Bias reduction in risk modelling*

M. Ivette Gomes†
DEIO and CEAUL, Faculty of Science of Lisbon,

and

Fernanda Figueiredo‡
CEAUL (University of Lisbon) and Faculty of Economics, Porto

Abstract. In *Statistics of Extremes* we are mainly interested in the estimation of quantities related to extreme events, and this is a relevant topic in many areas of application. For instance, in Insurance Mathematics and Finance, a typical requirement is to find a value, high enough, so that the chance of an exceedance of that value is small. We are then interested in the estimation of a *high quantile* χ_p, a value which is overpassed with a small probability p. Also in *Statistical Quality Control* the on-line region depends strongly on the estimation of quantiles χ_p, with p quite close either to 0 or to 1. In this paper we deal with the semi-parametric estimation of χ_p for heavy tails. Since the classical estimators exhibit a reasonably high bias for low thresholds, we shall deal with bias reduction techniques, trying to improve their performance.

AMS 2000 subject classification. Primary 62G32, 62E20; Secondary 65C05.
Keywords and phrases. *Heavy tails; High quantiles; Semi-parametric estimation; Statistics of Extremes.*

1 Introduction.

In the most diversified areas of research we are often interested in finding a value, high (or low) enough, so that the chance of an exceedance (or of a non-exceedance) of that value is small. For instance, in Insurance Mathematics and Finance, a typical requirement is the estimation of a level $\chi_p = VaR_p$, the *Value at Risk* at the level p, a value which is overpassed with a small probability $p = p_n \to 0$, as $n \to \infty$. Also in Statistical Quality Control the on-line region depends strongly on the estimation of quantiles χ_p, with p quite close either

*Research partially supported by FCT / POCTI / FEDER.

†DEIO, Faculdade de Ciências de Lisboa, Bloco C2, Piso 2, Campo Grande, 1749-016 Lisboa, Portugal; e-mail: ivette.gomes@fc.ul.pt

‡Faculdade de Economia do Porto, Rua Dr. Roberto Frias, 4200-464 Porto, Portugal; e-mail: otilia@fep.up.pt

to 0 or to 1. And in the design of dams, a typical requirement is that the sea wall must be high enough so that the chance of a flood is no more than once in ten thousand years. We are thus in the field of Statistics of Extremes, a field in which we are mainly interested in the estimation of quantities related to extreme or even rare events, being scarce the information from previous experiments.

Indeed, one of the main goals of extreme value theory is to do inference in the far tail of a probability distribution function (d.f.). For this it is necessary to extend the empirical d.f. beyond the range of the available data, estimating the tail parameter as accurately as possible. We shall assume that the d.f. F satisfies a regular variation condition, i.e.,

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-1/\gamma} \tag{1.1}$$

for all $x > 0$, where γ (> 0) is the *tail parameter* (see, for instance, Galambos (1987), for details on Extreme Value Theory). Equivalently, if we consider the quantile function $U(t) = F^{\leftarrow}(1 - 1/t)$, $t \geq 1$, with $F^{\leftarrow}$ denoting the generalized inverse function of F, i.e., $F^{\leftarrow}(t) = \inf\{x : F(x) \geq t\}$, (1.1) is equivalent to say that

$$\lim_{t\to\infty} \{\ln U(tx) - \ln U(t)\} = \gamma \ln x \tag{1.2}$$

for all $x > 0$. Several semi-parametric estimators of the tail parameter γ are available in the literature. In order to get asymptotic normality for those estimators, we need more than the first order condition (1.1) or (1.2). A convenient condition to achieve asymptotic normality is the second order condition

$$\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} = \frac{x^\rho - 1}{\rho} \tag{1.3}$$

for all $x > 0$, where A is a suitably chosen function of constant sign near infinity (positive or negative), and $\rho \leq 0$ is the second order parameter. The limit function in (1.3) is necessarily of this given form, and $|A| \in RV_\rho$ (Geluk and de Haan, 1987). The notation RV_β stands for the class of *regularly varying* functions at infinity with *index of regular variation* equal to β, i.e., positive measurable functions g such that $\lim_{t\to\infty} g(tx)/g(t) = x^\beta$, for all $x > 0$.

We want to estimate

$$\chi_p \equiv VaR_p : \; 1 - F(\chi_p) = p, \quad p \text{ small},$$

i.e.,

$$\chi_p := U\left(\frac{1}{p}\right), \quad p = p_n \to 0, \quad n\, p_n \to K \geq 0, \text{ as } n \to \infty, \tag{1.4}$$

and we shall assume to be working in Hall's class of models (Hall and Welsh, 1985), where $\rho < 0$ and $U(t) \sim C\ t^{\gamma}$. More specifically, we shall assume that

$$U(t) = Ct^{\gamma}\left(1 + Dt^{\rho} + o(t^{\rho})\right), \quad \rho < 0, \quad \text{as } t \to \infty. \tag{1.5}$$

We are going to base inference on the largest k top order statistics (o.s.), and as usual in semi-parametric estimation of parameters of rare events, we shall assume that k is an intermediate sequence of integers, i.e.,

$$k = k_n \to \infty, \quad k = o(n) \qquad \text{as } n \to \infty. \tag{1.6}$$

Our aim is essentially to present new estimators for χ_p in (1.4) and to prove their consistency and asymptotic normality under appropriate conditions. The main feature of these estimators is that for estimating χ_p more upper order statistics should be used than for the classical estimation of a quantile, because we are going to base our quantile estimation on an "asymptotically unbiased" estimator of the tail index γ, to be specified in section 2, getting thus an "asymptotically unbiased" estimator of the quantile.

We shall work with the statistics

$$M_n^{(\alpha)}(k) := \frac{1}{k}\sum_{i=1}^{k}\left[\ln X_{n-i+1:n} - \ln X_{n-k:n}\right]^{\alpha}, \ \alpha \in \mathbb{R}^{+}, \tag{1.7}$$

where $X_{i:n}$, $1 \le i \le n$, is the sample of ascending o.s. associated to our original, independent identically distributed (i.i.d.) sample $(X_1, X_2, \ldots, X_n)$. These statistics were introduced, and studied under a second order framework, by Dekkers et al. (1989). For more details on these statistics, and the way they may be used to build alternatives to the Hill estimator (Hill, 1975), given by $\alpha = 1$ in (1.7), see Gomes and Martins (2001) and also Caeiro and Gomes (2002). The starting point is a well-known expansion for $M_n^{\alpha}(k)$, for any $\alpha > 0$:

$$M_n^{\alpha}(k) = \gamma^{\alpha}\mu_{\alpha}^{(1)} + \left(\gamma^{\alpha}\sigma_{\alpha}^{(1)}\frac{1}{\sqrt{k}}P_k^{(\alpha)} + \alpha\gamma^{\alpha-1}\mu_{\alpha}^{(2)}(\rho)A(n/k)\right)\ (1 + o_p(1))$$

where $P_k^{(\alpha)}$ is a standard normal random variable (r.v.) and, with W denoting a unit exponential r.v.,

$$\mu_{\alpha}^{(1)} := E[W^{\alpha}] = \Gamma(\alpha+1),$$

$$\sigma_{\alpha}^{(1)} := \sqrt{Var[W^{\alpha}]} = \sqrt{\Gamma(2\alpha+1) - \Gamma^2(\alpha+1)},$$

and

$$\mu_{\alpha}^{(2)}(\rho) := E\left[W^{\alpha-1}\left(\frac{e^{\rho W}-1}{\rho}\right)\right] = \frac{\Gamma(\alpha)}{\rho}\frac{1-(1-\rho)^{\alpha}}{(1-\rho)^{\alpha}}.$$

Notice that since $\chi_p = U(1/p)$ and $U(t) \sim C\, t^\gamma$, an obvious estimator of χ_p is $\widehat{\chi}_p = \widehat{C} p^{-\widehat{\gamma}}$, with $\widehat{\gamma}$ any consistent estimator of γ. Also, with $Y_{i:n},\ 1 \le i \le n$, denoting the set of ascending o.s. associated to a standard Pareto i.i.d. sample from the model $F_Y(y) = 1 - y^{-1},\ y \ge 1,\ X_{n-k:n} = U(Y_{n-k:n}) \sim C\ Y_{n-k:n}^\gamma \sim C\left(\frac{n}{k}\right)^\gamma$. Consequently, an obvious estimator of C is $\widehat{C} = \left(\frac{k}{n}\right)^{\widehat{\gamma}} X_{n-k:n}$, and,

$$\widehat{\chi}_p(k) = X_{n-k:n} \left(\frac{k}{np}\right)^{\widehat{\gamma}} \tag{1.8}$$

is the obvious quantile estimator, to be considered in section 3 of this paper. Such an estimator depends strongly on an adequate estimation of the tail index γ, and we shall here propose the use of the "asymptotically unbiased" estimators discussed in section 2. Those estimators depend on an external estimation of the second order parameter $\rho < 0$ in (1.5), also addressed in this same section. Finally, in section 4, we present the simulated behaviour of our χ_p-estimators for parent distributions in Hall's class, like the Fréchet and the Burr.

2 The tail index estimation

Among the "asymptotically unbiased" estimators of the tail index already introduced and studied at sub-optimal levels, we shall select here the two ones with smaller asymptotic variances. Notice that these estimators have not yet been studied at their optimal levels, and this turns out to be an interesting topic of open research. These estimators require the estimation of the second order parameter ρ. Such an estimation will be briefly discussed before the discussion of the tail index estimators to be incorporated in the estimation of high quantiles.

2.1 Second order parameter's estimators

We shall consider here particular members of the class of estimators of the second order parameter ρ proposed by Fraga Alves et al. (2001). Under adequate general conditions, they are semi-parametric asymptotically normal estimators of ρ, which show highly stable sample paths as functions of k, the number of top o.s. used, for a wide range of large k-values. Such a class of estimators is parametrized in a tuning parameter τ and depends on the statistics

$$T_n^{(\tau)}(k) := \begin{cases} \dfrac{\left(M_n^{(1)}(k)\right)^\tau - \left(M_n^{(2)}(k)/2\right)^{\tau/2}}{\left(M_n^{(2)}(k)/2\right)^{\tau/2} - \left(M_n^{(3)}(k)/6\right)^{\tau/3}} & \text{if} \quad \tau > 0 \\[2ex] \dfrac{\ln\left(M_n^{(1)}(k)\right) - \frac{1}{2}\ln\left(M_n^{(2)}(k)/2\right)}{\frac{1}{2}\ln\left(M_n^{(2)}(k)/2\right) - \frac{1}{3}\ln\left(M_n^{(3)}(k)/6\right)} & \text{if} \quad \tau = 0, \end{cases}$$

which converge towards $3(1-\rho)/(3-\rho)$, independently of τ, whenever the second order condition (1.3) holds and k is such that $k \to \infty$, $k = o(n)$ and $\sqrt{k}\ A(n/k) \to \infty$, as $n \to \infty$. The estimators are thus given by

$$\widehat{\rho}_n^{(\tau)}(k) := \min\left(0,\ \frac{3(T_n^{(\tau)}(k)-1)}{T_n^{(\tau)}(k)-3}\right). \tag{2.1}$$

We shall formalize, without proofs, the main distributional results of the estimators in (2.1). Proofs may be found in Fraga Alves *et al.* (2001).

Proposition 2.1. *If the second order condition (1.3) holds, with $\rho < 0$, k is a sequence of intermediate integers, i.e., (1.6) holds, and*

$$\lim_{n\to\infty} \sqrt{k}\ A(n/k) = \infty, \tag{2.2}$$

then $\widehat{\rho}_n^{(\tau)}(k)$ in (2.1) converges in probability towards ρ, as $n \to \infty$.

Proposition 2.2. *If we further assume that a third order condition holds, which we here write as*

$$\lim_{t\to\infty} \frac{\frac{\ln U(tx)-\ln U(t)-\gamma\ln x}{A(t)} - \frac{x^\rho-1}{\rho}}{B(t)} = \frac{x^{2\rho}-1}{2\rho}, \tag{2.3}$$

then necessarily with $|B| \in RV_\rho$, $\rho < 0$, together with (2.2), and also the validity of the condition:

$$\lim_{n\to\infty} \sqrt{k}\ A^2(n/k) = \lambda_1,\ \textit{finite},$$

which implies

$$\lim_{n\to\infty} \sqrt{k}\ A(n/k)\ B(n/k) = \lambda_2,\ \textit{finite},$$

we may guaranteee asymptotic normality of the estimators $\widehat{\rho}_n^{(\tau)}(k)$ in (2.1). Moreover, $\widehat{\rho}_n^{(\tau)}(k) - \rho = O_p\left(\frac{1}{\sqrt{k}A(n/k)}\right)$.

Remark 2.1. *The theoretical and simulated results in Fraga Alves et al. (2001), together with the use of these estimators in the Generalized Jackknife statistics of Gomes et al. (2000) led us to advise in practice the consideration of the level*

$$k_1 = \min(n-1, [2n/\ln\ln n]) \tag{2.4}$$

(not chosen in an optimal way), and of the tuning parameters $\tau = 0$ for the region $\rho \in [-1, 0)$ and $\tau = 1$ for the region $\rho \in (-\infty, -1)$. In the simulations of section 4 we have always used $\tau = 1$ in (2.1) and the level k_1 in (2.4). Anyway, we advise practitioners not to choose blindly the value of τ in (2.1). It is sensible to draw a few sample paths of $\widehat{\rho}_n^{(\tau)}(k)$ in (2.1), as functions of k, electing the value of τ which provides higher stability for large k, by means of any stability criterion.

2.2 The "asymptotically unbiased" estimators of the tail index

We shall make first a review of some explicit "asymptotically unbiased" tail index estimators, $\widehat{\gamma}(k)$, so far proposed in the literature. Notice that they are all based on an external estimation of the second order parameter ρ. More than that: an asymptotic distributional representation of the type

$$\widehat{\gamma}(k) \stackrel{d}{=} \gamma + \frac{\sigma}{\sqrt{k}} Q_k + o_p(A(n/k)), \tag{2.5}$$

with Q_k an asymptotic standard normal r.v., holds true for any of the estimators $\widehat{\gamma}(k)$ herewith considered, i.e., $\sqrt{k}\,(\widehat{\gamma}(k) - \gamma)$ is asymptotically normal with null mean value not only when $\sqrt{k}\ A(n/k) \to 0$ (as usual), but also when $\sqrt{k}\ A(n/k) \to \lambda$, finite and non-null. Let us denote $\widehat{\rho} = \widehat{\rho}_n^{(\tau)}(k_1)$ any of the estimators in (2.1), with k_1 in (2.4).

1. Gomes and Martins (2001): If we consider the "asymptotically unbiased" estimator proposed in this paper, obtained through the external estimation of the second order parameter ρ, we get the estimator

   $$\widehat{\gamma}_n^{(1)} \equiv \widehat{\gamma}_n^{(\widehat{\alpha},1)}(k) := \frac{M_n^{(\widehat{\alpha})}(k)}{\Gamma(\widehat{\alpha}+1)\left[M_n^{(1)}(k)\right]^{\widehat{\alpha}-1}}, \tag{2.6}$$

   where $\widehat{\alpha}$ is such that:

   $$(1-\widehat{\rho})^{\widehat{\alpha}-1}[1+\widehat{\rho}(\widehat{\alpha}-2)] = 1. \tag{2.7}$$

   The asymptotic variance of the estimator in (2.6), i.e., the value σ^2 in (2.5) associated to this estimator, is

   $$\sigma_1^2 \equiv \varphi_1^2(\rho) = \gamma^2 \left(\frac{\Gamma(2\alpha_\rho^{(1)}+1)}{\Gamma^2(\alpha_\rho^{(1)}+1)} - \left(\alpha_\rho^{(1)}\right)^2 \right),$$

   with

   $$\alpha_\rho^{(1)}: \quad (1-\rho)^{\alpha_\rho^{(1)}-1}[1+\rho(\alpha_\rho^{(1)}-2)] = 1.$$

2. Gomes and Caeiro (2002): The estimator herewith considered is

$$\widehat{\gamma}_n^{(2)} \equiv \widehat{\gamma}_n^{(\widehat{\alpha},2)}(k) := \frac{\Gamma(\widehat{\alpha})}{M_n^{(\widehat{\alpha}-1)}(k)} \left(\frac{M_n^{(2\widehat{\alpha})}(k)}{\Gamma(2\widehat{\alpha}+1)} \right)^{1/2}, \tag{2.8}$$

with

$$\widehat{\alpha} = -\frac{\ln\left[1-\widehat{\rho}-\sqrt{(1-\widehat{\rho})^2-1}\right]}{\ln(1-\widehat{\rho})}. \tag{2.9}$$

The asymptotic variance σ_2^2 of the estimator in (2.8) is now ruled by

$$\varphi_2^2(\rho) = \frac{\gamma^2}{4}\left\{ \frac{\Gamma(4\alpha_\rho^{(2)})}{\alpha_\rho^{(2)}\Gamma^2(2\alpha_\rho^{(2)})} + \frac{4\,\Gamma(2\alpha_\rho^{(2)}-1)}{\Gamma^2(\alpha_\rho^{(2)})} - \frac{2\,\Gamma(3\alpha_\rho)^{(2)}}{\alpha_\rho\Gamma(\alpha_\rho^{(2)})\Gamma(\alpha_\rho)^{(2)}} - 1 \right\},$$

with

$$\alpha_\rho^{(2)} = \alpha_0^{(2)}(\rho) = -\frac{\ln\left[1-\rho-\sqrt{(1-\rho)^2-1}\right]}{\ln(1-\rho)}.$$

3. Gomes and Martins (2002b): Among the estimators considered in this paper, we have selected here the following two: the Generalized Jackknife estimator,

$$\widehat{\gamma}_n^{(3)} \equiv \gamma_n^{GJ(\widehat{\rho})}(k) := \frac{-(2-\widehat{\rho})\gamma_n^{(2)}(k) + 2\gamma_n^{(3)}(k)}{\widehat{\rho}}, \tag{2.10}$$

based on

$$\gamma_n^{(2)}(k) := M_n^{(2)}(k)/(2M_n^{(1)}(k))$$

and

$$\gamma_n^{(3)}(k) := \sqrt{M_n^{(2)}(k)/2};$$

and the "Maximum Likelihood" estimator

$$\widehat{\gamma}_n^{(4)} \equiv \gamma_n^{ML(\widehat{\rho})}(k) := \frac{1}{k}\sum_{i=1}^{k} U_i - \left(\frac{1}{k}\sum_{i=1}^{k} i^{-\widehat{\rho}}\, U_i\right) \frac{\left(\sum_{i=1}^{k} i^{-\widehat{\rho}}\right)\left(\sum_{i=1}^{k} U_i\right) - k\left(\sum_{i=1}^{k} i^{-\widehat{\rho}} U_i\right)}{\left(\sum_{i=1}^{k} i^{-\widehat{\rho}}\right)\left(\sum_{i=1}^{k} i^{-\widehat{\rho}} U_i\right) - k\left(\sum_{i=1}^{k} i^{-2\widehat{\rho}} U_i\right)}, \tag{2.11}$$

based on the scaled log-spacings

$$U_i := i\ \left[\ln X_{n-i+1:n} - \ln X_{n-i:n}\right]. \tag{2.12}$$

The asymptotic variances of the estimators in (2.10) and in (2.11) are ruled by

$$\varphi_3^2(\rho) = \frac{\gamma^2(2\rho^2 - 2\rho + 1)}{\rho^2}$$

and

$$\varphi_4^2(\rho) = \frac{\gamma^2(1-\rho)^2}{\rho^2}$$

respectively.

The quantile estimators associated to the tail index estimators $\widehat{\gamma}_n^{(j)}$ given before will be denoted $\widehat{\chi}_n^{(j)}$, $j = 1, 2, 3, 4$. The estimators $\widehat{\chi}_n^{(3)}$ and $\widehat{\chi}_n^{(4)}$ will also be denoted $\widehat{\chi}_n^{GJ}$ and $\widehat{\chi}_n^{ML}$, respectively. The quantile estimator associated to the Hill estimator will be denoted either $\widehat{\chi}_n^{(0)}$ or $\widehat{\chi}_n^{H}$. The standard deviations $\varphi_j(\rho)$, $j = 1, 2, 3, 4$, are drawn in Figure 1.

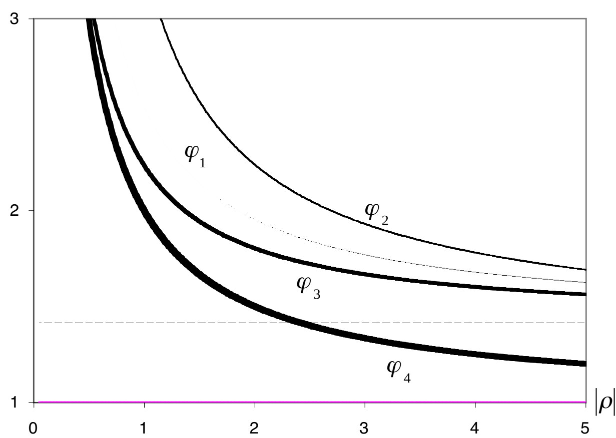

Figure 1: Asymptotic standard deviations of the tail index estimators under study, up to a factor γ

The comparison of those standard deviations enable us to draw the following conclusions:

1. The so-called ML estimator, first introduced in Beirlant et al. (1999) and Feuerverger and Hall (1999), studied with a mispecification of $\rho = -1$ in Gomes and Martins (2002a), and further studied under an external estimation of ρ in Gomes and Martins (2002b), is, among these four "asymptotically unbiased" estimators considered, the one which provides the smallest asymptotic variance for all values of ρ.

2. The one with the largest asymptotic variance is the one in Gomes and Caeiro (2002). This class of estimators was first considered in Caeiro and Gomes (2002), where the tuning parameter α was chosen through an

adequate stability criterion of the sample paths. It was then such a choice of α that would give rise to a possible estimate of ρ, useless for applications, due to its high variance.

3 How to estimate a high quantile

In a semi-parametric context, and for heavy tails, the most usual estimator of χ_p is, as previously mentioned,

$$\widehat{\chi}_p(k) := X_{n-k:n} \left(\frac{k}{np}\right)^{\widehat{\gamma}(k)} =: X_{n-k:n}\ a_n^{\widehat{\gamma}(k)}, \quad a_n = \frac{k}{np_n}. \tag{3.1}$$

where $\widehat{\gamma}(k)$ is any semi-parametric estimator of the tail index γ. Details on semi-parametric estimation of extremely high quantiles for a general tail index $\gamma \in \mathbb{R}$ may be found in Haan and Rootzén (1993) and more recently in Ferreira et al. (2002).

We state, without proof the following lemma.

Lemma 3.1. *Under the second order framework in (1.3) and for intermediate k, i.e., k such that (1.6) holds, we may write*

$$M_n^{(1)}(k) \stackrel{d}{=} \gamma + \left(\frac{\gamma\ P_k^{(1)}}{\sqrt{k}} + \frac{A(n/k)}{1-\rho}\right)(1 + o_p(1)), \tag{3.2}$$

where $P_k^{(1)}$ is a standard normal r.v. Also,

$$X_{n-k:n} \stackrel{d}{=} U(n/k)\left(1 + \frac{\gamma\ B_k}{\sqrt{k}} + o_p(A(n/k))\right), \tag{3.3}$$

with B_k standard normal, independent of $P_k^{(1)}$.

We may further state:

Theorem 3.1. *Under the conditions of Lemma 3.1, in Hall's class of models, where (1.5) holds, and whenever*

$$\ln np_n = o\left(\sqrt{k}\right), \tag{3.4}$$

if we consider a tail index estimator $\widehat{\gamma}_n(k)$ *such that* $\sqrt{k}\ \ (\widehat{\gamma}_n(k)-\gamma)$ *has an asymptotic variance equal to* σ^2 *and a null asymptotic bias, even when* $\sqrt{k}\ A(n/k) \to \lambda$*, non-null and finite, i.e., if*

$$\widehat{\gamma}(k)-\gamma \ \stackrel{d}{=}\ \frac{\sigma\ Q_k}{\sqrt{k}} + o_p(A(n/k)), \tag{3.5}$$

with Q_k *an asymptotically standard normal r.v.,*

$$\frac{\sqrt{k}\,(\widehat{\chi}_p(k)-\chi_p)}{\sigma\ a_n^{\gamma}\ \ln a_n\ U(n/k)}\ \stackrel{d}{=}\ Q_k + o_p(\sqrt{k}\ A(n/k)). \tag{3.6}$$

We may also write

$$\frac{\sqrt{k}}{\sigma\ln a_n}\left(\frac{\widehat{\chi}_p(k)}{\chi_p}-1\right)\ \stackrel{d}{=}\ Q_k + o_p(\sqrt{k}\ A(n/k)). \tag{3.7}$$

We thus have asymptotic normality with a null bias not only when $\sqrt{k}\ A(n/k)\to 0$*, as usually happens, but also when* $\sqrt{k}\ A(n/k)\to\lambda$*, non-null, but finite.*

Proof. The use of the δ-method, together with the fact that $\ln a_n = o\left(\sqrt{k}\right)$, enables us to write

$$a_n^{\widehat{\gamma}(k)} \stackrel{d}{=}\ a_n^{\gamma} + a_n^{\gamma}\ln a_n\ \ (\widehat{\gamma}(k)-\gamma)\,(1+o_p(1)).$$

Then, since $\chi_p = U(1/p) = U(na_n/k)$, and (3.3) holds,

$$\begin{aligned}\widehat{\chi}_p(k)-\chi_p &= X_{n-k:n}\left(a_n^{\widehat{\gamma}(k)} - \frac{U(na_n/k)}{U(n/k)}\ \frac{U(n/k)}{X_{n-k:n}}\right)\\ &= U(n/k)\left(1+\frac{\gamma\ B_k}{\sqrt{k}} + o_p(A(n/k))\right)\ a_n^{\gamma}\ln a_n(\widehat{\gamma}(k)-\gamma)(1+o_p(1)).\end{aligned}$$

Consequently, if we choose a tail index estimator with a null dominant component of asymptotic bias, i.e., if (3.5) holds, (3.6) follows, as well as the remaining of the theorem. □

Remark 3.1. *The result in Theorem 3.1 may be easily generalized to a real tail index* γ *provided we are able to get "asymptotically unbiased" estimators of* $\gamma\in\mathbb{R}$*, and this is still, as far as we know, an open question.*

Remark 3.2. *In order to have* $\widehat{\chi}_p - \chi_p \rightarrow 0$ *we need to guarantee that* $\frac{a_n^\gamma \ \ln a_n \ U(n/k)}{\sqrt{k}} \rightarrow 0$. *Consequently, even if K in (1.4) is non-null (and then* $a_n = O(k)$*, we need to have* $\ln k/\sqrt{k} = o(n^{-\gamma})$*, and this happens for* $k = n^\epsilon$*,* $\epsilon > 0$ *only if* $\gamma < 1/2$*.*

Remark 3.3. *However,* $\widehat{\chi}_p/\chi_p - 1 \rightarrow 0$*, as* $n \rightarrow \infty$ *whenever* $\ln a_n = o\left(\sqrt{k}\right)$*, and this happens whenever* $\ln np_n = o\left(\sqrt{k}\right)$*, one of the conditions in Theorem 3.1.*

Remark 3.4. *The simulation results obtained for a reasonable large class of heavy tail models, and partially presented in section 4, enable us to advance that the high stability of sample paths achieved with the "asymptotically unbiased" estimates of the tail index is no longer achieved for these high quantiles' estimates. Anyway, the proposed estimators perform better than the classical ones, and for most of the simulated models exhibit a much more stable mean value pattern, as a function of k, the number of top o.s. used.*

4 Simulated properties of the VaR_p-estimator

For the estimation of the second order parameter ρ we have here used the value $\tau = 1$ in (2.1), together with the level k_1 in (2.4). In Figures 2 and 3 we show the simulated patterns of mean value, $E[\cdot]$, and root mean squared error, $RMSE[\cdot]$, of $Q_p^{(j)}(k) := \widehat{\chi}_p^{(j)}(k)/\chi_p$, for $p = 1/n$ and $j = 0$, 3 and 4, based thus on the Hill (H), the Generalized Jackknife (GJ) and the "Maximum Likelihood" (ML), respectively. Results for the other two estimators are not a long way from the ones got for the GJ-estimator and are not pictured. The models underlying the original data are the Fréchet model with $\gamma = 0.25$ and the Burr model with $(\gamma, \rho) = (0.25, -1)$. The Fréchet($\gamma$) d.f. is $F(x) = \exp\left\{-x^{-1/\gamma}\right\}$, $x \geq 0, \quad \gamma > 0$, (for which $\rho = -1$), and the Burr d.f. is $F(x) = 1 - \left(1 + x^{-\rho/\gamma}\right)^{1/\rho}$, $x \geq 0, \quad \gamma > 0, \ \rho < 0$.

As may be seen from these figures we do not gain a lot regarding mean squared errod at the optimal level, but we gain in mean value stability along a wide range of k-values. This gives an obvious indication about the sample

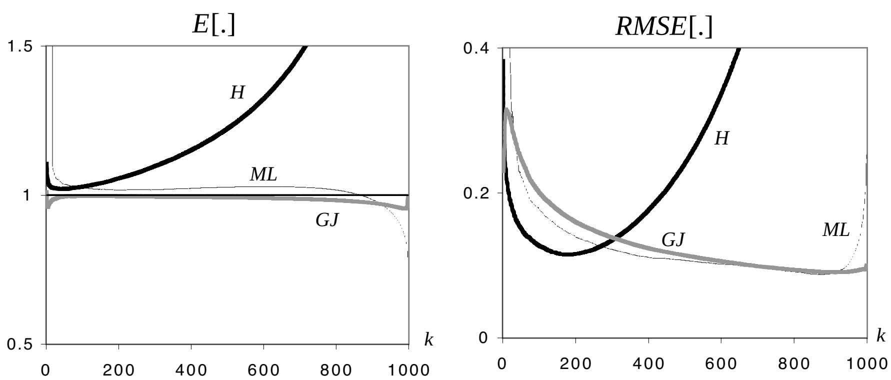

Figure 2: Fréchet parent with $\gamma = 0.25$

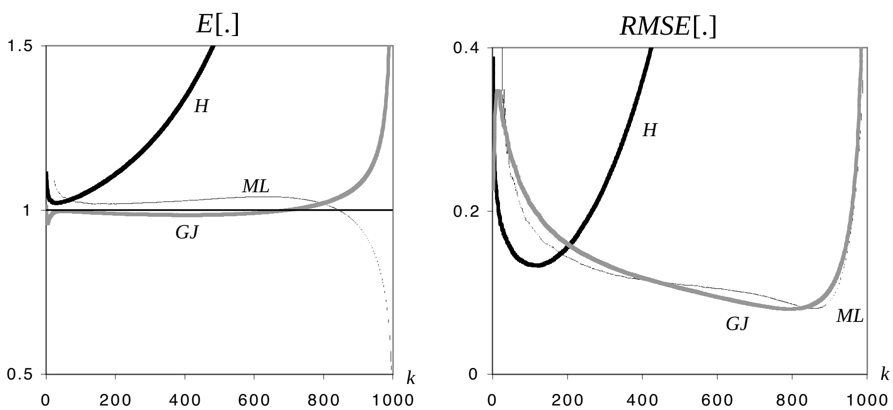

Figure 3: Burr parent with $(\gamma, \rho) = (0.25, -1)$

paths' stability of these new estimators; note that, under a different scale, the sample paths are identical to the corresponding mean values.

Figures 4 and 5 are equivalent to Figures 2 and 3, but now for Burr models with $(\gamma, \rho) = (0.25, -0.5)$ and $(\gamma, \rho) = (0.25, -2)$, respectively.

The behaviour exhibited here, by the "asymptotically unbiased" estimators of high quantiles, is quite different from the one we have got for the tail index, particularly in the region of ρ values close to 0. Indeed, in this region of ρ-values, a region where the "asymptotically unbiased" estimators of the tail index exhibit a highly stable sample path comparatively to the classical estimators, the associated estimators of high quantiles reveal a mean value pattern similar to that of the classical ones, although with a slightly better performance at the optimal level. For small values of ρ, here illustrated with

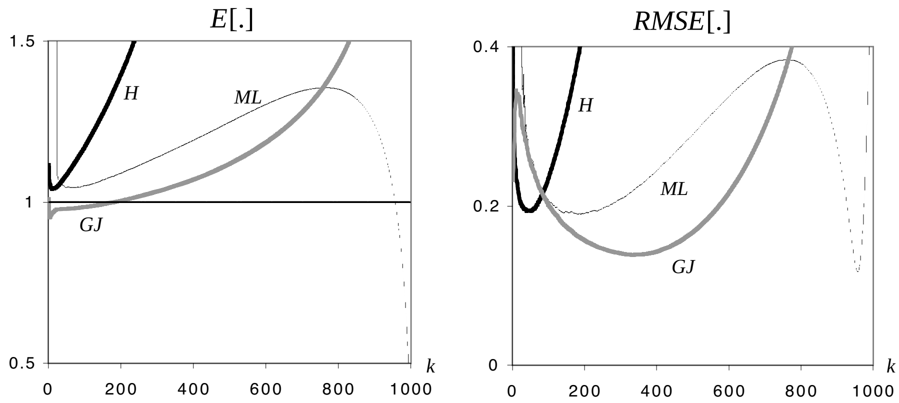

Figure 4: Burr parent with $(\gamma, \rho) = (0.25, -0.5)$

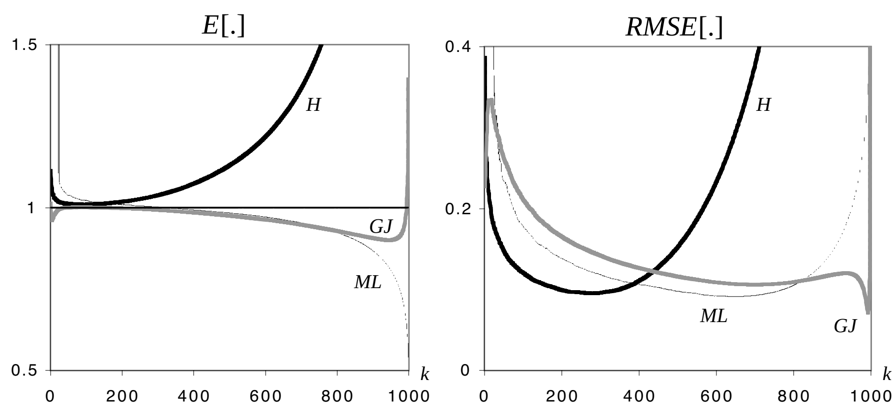

Figure 5: Burr parent with $(\gamma, \rho) = (0.25, -2)$

$\rho = -2$, the patterns of mean values of the new estimators exhibit clearly wider stability regions, but the gain in mean squared error, at the optimal level, is not at all significant.

We shall also present here, for finite values of n, more specifically for $n = 100$, 200, 500, 1000 and 2000, the simulated mean values and mean squared errors of $Q_p^{(j)}$, $j = 0$, 3 and 4, at their optimal levels, i.e. of $Q_{p0}^{(j)} := \widehat{\chi}_p^{(j)}(\widehat{k}_0^{(j)}(n))/\chi_p$, $\widehat{k}_0^{(j)}(n) := \arg\min_k MSE[\widehat{\chi}_n^{(j)}(k)]$, $j = 0$, 3, 4, $p = 1/n$. As said before, the distributional behaviour of $Q_p^{(1)}$ and $Q_p^{(2)}$ is quite similar to that of $Q_p^{(3)} \equiv Q_p^{GJ}$, and has been omitted from the tables. The peculiar behaviour of the mean squared error pattern, exhibited in Figure 4 by the ML-estimator and in Figure 5 by the GJ-estimator, led us to perform the search over the adequate region of k-values in order to obtain the relevant minimum value, the one we think to be underlying the developed theory.

The size $N \times r$ of the multi-sample simulation used, where N denotes the number of runs in each of r replicates of the experiment, is related to the computer time and not to the precision of the estimates. For samples of size $n \leq 500$ we have used a 5000×10 simulation, for $n = 1000$ a 5000×5 simulation and for $n = 2000$ only a 5000×2 simulation was performed. For any details on multi-sample simulation refer to Gomes and Oliveira (2001). Figures from 2 till 5 are based on the first replicate of 5000 runs. Among the estimators considered, and for every n, the one providing the smallest bias and the smallest mean squared error is underlined.

Tables 1 and 2 are related to underlying Fréchet and Burr parents, respectively.

Table 1: Simulated distributional behaviour of Q_{p0}^{H} $(j=0)$, Q_{p0}^{GJ} $(j=3)$ and Q_{p0}^{ML} $(j=4)$, for a Fréchet parent with $\gamma = .25$.

n	100	200	500	1000	2000
Mean values at the optimal level					
$E\left(Q_{p0}^{H}\right)$	1.0636 ±0.0049	1.0602 ±0.0047	1.0537 ±0.0028	1.0480 ±0.0030	1.0402 ±0.0034
$E\left(Q_{p0}^{GJ}\right)$	0.9058 ±0.0043	0.9299 ±0.0018	0.9539 ±0.0008	0.9697 ±0.0015	0.9774 ±0.0016
$E\left(Q_{p0}^{ML}\right)$	0.9508 ±0.0044	0.9719 ±0.0027	0.9897 ±0.0036	0.9957 ±0.0018	0.9991 ±0.0060
Mean squared errors at the optimal level					
$MSE\left(Q_{p0}^{H}\right)$	0.0361 ±0.0010	0.0272 ±0.0007	0.0184 ±0.0003	0.0135 ±0.0002	0.0097 ±0.0001
$MSE\left(Q_{p0}^{GJ}\right)$	0.0405 ±0.0012	0.0268 ±0.0005	0.0145 ±0.0003	0.0083 ±0.0001	0.0067 ±0.0002
$MSE\left(Q_{p0}^{ML}\right)$	0.0363 ±0.0007	0.0242 ±0.0007	0.0130 ±0.0002	0.0075 ±0.0002	0.0059 ±0.0005

In summary we may draw the following conclusions:

1. The new asymptotically unbiased quantile estimators have in general reasonably stable sample paths, which make less troublesome the choice of the optimal level k.

2. For Fréchet models, the ML-estimator is the best one, among the estimators considered, but not a long way from the other asymptotically unbiased estimators. The same happens for a Burr model with $\rho = -1$, provided n is large. For small n $(n < 1000)$ the GJ estimator overpasses the ML estimator, when both are computed at their optimal levels.

Table 2: Simulated distributional behaviour of Q^H_{p0} $(j = 0)$, Q^{GJ}_{p0} $(j = 3)$ and Q^{ML}_{p0} $(j = 4)$, for Burr parents with $\gamma = .25$.

n	100	200	500	1000	2000
$\rho = \mathbf{-0.5}$					
$E\left(Q^H_{p0}\right)$	1.0873 ± 0.0044	1.0966 ± 0.0033	1.1312 ± 0.0025	1.1717 ± 0.0014	1.2162 ± 0.0001
$E\left(Q^{GJ}_{p0}\right)$	$\underline{0.9347}$ ± 0.0039	$\underline{1.0138}$ ± 0.0030	$\underline{1.0379}$ ± 0.0034	$\underline{1.0477}$ ± 0.0023	$\underline{1.0496}$ ± 0.0000
$E\left(Q^{ML}_{p0}\right)$	0.9191 ± 0.0057	0.9687 ± 0.0046	1.0875 ± 0.0044	1.0833 ± 0.0124	1.0780 ± 0.0097
$MSE\left(Q^H_{p0}\right)$	0.0833 ± 0.0018	0.0657 ± 0.0019	0.0544 ± 0.0011	0.0560 ± 0.0007	0.0654 ± 0.0000
$MSE\left(Q^{GJ}_{p0}\right)$	$\underline{0.0284}$ ± 0.0004	$\underline{0.0205}$ ± 0.0004	$\underline{0.0211}$ ± 0.0005	$\underline{0.0199}$ ± 0.0004	$\underline{0.0163}$ ± 0.0000
$MSE\left(Q^{ML}_{p0}\right)$	0.0555 ± 0.0008	0.0350 ± 0.0007	0.0438 ± 0.0008	0.0363 ± 0.0011	0.0279 ± 0.0009
$\rho = \mathbf{-1}$					
$E\left(Q^H_{p0}\right)$	$\underline{1.0694}$ ± 0.0047	1.0674 ± 0.0072	1.0566 ± 0.0037	1.0544 ± 0.0016	1.0521 ± 0.0000
$E\left(Q^{GJ}_{p0}\right)$	0.8987 ± 0.0059	$\underline{0.9688}$ ± 0.0054	$\underline{1.0108}$ ± 0.0008	1.0203 ± 0.0011	1.0166 ± 0.0000
$E\left(Q^{ML}_{p0}\right)$	0.9265 ± 0.0036	0.9503 ± 0.0027	0.9778 ± 0.0011	$\underline{0.9906}$ ± 0.0019	$\underline{0.9929}$ ± 0.0000
$MSE\left(Q^H_{p0}\right)$	0.0464 ± 0.0007	0.0355 ± 0.0005	0.0243 ± 0.0004	0.0180 ± 0.0003	0.0133 ± 0.0000
$MSE\left(Q^{GJ}_{p0}\right)$	$\underline{0.0307}$ ± 0.0004	$\underline{0.0127}$ ± 0.0002	$\underline{0.0072}$ ± 0.0001	0.0065 ± 0.0001	0.0049 ± 0.0000
$MSE\left(Q^{ML}_{p0}\right)$	0.0364 ± 0.0005	0.0229 ± 0.0004	0.0113 ± 0.0001	$\underline{0.0064}$ ± 0.0001	$\underline{0.0044}$ ± 0.0000
$\rho = \mathbf{-2}$					
$E\left(Q^H_{p0}\right)$	$\underline{1.0503}$ ± 0.0025	$\underline{1.0463}$ ± 0.0027	$\underline{1.0388}$ ± 0.0030	$\underline{1.0334}$ ± 0.0019	$\underline{1.0259}$ ± 0.0000
$E\left(Q^{GJ}_{p0}\right)$	0.8991 ± 0.0035	0.9000 ± 0.0116	0.9337 ± 0.0019	0.9455 ± 0.0013	0.9598 ± 0.0034
$E\left(Q^{ML}_{p0}\right)$	0.9418 ± 0.0025	0.9475 ± 0.0033	0.9560 ± 0.0030	0.9647 ± 0.0017	0.9727 ± 0.0000
$MSE\left(Q^H_{p0}\right)$	$\underline{0.02931}$ ± 0.0007	$\underline{0.0211}$ ± 0.0004	0.0131 ± 0.0003	0.0090 ± 0.0002	0.0062 ± 0.0000
$MSE\left(Q^{GJ}_{p0}\right)$	0.0377 ± 0.0005	0.0265 ± 0.0002	0.0165 ± 0.0002	0.0113 ± 0.0003	0.0080 ± 0.0001
$MSE\left(Q^{ML}_{p0}\right)$	0.0314 ± 0.0005	0.0215 ± 0.0003	$\underline{0.0126}$ ± 0.0001	$\underline{0.0083}$ ± 0.0001	$\underline{0.0058}$ ± 0.0000

3. For values of ρ close to zero, here illustrated with $\rho = -0.5$ in a Burr parent, all the new estimators exhibit sample paths similar to the classical ones; however the GJ-estimator is the best one, in the sense of stability around the target value γ and of minimum mean squared error, at its optimal level.

4. For small values of ρ, here illustrated with $\rho = -2$ in a Burr model, and regarding mean squared errors at optimal levels, the ML-estimator is able to overpass slightly the H-estimator for large n. However, the H-estimator of a high quantile reveals a smaller bias at optimal levels.

References

[1] Beirlant, J., G. Diercks, Y.Goegebuur and G.Matthys (1999). Tail index estimation and an exponential regression model. *Extremes* **2**, 177-200.

[2] Caeiro, F. and M. I. Gomes (2002). A class of asymptotically unbiased semi-parametric estimators of the tail index. *Test* **11**:2, 345-364.

[3] Dekkers, A. L. M., J. H. J. Einmahl and L. de Haan (1989). A moment estimator for the index of an extreme-value distribution. *Ann. Statist.* 17, 1833-1855.

[4] Ferreira, A. Haan, L. de and L. Peng (2001). On optimising the estimation of high quantiles of a probability distribution. To appear in *Statistics*.

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

[6] Fraga Alves, M. I., Gomes M. I. and L. de Haan (2001). *A new class of semi-parametric estimators of the second order parameter.* Notas e Comunicações CEAUL 4/2001. To appear in *Portugaliae Mathematica*.

[7] Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics* (2nd edition). Krieger.

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

[9] Gomes, M. I. and F. Caeiro (2002) *Bias reduction of a heavy tail index estimator through an external estimation of the second order parameter.* Notas e Comunicações CEAUL 27/2002. Submitted.

[10] Gomes, M. I. and M. J. Martins (2001). Alternatives to Hill's estimator — asymptotic versus finite sample behaviour. *J. Statist. Planning and Inference* **93**, 161-180.

[11] Gomes, M. I. and M. J. Martins (2002a). Bias reduction and explicit estimation of the tail index. To appear in *J. Statist. Planning and Inference.*

[12] Gomes, M. I. and M. J. Martins (2002b). "Asymptotically unbiased" estimators of the tail index based on external estimation of the second order parameter. *Extremes* **5**:1, 5-31.

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

[14] Gomes, M. I. and O. Oliveira (2000). The bootstrap methodology in Statistical Extremes — choice of the optimal sample fraction. *Extremes* **4**:4, 331-358.

[15] Hann, L. de and H. Rootzén (1993). On the estimation of high quantiles. *J. Statist. Plann. Inference* **35**, 1-13.

[16] Hall, P. and Welsh, A.H. (1985). Adaptive estimates of parameters of regular variation. *Ann. Statist.* **13**, 331-341.

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