

Asymptotic behavior of the extrapolation error associated with the estimation of extreme quantiles

Clément Albert, Anne Dutfoy, Stéphane Girard

► **To cite this version:**

Clément Albert, Anne Dutfoy, Stéphane Girard. Asymptotic behavior of the extrapolation error associated with the estimation of extreme quantiles. 2018. <hal-01692544v2>

HAL Id: hal-01692544

<https://hal.archives-ouvertes.fr/hal-01692544v2>

Submitted on 26 Feb 2018

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.

Asymptotic behavior of the extrapolation error associated with the estimation of extreme quantiles

Clément Albert⁽¹⁾, Anne Dutfoy⁽²⁾ and Stéphane Girard^(1, *)

⁽¹⁾ *Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France*

⁽²⁾ *EDF R&D dept. Périclès, 91120 Palaiseau, France*

Abstract

We investigate the asymptotic behavior of the (relative) extrapolation error associated with some estimators of extreme quantiles based on extreme-value theory. It is shown that the extrapolation error can be interpreted as the remainder of a first order Taylor expansion. Necessary and sufficient conditions are then provided such that this error tends to zero as the sample size increases. Interestingly, in case of the so-called Exponential Tail estimator, these conditions lead to a subdivision of Gumbel maximum domain of attraction into three subsets. In contrast, the extrapolation error associated with Weissman estimator has a common behavior over the whole Fréchet maximum domain of attraction. First order equivalents of the extrapolation error are then derived and their accuracy is illustrated numerically.

Keywords: Extrapolation error, Extreme quantiles, Extreme-value theory.

AMS 2000 subject classification: 62G32, 62G20.

1 Introduction

The starting point of this work is the study of the asymptotic behavior of the Exponential Tail (ET) estimator, a nonparametric estimator of the extreme quantiles from an unknown distribution. Theoretical developments can be found in [4] while numerical aspects are investigated in [10]. Given a n -sample X_1, \dots, X_n from a cumulative distribution function F with associated survival distribution function \bar{F} , an extreme quantile is a $(1-p_n)$ th quantile $q(p_n)$ of F essentially larger than the maximal observation, *i.e.* such that $\bar{F}(q(p_n)) = p_n$ with $p_n \leq 1/n$. The estimation of extreme quantiles requires specific methods. Among them, the Peaks Over Threshold (POT) method relies on an approximation of the distribution of excesses over a given threshold [22]. More precisely, let us introduce a deterministic threshold u_n such that $\bar{F}(u_n) = \alpha_n$ or equivalently $u_n = q(\alpha_n)$ with $\alpha_n \rightarrow 0$ and $n\alpha_n > 1$ as $n \rightarrow \infty$. The excesses above u_n are defined as $Y_i = X_i - u_n$ for all $X_i > u_n$. The survival distribution function of an excess is given

*Corresponding author, Stephane.Girard@inria.fr

by $\bar{F}_{u_n}(x) = \bar{F}(u_n + x)/\bar{F}(u_n)$. Pickands theorem [12, 21] states that, under mild conditions, \bar{F}_{u_n} can be approximated by a Generalized Pareto Distribution (GPD). As a consequence, the extreme quantile $q(p_n)$ can be in turn approximated by the deterministic term

$$\tilde{q}_{\text{GPD}}(p_n; \alpha_n) = q(\alpha_n) + \frac{\sigma_n}{\gamma_n} \left[\left(\frac{\alpha_n}{p_n} \right)^{\gamma_n} - 1 \right], \quad (1)$$

where σ_n and γ_n are respectively the scale and shape parameters of the GPD distribution. Then, the POT method consists in estimating these two unknown parameters. The ET method corresponds to the important particular case where F belongs to Gumbel Maximum Domain of Attraction, MDA(Gumbel). In such a situation, $\gamma_n = 0$ and the GPD distribution reduces to an Exponential distribution with scale parameter σ_n . Thus, approximation (1) can be rewritten as

$$\tilde{q}_{\text{ET}}(p_n; \alpha_n) = q(\alpha_n) + \sigma_n \log(\alpha_n/p_n) \quad (2)$$

and the associated estimator [4] is

$$\hat{q}_{\text{ET}}(p_n; \alpha_n) = \hat{q}(\alpha_n) + \hat{\sigma}_n \log(\alpha_n/p_n)$$

where $\hat{q}(\alpha_n) = X_{n-k_n+1,n}$ with $k_n = \lfloor n\alpha_n \rfloor$ and

$$\hat{\sigma}_n = \frac{1}{k_n} \sum_{i=1}^{k_n} X_{n-i+1,n}.$$

Let us recall that $X_{1,n} \leq \dots \leq X_{n,n}$ denote the order statistics associated with X_1, \dots, X_n . The error $(q(p_n) - \hat{q}_{\text{ET}}(p_n; \alpha_n))$ can be expanded as a sum of two terms:

$$q(p_n) - \hat{q}_{\text{ET}}(p_n; \alpha_n) = (\tilde{q}_{\text{ET}}(p_n; \alpha_n) - \hat{q}_{\text{ET}}(p_n; \alpha_n)) + (q(p_n) - \tilde{q}_{\text{ET}}(p_n; \alpha_n)),$$

the first one being a random estimation error

$$\tilde{q}_{\text{ET}}(p_n; \alpha_n) - \hat{q}_{\text{ET}}(p_n; \alpha_n) = q(\alpha_n) - \hat{q}(\alpha_n) + (\sigma_n - \hat{\sigma}_n) \log(\alpha_n/p_n) \quad (3)$$

and the second one being a deterministic extrapolation error

$$q(p_n) - \tilde{q}_{\text{ET}}(p_n; \alpha_n) = q(p_n) - q(\alpha_n) - \sigma_n \log(\alpha_n/p_n). \quad (4)$$

The asymptotic behavior of the estimation error (3) is driven by the asymptotic distributions of $\hat{q}(\alpha_n)$ and $\hat{\sigma}_n$ which can be found for instance in [6], Theorem 2.4.1 and Theorem 3.4.2 respectively.

In this paper, we focus on the asymptotic behavior of the extrapolation error (4). Indeed, in view of (2), it appears that the ET method extrapolates in the distribution tail from $q(\alpha_n)$ to $q(p_n)$ thanks to an additive correction proportional to $\log(\alpha_n/p_n)$. Our goal is thus to quantify to which extent this extrapolation can be performed in a consistent way. More specifically, we provide necessary and sufficient conditions on the pair (p_n, α_n) such that the relative extrapolation error

$$\varepsilon_{\text{ET}}(p_n; \alpha_n) := (q(p_n) - \tilde{q}_{\text{ET}}(p_n; \alpha_n))/q(p_n) \quad (5)$$

tends to zero as $n \rightarrow \infty$. These conditions depend on the underlying distribution function F and they lead to a subdivision of MDA(Gumbel) into three sub-domains depending on the restrictions they impose on the extrapolation range. Related works include [5, 17] who exhibited penultimate approximations for F^n together with convergence rates for distributions in MDA(Gumbel). These results were extended to other maximum domains of attraction in [18, 19] while penultimate approximations were established for the distribution of the excesses F_{u_n} in [24].

Here, we focus on the approximation of quantiles rather than approximations of distribution functions. Let us also highlight that these investigations are not limited to the ET method. To illustrate this, let us introduce $x_n = \log(1/\alpha_n)$, $y_n = \log(1/p_n)$ and $\varphi(\cdot) = (\bar{F})^{-1}(1/\exp(\cdot))$. The extrapolation error (4) can thus be interpreted as the remainder of a first order Taylor expansion:

$$q(p_n) - \tilde{q}_{\text{ET}}(p_n; \alpha_n) = \varphi(y_n) - \varphi(x_n) - \sigma_n(y_n - x_n) \text{ where } \sigma_n = \varphi'(x_n). \quad (6)$$

We shall show that Weissman estimator [23] dedicated to MDA(Fréchet) can also enter this framework thanks to adapted definitions of x_n , y_n and φ . In this case, the necessary and sufficient conditions on the extrapolation range are automatically fulfilled for most distributions in MDA(Fréchet) which is a very different situation from MDA(Gumbel).

The paper is organized as follows: The asymptotic behavior of the remainder associated with the first order Taylor expansion (6) is investigated in Section 2. The applications to ET and Weissman approximations are detailed in Section 3 and Section 4 respectively. As a conclusion, some numerical illustrations are presented in Section 5. Proofs are postponed to Section 6 and auxiliary results can be found in the Appendix.

2 Theoretical framework

Let φ be an arbitrary twice differentiable, increasing function. Let (x_n) and (y_n) be two sequences such that $0 < x_n \leq y_n$ for all n large enough, $x_n \rightarrow \infty$ as $n \rightarrow \infty$ and

$$0 < \liminf x_n/y_n \leq \limsup x_n/y_n \leq 1. \quad (7)$$

Motivated by (5) and (6), we introduce

$$\Delta_n = \frac{\varphi(y_n) - \varphi(x_n) - (y_n - x_n)\varphi'(x_n)}{\varphi(y_n)}. \quad (8)$$

The goal of this section is to establish necessary and sufficient conditions on $\delta_n := (y_n - x_n)/y_n$ so that $\Delta_n \rightarrow 0$ as $n \rightarrow \infty$. The following two functions are of the utmost importance in this study:

$$K_1(x) = \frac{x\varphi'(x)}{\varphi(x)}, \quad K_2(x) = \frac{x^2\varphi''(x)}{\varphi(x)}, \quad x \in \mathbb{R}.$$

The limits of $K_1(x)$ and $K_2(x)$ when $x \rightarrow \infty$ are denoted by $\ell_1 \in \bar{\mathbb{R}}_+$ and $\ell_2 \in \bar{\mathbb{R}}$, when they exist. The study of Δ_n relies on the assumption that K_1 is regularly-varying at infinity with

index $\theta_1 \leq 1$. This property is denoted for short by $K_1 \in RV_{\theta_1}$ and means that K_1 is ultimately positive such that

$$\frac{K_1(tx)}{K_1(x)} \rightarrow t^{\theta_1} \text{ as } x \rightarrow \infty \text{ for all } t > 0.$$

We refer to [3] for a general account on regular variation theory. This assumption is discussed in Section 3 and Section 4 while applying this general framework to the particular cases of ET and Weissman estimators.

The first step is to establish the links between δ_n and Δ_n through the functions K_1 and K_2 .

Lemma 1 (Bounds on Δ_n)

(i) If $K_1 \in RV_{\theta_1}$ with $\theta_1 \leq 1$ and (7) holds then

$$\Delta_n = \delta_n^2 \int_0^1 \frac{K_2(y_n(1 - \delta_n u))}{(1 - \delta_n u)^2} \exp(K_1(y_n)L_{\theta_1}(1 - \delta_n u)(1 + o(1))) \, u \, du,$$

where $L_{\theta_1}(x) = \int_1^x u^{\theta_1-1} du$ for all $x \in \mathbb{R}$.

(ii) If, moreover, K_2 is ultimately of constant sign and monotone, then

$$\begin{aligned} |\Delta_n| &\leq \max(|K_2(y_n)|, |K_2(x_n)|) \frac{\delta_n^2}{(1 - \delta_n)^2} \Phi(\delta_n K_1(y_n)(1 + o(1))) \text{ and} \\ |\Delta_n| &\geq \min(|K_2(y_n)|, |K_2(x_n)|) \delta_n^2 \Phi\left(\delta_n K_1(y_n)(1 - \delta_n)^{\theta_1-1}(1 + o(1))\right), \end{aligned}$$

where $\Phi(t) = \int_0^1 u \exp(-ut) du$ for all $t \geq 0$.

Second, as a consequence of the above bounds, a sufficient condition as well as a necessary condition can be established such that $\Delta_n \rightarrow 0$ as $n \rightarrow \infty$.

Lemma 2 (Necessary and sufficient conditions for $\Delta_n \rightarrow 0$)

Assume (7) holds, $K_1 \in RV_{\theta_1}$ with $\theta_1 \leq 1$ and K_2 is ultimately of constant sign and monotone.

(i) If $\delta_n^2 \max(|K_2(y_n)|, |K_2(x_n)|) \rightarrow 0$ then $\Delta_n \rightarrow 0$ as $n \rightarrow \infty$.

(ii) If $\Delta_n \rightarrow 0$ then $\delta_n^2 \min(|K_2(y_n)|, |K_2(x_n)|) \rightarrow 0$ as $n \rightarrow \infty$.

It appears that the asymptotic behavior of Δ_n heavily relies on ℓ_2 . Since ℓ_2 can be derived from ℓ_1 , see Lemma 4 in the Appendix, we are now in position to state our first main result:

Proposition 1 (Role of ℓ_1 for $\Delta_n \rightarrow 0$)

Under the assumptions of Lemma 2,

(i) If $\ell_1 \in \{0, 1\}$ then $\ell_2 = 0$ and $\Delta_n \rightarrow 0$ as $n \rightarrow \infty$.

(ii) If $\ell_1 \in (0, \infty) \setminus \{1\}$ then $\ell_2 \in (0, \infty)$ and $\Delta_n \rightarrow 0$ if and only if $\delta_n \rightarrow 0$ as $n \rightarrow \infty$.

(iii) If $\ell_1 = \infty$ then $|\ell_2| = \infty$ and $\Delta_n \rightarrow 0$ if and only if $\delta_n^2 K_2(y_n) \rightarrow 0$ as $n \rightarrow \infty$.

Three cases appear. If $\ell_1 \in \{0, 1\}$ then $\Delta_n \rightarrow 0$ as $n \rightarrow \infty$ as soon as (7) holds. If $0 < \ell_1 < \infty$ and $\ell_1 \neq 1$ then a necessary and sufficient condition for $\Delta_n \rightarrow 0$ is $\delta_n \rightarrow 0$ as $n \rightarrow \infty$. If $\ell_1 = \infty$ then the necessary and sufficient condition for $\Delta_n \rightarrow 0$ is $\delta_n^2 K_2(y_n) \rightarrow 0$ as $n \rightarrow \infty$. Clearly, this condition implies $\delta_n \rightarrow 0$ since, in this situation, $|\ell_2| = \infty$.

Finally, first order approximations of Δ_n can be provided in each situation.

Proposition 2 (First order approximations of Δ_n)

Under the assumptions of Lemma 2,

(i) *Assume $\ell_1 = 0$ (and thus $\ell_2 = 0$).*

(a) *If $\delta_n \rightarrow 0$ then*

$$\frac{1}{2}\delta_n^2 |K_2(y_n)|(1 + o(1)) \leq |\Delta_n| \leq \frac{1}{2}\delta_n^2 |K_2(x_n)|(1 + o(1)).$$

(b) *If $\delta_n \rightarrow \delta \in (0, 1)$ then*

$$|K_2(y_n)| \frac{\delta^2}{2}(1 + o(1)) \leq |\Delta_n| \leq |K_2(x_n)| \frac{\delta^2}{2(1 - \delta)^2}(1 + o(1)).$$

(ii) *Assume $\ell_1 = 1$ (and thus $\ell_2 = 0$).*

(a) *If $\delta_n \rightarrow 0$ then*

$$\frac{1}{2}\delta_n^2 |K_2(y_n)|(1 + o(1)) \leq |\Delta_n| \leq \frac{1}{2}\delta_n^2 |K_2(x_n)|(1 + o(1)).$$

(b) *If $\delta_n \rightarrow \delta \in (0, 1)$ then*

$$|K_2(y_n)| \delta^2 \Phi(\delta/(1 - \delta))(1 + o(1)) \leq |\Delta_n| \leq |K_2(x_n)| \frac{\delta^2}{(1 - \delta)^2} \Phi(\delta)(1 + o(1)).$$

(iii) *Assume $0 < \ell_1 < \infty$ and $\ell_1 \neq 1$.*

(a) *If $\delta_n \rightarrow 0$ then*

$$\Delta_n = \frac{\ell_1(\ell_1 - 1)}{2} \delta_n^2 (1 + o(1)) \text{ as } n \rightarrow \infty.$$

(b) *If $\delta_n \rightarrow \delta \in (0, 1)$ then*

$$\Delta_n \rightarrow c(\delta, \ell_1) := \ell_1(\ell_1 - 1)\delta^2 \int_0^1 (1 - \delta u)^{\ell_1 - 2} u du \text{ as } n \rightarrow \infty.$$

(iv) *Assume $\ell_1 = \infty$.*

(a) *If $\delta_n K_1(y_n) \rightarrow 0$ then*

$$\Delta_n = \frac{1}{2}\delta_n^2 K_1^2(y_n) \sim \frac{1}{2}\delta_n^2 K_2(y_n)(1 + o(1)) \text{ as } n \rightarrow \infty.$$

(b) *If $\delta_n K_1(y_n) \rightarrow a > 0$ then*

$$\Delta_n \rightarrow a^2 \Phi(a) \text{ as } n \rightarrow \infty.$$

(c) If $\delta_n K_1(y_n) \rightarrow \infty$ then

$$\Delta_n \rightarrow 1 \text{ as } n \rightarrow \infty.$$

In situations (i) and (ii) where $\ell_1 \in \{0, 1\}$, $\Delta_n \rightarrow 0$ in both cases $\delta_n \rightarrow 0$ and $\delta_n \rightarrow \delta \neq 0$, and the convergence is the fastest in the case $\delta_n \rightarrow 0$. In situation (iii) where $0 < \ell_1 < \infty$ and $\ell_1 \neq 1$, Δ_n is asymptotically proportional to δ_n^2 . In situation (iv) where $\ell_1 = \infty$, $\Delta_n \rightarrow 0$ in the only case where $\delta_n K_1(y_n) \rightarrow 0$ and Δ_n is asymptotically proportional to $(\delta_n K_1(y_n))^2$ or equivalently to $\delta_n^2 K_2(y_n)$.

3 Application to the ET approximation

Recall that $y_n = \log(1/p_n)$, $x_n = \log(1/\alpha_n)$ with $0 < p_n \leq 1/n \leq \alpha_n < 1$. Introduce

$$\tau_n = \frac{\log(1/p_n)}{\log(n)} \text{ and } \tau'_n = \frac{\log(1/\alpha_n)}{\log(n)}$$

so that $p_n = n^{-\tau_n}$, $\tau_n \geq 1$, $\alpha_n = n^{-\tau'_n}$, $\tau'_n \leq 1$ and $\delta_n = 1 - \tau'_n/\tau_n$. In the sequel, F is assumed to be increasing and twice differentiable and the cumulative hazard rate function is denoted by $H(\cdot) = -\log \bar{F}(\cdot)$. Following the ideas of Section 1, we let $\varphi(\cdot) = (\bar{F})^{-1}(1/\exp(\cdot)) = H^{-1}(\cdot)$ so that $\varepsilon_{\text{ET}}(p_n; \alpha_n) = \Delta_n$. In this context, the assumption $K_1 \in RV_{\theta_1}$, $\theta_1 \in \mathbb{R}$ has been introduced and discussed by Cees de Valk *et. al.* in a series of papers [7, 8, 9]. The next result provides a characterization of the tail behavior of F according to the sign of θ_1 .

Proposition 3 (Characterizations)

Let $x^* := \sup\{x : F(x) < 1\}$ be the endpoint of F and assume that the density F' is ultimately monotone.

(i) If $H \in RV_{\beta}$, $\beta > 0$, then $K_1 \in RV_0$ and $\ell_1 = 1/\beta$.

(ii) $K_1 \in RV_{\theta_1}$, $\theta_1 > 0$ (and thus $\ell_1 = \infty$) if and only if $x^* = \infty$ and $H(\exp \cdot) \in RV_{1/\theta_1}$.

(iii) $K_1 \in RV_{\theta_1}$, $\theta_1 < 0$ (and thus $\ell_1 = 0$) if and only if $x^* < \infty$ and $H(x^*(1 - 1/\cdot)) \in RV_{-1/\theta_1}$.

In the case (i) where H is regularly varying with index $\beta > 0$, necessarily $\theta_1 = 0$ and F is referred to as a Weibull tail-distribution, see for instance [2, 13, 16]. Such distributions encompass Gaussian, Gamma, Exponential and strict Weibull distributions. In the case (ii) where $H(\exp \cdot)$ is regularly varying, F is called a log-Weibull tail-distribution, see [1, 11, 15], the most popular example being the lognormal distribution. The case (iii) corresponds to distributions with a Weibull tail behavior in the neighborhood of a finite endpoint.

Besides, let us highlight that the domain of attraction associated with F depends on the position of θ_1 with respect to 1:

Proposition 4 (Domains of attraction)

- (i) If $K_1 \in RV_{\theta_1}$, $\theta_1 < 1$, then $F \in MDA(\text{Gumbel})$.
- (ii) If $F \in MDA(\text{Fréchet})$ then $K_1 \in RV_1$.
- (iii) If $K_1 \in RV_{\theta_1}$, $\theta_1 > 1$, then F does not belong to any domain of attraction.

These results justify the assumption $\theta_1 \leq 1$ used in Lemma 2 and its consequences. The situation $\theta_1 > 1$ which does not correspond to any domain of attraction is sometimes referred to as super-heavy tails, see [1] or [3], Section 8.8 for further developments on this topic.

Applying Proposition 1 to the ET framework yields:

Theorem 1 (Necessary and sufficient conditions for $\varepsilon_{\text{ET}}(p_n; \alpha_n) \rightarrow 0$)

Assume $0 < \liminf \tau'_n/\tau_n \leq \limsup \tau'_n/\tau_n \leq 1$, $K_1 \in RV_{\theta_1}$ with $\theta_1 \leq 1$ and K_2 is ultimately of constant sign and monotone.

- (i) If $\ell_1 \in \{0, 1\}$ then $\varepsilon_{\text{ET}}(p_n; \alpha_n) \rightarrow 0$ as $n \rightarrow \infty$.
- (ii) If $\ell_1 \in (0, \infty) \setminus \{1\}$ then $\varepsilon_{\text{ET}}(p_n; \alpha_n) \rightarrow 0$ if and only if $\tau_n \rightarrow 1$ and $\tau'_n \rightarrow 1$ as $n \rightarrow \infty$.
- (iii) If $\ell_1 = \infty$ then $\varepsilon_{\text{ET}}(p_n; \alpha_n) \rightarrow 0$ if and only if $(\tau_n - \tau'_n)^2 K_2(\log n) \rightarrow 0$ as $n \rightarrow \infty$.

It appears that, from the extrapolation error point of view, $MDA(\text{Gumbel})$ can be splitted into three sub-domains of attraction:

- $MDA_1(\text{Gumbel})$ defined by $\ell_1 \in \{0, 1\}$ and where the relative extrapolation error tends to zero without restriction on the order p_n of the extreme quantile. As illustrated by Proposition 3(iii), the case $\ell_1 = 0$ includes distributions with a finite endpoint. The case $\ell_1 = 1$ encompasses Weibull tail-distributions with shape parameter $\beta = 1$ (Proposition 3(i)), *i.e.* close to the Exponential distribution (the Gamma distribution for instance) as well as the class E defined in [5].
- $MDA_2(\text{Gumbel})$ defined by $\ell_1 \in (0, \infty) \setminus \{1\}$ and where the relative extrapolation error tends to zero for extreme quantiles close to the maximal observation in the sense that $\log(p_n) \sim \log(1/n)$ as $n \rightarrow \infty$. Extreme orders such as $p_n = n^{-\tau}$, $\tau > 1$ are thus not permitted. As illustrated by Proposition 3(i), this situation encompasses Weibull tail-distributions with shape parameter $\beta \neq 1$ *i.e.* far from the Exponential distribution (the Gaussian distribution for instance).
- $MDA_3(\text{Gumbel})$ defined by $\ell_1 = \infty$ and where the relative extrapolation error tends to zero under strong restrictions on the order p_n of the extreme quantile: $\log(p_n)/\log(1/n) = 1 + o(|K_2(\log n)|^{1/2})$ as $n \rightarrow \infty$. As illustrated by Proposition 3(ii), this case corresponds to log-Weibull tail-distributions (including the lognormal distribution for instance).

We refer to Table 1 for examples of distributions in each sub-domain. To conclude this part, one may obtain first order approximations of the relative extrapolation error $\varepsilon_{ET}(p_n; \alpha_n)$ thanks to Proposition 2. For the sake of simplicity, it is assumed that $|K_2|$ is regularly varying. This assumption is in fact not restrictive. Indeed, in case of Weibull tail-distributions with shape parameter $\beta \neq 1$, Proposition 3(i) shows that $\ell_1 \neq 1$ and Lemma 4(iii) entails $|K_2| \in RV_0$. In case of log-Weibull tail-distributions, $\theta_1 > 0$ from Proposition 3(ii) and thus $K_2 \in RV_{2\theta_1}$ from Lemma 4(iv). Finally, for finite endpoint distributions, $\theta_1 < 0$ from Proposition 3(iii) and thus $|K_2| \in RV_{\theta_1}$ from Lemma 4(i).

The results are collected in Theorem 2 below. The following notations have been introduced: $u_n \sim v_n$ if and only if $u_n/v_n \rightarrow 1$ as $n \rightarrow \infty$ and $u_n \asymp v_n$ if and only if $0 < \liminf u_n/v_n \leq \limsup u_n/v_n < \infty$.

Theorem 2 (First order approximations of $\varepsilon_{ET}(p_n; \alpha_n)$)

Suppose $0 < \liminf \tau'_n/\tau_n \leq \limsup \tau'_n/\tau_n \leq 1$, $K_1 \in RV_{\theta_1}$ with $\theta_1 \leq 1$ and $|K_2| \in RV_{\theta_2}$ with $\theta_2 \in \mathbb{R}$. Then, as $n \rightarrow \infty$:

(i) Assume $F \in MDA_1(\text{Gumbel})$

(a) If $\delta_n \rightarrow 0$ then $\varepsilon_{ET}(p_n; \alpha_n) \sim \frac{1}{2}(\tau_n - \tau'_n)^2 K_2(\log n) \sim \frac{1}{2}\delta_n^2 K_2(\log n)$.

(b) If $\delta_n \rightarrow \delta \in (0, 1)$ then $|\varepsilon_{ET}(p_n; \alpha_n)| \asymp |K_2(\log n)|$.

(ii) Assume $F \in MDA_2(\text{Gumbel})$

(a) If $\delta_n \rightarrow 0$ then $\varepsilon_{ET}(p_n; \alpha_n) \sim \frac{\ell_1(\ell_1-1)}{2}(\tau_n - \tau'_n)^2 \sim \frac{\ell_1(\ell_1-1)}{2}\delta_n^2$.

(b) If $\delta_n \rightarrow \delta \in (0, 1)$ then $\varepsilon_{ET}(p_n; \alpha_n) \rightarrow c(\delta, \ell_1)$.

(iii) Assume $F \in MDA_3(\text{Gumbel})$

(a) If $\delta_n K_1(\log n) \rightarrow 0$ then $\varepsilon_{ET}(p_n; \alpha_n) \sim \frac{1}{2}(\tau_n - \tau'_n)^2 K_1^2(\log n) \sim \frac{1}{2}\delta_n^2 K_2(\log n)$.

(b) If $\delta_n K_1(\log n) \rightarrow a > 0$ then $\varepsilon_{ET}(p_n; \alpha_n) \rightarrow a^2 \Phi(a)$.

(c) If $\delta_n K_1(\log n) \rightarrow \infty$ then $\varepsilon_{ET}(p_n; \alpha_n) \rightarrow 1$.

Remark that the only situation where $\delta_n \rightarrow \delta \neq 0$ and $\varepsilon_{ET}(p_n; \alpha_n) \rightarrow 0$ as $n \rightarrow \infty$ occurs for $F \in MDA_1(\text{Gumbel})$. In this case, it is possible to choose extreme orders such that $p_n = n^{-\tau}$, $\tau > 1$, and the relative extrapolation error tends to zero at a logarithmic rate.

As expected, in the three situations (i,ii,iii)-(a) where $\delta_n \rightarrow 0$ and $\varepsilon_{ET}(p_n; \alpha_n) \rightarrow 0$ as $n \rightarrow \infty$, the convergence is the fastest in $MDA_1(\text{Gumbel})$ and the slowest in $MDA_3(\text{Gumbel})$. Let us also highlight that the rate of convergence is independent from the distribution in $MDA_2(\text{Gumbel})$. To illustrate these results, let us focus on the distributions introduced in Table 1. Clearly, in all six cases, $F \in DA(\text{Gumbel})$, K_1 and $|K_2|$ are regularly varying so that the assumptions of Theorem 2 are fulfilled. Let us consider the case where $p_n = 1/(n \log n)$ and $\alpha_n = (\log n)/n$ leading to

$$\tau_n = 1 + \frac{\log \log n}{\log n}, \tau'_n = 1 - \frac{\log \log n}{\log n} \text{ and } \delta_n \sim 2 \frac{\log \log n}{\log n}, \quad (9)$$

as $n \rightarrow \infty$. Let us stress that these choices entail $\delta_n \rightarrow 0$ and $\delta_n K_1(\log n) \rightarrow 0$ as $n \rightarrow \infty$ so that Theorem 2(i,ii,iii)-(a) can be applied and $\varepsilon_{\text{ET}}(p_n; \alpha_n) \rightarrow 0$ as $n \rightarrow \infty$ for all six distributions. The associated first order approximations of $\varepsilon_{\text{ET}}(p_n; \alpha_n)$ are provided in Table 2. It appears that, in most cases, the convergence of the relative extrapolation error to zero is rather slow. The log-Weibull($\beta > 1$) distribution corresponds to the worst case, since arbitrary low rates of convergence can be obtained by letting $\beta \xrightarrow{>} 1$. At the opposite, the Finite endpoint($\beta > 0$) distribution is the most favorable case, letting $\beta \xrightarrow{>} 0$ could lead to arbitrary high logarithmic rates of convergence.

4 Application to Weissman approximation

When $F \in \text{MDA}(\text{Fréchet})$, $\gamma_n > 0$ and the GPD approximation (1) can be simplified by letting $\sigma_n = \gamma_n q(\alpha_n)$, see [6], Theorem 1.2.5, leading to

$$\tilde{q}_W(p_n; \alpha_n) = q(\alpha_n) \left(\frac{\alpha_n}{p_n} \right)^{\gamma_n}, \quad (10)$$

which is called Weissman approximation in the sequel. Weissman estimator [23] is then obtained by replacing the intermediate quantile $q(\alpha_n)$ and the tail index γ_n by appropriate estimators:

$$\hat{q}_W(p_n; \alpha_n) = \hat{q}(\alpha_n) \left(\frac{\alpha_n}{p_n} \right)^{\hat{\gamma}_n}.$$

The most common choices are $\hat{q}(\alpha_n) = X_{n-k_n+1,n}$, see Section 1, and Hill estimator [20]:

$$\hat{\gamma}_n = \frac{1}{k_n} \sum_{i=1}^{k_n} \log X_{n-i+1,n} - \log X_{n-k_n+1,n}.$$

Taking the logarithm of (10) yields

$$\log q(p_n) - \log \tilde{q}_W(p_n; \alpha_n) = \log q(p_n) - \log q(\alpha_n) - \gamma_n \log(\alpha_n/p_n)$$

and thus, similarly to the ET case (4), the extrapolation error can be interpreted as a first order Taylor remainder. To this end, recall that $y_n = \log(1/p_n)$, $x_n = \log(1/\alpha_n)$ with $0 < p_n \leq 1/n \leq \alpha_n < 1$ and introduce $\varphi(\cdot) = \log(\bar{F})^{-1}(1/\exp(\cdot)) = \log U(\exp \cdot)$ where U is the tail quantile function, so that

$$\log q(p_n) - \log \tilde{q}_W(p_n; \alpha_n) = \varphi(y_n) - \varphi(x_n) - \gamma_n(y_n - x_n) \text{ where } \gamma_n = \varphi'(x_n).$$

The quantity of interest is

$$\varepsilon_W(p_n; \alpha_n) := (q(p_n) - \tilde{q}_W(p_n; \alpha_n))/q(p_n) = 1 - \exp(-\Delta_n \log q(p_n)), \quad (11)$$

where Δ_n is defined in (8). Here, the property $K_1 \in RV_{\theta_1}$ is a direct consequence of the assumption $F \in \text{MDA}(\text{Fréchet})$. Indeed, from [6], Corollary 1.2.1, $F \in \text{MDA}(\text{Fréchet})$ is equivalent to $U \in RV_\gamma$ for some $\gamma > 0$ which can be rewritten as

$$U(t) = t^\gamma L(t), \text{ with } L \in RV_0 \text{ and } \gamma > 0. \quad (12)$$

The function L is said to be slowly-varying [3]. Classical properties of slowly-varying functions yield, as $t \rightarrow \infty$,

$$\begin{aligned}\varphi(t) &= \gamma t + \log L(\exp t) \sim \gamma t, \\ \varphi'(t) &= \gamma + \eta(\exp t) \rightarrow \gamma,\end{aligned}\tag{13}$$

where $\eta(t) = tL'(t)/L(t)$ is called the auxiliary function associated with L . It follows that $K_1(t) \rightarrow 1$ as $t \rightarrow \infty$ and thus $\ell_1 = 1$ and $K_1 \in RV_0$. This means that only the case (i) of Proposition 1 and the case (ii) of Proposition 2 have to be considered. In particular $\Delta_n \rightarrow 0$ as $n \rightarrow \infty$ without further assumption, which is a very different situation from Section 3:

Theorem 3 (First order approximation of $\varepsilon_W(p_n; \alpha_n)$)

Assume (12) hold and let $\eta(t) = tL'(t)/L(t)$ be the auxiliary function associated with L . Suppose η is asymptotically monotone with a constant sign and $|\eta| \in RV_\rho$ with $\rho < 0$.

(i) If $\delta_n \rightarrow \delta \in (0, 1)$ then $\varepsilon_W(p_n; \alpha_n) \rightarrow 0$ as $n \rightarrow \infty$.

(ii) Moreover, as $n \rightarrow \infty$,

$$\varepsilon_W(p_n; \alpha_n) \sim -\frac{\delta}{1-\delta} \log(1/\alpha_n) \eta(1/\alpha_n).$$

The assumption $|\eta| \in RV_\rho$, $\rho < 0$, is recurrent in extreme-value statistics to control the bias of estimators, ρ being known as the second-order parameter, see *e.g.* [14]. This assumption holds for most heavy-tailed distributions such as Fréchet, Pareto or Student distributions.

Let us also remark that one can choose extreme orders such that $p_n = n^{-\tau}$, $\tau > 1$ as in DA_1 (Gumbel), see Theorem 2(i)-(b), and still obtain $\varepsilon_W(p_n; \alpha_n) \rightarrow 0$ as $n \rightarrow \infty$. However, here, the relative extrapolation error converges to zero at a polynomial rate, depending on ρ .

5 Numerical illustrations

To conclude, the quality of the first order approximations associated with the ET relative extrapolation error given in Table 2 is assessed graphically. Recall that these results are obtained by applying Theorem 2 to sequences (τ_n) and (τ'_n) given in (9) and distributions described in Table 1: Finite endpoint($\beta = 5$), Gamma($a = 0.1$), Weibull($\beta = 5$), Gaussian, log-Weibull($\beta = 3$) and lognormal($\sigma = 0.5$). The exact relative extrapolation error $\varepsilon_{ET}(p_n; \alpha_n)$ as well as the corresponding first order approximation provided by Theorem 2 are computed as functions of $\log n$. The results are displayed on Figures 1–3. It appears that, for all six distributions, the relative extrapolation error converges towards zero as predicted by Theorem 2, even though the convergence can be very slow in DA_3 (Gumbel), see Figure 3. In all cases, the asymptotic sign of $\varepsilon_{ET}(p_n; \alpha_n)$ is coherent with the first order equivalent given in Table 2: Positive for Gamma($a < 1$), log-Weibull($\beta > 1$) and lognormal distributions, negative for Finite endpoint($\beta > 0$), Weibull($\beta > 1$) and Gaussian distributions. Finally, the first order equivalent provides a reasonable approximation of the error behavior in all situations.

6 Proofs of main results

Proof of Lemma 1. (i) A second order Taylor expansion with integral remainder yields

$$\begin{aligned}\Delta_n &= \int_{x_n}^{y_n} \frac{K_2(t)}{t^2} \frac{\varphi(t)}{\varphi(y_n)} (y_n - t) dt \\ &= \delta_n^2 \int_0^1 \frac{K_2(y_n(1 - \delta_n u))}{(1 - \delta_n u)^2} \frac{\varphi(y_n(1 - \delta_n u))}{\varphi(y_n)} u du,\end{aligned}$$

thanks to the change of variable $u = (y_n - t)/(y_n - x_n)$. Besides,

$$\begin{aligned}\frac{\varphi(y_n(1 - \delta_n u))}{\varphi(y_n)} &= \exp(\log \varphi(y_n(1 - \delta_n u)) - \log \varphi(y_n)) \\ &= \exp\left(\int_{y_n}^{y_n(1 - \delta_n u)} (\log \varphi(t))' dt\right) \\ &= \exp\left(\int_{y_n}^{y_n(1 - \delta_n u)} \frac{K_1(t)}{t} dt\right) \\ &= \exp\left(\int_1^{1 - \delta_n u} \frac{K_1(vy_n)}{v} dv\right) \\ &= \exp\left(K_1(y_n) \int_1^{1 - \delta_n u} \frac{K_1(vy_n)}{K_1(y_n)} \frac{dv}{v}\right).\end{aligned}$$

Since $1 - \delta_n u \in [1 - \delta_n, 1]$, the regular variation property of K_1 yields $K_1(vy_n)/K_1(y_n) \rightarrow v^{\theta_1}$ uniformly locally as $y_n \rightarrow \infty$, and condition (7) leads to

$$\frac{\varphi(y_n(1 - \delta_n u))}{\varphi(y_n)} = \exp(K_1(y_n)L_{\theta_1}(1 - \delta_n u)(1 + o(1))).$$

It thus follows that

$$\Delta_n = \delta_n^2 \int_0^1 \frac{K_2(y_n(1 - \delta_n u))}{(1 - \delta_n u)^2} \exp(K_1(y_n)L_{\theta_1}(1 - \delta_n u)(1 + o(1))) u du$$

and the first part of the result is proved.

(ii) The sign of K_2 is supposed to be ultimately constant so that

$$|\Delta_n| = \delta_n^2 \int_0^1 \frac{|K_2(y_n(1 - \delta_n u))|}{(1 - \delta_n u)^2} \exp(K_1(y_n)L_{\theta_1}(1 - \delta_n u)(1 + o(1))) u du.$$

Let us remark that, for all $u \in [0, 1]$ and $\theta_1 \leq 1$, one has $1 - \delta_n \leq 1 - \delta_n u \leq 1$ and

$$-(1 - \delta_n)^{\theta_1 - 1} \delta_n u \leq L_{\theta_1}(1 - \delta_n u) \leq -\delta_n u.$$

It is thus clear that

$$\begin{aligned}|\Delta_n| &\leq \frac{\delta_n^2}{(1 - \delta_n)^2} \int_0^1 |K_2(y_n(1 - \delta_n u))| \exp(-\delta_n K_1(y_n)u(1 + o(1))) u du \text{ and} \\ |\Delta_n| &\geq \delta_n^2 \int_0^1 |K_2(y_n(1 - \delta_n u))| \exp\left(-\delta_n K_1(y_n)(1 - \delta_n)^{\theta_1 - 1} u(1 + o(1))\right) u du.\end{aligned}$$

Since $|K_2|$ is assumed to be ultimately monotone, it follows that, for n large enough,

$$\min(|K_2(y_n)|, |K_2(x_n)|) \leq |K_2(y_n(1 - \delta_n u))| \leq \max(|K_2(y_n)|, |K_2(x_n)|),$$

leading to

$$\begin{aligned} |\Delta_n| &\leq \max(|K_2(y_n)|, |K_2(x_n)|) \frac{\delta_n^2}{(1 - \delta_n)^2} \int_0^1 u \exp(-\delta_n K_1(y_n) u (1 + o(1))) du \text{ and} \\ |\Delta_n| &\geq \min(|K_2(y_n)|, |K_2(x_n)|) \delta_n^2 \int_0^1 u \exp\left(-\delta_n K_1(y_n) (1 - \delta_n)^{\theta_1 - 1} u (1 + o(1))\right) du. \end{aligned}$$

Introducing for all $t \geq 0$, $\Phi(t) = \int_0^1 u \exp(-ut) du$, the above bounds can be rewritten as

$$\begin{aligned} |\Delta_n| &\leq \max(|K_2(y_n)|, |K_2(x_n)|) \frac{\delta_n^2}{(1 - \delta_n)^2} \Phi(\delta_n K_1(y_n) (1 + o(1))) \text{ and} \\ |\Delta_n| &\geq \min(|K_2(y_n)|, |K_2(x_n)|) \delta_n^2 \Phi\left(\delta_n K_1(y_n) (1 - \delta_n)^{\theta_1 - 1} (1 + o(1))\right), \end{aligned}$$

which concludes the proof. \blacksquare

Proof of Lemma 2. Let us focus on the situation where $|K_2|$ is ultimately non decreasing and introduce $A_n = \delta_n K_1(y_n)$.

(i) Assume that $\delta_n^2 |K_2(y_n)| \rightarrow 0$. From Lemma 3(i) in the Appendix, $0 \leq \Phi(t) \leq 1/2$ for all $t \geq 0$ and thus Lemma 1(ii) entails

$$|\Delta_n| \leq \frac{\delta_n^2 |K_2(y_n)|}{2(1 - \delta_n)^2} \rightarrow 0 \tag{14}$$

as $n \rightarrow \infty$ in view of (7).

(ii) From Lemma 1(ii), one has

$$|\Delta_n| \geq |K_2(x_n)| \delta_n^2 \Phi\left(A_n (1 - \delta_n)^{\theta_1 - 1} (1 + o(1))\right) \geq |K_2(x_n)| \delta_n^2 \Phi(c A_n)$$

for n large enough and some $c > 0$ since Φ is non-increasing, see Lemma 3(i). For all $t \geq 0$, let $\psi(t) = \int_0^t x \exp(-x) dx = t^2 \Phi(t)$. Consider $t_0 \geq c(3 - 2\theta_1)$ with $\theta_1 \leq 1$ and remark that $\Phi(t) \geq \Phi(t_0)$ for all $0 \leq t \leq t_0$ and $\psi(t) \geq \psi(t_0)$ for all $t \geq t_0$. As a consequence, for all $t > 0$,

$$\Phi(t) \geq \frac{\psi(t_0)}{t_0^2} \mathbb{I}\{t \leq t_0\} + \frac{\psi(t_0)}{t^2} \mathbb{I}\{t \geq t_0\},$$

and thus

$$\begin{aligned} |\Delta_n| &\geq \frac{\psi(t_0)}{t_0^2} |K_2(x_n)| \delta_n^2 \mathbb{I}\{A_n \leq t_0/c\} + \frac{\psi(t_0)}{c^2} \frac{|K_2(x_n)|}{K_1^2(y_n)} \mathbb{I}\{A_n \geq t_0/c\} \\ &\geq \frac{\psi(t_0)}{t_0^2} |K_2(x_n)| \delta_n^2 \mathbb{I}\{A_n \leq t_0/c\} + \frac{\psi(t_0)}{c^2} \frac{|K_2(x_n)|}{K_1^2(x_n)} \frac{K_1^2(x_n)}{K_1^2(y_n)} \mathbb{I}\{A_n \geq t_0/c\}. \end{aligned} \tag{15}$$

Since K_1 is regularly varying, $K_1(x_n)/K_1(y_n) \sim (1 - \delta_n)^{\theta_1} \geq c' > 0$ in view of (7) and

$$|\Delta_n| \geq \frac{\psi(t_0)}{t_0^2} |K_2(x_n)| \delta_n^2 \mathbb{I}\{A_n \leq t_0/c\} + \psi(t_0) \left(\frac{c'}{c}\right)^2 \frac{|K_2(x_n)|}{K_1^2(x_n)} \mathbb{I}\{A_n \geq t_0/c\}.$$

Remarking that, (18) in the proof of Lemma 4 (see Appendix) implies that, for n large enough,

$$\frac{K_2(x_n)}{K_1^2(x_n)} = 1 + \frac{1}{K_1(x_n)} \left(\frac{x_n K_1'(x_n)}{K_1(x_n)} - 1 \right) = 1 + \frac{\delta_n}{A_n} (\theta_1 - 1 + o(1))$$

which yields when $A_n \geq t_0/c$,

$$\left| \frac{K_2(x_n)}{K_1^2(x_n)} - 1 \right| \leq \frac{c\delta_n}{t_0} |\theta_1 - 1 + o(1)| \leq \frac{c}{t_0} (3/2 - \theta_1) \leq \frac{1}{2}.$$

It thus follows that

$$\frac{|K_2(x_n)|}{K_1^2(x_n)} \mathbb{I}\{A_n \geq t_0/c\} \geq \frac{1}{2} \mathbb{I}\{A_n \geq t_0/c\}$$

and therefore,

$$|\Delta_n| \geq \frac{\psi(t_0)}{t_0^2} |K_2(x_n)| \delta_n^2 \mathbb{I}\{A_n \leq t_0/c\} + \frac{\psi(t_0)}{2} \left(\frac{c'}{c} \right)^2 \mathbb{I}\{A_n \geq t_0/c\}.$$

As a conclusion, $|\Delta_n| \rightarrow 0$ implies $|K_2(x_n)| \delta_n^2 \mathbb{I}\{A_n \leq t_0/c\} \rightarrow 0$ and $\mathbb{I}\{A_n \geq t_0/c\} \rightarrow 0$. Consequently, $A_n \leq t_0/c$ eventually and $\delta_n^2 K_2(x_n) \rightarrow 0$ as $n \rightarrow \infty$.

Let us now consider the situation where $|K_2|$ is ultimately non increasing.

(i) The proof is similar, the upper bound (14) is replaced by

$$|\Delta_n| \leq \frac{\delta_n^2 |K_2(x_n)|}{2(1 - \delta_n)^2}. \quad (16)$$

(ii) The lower bound (15) is replaced by

$$|\Delta_n| \geq \frac{\psi(t_0)}{t_0^2} |K_2(y_n)| \delta_n^2 \mathbb{I}\{A_n \leq t_0/c\} + \frac{\psi(t_0)}{c^2} \frac{|K_2(y_n)|}{K_1^2(y_n)} \mathbb{I}\{A_n \geq t_0/c\}$$

and the end of the proof is similar. ■

Proof of Proposition 1. (i) If $\ell_1 = 0$ then Lemma 4(i) in the Appendix shows that $\ell_2 = 0$.

If $\ell_1 = 1$ then, from Lemma 4(ii), $\ell_2 = 0$. Lemma 2(i) concludes the proof.

(ii) If $0 < \ell_1 < \infty$ and $\ell_1 \neq 1$ then Lemma 4(iii) entails that ℓ_2 is finite and non zero. Lemma 2(i,ii) concludes the proof.

(iii) If $\ell_1 = \infty$ then $K_2(x) \sim K_1^2(x)$ as $x \rightarrow \infty$, see Lemma 4(iv). Besides, K_2 is regularly varying of order $2\theta_1$ and $K_2(x_n) \asymp K_2(y_n)$. Lemma 2(i,ii) concludes the proof under the assumption (7). ■

Proof of Proposition 2. (i) Assume $\ell_1 = 0$ and $\delta_n \rightarrow \delta \in [0, 1)$. Then $\delta_n K_1(y_n) \rightarrow 0$ as $n \rightarrow \infty$, $\ell_2 = 0$ from Lemma 4(i) and, since $\Phi(t) \rightarrow 1/2$ as $t \rightarrow 0$, Lemma 1(ii) yields

$$\begin{aligned} |\Delta_n| &\leq \max(|K_2(y_n)|, |K_2(x_n)|) \frac{\delta_n^2}{2(1 - \delta)^2} (1 + o(1)) \text{ and} \\ |\Delta_n| &\geq \min(|K_2(y_n)|, |K_2(x_n)|) \frac{\delta_n^2}{2} (1 + o(1)). \end{aligned}$$

The conclusion follows by remarking that $|K_2|$ cannot be increasing when $\ell_2 = 0$.

(ii) Assume $\ell_1 = 1$ and $\delta_n \rightarrow \delta \in [0, 1)$. Then $\theta_1 = 0$, $\ell_2 = 0$ from Lemma 4(ii), $\delta_n K_1(y_n) \rightarrow \delta$ as $n \rightarrow \infty$ and, since Φ is continuous from Lemma 3(i), Lemma 1(ii) yields

$$\begin{aligned} |\Delta_n| &\leq \max(|K_2(y_n)|, |K_2(x_n)|) \frac{\delta_n^2}{(1-\delta)^2} \Phi(\delta)(1+o(1)) \text{ and} \\ |\Delta_n| &\geq \min(|K_2(y_n)|, |K_2(x_n)|) \delta_n^2 \Phi(\delta/(1-\delta))(1+o(1)), \end{aligned}$$

and the conclusion follows similarly to (i).

(iii) Assume $0 < \ell_1 < \infty$ and $\ell_1 \neq 1$. Then Lemma 4(iii) entails $\theta_1 = 0$, $\ell_2 = \ell_1(\ell_1 - 1)$, and Lemma 1(i) yields

$$\Delta_n = \delta_n^2 \int_0^1 K_2(y_n(1-\delta_n u))(1-\delta_n u)^{K_1(y_n)(1+o(1))-2} u du.$$

When $\delta_n \rightarrow \delta \in [0, 1)$, Lebesgue's dominated convergence theorem entails

$$\int_0^1 K_2(y_n(1-\delta_n u))(1-\delta_n u)^{K_1(y_n)(1+o(1))-2} u du \rightarrow \ell_1(\ell_1 - 1) \int_0^1 (1-\delta u)^{\ell_1-2} u du,$$

and the result is proved.

(iv) Assume $\ell_1 = \infty$. Then Lemma 4(iv) entails that $K_2(x) \sim K_1^2(x)$ as $x \rightarrow \infty$. As a consequence, Lemma 1(i) and Lebesgue's dominated convergence theorem yield

$$\Delta_n = \delta_n^2 \int_0^1 \frac{K_1^2(y_n(1-\delta_n u))}{(1-\delta_n u)^2} \exp(K_1(y_n)L_{\theta_1}(1-\delta_n u)(1+o(1))) u du(1+o(1)).$$

The regular variation property entails

$$\Delta_n = \delta_n^2 K_1^2(y_n) \int_0^1 (1-\delta_n u)^{2\theta_1-2} \exp(K_1(y_n)(L_{\theta_1}(1-\delta_n u)(1+o(1)))) u du(1+o(1)).$$

Two main situations are considered:

1. If $\delta_n \rightarrow 0$, then $L_{\theta_1}(1-\delta_n u) = -\delta_n u(1+o(1))$. Letting $A_n = \delta_n K_1(y_n)$, it follows

$$\Delta_n = A_n^2 \int_0^1 \exp(-A_n u(1+o(1))) u du(1+o(1)) = \Phi(A_n(1+o(1))) A_n^2(1+o(1)).$$

Three sub-cases arise: (a) If $A_n \rightarrow 0$ then $\Phi(A_n) \rightarrow 1/2$ in view of Lemma 3(i) and

$$\Delta_n \sim \frac{1}{2} \delta_n^2 K_1^2(y_n).$$

(b) If $A_n \rightarrow a > 0$ then $\Delta_n \rightarrow a^2 \Phi(a)$ in view of the continuity of Φ , see Lemma 3(i).

(c) If $A_n \rightarrow \infty$, then $\Phi(A_n) \sim 1/A_n^2$ from Lemma 3(ii) and therefore $\Delta_n \rightarrow 1$.

2. If $\delta_n \rightarrow \delta \in (0, 1)$, then $A_n \rightarrow \infty$ and we thus are in situation (c). Two successive change of variables yield

$$\begin{aligned} \Delta_n &= \delta^2 K_1^2(y_n) \int_0^1 (1-\delta u)^{2\theta_1-2} \exp(K_1(y_n)L_{\theta_1}(1-\delta u)(1+o(1))) u du(1+o(1)) \\ &= K_1^2(y_n) \int_{1-\delta}^1 (1-v)v^{2\theta_1-2} \exp(K_1(y_n)L_{\theta_1}(v)(1+o(1))) dv(1+o(1)) \\ &= K_1^2(y_n) \int_{L_{\theta_1}(1-\delta)}^0 \left(L_{\theta_1}^{-1}(w)\right)^{\theta_1-1} \left(1-L_{\theta_1}^{-1}(w)\right) \exp(K_1(y_n)w(1+o(1))) dw(1+o(1)). \end{aligned}$$

Let us introduce $\xi(w) = (L_{\theta_1}^{-1}(w))^{\theta_1-1}(1 - L_{\theta_1}^{-1}(w))$ for all $w \in [L_{\theta_1}(1 - \delta), 0]$. Routine calculations show that $\xi(0) = 0$ and $\xi'(0) = -1$. A second order Taylor expansion thus yields $\xi(w) = -w + w^2\xi''(\eta_w)/2$ with $\eta_w \in [w, 0] \subset [L_{\theta_1}(1 - \delta), 0]$. Replacing, we get

$$\begin{aligned}\Delta_n &= -K_1^2(y_n) \int_{L_{\theta_1}(1-\delta)}^0 w \exp(K_1(y_n)w(1 + o(1))) dw(1 + o(1)) + R_n \\ &= K_1^2(y_n)\Psi_1(K_1(y_n)(1 + o(1)); -L_{\theta_1}(1 - \delta)) + R_n,\end{aligned}$$

where Ψ_1 is defined in Lemma 3 and

$$R_n = \frac{1}{2}K_1^2(y_n) \int_{L_{\theta_1}(1-\delta)}^0 w^2\xi''(\eta_w) \exp(K_1(y_n)w(1 + o(1))) dw(1 + o(1))$$

Remarking that $|\xi''|$ is bounded on compact sets, there exists $M > 0$ such that

$$|R_n| \leq MK_1^2(y_n)\Psi_2(K_1(y_n)(1 + o(1)); -L_{\theta_1}(1 - \delta)),$$

where Ψ_2 is defined in Lemma 3. As a consequence of Lemma 3(ii), $R_n = O(1/K_1(y_n))$ and $\Delta_n \rightarrow 1$ as $n \rightarrow \infty$. Let us remark that this case is similar to the situation $\delta_n \rightarrow 0$ and $A_n \rightarrow \infty$. \blacksquare

Proof of Proposition 3. (i) If $H \in RV_\beta$, $\beta > 0$ then the monotone density theorem ([3], Proposition 1.7.2) yields

$$H(t) \sim \frac{1}{\beta}tH'(t) \text{ as } t \rightarrow \infty.$$

Letting $x = H(t)$, we have

$$x \sim \frac{1}{\beta} \frac{H^{-1}(x)}{(H^{-1})'(x)}$$

or equivalently $K_1(x) \rightarrow 1/\beta$ as $x \rightarrow \infty$. It follows that $\ell_1 = 1/\beta$ and $K_1 \in RV_0$.

(ii, \Leftarrow) Let us assume that $H(\exp \cdot) \in RV_{1/\theta_1}$, $\theta_1 > 0$. Then, $\log H^{-1} \in RV_{\theta_1}$ and the monotone density theorem ([3], Proposition 1.7.2) implies $(\log H^{-1})' \in RV_{\theta_1-1}$ i.e. $K_1 \in RV_{\theta_1}$.

(ii, \Rightarrow) Conversely, assume $K_1 \in RV_{\theta_1}$, $\theta_1 > 0$. Then, necessarily $\ell_1 = \infty$. From [3], Theorem 1.5.8, we have for all x_0 sufficiently large,

$$\log H^{-1}(x) - \log H^{-1}(x_0) = \int_{x_0}^x (\log H^{-1}(t))' dt = \int_{x_0}^x \frac{K_1(t)}{t} dt = \frac{1}{\theta_1}K_1(x)(1 + o(1)), \quad (17)$$

as $x \rightarrow \infty$. It is thus clear that $\log H^{-1} \in RV_{\theta_1}$ and therefore $H(\exp \cdot) \in RV_{1/\theta_1}$.

(iii, \Leftarrow) Let us assume that $x^* < \infty$ and $h(\cdot) := H(x^*(1 - 1/\cdot)) \in RV_{-1/\theta_1}$, $\theta_1 < 0$. As a consequence, $H^{-1}(\cdot) = x^*(1 - 1/h^{-1}(\cdot))$ where $h^{-1} \in RV_{-\theta_1}$ and $h^{-1}(x) \rightarrow \infty$ as $x \rightarrow \infty$. Straightforward calculations and the monotone density theorem ([3], Proposition 1.7.2) lead to

$$\begin{aligned}\log H^{-1}(x) &= \log x^* + \log \left(1 - \frac{1}{h^{-1}(x)}\right) \\ (\log H^{-1})'(x) &= \frac{(h^{-1})'(x)}{h^{-1}(x)(h^{-1}(x) - 1)} \\ K_1(x) &= \frac{x(h^{-1})'(x)}{h^{-1}(x)(h^{-1}(x) - 1)} \sim -\frac{\theta_1}{h^{-1}(x)},\end{aligned}$$

and therefore $K_1 \in RV_{\theta_1}$, $\theta_1 < 0$.

(iii, \implies) Conversely, assume $K_1 \in RV_{\theta_1}$, $\theta_1 < 0$. Thus $(\log H^{-1})' \in RV_{\theta_1-1}$ and [3], Theorem 1.5.8 yields first, for all x sufficiently large,

$$\log x^* - \log H^{-1}(x) = \int_x^\infty (\log H^{-1})'(t) dt < \infty$$

and thus $x^* < \infty$. Second, one also has

$$\frac{K_1(x)}{\int_x^\infty (\log H^{-1})'(t) dt} \rightarrow -\theta_1$$

as $x \rightarrow \infty$. Combining the two above results yield

$$\frac{K_1(x)}{\log x^* - \log H^{-1}(x)} = -\theta_1(1 + o(1))$$

and consequently

$$H^{-1}(x) = x^* \exp\left(\frac{1}{\theta_1} K_1(x)(1 + o(1))\right) = x^* \left(1 + \frac{1}{\theta_1} K_1(x)(1 + o(1))\right)$$

since $K_1(x) \rightarrow 0$ as $x \rightarrow \infty$. Applying [3], Theorem 1.5.12 yields $H(x^*(1 - 1/\cdot)) \in RV_{-1/\theta_1}$ and concludes the proof. \blacksquare

Proof of Proposition 4. (i) Assume $K_1 \in RV_{\theta_1}$, $\theta_1 < 1$ and let $U(\cdot) = H^{-1}(\log \cdot)$ be the tail quantile function. For all $x > 0$ and $t > 0$, consider

$$\frac{U'(tx)}{U'(t)} = \frac{1}{x} \frac{(H^{-1})'(\log tx)}{(H^{-1})'(\log t)} = \frac{1}{x} \frac{\log t}{\log tx} \frac{H^{-1}(\log tx)}{H^{-1}(\log t)} \frac{K_1(\log tx)}{K_1(\log t)}.$$

Since $K_1 \in RV_{\theta_1}$ and the logarithm is a slowly-varying function, $K_1(\log \cdot) \in RV_0$ and thus

$$\frac{U'(tx)}{U'(t)} = \frac{1}{x} \frac{H^{-1}(\log tx)}{H^{-1}(\log t)} (1 + o(1))$$

as $t \rightarrow \infty$. Besides,

$$\begin{aligned} \frac{H^{-1}(\log tx)}{H^{-1}(\log t)} &= \exp(\log H^{-1}(\log tx) - \log H^{-1}(\log t)) \\ &= \exp\left(\int_{\log t}^{\log tx} (\log H^{-1})'(u) du\right) \\ &= \exp\left(\int_{\log t}^{\log tx} \frac{K_1(u)}{u} du\right) \\ &= \exp\left(\log x \int_0^1 \frac{K_1(\log t + v \log x)}{\log t + v \log x} dv\right), \end{aligned}$$

and the regular variation property of K_1 implies that

$$\frac{K_1(\log t + v \log x)}{\log t + v \log x} = \frac{K_1(\log t)}{\log t} (1 + o(1))$$

as $t \rightarrow \infty$ uniformly locally on $v \in [0, 1]$. It follows that

$$\frac{H^{-1}(\log tx)}{H^{-1}(\log t)} = \exp\left(\log x \frac{K_1(\log t)}{\log t}(1 + o(1))\right) \rightarrow 1$$

as $t \rightarrow \infty$ since $K \in RV_{\theta_1}$ with $\theta_1 < 1$. As a conclusion, $U'(tx)/U'(t) \rightarrow 1/x$ as $t \rightarrow \infty$ for all $x > 0$ and thus $U' \in RV_{-1}$. This implies that $F \in \text{MDA}(\text{Gumbel})$, see [6], Corollary 1.1.10.

(ii) Assume $F \in \text{MDA}(\text{Fréchet})$. From [6], Corollary 1.2.10, there exists $\gamma > 0$ such that the tail quantile function $U \in RV_\gamma$. Since $H^{-1}(\cdot) = U(\exp \cdot)$, it follows that

$$K_1(x) = x \frac{\exp(x)U'(\exp x)}{U(\exp x)} \sim \gamma x$$

as $x \rightarrow \infty$ from the monotone density theorem [3], Theorem 1.7.2. It is thus clear that $K_1 \in RV_1$.

(iii) Assume $K_1 \in RV_{\theta_1}$, $\theta_1 > 1$. First, Proposition 3(ii) implies $x^* = \infty$ and thus $F \notin \text{MDA}(\text{Weibull})$. Second, Proposition 4(ii) shows that $F \in \text{MDA}(\text{Fréchet})$ entails $K_1 \in RV_1$. It is thus clear that $F \notin \text{MDA}(\text{Fréchet})$. Finally, it remains to show that $F \notin \text{MDA}(\text{Gumbel})$. To this end, consider for all $x > 0$ and $t \rightarrow \infty$,

$$\frac{U(tx)}{U(t)} = \frac{H^{-1}(\log tx)}{H^{-1}(\log t)} = \exp\left\{\frac{1}{\theta_1}(K_1(\log tx) - K_1(\log t))(1 + o(1))\right\}$$

from (17) in the proof of Proposition 3(ii, \implies). A first order Taylor expansion yields

$$\frac{U(tx)}{U(t)} = \exp\left\{\frac{\log x}{\theta_1}K_1'(\log t + \eta \log x)(1 + o(1))\right\} = \exp\left\{\frac{\log x}{\theta_1}K_1'(\log t)(1 + o(1))\right\}$$

where $\eta \in (0, 1)$ since $K_1' \in RV_{\theta_1-1}$. Recalling that $\theta_1 > 1$, it follows that $K_1'(\log t) \rightarrow \infty$ as $t \rightarrow \infty$ and therefore $U(tx)/U(t) \rightarrow 0$ as $t \rightarrow \infty$ if $x < 1$ and $U(tx)/U(t) \rightarrow \infty$ as $t \rightarrow \infty$ if $x > 1$. Finally [6], Lemma 1.2.9 shows that $F \notin \text{MDA}(\text{Gumbel})$ since $U(tx)/U(t)$ does not converge to 1 as $t \rightarrow \infty$. \blacksquare

Proof of Theorem 1. The proof relies on the application of Proposition 1. Condition (7) is fulfilled under the assumption $0 < \liminf \tau_n'/\tau_n \leq \limsup \tau_n'/\tau_n \leq 1$.

(i) is a straightforward consequence of Proposition 1(i).

(ii) is based on the remark that $\delta_n \rightarrow 0$ if and only if $\tau_n \rightarrow 1$ and $\tau_n' \rightarrow 1$ since, by assumption, $\tau_n' \leq 1 \leq \tau_n$.

(iii) Since $\ell_2 = \infty$, $\delta_n^2 K_2(\tau_n \log n) \rightarrow 0$ implies $\delta_n \rightarrow 0$ and thus $\tau_n \rightarrow 1$ and $\tau_n' \rightarrow 1$ in view of the above remark. Thus, $\delta_n^2 \sim (\tau_n - \tau_n')^2$ as $n \rightarrow \infty$. Besides, Lemma 4(iv) in the Appendix entails that K_2 is regularly varying when $\ell_2 = \infty$. As a consequence, $K_2(\tau_n \log n) \sim K_2(\log n)$ as $n \rightarrow \infty$ and the result is proved. \blacksquare

Proof of Theorem 2. The proof relies on the application of Proposition 2.

(i) is a consequence of Proposition 2(i, ii). Let us highlight that, when $\delta_n \rightarrow 0$ then $x_n \sim y_n$ and thus $|K_2(x_n)| \sim |K_2(y_n)| \sim |K_2(\log n)|$ in view of the assumption $|K_2|$ is regularly varying.

Moreover, $\delta_n^2 \sim (\tau_n - \tau'_n)^2$ as already seen in the proof of Theorem 1. Let us also remark that, since the sign of K_2 is ultimately constant, in view of Lemma 1(i), $\varepsilon_{\text{ET}}(p_n; \alpha_n)$ and $K_2(\log n)$ have ultimately the same sign. The absolute values can thus be dropped. Similarly, when $\delta_n \rightarrow \delta \in (0, 1)$ then $x_n \asymp y_n$ and thus $|K_2(x_n)| \asymp |K_2(y_n)| \asymp |K_2(\log n)|$.

(ii) is a straightforward consequence of Proposition 2(iii).

(iii) is a consequence of Proposition 2(iv). If $\delta_n K_1(\log n) \rightarrow a \geq 0$ then necessarily $\delta_n \rightarrow 0$ and thus $y_n \sim \log n$ leading to $K_1(y_n) \sim K_1(\log n)$. In the case where $\delta_n K_1(\log n) \rightarrow \infty$, one still has $1 \leq \liminf \tau_n \leq \limsup \tau_n < \infty$ and thus $K_1(y_n) \asymp K_1(\log n)$, the result is proved. \blacksquare

Proof of Theorem 3. The proof relies on the application of Lemma 1(i) and the fact that $K_1(t) \rightarrow 1$ as $t \rightarrow \infty$ entailing $K_1 \in RV_0$:

$$\Delta_n = \delta_n^2 \int_0^1 K_2(y_n(1 - \delta_n u)) (1 - \delta_n u)^{-1+o(1)} u du.$$

From (13), $\varphi''(t) = \exp(t)\eta'(\exp(t))$ and consequently

$$K_2(t) \sim \frac{1}{\gamma} t \exp(t) \eta'(\exp(t))$$

as $t \rightarrow \infty$. Moreover, η is asymptotically monotone with a constant sign and $|\eta| \in RV_\rho$ imply that $x\eta'(x)/\eta(x) \rightarrow \rho$ as $x \rightarrow \infty$, leading to, as $t \rightarrow \infty$,

$$K_2(t) \sim \frac{\rho}{\gamma} t \eta(\exp(t)).$$

It follows, when $\delta_n \rightarrow \delta \in (0, 1)$,

$$\Delta_n \sim y_n \frac{\delta^2 \rho}{\gamma} \int_0^1 u \eta(e^{y_n(1-\delta_n u)}) du.$$

Since $|\eta| \in RV_\rho$, Potter's bounds (see for instance [6], Proposition B.1.9 (5.)) entail that there exists $0 < \epsilon < |\rho|$ such that

$$(1 - \epsilon) e^{y_n \delta_n (1-u)(\rho-\epsilon)} \leq \frac{|\eta|(e^{y_n(1-\delta_n u)})}{|\eta|(e^{x_n})} \leq (1 + \epsilon) e^{y_n \delta_n (1-u)(\rho+\epsilon)}.$$

Recalling that η is asymptotically monotone with a constant sign yields

$$\Delta_n \sim \frac{\delta^2 \rho}{\gamma} \eta(e^{x_n}) I_n y_n,$$

where $I_n^- \leq I_n \leq I_n^+$ and

$$I_n^- = (1 - \epsilon) \int_0^1 u e^{y_n \delta_n (1-u)(\rho-\epsilon)} du$$

$$I_n^+ = (1 + \epsilon) \int_0^1 u e^{y_n \delta_n (1-u)(\rho+\epsilon)} du.$$

Straightforward calculations show that

$$\int_0^1 u e^{y_n \delta_n (1-u)x} du \sim -\frac{1}{y_n \delta x},$$

as $n \rightarrow \infty$, since $\delta_n \rightarrow \delta$ and thus $y_n \delta_n \rightarrow \infty$. Consequently, $I_n^- \sim (1 - \epsilon)/(y_n \delta (\epsilon - \rho))$, $I_n^+ \sim (1 + \epsilon)/(y_n \delta (-\epsilon - \rho))$ and thus $I_n y_n \in \left[\frac{(1 - \epsilon)}{\delta(\epsilon - \rho)}(1 + o(1)); \frac{(1 + \epsilon)}{\delta(-\epsilon - \rho)}(1 + o(1)) \right]$. Letting $\epsilon \rightarrow 0$ entails $I_n y_n \rightarrow -1/(\rho \delta)$ as $n \rightarrow \infty$ and thus

$$\Delta_n \sim -\frac{\delta}{\gamma} \eta(e^{x_n}).$$

Remarking that $\log q(p_n) = \varphi(y_n) \sim \gamma y_n \sim \frac{\gamma}{1 - \delta} x_n$ and taking account of (11) yield, when $\delta_n \rightarrow \delta \in (0, 1)$,

$$\varepsilon_W(p_n; \alpha_n) = 1 - \exp\left(-\frac{\gamma}{1 - \delta} \Delta_n x_n (1 + o(1))\right).$$

Finally, since $|\eta| \in RV_\rho$, $\rho < 0$, $\Delta_n x_n \sim -\frac{\delta}{\gamma} x_n \eta(e^{x_n}) \rightarrow 0$ as $n \rightarrow \infty$ and thus

$$\varepsilon_W(p_n; \alpha_n) \sim -\frac{\delta}{1 - \delta} x_n \eta(e^{x_n}) \sim -\frac{\delta}{1 - \delta} \log(1/\alpha_n) \eta(1/\alpha_n),$$

which concludes the proof of (i) and (ii). ■

References

- [1] I. Alves, L. de Haan, and C. Neves. A test procedure for detecting super-heavy tails. *Journal of Statistical Planning and Inference*, 139(2):213–227, 2009.
- [2] J. Beirlant, M. Broniatowski, J. Teugels, and P. Vynckier. The mean residual life function at great age: Applications to tail estimation. *Journal of Statistical Planning and Inference*, 45(1-2):21–48, 1995.
- [3] N.H. Bingham, C.M. Goldie, and J.L. Teugels. *Regular Variation*, volume 27 of *Encyclopedia of Mathematics and its application*. Cambridge University Press, 1987.
- [4] L. Breiman, C.J. Stone, and C. Kooperberg. Robust confidence bounds for extreme upper quantiles. *Journal of Statistical Computation and Simulation*, 37:127–149, 1990.
- [5] J. Cohen. Convergence rates for the ultimate and pentultimate approximations in extreme-value theory. *Advances in Applied Probability*, 14(4):833–854, 1982.
- [6] L. de Haan and A. Ferreira. *Extreme value theory: an introduction*. Springer Science & Business Media, 2007.
- [7] C. de Valk. Approximation and estimation of very small probabilities of multivariate extreme events. *Extremes*, 19(4):687–717, 2016.
- [8] C. de Valk. Approximation of high quantiles from intermediate quantiles. *Extremes*, 19(4):661–686, 2016.
- [9] C. de Valk and J.-J. Cai. A high quantile estimator based on the log-generalized Weibull tail limit. *Econometrics and Statistics*, 2017. <https://doi.org/10.1016/j.ecosta.2017.03.001>.
- [10] J. Diebolt, M.-A. El-Aroui, V. Durbec, and B. Villain. Estimation of extreme quantiles: Empirical tools for methods assessment and comparison. *International Journal of Reliability, Quality and Safety Engineering*, 7(01):75–94, 2000.
- [11] J. El Methni, L. Gardes, S. Girard, and A. Guillou. Estimation of extreme quantiles from heavy and light tailed distributions. *Journal of Statistical Planning and Inference*, 142(10):2735–2747, 2012.
- [12] J. Galambos. *The asymptotic theory of extreme order statistics*. R.E. Krieger publishing compagny, 1987.
- [13] L. Gardes and S. Girard. Estimation of the Weibull tail-coefficient with linear combination of upper order statistics. *Journal of Statistical Planning and Inference*, 138(5):1416–1427, 2008.
- [14] L. Gardes and S. Girard. Conditional extremes from heavy-tailed distributions: An application to the estimation of extreme rainfall return levels. *Extremes*, 13(2):177–204, 2010.

- [15] L. Gardes, S. Girard, and A. Guillo. Weibull tail-distributions revisited: a new look at some tail estimators. *Journal of Statistical Planning and Inference*, 141(1):429–444, 2011.
- [16] Y. Goegebeur, J. Beirlant, and T. De Wet. Generalized kernel estimators for the Weibull-tail coefficient. *Communications in Statistics-Theory and Methods*, 39(20):3695–3716, 2010.
- [17] M.I. Gomes. Penultimate limiting forms in extreme value theory. *Annals of the Institute of Statistical Mathematics*, 36(1):71–85, 1984.
- [18] M.I. Gomes and L. de Haan. Approximation by penultimate extreme value distributions. *Extremes*, 2(1):71–85, 1999.
- [19] M.I. Gomes and D.D. Pestana. Nonstandard domains of attraction and rates of convergence. *In: New Perspectives in Theoretical and Applied Statistics*, pages 467–477, 1987.
- [20] B. Hill. A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5):1163–1174, 1975.
- [21] J. Pickands. Statistical inference using extreme order statistics. *The Annals of Statistics*, 3:119–131, 1975.
- [22] R.L. Smith. Estimating tails of probability distributions. *The Annals of Statistics*, 15(3):1174–1207, 1987.
- [23] I. Weissman. Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, 73(364):812–815, 1978.
- [24] R. Worms. Penultimate approximation for the distribution of the excesses. *ESAIM: Probability and Statistics*, 6:21–31, 2002.

Appendix: Auxiliary results

Lemma 3 For all $(a, b, t) \in \mathbb{R}_+^3$, let

$$\Psi_a(t; b) = \int_0^b u^a \exp(-tu) du$$

and remark that $\Phi(\cdot) = \Psi_1(\cdot; 1)$.

(i) $\Psi_a(\cdot; b)$ is continuous, non-increasing on \mathbb{R}_+ , $\Psi_a(0; b) = b^{a+1}/(a+1)$ and $\Psi_a(t; b) \rightarrow 0$ as $t \rightarrow \infty$.

(ii) $\Psi_1(t, b) \sim 1/t^2$ and $\Psi_2(t, b) \sim 1/t^3$ as $t \rightarrow \infty$.

The proof is straightforward.

Lemma 4 Assume $K_1 \in RV_{\theta_1}$.

(i) If $\ell_1 = 0$ then $\theta_1 \leq 0$ and $\ell_2 = 0$. If, moreover, $\theta_1 < 0$, then $-K_2 \in RV_{\theta_1}$ and $K_2(t) \sim (\theta_1 - 1)K_1(t)$ as $t \rightarrow \infty$.

(ii) If $\ell_1 = 1$ then $\theta_1 = 0$ and $\ell_2 = 0$.

(iii) If $0 < \ell_1 < \infty$ and $\ell_1 \neq 1$ then $\theta_1 = 0$, $\ell_2 = \ell_1(\ell_1 - 1) \neq 0$ and $|K_2| \in RV_0$.

(iv) If $\ell_1 = \infty$ then $\theta_1 \geq 0$, $\ell_2 = \infty$, $K_2 \in RV_{2\theta_1}$ and $K_2(t) \sim K_1^2(t)$ as $t \rightarrow \infty$.

Proof. The proof relies on the following four facts: First, for all $x \in \mathbb{R}$,

$$K_2(x) = K_1^2(x) + K_1(x) \left(\frac{xK_1'(x)}{K_1(x)} - 1 \right),$$

or, equivalently,

$$\frac{K_2(x)}{K_1^2(x)} = 1 + \frac{1}{K_1(x)} \left(\frac{xK_1'(x)}{K_1(x)} - 1 \right). \quad (18)$$

Second, $xK_1'(x)/K_1(x) \rightarrow \theta_1$ as $x \rightarrow \infty$. Third, it straightforwardly follows that $\ell_2 = \ell_1(\ell_1 + \theta_1 - 1)$. Finally, for all positive function K , $K(x) \rightarrow c > 0$ as $x \rightarrow \infty$ implies $K \in RV_0$. ■

	$\bar{F}(x)$	θ_1	θ_2	$K_1(x)$	$K_2(x)$	ℓ_1
DA₁(Gumbel)						
Finite endpoint ($\beta > 0$)	$\exp(-(-\log x)^{-\beta})$ $x \in (0, 1)$	$-1/\beta$	$-1/\beta$	$\frac{1}{\beta}x^{-1/\beta}$	$-\frac{1+\beta}{\beta^2}x^{-1/\beta}(1+o(1))$	0
Gamma ($a > 0$)	$\frac{1}{\Gamma(a)} \int_x^\infty t^{a-1}e^{-t} dt$ $x \geq 0$	0	-1	$1+o(1)$	$\frac{1-a}{x}(1+o(1))$	1
DA₂(Gumbel)						
Weibull ($\beta \neq 1$)	$\exp(-x^\beta)$ $x \geq 0$	0	0	$\frac{1}{\beta}$	$\frac{1-\beta}{\beta^2}$	$1/\beta$
Gaussian	$\frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{t^2}{2}\right) dt$	0	0	$\frac{1}{2}+o(1)$	$-\frac{1}{4}+o(1)$	$1/2$
DA₃(Gumbel)						
Log-Weibull ($\beta > 1$)	$\exp(-(\log x)^\beta)$ $x \geq 1$	$1/\beta$	$2/\beta$	$\frac{1}{\beta}x^{1/\beta}$	$\frac{1}{\beta^2}x^{2/\beta}(1+o(1))$	$+\infty$
Lognormal ($\sigma > 0$)	$\frac{1}{\sigma\sqrt{2\pi}} \int_x^\infty \frac{1}{t} \exp\left(-\frac{(\log t)^2}{2\sigma^2}\right) dt$ $x \geq 0$	$1/2$	1	$\frac{\sigma}{\sqrt{2}}x^{1/2}(1+o(1))$	$\frac{\sigma^2}{2}x(1+o(1))$	$+\infty$

Table 1: Examples of distributions in DA(Gumbel).

Distribution	First order approximation of $\varepsilon_{\text{ET}}(p_n; \alpha_n)$
DA₁(Gumbel) Finite endpoint($\beta > 0$) Gamma($a > 0$)	$-\frac{2(1+\beta)}{\beta^2} \frac{(\log \log n)^2}{(\log n)^{2+1/\beta}}$ $2(1-a) \frac{(\log \log n)^2}{(\log n)^3}$
DA₂(Gumbel) Weibull($\beta \neq 1$) Gaussian	$\frac{2(1-\beta)}{\beta^2} \frac{(\log \log n)^2}{(\log n)^2}$ $-\frac{1}{2} \frac{(\log \log n)^2}{(\log n)^2}$
DA₃(Gumbel) Log-Weibull($\beta > 1$) Lognormal	$\frac{2}{\beta^2} \frac{(\log \log n)^2}{(\log n)^{2-2/\beta}}$ $\sigma^2 \frac{(\log \log n)^2}{\log n}$

Table 2: First order approximations of $\varepsilon_{\text{ET}}(p_n; \alpha_n)$ with $p_n = 1/(n \log n)$ and $\alpha_n = (\log n)/n$ associated with the distributions described in Table 1.

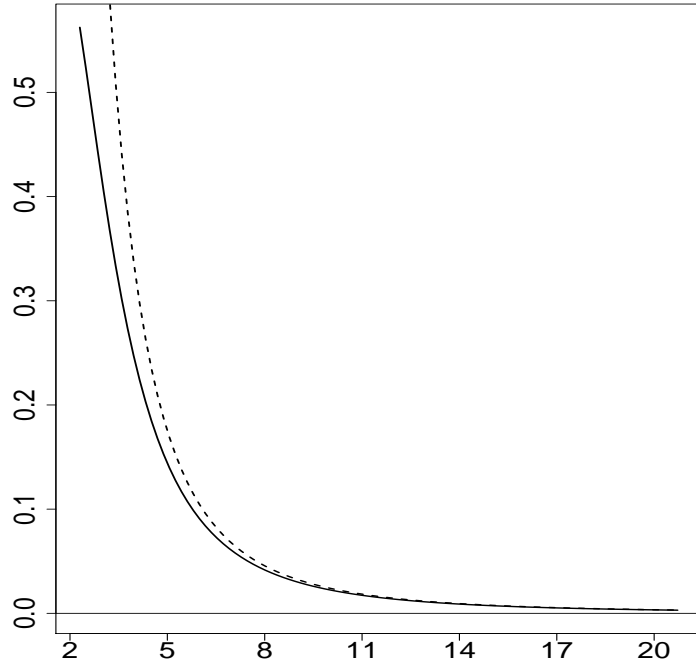
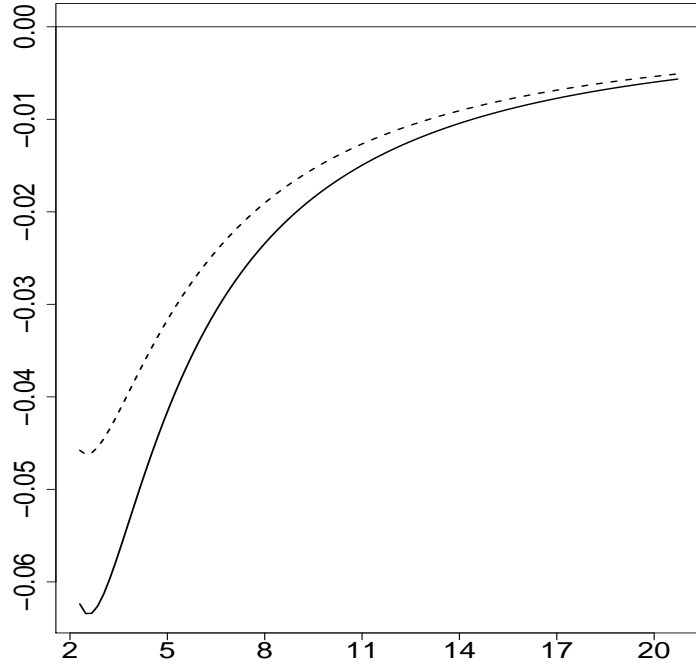


Figure 1: Extrapolation error in $DA_1(\text{Gumbel})$. Vertically: Extrapolation error $\varepsilon_{\text{ET}}(p_n; \alpha_n)$ (solid line) and its first order approximation $\frac{1}{2}\eta_n^2 K_2(\log n)$ (dashed line) provided by Theorem 2(i)-(a). Horizontally: $\log n$. Top: Finite endpoint($\beta = 5$) distribution, bottom: Gamma($a = 0.1$) distribution.

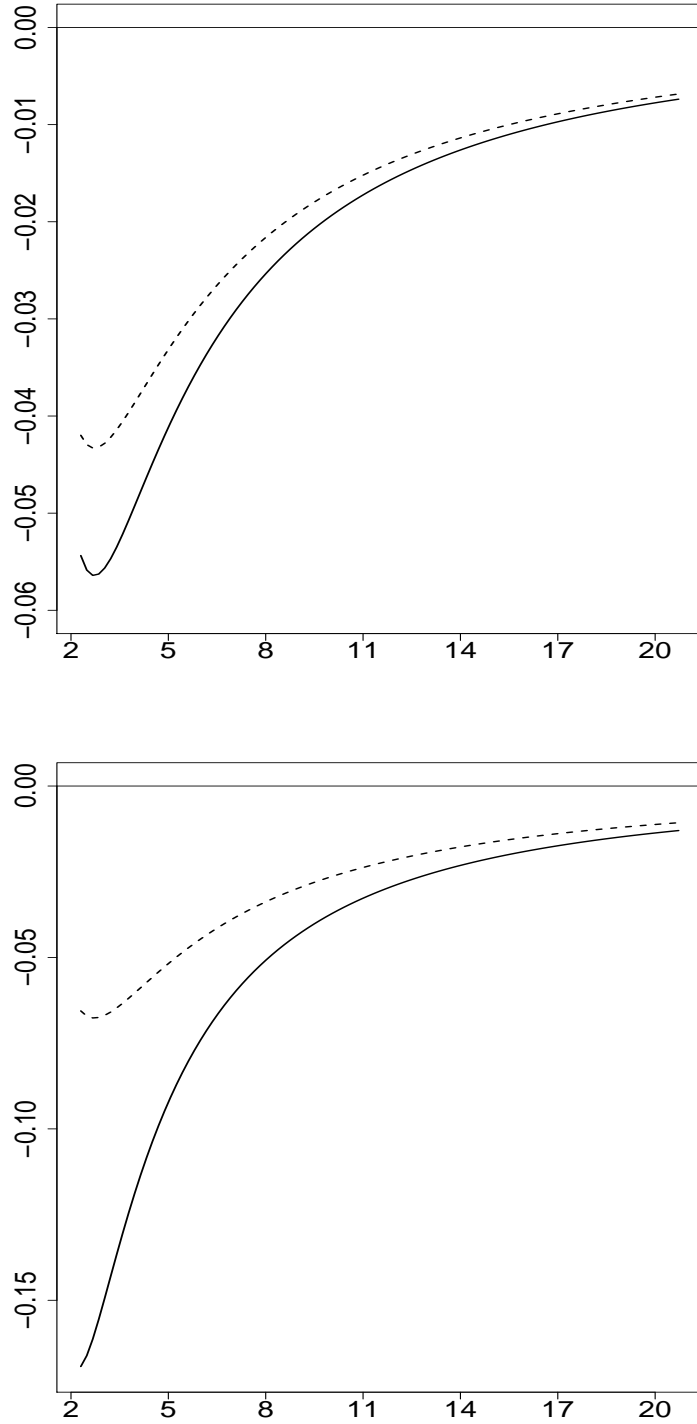


Figure 2: Extrapolation error in $DA_2(\text{Gumbel})$. Vertically: Extrapolation error $\varepsilon_{ET}(p_n; \alpha_n)$ (solid line) and its first order approximation $\frac{\ell_1(\ell_1-1)}{2}\eta_n^2$ (dashed line) provided by Theorem 2(ii)-(a). Horizontally: $\log n$. Top: Weibull($\beta = 5$) distribution, bottom: Gaussian distribution.

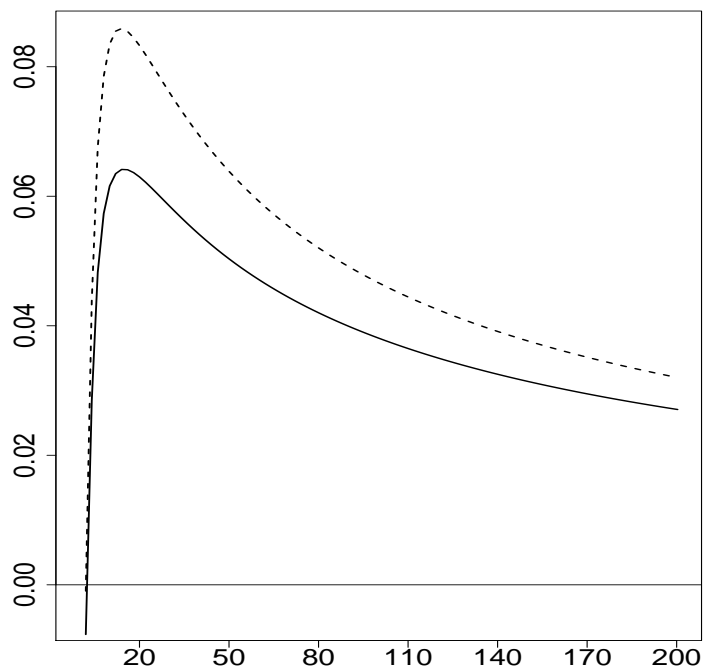
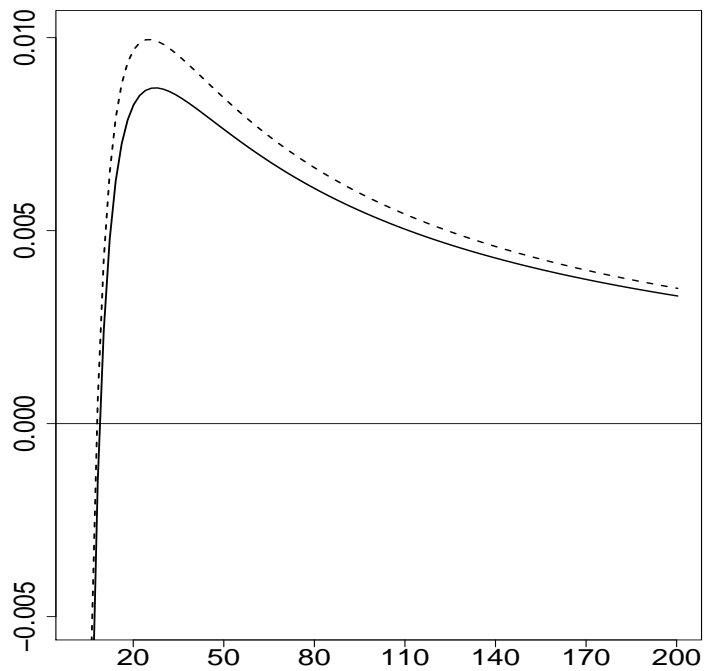


Figure 3: Extrapolation error in $DA_3(\text{Gumbel})$. Vertically: Extrapolation error $\varepsilon_{ET}(p_n; \alpha_n)$ (solid line) and its first order approximation $\frac{1}{2}\eta_n^2 K_2(\log n)$ (dashed line) provided by Theorem 2(iii)-(a). Horizontally: $\log n$. Top: log-Weibull($\beta = 3$) distribution, bottom: lognormal($\sigma = 0.5$) distribution.