REVSTAT – Statistical Journal
Volume 6, Number 1, March 2008, 1–20

MINIMUM-VARIANCE REDUCED-BIAS TAIL INDEX AND HIGH QUANTILE ESTIMATION

Authors: Frederico Caeiro
– F.C.T. (D.M.) and C.M.A.,
Universidade Nova de Lisboa, Portugal
fac@fct.unl.pt

M. Ivette Gomes
– F.C.U.L. (D.E.I.O.) and C.E.A.U.L.,
Universidade de Lisboa, Portugal
ivette.gomes@fc.ul.pt

Abstract:

- Heavy tailed-models are quite useful in many fields, like *insurance*, *finance*, *telecommunications*, *internet traffic*, among others, and it is often necessary to estimate a *high quantile*, i.e., a value that is exceeded with a probability p, small. The semi-parametric estimation of this parameter relies essentially on the estimation of the *tail index*, the primary parameter in *statistics of extremes*. Classical semi-parametric estimators of extreme parameters show usually a severe bias and are known to be very sensitive to the number k of top order statistics used in the estimation. For k small they have a high variance, and for large k a high bias. Recently, new second-order "shape" and "scale" estimators allowed the development of second-order reduced-bias estimators, which are much less sensitive to the choice of k. Here we shall study, under a third order framework, minimum-variance reduced-bias (MVRB) tail index estimators, recently introduced in the literature, and dependent on an adequate estimation of second order parameters. The improvement comes from the asymptotic variance, which is kept equal to the asymptotic variance of the classical Hill estimator, provided that we estimate the second order parameters at a level of a larger order than the level used for the estimation of the first order parameter. The use of those MVRB tail index estimators enables us to introduce new classes of reduced-bias high quantile estimators. These new classes are compared among themselves and with previous ones through the use of a small-scale Monte Carlo simulation.

Key-Words:

- *statistics of extremes; tail index; high quantiles; second-order reduced-bias semi-parametric estimation; third order framework.*

AMS Subject Classification:

- 62G32, 62E20, 65C05.

1. INTRODUCTION

Let $X_1, X_2, ..., X_n$ be a set of n independent and identically distributed (i.i.d.) random variables (r.v.'s), from a population with distribution function (d.f.) F, in the max domain of attraction of G_γ, $\gamma \in \mathbb{R}$, with

$$G_\gamma(x) = \begin{cases} \exp\Big[-(1+\gamma x)^{-\frac{1}{\gamma}}\Big], & 1+\gamma x > 0 \quad \text{if} \quad \gamma \neq 0\,, \\ \exp(-e^{-x})\,, & x \in \mathbb{R} \quad\quad \text{if} \quad \gamma = 0\,. \end{cases}$$

The parameter γ is the *extreme value index* and we then use the notation $F \in D(G_\gamma)$. In this paper we shall work only with heavy-tailed models, i.e., models $F \in D(G_\gamma)$ with $\gamma > 0$. Then γ is often called *tail index*.

Let us define $U(t) := F^{\leftarrow}(1-1/t)$, $t > 1$, with $F^{\leftarrow}(x) := \inf\{y\colon F(y) \geq x\}$ denoting the generalized inverse function of F. We have

$$(1.1) \qquad F \in D(G_\gamma),\ \gamma > 0 \quad \Longleftrightarrow \quad 1-F \in RV_{-1/\gamma} \quad \Longleftrightarrow \quad U \in RV_\gamma$$

(Gnedenko, 1943; de Haan, 1970), where, for any real a, RV_a stands for the class of regularly varying functions at infinity with index of regular variation a, i.e. positive measurable functions g such that $\lim_{t\to\infty} g(tx)/g(t) = x^a$, for all $x > 0$.

We are interested in the estimation of a high quantile, χ_{1-p}, a typical parameter in the most diversified areas of application. Such a quantile is a value exceeded with a small probability p, i.e., such that $F(\chi_{1-p}) = 1-p$. More specifically, we want to extrapolate beyond the sample, and to estimate

$$(1.2) \qquad \chi_{1-p} = U(1/p)\,, \qquad p = p_n \to 0, \quad np_n \to K \quad \text{as } n\to\infty, \quad K \in [0,1]\,.$$

Denoting by $X_{1:n} < ... < X_{n:n}$ the order statistics (o.s.'s) from the original sample, Weissman (1978) proposed, for heavy-tailed models, the following semi-parametric estimator of χ_{1-p} ,

$$(1.3) \qquad Q_{\hat\gamma}^{(p)}(k) := X_{n-k:n}\, c_n^{\hat\gamma}\,, \qquad c_n := \tfrac{k}{np} \to \infty, \quad \text{as} \quad n\to\infty\,,$$

where $\hat\gamma$ is any consistent estimator of γ. For $\gamma \in \mathbb{R}$, we can find semi-parametric high quantile estimators in de Haan and Rootzén (1983), Ferreira *et al.* (2003) and Matthys and Beirlant (2003). As usual in semi-parametric estimation of parameters from extreme value models, we shall assume that $k = k_n$ is an *intermediate* sequence, i.e., a sequence of integer values in $[1, n]$, such that

$$(1.4) \qquad k_n \to \infty\,, \qquad k_n = o(n), \quad \text{as} \quad n\to\infty\,.$$

For heavy tails, the classical tail index estimator, usually the one which is plugged in (1.3) for a semi-parametric quantile estimation, is the Hill estimator $\hat{\gamma} = \hat{\gamma}(k) =: H(k)$ (Hill, 1975),

$$H(k) := \frac{1}{k} \sum_{i=1}^{k} V_{ik} = \frac{1}{k} \sum_{i=1}^{k} U_i \ , \tag{1.5}$$

the average of the log-excesses $V_{ik} := \ln X_{n-i+1:n} - \ln X_{n-k:n}$, $1 \le i \le k < n$, as well as the average of the scaled log-spacings

$$U_i := i\big(\ln X_{n-i+1:n} - \ln X_{n-i:n}\big) \ , \qquad 1 \le i \le k < n \ . \tag{1.6}$$

We thus get the so-called classical quantile estimator, $Q_H^{(p)}(k)$, based on the Hill tail index estimator H. It is known that for intermediate k and if the first order condition (1.1) holds, $H(k)$ and $Q_H^{(p)}(k)$ are consistent for the estimation of γ and χ_{1-p}, respectively. The main problem with these semi-parametric estimators is a high variance for small k, i.e., high thresholds, and a high bias for large k.

To obtain information on the distributional behaviour of these estimators, we shall also assume a second order condition, that measures the rate of convergence of $\ln U(tx) - \ln U(t)$ to $\gamma \ln x$,

$$\begin{aligned} &\lim_{t\to\infty} \frac{\ln U(tx) - \ln U(t) - \gamma \ln x}{A(t)} = \frac{x^\rho - 1}{\rho} \iff \\ &\iff \lim_{t\to\infty} \frac{U(tx)/U(t) - x^\gamma}{A(t)} = x^\gamma \, \frac{x^\rho - 1}{\rho} \ , \end{aligned} \tag{1.7}$$

for all $x > 0$, where $\rho \le 0$ is the shape second order parameter and the function $|A|$ must be of regular variation with index ρ (Geluk and de Haan, 1987). To be able to reduce the bias of these estimators, it is quite useful to assume that we are working in Hall's class of heavy-tailed models (Hall, 1982; Hall and Welsh, 1985) where, with $\gamma > 0$, $\rho < 0$, $C > 0$ and $D_1 \ne 0$,

$$U(t) = C t^\gamma \Big(1 + D_1 t^\rho + o(t^\rho)\Big) \ , \qquad t \to \infty \ . \tag{1.8}$$

Then, the second order condition (1.7) holds with $A(t) = \rho D_1 t^\rho := \gamma \beta t^\rho$.

Proposition 1.1 (de Haan and Peng, 1998)**.** *Under the second order framework in* (1.7)*, and for intermediate k, i.e., whenever* (1.4) *holds, we may guarantee the asymptotic normality of* $H(k)$ *in* (1.5)*. Indeed, we may write,*

$$H(k) \overset{d}{=} \gamma + \frac{\gamma}{\sqrt{k}} Z_k + \frac{A(n/k)}{1-\rho} \big(1 + o_p(1)\big) \ , \tag{1.9}$$

with $Z_k = \sqrt{k} \left(\sum_{i=1}^k E_i/k - 1\right)$, *and* $\{E_i\}$ *i.i.d. standard exponential r.v.'s.*

Consequently, if we choose k such that $\sqrt{k}\,A(n/k) \to \lambda \neq 0$, finite, as $n \to \infty$, $\sqrt{k}\,\big(H(k)-\gamma\big)$ is asymptotically normal, with variance equal to γ^2 and a non-null mean value given by $\lambda/(1-\rho)$.

The result in (1.9) has recently led researchers to consider the possibility of dealing with the asymptotic bias dominant term in an appropriate way, building second-order reduced-bias estimators, discussed by Peng (1998), Beirlant *et al.* (1999), Feuerverger and Hall (1999), Gomes *et al.* (2000), among others. In the above mentioned papers, authors have been able to remove the dominant component of the asymptotic bias, but with an increase of the asymptotic variance. More recently, Gomes *et al.* (2004b), Caeiro *et al.* (2005) and Gomes *et al.* (2007a) proposed minimum-variance reduced-bias (MVRB) estimators, based on an external estimation of second order parameters, built in such way that they were able to reduce the bias without increasing the asymptotic variance, which is kept equal to γ^2, the asymptotic variance of the Hill estimator.

If we look at (1.9), we see that the dominant component of the bias of Hill's estimator is $A(n/k)/(1-\rho) = \gamma\beta\,(n/k)^\rho/\rho$, for models in (1.8). This component can be easily estimated and removed from Hill's estimator, leading to any of the asymptotically equivalent estimators (Caeiro *et al.*, 2005),

$$
\begin{aligned}
\overline{H}_{\hat\beta,\hat\rho}(k) \;&:=\; H(k)\left(1-\frac{\hat\beta}{1-\hat\rho}\left(\frac{n}{k}\right)^{\hat\rho}\right),\\
\overline{\overline{H}}_{\hat\beta,\hat\rho}(k) \;&:=\; H(k)\,\exp\!\left(-\frac{\hat\beta}{1-\hat\rho}\left(\frac{n}{k}\right)^{\hat\rho}\right),
\end{aligned}
\tag{1.10}
$$

where $\hat\rho$ and $\hat\beta$ need to be adequate consistent estimators of the second order parameters ρ and β, if we want to keep the asymptotic variance at γ^2. This requires an external estimation of the second order parameters using a number of top o.s.'s k_1, larger than the number of top o.s.'s, k, used for the tail index estimation, and an estimator $\hat\rho$ of ρ such that $\hat\rho-\rho = o_p(1/\ln n)$.

On the basis of the different papers dealing with high quantile semi-parametric estimation for heavy tails, among which we mention Gomes and Figueiredo (2006) and Caeiro and Gomes (2007), we can state the following result.

Proposition 1.2. *Under the conditions of Proposition 1.1, the validity of* (1.2)*, a known tail index γ and c_n defined in* (1.3),

$$
Q_\gamma^{(p)}(k) \;\stackrel{d}{=}\; \chi_{1-p}\left(1+\frac{\gamma}{\sqrt{k}}\,B_k+\frac{1-c_n^\rho}{\rho}\,A(n/k)\,\big(1+o_p(1)\big)\right),
\tag{1.11}
$$

with B_k an asymptotically standard normal r.v. Consequently, if $\sqrt{k}\,A(n/k) \to \lambda$, finite, $\sqrt{k}\,\big(Q_\gamma^{(p)}(k)/\chi_{1-p}-1\big)$ is asymptotically normal, with variance γ^2 and mean *value λ/ρ. If γ is unknown and is estimated by any consistent estimator $\hat\gamma$,*

$$(1.12)\qquad Q^{(p)}_{\hat\gamma}(k) \stackrel{d}{=} \chi_{1-p}\left(1+(\hat\gamma-\gamma)\ln c_n + \frac{\gamma}{\sqrt{k}}\,B_k + \frac{1-c_n^{\rho}}{\rho}\,A(n/k)\,\big(1+o_p(1)\big)\right).$$

Consequently, if $\sqrt{k}\,A(n/k)\to\lambda$, finite, and $\ln c_n/\sqrt{k}\to 0$, as $n\to\infty$, then $\frac{\sqrt{k}}{\ln c_n}\big(Q^{(p)}_{\hat\gamma}(k)/\chi_{1-p}-1\big)$ has asymptotically the same distribution as $\sqrt{k}\,(\hat\gamma-\gamma)$.

From (1.12) it is obvious that the behaviour of $\hat\gamma$ rules strongly the behaviour of $Q^{(p)}_{\hat\gamma}$. The summand $(1-c_n^{\rho})\,A(n/k)/\rho$, asymptotically equivalent to $A(n/k)/\rho$ and the dominant component of the bias of $Q^{(p)}_{\gamma}$ in (1.11), does not influence the limiting distribution of $Q^{(p)}_{\hat\gamma}$. But, as already noticed in Matthys *et al.* (2004), the removal of this term for finite samples, typically leads to an improvement in the overall stability of the quantile estimates as a function of k. Since $\chi_{1-p}/X_{n-k:n} \stackrel{p}{\sim} c_n^{\gamma}\big(1+(c_n^{\rho}-1)\,A(n/k)/\rho\big)$, we shall consider the new estimators,

$$(1.13)\qquad \overline{Q}^{(p)}_{\hat\gamma}(k) \;=\; \overline{Q}^{(p)}_{\hat\gamma}(k;\hat\beta,\hat\rho) \;:=\; X_{n-k:n}\; c_n^{\hat\gamma}\left(1+\hat\gamma\,\hat\beta\left(\frac{n}{k}\right)^{\hat\rho}\frac{c_n^{\hat\rho}-1}{\hat\rho}\right),$$

asymptotically equivalent, up to the second order, to the estimators already proposed before by Matthys *et al.* (2004), Beirlant *et al.* (2006) and Gomes and Pestana (2007b),

$$(1.14)\qquad \overline{\overline{Q}}^{(p)}_{\hat\gamma}(k) \;=\; \overline{\overline{Q}}^{(p)}_{\hat\gamma}(k;\hat\beta,\hat\rho) \;:=\; X_{n-k:n}\; c_n^{\hat\gamma}\,\exp\left(\hat\gamma\,\hat\beta\left(\frac{n}{k}\right)^{\hat\rho}\frac{c_n^{\hat\rho}-1}{\hat\rho}\right).$$

We shall replace $\hat\gamma$ by any of the MVRB estimators $\overline{H}(k)=\overline{H}_{\hat\beta,\hat\rho}(k)$ and $\overline{\overline{H}}(k)=\overline{\overline{H}}_{\hat\beta,\hat\rho}(k)$, generally denoted by $\widetilde{H}(k)$, with $\overline{H}_{\hat\beta,\hat\rho}(k)$ and $\overline{\overline{H}}_{\hat\beta,\hat\rho}(k)$ given in (1.10).

Remark 1.1. Since $c_n^{\rho}\ln c_n = o(1)$, the asymptotic behavior of (1.13) and (1.14) does not change if we replace $c_n^{\hat\rho}$ by 0. In the simulation study, we did not notice any change in the performance of the estimators with this replacement. Anyway, we shall keep working with the quantile estimators defined in (1.13).

Is section 2, and assuming a third order framework in order to get full information on the leading terms of asymptotic bias, we study the tail index estimators $\widetilde{H}(k)$ in (1.10), as well as $\overline{Q}^{(p)}_{\widetilde{H}}$, with $\overline{Q}^{(p)}_{\hat\gamma}$ given in (1.13). In Section 3, a small-scale simulation study helps us to identify the behaviour of the quantile estimators in (1.13) for finite samples. Finally, in Section 4, we draw a short final conclusion.

2. ASYMPTOTIC PROPERTIES

2.1. Third order framework

In order to derive the asymptotic bias of the MVRB estimators under study, we shall work with a sub-class of Hall's class such that

$$U(t) \;=\; C\,t^\gamma \Big(1 + D_1\,t^\rho + D_2\,t^{\rho+\rho^*} + o\big(t^{\rho+\rho^*}\big)\Big)\,, \qquad t\to\infty\;, \tag{2.1}$$

$C>0,\ D_1\neq 0,\ \rho<0,\ \rho^*<0$. Note that, compared to Hall's class in (1.8) we merely specify the summand $o(t^\rho)$. Note also that, with $h_\theta(x) := (x^\theta - 1)/\theta$, $\theta<0$, $A(t) = \rho\, D_1 t^\rho = \gamma\beta\, t^\rho,\ \rho' = \max(\rho,\rho^*) \geq \rho$ and

$$B(t) \;=\; \beta'\,t^{\rho'} = \begin{cases} \big((1+\rho^*/\rho)\,D_2/D_1\big)\,t^{\rho^*}, & \rho<\rho^*\,,\\ \big(2\,D_2/D_1 - D_1\big)\,t^\rho, & \rho=\rho^*\,,\\ -D_1\,t^\rho, & \rho>\rho^* \ \text{ or } \ D_2=0\;, \end{cases}$$

we may write for any $x>0$,

$$\ln\frac{U(tx)}{U(t)} - \gamma\ln x \;=\; A(t)\,h_\rho(x) + A(t)\,B(t)\,h_{\rho+\rho'}(x)\,\big(1+o(1)\big)\;, \tag{2.2}$$

which is, for arbitrary ρ and ρ', the third order condition used in the paper by Gomes *et al.* (2004a), equivalent to the ones assumed in Gomes *et al.* (2002) and Fraga Alves *et al.* (2003). As mentioned before, we shall essentially consider the validity of (2.1), which is equivalent to consider that (2.2) holds with $\rho\leq\rho'$ and $A(t) = \alpha\,t^\rho$ for some real α.

Remark 2.1. The class in (2.1) contains most of the heavy-tailed models used in applications, like the *Fréchet*, with $U(t) = \big(\ln(t/(t-1))\big)^{-\gamma}$, the *Burr*, with $U(t) = (t^{-\rho}-1)^{-\gamma/\rho},\ t>1$, the *Generalized Pareto* (*GP*), with $U(t) = (t^\gamma - 1)/\gamma,\ t>1$, and the *Student*'s-$t_\nu$, $\nu>0$, with d.f.

$$F(x) \;=\; F(x\,|\,\nu) \;=\; \frac{\Gamma\big((\nu+1)/2\big)}{\Gamma(\nu/2)\,\sqrt{\pi\nu}} \int_{-\infty}^x (1+z^2/\nu)^{-(\nu+1)/2}\,dz\,, \qquad x\in\mathbb{R},\quad \nu>0\;.$$

Although $\rho^* = \rho' = \rho$ for all these classical models, we have decided to work with a slight more general condition, the one in (2.1). Indeed, it is not so hard to find examples where $\rho'\neq\rho$. Gomes and Oliveira (2003) noticed that shifting the data can change the asymptotic behavior of the tail and the value of the second order parameters, i.e., if X is our original parent, and $Y = X + a$, then $U_Y(t) = U_X(t) + a$, and consequentially,

$$U_Y(t) \;=\; C\,t^\gamma \Big(1 + D_1\,t^\rho + a\,t^{-\gamma}/C + D_2\,t^{\rho+\rho^*} + o\big(t^{\rho+\rho^*}\big)\Big)\,, \qquad t\to\infty\;.$$

In Table 1 we present, for the above mentioned models, the values of the first, second and third order parameters in (2.1) and the values of β and β' in $A(t) = \gamma\beta\, t^\rho$ and $B(t) = \beta'\, t^\rho$. In this table, $c_\nu = (\nu\, \mathcal{B}(\nu/2, 1/2))^{1/\nu}$ ($c_1 = \pi$ leading to the usually called Cauchy d.f.), where $\mathcal{B}$ is the complete Beta function.

Table 1: Study of some distributions in Hall's class.

Distribution	C	D_1	D_2	γ	ρ	ρ^*	β	β'
Fréchet	1	$-\frac{\gamma}{2}$	$-\frac{\gamma}{12}$	γ	-1	-1	$\frac{1}{2}$	$\frac{5}{6}$
Burr	1	$\frac{\gamma}{\rho}$	$\frac{\gamma(\rho+\gamma)}{2\rho^2}$	γ	ρ	ρ	1	1
GP	$\frac{1}{\gamma}$	-1	0	γ	$-\gamma$	$-\gamma$	1	1
Student's t_ν	$\sqrt{v}\, c_\nu^{-1}$	$-\frac{(\nu+1)\, c_\nu^2}{2(\nu+2)}$	$-\frac{\nu\,(\nu+1)\,(\nu+3)\, c_\nu^4}{8\,(\nu+2)^2\,(\nu+4)}$	$\frac{1}{\nu}$	$-\frac{2}{\nu}$	$-\frac{2}{\nu}$	$\frac{(\nu+1)\, c_\nu^2}{\nu+2}$	$\frac{(\nu^2+4\nu+2)\, c_\nu^2}{(\nu+2)\,(\nu+4)}$

2.2. Estimation of second order parameters

The reduced-bias tail index and quantile estimators require the estimation of the second order parameters ρ and β, which will be now briefly discussed.

2.2.1. Estimation of the shape second order parameter ρ

We shall consider here particular members of the class of estimators of the second order parameter ρ proposed by Fraga Alves *et al.* (2003), but parameterized by a tuning real parameter τ (see Caeiro and Gomes, 2006). Denoting $M_n^{(j)}(k) := \frac{1}{k}\sum_{i=1}^k V_{ik}^j$ the j-moment of the log-excesses, $j=1,2,3$, these ρ-estimators depend on the statistics

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

which converge towards $3(1-\rho)/(3-\rho)$ for any real τ, whenever the second order condition (1.7) holds, k is such that (1.4) holds and $\sqrt{k}\,A(n/k)\to\infty$, as $n\to\infty$. The ρ-estimators considered have the functional expression,

$$\hat\rho_\tau(k) \;=\; \hat\rho(k;\tau) \;:=\; -\min\Big(0,\; 3\big(T_n^{(\tau)}(k)-1\big)\,/\,\big(T_n^{(\tau)}(k)-3\big)\Big)\;. \tag{2.3}$$

Proposition 2.1 (Fraga Alves *et al.*, 2003)**.** *If the second order condition* (1.7) *holds, with* $\rho<0$, (1.4) *holds and* $\sqrt{k}\,A(n/k)\to\infty$, *then* $\hat\rho(k;\tau)$ *in* (2.3) *converge in probability to* ρ, *as* $n\to\infty$. *Under the third order framework in* (2.2),

$$\hat\rho(k;\tau) \;\stackrel{d}{=}\; \rho+\left(\frac{\gamma\,\sigma_\rho W_k^\rho}{\sqrt{k}\,A(n/k)}+\upsilon_1 A(n/k)+\upsilon_2 B(n/k)\right)\big(1+o_p(1)\big)\;, \tag{2.4}$$

where W_k^ρ *is an asymptotically standard normal r.v.,* $\sigma_\rho=\frac{(1-\rho)^3}{\rho}\sqrt{(2\rho^2-2\rho+1)}$,

$$\upsilon_1 \;\equiv\; \upsilon_1(\gamma,\rho,\tau) \;=\; \frac{\rho\Big[\tau\,(1-2\rho)^2\,(3-\rho)\,(3-2\rho)\,+\,6\rho\,\big(4\,(2-\rho)\,(1-\rho)^2-1\big)\Big]}{12\,\gamma\,(1-\rho)^2\,(1-2\rho)^2}\;,$$

$$\upsilon_2 \;=\; \frac{\rho'(\rho+\rho')(1-\rho)^3}{\rho(1-\rho-\rho')^3}\;.$$

Consequently, if $\sqrt{k}\,A^2(n/k)\to\lambda_A$ *and* $\sqrt{k}\,A(n/k)\,B(n/k)\to\lambda_B$, *finite, then* $\sqrt{k}\,A(n/k)\,(\hat\rho(k;\tau)-\rho)\;\stackrel{d}{\longrightarrow}\;N(\lambda_A\upsilon_1+\lambda_B\upsilon_2,\,\gamma^2\sigma_\rho^2)$.

Corollary 2.1. *Under the third order framework in* (2.1), *if* (1.4) *holds,* $\sqrt{k}\,A(n/k)\to\infty$ *and* $\sqrt{k}\,A(n/k)\,B(n/k)\to\lambda_B$, *finite, then* $\hat\rho_n(k;\tau)-\rho=O_p\big(1/(\sqrt{k}\,A(n/k))\big)$. *But, if we chose* k *such that* $\sqrt{k}\,A(n/k)\,B(n/k)\to\infty$, *then* $1/(\sqrt{k}\,A(n/k))=o(B(n/k))$ *and* $\hat\rho_n(k;\tau)-\rho=O_p(B(n/k))$.

A comment on the choice of the tuning parameter τ. From Proposition 2.1, we can conclude that the tuning parameter τ only affects $\hat\rho(k;\tau)$ asymptotic bias. If $\rho'=\rho$, and consequentially $B(n/k)=O(A(n/k))$, we can always choose $\tau=\tau_0$ so that the asymptotic bias $\upsilon_1 A(n/k)+\upsilon_2 B(n/k)$ in (2.4) is null, even when $\sqrt{k}\,A^2(n/k)\to\lambda_A>0$ and $\sqrt{k}\,A(n/k)\,B(n/k)\to\lambda_B\neq 0$. It is enough to choose the value τ_0 which is the solution of $\upsilon_1\gamma\,\beta+\upsilon_2\,\beta'=0$. Such a value is independent of γ and, with $\xi=\beta'/\beta$, is given by

$$\tau_0 \;\equiv\; \tau_0(\rho,\xi) \;=\; \frac{-6\,\Big[4\xi\,(1-\rho^5)+\rho\,(1-2\rho)\,\big(4\,(2-\rho)\,(1-\rho)^2-1\big)\Big]}{(1-2\rho)^3\,(3-\rho)\,(3-2\rho)}\;. \tag{2.5}$$

Although τ_0, as a function of ρ, is not always monotone, it converges to $3(1-\xi/2)$, as $\rho\to-\infty$ and to $-8\xi/3$, as $\rho\to 0$.

Using the available values ρ, β and β', from Table 1, we have for the *Fréchet* model, $\rho = -1$, $\xi = 5/3$ and $\tau_0 = -217/270 \simeq -0.8$. For models like the *Burr* and the *GP*, where $\beta' = \beta$ and consequently $\xi = 1$, we present in Figure 1 (left) $\tau_0(\rho,1)$ as function of ρ. For *Student*'s t_ν distribution, ρ, β and β' are functions of ν, and the value τ_0 in (2.5) can also be written as a function of ν:

$$\tau_0(\nu) = \frac{12\left(384 + 1216\,\nu + 1440\,\nu^2 + 720\,\nu^3 + 72\,\nu^4 - 61\,\nu^5 - 21\,\nu^6 - 2\,\nu^7\right)}{(1+\nu)\,(4+\nu)^4\,(2+3\nu)\,(4+3\nu)} \,.$$

This function $\tau_0(\nu)$ is shown in Figure 1 (right), as a function of ν.

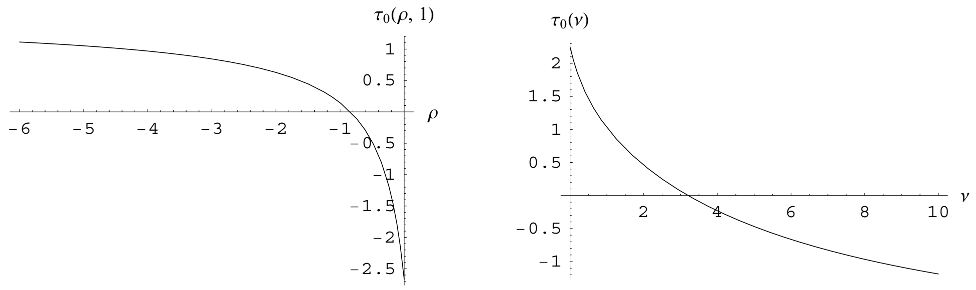

Figure 1: **Left:** $\tau_0(\rho,1)$ as function of ρ. **Right:** $\tau_0(\nu)$ for Student's t_ν.

As an example, for the $GP(\gamma = 0.5)$, we have $\tau_0(-0.5,1) = -213/448 \simeq -0.48$. In Figure 2 and to illustrate the comment above, we picture a sample path of $\hat\rho(k;\tau)$ with $\tau = \tau_0$ and $\tau = 0$, the value of τ most commonly suggested for models with $|\rho| < 1$. We conclude that $\hat\rho_{\tau_0}(k) = \hat\rho(k;\tau_0)$ is indeed more stable than $\hat\rho_0(k) = \hat\rho(k;0)$ around the true value $\rho = -0.5$.

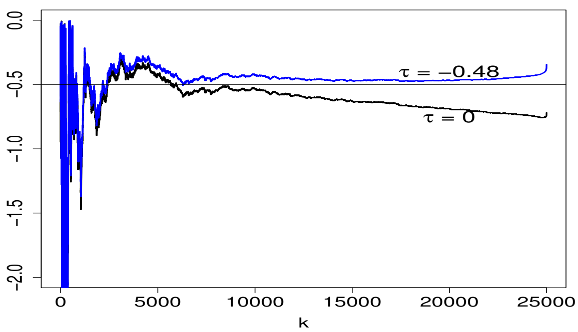

Figure 2: Sample path of the estimator $\widehat\rho(k;\tau)$, $\tau = 0, -0.48$, for one sample of size $n = 25000$ from the GP distribution with $\gamma = 0.5$.

Remark 2.2. Indeed, for an appropriate tuning parameter τ the ρ-estimators in (2.3) show highly stable sample paths as functions of k, the number of top o.s. used, for a wide range of large k-values. The theoretical and simulated results in Fraga Alves *et al.* (2003), together with the use of these estimators in different reduced-bias statistics, has led to advise in practice the estimation of ρ through the estimator in (2.3), computed at the value

$$k_1 := \left[n^{0.995}\right] , \tag{2.6}$$

not chosen in any optimal way, and the choice of the tuning parameter $\tau = 0$ for $\rho \in [-1, 0)$ and $\tau = 1$ for $\rho \in (-\infty, -1)$. As usual, $[x]$ denotes the integer part of x. However, practitioners should not choose blindly the value of τ in (2.3), and as pointed out in Caeiro and Gomes (2006), even negative values of τ should be possible candidates. It is indeed sensible to draw a few sample paths of $\hat\rho_\tau(k) = \hat\rho(k;\tau)$, as functions of k, electing the value of τ which provides the highest stability for large k, by means of any stability criterion, like the one suggested in Gomes *et al.* (2005) or Gomes and Pestana (2007a). For not too small n, we are frequently led to the above mentioned choice: $\hat\rho_0$ if $\rho \geq -1$ and $\hat\rho_1$ if $\rho < -1$, when we consider only the tuning parameters $\tau = 0$ and $\tau = 1$ as the possible alternatives. In practice, the adequate choice of τ is much more crucial than the choice of k_1, discussed in the following.

A few comments on the choice of the level k_1 for the ρ-estimation. On the basis of the results in Proposition 1.1 and Proposition 2.1, it seems sensible to estimate the second order ρ using a number k_1 of o.s.'s of a larger order than k, the number of o.s.'s used for the estimation of the tail index γ. We now make the following comments on the choice of the value k_1 that should be used for the estimation of the second order parameter ρ.

(**1**) The ideal situation would perhaps be the choice of an "optimal" k_1 for the estimation of ρ, in the sense of a value that enables the asymptotic normality of the ρ-estimator with a non-null asymptotic bias. For models in (2.1), k_1 is then such that $\sqrt{k_1}\, A(n/k_1)\, B(n/k_1) \to \lambda_{B_1}$, finite and non-null. We then get $k_1 = O\big(n^{-2(\rho+\rho')/(1-2(\rho+\rho'))}\big)$. Denoting $\hat\rho = \hat\rho(k_1;\tau)$ for any $\hat\rho(k;\tau)$ in (2.3), $\hat\rho - \rho$ is of the order of $1/\big(\sqrt{k_1}\, A(n/k_1)\big) = O\big(n^{\rho'/(1-2(\rho+\rho'))}\big) = o(1/\ln n)$, i.e.,

$$\hat\rho - \rho = o_p(1/\ln n), \qquad \text{as} \quad n \to \infty , \tag{2.7}$$

a condition needed later on. In practice, such a k_1 has only a "limited" interest at the current state-of-the-art. It is however of theoretical interest.

(**2**) Assume next the validity of the following condition:

Condition U: There exist a *tuning parameter* τ^* and a level k_1, with $\sqrt{k_1}\, A(n/k_1)\, B(n/k_1) \to \infty$, such that, with $\hat\rho(k;\tau)$ defined in (2.3), $\hat\rho^* - \rho = \hat\rho(k_1;\tau^*) - \rho = O_p\big(1/(\sqrt{k_1}\, A(n/k_1))\big)$.

This is obviously a strong assumption, practically equivalent to saying that for any specific model there is a τ^* and a k_1 such that $\hat\rho^* = \hat\rho(k_1;\tau^*)$ is an unbiased estimator for ρ, so that the bias has no influence in the rate of convergence, which is kept at $1/\big(\sqrt{k_1}\,A(n/k_1)\big)$. Indeed, such a claim is made on the basis of the high stability of sample paths of the ρ-estimates in (2.3) for a specific $\tau = \tau^*$ and large values of k (see Figure 2 and the comment made above on the choice of τ). Then, the use of a value k_1 larger than the so-called "optimal" level in item 1., but intermediate, like for instance, the one suggested in Gomes and Martins (2002),

$$k_1 := \min\big(n-1,\; 2n/\ln\ln n\big)\;, \tag{2.8}$$

enables us to guarantee that $\hat\rho^* - \rho = o_p(1/\ln n)$. Indeed, if we assume the validity of *Condition U* for k_1 in (2.8), we get $\hat\rho^* - \rho = O_p\big(1/(\sqrt{k_1}\,A(n/k_1))\big) = O_p\big((\ln\ln n)^{(1-2\rho)/2}/\sqrt{n}\big)$, which is obviously of smaller order than $\{1/\ln n\}$, i.e., (2.7) holds. This will be the unique situation under which we may work with the k_1 suggested in Gomes and Martins (2002), i.e, the one in (2.8), and still guarantee the above mentioned property on the ρ-estimator, and a possible generalization of the third-order results derived for $\widetilde H_{\beta,\rho}$ to $\widetilde H_{\hat\beta^*,\hat\rho^*}$, with $\hat\beta^*$ an adequate β-estimator, to be specified later on, in Section 2.2.2.

(3) If we consider a level k_1 of the order of $n^{1-\epsilon}$, for some small $\epsilon > 0$, we may also guarantee that (2.7) holds for a large class of models, without the need to assume a condition as strong as *Condition U*. This is the reason why, such as done in Caeiro *et al.* (2004b), Gomes and Pestana (2007a,b) and Gomes *et al.* (2004b, 2007a), we advise in practice, as a compromise between theoretical and practical considerations, the use of an intermediate level like the one in (2.6) or any other level $k_1 = [n^{1-\epsilon}]$ for some $\epsilon > 0$, small.

2.2.2. Estimation of the scale second order parameter β

Let us introduce the notation $N_n^{(\alpha)}(k) := \frac{1}{k}\sum_{i=1}^{k}\big(\frac{i}{k}\big)^{\alpha-1}U_i$, with U_i defined in (1.6). For the estimation of β we shall here consider the estimator in Gomes and Martins (2002), with the functional expression,

$$\hat\beta_{\hat\rho}(k) \;=\; \hat\beta(k;\hat\rho) \;:=\; \Big(\frac{k}{n}\Big)^{\hat\rho}\,\frac{\Big(\frac{1}{k}\sum\limits_{i=1}^{k}\big(\frac{i}{k}\big)^{-\hat\rho}\Big)\,N_n^{(1)}(k) \;-\; N_n^{(1-\hat\rho)}(k)}{\Big(\frac{1}{k}\sum\limits_{i=1}^{k}\big(\frac{i}{k}\big)^{-\hat\rho}\Big)\,N_n^{(1-\hat\rho)}(k) \;-\; N_n^{(1-2\hat\rho)}(k)}\;. \tag{2.9}$$

Theorem 2.1 (Gomes *et al.*, 2004b)**.** *If the second order condition* (1.7) *holds, with* $A(t) = \gamma\,\beta\,t^\rho$, $\rho < 0$, *if* (1.4) *holds, and if* $\sqrt{k}\,A(n/k) \to \infty$, *then, with* $\hat\rho_n(k;\tau)$ *and* $\hat\beta_{\hat\rho}(k)$ *given in* (2.3) *and* (2.9)*, respectively, and* $\hat\rho = \hat\rho_n(k;\tau)$ *such that* (2.7) *holds, i.e.,* $\hat\rho - \rho = o_p(1/\ln n)$*, as* $n \to \infty$, $\hat\beta_{\hat\rho}(k)$ *is consistent for the estimation of* β*. Moreover,*

$$\hat\beta_{\hat\rho}(k) - \beta \ \overset{p}{\sim}\ -\beta\,\ln(n/k)\,(\hat\rho-\rho) \ = \ o_p(1)\ . \tag{2.10}$$

2.3. Asymptotic properties of the tail index estimators, under a third order framework

We shall study now the asymptotic behaviour, under a third order framework, of the MVRB estimators $\overline{H}$ and $\overline{\overline{H}}$, generally denoted $\widetilde{H}$. We assume first that we know the two second order parameters β and ρ. Next we estimate both second-order parameters externally at a level k_1 of a larger order than the level k at which we compute the tail index.

Theorem 2.2.

(a) *Under the second order framework in* (1.8)*, and for intermediate* k*, i.e., whenever* (1.4) *holds, we may write,*

$$\widetilde{H}_{\beta,\rho}(k) \ \overset{d}{=}\ \gamma + \frac{\gamma}{\sqrt{k}}\,Z_k + o_p\big(A(n/k)\big)\ , \tag{2.11}$$

where Z_k *is the asymptotically standard normal r.v. in* (1.9)*. Also, if we choose* k *such that* $\sqrt{k}\,A(n/k) \to \lambda$*, finite, as* $n \to \infty$, $\sqrt{k}\,\big(\widetilde{H}_{\beta,\rho}(k) - \gamma\big)$ *are asymptotically normal, with variance* γ^2 *and a null mean value, even if* $\lambda \neq 0$*.*

(b) *If we further assume* (2.1)*, more information can be given for the term* $o_p(A(n/k))$*, and we get the asymptotic distributional representations:*

$$\overline{H}_{\beta,\rho}(k) \ \overset{d}{=}\ \gamma + \frac{\gamma}{\sqrt{k}}\,Z_k^* + \frac{A(n/k)\,B(n/k)}{1-\rho-\rho'}\left(1 - \frac{(1-\rho-\rho')\,A(n/k)}{\gamma\,(1-\rho)^2\,B(n/k)}\right)\big(1+o_p(1)\big)\ , \tag{2.12}$$

and

$$\overline{\overline{H}}_{\beta,\rho}(k) \ \overset{d}{=}\ \gamma + \frac{\gamma}{\sqrt{k}}\,Z_k^* + \frac{A(n/k)\,B(n/k)}{1-\rho-\rho'}\left(1 - \frac{(1-\rho-\rho')\,A(n/k)}{2\gamma\,(1-\rho)^2\,B(n/k)}\right)\big(1+o_p(1)\big)\ , \tag{2.13}$$

with Z_k^* *asymptotically standard normal. If* $\sqrt{k}\,A(n/k)\,B(n/k) \to \lambda_{_B}$*, finite (and then,* $\sqrt{k}\,A^2(n/k) \to \lambda_{_A}$*, also finite),* $\sqrt{k}\big(\overline{H}_{\beta,\rho}(k) - \gamma\big)$ *and* $\sqrt{k}\big(\overline{\overline{H}}_{\beta,\rho}(k) - \gamma\big)$ *are asymptotically normal with the same variance, equal to* γ^2*, and asymptotic bias* $b_{\overline{H}} = \lambda_{_B}/(1-\rho-\rho') - \lambda_{_A}/(\gamma(1-\rho)^2)$ *and* $b_{\overline{\overline{H}}} = \lambda_{_B}/(1-\rho-\rho') - \lambda_{_A}/(2\gamma\,(1-\rho)^2)$*, respectively.*

Proof: The first part of the theorem has been proved in Caeiro *et al.* (2005). Regarding the second part: from the third order set-up in (2.2), we get

$$H(k) \overset{d}{=} \gamma + \frac{\gamma}{\sqrt{k}} Z_k + \frac{A(n/k)}{1-\rho} + O_p\left(\frac{A(n/k)}{\sqrt{k}}\right) + \frac{A(n/k)\,B(n/k)}{1-\rho-\rho'} \big(1+o_p(1)\big) .$$

Consequently, as $\overline{H}_{\beta,\rho}(k) = H(k) \times \big(1 - A(n/k)/(\gamma(1-\rho))\big)$ for models in (2.1),

$$\overline{H}_{\beta,\rho}(k) \overset{d}{=} \gamma + \frac{\gamma}{\sqrt{k}} Z_k + \left(\frac{A(n/k)\,B(n/k)}{1-\rho-\rho'} - \frac{A^2(n/k)}{\gamma(1-\rho)^2} + O_p\left(\frac{A(n/k)}{\sqrt{k}}\right)\right) \big(1+o_p(1)\big) ,$$

$\overline{\overline{H}}_{\beta,\rho}(k) - \overline{H}_{\beta,\rho}(k) \overset{p}{\sim} A^2(n/k)/(2\gamma(1-\rho)^2)$, and the results in the theorem follow. Note that since $\sqrt{k}\, O_p\big(A(n/k)/\sqrt{k}\big) \to 0$, for the intermediate levels k considered, the term $O_p\big(A(n/k)/\sqrt{k}\big)$ is irrelevant for the asymptotic bias. □

Remark 2.3. Notice that $\overline{H}$ and $\overline{\overline{H}}$ have the same asymptotic variance and $b_{\overline{\overline{H}}} = b_{\overline{H}} + \lambda_A/(2\gamma(1-\rho)^2)$, with $\lambda_A \ge 0$. So if both bias are positive, $\overline{H}$ should have, asymptotically, a better performance than $\overline{\overline{H}}$.

Theorem 2.3.

(a) *Under the initial conditions of Theorem 2.2, let us consider the tail index estimators $\widetilde{H}_{\hat\beta,\hat\rho}$ with $\hat\beta$ and $\hat\rho$ consistent for the estimation of β and ρ, respectively, both computed at the level k_1 of a larger order than the level k at which we compute the tail index, and such that (2.7) holds. Then $\sqrt{k}\,\big(\widetilde{H}_{\hat\beta,\hat\rho}(k) - \gamma\big)$ are asymptotically normal, with variance equal to γ^2 and a null mean value, even if $\sqrt{k}\,A(n/k) \to \lambda \ne 0$, as $n \to \infty$.*

(b) *If we work under the third order framework in (2.1), consider $\hat\beta_{\hat\rho}(k)$ in (2.9), $\hat\beta = \hat\beta_{\hat\rho}(k_1)$, and choose k such that $\sqrt{k}\,A(n/k) \to \infty$, but $\sqrt{k}\,A(n/k)\,B(n/k) \to \lambda_B$, finite, then $\sqrt{k}\,\big(\overline{H}_{\hat\beta,\hat\rho}(k) - \gamma\big)$ and $\sqrt{k}\,\big(\overline{\overline{H}}_{\hat\beta,\hat\rho}(k) - \gamma\big)$ are asymptotically normal with variance γ^2 and asymptotic bias $b_{\overline{H}}$ and $b_{\overline{\overline{H}}}$, respectively, given in Theorem 2.2, provided that we can guarantee that $(\hat\rho-\rho)\ln n = o_p\big(1/\sqrt{k}\,A(n/k)\big)$. This last condition on $\hat\rho$ holds if we further assume the validity of Condition U for k_1 in (2.8).*

Proof: If we estimate consistently β and ρ through $\hat\beta$ and $\hat\rho$ under the conditions of the theorem, we may use Taylor's expansion series, and as $\partial\widetilde{H}_{\beta,\rho}/\partial\beta \overset{p}{\sim} A(n/k)/(\beta(1-\rho))$, $\partial\widetilde{H}_{\beta,\rho}/\partial\rho \overset{p}{\sim} -A(n/k)\,\big(\ln(n/k) + 1/(1-\rho)\big)/(1-\rho)$, we get

$$\widetilde{H}_{\hat\beta,\hat\rho}(k) - \widetilde{H}_{\beta,\rho}(k) \overset{p}{\sim} -\frac{A(n/k)}{1-\rho}\left\{\frac{\hat\beta-\beta}{\beta} + (\hat\rho-\rho)\left[\ln(n/k) + \frac{1}{1-\rho}\right]\right\} . \tag{2.14}$$

The first part of the theorem, related to levels k such that $\sqrt{k}\,A(n/k) \to \lambda$, finite, follows thus straightforwardly from (2.14).

Next, from (2.10), $(\hat\beta-\beta)/\beta \overset{p}{\sim} -\ln(n/k_1)\,(\hat\rho-\rho) = o_p\big(1/(\sqrt{k}\,A(n/k))\big)$, $\sqrt{k}\,\big(\widetilde H_{\hat\beta,\hat\rho}(k)-\widetilde H_{\beta,\rho}(k)\big) = o_p\big(1/\sqrt{k}\big)$ and the stated asymptotic normality of $\widetilde H_{\hat\beta,\hat\rho}$ follows as well. We may further write

$$\widetilde H_{\hat\beta,\hat\rho}(k)-\widetilde H_{\beta,\rho}(k) \overset{p}{\sim} -\frac{A(n/k)}{1-\rho}\,(\hat\rho-\rho)\left(\ln(k/k_1)+\frac{1}{1-\rho}\right). \tag{2.15}$$

If we assume the validity of *Condition U* for the level k_1 in (2.8) and consider $\widetilde H_{\hat\beta^*,\hat\rho^*}$, we straighforwardly guarantee that $\sqrt{k}\,(\hat\rho^*-\rho)\,A(n/k)\ln(k/k_1) = o_p(1)$. Consequently, the use of (2.15), with $(\hat\beta,\hat\rho)$ replaced by $(\hat\beta^*,\hat\rho^*)$, enables us to get the results in the theorem. □

2.4. Asymptotic properties of the reduced-bias quantile estimators, under a third order framework

We shall provide in theorems 2.4 and 2.5 the distributional behaviour of the quantile estimators under study, for models in (2.1).

Theorem 2.4. *Under the third order framework in* (2.1)*, for intermediate* k*, i.e., whenever* (1.4) *holds, and whenever* $\ln(np)=o(\sqrt{k})$*, we can write,*

$$\begin{aligned} Q^{(p)}_{H(k)}(k)/\chi_{1-p} &\overset{d}{=} 1+\big(H(k)-\gamma\big)\ln c_n+\frac{\gamma}{\sqrt{k}}\,B_k-h_\rho(c_n)\,A(n/k)+O_p\left(\frac{A(n/k)}{\sqrt{k}}\right)\\ &\quad-\left(h_{\rho+\rho'}(c_n)\,A(n/k)\,B(n/k)+\frac12\,h^2_\rho(c_n)\,A^2(n/k)\right)\big(1+o_p(1)\big)\,, \end{aligned} \tag{2.16}$$

where B_k *is an asymptotically standard normal r.v.,* $h_\theta(x)=(x^\theta-1)/\theta$*,* $\theta<0$*. Consequently, if* $\sqrt{k}\,A(n/k)\to\lambda$*, finite, and* $\ln c_n/\sqrt{k}\to 0$*, as* $n\to\infty$*, then* $\frac{\sqrt{k}}{\ln c_n}\big(Q^{(p)}_{H(k)}(k)/\chi_{1-p}-1\big)$ *has asymptotically the same distribution as* $\sqrt{k}\,(H(k)-\gamma)$*, i.e., it is asymptotically normal, with variance* γ^2 *and mean value* $\lambda/(1-\rho)$*.*

Proof: From (2.2), and as $t\to\infty$, we get,

$$\frac{U(tx)}{U(t)} = x^\gamma\left\{1+h_\rho(x)\,A(t)+\left(h_{\rho+\rho'}(x)\,A(t)\,B(t)+\frac12\,h^2_\rho(x)\,A^2(t)\right)\big(1+o(1)\big)\right\}. \tag{2.17}$$

Denoting by $\hat\gamma$ any consistent tail index estimator and since $X_{n-k:n} \overset{d}{=} U(Y_{n-k:n})$, where Y is a standard Pareto r.v., we can write

$$Q^{(p)}_{\hat\gamma(k)}(k)/\chi_{1-p} = \left(\frac{X_{n-k:n}}{U(1/p)}\right)c_n^{\hat\gamma(k)} = \left(\frac{X_{n-k:n}}{U(n/k)}\right)\left(\frac{U(n/k)}{U(nc_n/k)}\right)c_n^{\hat\gamma(k)}\,.$$

Using the delta method, together with the fact that $\ln c_n/\sqrt{k} \to 0$, as $n \to \infty$, $c_n^{\hat\gamma(k)} \overset{p}{\sim} c_n^{\gamma}\{1+(\hat\gamma(k)-\gamma)\ln c_n\}$. From (2.17), we obtain

$$
\begin{aligned}
Q^{(p)}_{H(k)}(k)/\chi_{1-p} &\overset{d}{=} \left(1+\frac{\gamma}{\sqrt{k}}B_k + O_p\left(\frac{A(n/k)}{\sqrt{k}}\right)\right)\\
&\times \left\{1-h_\rho(c_n)\,A(n/k) - \left(h_{\rho+\rho'}(c_n)\,A(n/k)\,B(n/k) + h_\rho^2(c_n)\,A^2(n/k)/2\right)\left(1+o_p(1)\right)\right\}\\
&\times \left(1+\left(\hat\gamma(k)-\gamma\right)\ln c_n\right)\left(1+o_p(1)\right),
\end{aligned}
$$

and, with $\hat\gamma$ replaced by H, (2.16) as well as the asymptotic normality follow. □

Theorem 2.5.

(**a**) *Under the conditions of Theorem 2.4, let us consider the tail index estimator* $\widetilde{H} = \widetilde{H}_{\hat\beta,\hat\rho}$ *with* $(\hat\beta,\hat\rho)$ *consistent estimators of* (β,ρ)*, both computed at* k_1*, with* $k=o(k_1)$ *and such that* $(\hat\rho-\rho)\ln n = o_p(1)$*. Then, if* $\sqrt{k}\,A(n/k)\to\lambda$*,* $\frac{\sqrt{k}}{\ln c_n}\left(\overline{Q}^{(p)}_{\widetilde{H}(k)}(k)/\chi_{1-p}-1\right)$ *has asymptotically the same distribution as* $\sqrt{k}\left(\widetilde{H}(k)-\gamma\right)$*, i.e., they are both asymptotically normal, with variance equal to* γ^2 *and a null mean value (even if* $\lambda\neq 0$*).*

(**b**) *If we choose* k *such that* $\sqrt{k}\,A(n/k)\to\infty$*, but* $\sqrt{k}\,A(n/k)\,B(n/k)\to\lambda_B$*, finite,* $\frac{\sqrt{k}}{\ln c_n}\left(\overline{Q}^{(p)}_{\widetilde{H}(k)}(k)/\chi_{1-p}-1\right)$ *and* $\sqrt{k}\left(\widetilde{H}(k)-\gamma\right)$ *also have asymptotically the same distributions, i.e., they are asymptotically normal, with variance equal to* γ^2 *and asymptotic bias given in Theorem 2.2, provided that we can guarantee that* $(\hat\rho-\rho)\ln n = o_p\left(1/\sqrt{k}\,A(n/k)\right)$*.*

Proof: Let as first assume to know β and ρ. Then, since $\overline{Q}^{(p)}_{\widetilde{H}_{\beta,\rho}}(k;\beta,\rho) = Q^{(p)}_{\widetilde{H}_{\beta,\rho}}(k)\left(1+\widetilde{H}_{\beta,\rho}(k)\,\beta\left(\frac{n}{k}\right)^{\rho}h_\rho(c_n)\right)$ for models in (2.1), we can use (2.16) and get

$$
\begin{aligned}
\overline{Q}^{(p)}_{\widetilde{H}_{\beta,\rho}}(k;\beta,\rho)/\chi_{1-p} &\overset{d}{=} 1+\left(H_{\beta,\rho}(k)-\gamma\right)\ln c_n + \frac{\gamma}{\sqrt{k}}B_k + O_p\left(\frac{A(n/k)}{\sqrt{k}}\right)\\
&\quad - \left(h_{\rho+\rho'}(c_n)\,A(n/k)\,B(n/k) + \frac{1}{2}\,h_\rho^2(c_n)\,A^2(n/k)\right)\left(1+o_p(1)\right),
\end{aligned}
\tag{2.18}
$$

Then $\frac{\sqrt{k}}{\ln c_n}\left(\overline{Q}^{(p)}_{\widetilde{H}_{\beta,\rho}(k;\beta,\rho)}(k)/\chi_{1-p}-1\right)$ has asymptotically the same distributions as $\sqrt{k}\left(\widetilde{H}_{\beta,\rho}(k)-\gamma\right)$. Since, $\widetilde{H}_{\hat\beta,\hat\rho}(k)=\gamma(1+o_p(1))$, $c_n^\rho\to 0$, $c_n^\rho\ln c_n\to 0$, for any intermediate k, we may use Cramer's delta-method, and write

$$
\widetilde{H}_{\hat\beta,\hat\rho}(k)\,\hat\beta\left(\frac{n}{k}\right)^{\hat\rho}h_{\hat\rho}(c_n) \overset{p}{\sim} h_\rho(c_n)\,A(n/k)\left\{1+\frac{\hat\beta-\beta}{\beta}+(\hat\rho-\rho)\ln(n/k)\right\}.
$$

Consequently,

$$
\left(\overline{Q}^{(p)}_{\widetilde{H}_{\hat\beta,\hat\rho}}(k;\beta,\rho)-\overline{Q}^{(p)}_{\widetilde{H}_{\beta,\rho}}(k;\beta,\rho)\right)/\chi_{1-p} \overset{p}{\sim} \left(\widetilde{H}_{\hat\beta,\hat\rho}(k)-\widetilde{H}_{\beta,\rho}(k)\right)\ln c_n
$$

and

$$\left(\overline{Q}^{(p)}_{\widetilde{H}_{\hat\beta,\hat\rho}}(k;\hat\beta,\hat\rho)-\overline{Q}^{(p)}_{\widetilde{H}_{\beta,\rho}}(k;\beta,\rho)\right)/\chi_{1-p}\ \overset{p}{\sim}$$
$$\overset{p}{\sim}\ \left(\widetilde{H}_{\hat\beta,\hat\rho}(k)-\widetilde{H}_{\beta,\rho}(k)\right)\ln c_n+h_\rho(c_n)\,A(n/k)\left\{\frac{\hat\beta-\beta}{\beta}+(\hat\rho-\rho)\,\ln(n/k)\right\}.$$

The remaining of the proof is analogous to the proof of Theorem 2.3. □

3. A SMALL-SCALE SIMULATION STUDY

We have implemented, for *Fréchet* underlying parents, a Monte Carlo simulation of size 5000 for $R_H\equiv Q^{(p)}_H/\chi_{1-p}$, $\overline{R}_{\overline{H}}\equiv\overline{Q}^{(p)}_{\overline{H}}/\chi_{1-p}$ and $\overline{R}_{\overline{\overline{H}}}\equiv\overline{Q}^{(p)}_{\overline{\overline{H}}}/\chi_{1-p}$. Results for $\overline{\overline{Q}}$, not presented, have also been simulated and almost overlap the ones for $\overline{Q}$. For every estimator $R=R(k)$, we have simulated for $p=1/n$ and $p=1/(n\ln n)$, the mean value, the root mean squared error (RMSE) and the optimal sample fraction, $OSF_R=k_0/n=\arg\min_k\{RMSE(R(k))\}/n$. The second order parameters were estimated through $\hat\rho_0=\hat\rho(k_1;0)$ and $\hat\beta_0=\hat\beta_{\hat\rho_0}(k_1)$, with $\hat\rho(k;\tau)$ and $\hat\beta_{\hat\rho}(k)$ defined in (2.3) and (2.9), respectively, and k_1 given in (2.6).

Table 2: Simulated mean values /RMSE at optimal levels.

n	100	500	1000	5000
Fréchet parent with $\gamma=0.25$ and $p=1/n$				
R_H	1.056 / 0.191	1.053 / 0.136	1.053 / 0.118	1.037 / 0.080
$\overline{R}_{\overline{H}}$	0.969 / 0.164	0.984 / 0.116	0.988 / 0.099	0.992 / 0.061
$\overline{R}_{\overline{\overline{H}}}$	1.007 / 0.154	1.006 / 0.108	1.004 / 0.092	1.004 / 0.057
Fréchet parent with $\gamma=0.25$ and $p=1/(n\ln n)$				
R_H	1.106 / 0.298	1.089 / 0.259	1.085 / 0.172	1.057 / 0.112
$\overline{R}_{\overline{H}}$	0.960 / 0.236	0.984 / 0.162	0.988 / 0.135	0.991 / 0.080
$\overline{R}_{\overline{\overline{H}}}$	1.009 / 0.224	1.013 / 0.152	1.009 / 0.127	1.009 / 0.076

A few remarks for Fréchet parents:

- For *Fréchet* parents, the RMSE of $\overline{R}_{\overline{H}(k)}(k)$ and $\overline{R}_{\overline{\overline{H}}(k)}(k)$ is always smaller (or equal) than the RMSE of the classical quantil estimator, $R_{H(k)}(k)$.
- Also, the normalized quantile estimator $\overline{R}_{\overline{\overline{H}}(k)}(k)$ has always the smallest mean squared error.

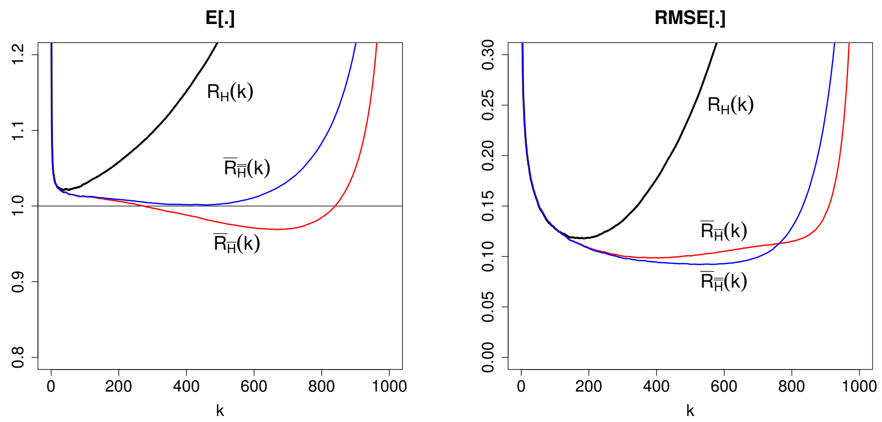

Figure 3: Underlying Fréchet parent with $\gamma = 0.25$, $p = 1/n$, and $n = 1000$.

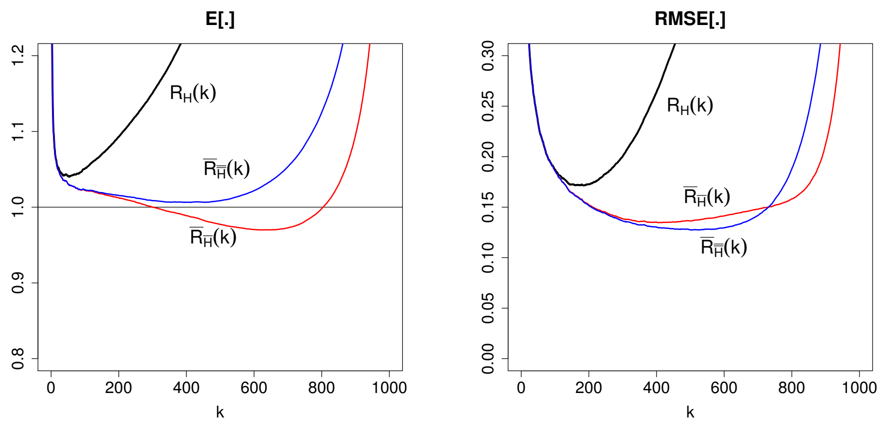

Figure 4: Underlying Fréchet parent with $\gamma = 0.25$, $p = 1/(n \ln n)$ and $n = 1000$.

4. CONCLUSION

The MVRB estimators proposed in this paper are bias-corrected Hill estimators which perform better than the classical Hill estimator for all k, the number of top o.s.'s used in the estimation of the tail index γ. Despite of this, it is sensible to understand their comparative behaviour at optimal levels, not only for finite sample size, but also asymptotically, as recently done in Gomes and Neves (2007) for some of the classical estimators, like the well-known Hill, moment, maximum likelihood and the recently introduced mixed moment estimator (Fraga Alves *et al.*, 2007). It is thus crucial to have information on the order of the dominant component(s) of their asymptotic bias, the main contribution in this paper, for the MVRB tail index estimators in (1.10) and the associated quantile estimators in (1.13). The adaptive choice of the threshold is now becoming feasible for a wide class of models, but it is outside of the scope of this paper.

ACKNOWLEDGMENTS

Research partially supported by FCT/POCTI and POCI/FEDER.

REFERENCES

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

[2] Beirlant, J.; Figueiredo, F.; Gomes, M.I. and Vandewalle, B. (2006). Improved reduced-bias tail index and quantile estimators, *J. Statist. Plann. and Inference*, DOI:10.1016/j.jspi.2007.07.015, 2007.

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

[4] Caeiro, C. and Gomes, M.I. (2007). *Semi-parametric second-order reduced-bias high quantile estimation*, Pré-Publicações CMA-UNL 8/2007 (submitted).

[5] Caeiro, C.; Gomes, M.I. and Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator, *Revstat*, **3**(2), 113–136.

[6] Ferreira, A.; Haan, L. de and Peng, L. (2003). On optimising the estimation of high quantiles of a probability distribution, *Statistics*, **37**(5), 401–434.

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

[8] Fraga Alves, M.I.; Gomes, M.I. and Haan, L. de (2003). A new class of semi-parametric estimators of the second order parameter, *Portugaliae Mathematica*, **60**(2), 193–213.

[9] Fraga Alves, M.I.; Gomes, M.I.; Haan, L. de and Neves, C. (2007). *Mixed moment estimator and location invariant alternatives*, Notas e Comunicações CEAUL 14/2007 (submitted).

[10] Geluk, J. and Haan, L. de (1980). *Regular Variation, Extensions and Tauberian Theorems*, CWI Tract 40, Centre for Mathematics and Computer Science, Amsterdam, The Netherlands.

[11] Gnedenko, B.V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire, *Ann. Math.*, **44**(6), 423–453.

[12] Gomes, M.I.; Caeiro, F. and Figueiredo, F. (2004a). Bias reduction of a extreme value index estimator trough an external estimation of the second order parameter, *Statistics*, **38**(6), 497–510.

[13] Gomes, M.I. and Figueiredo, F. (2006). Bias reduction in risk modelling: semi-parametric quantile estimation, *Test*, **15**(2), 375–396.

[14] Gomes, M.I.; Figueiredo, F. and Mendonça, S. (2005). Asymptotically best linear unbiased tail estimators under a second order regular variation, *J. Statist. Planning and Inference*, **134**(2), 409–433.

[15] Gomes, M.I.; de Haan, L. and Henriques Rodrigues, L. (2004b). Tail Index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses, *J. Royal Statistical Society B*, DOI: 10.1111/j.1467-9869.2007.00620.x, 2007.

[16] Gomes, M.I.; Haan, L. de and Peng, L. (2002). Semi-parametric estimation of the second order parameter — asymptotic and finite sample behaviour, *Extremes*, **5**(4), 387–414.

[17] 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.

[18] 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.

[19] Gomes, M.I.; Martins, M.J. and Neves, M. (2007a). Improving second order reduced bias extreme value index estimation, *Revstat*, **5**(2), 177–207.

[20] Gomes, M.I. and Neves, C. (2007). Asymptotic comparison of the mixed moment and classical extreme value index estimators, *Statistics and Probability Letters*, DOI: 10.1016/j,spl.2007.07.026, 2007.

[21] Gomes, M.I. and Oliveira, O. (2003). How can non-invariant statistics work in our benefit in the semi-parametric estimation of parameters of rare events, *Comm. in Statist. — Simulation and Computation*, **32**(4), 1005–1028.

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

[23] Gomes, M.I. and Pestana, D. (2007b). A sturdy reduced bias extreme quantile (VaR) estimator, *J. Amer. Statist. Assoc.*, **102**, 477, 280–292.

[24] Haan, L. de (1970). *On Regular Variation and its Application to the Weak Convergence of Sample Extremes*, Mathematical Centre Tract 32, Amsterdam.

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

[26] Haan, L. de and Rootzén, H. (1983). On the estimation of high quantiles, *J. Statist. Planning and Inference*, **35**, 1–13.

[27] Hall, P. (1982). On some simple estimates of an exponent of regular variation, *J. R. Statist. Soc.*, **44**(1), 37–42.

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

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

[30] Matthys, G. and Beirlant, J. (2003). Estimating the extreme value index and high quantiles with exponential regression models, *Statistica Sinica*, **13**, 853–880.

[31] Matthys, G.; Delafosse, M.; Guillou, A. and Beirlant, J. (2004). Estimating catastrophic quantile levels for heavy-tailed distributions, *Insurance: Mathematics and Economics*, **34**, 517–537.

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

[33] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations, *J. Amer. Statist. Assoc.*, **73**, 812–815.