Comparison of Weibull tail-coefficient estimators

Laurent Gardes, Stephane Girard

Comparison of Weibull tail-coefficient estimators

Laurent Gardes[(1)] & Stéphane Girard[(2,⋆)]

[(1)]Université Grenoble 2, LabSAD, BP 47, 38040 Grenoble cedex 9, France,
Laurent.Gardes@upmf-grenoble.fr
[(2,⋆)]Université Grenoble 1, LMC-IMAG, BP 53, 38041 Grenoble cedex 9, France,
Stephane.Girard@imag.fr (corresponding author)

Abstract

We address the problem of estimating the Weibull tail-coefficient which is the regular variation exponent of the inverse failure rate function. We propose a family of estimators of this coefficient and an associate extreme quantile estimator. Their asymptotic normality are established and their asymptotic mean-square errors are compared. The results are illustrated on some finite sample situations.

Keywords: Weibull tail-coefficient, extreme quantile, extreme value theory, asymptotic normality.

AMS 2000 subject classification: 62G32, 62F12, 62G30.

1 Introduction

Let $X_1, X_2, \dots, X_n$ be a sequence of independent and identically distributed random variables with cumulative distribution function F. We denote by $X_{1,n} \leq \dots \leq X_{n,n}$ their associated order statistics. We address the problem of estimating the Weibull tail-coefficient $\theta > 0$ defined when the distribution tail satisfies

(A.1) $1 - F(x) = \exp(-H(x))$, $x \geq x_0 \geq 0$, $H^{\leftarrow}(t) = \inf\{x,\ H(x) \geq t\} = t^{\theta}\ell(t)$,

where ℓ is a slowly varying function *i.e.*

$$\ell(\lambda x)/\ell(x) \to 1 \text{ as } x \to \infty \text{ for all } \lambda > 0.$$

The inverse cumulative hazard function $H^{\leftarrow}$ is said to be regularly varying at infinity with index θ and this property is denoted by $H^{\leftarrow} \in \mathcal{R}_{\theta}$, see [7] for more details on this topic. As a comparison, Pareto type distributions satisfy $(1/(1-F))^{\leftarrow} \in \mathcal{R}_\gamma$, and $\gamma > 0$ is the so-called extreme value index. Weibull tail-distributions include for instance Gamma, Gaussian and, of course, Weibull distributions.

Let (k_n) be a sequence of integers such that $1 \leq k_n < n$ and (T_n) be a positive sequence. We examine the asymptotic behavior of the following family of estimators of θ :

$$\hat{\theta}_n = \frac{1}{T_n}\frac{1}{k_n}\sum_{i=1}^{k_n}(\log(X_{n-i+1,n}) - \log(X_{n-k_n+1,n})). \tag{1}$$

Following the ideas of [10], an estimator of the extreme quantile x_{p_n} can be deduced from (1) by:

$$\hat{x}_{p_n} = X_{n-k_n+1,n}\left(\frac{\log(1/p_n)}{\log(n/k_n)}\right)^{\hat{\theta}_n} =: X_{n-k_n+1,n}\tau_n^{\hat{\theta}_n}. \tag{2}$$

Recall that an extreme quantile x_{p_n} of order p_n is defined by the equation

$$1 - F(x_{p_n}) = p_n, \text{ with } 0 < p_n < 1/n.$$

The condition $p_n < 1/n$ is very important in this context. It usually implies that x_{p_n} is larger than the maximum observation of the sample. This necessity to extrapolate sample results to areas where no data are observed occurs in reliability [8], hydrology [21], finance [9],... We establish in Section 2 the asymptotic normality of $\hat{\theta}_n$ and $\hat{x}_{p_n}$. The asymptotic mean-square error of some particular members of (1) are compared in Section 3. In particular, it is shown that family (1) encompasses the estimator introduced in [12] and denoted by $\hat{\theta}_n^{(2)}$ in the sequel. In this paper, the asymptotic normality of $\hat{\theta}_n^{(2)}$ is obtained under weaker conditions. Furthermore, we show that other members of family (1) should be preferred in some typical situations. We also quote some other estimators of θ which do not belong to family (1): [4, 3, 6, 19]. We refer to [12] for a comparison with $\hat{\theta}_n^{(2)}$. The asymptotic results are illustrated in Section 4 on finite sample situations. Proofs are postponed to Section 5.

2 Asymptotic normality

To establish the asymptotic normality of $\hat{\theta}_n$, we need a second-order condition on ℓ:

(A.2) There exist $\rho \leq 0$ and $b(x) \to 0$ such that uniformly locally on $\lambda \geq 1$

$$\log\left(\frac{\ell(\lambda x)}{\ell(x)}\right) \sim b(x)K_\rho(\lambda), \text{ when } x \to \infty,$$

with $K_\rho(\lambda) = \int_1^\lambda u^{\rho-1}du$.

It can be shown [11] that necessarily $|b| \in \mathcal{R}_\rho$. The second order parameter $\rho \leq 0$ tunes the rate of convergence of $\ell(\lambda x)/\ell(x)$ 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. It is used in [18, 17, 5] to prove the asymptotic normality of estimators of the extreme value index γ. In regular case, as noted in [13], one can choose $b(x) = x\ell'(x)/\ell(x)$ leading to

$$b(x) = \frac{xe^{-x}}{F^{-1}(1-e^{-x})f(F^{-1}(1-e^{-x}))} - \theta, \tag{3}$$

where f is the density function associated to F.
Let us introduce the following functions : for $t > 0$ and $\rho \leq 0$,

$$\begin{aligned}\mu_\rho(t) &= \int_0^\infty K_\rho\left(1+\frac{x}{t}\right)\mathrm{e}^{-x}dx \\ \sigma_\rho^2(t) &= \int_0^\infty K_\rho^2\left(1+\frac{x}{t}\right)\mathrm{e}^{-x}dx - \mu_\rho^2(t),\end{aligned}$$

and let $a_n = \mu_0(\log(n/k_n))/T_n - 1$. As a preliminary result, we propose an asymptotic expansion of $(\hat{\theta}_n - \theta)$:

Proposition 1 *Suppose* **(A.1)** *and* **(A.2)** *hold. If* $k_n \to \infty$*,* $k_n/n \to 0$*,* $T_n \log(n/k_n) \to 1$ *and* $k_n^{1/2} b(\log(n/k_n)) \to \lambda \in \mathbb{R}$ *then,*

$$\begin{aligned}k_n^{1/2}(\hat{\theta}_n - \theta) &= \theta\xi_{n,1} + \theta\mu_0(\log(n/k_n))\xi_{n,2} + k_n^{1/2}\theta a_n \\ &+ k_n^{1/2} b(\log(n/k_n))(1 + o_{\mathrm{P}}(1)),\end{aligned}$$

where $\xi_{n,1}$ *and* $\xi_{n,2}$ *converge in distribution to a standard normal distribution.*

Similar distributional representations exist for various estimators of the extreme value index γ. They are used in [16] to compare the asymptotic properties of several tail index estimators. In [15], a bootstrap selection of k_n is derived from such a representation. It is also possible to derive bias reduction method as in [14]. The asymptotic normality of $\hat{\theta}_n$ is a straightforward consequence of Proposition 1.

Theorem 1 *Suppose* **(A.1)** *and* **(A.2)** *hold. If* $k_n \to \infty$*,* $k_n/n \to 0$*,* $T_n \log(n/k_n) \to 1$ *and* $k_n^{1/2} b(\log(n/k_n)) \to \lambda \in \mathbb{R}$ *then,*

$$k_n^{1/2}(\hat{\theta}_n - \theta - b(\log(n/k_n)) - \theta a_n) \xrightarrow{d} \mathcal{N}(0, \theta^2).$$

Theorem 1 implies that the Asymptotic Mean Square Error (AMSE) of $\hat{\theta}_n$ is given by :

$$AMSE(\hat{\theta}_n) = (\theta a_n + b(\log(n/k_n)))^2 + \frac{\theta^2}{k_n}. \tag{4}$$

It appears that all estimators of family (1) share the same variance. The bias depends on two terms $b(\log(n/k_n))$ and θa_n. A good choice of T_n (depending on the function b) could lead to a sequence a_n cancelling the bias. Of course, in the general case, the function b is unknown making difficult the choice of a "universal" sequence T_n. This is discussed in the next section.
Clearly, the best rate of convergence in Theorem 1 is obtained by choosing $\lambda \neq 0$. In this case, the expression of the intermediate sequence (k_n) is known.

Proposition 2 *If* $k_n \to \infty$, $k_n/n \to 0$ *and* $k_n^{1/2} b(\log(n/k_n)) \to \lambda \neq 0$,

$$k_n \sim \left(\frac{\lambda}{b(\log(n))}\right)^2 = \lambda^2 (\log(n))^{-2\rho} L(\log(n)),$$

where L *is a slowly varying function.*

The "optimal" rate of convergence is thus of order $(\log(n))^{-\rho}$, which is entirely determined by the second order parameter ρ: small values of $|\rho|$ yield slow convergence. The asymptotic normality of the extreme quantile estimator (2) can be deduced from Theorem 1:

Theorem 2 *Suppose* **(A.1)** *and* **(A.2)** *hold. If moreover,* $k_n \to \infty$, $k_n/n \to 0$, $T_n \log(n/k_n) \to 1$, $k_n^{1/2} b(\log(n/k_n)) \to 0$ *and*

$$1 \leq \liminf \tau_n \leq \limsup \tau_n < \infty \tag{5}$$

then,

$$\frac{k_n^{1/2}}{\log \tau_n} \left(\frac{\hat{x}_{p_n}}{x_{p_n}} - \tau_n^{\theta a_n} \right) \xrightarrow{d} \mathcal{N}(0, \theta^2).$$

3 Comparison of some estimators

First, we propose some choices of the sequence (T_n) leading to different estimators of the Weibull tail-coefficient. Their asymptotic distributions are provided, and their AMSE are compared.

3.1 Some examples of estimators

– The natural choice is clearly to take

$$T_n = T_n^{(1)} =: \mu_0(\log(n/k_n)),$$

in order to cancel the bias term a_n. This choice leads to a new estimator of θ defined by :

$$\hat{\theta}_n^{(1)} = \frac{1}{\mu_0(\log(n/k_n))} \frac{1}{k_n} \sum_{i=1}^{k_n} (\log(X_{n-i+1,n}) - \log(X_{n-k_n+1,n})).$$

Remarking that

$$\mu_\rho(t) = \mathrm{e}^t \int_1^\infty \mathrm{e}^{-tu} u^{\rho-1} du$$

provides a simple computation method for $\mu_0(\log(n/k_n))$ using the Exponential Integral (EI), see for instance [1], Chapter 5, pages 225–233.

– Girard [12] proposes the following estimator of the Weibull tail-coefficient:

$$\hat{\theta}_n^{(2)} = \sum_{i=1}^{k_n} (\log(X_{n-i+1,n}) - \log(X_{n-k_n+1,n})) \Bigg/ \sum_{i=1}^{k_n} (\log_2(n/i) - \log_2(n/k_n)),$$

where $\log_2(x) = \log(\log(x))$, $x > 1$. Here, we have

$$T_n = T_n^{(2)} =: \frac{1}{k_n} \sum_{i=1}^{k_n} \log\left(1 - \frac{\log(i/k_n)}{\log(n/k_n)}\right).$$

It is interesting to remark that $T_n^{(2)}$ is a Riemann's sum approximation of $\mu_0(\log(n/k_n))$ since an integration by parts yields:

$$\mu_0(t) = \int_0^1 \log\left(1 - \frac{\log(x)}{t}\right) dx.$$

– Finally, choosing T_n as the asymptotic equivalent of $\mu_0(\log(n/k_n))$,

$$T_n = T_n^{(3)} =: 1/\log(n/k_n)$$

leads to the estimator :

$$\hat{\theta}_n^{(3)} = \frac{\log(n/k_n)}{k_n} \sum_{i=1}^{k_n} (\log(X_{n-i+1,n}) - \log(X_{n-k_n+1,n})).$$

For $i = 1, 2, 3$, let us denote by $\hat{x}_{p_n}^{(i)}$ the extreme quantile estimator built on $\hat{\theta}_n^{(i)}$ by (2). Asymptotic normality of these estimators is derived from Theorem 1 and Theorem 2. To this end, we introduce the following conditions:

(C.1) $k_n/n \to 0$,

(C.2) $\log(k_n)/\log(n) \to 0$,

(C.3) $k_n/n \to 0$ and $k_n^{1/2}/\log(n/k_n) \to 0$.

Our result is the following:

Corollary 1 *Suppose* **(A.1)** *and* **(A.2)** *hold,* $k_n \to \infty$ *and* $k_n^{1/2} b(\log(n/k_n)) \to 0$. *For* $i = 1, 2, 3$*:*

i) If **(C.i)** *hold then*

$$k_n^{1/2}(\hat{\theta}_n^{(i)} - \theta) \xrightarrow{d} \mathcal{N}(0, \theta^2).$$

ii) If **(C.i)** *and (5) hold, then*

$$\frac{k_n^{1/2}}{\log \tau_n} \left(\frac{\hat{x}_{p_n}^{(i)}}{x_{p_n}} - 1 \right) \xrightarrow{d} \mathcal{N}(0, \theta^2).$$

In view of this corollary, the asymptotic normality of $\hat{\theta}_n^{(1)}$ is obtained under weaker conditions than $\hat{\theta}_n^{(2)}$ and $\hat{\theta}_n^{(3)}$, since **(C.2)** implies **(C.1)**. Let us also highlight that the asymptotic distribution of $\hat{\theta}_n^{(2)}$ is obtained under less assumptions than in [12], Theorem 2, the condition $k_n^{1/2}/\log(n/k_n) \to 0$ being not necessary here. Finally, note that, if b is not ultimately zero, condition $k_n^{1/2} b(\log(n/k_n)) \to 0$ implies **(C.2)** (see Lemma 1).

3.2 Comparison of the AMSE of the estimators

We use the expression of the AMSE given in (4) to compare the estimators proposed previously.

Theorem 3 *Suppose* **(A.1)** *and* **(A.2)** *hold,* $k_n \to \infty$, $\log(k_n)/\log(n) \to 0$ *and* $k_n^{1/2} b(\log(n/k_n)) \to \lambda \in \mathbb{R}$. *Several situations are possible:*

i) b is ultimately non-positive. Let us introduce $\alpha = -4 \lim_{n\to\infty} b(\log n) \dfrac{k_n}{\log k_n} \in [0, +\infty]$.
If $\alpha > \theta$, *then, for n large enough,*

$$AMSE(\hat{\theta}_n^{(2)}) < AMSE(\hat{\theta}_n^{(1)}) < AMSE(\hat{\theta}_n^{(3)}).$$

If $\alpha < \theta$, *then, for n large enough,*

$$AMSE(\hat{\theta}_n^{(1)}) < \min(AMSE(\hat{\theta}_n^{(2)}), AMSE(\hat{\theta}_n^{(3)})).$$

ii) b is ultimately non-negative. Let us introduce $\beta = 2 \lim_{x\to\infty} x b(x) \in [0, +\infty]$.

If $\beta > \theta$ *then, for n large enough,*

$$AMSE(\hat{\theta}_n^{(3)}) < AMSE(\hat{\theta}_n^{(1)}) < AMSE(\hat{\theta}_n^{(2)}).$$

If $\beta < \theta$ *then, for n large enough,*

$$AMSE(\hat{\theta}_n^{(1)}) < \min(AMSE(\hat{\theta}_n^{(2)}), AMSE(\hat{\theta}_n^{(3)})).$$

It appears that, when b is ultimately non-negative (case ii)), the conclusion does not depend on the sequence (k_n). The relative performances of the estimators is entirely determined by the nature of the distribution: $\hat{\theta}_n^{(1)}$ has the best behavior, in terms of AMSE, for distributions close to the Weibull distribution (small b and thus, small β). At the opposite, $\hat{\theta}_n^{(3)}$ should be preferred for distributions far from the Weibull distribution.
The case when b is ultimately non-positive (case i)) is different. The value of α depends on k_n, and thus, for any distribution, one can obtain $\alpha = 0$ by choosing small values of k_n(for instance $k_n = -1/b(\log n)$) as well as $\alpha = +\infty$ by choosing large values of k_n (for instance $k_n = (1/b(\log n))^2$ as in Proposition 2).

4 Numerical experiments

4.1 Examples of Weibull tail-distributions

Let us give some examples of distributions satisfying assumptions **(A.1)** and **(A.2)**.

Absolute Gaussian distribution $|\mathcal{N}(\mu, \sigma^2)|$, $\sigma > 0$. From [9], Table 3.4.4, we have $H^{\leftarrow}(x) = x^\theta \ell(x)$, where $\theta = 1/2$ and an asymptotic expansion of the slowly varying function is given by:

$$\ell(x) = 2^{1/2}\sigma - \frac{\sigma}{2^{3/2}} \frac{\log x}{x} + O(1/x).$$

Therefore $\rho = -1$ and $b(x) = \log(x)/(4x) + O(1/x)$. b is ultimately positive, which corresponds to case ii) of Theorem 3 with $\beta = +\infty$. Therefore, one always has, for n large enough:

$$AMSE(\hat{\theta}_n^{(3)}) < AMSE(\hat{\theta}_n^{(1)}) < AMSE(\hat{\theta}_n^{(2)}). \tag{6}$$

Gamma distribution $\Gamma(a, \lambda)$, $a, \lambda > 0$. We use the following parameterization of the density

$$f(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} \exp(-\lambda x).$$

From [9], Table 3.4.4, we obtain $H^{\leftarrow}(x) = x^{\theta}\ell(x)$ with $\theta = 1$ and

$$\ell(x) = \frac{1}{\lambda} + \frac{a-1}{\lambda}\frac{\log x}{x} + O(1/x).$$

We thus have $\rho = -1$ and $b(x) = (1-a)\log(x)/x + O(1/x)$. If $a > 1$, b is ultimately negative, corresponding to case i) of Theorem 3. The conclusion depends on the value of k_n as explained in the preceding section. If $a < 1$, b is ultimately positive, corresponding to case ii) of Theorem 3 with $\beta = +\infty$. Therefore, we are in situation (6).

Weibull distribution $\mathcal{W}(a, \lambda)$, $a, \lambda > 0$. The inverse failure rate function is $H^{\leftarrow}(x) = \lambda x^{1/a}$, and then $\theta = 1/a$, $\ell(x) = \lambda$ for all $x > 0$. Therefore $b(x) = 0$ and we use the usual convention $\rho = -\infty$. One may apply either i) or ii) of Theorem 3 with $\alpha = \beta = 0$ to get for n large enough,

$$AMSE(\hat{\theta}_n^{(1)}) < \min(AMSE(\hat{\theta}_n^{(2)}), AMSE(\hat{\theta}_n^{(3)})). \tag{7}$$

4.2 Numerical results

The finite sample performance of the estimators $\hat{\theta}_n^{(1)}$, $\hat{\theta}_n^{(2)}$ and $\hat{\theta}_n^{(3)}$ are investigated on 5 different distributions: $\Gamma(0.5, 1)$, $\Gamma(1.5, 1)$, $|\mathcal{N}(0, 1)|$, $\mathcal{W}(2.5, 2.5)$ and $\mathcal{W}(0.4, 0.4)$. In each case, $N = 200$ samples $(\mathcal{X}_{n,i})_{i=1,\dots,N}$ of size $n = 500$ were simulated. On each sample $(\mathcal{X}_{n,i})$, the estimates $\hat{\theta}_{n,i}^{(1)}(k)$, $\hat{\theta}_{n,i}^{(2)}(k)$ and $\hat{\theta}_{n,i}^{(3)}(k)$ are computed for $k = 2, \dots, 150$. Finally, the associated Mean Square Error (MSE) plots are built by plotting the points

$$\left(k, \frac{1}{N}\sum_{i=1}^{N}\left(\hat{\theta}_{n,i}^{(j)}(k) - \theta\right)^2\right) \quad j = 1,\ 2,\ 3.$$

They are compared to the AMSE plots (see (4) for the definition of the AMSE):

$$\left(k, (\theta a_n^{(j)} + b(\log(n/k)))^2 + \frac{\theta^2}{k}\right) \quad j = 1,\ 2,\ 3,$$

and where b is given by (3). It appears on Figure 1 – Figure 5 that, for all the above mentioned distributions, the MSE and AMSE have a similar qualitative behavior. Figure 1 and Figure 2 illustrate situation (6) corresponding to ultimately positive bias functions. The case of an ultimately negative bias function is presented on Figure 3 with the $\Gamma(1.5, 1)$ distribution. It clearly appears that the MSE associated to $\hat{\theta}_n^{(3)}$ is the largest. For small values of k, one has $MSE(\hat{\theta}_n^{(1)}) < MSE(\hat{\theta}_n^{(2)})$ and $MSE(\hat{\theta}_n^{(1)}) > MSE(\hat{\theta}_n^{(2)})$ for large value of k. This phenomenon is the illustration of the asymptotic result presented in Theorem 3i). Finally, Figure 4 and Figure 5 illustrate situation (7) of asymptotically null bias functions. Note that, the MSE of $\hat{\theta}_n^{(1)}$ and $\hat{\theta}_n^{(2)}$ are very similar. As a conclusion, it appears that, in all situations, $\hat{\theta}_n^{(1)}$ and $\hat{\theta}_n^{(2)}$ share a similar behavior, with a small advantage to $\hat{\theta}_n^{(1)}$. They provide good results for null and negative bias functions. At the opposite, $\hat{\theta}_n^{(3)}$ should be preferred for positive bias functions.

5 Proofs

For the sake of simplicity, in the following, we note k for k_n. We first give some preliminary lemmas. Their proofs are postponed to the appendix.

5.1 Preliminary lemmas

We first quote a technical lemma.

Lemma 1 *Suppose that b is ultimately non-zero. If $k \to \infty$, $k/n \to 0$ and $k^{1/2}b(\log(n/k)) \to \lambda \in \mathbb{R}$, then $\log(k)/\log(n) \to 0$.*

The following two lemmas are of analytical nature. They provide first-order expansions which will reveal useful in the sequel.

Lemma 2 *For all $\rho \leq 0$ and $q \in \mathbb{N}^*$, we have*

$$\int_0^\infty K_\rho^q \left(1 + \frac{x}{t}\right) \mathrm{e}^{-x} dx \sim \frac{q!}{t^q} \text{ as } t \to \infty.$$

Let $a_n^{(i)} = \mu_0(\log(n/k_n))/T_n^{(i)} - 1$, for $i = 1, 2, 3$.

Lemma 3 *Suppose $k \to \infty$ and $k/n \to 0$.*

i) $T_n^{(1)} \log(n/k) \to 1$ *and* $a_n^{(1)} = 0$.

ii) $T_n^{(2)} \log(n/k) \to 1$. *If moreover* $\log(k)/\log(n) \to 0$ *then* $a_n^{(2)} \sim \log(k)/(2k)$.

iii) $T_n^{(3)} \log(n/k) = 1$ *and* $a_n^{(3)} \sim -1/\log(n/k)$.

The next lemma presents an expansion of $\hat{\theta}_n$.

Lemma 4 *Suppose $k \to \infty$ and $k/n \to 0$. Under* **(A.1)** *and* **(A.2)**, *the following expansions hold:*

$$\hat{\theta}_n = \frac{1}{T_n} \left(\theta U_n^{(0)} + b(\log(n/k)) U_n^{(\rho)} (1 + o_{\mathrm{P}}(1)) \right),$$

where

$$U_n^{(\rho)} = \frac{1}{k} \sum_{i=1}^{k-1} K_\rho \left(1 + \frac{F_i}{E_{n-k+1,n}}\right), \ \rho \leq 0$$

and where $E_{n-k+1,n}$ is the $(n-k+1)$th order statistics associated to n independent standard exponential variables and $\{F_1, \dots, F_{k-1}\}$ are independent standard exponential variables and independent from $E_{n-k+1,n}$.

The next two lemmas provide the key results for establishing the asymptotic distribution of $\hat{\theta}_n$. Their describe they asymptotic behavior of the random terms appearing in Lemma 4.

Lemma 5 *Suppose $k \to \infty$ and $k/n \to 0$. Then, for all $\rho \le 0$,*

$$\mu_\rho(E_{n-k+1,n}) \overset{P}{\sim} \sigma_\rho(E_{n-k+1,n}) \overset{P}{\sim} \frac{1}{\log(n/k)}.$$

Lemma 6 *Suppose $k \to \infty$ and $k/n \to 0$. Then, for all $\rho \le 0$,*

$$\frac{k^{1/2}}{\sigma_\rho(E_{n-k+1,n})}(U_n^{(\rho)} - \mu_\rho(E_{n-k+1,n})) \xrightarrow{d} \mathcal{N}(0,1).$$

5.2 Proofs of the main results

Proof of Proposition 1 – Lemma 6 states that for $\rho \le 0$,

$$\frac{k^{1/2}}{\sigma_\rho(E_{n-k+1,n})}(U_n^{(\rho)} - \mu_\rho(E_{n-k+1,n})) = \xi_n(\rho),$$

where $\xi_n(\rho) \xrightarrow{d} \mathcal{N}(0,1)$ for $\rho \le 0$. Then, by Lemma 4

$$\begin{aligned} k^{1/2}(\hat{\theta}_n - \theta) &= \theta \frac{\sigma_0(E_{n-k+1,n})}{T_n}\xi_n(0) + k^{1/2}\theta\left(\frac{\mu_0(E_{n-k+1,n})}{T_n} - 1\right) \\ &+ k^{1/2} b(\log(n/k))\left(\frac{\sigma_\rho(E_{n-k+1,n})}{T_n}\frac{\xi_n(\rho)}{k^{1/2}} + \frac{\mu_\rho(E_{n-k+1,n})}{T_n}\right)(1 + o_{\mathrm{P}}(1)). \end{aligned}$$

Since $T_n \sim 1/\log(n/k)$ and from Lemma 5, we have

$$k^{1/2}(\hat{\theta}_n - \theta) = \theta\xi_{n,1} + k^{1/2}\theta\left(\frac{\mu_0(E_{n-k+1,n})}{T_n} - 1\right) + k^{1/2} b(\log(n/k))(1 + o_{\mathrm{P}}(1)), \tag{8}$$

where $\xi_{n,1} \xrightarrow{d} \mathcal{N}(0,1)$. Moreover, a first-order expansion of μ_0 yields

$$\frac{\mu_0(E_{n-k+1,n})}{\mu_0(\log(n/k))} = 1 + (E_{n-k+1,n} - \log(n/k))\frac{\mu_0^{(1)}(\eta_n)}{\mu_0(\log(n/k))},$$

where $\eta_n \in]\min(E_{n-k+1,n}, \log(n/k)), \max(E_{n-k+1,n}, \log(n/k))[$ and

$$\mu_0^{(1)}(t) = \frac{d}{dt}\int_0^\infty \log\left(1 + \frac{x}{t}\right) \mathrm{e}^{-x} dx =: \frac{d}{dt}\int_0^\infty f(x,t)dx.$$

Since for $t \ge T > 0$, $f(.,t)$ is integrable, continuous and

$$\left|\frac{\partial f(x,t)}{\partial t}\right| = \frac{x}{t^2}\left(1 + \frac{x}{t}\right)^{-1} \mathrm{e}^{-x} \le x\frac{\mathrm{e}^{-x}}{T^2},$$

we have that

$$\mu_0^{(1)}(t) = -\int_0^\infty \frac{x}{t^2}\left(1 + \frac{x}{t}\right)^{-1} \mathrm{e}^{-x} dx.$$

Then, Lebesgue Theorem implies that $\mu_0^{(1)}(t) \sim -1/t^2$ as $t \to \infty$. Therefore, $\mu_0^{(1)}$ is regularly varying at infinity and thus

$$\frac{\mu_0^{(1)}(\eta_n)}{\mu_0(\log(n/k))} \overset{P}{\sim} \frac{\mu_0^{(1)}(\log(n/k))}{\mu_0(\log(n/k))} \sim -\frac{1}{\log(n/k)}.$$

Since $k^{1/2}(E_{n-k+1,n} - \log(n/k)) \xrightarrow{d} \mathcal{N}(0,1)$ (see [12], Lemma 1), we have

$$\frac{\mu_0(E_{n-k+1,n})}{\mu_0(\log(n/k))} = 1 - \frac{k^{-1/2}}{\log(n/k)}\xi_{n,2}, \tag{9}$$

where $\xi_{n,2} \xrightarrow{d} \mathcal{N}(0,1)$. Collecting (8), (9) and taking into account that $T_n \log(n/k) \to 1$ concludes the proof. ■

Proof of Proposition 2 − Lemma 1 entails $\log(n/k) \sim \log(n)$. Since $|b|$ is a regularly varying function, $b(\log(n/k)) \sim b(\log(n))$ and thus, $k^{1/2} \sim \lambda/b(\log(n))$. ■

Proof of Theorem 2 − The asymptotic normality of $\hat{x}_{p_n}$ can be deduced from the asymptotic normality of $\hat{\theta}_n$ using Theorem 2.3 of [10]. We are in the situation, denoted by **(S.2)** in the above mentioned paper, where the limit distribution of $\hat{x}_{p_n}/x_{p_n}$ is driven by $\hat{\theta}_n$. Following, the notations of [10], we denote by $\alpha_n = k_n^{1/2}$ the asymptotic rate of convergence of $\hat{\theta}_n$, by $\beta_n = \theta a_n$ its asymptotic bias, and by $\mathcal{L} = \mathcal{N}(0, \theta^2)$ its asymptotic distribution. It suffices to verify that

$$\log(\tau_n)\log(n/k) \to \infty. \tag{10}$$

To this end, note that conditions (5) and $p_n < 1/n$ imply that there exists $0 < c < 1$ such that

$$\log(\tau_n) > c(\tau_n - 1) > c\left(\frac{\log(n)}{\log(n/k)} - 1\right) = c\frac{\log(k)}{\log(n/k)},$$

which proves (10). We thus have

$$\frac{k^{1/2}}{\log \tau_n}\tau_n^{-\theta a_n}\left(\frac{\hat{x}_{p_n}}{x_{p_n}} - \tau_n^{\theta a_n}\right) \xrightarrow{d} \mathcal{N}(0, \theta^2).$$

Now, remarking that, from Lemma 2, $\mu_0(\log(n/k)) \sim 1/\log(n/k) \sim T_n$, and thus $a_n \to 0$ gives the result. ■

Proof of Corollary 1 − Lemma 3 shows that the assumptions of Theorem 1 and Theorem 2 are verified and that, for $i = 1, 2, 3$, $k^{1/2}a_n^{(i)} \to 0$. ■

Proof of Theorem 3 −

i) First, from (4) and Lemma 3 iii), since b is ultimately non-positive,

$$AMSE(\hat{\theta}_n^{(1)}) - AMSE(\hat{\theta}_n^{(3)}) = -\theta(a_n^{(3)})^2\left(\theta + 2\frac{b(\log(n/k))}{a_n^{(3)}}\right) < 0. \tag{11}$$

Second, from (4),

$$AMSE(\hat{\theta}_n^{(2)}) - AMSE(\hat{\theta}_n^{(1)}) = \theta(a_n^{(2)})^2\left(\theta + 2\frac{b(\log(n/k))}{a_n^{(2)}}\right). \tag{12}$$

If b is ultimately non-zero, Lemma 1 entails that $\log(n/k) \sim \log(n)$ and consequently, since $|b|$ is regularly varying, $b(\log(n/k)) \sim b(\log(n))$. Thus, from Lemma 3 ii),

$$2\frac{b(\log(n/k))}{a_n^{(2)}} \sim 4b(\log n)\frac{k}{\log(k)} \to -\alpha. \tag{13}$$

Collecting (11)–(13) concludes the proof of i).

ii) First, (12) and Lemma 3 ii) yields

$$AMSE(\hat{\theta}_n^{(2)}) - AMSE(\hat{\theta}_n^{(1)}) > 0, \tag{14}$$

since b is ultimately non-negative. Second, if b is ultimately non-zero, Lemma 1 entails that $\log(n/k) \sim \log(n)$ and consequently, since $|b|$ is regularly varying, $b(\log(n/k)) \sim b(\log(n))$. Thus, observe that in (11),

$$2\frac{b(\log(n/k))}{a_n^{(3)}} \sim -2b(\log n)(\log n) \to -\beta. \tag{15}$$

Collecting (11), (14) and (15) concludes the proof of ii). The case when b is ultimately zero is obtained either by considering $\alpha = 0$ in (13), or $\beta = 0$ in (15). ■

References

[1] Abramowitz, M. and Stegan, J.A. (Eds), (1972), Exponential integral and related functions. *Handbook of Mathematical Functions with Formulas, Graphs and mathematical Tables, 9th printing*. New-York, Bover.

[2] Arnold, B.C, Balakrishnan, N., Nagaraja H.N., (1992), *A First course in order statistics*, Wiley and sons.

[3] Beirlant, J., Bouquiaux, C., Werker, B., (2005), Semiparametric lower bounds for tail index estimation, *Journal of Statistical Planning and Inference*, to appear.

[4] Beirlant, J., Broniatowski, M., Teugels, J.L., Vynckier, P., (1995), The mean residual life function at great age: Applications to tail estimation, *Journal of Statistical Planning and Inference*, **45**, 21–48.

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

[6] Broniatowski, M., (1993), On the estimation of the Weibull tail coefficient, *Journal of Statistical Planning and Inference*, **35**, 349–366.

[7] Bingham, N.H., Goldie, C.M. and Teugels, J.L., (1987), Regular Variation. *Cambridge University Press*.

[8] Ditlevsen, O., (1994), Distribution Arbitrariness in Structural Reliability, *Structural Safety and Reliability*, 1241–1247, Balkema, Rotterdam.

[9] Embrechts, P., Klüppelberg, C., Mikosch, T., (1997), *Modelling extremal events*, Springer.

[10] Gardes, L. and Girard, S., (2005), Estimating extreme quantiles of Weibull tail-distributions, *Communication in Statistics - Theory and Methods*, **34**, 1065–1080.

[11] Geluk, J.L. and de Haan, L., (1987), Regular Variation, Extensions and Tauberian Theorems. *Math Centre Tracts*, **40**, Centre for Mathematics and Computer Science, Amsterdam.

[12] Girard, S., (2004), A Hill type estimate of the Weibull tail-coefficient, *Communication in Statistics - Theory and Methods*, **33**(2), 205–234.

[13] Gomes, I., de Haan, L. and Peng, L., (2003), Semi-parametric estimation of the second order parameter in statistics of extremes. *Extremes*, **5**(4), 387–414.

[14] Gomes, I. and Caeiro, F., (2002), A class of asymptotically unbiased semi-parametric estimators of the tail index, *Test*, **11**(2), 345–364.

[15] Gomes, I. and Oliveira, O., (2001), The bootstrap methodology in statistics of extremes. Choice of the optimal sample fraction, *Extremes*, **4**(4), 331–358.

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

[17] Häusler, E., Teugels, J.L. (1985), On asymptotic normality of Hill's estimator for the exponent of regular variation, *The Annals of Statistics*, **13**, 743–756.

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

[19] Klüppelberg, C., Villaseñor, J.A., (1993), Estimation of distribution tails - a semiparametric approach, *Deutschen Gesellschaft für Versicherungsmathematik*, **XXI(2)**, 213–235.

[20] Petrov, V.V., (1975), Sums of independent random variables, *Springer-Verlag. Berlin, Eidelberg, New York*.

[21] Smith, J., (1991), Estimating the upper tail of flood frequency distributions, *Water Resour. Res.* , **23**(8), 1657–1666.

Appendix: proof of lemmas

Proof of Lemma 1 – Remark that, for n large enough,

$$|k^{1/2}b(\log(n/k))| \leq |k^{1/2}b(\log(n/k)) - \lambda| + |\lambda| \leq 1 + |\lambda|,$$

and thus, if b is ultimately non-zero,

$$0 \leq \frac{1}{2}\frac{\log(k)}{\log(n/k)} \leq \frac{\log(1+|\lambda|)}{\log(n/k)} - \frac{\log|b(\log(n/k))|}{\log(n/k)}. \tag{16}$$

Since $|b|$ is a regularly varying function, we have that (see [7], Proposition 1.3.6.)

$$\frac{\log|b(\log(x))|}{\log(x)} \to 0 \text{ as } x \to \infty.$$

Then, (16) implies $\log(k)/\log(n/k) \to 0$ which entails $\log(k)/\log(n) \to 0$. ■

Proof of Lemma 2 – Since for all $x, t > 0$, $tK_\rho(1 + x/t) < x$, Lebesgue Theorem implies that

$$\lim_{t\to\infty} \int_0^\infty \left(tK_\rho\left(1+\frac{x}{t}\right)\right)^q \mathrm{e}^{-x}dx = \int_0^\infty \lim_{t\to\infty} \left(tK_\rho\left(1+\frac{x}{t}\right)\right)^q \mathrm{e}^{-x}dx = \int_0^\infty x^q \mathrm{e}^{-x}dx = q!,$$

which concludes the proof. ■

Proof of Lemma 3 –

i) Lemma 2 shows that $\mu_0(t) \sim 1/t$ and thus $T_n^{(1)}\log(n/k) \to 1$. By definition, $a_n^{(1)} = 0$.

ii) The well-known inequality $-x^2/2 \leq \log(1+x) - x \leq 0$, $x > 0$ yields

$$-\frac{1}{2}\frac{1}{\log(n/k)}\frac{1}{k}\sum_{i=1}^{k}\log^2(k/i) \leq \log(n/k)T_n^{(2)} - \frac{1}{k}\sum_{i=1}^{k}\log(k/i) \leq 0. \tag{17}$$

Now, since when $k \to \infty$,

$$\frac{1}{k}\sum_{i=1}^{k}\log^2(k/i) \to \int_0^1 \log^2(x)dx = 2 \text{ and } \frac{1}{k}\sum_{i=1}^{k}\log(k/i) \to -\int_0^1 \log(x)dx = 1,$$

it follows that $T_n^{(2)}\log(n/k) \to 1$. Let us now introduce the function defined on $(0,1]$ by:

$$f_n(x) = \log\left(1 - \frac{\log(x)}{\log(n/k)}\right).$$

We have:

$$\begin{aligned} a_n^{(2)} &= -\frac{1}{T_n^{(2)}}(T_n^{(2)} - \mu_0(\log(n/k))) = -\frac{1}{T_n^{(2)}}\left(\frac{1}{k}\sum_{i=1}^{k-1} f_n(i/k) - \int_0^1 f_n(t)dt\right) \\ &= -\frac{1}{T_n^{(2)}}\sum_{i=1}^{k-1}\int_{i/k}^{(i+1)/k}(f_n(i/k) - f_n(t))dt + \frac{1}{T_n^{(2)}}\int_0^{1/k} f_n(t)dt. \end{aligned}$$

Since

$$f_n(t) = f_n(i/k) + (t - i/k)\, f_n^{(1)}(i/k) + \int_{i/k}^{t}(t-x)f_n^{(2)}(x)dx,$$

where $f_n^{(p)}$ is the pth derivative of f_n, we have:

$$
\begin{aligned}
a_n^{(2)} &= \frac{1}{T_n^{(2)}}\sum_{i=1}^{k-1}\int_{i/k}^{(i+1)/k}(t-i/k)f_n^{(1)}(i/k)dt \\
&+ \frac{1}{T_n^{(2)}}\sum_{i=1}^{k-1}\int_{i/k}^{(i+1)/k}\int_{i/k}^{t}(t-x)f_n^{(2)}(x)dxdt + \frac{1}{T_n^{(2)}}\int_0^{1/k} f_n(t)dt =: \Psi_1+\Psi_2+\Psi_3.
\end{aligned}
$$

Let us focus first on the term Ψ_1:

$$
\begin{aligned}
\Psi_1 &= \frac{1}{T_n^{(2)}}\frac{1}{2k^2}\sum_{i=1}^{k-1} f_n^{(1)}(i/k) \\
&= \frac{1}{2kT_n^{(2)}}\int_{1/k}^{1} f_n^{(1)}(x)dx + \frac{1}{2kT_n^{(2)}}\left(\frac{1}{k}\sum_{i=1}^{k-1} f_n^{(1)}(i/k) - \int_{1/k}^{1} f_n^{(1)}(x)dx\right) \\
&= \frac{1}{2kT_n^{(2)}}(f_n(1)-f_n(1/k)) - \frac{1}{2kT_n^{(2)}}\sum_{i=1}^{k-1}\int_{i/k}^{(i+1)/k}(f_n^{(1)}(x)-f_n^{(1)}(i/k))dx =: \Psi_{1,1}-\Psi_{1,2}.
\end{aligned}
$$

Since $T_n^{(2)} \sim 1/\log(n/k)$ and $\log(k)/\log(n) \to 0$, we have:

$$
\Psi_{1,1} = -\frac{1}{2kT_n^{(2)}}\log\left(1+\frac{\log(k)}{\log(n/k)}\right) = -\frac{\log(k)}{2k}(1+o(1)).
$$

Furthermore, since, for n large enough, $f_n^{(2)}(x) > 0$ for $x \in [0,1]$,

$$
\begin{aligned}
O \le \Psi_{1,2} &\le \frac{1}{2kT_n^{(2)}}\sum_{i=1}^{k-1}\int_{i/k}^{(i+1)/k}(f_n^{(1)}((i+1)/k)-f_n^{(1)}(i/k))dx = \frac{1}{2k^2T_n^{(2)}}(f_n^{(1)}(1)-f_n^{(1)}(1/k)) \\
&= \frac{1}{2k^2T_n^{(2)}}\left(-\frac{1}{\log(n/k)}+\frac{k}{\log(n/k)}\left(1+\frac{\log(k)}{\log(n/k)}\right)^{-1}\right) \sim \frac{1}{2k} = o\left(\frac{\log(k)}{k}\right).
\end{aligned}
$$

Thus,

$$
\Psi_1 = -\frac{\log(k)}{2k}(1+o(1)). \tag{18}
$$

Second, let us focus on the term Ψ_2. Since, for n large enough, $f_n^{(2)}(x) > 0$ for $x \in [0,1]$,

$$
\begin{aligned}
0 \le \Psi_2 &\le \frac{1}{T_n^{(2)}}\sum_{i=1}^{k-1}\int_{i/k}^{(i+1)/k}\int_{i/k}^{(i+1)/k}(t-i/k)f_n^{(2)}(x)dxdt \\
&= \frac{1}{2k^2T_n^{(2)}}(f_n^{(1)}(1)-f_n^{(1)}(1/k)) = o\left(\frac{\log(k)}{k}\right).
\end{aligned} \tag{19}
$$

Finally,

$$
\Psi_3 = \frac{1}{T_n^{(2)}}\int_0^{1/k} -\frac{\log(t)}{\log(n/k)}dt + \frac{1}{T_n^{(2)}}\int_0^{1/k}\left(f_n(t)+\frac{\log(t)}{\log(n/k)}\right)dt =: \Psi_{3,1}+\Psi_{3,2},
$$

and we have:

$$
\Psi_{3,1} = \frac{1}{\log(n/k)T_n^{(2)}}\frac{1}{k}(\log(k)+1) = \frac{\log(k)}{k}(1+o(1)).
$$

Furthermore, using the well known inequality: $|\log(1+x)-x| \le x^2/2$, $x>0$, we have:

$$
\begin{aligned}
|\Psi_{3,2}| &\le \frac{1}{2T_n^{(2)}}\int_0^{1/k}\left(\frac{\log(t)}{\log(n/k)}\right)^2 dt = \frac{1}{2T_n^{(2)}}\frac{1}{k(\log(n/k))^2}((\log(k))^2+2\log(k)+2) \\
&\sim \frac{(\log(k))^2}{2k\log(n/k)} = o\left(\frac{\log(k)}{k}\right),
\end{aligned}
$$

since $\log(k)/\log(n) \to 0$. Thus,

$$\Psi_3 = \frac{\log(k)}{k}(1 + o(1)). \tag{20}$$

We conclude the proof of i) by collecting (18)-(20).

ii) First, $T_n^{(3)} \log(n/k) = 1$ by definition. Besides, we have

$$\begin{aligned}
a_n^{(3)} &= \frac{\mu_0(\log(n/k))}{T_n^{(3)}} - 1 = \log(n/k)\mu_0(\log(n/k)) - 1 \\
&= \int_0^\infty \log(n/k) \log\left(1 + \frac{x}{\log(n/k)}\right) \mathrm{e}^{-x} dx - 1 \\
&= \int_0^\infty x\mathrm{e}^{-x} dx - \frac{1}{2}\int_0^\infty \frac{x^2}{\log(n/k)} \mathrm{e}^{-x} dx - 1 + R_n = -\frac{1}{\log(n/k)} + R_n,
\end{aligned}$$

where

$$R_n = \int_0^\infty \log(n/k) \left(\log\left(1 + \frac{x}{\log(n/k)}\right) - \frac{x}{\log(n/k)} + \frac{x^2}{2(\log(n/k))^2}\right) \mathrm{e}^{-x} dx.$$

Using the well known inequality: $|\log(1+x) - x + x^2/2| \le x^3/3$, $x > 0$, we have,

$$|R_n| \le \frac{1}{3}\int_0^\infty \frac{x^3}{(\log(n/k))^2} \mathrm{e}^{-x} dx = o\left(\frac{1}{\log(n/k)}\right),$$

which finally yields $a_n^{(3)} \sim -1/\log(n/k)$. ■

Proof of Lemma 4 – Recall that

$$\hat{\theta}_n =: \frac{1}{T_n}\frac{1}{k}\sum_{i=1}^{k-1} (\log(X_{n-i+1,n}) - \log(X_{n-k+1,n})),$$

and let $E_{1,n}, \ldots, E_{n,n}$ be ordered statistics generated by n independent standard exponential random variables. Under **(A.1)**, we have

$$\begin{aligned}
\hat{\theta}_n &\stackrel{d}{=} \frac{1}{T_n}\frac{1}{k}\sum_{i=1}^{k-1} (\log H^{\leftarrow}(E_{n-i+1,n}) - \log H^{\leftarrow}(E_{n-k+1,n})) \\
&\stackrel{d}{=} \frac{1}{T_n}\left(\theta \frac{1}{k}\sum_{i=1}^{k-1} \log\left(\frac{E_{n-i+1,n}}{E_{n-k+1,n}}\right) + \frac{1}{k}\sum_{i=1}^{k-1} \log\left(\frac{\ell(E_{n-i+1,n})}{\ell(E_{n-k+1,n})}\right)\right).
\end{aligned}$$

Define $x_n = E_{n-k+1,n}$ and $\lambda_{i,n} = E_{n-i+1,n}/E_{n-k+1,n}$. It is clear, in view of [12], Lemma 1 that $x_n \xrightarrow{P} \infty$ and $\lambda_{i,n} \xrightarrow{P} 1$. Thus, **(A.2)** yields that uniformly in $i = 1, \ldots, k-1$:

$$\hat{\theta}_n \stackrel{d}{=} \frac{1}{T_n}\left(\theta \frac{1}{k}\sum_{i=1}^{k-1} \log\left(\frac{E_{n-i+1,n}}{E_{n-k+1,n}}\right) + (1 + o_p(1)) b(E_{n-k+1,n}) \frac{1}{k}\sum_{i=1}^{k-1} K_\rho\left(\frac{E_{n-i+1,n}}{E_{n-k+1,n}}\right)\right).$$

The Rényi representation of the Exp(1) ordered statistics (see [2], p. 72) yields

$$\left\{\frac{E_{n-i+1,n}}{E_{n-k+1,n}}\right\}_{i=1,\ldots,k-1} \stackrel{d}{=} \left\{1 + \frac{F_{k-i,k-1}}{E_{n-k+1,n}}\right\}_{i=1,\ldots,k-1}, \tag{21}$$

where $\{F_{1,k-1}, \ldots, F_{k-1,k-1}\}$ are ordered statistics independent from $E_{n-k+1,n}$ and generated by $k-1$ independent standard exponential variables $\{F_1, \ldots, F_{k-1}\}$. Therefore,

$$\begin{aligned}
\hat{\theta}_n \stackrel{d}{=} & \frac{1}{T_n}\left(\theta \frac{1}{k}\sum_{i=1}^{k-1} \log\left(1 + \frac{F_i}{E_{n-k+1,n}}\right)\right. \\
& + \left.(1 + o_p(1)) b(E_{n-k+1,n}) \frac{1}{k}\sum_{i=1}^{k-1} K_\rho\left(1 + \frac{F_i}{E_{n-k+1,n}}\right)\right).
\end{aligned}$$

Remarking that $K_0(x) = \log(x)$ concludes the proof. ■

Proof of Lemma 5 – Lemma 2 implies that,

$$\mu_\rho(E_{n-k+1,n}) \overset{P}{\sim} \frac{1}{E_{n-k+1,n}} \overset{P}{\sim} \frac{1}{\log(n/k)},$$

since $E_{n-k+1,n}/\log(n/k) \xrightarrow{P} 1$ (see [12], Lemma 1). Next, from Lemma 2,

$$\begin{aligned}\sigma_\rho^2(E_{n-k+1,n}) &= \frac{2}{E^2_{n-k+1,n}}(1+o_{\mathrm{P}}(1)) - \frac{1}{E^2_{n-k+1,n}}(1+o_{\mathrm{P}}(1)) \\ &= \frac{1}{E^2_{n-k+1,n}}(1+o_{\mathrm{P}}(1)) = \frac{1}{(\log(n/k))^2}(1+o_{\mathrm{P}}(1)),\end{aligned}$$

which concludes the proof. ■

Proof of Lemma 6 – Remark that

$$\begin{aligned}\frac{k^{1/2}}{\sigma_\rho(E_{n-k+1,n})}\left(U_n^{(\rho)} - \mu_\rho(E_{n-k+1,n})\right) &= \frac{k^{-1/2}}{\sigma_\rho(E_{n-k+1,n})}\sum_{i=1}^{k-1}\left(K_\rho\left(1+\frac{F_i}{E_{n-k+1,n}}\right) - \mu_\rho(E_{n-k+1,n})\right) \\ &- k^{-1/2}\frac{\mu_\rho(E_{n-k+1,n})}{\sigma_\rho(E_{n-k+1,n})}.\end{aligned}$$

Let us introduce the following notation:

$$S_n(t) = \frac{(k-1)^{-1/2}}{\sigma_\rho(t)}\sum_{i=1}^{k-1}\left(K_\rho\left(1+\frac{F_i}{t}\right) - \mu_\rho(t)\right).$$

Thus,

$$\frac{k^{1/2}}{\sigma_\rho(E_{n-k+1,n})}\left(U_n^{(\rho)} - \mu_\rho(E_{n-k+1,n})\right) = S_n(E_{n-k+1,n})(1+o(1)) + o_{\mathrm{P}}(1),$$

from Lemma 5. It remains to prove that for $x \in \mathbb{R}$,

$$\mathrm{P}(S_n(E_{n-k+1,n}) \le x) - \Phi(x) \to 0 \text{ as } n \to \infty,$$

where Φ is the cumulative distribution function of the standard Gaussian distribution. Lemma 2 implies that for all $\varepsilon \in]0,1[$, there exists T_ε such that for all $t \ge T_\varepsilon$,

$$\frac{q!}{t^q}(1-\varepsilon) \le \mathbb{E}\left(\left(K_\rho\left(1+\frac{F_1}{t}\right)\right)^q\right) \le \frac{q!}{t^q}(1+\varepsilon). \tag{22}$$

Furthermore, for $x \in \mathbb{R}$,

$$\begin{aligned}\mathrm{P}(S_n(E_{n-k+1,n}) \le x) - \Phi(x) &= \int_0^{T_\varepsilon}(\mathrm{P}(S_n(t) \le x) - \Phi(x))h_n(t)dt \\ &+ \int_{T_\varepsilon}^{\infty}(\mathrm{P}(S_n(t) \le x) - \Phi(x))h_n(t)dt =: A_n + B_n,\end{aligned}$$

where h_n is the density of the random variable $E_{n-k+1,n}$. First, let us focus on the term A_n. We have,

$$|A_n| \le 2\mathrm{P}(E_{n-k+1,n} \le T_\varepsilon).$$

Since $E_{n-k+1,n}/\log(n/k) \xrightarrow{P} 1$ (see [12], Lemma 1), it is easy to show that $A_n \to 0$. Now, let us consider the term B_n. For the sake of simplicity, let us denote:

$$\left\{ Y_i = K_\rho\left(1+\frac{F_i}{t}\right) - \mu_\rho(t),\ i = 1,\dots,k-1 \right\}.$$

Clearly, $Y_1,\dots,Y_{k-1}$ are independent, identically distributed and centered random variables. Furthermore, for $t \geq T_\varepsilon$,

$$\begin{aligned}
\mathbb{E}(|Y_1|^3) &\leq \mathbb{E}\left(\left(K_\rho\left(1+\frac{F_1}{t}\right)+\mu_\rho(t)\right)^3\right) \\
&= \mathbb{E}\left(\left(K_\rho\left(1+\frac{F_1}{t}\right)\right)^3\right) + (\mu_\rho(t))^3 + 3\mathbb{E}\left(\left(K_\rho\left(1+\frac{F_1}{t}\right)\right)^2\right)\mu_\rho(t) \\
&+ 3\mathbb{E}\left(K_\rho\left(1+\frac{F_1}{t}\right)\right)(\mu_\rho(t))^2 \\
&\leq \frac{1}{t^3}C_1(q,\varepsilon) < \infty,
\end{aligned}$$

from (22) where $C_1(q,\varepsilon)$ is a constant independent of t. Thus, from Esseen's inequality (see [20], Theorem 3), we have:

$$\sup_x |\mathrm{P}(S_n(t) \leq x) - \Phi(x)| \leq C_2 L_n,$$

where C_2 is a positive constant and

$$L_n = \frac{(k-1)^{-1/2}}{(\sigma_\rho(t))^3}\mathbb{E}(|Y_1|^3).$$

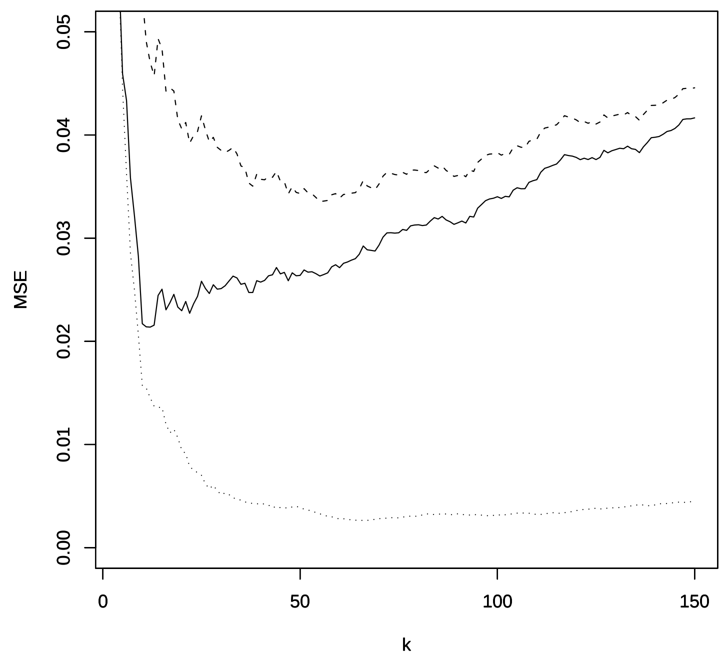
From (22), since $t \geq T_\varepsilon$,

$$(\sigma_\rho(t))^2 = \mathbb{E}\left(\left(K_\rho\left(1+\frac{F_1}{t}\right)\right)^2\right) - \left(\mathbb{E}\left(K_\rho\left(1+\frac{F_1}{t}\right)\right)\right)^2 \geq \frac{1}{t^2}C_3(\varepsilon),$$

where $C_3(\varepsilon)$ is a constant independent of t. Thus, $L_n \leq (k-1)^{-1/2}C_4(q,\varepsilon)$ where $C_4(q,\varepsilon)$ is a constant independent of t, and therefore

$$|B_n| \leq C_4(q,\varepsilon)(k-1)^{-1/2}\mathrm{P}(E_{n-k+1,n} \geq T_\varepsilon) \leq C_4(q,\varepsilon)(k-1)^{-1/2} \to 0,$$

which concludes the proof. ∎



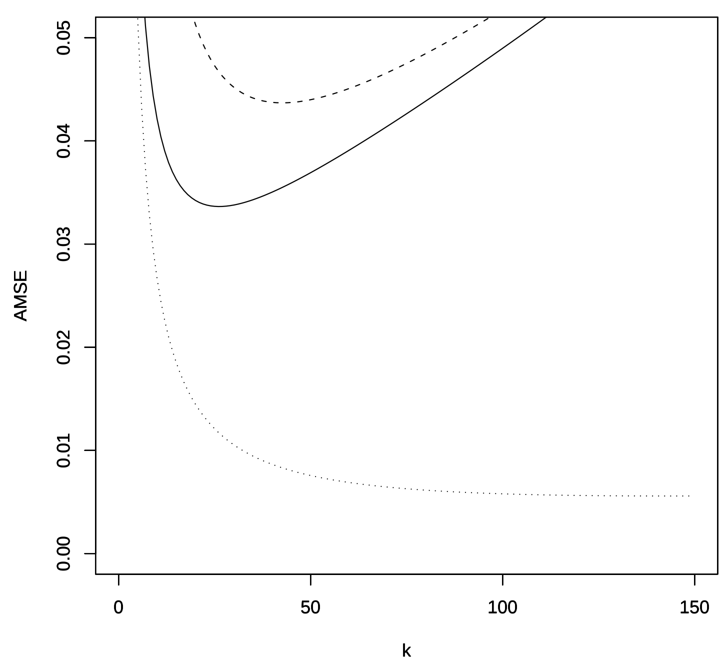
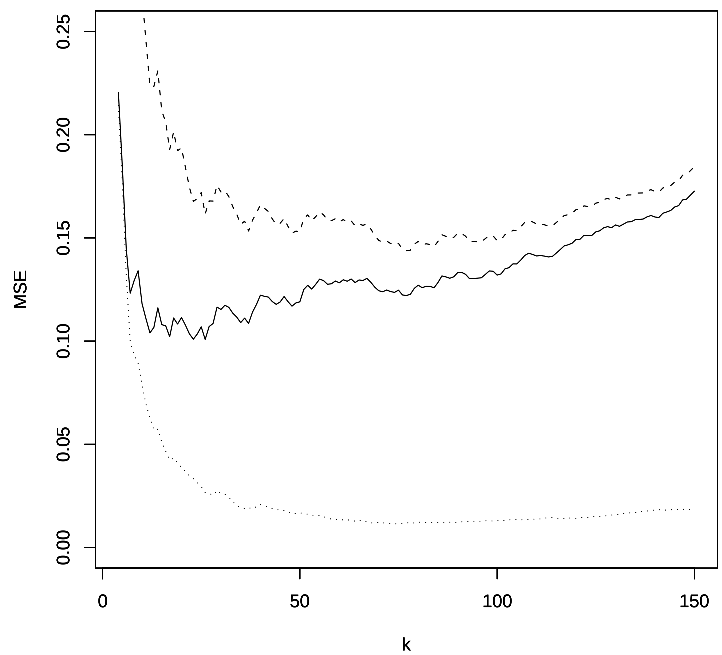


Figure 1: Comparison of estimates $\hat{\theta}_n^{(1)}$ (solid line), $\hat{\theta}_n^{(2)}$ (dashed line) and $\hat{\theta}_n^{(3)}$ (dotted line) for the $|\mathcal{N}(0,1)|$ distribution. Up: MSE, down: AMSE.



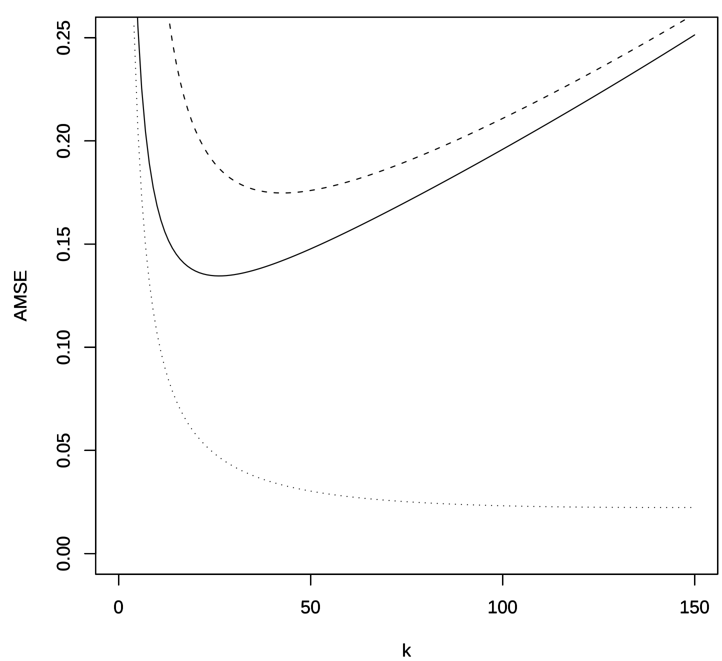
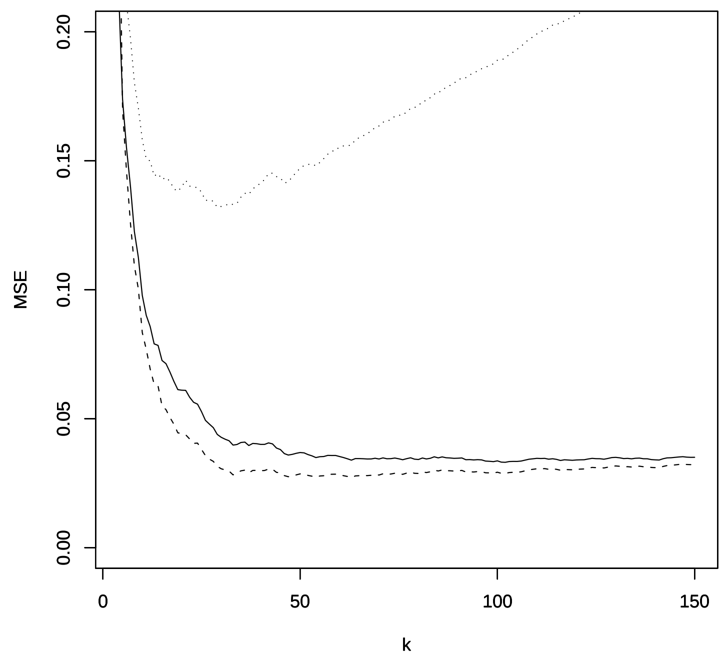


Figure 2: Comparison of estimates $\hat{\theta}_n^{(1)}$ (solid line), $\hat{\theta}_n^{(2)}$ (dashed line) and $\hat{\theta}_n^{(3)}$ (dotted line) for the $\Gamma(0.5, 1)$ distribution. Up: MSE, down: AMSE.



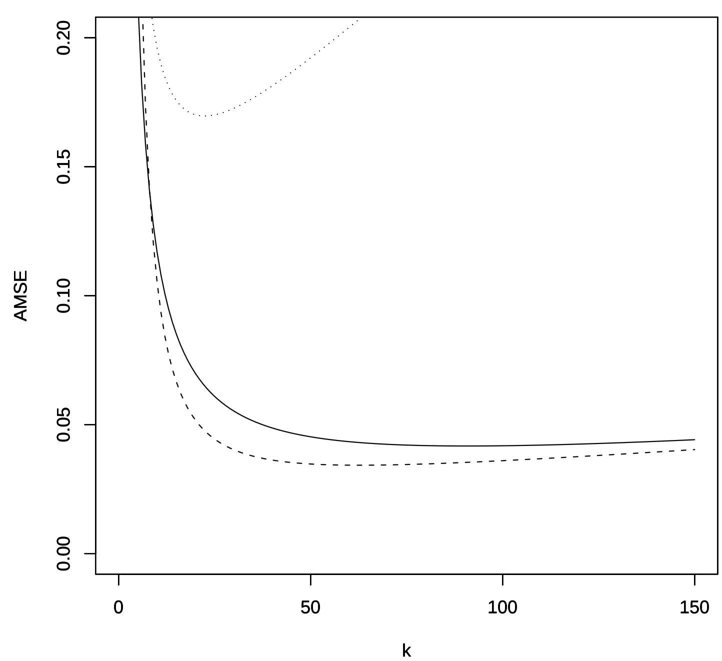


Figure 3: Comparison of estimates $\hat{\theta}_n^{(1)}$ (solid line), $\hat{\theta}_n^{(2)}$ (dashed line) and $\hat{\theta}_n^{(3)}$ (dotted line) for the $\Gamma(1.5, 1)$ distribution. Up: MSE, down: AMSE.

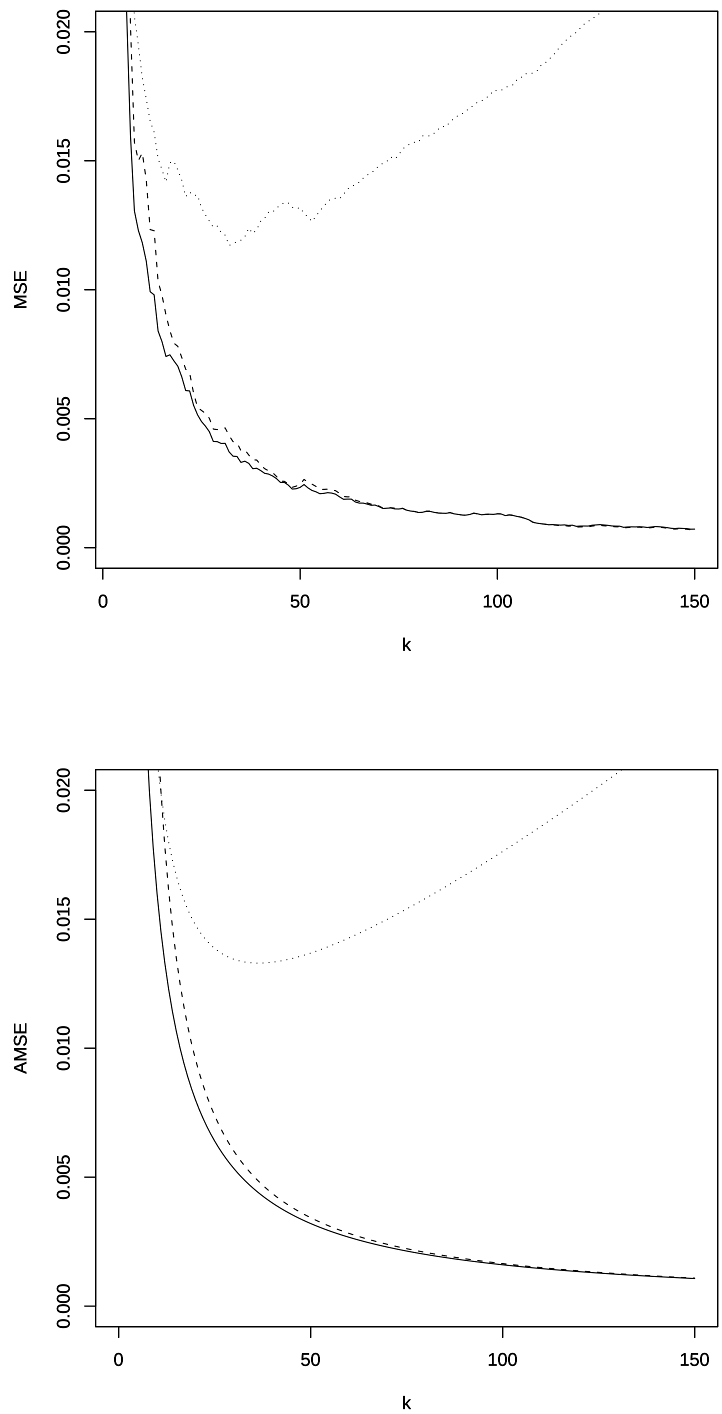
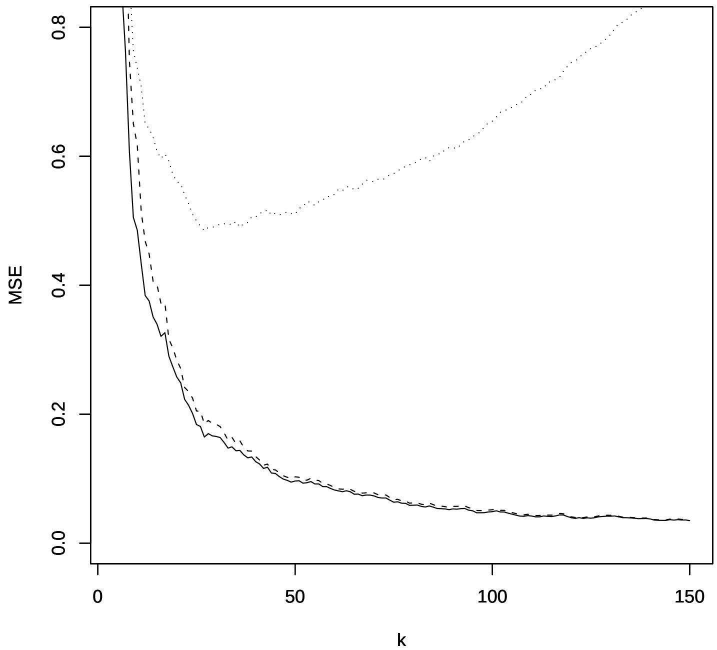


Figure 4: Comparison of estimates $\hat{\theta}_n^{(1)}$ (solid line), $\hat{\theta}_n^{(2)}$ (dashed line) and $\hat{\theta}_n^{(3)}$ (dotted line) for the $\mathcal{W}(2.5, 2.5)$ distribution. Up: MSE, down: AMSE.



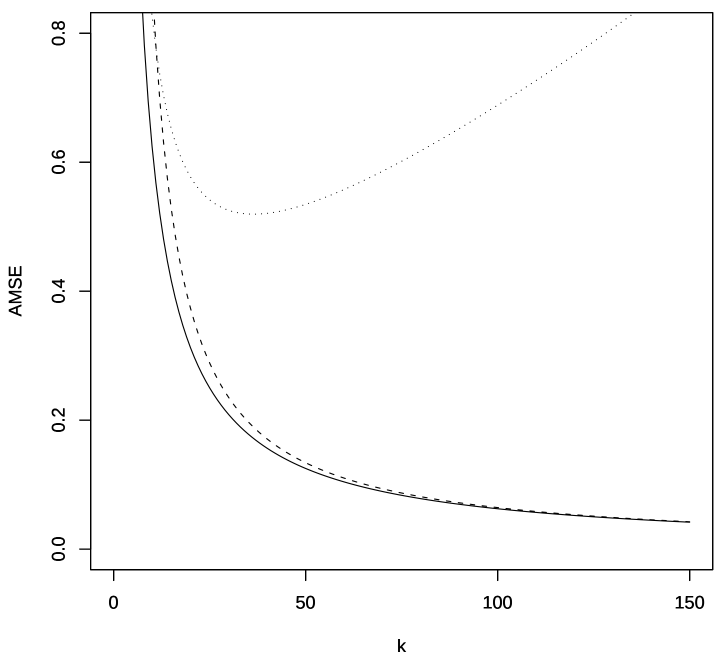


Figure 5: Comparison of estimates $\hat{\theta}_n^{(1)}$ (solid line), $\hat{\theta}_n^{(2)}$ (dashed line) and $\hat{\theta}_n^{(3)}$ (dotted line) for the $\mathcal{W}(0.4, 0.4)$ distribution. Up: MSE, down: AMSE.