An L^p-quantile methodology for tail index estimation

Stéphane Girard[a], Gilles Stupfler[b] & Antoine Usseglio-Carleve[a]

[a] Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France
[b] Univ. Rennes, Ensai, CNRS, CREST - UMR 9194, F-35000 Rennes, France

Abstract. Quantiles are a fundamental concept in extreme-value theory. They can be obtained from a minimization framework using an absolute error loss criterion. The companion notion of expectiles, based on squared rather than absolute error loss minimization, has recently been receiving substantial attention from the fields of actuarial science, finance and econometrics. Both of these notions can actually be embedded in a common framework of L^p-quantiles, whose extreme value properties have been explored very recently. However, and even though this generalized notion of quantiles has shown potential for the estimation of extreme quantiles and expectiles, it has so far not been used in the estimation of extreme value parameters of the underlying distribution of interest. In this paper, we work in a context of heavy tails, which is especially relevant to actuarial science, finance, econometrics and natural sciences, and we construct an estimator of the tail index of the underlying distribution based on extreme L^p-quantiles. We establish the asymptotic normality of such an estimator and in doing so, we extend very recent results on extreme expectile and L^p-quantile estimation. We provide a discussion of the choice of p in practice, as well as a methodology for reducing the bias of our estimator. Its finite-sample performance is evaluated on simulated data and on a set of real hydrological data.

Keywords. Extrapolation, extremes, heavy tails, L^p-quantile estimation, tail index.

1 Introduction

Carrying out inference about the extremes of a random phenomenon is an important goal in several fields of statistical applications. The first motivating such problem was to determine how high dikes constructed in the Netherlands should be so that areas below sea level can be adequately protected from flooding (de Haan and Ferreira, 2006). Other related examples are inferring extreme rainfall levels at a given location (Koutsoyiannis, 2004) and estimating extreme daily wind speeds (Beirlant et al., 1996). Extreme phenomena may also negatively affect financial institutions or insurance companies; examples of studies in this direction include the analysis of extreme log-returns of financial time series (Drees, 2003) and the inference about extreme risks linked to large losses in insurance (Rootzén and Tajvidi, 1997). Providing a solution to such problems typically relies on estimating an extreme quantile of a well-chosen univariate random variable.

It is however well-known that the straightforward empirical estimator of an extreme quantile, that is, a quantile exceeding the range covered by the available data, is inconsistent. In many of the above applications, this problem can be solved by modeling the variable of interest using a heavy-tailed distribution: a random variable Y is said to be heavy-tailed when its survival function can be written $\mathbb{P}(Y > y) = y^{-1/\gamma}\ell(y)$, where $\gamma > 0$ and ℓ is a slowly varying function, *i.e.* such that $\ell(ty)/\ell(y) \to 1$ as $y \to \infty$ for all $t > 0$. The parameter γ is the so-called tail index of Y. It can then be shown that the knowledge of this parameter allows one to successfully estimate extreme quantiles by extrapolating beyond the range of the data: a standard example is Weissman's estimator (Weissman, 1978), with more recent alternatives introduced by de Haan and Rootzén (1993) and Li et al. (2010).

The estimation of the tail index γ is therefore a central problem in the extreme value literature. A very popular semiparametric estimator was proposed by Hill (1975). This estimator has several

interpretations, most notably in terms of quasi-maximum likelihood estimation in a Pareto model. This estimator has found a number of competitors over the last forty years: although this is a non-exhaustive list, we refer to the quasi-maximum likelihood estimator in a Generalized Pareto model (Smith, 1987; Drees et al., 2004), probability-weighted moment estimators (Hosking et al., 1987) and the general moment estimators of Dekkers et al. (1989) and de Haan and Peng (1998). More recent, perhaps lesser-known, attempts include McElroy and Politis (2007), Paulauskas and Vaičiulis (2013), Brilhante et al. (2013), Beran et al. (2014) and McElroy and Nagaraja (2016).

The common point of all these estimators is that they are based on quantiles, in the sense that they use high order statistics of the available sample for their calculation. Even though quantiles are straightforward to calculate and interpret, they are not devoid of drawbacks: for instance, a quantile only takes into account information on the frequency of a tail event, ignoring information about its magnitude. It is the goal of this paper to construct a methodology based on a class of extreme value indicators more general than high quantiles and whose calculation is driven both by the frequency and actual magnitude of high values. To this end, we first note that quantiles may equivalently be seen as the solution of the following minimization problem:

$$q(\alpha) \in \arg\min_{t\in\mathbb{R}} \mathbb{E}\left[\rho_\alpha^{(1)}(Y-t) - \rho_\alpha^{(1)}(Y)\right], \tag{1.1}$$

where $\rho_\alpha^{(1)}(y) = \left|\alpha - \mathbb{1}_{\{y<0\}}\right| |y|$ is the quantile loss (or check) function, see Koenker and Bassett (1978). By replacing such an absolute loss with a more general L^p-loss, one obtains the family of L^p-quantiles: if $p \geq 1$ and $\alpha \in (0,1)$, the L_p-quantile $q^{(p)}(\alpha)$ of Y is given by

$$q^{(p)}(\alpha) \in \arg\min_{t\in\mathbb{R}} \mathbb{E}\left[\rho_\alpha^{(p)}(Y-t) - \rho_\alpha^{(p)}(Y)\right], \tag{1.2}$$

where $\rho_\alpha^{(p)}(y) = \left|\alpha - \mathbb{1}_{\{y<0\}}\right| |y|^p$. Clearly $p=1$ leads to the quantile; the case $p=2$, initially studied in Newey and Powell (1987), leads to the so-called expectile. Inference on, and using, extreme expectiles has recently been studied by, among others, Daouia et al. (2018a), Daouia et al. (2018b) and Daouia et al. (2019b). Expectiles have the advantage to be a coherent risk measure (Artzner et al., 1999), contrary to quantiles, making them particularly appealing in actuarial and financial applications (see, among others, Taylor, 2008; Cai and Weng, 2016). The estimation of extreme L^p-quantiles, as a way to recover extreme quantiles and expectiles, is investigated in Daouia et al. (2019a).

The goal of the present paper is to take that idea one step further and use L^p-quantiles as a way to estimate the tail index γ. This was done, specifically in the case $p=2$ of expectiles, by Daouia et al. (2018b), who proposed to replace empirical quantiles by empirical expectiles in the Hill estimator, resulting in an estimator called the expectHill estimator. Such an approach is motivated by general results of Bellini et al. (2014) and Daouia et al. (2019a) showing that high quantiles and L_p-quantiles are asymptotically proportional, with a proportionality coefficient only depending on the tail index γ and the value of p. However, the expectHill estimator suffers from a couple of important drawbacks, among which a tendency to quickly drift away from the true value of γ due to finite-sample bias. Our idea here is to move away from the Hill estimator and exploit the asymptotic relationship between the high quantiles and L_p-quantiles, through what we will show is actually a decreasing function of γ, in order to define a different type of estimator of the tail index. We shall in particular explain why our estimator performs better than the Hill estimator in certain situations that are known to be difficult in the analysis of heavy tails, making it a worthy addition to the extreme value toolbox.

The paper is organized as follows. Section 2 introduces the setting, our assumptions, the Hill and expectHill estimators as well as our newly proposed L^p-quantile based estimator. Section 3 gives the asymptotic properties of this estimator, introduces a correction of the bias purely due to the use of the L^p-quantile methodology, and discusses the choice of the parameter p. Section 4 examines the finite-sample performance of our estimator on a simulation study, and Section 5 showcases our estimator on a set of real data from hydrology. Proofs are postponed to the Appendix.

2 Setting, notations and assumptions

Let $p \geq 1$ and let Y be a real random variable such that $\mathbb{E}[|Y|^{p-1}] < \infty$. Differentiating the cost function (1.2), the L^p-quantile $q^{(p)}(\alpha)$ of level $\alpha \in (0,1)$ of Y is equivalently defined as follows:

$$\bar{F}^{(p)}(y) = \frac{\mathbb{E}\left[|Y-y|^{p-1}\mathbb{1}_{\{Y>y\}}\right]}{\mathbb{E}\left[|Y-y|^{p-1}\right]}\ ,\ q^{(p)}(\alpha) = \inf\left\{y \in \mathbb{R}\colon \bar{F}^{(p)}(y) \leq 1-\alpha\right\}. \tag{2.1}$$

This is justified because $\bar{F}^{(p)}$ is a decreasing function: see Lemma 1 in the case $p > 1$. Let us recall that the cases $p = 1$ and $p = 2$ respectively lead to the notions of quantile and expectile, see Koenker and Bassett (1978) and Newey and Powell (1987). In view of (2.1), if copies $Y_1, \ldots, Y_n$ of Y are available, an estimator of an L^p-quantile can be obtained by replacing expectations by their empirical counterparts:

$$\hat{\bar{F}}_n^{(p)}(y) = \frac{\sum_{i=1}^n |Y_i-y|^{p-1}\mathbb{1}_{\{Y_i>y\}}}{\sum_{i=1}^n |Y_i-y|^{p-1}}\ ,\ \hat{q}_n^{(p)}(\alpha) = \inf\left\{y \in \mathbb{R}\colon \hat{\bar{F}}_n^{(p)}(y) \leq 1-\alpha\right\}. \tag{2.2}$$

The asymptotic properties of $\hat{q}_n^{(p)}(\alpha)$ are established in Daouia et al. (2019a) for high levels $\alpha \to 1$. For that purpose, an assumption of heavy tails is required on the underlying distribution. Besides, we assume hereafter for the sake of simplicity that **the function $\bar{F}^{(1)}$ is continuous and strictly decreasing on the range of values of Y**. This is not a strong regularity condition, and it can actually be removed altogether using the heavy tails assumption, *via* results such as Lemma 1 in Daouia et al. (2019b), but this adds further technical complications to the proofs.

We thus introduce the following first-order condition:

$\mathcal{C}_1(\gamma)$ The survival function $\bar{F}^{(1)}$ is regularly varying at infinity with index $-1/\gamma < 0$, that is

$$\lim_{y\to\infty} \frac{\bar{F}^{(1)}(ty)}{\bar{F}^{(1)}(y)} = t^{-1/\gamma},\ \forall t > 0. \tag{2.3}$$

According to de Haan and Ferreira (2006, Theorem 1.2.1), such an assumption is equivalent to $\bar{F}^{(1)}$ belongs to the Fréchet maximum domain of attraction, with extreme-value index $\gamma > 0$. Several estimators of γ can be found in the literature, the most popular one being introduced in Hill (1975), and based on ratios of empirical quantiles:

$$\hat{\gamma}_{\alpha_n}^{(H)} = \frac{1}{\lfloor n(1-\alpha_n)\rfloor}\sum_{i=1}^{\lfloor n(1-\alpha_n)\rfloor} \log\left(\frac{\hat{q}_n^{(1)}\left(1-(i-1)/n\right)}{\hat{q}_n^{(1)}\left(1-\lfloor n(1-\alpha_n)\rfloor/n\right)}\right). \tag{2.4}$$

Here, (α_n) is an *intermediate sequence*, *i.e.* such that $\alpha_n \in (0,1)$, $\alpha_n \to 1$ and $n(1-\alpha_n) \to \infty$ as $n \to \infty$. The asymptotic properties of $\hat{\gamma}_{\alpha_n}^{(H)}$ are discussed in *e.g.* de Haan and Resnick (1998) and de Haan and Ferreira (2006). More recently, Daouia et al. (2018b) proposed to replace empirical quantiles by empirical expectiles, leading to the so-called expectHill estimator:

$$\hat{\gamma}_{\alpha_n}^{(EH)} = \frac{1}{\lfloor n(1-\alpha_n)\rfloor}\sum_{i=1}^{\lfloor n(1-\alpha_n)\rfloor} \log\left(\frac{\hat{q}_n^{(2)}\left(1-(i-1)/n\right)}{\hat{q}_n^{(2)}\left(1-\lfloor n(1-\alpha_n)\rfloor/n\right)}\right). \tag{2.5}$$

In this paper, we propose to use the larger class of L^p-quantiles to estimate the tail index. Our motivation is that, while the expectHill estimator has very smooth sample paths, its finite-sample bias tends to set in fairly quickly; it is our intention here to show that an L^p-quantile methodology, for $p \in (1,2)$, can provide a good complement to the Hill and expectHill estimators. For that purpose,

let us recall that, under $\boldsymbol{\mathcal{C}_1(\gamma)}$, it is proved in Daouia et al. (2019a, Proposition 1) that for all $p > 1$ and $0 < \gamma < 1/(p-1)$,

$$\lim_{\alpha \to 1} \frac{\bar{F}^{(1)}\left(q^{(p)}(\alpha)\right)}{1-\alpha} = \frac{\gamma}{B\left(p, \gamma^{-1} - p + 1\right)} =: g_p(\gamma), \tag{2.6}$$

where $B(x,y) = \int_0^1 u^{x-1}(1-u)^{y-1}du$ is the Beta function. The function g_p being strictly decreasing for all $p > 1$ (see Lemma 2 in the Appendix), a natural idea may be to estimate the tail index γ using the asymptotic relationship (2.6). To this end, let us consider an intermediate sequence (α_n) and define the tail index estimator $\hat{\gamma}_{\alpha_n}^{(p)}$ for all $p > 1$ as

$$\hat{\gamma}_{\alpha_n}^{(p)} = \inf\left\{\gamma > 0 \colon g_p(\gamma) \leq \frac{\hat{\bar{F}}_n^{(1)}\left(\hat{q}_n^{(p)}(\alpha_n)\right)}{1-\alpha_n}\right\}. \tag{2.7}$$

From known properties of L^p-quantiles (see Bellini et al. (2014)), it is noteworthy that the new estimator $\hat{\gamma}_{\alpha_n}^{(p)}$ is both translation and scale invariant, by contrast to the Hill estimator which is not location invariant. The particular cases $p \in \{2,3,4,5\}$ yield closed formulas for $\hat{\gamma}_{\alpha_n}^{(p)}$. For instance, $p = 2$ leads to $g_p(\gamma) = \gamma^{-1} - 1$ and therefore to the expectile-based estimator

$$\hat{\gamma}_{\alpha_n}^{(2)} = \left(1 + \frac{\hat{\bar{F}}_n^{(1)}\left(\hat{q}_n^{(2)}(\alpha_n)\right)}{1-\alpha_n}\right)^{-1}. \tag{2.8}$$

This estimator was introduced by Girard et al. (2019) in a conditional setting. The asymptotic normality of $\hat{\gamma}_{\alpha_n}^{(2)}$ is established therein under the condition $\gamma < 1/2$.

Here, we deal with the case $p > 1$ in its full generality. In the next section, we shall establish an asymptotic normality result under the condition $\gamma < 1/(2p-2)$ which is weaker than the condition $\gamma < 1/2$ required for the asymptotic normality of the expectile-based estimator (2.8) as soon as $p \in (1,2)$. To this end, we introduce the following second-order refinement of the first-order regular variation condition:

$\boldsymbol{\mathcal{C}_2(\gamma, \rho, A)}$ The survival function $\bar{F}^{(1)}$ is second-order regularly varying with index $-1/\gamma < 0$, second-order parameter $\rho \leq 0$ and auxiliary function A having constant sign and converging to 0 at infinity, *i.e.*

$$\lim_{y \to \infty} \frac{1}{A\left(\frac{1}{\bar{F}^{(1)}(y)}\right)}\left(\frac{\bar{F}^{(1)}(ty)}{\bar{F}^{(1)}(y)} - t^{-1/\gamma}\right) = t^{-1/\gamma}\frac{t^{\rho/\gamma}-1}{\gamma\rho}, \ \forall t > 0. \tag{2.9}$$

We note that $(t^x - 1)/x$ should be read as $\log(t)$ when $x = 0$.

It is clear that $\boldsymbol{\mathcal{C}_2(\gamma, \rho, A)}$ implies $\boldsymbol{\mathcal{C}_1(\gamma)}$. Let us also highlight that (2.9) is equivalent to de Haan and Ferreira (2006, Theorem 2.3.9)

$$\lim_{y \to \infty} \frac{1}{A(y)}\left(\frac{q^{(1)}(1-1/(ty))}{q^{(1)}(1-1/y)} - t^{\gamma}\right) = t^{\gamma}\frac{t^{\rho}-1}{\rho}, \ \forall t > 0. \tag{2.10}$$

Such a second-order condition allows the control of bias terms in statistical extreme value procedures through the function A, and is thus the cornerstone for asymptotic normality results in extreme-value theory. For instance, it is established in de Haan and Ferreira (2006, Theorem 3.2.5) that, under $\boldsymbol{\mathcal{C}_2(\gamma, \rho, A)}$ and the additional condition $\sqrt{n(1-\alpha_n)}A\left((1-\alpha_n)^{-1}\right) \to \lambda \in \mathbb{R}$ as $n \to \infty$, one has

$$\sqrt{n(1-\alpha_n)}\left(\hat{\gamma}_{\alpha_n}^{(H)} - \gamma\right) \xrightarrow{d} \mathcal{N}\left(b_H(\rho), v_H(\gamma)\right), \tag{2.11}$$

with $b_H(\rho) = \lambda/(1-\rho)$ and $v_H(\gamma) = \gamma^2$. Similarly, the asymptotic normality of the expectHill estimator (2.5) is established in Daouia et al. (2018b, Theorem 2) for $\gamma < 1/2$ under the two additional conditions $\mathbb{E}(Y^2) < \infty$ and $\sqrt{n(1-\alpha_n)}/q^{(1)}(\alpha_n) \to \mu \in \mathbb{R}$ as $n \to \infty$:

$$\sqrt{n(1-\alpha_n)}\left(\hat{\gamma}_{\alpha_n}^{(EH)} - \gamma\right) \xrightarrow{d} \mathcal{N}\left(b_{EH}(\gamma,\rho), v_{EH}(\gamma)\right), \tag{2.12}$$

with

$$b_{EH}(\gamma,\rho) = \frac{(1-\gamma)\left(\gamma^{-1}-1\right)^{-\rho}}{(1-\rho)(1-\gamma-\rho)}\lambda - \mathbb{E}(Y)\frac{\gamma^2\left(\gamma^{-1}-1\right)^{\gamma}}{\gamma+1}\mu \text{ and } v_{EH}(\gamma) = \frac{2\gamma^3}{1-2\gamma}.$$

The objective of the next section is to establish a general asymptotic normality result for $\hat{\gamma}_{\alpha_n}^{(p)}$. We shall also design a method for the reduction of the asymptotic bias and derive the associated asymptotic mean squared error of the estimator.

3 Main results

3.1 Asymptotic analysis

In the sequel, we denote by $IB(t,x,y) = \int_0^t u^{x-1}(1-u)^{y-1}du$ the incomplete Beta function and by Ψ the digamma function *i.e.* the log-derivative of Euler's Gamma function. Our first result provides the joint asymptotic distribution of the L^1- and L^p-empirical quantiles, $p > 1$. Since the calculation of L^p-quantiles depends on the whole of the distribution of Y and in particular requires the assumption $\mathbb{E}(|Y|^{p-1}) < \infty$, an assumption on the left tail of Y is needed to ensure that $\mathbb{E}(|Y|^{p-1})$ exists. We thus assume that $\bar{F}_-^{(1)}$, the survival function of $-Y$, is either light-tailed (i.e. $y^a\bar{F}_-^{(1)}(y) \to 0$ as $y \to +\infty$ for all $a > 0$), or satisfies a first-order condition of the form $\boldsymbol{\mathcal{C}_1(\gamma_\ell)}$. In the case where $\bar{F}_-^{(1)}$ is light-tailed, our subsequent conditions and results may be read with $\gamma_\ell = 0$, and we make the convention $1/0 = \infty$.

Proposition 1. *Assume $\bar{F}^{(1)}$ satisfies $\boldsymbol{\mathcal{C}_2(\gamma,\rho,A)}$, $\bar{F}_-^{(1)}$ is either light-tailed or satisfies $\boldsymbol{\mathcal{C}_1(\gamma_\ell)}$. Let (α_n) and (β_n) be two intermediate sequences such that $(1-\beta_n)/(1-\alpha_n) \to \theta > 0$ as $n \to \infty$. Then, for all $1 < p < (\gamma \vee \gamma_\ell)^{-1}/2 + 1$,*

$$\sqrt{n(1-\alpha_n)}\left(\frac{\hat{q}_n^{(p)}(\alpha_n)}{q^{(p)}(\alpha_n)} - 1, \frac{\hat{q}_n^{(1)}(\beta_n)}{q^{(1)}(\beta_n)} - 1\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \gamma^2\boldsymbol{\Lambda}\right), \tag{3.1}$$

where $\boldsymbol{\Lambda}$ is a symmetric matrix having entries

$$\begin{cases} \Lambda_{1,1} &= \dfrac{B\left(2p-1, \gamma^{-1}-2p+2\right)}{B\left(p, \gamma^{-1}-p+1\right)}, \\ \Lambda_{1,2} &= (p-1)\gamma(g_p(\gamma)/\theta)IB\left((g_p(\gamma)/\theta)^{-\gamma} \wedge 1, \gamma^{-1}-p+1, p-1\right) + \left((g_p(\gamma)/\theta)^{\gamma} \vee 1 - 1\right)^{p-1}, \\ \Lambda_{2,2} &= 1/\theta. \end{cases}$$

This result extends both Daouia et al. (2019b, Theorem 3) which is dedicated to the case $p = 2$, and Daouia et al. (2019a, Theorem 1) which focuses on the marginal asymptotic distribution of $\hat{q}_n^{(p)}(\alpha_n)$ only. Using Proposition 1, the asymptotic normality of $\hat{\gamma}_{\alpha_n}^{(p)}$ can then be established. A simplified version of this result, in the important case $\mathbb{E}(|Y|) < \infty$, is provided by Corollary 1 below.

Theorem 1. *Assume $\bar{F}^{(1)}$ satisfies $\boldsymbol{\mathcal{C}_2(\gamma,\rho,A)}$ and $\bar{F}_-^{(1)}$ is either light-tailed or satisfies $\boldsymbol{\mathcal{C}_1(\gamma_\ell)}$. Let (α_n) be an intermediate sequence such that, as $n \to \infty$,*

i) $\sqrt{n(1-\alpha_n)}A\left((1-\alpha_n)^{-1}\right) \to \lambda_1 \in \mathbb{R}$,

ii) $\sqrt{n(1-\alpha_n)}\left(\frac{\mathbb{E}\left[Y\mathbb{1}_{\{0<Y<q^{(1)}(\alpha_n)\}}\right]}{q^{(1)}(\alpha_n)} \vee (1-\alpha_n)\right) \to \lambda_2 \in \mathbb{R}$,

iii) $\sqrt{n(1-\alpha_n)}\left(\frac{\mathbb{E}\left[-Y\mathbb{1}_{\{0<-Y<q^{(1)}(\alpha_n)\}}\right]}{q^{(1)}(\alpha_n)} \vee F^{(1)}\left(-q^{(1)}(\alpha_n)\right)\right) \to \lambda_3 \in \mathbb{R}$.

Then, for all $1 < p < (\gamma \vee \gamma_\ell)^{-1}/2 + 1$,

$$\sqrt{n(1-\alpha_n)}\left(\hat{\gamma}^{(p)}_{\alpha_n} - \gamma\right) \xrightarrow{d} \mathcal{N}\left(b_p(\gamma,\rho), v_p(\gamma)\right), \tag{3.2}$$

where

$$\begin{cases} b_p(\gamma,\rho) = & \dfrac{-\gamma}{1+\frac{1}{\gamma}\left(\Psi\left(\gamma^{-1}-p+1\right)-\Psi\left(\gamma^{-1}+1\right)\right)}\left[\lambda_1 b_p^{(1)}(\gamma,\rho)+\lambda_2 b_p^{(2)}(\gamma)+\lambda_3 b_p^{(3)}(\gamma,\gamma_\ell)\right], \\ v_p(\gamma) = & \dfrac{\gamma B\left(p,\gamma^{-1}-p+1\right)}{\left(1+\frac{1}{\gamma}\left(\Psi\left(\gamma^{-1}-p+1\right)-\Psi\left(\gamma^{-1}+1\right)\right)\right)^2}\left[\gamma\dfrac{B\left(2p-1,\gamma^{-1}-2p+2\right)}{B\left(p,\gamma^{-1}-p+1\right)^2}-1\right], \end{cases}$$

with

$$\begin{cases} b_p^{(1)}(\gamma,\rho) &= g_p(\gamma)K(p,\gamma,\rho), \\ b_p^{(2)}(\gamma) &= (p-1)g_p(\gamma)^\gamma \mathbb{1}_{\{\gamma\le 1\}} + g_p(\gamma)\left[1+(p-1)B\left(p-1,1-\gamma^{-1}\right)\right]\mathbb{1}_{\{\gamma>1\}}, \\ b_p^{(3)}(\gamma,\gamma_\ell) &= -(p-1)\left[g_p(\gamma)^\gamma \mathbb{1}_{\{\gamma_\ell\le 1\}} + g_p(\gamma)^{\gamma/\gamma_\ell} B\left(\gamma_\ell^{-1}-p+1,1-\gamma_\ell^{-1}\right)\mathbb{1}_{\{\gamma_\ell>1\}}\right], \end{cases}$$

and

$$K(p,\gamma,\rho) = \begin{cases} \dfrac{g_p(\gamma)^{-\rho}}{\gamma^2\rho}\left[(1-\rho)B\left(p,(1-\rho)\gamma^{-1}-p+1\right)-B\left(p,\gamma^{-1}-p+1\right)\right] & \text{if } \rho<0, \\ \dfrac{p-1}{\gamma^2}\displaystyle\int_1^\infty (u-1)^{p-2}u^{-1/\gamma}\log(u)du & \text{if } \rho=0. \end{cases} \tag{3.3}$$

The asymptotic variance $v_p(\gamma)$ is compared to the asymptotic variances $v_H(\gamma)$ and $v_{EH}(\gamma)$ of Hill and expectHill estimators in Figure 1 for $p \in \{1.2, 1.4, 1.6, 1.8, 2\}$ and $\gamma \in (0,1)$. It appears that $v_p(\gamma)$ is smaller than $v_{EH}(\gamma)$ for $p = 2$ and is similar to $v_H(\gamma)$ when p is close to 1. Moreover, the asymptotic variance of $\hat{\gamma}^{(p)}_{\alpha_n}$ is smaller than $v_H(\gamma)$ when γ is small, for all considered values of p. On the contrary, $v_p(\gamma)$ increases faster than $v_H(\gamma)$ as γ increases. Finally, $p = 1.4$ seems to be a good compromise when γ is less than $\simeq 2/3$. This includes the important case $\gamma < 1/2$ of a finite variance, which is widespread in insurance and finance. A detailed discussion on the choice of p in practice is proposed in Section 4.

The asymptotic bias component involves three parts. The first one, $b_p^{(1)}(\gamma,\rho)$, is proportional to the auxiliary function A. Similar terms can be found in most extreme-value estimators, such as the Hill (see (2.11)) or expectHill estimators (see (2.12)). The second and third terms are specific to the use of L_p-quantiles. Since these terms do not depend neither on A nor on ρ, they can be easily estimated and removed, as we shall explain in the next paragraph. Before that, we propose a substantial simplification of $b_p^{(2)}(\gamma)$ and $b_p^{(3)}(\gamma,\gamma_\ell)$ under the assumption that $\gamma_\ell < 1$ and $\gamma < 1$. In this situation, the random variable Y is integrable, *i.e.* $\mathbb{E}|Y| < \infty$ and Theorem 1 can be simplified as follows.

Corollary 1. *Assume* $\bar{F}^{(1)}$ *satisfies* $\boldsymbol{\mathcal{C}_2(\gamma,\rho,A)}$ *with* $\gamma < 1$ *and* $\bar{F}^{(1)}_-$ *is either light-tailed or satisfies* $\boldsymbol{\mathcal{C}_1(\gamma_\ell)}$ *with* $\gamma_\ell < 1$. *Let* (α_n) *be an intermediate sequence such that, as* $n \to \infty$,

i) $\sqrt{n(1-\alpha_n)}A\left((1-\alpha_n)^{-1}\right) \to \lambda \in \mathbb{R}$,

ii) $\sqrt{n(1-\alpha_n)}/q^{(1)}(\alpha_n) \to \mu \in \mathbb{R}$.

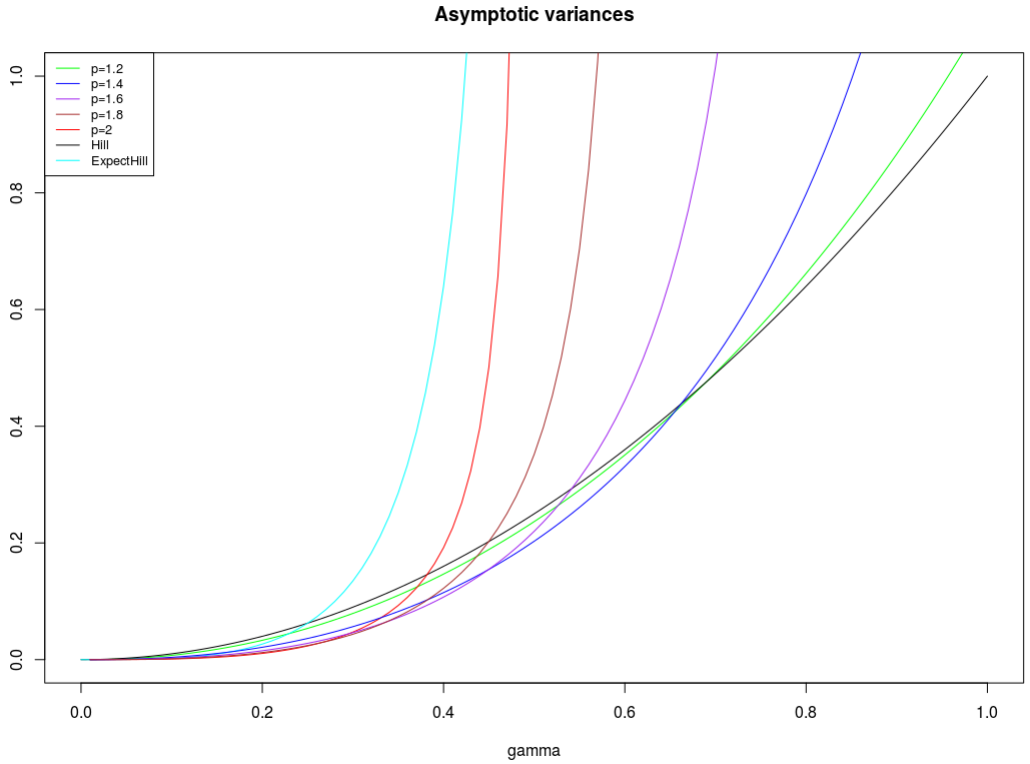


Figure 1: Asymptotic variance $v_p(\gamma)$ for $p = 1.2$ (green), 1.4 (blue), 1.6 (violet), 1.8 (brown) and 2 (red), as a function of $\gamma \in (0,1)$. The black and cyan curves respectively represent $v_H(\gamma)$, the asymptotic variance of Hill estimator, and $v_{EH}(\gamma)$, the asymptotic variance of expectHill estimator.

Then, for all $1 < p < (\gamma \vee \gamma_\ell)^{-1}/2 + 1$,

$$\sqrt{n(1-\alpha_n)}\left(\hat{\gamma}_{\alpha_n}^{(p)} - \gamma\right) \xrightarrow{d} \mathcal{N}\left(b_p(\gamma,\rho), v_p(\gamma)\right), \tag{3.4}$$

with

$$\begin{cases} b_p(\gamma,\rho) &= -\gamma \dfrac{g_p(\gamma)K(p,\gamma,\rho)\lambda + (p-1)\mathbb{E}(Y)g_p(\gamma)^\gamma \mu}{1+\frac{1}{\gamma}\left(\Psi\left(\gamma^{-1}-p+1\right)-\Psi\left(\gamma^{-1}+1\right)\right)}, \\ v_p(\gamma) &= \dfrac{\gamma B\left(p,\gamma^{-1}-p+1\right)}{\left(1+\frac{1}{\gamma}\left(\Psi\left(\gamma^{-1}-p+1\right)-\Psi\left(\gamma^{-1}+1\right)\right)\right)^2}\left[\gamma\dfrac{B\left(2p-1,\gamma^{-1}-2p+2\right)}{B\left(p,\gamma^{-1}-p+1\right)^2}-1\right]. \end{cases} \tag{3.5}$$

Let us highlight that the assumption $\gamma < 1$ is popular in risk management, see for instance Cai et al. (2015) and Daouia et al. (2018a). In these studies, the estimates of γ were found below 1/2. We also refer the reader to the R package `CASdatasets` which contains several datasets where estimates of γ typically vary between 1/4 and 1/2. For this kind of data, the expectHill estimator will tend to be unstable and/or will suffer from strong finite-sample bias, which is why our general L^p-quantile approach, with its reduced integrability assumptions, will be interesting in such cases. Finally, in the particular case $p = 2$, Corollary 1 provides an unconditional refinement of Girard et al. (2019, Theorem 4).

3.2 Bias correction

In view of (3.5) and conditions $i)$, $ii)$ in Corollary 1, the bias term due to the auxiliary function $A\left((1-\alpha_n)^{-1}\right)$ is negligible compared to the bias term due to $1/q^{(1)}(\alpha_n)$ when $\gamma < -\rho$. The converse case $|\rho| \leq \gamma$ is, in the most frequently considered situation when $\gamma < 1$, a case when the second-order

parameter ρ is "close" to 0. In such a case, the Pareto tail is, in finite samples, an unreliable representation of the tail of the underlying distribution; this is known to be very difficult to handle in the heavy tails setup we focus on here and finite-sample results are then bound to be disappointing. This is our rationale for concentrating on the case $\gamma < -\rho$, and for specifically correcting for the dominant bias component proportional to $1/q^{(1)}(\alpha_n)$, which is purely due to the L^p-quantile methodology.

To this end, we note that $g_p(\gamma)^\gamma \mu$ is asymptotically proportional to $\sqrt{n(1-\alpha_n)}g_p(\gamma)^\gamma/q^{(1)}(\alpha_n)$, itself equivalent to $\sqrt{n(1-\alpha_n)}/q^{(p)}(\alpha_n)$ in view of Daouia et al. (2019a, Proposition 3). Therefore, the associated bias term

$$-\gamma\frac{(p-1)\mathbb{E}(Y)g_p(\gamma)^\gamma\mu}{1+\frac{1}{\gamma}\left(\Psi\left(\gamma^{-1}-p+1\right)-\Psi\left(\gamma^{-1}+1\right)\right)} \approx -\gamma\frac{(p-1)\mathbb{E}(Y)}{1+\frac{1}{\gamma}\left(\Psi\left(\gamma^{-1}-p+1\right)-\Psi\left(\gamma^{-1}+1\right)\right)} \times \frac{\sqrt{n(1-\alpha_n)}}{q^{(p)}(\alpha_n)}$$

can be estimated by replacing $\mathbb{E}(Y)$ by the empirical mean, $q^{(p)}(\alpha_n)$ by its estimator $\hat{q}_n^{(p)}(\alpha_n)$, and by plugging any consistent estimator of γ in the denominator. This results in the following bias-corrected estimator:

$$\tilde{\gamma}_{\alpha_n}^{(p)} = \hat{\gamma}_{\alpha_n}^{(p)}\left(1+\frac{(p-1)\left(\frac{1}{n}\sum_{i=1}^n Y_i\right)\hat{q}_n^{(p)}(\alpha_n)^{-1}}{1+\frac{1}{\hat{\gamma}_{\alpha_n}^{(H)}}\left(\Psi\left(1/\hat{\gamma}_{\alpha_n}^{(H)}-p+1\right)-\Psi\left(1/\hat{\gamma}_{\alpha_n}^{(H)}+1\right)\right)}\right). \tag{3.6}$$

In view of (2.12), a similar technique can be used to reduce the bias of the expectHill estimator:

$$\tilde{\gamma}_{\alpha_n}^{(EH)} = \hat{\gamma}_{\alpha_n}^{(EH)}\left(1+\frac{\hat{\gamma}_{\alpha_n}^{(H)}}{\hat{\gamma}_{\alpha_n}^{(H)}+1}\frac{\frac{1}{n}\sum_{i=1}^n Y_i}{\hat{q}_n^{(2)}(\alpha_n)}\right). \tag{3.7}$$

The asymptotic normality of both of these bias-corrected estimators can be established as straightforward consequences of Corollary 1 and Daouia et al. (2018b, Theorem 2).

Corollary 2. *Under the assumptions of Corollary 1:*

$$\sqrt{n(1-\alpha_n)}\left(\tilde{\gamma}_{\alpha_n}^{(p)}-\gamma\right) \xrightarrow{d} \mathcal{N}\left(b_p^*(\gamma,\rho), v_p(\gamma)\right), \tag{3.8}$$

where

$$b_p^*(\gamma,\rho) = -\frac{\gamma g_p(\gamma)K(p,\gamma,\rho)}{1+\frac{1}{\gamma}\left(\Psi\left(\gamma^{-1}-p+1\right)-\Psi\left(\gamma^{-1}+1\right)\right)}\lambda.$$

If moreover $\gamma < 1/2$ *and* $\mathbb{E}[Y^2] < \infty$ *then*

$$\sqrt{n(1-\alpha_n)}\left(\tilde{\gamma}_{\alpha_n}^{(EH)}-\gamma\right) \xrightarrow{d} \mathcal{N}\left(b_{EH}^*(\gamma,\rho), v_{EH}(\gamma)\right),$$

where

$$b_{EH}^*(\gamma,\rho) = \frac{(1-\gamma)\left(\gamma^{-1}-1\right)^{-\rho}}{(1-\rho)(1-\gamma-\rho)}\lambda. \tag{3.9}$$

The remaining bias terms $b_p^*(\gamma,\rho)$ and $b_{EH}^*(\gamma,\rho)$ may also be removed using estimators of ρ and $A(\cdot)$. We refer the reader to Matthys and Beirlant (2003) and more recently Paulauskas and Vaičiulis (2017) for examples, and Gomes and Guillou (2015, Section 5.3) for a review. We did try implementing such a correction on top of our proposed bias reduction method; the numerical results giving approximately the same trends in terms of relative performance between the different estimators we shall be testing, we choose here to skip this correction and the discussion of its implementation. The next paragraph focuses on the asymptotic mean squared error of our bias-reduced estimator and on a related criterion for the choice of the parameter p.

3.3 Asymptotic mean squared error

Following de Haan and Peng (1998), the estimators $\hat{\gamma}_{\alpha_n}^{(H)}$, $\tilde{\gamma}_{\alpha_n}^{(EH)}$ and $\tilde{\gamma}_{\alpha_n}^{(p)}$ are compared in terms of their asymptotic mean squared errors, written as functions of α_n. In the case of $\tilde{\gamma}_{\alpha_n}^{(p)}$, one has

$$AMSE_p(\gamma,\rho) := \frac{b_p^*(\gamma,\rho)^2}{\lambda^2}[A\left((1-\alpha_n)^{-1}\right)]^2 + \frac{v_p(\gamma)}{n(1-\alpha_n)}.$$

Let us consider the typical case of an auxiliary function $A(t) = ct^\rho$, $c \neq 0$, $\rho < 0$. Minimizing the AMSE with respect to $1-\alpha_n$ yields an optimal α_n^* satisfying

$$1-\alpha_n^* = \left(\frac{v_p(\gamma)\lambda^2}{-2\rho c^2 b_p^*(\gamma,\rho)^2}\right)^{\frac{1}{1-2\rho}} n^{-\frac{1}{1-2\rho}}, \tag{3.10}$$

and thus, the optimal AMSE is given by

$$AMSE^*(\gamma,\rho) = \left(\frac{v_p(\gamma)^{-\rho} b_p^*(\gamma,\rho)}{\lambda}\right)^{\frac{2}{1-2\rho}} c^{\frac{2}{1-2\rho}} (-2\rho)^{\frac{2\rho}{1-2\rho}} [1-2\rho]\, n^{\frac{2\rho}{1-2\rho}}.$$

Let us stress that this is actually a general calculation of the optimal AMSE; the only term that is specific to the estimator $\tilde{\gamma}_{\alpha_n}^{(p)}$ is $v_p(\gamma)^{-\rho} b_p^*(\gamma,\rho)$. Equivalent formulas are then obtained for the Hill (de Haan and Ferreira, 2006, Section 3.2) or expectHill estimators by replacing $v_p(\gamma)^{-\rho} b_p^*(\gamma,\rho)$ by $v_H(\gamma)^{-\rho} b_H(\rho)$ and $v_{EH}(\gamma)^{-\rho} b_{EH}^*(\gamma,\rho)$, respectively. It can be observed numerically that the optimal AMSE associated to the expectHill estimator is the smallest for small values of γ and $|\rho|$ (blue areas). The difference in terms of AMSE between $\tilde{\gamma}_{\alpha_n}^{(p)}$ and the Hill estimator is strongly related to the value of p. The set of parameters (γ,ρ) where $\tilde{\gamma}_{\alpha_n}^{(p)}$ has a lower optimal AMSE than the Hill and expectHill estimators is depicted in red on Figure 2. It appears that $\tilde{\gamma}_{\alpha_n}^{(p)}$ has a smaller optimal AMSE for small values of γ when p is close to 2. At the opposite, $\tilde{\gamma}_{\alpha_n}^{(p)}$ is better for γ between 1/2 and 1 (with small values of $|\rho|$) when p is smaller. Note that if p is too close to 1, our L^p-quantile based estimator is never better than the others. Our numerical experiments thus indicate the following methodology: take $p = 2$ for small values of γ (less than 0.25), a value of $p = 1.7$ for intermediate values of γ (between 0.25 and 0.5), and $p = 1.4$ for larger values of the tail index (greater than 0.5). Obviously, such a methodology requires a prior estimation of γ, which may for instance be done using the Hill estimator.

4 Simulation study

4.1 Experimental design

The L^p-quantile based estimator is compared with the Hill and expectHill estimators on three heavy-tailed distributions:

- The Generalized Pareto Distribution GPD(γ) with shape parameter $\gamma > 0$, whose survival function is defined for all $y \geq 0$ by

  $$\bar{F}^{(1)}(y) = (1+\gamma y)^{-1/\gamma}.$$

  The second-order condition $\boldsymbol{\mathcal{C}_2(\gamma, -\gamma, A)}$ holds with $A(y)$ proportional to $y^{-\gamma}$.

- A Burr-type distribution Burr(γ,ρ) with $\gamma > 0$ and $\rho < 0$, whose survival function is defined for all $y \geq 0$ by

  $$\bar{F}^{(1)}(y) = \left(1 + y^{-\rho/\gamma}\right)^{1/\rho}.$$

  The second-order condition $\boldsymbol{\mathcal{C}_2(\gamma, \rho, A)}$ holds with $A(y)$ proportional to y^ρ.

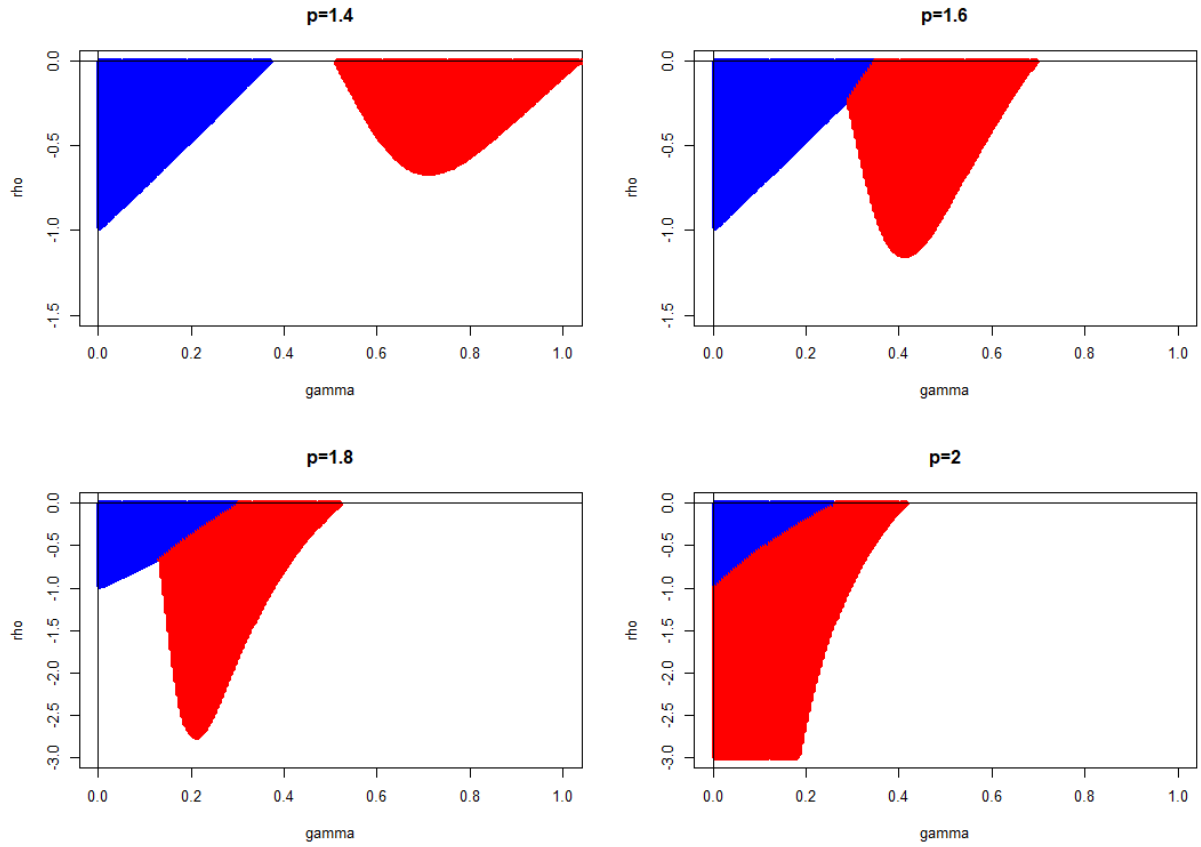


Figure 2: Areas (in red) where the L^p-quantile based tail index estimator has a smaller optimal AMSE than the Hill and expectHill estimators, for $p \in \{1.4, 1.6, 1.8, 2\}$. The Hill and expectHill estimators have the smallest AMSE in the white and blue areas respectively.

- The Student distribution Student(γ) with $1/\gamma > 0$ degrees of freedom, whose survival function is defined for all $y \in \mathbb{R}$ by

$$\bar{F}^{(1)}(y) = \sqrt{\frac{\gamma}{\pi}} \frac{\Gamma\left(\frac{\gamma^{-1}+1}{2}\right)}{\Gamma\left(\frac{\gamma^{-1}}{2}\right)} \int_y^\infty \left(1+\gamma t^2\right)^{-\frac{\gamma^{-1}+1}{2}} dt.$$

  The second-order condition $\boldsymbol{\mathcal{C}_2(\gamma, -2\gamma, A)}$ holds with $A(y)$ proportional to $y^{-2\gamma}$.

For each distribution, $N = 100$ samples of size $n = 2{,}500$ are simulated. The tail index γ is estimated successively by the Hill estimator $\hat{\gamma}_{\alpha_n}^{(H)}$, the expectHill estimator $\tilde{\gamma}_{\alpha_n}^{(EH)}$ and the L^p-quantile based $\tilde{\gamma}_{\alpha_n}^{(p)}$ estimator (with $p = 1.4$, 1.7 or 2, see the discussion in Subsection 3.3). The finite sample properties of the estimators are assessed by computing the empirical mean of the estimates as well as the empirical mean squared error:

$$\bar{\hat{\gamma}}_n = \frac{1}{N}\sum_{j=1}^{N} \hat{\gamma}_j \quad \text{and} \quad \widehat{MSE}_n\left(\hat{\gamma}\right) = \frac{1}{N}\sum_{j=1}^{N} \left(\hat{\gamma}_j - \gamma\right)^2,$$

where $\hat{\gamma} \in \{\hat{\gamma}_{\alpha_n}^{(H)}, \tilde{\gamma}_{\alpha_n}^{(EH)}, \tilde{\gamma}_{\alpha_n}^{(p)}\}$ and $\hat{\gamma}_j$ denotes the estimate obtained by $\hat{\gamma}$ on the jth sample.
All the considered estimators require to select an intermediate sequence (α_n). This is a well-known but difficult problem for which many approaches have been developed, see *e.g.* Gomes and Guillou (2015, Section 5.4). In the next paragraph, we propose a technique following the ideas of Drees and Kaufmann (1998) based on the minimization of the AMSE.

4.2 Choice of α_n

For all three considered estimators, the intermediate sequence is chosen by minimizing the AMSE, see (3.10) in the case of the L^p-quantile based estimator, with similar formulae for the Hill and expectHill estimators. The optimal value α_n^* depends on the three unknown parameters c, ρ and γ though. While the tail index γ can be replaced by the estimator under consideration, the estimation of the second-order parameters ρ and c is much more delicate in practice. We thus chose to make the canonical choices $\rho = -\gamma$ and $c = 1$: this technique is widely used in the bias reduction of extreme-value estimators, see for instance Drees and Kaufmann (1998) or Beirlant et al. (2005). Its rationale is that, while estimating ρ and c using adapted estimators is a sensible solution, such estimators will tend to have a low rate of convergence and therefore a large variance. Meanwhile, making a canonical choice might induces a misspecification, thus failing to eliminate some of the bias, but such a choice has variance 0. It is then typically observed that the finite-sample performance of bias-reduction methods using such canonical choices is usually good because the potential bias incurred is more than compensated by the fact that the constructed bias-reduced estimators have a low variance.

4.3 Results

An overview of the results is presented in Figure 3 (empirical means) and Figure 4 (empirical mean squared errors). The three considered estimators are compared on the cases of the GPD(γ), Burr(γ, ρ) and Student(γ) distributions, with $\rho \in \{-1, -0.5\}$ and $\gamma \in \{0.15, 0.3, 0.45, 0.6, 0.75\}$. It appears that the results are coherent with the AMSE plots of Figure 4: the L^2- and expectHill estimators perform very well for small values of γ, the $L^{1.7}-$estimator shows very good performance for γ between 0.3 and 0.45 (note that it is clearly better than the others for small values of $|\rho|$, and performs comparably otherwise), and the $L^{1.4}-$estimator outperforms the Hill and expectHill estimators for $\gamma = 0.6$ and $\gamma = 0.75$ when $|\rho|$ is small enough. Obviously, when $|\rho|$ is larger, the underlying distribution is close to a Pareto, and the finite-sample performance of our estimator is not better than that of the Hill estimator. In summary, a suitable choice of p in the L^p-quantile based estimator $\tilde{\gamma}_{\alpha_n}^{(p)}$ may improve the estimation of the tail index in the challenging situations where $|\rho|$ and/or γ are small.

5 Real data example

In this section, we consider a dataset of daily river flows of the Salt River near Roosevelt, Arizona. This dataset is available at `https://waterdata.usgs.gov/nwis/inventory` (station 09497500) and has already been studied in Bowers et al. (2012) and Fonseca et al. (2019) to cite a few. We also refer to Anderson and Meerschaert (1998) for a similar study of the Salt River with monthly flows. In our context, $T = 31{,}755$ discharges $X_1, \ldots, X_T$ (in cubic feet per second) have been recorded during 87 years over the period 1924 to 2010, see the left panel of Figure 5 for an illustration. Similarly to Fonseca et al. (2019), the daily flow rates $R_t = (X_t - X_{t-1})/X_{t-1}$ are computed and we focus on the upper 10% of the order statistics of R, leading to a sample $Y_1, \ldots, Y_n$ of large flow rates of size $n = 3{,}176$. Following Fonseca et al. (2019), the resulting dataset is a set of peaks over a high threshold which we model using a heavy-tailed distribution, and $Y_1, \ldots, Y_n$ are observed not to be strongly dependent, thus justifying the use of our estimators in this context.
Our goal is the estimation of the 100-year flow rate, that is, the flow rate reached by the Salt River once per century on average, or equivalently the quantile $q_R(\beta_n)$ such that $\mathbb{P}(R > q_R(\beta_n)) = 1 - \beta_n = 1/(365 \times 100)$. Clearly

$$\mathbb{P}(Y > q_R(\beta_n)) = \mathbb{P}(R > q_R(\beta_n) | R > q_R(0.9)) = \frac{\mathbb{P}(R > q_R(\beta_n))}{\mathbb{P}(R > q_R(0.9))} = 10\,\mathbb{P}(R > q_R(\beta_n)),$$

hence $q_R(\beta_n) = q_Y(\beta_n')$, where $1 - \beta_n' = 1/(10 \times 365) < 1/n$. For this reason, such an extreme quantile cannot be estimated by a standard empirical quantile estimator (*i.e.* a high order statistic), and its estimation therefore requires an extrapolation technique. Weissman's estimator, see Weissman (1978),

allows us to estimate this quantity as follows:

$$\hat{q}_{R,n}^{(W)}(\beta_n) = \left(\frac{1-\alpha_n}{1-\beta_n'}\right)^{\hat{\gamma}_{\alpha_n}} \hat{q}_{Y,n}(\alpha_n),$$

where (α_n) is an intermediate sequence, $\hat{q}_{Y,n}(\alpha_n) = Y_{\lfloor n\alpha_n \rfloor,n}$ with $Y_{1,n} \leq \ldots \leq Y_{n,n}$ being the order statistics from the sample and $\hat{\gamma}_{\alpha_n}$ may be either $\hat{\gamma}_{\alpha_n}^{(H)}$, $\tilde{\gamma}_{\alpha_n}^{(EH)}$ or $\tilde{\gamma}_{\alpha_n}^{(p)}$. Moreover, from de Haan and Ferreira (2006, Theorem 4.3.8), the quantities

$$\frac{\sqrt{n(1-\alpha_n)}}{\log\left((1-\alpha_n)/(1-\beta_n')\right)}\left(\hat{q}_{R,n}^{(W)}(\beta_n)/q_R(\beta_n) - 1\right) \text{ and } \sqrt{n(1-\alpha_n)}\,(\hat{\gamma}_{\alpha_n} - \gamma)$$

have the same asymptotic distribution. Hence, the asymptotic results established in Section 3 can be used to provide approximate confidence intervals for the 100-year flow rate.

Figure 6 gives an overview of the results obtained with $\hat{\gamma}_{\alpha_n}^{(H)}$, $\tilde{\gamma}_{\alpha_n}^{(EH)}$ and $\tilde{\gamma}_{\alpha_n}^{(p)}$ ($p = 1.4$ in this case) for the tail index and the associated 100-year flow rate estimates. There is no apparent stability region for the expectHill estimate $\tilde{\gamma}_{\alpha_n}^{(EH)}$, which underlines the strong influence of the bias due to the second-order function A on this estimator. The selection procedure described in Subsection 4.2 yields $k_n^* = \lfloor n(1-\alpha_n^*) \rfloor = 72$ for the Hill estimator and 65 for the L^p-quantile approach. It appears that both the Hill and L^p-quantile estimators provide an estimation of the tail index around 0.65 around these values of $n(1-\alpha_n)$. The right panel of Figure 5 displays a quantile-quantile plot of the log-spacings $\log\left(Y_{n-i+1,n}/Y_{n-k_n^*+1,n}\right)$ against the quantiles $\log(k_n^*/i)$ of the standard exponential distribution, $1 \leq i \leq k_n^* = 65$. The set of points is close to the line of slope $\tilde{\gamma}_{\alpha_n^*}^{(p)}$: this graphical diagnostic confirms that the assumption of heavy tails is reasonable for the considered dataset (see for instance Beirlant et al., 1999).

The estimated 100-year flow rates are broadly similar at the indicated values k_n^*. More precisely, the associated 95% confidence intervals on the 100-year flow rate are respectively $[40, 210]$ ($\hat{q}_{R,n}^{(W)}(\beta_n) = 125$, for the Hill estimator) and $[37, 176]$ ($\hat{q}_{R,n}^{(W)}(\beta_n) = 107$, for the L^p-quantile estimator). The asymptotic normality result for the expectHill estimator, meanwhile, cannot be used to provide sensible confidence intervals since the estimated tail index γ is above $1/2$. It appears that, comparing to the Hill approach, the L^p-quantile estimator provides a 19% reduction on the size of the 100-year flow rate confidence interval. The latter and its associated point estimate also appear to be more in agreement with the maximum flow rate observed over 87 years $Y_{n,n} = R_{T,T} = 85$.

6 Appendix

6.1 Preliminary results

The first result states that the function $\bar{F}^{(p)}$ defined in Equation (2.1) is continuous and strictly decreasing on the range of values of Y, and is therefore in particular a survival function. A general result for M-quantiles has been proved in Jones (1994) under conditions whose validity for L^p-quantiles is not completely clear.

Lemma 1. *Assume $\bar{F}^{(1)}$ satisfies $\boldsymbol{\mathcal{C}_1(\gamma)}$, and let $y^* = \inf\left\{y \in \mathbb{R} \colon \bar{F}^{(1)}(y) < 1\right\}$. Then for all $p > 1$ such that $\mathbb{E}\left[|Y|^{p-1}\right] < +\infty$, $\bar{F}^{(p)} : (y^*, +\infty) \to (0,1)$ is continuous and strictly decreasing.*

Proof. For all $y > y^*$, the integration by parts formula easily leads to

$$\frac{1}{\bar{F}^{(p)}(y)} = 1 + \frac{\mathbb{E}\left[(y-Y)^{p-1}\mathbb{1}_{\{Y \leq y\}}\right]}{\mathbb{E}\left[(Y-y)^{p-1}\mathbb{1}_{\{Y > y\}}\right]} = 1 + \frac{\int_0^{+\infty} t^{p-2} F^{(1)}(y-t)dt}{\int_0^{+\infty} t^{p-2}\bar{F}^{(1)}(y+t)dt}.$$

Then for all $h > 0$, $\bar{F}^{(p)}(y+h)^{-1} - \bar{F}^{(p)}(y)^{-1} =$

$$\frac{\int_{(0,+\infty)^2} s^{p-2}t^{p-2}\left\{F^{(1)}(y+h-t)\bar{F}^{(1)}(y+s) - F^{(1)}(y-t)\bar{F}^{(1)}(y+h+s)\right\} ds\, dt}{\int_0^{+\infty} t^{p-2}\bar{F}^{(1)}(y+h+t)dt \int_0^{+\infty} t^{p-2}\bar{F}^{(1)}(y+t)dt}.$$

Notice that

$$\begin{aligned} & F^{(1)}(y+h-t)\bar{F}^{(1)}(y+s) - F^{(1)}(y-t)\bar{F}^{(1)}(y+h+s) \\ = \; & \bar{F}^{(1)}(y+s)\left\{F^{(1)}(y+h-t) - F^{(1)}(y-t)\right\} + F^{(1)}(y-t)\left\{\bar{F}^{(1)}(y+s) - \bar{F}^{(1)}(y+h+s)\right\} \geq 0. \end{aligned}$$

Let us suppose now that there exist $y > y^*$ and $h > 0$ such that $\bar{F}^{(p)}(y+h) = \bar{F}^{(p)}(y)$. In that case, we necessarily have

$$\bar{F}^{(1)}(y+s)\left\{F^{(1)}(y+h-t) - F^{(1)}(y-t)\right\} = 0 \text{ for almost all } s,t > 0.$$

Since $\bar{F}^{(1)}$ fulfills $\boldsymbol{\mathcal{C}_1(\gamma)}$, $\bar{F}^{(1)}(y+s) > 0$ for s large enough, hence

$$F^{(1)}(y+h-t) = F^{(1)}(y-t) \text{ for almost all } t > 0.$$

By taking a $t \in (y-y^*, y+h-y^*)$ satisfying this identity, we get $0 < F^{(1)}(y+h-t) = F^{(1)}(y-t) = 0$ which is an obvious contradiction. Hence $\bar{F}^{(p)}(y+h)^{-1} - \bar{F}^{(p)}(y)^{-1} > 0$, proving that $\bar{F}^{(p)}$ is strictly decreasing on $(y^*, +\infty)$. Continuity of this mapping can be shown by proving the continuity of both

$$y \mapsto \mathbb{E}\left[(y-Y)^{p-1}\mathbb{1}_{\{Y \leq y\}}\right] \text{ and } y \mapsto \mathbb{E}\left[(Y-y)^{p-1}\mathbb{1}_{\{Y > y\}}\right]$$

which is a straightforward consequence of the dominated convergence theorem. □

Our second lemma establishes that the function g_p defined in (2.6) is indeed monotonic.

Lemma 2. *For all $p > 1$, $g_p(\cdot)$ is strictly decreasing on $(0, 1/(p-1))$ and has derivative*

$$g_p'(\gamma) = \frac{1 - \gamma^{-1}\left(\Psi\left(\gamma^{-1}+1\right) - \Psi\left(\gamma^{-1}-p+1\right)\right)}{B\left(p, \gamma^{-1}-p+1\right)}, \text{ for all } \gamma \in (0, 1/(p-1)).$$

Proof. The expression of $g_p'(\cdot)$ follows from direct calculations. Since $p > 1$, we have $\Psi\left(\gamma^{-1}+1\right) - \Psi\left(\gamma^{-1}-p+1\right) > \Psi\left(\gamma^{-1}+1\right) - \Psi\left(\gamma^{-1}\right) = \gamma$, see Abramowitz and Stegun (1972, Chapter 6). Thus $g_p'(\gamma) < 0$ and g_p is strictly decreasing as required. □

For all $y \in \mathbb{R}$, let us denote by

$$\varphi^{(k)}(y) = \mathbb{E}\left[|Y-y|^k \mathbb{1}_{\{Y>y\}}\right] \text{ and } m^{(k)}(y) = \mathbb{E}\left[|Y-y|^k\right]$$

the truncated and total moments of $|Y-y|$. The associated empirical estimators are

$$\widehat{\varphi}_n^{(k)}(y) = \frac{1}{n}\sum_{i=1}^n |Y_i - y|^k \mathbb{1}_{\{Y_i > y\}} \text{ and } \widehat{m}_n^{(k)}(y) = \frac{1}{n}\sum_{i=1}^n |Y_i - y|^k.$$

Note that, for all $p \geq 1$, $\bar{F}^{(p)}(y) = \varphi^{(p-1)}(y)/m^{(p-1)}(y)$ and in particular $\bar{F}^{(1)}(y) = \varphi^{(0)}(y)$. The next lemma states the properties of $\varphi^{(k)}(y)$, $m^{(k)}(y)$ and $\bar{F}^{(p)}(y)$ as $y \to \infty$.

Lemma 3. *Assume $\bar{F}^{(1)}$ satisfies $\boldsymbol{\mathcal{C}_1(\gamma)}$.*

i) For all $k \in [0, 1/\gamma)$,

$$\varphi^{(k)}(y) = \frac{B\left(k+1, \gamma^{-1}-k\right)}{\gamma} y^k \bar{F}^{(1)}(y)(1+o(1)) \text{ as } y \to \infty.$$

ii) If $\bar{F}_-^{(1)}$ is either light-tailed or satisfies $\boldsymbol{\mathcal{C}_1(\gamma_\ell)}$, then, for all $k \in [0, 1/(\gamma \vee \gamma_\ell))$,

$$m^{(k)}(y) = y^k(1+o(1)), \text{ as } y \to \infty.$$

iii) If $\bar{F}_-^{(1)}$ is either light-tailed or satisfies $\boldsymbol{\mathcal{C}_1(\gamma_\ell)}$, then, for all $p \in [1, 1 + 1/(\gamma \vee \gamma_\ell))$,

$$\bar{F}^{(p)}(y) = \frac{B\left(p, \gamma^{-1} - p + 1\right)}{\gamma} \bar{F}^{(1)}(y)(1 + o(1)) \text{ as } y \to \infty,$$

and the function $\bar{F}^{(p)}$ is regularly varying:

$$\forall t > 0, \lim_{y\to\infty} \frac{\bar{F}^{(p)}(ty)}{\bar{F}^{(p)}(y)} = t^{-1/\gamma} \text{ or equivalently } \lim_{y\to\infty} \frac{q^{(p)}\left(1 - 1/(ty)\right)}{q^{(p)}\left(1 - 1/y\right)} = t^{\gamma}.$$

Proof. i) The case $k = 0$ is straightforward, since $\varphi^{(0)}(y) = \bar{F}^{(1)}(y)$. If $k > 0$, remark that

$$\varphi^{(k)}(y) = y^k \mathbb{E}\left[\left|\frac{Y}{y} - 1\right|^k \mathbb{1}_{\{Y>y\}}\right],$$

and apply Daouia et al. (2019a, Lemma 1(i)) with $H(x) = (x-1)^k$ and $b = 1$. ii) follows from writing $m^{(k)}(y) = y^k \mathbb{E}\left[|Y/y - 1|^k\right]$, and noting that $\mathbb{E}\left[|Y/y - 1|^k\right] \to 1$ as $y \to \infty$ by the dominated convergence theorem. iii) follows from i) and ii). □

In the following two results, some (joint) asymptotic normality results are established for the estimators $\widehat{\varphi}_n^{(k)}(y_n)$ and $\hat{\bar{F}}_n^{(k)}(y_n)$ where $y_n \to \infty$ and $n \to \infty$.

Lemma 4. *Assume $\bar{F}^{(1)}$ satisfies $\boldsymbol{\mathcal{C}_2(\gamma, \rho, A)}$, and $\bar{F}_-^{(1)}$ is either light-tailed or satisfies $\boldsymbol{\mathcal{C}_1(\gamma_\ell)}$. Let (y_n) and (y'_n) be two sequences such that $y_n \to \infty$, $n\bar{F}^{(1)}(y_n) \to \infty$ and $y'_n/y_n \to \lambda > 0$ as $n \to \infty$. Then, for all $1 \le p < 1 + (\gamma \vee \gamma_\ell)^{-1}/2$,*

$$\sqrt{n\bar{F}^{(1)}(y_n)} \left(\frac{\widehat{\varphi}_n^{(p-1)}(y_n)}{\varphi^{(p-1)}(y_n)} - 1, \frac{\widehat{\varphi}_n^{(0)}(y'_n)}{\varphi^{(0)}(y'_n)} - 1\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Sigma}\right), \tag{6.1}$$

where $\boldsymbol{\Sigma}$ is a symmetric matrix having entries

$$\begin{cases} \Sigma_{1,1} &= \gamma \dfrac{B\left(2p-1, \gamma^{-1} - 2p + 2\right)}{B\left(p, \gamma^{-1} - p + 1\right)^2}, \\ \Sigma_{1,2} &= \gamma \dfrac{(p-1)\lambda^{1/\gamma} IB\left((\lambda \vee 1)^{-1}, \gamma^{-1} - p + 1, p - 1\right) + (\lambda \vee 1 - 1)^{p-1}}{B\left(p, \gamma^{-1} - p + 1\right)}, \\ \Sigma_{2,2} &= \lambda^{1/\gamma}. \end{cases} \tag{6.2}$$

Proof. Let $\boldsymbol{\beta} = (\beta_1, \beta_2) \in \mathbb{R}^2$ and focus on the asymptotic distribution of

$$Z_n = \sqrt{n\bar{F}^{(1)}(y_n)} \left\{\beta_1 \left(\frac{\widehat{\varphi}_n^{(p-1)}(y_n)}{\varphi^{(p-1)}(y_n)} - 1\right) + \beta_2 \left(\frac{\widehat{\varphi}_n^{(0)}(y'_n)}{\varphi^{(0)}(y'_n)} - 1\right)\right\}.$$

We clearly have $\mathbb{E}[Z_n] = 0$ and $\mathbb{V}\mathrm{ar}[Z_n] = \bar{F}^{(1)}(y_n) \boldsymbol{\beta}^\top \boldsymbol{S}^{(n)} \boldsymbol{\beta}$, where $\boldsymbol{S}^{(n)}$ is the symmetric matrix having entries

$$\begin{cases} S_{1,1}^{(n)} &= \dfrac{\mathbb{V}\mathrm{ar}\left[|Y - y_n|^{p-1} \mathbb{1}_{\{Y>y_n\}}\right]}{\varphi^{(p-1)}(y_n)^2}, \\ S_{1,2}^{(n)} &= \dfrac{\mathrm{cov}\left(|Y - y_n|^{p-1} \mathbb{1}_{\{Y>y_n\}}, \mathbb{1}_{\{Y>y'_n\}}\right)}{\varphi^{(p-1)}(y_n)\varphi^{(0)}(y'_n)}, \\ S_{2,2}^{(n)} &= \dfrac{\mathbb{V}\mathrm{ar}\left[\mathbb{1}_{\{Y>y'_n\}}\right]}{\varphi^{(0)}(y'_n)^2}. \end{cases}$$

Let us first focus on $S_{1,1}^{(n)}$. Lemma 3i) yields

$$S_{1,1}^{(n)} = \frac{\varphi^{(2p-2)}(y_n)}{\varphi^{(p-1)}(y_n)^2} - 1 = \gamma \frac{B\left(2p-1, \gamma^{-1}-2p+2\right)}{B\left(p, \gamma^{-1}-p+1\right)^2} \bar{F}^{(1)}(y_n)^{-1}(1+o(1)).$$

The calculation of the covariance term $S_{1,2}^{(n)}$ relies on Lemma 3i) and the decomposition

$$S_{1,2}^{(n)} = \frac{\mathbb{E}\left[|Y-y_n|^{p-1}\mathbb{1}_{\{Y>by_n\}}\right]}{\varphi^{(p-1)}(y_n)\varphi^{(0)}(y_n')} + \frac{\mathbb{E}\left[|Y-y_n|^{p-1}\left(\mathbb{1}_{\{Y>y_n\vee y_n'\}} - \mathbb{1}_{\{Y>by_n\}}\right)\right]}{\varphi^{(p-1)}(y_n)\varphi^{(0)}(y_n')} - 1$$

where $b = \lambda \vee 1$. The first term above is evaluated by using Lemma 3i) to write

$$\frac{\mathbb{E}\left[|Y-y_n|^{p-1}\mathbb{1}_{\{Y>by_n\}}\right]}{\varphi^{(p-1)}(y_n)\varphi^{(0)}(y_n')} = \frac{\mathbb{E}\left[\left\{\left|\frac{Y}{y_n}-1\right|^{p-1} - (b-1)^{p-1}\right\}\mathbb{1}_{\{Y>by_n\}}\right] + (b-1)^{p-1}\bar{F}^{(1)}(by_n)}{\frac{B(p,\gamma^{-1}-p+1)}{\gamma}\bar{F}^{(1)}(y_n)\bar{F}^{(1)}(y_n')(1+o(1))}.$$

Applying Daouia et al. (2019a, Lemma 1(i)) with $H(x) = (x-1)^{p-1}$ and carrying out straightforward calculations based on the change of variable $t = 1/x$ and the regular variation property of $\bar{F}^{(1)}$, we obtain

$$\begin{aligned}
&\frac{\mathbb{E}\left[|Y-y_n|^{p-1}\mathbb{1}_{\{Y>by_n\}}\right]}{\varphi^{(p-1)}(y_n)\varphi^{(0)}(y_n')} \\
= \;& \gamma \frac{\bar{F}^{(1)}(y_n)\int_b^\infty (p-1)(x-1)^{p-2}x^{-1/\gamma}dx(1+o(1)) + (b-1)^{p-1}\bar{F}^{(1)}(by_n)}{B\left(p,\gamma^{-1}-p+1\right)\bar{F}^{(1)}(y_n)\bar{F}^{(1)}(y_n')(1+o(1))} \\
= \;& \left[(p-1)\gamma\lambda^{1/\gamma}\frac{IB\left(b^{-1},\gamma^{-1}-p+1,p-1\right)}{B\left(p,\gamma^{-1}-p+1\right)} + \gamma\frac{(b-1)^{p-1}}{B\left(p,\gamma^{-1}-p+1\right)}\right]\bar{F}^{(1)}(y_n)^{-1}(1+o(1)).
\end{aligned}$$

Besides, by Lemma 3i) and the regular variation property of $\bar{F}^{(1)}$ again,

$$\left|\frac{\mathbb{E}\left[|Y-y_n|^{p-1}\left(\mathbb{1}_{\{Y>y_n\vee y_n'\}} - \mathbb{1}_{\{Y>by_n\}}\right)\right]}{\varphi^{(p-1)}(y_n)\varphi^{(0)}(y_n')}\right| = O\left(\bar{F}^{(1)}(y_n)^{-1}\left|\frac{\bar{F}^{(1)}(y_n\vee y_n')}{\bar{F}^{(1)}(y_n)} - \frac{\bar{F}^{(1)}(by_n)}{\bar{F}^{(1)}(y_n)}\right|\right).$$

Noting that $(y_n \vee y_n')/y_n \to b$ and using local uniformity of the regular variation property of $\bar{F}^{(1)}$ (Bingham et al., 1989, Theorem 1.2.1), we find

$$\left|\frac{\mathbb{E}\left[|Y-y_n|^{p-1}\left(\mathbb{1}_{\{Y>y_n\vee y_n'\}} - \mathbb{1}_{\{Y>by_n\}}\right)\right]}{\varphi^{(p-1)}(y_n)\varphi^{(0)}(y_n')}\right| = o\left(\bar{F}^{(1)}(y_n)^{-1}\right).$$

Conclude that

$$S_{1,2}^{(n)} = \left[(p-1)\gamma\lambda^{1/\gamma}\frac{IB\left(b^{-1},\gamma^{-1}-p+1,p-1\right)}{B\left(p,\gamma^{-1}-p+1\right)} + \gamma\frac{(b-1)^{p-1}}{B\left(p,\gamma^{-1}-p+1\right)}\right]\bar{F}^{(1)}(y_n)^{-1}(1+o(1)).$$

Finally, using regular variation again,

$$S_{2,2}^{(n)} = \frac{\bar{F}^{(1)}(y_n')}{\bar{F}^{(1)}(y_n')^2} - 1 = \lambda^{1/\gamma}\bar{F}^{(1)}(y_n)^{-1}(1+o(1)),$$

and therefore, $\mathbb{V}\mathrm{ar}[Z_n] \to \boldsymbol{\beta}^\top\boldsymbol{\Sigma}\boldsymbol{\beta}$ as $n \to \infty$, where $\boldsymbol{\Sigma}$ is given in (6.2). To prove the asymptotic normality of Z_n, write $Z_n = \sum_{i=1}^n Z_{i,n}$, where

$$Z_{i,n} = \frac{\sqrt{n\bar{F}^{(1)}(y_n)}}{n}\left\{\beta_1\left(\frac{|Y_i-y_n|^{p-1}\mathbb{1}_{\{Y_i>y_n\}}}{\varphi^{(p-1)}(y_n)} - 1\right) + \beta_2\left(\frac{\mathbb{1}_{\{Y_i>y_n'\}}}{\varphi^{(0)}(y_n')} - 1\right)\right\}.$$

Since $Z_{1,n}, \ldots, Z_{n,n}$ are independent and identically distributed centered random variables, according to the Lyapunov central limit theorem, a sufficient condition for asymptotic normality (6.1) is the existence of $\delta > 0$ such that $n\mathbb{E}\left[|Z_{1,n}|^{2+\delta}\right] \to 0$ as $n \to \infty$. To this end, note that, if $T_1, \ldots, T_q$ have finite $(2+\delta)-$moments, then

$$\mathbb{E}\left[\left|\sum_{j=1}^{q}(T_j - \mathbb{E}[T_j])\right|^{2+\delta}\right]^{1/(2+\delta)} \leq 2q \max_{1\leq j\leq q}\mathbb{E}\left[|T_j|^{2+\delta}\right]^{1/(2+\delta)},$$

by the triangle inequality. Therefore, letting $\delta > 0$ be so small that $\delta < \gamma^{-1}/(p-1) - 2$, one has

$$\mathbb{E}\left[|Z_{1,n}|^{2+\delta}\right] \leq 4^{2+\delta}\bar{F}^{(1)}(y_n)^{1+\delta/2}n^{-1-\delta/2}\max\left\{\beta_1^{2+\delta}\frac{\varphi^{(p-1)(2+\delta)}(y_n)}{\varphi^{(p-1)}(y_n)^{2+\delta}}, \beta_2^{2+\delta}\frac{\bar{F}^{(1)}(y_n')}{\bar{F}^{(1)}(y_n')^{2+\delta}}\right\}.$$

Lemma 3i) entails $n\mathbb{E}\left[|Z_{1,n}|^{2+\delta}\right] = O\left(\left(n\bar{F}^{(1)}(y_n)\right)^{-\delta/2}\right) = o(1)$ and the result is proved. □

Lemma 5. *Under the conditions of Lemma 4 and with* $\boldsymbol{\Sigma}$ *given in* (6.2),

$$\sqrt{n\bar{F}^{(1)}(y_n)}\left(\frac{\hat{\bar{F}}_n^{(p)}(y_n)}{\bar{F}^{(p)}(y_n)} - 1, \frac{\hat{\bar{F}}_n^{(1)}(y_n')}{\bar{F}^{(1)}(y_n')} - 1\right) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}). \tag{6.3}$$

Proof. Let $\boldsymbol{\beta} = (\beta_1, \beta_2) \in \mathbb{R}^2$, and consider

$$\begin{aligned} Z_n &= \sqrt{n\bar{F}^{(1)}(y_n)}\beta_1\left(\frac{\hat{\bar{F}}_n^{(p)}(y_n)}{\bar{F}^{(p)}(y_n)} - 1\right) + \sqrt{n\bar{F}^{(1)}(y_n)}\beta_2\left(\frac{\hat{\bar{F}}_n^{(1)}(y_n')}{\bar{F}^{(1)}(y_n')} - 1\right) \\ &= \sqrt{n\bar{F}^{(1)}(y_n)}\left\{\beta_1\left(\frac{\widehat{\varphi}_n^{(p-1)}(y_n)}{\varphi^{(p-1)}(y_n)} - 1\right)\frac{m^{(p-1)}(y_n)}{\widehat{m}_n^{(p-1)}(y_n)} + \beta_1\left(\frac{m^{(p-1)}(y_n)}{\widehat{m}_n^{(p-1)}(y_n)} - 1\right) + \beta_2\left(\frac{\widehat{\varphi}_n^{(0)}(y_n')}{\varphi^{(0)}(y_n')} - 1\right)\right\}. \end{aligned}$$

Straightforward calculations yield $\mathbb{E}\left[\widehat{m}_n^{(p-1)}(y_n)/m^{(p-1)}(y_n)\right] = 1$, and $\mathbb{V}\mathrm{ar}\left[\widehat{m}_n^{(p-1)}(y_n)/m^{(p-1)}(y_n)\right] = O(1/n)$ by Lemma 3ii). Thus

$$\sqrt{n\bar{F}^{(1)}(y_n)}\left(\frac{m^{(p-1)}(y_n)}{\widehat{m}_n^{(p-1)}(y_n)} - 1\right) = O_{\mathbb{P}}\left(\bar{F}^{(1)}(y_n)\right) = o_{\mathbb{P}}(1),$$

and the result follows from Lemma 4. □

6.2 Proofs of main results

6.2.1 Proof of Proposition 1

Let us introduce $(z_1, z_p) \in \mathbb{R}^2$, $\sigma_n = 1/\sqrt{n(1-\alpha_n)}$, and, for all $p > 1$, focus on the probability:

$$\begin{aligned} \Phi_n(z_1, z_p) &= \mathbb{P}\left(\left\{\sigma_n^{-1}\left(\frac{\hat{q}_n^{(p)}(\alpha_n)}{q^{(p)}(\alpha_n)} - 1\right) \leq z_p\right\}\bigcap\left\{\sigma_n^{-1}\left(\frac{\hat{q}_n^{(1)}(\beta_n)}{q^{(1)}(\beta_n)} - 1\right) \leq z_1\right\}\right) \\ &= \mathbb{P}\left(\left\{\hat{q}_n^{(p)}(\alpha_n) \leq q^{(p)}(\alpha_n)(1 + z_p\sigma_n)\right\}\bigcap\left\{\hat{q}_n^{(1)}(\beta_n) \leq q^{(1)}(\beta_n)(1 + z_1\sigma_n)\right\}\right). \end{aligned}$$

Remarking that $1-\alpha_n = \bar{F}^{(p)}\left(q^{(p)}(\alpha_n)\right)$ by continuity of $\bar{F}^{(p)}$ (see Lemma 1) and $1-\beta_n = \bar{F}^{(1)}\left(q^{(1)}(\beta_n)\right)$ by continuity of $\bar{F}^{(1)}$, and using the fact that, for all y, α and k, $\hat{q}_n^{(k)}(\alpha) \leq y \Leftrightarrow \hat{\bar{F}}_n^{(k)}(y) \leq 1-\alpha$, it follows that $\Phi_n(z_1, z_p)$ can be rewritten as

$$\mathbb{P}\left(\left\{\hat{\bar{F}}_n^{(p)}\left(q^{(p)}(\alpha_n)(1 + z_p\sigma_n)\right) \leq \bar{F}^{(p)}\left(q^{(p)}(\alpha_n)\right)\right\}\bigcap\left\{\hat{\bar{F}}_n^{(1)}\left(q^{(1)}(\beta_n)(1 + z_1\sigma_n)\right) \leq \bar{F}^{(1)}\left(q^{(1)}(\beta_n)\right)\right\}\right).$$

Equivalently, letting $y_n = q^{(p)}(\alpha_n)(1+z_p\sigma_n)$ and $y'_n = q^{(1)}(\beta_n)(1+z_1\sigma_n)$, one has

$$\Phi_n(z_1,z_p) = \mathbb{P}\left(\left\{\sqrt{n\bar{F}^{(1)}(y_n)}\left(\frac{\hat{\bar{F}}_n^{(p)}(y_n)}{\bar{F}^{(p)}(y_n)}-1\right) \le \sqrt{n\bar{F}^{(1)}(y_n)}\left(\frac{\bar{F}^{(p)}\left(q^{(p)}(\alpha_n)\right)}{\bar{F}^{(p)}\left(q^{(p)}(\alpha_n)(1+z_p\sigma_n)\right)}-1\right)\right\}\right.$$
$$\left.\bigcap\left\{\sqrt{n\bar{F}^{(1)}(y_n)}\left(\frac{\hat{\bar{F}}_n^{(1)}(y'_n)}{\bar{F}^{(1)}(y'_n)}-1\right) \le \sqrt{n\bar{F}^{(1)}(y_n)}\left(\frac{\bar{F}^{(1)}\left(q^{(1)}(\beta_n)\right)}{\bar{F}^{(1)}\left(q^{(1)}(\beta_n)(1+z_1\sigma_n)\right)}-1\right)\right\}\right).$$

Since $\bar{F}^{(p)}$ is regularly varying at infinity from Lemma 3iii), its local uniform convergence properties (Bingham et al., 1989, Theorem 1.2.1) entail that $\bar{F}^{(p)}\left(q^{(p)}(\alpha_n)(1+z_p\sigma_n)\right)/\bar{F}^{(p)}\left(q^{(p)}(\alpha_n)\right) \to 1$. In addition, condition $\boldsymbol{\mathcal{C}_2(\gamma,\rho,A)}$ yields:

$$\bar{F}^{(1)}\left(q^{(1)}(\beta_n)(1+z_1\sigma_n)\right) = \bar{F}^{(1)}\left(q^{(1)}(\beta_n)\right)\left(1-\frac{z_1\sigma_n}{\gamma}(1+o(1))\right).$$

In view of (2.10), $q^{(1)}(\cdot)$ verifies a second-order condition too. The asymptotic equivalence in (Daouia et al., 2019a, Proposition 3) thus shows that $q^{(p)}(\cdot)$ also verifies a second-order condition which in turn implies that $\bar{F}^{(p)}$ verifies a second-order condition analogous to $\boldsymbol{\mathcal{C}_2(\gamma,\rho,A)}$. As a consequence, one has

$$\bar{F}^{(p)}\left(q^{(p)}(\alpha_n)(1+z_p\sigma_n)\right) = \bar{F}^{(p)}\left(q^{(p)}(\alpha_n)\right)\left(1-\frac{z_p\sigma_n}{\gamma}(1+o(1))\right),$$

and therefore:

$$\begin{cases} \dfrac{\bar{F}^{(p)}\left(q^{(p)}(\alpha_n)\right)}{\bar{F}^{(p)}\left(q^{(p)}(\alpha_n)(1+z_p\sigma_n)\right)}-1 &= \dfrac{z_p}{\gamma}\sigma_n(1+o(1)), \\ \dfrac{\bar{F}^{(1)}\left(q^{(1)}(\beta_n)\right)}{\bar{F}^{(1)}\left(q^{(1)}(\beta_n)(1+z_1\sigma_n)\right)}-1 &= \dfrac{z_1}{\gamma}\sigma_n(1+o(1)). \end{cases}$$

Since $\sqrt{n\bar{F}^{(1)}(y_n)} = \sigma_n^{-1}\sqrt{g_p(\gamma)}(1+o(1))$ from (2.6) and local uniformity of the regular variation condition, it follows that

$$\Phi_n(z_1,z_p) = \mathbb{P}\left(\left\{\sqrt{n\bar{F}^{(1)}(y_n)}\left(\frac{\hat{\bar{F}}_n^{(p)}(y_n)}{\bar{F}^{(p)}(y_n)}-1\right) \le \frac{\sqrt{g_p(\gamma)}}{\gamma}z_p(1+o(1))\right\}\right.$$
$$\left.\bigcap\left\{\sqrt{n\bar{F}^{(1)}(y_n)}\left(\frac{\hat{\bar{F}}_n^{(1)}(y'_n)}{\bar{F}^{(1)}(y'_n)}-1\right) \le \frac{\sqrt{g_p(\gamma)}}{\gamma}z_1(1+o(1))\right\}\right).$$

Another direct consequence of (2.6), using the regular variation property of $\bar{F}^{(1)}$, is $q^{(1)}(\alpha_n) = g_p(\gamma)^\gamma q^{(p)}(\alpha_n)(1+o(1))$, hence $y'_n = \theta^{-\gamma}q^{(1)}(\alpha_n)(1+o(1)) = \theta^{-\gamma}g_p(\gamma)^\gamma y_n(1+o(1))$. It only remains to apply Lemma 5 with $\lambda = \theta^{-\gamma}g_p(\gamma)^\gamma$ to conclude the proof.

6.2.2 Proof of Theorem 1

Let again $\sigma_n = 1/\sqrt{n(1-\alpha_n)}$ and for any $z \in \mathbb{R}$, focus on the probability

$$\Phi_n(z) = \mathbb{P}\left(\sigma_n^{-1}\left(\frac{\hat{\bar{F}}_n^{(1)}\left(\hat{q}_n^{(p)}(\alpha_n)\right)}{1-\alpha_n}-g_p(\gamma)\right) \le z\right) = \mathbb{P}\left(\hat{\bar{F}}_n^{(1)}\left(\hat{q}_n^{(p)}(\alpha_n)\right) \le (1-\alpha_n)\left(g_p(\gamma)+z\sigma_n\right)\right).$$

Equivalently, $\Phi_n(z) = \mathbb{P}\left(\hat{q}_n^{(p)}(\alpha_n) \ge \hat{q}_n^{(1)}(\beta_n)\right)$, where $\beta_n = 1-(1-\alpha_n)\left(g_p(\gamma)+z\sigma_n\right)$ and therefore,

$$\Phi_n(z) = \mathbb{P}\left(\sigma_n^{-1}\left(\frac{\hat{q}_n^{(p)}(\alpha_n)}{q^{(p)}(\alpha_n)}-1\right) \ge \sigma_n^{-1}\left(\frac{\hat{q}_n^{(1)}(\beta_n)}{q^{(1)}(\beta_n)}-1\right)\frac{q^{(1)}(\beta_n)}{q^{(p)}(\alpha_n)} + \sigma_n^{-1}\left(\frac{q^{(1)}(\beta_n)}{q^{(p)}(\alpha_n)}-1\right)\right).$$

From de Haan and Ferreira (2006, Theorem 2.3.9), condition $\boldsymbol{\mathcal{C}_2(\gamma,\rho,A)}$ entails

$$\frac{q^{(1)}(\beta_n)}{q^{(1)}(\alpha_n)} = g_p(\gamma)^{-\gamma}\left[1+\frac{g_p(\gamma)^{-\rho}-1}{\rho}A\left((1-\alpha_n)^{-1}\right)(1+o(1))-\frac{\gamma z\sigma_n}{g_p(\gamma)}(1+o(1))\right]. \qquad (6.4)$$

In addition, under $\boldsymbol{\mathcal{C}_2(\gamma, \rho, A)}$, Daouia et al. (2019a, Proposition 3) yields

$$\frac{q^{(p)}(\alpha_n)}{q^{(1)}(\alpha_n)} = g_p(\gamma)^{-\gamma}\left[1-\gamma R(\alpha_n,p)(1+o(1)) + \frac{g_p(\gamma)^{-\rho}-1}{\rho}A\left((1-\alpha_n)^{-1}\right)(1+o(1))\right], \tag{6.5}$$

with

$$\begin{aligned}
R(\alpha_n,p) &= -g_p(\gamma)\left[(1-\alpha_n)(1+o(1)) + K(p,\gamma,\rho)A\left((1-\alpha_n)^{-1}\right)(1+o(1))\right] \\
&- (p-1)g_p(\gamma)^{\gamma\wedge 1}\left(\frac{\mathbb{E}\left[Y\mathbb{1}_{\{0<Y<q^{(1)}(\alpha_n)\}}\right]}{q^{(1)}(\alpha_n)}\mathbb{1}_{\{\gamma\le 1\}} + (1-\alpha_n)B\left(p-1,1-\gamma^{-1}\right)\mathbb{1}_{\{\gamma>1\}}\right) \\
&+ (p-1)g_p(\gamma)^{\frac{\gamma}{\gamma_\ell\vee 1}}\frac{\mathbb{E}\left[-Y\mathbb{1}_{\{0<-Y<q^{(1)}(\alpha_n)\}}\right]}{q^{(1)}(\alpha_n)}\mathbb{1}_{\{\gamma_\ell\le 1\}} \\
&+ (p-1)g_p(\gamma)^{\frac{\gamma}{\gamma_\ell\vee 1}}F^{(1)}\left(-q^{(1)}(\alpha_n)\right)B\left(\gamma_\ell^{-1}-p+1,1-\gamma_\ell^{-1}\right)\mathbb{1}_{\{\gamma_\ell>1\}},
\end{aligned}$$

and where $K(p,\gamma,\rho)$ is given in (3.3). Combining (6.4) and (6.5) yields

$$\sigma_n^{-1}\left(\frac{q^{(1)}(\beta_n)}{q^{(p)}(\alpha_n)}-1\right) = -z\frac{\gamma}{g_p(\gamma)} + o(1) + \gamma\sigma_n^{-1}R(\alpha_n,p)(1+o(1)).$$

In view of conditions $i)$, $ii)$ and $iii)$, one has

$$\sigma_n^{-1}R(\alpha_n,p) \to \lambda = -\left[\lambda_1 b_p^{(1)}(\gamma,\rho) + \lambda_2 b_p^{(2)}(\gamma) + \lambda_3 b_p^{(3)}(\gamma,\gamma_\ell)\right] \text{ as } n\to\infty.$$

Then,

$$\Phi_n(z) = \mathbb{P}\left(\frac{g_p(\gamma)}{\gamma}\left[-\sigma_n^{-1}\left(\frac{\hat{q}_n^{(p)}(\alpha_n)}{q^{(p)}(\alpha_n)}-1\right) + \sigma_n^{-1}\left(\frac{\hat{q}_n^{(1)}(\beta_n)}{q^{(1)}(\beta_n)}-1\right)(1+o(1))\right] + \lambda g_p(\gamma) + o(1) \le z\right),$$

and the asymptotic distribution of $\sigma_n^{-1}\left(\frac{\hat{\bar{F}}_n^{(1)}\left(\hat{q}_n^{(p)}(\alpha_n)\right)}{1-\alpha_n} - g_p(\gamma)\right)$ is the same as that of

$$\frac{g_p(\gamma)}{\gamma}(-1,1)\,\sigma_n^{-1}\begin{pmatrix}\frac{\hat{q}_n^{(p)}(\alpha_n)}{q^{(p)}(\alpha_n)}-1\\ \frac{\hat{q}_n^{(1)}(\beta_n)}{q^{(1)}(\beta_n)}-1\end{pmatrix} + \lambda g_p(\gamma).$$

Finally, $\hat{\gamma}_{\alpha_n}^{(p)}$ being obtained by inverting $g_p(\cdot)$ at $\hat{\bar{F}}_n^{(1)}(\hat{q}_n^{(p)}(\alpha_n))/(1-\alpha_n)$, its asymptotic distribution is the limit in distribution of

$$\frac{g_p(\gamma)}{\gamma g_p'(\gamma)}(-1,1)\,\sigma_n^{-1}\begin{pmatrix}\frac{\hat{q}_n^{(p)}(\alpha_n)}{q^{(p)}(\alpha_n)}-1\\ \frac{\hat{q}_n^{(1)}(\beta_n)}{q^{(1)}(\beta_n)}-1\end{pmatrix} + \frac{\lambda g_p(\gamma)}{g_p'(\gamma)}$$

by the delta-method. Apply Lemma 2 and Proposition 1 with $\theta = g_p(\gamma)$ to conclude the proof.

Acknowledgements

S. Girard gratefully acknowledges the support of the Chair Stress Test, Risk Management and Financial Steering, led by the French Ecole Polytechnique and its Foundation and sponsored by BNP Paribas, as well as the support of the French National Research Agency in the framework of the Investissements d'Avenir program (ANR-15-IDEX-02).

References

Abramowitz, M. and Stegun, I. A. (1972). *Handbook of Mathematical Functions (10th printing)*. U.S. National Bureau of Standards.

Anderson, P. L. and Meerschaert, M. M. (1998). Modeling river flows with heavy tails. *Water Resources Research*, 34(9):2271–2280.

Artzner, P., Delbaen, F., Eber, J., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3):203–228.

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

Beirlant, J., Dierckx, G., and Guillou, A. (2005). Estimation of the extreme-value index and generalized quantile plots. *Bernoulli*, 11(6):949–970.

Beirlant, J., Vynckier, P., and Teugels, J. (1996). Excess functions and estimation of the extreme value index. *Bernoulli*, 2(4):293–318.

Bellini, F., Klar, B., Muller, A., and Gianin, E. R. (2014). Generalized quantiles as risk measures. *Insurance: Mathematics and Economics*, 54:41–48.

Beran, J., Schell, D., and Stehlík, M. (2014). The harmonic moment tail index estimator: asymptotic distribution and robustness. *Annals of the Institute of Statistical Mathematics*, 66(1):193–220.

Bingham, N. H., Goldie, C. M., and Teugels, J. L. (1989). *Regular Variation*. Cambridge University Press.

Bowers, M. C., Tung, W. W., and Gao, J. B. (2012). On the distributions of seasonal river flows: Lognormal or power law? *Water Resources Research*, 48(5):W05536.

Brilhante, M. F., Gomes, M. I., and Pestana, D. (2013). A simple generalisation of the Hill estimator. *Computational Statistics & Data Analysis*, 57(1):518–535.

Cai, J., Einmahl, J., de Haan, L., and Zhou, C. (2015). Estimation of the marginal expected shortfall: the mean when a related variable is extreme. *Journal of the Royal Statistical Society: Series B*, 2015(77):417–442.

Cai, J. and Weng, C. (2016). Optimal reinsurance with expectile. *Scandinavian Actuarial Journal*, 2016(7):624–645.

Daouia, A., Girard, S., and Stupfler, G. (2018a). Estimation of tail risk based on extreme expectiles. *Journal of the Royal Statistical Society: Series B*, 80(2):263–292.

Daouia, A., Girard, S., and Stupfler, G. (2018b). ExpectHill estimation, extreme risk and heavy tails, preprint. `https://hal.archives-ouvertes.fr/hal-01856212`.

Daouia, A., Girard, S., and Stupfler, G. (2019a). Extreme M-quantiles as risk measures: from L^1 to L^p optimization. *Bernoulli*, 25(1):264–309.

Daouia, A., Girard, S., and Stupfler, G. (2019b). Tail expectile process and risk assessment. *Bernoulli*, to appear.

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

de Haan, L. and Peng, L. (1998). Comparison of tail index estimators. *Statistica Neerlandica*, 52(1):60–70.

de Haan, L. and Resnick, S. (1998). On asymptotic normality of the Hill estimator. *Communications in Statistics*, 14(4):849–866.

de Haan, L. and Rootzén, H. (1993). On the estimation of high quantiles. *Journal of Statistical Planning and Inference*, 35(1):1–13.

Dekkers, A. L., Einmahl, J. H., and de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. *The Annals of Statistics*, 17(4):1833–1855.

Drees, H. (2003). Extreme quantile estimation for dependent data, with applications to finance. *Bernoulli*, 9(1):617–657.

Drees, H., Ferreira, A., and de Haan, L. (2004). On maximum likelihood estimation of the extreme value index. *Annals of Applied Probability*, 14:1179–1201.

Drees, H. and Kaufmann, E. (1998). Selecting the optimal sample fraction in univariate extreme value estimation. *Stochastic Processes and their Applications*, 75(2):149–172.

Fonseca, R. V., Mondal, D., and Zhang, L. (2019). Wavelet variances for heavy-tailed time series. *Environmetrics*, 30(6):e2563.

Girard, S., Stupfler, G., and Usseglio-Carleve, A. (2019). Nonparametric extreme conditional expectile estimation, preprint. `https://hal.archives-ouvertes.fr/hal-02114255`.

Gomes, M. and Guillou, A. (2015). Extreme value theory and statistics of univariate extremes: a review. *International Statistical Review*, 83:263–292.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5):1163–1174.

Hosking, J., Wallis, J., and Wood, E. (1987). Estimation of the generalized extreme value distribution by the method of probability-weighted moments. *Technometrics*, 27:251–261.

Jones, M. (1994). Expectiles and M-quantiles are quantiles. *Statistics & Probability Letters*, 20:149–153.

Koenker, R. and Bassett, G. J. (1978). Regression quantiles. *Econometrica*, 46(1):33–50.

Koutsoyiannis, D. (2004). Statistics of extremes and estimation of extreme rainfall: Ii. empirical investigation of long rainfall records. *Hydrological Sciences Journal*, 49(4):591–610.

Li, D., Peng, L., and Yang, J. (2010). Bias reduction for high quantiles. *Journal of Statistical Planning and Inference*, 140(9):2433–2441.

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

McElroy, T. and Nagaraja, C. (2016). Tail index estimation with a fixed tuning parameter fraction. *Journal of Statistical Planning and Inference*, 170:27–45.

McElroy, T. and Politis, D. (2007). Moment-based tail index estimation. *Journal of Statistical Planning and Inference*, 137(4):1389–1406.

Newey, W. and Powell, J. (1987). Asymmetric least squares estimation and testing. *Econometrica*, 55(4):819–847.

Paulauskas, V. and Vaičiulis, M. (2013). On an improvement of Hill and some other estimators. *Lithuanian Mathematical Journal*, 53(3):336–355.

Paulauskas, V. and Vaičiulis, M. (2017). A class of new tail index estimators. *Annals of the Institute of Statistical Mathematics*, 69(2):461–487.

Rootzén, H. and Tajvidi, N. (1997). Extreme value statistics and wind storm losses: a case study. *Scandinavian Actuarial Journal*, 1997(1):70–94.

Smith, R. (1987). Estimating tails of probability distributions. *Annals of Statistics*, 15:1174–1207.

Taylor, J. W. (2008). Estimating Value at Risk and Expected Shortfall Using Expectiles. *Journal of Financial Econometrics*, 6(2):231–252.

Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, 73(364):812–815.

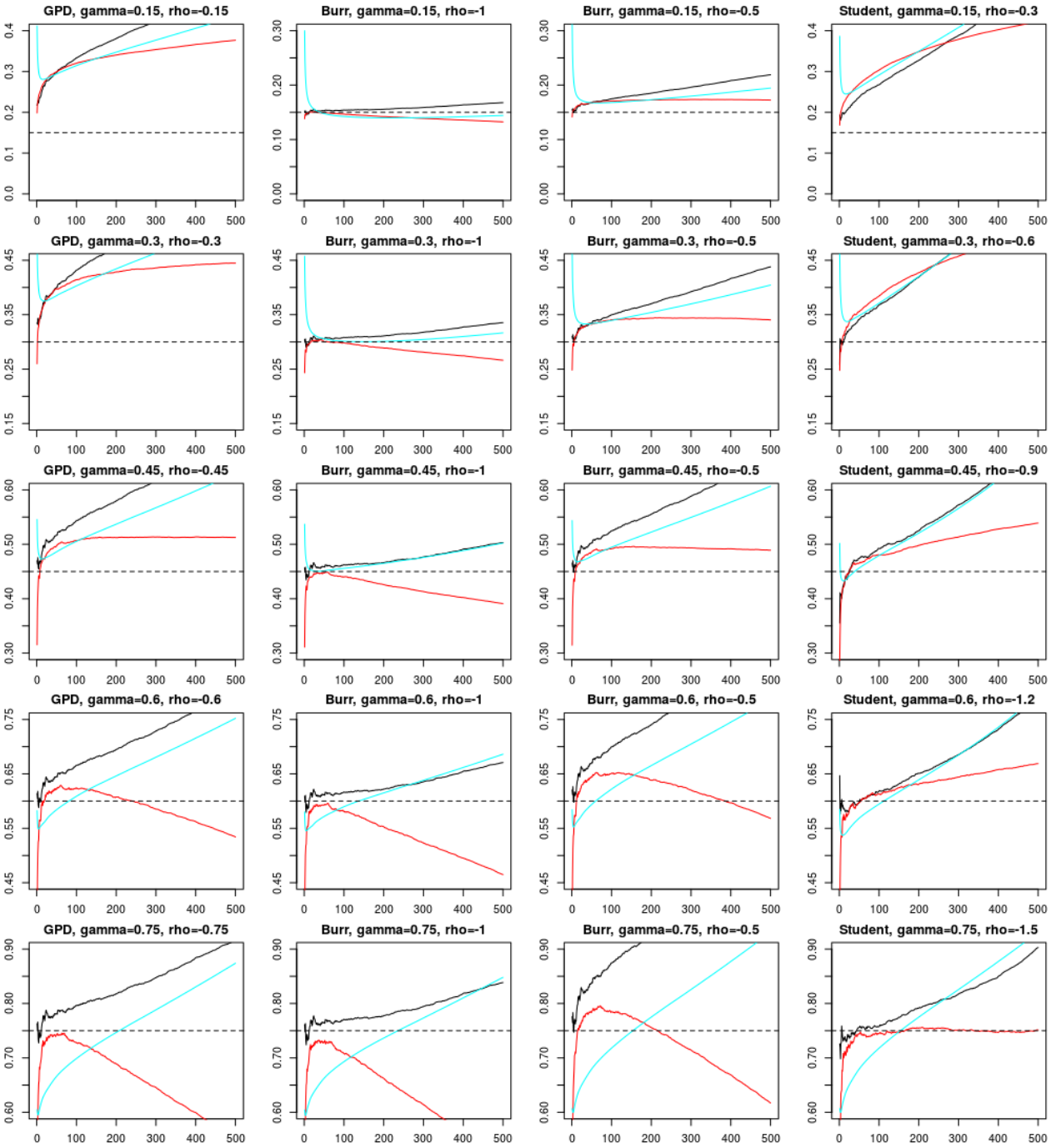


Figure 3: Empirical mean of the estimators: Hill (black), expectHill (cyan) and L^p-quantile based (red). Horizontally $n(1-\alpha_n)$, vertically $\bar{\tilde{\gamma}}_n$. From left to right: GPD(γ), Burr($\gamma,-1$), Burr($\gamma,-0.5$) and Student(γ). From top to bottom: $\gamma = 0.15$, 0.3, 0.45, 0.6 and 0.75. The true value of γ is depicted by the horizontal dotted line. The value of p in the L^p-quantile based estimator is $p = 2$ in the first row, 1.7 in the second and third rows, and 1.4 in the two bottom rows.

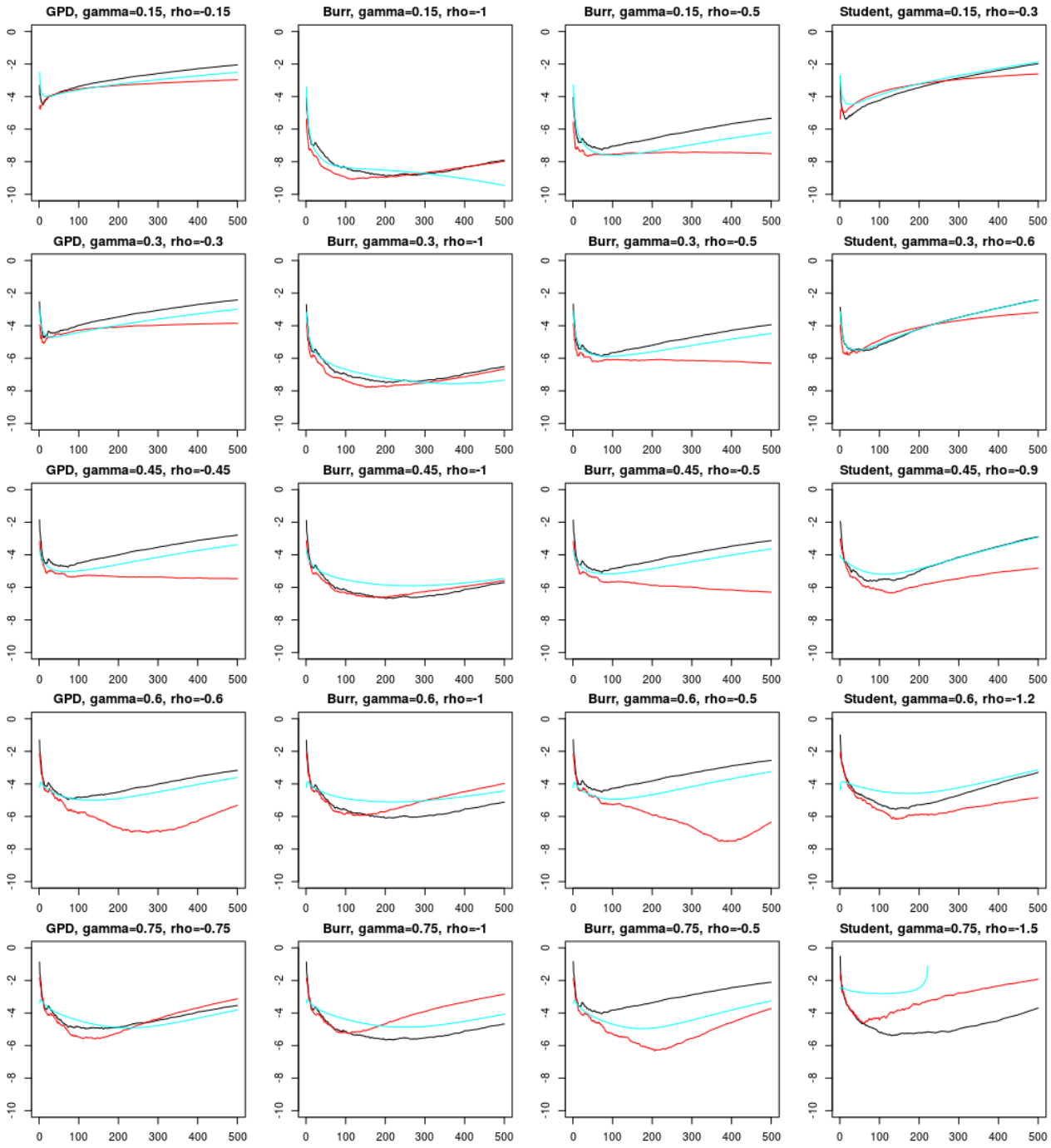


Figure 4: Log-empirical mean squared errors of the estimators: Hill (black), expectHill (cyan) and L^p-quantile based (red). Horizontally $n(1-\alpha_n)$, vertically $\log \widehat{MSE}_n$. From left to right: GPD(γ), Burr($\gamma, -1$), Burr($\gamma, -0.5$) and Student(γ). From top to bottom: $\gamma = 0.15$, 0.3, 0.45, 0.6 and 0.75. The value of p in the L^p-quantile based estimator is $p = 2$ in the first row, 1.7 in the second and third rows, and 1.4 in the two bottom rows.

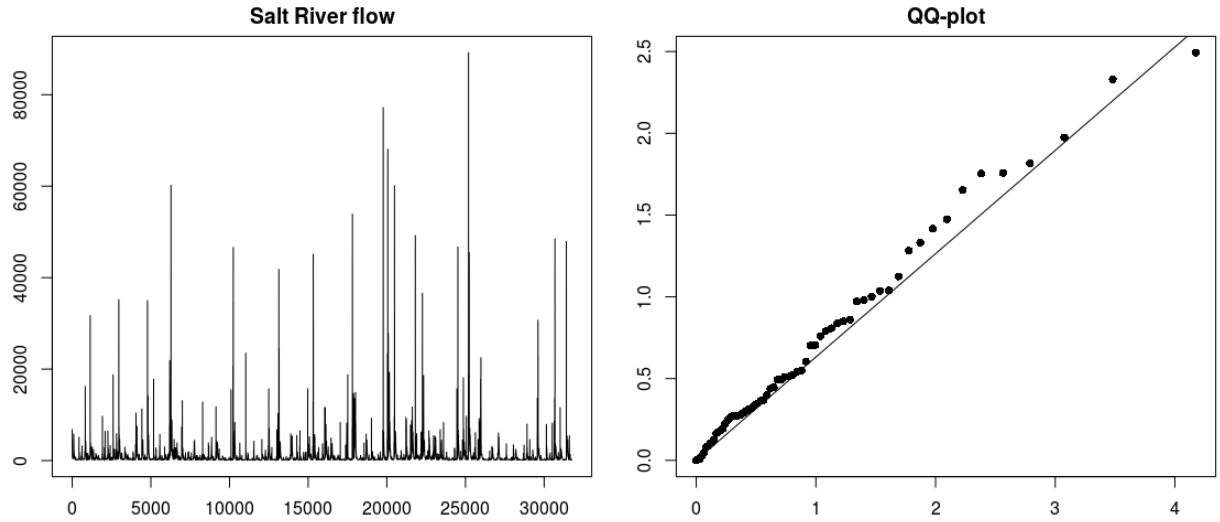


Figure 5: Daily river flows (in cubic feet per second) of the Salt River near Roosevelt, Arizona (left panel) and exponential quantile-quantile plot (right panel). Vertically: log-spacings $\log\left(Y_{n-i+1,n}/Y_{n-k_n^*+1,n}\right)$. Horizontally: $\log(k_n^*/i)$, $1 \le i \le k_n^* = 65$. The straight line has slope $\tilde{\gamma}^{(p)}_{\alpha_n^*}$.

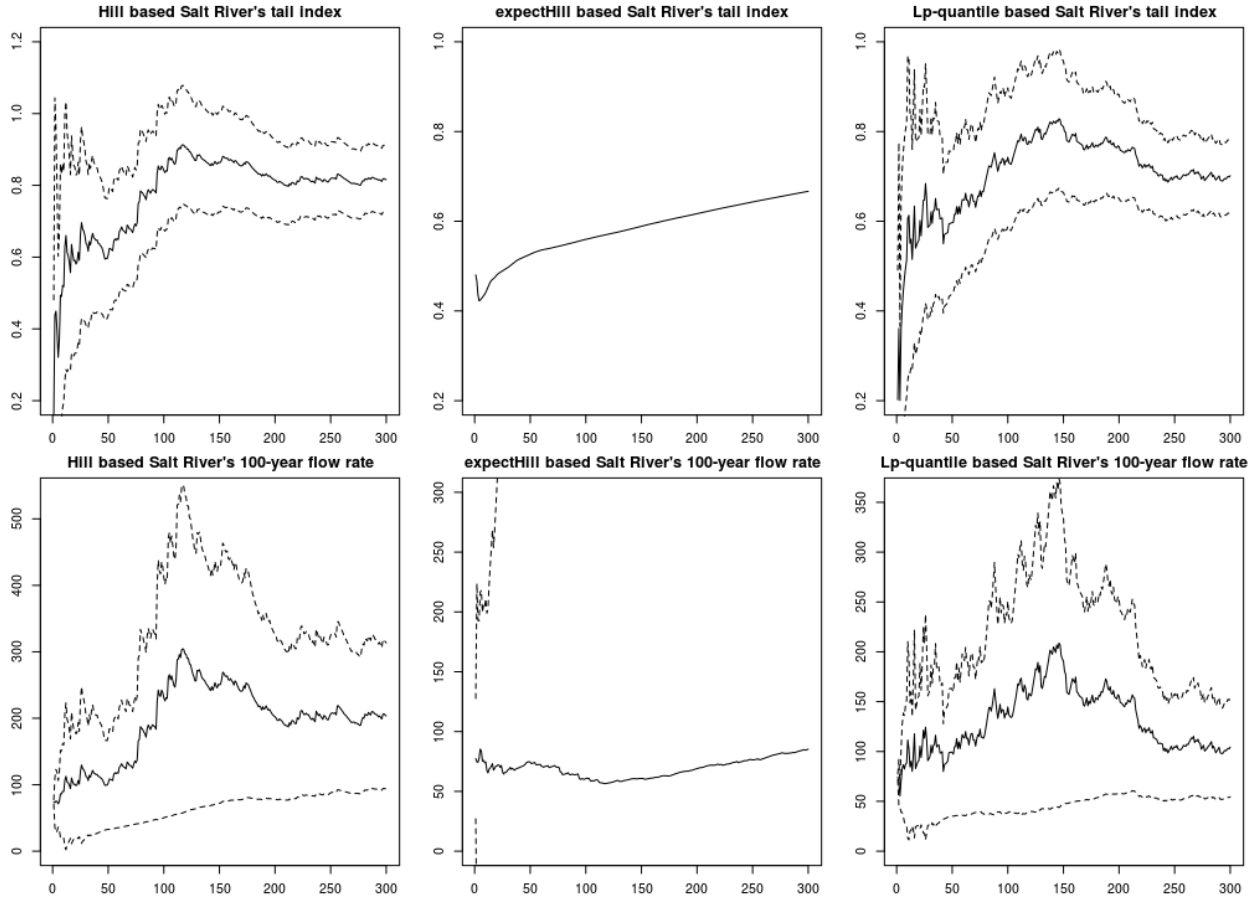


Figure 6: Daily river flow rates of the Salt River near Roosevelt, Arizona. Horizontally: $n(1-\alpha_n)$. Top panels: tail index estimators (continuous lines) $\hat{\gamma}^{(H)}_{\alpha_n}$, $\tilde{\gamma}^{(EH)}_{\alpha_n}$, $\tilde{\gamma}^{(p)}_{\alpha_n}$ (from left to right), and associated 95% asymptotic confidence intervals (dotted lines). Bottom panels: 100-year flow rate estimators based on $\hat{\gamma}^{(H)}_{\alpha_n}$, $\tilde{\gamma}^{(EH)}_{\alpha_n}$, $\tilde{\gamma}^{(p)}_{\alpha_n}$ (from left to right), and asymptotic 95% confidence intervals.