# Peaks Over Random Threshold Methodology for Tail Index and High Quantile Estimation

Paulo Araújo Santos
Instituto Politécnico de Santarém, ESGS
paulo.santos@esgs.pt

Isabel Fraga Alves
Universidade de Lisboa, CEAUL and DEIO (FCUL)
isabel.alves@fc.ul.pt

M. Ivette Gomes
Universidade de Lisboa, CEAUL and DEIO (FCUL)
ivette.gomes@fc.ul.pt

May 15, 2006

**Abstract:** In this paper we present a class of semi-parametric *high quantile* estimators which enjoy a desirable property in the presence of linear transformations of the data. Such feature is in accordance with the empirical counterpart of the theoretical linearity of a quantile, $\chi_p$: $\chi_p(\delta X + \lambda) = \delta\chi_p(X) + \lambda$, for any real $\lambda$ and positive $\delta$. This class of estimators is based on the sample of excesses over a random threshold, originating what we denominate PORT (Peaks Over Random Threshold) methodology. We prove consistency and asymptotic normality of two high quantile estimators in this class, associated with the PORT-estimators for the tail index. The exact performance of the new tail index and quantile PORT-estimators is compared with the original semi-parametric estimators, through a simulation study. As an empirical example we estimate the Value at Risk (VaR) for the Nasdaq Composite index, for investors betting on a fall in the index.

**Keywords:** High Quantiles; Semi-parametric Estimation; Linear property; Sample of Excesses

## 1 INTRODUCTION

In this paper we deal with semi-parametric estimators of the *tail index* $\gamma$ and *high quantiles* $\chi_p$, which enjoy desirable properties in the presence of linear transformations of the available data. We recall that a *high quantile* is a value exceeded with a small probability. Formally, we denote by $F$ the heavy-tailed distribution function (d.f.) of a random variable (r.v.) $X$, the common d.f. of the i.i.d. sample

$\underline{X} := \{X_i\}_{i=1}^n$, for which the high quantile

$$\chi_p(X) := F^{\leftarrow}(1-p),\ p = p_n \to 0,\ \text{as } n \to \infty,\ np_n \to c \geq 0, \tag{1.1}$$

has to be estimated. Here $F^{\leftarrow}(t) := \inf\{x : F(x) \geq t\}$ denotes the generalized inverse function of $F$.

We consider estimators based on the $k+1$ top order statistics (o.s.), $X_{n:n} \geq \cdots \geq X_{n-k:n}$, where $X_{n-k:n}$ is an intermediate o.s., i.e., $k$ is an intermediate sequence of integers such that

$$k = k_n \to \infty, \quad k_n/n \to 0, \quad \text{as } n \to \infty. \tag{1.2}$$

We assume that we are working in a context of heavy tails, i.e., $\gamma > 0$ in the extreme value distribution

$$G_\gamma(x) = \begin{cases} \exp\left\{-(1+\gamma x)^{-1/\gamma}\right\}, & 1+\gamma x > 0,\ \gamma \neq 0 \\ \exp\left(-e^{-x}\right), & x \in \mathbb{R},\ \gamma = 0, \end{cases} \tag{1.3}$$

the non-degenerate d.f. to which the maximum $X_{n:n}$ is attracted, after a suitable linear normalization. When this happens we say that the d.f. $F$ is in the Fréchet domain of attraction and we write $F \in D(G_\gamma)_{\gamma>0}$.

The paper is developed under the first order regular variation condition, which allows the extension of the empirical d.f. beyond the range of the available data, assuming a polynomial decay of the tail. This condition can be expressed by

$$F \in D(G_\gamma)_{\gamma>0} \quad \text{iff} \quad \overline{F} \in RV_{-1/\gamma} \quad \text{iff} \quad U \in RV_\gamma, \tag{1.4}$$

where $U$ is the quantile function defined as $U(t) := F^{\leftarrow}(1-1/t)$, $t \geq 1$; the notation $RV_\alpha$ stands for the class of regularly functions at infinity with index of regular variation $\alpha$, i.e., positive measurable functions $h$ such that $\lim_{t\to\infty} h(tx)/h(t) = x^\alpha$, for all $x > 0$.

It is interesting to note that the $p$-quantile can be expressed as $\chi_{p_n} = U\left(1/p_n\right)$.

To get asymptotic normality of estimators of parameters of extreme events, it is usual to assume the following extra second regular variation condition, that involves a non-positive parameter $\rho$:

$$\lim_{t\to\infty} \frac{U(tx)/U(t) - x^\gamma}{A(t)} = x^\gamma \frac{x^\rho - 1}{\rho}, \tag{1.5}$$

for all $x > 0$, where $A$ is a suitable chosen function of constant sign near infinity. Then, $|A| \in RV_\rho$ and $\rho$ is called the second order parameter (Geluk and de Haan, 1987).

More specifically, we consider that $F$ belongs to the wide class of Hall (Hall, 1982), that is, the associated quantile function $U$ satisfies

$$U(t) = Ct^\gamma(1 + Dt^\rho + o(t^\rho)), \quad \rho < 0, \quad C > 0, \quad D \neq 0 \quad \text{as} \quad t \to \infty, \tag{1.6}$$

or equivalently, (1.5) holds, with $A(t) = D\rho t^\rho$.

Returning to the problem of high quantile estimation, we recall the classical semi-parametric Weissman-type estimator of $\chi_{p_n}$ (Weissman, 1978),

$$\widehat{\chi}^W_{p_n} = \widehat{\chi}^W_{p_n}(\underline{X}) = X_{n-k_n:n}\left(\frac{k_n}{np_n}\right)^{\hat{\gamma}_n}, \tag{1.7}$$

with $\hat{\gamma}_n = \hat{\gamma}_n(\underline{X})$ some consistent estimator of the tail parameter $\gamma$.

In the classical approach one considers for $\hat{\gamma}_n$ the well known Hill estimator (Hill, 1975),

$$\hat{\gamma}^H_n = \hat{\gamma}^H_n(\underline{X}) = \frac{1}{k_n}\sum_{j=1}^{k_n}\log\frac{X_{n-j+1:n}}{X_{n-k_n:n}}, \tag{1.8}$$

or the Moment estimator (Dekkers et al., 1989),

$$\hat{\gamma}^M_n = \hat{\gamma}^M_n(\underline{X}) = M^{(1)}_n + 1 - \frac{1}{2}\left\{1 - \frac{(M^{(1)}_n)^2}{M^{(2)}_n}\right\}^{-1}, \tag{1.9}$$

with $M^{(r)}_n$, the $r$-Moment of the log-excesses, defined by

$$M^{(r)}_n = M^{(r)}_n(\underline{X}) = \frac{1}{k_n}\sum_{j=1}^{k_n}\left(\log\frac{X_{n-j+1:n}}{X_{n-k_n:n}}\right)^r,\ r = 1, 2. \tag{1.10}$$

We use the following notation:

$$\widehat{\chi}^H_{p_n} = X_{n-k_n:n}\left(\frac{k_n}{np_n}\right)^{\hat{\gamma}^H_n}, \qquad \widehat{\chi}^M_{p_n} = X_{n-k_n:n}\left(\frac{k_n}{np_n}\right)^{\hat{\gamma}^M_n}. \tag{1.11}$$

A very simple motivation of the estimator (1.7) may be found in Gomes and Figueiredo (2003). Note that condition (1.4) is equivalent to $\lim_{t\to\infty} U(tx)/U(t) = x^\gamma$. Assume that $U(t) \sim Ct^\gamma$, as $t \to \infty$; then, a natural estimator of $\chi_{p_n}$ is $\widehat{\chi}_{p_n} = \hat{C}p_n^{-\hat{\gamma}}$, with $\hat{\gamma}_n$ any consistent estimator of $\gamma$ and $\hat{C} = (k_n/n)^{\hat{\gamma}_n} X_{n-k_n:n}$. To justify the expression for $\hat{C}$, recall that

$$X_{j:n} \stackrel{d}{=} U(Y_{j:n}), \quad \text{for } j = 1, \cdots, n,$$

with $Y_{j:n}$ the o.s. associated to an i.i.d. sample from the standard Pareto model $F_Y(y) = 1 - y^{-1}$, $y \geq 1$; for the intermediate o.s. we have

$$Y_{n-k_n:n} \stackrel{p}{\sim} \frac{n}{k_n} \quad \text{and} \quad X_{n-k_n:n} \stackrel{d}{=} U(Y_{n-k_n:n}) \sim CY^\gamma_{n-k_n:n} \stackrel{p}{\sim} C\left(\frac{n}{k_n}\right)^\gamma,$$

which leads to the mentioned expression for $\hat{C}$.

Finally, we explain the question that motivated this paper. It is well known that scale transformations to the data do not interfere with the stochastic behaviour of the tail index estimators (1.8) and (1.9), i.e., we can say that they enjoy scale invariance. The incorporation of (1.8) or (1.9) in the Weissman-type estimator in (1.7), allows us to obtain the following desirable exact property for quantile estimators: for any real positive $\delta$,

$$\widehat{\chi}^{W}_{p_n}(\delta \underline{X}) = \delta X_{n-k_n:n} \left(\frac{k_n}{np_n}\right)^{\hat{\gamma}_n} = \delta \widehat{\chi}^{W}_{p_n}(\underline{X}). \tag{1.12}$$

But we want a similar linear property in the case of location transformations to the data, $Z_j := X_j + \lambda,\ j = 1, \cdots, n$, for any real $\lambda$. That is, our main goal is that, for the transformed data $\underline{Z} := \{Z_j\}_{j=1}^{n}$, the quantile estimator satisfies

$$\widehat{\chi}_{p_n}(\underline{Z}) = \widehat{\chi}_{p_n}(\underline{X}) + \lambda. \tag{1.13}$$

Altogether, this represents the empirical counterpart of the following theoretical linear property for quantiles,

$$\chi_p(\delta X + \lambda) = \delta \chi_p(X) + \lambda, \text{ for any real } \lambda \text{ and real positive } \delta. \tag{1.14}$$

Fraga Alves and Araújo Santos (2004) proposed a simple modification of (1.7) that enjoys the property (1.13) approximately. Here we present a class of high quantile-estimators for which (1.12) and (1.13) hold exactly, pursuing the empirical counterpart of the theoretical linear property (1.14).

Denote by $x_F := \inf\{x : F(x) > 0\}$ the left endpoint of a d.f. $F$. The class of estimators suggested here is function of a sample of excesses over a random threshold $X_{n_q:n}$,

$$\underline{X}^{(q)} := \left(X_{n:n} - X_{n_q:n}, X_{n-1:n} - X_{n_q:n}, \cdots, X_{n_q+1:n} - X_{n_q:n}\right), \tag{1.15}$$

where $n_q := [nq] + 1$, with:

- $0 < q < 1$, for d.f.'s with finite or infinite left endpoint $x_F$ (*the random threshold is an empirical quantile*);
- $q = 0$, for d.f.'s with finite left endpoint $x_F$ (*the random threshold is the minimum*).

A statistical inference method based on the sample of excesses $\underline{X}^{(q)}$ defined in (1.15) will be called a PORT-methodology, with PORT standing for Peaks Over Random Threshold. We propose the following **PORT-Weissman estimators**:

$$\widetilde{\chi}^{(q)}_{p_n} = (X_{n-k_n:n} - X_{n_q:n}) \left(\frac{k_n}{np_n}\right)^{\hat{\gamma}_n^{(q)}} + X_{n_q:n}, \tag{1.16}$$

where $\hat{\gamma}_n^{(q)}$ is any consistent estimator of the tail parameter $\gamma$, made location/scale invariant by using the transformed sample $\underline{X}^{(q)}$. Indeed, the incorporation in the PORT-Weissman estimator in (1.16), of tail index estimators, as function of the sample of excesses, allows us to obtain exactly the linear property (1.13).

In Section 2, we study the behavior of the classical tail index and quantile estimators in the presence of shifts, for data generated from the Pareto Model. In Section 3, we derive asymptotic properties for the Hill and Moment PORT-estimators, i.e., as functions of the sample of excesses (1.15). In Section 4, we propose two estimators for $\chi_p$ that belong to the class (1.16) and prove their asymptotic normality. In Section 5, through simulation experiments, we compare the performance of the new estimators with the classical ones. Finally, in Section 6 we estimate the Value at Risk (VaR) for a Nasdaq Composite index data set.

## 2 SPECIAL CASE: SHIFT IN THE PARETO MODEL

It is worth looking at the special case of the Pareto$(\gamma, \delta)$ model,

$$F_{\gamma,\delta}(x) = 1 - (x/\delta)^{-1/\gamma}\,,\; x > \delta,\; \delta > 0. \tag{2.1}$$

With this underlying d.f., the Hill estimator $\hat{\gamma}_n^H$ is unbiased for any sequence of integers $k_n$, $1 \le k_n < n$ and the quantile function is $U(t) = \delta t^\gamma$, $t \ge 1$. So, for this special model, Weissman-type estimators in (1.7) perform reasonably well, with $\hat{\gamma}_n = \hat{\gamma}_n^H$; indeed, for the model (2.1), the two errors due to the estimation of $\gamma$ and to the approximation taken for $U(t)$ do not intervene. However, a small shift in the data may lead to disastrous results. Notice that for any r.v. $X$, with quantile function $U_X(t)$, the transformed r.v. $Z = X + \lambda$ has an associated quantile function given by $U_Z(t) = U_X(t) + \lambda$. In the Pareto$(\gamma, \delta)$ model (2.1), this means that if $X$ has a d.f. $F_{\gamma,\delta}(x)$ then the r.v. $Z = X + \lambda$ has d.f.

$$F_{\gamma,\delta,\lambda}(z) = 1 - \left(\frac{z-\lambda}{\delta}\right)^{-1/\gamma},\; z > \lambda + \delta,\; \delta > 0. \tag{2.2}$$

The quantile function of the r.v. $Z$ is

$$U_Z(t) = \delta t^\gamma + \lambda = C t^\gamma (1 + D t^\rho), \quad C := \delta, \quad D := \lambda/\delta, \quad \rho := -\gamma.$$

The extra term $Dt^\rho$, in the slowly varying function $1 + Dt^\rho$, affects the first approximation $U_Z(t) \sim Ct^\gamma$, with severe consequences in the two above mentioned errors, as illustrated in the sequel by simulation.

We have generated a sample of size $n = 1000$ of i.i.d. Pareto$(1,1)$ r.v.'s; since the scale does not play a role for our purpose here, our aim is to estimate the upper $p_n$-quantile, with $p_n = 0.001$, i.e., $\chi_{p_n}(X) = F^{\leftarrow}_{\gamma,1}(0.999)$, with $\gamma = 1$. That is, $np_n = 1$, i.e., we are actually looking for an extreme quantile. Notice that in a Pareto$(1,1)$ model, we thus want to estimate $\chi_{p_n}(X) = 1000$.

On the left side of Figure 2.1 we compare two classical semi-parametric procedures for the estimation of the tail parameter $\gamma$ — the Hill estimator, $\hat{\gamma}_n^H$, in (1.8) and the Moment estimator, $\hat{\gamma}_n^M$, in (1.9) — through the mean of $N = 200$ replications, for $6 \le k \le 800$.

Similarly, on the right side are plotted the graphics of the Weissman quantile estimators $\hat{\chi}^H_{p_n}$ and $\widehat{\chi}^M_{p_n}$ in (1.11), correspondent to the means of $N = 200$ replications for the same range of $k$.

At a second stage, we introduce a positive shift and a similar analysis has been accomplished. Namely, in Figure 2.2 we have considered a random generation from $Z = X + \lambda$. Our main objective is to estimate here the target quantile $\chi_{p_n}(Z) = \chi_{p_n}(X) + \lambda = F^{\leftarrow}_{\gamma,1}(0.999) + \lambda$. We set $\lambda = \chi_{.01}(X) = 100$, so that the target quantile to be estimated for the shifted model is $\chi_{p_n}(Z) = 1100$.

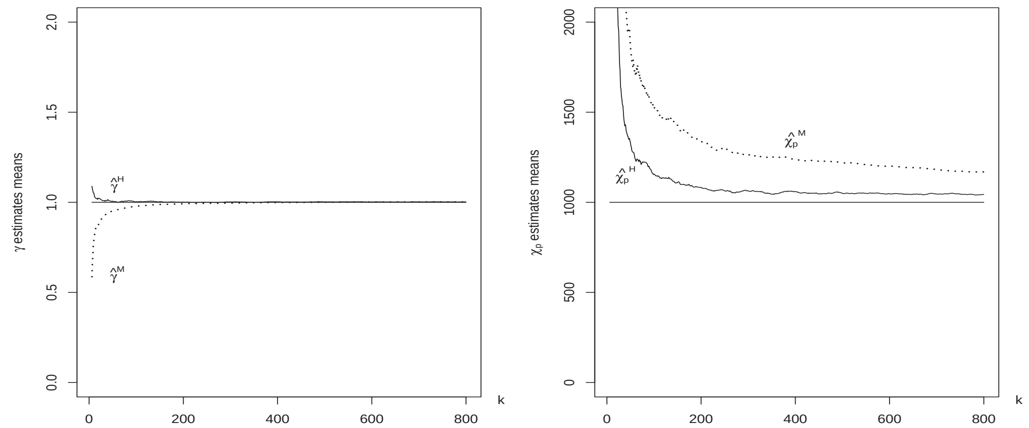


Figure 2.1: Mean values of $\hat{\gamma}^H$, $\hat{\gamma}^M$ (left) and $\widehat{\chi}^H_{p_n}$, $\widehat{\chi}^M_{p_n}$, $p_n = 0.001$ (right), for sample size $n = 1000$, from the Pareto model in (2.2) with $\gamma = 1$, $\delta = 1$ and $\lambda = 0$ (target quantile $\chi_{p_n} = 1000$).

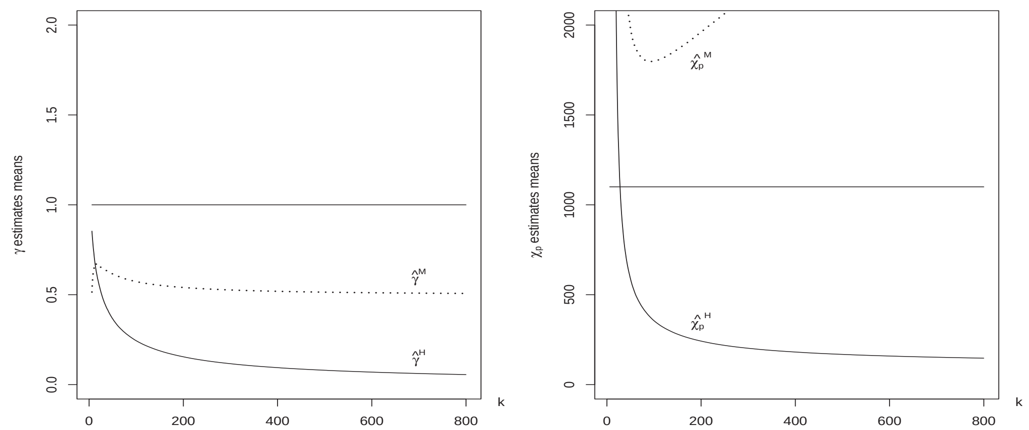


Figure 2.2: Mean values of $\hat{\gamma}^H$, $\hat{\gamma}^M$ (left) and $\widehat{\chi}^H_{p_n}$, $\widehat{\chi}^M_{p_n}$, $p_n = 0.001$ (right), for sample size $n = 1000$, from the Pareto model in (2.2) with $\gamma = 1$, $\delta = 1$ and $\lambda = \chi_{0.01} = 100$ (target quantile $\chi_{p_n} = 1100$).

It is evident from the left graphic of Figure 2.1 the unbiased property of the Hill estimator for the Pareto model (2.1). On the other hand, Figure 2.2 illustrates the disastrous results we may achieve with shifted data. In particular, we enhance the fact that the flat stable zones of the graphs, based on the shifted data, may lead us to dangerous misleading conclusions: wrong underestimated tail index and quantile, for instance. Indeed, the "stable zones" need to be carefully identified.

This clearly indicates that in practice we should take care with the Pareto approximation $U(t) \sim Ct^\gamma$. A usual solution to overcome this problem is to make statistical inference only after a suitable shift of the data. In the literature, it has been sometimes suggested to subtract a random quantity, usually the minimum of the sample. This shifted data set has the advantage of working out with non-negative values, a desirable property for classical semi-parametric estimators. An extensive discussion about this type of shifted procedures can be found in Drees (2003). Therein, it is studied the effect of subtracting the minimum of the sample, previously to the subsequent analysis of the Nasdaq Composite index log-returns data set, in the context of VaR estimation. In fact, for that particular data, it is therein observed that this procedure constitutes a considerable improvement, arising for Hill $\gamma$-estimates a larger flat zone in the associated sample path, after transforming the original data through the subtraction of the smallest observation. In this paper, we draw the attention for the erroneous conclusions we may take from such a data analysis.

As far as we know, no systematic study has been done concerning asymptotic and exact properties of semi-parametric methodologies for tail index and high quantile estimation, using the transformed sample by subtraction of a random threshold. Somehow related with this subject, Gomes and Oliveira (2003), in a context of regularly varying tails, suggested a simple generalization of the classical Hill estimator associated to artificially shifted data. The shift imposed to the data is deterministic, with the aim of reducing the main component of the bias of Hill's estimator, getting thus estimates with stable sample paths around the target value. A preliminary study has also been carried out, by the same authors, replacing the artificial deterministic shift by a random shift, which in practice represents a transformation of the original data through the subtraction of the smallest observation, added by one, whenever we are aware that the underlying heavy-tailed model has a finite left endpoint.

With the purpose of tail index and high quantile estimation there is, in our opinion, a gap in the literature regarding classical semi-parametric estimation methodologies adapted for shifted data, the main topic of this paper.

# 3 ASYMPTOTIC BEHAVIOR OF TAIL INDEX PORT-ESTIMATORS

In this section we present asymptotic results for the classical Hill estimator in (1.8) and the Moment estimator in (1.9), both based on the sample of excesses $\underline{X}^{(q)}$ in (1.15), which will be denoted respectively, by

$$\hat{\gamma}_n^{H(q)} := \hat{\gamma}_n^{H}(\underline{X}^{(q)}) \quad \text{and} \quad \hat{\gamma}_n^{M(q)} := \hat{\gamma}_n^{M}(\underline{X}^{(q)}), \quad 0 \leq q < 1. \tag{3.1}$$

In the following, $\chi_q^*$ denotes the $q$-quantile of $F$: $F(\chi_q^*) = q$ (by convention $\chi_0^* := x_F$), so that

$$X_{n_q:n} \xrightarrow{p} \chi_q^*, \qquad \text{as } n \to \infty, \qquad \text{for } 0 \le q < 1.$$

For the estimators in (3.1) we have the asymptotic distributional representations expressed in Theorems 3.1 and 3.2.

**Theorem 3.1 (PORT-Hill).** *For any intermediate sequence $k$ as in* (1.2)*, under the validity of the second order condition in* (1.5) *and for any real $q$, $0 \le q < 1$, the asymptotic distributional representation*

$$\hat{\gamma}_n^{H(q)} \stackrel{d}{=} \gamma + \frac{\sigma_H}{\sqrt{k}} P_k^H + \left( c_H A(n/k) + d_H \frac{\chi_q^*}{U(n/k)} \right) \big(1 + o_p(1)\big) \tag{3.2}$$

*holds, where $P_k^H$ is an asymptotically standard normal r.v.,*

$$\sigma_H^2 := \gamma^2, \quad c_H := \frac{1}{1-\rho} \quad \text{and} \quad d_H := \frac{\gamma}{\gamma+1}\,. \tag{3.3}$$

*Proof:* The r.v. $\hat{\gamma}_n^{H(q)}$ may be written as

$$\hat{\gamma}_n^H + \frac{1}{k} \sum_{j=1}^{k} \ln \frac{1 - \frac{X_{n_q:n}}{X_{n-j+1:n}}}{1 - \frac{X_{n_q:n}}{X_{n-k:n}}}$$

and if we use the first order approximation $\ln(1+x) \sim x$, as $x \to 0$, we get

$$\begin{aligned} \frac{1}{k} \sum_{j=1}^{k} \ln \frac{1 - \frac{X_{n_q:n}}{X_{n-j+1:n}}}{1 - \frac{X_{n_q:n}}{X_{n-k:n}}} &\sim \frac{1}{k} \sum_{j=1}^{k} \left[ \frac{X_{n_q:n}}{X_{n-k:n}} - \frac{X_{n_q:n}}{X_{n-j+1:n}} \right] \\ &= \frac{1}{k} \sum_{j=1}^{k} \left[ \frac{\chi_q^*}{X_{n-k:n}} \Big( 1 - \frac{X_{n-k:n}}{X_{n-j+1:n}} \Big) \right] \big(1 + o_p(1)\big). \end{aligned}$$

Let $Y_1, Y_2, ..., Y_n$ be i.i.d. standard Pareto. Then

$$\left( \frac{k}{n} \right) Y_{n-k:n} \xrightarrow{p} 1,$$

for any intermediate sequence $k$, and

$$\begin{aligned} \hat{\gamma}_n^{H(q)} &= \hat{\gamma}_n^H + \frac{\chi_q^*}{U(n/k)} \frac{1}{k} \sum_{j=1}^{k} \Big[ 1 - Y_{k-j+1:n}^{-\gamma} \Big] \big(1 + o_p(1)\big) \\ &= \hat{\gamma}_n^H + \frac{\chi_q^*}{U(n/k)} \frac{1}{k} \sum_{j=1}^{k} \Big[ 1 - Y_j^{-\gamma} \Big] \big(1 + o_p(1)\big). \end{aligned}$$

We know that $E\big[Y^{-\gamma}\big] = \frac{1}{\gamma+1}$ and by the weak law of large numbers

$$\hat{\gamma}_n^{H(q)} = \hat{\gamma}_n^H + \frac{\chi_q^*}{U(n/k)} \frac{\gamma}{\gamma+1} \big(1 + o_p(1)\big). \tag{3.4}$$

Now, considering the well know asymptotic distributional representation for $\hat{\gamma}_n^H$,

$$\hat{\gamma}_n^H \stackrel{d}{=} \gamma + \frac{\gamma}{\sqrt{k}} P_k^H + \frac{1}{1-\rho} A(n/k)\big(1 + o_p(1)\big), \tag{3.5}$$

where $P_k^H$ is an asymptotically standard normal r.v., the representation (3.2) follows immediately. □

**Theorem 3.2 (PORT-Moment).** *For any intermediate sequence $k$, as in* (1.2), *under the validity of the second order condition in* (1.5) *and for any real $q$, $0 \le q < 1$, the asymptotic distributional representation*

$$\hat{\gamma}_n^{M(q)} \stackrel{d}{=} \gamma + \frac{\sigma_M}{\sqrt{k}} P_k^M + \left( c_M A(n/k) + d_M \frac{\chi_q^*}{U(n/k)} \right) \Big(1 + o_p(1)\Big), \tag{3.6}$$

*holds, where $P_k^M$ is asymptotically standard normal,*

$$\sigma_M^2 := \gamma^2 + 1, \quad c_M := \frac{\gamma(1-\rho)+\rho}{\gamma(1-\rho)^2} \quad \textit{and} \quad d_M := \left(\frac{\gamma}{\gamma+1}\right)^2. \tag{3.7}$$

*Proof:* From now on, we denote by $M_n^{(r,q)}$ the $M_n^{(r)}$ statistics in (1.10), as functions of the transformed sample $\underline{X}^{(q)}$, $0 \le q < 1$ in (1.15); that is, $M_n^{(r,q)} := M_n^{(r)}(\underline{X}^{(q)})$. Developments similar to the ones in the proof of Theorem 3.1, enable us to write

$$M_n^{(2,q)} = \frac{1}{k}\sum_{j=1}^{k} \Big[ \ln \frac{X_{n-j+1:n}}{X_{n-k:n}} + \frac{\chi_q^*}{U(n/k)} \big(1 - Y_{k-j+1}^{-\gamma}\big)\big(1 + o_p(1)\big)\Big]^2 = M_n^{(2)} + V_1,$$

where

$$\begin{aligned} V_1 &= \tfrac{2\chi_q^*}{U(n/k)} \tfrac{1}{k} \sum_{j=1}^{k} \Big[ \ln \frac{X_{n-j+1:n}}{X_{n-k:n}} \big(1 - Y_{k-j+1}^{-\gamma}\big)\big(1 + o_p(1)\big)\Big] + \frac{1}{k}\sum_{j=1}^{k} O_p\Big(\tfrac{1}{U^2(n/k)}\Big) \\ &\stackrel{d}{\sim} \tfrac{2\chi_q^*}{U(n/k)} \Bigg[ M_n^{(1)} - \tfrac{1}{k}\sum_{j=1}^{k} \Big[ \Big( \ln \frac{X_{n-j+1:n}}{X_{n-k:n}} \Big) Y_{k-j+1}^{-\gamma} \Big] \Bigg]. \end{aligned}$$

Under the first order condition (1.4), we can write the last expression as

$$\frac{2\chi_q^*}{U(n/k)} \Bigg[ M_n^{(1)} - \frac{1}{k}\sum_{j=1}^{k} \Big[ \gamma\big(\ln Y_j\big) Y_j^{-\gamma} + \frac{Y_j^{\rho}-1}{\rho} Y_j^{-\gamma} A(n/k) + o_p\big(A(n/k)\big)\Big]\Bigg].$$

We know that $E\big[\big(\ln Y\big) Y^{-\gamma}\big] = \frac{1}{(\gamma+1)^2}$ and by the weak law of large numbers,

$$\begin{aligned} M_n^{(2,q)} &= M_n^{(2)} + \tfrac{2\chi_q^*}{U(n/k)} \big[\gamma - \tfrac{\gamma}{(\gamma+1)^2} + O_p\big(\tfrac{1}{\sqrt{k}}\big) + O_p\big(A(n/k)\big) + o_p(1)\big] \\ &= M_n^{(2)} + \tfrac{2\chi_q^*}{U(n/k)} \big[\gamma - \tfrac{\gamma}{(\gamma+1)^2} + o_p(1)\big]. \end{aligned} \tag{3.8}$$

For $\big(M_n^{(1,q)}\big)^2$ we use the representation (3.4) proved in Theorem 3.1, and write,

$$\begin{aligned} \big(M_n^{(1,q)}\big)^2 &= \Big[ M_n^{(1)} + \tfrac{\gamma\chi_q^*}{(1+\gamma)U(n/k)} \big(1 + o_p(1)\big)\Big]^2 \\ &= \big(M_n^{(1)}\big)^2 + \tfrac{2\gamma^2\chi_q^*}{(1+\gamma)U(n/k)} \big(1 + o_p(1)\big). \end{aligned} \tag{3.9}$$

Combining (3.8), (3.9) and the definition of the Moment estimator (1.9):

$$\hat{\gamma}^{M(q)} = M_n^{(1,q)} + \frac{\frac{1}{2}M_n^{(2,q)} - \left(M_n^{(1,q)}\right)^2}{M_n^{(2,q)} - \left(M_n^{(1,q)}\right)^2} = M_n^{(1)} + \frac{\frac{1}{2}M_n^{(2)} - \left(M_n^{(1)}\right)^2 + V_2}{M_n^{(2)} - \left(M_n^{(1)}\right)^2 + V_3}, \qquad (3.10)$$

where

$$V_2 = \frac{2\chi_p}{U(n/k)}\Big[\frac{\gamma}{2} + \frac{\gamma}{2(1+\gamma)^2} - \frac{\gamma^2}{(1+\gamma)}\Big]\big(1+o_p(1)\big)$$

and

$$V_3 = \frac{2\chi_p}{U(n/k)}\Big[\gamma + \frac{\gamma}{(1+\gamma)^2} - \frac{\gamma^2}{(1+\gamma)}\Big]\big(1+o_p(1)\big).$$

It has been proved in Martins (2000) that, for $r = 1, 2$,

$$M_n^{(r)} \stackrel{d}{=} \gamma^r\Gamma(r+1)\Big[1 + \frac{P_n^{(r)}}{\Gamma(r+1)\sqrt{k}} + \frac{(1-\rho)^{-r}-1}{\gamma\rho}A(n/k)\Big] + o_p(A(n/k)), \quad (3.11)$$

where $\big(P_n^{(1)}, P_n^{(2)}\big)$ is an asymptotically binormal r.v. with null mean value and covariance matrix

$$\Sigma_{1,2} = \begin{bmatrix} 1 & 4 \\ 4 & 20 \end{bmatrix}.$$

Now, considering (3.5), (3.10) and (3.11), one obtains

$$\hat{\gamma}_n^{M(q)} = M_n^{(1)} + \frac{\frac{P_n^{(2)}}{2\sqrt{k}} - \frac{2P_n^{(1)}}{\sqrt{k}} + \frac{A(n/k)}{\gamma}\big[\frac{1}{\rho(1-\rho)^2} - \frac{1}{\rho} - \frac{2}{1-\rho}\big] + V_2 + o_p\big(A(n/k)\big)}{1 + \frac{P_n^{(2)}}{\sqrt{k}} - \frac{2P_n^{(1)}}{\sqrt{k}} + \frac{2A(n/k)}{\gamma}\big[\frac{1}{\rho(1-\rho)^2} - \frac{1}{\rho} - \frac{1}{1-\rho}\big] + V_3 + o_p\big(A(n/k)\big)}.$$

Finally, using the first order approximation, $1-x$, for $1/(1+x)$, as $x \longrightarrow 0$, the representation (3.6) follows. □

**Remark 3.1.** Notice that $\sigma_M^2 = \sigma_H^2 + 1$, $c_M = c_H + \frac{\rho}{\gamma(1-\rho)^2}$ and $d_M = (d_H)^2$. Consequently, $\sigma_M > \sigma_H$, $c_M < c_H$ and $d_M < d_H$.

Under the conditions of Theorems 3.1 and 3.2 and with the notations defined in (3.3) and in (3.7), the following results hold:

**Corollary 3.1.** *Let $\mu_1$ and $\mu_2$ be finite constants and let $T$ generically denote either $H$ or $M$.*

**i)** *For $\gamma > -\rho$,*

$$\hat{\gamma}_n^{T(q)} \stackrel{d}{=} \gamma + \frac{\sigma_T}{\sqrt{k}}P_k^T + c_T A(n/k)\Big(1+o_p(1)\Big).$$

*If* $\sqrt{k}\ A\,(n/k) \to \mu_1$, *then*

$$\sqrt{k}\left(\hat{\gamma}_n^{T(q)} - \gamma\right) \underset{n\to\infty}{\overset{d}{\longrightarrow}} Normal\left(\mu_1 c_{_T}, \sigma_{_T}^2\right).$$

**ii)** *For* $\gamma < -\rho$,

$$\hat{\gamma}_n^{T(q)} \overset{d}{=} \gamma + \frac{\sigma_{_T}}{\sqrt{k}} P_k^T + d_{_T} \frac{\chi_q^*}{U(n/k)}\Big(1 + o_p(1)\Big).$$

*If* $\sqrt{k}/U\,(n/k) \to \mu_2$, *then*

$$\sqrt{k}\left(\hat{\gamma}_n^{T(q)} - \gamma\right) \underset{n\to\infty}{\overset{d}{\longrightarrow}} Normal\left(\mu_2 d_{_T} \chi_q^*, \sigma_{_T}^2\right).$$

**iii)** *For* $\gamma = -\rho$,

$$\hat{\gamma}_n^{T(q)} \overset{d}{=} \gamma + \frac{\sigma_{_T}}{\sqrt{k}} P_k^T + \left[c_{_T} A(n/k) + d_{_T} \frac{\chi_q^*}{U(n/k)}\right]\Big(1 + o_p(1)\Big).$$

*If* $\sqrt{k}A\,(n/k) \to \mu_1$ *and* $\sqrt{k}/U\,(n/k) \to \mu_2$, *then*

$$\sqrt{k}\left(\hat{\gamma}_n^{T(q)} - \gamma\right) \underset{n\to\infty}{\overset{d}{\longrightarrow}} Normal\left(\mu_1 c_{_T} + \mu_2 d_{_T} \chi_q^*, \sigma_{_T}^2\right).$$

# 4 HIGH QUANTILE PORT-ESTIMATORS

On the basis of (1.16), we shall now consider the following estimators of $\chi_{p_n}$, functions of the sample of excesses over $X_{n_q:n}$, i.e., of the sample $\underline{X}^{(q)}$ in (1.15):

$$\widetilde{\chi}_{p_n,H}^{(q)} := (X_{n-k_n:n} - X_{n_q:n}) \left(\frac{k_n}{np_n}\right)^{\hat{\gamma}_n^{H(q)}} + X_{n_q:n}, \quad 0 \le q < 1, \tag{4.1}$$

$$\widetilde{\chi}_{p_n}^{M(q)} := (X_{n-k_n:n} - X_{n_q:n}) \left(\frac{k_n}{np_n}\right)^{\hat{\gamma}_n^{M(q)}} + X_{n_q:n}, \quad 0 \le q < 1. \tag{4.2}$$

For these estimators we have the asymptotic distributional representations presented in Theorem 4.1.

**Theorem 4.1.** *In Hall's class* (1.6), *for intermediate sequences* $k_n$ *that satisfy*

$$\log(np_n)/\sqrt{k_n} \to 0, \ \ as\ n \to \infty, \tag{4.3}$$

*with* $p_n$ *such that* (1.1) *holds, then, with* $T$ *denoting either* $H$ *or* $M$, $(c_{_H}, d_{_H}, \sigma_{_H})$ *and* $(c_{_M}, d_{_M}, \sigma_{_M})$ *defined in* (3.3) *and* (3.7), *respectively, and for any real* $q$, $0 \le q < 1$,

$$\frac{\sqrt{k_n}}{\sigma_T \log(k_n/(np_n))}\left(\frac{\widetilde{\chi}_{p_n}^{T(q)}}{\chi_{p_n}} - 1\right) = P_k^T + \sqrt{k_n}\left(c_{_T} A(n/k) + d_{_T} \frac{\chi_q^*}{U(n/k)}\right)\Big(1 + o_p(1)\Big),$$

*where* $P_k^T$ *is an asymptotically standard normal r.v.*

*Proof:* From now on, we denote $a_n := \frac{k_n}{np_n}$. With the underlying conditions in (1.1), $a_n$ tends to infinity, as $n \to \infty$, and the quantile to be estimated can be expressed as

$$\chi_{p_n} = U\left(\frac{1}{p_n}\right) = U\left(\frac{na_n}{k_n}\right).$$

We will present the proof for $T = H$, since for $T = M$ the proof follows the same steps with the use of the representation (3.6) instead of the representation (3.2). First notice that

$$\begin{aligned} \widetilde{\chi}_{p_n}^{H(q)} &= (X_{n-k_n:n} - X_{n_q:n})\, a_n^{\hat{\gamma}_n^{H(q)}} + X_{n_q:n} \\ &= X_{n-k_n:n}\left[\left(1 - \frac{X_{n_q:n}}{X_{n-k_n:n}}\right) a_n^{\hat{\gamma}_n^{H(q)}} + \frac{X_{n_q:n}}{X_{n-k_n:n}}\right]. \end{aligned}$$

Now, since $X_{n_q:n} \xrightarrow{p} \chi_q^*$, we have $\frac{X_{n_q:n}}{X_{n-k_n:n}} = o_p(1)$. Then

$$\widetilde{\chi}_{p_n}^{H(q)} = X_{n-k_n:n}\left[a_n^{\hat{\gamma}_n^{H(q)}}(1 + o_p(1))\right],$$

which means that the proposed estimator $\widetilde{\chi}_{p_n}^{H(q)}$ is asymptotically equivalent to the Weissman type estimator (1.7), whenever we use the consistent estimator $\hat{\gamma}_n \equiv \hat{\gamma}_n^{H(q)}$.

Consider now a convenient representation for the difference,

$$\widetilde{\chi}_{p_n}^{H(q)} - \chi_{p_n} = X_{n-k_n:n}\left\{a_n^{\hat{\gamma}_n^{H(q)}} - a_n^{\hat{\gamma}_n^{H(q)}}\left(\frac{X_{n_q:n}}{X_{n-k_n:n}}\right) + \frac{X_{n_q:n}}{X_{n-k_n:n}} - \frac{\chi_{p_n}}{X_{n-k_n:n}}\right\},$$

and recall that we may write

$$\frac{\chi_{p_n}}{X_{n-k_n:n}} = \frac{U(\frac{n}{k_n}a_n)}{U(\frac{n}{k_n})}\frac{U(\frac{n}{k_n})}{U(Y_{n-k_n:n})}.$$

According to (1.5), for $\rho < 0$, $U(\frac{n}{k_n}a_n)/U(\frac{n}{k_n}) = a_n^\gamma\,(1 - A(n/k_n)/\rho)\,(1 + o_p(1))$.

Considering that for the estimator $\hat{\gamma}_n^{H(q)}$, the representation (3.2) holds, we get successively, for sequences $k_n$ that verify (4.3),

$$a_n^{\hat{\gamma}_n^{H(q)}} = a_n^\gamma\Big(1 + \log a_n\left(\hat{\gamma}_n^{H(q)} - \gamma\right)\Big)(1 + o_p(1))$$

and

$$\begin{aligned} \widetilde{\chi}_{p_n}^{H(q)} - \chi_{p_n} &= a_n^\gamma X_{n-k_n:n}\Big\{1 + \log a_n\left(\hat{\gamma}_n^{H(q)} - \gamma\right)(1 + o_p(1)) \\ &\qquad - (1 - A(n/k_n)/\rho)\,(1 + o_p(1))\Big\} \\ &= a_n^\gamma X_{n-k_n:n}\Big\{\log a_n\left(\hat{\gamma}_n^{H(q)} - \gamma\right) + A(n/k_n)/\rho\Big\}(1 + o_p(1)). \end{aligned}$$

Now, we consider the following representation for intermediate statistics, proved in Ferreira et al. (2003),

$$X_{n-k_n:n} \stackrel{d}{=} U\left(\frac{n}{k_n}\right)\left(1 + \frac{\gamma B_k}{\sqrt{k_n}} + o_p\left(\frac{1}{\sqrt{k_n}}\right) + o_p(A(n/k_n))\right), \tag{4.4}$$

with $B_k$ an asymptotically standard normal r.v.

Using (3.2) and (4.4), we may write

$$\widetilde{\chi}_{p_n}^{H(q)} - \chi_{p_n} = U\left(\frac{n}{k_n}\right) a_n^{\gamma} \left(1 + O_p(1/\sqrt{k_n}\right) \left\{W_n + A\left(\frac{n}{k_n}\right)/\rho\right\} (1 + o_p(1)),$$

where

$$\begin{aligned} W_n &= \log a_n \left(\hat{\gamma}_n^{H(q)} - \gamma\right) \\ &= \log a_n \left( \frac{\sigma_H}{\sqrt{k_n}} P_k^H + \left( c_H A(n/k) + d_H \frac{\chi_q^*}{U(n/k)}\right)\left(1 + o_p(1)\right)\right), \end{aligned}$$

with $P_k^H$ independent of the random sequence $B_k$ in (4.4).

Consequently,

$$\frac{\widetilde{\chi}_{p_n}^{H(q)} - \chi_{p_n}}{a_n^{\gamma} U(\frac{n}{k_n})} = \{W_n + A(n/k)/\rho\} (1 + o_p(1))$$

and

$$\frac{\sqrt{k_n}}{\sigma_H \log a_n} \left( \frac{\widetilde{\chi}_{p_n}^{H(q)}}{\chi_{p_n}} - 1\right) = P_k^H + \sqrt{k_n}\left( c_H A(n/k) + d_H \frac{\chi_q^*}{U(n/k)}\right)\left(1 + o_p(1)\right).$$

□

The following result is a direct consequence of Corollary 3.1 and Theorem 4.1.

**Corollary 4.1.** *Under the same conditions of Theorem 4.1, then, with $T$ replaced by $H$ or $M$, and $(c_H, d_H, \sigma_H)$ and $(c_M, d_M, \sigma_M)$ defined in* (3.3) *and* (3.7)*, respectively, the following results hold.*

**i)** *For $\gamma > -\rho$,*

$$\frac{\sqrt{k_n}}{\sigma_T \log(k_n/(np_n))} \left( \frac{\widetilde{\chi}_{p_n}^{T(q)}}{\chi_{p_n}} - 1\right) = P_k^T + \sqrt{k_n}\left(c_T A(n/k)\right)\left(1 + o_p(1)\right),$$

*If $\sqrt{k_n} A(n/k_n) \to \mu_1$, finite, as $n \to \infty$, then the mean value is $\mu_1 c_T$.*

**ii)** *For $\gamma < -\rho$,*

$$\frac{\sqrt{k_n}}{\sigma_T \log(k_n/(np_n))} \left( \frac{\widetilde{\chi}_{p_n}^{T(q)}}{\chi_{p_n}} - 1\right) = P_k^T + \sqrt{k_n}\left(d_T \frac{\chi_q^*}{U(n/k_n)}\right)\left(1 + o_p(1)\right),$$

*If $\sqrt{k_n}/U(n/k_n) \to \mu_2$, finite, as $n \to \infty$, then the mean values is $\mu_2 d_T \chi_q^*$ .*

**iii)** *For* $\rho = -\gamma$,

$$\frac{\sqrt{k_n}}{\sigma_T \log(k_n/(np_n))}\left(\frac{\widetilde{\chi}_{p_n}^{T(q)}}{\chi_{p_n}} - 1\right) = P_k^T$$
$$+ \sqrt{k_n}\Big(c_T A(n/k) + d_T \frac{\chi_q^*}{U(n/k_n)}\Big)\Big(1 + o_p(1)\Big),$$

*If* $\sqrt{k_n}A(n/k_n) \to \mu_1$, *finite, and* $\sqrt{k_n}/U(n/k_n) \to \mu_2$, *finite, as* $n \to \infty$, *then the mean value is* $\mu_1 c_T + \mu_2 d_T \chi_q^*$.

# 5 SIMULATIONS

Here, we compare the finite sample behavior of the proposed high quantile estimators $\widetilde{\chi}_{p_n}^{H(q)}$ in (4.1) and $\widetilde{\chi}_{p_n}^{M(q)}$ in (4.2) with the classical semi-parametric estimators $\widehat{\chi}_{p_n}^{H}$ and $\widehat{\chi}_{p_n}^{M}$ in (1.11). We have generated $N = 200$ independent replicates of sample size $n = 1000$ from the following models:

- Generalized Pareto Model: $X \frown \mathrm{GP}(\gamma)$ $(\rho = -\gamma)$, $\gamma = 1$, with d.f.

$$F(x) = 1 - (1 + \gamma x)^{-1/\gamma}, \quad x \geq 0.$$

- Burr Model: $X \frown \mathrm{Burr}(\gamma, \rho)$, $\gamma = 1$, $\rho = -2, -0.5$, with d.f.

$$F(x) = 1 - (1 + x^{-\rho/\gamma})^{1/\rho}, \quad x \geq 0.$$

- Cauchy Model: $X \frown$ Cauchy, $\gamma = 1$, $\rho = -2$, with d.f.

$$F(x) = \frac{1}{2} + \frac{1}{\pi}\text{arctang } x, \quad x \in \mathbb{R}.$$

At a first stage, we generate samples from the standard model. At a second stage, we introduce a positive shift $\lambda = \chi_{0.01}$, i.e., a new location chosen in a comparable basis as the percentile 99% of the starting point distribution.

We estimate a high quantile $\chi_{0.001}$, for each model, and we present patterns of Mean Values and Root of Mean Squared Errors, plotted against $k = 6, \cdots, 800$.

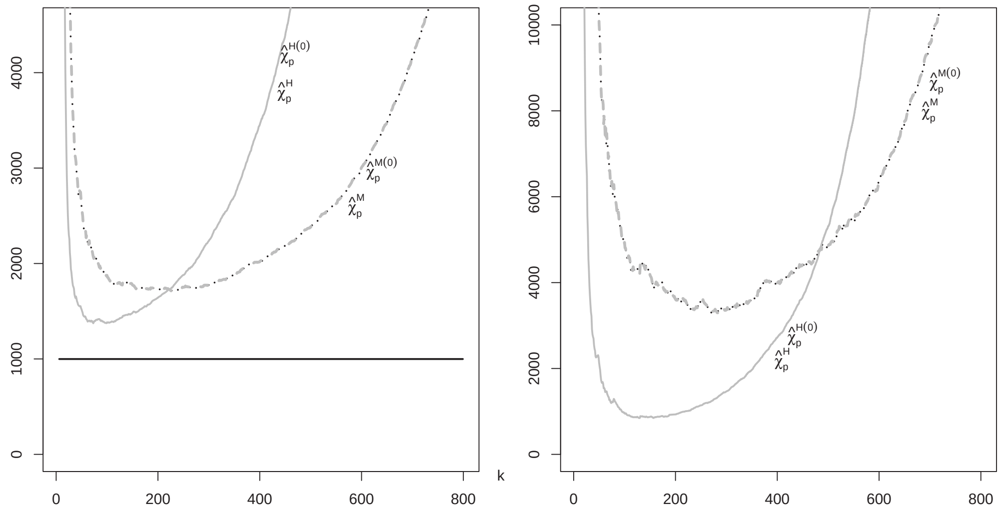


Figure 5.1: Mean values (left) and root mean squared errors (right), of $\widetilde{\chi}_{p_n}^{H(0)}$, $\widetilde{\chi}_{p_n}^{M(0)}$, $\widehat{\chi}_{p_n}^{H}$ and $\widehat{\chi}_{p_n}^{M}$, for a sample size $n = 1000$, from a Generalized Pareto model with $\gamma = 1$, $\rho = -1$ and $\lambda = 0$ (target quantile $\chi_{0.001} = 999$).

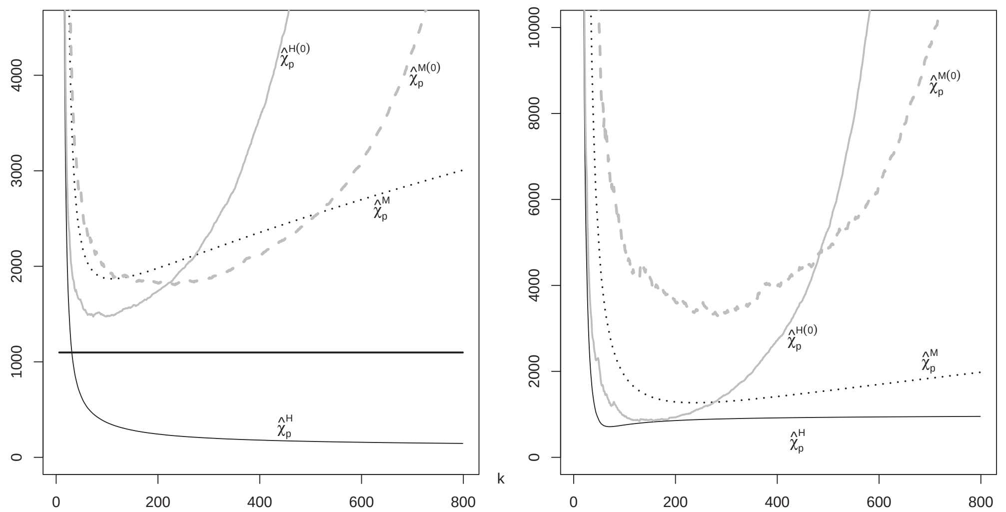


Figure 5.2: Mean values (left) and root mean squared errors (right), of $\widetilde{\chi}_{p_n,H}^{(0)}$, $\widetilde{\chi}_{p_n}^{M(0)}$, $\widehat{\chi}_{p_n}^{H}$ and $\widehat{\chi}_{p_n}^{M}$, for a sample size $n = 1000$, from a Generalized Pareto model with $\gamma = 1$, $\rho = -1$ and $\lambda = 99$ (target quantile $\chi_{0.001} = 1098$).

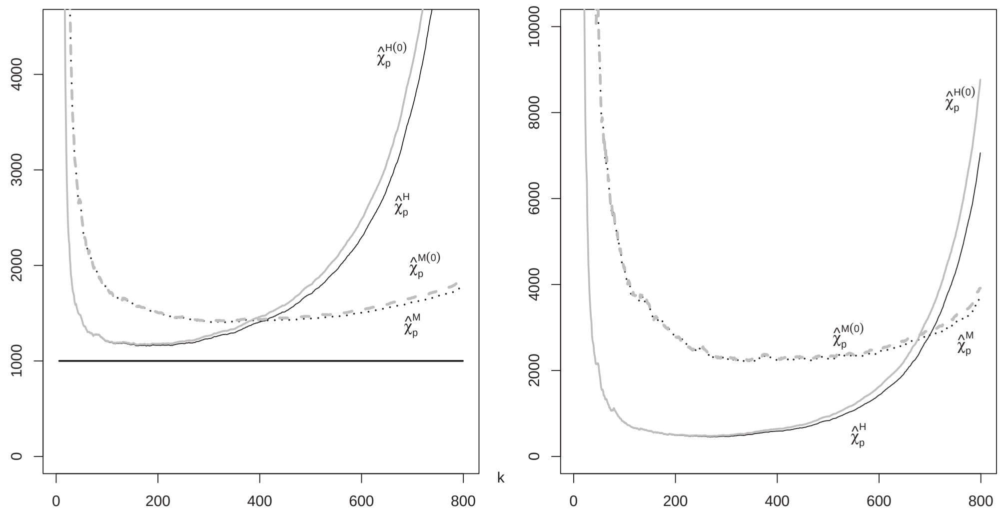


Figure 5.3: Mean values (left) and root mean squared errors (right), of $\widetilde{\chi}^{H(0)}_{p_n}$, $\widetilde{\chi}^{M(0)}_{p_n}$, $\widehat{\chi}^{H}_{p_n}$ and $\widehat{\chi}^{M}_{p_n}$, for a sample size $n = 1000$, from a Burr model with $\gamma = 1$, $\rho = -2$ and $\lambda = 0$ (target quantile $\chi_{0.001} = 1000$).

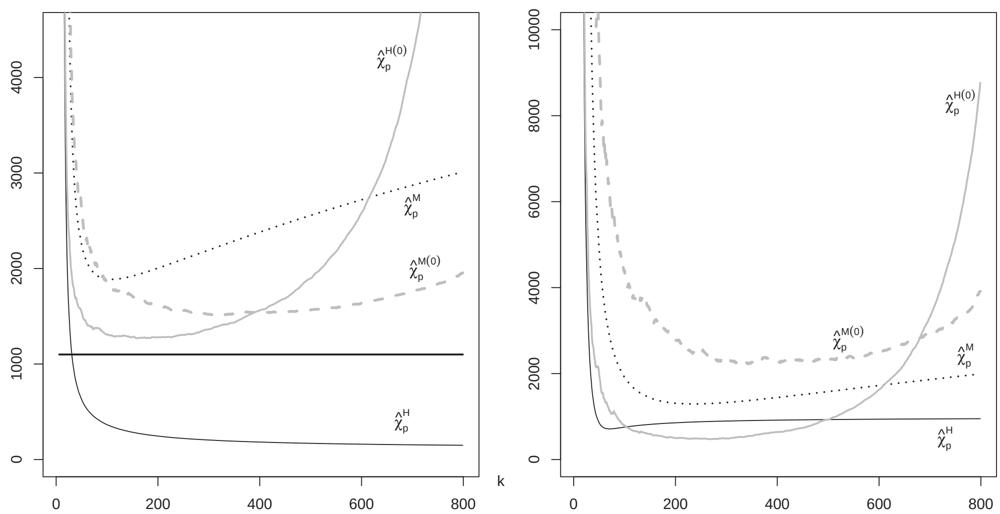


Figure 5.4: Mean values (left) and root mean squared errors (right), of $\widetilde{\chi}^{H(0)}_{p_n}$, $\widetilde{\chi}^{M(0)}_{p_n}$, $\widehat{\chi}^{H}_{p_n}$ and $\widehat{\chi}^{M}_{p_n}$, for a sample size $n = 1000$, from a Burr model with $\gamma = 1$, $\rho = -2$ and $\lambda = 99.99$ (target quantile $\chi_{0.001} = 1099.99$).

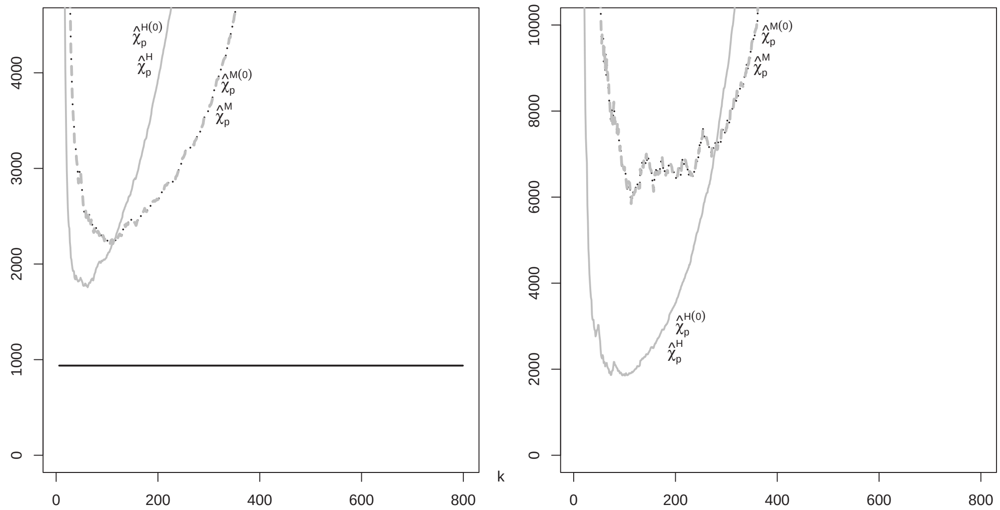


Figure 5.5: Mean values (left) and root mean squared errors (right), of $\widetilde{\chi}_{p_n}^{H(0)}$, $\widetilde{\chi}_{p_n}^{M(0)}$, $\widehat{\chi}_{p_n}^{H}$ and $\widehat{\chi}_{p_n}^{M}$, for a sample size $n = 1000$, from a Burr model with $\gamma = 1$, $\rho = -0.5$ and $\lambda = 0$ (target quantile $\chi_{0.001} = 937.731$).

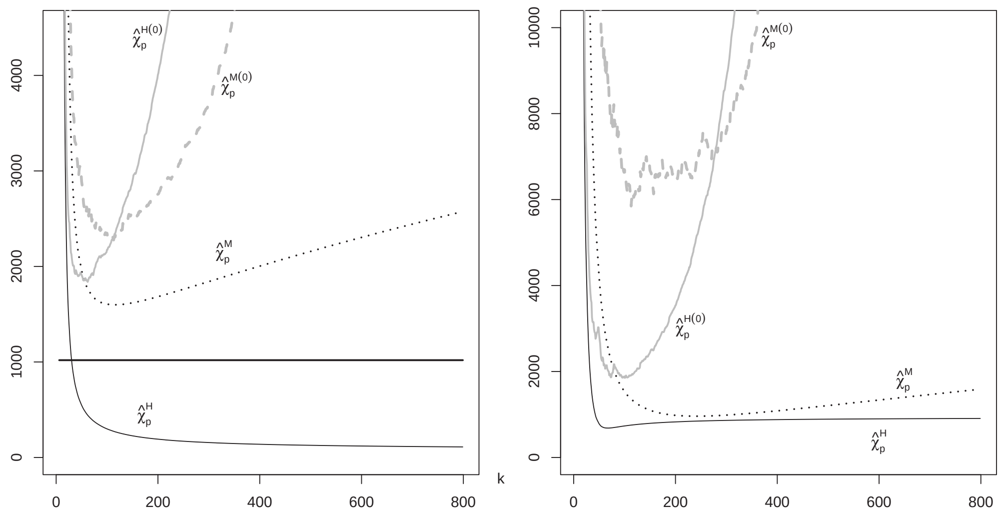


Figure 5.6: Mean values (left) and root mean squared errors (right), of $\widetilde{\chi}_{p_n}^{H(0)}$, $\widetilde{\chi}_{p_n}^{M(0)}$, $\widehat{\chi}_{p_n}^{H}$ and $\widehat{\chi}_{p_n}^{M}$, for a sample size $n = 1000$, from a Burr model with $\gamma = 1$, $\rho = -0.5$ and $\lambda = 81.023$ (target quantile $\chi_{0.001} = 1018.754$).

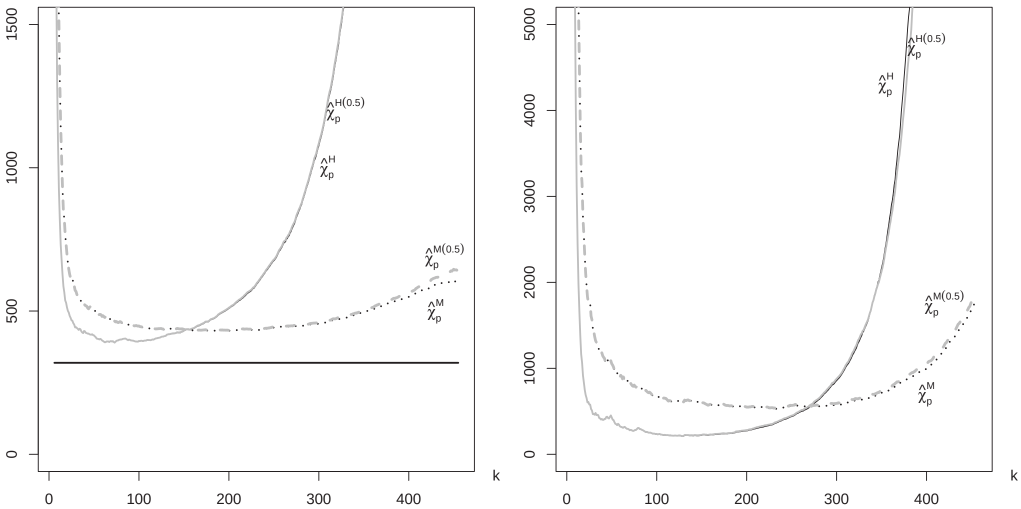


Figure 5.7: Mean values (left) and root mean squared errors (right), of $\widetilde{\chi}^{H(0.5)}_{p_n}$, $\widetilde{\chi}^{M(0.5)}_{p_n}$, $\widehat{\chi}^{H}_{p_n}$ and $\widehat{\chi}^{M}_{p_n}$, for a sample size $n = 1000$, from a Cauchy model with $\gamma = 1$, $\rho = -2$ and $\lambda = 0$ (target quantile $\chi_{0.001} = 319.309$).

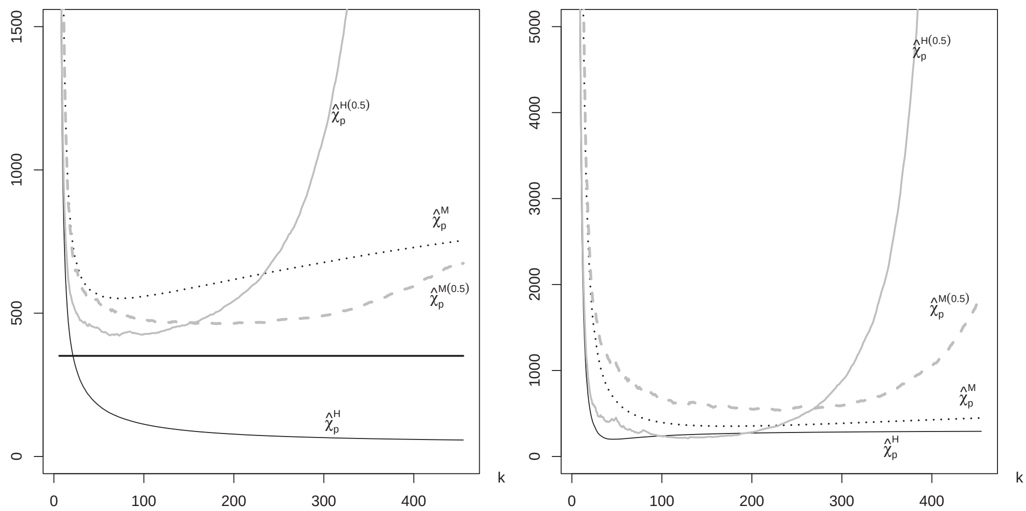


Figure 5.8: Mean values (left) and root mean squared errors (right), of $\widetilde{\chi}^{H(0.5)}_{p_n}$, $\widetilde{\chi}^{M(0.5)}_{p_n}$, $\widehat{\chi}^{H}_{p_n}$ and $\widehat{\chi}^{M}_{p_n}$, for a sample size $n = 1000$, from a Cauchy model with $\gamma = 1$, $\rho = -2$ and $\lambda = 31.821$ (target quantile $\chi_{0.001} = 351.13$).

The simulations illustrate the dramatic disturbance on the behavior of the classical quantile estimators in (1.11), when a shift is introduced. We, again, enhance that the flat stable zones achieved with these estimators, in the presence of shifts, could lead us to dangerous misleading conclusions, unless we are aware of the suitable threshold $k$ or of specific properties of the underlying model.

From the figures, in this section, we observe that the classical quantile estimators diverge a lot from the important linear property (1.13). On the other hand, the estimators we propose, (4.1) and (4.2), enjoy exactly this property.

# 6 APPLICATION TO A FINANCIAL MARKET

As an empirical example, we place ourselves in a context from finance. We analyze the risk for investors holding short positions in the Nasdaq Composite index, i.e., for investors betting on a fall in the index.

It is a common risk measure for large losses the Value at Risk (VaR) — defined as a large quantile of negative log-returns, i.e., of $L_i = -\ln(S_{i+1}/S_i)$, $1 \le i \le n-1$, with $S_i$, $1 \le i \le n$, a sample of consecutive close prices. For details about VaR see for example Holton (2003). Here, since we are interested in the analysis of the risk of holding short positions, we begin with the positive log-returns, i.e., with $X_i = \ln(S_{i+1}/S_i) = -L_i$, $1 \le i \le n-1$, assumed to be stationary and weakly dependent. With the purpose of comparison with a case study from Drees (2003), we have used the daily log-returns from 1997 to 2000, which corresponds to a sample size $n = 1037$. We present a scatterplot of the returns $X_i$ in Figure 6.1. Although there is some increasing trend in the volatility, stationarity is assumed, under the same considerations as in Drees (2003).

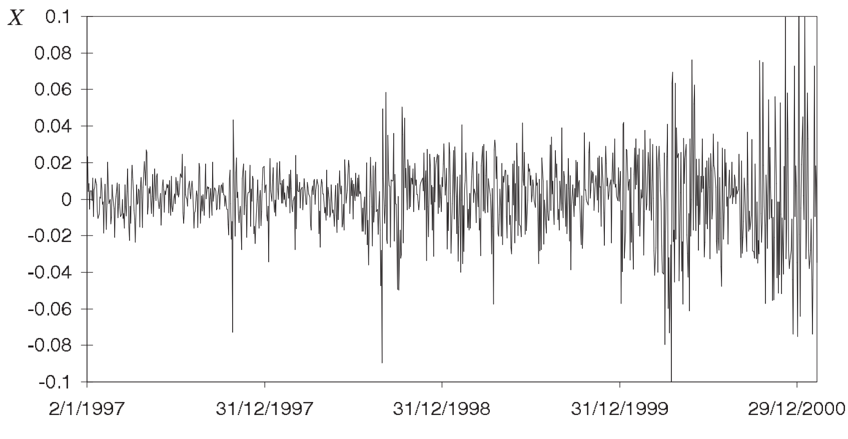


Figure 6.1: Log-returns of the Nasdaq Composite index, 1997-2000.

In Figure 6.2 we display the estimates for the tail index (left), associated to $\hat{\gamma}_n^{H(q)}$, $\hat{\gamma}_n^{M(q)}$, $\hat{\gamma}_n^{H}$, $\hat{\gamma}_n^{M}$ and for the high quantile $\chi_{0.001}$ (right), associated to $\widetilde{\chi}_{p_n}^{H(q)}$, $\widetilde{\chi}_{p_n}^{M(q)}$, $\widehat{\chi}_{p_n}^{H}$ and $\widehat{\chi}_{p_n}^{M}$, with $q = 0, 0.05, 0.25, 0.5$.

It is clear from the analysis of the $\gamma$-scatterplots that all estimates are positive for $k$ from about 50 up to 450, i.e., there is a strong evidence for a heavy-tailed underlying distribution. However, the patterns exhibited by the different estimators $\hat{\gamma}_n^{H(q)}$ are significantly different for different values of the *tuning* parameter $q$.

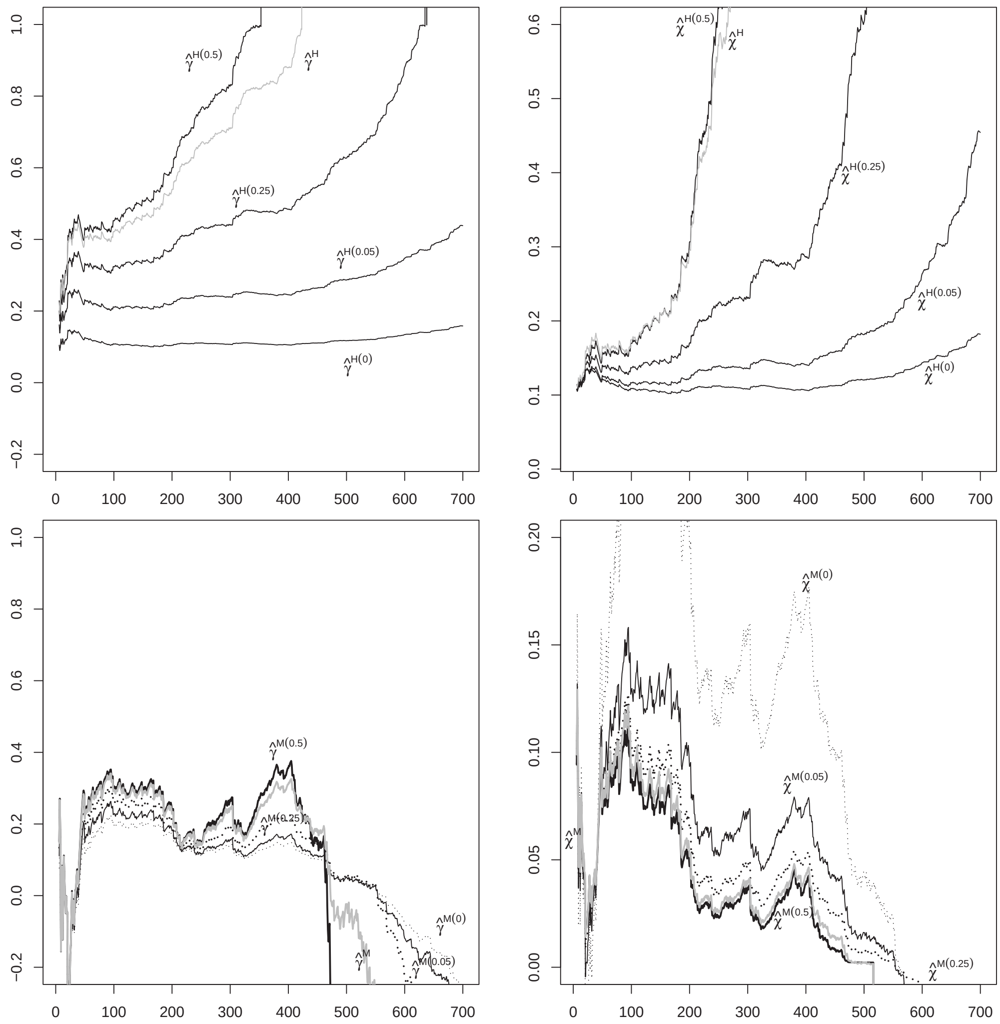


Figure 6.2: Tail index estimates (left) from $\hat{\gamma}_n^{H(q)}$, $\hat{\gamma}_n^{M(q)}$, $\hat{\gamma}_n^{H}$, $\hat{\gamma}_n^{M}$ and high quantile $\chi_{0.001}$ estimates (right) from $\widetilde{\chi}_{p_n}^{H(q)}$, $\widetilde{\chi}_{p_n}^{M(q)}$, $\widehat{\chi}_{p_n}^{H}$ and $\widehat{\chi}_{p_n}^{M}$, with $q = 0, 0.05, 0.25, 0.5$.

We have been particularly puzzled with the sample path of $\hat{\gamma}_n^{H(0)}$, and such a sample path immediately suggested us the possible non-consistency of $\hat{\gamma}_n^{H(0)}$ due to an infinite left endpoint of the underlying model. We have thus decided to analyze more deeply both tails of the model underlying the sample $X_i, 1 \le i \le 1036$. For that we have used not only the Hill estimator, but also two second order reduced bias estimators, recently introduced in the literature, which are alternatives to the Hill estimator not only at the optimal levels or for large $k$, as happens with the "classical" second order reduced bias tail index estimators, but for all $k$. Those estimators are valid in Hall's class of models in (1.6), where we may choose $A(t) = \rho\, D\, t^\rho =: \gamma\, \beta\, t^\rho$. The first of those estimators has been considered in Caeiro *et al.* (2005), and is given by

$$\overline{H}(k) \equiv \overline{H}_{\widehat{\beta},\ \widehat{\rho}}(k) := \widehat{\gamma}_n^H(k) \left(1 - \frac{\widehat{\beta}}{1-\widehat{\rho}} \left(\frac{n}{k}\right)^{\widehat{\rho}}\right), \tag{6.1}$$

where $(\widehat{\beta},\ \widehat{\rho})$ is an adequate consistent estimator of $(\beta,\ \rho)$, with both $\widehat{\beta}$ and $\widehat{\rho}$ based on a number of top o.s. $k_1$ of a larger order than the number of top o.s. $k$ used for the tail index estimation. In a similar way, and denoting the scaled log-spacings by $U_i := i\,(\ln X_{n-i+1:n} - \ln X_{n-i:n})$, $1 \le i \le k$, Gomes *et al.* (2005) considered the tail index estimator

$$\overline{M}(k) \equiv \overline{M}_{\widehat{\beta},\widehat{\rho}}(k) := \widehat{\gamma}_n^H(k) - \frac{\widehat{\beta}}{k}\left(\frac{n}{k}\right)^{\widehat{\rho}} \sum_{i=1}^{k} \left(\frac{i}{k}\right)^{-\widehat{\rho}} U_i. \tag{6.2}$$

Notice that the estimators in (6.1) and (6.2) are both bias-corrected Hill estimators: the dominant component of the bias of Hill's estimator, provided in (3.5) and given by $\gamma\,\beta(n/k)^{\rho}/(1-\rho)$, is estimated in two different ways and directly removed from the Hill estimator $\widehat{\gamma}_n^H(k)$ in (1.8). The study of adapted second order reduced bias tail index and quantile estimators will also be an interesting topic of research, surpassing the scope of this paper.

*Right Tail analysis of Nasdaq data in Figure 6.1* — In Figure 6.3, and working with the $n_0 = 570$ positive values of the log-returns $X_i$ on NASDAQ data, we picture the sample paths of the two most common estimators of the second order parameter $\rho$ (see, for instance, Gomes and Pestana (2005), for an algorithm related to the second order parameters' estimation). The two sample paths of the $\rho$-estimates provided, $\widehat{\rho}_0(k)$ and $\widehat{\rho}_1(k)$, lead us to choose, on the basis of any stability criterion for large values of $k$, the estimate associated to $\tau = 0$. We have considered $\widehat{\rho} = \widehat{\rho}_0(k_1)$, with $k_1 = n_0^{0.995}$. We have got $\widehat{\rho}_0 = \widehat{\rho}_0(552) = -0.71$. The use of the $\beta$-estimate suggested in the above mentioned algorithm, leads us to the estimate $\widehat{\beta}_0 = 1.04$. For the estimation of $\gamma$ through the reduced bias tail index estimators, we have used the heuristic estimate of the level provided in Gomes and Pestana (2005), i.e., the value

$$k_{01} \equiv k_{01}(n;\beta,\rho) = \left(\frac{1.96(1-\rho)n^{-\rho}}{|\beta|}\right)^{2/(1-2\rho)}.$$

Levels of this type are still levels such that $\sqrt{k}\ (n/k)^{\rho} \to \lambda$, finite, and are not yet optimal for the tail index estimation through second order reduced bias' tail index estimators. However, do not forget that with tail index estimators like $\overline{H}$ and $\overline{M}$ in (6.1) and (6.2), respectively, we are always safe and able to provide a more reliable estimation. We came to $\widehat{k}_{01} = 109$ and to the estimate $\widehat{\gamma} = \overline{M}(109) = 0.34$. Note that the estimation of the optimal threshold for the estimation through the Hill estimator in (1.8), leads us to

$$\widehat{k}_0 = \left(\frac{(1-\widehat{\rho})\ n^{-\widehat{\rho}}}{\widehat{\beta}\ \sqrt{-2\widehat{\rho}}}\right)^{2/(1-2\widehat{\rho})} = 55 \quad \Longrightarrow \quad \widehat{\gamma}_n^H(\widehat{k}_0) = 0.33.$$

*Left Tail analysis of Nasdaq data in Figure 6.1* — Figure 6.4 is related to a similar data analysis, carried on the $n_0 = 466$ positive values of negative log-returns $L_i = -\log(S_i/S_{i-1})$. We have now obtained $\widehat{\rho} = -0.71$, $\widehat{\beta} = 1.05$, $\widehat{\gamma}_n^H(\widehat{k}_0) = 0.35$ and $\widehat{\gamma} = \overline{M}(97) = 0.30$.

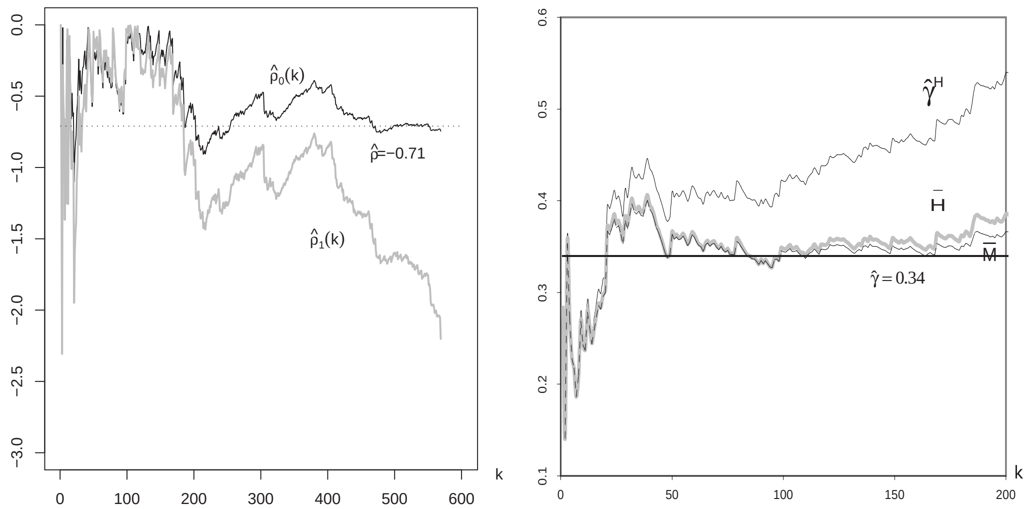


Figure 6.3: Estimates of the second order parameters $\rho$, through $\widehat{\rho}_0(k)$ and $\widehat{\rho}_1(k)$ (*left*), and the tail index $\gamma$ (*right*), for the positive log-returns $X_i = \log(S_i/S_{i-1})$ — *Right Tail of NASDAQ data.*

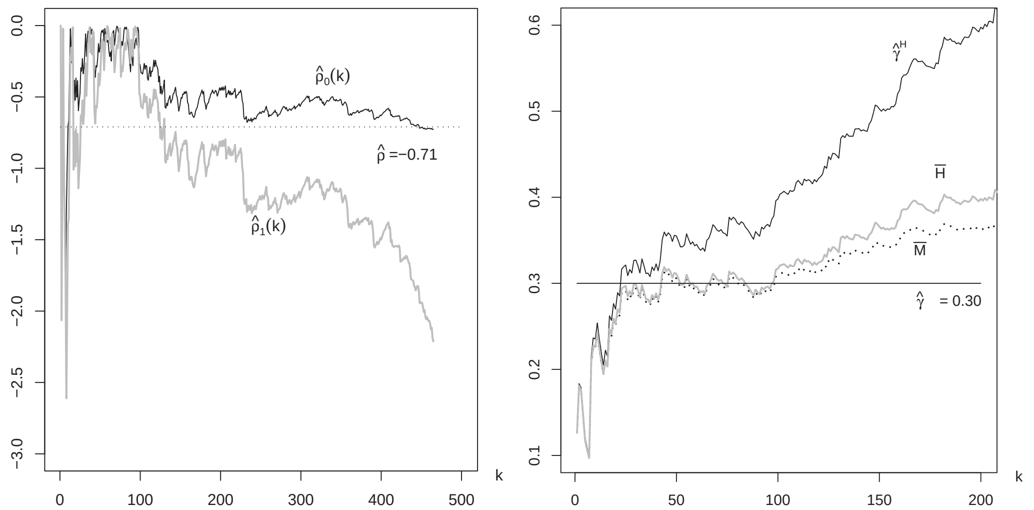


Figure 6.4: Estimates of the second order parameters $\rho$, through $\widehat{\rho}_0(k)$ and $\widehat{\rho}_1(k)$ (*left*), and the tail index $\gamma$ (*right*), for negative log-returns $L_i = -\log(S_i/S_{i-1})$ — *Left Tail of NASDAQ data.*

This data analysis leads us immediately to the conclusion that the underlying model is almost symmetric around 0, in what concerns the heaviness of the tails. Indeed, when we induce a shift associated to the tuning parameter $q = 0.5$, we get a sample path almost overlapping that of the Hill estimator. Both tails are quite similar, with the right tail underlying to the $X_i$ slightly heavier than the left tail ($\widehat{\gamma} = 0.35$ for the right tail versus $\widehat{\gamma} = 0.30$ for the left tail). This obviously implies an underlying model with support $(-\infty, +\infty)$. It is then not at all sensible to induce a shift $X_{1,n}$, like it is suggested in Drees (2003). Such a shift is appealing, because it induces for the Hill estimator an almost flat sample path (see Figure 6.2), but as mentioned before, the flat zone leads, in this case, to a severe underestimation of the tail index $\gamma$.

To support this statement, we still present in Figure 6.5 the pattern of mean values and mean squared errors of the Hill and Moment PORT-estimators, for models from a Student-$t$ parent with $\nu = 4$ degrees of freedom.

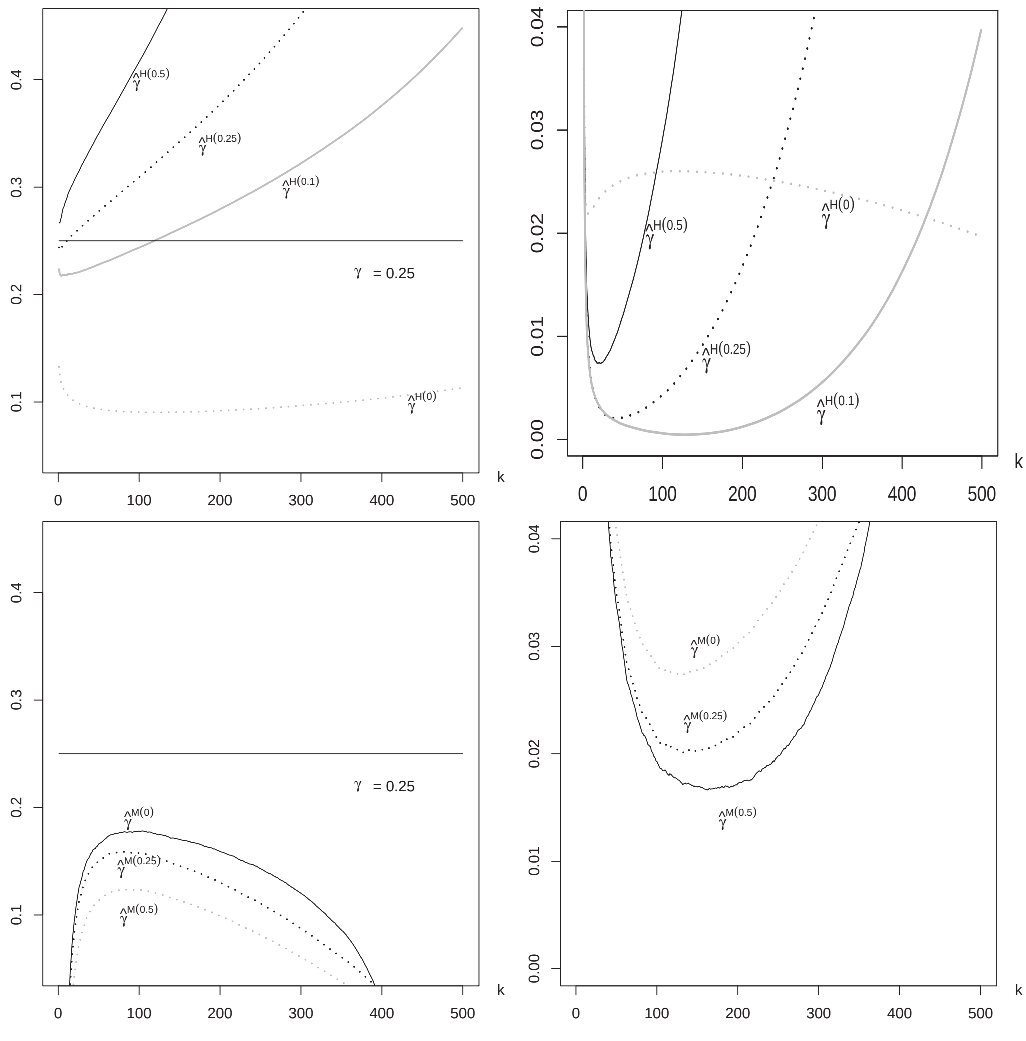


Figure 6.5: Mean values ($left$) and mean squared errors ($right$) of Hill and Moment PORT-estimators associated to a Student-$t$ d.f., with $\nu = 4$ ($\gamma = 0.25$, $\rho = -0.5$).

The similarities between the behavior of the mean value patterns in Figure 6.5 and the sample paths of the Hill and Moment PORT-estimators in Figure 6.2, suggests that the d.f. underlying these returns is not a long way from a Student d.f., a very common model in the area of extremes and finance (for a recent reference see McNeil *et al.* (2005)). However, a parametric data analysis of this data is totally outside the scope of the present paper.

In this application, and taking into account the previous analysis, it seems sensible to consider as a compromise choice, the shift induced by the first empirical quartile, i.e., to pick up the value $q = 0.25$.

## References


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

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

[3] Drees, H. (2003). Extreme quantile estimation for dependent data, with applications to finance. *Bernoulli* **9**(4), 617-657.

[4] Ferreira, A. de Haan, L. and Peng, L. (2003). On optimizing the estimation of high quantiles of a probability distribution. *Statistics* **37**, 401-434.

[5] Fraga Alves, M.I. and Araújo Santos, P. (2004). *Extreme Quantiles Estimation with Shifted Data From Heavy Tails.* Notas e Comunicações CEAUL 11/2004.

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

[7] Gomes, M.I. and Figueiredo, F. (2002) Bias Reduction in risk modelling: semi-parametric quantile estimation, (to appear in *Test*).

[8] Gomes, M. I., Martins, M. J. and Neves, M. (2005). *Revisiting the second order reduced bias "maximum likelihood" tail index estimators.* Notas e Comunicações CEAUL 10/2005. Submitted.

[9] 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. *Commun. Stat., Simulation Comput.* **32**(4), 1005-1028.

[10] Gomes, M. I. and Pestana, D. (2005). A sturdy second order reduced bias' Value at Risk estimator. (to appear in *J. Amer. Statist. Assoc.*).

[11] Hall, P. (1982). On some Simple Estimates of an Exponent of Regular Variation. *J. R. Statist. Soc.* **44**, no. 1, 37-42.

[12] Hill, B.M. (1975). A Simple General Approach to Inference about the Tail of a Distribution. *Ann. Statist.* **3**, no. 5, 1163-1174.

[13] Holton, G.A. (2003). *Value-at-Risk Theory and Practice.* Academic Press.

[14] Martins, M.J.(2000). *Estimacão de Caudas Pesadas - Variantes ao Estimador de Hill.* Tese de Doutoramento, D.E.I.O., Faculdade de Ciências da Universidade de Lisboa.

[15] McNeil, A. J., Frey, R. and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools.* Princeton University Press, Princeton.

[16] Weissman, I. (1978). Estimation of Parameters and Large Quantiles Based on the $k$ Largest Observations. *J. Amer. Statist. Assoc.*, **73**, 812-815.