Bias-reduced estimators of the Weibull tail-coefficient

Jean Diebolt, Laurent Gardes, Stephane Girard, Armelle Guillou

BIAS-REDUCED ESTIMATORS OF THE WEIBULL TAIL-COEFFICIENT

Jean Diebolt[(1)], Laurent Gardes[(2)], Stéphane Girard[(3,*)] and Armelle Guillou[(4)]

[(1)] *CNRS, Université de Marne-la-Vallée*
Équipe d'Analyse et de Mathématiques Appliquées
5, boulevard Descartes, Batiment Copernic
Champs-sur-Marne
77454 Marne-la-Vallée Cedex 2

[(2)] *Université Grenoble 2,*
LabSAD, 1251 Avenue centrale
B.P. 47, 38040 Grenoble Cedex 9

[(3,*)] *Corresponding author.*
Université Grenoble 1,
LMC-IMAG, 51 rue des Mathématiques
B.P. 53, 38041 Grenoble Cedex 9,
Phone: +(33) 4 76 51 45 53, Fax: +(33) 4 76 63 12 63
`Stephane.Girard@imag.fr`

[(4)] *Université Paris VI*
Laboratoire de Statistique Théorique et Appliquée
Boîte 158
175 rue du Chevaleret
75013 Paris

Abstract. *In this paper, we consider the problem of the estimation of a Weibull tail-coefficient θ. In particular, we propose a regression model, from which we derive a bias-reduced estimator of θ. This estimator is based on a least-squares approach. The asymptotic normality of this estimator is also established. A small simulation study is provided in order to prove its efficiency.*

Key words and phrases. Weibull tail-coefficient, Bias-reduction, least-squares approach, asymptotic normality.

AMS Subject classifications. 62G05, 62G20, 62G30.

1. Introduction.

Let $X_1, ..., X_n$ be a sequence of independent and identically distributed random variables with distribution function F, and let $X_{1,n} \leq ... \leq X_{n,n}$ denote the order statistics associated to this sample.

In the present paper, we address the problem of estimating the Weibull tail-coefficient $\theta > 0$ defined as

$$1 - F(x) = \exp(-H(x)) \qquad \text{with} \qquad H^{-1}(x) := \inf\{t : H(t) \geq x\} = x^{\theta}\ell(x), \tag{1}$$

where ℓ is a slowly varying function at infinity satisfying

$$\frac{\ell(\lambda x)}{\ell(x)} \longrightarrow 1, \text{ as } x \to \infty, \text{ for all } \lambda > 0. \tag{2}$$

Girard (2004) investigated this estimation problem and proposed the following estimator of θ:

$$\widetilde{\theta}_n = \frac{\displaystyle\sum_{i=1}^{k_n} \Big(\log X_{n-i+1,n} - \log X_{n-k_n+1,n} \Big)}{\displaystyle\sum_{i=1}^{k_n} \Big(\log\log \frac{n}{i} - \log\log \frac{n}{k_n} \Big)}, \tag{3}$$

where k_n is an intermediate sequence, i.e. a sequence such that $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$.

We refer to Beirlant *et al.* (1995) and Broniatowski (1993) for other propositions and to Beirlant *et al.* (2005) for Local Asymptotic Normality (LAN) results. Estimator (3) is closed in spirit to the Hill estimator (see Hill, 1975) in the case of Pareto-type distributions. In Girard (2004), the asymptotic normality of $\widetilde{\theta}_n$ is established under suitable assumptions. To prove such a result, a second-order condition is required in order to specify the bias-term. This assumption can be expressed in terms of the slowly varying function ℓ as follows:

Assumption $(R_\ell(b, \rho))$ *There exists a constant $\rho \leq 0$ and a rate function b satisfying $b(x) \to 0$ as $x \to \infty$, such that for all $\varepsilon > 0$ and $1 < A < \infty$, we have*

$$\sup_{\lambda \in [1,A]} \left| \frac{\log \frac{\ell(\lambda x)}{\ell(x)}}{b(x) K_\rho(\lambda)} - 1 \right| \leq \varepsilon, \quad \textit{for } x \textit{ sufficiently large,}$$

with $K_\rho(\lambda) = \displaystyle\int_1^\lambda t^{\rho-1} dt$.

It can be shown that necessarily $|b|$ is regularly varying with index ρ (see e.g. Geluk and de Haan, 1987). Moreover, we focus on the case where the convergence (2) is slow,

and thus when the bias term in $\widetilde{\theta}_n$ is large. This situation is described by the following assumption:

$$xb(x) \to \infty \text{ as } x \to \infty. \tag{4}$$

Let us note that this condition implies $\rho \geq -1$. Gamma and Gaussian distributions fulfill (4), whereas Weibull distributions do not (see Table 1) since, in this case, the bias term vanishes.
Using this framework, we will establish rigorously in Section 2 the following approximation for the log-spacings of upper order statistics:

$$Z_j := j \log \frac{n}{j}\Big(\log X_{n-j+1,n} - \log X_{n-j,n} \Big) \approx \left(\theta + b\Big(\log \frac{n}{k_n} \Big) \left(\frac{\log \frac{n}{j}}{\log \frac{n}{k_n}} \right)^{\rho} \right) f_j, \tag{5}$$

for $1 \leq j \leq k_n$, where $(f_1, ..., f_{k_n})$ is a vector of independent and standard exponentially distributed random variables.

This exponential regression model is similar to the ones proposed by Beirlant *et al.* (1999, 2002) and Feuerverger and Hall (1999) in the case of Pareto-type distributions.

Ignoring $b\Big(\log \frac{n}{k_n} \Big) \left(\frac{\log \frac{n}{j}}{\log \frac{n}{k_n}} \right)^{\rho}$ in (5) leads to the maximum likelihood estimator

$$\check{\theta}_n = \frac{1}{k_n} \sum_{j=1}^{k_n} Z_j,$$

which turns out to be an alternative estimator of $\widetilde{\theta}_n$.

The full model (5) allows us to generate bias-corrected estimates $\widehat{\theta}_n$ for θ through maximum likelihood estimation of θ, $b(\log n/k_n)$ and ρ for each $1 \leq k_n \leq n-1$. An alternative to this approach consists in using a canonical choice for ρ and to estimate the two other parameters by a least-squares method (LS). For the canonical choice of ρ, we can use for instance the value -1, which is the same as the one proposed by Feuerverger and Hall (1999) for the regression model in the case of Pareto-type distributions. The asymptotic normality of the resulting LS-estimator is established in Section 3. In order to illustrate the usefulness of the bias-term, we provide a small simulation study in Section 4. The proofs of our results are postponed to Section 6.

2. Exponential regression model

In this section, we formalize (5). First, remark that

$$F^{-1}(x) = \Big[-\log(1-x) \Big]^{\theta} \ell\Big(-\log(1-x) \Big).$$

Since $X_{n-j+1,n} \stackrel{d}{=} F^{-1}(U_{n-j+1,n}), 1 \leq j \leq k_n$, where $U_{j,n}$ denotes the j-th order statistic of a

uniform sample of size n, we have

$$X_{n-j+1,n} \stackrel{d}{=} \Big[-\log(1-U_{n-j+1,n})\Big]^{\theta} \ell\Big(-\log(1-U_{n-j+1,n})\Big)$$

which implies that

$$\log X_{n-j+1,n} \stackrel{d}{=} \theta \log\Big[-\log(1-U_{n-j+1,n})\Big] + \log\Big[\ell\Big(-\log(1-U_{n-j+1,n})\Big)\Big].$$

Moreover, considering the order statistics from an independent standard exponential sample, $E_{n-j+1,n} \stackrel{d}{=} -\log(1-U_{n-j+1,n})$. Therefore

$$\begin{aligned} \log X_{n-j+1,n} &\stackrel{d}{=} \theta \log(E_{n-j+1,n}) + \log\Big[\ell(E_{n-j+1,n})\Big] \\ &=: A_n(j) + B_n(j). \end{aligned}$$

Our basic result now reads as follows.

Theorem 1 *Suppose (1) holds together with* $(R_\ell(b,\rho))$ *and (4). Then, if* $k_n \to \infty$ *and* $\log k_n / \log n \to 0$*, we have*

$$\begin{aligned} &\sup_{1\le j\le k_n} \left| j \log \frac{n}{j}\Big(\log X_{n-j+1,n} - \log X_{n-j,n}\Big) - \left(\theta + b\Big(\log \frac{n}{k_n}\Big)\left(\frac{\log \frac{n}{j}}{\log \frac{n}{k_n}}\right)^{\rho}\right) f_j \right| \\ &= o_{\mathbb{P}}\Big(b\Big(\log \frac{n}{k_n}\Big)\Big), \end{aligned} \tag{6}$$

where $(f_1, ..., f_{k_n})$ *is a vector of independent and standard exponentially distributed random variables.*

The proof of this theorem is based on the following two lemmas:

Lemma 1 *Suppose (1) holds together with* $(R_\ell(b,\rho))$ *and (4). Then, if* $k_n \to \infty$ *and* $k_n/n \to 0$*, we have*

$$\sup_{1\le j\le k_n} \left| j \log \frac{n}{j}\Big[A_n(j) - A_n(j+1)\Big] - \theta\, f_j \right| = o_{\mathbb{P}}\Big(b\Big(\log \frac{n}{k_n}\Big)\Big), \tag{7}$$

and

Lemma 2 *Suppose (1) holds together with* $(R_\ell(b,\rho))$*. Then, if* $k_n \to \infty$ *and* $\log k_n / \log n \to 0$*, we have*

$$\sup_{1\le j\le k_n} \left| j \log \frac{n}{j}\Big[B_n(j) - B_n(j+1)\Big] - b\Big(\log \frac{n}{k_n}\Big)\left(\frac{\log \frac{n}{j}}{\log \frac{n}{k_n}}\right)^{\rho} f_j \right| = o_{\mathbb{P}}\Big(b\Big(\log \frac{n}{k_n}\Big)\Big). \tag{8}$$

The proof of these lemmas is postponed to Section 6.

Remark 1 *Under the assumptions of Theorem 1, we also have*

$$\sup_{1\le j\le k_n}\left| j\log\frac{n}{j}\Big(\log X_{n-j+1,n}-\log X_{n-j,n}\Big)-\left(\theta+b\Big(\log\frac{n}{k_n}\Big)\left(\frac{\log\frac{n}{j}}{\log\frac{n}{k_n}}\right)^{-1}\right)f_j\right| = o_{\mathbb{P}}\Big(b\Big(\log\frac{n}{k_n}\Big)\Big),$$

where $(f_1,...,f_{k_n})$ *is a vector of independent and standard exponentially distributed random variables.*

This implies that one can plug the canonical choice $\rho=-1$ in the regression model (6) without perturbing the approximation. From model (6) we can easily deduce the asymptotic normality of the estimator $\check{\theta}_n$, given in the next theorem:

Theorem 2 *Suppose (1) holds together with* $(R_\ell(b,\rho))$ *and (4). Then, if* $k_n\to\infty$, $\sqrt{k_n}b(\log(n/k_n))\to\lambda\in\mathbb{R}$ *and if* $\lambda=0$*:* $\log k_n/\log n\to 0$*, we have*

$$\sqrt{k_n}\left(\check{\theta}_n-\theta-b\Big(\log\frac{n}{k_n}\Big)\frac{1}{k_n}\sum_{j=1}^{k_n}\left(\frac{\log\frac{n}{j}}{\log\frac{n}{k_n}}\right)^{\rho}\right)\xrightarrow{d}\mathcal{N}(0,\theta^2).$$

This model (6) now plays the central role in the remainder of this paper. It allows us to generate bias-corrected estimates of θ as we will show in the next section.

3. Bias-reduced estimates of θ

In order to reduce the bias of the estimator $\check{\theta}_n$, we can either estimate simultaneously $\theta, b(\log n/k_n)$ and ρ by a maximum likelihood method or estimate θ and b by a least-squares approach after substituting a canonical choice for ρ. In fact, this second-order parameter is difficult to estimate in practice and we can easily check by simulations that fixing its value does not much influence the result. This problem has already been discussed in Beirlant *et al.* (1999, 2002) and Feuerverger and Hall (1999) where similar observations have been made in the case of Pareto-type distributions. The canonical choice $\rho=-1$ is often used although other choices could be motivated performing a model selection.

In all the sequel, we will estimate θ and $b(\log\frac{n}{k_n})$ by a LS-method after substituting ρ with the value -1. In that case, we find the following LS-estimators:

$$\begin{cases}\widehat{\theta}_n=\overline{Z}_{k_n}-\widehat{b}\Big(\log\frac{n}{k_n}\Big)\overline{x}_{k_n}\\ \widehat{b}\Big(\log\frac{n}{k_n}\Big)=\dfrac{\sum_{j=1}^{k_n}(x_j-\overline{x}_{k_n})Z_j}{\sum_{j=1}^{k_n}(x_j-\overline{x}_{k_n})^2}\end{cases}$$

where $x_j = \left(\frac{\log \frac{n}{j}}{\log \frac{n}{k_n}}\right)^{-1}$, $\overline{x}_{k_n} = \frac{1}{k_n}\sum_{j=1}^{k_n} x_j$ and $\overline{Z}_{k_n} = \frac{1}{k_n}\sum_{j=1}^{k_n} Z_j$.

Our next goal is to establish, under suitable assumptions, the asymptotic normality of $\widehat{\theta}_n$. This is done in the following theorem.

Theorem 3 *Suppose (1) holds together with* $(R_\ell(b,\rho))$ *and (4). Then, if* k_n *is such that*

$$k_n \to \infty, \frac{\sqrt{k_n}}{\log \frac{n}{k_n}} b\Big(\log \frac{n}{k_n}\Big) \to \Lambda \in \mathbb{R} \text{ and, if } \Lambda = 0,\ \frac{\log^2 k_n}{\log \frac{n}{k_n}} \to 0 \text{ and} \frac{\sqrt{k_n}}{\log \frac{n}{k_n}} \to \infty, \tag{9}$$

we have

$$\frac{\sqrt{k_n}}{\log \frac{n}{k_n}}\Big(\widehat{\theta}_n - \theta\Big) \xrightarrow{d} \mathcal{N}(0,\theta^2).$$

Remark that the rate of convergence of $\check{\theta}_n$ is the same as the one of $\widehat{\theta}_n$ in the cases where both λ and Λ are not equal to 0.

The proof of this theorem is postponed to Section 6.

In order to illustrate the usefulness of the bias-term in the model (6), we will provide a small simulation study in the next section.

4. A small simulation study

The finite sample performances of the estimators $\widehat{\theta}_n$, $\widetilde{\theta}_n$ and $\check{\theta}_n$ are investigated on 5 different distributions: $\Gamma(0.25,1)$, $\Gamma(4,1)$, $\mathcal{N}(1.1,1)$, $\mathcal{W}(0.25,0.25)$ and $\mathcal{W}(4,4)$. We limit ourselves to these three estimators, since it is shown in Girard (2004) that $\widetilde{\theta}_n$ gives better results than the other approaches (Beirlant *et al.*, 1995; Broniatowski, 1993). 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 estimates $\widehat{\theta}_{n,i}(k_n)$, $\widetilde{\theta}_{n,i}(k_n)$ and $\check{\theta}_{n,i}(k_n)$ were computed for $k_n = 2,\dots,360$. Finally, the Hill-type plots were built by drawing the points

$$\left(k_n, \frac{1}{N}\sum_{i=1}^{N} \widehat{\theta}_{n,i}(k_n)\right), \left(k_n, \frac{1}{N}\sum_{i=1}^{N} \widetilde{\theta}_{n,i}(k_n)\right) \text{ and } \left(k_n, \frac{1}{N}\sum_{i=1}^{N} \check{\theta}_{n,i}(k_n)\right).$$

We also present the associated MSE (mean square error) plots obtained by plotting the points

$$\left(k_n, \frac{1}{N}\sum_{i=1}^{N} \Big(\widehat{\theta}_{n,i}(k_n) - \theta\Big)^2\right), \left(k_n, \frac{1}{N}\sum_{i=1}^{N} \Big(\widetilde{\theta}_{n,i}(k_n) - \theta\Big)^2\right), \text{ and } \left(k_n, \frac{1}{N}\sum_{i=1}^{N} \Big(\check{\theta}_{n,i}(k_n) - \theta\Big)^2\right).$$

The results are presented on figures 1–5. In all the plots, the graphs associated to $\widetilde{\theta}_n$

and $\check{\theta}_n$ are similar, with a slightly better behaviour of $\check{\theta}_n$. The bias corrected estimator $\widehat{\theta}_n$ always yields a smaller bias than the two previous ones leading to better results for Gamma and Gaussian distributions (figures 1–3). On Weibull distributions (figures 4–5), it presents a larger variance.

5. Concluding remarks

In this paper, we introduce a regression model, from which we derive a bias-reduced estimator for the Weibull tail-coefficient θ. Its asymptotic normality is established and its efficiency is illustrated in a small simulation study. However, in many cases of practical interest, the problem of estimating a quantile $x_{p_n} = F^{-1}(1-p_n)$, with $p_n < 1/n$, is much more important. Such a problem has already been studied in Gardes and Girard (2005) where the following Weissman-type estimator has been introduced

$$\widetilde{x}_{p_n} = X_{n-k_n+1,n}\left(\frac{\log\frac{1}{p_n}}{\log\frac{n}{k_n+1}}\right)^{\widetilde{\theta}_n}.$$

It is, however, desirable to refine $\widetilde{x}_{p_n}$ with the additional information about the slowly varying function ℓ that is provided by the LS-estimates for θ and b. To this aim, condition $(R_\ell(b,\rho))$ is used to approximate the ratio $F^{-1}(1-p_n)/X_{n-k_n+1,n}$, noting that

$$X_{n-k_n+1,n} \stackrel{d}{=} F^{-1}(U_{n-k_n+1,n}),$$

with $U_{1,n} \le ... \le U_{n,n}$ the order statistics of a uniform $(0,1)$ sample of size n,

$$\begin{aligned}
\frac{x_{p_n}}{X_{n-k_n+1,n}} &\stackrel{d}{=} \frac{F^{-1}(1-p_n)}{F^{-1}(U_{n-k_n+1,n})} \\
&\stackrel{d}{=} \frac{(-\log p_n)^\theta}{(-\log(1-U_{n-k_n+1,n}))^\theta}\,\frac{\ell(-\log p_n)}{\ell(-\log(1-U_{n-k_n+1,n}))} \\
&\stackrel{d}{\simeq} \left(\frac{\log\frac{1}{p_n}}{\log\frac{n}{k_n+1}}\right)^\theta \exp\left[b\Big(\log\frac{n}{k_n}\Big)\frac{\Big(\frac{\log\frac{1}{p_n}}{\log\frac{n}{k_n+1}}\Big)^\rho - 1}{\rho}\right].
\end{aligned}$$

The last step follows from replacing $U_{k_n+1,n}$ (resp. $E_{n-k_n+1,n}$) by $(k_n+1)/n$ (resp. $\log n/(k_n+1)$). Hence, we arrive at the following estimator for extreme quantiles

$$\widehat{x}_{p_n} = X_{n-k_n+1,n}\left(\frac{\log\frac{1}{p_n}}{\log\frac{n}{k_n+1}}\right)^{\widehat{\theta}_n} \exp\left[\widehat{b}\Big(\log\frac{n}{k_n}\Big)\frac{\Big(\frac{\log\frac{1}{p_n}}{\log\frac{n}{k_n+1}}\Big)^{\widehat{\rho}} - 1}{\widehat{\rho}}\right].$$

Here, the LS-estimators of θ and b can be used after substituting ρ by the canonical choice -1. The study of the asymptotic properties of such an estimator is beyond the scope of the present paper, but it will lead to further investigations.

6. Proofs of our results

6.1 Preliminary lemmas

Lemma 3 *For all $1 \le j \le k_n$ such that $k_n \to \infty$ and $\frac{k_n}{n} \to 0$, we have*

$$\frac{E_{n-j,n}}{\log \frac{n}{j}} = 1 + O_{\mathbb{P}}\Big(\frac{1}{\log \frac{n}{k_n}}\Big) \quad \textit{uniformly in } j.$$

Proof of Lemma 3. According to Rényi's representation, we have

$$E_{n-j,n} \stackrel{d}{=} \sum_{\ell=1}^{n-j+1} \frac{f_{n-\ell-j+1}}{\ell + j - 1}$$

where $f_j \stackrel{i.i.d.}{\sim} \text{Exp}(1)$. Since

$$\text{Var}\Big(\sum_{\ell=1}^{n-j+1} \frac{f_{n-\ell-j+1}}{\ell + j - 1}\Big) = O(1),$$

denoting

$$T_{j,n} := \sum_{\ell=1}^{n-j+1} \Big[\frac{f_{n-\ell-j+1}}{\ell + j - 1} - \mathbb{E}\frac{f_{n-\ell-j+1}}{\ell + j - 1}\Big],$$

we have, using Kolmogorov's inequality (see e.g. Shorack and Wellner, 1986, p. 843), that

$$\mathbb{P}\Big(\max_{1\le j\le k_n} |T_{j,n}| \ge \lambda\Big) \le \frac{\text{Var}(T_{1,n})}{\lambda^2}, \qquad \lambda > 0.$$

This implies that $T_{j,n} = O_{\mathbb{P}}(1)$ uniformly in j. Taking into account the fact that

$$\Big|\sum_{\ell=j}^{n} \frac{1}{\ell} - \log \frac{n}{j}\Big| = O(1) \qquad \text{uniformly in } j, 1 \le j \le k_n,$$

it is easy to deduce Lemma 3. □

Let us introduce the E_m-function defined by the integral

$$E_m(x) := \int_1^\infty \frac{e^{-xt}}{t^m} dt$$

for a positive integer m. The asymptotic expansion of this integral is given in the following lemma.

Lemma 4 *As $x \to \infty$, for any fixed positive integers m and p, we have*

$$E_m(x) = \frac{e^{-x}}{x}\left\{1 - \frac{m}{x} + \frac{m(m+1)}{x^2} + ... + (-1)^p \frac{m(m+1)...(m+p-1)}{x^p} + O\Big(\frac{1}{x^{p+1}}\Big)\right\}.$$

The proof of this lemma is straightforward from Abramowitz and Stegun (1972, p. 227-233) and the $O-$term can be obtained by a Taylor expansion with an integral remainder. Denote

$$\mu_p := \frac{1}{k_n}\sum_{j=1}^{k_n}\Big(x_j - \overline{x}_{k_n}\Big)^p,\ p \in \mathbb{N}^*.$$

The next lemma provides a first order expansion of this Riemman sum.

Lemma 5 *If $k_n \to \infty$, $\frac{k_n}{n} \to 0$, $\frac{k_n}{\log \frac{n}{k_n}} \to \infty$ and $\frac{\log^2 k_n}{\log \frac{n}{k_n}} \to 0$, then*

$$\mu_p \sim C_p\left(\log \frac{n}{k_n}\right)^{-p} \qquad \text{as } n \to \infty, \text{ where} \qquad C_p = \int_0^1 \Big(\log x + 1\Big)^p dx < \infty.$$

Proof of Lemma 5. Denote $\alpha_n = \frac{1}{\log n/k_n}$. Then $\overline{x}_{k_n}$ can be rewritten as

$$\overline{x}_{k_n} = \frac{1}{k_n} + \left(\frac{1}{k_n}\sum_{j=1}^{k_n-1} f_n(j/k_n) - \int_0^1 f_n(x)dx\right) + \int_0^1 f_n(x)dx =: \frac{1}{k_n} + T_1 + T_2,$$

where $f_n(x) = (1 - \alpha_n \log x)^{-1}, x \in [0,1]$.

Denoting by $f_n^{(i)}$, $i \in \{1,2\}$, the ith derivative of f_n, we infer that

$$\begin{aligned} T_1 &= \sum_{j=1}^{k_n-1}\int_{j/k_n}^{(j+1)/k_n}\Big(\frac{j}{k_n} - t\Big)f_n^{(1)}\Big(\frac{j}{k_n}\Big)dt + \sum_{j=1}^{k_n-1}\int_{j/k_n}^{(j+1)/k_n}\int_{j/k_n}^{t}(x - t)f_n^{(2)}(x)dxdt \\ &+ \int_0^{1/k_n} f_n(x)dx \\ &=: T_3 + T_4 + T_5. \end{aligned}$$

Remark that

$$\begin{aligned} T_3 &= -\frac{1}{2k_n}\left(\frac{1}{k_n}\sum_{j=1}^{k_n-1} f_n^{(1)}\Big(\frac{j}{k_n}\Big) - \int_{1/k_n}^1 f_n^{(1)}(t)dt\right) - \frac{1}{2k_n}\int_{1/k_n}^1 f_n^{(1)}(t)dt \\ &=: -\frac{1}{2k_n}T_6 + T_7. \end{aligned}$$

Since $f_n^{(1)}$ is positive and decreasing on $\left[\frac{1}{k_n}, 1\right]$ for n sufficiently large, we can prove that

$$\begin{cases} |T_4| \leq \frac{1}{2k_n^2}\left|f_n^{(1)}\left(\frac{1}{k_n}\right) - f_n^{(1)}(1)\right| = o\left(\frac{1}{k_n}\right) \\ T_5 = O\left(\frac{1}{k_n}\right), \\ |T_6| \leq \frac{1}{k_n}\left|f_n^{(1)}\left(\frac{1}{k_n}\right) - f_n^{(1)}(1)\right| = o(1) \\ T_7 = -\frac{1}{2k_n}\left(f_n(1) - f_n\left(\frac{1}{k_n}\right)\right) = o\left(\frac{1}{k_n}\right) \end{cases}$$

and consequently $T_1 = O\left(\frac{1}{k_n}\right)$. Besides, a direct application of Lemma 4 provides

$$T_2 = 1 - \alpha_n + O(\alpha_n^2).$$

Therefore $\overline{x}_{k_n} = 1 - \alpha_n + O\left(\frac{1}{k_n}\right) + O(\alpha_n^2)$. Now, we can check that

$$\mu_p = \alpha_n^p \left\{ \frac{1}{k_n} \sum_{j=1}^{k_n} \left(\log\left(\frac{j}{k_n}\right) + 1 \right)^p + R_n \right\}$$

where

$$R_n = \frac{1}{k_n} \sum_{j=1}^{k_n - 1} \left\{ \left(\log\left(\frac{j}{k_n}\right) + 1 + \varepsilon_n \right)^p - \left(\log\left(\frac{j}{k_n}\right) + 1 \right)^p \right\}$$

with $\varepsilon_n = O\left(\alpha_n \log^2 k_n\right) + O\left(\frac{1}{k\alpha_n}\right)$ which tends to 0 by assumption.

Since $\frac{1}{C_p}\frac{1}{k_n}\sum_{j=1}^{k_n}\left(\log\left(\frac{j}{k_n}\right)+1\right)^p \to 1$, in order to conclude the proof of Lemma 5, we only have to remark that $R_n \to 0$. □

6.2 Proof of Lemma 1

Remark that

$$\begin{aligned} \alpha_{j,n} &:= j \log \frac{n}{j}\left[A_n(j) - A_n(j+1)\right] \\ &= \theta j \log \frac{n}{j} \log \frac{E_{n-j+1,n}}{E_{n-j,n}} \\ &= \theta \log \frac{n}{j} \frac{j(E_{n-j+1,n} - E_{n-j,n})}{E^*_{n-j,n}} \\ &\stackrel{d}{=} \theta f_j \frac{\log \frac{n}{j}}{E^*_{n-j,n}} \end{aligned}$$

where $E^*_{n-j,n} \in \Big[E_{n-j,n}; E_{n-j+1,n}\Big]$. Consequently, from Lemma 3,

$$\begin{aligned} \alpha_{j,n} &= \theta f_j + O_{\mathbb{P}}\Big(\frac{1}{\log \frac{n}{k_n}}\Big) \\ &= \theta f_j + o_{\mathbb{P}}\Big(b\Big(\log \frac{n}{k_n}\Big)\Big), \end{aligned} \tag{10}$$

by the assumption $xb(x) \to \infty$ as $x \to \infty$ with a $o_{\mathbb{P}}-$term which is uniform in j. Lemma 1 is therefore proved. □

6.3 Proof of Lemma 2

We consider

$$\beta_{j,n} := j \log \frac{n}{j}\Big[B_n(j) - B_n(j+1)\Big].$$

In order to study this term, we will use the notations $\lambda_{1j} = \frac{E_{n-j+1,n}}{E_{n-k_n+1,n}}$, $\lambda_{2j} = \frac{E_{n-j,n}}{E_{n-k_n+1,n}}$ and $y_{k_n} = E_{n-k_n+1,n}$, and we rewrite $\beta_{j,n}$ as

$$\beta_{j,n} = j \log \frac{n}{j}\Big\{ \log \ell\Big(\lambda_{2j} \frac{\lambda_{1j}}{\lambda_{2j}} y_{k_n}\Big) - \log \ell\Big(\lambda_{2j}\, y_{k_n}\Big)\Big\}.$$

It is clear that $1 \le \frac{\lambda_{1j}}{\lambda_{2j}} \xrightarrow{\mathbb{P}} 1$ uniformly in j by Lemma 3 and therefore for $n \ge N_0$, $\frac{\lambda_{1j}}{\lambda_{2j}} \in [1,2]$ in probability. Under our assumption $(R_\ell(b,\rho))$ on the slowly varying function, we deduce that

$$\beta_{j,n} = j \log \frac{n}{j}\Big\{ b(\lambda_{2j} y_{k_n}) K_\rho\Big(\frac{\lambda_{1j}}{\lambda_{2j}}\Big)(1 + o_{\mathbb{P}}(1))\Big\}.$$

Now, since $\lambda_{2j} \xrightarrow{\mathbb{P}} 1$ uniformly in j and $b(.)$ is regularly varying with index ρ, $b(\lambda_{2j} y_{k_n}) = \lambda_{2j}^{\rho} b(y_{k_n})(1 + o_{\mathbb{P}}(1))$ with a $o_{\mathbb{P}}(1)$-term uniform in j.

Therefore

$$\beta_{j,n} = j \log \frac{n}{j}\, b(y_{k_n})\Big\{ \lambda_{2j}^{\rho} K_\rho\Big(\frac{\lambda_{1j}}{\lambda_{2j}}\Big)(1 + o_{\mathbb{P}}(1))\Big\}.$$

Again, uniformly in j,

$$K_\rho\Big(\frac{\lambda_{1j}}{\lambda_{2j}}\Big) = \Big(\frac{\lambda_{1j}}{\lambda_{2j}} - 1\Big)(1 + o_{\mathbb{P}}(1)),$$

which implies that $\beta_{j,n}$ can be rewritten as follows:

$$\beta_{j,n} = -\, j \log \frac{n}{j}\, b(y_{k_n})(\lambda_{2j} - \lambda_{1j})\lambda_{2j}^{\rho-1}(1 + o_{\mathbb{P}}(1)).$$

Therefore, we have

$$\beta_{j,n} = f_j\left(\frac{\log \frac{n}{j}}{\log \frac{n}{k_n}}\right)^{\rho} b(y_{k_n})(1 + o_{\mathbb{P}}(1)),$$

with a $o_{\mathbb{P}}(1)$-term which is uniform in j. This achieves the proof of Lemma 2. □

Remark that, since $\frac{\log(n/j)}{\log(n/k_n)} \to 1$ uniformly in j, one also has

$$\beta_{j,n} = f_j\left(\frac{\log \frac{n}{j}}{\log \frac{n}{k_n}}\right)^{-1} b(y_{k_n})(1 + o_{\mathbb{P}}(1)),$$

with a $o_{\mathbb{P}}(1)$-term which is uniform in j, and this proves Remark 1.

6.4 Proof of Theorem 2

From model (6), we infer that

$$\sqrt{k_n}\left(\check{\theta}_n - \theta - b\Big(\log \frac{n}{k_n}\Big)\frac{1}{k_n}\sum_{j=1}^{k_n}\left(\frac{\log \frac{n}{j}}{\log \frac{n}{k_n}}\right)^{\rho}\right)$$

$$= \sqrt{k_n}\,\theta\,\frac{1}{k_n}\sum_{j=1}^{k_n}(f_j - 1) + \sqrt{k_n} b\Big(\log \frac{n}{k_n}\Big)\frac{1}{k_n}\sum_{j=1}^{k_n}\left(\frac{\log \frac{n}{j}}{\log \frac{n}{k_n}}\right)^{\rho}(f_j - 1) + o_{\mathbb{P}}\Big(\sqrt{k_n}\, b\Big(\log \frac{n}{k_n}\Big)\Big).$$

Now, an application of Tchebychev's inequality gives that

$$\frac{1}{k_n}\sum_{j=1}^{k_n}\left(\frac{\log \frac{n}{j}}{\log \frac{n}{k_n}}\right)^{\rho}(f_j - 1) = o_{\mathbb{P}}(1).$$

Then, under our assumptions, Theorem 2 follows by an application of the Central Limit Theorem. □

6.5 Proof of Theorem 3

From Remark 1, we have

$$\begin{aligned}\frac{\sqrt{k_n}}{\log \frac{n}{k_n}}\Big(\widehat{\theta}_n - \theta\Big) &= \frac{\sqrt{k_n}}{\log \frac{n}{k_n}}\frac{1}{k_n}\sum_{j=1}^{k_n}\left(\theta + b\Big(\log \frac{n}{k_n}\Big)x_j\right)\left(1 - \frac{x_j - \overline{x}_{k_n}}{\mu_2}\overline{x}_{k_n}\right)(f_j - 1)\\ &\quad + o_{\mathbb{P}}\left(\frac{\sqrt{k_n}}{\log \frac{n}{k_n}} b\Big(\log \frac{n}{k_n}\Big)\right).\end{aligned}$$

Since we have (9), the $o_{\mathbb{P}}$-term is negligible. The first term can be viewed as a sum of a weighted mean of independent and identically distributed variables. Now, using

Lyapounov's theorem, we only have to show that

$$\lim_{k_n\to\infty}\frac{1}{s_{k_n}^4}\sum_{j=1}^{k_n}\mathbb{E}X_j^4=0,$$

where $X_j=\Big(\theta+b\Big(\log\frac{n}{k_n}\Big)x_j\Big)\Big(1-\frac{x_j-\overline{x}_{k_n}}{\mu_2}\overline{x}_{k_n}\Big)(f_j-1)$, $j=1,...,k_n$ and $s_{k_n}^2=\sum_{j=1}^{k_n}\mathrm{Var}X_j$. We remark that

$$s_{k_n}^2\sim\theta^2\sum_{j=1}^{k_n}\left(1-\frac{x_j-\overline{x}_{k_n}}{\mu_2}\overline{x}_{k_n}\right)^2\quad\text{as } n\to\infty$$

and

$$\sum_{j=1}^{k_n}\mathbb{E}X_j^4\sim 9\theta^4\sum_{j=1}^{k_n}\left(1-\frac{x_j-\overline{x}_{k_n}}{\mu_2}\overline{x}_{k_n}\right)^4\quad\text{as } n\to\infty$$

from which we deduce by direct computations that

$$\begin{aligned}\frac{1}{s_{k_n}^4}\sum_{j=1}^{k_n}\mathbb{E}X_j^4 &\sim \frac{9}{k_n}\frac{\mu_2^4+6(\overline{x}_{k_n})^2\mu_2^3-4(\overline{x}_{k_n})^3\mu_2\mu_3+(\overline{x}_{k_n})^4\mu_4}{[\mu_2^2+(\overline{x}_{k_n})^2\mu_2]^2}\\ &\sim \frac{9C_4}{k_n}\end{aligned}$$

by Lemma 5. Our Theorem 3 now follows from the fact that

$$s_{k_n}^2\sim\theta^2k_n\log^2(n/k_n).$$

□

References

[1] Abramowitz, M., Stegun, I., (1972), *Handbook of Mathematical Functions*, Dover.

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

[3] 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.

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

[5] Beirlant, J., Dierckx, G., Guillou, A., Starica, C., (2002), On exponential representations of log-spacings of extreme order statistics, *Extremes*, **5** (2), 157–180.

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

[7] Feuerverger, A., Hall, P., (1999), Estimating a Tail Exponent by Modelling Departure from a Pareto Distribution, *Annals of Statistics*, **27**, 760–781.

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

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

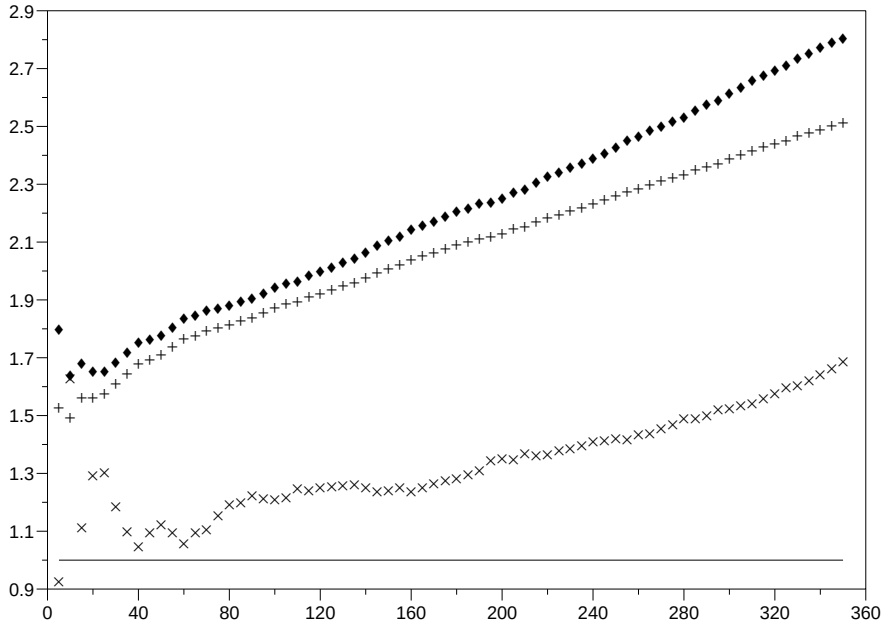
[10] Girard, S., (2004), A Hill type estimate of the Weibull tail-coefficient, *Communication in Statistics - Theory and Methods*, **33**(2), 205–234.

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

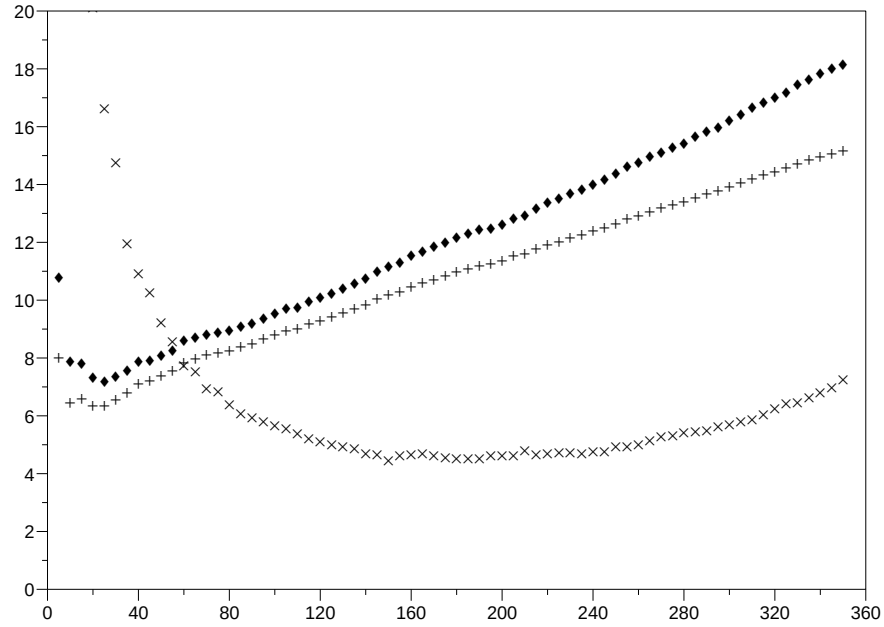
[12] Shorack, G.R., Wellner, J.A., (1986), *Empirical Processes with Applications to Statistics*, Wiley New York.

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

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

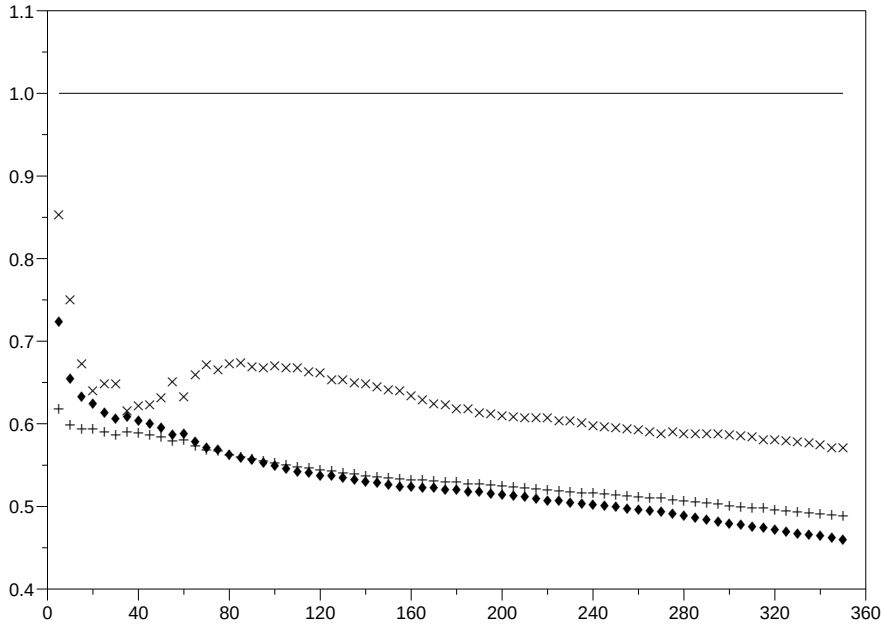


(a) Mean as a function of k_n

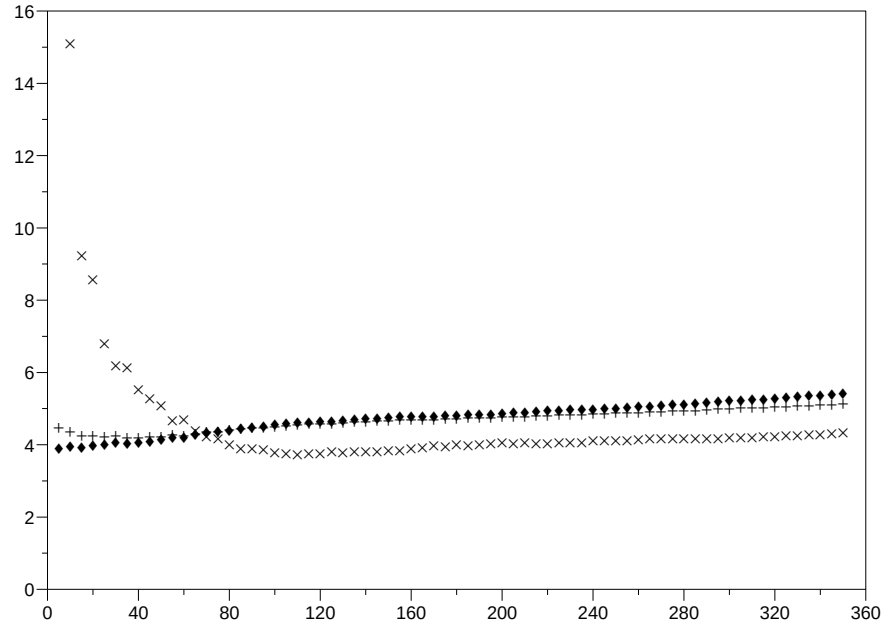


(b) Mean square error as a function of k_n

Figure 1: Comparison of estimates $\widehat{\theta}_n$ ($\times\times\times$), $\widetilde{\theta}_n$ ($\blacklozenge\blacklozenge\blacklozenge$) and $\check{\theta}_n$ ($+++$) for the $\Gamma(0.25, 1)$ distribution. In (a), the straight line is the true value of θ.

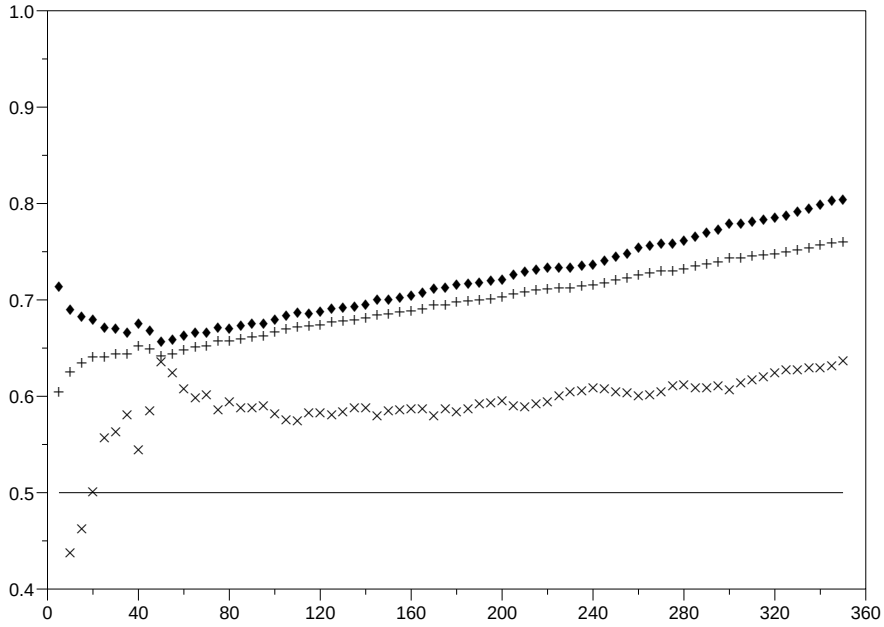


(a) Mean as a function of k_n

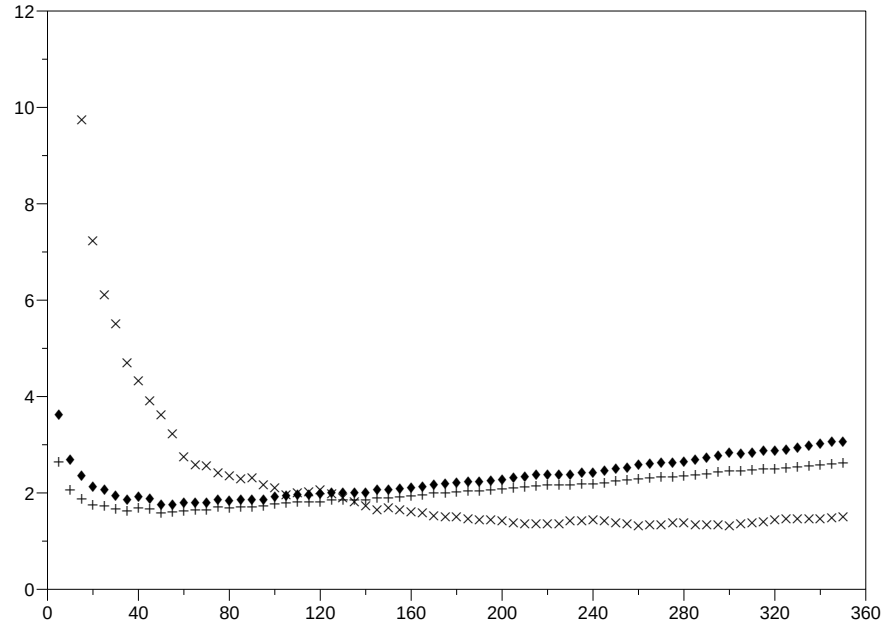


(b) Mean square error as a function of k_n

Figure 2: Comparison of estimates $\widehat{\theta}_n$ (×××), $\widetilde{\theta}_n$ (◆◆◆) and $\check{\theta}_n$ (+++) for the $\Gamma(4, 1)$ distribution. In (a), the straight line is the true value of θ.

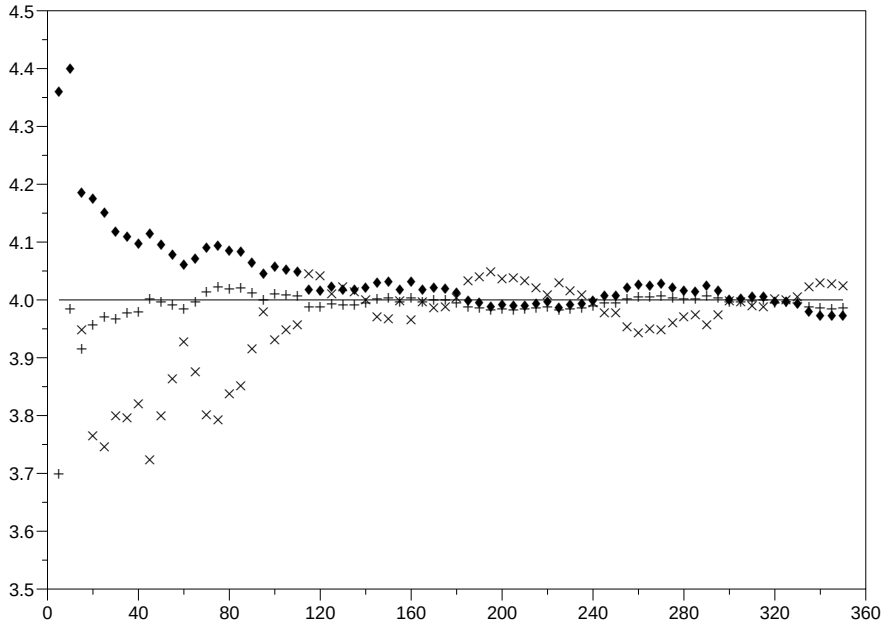


(a) Mean as a function of k_n

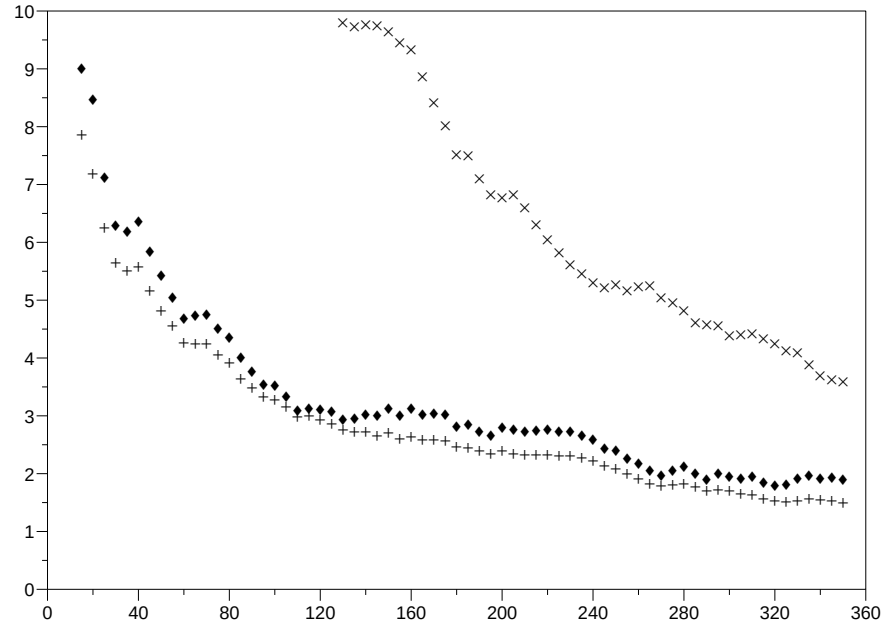


(b) Mean square error as a function of k_n

Figure 3: Comparison of estimates $\widehat{\theta}_n$ ($\times\times\times$), $\widetilde{\theta}_n$ ($\blacklozenge\blacklozenge\blacklozenge$) and $\check{\theta}_n$ ($+++$) for the $\mathcal{N}(1.1,1)$ distribution. In (a), the straight line is the true value of θ.

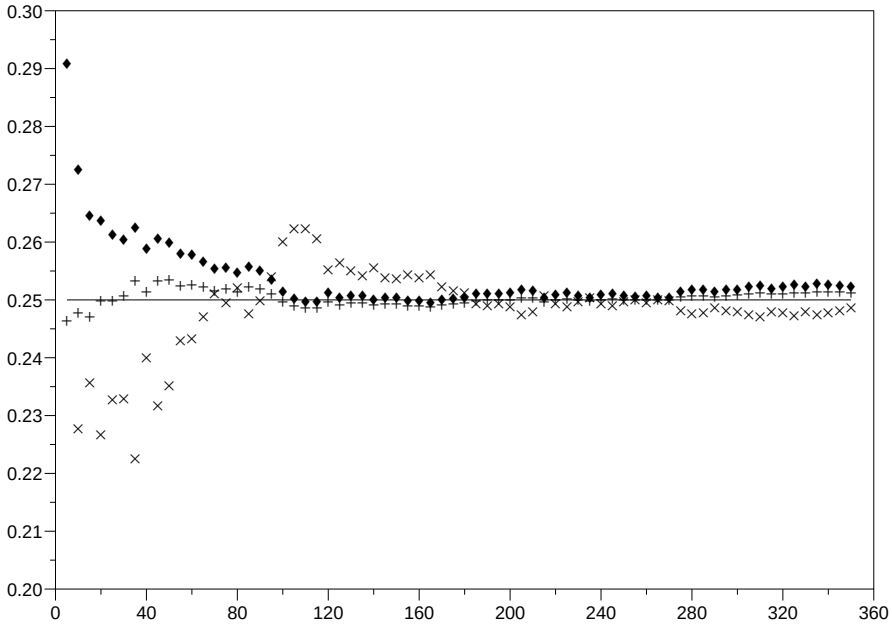


(a) Mean as a function of k_n

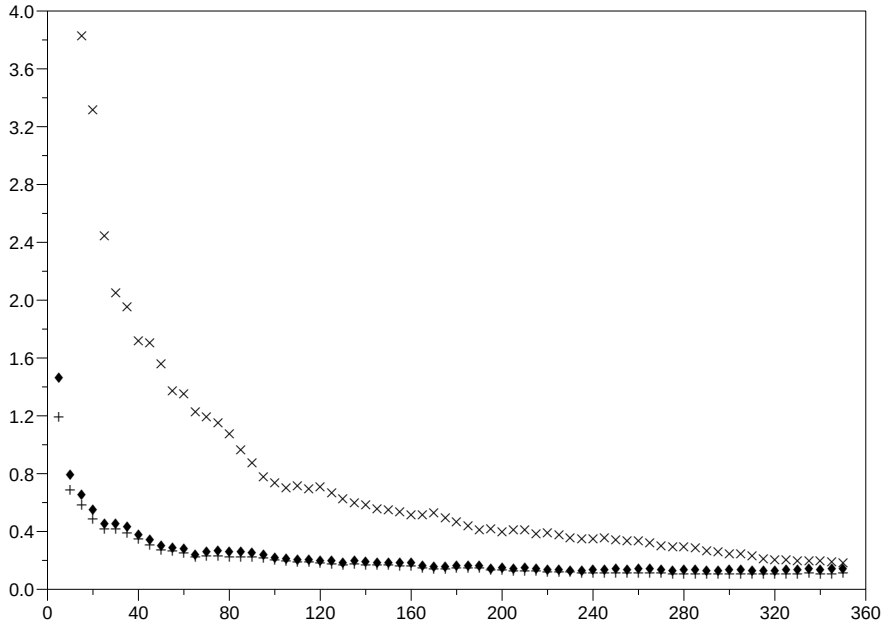


(b) Mean square error as a function of k_n

Figure 4: Comparison of estimates $\widehat{\theta}_n$ ($\times\times\times$), $\widetilde{\theta}_n$ ($\blacklozenge\blacklozenge\blacklozenge$) and $\check{\theta}_n$ ($+++$) for the $\mathcal{W}(0.25, 0.25)$ distribution. In (a), the straight line is the true value of θ.



(a) Mean as a function of k_n



(b) Mean square error as a function of k_n

Figure 5: Comparison of estimates $\widehat{\theta}_n$ (×××), $\widetilde{\theta}_n$ (◆◆◆) and $\check{\theta}_n$ (+++) for the $\mathcal{W}(4,4)$ distribution. In (a), the straight line is the true value of θ.