Resampling Methodologies and Reliable Tail Estimation

M. Ivette Gomes

Universidade de Lisboa, Portugal

Research partially supported by National Funds through **FCT** — Fundação para a Ciência e a Tecnologia, projects PEst-OE/MAT/UI0006/2011 and EXTREMA, PTDC/MAT/101736/2008.

To ROSS, as a first token of friendship: a photo of *Liseberg Amusement Park*—Gothenburg, Sweden, we both visited in 2005.

Vimeiro — 1983

Oberwolfach — 1987

But, let's go into Science . . .

We are both "**extremists**", and my main scientific connections with ROSS are related to:

- **Dependence** conditions. My 1st Ph.D. student had Ross has a mentor, and has even stayed for a while in North Carolina, in the late eighties. And I had several Ph.D. students who worked and still work deeply in this area . . . where Ross is a King.
- **Extremal index**. When I was a Ph.D. student at Sheffield, I have enthusiastically read Ross' 1973 paper entitled "On extreme values in stationary sequences" [*ZWT*], and I got very much interested in the theme, despite of having worked only sporadically in it.

This was the main reason for the choice of the topic of this presentation, where I give some emphasis on the use of the jackknife in the estimation of the **extremal index**.

1. OUTLINE

- Resampling methodologies have recently revealed to be very fruitful in the field of *statistics of extremes*.
- We mention the importance of
  - the Generalized Jackknife and
  - the Bootstrap

  in the obtention of a reliable semi-parametric estimate of any parameter of extreme or even rare events, like a *high quantile*, the expected shortfall, the *return period* of a high level or the two primary parameters of extreme events, the *extreme value index* (EVI) and the *extremal index* (EI).
- In order to illustrate such topics, we shall consider minimum-variance reduced-bias (MVRB) estimators of a positive EVI and a jackknife Leadbetter-Nandagopalan EI-estimator.

2. EXTREME VALUE THEORY (EVT) – A BRIEF INTRODUCTION

2.1. The extreme value index (EVI)

- We use the notation γ for the EVI, the shape parameter in the *Extreme Value* d.f.,

$$\boxed{EV_\gamma(x) = \begin{cases} \exp(-(1+\gamma x)^{-1/\gamma}),\ 1+\gamma x > 0 & \text{if} \quad \gamma \neq 0 \\ \exp(-\exp(-x)),\ x \in \mathbb{R} & \text{if} \quad \gamma = 0, \end{cases}}$$

and we now consider models with a heavy right-tail, i.e.

$$\boxed{\overline{F} := 1 - F \in RV_{-1/\gamma}, \quad \text{for some} \quad \gamma > 0,}$$

where the notation RV_α stands for the class of regularly-varying functions with an index $\alpha \in \mathbb{R}$, i.e., positive measurable functions $g(\cdot)$ such that $\forall x > 0$, $g(tx)/g(t) \to x^\alpha$, as $t \to \infty$.

2.2. The extremal index (EI)

- The EI is a parameter of extreme events related to the clustering of exceedances of high thresholds, a situation that occurs for stationary sequences [Leadbetter (1973), *ZWT*].
- We thus assume to be working with a strictly stationary sequence of r.v.'s, $\{X_n\}_{n\geq 1}$, from F, under the long range dependence condition **D** [Leadbetter, Lindgren & Rootzén, 1983] and the local dependence condition **D**'' [Leadbetter & Nandagopalan, 1989], straightforwardly true for i.i.d. data.

Definition 1. *The stationary sequence $\{X_n\}_{n\geq 1}$ is said to have an extremal index θ $(0 < \theta \leq 1)$ if, for all $\tau > 0$, we can find a sequence of levels $u_n = u_n(\tau)$ such that, with $\{Y_n\}_{n\geq 1}$ the associated i.i.d. sequence (i.e., an i.i.d. sequence from the same F),*

$$\boxed{\mathbb{P}\left(Y_{n:n} \leq u_n\right) = F^n(u_n) \underset{n\to\infty}{\longrightarrow} e^{-\tau} \text{ and } \mathbb{P}\left(X_{n:n} \leq u_n\right) \underset{n\to\infty}{\longrightarrow} e^{-\theta\tau}.}$$

- For dependent sequences there can thus appear a *"shrinkage"* of maximum values, but the limiting d.f. of $X_{n:n}$, linearly normalized, is still an *Extreme Value* d.f., EV_γ.
- Following Leadbetter (1983), *ZWT*, the *extremal index* can also be defined as:

$$\begin{aligned} \theta &= \frac{1}{\text{limiting mean size of clusters}} \\ &= \lim_{n\to\infty} P(X_2 \le u_n | X_1 > u_n) = \lim_{n\to\infty} P(X_1 \le u_n | X_2 > u_n), \end{aligned}$$

$$\boxed{u_n: \; F(u_n) = 1 - \tau/n + o(1/n), \text{ as } n \to \infty, \text{ with } \tau > 0, \text{ fixed.}}$$

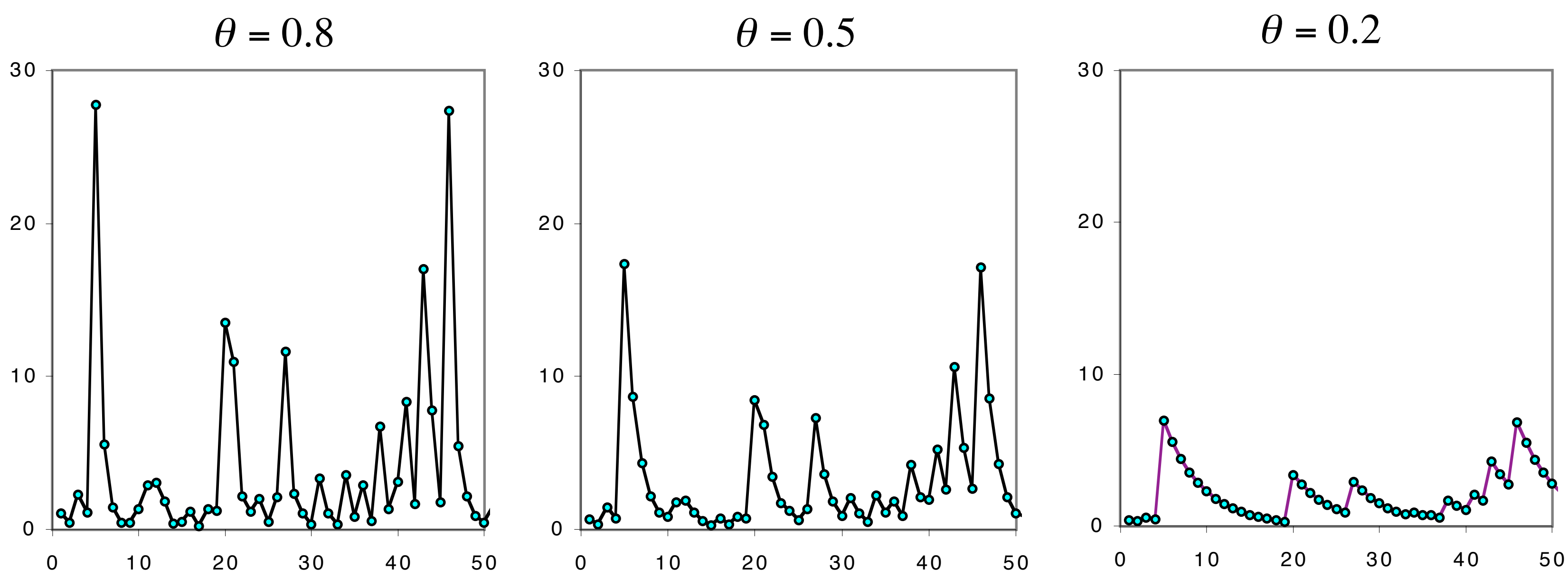
- The $ARMAX$ processes, will be the ones used here for illustration. Such processes are based on an i.i.d. sequence of innovations $\{Z_i\}_{i \geq 1}$, with d.f. H, and are defined through the relation,

$$\boxed{X_i = \beta \max(X_{i-1}, Z_i), \quad i \geq 1, \quad 0 < \beta < 1.}$$

- The $ARMAX$ sequence has a stationary distribution F, dependent on H through the relation $F(\beta x)/F(x) = H(x)$ [Alpuim, 1989, *JAP*].

- Conditions **D** e **D**'' hold for these sequences and stationary $ARMAX$ sequences may possess an extremal index $\theta < 1$.

- For illustration, we shall consider ARMAX processes with Fréchet innovations. If $H(x) = \Phi_\gamma^{\beta^{-1/\gamma}-1}(x)$, $F(x) = \Phi_\gamma(x) = \exp\left(-x^{-1/\gamma}\right)$, $x \geq 0$, and $\theta = 1 - \beta^{1/\gamma}$.



Notice the the richness of these processes, regarding clustering of exceedances. Note also that there is a "shrinkage" of maximum values, together with the exhibition of larger and larger "clusters" of exceedances of high values, as θ decreases.

2.3. First, second and third-order frameworks

- If $\overline{F} \in RV_{-1/\gamma}$, $\gamma > 0$, then [Gnedenko, 1943, *AM*], F is in the domain of attraction for maxima of a Fréchet-type *Extreme Value* d.f., and we write

$$\boxed{F \in \mathcal{D}_{\mathcal{M}}(EV_{\gamma>0}) =: \mathcal{D}_{\mathcal{M}}{}^{+}.}$$

- In this same context of heavy right-tails, and with the notation

$$\boxed{U(t) = F^{\leftarrow}(1 - 1/t), \ t \geq 1,}$$

with $F^{\leftarrow}(y) = \inf\{x : F(x) \geq y\}$ the *generalized inverse function* of the underlying model F, we can further say that

$$\boxed{F \in \mathcal{D}_{\mathcal{M}}{}^{+} \quad \Longleftrightarrow \quad \overline{F} \in RV_{-1/\gamma} \quad \Longleftrightarrow \quad U \in RV_{\gamma},}$$

the so-called *first-order conditions*.

- For consistent semi-parametric EVI-estimation, in the whole $\mathcal{D}^+_{\mathcal{M}}$, we merely need to assume the validity of the *first-order condition*, $U \in RV_\gamma$, and to work with adequate functionals, dependent on an *intermediate tuning* parameter k, the number of top o.s.'s involved in the estimation. This means that k needs to be such that

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

- To obtain information on the non-degenerate asymptotic behaviour of semi-parametric EVI-estimators, we need further assuming a *second-order condition*, ruling the rate of convergence in the *first-order condition*. The *second-order parameter*, ρ (≤ 0), rules such a rate of convergence, and it is the parameter appearing in the limiting result,

$$\boxed{\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} = \frac{x^\rho - 1}{\rho},}$$

which we often assume to hold for every $x > 0$, and where $|A|$ must be in RV_ρ [Geluk and de Haan, 1987]. For technical simplicity, we usually further assume that $\rho < 0$, writing $A(t) =: \gamma\beta t^\rho$.

- In order to obtain full information on the asymptotic bias of any corrected-bias EVI-estimator, it is usual to consider a *Pareto third-order condition*, i.e., a Pareto-type class of models, with a tail function

$$\boxed{1 - F(x) = Cx^{-1/\gamma}\Big(1 + D_1 x^{\rho/\gamma} + D_2 x^{2\rho/\gamma} + o\big(x^{2\rho/\gamma}\big)\Big),}$$

as $x \to \infty$, with $C > 0$, $D_1,\ D_2 \neq 0$, $\rho < 0$.

3. EVI and EI-ESTIMATORS

3.1. Classical EVI-estimators

- For models in $\mathcal{D}_{\mathcal{M}}^{+}$, the classical EVI-estimators are the Hill estimators [Hill, 1975, *AS*], averages of the *log-excesses*,

$$\boxed{V_{ik} := \ln X_{n-i+1:n} - \ln X_{n-k:n}, \qquad 1 \le i \le k < n,}$$

i.e.,

$$\boxed{H_n(k) \equiv H(k) := \frac{1}{k}\sum_{i=1}^{k} V_{ik}, \quad 1 \le k < n.}$$

- But these EVI-estimators have often a strong asymptotic bias for moderate up to large values of k, of the order of $A(n/k)$, and the adequate accommodation of this bias has recently been extensively addressed.

3.2. Second-order reduced-bias (SORB) EVI-estimators

- We mention the pioneering papers by Peng (1998) [*SN*], Beirlant, Dierckx, Goegebeur and Matthys (1999) [*Extremes*], Feuerverger and Hall (1999) [*AS*], and Gomes, Martins and Neves (2000) [*Extremes*], among others.

- In these papers, authors are led to SORB EVI-estimators, with asymptotic variances larger than or equal to $(\gamma\ (1-\rho)/\rho)^2$, where $\rho(<0)$ is the aforementioned "shape" second-order parameter, ruling the rate of convergence of the normalized sequence of maximum values towards the limiting law EV_γ.

3.3. MVRB EVI-estimators

- Later on, Caeiro, Gomes & Pestana (2005) [*Revstat*], Gomes, Martins & Neves (2007) [*Revstat*] and Gomes, de Haan and Henriques-Rodrigues (2008) [*JRSS*] have been able to *reduce the bias without increasing the asymptotic variance*, kept at γ^2.
- Those estimators, called *minimum-variance reduced-bias* (MVRB) EVI-estimators, are all based on an adequate "external" consistent estimation of the pair of second-order parameters, $(\beta, \rho) \in (\mathbb{R}, \mathbb{R}^-)$, done through estimators denoted $(\hat\beta, \hat\rho)$, and outperform the classical estimators for all k.
- We now consider the simplest class of MVRB EVI-estimators:

$$\boxed{\overline{H}(k) \equiv \overline{H}_{\hat\beta,\hat\rho}(k) := H(k)\left(1 - \hat\beta\,(n/k)^{\hat\rho}/(1-\hat\rho)\right).}$$

3.4. Asymptotic comparison of classical and MVRB EVI-estimators

- The Hill estimator reveals usually a high asymptotic bias. Indeed, it follows from the results of de Haan & Peng (1998) that under the *general second-order condition*,

$$\boxed{\sqrt{k}\,(H(k)-\gamma) \;\stackrel{d}{=}\; \text{Normal}_{0,\gamma^2} + b_H\sqrt{k}A(n/k) + o_p(\sqrt{k}A(n/k)),}$$

  where the bias $b_H\sqrt{k}A(n/k) = \gamma\;\beta\;\sqrt{k}\;(n/k)^\rho/(1-\rho)$ can be very large, moderate or small (i.e. go to ∞, 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 MSE pattern, as a function of k.

- Under the same conditions as before, $\sqrt{k}\left(\overline{H}(k)-\gamma\right)$ is asymptotically normal with variance also equal to γ^2 but with a null mean value. Indeed, under the validity of the aforementioned third-order condition related to Pareto-type class of models, we can then adequately estimate the vector of second-order parameters, (β,ρ), and write [Caeiro, Gomes & Henriques-Rodrigues, 2009, *CSTM*]

$$\boxed{\sqrt{k}\left(\overline{H}(k)-\gamma\right) \stackrel{d}{=} \mathrm{Normal}_{0,\gamma^2} + b_{\overline{H}}\sqrt{k}A^2(n/k) + o_p(\sqrt{k}A^2(n/k)).}$$

- Consequently, $\overline{H}(k)$ outperforms $H(k)$ for all k.

3.5. Classical EI-estimators

- Given a sample $(X_1, X_2, \ldots, X_n)$ and chosen a suitable threshold u, with I_A the indicator function of A, a possible estimator of θ [Leadbetter and Nandagopalan, 1989] is given by

$$\boxed{\widehat{\theta}_n^N = \widehat{\theta}_n^N(u) := \frac{\sum_{j=1}^{n-1} I_{[X_j>u,X_{j+1}\le u]}}{\sum_{j=1}^{n} I_{[X_j>u]}} = \frac{\sum_{j=1}^{n-1} I_{[X_j\le u<X_{j+1}]}}{\sum_{j=1}^{n} I_{[X_j>u]}}.}$$

- To have consistency, the high level u must be: $n(1-F(u_n)) = c_n\tau = \tau_n$, $\tau_n \to \infty$ and $\tau_n/n \to 0$ [Nandagopalan, 1990].

- To make the semi-parametric EI-estimation closer to the semi-parametric EVI-estimation, we consider [Gomes, Hall & Miranda, 2008, *CSDA*] $u \in \left[X_{n-k:n}, X_{n-k+1:n}\right)$ and the estimator

$$\boxed{\widehat{\theta}_n^N(k) \equiv \theta_n^N(u) := \frac{1}{k} \sum_{j=1}^{n-1} I_{[X_j \le X_{n-k:n} < X_{j+1}]}.}$$

Bias assumption on the data structures.

- For independent, identically distributed data ($\theta = 1$):

$$\boxed{\mathbb{E}[\widehat{\theta}_n^N(k)] = 1 + \left(\frac{1}{2\ k} - \frac{k}{n}\right)(1 + o(1)).}$$

- Moreover, for ARMAX processes, we get

$$\boxed{\mathbb{E}\left[\widehat{\theta}_n^N(k)\right] = \theta - \left(\frac{\theta(\theta+1)}{2}\left(\frac{k}{n}\right) - \frac{3 - 2\ \theta}{2\ k}\right)(1 + o(1))}.$$

- We shall thus consider the EI-estimator as a function of k, the number of o.s.'s higher than the chosen threshold. We further assume that, as $n \to \infty$, and for intermediate k,

$$\boxed{Bias\left[\widehat{\theta}_n^N(k)\right] = \varphi_1(\theta)\left(\frac{k}{n}\right) + \varphi_2(\theta)\left(\frac{1}{k}\right) + o\left(\frac{1}{k}\right) + o\left(\frac{k}{n}\right).}$$

- In the semi-parametric EI-estimation we have thus to cope with problems similar to the ones appearing in the EVI-estimation: *increasing bias, as the threshold decreases and a high variance for high thresholds*.

Is it possible to improve the performance of estimators through the use of computer intensive methods?

4. RESAMPLING METHODOLOGIES

- The use of resampling methodologies [Efron, 1979, *AS*] has revealed to be promising in the estimation of the nuisance parameter k, and in the reduction of bias of any estimator of a parameter of extreme events.
- If we ask how to choose the tuning parameter k in the estimation of a parameter of extreme events, η, through $T(k)$, we usually consider the estimation of $k_0^T := \arg\min_k MSE(T(k))$.
- To obtain estimates of k_0^T one can then use a *double-bootstrap* method applied to an adequate *auxiliary statistic* like $A(k) := T(k) - T(\lfloor k/2 \rfloor)$, where $\lfloor x \rfloor$ stands as usual to the integer part of x, and which tends to **zero** and has an asymptotic behaviour similar to the one of $T(k)$ (Gomes and Oliveira, 2001, *Extremes*, among others). We shall not sketch such a *double-bootstrap* algorithm.

- At such optimal levels, we have a non-null asymptotic bias.
- If we still want to remove such a bias, we can then make use of the *generalized jackknife* methodology.
- The main objectives of the *Jackknife methodology* are:
  1. Bias and variance estimation of a certain estimator, only through manipulation of observed data $\underline{x}$.
  2. The building of estimators with bias and mean squared error smaller than those of an initial set of estimators.
- The **Jackknife** or **Generalized Jacknife** (GJ) are resampling methodologies, which usually give a positive answer to the question: *"May the combination of information improve the quality of estimators of a certain parameter or functional?"* .

- It is then enough to consider an adequate pair of estimators of the parameter of extreme events under consideration, possibly also $T(k)$ and $T(\lfloor k/2 \rfloor)$, and to built a *reduced-bias affine combination* of them. In Gomes, Martins & Neves, 2000, also among others, we can find an application of this technique to the Hill estimator.
- In order to illustrate the use of these methodologies in EVT, we shall essentially consider, just as performed in Gomes, Martins & Neves, 2013, *CSTM*, the aforementioned MVRB EVI-estimators $\overline{H}(k)$ in Caeiro *et al.* (2005), and the classical EI-estimators, as performed in Gomes, Martins & Neves, 2007.

4.1. The jackknife methodology and bias reduction

- The pioneering EVI reduced-bias estimators are, in a certain sense, *generalized jackknife* (GJ) estimators, i.e., affine combinations of well-known estimators of γ.
- The *generalized jackknife* statistic was introduced by Gray and Shucany (1972): Let $T_n^{(1)}$ and $T_n^{(2)}$ be two biased estimators of γ, with similar bias properties, i.e.,

$$\boxed{\text{Bias}(T_n^{(i)}) = \gamma + \phi(\gamma) d_i(n), \quad i = 1, 2.}$$

Then, if $q = q_n = d_1(n)/d_2(n) \neq 1$, the affine combination

$$\boxed{T_n^G := \left(T_n^{(1)} - q T_n^{(2)}\right) / (1 - q)}$$

is an unbiased estimator of γ.

4.2. A GJ corrected-bias EVI-estimator

- Given $\overline{H}$, the most natural *GJ* r.v. is the one associated to the random pair $\left(\overline{H}(k), \overline{H}(\lfloor\theta k\rfloor)\right)$, $0 < \theta < 1$, is

$$\boxed{\overline{H}^{GJ(q,\theta)}(k) := \frac{\overline{H}(k) - q\ \overline{H}(\lfloor\theta k\rfloor)}{1-q},\ \ 0 < \theta < 1,}$$

with

$$\boxed{q = q_n = \frac{Bias_\infty[\overline{H}(k)]}{Bias_\infty[\overline{H}(\lfloor\theta k\rfloor)]} = \frac{A^2(n/k)}{A^2(n/\lfloor\theta k\rfloor)} \underset{n/k\to\infty}{\longrightarrow} \theta^{2\rho}.}$$

It is thus sensible to consider $q = \theta^{2\rho}$, $\theta = 1/2$, and, with $\hat\rho$ a consistent estimator of ρ, the GJ estimator,

$$\boxed{\overline{H}^{GJ}(k) := \frac{2^{2\hat\rho}\ \overline{H}(k) - \overline{H}(\lfloor k/2\rfloor)}{2^{2\hat\rho} - 1}.}$$

- Then, and provided that $\widehat{\rho} - \rho = o_p(1)$,

$$\boxed{\sqrt{k}\left(\overline{H}^{GJ}(k) - \gamma\right) \stackrel{d}{=} \text{Normal}_{0,\sigma^2_{GJ}} + o_p(\sqrt{k}A^2(n/k)),}$$

  with

$$\boxed{\sigma^2_{GJ} = \gamma^2(1 + 1/(2^{-2\rho} - 1)^2.}$$

- We have thus a trade-off between variance and bias . . .
  The bias decreases, but the variance increases . . .
  But we are able to reach a better performance at optimal levels.

4.3. A GJ corrected-bias EI-estimator

- Since the bias term of the aforementioned classical EI-estimator reveals 2 main components of $\neq$ orders, we need to use an affine combination of 3 EI-estimators and a order-2 GJ-statistic.
- Let $\underline{X} = (X_1, \ldots, X_n)$ be a sample from F, and let $T_n = T_n(\underline{X}, F)$ be an estimator of a functional $\theta(F)$, or of a parameter θ.
- If the bias of our estimator reveals 2 main terms that we would like to remove, the GJ methodology advises us to deal with 3 estimators with the same type of bias:

Definition 2. *Given 3 estimators $T_n^{(1)}$, $T_n^{(2)}$ and $T_n^{(3)}$ of θ:*

$$E\left[T_n^{(i)} - \theta\right] = d_1(\theta)\ \varphi_1^{(i)}(n) + d_2(\theta)\ \varphi_2^{(i)}(n),\ \ i = 1, 2, 3,$$

the GJ statistic (of order 2) is given by

$$\boxed{T_n^{GJ} := \left\| \begin{matrix} T_n^{(1)} & T_n^{(2)} & T_n^{(3)} \\ \varphi_1^{(1)} & \varphi_1^{(2)} & \varphi_1^{(3)} \\ \varphi_2^{(1)} & \varphi_2^{(2)} & \varphi_2^{(3)} \end{matrix} \right\| \Bigg/ \left\| \begin{matrix} 1 & 1 & 1 \\ \varphi_1^{(1)} & \varphi_1^{(2)} & \varphi_1^{(3)} \\ \varphi_2^{(1)} & \varphi_2^{(2)} & \varphi_2^{(3)} \end{matrix} \right\|,}$$

with $||A||$ denoting, as usual, the determinant of the matrix A.

- Straightforwardly, one may state:

Proposition 1. *T_n^{GJ} is unbiased for the estimation of θ.*

- Moreover, although the variance of T_n^{GJ} is always larger than the variance of the original estimators, the MSE of T_n^{GJ} is often smaller than that of any of the statistics $T_n^{(i)}$, $i = 1, 2, 3$.
- The information on the bias of the EI-estimator $\widehat{\theta}_n^N(k)$ led us to consider first the GJ EI-estimator of order 2, based on the estimator $\widehat{\theta}_n^N(k)$ computed at the three levels, k, $\lfloor k/2 \rfloor + 1$ and $\lfloor k/4 \rfloor + 1$ [Gomes and Miranda, 2003]:

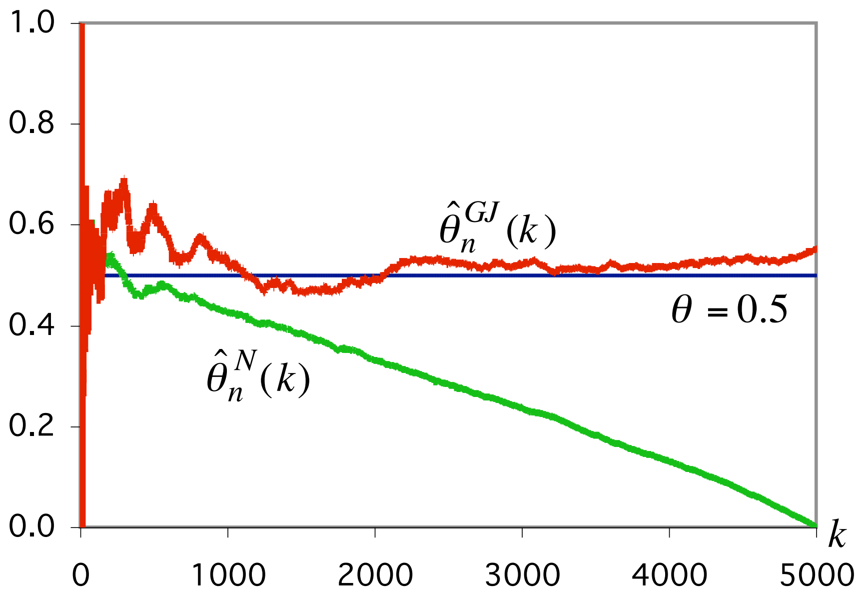
$$\boxed{\widehat{\theta}_n^{GJ}(k) = 5\ \widehat{\theta}_n^N\left(\lfloor k/2 \rfloor + 1\right) - 2\left(\widehat{\theta}_n^N\left(\lfloor k/4 \rfloor + 1\right) + \widehat{\theta}_n^N(k)\right).}$$

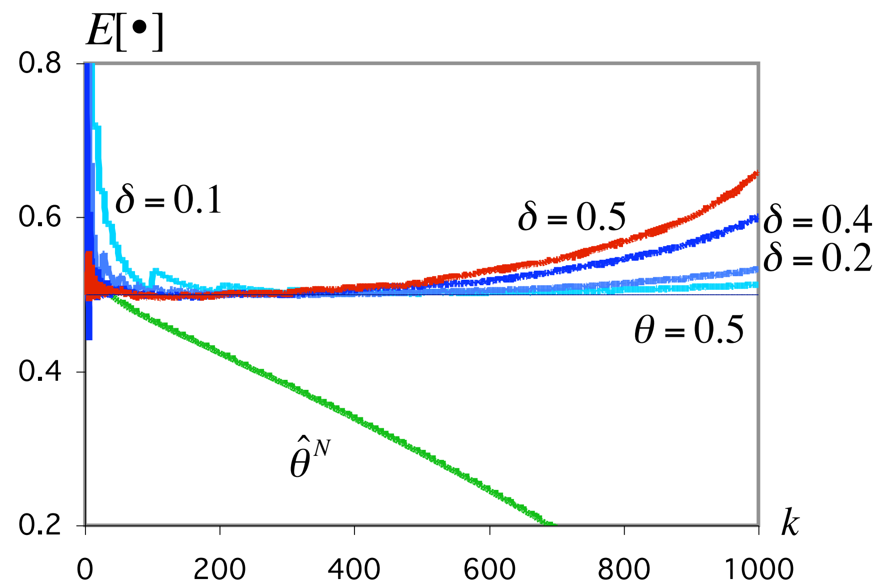
- This estimator has very stable sample paths, around the target value θ, BUT at expenses of a very high variance, which does not enable it to overpass the original estimator, regarding MSE at optimal levels.

- We thus think sensible to consider, more generally, the levels k, $\lfloor \delta k \rfloor + 1$ and $\lfloor \delta^2 k \rfloor + 1$, dependent of a *tuning parameter* δ, $0 < \delta < 1$, and the class of estimators,

$$\boxed{\widehat{\theta}_n^{GJ(\delta)}(k) := \frac{(\delta^2+1)\ \widehat{\theta}_n^N\left(\lfloor \delta k \rfloor + 1\right) - \delta\left(\widehat{\theta}_n^N\left(\lfloor \delta^2 k \rfloor + 1\right) + \widehat{\theta}_n^N(k)\right)}{(1-\delta)^2}.}$$

- Note that $\widehat{\theta}_n^{GJ}(k) \equiv \widehat{\theta}_n^{GJ(1/2)}(k)$.
- For a stationary Fréchet(1) ARMAX sample of size $n = 5000$, with $\theta = 0.5$, we next present
  - sample paths of $\widehat{\theta}_n^N(k)$ and $\widehat{\theta}_n^{GJ}(k)$ (*left*), and
  - the expected values of such an estimator, associated to $\delta =$ 0.1, 0.2, 0.4 e 0.5(*right*).





- Note the reasonably high stability around the target value $\theta = 0.5$, of the sample path and mean value of the GJ EI-estimator for a wide range of k-values, comparatively to that of Nandagopalan's estimator.

Remark 1. *The mean value stability around the target value θ, for a wide range of k-levels, is true for all θ and for all simulated models.*

But the GJ-estimator, $\widehat{\theta}_n^{GJ}$, may not overpass, for $n = 1000$ (and small θ), the original estimator, $\widehat{\theta}_n^{N}$, regarding MSE at optimal levels. Extra investment is thus needed on the "optimal" choice of the 3 levels to be used in the building of a GJ extremal index estimator or on the use of extra resampling or sub-sampling techniques, as performed in Gomes, Hall & Miranda (2008), who have used simple subsampling techniques, in order to attain a smaller mean squared error (MSE) at optimal levels.

5. A CASE STUDY

5.1. The GJ EVI-estimation applied to insurance data

We consider an illustration of the performance of the EVI-estimates under study, through the analysis of 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). This data set was already studied in Beirlant, Goegebeur, Segers & Teugels (2004), WILEY, Vandewalle and Beirlant (2006), *IME* and Beirlant, Figueiredo, Gomes & Vandewalle (2008), *JSPI*, as an example to excess-of-loss reinsurance rating and heavy-tailed distributions in car insurance.

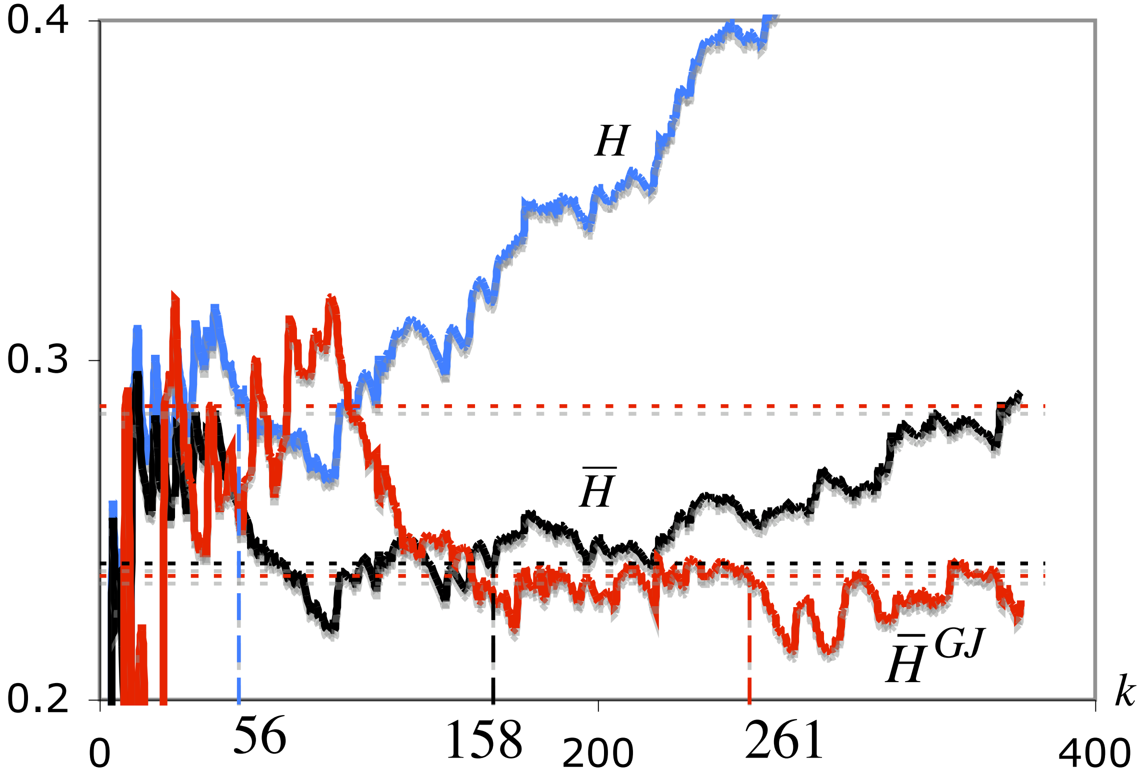
- Regarding the EVI-estimation, note that whereas the Hill estimator is unbiased for the estimation of γ when the underlying model is a strict Pareto model, it always exhibits a relevant bias when we have only Pareto-like tails, as happens here.
- The corrected-bias estimators, which are "asymptotically unbiased", have a smaller bias, exhibit more stable sample paths as functions of k, and enable us to take a decision upon the estimate of γ to be used, even with the help of any heuristic stability criterion, like the "largest run" suggested in Gomes and Figueiredo (2006), *Test*.
- For the Hill estimator, as we know how to estimate β and ρ, and we have simple techniques to estimate the OSF. Indeed, we get $\hat{k}_0^H = \left((1-\hat{\rho})^2 n^{-2\hat{\rho}} / \left(-2\ \hat{\rho}\ \hat{\beta}^2\right)\right)^{1/(1-2\hat{\rho})} = 58.$

- The aforementioned bootstrap algorithm, not detailed here, helps us to provide an adaptive choice for corrected-bias EVI-estimators.
- We have got $\hat{k}_{0|H} = 56$, $\hat{k}_{0|\overline{H}} = 158$, $\hat{k}_{0|\overline{H}^{GJ}} = 261$, and the EVI-estimates

$$H^* = 0.286, \quad \overline{H}^* = 0.240 \quad \text{and} \quad \overline{H}^{GJ^*} = 0.236,$$

the values pictured in the following Figure.

Remark 2. *Note that bootstrap confidence intervals as well as asymptotic confidence intervals are easily associated with the estimates presented, the smallest size (with a high coverage probability) being related with* $\overline{H}^{GJ^*}$.



$$H^* = .286\,(.236, .346)$$

$$\overline{H}^* = .240\,(.205, .275)$$

$$\overline{H}^{GJ^*} = .236\,(.208, .264)$$

Estimates of the extreme value index γ for the Secura Belgian Re data.

6. SOME OVERALL CONCLUSIONS

1. The most attractive features of the GJ estimators are their stable sample paths (for a wide region of k values), close to the target value, and the "bath-tube" MSE patterns.
   The insensitivity of the mean value (and sample path) to changes in k is indeed the nicest feature of these GJ-estimators.

2. Regarding MSE at optimal levels, the simplest GJ EI-estimator does not overpass the original one. To obtain relative efficiencies greater than 1, we had to proceed to a $\neq$ choice of the 3 levels under play. Even with such a choice, and for θ small, such an objective is often attained only with the extra use of a subsampling algorithm. Further investment is thus welcome.

REFERENCES

1. Alpuim, M.T. (1989). An extremal markovian sequence. *J. Appl. Probab.* **26**, 219–232.
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., Goegebeur, Y., Segers, J. and Teugels, J. (2004). *Statistics of Extremes. Theory and Applications*. Wiley.
4. Beirlant, J., Figueiredo, F., Gomes, M.I. and Vandewalle, B. (2008). Improved reduced-bias tail index and quantile estimators. *J. Statist. Plann. and Inference* **138**:6, 1851–1870.
5. Caeiro, F., Gomes, M.I. and Pestana, D.D. (2005). Direct reduction of bias of the classical Hill estimator. *Revstat* **3**(2), 113–136.
6. Caeiro, F., Gomes, M.I. and Henriques-Rodrigues, L. (2009). Reduced-bias tail index estimators under a third order framework. *Communications in Statistics – Theory & Methods* **38**:7, 1019–1040.
7. Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife *Ann. Statist.* **7**:1, 1–26.
8. Feuerverger, A. and Hall, P. (1999). Estimating a tail exponent by modelling departure from a Pareto distribution. *Annals Statistics* **27**, 760–781.
9. Geluk, J. and de Haan, L. (1987). *Regular Variation, Extensions and Tauberian Theorems*. CWI Tract 40, Center for Mathematics and Computer Science, Amsterdam, Netherlands.
10. Gnedenko, B.V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. Math.* **44**, 423–453.

11. Gomes, M.I. and Figueiredo, F. (2006). Bias reduction in risk modelling: semi-parametric quantile estimation. *Test* **15**:2, 375–396.
12. Gomes, M.I. and Miranda, M.C. (2003). A metodologia jackknife na estimação do índice extremal. In Brito, P. et al. (eds.), *Literacia e Estatística*, Edições S.P.E., 299–310.
13. 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, 2002.
14. Gomes, M.I., Martins, M.J. and Neves, M. (2000). Alternatives to a semi-parametric estimator of parameters of rare events: the Jackknife methodology. *Extremes* **3**(3), 207–229.
15. Gomes, M. I., Martins, M. J. and Neves, M. (2007). Improving second order reduced bias extreme value index estimation. *Revstat* **5**(2), 177–207.
16. Gomes, M. I., de Haan, L. and Henriques Rodrigues, L. (2008). Tail index estimation through accommodation of bias in the weighted log-excesses. *J. Royal Statistical Society* **B70**, Issue 1, 31–52.
17. Gomes, M.I., Hall, A. and Miranda, C. (2008). Subsampling techniques and the Jackknife methodology in the estimation of the extremal index. *J. Comput. Statist. and Data Analysis* **52**:4, 2022–2041.
18. Gomes, M.I., Martins, M.J. and Neves, M.M. (2013). Generalised Jackknife-Based Estimators for Univariate Extreme-Value Modelling. *Communications in Statistics: Theory and Methods* **42**:7, 1227–1245.
19. Gray, H.L., and Schucany, W.R. (1972). *The Generalized Jackknife Statistic*. Marcel Dekker.
20. de Haan, L. and Peng, L. (1998). Comparison of tail index estimators. *Statistica Neerlandica* **52**, 60–70.

21. Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3**, 1163–1174.
22. Leadbetter, M.R. (1973). On extreme values in stationary sequences. *Z. Wahrsch. Verw. Gebiete* **28**, 289–303.
23. Leadbetter, M.R. (1983). Extremes and local dependence in stationary sequences. *Z. Wahrsch. Verw. Gebiete* **65**:2, 291–306.
24. Leadbetter, M.R. and Nandagopalan, S. (1989). On exceedance point processes for stationary sequences under mild oscillation restrictions. In Hüsler, J. and R.-D. Reiss (eds.), *Extreme Value Theory*, Springer-Verlag, 69–80.
25. Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Series*. Springer-Verlag, New York.
26. Nandagopalan, S. (1990). *Multivariate Extremes and Estimation of the Extremal Index*. Ph.D. Thesis, Univ. North Carolina at Chapel Hill.
27. Peng, L. (1998). Asymptotically unbiased estimator for the extreme-value index. *Statist. and Probab. Letters* **38**(2), 107–115.
28. Vandewalle, B. and Beirlant, J. (2006). On univariate extreme value statistics and the estimation of reinsurance premiums. *Insurance: Mathematics and Economics* **38**, 441–459.

THAT's ALL and THANKS . . .

To Ross: a photo of Lisbon, we both love, as another token of friendship.