On the bootstrap methodology for the estimation of the tail sample fraction

Frederico Caeiro[1] and M.Ivette Gomes[2]

[1] *CMA and FCT, Universidade Nova de Lisboa, PORTUGAL*
[2] *CEAUL and FCUL, Universidade de Lisboa, PORTUGAL*

Abstract: In statistics of extremes we are usually interested in the estimation of parameters of extreme events. Such estimation is usually based on the largest $k + 1$ order statistics or on the excesses over a high level u. In this paper, we consider the adaptive estimation of either k or u through the nonparametric bootstrap methodology. We shall introduce an improved version of Hall's bootstrap methodology and compare it with the double bootstrap methodology. The comparison of such methodologies is performed for simulated data sets.

Key Words: Bootstrap Methodology, Extreme Value Index, Heavy tail, Tail sample fraction.

This is an electronic reprint of the original article: Caeiro, F. and Gomes, M.I. (2014). On the bootstrap methodology for the estimation of the tail sample fraction. In Gilli, M., Gonzalez-Rodriguez, G. and Nieto-Reyes, A. (Eds.), Proceedings of COMPSTAT 2014: 21th International Conference on Computational Statistics, 545–552. This reprint differs from the original in typographic details.

1 Introduction

Let $\underline{X}_n = (X_1, \ldots, X_n)$ denote a sample of either independent, identically distributed (i.i.d.) or even weakly dependent random variables (r.v.'s) from an underlying distribution function F. We shall assume that we are in the max-domain of attraction of the Extreme Value distribution $EV_\xi(x) := \exp(-(1 + \xi x)^{-1/\xi})$, $1 + \xi x > 0$, where the shape parameter ξ is the well known extreme value index (EVI). We shall consider $\xi > 0$, i.e., models with a heavy right tail. Then, the quantile function $U(t) := F^{\leftarrow}(1 - 1/t) = \inf\{x : F(x) \geq 1 - 1/t\}$, $t > 1$, is a regularly varying function with a positive index of regular variation equal to ξ, i.e.,

$$\lim_{t\to\infty} \frac{U(tx)}{U(t)} = x^\xi \,. \tag{1}$$

For a heavy tailed model, the classic semi-parametric Hill estimator of ξ, introduced in [17], is

$$H(k) \equiv \hat{\xi}_n^H(k) := \frac{1}{k}\sum_{i=1}^{k} \left(\ln X_{n-i+1:n} - \ln X_{n-k:n}\right), \quad k = 1, 2, \ldots, n-1, \tag{2}$$

the average of the log excesses over the high threshold $X_{n-k:n}$, where $X_{i:n}$ denotes the i-th ascending order statistic of the sample of size n. Consistency is achieved for intermediate k, i.e. for sequences of integers $k = k_n,\ 1 \le k < n$, such that

$$k \to \infty \qquad \text{and} \qquad k/n \to 0, \quad \text{as} \quad n \to \infty. \tag{3}$$

To obtain the asymptotic distributional behaviour of the Hill and other semi-parametric EVI-estimators, we need to assume a second-order condition, that measures the rate of convergence in the first-order condition, i.e. the way $\ln U(tx) - \ln U(t)$ approaches $\xi \ln x$,

$$\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \xi \ln x}{A(t)} = \begin{cases} (x^\rho - 1)/\rho, & \text{if } \rho < 0, \\ \ln x, & \text{if } \rho = 0, \end{cases} \tag{4}$$

for every $x > 0$, where $\rho\ (\le 0)$ is a second-order parameter that rules the rate of convergence and $|A|$ is compulsory a regular varying function with index ρ. For technical simplicity, we shall assume $\rho < 0$. Under the second-order condition in (4) the Hill estimator has usually a high asymptotic bias and recently several authors have considered different ways of reducing the bias. A simple class of second-order minimum-variance reduced-bias (MVRB) EVI estimators is the one in [2], given by

$$CH(k) \equiv \hat{\xi}_n^{CH}(k) := \hat{\xi}_n^H(k)\left(1 - \frac{\hat{\beta}(n/k)^{\hat{\rho}}}{(1-\hat{\rho})}\right), \quad k = 1, 2, \ldots, n-1, \tag{5}$$

with $(\hat{\beta}, \hat{\rho})$ adequate estimators of the second-order parameters (β, ρ) such that $A(t) = \gamma\beta t^\rho$, $\rho < 0$. This estimator has an asymptotic variance equal to that of the Hill EVI-estimator, but an asymptotic bias of smaller order, and thus beats the classical estimators for all k. For a reliable estimation of the EVI, some attention should be given to the choice of the number k, or equivalently to the threshold $X_{n-k:n}$. Recent overviews of statistics of univariate extremes were recently published (see [1, 6, 12], among others).

In section 2 of this paper we present several known results that allow us to compute the theoretical optimal level of the EVI-estimators in (2) and (5). In Section 3, we discuss the estimation of the second-order parameters ρ and β. In Section 4 we shall use bootstrap computer-intensive resampling methods for the choice of k, not only for the use of $H(k)$, but also for the use of $CH(k)$. We introduce a new bootstrap method, based on Hall's methodology, and present the double bootstrap algorithm. Finally, we provide an application to simulated data sets.

2 Asymptotic Properties

If we assume the validity of the second-order framework in (4), $\hat{\xi}_n^H(k)$ is asymptotically normal, provided that $\sqrt{k}A(n/k) \to \lambda$, finite, as $n \to \infty$. Indeed, we have, with $\mathcal{N}_{\mu,\sigma^2}$ denoting a normal random variable with mean value μ and variance σ^2, and $b_1 = 1/(1-\rho)$,

$$\sqrt{k}\big(\hat{\xi}_n^H(k) - \xi\big) \stackrel{d}{=} \mathcal{N}_{0,\xi^2} + b_1\sqrt{k}A(n/k) + o_p\big(\sqrt{k}A(n/k)\big), \quad \text{as} \quad n \to \infty. \tag{6}$$

The bias $b_1\sqrt{k}A(n/k) = \xi\,\beta\,\sqrt{k}\,(n/k)^{\rho}/(1-\rho)$ can be very large, moderate or small, and increases as k increases. And since the variance decreases with k, we have usually a very sharp mean square error (MSE) pattern, as a function of k. Under the same conditions as before, $\sqrt{k}\big(\hat\xi_n^{CH}(k)-\xi\big)$ is asymptotically normal with variance also equal to ξ^2 but with a null mean value.

To obtain information on the bias of MVRB EVI-estimators it is common to slightly restrict the class of models in (4), further assuming a third-order condition, ruling now the rate of convergence in the second-order condition in (4). We shall consider the third-order condition used in [3], which guarantees that for all $x > 0$,

$$\lim_{t\to\infty} \frac{\frac{\ln U(tx)-\ln U(t)-\xi\ln x}{A(t)} - \frac{x^{\rho}-1}{\rho}}{B(t)} = \frac{x^{\rho+\rho'}-1}{\rho+\rho'}, \tag{7}$$

where $|B|$ is a regular varying function with index ρ'. Further details can be found in [11].

The full asymptotic behaviour of $\hat\xi_n^{CH}(k)$ is provided in the following theorem.

Theorem 1. *If under the validity of the second-order condition in* (4)*, 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 $\sqrt{k}\big(\hat\xi_n^{CH}(k)-\xi\big) \overset{d}{=} \mathcal{N}_{0,\xi^2}+o_p\big(\sqrt{k}A(n/k)\big)$ holds. Under the validity of equation* (7)*, we can guarantee*

$$\sqrt{k}\big(\hat\xi_n^{CH}(k)-\xi\big) \overset{d}{=} \mathcal{N}_{0,\xi^2} + b_2\sqrt{k}A^2(n/k)\,(1+o_p(1)), \tag{8}$$

for adequate k values such that $\sqrt{k}A^2(n/k)\to\lambda_A$, finite and $b_2 = (\omega/(1-2\rho)-(1-\rho)^{-2})/\xi$ with $\omega = B(n/k)/A(n/k)$.

Regarding the choice of k, an usual approach is to minimize the MSE of the EVI-estimator. With AMSE standing for 'asymptotic MSE', on the basis of (6) and (8), and with the notation $\hat\xi_n^{(1)} = \hat\xi_n^{H}$, $\hat\xi_n^{(2)} = \hat\xi_n^{CH}$, we get $\mathrm{AMSE}\big(\hat\xi_n^{(c)}(k)\big) = \xi^2/k + b_c^2\,A^{2c}(n/k),\ c=1,2$ and

$$k_0^{(c)}(n) := \arg\min_k \mathrm{AMSE}\big(\hat\xi_n^{(c)}(k)\big) = \left(\frac{n^{-2c\rho}}{(-2c\rho)b_c^2\,\xi^{2(1-c)}\,\beta^{2c}}\right)^{1/(1-2c\rho)}, \quad c=1,2. \tag{9}$$

3 Estimation of the second-order parameters

We have used particular members of the class of estimators of the second-order parameter ρ proposed in [10]. Such a class of estimators has been first parameterized by a tuning parameter $\tau\geq 0$, that can be straightforwardly considered as a real number, and is defined as

$$\hat\rho_\tau(k) := \min\left\{0\,,\,\frac{3(T_n^{(\tau)}(k)-1)}{T_n^{(\tau)}(k)-3}\right\}, \qquad T_n^{(\tau)}(k) := \frac{\big(M_n^{(1)}(k)\big)^{\tau} - \big(M_n^{(2)}(k)/2\big)^{\tau/2}}{\big(M_n^{(2)}(k)/2\big)^{\tau/2} - \big(M_n^{(3)}(k)/6\big)^{\tau/3}}, \quad \tau\in\mathbb{R},$$

with the notation $a^{b\tau} = b\ln a$ if $\tau = 0$ and where $M_n^{(j)}(k) := \frac{1}{k}\sum_{i=1}^{k}\{\ln X_{n-i+1:n}-\ln X_{n-k:n}\}^j,\quad j=1,2,3$. Interesting alternative ρ-estimators have recently been introduced in [5, 8]. Here we consider the same type of criterion used in [15] for the adaptive estimation of ρ: Consider a sample with n positive values, compute $\{\hat\rho_\tau(k)\}_{k\in\mathcal{K}}$, with $\mathcal{K} = (\lfloor n^{0.995}\rfloor, \lfloor n^{0.999}\rfloor)$, compute their median, denoted η_τ, and compute $I_\tau := \sum_{k\in\mathcal{K}} (\hat\rho_\tau(k) - \eta_\tau)^2$, $\tau = 0, 1$. Next choose $\tau^* = 0$ if $I_0 \le I_1$; otherwise, choose $\tau^* = 1$ and compute $\hat\rho \equiv \hat\rho_{\tau^*} = \hat\rho_{\tau^*}(k_1)$, with $k_1 = \lfloor n^{0.995}\rfloor$.

The estimate of the scale second-order parameter β is given by $\hat\beta = \hat\beta_{\hat\rho}(k_1)$ with $\hat\beta_{\hat\rho}(k)$ the estimator in [13], given by

$$\hat\beta_{\hat\rho}(k) := \left(\frac{k}{n}\right)^{\hat\rho} \frac{D_{\hat\rho,0}(k)\, D_{0,1}(k) - D_{\hat\rho,1}(k)}{D_{\hat\rho,0}(k)\, D_{\hat\rho,1}(k) - D_{2\hat\rho,1}(k)}, \qquad D_{\alpha_1,\alpha_2}(k) := \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\alpha_1} U_i^{\alpha_2},$$

with $U_i := i(\ln X_{n-i+1:n} - \ln X_{n-i:n})$ the the rescaled log-spacings and dependent on the estimator $\hat\rho$, suggested before.

4 The bootstrap methodology

4.1 A method based on Hall's single bootstrap

The bootstrap methodology for the selection of the threshold k was first introduced by Hall ([16]). To avoid the underestimation of the bias, it is necessary to use smaller resamples of size $n_1 = o(n)$, where n is the size of the initial sample. Let $\underline{X}^*_{n_1} = \{X_1, \ldots, X_{n_1}\}$ denote a resample of size $n_1 = o(n)$ taken with replacement. Hall's considered the minimization of the bootstrap estimate of the MSE of $\hat\xi^H_{n_1}(k)$,

$$MSE(n_1, k) = E\left[\left\{\hat\xi^H_{n_1}(k) - \hat\xi^H_{n_1}(k_{aux})\right\}^2 | \underline{X}^*_{n_1}\right] \tag{10}$$

where k_{aux} is an initial threshold such that $\hat\xi^H_{n_1}(k_{aux})$ is consistent for ξ. Next we choose the value $k^*_0(n_1)$ that minimizes (10). The bootstrap estimate of the tail fraction is then

$$k^*_0(n) = k^*_0(n_1)(n/n_1)^\alpha.$$

Hall suggested $\alpha = 2/3$, which is equivalent to say that our model is under the second-order condition with $\rho = -1$. Also, as noticed by [14], the method is very sensitive to the choice of k_{aux}. Here we shall consider again an auxiliary statistic of the type of the one considered in [14], directly related to the EVI-estimator under consideration, but going to the known value zero,

$$T^{(c)}_n(k) := \hat\xi^{(c)}_n(\lfloor k/2\rfloor) - \hat\xi^{(c)}_n(k), \quad k = 2, \ldots, n-1, \quad c = 1, 2. \tag{11}$$

Notice that if $c = 1$ this approach is equivalent to consider $k_{aux} = \lfloor k/2\rfloor$ in (10). On the basis of the results similar to the ones in [14], we can get for $T^{(c)}_n(k)$, in (11), the asymptotic distributional representation,

$$T^{(c)}_n(k) \overset{d}{=} \frac{\xi^2}{\sqrt{k}} Q_k + b_c(2^{c\rho} - 1)\, A(n/k)(1 + o_p(1)),$$

with Q_k asymptotically $\mathcal{N}_{0,1}$, and b_c, $c = 1, 2$ given in Section 2. Then, the AMSE of $T^{(c)}_n(k)$ is minimal at a level $k^{(c)}_{0|T}(n)$, such that

$$k^{(c)}_0(n) = k^{(c)}_{0|T}(n)(2^{c\rho} - 1)^{\frac{2}{1-2c\rho}}$$

Based on Hall's method we now introduce a new bootstrap algorithm:

Algorithm 1. *Let* $\hat{\xi}_n^{(c)}(k)$ *denote any of the EVI-estimators in* (2) $(c = 1)$ *or in* (5) $(c = 2)$. *We now proceed with the description of the algorithm for the adaptive estimation of the optimal threshold* $k_0^{(c)}(n)$ *and the adaptive estimation of* ξ.

Step 1 *Given a sample* $(x_1, x_2, \ldots, x_n)$, *compute the estimates* $\hat{\rho}$ *and* $\hat{\beta}$ *of the second-order parameters* ρ *and* β *as described in Section 3.*

Step 2 *Next, consider a sub-sample size* $n_1 = o(n)$. *For* l *from 1 until* B, *generate independently* B *bootstrap samples* $(x_1^*, x_2^*, \ldots, x_{n_1}^*)$ *of size* 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* $(x_1, x_2, \ldots, x_n)$.

Step 3 *Denoting* $T_{n_1}^{(c)*}(k)$ *the bootstrap counterpart of* $T_{n_1}^{(c)}(k)$, *defined in* (11), *obtain* $t_{n_1,l}^*(k)$, $1 \le l \le B$, *the observed values of* $T_{n_1}^{(c)*}(k)$. *For* $k = 2, \ldots, n_1 - 1$, *compute* $MSE^*(n_1, k) = \frac{1}{B}\sum_{l=1}^{B}\left(t_{n_1,l}^*(k)\right)^2$, *and obtain* $\hat{k}_{0|T}^*(n_1) := \arg\min_{1<k<n_1} MSE^*(n_1, k)$.

Step 4 *Compute the threshold estimate*

$$\hat{k}_0^*(n) \equiv \left\lfloor (1 - 2^{c\hat{\rho}})^{\frac{2}{1-2c\hat{\rho}}}\, \hat{k}_{0|T}^*(n_1)\, (n/n_1)^{\frac{-2c\hat{\rho}}{1-2c\hat{\rho}}} \right\rfloor + 1.$$

If $\hat{k}_0^*(n) > n-1$ *go back to* ***Step 2***, *being careful not to generate the same samples.*

Step 5 *Obtain* $\hat{\xi}^* \equiv \hat{\xi}_n^{(c)}(\hat{k}_0^*(n))$.

4.2 The double bootstrap method

The assumptions in Hall's methodology were overpassed with the use of a double bootstrap method. This method was first used in [9] for the general max-domain of attraction and in [7, 14] for heavy tailed models. More recently, [15] modified the double bootstrap algorithm for an adaptive choice of the thresholds for second-order corrected-bias estimators. The next algorithm follows closely the bootstrap method in [15].

Algorithm 2. *Let* $\hat{\xi}_n^{(c)}(k)$ *denote any of the EVI-estimators in* (2) $(c = 1)$ *or in* (5) $(c = 2)$. *We now proceed with the description of the algorithm for the adaptive estimation of the optimal threshold* $k_0^{(c)}(n)$ *and the adaptive estimation of* ξ.

Step 1 *Equal to* ***Step 1*** *in Algorithm 1.*

Step 2 *Next, consider a sub-sample size* $n_1 = o(n)$ *and* $n_2 = \lfloor n_1^2/n \rfloor + 1$. *For* l *from 1 until* B, *generate independently* B *bootstrap samples* $(x_1^*, \ldots, x_{n_2}^*)$ *and* $(x_1^*, \ldots, x_{n_2}^*, x_{n_2+1}^*, \ldots, x_{n_1}^*)$, *of sizes* n_2 *and* n_1, *respectively, 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* $(x_1, x_2, \ldots, x_n)$.

Step 3 *Denoting* $T_{n_i}^{(c)*}(k)$ *the bootstrap counterpart of* $T_{n_i}^{(c)}(k)$, *in* (11), *obtain* $t_{n_i,l}^*(k)$, $1 \le l \le B$, $i = 1, 2$ *the observed values of* $T_n^{(c)*}(k)$. *For* $k = 2, \ldots, n_i - 1$, *and* $i = 1, 2$ *compute* $MSE^*(n_i, k) = \frac{1}{B}\sum_{l=1}^{B}\left(t_{n_1,l}^*(k)\right)^2$, *and obtain* $\hat{k}_{0|T}^*(n_i) := \arg\min_{1<k<n_i} MSE^*(n_i, k)$.

Step 4 *Compute the threshold estimate*

$$\hat{k}_0^*(n) \equiv \left\lfloor (1 - 2^{c\hat{\rho}})^{\frac{2}{1-2c\hat{\rho}}} \, (\hat{k}_{0|T}^*(n_1))^2 / \hat{k}_{0|T}^*(n_2) \right\rfloor + 1.$$

If $\hat{k}_0^*(n) > n-1$ *go back to* ***Step 2****, being careful not to generate the same samples.*

Step 5 *Obtain* $\hat{\xi}^* \equiv \hat{\xi}_n^{(c)}(\hat{k}_0^*(n))$.

Remarks:

- The use of the sample $(x_1^*, x_2^*, \ldots, x_{n_2}^*)$, and of the extended sample $(x_1^*, \ldots, x_{n_2}^*, \ldots, x_{n_1}^*)$, $n_2 < n_1$, lead us to an increased precision of the result with the same number B of bootstrap samples generated in **Step 2**. This is quite similar to the use of the simulation technique of "Common Random Numbers" in comparison problems.
- Bootstrap confidence intervals are easily obtained, through the replication of this algorithm r times. The replication can also provide us more precise estimates, if we consider the estimate given by the mean or the median of the r bootstrap estimates.
- A few practical questions may be raised under the set-up developed: How does the asymptotic method work for moderate sample sizes? Is the method strongly dependent on the choice of n_1? Although aware of the theoretical need to have $n_1 = o(n)$, what happens if we choose $n_1 = n$? We will try to answer those questions in the next section.

4.3 Applications to Simulated Data Sets

Here we shall present an illustration of the performance of the algorithms to simulated samples, cases where we know the value of ξ. We have simulated one random sample of size $n = 500$, from a Burr model with d.f. $F(x) = 1 - (1 + x^{-\rho/\xi})^{1/\rho}$, $x > 0$, $\xi > 0$, $\rho < 0$ with $\xi = 0.25$ and $\rho = -0.75$ and one Student's-t_4 sample of size $n = 1000$ ($\xi = 0.25$, $\rho = -0.5$).

Conclusions: Bootstrap estimates of the optimal sample fractions, $\hat{k}_0^*(n)/n$ and of the EVI, $\hat{\xi}^*$, as functions of n_1, for $\lfloor n^{0.85} \rfloor \le n_1 \le n$, are pictured in Figs. 1-2. Since we know the true value of ξ, and we can easily assess the reliability of the estimates provided by the Algorithms, immediately coming to the conclusion that Algorithm 2 provides a quite reliable EVI-estimation, even with $n_1 = n$. Algorithm 2 can be very sensitive to the choice of n_1 (see Fig. 1). We noticed that we can have some volatility in the estimates as function of n_1 and such volatility only decreases substantially with the replication of the algorithm $r = 25$ times. For the Burr sample, Algorithm 4.1 is sensitive to the choice of n_1. These results claim obviously for a simulation study of the Algorithms and its application to real data sets. These are however topics that can only be covered in a full-length paper.

Acknowledgement

Research partially supported by FCT - Fundação para a Ciência e a Tecnologia, projects PEst-OE/MAT/UI0006/2014 (CEAUL), PEst-OE/MAT/UI0297/2014 (CMA/UNL).

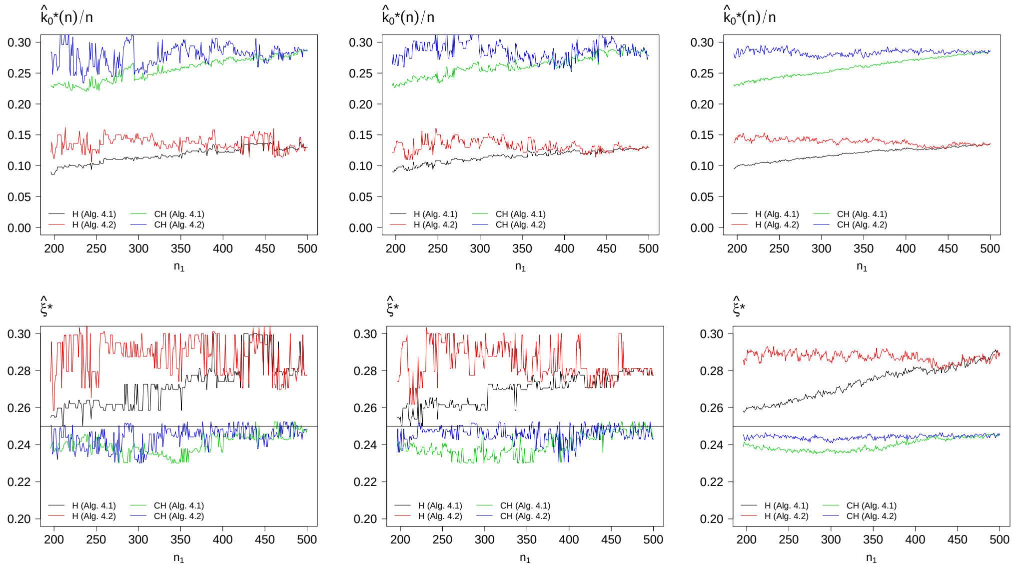

Figure 1: Adaptive estimates of $\hat{k}_0^*(n)/n$ (above) and $\hat{\xi}^*$ (below), as function of n_1, with $B = 250$ (*left*), $B = 1000$ (*center*) and mean of $r \times B = 25 \times 250$ (*right*), for the Burr simulated sample.

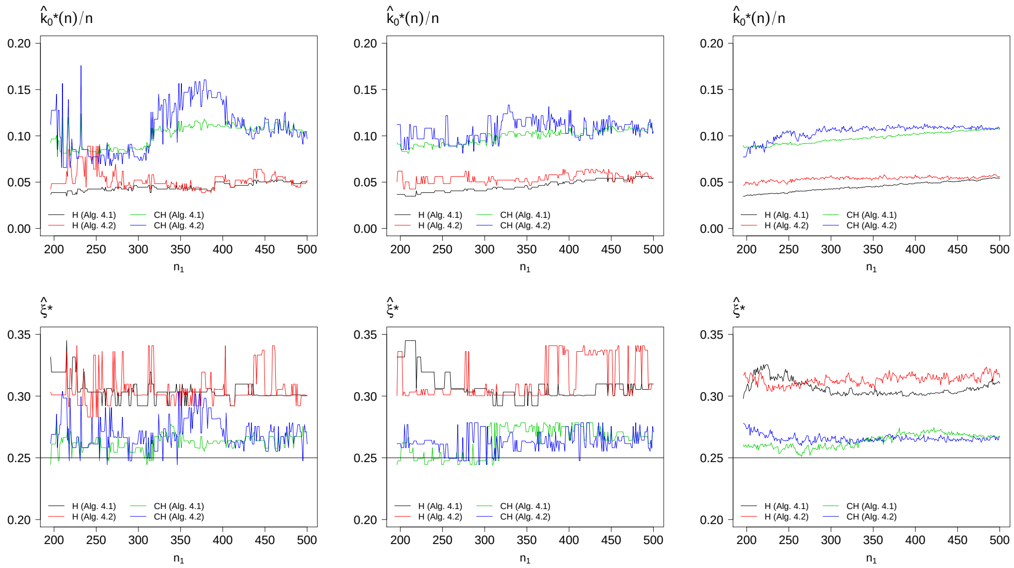

Figure 2: Adaptive estimates of $\hat{k}_0^*(n)/n$ (above) and $\hat{\xi}^*$ (below), as function of n_1, with $B = 250$ (*left*), $B = 1000$ (*center*) and mean of $r \times B = 25 \times 250$ (*right*), for Student's-t sample.

References

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

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

[3] Caeiro, F. and Gomes, M.I. (2011) *Asymptotic comparison at optimal levels of reduced-bias extreme value index estimators*. Statistica Neerlandica, **65**:4, 462–

488.

[4] Caeiro, F. and Gomes, M.I. (2013) *Asymptotic comparison at optimal levels of minimum-variance reduced-bias tail index estimators*. In Lita da Silva, J. *et al.* (Eds.), Advances in Regression, Survival Analysis, Extreme Values, Markov Processes and Other Statistical Applications, Studies in Theoretical and Applied Statistics, 83–91, Springer Berlin Heidelberg.

[5] Caeiro, F. and Gomes, M.I. (2014) *A semi-parametric estimator of a shape second order parameter*. In Pacheco, A., Santos, R., Rosário Oliveira, M. and Paulino, C.D. (Eds.), New Advances in Statistical Modeling and Applications, 137–144, Springer.

[6] Charras-Garrido, M. and Lezaud, P. (2013) *Extreme Value Analysis: an Introduction*, Journal de la Société Française de Statistique, **154**:2, 66–97.

[7] Danielsson, J., Haan, L. de, Peng, L., and de Vries, C.G. (2001) *Using a bootstrap method to choose the sample fraction in the tail index estimation*. J. Multivariate Anal., **76**, 226–248.

[8] Deme, E.H., Gardes, L. and Girard, S. (2013) *On the estimation of the second order parameter for heavy-tailed distributions*. Revstat **11**:3, 277–299.

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

[10] Fraga Alves, M.I., Gomes, M.I. and de Haan, L. (2003) *A new class of semiparametric estimators of the second order parameter*. *Portugaliae Math.*, **60**:2, 193–213.

[11] de Haan, L. and Ferreira, A. (2006) *Extreme Value Theory: An Introduction*. Springer, New York.

[12] Gomes, M.I. and Guillou, A. (2014) *Extreme Value Theory and Statistics of Univariate Extremes: A Review*. International Statistical Review, doi:10.1111/insr.12058

[13] 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.

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

[15] Gomes, M.I., Figueiredo, F. and Neves M.M. (2012) *Adaptive estimation of heavy right tails: resampling-based methods in action*. Extremes 2012, **15**:4, 463–489.

[16] Hall, P. (1990) *Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems*. J. Multivariate Anal., **32**, 177–203.

[17] Hill, B.M. (1975) *A simple general approach to inference about the tail of a distribution*. Ann. Statist. **3**:5, 1163–1174.