $H^{(q)}_{k,n}$ by $\overline{H}^{(q)}_{k,n}$, in* (2.16)*.*

2.4 Asymptotic properties of the EVI-estimators

The Hill estimator reveals usually a high asymptotic bias. Indeed, from the results of de Haan and Peng (1998), and with $\mathcal{N}_{\mu,\sigma^2}$ denoting a normal RV with mean value μ and variance σ^2,

$$\sqrt{k}\,(H_{k,n}-\xi) \;\overset{d}{=}\; \mathcal{N}_{0,\xi^2} + b_H\sqrt{k}A(n/k) + o_p\big(\sqrt{k}A(n/k)\big), \tag{2.19}$$

where the bias $b_H\sqrt{k}A(n/k) = \xi\ \beta\ \sqrt{k}\ (n/k)^\rho/(1-\rho)$ under condition (2.7) can be very large, moderate or small, going respectively to ∞, a non-null constant or 0, as $n\to\infty$.

This non-null asymptotic bias, together with a rate of convergence of the order of $1/\sqrt{k}$, leads to sample paths with a high variance for small k, a high bias for large k, and a very sharp mean square error (MSE) pattern, as a function of k. Under the same conditions as before, $\sqrt{k}\,(\overline{H}_{k,n} - \xi)$ is asymptotically normal with variance also equal to ξ^2 but with a null mean value. Indeed, under the validity of the aforementioned third-order condition in (2.7), related to Pareto-type class of models, we can adequately estimate the vector of second-order parameters, (β, ρ) so that $\overline{H}_{k,n}$ outperforms $H_{k,n}$ for all k. Indeed, and for an adequate $b_{\overline{H}}$, computed in Caeiro *et al.* (2009), we can write

$$\sqrt{k}\,(\overline{H}_{k,n} - \xi) \stackrel{d}{=} \mathcal{N}_{0,\xi^2} + b_{\overline{H}}\sqrt{k}A^2(n/k) + o_p\big(\sqrt{k}A^2(n/k)\big). \tag{2.20}$$

We can further summarize the aforementioned results in the following theorem.

Theorem 2.1. *Assume that condition* (2.5) *holds, and let $k \equiv k_n$ be an intermediate sequence, i.e.,* (2.4) *holds. Then $H_{k,n}$, in* (2.3)*, is consistent for the estimation of ξ. Moreover, there exist a sequence Z_k^H of asymptotically standard normal* RVs*, and real numbers $\sigma_H = \xi > 0$ and $b_{H,1} = b_H$, such that the asymptotic distributional representation in* (2.19) *holds.*

If we further assume that (2.7) *holds, there exists an extra real number $b_{H,2}$, such that we can write*

$$\sqrt{k}\big(H_{k,n} - \xi\big) \stackrel{d}{=} \mathcal{N}_{0,\xi^2} + b_{H,1}\sqrt{k}A(n/k) + b_{H,2}\sqrt{k}A^2(n/k)\big(1 + o_p(1)\big).$$

If under the validity of the second-order condition in (2.5)*, we estimate β and ρ consistently through $\hat{\beta}$ and $\hat{\rho}$, in such a way that $\hat{\rho} - \rho = o_p(1/\ln n)$, the asymptotic distributional representation*

$$\sqrt{k}\big(\overline{H}_{k,n} - \xi\big) \stackrel{d}{=} \mathcal{N}_{0,\xi^2} + o_p\big(\sqrt{k}A(n/k)\big)$$

holds. Under the validity of equation (2.7)*, we can guarantee that there exists a pair of real numbers $(b_{\overline{H},1}, b_{\overline{H},2})$, but with $b_{\overline{H},1} = 0$ and $b_{\overline{H},2} = b_{\overline{H}}$, given in* (2.20)*, and*

$$\begin{aligned}\sqrt{k}\big(\overline{H}_{k,n} - \xi\big) &\stackrel{d}{=} \mathcal{N}_{0,\xi^2} + b_{\overline{H},1}\sqrt{k}A(n/k) + b_{\overline{H},2}\sqrt{k}A^2(n/k)\ (1 + o_p(1)) \\ &\stackrel{d}{=} \mathcal{N}_{0,\xi^2} + b_{\overline{H},2}\sqrt{k}A^2(n/k)\ (1 + o_p(1)),\end{aligned}$$

for adequate k values such that $\sqrt{k}A^2(n/k) \to \lambda_A$, finite.

For the asymptotic behaviour of the PORT-Hill EVI-estimators, we refer Araújo Santos *et al.* (2006). The full asymptotic behaviour of the PORT MVRB EVI-estimators is still under development. It is known that the rate of convergence and asymptotic variance do not change. There are however big changes in the bias but for adequate q-values the PORT-MVRB EVI-estimators are indeed MVRB EVI-estimators. Contrarily to what has been done in Gomes *et al.* (2014), we shall thus consider for them the same double-bootstrap **Algorithm** we use for the MVRB EVI-estimation.

3 The bootstrap methodology in SUE

The use of bootstrap resampling methodologies has revealed to be promising in the choice of the nuisance parameter k, or equivalently of the OSF, k/n, in the semi-parametric estimation of any parameter of extreme events. If we ask how to choose the tuning parameter k in the EVI-estimation, either through $H_{k,n}$ or $\overline{H}_{k,n}$ or $H^{(q)}_{k,n}$ or $\overline{H}^{(q)}_{k,n}$, $0 \le q < 1$, generally denoted $E_{k,n}$, we usually consider the estimation of

$$k_{0|E}(n) := \arg\min_k \mathrm{MSE}(E_{k,n}). \tag{3.1}$$

To obtain estimates of $k_{0|E}(n)$, one can use a *double-bootstrap* method applied to an adequate *auxiliary statistic* like

$$T_{k,n} \equiv T_{k,n|E} := E_{\lfloor k/2 \rfloor,n} - E_{k,n}, \quad k = 2, \dots, n-1, \tag{3.2}$$

which tends to the well-known value **zero** and has an asymptotic behaviour similar to the one of $E_{k,n}$ (see Gomes and Oliveira, 2001, among others, for the estimation through $H_{k,n}$ and Gomes *et al.*, 2012, for the the estimation through $\overline{H}_{k,n}$). See also Gomes *et al.* (2014) and Section 3.2 of this chapter.

With AMSE standing for 'asymptotic MSE', and on the basis of (2.19), and (2.20), we get

$$\begin{aligned} k_{A|E}(n) &:= \arg\min_k \mathrm{AMSE}\big(E_{k,n}\big) \\ &= \arg\min_k \begin{cases} \xi^2/k + b_{_E}^2\ A^2(n/k), & \text{if} \quad E = H, \\ \xi^2/k + b_{_E}^2\ A^4(n/k), & \text{if} \quad E = \overline{H}, \end{cases} \\ &= k_{0|E}(n)(1 + o(1)), \end{aligned} \tag{3.3}$$

with $k_{0|E}(n)$ defined in (3.1). See Theorem 1 of Draisma *et al.*, 1999, for a proof of this result, in the case of H. The proof is similar for the cases of $\overline{H}$, as already mentioned in Gomes *et al.* (2012). Things work more intricately for the PORT-MVRB EVI-estimators, and as mentioned above we shall consider an algorithm similar to the one devised for the MVRB EVI-estimators in case we are working with $\overline{H}^{(q)}$, $0 \le q < 1$, since we are interested in the possible specific value of q that makes these PORT estimators MVRB EVI-estimators. The bootstrap methodology enables us to estimate $k_{0|E}(n)$, in (3.1), in a way similar to the one used for the classical EVI-estimators, on the basis of a consistent estimator of $k_{A|E}(n)$, in (3.3), and now through the use of an auxiliary statistic like the one in (3.2), a method detailed in Gomes *et al.* (2011b; 2012) for the MVRB EVI-estimation. For sake of simplicity, we shall next describe the methodology for $\big(H, \overline{H}\big)$, but similar formulas work for $\big(H^{(q)}, \overline{H}^{(q)}\big)$ provided that we replace $(n, \underline{X}_n, \beta, \rho)$ by $(n^{(q)}, \underline{X}_n^{(q)}, \beta_q, \rho_q)$, $0 \le q \le 1$. Indeed, under the above-mentioned third-order framework in (2.7),

$$T_{k,n|E} \overset{d}{=} \frac{\xi\ P_k^E}{\sqrt{k}} + \begin{cases} b_{E,1}(2^\rho - 1)\ A(n/k)(1+o_p(1)), & \text{if} \quad E = H, \\ b_{E,2}(2^{2\rho} - 1)\ A^2(n/k)(1+o_p(1)), & \text{if} \quad E = \overline{H}, \end{cases}$$

with P_k^E asymptotically standard normal.

Consequently, denoting $k_{0|T,E}(n) := \arg\min_k \text{MSE}(T_{k,n|E})$, we have

$$k_{0|E}(n) = k_{0|T,E}(n) \times \begin{cases} (1-2^\rho)^{\frac{2}{1-2\rho}}\,(1+o(1)), & \text{if} \quad E = H, \\ (1-2^{2\rho})^{\frac{2}{1-4\rho}}\,(1+o(1)), & \text{if} \quad E = \overline{H}. \end{cases} \tag{3.4}$$

3.1 The resampling methodology in action

How does the resampling methodology then work? Given the sample $\underline{X}_n = (X_1, \ldots, X_n)$ from an unknown model F, and the functional in (3.2), $T_{k,n|E} =: \phi_k(\underline{X}_n)$, $1 < k < n$, consider for any $m_1 = O(n^{1-\epsilon})$, $0 < \epsilon < 1$, the bootstrap sample

$$\underline{X}^*_{m_1} = (X^*_1, \ldots, X^*_{m_1}),$$

from F^*_n, in (1.1), the EDF associated with the available random sample, $\underline{X}_n$.

Next, associate to the bootstrap sample the corresponding bootstrap auxiliary statistic, $T^*_{k_1,m_1|E} := \phi_{k_1}(\underline{X}^*_{m_1})$, $1 < k_1 < m_1$. Then, with $k^*_{0|T,E}(m_1) = \arg\min_{k_1} \text{MSE}\big(T^*_{k_1,m_1|E}\big)$,

$$\frac{k^*_{0|T,E}(m_1)}{k_{0|T,E}(n)} = \left(\frac{m_1}{n}\right)^{-\frac{c\rho}{1-c\rho}}(1+o(1)), \qquad c = \begin{cases} 2 & \text{if} \quad E = H, \\ 4 & \text{if} \quad E = \overline{H}. \end{cases}$$

Consequently, for another sample size m_2, and for every $a > 1$,

$$\frac{\big(k^*_{0|T,E}(m_1)\big)^a}{k^*_{0|T,E}(m_2)} = \left(\frac{m_1^a}{n^a}\frac{n}{m_2}\right)^{-\frac{c\rho}{1-c\rho}}\big(k_{0|T,E}(n)\big)^{a-1}(1+o(1)).$$

It is then enough to choose $m_2 = \lfloor n\,(m_1/n)^a \rfloor + 1$, in order to have independence of ρ. If we consider $a = 2$, i.e. $m_2 = \lfloor m_1^2/n \rfloor + 1$, we have

$$\big(k^*_{0|T,E}(m_1)\big)^2 / k^*_{0|T,E}(m_2) = k_{0|T,E}(n)(1+o(1)), \text{ as } n \to \infty. \tag{3.5}$$

On the basis of (3.5), we are now able to consistently estimate $k_{0|T,E}$ and next $k_{0|E}$ through (3.4), on the basis of any estimate $\hat\rho$ of the second-order parameter ρ. With $\hat{k}^*_{0|T,E}$ denoting the sample counterpart of $k^*_{0|T,E}$, and $\hat\rho = \rho$ an adequate ρ-estimate, we thus have the k_0-estimate

$$\hat{k}^*_{0|E} \equiv \hat{k}_{0|E}(n; m_1) := \min\left(n-1,\ \left\lfloor \frac{c_{\hat\rho}\ (\hat{k}^*_{0|T,E}(m_1))^2}{\hat{k}^*_{0|T,E}([m_1^2/n])} \right\rfloor + 1\right), \tag{3.6}$$

with

$$c_\rho = \begin{cases} (1-2^\rho)^{\frac{2}{1-2\rho}} & \text{if} \quad E = H, \\ (1-2^{2\rho})^{\frac{2}{1-4\rho}} & \text{if} \quad E = \overline{H}. \end{cases}$$

The adaptive estimate of ξ is then given by

$$E^* \equiv E^*_{n,m_1|T} := E_{\hat{k}^*_{0|E},n}.$$

3.2 Adaptive EVI-estimation

In the following **Algorithm** we include the Hill, the MVRB, the PORT-Hill and the PORT-MVRB EVI-estimators in the overall selection.

Algorithm—Adaptive bootstrap estimation of ξ

1. Consider a finite set $\mathcal{Q}$ with values in $[0, 1)$ and define $\mathcal{Q}_1 := \mathcal{Q} \cup \{1\}$. For example, if F has a finite left endpoint, we can select $\mathcal{Q} = \{0(0.05)0.95\}$. On the other hand, if F has a infinite left endpoint, we should not select values close to zero.

2. Given an observed sample $\underline{\mathbf{x}}_n = (x_1, \dots, x_n)$, execute the following Steps, for each $q \in \mathcal{Q}_1$:

2.1 Obtain the sample $\underline{\mathbf{x}}_n^{(q)}$, in (2.14). If $q = 1$, $\underline{\mathbf{x}}_n^{(1)} \equiv \underline{\mathbf{x}}_n$.

2.2 Compute, for the tuning parameter $\tau = 0$ the observed values of $\hat{\rho}^{(q)}(k) := \hat{\rho}_\tau(k; \underline{\mathbf{X}}_n^{(q)})$, with $\hat{\rho}_\tau(k)$ defined in (2.11).

2.3 Work with $\hat{\rho}^{(q)} \equiv \hat{\rho}_0^{(q)}(k_1)$ and $\hat{\beta}^{(q)} = \hat{\beta}^{(q)}_{\hat{\rho}^{(q)}}(k_1)$, with $\hat{\beta}_{\hat{\rho}}(k)$ and k_1 given in (2.12) and (2.13), respectively.

2.4 Compute $H^{(q)}_{k,n}$, in (2.15), and $\overline{H}^{(q)}_{k,n}$, in (2.16), for $k = 1, 2, \dots$.

2.5 Consider sub-samples of size $m_1 = o(n)$ and $m_2 = \lfloor m_1^2/n \rfloor + 1$.

2.6 For l from 1 until B, independently generate from the observed EDF, $F_n^*(x)$, associated with the observed sample $(x_1, \dots, x_n)$, B bootstrap samples

$$(x_1^*, \dots, x_{m_2}^*) \quad \text{and} \quad (x_1^*, \dots, x_{m_2}^*, x_{m_2+1}^*, \dots, x_{m_1}^*),$$

with sizes m_2 and m_1, respectively.

2.7 Generally denoting by $E_{k,n}$ any of the estimators under study, let us denote by $T^*_{k,n} \equiv T^*_{k,n|E}$ the bootstrap counterpart of the auxiliary statistic in (3.2), obtain

$$t^*_{k,m_1,l|E}, \ 1 < k < m_1, \quad t^*_{k,m_2,l|E}, \ 1 < k < m_2, 1 \le l \le B,$$

the observed values of the statistics T^*_{k,m_i}, $i=1,2$, compute

$$\mathrm{MSE}^*_E(m_i,k) = \frac{1}{B}\sum_{l=1}^{B}\left(t^*_{k,m_i,l|E}\right)^2, \quad k=2,\ldots,m_i-1, \quad i=1,2,$$

and obtain $\hat{k}^*_{0|T,E}(m_i) := \arg\min_{1<k\leq m_i-1} \mathrm{MSE}^*_E(m_i,k)$, $i=1,2$.

2.8 Compute $\hat{k}^*_{0|E}$ in (3.6).

2.9 Obtain $H^{*(q)} := H^{(q)}_{\hat{k}^*_{0|H^{(q)}},n}$ and $\overline{H}^{*(q)} := \overline{H}^{(q)}_{\hat{k}^*_{0|\overline{H}^{(q)}},n}$.

3. With $B^*_E(m_i,k) = \frac{1}{B}\sum_{l=1}^{B} t^*_{k,m_i,l|E}$, $\;k=2,\ldots,m_i-1$, $\;i=1,2$, compute for $k=\hat{k}^*_{0|H^{(q)}}$ and all values $q\in\mathcal{Q}_1$,

$$\widehat{\mathrm{RMSE}}_{\mathrm{H}}(k;q) := \sqrt{\frac{\left(H^{*(q)}\right)^2}{k} + \left(\frac{\left(B^*_{H^{(q)}}(m_1,k)\right)^2}{\left(2^{\hat{\rho}^{(q)}}-1\right)B^*_{H^{(q)}}(m_2,k)}\right)^2},$$

as well as

$$\widehat{\mathrm{RMSE}}_{\overline{\mathrm{H}}}(k;q) := \sqrt{\frac{\left(\overline{H}^{*(q)}\right)^2}{k} + \left(\frac{\left(B^*_{\overline{H}^{(q)}}(m_1,k)\right)^2}{\left(2^{2\hat{\rho}^{(q)}}-1\right)B^*_{\overline{H}^{(q)}}(m_2,k)}\right)^2}.$$

4. Compute $\hat{q}_{H} := \arg\min_q \widehat{\mathrm{RMSE}}_{H}(\hat{k}^*_{0|H^{(q)}};q)$ and $\hat{q}_{\overline{H}} := \arg\min_q \widehat{\mathrm{RMSE}}_{\overline{H}}(\hat{k}^*_{0|\overline{H}^{(q)}};q)$.

5. Obtain the adaptive EVI-estimates,

$$H^{**} \equiv H^{**}|\hat{q}_H \equiv H^{*(\hat{q}_H)}_{n,m_1} := H^{(\hat{q}_H)}_{\hat{k}_0^{(\hat{q}_H)},n} \quad \text{and} \quad \overline{H}^{**} \equiv \overline{H}^{**}|\hat{q}_{\overline{H}} \equiv \overline{H}^{*(\hat{q}_{\overline{H}})}_{n,m_1} := \overline{H}^{(\hat{q}_{\overline{H}})}_{\hat{k}_0^{(\hat{q}_{\overline{H}})},n}.$$

Remark 3.1. *Instead of* **Steps 2.2** *and* **2.3**, *we could have considered* **Steps 2.2'** *and* **2.3'** *below, reproduced from the algorithm provided in* Gomes and Pestana (2007b) *for the estimation of the second-order parameters β and ρ.*

2.2' *Compute, for the tuning parameters $\tau=0$ and $\tau=1$ the observed values of $\hat{\rho}^{(q)}(k) := \hat{\rho}_\tau(k;\underline{\mathbf{X}}_n^{(q)})$, with $\hat{\rho}_\tau(k)$ defined in* (2.11). *Consider $\left\{\hat{\rho}^{(q)}(k)\right\}_{k\in\mathcal{K}}$, with $\mathcal{K} = \left(\lfloor n_q^{0.995}\rfloor, \lfloor n_q^{0.999}\rfloor\right)$, compute their median, denoted χ_τ, and further compute $I_\tau := \sum_{k\in\mathcal{K}}\left(\hat{\rho}^{(q)}(k)-\chi_\tau\right)^2$, $\tau=0,1$. Next choose the tuning parameter $\tau^*=0$ if $I_0\leq I_1$; otherwise, choose $\tau^*=1$.*

2.3' *Work with $\hat{\rho}^{(q)} \equiv \hat{\rho}^{(q)}_{\tau^*}(k_1)$ and $\hat{\beta}^{(q)} = \hat{\beta}^{(q)}_{\hat{\rho}^{(q)}}(k_1)$, with $\hat{\beta}_{\hat{\rho}}(k)$ and k_1 given in* (2.12) *and* (2.13), *respectively.*

Remark 3.2. *If there are negative elements in any of the samples in the* **Algorithm***, the sample size must be replaced by the number of positive elements in the sample.*

Remark 3.3. *An analogue procedure can be used for any other parameter of rare events.*

Remark 3.4. *A few practical questions may be raised under the set-up developed: How does the asymptotic method work for moderate sample sizes? What is the type of the sample path of the new estimator for different values of* m_1*? What is the dependence of the method on the choice of* m_1*? What is the sensitivity of the method with respect to the choice of the* ρ*–estimator? Although aware of the theoretical need of* $m_1 = o(n)$*, what happens if we choose* $m_1 = n$*? Answers to these questions were given in* Gomes and Oliveira (2001) *for the estimation of* ξ *through the* Hill *estimator, and can be addressed here. Quite often, the method is only moderately dependent on the choice of the nuisance parameter* m_1*, in* Step 5. *of the* **Algorithm***, particularly for the* MVRB EVI-*estimators. This enhances the practical value of the method. Moreover, although aware of the need of* $m_1 = o(n)$*, it seems that we get good results up till* n*, again particularly for the* MVRB EVI-*estimator,* $\overline{H}_{k,n}$*, in* (2.9)*. To detect the sensitivity of the algorithm to changes of* m_1*, we have run it for* $q = 1$ *and values of* $m_1 = \lfloor n^b \rfloor$*,* $b = 0.950(0.005)0.995$*, different values of* n *and different models. In* Figure 2*, as an illustration, we present for a Fréchet underlying parent, from a* CDF $F(x) = \exp(-x^{-1/\xi})$*,* $x \geq 0$*, with* $\xi = 0.25$*, the bootstrap* ξ*-estimates* H^* *and* $\overline{H}^*$ *as a function of* b*, for* $n = 100$ *and* $n = 1000$*.*

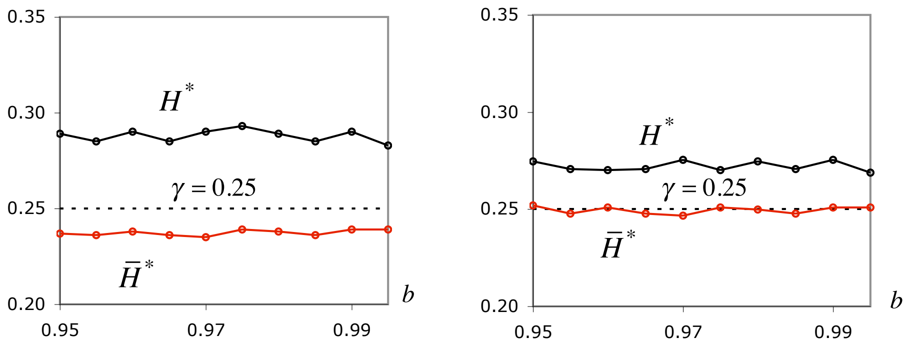

Figure 2: Bootstrap adaptive EVI-estimates, H^* and $\overline{H}^*$, as a function of b, in $m_1 = \lfloor n^b \rfloor$, for $n = 100$ (*left*) and $n = 1000$ (*right*)

A few comments on the results:

- *As expected, and due to the fact that the method works asymptotically, there is a general improvement in the estimation as the sample size,* n*, increases.*
- *The sensitivity of the* **Algorithm** *in* Section 3.2 to the nuisance parameter m_1 is quite weak for both H and $\overline{H}$, particularly if n is large.

Remark 3.5. *Note that bootstrap confidence intervals associated with the adaptive* EVI-*estimates are easily computed on the basis of the replication of the* **Algorithm** R *times, for an adequate* R. *The value of* B *can also be adequately chosen.*

Remark 3.6. *We would like to stress again that the use of the random sample of size* m_2, $(x_1^*, \ldots, x_{m_2}^*)$, *and the extended sample of size* m_1, $(x_1^*, \ldots, x_{m_2}^*, x_{m_2+1}^*, \ldots, x_{m_1}^*)$, *leads to a higher precision of the result with a smaller* B. *Indeed, if we had generated the sample of size* m_1 *independently of the sample of size* m_2, *just as done in* Draisma *et al.* (1999), *we would have got a wider confidence interval for the* EVI, *should we have kept the same value for* B. *This is quite similar to the use of the simulation technique of "common random numbers" in comparison algorithms, when we want to decrease the variance of a final answer to* $z = y_1 - y_2$, *inducing a positive dependence between* y_1 *and* y_2.

Remark 3.7. *An* **R** *package named* **evt0** *for the implementation of the algorithm above, among other methodologies in the field of* SUE, *is being developed by the authors, whose "in progress" version is already available in* Manjunath and Caeiro (2014).

4 Applications to simulated data

To enhance the importance of the PORT-Hill and PORT-MVRB EVI-estimation in the field of finance, we refer Gomes and Pestana (2007b) and Gomes *et al.* (2013), where respectively the MVRB and the PORT-MVRB EVI-estimation has been applied to log-returns associated with a few sets of financial data. Due to the specificity of such real data sets, and to the fact that log-returns have often been modelled by a Student-t or its skewed versions (see Jones and Faddy, 2003, among others), we have sequentially simulated three random samples of size $n = 1000$, from a Student's t_ν-model with $\nu = 4$ degrees of freedom ($\xi = 0.25$ and $\rho = -0.5$). Due to the specificity of the data (infinite left endpoint), we have considered for both the PORT-Hill and the PORT-MVRB EVI-estimation, q-values from 0.15 until 1, with step 0.05. When $q = 1$, we elect the Hill or the MVRB EVI-estimates. If $q < 1$, the PORT methodology is elected. We have further considered $m_1 = \lfloor n^b \rfloor$, with b from 0.950 until 0.995, with step 0.0025, and $B = 400$.

Figure 3 is related to the first Student-t_4 generated sample, and we there present the PORT-Hill/Hill and PORT-MVRB/MVRB EVI-estimates (*left*), the q-estimates (*center*) and the RMSE-estimates (*right*).

Figure 4 and Figure 5 are similar to Figure 3, but for the two other sequentially generated Student-t_4 samples.

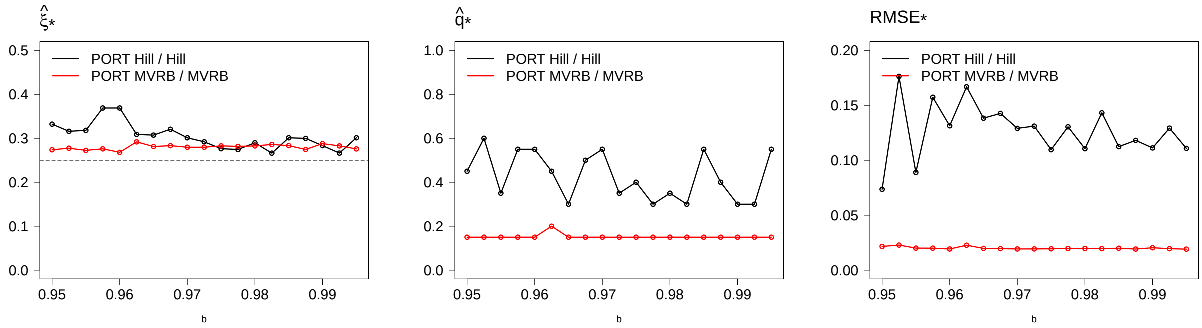

Figure 3: PORT-Hill/Hill and PORT-MVRB/MVRB adaptive EVI-estimates (*top*), the q-estimates (*center*) and the RMSE-estimates (*left*), for the first generated Student-t_4 sample

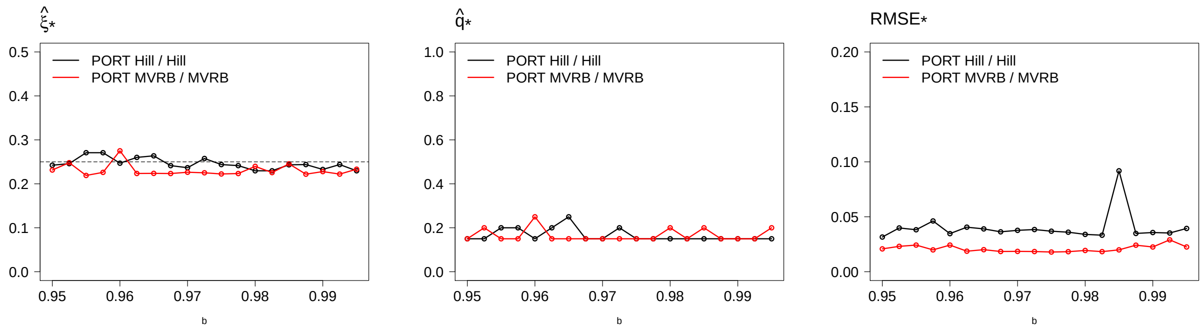

Figure 4: PORT-Hill/Hill and PORT-MVRB/MVRB adaptive EVI-estimates (*top*), the q-estimates (*center*) and the RMSE-estimates (*left*), for the second generated Student-t_4 sample

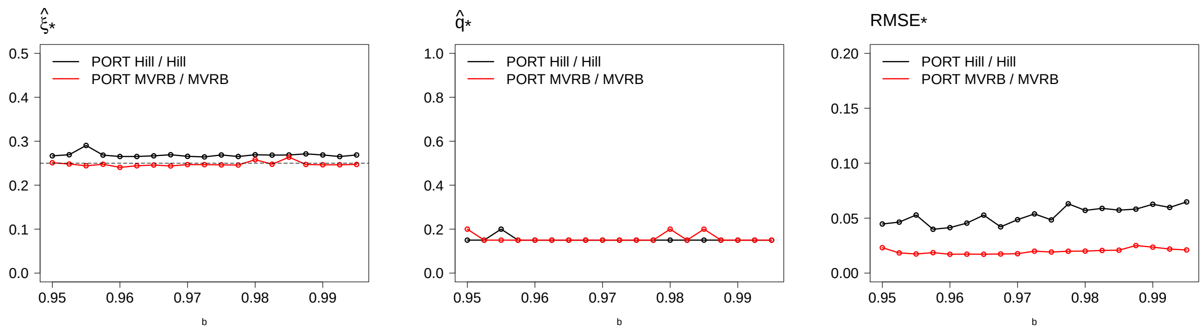

Figure 5: PORT-Hill/Hill and PORT-MVRB/MVRB adaptive EVI-estimates (*top*), the q-estimates (*center*) and the RMSE-estimates (*left*), for the third generated Student-t_4 sample

5 Concluding remarks

- For these simulated samples, we know the true value of ξ, the value 0.25 and we can easily assess the reliability of the estimates provided by the **Algorithm** in Section 3.2, immediately coming to the conclusion that, as expected, the PORT-MVRB methodology provides the more reliable EVI-estimation.

- It is clear that similarly to what usually happens with the Hill EVI-estimators, even the PORT-Hill EVI-estimation leads to an over-estimation of the EVI. The adaptive PORT-MVRB are closer to the target.
- Moreover, the RMSE-estimates associated to the adaptive PORT-MVRB EVI-estimates are always below the RMSE-estimates associated to the adaptive PORT-Hill, another point in favour of the PORT-MVRB methodology.
- These case studies claim obviously for a simulation study of the **Algorithm** and its application to real data sets. These are however topics out of the scope of this chapter.

References

[1] Araújo Santos, P., Fraga Alves, M.I. and Gomes, M.I. (2006). Peaks over random threshold methodology for tail index and quantile estimation. *Revstat* **4**:3, 227–247.

[2] Beirlant, J., Dierckx, G., Goegebeur, Y. and Matthys, G. (1999). Tail index estimation and an exponential regression model. *Extremes* **2**, 177–200.

[3] Beirlant, J., Caeiro, F. and Gomes, M.I. (2012). An overview and open research topics in the field of statistics of univariate extremes. *Revstat* **10**:1, 1–31.

[4] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation.* Cambridge Univ. Press.

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

[6] Caeiro, F. and Gomes, M.I. (2012a). A semi-parametric estimator of a shape second order parameter. In Pacheco, A., Oliveira, M.R., Santos, R. and Paulino, C.D. (eds.), *New Advances in Statistical Modeling and Application. Studies in Theoretical and Applied Statistics*, Selected Papers of the Statistical Societies, Springer-Verlag, Berlin and Heidelberg, in press.

[7] Caeiro, F. and Gomes, M.I. (2012b). *Bias Reduction in the Estimation of a Shape Second-Order parameter of a Heavy Right Tail Model.* Preprint, CMA 22-2012.

[8] Caeiro, F., Gomes, M.I. and Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator. *Revstat* **3**:2, 111–136.

[9] Caeiro, F., Gomes, M.I. and Henriques-Rodrigues, L. (2009). Reduced-bias tail index estimators under a third order framework. *Commun. in Statistics—Theory and Methods* **38**:7, 1019–1040.

[10] Ciuperca, G. and Mercadier, C. (2010). Semi-parametric estimation for heavy tailed distributions. *Extremes* **13**:1, 55–87.

[11] Davison, A. and Hinkley, D.V. (1997). *Bootstrap Methods and their Application*, Cambridge University Press.

[12] Draisma, G., de Haan, L., Peng, L. and Pereira, M.T. (1999). A bootstrap-based method to achieve optimality in estimating the extreme value index. *Extremes* **2**:4, 367–404.

[13] Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist.* **7**:1, 1–26.

[14] Efron, B. and Tibshirani, R.J. (1994). *An Introduction to the Bootstrap.* Boca Raton, FL: CRC Press.

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

[16] Fisher, R.A. and Tippett, L.H.C. (1928). Limiting forms of the frequency of the largest or smallest member of a sample. *Proc. Cambridge Phil. Soc.* **24**, 180–190.

[17] 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**:2, 194–213.

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

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

[20] Goegebeur, Y., Beirlant, J. and de Wet, T. (2008). Linking Pareto-tail kernel goodness-of-fit statistics with tail index at optimal threshold and second order estimation. *Revstat* **6**:1, 5169.

[21] Goegebeur, Y., Beirlant, J. and de Wet, T. (2010). Kernel estimators for the second order parameter in extreme value statistics. *J. Statist. Plann. Inference* **140**:9, 2632–2652.

[22] 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.

[23] Gomes, M.I. and Oliveira, O. (2001). The bootstrap methodology in Statistical Extremes — choice of the optimal sample fraction. *Extremes* **4**:4, 331–358.

[24] Gomes, M.I. and Pestana, D. (2007a). A simple second order reduced-bias' tail index estimator. *J. Statist. Comp. and Simul.* **77**:6, 487–504.

[25] Gomes, M.I. and Pestana, D. (2007b). A sturdy reduced-bias extreme quantile (VaR) estimator. *J. American Statistical Association* **102**:477, 280–292.

[26] Gomes, M.I., Martins, M.J. and Neves, M.M. (2000). Alternatives to a semi-parametric estimator of parameters of rare events—the Jackknife methodology. *Extremes* **3**:3, 207–229.

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

[28] Gomes, M.I., Reiss, R.-D. and Thomas, M. (2007b). Reduced-bias estimation. In Reiss, R.-D. & Thomas, M., *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*, 3rd Ed., Chapter 6, 189–204, Birkhäuser Verlag, Basel-Boston-Berlin.

[29] Gomes, M.I., Canto e Castro, L., Fraga Alves, M.I. and Pestana, D. (2008a). Statistics of extremes for iid data and breakthroughs in the estimation of the extreme value index: Laurens de Haan leading contributions. *Extremes* **11**:1, 3–34.

[30] Gomes, M.I., Fraga Alves, M.I. and Araújo Santos, P. (2008b). PORT Hill and moment estimators for heavy-tailed models. *Commun. in Statist.—Simul. and Comput.* **37**, 1281–1306.

[31] Gomes, M.I., de Haan, L. and Henriques-Rodrigues, L. (2008c). Tail index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses. *J. Royal Statistical Society B* **70**:1, 31–52.

[32] Gomes, M.I., Pestana, D. and Caeiro, F. (2008d). A note on the asymptotic variance at optimal levels of a bias-corrected Hill estimator. *Statist. and Probab. Letters* **79**, 295–303.

[33] Gomes, M.I., Henriques-Rodrigues, L., Pereira, H. and Pestana, D. (2010). Tail index and second order parameters' semi-parametric estimation based on the log-excesses. *J. Statist. Comput. and Simul.* **80**:6, 653–666.

[34] Gomes, M.I., Henriques-Rodrigues, L. and Miranda, C. (2011a). Reduced-bias location-invariant extreme value index estimation: a simulation study. *Commun. in Statist.—Simul. and Comput.* **40**:3, 424–447.

[35] Gomes, M.I., Mendonça, S. and Pestana, D. (2011b). Adaptive reduced-bias tail index and VaR estimation via the bootstrap methodology. *Commun. in Statistics—Theory and Methods* **40**:16, 2946–2968.

[36] Gomes, M.I., Figueiredo, F. and Neves, M.M. (2012). Adaptive estimation of heavy right tails: resampling-based methods in action. *Extremes* **15**, 463–489.

[37] Gomes, M.I., Henriques-Rodrigues, L., Fraga Alves, M.I. and Manjunath, B.G. (2013). Adaptive PORT-MVRB estimation: an empirical comparison of two heuristic algorithms. *J. Statist. Comput. and Simul.* **83**:6, 1129–1144.

[38] Gomes, M.I., Henriques-Rodrigues, L. and Figueiredo, F. (2014). *Resampling-Based Methodologies in Statistics of Extremes: Environmental and Financial Applications.* Notas e Comunicações CEAUL 01/2014.

[39] Gumbel, E.J. (1958). *Statistics of Extremes.* Columbia Univ. Press, New York.

[40] de Haan, L. (1984). Slow variation and characterization of domains of attraction. In Tiago de Oliveira, ed., *Statistical Extremes and Applications*, D. Reidel, Dordrecht, 31–48.

[41] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: an Introduction.* Springer Science+Business Media, LLC, New York.

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

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

[44] Jones, M.C. and Faddy M.J. (2003). A skew extension of the t-distribution, with applications. *J. Royal Statistical Society B* **65**:1, 159–174.

[45] Manjunath, B.G. and Caeiro, F. (2014). **evt0**: *Mean of order p, peaks over random threshold Hill and high quantile estimates*, R package version 1.2., `URLhttp://CRAN.R-project.org/package=evt0`

[46] Peng, L. (1998). Asymptotically unbiased estimator for the extreme-value index. *Statist. and Probab. Letters* **38**(2), 107–115.

[47] 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.

[48] Reiss, R.-D. and Thomas, M. (2007). *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*, 3rd edition, Birkhuser Verlag, Basel-Boston-Berlin, 2007.

[49] Resnick, S. (1997). Heavy tail modelling and teletraffic data. *Ann. Statist.* **25**:5, 1805–1869.

[50] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations, *J. American Statististical Association*, **73**, 812–815.