



HAL
open science

A refined extreme quantiles estimator for Weibull tail-distributions

Jonathan El Methni, Stéphane Girard

► **To cite this version:**

Jonathan El Methni, Stéphane Girard. A refined extreme quantiles estimator for Weibull tail-distributions. 2023. hal-04022982v1

HAL Id: hal-04022982

<https://hal.science/hal-04022982v1>

Preprint submitted on 10 Mar 2023 (v1), last revised 10 Jan 2024 (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 extreme quantiles estimator for Weibull tail-distributions

Jonathan El Methni⁽¹⁾ and Stéphane Girard⁽²⁾

⁽¹⁾ Université Paris Cité, CNRS, MAP5, UMR 8145, 75006 Paris, France

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

Abstract

We address the problem of estimation extreme quantiles of Weibull tail-distributions. Since such quantiles are asymptotically larger than the sample maxima, their estimation requires extrapolation methods. In the case of Weibull tail-distributions, classical extreme-value estimators are numerically outperformed by estimators dedicated to this set of light-tailed distributions. The latter estimators of extreme quantiles are based on two estimators: an order statistic to estimate an intermediate quantile and an estimator of the Weibull tail-coefficient. The common practice is to select the same intermediate sequence for both estimators. We show how an adapted choice of two different intermediate sequences leads to a reduction of the asymptotic bias associated with the resulting refined estimator. The asymptotic normality of the refined estimator is established, and a data-driven method is introduced for the practical selection of the intermediate sequences. Our approach is compared to three estimators of extreme quantiles dedicated to Weibull tail-distributions in a simulation study.

Key words: extreme quantile, bias reduction, Weibull tail-distribution, extreme-value statistics, asymptotic normality.

Mathematics Subject Classification: 60G70, 62G32, 62G20.

Acknowledgments: 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.

The code is available from the corresponding author upon request.

Declaration - Conflict of interest: The authors declare that they have no conflict of interest.

1 Introduction

Let X_1, X_2, \dots, X_n be independent and identically distributed random variables with cumulative distribution function F and let $X_{1,n} \leq \dots \leq X_{n,n}$ denote the associated order statistics. We consider the case where F belongs to the family of Weibull tail-distributions [10]:

(A.1) F is twice differentiable and $F(\cdot) = 1 - \exp(-H(\cdot))$ with $V(t) := H^\leftarrow(t) = t^\theta \ell(t)$ for all $t > 0$, where $\theta > 0$ is called the Weibull tail-coefficient and where ℓ is a (positive) slowly-varying function *i.e.* $\ell(tx)/\ell(x) \rightarrow 1$ as $x \rightarrow \infty$ for all $t > 0$.

Here, and in the following, $f^\leftarrow(\cdot) = \inf\{x \in \mathbb{R}, f(x) > \cdot\}$ denotes the generalized inverse of the function f . The inverse cumulative hazard function V is said to be regularly-varying at infinity with index θ and this property is denoted by $V \in \mathcal{RV}_\theta$, see [9] for a detailed account on this topic. The shape parameter θ is referred to as the Weibull tail-coefficient. Weibull tail-distributions include for instance exponential ($\theta = 1$), Gamma ($\theta = 1$), logistic ($\theta = 1$), Gaussian ($\theta = 1/2$) and Weibull distributions (θ is the inverse of the shape parameter), see [18, Table 1]. Weibull tail-distributions are included in the Gumbel maximum domain of attraction, *i.e.* with extreme-value index $\gamma = 0$, see [17, Proposition 2(ii)], and as such, are light-tailed distributions. We refer to [7] for an application to the modeling of large claims in non-life insurance and to [25] for an analysis of neural networks distributional properties.

Dedicated methods have been proposed to estimate the Weibull tail-coefficient θ since the relevant information is localised in the extreme upper part of the sample. Most approaches rely on the k_n upper order statistics $X_{n-k_n+1,n}, \dots, X_{n,n}$ where $k_n \rightarrow \infty$ as $n \rightarrow \infty$. Note that, since θ is defined through a tail behavior, the associated estimator should only use the extreme-values of the sample and thus the extra condition $k_n/n \rightarrow 0$ is required. More specifically, recent estimators are based on the log-spacings between the k_n upper order statistics [8, 15, 17, 18] or on the mean excess function [4, 5, 6]. The introduction of kernel based weights has been investigated for both approaches, see [16, 19] for the log-spacings case and [20] for the mean excess function framework. A bias reduction method adapted to the estimation of the Weibull tail-coefficient is proposed in [12] and the adaptation to random censoring is achieved in [26].

We address the problem of estimating extreme quantiles of Weibull tail-distributions. Recall that an extreme quantile $q(\alpha_n)$ of order α_n is defined by the equation $1 - F(q(\alpha_n)) = \alpha_n$ with $n\alpha_n \rightarrow 0$ as $n \rightarrow \infty$. The latter condition implies that $q(\alpha_n)$ is almost surely asymptotically larger than $X_{n,n}$, the sample maximum. It is shown in [14] that classical extreme-value estimators of such large quantiles are numerically outperformed by estimators dedicated to Weibull tail-distributions [13], see also Lemma 2 in the Appendix for a theoretical argument. The latter methods estimate $q(\alpha_n)$ by combining two ingredients: an order statistic $X_{n-k_n+1,n}$ and an estimator of the Weibull tail-coefficient θ .

In this work, we show that the biases associated with the previous extrapolation method and the estimator of θ may asymptotically cancel out in the extreme quantile estimator thanks to an appropriate tuning of the number of upper order statistics involved in the Weibull tail-coefficient estimator. The construction of the resulting estimator is presented in Section 2 and an asymptotic normality result is provided, emphasizing that the proposed extreme quantile estimator is asymptotically less biased than the original one [14]. Its performances are illustrated on simulated data

in Section 3 and compared to state-of-the-art competitors [6, 13, 14]. Proofs are postponed to the Appendix.

2 A refined estimator of the extreme quantile

2.1 Extreme quantile estimators

Weibull-tail estimators of the extreme quantile $q(\alpha_n)$ rely on an intermediate quantile $q(k_n/n)$ where (k_n) is an intermediate sequence of integers *i.e.* such that $k_n \in \{1, \dots, n-1\}$, $k_n \rightarrow \infty$ and $k_n/n \rightarrow 0$ as $n \rightarrow \infty$, see *i.e.* [13, 14]. Indeed, in view of (A.1), one has

$$\frac{q(\alpha_n)}{q(k_n/n)} = \frac{V(\log(1/\alpha_n))}{V(\log(n/k_n))} \simeq \left(\frac{\log(1/\alpha_n)}{\log(n/k_n)} \right)^\theta =: \tau_n^\theta,$$

as $n \rightarrow \infty$, where $\tau_n = \log(1/\alpha_n)/\log(n/k_n)$ is the (logarithmic) extrapolation factor. Following the ideas of [3], we propose a refined Weissman type estimator defined by:

$$\hat{q}_n(\alpha_n, k_n, k'_n) = X_{n-k_n+1, n} \left(\frac{\log(1/\alpha_n)}{\log(n/k_n)} \right)^{\hat{\theta}_n(k'_n)} = X_{n-k_n+1, n} \tau_n^{\hat{\theta}_n(k'_n)}, \quad (1)$$

with $\hat{\theta}_n(k'_n)$ an estimator of the Weibull tail-coefficient θ depending on another intermediate sequence (k'_n) . We focus on the classical estimator of θ introduced in [15]:

$$\hat{\theta}_n^{\text{RSH}}(k'_n) = \frac{1}{\mu(\log(n/k'_n))} \frac{1}{k'_n} \sum_{i=1}^{k'_n} (\log(X_{n-i+1, n}) - \log(X_{n-k'_n+1, n})), \quad (2)$$

with, for $t > 0$,

$$\mu(t) = \int_0^{+\infty} \log\left(1 + \frac{x}{t}\right) e^{-x} dx = e^t E_1(t),$$

where E_1 is the Exponential integral function [2, Page 228]. Clearly, $\hat{\theta}_n^{\text{RSH}}$ can be interpreted as a rescaled Hill estimator since

$$\hat{\theta}_n^{\text{RSH}}(k'_n) = \frac{\hat{\gamma}_n^{\text{H}}(k'_n)}{\mu(\log(n/k'_n))},$$

where $\hat{\gamma}_n^{\text{H}}(k'_n)$ is the well-known Hill estimator [24] of the extreme-value index $\gamma > 0$.

Let us note that $k'_n = k_n$ yields back the extreme quantile estimator for Weibull tail-distributions introduced in [14]. In the next paragraph, the asymptotic normality of $\hat{q}_n(\alpha_n, k_n, k'_n)$ is established, and we show that choosing $k'_n \neq k_n$ can yield better results from an asymptotic point of view. A similar phenomenon occurs in the estimation of the endpoint of a distribution, see [1] for details.

2.2 Asymptotic analysis

The study of the limit distribution of $\hat{q}_n(\alpha_n, k_n, k'_n)$ requires a second-order condition on ℓ :

(A.2) There exist $\rho \leq 0$ and $b(t) \rightarrow 0$ as $t \rightarrow \infty$, with ultimately constant sign, such that uniformly locally on $x \geq 1$,

$$\lim_{t \rightarrow \infty} \frac{1}{b(t)} \log\left(\frac{\ell(tx)}{\ell(t)}\right) = K_\rho(x) := \int_1^x u^{\rho-1} du.$$

It can be shown that necessarily $|b| \in \mathcal{RV}_\rho$. The second-order Weibull parameter $\rho \leq 0$ tunes the rate of convergence of the ratio $\ell(tx)/\ell(t)$ to 1. The closer ρ is to 0, the slower is the convergence. Condition **(A.2)** is the cornerstone in all proofs of asymptotic normality for extreme-value estimators. Again, we refer to [18, Table 1] for ρ parameters associated with usual Weibull tail-distributions.

Our first result is a refinement of [15, Corollary 3.1]. It provides an asymptotic normality result for the extreme quantile estimator (1) based on two intermediate sequences (k_n) and (k'_n) .

Theorem 1. *Assume **(A.1)** and **(A.2)** hold. Let (k_n) and (k'_n) be two intermediate sequences such that $k'_n \leq k_n$ and introduce (α_n) a probability sequence such that $\alpha_n \rightarrow 0$ as $n \rightarrow \infty$. Suppose, as $n \rightarrow \infty$,*

- (i) $\sqrt{k'_n} b(\log(n/k'_n)) \rightarrow \lambda \in \mathbb{R}$,
- (ii) $\log(n/k'_n)/\log(n/k_n) \rightarrow \beta \geq 1$,
- (iii) $\tau_n \rightarrow \tau > \beta$.

Then, as $n \rightarrow \infty$,

$$\sqrt{k'_n} \left(\frac{\hat{q}_n(\alpha_n, k_n, k'_n)}{q(\alpha_n)} - 1 \right) \xrightarrow{d} \mathcal{N}(\lambda(\log(\tau) - \beta^{-\rho} K_\rho(\tau)), (\theta \log \tau)^2). \quad (3)$$

Let us first remark that condition (i) implies that $\log(n/k'_n) \sim \log(n)$ as $n \rightarrow \infty$ (see [15, Lemma 5.1]), then condition (ii) yields that $\log(n/k_n) \sim \log(n)/\beta$ and therefore condition (iii) can be rewritten as $\log(1/\alpha_n) \sim (\tau/\beta) \log(n)$ as $n \rightarrow \infty$. As a consequence, the condition $\tau > \beta$ in (iii) implies $n\alpha_n \rightarrow 0$ as $n \rightarrow \infty$ which, in turns, implies that $q(\alpha_n)$ is an extreme quantile.

It follows from (3) that the asymptotic bias associated with $\hat{q}_n(\alpha_n, k_n, k'_n)$ is given by

$$(\log \tau - \beta^{-\rho} K_\rho(\tau)) b(\log(n/k'_n)) \sim (\beta^\rho \log(\tau) - K_\rho(\tau)) b(\log(n/k_n)) =: B(\beta, \tau, \rho) b(\log(n/k_n)).$$

It thus appears that each choice of k'_n yields an associated constant β and thus a corresponding bias factor $B(\beta, \tau, \rho)$. From the theoretical point of view, two cases can be considered.

- The usual choice $k'_n = k_n$ yields $\beta = 1$ and

$$B(1, \tau, \rho) = \log(\tau) - K_\rho(\tau), \quad (4)$$

which is the bias factor associated with the extreme quantile estimator $\hat{q}_n(\alpha_n, k_n, k_n)$ investigated in [14].

- The choice $\beta^*(\tau, \rho) = (K_\rho(\tau)/\log(\tau))^{1/\rho}$ yields

$$B(\beta^*(\tau, \rho), \tau, \rho) = 0. \quad (5)$$

The associated intermediate sequence is given by $k_n^*(\tau, \rho) = \lfloor n(k_n/n)^{\beta^*(\tau, \rho)} \rfloor$ and therefore the extreme quantile estimator $\hat{q}_n(\alpha_n, k_n, k_n^*(\tau, \rho))$ is asymptotically unbiased. Note that this estimator cannot be used in practice since the second-order Weibull parameter ρ is unknown.

Up to our knowledge, there is no estimator of the second-order Weibull parameter in the statistical literature. In practice, one can replace ρ by a misspecified value $y < 0$ in the above $\beta^*(\tau, \rho)$ leading to

$$\beta^*(\tau, y) = (K_y(\tau)/\log(\tau))^{1/y}, \quad (6)$$

$$k_n^*(\tau, y) = \lfloor n(k_n/n)^{\beta^*(\tau, y)} \rfloor, \quad (7)$$

$$B(\beta^*(\tau, y), \tau, \rho) = (K_y(\tau)/\log(\tau))^{\rho/y} \log(\tau) - K_\rho(\tau). \quad (8)$$

This misspecification technique has been used both to deal with Pareto-type distributions ($\gamma > 0$), see for instance [11], and Weibull tail-distributions ($\gamma = 0$) [13]. Some properties of the intermediate sequence $k_n^*(\tau, y)$ are given in the next Lemma.

Lemma 1. *Let $\beta^*(\tau, y)$ and $k_n^*(\tau, y)$ be defined by (6) and (7) respectively. Then, for all $\tau > 1$:*

- (i) $\beta^*(\tau, \cdot)$ can be extended by continuity by setting $\beta^*(\tau, 0) := \sqrt{\tau}$.
- (ii) $1 < \beta^*(\tau, y) < \tau$ for all $y \leq 0$.
- (iii) For all $y \leq 0$, $k_n^*(\tau, y)$ is an increasing function of k_n , $k_n^*(\tau, y) \leq k_n$ and $k_n^*(\tau, y)/k_n \rightarrow 0$ as $n \rightarrow \infty$.

In particular, it appears in (iii) that the number of upper order statistics $k_n^*(\tau, y)$ used in the Weibull tail-coefficient estimator should be asymptotically small compared to k_n . The next Corollary shows that these choices indeed lead to a bias reduction in the estimation of the extreme quantile.

Corollary 1. *Assume (A.1) and (A.2) hold. Let $c > 0$, $\tau > 1$, $y < 0$, $\lambda \neq 0$ such that $\lambda b(\cdot)$ is ultimately positive, and $\beta^*(\tau, y)$ be defined as in (6). Let $\alpha_n = cn^{-\tau/\beta^*(\tau, y)}$, $k_n = \lfloor n\{\lambda^2/(nb^2(\log n))\}^{1/\beta^*(\tau, y)} \rfloor$ and define $k_n^*(\tau, y)$ as in (7).*

(i) *Then, as $n \rightarrow \infty$,*

$$\sqrt{k_n^*(\tau, y)} \left(\frac{\hat{q}_n(\alpha_n, k_n, k_n^*(\tau, y))}{q(\alpha_n)} - 1 \right) \xrightarrow{d} \mathcal{N} \left(\lambda(\log(\tau) - (K_y(\tau)/\log(\tau))^{-\rho/y} K_\rho(\tau)), (\theta \log \tau)^2 \right). \quad (9)$$

(ii) *Moreover, for all $\tau > 1$, $\rho \leq 0$ and $y \leq 0$, $|B(\beta^*(\tau, y), \tau, \rho)| \leq |B(1, \tau, \rho)|$.*

Let us first highlight that $\sqrt{k_n^*} \sim \lambda/b(\log n)$ as $n \rightarrow \infty$ (see the proof of Corollary 1 in the Appendix) which is the rate of convergence of usual extreme quantile estimators dedicated to Weibull tail-distributions, see for instance [13, Theorem 1]. Surprisingly, as a consequence of Corollary 1(ii), the extreme quantile estimator $\hat{q}_n(\alpha_n, k_n, k_n^*(\tau, y))$ computed with $k_n^*(\tau, y) = \lfloor n(k_n/n)^{\beta^*(\tau, y)} \rfloor$ and $\beta^*(\tau, y) = (K_y(\tau)/\log(\tau))^{1/y}$ has a smaller asymptotic bias than the usual one $\hat{q}_n(\alpha_n, k_n, k_n)$ whatever the chosen value $y \leq 0$. Let us recall that, from (5), the theoretical best choice would be $y = \rho$. In practice, we use $y = \rho^\# = -1$ leading to $\beta^*(\tau, -1) = \tau \log(\tau)/(\tau - 1)$. This "canonical" choice is also used in [13], see Section 3.2 hereafter. The performance of the associated estimator is assessed on simulated data in the next Section.

3 Validation on simulated data

Our estimator is compared on simulated data to the original estimator and to three other bias-reduced estimators of the extreme quantile.

3.1 Experimental design

We focus on the class of $\mathcal{D}(\alpha, \beta, a)$ -distributions which is an adaptation of Hall's class ([22, 23]) to the Weibull-tail framework. In this family of distributions, the inverse cumulative hazard function is defined for all $x > 0$ by

$$V(x) := x^{1/\alpha} \left(1 + \frac{a}{\beta} x^{-\beta} \right),$$

with $a, \alpha, \beta > 0$ and $\alpha\beta \leq 1$. Under these conditions, the above class of distributions fulfills assumptions (A.1) and (A.2) with Weibull tail-coefficient $\theta = 1/\alpha$, second-order parameter $\rho = -\beta$, slowly-varying function $\ell(x) = 1 + (a/\beta)x^{-\beta}$ and $b(x) = -ax^{-\beta}$. Unlike classical distributions such as exponential ($\theta = 1, \rho = -\infty$), Gamma ($\theta = 1, \rho = -1$), Gaussian ($\theta = 1/2, \rho = -1$) and Weibull distributions ($\rho = -\infty$), it is thus possible to obtain \mathcal{D} -distributions with arbitrary Weibull tail-coefficient $\theta > 0$ and second-order parameter $\rho \in (0, -\theta]$.

In the following, we set $a = 10$, $\theta \in \{1/2, 3/4, \dots, 5/2\}$, $\rho \in \{-5/2, -2, \dots, -1/2\}$ and we only consider the 25 situations where $\rho \leq -\theta$ to fulfill the constraint $\alpha\beta \leq 1$, see Table 1. In each case, we simulate $N = 1,000$ replications of a data set of $n = 500$ i.i.d. realisations from the 25 considered parametric models. Finally, same two cases as in [13] are investigated for the order of the extreme quantile: $\alpha_n \in \{1/n^2, 1/n^4\}$. Summarizing, this experimental design includes $25 \times 2 = 50$ configurations.

(ρ, θ)	1/2	3/4	1	5/4	3/2	7/4	2	9/4	5/2
-5/2									yes
-2							yes	yes	yes
-3/2					yes	yes	yes	yes	yes
-1			yes	yes	yes	yes	yes	yes	yes
-1/2	yes	yes	yes	yes	yes	yes	yes	yes	yes

Table 1: Considered (ρ, θ) configurations in the $\mathcal{D}(\alpha = 1/\theta, \beta = -\rho, a = 10)$ -distribution.

3.2 Competitors

Our refined estimator dedicated to the estimation of extreme quantiles for Weibull tail-distributions is compared to three competitors. All three estimators share the same structure and rely on three terms, *i.e.* an order statistic, an extrapolation factor and an estimator of the Weibull tail-coefficient:

$$\hat{q}_n(\alpha_n, k_n) = X_{n-k_n+1, n} \tau_n^{\hat{\theta}_n(k_n)}. \quad (10)$$

Since all considered estimators depend on an estimator of the Weibull tail-coefficient, we will start by describing the latter.

1. We first consider the classical estimator of the Weibull tail-coefficient given in (2) and introduced in [15]. The authors constructed in [14] an estimator of the extreme quantile by plugging $\hat{\theta}_n^{\text{RSH}}$ in (10) to obtain

$$\hat{q}_n^{\text{RSH}}(\alpha_n, k_n) = X_{n-k_n+1, n} \tau_n^{\hat{\theta}_n^{\text{RSH}}(k_n)}. \quad (11)$$

The resulting estimator does not include a bias reduction.

2. In [6] the authors proposed an estimator of the Weibull tail-coefficient based on the mean excess function defined by $t \mapsto m(t) = \mathbb{E}(X - t \mid X > t)$. In practice, they consider $t = X_{n-j,n}$ and estimate $m(X_{n-j,n})$ by

$$\hat{m}_n(X_{n-j,n}) = \frac{1}{j} \sum_{i=1}^j X_{n-i+1,n} - X_{n-j,n},$$

for $j \in \{1, \dots, k_n\}$, which leads to the following estimator of θ based on log-spacings between the mean excesses:

$$\hat{\theta}_n^{\text{MEF}}(k_n) = \left(1 - \frac{\frac{1}{k_n} \sum_{j=1}^{k_n} \log(\hat{m}_n(X_{n-j,n})) - \log(\hat{m}_n(X_{n-k_n-1,n}))}{\frac{1}{k_n} \sum_{j=1}^{k_n} \log(X_{n-j,n}) - \log(X_{n-k_n-1,n})} \right)^{-1}.$$

We can then build an estimator based on the mean excess function and dedicated to the extreme quantiles by replacing $\hat{\theta}_n^{\text{MEF}}$ in (10), which yields

$$\hat{q}_n^{\text{MEF}}(\alpha_n, k_n) = X_{n-k_n+1,n} \tau_n^{\hat{\theta}_n^{\text{MEF}}(k_n)}. \quad (12)$$

3. Up to our knowledge there exists only one bias-reduced extreme quantile estimator [13] dedicated to Weibull tail-distributions. This estimator is based on a bias-reduced estimator of the Weibull tail-coefficient proposed by the same authors in [12]. They estimated the Weibull tail-coefficient θ through a least-squares estimator given by

$$\hat{\theta}_n^{\text{LSE}}(k_n) = \bar{Y}_{k_n} - \hat{b}(\log(n/k_n)) \bar{x}_{k_n},$$

where

$$\bar{Y}_{k_n} = \frac{1}{k_n} \sum_{j=1}^{k_n} Y_j \quad \text{and} \quad Y_j = j \log(n/j) (\log(X_{n-j+1,n}) - \log(X_{n-j,n})),$$

and with

$$\hat{b}(\log(n/k_n)) = \frac{\sum_{j=1}^{k_n} (x_j - \bar{x}_{k_n}) Y_j}{\sum_{j=1}^{k_n} (x_j - \bar{x}_{k_n})^2},$$

where

$$\bar{x}_{k_n} = \frac{1}{k_n} \sum_{j=1}^{k_n} x_j \quad \text{and} \quad x_j = \log(n/k_n) / \log(n/j).$$

Finally, they obtain

$$\hat{q}_n^{\text{LSE}}(\alpha_n, k_n) = X_{n-k_n+1,n} \tau_n^{\hat{\theta}_n^{\text{LSE}}(k_n)} \exp\left(\hat{b}(\log(n/k_n)) K_{\hat{\rho}_n}(\tau_n)\right). \quad (13)$$

From a practical point of view, the authors suggested to choose $\hat{\rho}_n = \rho^\# = -1$. This estimator features two bias corrections: a first one in the estimator of the Weibull tail-coefficient and a second one in the extrapolation factor. This estimator is built under the assumption $x|b(x)| \rightarrow \infty$ as $x \rightarrow \infty$ which leads to the constraint $\rho \geq -1$. The latter assumption is fulfilled by the class of $\mathcal{D}(\alpha, \beta, a)$ -distributions when $\beta \leq 1$.

Recall that our estimator is given by

$$\hat{q}_n^{\text{RWT}}(\alpha_n, k_n, k_n^*(\tau_n, -1)) = X_{n-k_n+1,n} \tau_n^{\hat{\theta}_n^{\text{RSH}}(k_n^*(\tau_n, -1))}, \quad (14)$$

where $k_n^*(\tau_n, -1) = \lfloor n(k_n/n)^{\beta^*(\tau_n, -1)} \rfloor$ and $\beta^*(\tau_n, -1) = \tau_n \log(\tau_n) / (\tau_n - 1)$. For the sake of simplicity, the above extreme quantile estimators (11)–(14) are respectively referred to as RSH, MEF, LSE and RWT in the sequel.

3.3 Selection of the thresholds

All considered extreme quantile estimators (RWT, RSH, LSE, MEF) 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 new algorithm for the selection of k_n is proposed in [3], basing on a bisection method inspired from random forests. The objective is to find the region with the smallest variance in a given series $\{Z_1, \dots, Z_m\}$. The proposed method starts by randomly splitting the series into two, compute the variance in each sub-region and repeat the action in the one with smallest variance until getting a final single point, see [3, Algorithm 2]. The above procedure is embedded in a bootstrap technique, see [3, Algorithm 1], and the final k_n^\dagger is selected by taking the median across the $T = 10,000$ bootstrap samples. In the simulations, we used $Z_j = \hat{q}(\alpha_n, k_{j,n})$ an estimator of the extreme quantile at level α_n computed with the intermediate sequence $k_{j,n} \in \{15, \dots, 3n/4\}$.

3.4 Results

The performance of the four extreme quantile estimators is assessed using the Mean absolute relative error:

$$\text{MARE}(\hat{q}_n(\alpha_n)) = \frac{1}{N} \sum_{i=1}^N \left| \frac{\hat{q}_n^{(i)}(\alpha_n, k_n^{(i,\dagger)})}{q_n(\alpha_n)} - 1 \right|, \quad (15)$$

where $\hat{q}_n^{(i)}(\alpha_n, k_n^{(i,\dagger)})$ denotes the estimator computed on the i th replication, $i \in \{1, \dots, N\}$ with the selected intermediate sequence $k_n^{(i,\dagger)}$. The results are provided in Table 2 and Table 3 for respectively, $\alpha_n = 1/n^2$ and $\alpha_n = 1/n^4$. We start by remarking that, as expected, the smaller the order α_n of the extreme quantile is, *i.e.* the more we extrapolate, the larger the error is. This is true for the four considered estimators.

The proposed RWT estimator is the most accurate one since it provides the best results in 44% of cases, especially when $\rho \leq -1$, except for the case $\theta = -\rho = 1/2$. Let us remark that, since we fixed $\rho^\# = -1$ in our estimator, it performs globally better, for a given value of the Weibull tail-coefficient θ , when ρ is close to -1 . Surprisingly, the LSE estimator, which also uses the same fixed value $\rho^\# = -1$, behaves differently: It performs best when ρ is large. The second most accurate estimator is RSH which provides the best results in 28% of the considered cases. It is only efficient when $\rho = -1/2$ where it obtains all its 14 best results. Unexpectedly, RSH performs well in this difficult case despite the fact that it does not benefit from a bias reduction. The four cases where it fails to obtain the best results when $\rho = -1/2$ correspond to a Weibull tail-coefficient smaller than 1. The LSE estimator obtains similar results with 22% of best results (11 out of 50 situations). As expected, the constraint $\rho \geq -1$, inherent in the design of LSE, corresponds to 9 out of its 11 cases where it obtains the best results. Finally, MEF yields poor estimations, with only 6% of best results. It does not seem to adapt well to the presence of bias in the distribution. In particular, when θ is close to 1 it does not give acceptable results, in the sense that $\text{MARE} \geq 1$.

To conclude, it appears on these experiments on simulated data that, in average, the RWT estimator performs the best within the four considered estimators. One of its main competitors is LSE, which, similarly to RWT, considers the two sources of bias (associated with the Weibull tail coefficient estimator and the extrapolation term).

References

- [1] Aarssen, K., and de Haan, L. (1994). On the maximal life span of humans. *Mathematical Population Studies*, 4(4), 259–281.
- [2] Abramowitz, M., and Stegun, I. A. (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables* (Vol. 55), US Government printing office.
- [3] Allouche, M., El Methni, J., and Girard, S. (2022). A refined Weissman estimator for extreme quantiles, *Extremes*, <https://doi.org/10.1007/s10687-022-00452-8>.
- [4] Beirlant, J., Bouquiaux, C., and Werker, B. (2006). Semiparametric lower bounds for tail index estimation, *Journal of Statistical Planning and Inference*, 136, 705–729.
- [5] Beirlant, J., Broniatowski, M., Teugels, J. L., and Vynckier, P. (1995). The mean residual life function at great age: Applications to tail estimation. *Journal of Statistical Planning and Inference*, 45, 21–48.
- [6] Beirlant, J., Dierckx, G., Guillou, A., and de Waal, D. (2009). A new estimation method for Weibull-type tails based on the mean excess function, *Journal of Statistical Planning and Inference*, 139(6), 1905–1920.
- [7] Beirlant, J., Teugels, J. L. (1992). Modeling large claims in non-life insurance. *Insurance: Mathematics and Economics*, 11, 17–29.
- [8] Beirlant, J., Teugels, J. L., and Vynckier, P. (1996). *Practical analysis of extreme values*, Leuven University Press, Leuven, Belgium.
- [9] Bingham, N. H., Goldie, C. M., and Teugels, J. L. (1989). *Regular variation*, Cambridge university press.
- [10] Broniatowski, M. (1993). On the estimation of the Weibull tail coefficient. *Journal of Statistical Planning and Inference*, 35, 349–366.
- [11] Feuerverger, A., and Hall, P. (1999). Estimating a tail exponent by modelling departure from a Pareto distribution, *The Annals of Statistics*, 27, 760–781.
- [12] Diebolt, J., Gardes, L., Girard, S., and Guillou, A. (2008). Bias-reduced estimators of the Weibull tail-coefficient. *Test*, 17(2), 311–331.
- [13] Diebolt, J., Gardes, L., Girard, S., and Guillou, A. (2008). Bias-reduced extreme quantile estimators of Weibull tail-distributions. *Journal of Statistical Planning and Inference*, 138(5), 1389–1401.
- [14] Gardes, L., and Girard, S. (2005). Estimating extreme quantiles of Weibull tail distributions. *Communications in Statistics—Theory and Methods*, 34(5), 1065–1080.
- [15] Gardes, L., and Girard, S. (2006). Comparison of Weibull tail-coefficient estimators. *REVSTAT—Statistical Journal*, 4(2), 163–188.

- [16] Gardes, L., and Girard, S. (2008). Estimation of the Weibull tail-coefficient with linear combination of upper order statistics, *Journal of Statistical Planning and Inference*, 138, 1416–1427.
- [17] Gardes, L., Girard, S., and Guillou, A. (2011). Weibull tail-distributions revisited: a new look at some tail estimators, *Journal of Statistical Planning and Inference*, 141(1), 429–444.
- [18] Girard, S. (2004). A Hill type estimate of the Weibull tail-coefficient. *Communications in Statistics–Theory and Methods*, 33(2), 205–234.
- [19] Goegebeur, Y., and Guillou, A. (2010). Goodness-of-fit testing for Weibull-type behavior. *Journal of Statistical Planning and Inference*, 140, 1417–1436.
- [20] Goegebeur, Y., and Guillou, A. (2011). A weighted mean excess function approach to the estimation of Weibull-type tails, *Test*, 20(1), 138–162.
- [21] de Haan, L., and Ferreira, A. (2007). *Extreme value theory: an introduction*. Springer Science and Business Media.
- [22] 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.
- [23] Hall, P., and Welsh, A. W. (1985). Adaptive estimates of parameters of regular variation. *The Annals of Statistics*, 13, 331–341.
- [24] Hill, B. M., (1975). A simple general approach to inference about the tail of a distribution, *The Annals of Statistics*, 3, 1163–1174.
- [25] Vladimirova, M., Girard, S., Hien, N., and Arbel, J. (2020). Sub-Weibull distributions: generalizing sub-Gaussian and sub-Exponential properties to heavier-tailed distributions, *Stat*, 9, e318.
- [26] Worms, J., and Worms, R. (2019). Estimation of extremes for Weibull-tail distributions in the presence of random censoring. *Extremes*, 22, 667–704.

A Appendix: Proofs

Proofs of main results are collected in Subsection A.1. Some auxiliary results are provided in Subsection A.2 and proved in Subsection A.3.

A.1 Proofs of main results

Proof of Theorem 1. Clearly,

$$\sqrt{k'_n} \log \left(\frac{\hat{q}_n(\alpha_n, k_n, k'_n)}{q(\alpha_n)} \right) = T_n^{(1)} + T_n^{(2)} + T_n^{(3)}$$

with

$$\begin{aligned} T_n^{(1)} &= \sqrt{k'_n} \left(\frac{X_{n-k_n+1, n}}{V(\log(n/k_n))} \right), \\ T_n^{(2)} &= \sqrt{k'_n} (\log \tau_n) (\hat{\theta}_n^{\text{RSH}}(k'_n) - \theta), \\ T_n^{(3)} &= \sqrt{k'_n} \left(\frac{\ell(\log(n/k_n))}{\ell(\log(1/\alpha_n))} \right). \end{aligned}$$

Let us consider the three terms separately. First, [14, Lemma 1] shows that, under **(A.1)**, $k_n \rightarrow \infty$, $k_n/n \rightarrow 0$ and (iii):

$$T_n^{(1)} = \frac{\sqrt{k'_n/k_n}}{\log(n/k_n)} \theta \xi'_n + O_P \left(\frac{\sqrt{k'_n}}{k_n \log^2(n/k_n)} \right), \quad (16)$$

where $\xi'_n \xrightarrow{d} \mathcal{N}(0, 1)$. Second, [15, Proposition 2.1] entails that, under assumptions **(A.1)**, **(A.2)**, $k'_n \rightarrow \infty$ and $k'_n/n \rightarrow 0$, the following expansion holds:

$$T_n^{(2)} = \theta \log(\tau) \xi_n + \theta \log(\tau) \mu(\log(n/k'_n)) \xi''_n + \log(\tau) \sqrt{k'_n} b(\log(n/k'_n))(1 + o(1)), \quad (17)$$

where $\xi_n \xrightarrow{d} \mathcal{N}(0, 1)$ and $\xi''_n \xrightarrow{d} \mathcal{N}(0, 1)$. Third, using **(A.2)** and (iii), one obtains

$$T_n^{(3)} = -K_\rho(\tau) \sqrt{k'_n} b(\log(n/k_n))(1 + o(1)). \quad (18)$$

Collecting (16), (17) and (18), we obtain

$$\begin{aligned} \sqrt{k'_n} \log \left(\frac{\hat{q}_n(\alpha_n, k_n, k'_n)}{q(\alpha_n)} \right) &= \theta \log(\tau) \xi_n + \theta \log(\tau) \mu(\log(n/k'_n)) \xi''_n + \frac{\sqrt{k'_n/k_n}}{\log(n/k_n)} \theta \xi'_n \\ &+ \sqrt{k'_n} \{ \log(\tau) b(\log(n/k'_n))(1 + o(1)) - K_\rho(\tau) b(\log(n/k_n))(1 + o(1)) \} \\ &+ O_P \left(\frac{\sqrt{k'_n}}{k_n \log^2(n/k_n)} \right). \end{aligned}$$

Recalling that, from [15, Lemma 5.3], $\mu(t) \sim 1/t$ as $t \rightarrow \infty$, the above expansion can be simplified as

$$\begin{aligned} \sqrt{k'_n} \log \left(\frac{\hat{q}_n(\alpha_n, k_n, k'_n)}{q(\alpha_n)} \right) &= \theta \log(\tau) \xi_n + \frac{\sqrt{k'_n/k_n}}{\log(n/k_n)} \theta \xi'_n \\ &+ \sqrt{k'_n} \{ \log(\tau) b(\log(n/k'_n))(1 + o(1)) - K_\rho(\tau) b(\log(n/k_n))(1 + o(1)) \} \\ &+ O_P \left(\frac{\sqrt{k'_n}}{k_n \log^2(n/k_n)} \right), \end{aligned}$$

Finally, remark that assumption (ii) implies $k'_n \leq k_n$ eventually and $b(\log(n/k_n)) \sim \beta^{-\rho} b(\log(n/k'_n))$ so that

$$\begin{aligned} \sqrt{k'_n} \log \left(\frac{\hat{q}_n(\alpha_n, k_n, k'_n)}{q(\alpha_n)} \right) &= \theta \log(\tau) \xi_n + \sqrt{k'_n} b(\log(n/k'_n)) (\log(\tau) - \beta^{-\rho} K_\rho(\tau) + o(1)) (1 + o(1)) \\ &+ o_P(1). \end{aligned}$$

Assumption (i) then yields

$$\sqrt{k'_n} \log \left(\frac{\hat{q}_n(\alpha_n, k_n, k'_n)}{q(\alpha_n)} \right) \xrightarrow{d} \mathcal{N}(\lambda(\log \tau - \beta^{-\rho} K_\rho(\tau)), (\theta \log \tau)^2)$$

and a first order Taylor expansion proves the result. \blacksquare

Proof of Lemma 1 (i) The result follows from a second-order Taylor expansion.

(ii) First, Lemma 4(iii) implies that, for all $\tau > 1$ and $y \leq 0$, $h_y(\tau) > 1$ and thus $\beta^*(\tau, y) = h_y(\tau)^{-1/y} > 1$. Second, Lemma 4(iii) implies that, for all $\tau \geq 1$ and $y \leq 0$, $h_y(\tau) \leq \tau^{-y/2} \leq \tau^{-y}$ when $y < 0$. This straightforwardly implies that $K_y(\tau)/\log(\tau) > \tau^y$ which is equivalent to $\beta^*(\tau, y) < \tau$. In the particular case where $y = 0$, from (i), one can take $\beta^*(\tau, 0) := \sqrt{\tau} < \tau$ since $\tau > 1$ and the result is proved.

(iii) Let us first consider $\tilde{k}_n(\tau, y) = n(k_n/n)^{\beta^*(\tau, y)}$ such that $k_n^*(\tau, y) = \lfloor \tilde{k}_n(\tau, y) \rfloor$. Clearly, $\tilde{k}_n(\tau, y)/k_n = (k_n/n)^{\beta^*(\tau, y)-1}$ and $\beta^*(\tau, y) > 1$ in view of (ii). As a consequence, for all $y \leq 0$, $\tilde{k}_n(\tau, y)$ is an increasing function of k_n , $\tilde{k}_n(\tau, y) \leq k_n$ and $\tilde{k}_n(\tau, y)/k_n \rightarrow 0$ as $n \rightarrow \infty$. These properties can be extended to $k_n^*(\tau, y)$ without difficulty since the integer part is an increasing function and $k_n^*(\tau, y) \leq \tilde{k}_n(\tau, y)$. \blacksquare

Proof of Corollary 1. (i) To prove the convergence in distribution (9), it is sufficient to show that conditions (i), (ii) and (iii) hold. First, one can easily check that $k_n^*(\tau, y) b^2(\log n) \rightarrow \lambda^2$ as $n \rightarrow \infty$ and thus $\sqrt{k_n^*(\tau, y)} b(\log n) \rightarrow \lambda$ in view of the sign assumption on λ . Besides,

$$\log(n/k_n^*(\tau, y)) = \log n + 2 \log |b(\log n)| - 2 \log |\lambda| + o(1) \sim \log n, \quad (19)$$

since $b(\cdot)$ is regularly-varying so that $b(\log(n/k_n^*(\tau, y))) \sim b(\log n)$ and thus

$$\sqrt{k_n^*(\tau, y)} b(\log(n/k_n^*(\tau, y))) \rightarrow \lambda,$$

as $n \rightarrow \infty$. (i) is thus proved. Second, observe that

$$\tau_n = \frac{\log(1/\alpha_n)}{\log(n/k_n)} = \frac{\log(1/c)}{\log(n/k_n)} + \frac{\tau}{\beta^*(\tau, \rho)} \frac{\log n}{\log(n/k_n)}$$

and

$$\log(n/k_n) = \frac{1}{\beta^*(\tau, \rho)} (\log n + 2 \log |b(\log n)| - 2 \log |\lambda|) + o(1) \sim \frac{1}{\beta^*(\tau, \rho)} \log n, \quad (20)$$

as $n \rightarrow \infty$. It is thus clear that $\tau_n \rightarrow \tau$ as $n \rightarrow \infty$, which is (ii). Third, (iii) is a straightforward consequence of (19) and (20).

(ii) Proposition 1 concludes the proof. \blacksquare

A.2 Auxiliary results

We begin with a Lemma that establishes that the strict Weibull distribution belongs to the Gumbel maximum domain of attraction ($\gamma = 0$), and more importantly, with a second-order parameter $\psi = 0$. This result illustrates why inference on Weibull-tail distributions may be difficult since the situation $\gamma = \psi = 0$ is the most complicated one for classical extreme-value estimators. Let us also recall that, in contrast, the second-order Weibull parameter is $\rho = -\infty$, see [18, Table 1] and therefore strict Weibull distributions are an easy situation for dedicated Weibull-tail estimators.

Lemma 2. *Suppose F is the cumulative distribution function of a strict Weibull distribution with shape parameter $\beta > 0$, $\beta \neq 1$ and scale parameter $\lambda > 0$. Then, the associated tail quantile function $U(\cdot) := F^{\leftarrow}(1 - 1/\cdot)$ verifies the second-order condition*

$$\frac{\frac{U(tx) - U(t)}{a(t)} - K_\gamma(x)}{A(t)} \rightarrow \int_1^x s^{\gamma-1} K_\psi(s) ds,$$

as $t \rightarrow \infty$, for all $x > 0$, see [21, Equation (3.4.5)], with $\gamma = 0$, $\psi = 0$, $a(t) = (\lambda/\beta)(\log t)^{1/\beta-1}$ and $A(t) = (1 - \beta)/(\beta \log t)$.

The following three analytical results are used to prove Proposition 1 below.

Lemma 3. $\forall u \in [0, 1]$ and $\forall \beta \in [0, 1/2]$ we have $g_\beta(u) := \log(u)(1 + u^\beta) - 2(u - 1) \leq 0$.

Lemma 4. *Let us define, for all $\tau > 1$ and $y \leq 0$,*

$$h_y(\tau) := \frac{K_0(\tau)}{K_y(\tau)} = \frac{y \log(\tau)}{\tau^y - 1} \text{ if } y < 0 \text{ and } h_0(\tau) = 1 \text{ otherwise.}$$

Then, for all $y \leq 0$:

(i) $h_y(\cdot)$ can be extended by continuity letting $h_y(1) = 1$.

(ii) $h_y(\cdot)$ is an increasing function on $[1, \infty)$.

(iii) $1 < h_y(\tau) \leq \tau^{-y/2}$ for all $\tau > 1$.

Lemma 5. *Let us define, for all $\tau > 1$ and $y \leq 0$,*

$$f_y(\tau) := -\frac{\log(h_y(\tau))}{y \log(\tau)} \text{ if } y < 0 \text{ and } f_0(\tau) = 1/2 \text{ otherwise.}$$

Then, for all $y \leq 0$:

(i) $f_y(\cdot)$ can be extended by continuity letting $f_y(1) = 1/2$.

(ii) $f_y(\cdot)$ is a decreasing function on $[1, \infty)$.

(iii) $0 \leq f_y(\tau) \leq 1/2$ for all $\tau \geq 1$.

The next Proposition establishes that the bias associated with the refined estimator of extreme quantiles is smaller than the bias associated with the original one, even though a misspecification of the second-order Weibull parameter is used.

Proposition 1. *For all $\tau > 1$, $\rho \leq 0$ and $y \leq 0$, $|B(\beta^*(\tau, y), \tau, \rho)| \leq |B(1, \tau, \rho)|$ where both quantities are defined in (4) and (8).*

A.3 Proofs of preliminary results

Proof of Lemma 2. Let $\theta = 1/\beta$. A second-order Taylor expansion yields, as $t \rightarrow \infty$,

$$\begin{aligned} U(tx) - U(t) &= \lambda(\log t)^\theta \left(\left(1 + \frac{\log x}{\log t}\right)^\theta - 1 \right) \\ &= \lambda\theta(\log x)(\log t)^{\theta-1} \left(1 + \frac{\theta-1}{2} \frac{\log x}{\log t} (1 + o(1)) \right). \end{aligned}$$

Remarking that $K_0(x) = \log x$ and letting $a(t) = \lambda\theta(\log t)^{\theta-1}$, it follows

$$\frac{U(tx) - U(t)}{a(t)} - \log x = \frac{\theta-1}{2} \frac{(\log x)^2}{\log t} (1 + o(1)) = A(t) \int_1^x \frac{\log s}{s} ds (1 + o(1)),$$

where we have defined $A(t) = (\theta-1)/\log t$, the result is thus proved. \blacksquare

Proof of Lemma 3. Differentiating three times, one gets

$$\begin{aligned} g'_\beta(u) &= \frac{1}{u} (1 - 2u + u^\beta (1 + \beta \log(u))) =: \tilde{g}_\beta(u), \\ \tilde{g}'_\beta(u) &= -2 + \beta u^{\beta-1} (2 + \beta \log(u)), \\ \tilde{g}''_\beta(u) &= \beta u^{\beta-2} (3\beta - 2 + (\beta^2 - \beta) \log(u)). \end{aligned}$$

Let $u_0(\beta) = \exp\left(\frac{2-3\beta}{\beta(\beta-1)}\right)$ be the unique point in $u \in [0, 1]$ such that $\tilde{g}''_\beta(u) = 0$ when $\beta \in [0, 1/2]$. It is easily checked that $\tilde{g}''_\beta(u) \geq 0$ when $u \in [0, u_0(\beta)]$ while $\tilde{g}''_\beta(u) \leq 0$ when $u \in [u_0(\beta), 1]$. As a consequence, $\tilde{g}'_\beta(\cdot)$ has a global maxima $m(\beta)$ on $[0, 1]$ at $u_0(\beta)$ given by

$$m_0(\beta) := \tilde{g}'_\beta(u_0(\beta)) = -2 - \frac{\beta^2}{\beta-1} \exp\left(\frac{2-3\beta}{\beta}\right).$$

The sign of $m_0(\beta)$ depends on $\beta \in [0, 1/2]$. Observe that $m_0(\beta) \rightarrow +\infty$ as $\beta \rightarrow 0$, $m_0(1/2) = -2 + e/2 < 0$ and

$$m'_0(\beta) = -\frac{\beta^2 - 4\beta + 2}{(\beta-1)^2} \exp\left(\frac{2-3\beta}{\beta}\right) < 0$$

for all $\beta \in [0, 1/2]$. As a consequence, there exists a unique $\beta_0 \in [0, 1/2]$ such that $m_0(\beta) = 0$ and $m_0(\beta) \geq 0$ for all $\beta \in [0, \beta_0]$ while $m_0(\beta) \leq 0$ when $\beta \in [\beta_0, 1/2]$. Two cases appear:

- If $\beta \in [\beta_0, 1/2]$, then $m_0(\beta) \leq 0$ and consequently $\tilde{g}'_\beta(u) \leq 0$ for all $u \in [0, 1]$. The function $\tilde{g}_\beta(\cdot)$ is thus decreasing on $[0, 1]$, and taking account of $\tilde{g}_\beta(1) = 0$ yields $\tilde{g}_\beta(u) \geq 0$ for all $u \in [0, 1]$.
- If $\beta \in [0, \beta_0]$, then $m_0(\beta) \geq 0$ and there exists two unique points $u_1(\beta) \in [0, u_0(\beta)]$ and $u_2(\beta) \in [u_0(\beta), 1]$ such that $\tilde{g}'_\beta(u) = 0$. The function $\tilde{g}_\beta(\cdot)$ is thus decreasing on $[0, u_1(\beta)]$, increasing on $[u_1(\beta), u_2(\beta)]$ and decreasing on $[u_2(\beta), 1]$ see Figure 1 for an illustration. It has two local minima given by

$$\tilde{g}_\beta(u_1(\beta)) = 1 - u_1^\beta + 2u_1 \left(\frac{1-\beta}{\beta} \right) > 0$$

and $\tilde{g}_\beta(1) = 0$. This proves that $\tilde{g}_\beta(u) \geq 0$ for all $u \in [0, 1]$.

As a conclusion, in both cases $\tilde{g}_\beta(u) \geq 0$ and $g'_\beta(u) > 0$ for all $u \in [0, 1]$ which implies that $g_\beta(\cdot)$ is an increasing function on $[0, 1]$. Since $g_\beta(1) = 0$ then $g_\beta(u) \leq 0$ for all $u \in [0, 1]$ and $\beta \in [0, 1/2]$. \blacksquare

u	0	$u_1(\beta)$	$u_0(\beta)$	$u_2(\beta)$	1
$\tilde{g}''_{\beta}(u)$		+	+	0	-
$\tilde{g}'_{\beta}(u)$			$m_0(\beta) \geq 0$		
	$-\infty$	0		0	$2\beta - 2$
$\tilde{g}'_{\beta}(u)$		-	0	+	0
$\tilde{g}_{\beta}(u)$	1			$\tilde{g}_{\beta}(u_2(\beta)) > 0$	0
		$\tilde{g}_{\beta}(u_1(\beta)) > 0$			

Figure 1: Variations of the function $\tilde{g}_{\beta}(\cdot)$, $\beta \in [0, 1/2]$ studied in the proof of Lemma 3.

Proof of Lemma 4. (i) A first-order Taylor expansion yields $h_y(1) = 1$.

(ii) Differentiating twice, one gets for all $y \leq 0$ and $\tau \geq 1$:

$$a'_y(\tau) = \frac{y}{(\tau^y - 1)^2} \left(\frac{\tau^y - 1}{\tau} - y\tau^{y-1} \log(\tau) \right) = \frac{y}{\tau(\tau^y - 1)^2} (\tau^y - 1 - y\tau^y \log(\tau)) =: \frac{y}{\tau(\tau^y - 1)^2} \tilde{a}_y(\tau).$$

with $\tilde{a}'_y(\tau) = -y^2\tau^{y-1} \log(\tau) \leq 0$, for all $y \leq 0$ and $\tau \geq 1$. As a consequence $\tilde{a}_y(\cdot)$ is a decreasing function for all $y \leq 0$. Besides, since $\tilde{a}_y(1) = 0$, it follows that $\tilde{a}_y(\tau) \leq 0$ and $a'_y(\tau) > 0$ for all $y \leq 0$ and $\tau \geq 1$. This proves that $h_y(\cdot)$ is an increasing function for all $y \leq 0$.

(iii) $h_y(\tau) \geq 1$ is a direct consequence of (i) and (ii). To prove the second inequality, let $\Delta_y(\tau) := h_y(\tau)\tau^{y/2} - 1$ for all $\tau \geq 1$ and $y \leq 0$. Differentiating three times yields

$$\begin{aligned} \Delta'_y(\tau) &= \frac{y\tau^{y/2-1}}{(\tau^y - 1)^2} \left(\tau^y - 1 - \frac{1}{2}y \log \tau (\tau^y + 1) \right) =: \frac{y\tau^{y/2-1}}{(\tau^y - 1)^2} \tilde{\Delta}_y(\tau), \\ \tilde{\Delta}'_y(\tau) &= \frac{y}{2\tau} (\tau^y(1 - y \log \tau) - 1) =: \frac{y}{2\tau} \tilde{\tilde{\Delta}}_y(\tau), \\ \tilde{\tilde{\Delta}}'_y(\tau) &= -y^2(\log \tau)\tau^{y-1} \leq 0. \end{aligned}$$

It thus appears that $\tilde{\tilde{\Delta}}_y(\cdot)$ is a decreasing function on $[1, \infty)$ with $\tilde{\tilde{\Delta}}_y(1) = 0$ so that $\tilde{\tilde{\Delta}}_y(\tau) \leq 0$ for all $\tau \geq 1$. As a consequence $\tilde{\Delta}'_y(\tau) \geq 0$ and $\tilde{\Delta}_y(\cdot)$ is an increasing function on $[1, \infty)$ with $\tilde{\Delta}_y(1) = 0$ so that $\tilde{\Delta}_y(\tau) \geq 0$ for all $\tau \geq 1$. Finally, $\Delta'_y(\tau) \leq 0$ and $\Delta_y(\cdot)$ is thus a decreasing function on $[1, \infty)$ with $\Delta_y(1) = 0$ so that $\Delta_y(\tau) \leq 0$ for all $\tau \geq 1$. The result is proved. \blacksquare

Proof of Lemma 5. (i) A second-order Taylor expansion yields $f_y(1) = 1/2$.

(ii) Differentiating twice, one has:

$$\begin{aligned} f'_y(\tau) &= -\frac{1}{\tau y \log(\tau)^2} (1 - \log h_y(\tau) - \tau^y h_y(\tau)) =: -\frac{\tilde{f}_y(\tau)}{\tau y \log(\tau)^2}, \\ \tilde{f}'_y(\tau) &= \frac{a_y^2(\tau) \tau^y - 1}{\tau \log \tau} \leq 0, \end{aligned}$$

in view of Lemma 4(iii). This implies that $\tilde{f}_y(\cdot)$ is a decreasing function on $[1, \infty)$. Besides $\tilde{f}_y(1) = 0$ since $h_y(1) = 1$ from Lemma 4(i)), and therefore $\tilde{f}_y(\tau) < 0$ and $f'_y(\tau) < 0$ for all $\tau \geq 1$ which implies that $f_y(\cdot)$ is a decreasing function on $[1, \infty)$.

(iii) $f_y(\tau) \leq 1/2$ is a direct consequence of (i) and (ii) while $f_y(\tau) \geq 0$ follows from Lemma 4(iii). ■

Proof of Proposition 1. First, note that for all $\tau > 1$, $\rho \leq 0$ and $y \leq 0$, one has

$$\begin{aligned} B(1, \tau, \rho) - B(\beta^*(\tau, y), \tau, \rho) &= \log(\tau) - K_\rho(\tau) - \log(\tau) \left(\frac{K_y(\tau)}{\log(\tau)} \right)^{\rho/y} + K_\rho(\tau) \\ &= \log(\tau) (1 - h_y(\tau)^{-\rho/y}) \\ &\geq 0, \end{aligned} \tag{21}$$

since $\log(\tau) \geq 0$, $\rho/y \geq 0$ and in view of Lemma 4(iii). Second,

$$\rho \{B(1, \tau, \rho) + B(\beta^*(\tau, y), \tau, \rho)\} = \rho \left(\log(\tau) + \log(\tau) \left(\frac{K_y(\tau)}{\log(\tau)} \right)^{\rho/y} - 2K_\rho(\tau) \right).$$

for all $\tau \geq 1$, $\rho < 0$ and $y \leq 0$. The change of variable $\rho \mapsto u = \tau^\rho$ yields

$$\begin{aligned} \rho \{B(1, \tau, \rho) + B(\beta^*(\tau, y), \tau, \rho)\} &= \log(u) + \log(u) \left(\frac{K_y(\tau)}{\log(\tau)} \right)^{\log(u)/(y \log(\tau))} - 2(u - 1) \\ &= \log(u) (1 + u^{f_y(\tau)}) - 2(u - 1), \\ &= g_{f_y(\tau)}(u), \end{aligned}$$

with

$$\begin{aligned} g_\beta(u) &= \log(u)(1 + u^\beta) - 2(u - 1), \text{ (see Lemma 3),} \\ f_y(\tau) &= \frac{1}{y \log(\tau)} \log \left(\frac{K_y(\tau)}{\log(\tau)} \right) = -\frac{\log(a(\tau))}{y \log(\tau)}, \text{ (see Lemma 5),} \\ h_y(\tau) &= \frac{K_0(\tau)}{K_y(\tau)}, \text{ (see Lemma 4).} \end{aligned}$$

Lemma 5(iii) shows that $0 \leq f_y(\tau) \leq 1/2$ and thus one can apply Lemma 3 to obtain $g_{f_y(\tau)}(u) \leq 0$ and consequently

$$B(1, \tau, \rho) + B(\beta^*(\tau, y), \tau, \rho) = \frac{g_{f_y(\tau)}(u)}{\rho} \geq 0. \tag{22}$$

Collecting (21) and (22) concludes the proof. ■

B Appendix: Tables

Diebolt	RWT	RSH	LSE	MEF
$\theta = 1/2$				
$\rho = -1/2$	0.0134	0.0442	0.0204	0.3601
$\theta = 3/4$				
$\rho = -1/2$	0.0705	0.0461	0.0413	-
$\theta = 1$				
$\rho = -1$	0.1133	0.3455	0.1819	-
$\rho = -1/2$	0.1318	0.0639	0.0717	-
$\theta = 5/4$				
$\rho = -1$	0.1427	0.3391	0.1852	-
$\rho = -1/2$	0.1919	0.0853	0.1024	0.8001
$\theta = 3/2$				
$\rho = -3/2$	0.2118	0.7104	0.3858	0.4408
$\rho = -1$	0.1889	0.3362	0.1893	0.8957
$\rho = -1/2$	0.2486	0.1079	0.1330	0.1175
$\theta = 7/4$				
$\rho = -3/2$	0.2474	0.6997	0.3843	0.1889
$\rho = -1$	0.2406	0.3337	0.1990	0.1022
$\rho = -1/2$	0.3099	0.1351	0.1655	0.6295
$\theta = 2$				
$\rho = -2$	0.3203	0.8839	0.4438	0.8286
$\rho = -3/2$	0.2866	0.6953	0.3846	0.5759
$\rho = -1$	0.2941	0.3317	0.2135	0.5930
$\rho = -1/2$	0.3734	0.1603	0.1971	0.8609
$\theta = 9/4$				
$\rho = -2$	0.3784	0.8781	0.4388	0.8840
$\rho = -3/2$	0.3350	0.6919	0.3861	0.8253
$\rho = -1$	0.3495	0.3304	0.2339	0.8536
$\rho = -1/2$	0.4488	0.1872	0.2315	0.9509
$\theta = 5/2$				
$\rho = -5/2$	0.8869	0.9498	0.3756	0.9611
$\rho = -2$	0.4341	0.8745	0.4384	0.9376
$\rho = -3/2$	0.3783	0.6901	0.3884	0.9221
$\rho = -1$	0.4045	0.3298	0.2529	0.9384
$\rho = -1/2$	0.5098	0.2095	0.2623	0.9776

Table 2: MARE associated with 4 estimators of the extreme quantile $q(\alpha_n = 1/n^2)$ computed on simulated data from the $\mathcal{D}(\alpha = 1/\theta, \beta = -\rho, a = 10)$ -distribution. The best result is emphasized in bold. MAREs larger than 1 are not reported.

Diebolt	RWT	RSH	LSE	MEF
$\theta = 1/2$				
$\rho = -1/2$	0.0364	0.0836	0.0500	0.4395
$\theta = 3/4$				
$\rho = -1/2$	0.1040	0.0825	0.0683	-
$\theta = 1$				
$\rho = -1$	0.2803	0.5477	0.3753	-
$\rho = -1/2$	0.1916	0.0961	0.1070	-
$\theta = 5/4$				
$\rho = -1$	0.2954	0.5413	0.3765	-
$\rho = -1/2$	0.2765	0.1199	0.1491	-
$\theta = 3/2$				
$\rho = -3/2$	0.3994	0.8696	0.6239	0.6051
$\rho = -1$	0.3338	0.5390	0.3752	0.7545
$\rho = -1/2$	0.3531	0.1466	0.1925	0.1618
$\theta = 7/4$				
$\rho = -3/2$	0.4250	0.8634	0.6149	0.5039
$\rho = -1$	0.3799	0.5365	0.3749	0.3113
$\rho = -1/2$	0.4400	0.1775	0.2412	0.7275
$\theta = 2$				
$\rho = -2$	0.4661	0.9612	0.6624	0.9438
$\rho = -3/2$	0.4610	0.8606	0.6124	0.7907
$\rho = -1$	0.4327	0.5345	0.3800	0.7539
$\rho = -1/2$	0.5179	0.2075	0.2869	0.9194
$\theta = 9/4$				
$\rho = -2$	0.5578	0.9589	0.6436	0.9631
$\rho = -3/2$	0.5078	0.8588	0.6094	0.9310
$\rho = -1$	0.4949	0.5323	0.3911	0.9324
$\rho = -1/2$	0.6062	0.2408	0.3414	0.9767
$\theta = 5/2$				
$\rho = -5/2$	-	0.9866	0.5616	0.9907
$\rho = -2$	0.6339	0.9574	0.6336	0.9828
$\rho = -3/2$	0.5578	0.8576	0.6060	0.9745
$\rho = -1$	0.5529	0.5305	0.4063	0.9761
$\rho = -1/2$	0.6987	0.2700	0.3833	0.9910

Table 3: MARE associated with 4 estimators of the extreme quantile $q(\alpha_n = 1/n^4)$ computed on simulated data from the $\mathcal{D}(\alpha = 1/\theta, \beta = -\rho, a = 10)$ -distribution. The best result is emphasized in bold. MAREs larger than 1 are not reported.