

A HILL TYPE ESTIMATOR OF THE WEIBULL TAIL-COEFFICIENT

Stéphane Girard

► **To cite this version:**

Stéphane Girard. A HILL TYPE ESTIMATOR OF THE WEIBULL TAIL-COEFFICIENT. Communication in Statistics - Theory and Methods, Taylor

Francis, 2004, 33 (2), pp.205-234. <hal-00724602>

HAL Id: hal-00724602

<https://hal.inria.fr/hal-00724602>

Submitted on 21 Aug 2012

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.

Stéphane Girard

SMS/LMC, Université Grenoble 1,

BP 53, 38041 Grenoble cedex 9, France

Stephane.Girard@imag.fr

Key Words: Weibull tail coefficient; extreme value index; order statistics.

ABSTRACT

We present a new estimator of the Weibull tail-coefficient. The Weibull tail-coefficient is defined as the regular variation coefficient of the inverse cumulative hazard function. Our estimator is based on the log-spacings of the upper order statistics. Therefore, it is very similar to the Hill estimator for the extreme value index. We prove the weak consistency and the asymptotic normality of our estimator. Its asymptotic as well as its finite sample performances are compared to classical ones.

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 address the problem of estimating the Weibull tail-coefficient $\theta > 0$ defined when the distribution tail satisfies

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

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

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

The inverse cumulative hazard function H^\leftarrow is said to be regularly varying at infinity with index θ and this property is denoted by $H^\leftarrow \in \mathcal{R}_\theta$.

When ℓ is a constant function, this problems reduces to estimating the shape parameter of a Weibull distribution. In this context, simple and efficient methods exist, see for instance [1] for a moment based estimator. Distributions with non-constant slowly varying functions include for instance normal, gamma and extended Weibull distributions (see Section 3 for their definitions). Such distributions are of great use to model large claims in non-life insurance [2]. Dedicated estimation methods have been proposed since the relevant information on the Weibull tail-coefficient is only contained in the extreme upper part of the sample. A first direction was investigated in [3] where an estimator based on the record values is proposed. Another approach consists of using the k_n upper order statistics $X_{n-k_n+1,n} \leq \dots \leq X_{n,n}$ where (k_n) is a sequence of integers such that $1 \leq k_n < n$. Our estimator belongs to this family. It is defined by

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

where $\log_2(t) = \log(\log(t))$, $t > 1$. Some related estimators [4, 5, 6] will be presented below and a precise comparison will be done from both theoretical and practical points of view. Formally, (1) is similar to the expression of the Hill estimator [7]:

$$\hat{\gamma}_n = \frac{1}{k_n - 1} \sum_{i=1}^{k_n-1} (\log(X_{n-i+1,n}) - \log(X_{n-k_n+1,n})) \quad (2)$$

of the extreme value index γ for Pareto type distributions, *i.e.* such that

$$(1 - F)(\lambda x)/(1 - F)(x) \rightarrow \lambda^{-1/\gamma} \text{ as } x \rightarrow \infty \text{ for all } \lambda > 0.$$

We refer to [8] for a review on this topic. Although estimators (1) and (2) do not address the same problem, it will appear that they share similar properties. In Section 2 we state the main asymptotic properties of our estimator. These results are illustrated in Section 3 on some examples of distributions and are compared in Section 4 with the results obtained by other estimators. The behavior of our estimator on finite sample situations is presented in Section 5. Finally, Section 6 is dedicated to the proofs.

2. ASYMPTOTIC RESULTS

In this section, we state our main results on the limiting behavior of the estimator $\hat{\theta}_n$. Its weak consistency is established in Theorem 1 and its asymptotic normality in Theorem 2. The proofs are postponed to Section 6. The weak consistency of $\hat{\theta}_n$ is proved under the usual conditions on the (k_n) sequence (which are also necessary for the weak consistency of the Hill estimator, see [9], Theorem 2):

Theorem 1 *Suppose (A.1) holds. If $k_n \rightarrow \infty$ and $k_n/n \rightarrow 0$ then $\hat{\theta}_n \xrightarrow{P} \theta$.*

The study of the limit distribution of $\hat{\theta}_n$ requires a second order condition on ℓ : There exist $\rho \leq 0$ and $b(x) \rightarrow 0$ such that uniformly locally on $\lambda \geq 1$ when $x \rightarrow \infty$,

$$\text{(A.2)} : \log \left(\frac{\ell(\lambda x)}{\ell(x)} \right) \sim b(x) K_\rho(\lambda),$$

with $K_\rho(\lambda) = \int_1^\lambda u^{\rho-1} du$. It can be shown [10] 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 [11] to prove the asymptotic normality of the Hill estimator and in [12] for one of its refinements.

Theorem 2 *Suppose (A.1) and (A.2) hold. Then*

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

for any sequence (k_n) satisfying

$$k_n \rightarrow \infty, k_n^{1/2}b(\log(n/k_n)) \rightarrow 0 \text{ and } k_n^{1/2}/\log(n/k_n) \rightarrow 0.$$

Note that condition $k_n^{1/2}/\log(n/k_n) \rightarrow 0$ implies condition $k_n/n \rightarrow 0$. If $\rho < -1$, condition $k_n^{1/2}/\log(n/k_n) \rightarrow 0$ implies $k_n^{1/2}b(\log(n/k_n)) \rightarrow 0$ and therefore all the corresponding distributions require the same conditions on the sequence (k_n) . Conversely, if $-1 < \rho \leq 0$, the condition $k_n^{1/2}b(\log(n/k_n)) \rightarrow 0$ is the stronger one and the convergence is slower. If $\rho = -1$, the condition $k_n^{1/2}b(\log(n/k_n)) \rightarrow 0$ may reveal necessary depending on the function b . Some examples are now provided.

3. EXAMPLES

In this section, we give some examples of distributions satisfying assumptions (A.1) and (A.2).

- Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$, $\sigma > 0$. From [13], Table 3.4.4, we have $H^\leftarrow(x) = x^{1/2}\ell(x)$ 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 $\theta = 1/2$, $\rho = -1$ and $b(x) = \log(x)/(4x)$.

- Gamma distribution $\Gamma(\beta, \alpha)$, $\alpha, \beta > 0$. We use the following parameterization of the density

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

From [13], Table 3.4.4, we obtain $H^\leftarrow(x) = x\ell(x)$ with

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

We thus have $\theta = 1$, $\rho = -1$ and $b(x) = (1 - \alpha) \log(x)/x$.

- Weibull distribution $\mathcal{W}(\alpha, \lambda)$, $\alpha, \lambda > 0$. The inverse cumulative hazard function is $H^-(x) = \lambda x^{1/\alpha}$, and then $\ell(x) = \lambda$ for all $x > 0$. Therefore $b(x) = 0$ and we use the usual convention $\rho = -\infty$. A simplified version of Theorem 2 dedicated to this particular case is given in Corollary 1 below.
- Extended Weibull distribution $\mathcal{EW}(\alpha, \beta)$, $\alpha \in (0, 1)$, $\beta \in \mathbb{R}$. This distribution is introduced in [6]. Its distribution tail is given by:

$$1 - F(x) = r(x) \exp(-x^\alpha), \quad (3)$$

where $r \in \mathcal{R}_\beta$. It follows that $H^-(x) = x^{1/\alpha} \ell(x)$ and the following asymptotic expansion holds:

$$\ell(x) = 1 + \frac{\beta \log x}{\alpha^2 x} + O(1/x).$$

It is easily seen that $\theta = 1/\alpha$, $\rho = -1$ and $b(x) = -\beta \log(x)/(\alpha^2 x)$. In [6], it is remarked that this model encompasses the normal and gamma distributions. Then, an estimator of α dedicated to the model (3) is proposed. Its properties are briefly reviewed and compared to those of $\hat{\theta}_n$ in the next section. In particular, we give in Corollary 2 a specific version of Theorem 2 for this family of distributions.

- Modified Weibull distribution $\mathcal{MW}(\alpha)$, $\alpha > 0$. Let $Y \sim \mathcal{W}(\alpha, 1)$, $\alpha > 0$, and introduce $X = Y \log Y$. Then, the inverse cumulative hazard function of X can be written $H^-(x) = x^{1/\alpha} \ell(x)$ with $\ell(x) = \alpha \log(x)$. Therefore, we have $\theta = 1/\alpha$, $\rho = 0$ and $b(x) = 1/\log(x)$. Let us note that this distribution does not belong to the Extended Weibull family.

The parameters θ and ρ as well as the auxiliary function $b(x)$ associated to these distributions are summarized in Table 1. We compare theoretically in the next section our estimator to other ones proposed in the literature on some of these examples.

Our estimator is now compared to other proposals by studying their asymptotic normality properties obtained in some simple situations.

4.1 Comparison with Broniatowski estimator.

In [5], another estimator of the Weibull tail coefficient is proposed. It is defined by

$$\tilde{\theta}_n = \frac{1}{k_n} \sum_{i=1}^{k_n-1} \frac{\log(X_{n-i+1,n})}{\log_2(n/i)}. \quad (4)$$

Both estimators $\hat{\theta}_n$ and $\tilde{\theta}_n$ are based on a similar principle: Introduce $q(t)$, the quantile function defined by

$$q(t) = (1 - F)^-(t) = H^-(\log(1/t)) = (\log(1/t))^\theta \ell(\log(1/t)),$$

under **(A.1)**. The estimator $\tilde{\theta}_n$ is motivated by the approximation

$$\frac{\log(q(t))}{\log_2(1/t)} = \theta + \frac{\log(\ell(\log(1/t)))}{\log_2(1/t)} \simeq \theta. \quad (5)$$

The estimator $\hat{\theta}_n$, as for itself, is motivated by the approximation

$$\begin{aligned} \log(q(t)) - \log(q(s)) &= \theta(\log_2(1/t) - \log_2(1/s)) + \log\left(\frac{\ell(\log(1/t))}{\ell(\log(1/s))}\right) \\ &\simeq \theta(\log_2(1/t) - \log_2(1/s)), \end{aligned} \quad (6)$$

since ℓ is slowly varying. This approximation is similar to the one used in the Hill estimator [7] for estimating the extreme value index. It is important to note that, when ℓ is a constant function, the approximation (6) is exact but not the approximation (5). This remark has an important consequence on the asymptotic behavior of the estimator. From Theorem 2, we easily obtain that, when ℓ is a constant function, the asymptotic normality of $(\hat{\theta}_n - \theta)$ holds under weaker assumptions on k_n :

Corollary 1 *Suppose **(A.1)** holds and ℓ is a constant function. Then,*

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

for any sequence (k_n) satisfying

$$k_n \rightarrow \infty \text{ and } k_n^{1/2}/\log(n/k_n) \rightarrow 0. \quad (7)$$

Obtaining a similar result for $(\theta_n - \theta)$ is not straightforward. In particular, Theorem 2 in [5] requires to verify the condition

$$k_n^{1/2} \log(n/k_n) \log_2(n/k_n) \beta_n \rightarrow 0, \quad (8)$$

where β_n is a bias term defined by

$$\beta_n = \frac{n}{k_n} \int_0^{k_n/n} \frac{du}{\log_2(1/u)}. \quad (9)$$

It can be shown (see Lemma 2) that there exist no sequence (k_n) verifying condition (8), and thus, it is not possible to make the bias term vanish. This comparison is completed on Section 5 by some simulations.

4.2 Comparison with Beirlant *et al.* estimator.

Beirlant *et al.* [4] propose to use the mean residual life function $e(x) = E(X - x | X > x)$ to estimate the parameter θ . They remark that, under assumption **(A.1)**, one have

$$(\log x) \frac{e(H^{\leftarrow}(\log x))}{H^{\leftarrow}(\log x)} \sim \theta \quad (10)$$

when $x \rightarrow \infty$. They thus introduce the following estimator

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

A central limit theorem for $\check{\theta}_n$ is also established (see [4], Theorem 3.2(i)), with the same hypothesis and conclusions as Theorem 2. Consequently, $\hat{\theta}_n$ and $\check{\theta}_n$ share the same asymptotic properties. Nevertheless, we can again remark that the approximation (10) is not exact when ℓ is a constant function. The performance of $\hat{\theta}_n$ and $\check{\theta}_n$ on finite sample situations will be compared in Section 5.

4.3 Estimation within the Extended Weibull family.

In [6], an asymptotic maximum likelihood estimator (AMLE) for the parameter α of the $\mathcal{EW}(\alpha, \beta)$ distribution is proposed. This estimator, which will be denoted by $\hat{\alpha}_n$, is defined by an implicit equation that we do not reproduce here. The authors proved that ([6], Theorem 4.5)

$$k_n^{1/2} \log_2(n/k_n) \left(\hat{\alpha}_n^{-1} - \alpha^{-1} - \frac{\log(\hat{\alpha}_n/\alpha)}{\log_2(n/k_n)} \right) \xrightarrow{d} \mathcal{N}(0, 1), \quad (11)$$

under conditions

$$k_n \rightarrow \infty, k_n/n \rightarrow 0 \text{ and } k_n^{1/2} \frac{\log_2(n/k_n)}{\log(n/k_n)} \rightarrow 0. \quad (12)$$

As a comparison, the application of Theorem 2 to the particular case of $\mathcal{EW}(\alpha, \beta)$ distributions lead to the following

Corollary 2 *Suppose F is a $\mathcal{EW}(\alpha, \beta)$ distribution. Then,*

$$k_n^{1/2}(\hat{\theta}_n - \theta) \xrightarrow{d} \mathcal{N}(0, \theta^2), \quad (13)$$

for any sequence (k_n) satisfying (7).

We can note that, for a fixed value of k_n , the asymptotic convergence rate in (11) is slightly better than in (13). At the opposite, condition (12) is more severe than condition (7) and then the optimal convergence rates are the same.

5. NUMERICAL EXPERIMENTS

The finite sample performance of our estimator is investigated on 7 different distributions: $\Gamma(0.5, 1)$, $\Gamma(1.5, 1)$, $\mathcal{N}(1.2, 1)$, $\mathcal{W}(2.5, 2.5)$, $\mathcal{W}(0.4, 0.4)$, $\mathcal{MW}(2.5)$, and $\mathcal{MW}(0.4)$. In each case, $N = 100$ samples $(\mathcal{X}_{n,i})_{i=1,\dots,N}$ of size $n = 500$ were simulated. On each sample $(\mathcal{X}_{n,i})$, the estimate $\hat{\theta}_{n,i}(k_n)$ is computed for $k_n = 2, \dots, 120$. Finally, the Hill-type plots are built by drawing the points

$$\left(k_n, \frac{1}{N} \sum_{i=1}^N \hat{\theta}_{n,i}(k_n) \right)$$

as well as the associated empirical 90% confidence interval. We also present the associated MSE (mean square error) plots obtained by plotting the points

$$\left(k_n, \sum_{i=1}^N \left(\hat{\theta}_{n,i}(k_n) - \theta \right)^2 \right).$$

The same work is achieved with the estimators $\tilde{\theta}_n$ and $\check{\theta}_n$ for comparison. The estimator $\hat{\alpha}_n$ is not implemented here since it does not address the same problem as $\hat{\theta}_n$, $\tilde{\theta}_n$ and $\check{\theta}_n$. For instance, it is not suited for the modified Weibull distribution. Nevertheless, the reader

Results are presented on Figures 2–8. It appears that, in all the simulated cases, $\hat{\theta}_n$ gives better results than $\check{\theta}_n$. It is also seen that $\hat{\theta}_n$ has a smaller bias and a larger variance than $\check{\theta}_n$. For this reason, $\hat{\theta}_n$ should be preferred to $\check{\theta}_n$ even if the MSE associated to $\hat{\theta}_n$ is larger than $\check{\theta}_n$'s one. Indeed, in many cases, the true value θ belongs to the empirical confidence interval associated to $\hat{\theta}_n$ and is outside the interval associated to $\check{\theta}_n$.

Let us note that the sign of the bias of $\hat{\theta}_n$ and $\check{\theta}_n$ is given by the function b . It is positive for the normal, modified Weibull and gamma distributions (with shape parameter less than one) and negative for the gamma distribution with shape parameter more than one. The bias term vanishes for the Weibull distribution. The worst results are obtained for the modified Weibull distribution. In this situation $\rho = 0$ and thus the bias term which is driven by b decreases very slowly.

In practice, the choice of the parameter k_n is the key problem to obtain a correct estimation of θ . If k_n is too small, the variance of $\hat{\theta}_n$ is huge and conversely, if k_n is too large, the bias is very important. A first solution to select the optimal k_n can be based on the drawing of a quantile plot (QQ-plot) adapted to our situation. It consists of drawing the pairs $(\log_2(n/i), \log(X_{n-i+1,n}))$ for $i = 1, \dots, n - 1$. From (6), the resulting graph should be approximatively linear (with slope θ), at least for the large values of i . On Figure 1 two cases are presented: The QQ-plot associated to a Weibull distribution which is perfectly linear, and the QQ-plot associated to a normal distribution which is only ultimately linear. A graphical method to select k_n could consist of choosing the point where the graph becomes linear. Of course, this method is not very satisfying and an automatic selection method should be developed. A first step could be to adapt the procedure proposed in [15] for the Hill estimator to our framework.

6. PROOFS

Consider $\{E_{1,n}, \dots, E_{n,n}\}$ the order statistics generated by n independent standard exponential random variables. We collect in the next lemma classical results on the asymptotic behavior of these order statistics.

Lemma 1 *Suppose $k_n \rightarrow \infty$ and $k_n/n \rightarrow 0$. Then,*

- (i) $E_{n-i+1,n}/\log(n/i) \xrightarrow{P} 1$, uniformly on $i = 1, \dots, k_n$, and
- (ii) $k_n^{1/2} (E_{n-k_n+1,n} - \log(n/k_n)) \xrightarrow{d} \mathcal{N}(0, 1)$.

Proof : Both results are proved using the representation of the $\text{Exp}(1)$ ordered statistics (see [16], p. 72):

$$\{E_{j,n}\}_{j=1,\dots,n} \stackrel{d}{=} \left\{ \sum_{r=1}^j \frac{E_r}{n-r+1} \right\}_{j=1,\dots,n},$$

where $\{E_1, \dots, E_n\}$ are independent standard exponential variables.

- (i) Introducing for $i = 1, \dots, k_n$ the random variables

$$Z_{n,i} = \frac{E_{n-i+1,n}}{\log(n/i)} - 1,$$

straightforward calculations lead to the uniform bounds $0 \leq E(Z_{n,i}) \leq 1/\log(n/k_n)$ and $\text{var}(Z_{n,i}) \leq 2/\log^2(n/k_n)$.

- (ii) This result is also found in [4], p. 32. The Lyapounov theorem for triangular array yields

$$k_n^{1/2} \left(E_{n-k_n+1,n} - \sum_{s=k_n}^n \frac{1}{s} \right) \xrightarrow{d} \mathcal{N}(0, 1).$$

The remaining term is controlled with the well-known equality

$$\sum_{s=1}^n \frac{1}{s} = \gamma_E + \log n + O(1/n),$$

where γ_E is the Euler's constant. ■

The next lemma is used to prove that there exist no sequence (k_n) such that $k_n \rightarrow \infty$, $k_n/n \rightarrow 0$ and $k_n^{1/2} \log(n/k_n) \log_2(n/k_n) \beta_n \rightarrow 0$, with β_n given in (9).

Lemma 2 *If $k_n \rightarrow \infty$ and $k_n/n \rightarrow 0$, then, for n large enough, $\beta_n \log_2(n/k_n) > 1/2$.*

Proof : Introducing $t = k_n/(nu)$ and $\mu_n = \log(n/k_n)$ in (9) yields

$$\beta_n = \int_1^{+\infty} \frac{dt}{t^2 \log(\log(t) + \mu_n)}.$$

Since $\mu_n \rightarrow \infty$, for n large enough $\log(\log(t) + \mu_n) > 0$ and therefore

$$\beta_n > \int_1^{e^{\mu_n}} \frac{dt}{t^2 \log(\log(t) + \mu_n)} > \frac{1}{\log(2\mu_n)} \int_1^{e^{\mu_n}} \frac{dt}{t^2} \sim \frac{1}{\log_2(n/k_n)}.$$

Thus $\beta_n \log_2(n/k_n)$ is larger than a sequence tending to 1. As a conclusion, for n large enough, $\beta_n \log_2(n/k_n) > 1/2$. ■

Proof of Theorem 1: For the sake of simplicity, in this proof, we note k for k_n . We also write $\hat{\theta}_n = \tau_n^{(2)}/\tau_n^{(1)}$ with

$$\tau_n^{(1)} = \frac{1}{k-1} \sum_{i=1}^{k-1} (\log_2(n/i) - \log_2(n/k)) = \frac{1}{k-1} \sum_{i=1}^{k-1} \log \left(1 + \frac{\log(k/i)}{\log(n/k)} \right).$$

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

$$-\frac{1}{2} \frac{1}{\log(n/k)} \frac{1}{k-1} \sum_{i=1}^{k-1} \log^2(k/i) \leq \log(n/k) \tau_n^{(1)} - \frac{1}{k-1} \sum_{i=1}^{k-1} \log(k/i) \leq 0. \quad (14)$$

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

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

it follows that

$$\tau_n^{(1)} = \frac{1}{\log(n/k)} + O \left(\frac{1}{\log^2(n/k)} \right). \quad (15)$$

Let us consider now

$$\tau_n^{(2)} = \frac{1}{k-1} \sum_{i=1}^{k-1} (\log(X_{n-i+1,n}) - \log(X_{n-k_n+1,n})).$$

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

$$\begin{aligned}\tau_n^{(2)} &\stackrel{d}{=} \frac{1}{k-1} \sum_{i=1}^{k-1} (\log H^{\leftarrow}(E_{n-i+1,n}) - \log H^{\leftarrow}(E_{n-k+1,n})) \\ &\stackrel{d}{=} \theta \frac{1}{k-1} \sum_{i=1}^{k-1} \log \left(\frac{E_{n-i+1,n}}{E_{n-k+1,n}} \right) + \frac{1}{k-1} \sum_{i=1}^{k-1} \log \left(\frac{\ell(E_{n-i+1,n})}{\ell(E_{n-k+1,n})} \right) \\ &=: \tau_n^{(3)} \theta + \tau_n^{(4)}.\end{aligned}$$

In order to rewrite $\tau_n^{(3)}$ into a simpler form, we refer to [17], Lemma 1.4.3, which implies that

$$\begin{aligned}\{E_{n-i+1,n}\}_{i=1,\dots,k-1} &= \{E_{n-k+1,n} + (E_{n-i+1,n} - E_{n-k+1,n})\}_{i=1,\dots,k-1} \\ &\stackrel{d}{=} \{E_{n-k+1,n} + E_{k-i,k-1}\}_{i=1,\dots,k-1}\end{aligned}\tag{16}$$

where $\{E_{1,k-1}, \dots, E_{k-1,k-1}\}$ are ordered statistics independent from $E_{n-k+1,n}$ and generated by $k-1$ independent standard exponential variables $\{E_1, \dots, E_{k-1}\}$. Therefore,

$$\tau_n^{(3)} \stackrel{d}{=} \frac{1}{k-1} \sum_{i=1}^{k-1} \log \left(1 + \frac{E_{k-i,k-1}}{E_{n-k+1,n}} \right) \stackrel{d}{=} \frac{1}{k-1} \sum_{i=1}^{k-1} \log \left(1 + \frac{E_i}{E_{n-k+1,n}} \right),$$

and, similarly to (14),

$$-\frac{1}{2} \frac{1}{E_{n-k+1,n}} \frac{1}{k-1} \sum_{i=1}^{k-1} E_i^2 \leq \tau_n^{(3)} E_{n-k+1,n} - \frac{1}{k-1} \sum_{i=1}^{k-1} E_i \leq 0.$$

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

$$\frac{1}{k-1} \sum_{i=1}^{k-1} E_i^2 \xrightarrow{a.s.} 2 \text{ and } \frac{1}{k-1} \sum_{i=1}^{k-1} E_i \xrightarrow{a.s.} 1,$$

by the strong law of large numbers and Lemma 1(i), it yields

$$\tau_n^{(3)} \stackrel{d}{=} \frac{1}{E_{n-k+1,n}} \frac{1}{k-1} \sum_{i=1}^{k-1} E_i + O_P \left(\frac{1}{\log^2(n/k)} \right),\tag{17}$$

and, in view of (15),

$$\frac{\tau_n^{(3)}}{\tau_n^{(1)}} \stackrel{d}{=} \frac{\log(n/k)}{E_{n-k+1,n}} \frac{1}{k-1} \sum_{i=1}^{k-1} E_i + O_P \left(\frac{1}{\log(n/k)} \right) \xrightarrow{P} 1,\tag{18}$$

with Lemma 1(i) and the strong law of large numbers.

Finally, $\tau_n^{(4)}$ can be rewritten using the Karamata representation of ℓ (see e.g. [18]):

$$\ell(x) = c(x) \exp\left(\int_{x_0}^x \frac{\delta(u)}{u} du\right), \quad x \geq x_0,$$

where $c(x) \rightarrow c_0 > 0$ and $\delta(x) \rightarrow 0$ as $x \rightarrow \infty$. This entails

$$\tau_n^{(4)} = \frac{1}{k-1} \sum_{i=1}^{k-1} \log\left(\frac{c(E_{n-i+1,n})}{c(E_{n-k+1,n})}\right) + \frac{1}{k-1} \sum_{i=1}^{k-1} \int_{E_{n-k+1,n}}^{E_{n-i+1,n}} \frac{\delta(u)}{u} du =: \tau_n^{(5)} + \tau_n^{(6)},$$

and the end of the proof is similar to [9], p. 757: $\tau_n^{(5)}$ converges to 0 in probability in view of Lemma 1(i). Besides, $\tau_n^{(6)}$ can be controlled by introducing the random variable

$$\xi_n = \sup\{|\delta(u)|, u \geq E_{n-k+1,n}\},$$

and remarking that

$$|\tau_n^{(6)}| \leq \xi_n \frac{1}{k-1} \sum_{i=1}^{k-1} \int_{E_{n-k+1,n}}^{E_{n-i+1,n}} \frac{1}{u} du = \xi_n \tau_n^{(3)} = o_P(\tau_n^{(3)}),$$

with Lemma 1(i). From (15) and (18), we obtain that $\tau_n^{(4)}/\tau_n^{(1)}$ converges to 0 in probability and the conclusion follows with (18): $\hat{\theta}_n = \tau_n^{(2)}/\tau_n^{(1)} \xrightarrow{P} \theta$. ■

Proof of Theorem 2: Keeping in mind the notations introduced in the previous proof, we have

$$k^{1/2}(\hat{\theta}_n - \theta) = \theta k^{1/2} \left(\frac{\tau_n^{(3)}}{\tau_n^{(1)}} - 1 \right) + k^{1/2} \frac{\tau_n^{(4)}}{\tau_n^{(1)}}.$$

In view of (15), (17) and Lemma 1(i),

$$\begin{aligned} k^{1/2} \left(\frac{\tau_n^{(3)}}{\tau_n^{(1)}} - 1 \right) &\stackrel{d}{=} \left[k^{1/2} \left(\frac{1}{k-1} \sum_{i=1}^{k-1} E_i - 1 \right) - k^{1/2} \left(\frac{E_{n-k+1,n}}{\log(n/k)} - 1 \right) \right] \frac{\log(n/k)}{E_{n-k+1,n}} \\ &\quad + O_P\left(\frac{k^{1/2}}{\log(n/k)}\right) \\ &\stackrel{d}{=} \left[k^{1/2} \left(\frac{1}{k-1} \sum_{i=1}^{k-1} E_i - 1 \right) - k^{1/2} \left(\frac{E_{n-k+1,n}}{\log(n/k)} - 1 \right) \right] (1 + o_P(1)) \end{aligned}$$

$$\begin{aligned}
& + O_P\left(\frac{k^{1/2}}{\log(n/k)}\right) \\
& \stackrel{d}{=} k^{1/2}\left(\frac{1}{k-1}\sum_{i=1}^{k-1}E_i - 1\right)(1 + o_P(1)) + O_P\left(\frac{k^{1/2}}{\log(n/k)}\right),
\end{aligned}$$

with Lemma 1(ii). The central limit theorem together with the restrictions imposed on k yield

$$\theta k^{1/2}\left(\frac{\tau_n^{(3)}}{\tau_n^{(1)}} - 1\right) \xrightarrow{d} \mathcal{N}(0, \theta^2).$$

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 Lemma 1(i) that $x_n \xrightarrow{P} \infty$ and $\lambda_{i,n} \xrightarrow{P} 1$. Thus, **(A.2)** yields that uniformly in $i = 1, \dots, k-1$:

$$\log\left(\frac{\ell(E_{n-i+1,n})}{\ell(E_{n-k+1,n})}\right) \sim b(E_{n-k+1,n})K_\rho\left(\frac{E_{n-i+1,n}}{E_{n-k+1,n}}\right)(1 + o_P(1)).$$

Using the representation (16), we get

$$\begin{aligned}
\tau_n^{(4)} & \stackrel{d}{=} (1 + o_P(1))b(E_{n-k+1,n})\frac{1}{k-1}\sum_{i=1}^{k-1}K_\rho\left(1 + \frac{E_{k-i,k-1}}{E_{n-k+1,n}}\right) \\
& \stackrel{d}{=} (1 + o_P(1))b(E_{n-k+1,n})\frac{1}{k-1}\sum_{i=1}^{k-1}K_\rho\left(1 + \frac{E_i}{E_{n-k+1,n}}\right) \\
& \stackrel{d}{=} (1 + o_P(1))\frac{b(E_{n-k+1,n})}{E_{n-k+1,n}}\frac{1}{k-1}\sum_{i=1}^{k-1}E_i,
\end{aligned}$$

similarly to (17), since $(\rho - 1)x^2/2 \leq K_\rho(1 + x) - x \leq 0$ when $x \geq 0$. Now, (15) and Lemma 1(i) entail

$$k^{1/2}\frac{\tau_n^{(4)}}{\tau_n^{(1)}} \stackrel{d}{=} (1 + o_P(1))k^{1/2}b(E_{n-k+1,n})\frac{1}{k-1}\sum_{i=1}^{k-1}E_i \stackrel{d}{=} (1 + o_P(1))k^{1/2}b(E_{n-k+1,n}),$$

by the strong law of large numbers. Finally, $b(E_{n-k+1,n})/b(\log(n/k))$ converges to 1 in probability, since $E_{n-k+1,n}/\log(n/k)$ converges to 1 in probability and $|b| \in \mathcal{R}_\rho$. As a conclusion, and again due to the restrictions on k ,

$$k^{1/2}\frac{\tau_n^{(4)}}{\tau_n^{(1)}} \stackrel{d}{=} (1 + o_P(1))k^{1/2}b(\log(n/k)) = o_P(1),$$

and the result is proved. ■

- [1] Cran, G.W., (1988), Moment estimators for the 3-parameter Weibull distribution, *IEEE Transactions on Reliability*, **37**(4), 360–363.
- [2] Beirlant, J., Teugels, J.L., (1992), Modeling large claims in non-life insurance. *Insurance: Math. Econom.*, **11**, 17–29.
- [3] Berred, M., (1991), Record values and the estimation of the Weibull tail-coefficient. *Comptes-Rendus de l'Académie des Sciences t. 312*, Série I, 943–946.
- [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] Broniatowski, M., (1993), On the estimation of the Weibull tail coefficient, *Journal of Statistical Planning and Inference*, **35**, 349–366.
- [6] Klüppelberg, C., Villaseñor, J.A., (1993), Estimation of distribution tails - a semi-parametric approach, *Deutschen Gesellschaft für Versicherungsmathematik*, **XXI**(2), 213–235.
- [7] Hill, B.M., (1975), A simple general approach to inference about the tail of a distribution, *The Annals of Statistics*, **3**, 1163–1174.
- [8] Csörgö, S., Viharos, L., (1998), Estimating the tail index, *Asymptotic methods in probability and statistics*, North-Holland, Amsterdam, 833–881.
- [9] Mason, D.M., (1982), Laws of large numbers for sums of extreme values, *Annals of Probability*, **10**, 756–764.
- [10] Geluk, J.L., de Haan, L., (1987), Regular Variation, Extensions and Tauberian Theorems. *Math Centre Tracts*, **40**, Centre for Mathematics and Computer Science, Amsterdam.

- [11] 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.
- [12] Beirlant, J., Dierckx, G., Goegebeur, Y., Matthys, G., (1999), Tail index estimation and an exponential regression model, *Extremes*, **2**, 177–200.
- [13] Embrechts, P., Klüppelberg, C., Mikosch, T., (1997), *Modelling extremal events*, Springer.
- [14] Keller, B., Klüppelberg, C., (1991), Statistical estimation of large claims distributions. *Mitteilungen der Schweizerischen Vereinigung der Versicherungsmathematiker*, 203–216.
- [15] Guillou, A., Hall, P., (2001), A diagnostic for selecting the threshold in extreme-value analysis, *J. Roy. Statist. Soc. Ser. B*, **63**(2), 293–305.
- [16] Arnold, B.C., Balakrishnan, N., Nagaraja H.N., (1992), *A First course in order statistics*, Wiley and sons.
- [17] Reiss, R.-D., (1989), *Approximate distribution of order statistics*, Springer-Verlag.
- [18] Bingham, N.H., Goldie, C.M., Teugels, J.L., (1987), *Regular variation*, Encyclopedia of Mathematics and its Applications, **27**, Cambridge University Press.

	θ	$b(x)$	ρ
$\mathcal{N}(\mu, \sigma^2)$	$1/2$	$\frac{1}{4} \frac{\log x}{x}$	-1
$\Gamma(\beta, \alpha \neq 1)$	1	$(1 - \alpha) \frac{\log x}{x}$	-1
$\mathcal{W}(\alpha, \lambda)$	$1/\alpha$	0	$-\infty$
$\mathcal{EW}(\alpha, \beta \neq 0)$	$1/\alpha$	$-\frac{\beta}{\alpha^2} \frac{\log x}{x}$	-1
$\mathcal{MW}(\alpha)$	$1/\alpha$	$\frac{1}{\log x}$	0

Table 1: Parameters θ , ρ and the function $b(x)$ associated to some usual distributions

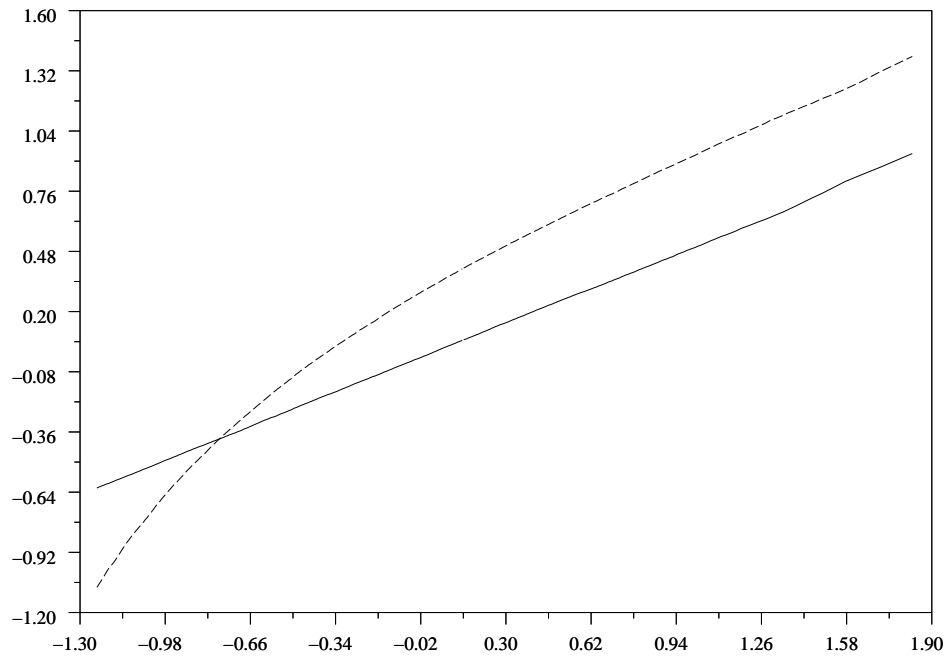
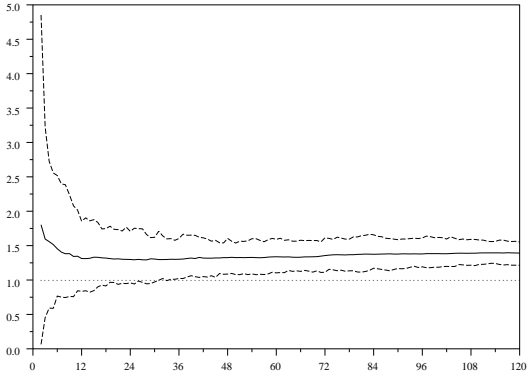
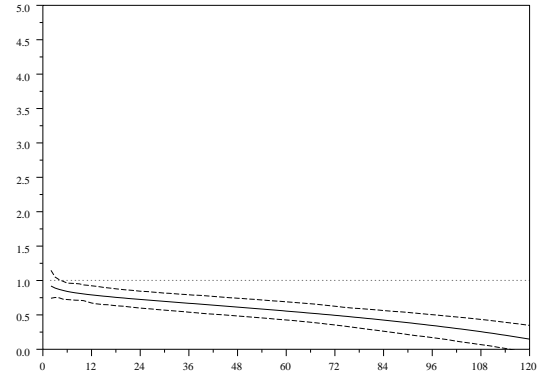


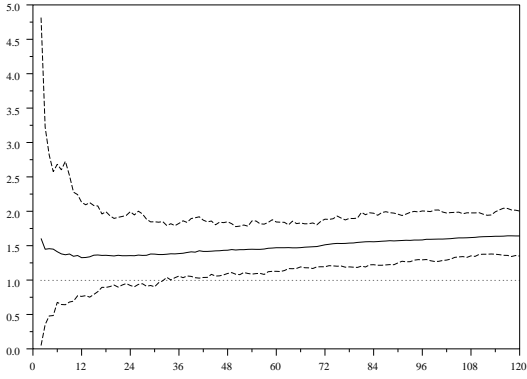
Figure 1: Two examples of QQ-plots, $\log(X_{n-i+1,n})$ is plotted versus $\log_2(n/i)$. Solid line: $\mathcal{W}(2,1)$ distribution, dashed line: $\mathcal{N}(1,1)$ distribution



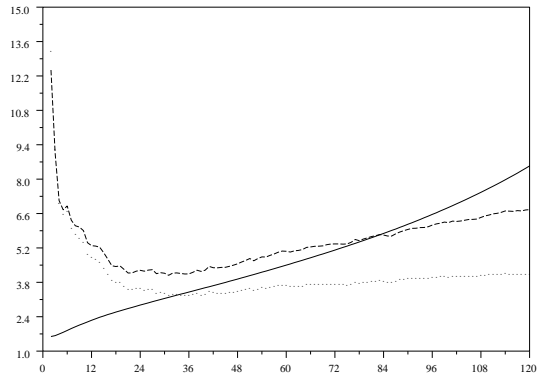
(a) Estimate $\hat{\theta}_n$ as a function of k_n .



(b) Estimate $\tilde{\theta}_n$ as a function of k_n .

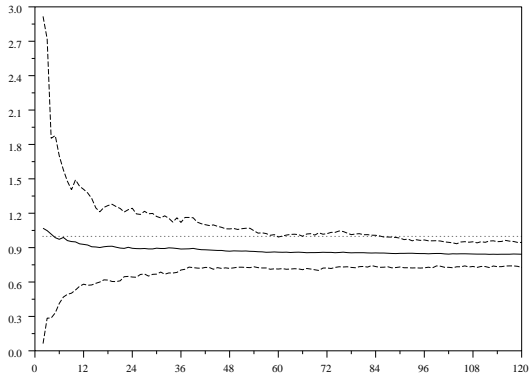


(c) Estimate $\check{\theta}_n$ as a function of k_n .

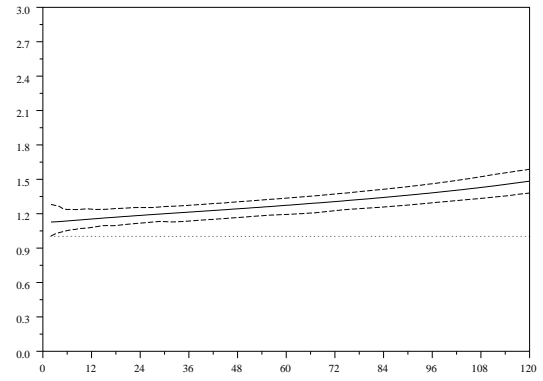


(d) Mean square error as a function of k_n .

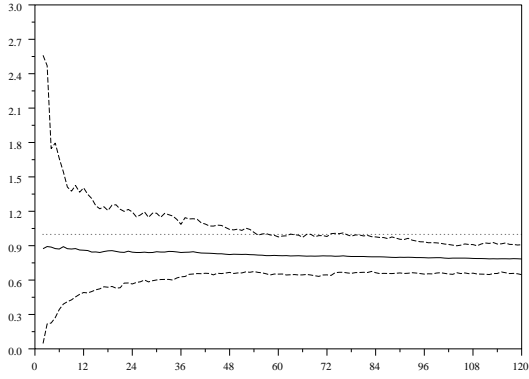
Figure 2: Comparison of estimates $\hat{\theta}_n$, $\tilde{\theta}_n$ and $\check{\theta}_n$ for the $\Gamma(0.5, 1)$ distribution. (a)-(c) Dotted line: true value, solid line: mean over 100 samples, dashed lines: empirical 90% confidence interval. (d) Dotted line: $\hat{\theta}_n$, solid line: $\tilde{\theta}_n$, dashed line: $\check{\theta}_n$.



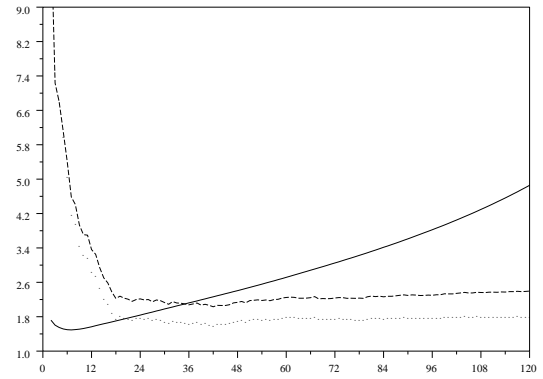
(a) Estimate $\hat{\theta}_n$ as a function of k_n .



(b) Estimate $\tilde{\theta}_n$ as a function of k_n .

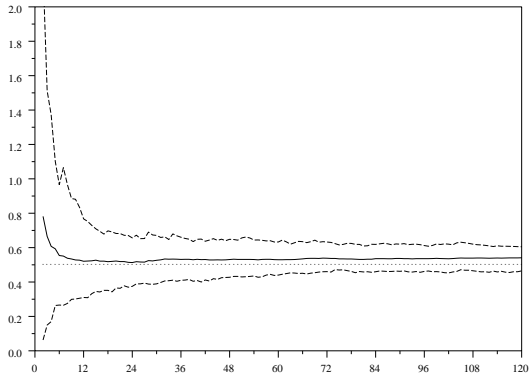


(c) Estimate $\check{\theta}_n$ as a function of k_n .

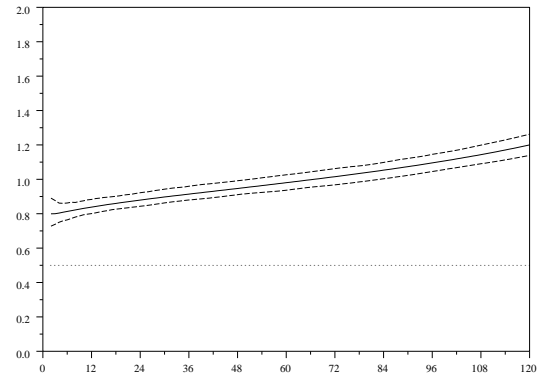


(d) Mean square error as a function of k_n .

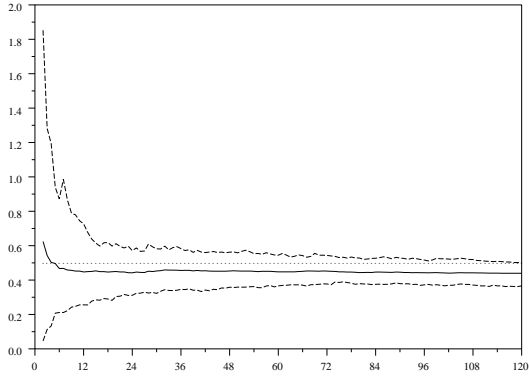
Figure 3: Comparison of estimates $\hat{\theta}_n$, $\tilde{\theta}_n$ and $\check{\theta}_n$ for the $\Gamma(1.5, 1)$ distribution. (a)-(c) Dotted line: true value, solid line: mean over 100 samples, dashed lines: empirical 90% confidence interval. (d) Dotted line: $\hat{\theta}_n$, solid line: $\tilde{\theta}_n$, dashed line: $\check{\theta}_n$.



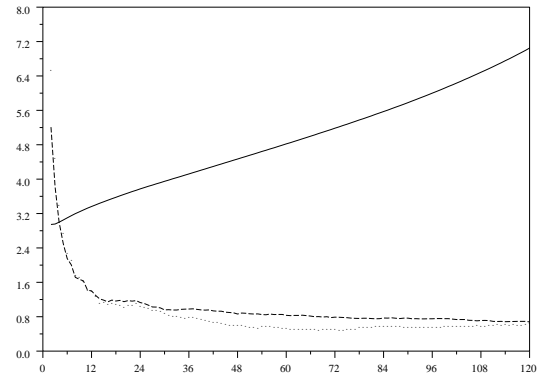
(a) Estimate $\hat{\theta}_n$ as a function of k_n .



(b) Estimate $\tilde{\theta}_n$ as a function of k_n .

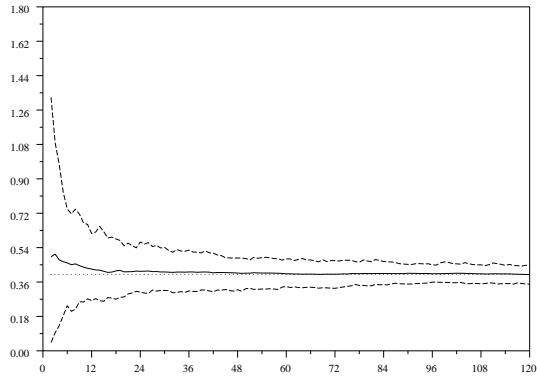


(c) Estimate $\check{\theta}_n$ as a function of k_n .

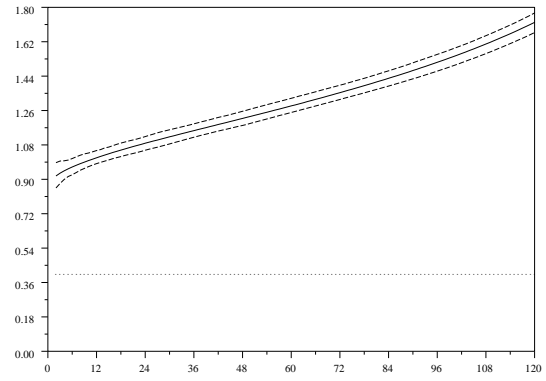


(d) Mean square error as a function of k_n .

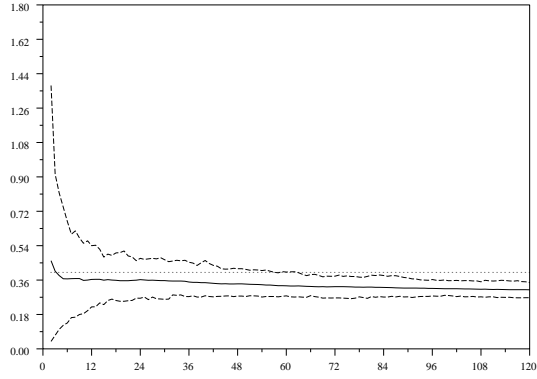
Figure 4: Comparison of estimates $\hat{\theta}_n$, $\tilde{\theta}_n$ and $\check{\theta}_n$ for the $\mathcal{N}(1.2, 1)$ distribution. (a)-(c) Dotted line: true value, solid line: mean over 100 samples, dashed lines: empirical 90% confidence interval. (d) Dotted line: $\hat{\theta}_n$, solid line: $\tilde{\theta}_n$, dashed line: $\check{\theta}_n$.



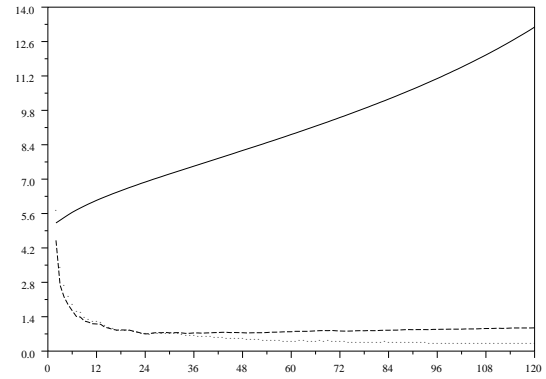
(a) Estimate $\hat{\theta}_n$ as a function of k_n .



(b) Estimate $\tilde{\theta}_n$ as a function of k_n .

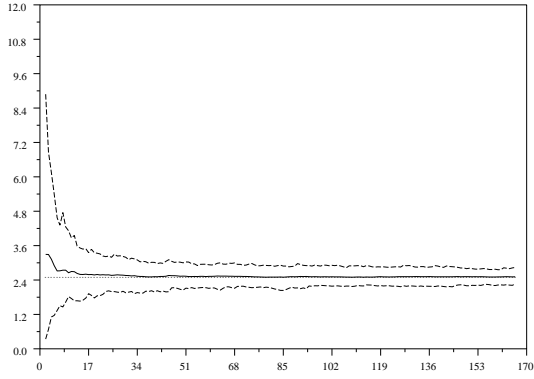


(c) Estimate $\check{\theta}_n$ as a function of k_n .

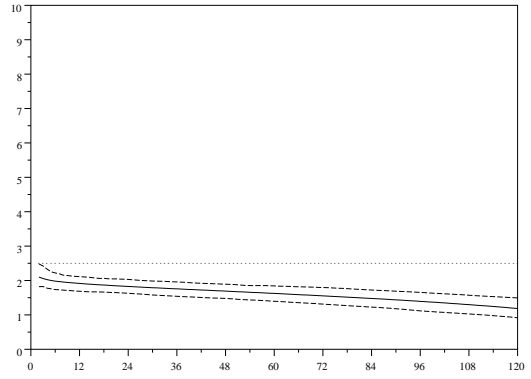


(d) Mean square error as a function of k_n .

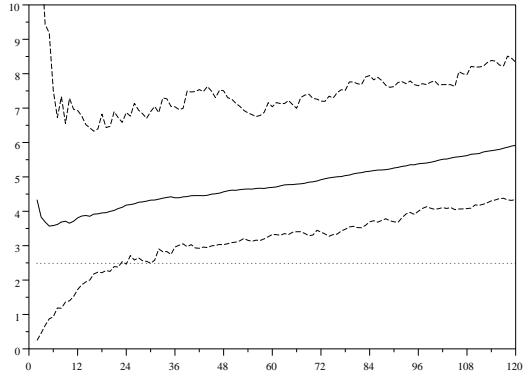
Figure 5: Comparison of estimates $\hat{\theta}_n$, $\tilde{\theta}_n$ and $\check{\theta}_n$ for the $\mathcal{W}(2.5, 2.5)$ distribution. (a)-(c) Dotted line: true value, solid line: mean over 100 samples, dashed lines: empirical 90% confidence interval. (d) Dotted line: $\hat{\theta}_n$, solid line: $\tilde{\theta}_n$, dashed line: $\check{\theta}_n$.



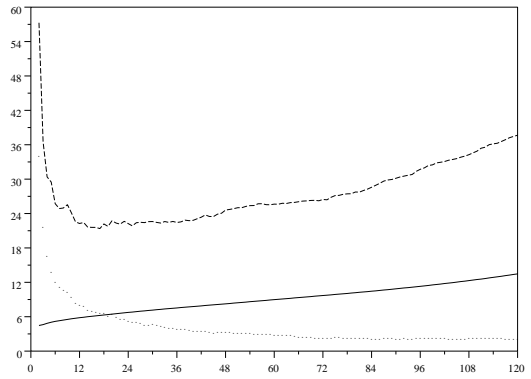
(a) Estimate $\hat{\theta}_n$ as a function of k_n .



(b) Estimate $\tilde{\theta}_n$ as a function of k_n .

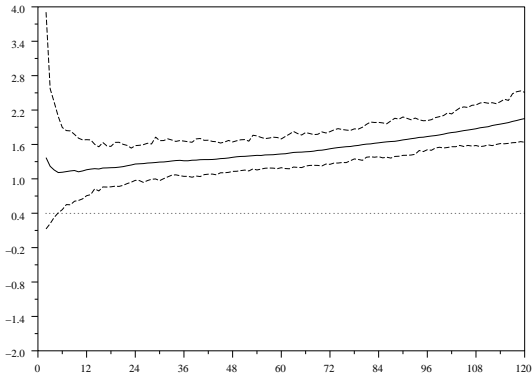


(c) Estimate $\check{\theta}_n$ as a function of k_n .

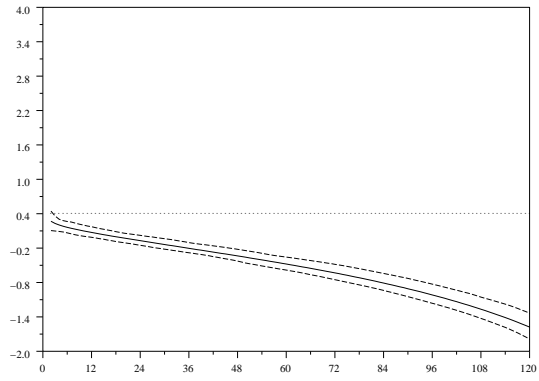


(d) Mean square error as a function of k_n .

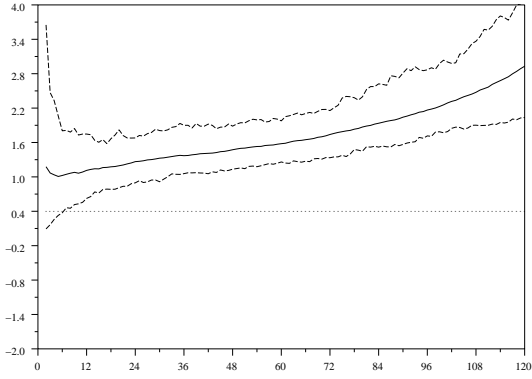
Figure 6: Comparison of estimates $\hat{\theta}_n$, $\tilde{\theta}_n$ and $\check{\theta}_n$ for the $\mathcal{W}(0.4, 0.4)$ distribution. (a)-(c) Dotted line: true value, solid line: mean over 100 samples, dashed lines: empirical 90% confidence interval. (d) Dotted line: $\hat{\theta}_n$, solid line: $\tilde{\theta}_n$, dashed line: $\check{\theta}_n$.



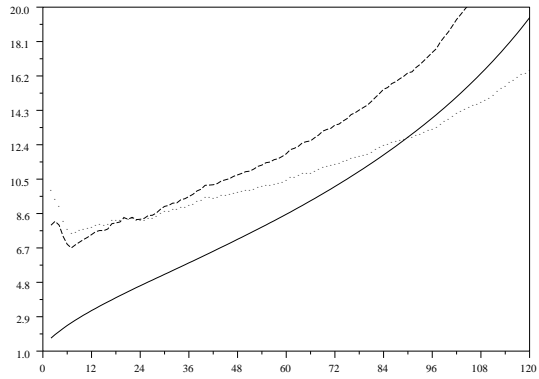
(a) Estimate $\hat{\theta}_n$ as a function of k_n .



(b) Estimate $\tilde{\theta}_n$ as a function of k_n .

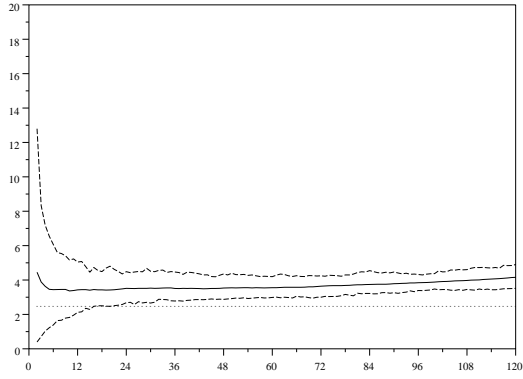


(c) Estimate $\check{\theta}_n$ as a function of k_n .

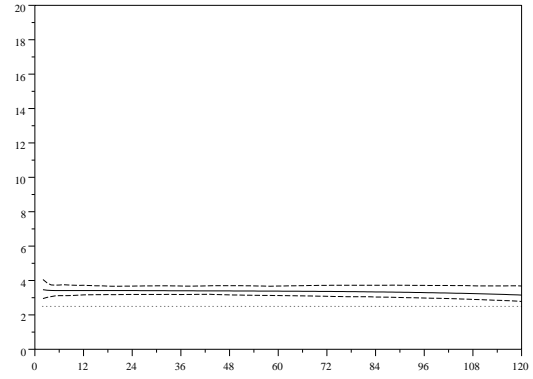


(d) Mean square error as a function of k_n .

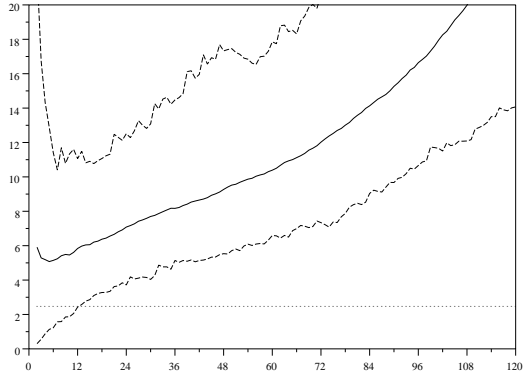
Figure 7: Comparison of estimates $\hat{\theta}_n$, $\tilde{\theta}_n$ and $\check{\theta}_n$ for the $\mathcal{MW}(2.5)$ distribution. (a)-(c) Dotted line: true value, solid line: mean over 100 samples, dashed lines: empirical 90% confidence interval. (d) Dotted line: $\hat{\theta}_n$, solid line: $\tilde{\theta}_n$, dashed line: $\check{\theta}_n$.



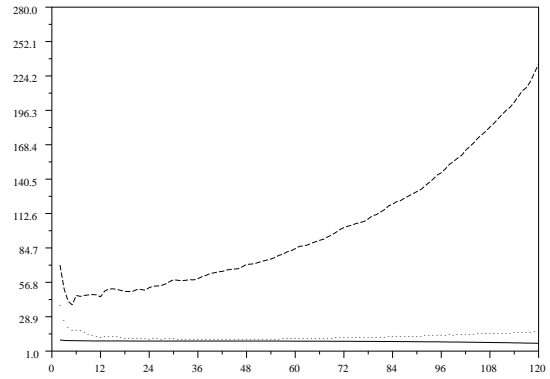
(a) Estimate $\hat{\theta}_n$ as a function of k_n .



(b) Estimate $\tilde{\theta}_n$ as a function of k_n .



(c) Estimate $\check{\theta}_n$ as a function of k_n .



(d) Mean square error as a function of k_n .

Figure 8: Comparison of estimates $\hat{\theta}_n$, $\tilde{\theta}_n$ and $\check{\theta}_n$ for the $\mathcal{MW}(0.4)$ distribution. (a)-(c) Dotted line: true value, solid line: mean over 100 samples, dashed lines: empirical 90% confidence interval. (d) Dotted line: $\hat{\theta}_n$, solid line: $\tilde{\theta}_n$, dashed line: $\check{\theta}_n$.