ADAPTIVE PROBABILITY WEIGHTED MOMENTS ESTIMATION

M. Ivette Gomes, Frederico Caeiro
CEAUL and DEIO (FCUL), Universidade de Lisboa
Lisbon, Portugal
e-mail: ivette.gomes@fc.ul.pt

Abstract

In this paper, we make use of *probability weighted moments* of largest observations, in order to build classes of estimators of the *extreme value index*, the primary parameter in *statistics of extremes*. Due to the specificity of the estimators, we propose the use of bootstrap computer intensive methods for an adaptive choice of the optimal number of order statistics to be used in the estimation. The developed methodology is applied to a data set in the field of insurance.

1 Introduction and preliminaries

The *extreme value index* (EVI) is the parameter $\gamma \in \mathbb{R}$ in the general *extreme value* (EV) distribution function (d.f.), $EV_\gamma(x) := \exp(-(1+\gamma x)^{-1/\gamma})$, $1+\gamma x > 0$. Let $\underline{X}_n = (X_1, \ldots, X_n)$ denote a sample of size n from either independent, identically distributed or even weakly dependent random variables and consider the associated sample of ascending order statistics (o.s.'s) $(X_{1:n} \le \cdots \le X_{n:n})$. The EV d.f. appears as the limiting d.f. of the maximum $X_{n:n}$, suitably linearly normalized, whenever a non-degenerate limit exists. We then say that F is in the *domain of attraction* for maxima of the EV d.f., and use the notation $F \in \mathcal{D}_{\mathcal{M}}(EV_\gamma)$. We shall deal with heavy-tails, i.e. a positive EVI. Then the right-tail function is of *regular variation* with an index $-1/\gamma$, and with the notations $U(t) := \inf\{x : F(x) \ge 1 - 1/t\}$, $t \ge 1$, and RV_α standing for the class of regularly varying functions at infinity with an index of regular variation α, $F \in \mathcal{D}_{\mathcal{M}}(EV_\gamma)_{\gamma>0} \iff \overline{F} := 1 - F \in RV_{-1/\gamma} \iff U \in RV_\gamma$.

One of the first classes of semi-parametric estimators of a positive EVI, considered in [7], is given by

$$\hat{\gamma}^H_{k,n} := \frac{1}{k}\sum_{i=1}^{k} \{\ln X_{n-i+1:n} - \ln X_{n-k:n}\}, \quad k = 1, 2, \ldots, n-1. \tag{1}$$

We shall also deal with the *Pareto probability weighted moments* (PPWM) EVI-estimators, recently introduced in [1]. They are valid for heavy right-tails, compare favourably with the Hill estimator, in (1), and are given by

$$\widehat{\gamma}^{PPWM}_{k,n} := 1 - \hat{a}_1(k)/\big(\hat{a}_0(k) - \hat{a}_1(k)\big), \tag{2}$$

with $\hat{a}_0(k) := \frac{1}{k}\sum_{i=1}^{k} X_{n-i+1:n}$ and $\hat{a}_1(k) := \frac{1}{k}\sum_{i=1}^{k} \frac{i}{k} X_{n-i+1:n}$. Consistency of the EVI-estimators in (1) and (2) is achieved in the whole $\mathcal{D}_{\mathcal{M}}(EV_\gamma)_{\gamma\ge 0}$ provided that $X_{n-k:n}$ is an *intermediate* o.s., i.e., if $k = k_n \to \infty$ and $k/n \to 0$, as $n \to \infty$.

In order to derive the asymptotic normality of these EVI-estimators, it is often assumed the validity of a second-order condition either on $\overline{F}$ or on U, like $\lim_{t\to\infty}\big(\ln U(tx)-\ln U(t)-\gamma\ln x\big)/A(t)=(x^\rho-1)/\rho$, where $\rho\le 0$ is a second-order parameter and $|A|\in RV_\rho$ ([3]). If we assume the validity of such a second-order framework, these EVI-estimators are asymptotically normal, provided that $\sqrt{k}A(n/k)\to\lambda_{_A}$, finite, as $n\to\infty$. Indeed, if we denote $\hat\gamma^\bullet_{k,n}$, either the Hill or the PPWM estimator, we have, with $Z^\bullet_k$ asymptotically standard normal and for adequate $(b_\bullet,\ \sigma_\bullet)\in(\mathbb{R},\ \mathbb{R}^+)$, the validity of the asymptotic distributional representation

$$\hat\gamma^\bullet_{k,n}\ \stackrel{d}{=}\ \gamma+\sigma_\bullet Z^\bullet_k/\sqrt{k}+b_\bullet\ A(n/k)(1+o_p(1)),\quad \text{as}\quad n\to\infty. \tag{3}$$

In this article, after a review, in Section 2, of the role of the bootstrap methodology in the estimation of optimal sample fractions, we provide an algorithm for the adaptive estimation through the PPWM EVI-estimators, also valid for Hill estimators. In Section 3, we apply such a data-driven estimation to a data set in the field of insurance.

2 The bootstrap methodology and optimal levels

Under the above mentioned second-order framework, but with $\rho<0$, let us use the parameterization $A(t)=\gamma\beta t^\rho$, where β and ρ are generalized scale and shape second-order parameters. Given the EVI-estimator, $\hat\gamma^\bullet_{k,n}$, let us denote $k_0^{\hat\gamma^\bullet}(n):=\arg\min_k \mathrm{MSE}(\hat\gamma^\bullet_{k,n})$, with MSE standing for *mean squared error*. With $\mathbb{E}$ denoting the mean value operator and AMSE standing for *asymptotic mean squared error*, a possible substitute for $\mathrm{MSE}(\hat\gamma^\bullet_{k,n})$ is $\mathrm{AMSE}(\hat\gamma^\bullet_{k,n}):=\mathbb{E}\big(\sigma_\bullet\ \overline{Z}_k/\sqrt{k}+b_\bullet A(n/k)\big)^2=\sigma^2_\bullet/k+b^2_\bullet\ \gamma^2\ \beta^2\ (n/k)^{2\rho}$, cf. equation (3). Then, with the notation $k_{0|\hat\gamma^\bullet}(n):=\arg\min_k \mathrm{AMSE}\big(\hat\gamma^\bullet_{k,n}\big)$, we get

$$k_{0|\hat\gamma^\bullet}(n)=\big((-2\rho)\ b^2_\bullet\ \gamma^2\beta^2\ n^{2\rho}/\sigma^2_\bullet\big)^{-1/(1-2\rho)}\ =k_0^{\hat\gamma^\bullet}(n)(1+o(1)). \tag{4}$$

For the Hill estimator, we have, in (3), $(b_{_H},\sigma_{_H})=(1/(1-\rho),\gamma)$. Consequently, with $(\hat\beta,\hat\rho)$ a consistent estimator of (β,ρ) and $[x]$ denoting the integer part of x, (4) justifies asymptotically the estimator $\hat k_0^H:=[((1-\hat\rho)^2 n^{-2\hat\rho}/(-2\hat\rho\hat\beta^2))^{1/(1-2\hat\rho)}]$. The same does not happen with the PPWM EVI-estimators, due to the fact that $\sigma_{_{PPWM}}$ and $b_{_{PPWM}}$ depend both on γ. In this situation, it is sensible to use the bootstrap methodology for the adaptive PPWM EVI-estimation. Similarly to [6] and [5], let us consider the auxiliary statistic, $T^\bullet_{k,n}:=\hat\gamma^\bullet_{[k/2],n}-\hat\gamma^\bullet_{k,n},\ k=2,\ldots,n-1$. On the basis of results similar to the ones in [6], we get the asymptotic distributional representation, $T^\bullet_{k,n}=^d \sigma_\bullet\ Q^\bullet_k/\sqrt{k}+b_\bullet\ (2^\rho-1)\ A(n/k)+o_p(A(n/k))$, with $Q^\bullet_k$ asymptotically standard normal, and $(b_\bullet,\sigma_\bullet)$ given in (3). The AMSE of $T^\bullet_{k,n}$ is thus minimal at a level $k_{0|T^\bullet}(n)$ such that $\sqrt{k}\ A(n/k)\to\lambda'_{_A}\neq 0$, of the type of the one in (4), with $b_\bullet$ replaced by $b_\bullet(2^\rho-1)$. We consequently have $k_{0|\hat\gamma^\bullet}(n)=k_{0|T^\bullet}(n)\,(1-2^\rho)^{\frac{1}{1-2\rho}}\,(1+o(1))$. Given the random sample $\underline{X}_n$, consider for any $n_1=O(n^{1-\epsilon})$, $0<\epsilon<1$, the bootstrap sample $\underline{X}^*_{n_1}=(X^*_1,\ldots,X^*_{n_1})$, from $F^*_n(x)=\frac{1}{n}\sum_{i=1}^n I_{\{X_i\le x\}}$, the empirical d.f. associated with $\underline{X}_n$, and associate to that bootstrap sample the corresponding bootstrap auxiliary statistic, $T^*_{k_1,n_1}$. Then, with the obvious notation $k^*_{0|T}(n_1)=\arg\min_{k_1}\mathrm{AMSE}\big(T^*_{k_1,n_1}\big)$, with $\hat k^*_{0|T}$

denoting the sample counterpart of $k^*_{0|T}$, and taking into account (4), we can build the k_0-estimate,

$$\hat{k}^\bullet_{0*} \equiv \hat{k}^\bullet_{0*}(n;n_1) := \min\left(n-1,\ \left[\left(1-2^{\hat\rho}\right)^{\frac{1}{1-2\hat\rho}}\left(\hat{k}^*_{0|T}(n_1)\right)^2/\hat{k}^*_{0|T}([n_1^2/n]+1)\right]+1\right). \quad (5)$$

2.1 An algorithm for the adaptive EVI-estimation

Now, with $\hat\gamma^{PPWM}_{k,n}$ defined in (2), the algorithm is the following:

1. Given a sample $(x_1, x_2, \ldots, x_n)$, compute $\hat\gamma^{PPWM}_{k,n}$, $k = 1, 2, \ldots, n-1$, and plot, for *tuning* parameters $\tau = 0$ and $\tau = 1$, the observed values of $\hat\rho_\tau(k)$ introduced and studied in [2].
2. Consider $\{\hat\rho_\tau(k)\}_{k\in\mathcal{K}}$, with $\mathcal{K} = ([n^{0.995}],\ [n^{0.999}])$, compute their median, denoted η_τ, and compute $I_\tau := \sum_{k\in\mathcal{K}}\left(\hat\rho_\tau(k)-\eta_\tau\right)^2$, $\tau = 0, 1$. Next choose the *tuning parameter* $\tau^* = 0$ if $I_0 \le I_1$; otherwise, choose $\tau^* = 1$.
3. Work with the second-order parameters' estimates $\hat\rho \equiv \hat\rho_{\tau^*} = \hat\rho_{\tau^*}(k_1)$ and $\hat\beta \equiv \hat\beta_{\tau^*} := \hat\beta_{\hat\rho_{\tau^*}}(k_1)$, $k_1 = [n^{0.999}]$ and $\hat\beta_{\hat\rho}(k)$ given in [4].
4. Next, consider a sub-sample size $n_1 = o(n)$, and $n_2 = [n_1^2/n] + 1$.
5. For l from 1 until B, generate independently B bootstrap samples $(x_1^*, \ldots, x_{n_2}^*)$ and $(x_1^*, \ldots, x_{n_2}^*, x_{n_2+1}^*, \ldots, x_{n_1}^*)$, from the empirical d.f. $F_n^*(x) = \frac{1}{n}\sum_{i=1}^n I_{\{X_i\le x\}}$ associated with the observed sample $(x_1, \ldots, x_n)$.
6. Denoting $T^*_{k,n}$ the bootstrap counterpart of $T^{PPWM}_{k,n}$, obtain $(t^*_{k,n_1,l},\ t^*_{k,n_2,l})$, $1 \le l \le B$, the observed values of the statistic T^*_{k,n_i}, $i = 1, 2$. For $k = 2, \ldots, n_i - 1$, compute $\mathrm{MSE}^*(n_i, k) = \frac{1}{B}\sum_{l=1}^B\left(t^*_{k,n_i,l}\right)^2$, and obtain $\hat{k}^*_{0|T}(n_i) := \arg\min_{1<k<n_i}\mathrm{MSE}^*(n_i, k)$, $i = 1, 2$.
7. Compute the threshold estimate $\hat{k}_{0*} \equiv \hat{k}^{PPWM}_{0*}$, already defined in (5).
8. Obtain $PPWM^* \equiv \hat\gamma^{PPWM}_* \equiv \hat\gamma^{PPWM}_*(n;n_1) := \hat\gamma_{\hat{k}_{0*},n}$.

A similar procedure can be used for the bootstrap data-driven estimation through the Hill estimator, in (1). Note also that bootstrap confidence intervals are easily associated with the bootstrap EVI-estimates, through the replication of this algorithm r times.

3 A case study in the field of insurance

We shall next consider an illustration of the performance of the adaptive PPWM EVI-estimates under study, comparatively with the same methodology applied to the Hill EVI-estimates, again through the analysis of $n = 371$ automobile claim amounts exceeding 1,200,000 Euro over the period 1988-2001, gathered from several European insurance companies co-operating with the same re-insurer, Secura Belgian Re (see [5], and references therein). The algorithm in Section 2.1 led us to $\hat\rho_0 = -0.74$ and $\hat\beta_0 = 0.80$. For a sub-sample size $n_1 = [n^{0.955}] = 284$, and $B = 250$ bootstrap generations, we were led to $\hat{k}^{PPWM}_{0*} = 58$ and to $PPWM^* = 0.272$. This same algorithm applied to the Hill estimates leads us to $\hat{k}^H_{0*} = 52$ and to $H^* = 0.299$.

In Figure 1, as a function of the sub-sample size n_1, ranging from $n_1 = [n^{0.95}] = 275$ until $n_1 = [n^{0.9999}] = 370$, we picture, at the left, the estimates of the *optimal sample* *fraction* (OSF), $k_0^\bullet/n$, for the adaptive bootstrap estimation of γ through the Hill and the PPWM estimators, in (1) and (2), respectively. Associated bootstrap EVI-estimates are pictured at the right.

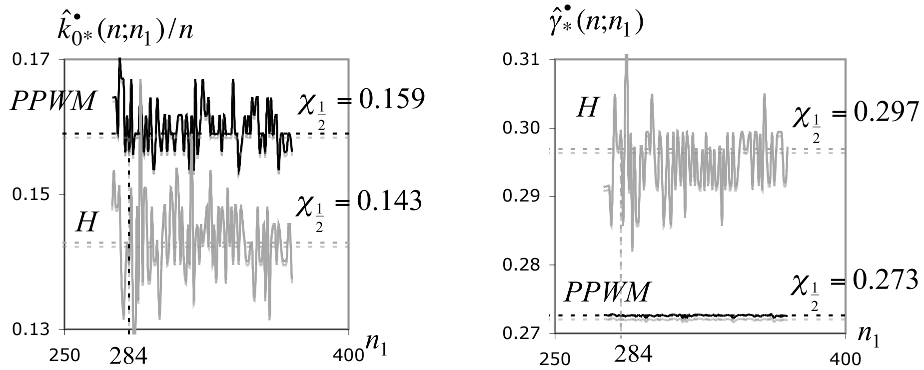

Figure 1: Estimates of the OSF's $\hat{k}_0^\bullet/n$ (*left*) and the bootstrap adaptive extreme value index estimates $\hat{\gamma}_*^\bullet$ (*right*), as functions of the sub-sample size n_1, for the SECURA data.

Contrarily to the bootstrap Hill, the bootstrap PPWM EVI-estimates are quite stable as a function of the sub-sample size n_1 (see Figure 1, *right*).

The running of the above mentioned algorithm $r = 100$ times, for $n_1 = [n^{0.955}]$, provided, for the PPWM estimates, a median value 0.2726, an average 0.2725, and a 95% bootstrap confidence interval for γ given by $(0.2715, 0.2728)$. The equivalent indicators for the bootstrap Hill estimates were 0.2969, 0.2949 and $(0.2826, 0.3133)$. The size of the confidence intervals are in favour of the PPWM estimation. As already detected in previous papers, and in the most diversified comparisons, the Hill estimates are clearly over-estimating the true value of the EVI, and should be used with care.

References

[1] Caeiro F., Gomes M.I. (2009). *Semi-Parametric Tail Inference through Probability-Weighted Moments*. UNL preprint. Submitted.

[2] Fraga Alves M.I., Gomes M.I., de Haan L. (2003). A new class of semi-parametric estimators of the second order parameter. *Portugaliae Mathematica* **60**:2, 194–213.

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

[4] Gomes M.I., 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.

[5] Gomes M.I., Mendonça S., Pestana D. (2009). Adaptive reduced-bias tail index and value-at-risk estimation. In Sakalauskas L., Skiadas C., Zavadskas E.K. (eds.) *Applied Stochastic Models and Data Analysis ASMDA* 2009, IMI & VGTU Editions, 41–44.

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

[7] Hill, B.M. (1975). A Simple General Approach to Inference about the Tail of a Distribution. *Ann. Statist.* **3**(5), 1163–1174.