HAL open science

A refined Weissman estimator for extreme quantiles

Jonathan El Methni, Stéphane Girard

▶ To cite this version:

Jonathan El Methni, Stéphane Girard. A refined Weissman estimator for extreme quantiles. 2021. hal-03266676v2

HAL Id: hal-03266676
https://inria.hal.science/hal-03266676v2

Preprint submitted on 8 Sep 2021 (v2), last revised 15 Mar 2022 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A refined Weissman estimator for extreme quantiles

Jonathan El Methni[(1,⋆)] and Stéphane Girard[(2)]

[(1)] Université de Paris, CNRS, MAP5, UMR 8145, 75006 Paris, France

[(⋆)] Corresponding author: jonathan.el-methni@parisdescartes.fr

[(2)] Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France.

Abstract

Weissman extrapolation device for estimating extreme quantiles from heavy-tailed distribution is based on two estimators: an order statistic to estimate an intermediate quantile and an estimator of the tail-index. The common practice is to select the same intermediate sequence for both estimators. In this work, we show how an adapted choice of two different intermediate sequences leads to a reduction of the asymptotic bias associated with the resulting refined Weissman estimator. This new bias reduction method is fully automatic and does not involve the selection of extra parameters. Our approach is compared to other bias reduced estimators of extreme quantiles both on simulated and real data.

Key words: extreme quantile, bias reduction, heavy-tailed distribution, extreme-value statistics, asymptotic normality.

1 Introduction

Assessing the extreme behaviour of a random phenomenon is a major issue in quantitative finance, assurance and environmental science. For instance, extreme weather events may have strong negative and simultaneous impacts, including loss of life, damages to buildings, decrease of agricultural production, as well as longer term economic consequences [34, 41]. Assuming the phenomenon of interest is modelled by a quantitative random variable X, the associated risk is usually represented by a quantile $q(\alpha)$ such that $\mathbb{P}(X > q(\alpha)) = \alpha$. In finance or insurance, $q(\alpha)$ is referred to as the Value at Risk (VaR), while in environmental sciences, $q(\alpha)$ is referred to as the return level. Focusing on extreme risks, when α is small, the quantile of interest may be larger than the maximal observation. Indeed, denoting by n the size of an independent and identically distributed (i.i.d.) sample and by $X_{n,n}$ the maximal observation, it is easily seen that $\mathbb{P}(X_{n,n} \leq q(\alpha)) = \exp(-n\alpha(1 + o(1))) \to 1$ provided that $n\alpha \to 0$. The empirical cumulative distribution function of X being non consistent in such a situation, dedicated extrapolation methods are then necessary to estimate the so-called *extreme* quantile $q(\alpha)$.

The celebrated Weissman estimator [44] assumes that the distribution of X is heavy-tailed, *i.e.* the associated survival function $\bar{F}(x)$ decays like a power function $x^{-1/\gamma}$ as $x \to \infty$, with $\gamma > 0$, see (1) in the next section for a formal definition. As a consequence, $q(\alpha)$ can be estimated by combining two ingredients: an order statistic and an estimator of the tail-index γ. This extrapolation principle has been adapted to a variety of situations: light-tailed distributions [17], conditional distributions (to account for covariates) [9, 23] and other risk measures including expectiles [11], M-quantiles [12], Wang risk measures [18], extremiles [10], marginal expected shortfall [7], to cite a few.

Since the reliability of the extrapolations provided by Weissman device heavily depends on the quality of the estimation of the tail-index, a lot of efforts have been made to improve the original Hill estimator [33]. A number of bias reduction techniques for estimating γ have been introduced [6, 24, 25] and their consequences on Weissman estimator have been investigated in [26, 27, 29]. These estimators are described in further details in Section 3. Let us note that all of them are dedicated to the particular Hall-Welsh class of heavy-tailed distributions, see (9) below and [31, 32].

In this work, a different direction is explored to reduce the bias of Weissman estimator. We show that the biases associated with Weissman extrapolation device and the tail-index estimator may asymptotically cancel out in the extreme quantile estimator thanks to an appropriate tuning of the number of upper order statistics involved in the tail-index estimator. The construction of the resulting estimator is presented in Section 2 both from a theoretical and a practical point of view. In particular, an asymptotic normality result is provided, emphasizing that the proposed extreme quantile estimator is asymptotically unbiased in contrast to the original Weissman estimator. Its performances are illustrated on simulated data in Section 3 and compared to state-of-the-art competitors. An illustration on an actuarial real data set is provided in Section 4. Finally, a small discussion is proposed in Section 5 and the proofs are postponed to the Appendix.

2 A refined Weissman estimator

2.1 Statistical framework

Let $X_1, \dots, X_n$ be an i.i.d. sample from a cumulative distribution function F and let $X_{1,n} \leq \dots \leq X_{n,n}$ denote the order statistics associated with this sample. We denote by U the associate tail quantile function defined as $U(t) = F^{\leftarrow}(1 - 1/t)$ for all $t > 1$, where $F^{\leftarrow}(\cdot) = \inf\{x \in \mathbb{R},\ F(x) > \cdot\}$ denotes the generalized inverse of F. In the following it is assumed that F belongs to the maximum domain attraction of Fréchet, which is equivalent to assuming that U is regularly varying with index $\gamma > 0$:

$$\lim_{t\to\infty} \frac{U(ty)}{U(t)} = y^{\gamma}, \tag{1}$$

for all $y > 0$. Recall that, equivalently, U can be rewritten as $U(t) = t^{\gamma} L(t)$ where L is a slowly-varying function, *i.e.* a regularly-varying function with index zero. See [4] for more details on regular variation theory. In such a situation, the distribution associated with U is said to be heavy-tailed and γ is called the tail-index. The goal is to estimate the extreme quantile $q(\alpha_n) = U(1/\alpha_n)$ where $\alpha_n \to 0$ as $n \to \infty$ basing on an intermediate quantile $U(n/k_n)$ where (k_n) is an intermediate sequence *i.e.* such that $k_n \in \{1, \dots, n-1\}$, $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$. In view of the regular variation property (1), one has

$$\frac{U(1/\alpha_n)}{U(n/k_n)} \simeq \left(\frac{k_n}{n\alpha_n}\right)^{\gamma} =: d_n^{\gamma},$$

as $n \to \infty$, where $d_n = k_n/(n\alpha_n)$ is the extrapolation factor. Estimating $U(n/k_n)$ by its empirical counterpart $X_{n-k_n,n}$ and γ by a convenient estimator $\hat{\gamma}_n(k_n')$ depending on another intermediate sequence (k_n') yields

$$\hat{q}_n(\alpha_n, k_n, k_n') = X_{n-k_n,n}\, d_n^{\hat{\gamma}_n(k_n')}. \tag{2}$$

One can for instance use Hill estimator [33] defined as

$$\mathrm{H}(k_n') = \frac{1}{k_n'} \sum_{i=1}^{k_n'} \log(X_{n-i+1,n}) - \log(X_{n-k_n',n}), \tag{3}$$

and choose $k_n' = k_n$ to get the original Weissman estimator [44]:

$$\hat{q}_n(\alpha_n, k_n, k_n) = X_{n-k_n,n}\, d_n^{\mathrm{H}(k_n)}. \tag{4}$$

The asymptotic properties of $\hat{q}_n(\alpha_n, k_n, k_n)$ are established for instance in [30, Theorem 4.3.8]. In the next paragraph, we show that choosing $k_n' \neq k_n$ can yield better results from an asymptotic point of view.

2.2 Asymptotic analysis

A classical device in extreme-value analysis for bias assessment is the following second-order condition that refines the initial heavy-tail assumption (1). The tail quantile function U is assumed to be second-order regularly varying with index $\gamma > 0$, second-order parameter $\rho < 0$ and an auxiliary function A having constant sign and converging to 0 at infinity, *i.e.*

$$\lim_{t\to\infty} \frac{1}{A(t)} \left(\frac{U(ty)}{U(t)} - y^\gamma \right) = y^\gamma \frac{y^\rho - 1}{\rho}, \tag{5}$$

for all $y > 0$. The auxiliary function A drives the bias of most extreme-value estimators. In particular, the asymptotic sign of A determines the sign of the asymptotic bias. Besides, necessarily, $|A|$ is regularly varying with index ρ, see for instance [30, Theorem 2.3.9], so that $|A(t)| = t^\rho \ell(t)$ with ℓ a slowly-varying function. The larger ρ is, the larger the (absolute) asymptotic bias. Numerous distributions satisfying assumption (5) can be found in [2], see also Table 1 for examples.

Our first result is a refinement of [30, Theorem 4.3.8]. It provides an asymptotic normality result for the extreme quantile estimator (2) based on two intermediate sequences (k_n) and (k_n').

Theorem 1. *Assume the second-order condition (5) holds. Let (k_n) and (k_n') be two intermediate sequences and (α_n) a sequence in $(0,1)$ such that $\alpha_n \to 0$ as $n \to \infty$. Suppose, as $n \to \infty$,*

(i) $\sqrt{k_n'} A(n/k_n') \to \lambda \in \mathbb{R}$,

(ii) *$\hat{\gamma}_n(k_n')$ is an estimator of γ such that $\sqrt{k_n'}(\hat{\gamma}_n(k_n') - \gamma) \xrightarrow{d} \mathcal{N}(\lambda\mu, \sigma^2)$ where $\mu, \sigma > 0$,*

(iii) *$d_n \to \infty$, $(\log d_n)/\sqrt{k_n'} \to 0$ and $(k_n'/k_n)^\rho / \log d_n \to c \geq 0$.*

Then, as $n \to \infty$,

$$\frac{\sqrt{k_n'}}{\log d_n} \left(\frac{\hat{q}_n(\alpha_n, k_n, k_n')}{q(\alpha_n)} - 1 \right) \xrightarrow{d} \mathcal{N}(\lambda(\mu + c/\rho), \sigma^2). \tag{6}$$

Distribution (parameters)	Density function	γ	ρ
Generalised Pareto ($\sigma, \xi > 0$)	$\sigma^{-1}\left(1+\frac{\xi}{\sigma}t\right)^{-1-1/\xi}$, $t>0$	ξ	$-\xi$
Burr ($\zeta, \theta > 0$)	$\zeta\theta t^{\zeta-1}\left(1+t^{\zeta}\right)^{-\theta-1}$, $t>0$	$1/(\zeta\theta)$	$-1/\theta$
Fréchet ($\zeta > 0$)	$\zeta t^{-\zeta-1}\exp\left(-t^{-\zeta}\right)$, $t>0$	$1/\zeta$	-1
Fisher ($\nu_1, \nu_2 > 0$)	$\frac{(\nu_1/\nu_2)^{\nu_1/2}}{B(\nu_1/2,\nu_2/2)}t^{\nu_1/2-1}\left(1+\frac{\nu_1}{\nu_2}t\right)^{-(\nu_1+\nu_2)/2}$, $t>0$	$2/\nu_2$	$-2/\nu_2$
Inverse Gamma ($\zeta, \theta > 0$)	$\frac{\theta^{\zeta}}{\Gamma(\zeta)}t^{-\zeta-1}\exp(-\theta/t)$, $t>0$	$1/\zeta$	$-1/\zeta$
Student ($\nu > 0$)	$\frac{1}{\sqrt{\nu}B\left(\nu/2,1/2\right)}\left(1+\frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}$	$1/\nu$	$-2/\nu$

Table 1: Examples of heavy-tailed distributions satisfying the second-order condition (5) with the associated values of γ and ρ. Here, $\Gamma(\cdot)$ and $B(\cdot,\cdot)$ denote respectively the Gamma and Beta functions.

If, moreover, $(\log d_n)/\sqrt{k_n'} = o(k_n^{-1/4})$, *then the following asymptotic expansion holds:*

$$\begin{aligned}\frac{\sqrt{k_n'}}{\log d_n}\left(\frac{\hat{q}_n(\alpha_n,k_n,k_n')}{q(\alpha_n)}-1\right) &= \lambda\mu + \lambda(k_n'/k_n)^{\rho}\left(\frac{1-d_n^{\rho}}{\rho\log d_n}\right)\frac{\ell(n/k_n)}{\ell(n/k_n')}(1+o(1)) \\ &+ \sigma\xi_n + \frac{\sqrt{k_n'/k_n}}{\log d_n}\gamma\xi_n', \end{aligned} \tag{7}$$

where ξ_n *and* ξ_n' *are asymptotically standard Gaussian random variables.*

Assumptions (i) and (ii) are inherited from [30, Theorem 4.3.8]: they ensure the asymptotic normality of $\hat{\gamma}_n(k_n')$ with balanced bias and variance. The role of (iii) is to control the extrapolation factor d_n appearing in (2). Compared to [30, Theorem 4.3.8], it involves the extra condition $(k_n'/k_n)^{\rho}/\log d_n \to c \geq 0$ which is used to balance the extrapolation bias with the bias of $\hat{\gamma}_n(k_n')$. In view of (6), the refined Weissman estimator inherits its asymptotic distribution from $\hat{\gamma}_n(k_n')$ with an additional bias component $\lambda c/\rho$ compared to the original Weissman estimator. Let us note that, in the particular case where $k_n' = k_n$, the above extra condition is satisfied with $c = 0$, and we recover the classical Weissman asymptotic normality result.

Let us focus on the case where the auxiliary function in (5) is given by

$$A(t) = \beta\gamma t^{\rho}(1+o(1)), \tag{8}$$

as $t\to\infty$ with $\beta\neq 0$, *i.e.* $\ell(t)\to\beta\gamma$ as $t\to\infty$. This situation arises for instance in the Hall-Welsh class of heavy-tailed distributions [31, 32] defined by

$$U(t) = Ct^{\gamma}(1+\gamma\beta t^{\rho}/\rho + o(t^{\rho})), \text{ with } C>0. \tag{9}$$

In view of (7), the asymptotic squared bias of $\hat{q}_n(\alpha_n, k_n, k'_n)$ is then given by

$$AB^2(k_n, k'_n, \alpha_n) = (\log d_n)^2 A^2(n/k'_n) \left(\mu + (k'_n/k_n)^\rho \left(\frac{1 - d_n^\rho}{\rho \log d_n} \right) \right)^2 .$$

The crucial point is that the asymptotic bias can be cancelled by letting

$$k_n^\star := k'_n(\rho, \alpha_n, k_n) = k_n \left(-\rho\mu \frac{\log d_n}{1 - d_n^\rho} \right)^{1/\rho} . \tag{10}$$

The next lemma describes the behavior of $k_n^\star$ as a function of k_n.

Lemma 1.

(i) *For all $d_n \geq 1$ and $\rho < 0$, $k_n^\star$ is an increasing function of k_n and $k_n^\star \leq \mu^{1/\rho} k_n$.*

(ii) *For all $\rho < 0$, $k_n^\star \sim \tau k_n (\log d_n)^{1/\rho}$ as $n \to \infty$, where $\tau := (-\rho\mu)^{1/\rho}$.*

(iii) *If, moreover, $\log(n\alpha_n)/\log(k_n) \to c' \leq 0$ as $n \to \infty$, then $k_n^\star \sim \tau' k_n (\log k_n)^{1/\rho}$, where $\tau' = \tau(1 - c')^{1/\rho}$.*

From (ii), it appears that $k_n^\star/k_n \to 0$ as $n \to \infty$, meaning that the number of upper order statistics used in the tail-index estimator should be asymptotically small compared to k_n. See Paragraph 2.3 for a detailed discussion in the case of Hill estimator.

The next result provides the asymptotic distribution of the extreme quantile estimator (2) computed with $k'_n = k_n^\star$.

Corollary 1. *Assume the second-order condition (5) holds with auxiliary function A given by (8). Let (k_n) be an intermediate sequence and (α_n) a sequence in $(0,1)$ such that $\alpha_n \to 0$ as $n \to \infty$. Let $d_n = k_n/(n\alpha_n) \to \infty$ such that*

$$\sqrt{k_n} (\log d_n)^{1/(2\rho)-1} A(n/k_n) \to \lambda' \in \mathbb{R} \text{ and } (\log d_n)^{1-1/(2\rho)}/\sqrt{k_n} \to 0, \tag{11}$$

as $n \to \infty$. Define $k_n^\star$ as in (10) for some $\mu > 0$ and let $\hat{\gamma}_n(k_n^\star)$ be an estimator of γ such that $\sqrt{k_n^\star}(\hat{\gamma}_n(k_n^\star) - \gamma) \xrightarrow{d} \mathcal{N}(\lambda'\mu, \sigma^2)$ where $\sigma > 0$. Then, as $n \to \infty$,

$$\frac{\sqrt{k_n^\star}}{\log d_n} \left(\frac{\hat{q}_n(\alpha_n, k_n, k_n^\star)}{q(\alpha_n)} - 1 \right) \xrightarrow{d} \mathcal{N}(0, \sigma^2). \tag{12}$$

As expected, the resulting asymptotic Gaussian distribution in (12) is centered, in contrast to (6). A possible choice of sequences is $k_n = n^{-2\rho/(1-2\rho)}$ and $\alpha_n = n^{-a}$ for all $a > 1/(1-2\rho)$, leading to $c' = (1-a)(1-2\rho)/(-2\rho)$ in Lemma 1(iii). These sequences yield the usual rate of convergence of order $n^{-\rho/(1-2\rho)}$ in (12), up to a logarithmic factor.

In practice, the refined Weissman estimator is computed using $\hat{k}_n^\star$, an estimation of the intermediate sequence given in (10):

$$\hat{k}_n^\star := \left\lfloor k_n \left(-\hat{\rho}_n \mu(\hat{\rho}_n) \frac{\log d_n}{1 - d_n^{\hat{\rho}_n}} \right)^{1/\hat{\rho}_n} \right\rfloor , \tag{13}$$

with $\hat{\rho}_n$ an estimator of the second order parameter ρ and where $\lfloor \cdot \rfloor$ denotes the integer part. The estimation of ρ has been extensively discussed in the extreme-value literature, we refer to [15] for a review in the particular case of heavy-tailed distributions (which is our situation there) and to Paragraph 2.4 for implementation details. In (13), we assumed that the asymptotic bias μ of $\hat{\gamma}_n$ only depends on ρ. It can then be shown that the estimated intermediate sequence $\hat{k}_n^\star$ is asymptotically equivalent to the theoretical one $k_n^\star$.

Lemma 2. *Let $\log(n\alpha_n)/\log(k_n) \to c' \leq 0$ as $n \to \infty$ and consider $\hat{\rho}_n$ an estimator of $\rho < 0$ such that $(\log k_n)(\hat{\rho}_n - \rho) = O_P(1)$. Assume that $\mu(\cdot)$ in (13) is a positive continuous function. Then, the estimated intermediate sequence verifies $\hat{k}_n^\star = k_n^\star(1 + o_P(1))$.*

The required consistency condition on $\hat{\rho}_n$ is rather weak. It is fulfilled for instance by estimators from [22, 32, 45] as a consequence of [15, Lemma 1 and Theorem 2] and by estimators introduced in [8, 20, 21, 28, 38] as a consequence of [15, Lemma 2 and Theorem 2]. The refined Weissman estimator $\hat{q}_n(\alpha_n, k_n, \hat{k}_n^\star)$ using the estimated intermediate sequence thus inherits its asymptotic distribution from its theoretical counterpart $\hat{q}_n(\alpha_n, k_n, k_n^\star)$.

Corollary 2. *Assume the second-order condition (5) holds with auxiliary function A given by (8). Let (k_n) be an intermediate sequence and (α_n) in $(0, 1)$ such that $\alpha_n \to 0$, $\log(n\alpha_n)/\log(k_n) \to c' \leq 0$ and $\sqrt{k_n}(\log k_n)^{1/(2\rho)-1}A(n/k_n) \to \lambda' \in \mathbb{R}$ as $n \to \infty$. Define $k_n^\star$, $\hat{k}_n^\star$ as in (10), (13) respectively (with $\mu(\cdot)$ a positive continuous function) and assume that*

- *$\hat{\gamma}_n(k_n^\star)$ is an estimator of γ such that $\sqrt{k_n^\star}(\hat{\gamma}_n(k_n^\star) - \gamma) \xrightarrow{d} \mathcal{N}(\lambda'\mu(\rho), \sigma^2)$ where $\sigma > 0$,*
- *$\hat{\rho}_n$ is an estimator of $\rho < 0$ such that $(\log k_n)(\hat{\rho}_n - \rho) = O_P(1)$.*

Then, as $n \to \infty$,

$$\frac{\sqrt{k_n^\star}}{\log d_n}\left(\frac{\hat{q}_n(\alpha_n, k_n, \hat{k}_n^\star)}{q(\alpha_n)} - 1\right) \xrightarrow{d} \mathcal{N}(0, \sigma^2). \tag{14}$$

Let us highlight that convergence (14) also holds with the random rate of convergence $\sqrt{\hat{k}_n^\star}/\log d_n$ (see (29) in the Appendix) so that asymptotic confidence intervals on $q(\alpha_n)$ can easily be derived. Some examples of tail-index estimators satisfying the conditions of the above theoretical results are provided in the next paragraph, as well as their associated asymptotic mean $\mu(\cdot)$ and variance σ^2.

2.3 Examples

Hill estimator. Let us first focus on the case where $\hat{\gamma}_n(\cdot)$ is Hill estimator (3). The assumptions of the above results are satisfied with $\mu(\rho) = 1/(1-\rho)$ and $\sigma = \gamma$, see for instance [30, Theorem 3.2.5], leading to

$$k_n^{\mathrm{H},\star} = k_n^{\mathrm{H},\star}(\rho, \alpha_n, k_n) = k_n \left(\frac{-\rho}{1-\rho}\frac{\log d_n}{1-d_n^\rho}\right)^{1/\rho}. \tag{15}$$

Remark that, when $\rho \to -\infty$, then $k_n^{\mathrm{H},\star}(-\infty, \alpha_n, k_n) = k_n$ and we find back the original Weissman estimator (4). At the opposite, when $\rho \to 0$, one has $k_n^{\mathrm{H},\star}(0, \alpha_n, k_n) = ek_n/\sqrt{d_n} \sim ek_n^{\frac{1+c'}{2}+o(1)}$ as $n \to \infty$ under the condition $\log(n\alpha_n)/\log(k_n) \to c' \leq 0$. It thus appear that, in the low bias situation, $k_n^{\mathrm{H},\star}$ and k_n are of the same order. Conversely, in the high bias situation, $k_n^{\mathrm{H},\star}$ is significantly smaller than k_n, see Figure 1 for an illustration in the case $\alpha = 1/n$ *i.e.* $d_n = k_n$ and $n = 500$. It appears that, the larger ρ is, the smaller $k_n^{\mathrm{H},\star}$ is, in order to dampen the extrapolation bias.

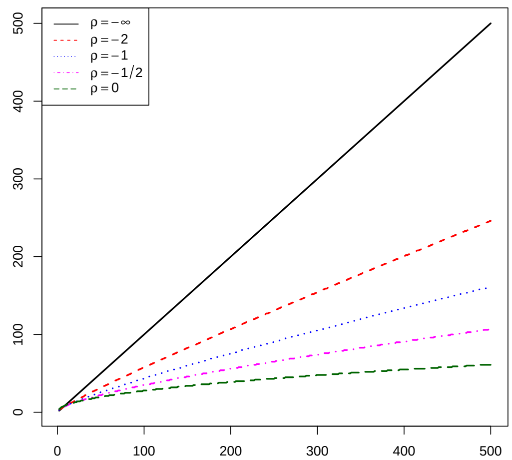

Figure 1: $k_n^{\mathrm{H},\star}$ as a function of k_n for $\rho \in \{-\infty, -2, -1, -1/2, 0\}$ when $\alpha_n = 1/n$ *i.e.* $d_n = k_n$ and $n = 500$.

Other examples. Similarly, Zipf estimator, based on a least-squares regression on the quantile-quantile plot and proposed simultaneously by [35, 42], fulfills the assumptions of Theorem 1 with $\mu(\rho) = 1/(1-\rho)^2$ and $\sigma^2 = 2\gamma^2$, see [1]. Finally, the maximum likelihood estimator, the moment [14] and Pickands estimator [39] all satisfy the assumptions of the above theorem, we refer respectively to [30, Theorem 3.4.2], [30, Theorem 3.5.4] and [30, Theorem 3.3.5] for the associated values of (μ, σ^2).

In the sequel, we focus on Hill estimator, which is the tail-index estimator with smallest variance among the above mentioned ones.

2.4 Implementation

In practice, the refined Weissman estimator is computed using Hill estimator $\mathrm{H}(\hat{k}_n^{\mathrm{H},\star})$ where

$$\hat{k}_n^{\mathrm{H},\star} := k_n^{\mathrm{H},\star}(\hat{\rho}_n, \alpha_n, k_n) = k_n \left(\frac{-\hat{\rho}_n}{1-\hat{\rho}_n} \frac{\log d_n}{1 - d_n^{\hat{\rho}_n}} \right)^{1/\hat{\rho}_n} \tag{16}$$

is an estimation of the intermediate sequence given in (15). Here, we adopt the estimator $\hat{\rho}_n$ of the second-order parameter introduced in [20, Equation (2.18)] and implemented in the package evt0 of the R software [36] which features a satisfying behaviour in practice. The resulting refined Weissman estimator is denoted by

$$\mathrm{RW}(k_n) = \hat{q}_n(\alpha_n, k_n, \hat{k}_n^{\mathrm{H},\star}) = \hat{q}_n(\alpha_n, k_n, k_n^{\mathrm{H},\star}(\hat{\rho}_n, \alpha_n, k_n)).$$

This data-driven choice of $\hat{k}_n^{\mathrm{H},\star}$ as a function of k_n is illustrated on Figure 2 on data sets of size $n = 500$ together with its consequences on the estimation of the tail-index (left panel) and extreme quantiles (right panel). Two Burr distributions are considered with $\gamma = 1/4$ and $\rho \in \{-2, -3/4\}$, see Table 1. The Relative mean-squared error (RMSE) is computed on $N = 1000$ replications, see (21) below. It appears that, in the low bias situation ($\rho = -2$, top panel), $\hat{k}_n^{\mathrm{H},\star}$ is automatically limited to the range $\{1, \dots, 252\}$ as $k_n \in \{1, \dots, n-1\}$. In the high bias situation ($\rho = -3/4$, bottom panel), $\hat{k}_n^{\mathrm{H},\star}$ is further limited to the range $\{1, \dots, 118\}$. In both cases, the bias associated with Hill estimator $\mathrm{H}(\hat{k}_n^{\mathrm{H},\star})$ remains acceptable. As a consequence, the relative mean-squared error associated with $\mathrm{RW}(k_n)$ is small for a wide range of values of k_n, in contrast to the original Weissman estimator. These preliminary numerical experiments are conducted in a more systematic way in the following section.

3 Validation on simulated data

The proposed refined Weissman estimator is compared on simulated data to the original Weissman estimator and to six other bias-reduced estimators of the extreme quantile.

Experimental design. The comparison is achieved on the following six heavy-tailed distributions: Burr, Fréchet, Fisher, generalized Pareto distribution (GPD), Inverse Gamma, and Student. All of them satisfy the second-order condition (5) with (8), see Table 1 for their definitions and associated values of γ and ρ. For all distributions, four tail-index values $\gamma \in \{1/8, 1/4, 1/2, 1\}$ are investigated. The choice of the second-order parameter ρ depends on the considered distribution:

- Burr distribution: the second order parameter can be chosen independently from the tail-index, five values are tested $\rho \in \{-1/8, -1/4, -1/2, -1, -2\}$.
- Fréchet distribution: the second order parameter is fixed to $\rho = -1$.
- Fisher, Generalized Pareto Distribution and Inverse Gamma: the second order parameter is fixed to $\rho = -\gamma$.
- Student distribution: the second order parameter is fixed to $\rho = -2\gamma$.

In each case, we simulate $N = 1000$ replications of a data set of $n = 500$ i.i.d. realisations from the $4 \times (5 + 5) = 40$ considered parametric models. Finally, two cases are investigated for the order of the extreme quantile: $\alpha_n \in \{1/n, 1/(2n)\}$. Summarizing, this experimental design includes $40 \times 2 = 80$ configurations.

Competitors. Since the original Weissman estimator inherits its asymptotic distribution from $\mathrm{H}(k_n)$, the main idea of most of reduced bias estimators of the extreme quantile is to replace Hill estimator $\mathrm{H}(k_n)$ in (4) by a bias-reduced version. We shall first consider the Corrected-Hill (CH) [6]:

$$\mathrm{CH}(k_n) = \mathrm{H}(k_n)\left(1 - \frac{\hat{\beta}_n}{1 - \hat{\rho}_n}\left(\frac{n}{k_n}\right)^{\hat{\rho}_n}\right), \tag{17}$$

where $\hat{\rho}_n$ and $\hat{\beta}_n$ are estimators of the second-order parameters ρ and β, see (8). The associated bias reduced Weissman estimator is studied in [29]. Second, let us introduce

$$\mathrm{H}_p(k_n) = \begin{cases} \dfrac{1}{p}\left(1 - \left(\dfrac{1}{k_n}\displaystyle\sum_{i=1}^{k_n}\left(\frac{X_{n-i+1,n}}{X_{n-k_n,n}}\right)^p\right)^{-1}\right) & \text{if } p < 1/\gamma \text{ and } p \neq 0, \\ \mathrm{H}(k_n) & \text{if } p = 0, \end{cases} \tag{18}$$

the Mean-of-order-p estimator of γ proposed almost simultaneously in [3, 5, 37], where p is some tuning parameter. A bias-reduced version of the previous estimator (18), referred to as reduced-bias mean-of-order-p and denoted by CH_p, is considered in [25]:

$$\mathrm{CH}_p(k_n) = \mathrm{H}_p(k_n)\left(1 - \frac{\hat\beta_n(1-p\mathrm{H}_p(k_n))}{1-\hat\rho_n - p\mathrm{H}_p(k_n)}\left(\frac{n}{k_n}\right)^{\hat\rho_n}\right), \tag{19}$$

following the same principle as in (17), so that $\mathrm{CH}_0(k_n) = \mathrm{CH}(k_n)$. An alternative bias-reduced version of (18) is proposed in [24] by replacing in the bias correction term of (19) an optimal value of p in terms of asymptotic efficiency. This gives rise to the Partially Reduced-Bias mean-of-order-p (PRB_p) estimator defined by

$$\mathrm{PRB}_p(k_n) = \mathrm{H}_p(k_n)\left(1 - \frac{\hat\beta_n(1-\varphi_{\hat\rho_n})}{1-\hat\rho_n - \varphi_{\hat\rho_n}}\left(\frac{n}{k_n}\right)^{\hat\rho_n}\right), \tag{20}$$

with $\varphi_\rho = 1-\rho/2-\sqrt{(1-\rho/2)^2-1/2}$. Both CH_p and PRB_p estimators are plugged in Weissman estimator to obtain extreme quantile estimators, see [26]. In the following, for the sake of simplicity, the extreme quantile estimators derived from (17), (19) and (20) are denoted by their associated tail-index estimators, namely CH, CH_p and PRB_p. Finally, a Corrected Weissman (CW) estimator is introduced in [29] implementing two bias corrections: a first one in the tail-index estimator and a second one in the extrapolation factor:

$$\mathrm{CW}(k_n) = X_{n-k_n,n}\left(\frac{k_n}{n\alpha_n}\exp\left(\hat\beta_n\left(\frac{n}{k_n}\right)^{\hat\rho_n}\frac{(k_n/(n\alpha_n))^{\hat\rho_n}-1}{\hat\rho_n}\right)\right)^{\mathrm{CH}(k_n)}.$$

Selection of hyperparameters. All considered extreme quantile estimators (Weissman, RW, CH, CHP_p, PRB_p and CW) depend on the intermediate sequence k_n. The selection of k_n is a crucial point which has been widely discussed in the extreme-value literature. A standard practice is to pick out a value of k_n in the first stable part of the plot $k_n \mapsto \hat\gamma(k_n)$ where $\hat\gamma(\cdot)$ is the tail-index estimator of interest, see [30, Chapter 3]. Some attempts at formalizing this procedure can be found in [16, 40] and, more recently, in [18, 19]. Here, we adopt the sample path stability criterion described in [26, Algorithm 4.1]. Besides, CH_p and PRB_p involve an extra parameter p which also has to be selected. Two solutions are possible. First, following [26, p. 1739], one can choose the optimal value of p in terms of efficiency:

$$p^\star = \varphi_{\hat\rho_n}/\mathrm{CH}(\hat k_0) \quad \text{where} \quad \hat k_0 = \min\left(n-1, \left\lfloor\left((1-\hat\rho_n)^2 n^{-2\hat\rho_n}/(-2\hat\rho_n\hat\beta_n^2)\right)^{1/(1-2\hat\rho_n)}\right\rfloor + 1\right),$$

which gives rise to two estimators $\mathrm{CH}_{p^\star}$ and $\mathrm{PRB}_{p^\star}$. Second, one may select simultaneously k_n and p using the sample path stability criterion of [26, Algorithm 4.2]. The resulting estimators are still denoted by CH_p and PRB_p. To summarize, eight extreme quantile estimators are compared in the following: Weissman, RW, CH, CHP_p, PRB_p, $\mathrm{CHP}_{p^\star}$, $\mathrm{PRB}_{p^\star}$ and CW.

Results. The performance of the extreme quantile estimators is assessed using the Relative mean-squared error (RMSE):

$$\mathrm{RMSE}\left(\hat q_n(\alpha_n)\right) = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\hat q_n(\alpha_n)}{q_n(\alpha_n)} - 1\right)^2. \tag{21}$$

The results are provided in Table 3 ($\alpha_n = 1/n$), Table 5 ($\alpha_n = 1/(2n)$), for the Burr distribution and in Table 4 ($\alpha_n = 1/n$), Table 6 ($\alpha_n = 1/(2n)$) for Fréchet, Fisher, GPD, Inverse Gamma, and Student distributions. Let us first remark that, in the case $\gamma = 1$, for all distributions with $\rho > -2$, all estimators fail in estimating the extreme quantiles $q(1/n)$ and $q(1/(2n))$: this corresponds to the most difficult situation where both γ and ρ are large. Besides, on the 80 considered situations, the original Weissman estimator never yields the best result, indicating that bias reduction is useful in general. At the opposite, the proposed RW estimator is the most accurate one since it provides the best result 41 out of 80 times. The second main accurate estimator is CH_p which provides the best result only 10 out of 80 times. As a conclusion, it appears on these experiments on simulated data that, in average, the RW estimator performs better than the seven considered competitors. The behaviour of these estimators on real data is illustrated in the next section.

4 Illustration on an actuarial data set

We consider here the Secura Belgian reinsurance data set on automobile claims from 1998 until 2001, introduced in [2] and further analyzed in [18] from an extreme risk measures perspective. This data set consists of $n = 371$ claims which were at least as large as 1.2 million Euros and were corrected for inflation. Our goal is to estimate the extreme quantile $q(1/n)$ (with $1/n \approx 0.0027$) and to compare it to the maximum of the sample $x_{n,n} = 7.898$ million Euros.

The first step is to estimate the second order parameter. We get $\hat{\rho}_n = -0.756$ (see Paragraph 2.4 for implementation details) which corresponds to a relatively high bias situation. We refer to Figure 2 for an illustration in a similar simulated situation with $\rho = -3/4$. Second, the estimated intermediate sequence is then computed from (16): $\hat{k}_n^{\mathrm{H},\star} \in \{1, \dots, 107\}$ as $k_n \in \{1, \dots, n-1 = 370\}$. The associated Hill plots $\mathrm{H}(k_n)$ and $\mathrm{H}(\hat{k}_n^{\mathrm{H},\star})$ are displayed on the left panel of Figure 3. The sample path stability criterion [26, Algorithm 4.1] selects $k_n = 209$ leading to $\hat{k}_n^{\mathrm{H},\star} = 68$ and $\mathrm{H}(\hat{k}_n^{\mathrm{H},\star}) = 0.2802$ as estimated tail-index. As a visual check, a quantile-quantile plot of the log-excesses $\log(X_{n-i+1,n}) - \log(X_{n-\hat{k}_n^{\mathrm{H},\star},n})$ against the quantiles of the unit exponential distribution $\log(\hat{k}_n^{\mathrm{H},\star}/i)$ for $i = 1, \dots, \hat{k}_n^{\mathrm{H},\star}$ is drawn on the right panel of Figure 3. The relationship appearing in this plot is approximately linear, which constitutes an empirical evidence that the heavy-tail assumption makes sense and that $\hat{k}_n^{\mathrm{H},\star} = 68$ is a reasonable choice to estimate the tail-index.

The eight estimates of the tail-index (see Section 3) are reported in Table 2. For the last two estimators, the automatic selection procedure provided $p^\star = 0.765$. The original Hill and CW estimators point towards values similar to our result while the remaining five estimators provide smaller estimated tail-indices. The corresponding estimated extreme quantiles $\hat{q}_n(1/n)$ are also reported. Note that in [18], the authors obtained $\hat{q}_n(0.005) = 7.163$ and $\hat{q}_n(0.001) = 10.899$. It appears from Table 2 that the estimations provided by CH_p, PRB_p and $\mathrm{PRB}_{p^\star}$ are not coherent with the previous results since, in these cases $\hat{q}_n(1/n) \leq \hat{q}_n(0.005)$ while $1/n < 0.005$. Underestimation can then be suspected for these three estimators. The proposed refined Weissman estimator gives the closest estimation of the maximum value of the sample: $\mathrm{RW}(1/n) = 8.328$ and $x_{n,n} = 7.898$ (millions Euros). Besides, the maximum value does belong to the asymptotic 95% confidence interval $[5.364, 11.291]$ associated with $\mathrm{RW}(1/n)$, which is thus a reasonable estimation of $q(1/n)$.

As a conclusion, according to $\mathrm{RW}(1/n)$ estimate, one can expect a claim larger than 8.328 million Euros to occur in average once every four years.

	Weissman	RW	CW	CH	CH_p	PRB_p	$\mathrm{CH}_{p^\star}$	$\mathrm{PRB}_{p^\star}$
$\hat\gamma_n$	0.2855	0.2802	0.2819	0.2500	0.2191	0.2351	0.2493	0.2471
$\hat q_n(1/n)$	9.362	8.328	9.070	7.358	6.575	6.585	7.354	7.294

Table 2: Comparison of eight estimators on the Secura Belgian reinsurance actuarial data set. Estimates of the tail-index γ and of the extreme quantile $q(1/n)$ (in million Euros).

5 Conclusion

As a conclusion, it appears that the refined Weissman estimator is an efficient tool for estimating extreme quantiles in a variety of heavy-tailed situations. In contrast to usual bias reduced estimators, our proposition is not based on a preliminary reduction of the bias associated with some tail-index estimator. In constrast, it relies on an original idea consisting in selecting carefully two intermediate sequences to make the asymptotic bias vanish. This methodology requires an estimator of the second order parameter, but, unlike some other estimators, it does not involve extra tuning parameters, making the refined Weissman estimator fully data driven.

Our further work will consist in extending this bias reduction principle in the more general context of an arbitrary maximum domain of attraction.

Appendix: proofs

Proof of Theorem 1. It follows the same lines as the one of [30, Theorem 4.3.8]. Let us consider the expansion

$$\frac{\sqrt{k_n'}}{\log d_n}\left(\frac{\hat q_n(\alpha_n,k_n,k_n')}{q(\alpha_n)}-1\right)=\frac{\sqrt{k_n'}}{\log d_n}\left(\frac{X_{n-k_n,n}\, d_n^{\hat\gamma_n(k_n')}}{U(1/\alpha_n)}-1\right)=\frac{T_{1,n}+T_{2,n}+T_{3,n}}{T_{0,n}},$$

where we have introduced:

$$\begin{aligned}
T_{0,n} &= d_n^{-\gamma}\,\frac{U(1/\alpha_n)}{U(n/k_n)},\\
T_{1,n} &= \frac{\sqrt{k_n'}}{\log d_n}\left(\frac{X_{n-k_n,n}}{U(n/k_n)}-1\right)d_n^{\hat\gamma_n(k_n')-\gamma},\\
T_{2,n} &= \frac{\sqrt{k_n'}}{\log d_n}\left(d_n^{\hat\gamma_n(k_n')-\gamma}-1\right),\\
T_{3,n} &= \frac{\sqrt{k_n'}}{\log d_n}(1-T_{0,n}).
\end{aligned}$$

Let us first focus on $T_{0,n}$. From [30, Theorem 2.3.9], it follows from the second-order condition (5) that, for any ε, $\delta>0$, there exists $t_0>1$ such that for all $t\ge t_0$ and $x\ge 1$,

$$\left|\frac{1}{A_0(t)}\left(\frac{U(tx)}{U(t)}-x^\gamma\right)-x^\gamma\frac{x^\rho-1}{\rho}\right|\le \varepsilon x^{\gamma+\rho+\delta},$$

where A_0 is asymptotically equivalent to A. Letting $x=d_n$ and $t=n/k_n$ then yields

$$\left|\frac{T_{0,n}-1}{A_0(n/k_n)}-\frac{d_n^\rho-1}{\rho}\right|\le \varepsilon d_n^{\rho+\delta},$$

or equivalently,

$$T_{0,n} = 1 + A_0(n/k_n)\left(\frac{d_n^\rho - 1}{\rho} + \varepsilon R_n\right),$$

where $|R_n| \le d_n^{\rho+\delta}$. Now, writing $|A_0|(t) = t^\rho \ell(t)$, where ℓ is a slowly-varying function, it follows,

$$A_0(n/k_n) = A_0(n/k_n')(k_n'/k_n)^\rho \frac{\ell(n/k_n)}{\ell(n/k_n')},$$

as $n \to \infty$. As a consequence, we obtain

$$T_{0,n} = 1 + (k_n'/k_n)^\rho A_0(n/k_n')\left(\frac{d_n^\rho - 1}{\rho} + \varepsilon R_n\right)\frac{\ell(n/k_n)}{\ell(n/k_n')},$$

and letting $\varepsilon \to 0$ yields

$$T_{0,n} = 1 + (k_n'/k_n)^\rho A_0(n/k_n')\left(\frac{d_n^\rho - 1}{\rho}\right)\frac{\ell(n/k_n)}{\ell(n/k_n')}. \tag{22}$$

Second, under the assumption $\sqrt{k_n'}(\hat\gamma_n(k_n') - \gamma) \xrightarrow{d} \mathcal{N}(\lambda\mu, \sigma^2)$ as $\sqrt{k_n'}A(n/k_n') \to \lambda$, we get

$$d_n^{\hat\gamma_n(k_n')-\gamma} = \exp\left(\frac{(\log d_n)}{\sqrt{k_n'}}\sqrt{k_n'}(\hat\gamma_n(k_n') - \gamma)\right) = \exp\left(\frac{(\log d_n)}{\sqrt{k_n'}}(\lambda\mu + \sigma\xi_n)\right),$$

where $\xi_n \xrightarrow{d} \mathcal{N}(0,1)$. Recalling that $(\log d_n)/\sqrt{k_n'} \to 0$ as $n \to \infty$, the following first order expansion holds

$$d_n^{\hat\gamma_n(k_n')-\gamma} = 1 + \frac{(\log d_n)}{\sqrt{k_n'}}(\lambda\mu + \sigma\xi_n) + O_P\left(\frac{(\log d_n)^2}{k_n'}\right). \tag{23}$$

In particular, $d_n^{\hat\gamma_n(k_n')-\gamma} \xrightarrow{\mathbb{P}} 1$ and therefore,

$$T_{1,n} = \frac{\sqrt{k_n'/k_n}}{\log d_n}\sqrt{k_n}\left(\frac{X_{n-k_n,n}}{U(n/k_n)} - 1\right)(1 + o_P(1)) = \frac{\sqrt{k_n'/k_n}}{\log d_n}\gamma\xi_n''(1 + o_P(1)), \tag{24}$$

where $\xi_n'' \xrightarrow{d} \mathcal{N}(0,1)$, from [30, Theorem 2.2.1]. Third, it immediately follows from (23) that

$$T_{2,n} = \lambda\mu + \sigma\xi_n + O_P\left(\frac{\log d_n}{\sqrt{k_n'}}\right). \tag{25}$$

Finally, in view of (22) and recalling that $\sqrt{k_n'}A_0(n/k_n') \to \lambda$, one has

$$T_{3,n} = \lambda(k_n'/k_n)^\rho\left(\frac{1 - d_n^\rho}{\rho\log d_n}\right)\frac{\ell(n/k_n)}{\ell(n/k_n')}(1 + o(1)). \tag{26}$$

Collecting (22), (24), (25) and (26) yields

$$\begin{aligned}\frac{\sqrt{k_n'}}{\log d_n}\left(\frac{\hat q_n(\alpha_n, k_n, k_n')}{q(\alpha_n)} - 1\right) &= \lambda\mu + \lambda(k_n'/k_n)^\rho\left(\frac{1 - d_n^\rho}{\rho\log d_n}\right)\frac{\ell(n/k_n)}{\ell(n/k_n')}(1 + o(1)) \\ &+ \sigma\xi_n + \frac{\sqrt{k_n'/k_n}}{\log d_n}\gamma\xi_n''(1 + o_P(1)) + O_P\left(\frac{\log d_n}{\sqrt{k_n'}}\right),\end{aligned}$$

since $T_{0,n} = 1 + o(T_{3,n})$. Besides, assumptions $(k_n'/k_n)^\rho/(\log d_n) \to c \ge 0$ and $d_n \to \infty$ as $n \to \infty$ imply

$$\frac{\sqrt{k_n'}}{\log d_n}\left(\frac{\hat q_n(\alpha_n, k_n, k_n')}{q(\alpha_n)} - 1\right) \xrightarrow{d} \mathcal{N}(\lambda(\mu + c/\rho), \sigma^2),$$

and the first part of the result is proved. If, moreover, $(\log d_n)/\sqrt{k_n'} = o(k_n^{-1/4})$, then

$$\frac{\sqrt{k_n'}}{\log d_n}\left(\frac{\hat{q}_n(\alpha_n, k_n, k_n')}{q(\alpha_n)} - 1\right) = \lambda\mu + \lambda(k_n'/k_n)^\rho\left(\frac{1-d_n^\rho}{\rho\log d_n}\right)\frac{\ell(n/k_n)}{\ell(n/k_n')}(1+o(1)) + \sigma\xi_n + \frac{\sqrt{k_n'/k_n}}{\log d_n}\gamma\xi_n',$$

by letting $\xi_n' = \xi_n''(1+o_P(1))$, which proves the second part of the result. ■

Proof of Lemma 1. **(i)** Letting $f(x) = -\rho(\log x)/(1-x^\rho)$ for all $x \geq 1$ and $\rho < 0$, from (10) one has $k_n^\star = \mu^{1/\rho}k_n(f(d_n))^{1/\rho}$ with $d_n = k_n/(n\alpha_n) \geq 1$. First, routine calculations give:

$$\frac{\partial k_n^\star}{\partial k_n} = \mu^{1/\rho}(f(d_n))^{1/\rho}\left(1 + \frac{d_n}{\rho}(\log f)'(d_n)\right) = \mu^{1/\rho}(f(d_n))^{1/\rho}\left(\frac{1}{\rho\log d_n} + \frac{1}{1-d_n^\rho}\right) \geq 0,$$

for all $d_n \geq 1$. As a conclusion, $\partial k_n^\star/\partial k_n \geq 0$ which proves that $k_n^\star$ is an increasing function of k_n. Second, it is easily shown that f is increasing and $f(1) = 1$, leading to $f(d_n) \geq 1$ and thus $k_n^\star \leq \mu^{1/\rho}k_n$.
(ii) is a consequence of $f(x) \sim -\rho\log x$ as $x \to \infty$.
(iii) Remark that assumption $\log(n\alpha_n)/\log(k_n) \to c' \leq 0$ implies that $\log d_n \sim (1-c')\log k_n$ as $n \to \infty$. The conclusion follows. ■

Proof of Corollary 1. It is sufficient to prove that assumptions (i) and (iii) of Theorem 1 hold true. First, Lemma 1(ii) entails that $k_n^\star/k_n \sim \tau(\log d_n)^{1/\rho}$ as $n \to \infty$. Besides, from (8), we have $A(n/k_n^\star)/A(n/k_n) \sim (k_n^\star/k_n)^{-\rho}$, so that

$$\sqrt{k_n^\star}A(n/k_n^\star) \sim (k_n^\star/k_n)^{1/2-\rho}\sqrt{k_n}A(n/k_n) \sim \tau^{1/2-\rho}(\log d_n)^{1/(2\rho)-1}\sqrt{k_n}A(n/k_n) \to \lambda'\tau^{1/2-\rho}$$

as $n \to \infty$ in view of the first part of (11). Assumption (i) of Theorem 1 thus holds true with $\lambda = \lambda'\tau^{1/2-\rho}$. Second,

$$\frac{\log d_n}{\sqrt{k_n^\star}} = \frac{\log d_n}{\sqrt{k_n}}(k_n^\star/k_n)^{-1/2} \sim \tau^{-1/2}\frac{(\log d_n)^{1-1/(2\rho)}}{\sqrt{k_n}} \to 0 \tag{27}$$

as $n \to \infty$ in view of the second part of (11). Third,

$$\frac{(k_n^\star/k_n)^\rho}{\log d_n} \to \tau^\rho \tag{28}$$

as $n \to \infty$. Collecting (27) and (28) proves that assumption (iii) of Theorem 1 thus holds true with $c = \tau^\rho$. ■

Proof of Lemma 2. Recall that, from the proof of Lemma 1(iii), $\log d_n \sim (1-c')\log k_n$ as $n \to \infty$. Let us then observe that

$$\frac{\log d_n}{1-d_n^{\hat{\rho}_n}} = \frac{\log d_n}{1-\exp(\rho\log d_n + O_P(1))} = \frac{\log d_n}{1-O_P(d_n^\rho)} = (\log d_n)(1+O_P(d_n^\rho))$$

and consequently

$$\begin{aligned}\left(\frac{\log d_n}{1-d_n^{\hat{\rho}_n}}\right)^{1/\hat{\rho}_n} &= \exp\left(\frac{\log\log d_n + O_P(d_n^\rho)}{\rho + O_P(1/\log d_n)}\right)\\ &= \exp\left(\frac{\log\log d_n}{\rho} + O_P\left(\frac{\log\log d_n}{\log d_n}\right) + O_P(d_n^\rho)\right)\\ &= (\log d_n)^{1/\rho}(1+o_P(1)).\end{aligned}$$

Besides, since $\hat{\rho}_n$ is a consistent estimator of ρ and $\mu(\cdot)$ is continuous, it follows that $(-\hat{\rho}_n\mu(\hat{\rho}_n))^{1/\hat{\rho}_n} \xrightarrow{\mathbb{P}} (-\rho\mu(\rho))^{1/\rho}$ and therefore

$$\begin{aligned} k_n\left(-\hat{\rho}_n\mu(\hat{\rho}_n)\frac{\log d_n}{1-d_n^{\hat{\rho}_n}}\right)^{1/\hat{\rho}_n} &= k_n\left(-\rho\mu(\rho)(\log d_n)\right)^{1/\rho}(1+o_P(1)) \\ &= k_n\left(-\rho\mu(\rho)(1-c')(\log k_n)\right)^{1/\rho}(1+o_P(1)) \\ &= k_n^\star(1+o_P(1)), \end{aligned}$$

in view of Lemma 1(iii). Remarking that the right hand side term tends to infinity in probability, one immediately has

$$\hat{k}_n^\star = \left\lfloor k_n\left(-\hat{\rho}_n\mu(\hat{\rho}_n)\frac{\log d_n}{1-d_n^{\hat{\rho}_n}}\right)^{1/\hat{\rho}_n}\right\rfloor = k_n^\star(1+o_P(1)),$$

and the result is proved. ∎

Proof of Corollary 2. The first step is to prove that

$$\frac{\sqrt{\hat{k}_n^\star}}{\log d_n}\left(\frac{\hat{q}_n(\alpha_n,k_n,\hat{k}_n^\star)}{q(\alpha_n)}-1\right) \xrightarrow{d} \mathcal{N}(0,\sigma^2). \tag{29}$$

To this end, recall that $\log(n\alpha_n)/\log(k_n) \to c' \le 0$ implies that $\log d_n \sim (1-c')\log k_n$ as $n\to\infty$ and therefore condition (11) of Corollary 1 is fulfilled under the assumptions of Corollary 2. Besides, recalling that $k_n/n \to 0$ as $n\to\infty$, Lemma 2 entails that $\hat{k}_n^\star \xrightarrow{\mathbb{P}} \infty$ and $\hat{k}_n^\star/n \xrightarrow{\mathbb{P}} 0$. Therefore, for n large enough, $\hat{k}_n^\star < n$ almost surely. Besides, for all $m_n \in \{1,\dots,n\}$,

$$\frac{\sqrt{\hat{k}_n^\star}}{\log d_n}\left(\frac{\hat{q}_n(\alpha_n,k_n,\hat{k}_n^\star)}{q(\alpha_n)}-1\right)|\{\hat{k}_n^\star=m_n\} \stackrel{d}{=} \frac{\sqrt{m_n}}{\log d_n}\left(\frac{\hat{q}_n(\alpha_n,k_n,m_n)}{q(\alpha_n)}-1\right).$$

By [43, Lemma 8], since $\hat{k}_n^\star \in \{1,\dots,n-1\}$ and $\hat{k}_n^\star \xrightarrow{\mathbb{P}} \infty$, it is enough to show that the desired convergence (29) holds with $\hat{q}_n(\alpha_n,k_n,\hat{k}_n^\star)$ replaced by its de-conditioned version $\hat{q}_n(\alpha_n,k_n,m_n)$: this is a direct consequence of Corollary 1. The second and final step consists in replacing $\hat{k}_n^\star$ by its non random version $k_n^\star$ in the rate of convergence of (29). This can be achieved using Lemma 2 and Slutsky's lemma. ∎

References

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

[2] Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications*, Wiley.

[3] 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.

[4] Bingham, N. H., Goldie, C. M., and Teugels, J. L. (1989). *Regular variation*, Cambridge university press.

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

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

[7] Cai, J. J., Einmahl, J. H., 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*, 417–442.

[8] Ciuperca, G., and Mercadier, C. (2010). Semi-parametric estimation for heavy tailed distributions. *Extremes*, 13, 55–87.

[9] Daouia, A., Gardes, L., Girard, S., and Lekina, A. (2011). Kernel estimators of extreme level curves, *Test*, 20(2), 311–333.

[10] Daouia, A., Gijbels, I., and Stupfler, G. (2019). Extremiles: A new perspective on asymmetric least squares. *Journal of the American Statistical Association*, 114(527), 1366–1381.

[11] Daouia, A., Girard, S., and Stupfler, G. (2018). Estimation of Tail Risk based on Extreme Expectiles, *Journal of the Royal Statistical Society: Series B*, 80, 262–292.

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

[13] Daouia, A., Girard, S., and Stupfler, G. (2020). Tail expectile process and risk assessment. *Bernoulli*, 26(1), 531–556.

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

[15] Deme, E., Gardes, L., and Girard, S. (2013). On the estimation of the second order parameter for heavy-tailed distributions, *Revstat*, 11, 277–299.

[16] Drees, H., de Haan, L., and Resnick, S. (2000). How to make a Hill plot, *Annals of Statistics*, 28, 254–274.

[17] El Methni, J., Gardes, L., Girard, S., and Guillou, A. (2012). Estimation of extreme quantiles from heavy and light tailed distributions, *Journal of Statistical Planning and Inference*, 142(10), 2735–2747.

[18] El Methni, J., and Stupfler, G. (2017). Extreme versions of Wang risk measures and their estimation for heavy-tailed distributions, *Statistica Sinica*, 27, 907–930.

[19] El Methni, J., and Stupfler, G. (2018). Improved estimators of extreme Wang distortion risk measures for very heavy-tailed distributions, *Econometrics and Statistics*, 6, 129–148.

[20] Fraga Alves, M. I., Gomes, M. I., and de Haan, L. (2003). A new class of semi-parametric estimators of the second order parameter. *Portugaliae Mathematica*, 60(2), 193–214.

[21] Fraga Alves, M. I., de Haan, L., and Lin, T. (2003). Estimation of the parameter controlling the speed of convergence in extreme value theory. *Mathematical Methods of Statistics*, 12(2), 155–176.

[22] Goegebeur, Y., Beirlant, J., and de Wet, T. (2010). Kernel estimators for the second order parameter in extreme value statistics. *Journal of Statistical Planning and Inference*, 140, 2632–2652.

[23] Goegebeur, Y., Guillou, A., and Schorgen, A. (2014). Nonparametric regression estimation of conditional tails: the random covariate case. *Statistics*, 48(4), 732–755.

[24] Gomes, M. I., Brilhante, M. F., Caeiro, F., and Pestana, D. (2015). A new partially reduced-bias mean-of-order p class of extreme value index estimators. *Computational Statistics and Data Analysis*, 82, 223–237.

[25] Gomes, M. I., Brilhante, M. F., and Pestana, D. (2016). New reduced-bias estimators of a positive extreme value index. *Communications in Statistics-Simulation and Computation*, 45(3), 833–862.

[26] Gomes, M. I., Caeiro, F., Figueiredo, F., Henriques-Rodrigues, L., and Pestana, D. (2020). Reduced-bias and partially reduced-bias mean-of-order-p value-at-risk estimation: a Monte-Carlo comparison and an application. *Journal of Statistical Computation and Simulation*, 90(10), 1735–1752.

[27] Gomes, M. I., Caeiro, F., Figueiredo, F., Henriques-Rodrigues, L., and Pestana, D. (2020). Corrected-Hill versus partially reduced-bias value-at-risk estimation. *Communications in Statistics-Simulation and Computation*, 49(4), 867–885.

[28] Gomes, M. I., de Haan, L., and Peng, L. (2002). Semi-parametric estimation of the second order parameter in statistics of extremes. *Extremes*, 5, 387–414.

[29] Gomes, M. I., and Pestana, D. (2007). A sturdy reduced-bias extreme quantile (VaR) estimator. *Journal of the American Statistical Association*, 102(477), 280–292.

[30] de Haan, L., and Ferreira, A. (2007). Extreme value theory: an introduction. Springer Science and Business Media.

[31] Hall, P. (1982). On some simple estimates of an exponent of regular variation. *Journal of the Royal Statistical Society: Series B*, 44(1), 37–42.

[32] Hall, P., and Welsh, A. W. (1985). Adaptive estimates of parameters of regular variation *The Annals of Statistics*, 13, 331–341.

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

[34] Kazama, S., Sato, A., and Kawagoe, S. (2009). Evaluating the cost of flood damage based on changes in extreme rainfall in Japan. *Sustainability Science*, 4(61), `https://doi.org/10.1007/s11625-008-0064-y`

[35] Kratz, M., and Resnick, S. I. (1996). The QQ-estimator and heavy tails. *Stochastic Models*, 12(4), 699–724.

[36] Manjunath, B. G., and Caeiro, F. (2013). `evt0`: Mean of order p, peaks over random threshold Hill and high quantile estimates. *R package version 1.1-3*.

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

[38] Peng, L. (1998). Asymptotic unbiased estimators for the extreme value index. *Statistics and Probability Letters*, 38, 107–115.

[39] Pickands, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics*, 3(1), 119–131.

[40] Resnick, S., and Stărică, C. (1997). Smoothing the Hill estimator, *Advances in Applied Probability*, 29, 271—293.

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

[42] Schultze, J., and Steinebach, J. (1996). On least squares estimates of an exponential tail coefficient. *Statistics and Risk Modeling*, 14(4), 353–372.

[43] Stupfler, G. (2019). On a relationship between randomly and non-randomly thresholded empirical average excesses for heavy tails. *Extremes*, 22, 749–769.

[44] 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.

[45] de Wet, T., Goegebeur, Y., and Munch, M. R. (2012). Asymptotically unbiased estimation of the second order tail parameter. *Statistics and Probability Letters*, 82, 565–573.

Acknowledgments

The authors would like to thank Frederico Caeiro for providing the R code associated with Algorithm 4.1 and Algorithm 4.2 from [26]. This work is supported by the French National Research Agency (ANR) in the framework of the Investissements d'Avenir Program (ANR-15-IDEX-02). S. Girard also 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.

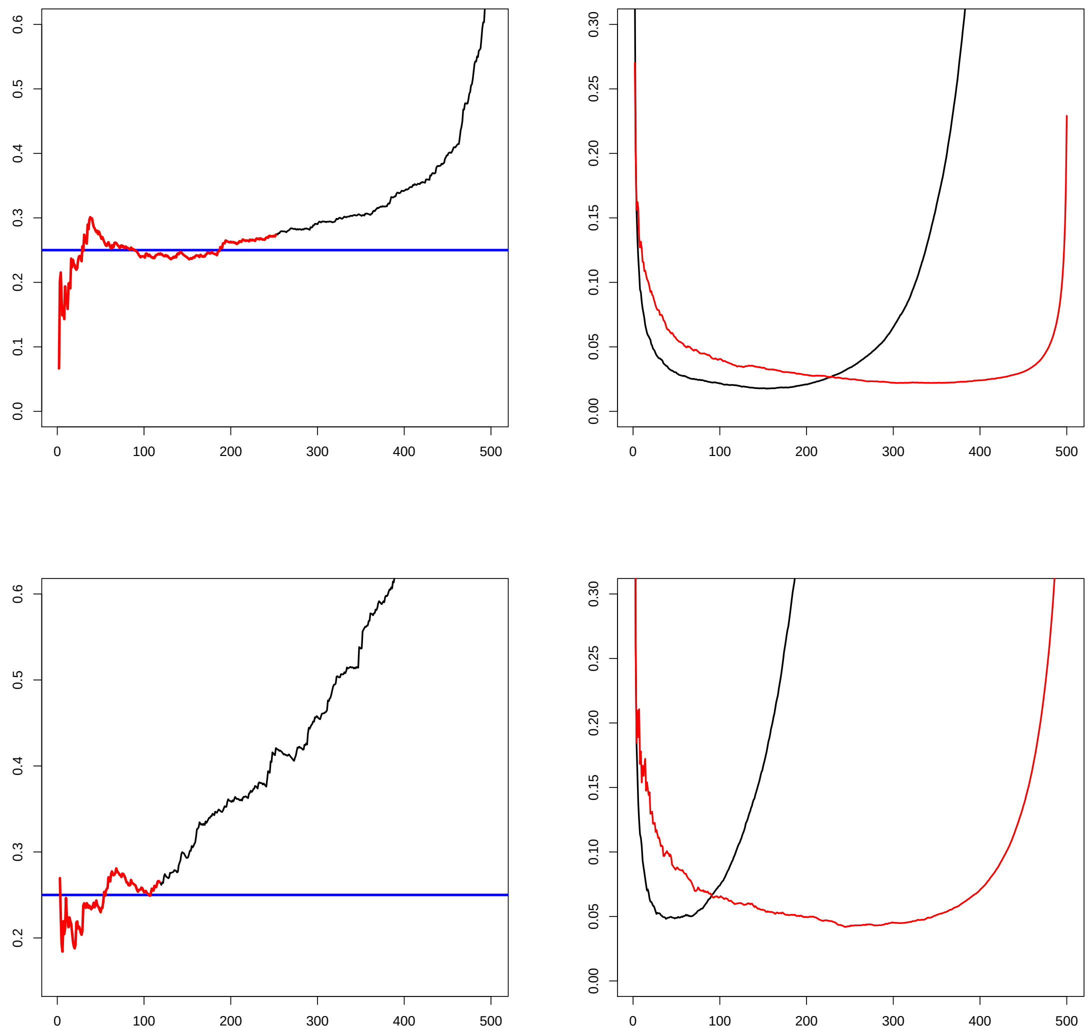

Figure 2: Illustration on simulated data sets of size $n = 500$ from a Burr distribution with $\gamma = 1/4$. Top: $\rho = -2$ and bottom: $\rho = -3/4$. Left panel: Hill estimators $\mathrm{H}(k_n)$ (black) and $\mathrm{H}(\hat{k}_n^{\mathrm{H},\star})$ (red) as functions of k_n. The true value of $\gamma = 1/4$ is depicted by a blue horizontal line. Right panel: RMSEs as functions of k_n computed on $N = 1000$ replications associated with Weissman estimator $\hat{q}_n(\alpha_n = 1/n, k_n)$ (black) and the refined version $\hat{q}_n(\alpha_n = 1/n, k_n, \hat{k}_n^{\mathrm{H},\star})$ (red).

Burr	Weissman	RW	CW	CH	CH_p	PRB_p	$CH_{p^\star}$	$PRB_{p^\star}$
$\gamma = 1/8$								
$\rho = -1/8$	0.1985	0.0205	0.1363	0.0163	0.0060	0.0077	0.0114	**0.0055**
$\rho = -1/4$	0.1827	0.0224	0.1322	0.0154	0.0058	0.1354	0.0108	**0.0055**
$\rho = -1/2$	0.0775	0.0100	0.0138	**0.0035**	0.0048	0.0091	0.0039	0.0066
$\rho = -1$	0.0216	**0.0044**	0.0067	0.0137	0.0118	0.0148	0.0114	0.0181
$\rho = -2$	0.0420	**0.0045**	0.0055	0.0111	0.0101	0.0160	0.0096	0.0116
$\gamma = 1/4$								
$\rho = -1/8$	-	**0.0389**	-	0.6875	0.0800	0.9894	0.4508	0.1355
$\rho = -1/4$	0.7886	**0.0131**	0.4584	0.0811	0.0240	0.5226	0.0536	0.0213
$\rho = -1/2$	0.2265	**0.0104**	0.0593	0.0164	0.0239	0.0498	0.0147	0.0295
$\rho = -1$	0.0612	**0.0158**	0.0287	0.0443	0.0469	0.0723	0.0392	0.0551
$\rho = -2$	0.0195	**0.0186**	0.0212	0.0359	0.0408	0.0531	0.0322	0.0360
$\gamma = 1/2$								
$\rho = -1/8$	-	0.7514	-	-	0.7702	0.9994	-	**0.5446**
$\rho = -1/4$	-	**0.1055**	-	0.2664	0.2233	0.9985	0.2551	0.1059
$\rho = -1/2$	0.9552	0.2893	0.1850	0.0537	0.1006	0.4483	**0.0513**	0.1386
$\rho = -1$	0.3199	**0.0625**	0.1267	0.1856	0.1691	0.2741	0.1520	0.2206
$\rho = -2$	0.1448	**0.0736**	0.0871	0.1387	0.1381	0.1968	0.1205	0.1385
$\gamma = 1$								
$\rho = -1/8$	-	**0.9883**	-	-	-	-	-	0.9994
$\rho = -1/4$	-	**0.9998**	-	-	-	-	-	-
$\rho = -1/2$	-	-	-	**0.2916**	0.9493	0.9986	0.3260	-
$\rho = -1$	-	**0.5451**	0.8959	0.9399	0.6693	-	0.8286	0.5930
$\rho = -2$	0.7843	0.5035	**0.3657**	0.4943	0.5485	0.5984	0.4098	0.6127

Table 3: RMSE associated with eight estimators of the extreme quantile $q(\alpha_n = 1/n)$ on a Burr distribution. The best result is emphasized in bold. RMSEs larger than 1 are not reported.

	Weissman	RW	CW	CH	CH_p	PRB_p	$\mathrm{CH}_{p^\star}$	$\mathrm{PRB}_{p^\star}$
Fréchet ($\rho = -1$)								
$\gamma = 1/8$	0.0088	0.0052	**0.0039**	0.0083	0.0097	0.0118	0.0072	0.0069
$\gamma = 1/4$	0.0400	**0.0165**	0.0206	0.0297	0.0341	0.0388	0.0270	0.0373
$\gamma = 1/2$	0.2583	**0.0699**	0.0831	0.1162	0.1450	0.2047	0.1030	0.1799
$\gamma = 1$	-	**0.4469**	-	-	0.8632	0.9088	0.4529	-
Fisher ($\rho = -\gamma$)								
$\gamma = 1/8$	0.2559	0.0624	0.7070	**0.0081**	0.0092	0.0745	0.0084	0.0109
$\gamma = 1/4$	0.5555	**0.0142**	0.1044	0.0170	0.0279	0.0694	0.0154	0.0410
$\gamma = 1/2$	0.9255	0.1234	0.1973	0.0984	0.2037	0.3109	**0.0871**	0.6890
$\gamma = 1$	-	**0.5410**	-	0.6597	0.7823	-	0.5993	-
GPD ($\rho = -\gamma$)								
$\gamma = 1/8$	0.6379	0.1075	0.4635	0.1465	**0.0163**	0.8355	0.1079	0.1357
$\gamma = 1/4$	0.7184	0.2021	0.3716	0.0715	**0.0288**	0.7292	0.0693	0.5801
$\gamma = 1/2$	-	**0.0276**	0.2770	0.0597	0.1060	0.4886	0.0627	0.9440
$\gamma = 1$	-	**0.4261**	-	0.5796	0.8320	0.9566	0.5130	-
Inverse Gamma ($\rho = -\gamma$)								
$\gamma = 1/8$	0.1768	0.0193	0.0293	0.0070	0.0097	0.0215	**0.0062**	0.0158
$\gamma = 1/4$	0.2417	0.0316	0.0325	0.0196	0.0342	0.0770	**0.0184**	0.0313
$\gamma = 1/2$	0.6804	**0.1020**	0.1263	0.1396	0.1543	0.4251	0.1155	0.1791
$\gamma = 1$	-	0.6597	0.8210	**0.4420**	-	0.9985	0.4513	0.8334
Student ($\rho = -2\gamma$)								
$\gamma = 1/8$	0.3029	0.0632	0.2603	0.0936	0.0111	0.0228	0.0455	**0.0100**
$\gamma = 1/4$	0.3747	0.0963	0.1737	0.0283	0.0195	0.0196	0.0196	**0.0150**
$\gamma = 1/2$	0.4355	0.2920	**0.0494**	0.1109	0.1347	0.2383	0.0945	0.4609
$\gamma = 1$	0.7909	**0.0906**	0.4674	0.5170	0.6569	0.8597	0.5224	0.8059

Table 4: RMSE associated with eight estimators of the extreme quantile $q(\alpha_n = 1/n)$ on five heavy-tailed distributions. The best result is emphasized in bold. RMSEs larger than 1 are not reported.

Burr	Weissman	RW	CW	CH	CH_p	PRB_p	$CH_{p^\star}$	$PRB_{p^\star}$
$\gamma = 1/8$								
$\rho = -1/8$	-	**0.0591**	1.1401	0.4347	0.0595	0.8671	0.2633	0.1493
$\rho = -1/4$	0.3292	0.0158	0.2110	0.0724	**0.0088**	0.1241	0.0454	0.0177
$\rho = -1/2$	0.0830	0.0172	0.0287	0.0450	0.0510	0.0094	0.0480	**0.0064**
$\rho = -1$	0.0301	0.0083	**0.0064**	0.0149	0.0140	0.0188	0.0127	0.0191
$\rho = -2$	0.0096	**0.0069**	0.0077	0.0135	0.0119	0.0201	0.0114	0.0120
$\gamma = 1/4$								
$\rho = -1/8$	-	**0.1515**	-	-	0.1582	0.9921	-	0.9848
$\rho = -1/4$	-	0.1022	0.7320	0.2385	**0.0365**	0.5502	0.1860	0.1791
$\rho = -1/2$	0.4938	0.0252	0.1031	**0.0155**	0.0270	0.0495	0.0175	0.0206
$\rho = -1$	0.1103	**0.0262**	0.0424	0.0563	0.0523	0.0802	0.0455	0.0661
$\rho = -2$	0.0275	**0.0201**	0.0270	0.0459	0.0457	0.0544	0.0406	0.0500
$\gamma = 1/2$								
$\rho = -1/8$	-	**0.8734**	-	-	0.9123	-	-	0.9996
$\rho = -1/4$	-	0.6167	-	0.8482	**0.3754**	0.9989	0.8038	0.9991
$\rho = -1/2$	-	0.1236	0.3795	**0.0622**	0.1188	0.4713	0.0666	0.9448
$\rho = -1$	0.6053	**0.1123**	0.1871	0.2122	0.1896	0.3101	0.1800	0.3293
$\rho = -2$	0.1849	**0.0975**	0.1057	0.1520	0.1619	0.1985	0.1395	0.1974
$\gamma = 1$								
$\rho = -1/8$	-	-	-	-	-	-	-	-
$\rho = -1/4$	-	-	-	-	-	-	-	-
$\rho = -1/2$	-	-	-	-	**0.9600**	-	-	-
$\rho = -1$	-	-	-	-	**0.8118**	-	-	-
$\rho = -2$	-	**0.5719**	-	-	0.5721	0.9563	-	-

Table 5: RMSE associated with eight estimators of the extreme quantile $q(\alpha_n = 1/(2n))$ on a Burr distribution. The best result is emphasized in bold. RMSEs larger than 1 are not reported.

	Weissman	RW	CW	CH	CH_p	PRB_p	$CH_{p^\star}$	$PRB_{p^\star}$
Fréchet ($\rho = -1$)								
$\gamma = 1/8$	0.0179	**0.0081**	0.0096	0.0095	0.0090	0.0157	0.0087	0.0099
$\gamma = 1/4$	0.0567	**0.0209**	0.0213	0.0343	0.0444	0.0529	0.0312	0.2805
$\gamma = 1/2$	0.3538	0.1081	**0.0886**	0.1347	0.1826	0.2294	0.1252	0.2125
$\gamma = 1$	0.9850	**0.9460**	0.9864	0.9877	-	-	0.9870	-
Fisher ($\rho = -\gamma$)								
$\gamma = 1/8$	0.4576	**0.0101**	0.1336	0.0203	0.0105	0.0631	0.0206	0.0146
$\gamma = 1/4$	0.9311	0.0291	0.1901	**0.0232**	0.0301	0.0695	0.0273	0.0411
$\gamma = 1/2$	-	0.2334	0.2598	**0.0944**	0.3213	0.4579	0.0986	0.8496
$\gamma = 1$	-	**0.9916**	-	-	-	-	-	-
GPD ($\rho = -\gamma$)								
$\gamma = 1/8$	0.9387	0.0809	0.7988	0.3519	**0.0328**	0.8577	0.8301	0.3659
$\gamma = 1/4$	-	0.2491	0.6321	0.2094	**0.0383**	0.7582	0.1990	0.7896
$\gamma = 1/2$	-	**0.0325**	0.5436	0.0824	0.1202	0.5043	0.0948	0.9636
$\gamma = 1$	-	-	-	-	**0.8118**	-	-	-
Inverse Gamma ($\rho = -\gamma$)								
$\gamma = 1/8$	0.1821	0.0130	0.0404	0.0190	0.0156	0.0750	**0.0093**	0.0449
$\gamma = 1/4$	0.4147	**0.0215**	0.0630	0.0254	0.0367	0.0435	0.0226	0.0394
$\gamma = 1/2$	0.8202	0.1955	**0.1214**	0.1395	0.3718	0.5781	0.1214	0.2517
$\gamma = 1$	-	-	-	-	-	-	**0.9215**	-
Student ($\rho = -2\gamma$)								
$\gamma = 1/8$	0.5643	**0.0103**	0.4662	0.2148	0.0180	0.3625	0.1251	0.0373
$\gamma = 1/4$	0.6091	**0.0135**	0.2795	0.0752	0.0255	0.0560	0.0502	0.0345
$\gamma = 1/2$	0.6488	0.2739	**0.0777**	0.1120	0.1684	0.2685	0.0983	0.7009
$\gamma = 1$	-	**0.5982**	-	0.6812	0.8226	-	0.7982	-

Table 6: RMSE associated with eight estimators of the extreme quantile $q(\alpha_n = 1/(2n))$ on five heavy-tailed distributions. The best result is emphasized in bold. RMSEs larger than 1 are not reported.

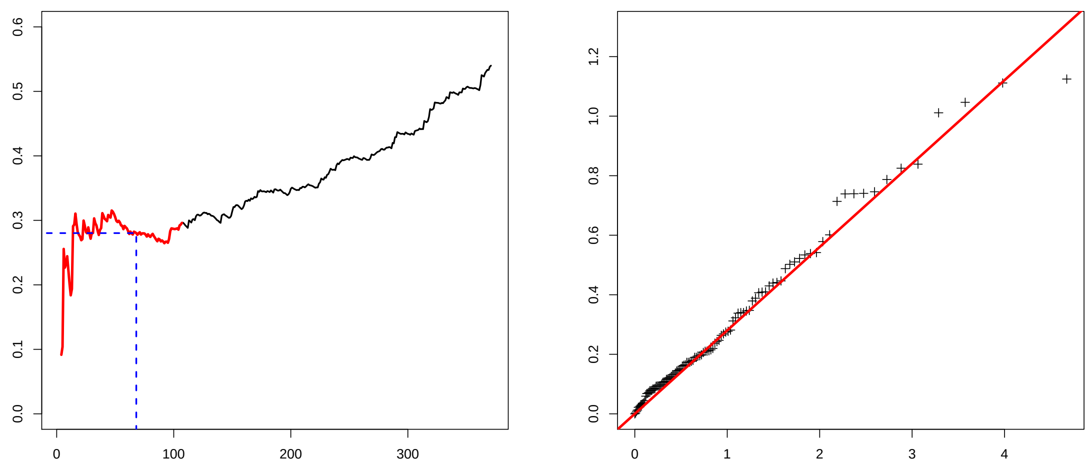

Figure 3: Illustration on the Secura Belgian reinsurance actuarial data set. Left panel: Hill estimators $\mathrm{H}(k_n)$ (black) and $\mathrm{H}(\hat{k}_n^{\mathrm{H},\star})$ (red) as functions of k_n. The pair $(\hat{k}_n^{\mathrm{H},\star}, \mathrm{H}(\hat{k}_n^{\mathrm{H},\star}))$ associated with the value of k_n selected by the sample path stability criterion is emphasized in blue. Right panel: quantile-quantile plot (horizontally: $\log(\hat{k}_n^{\mathrm{H},\star}/i)$, vertically: $\log(X_{n-i+1,n}) - \log(X_{n-\hat{k}_n^{\mathrm{H},\star},n})$ for $i = 1, \ldots, \hat{k}_n^{\mathrm{H},\star}$). The regression line with the estimated value of γ as slope is superimposed in red.