Estimating an endpoint with high order moments in the Weibull domain of attraction

Stéphane Girard, Armelle Guillou, Gilles Stupfler

Estimating an endpoint with high order moments in the Weibull domain of attraction

Stéphane Girard[(1)], Armelle Guillou[(2)] & Gilles Stupfler[(2)]

[(1)] Team Mistis, INRIA Rhône-Alpes & LJK, Inovallée, 655, av. de l'Europe, Montbonnot, 38334 Saint-Ismier cedex, France

[(2)] Université de Strasbourg & CNRS, IRMA, UMR 7501, 7 rue René Descartes, 67084 Strasbourg cedex, France

Abstract. We present a method for estimating the endpoint of a unidimensional sample when the distribution function belongs to the Weibull-max domain of attraction. The approach relies on transforming the variable of interest and then using high order moments of the positive variable obtained this way. It is assumed that the order of the moments goes to infinity. We give conditions on the rate of divergence to get the weak and strong consistency as well as the asymptotic normality of the estimator. The good performance of the estimator is illustrated on some finite sample situations.

AMS Subject Classifications: 62G32, 62G05.

Keywords: Endpoint estimation, high order moments, consistency, asymptotic normality.

1 Introduction

Let $(X_1, \ldots, X_n)$ be independent copies of a random variable X, where the distribution of X has a finite right endpoint θ, with θ being unknown. We are interested in estimating θ. Recent work on endpoint estimation includes a Bayesian likelihood approach (Hall and Wang, 2005), censored likelihood estimators (Li *et al.*, 2011a) and the empirical likelihood method (Li *et al.*, 2011b). In Girard *et al.* (2011), a new estimator of θ, based upon the use of high order moments of the X_k, is introduced. From a practical point of view, taking high order moments gives (exponentially) more weight to the X_k close to θ. A thorough study of the estimator is carried out when X is a positive random variable and its distribution function belongs to the Weibull max-domain of attraction. In this paper, we address the problem of estimating the endpoint θ with high order moments when the positivity assumption on X is dropped. We cannot use moments of the variable of interest X, since $|X|$ could have an infinite mean. To overcome this problem, it can be noted that the random

variable e^X has a bounded support $\left[0,\, e^\theta\right]$. Moreover, letting $\mu_{p_n} := \mathbb{E}\left(e^{p_n X}\right)$, $p_n \to \infty$ be the p_nth order moment of e^X, yields, for all $u \geq 1$, $\mu_{p_n}/\mu_{p_n+u} \to e^{-u\theta}$ as $n \to \infty$ (see Lemma 1 in the Appendix). For all $a > 0$, it follows that

$$\Theta_n := \frac{1}{a}\left\{\ln\left[\frac{\mu_{p_n}}{\mu_{p_n+1}}\right] - \ln\left[\frac{\mu_{(a+1)p_n}}{\mu_{(a+1)p_n+a+1}}\right]\right\} = \theta(1+\mathrm{o}(1)). \tag{1}$$

We therefore introduce an estimator using high order moments of the variable e^X. Replacing the true moment μ_{p_n} with its empirical counterpart $\widehat{\mu}_{p_n}$ in the expression of Θ_n yields

$$\widehat{\theta}_n := \frac{1}{a}\left\{\ln\left[\frac{\widehat{\mu}_{p_n}}{\widehat{\mu}_{p_n+1}}\right] - \ln\left[\frac{\widehat{\mu}_{(a+1)p_n}}{\widehat{\mu}_{(a+1)p_n+a+1}}\right]\right\}$$

where (p_n) is a positive, nonrandom sequence such that $p_n \to \infty$, $a > 0$ and

$$\widehat{\mu}_{p_n} := \frac{1}{n}\sum_{i=1}^{n} e^{p_n X_i}$$

is the classical moment estimator of μ_{p_n}. It is shown in Section 2 that $\widehat{\theta}_n$ is consistent without any parametric assumption on the distribution of X. Moreover, we state and prove that $\widehat{\theta}_n$ is asymptotically Gaussian when the distribution function of X belongs to the Weibull max-domain of attraction. Some simulations are proposed in Section 3 to illustrate the efficiency of our estimator, and to compare it with estimators of the endpoint estimation literature. Auxiliary results are postponed to the Appendix.

2 Main results

Let us first state the consistency of the estimator. The only assumption is

(A_0) X has a finite right endpoint θ.

Theorem 1. *If* (A_0) *holds and* $n\,\mu_{(a+1)p_n}/e^{(a+1)p_n\theta} \to \infty$ *as* $n \to \infty$, *then* $\widehat{\theta}_n \xrightarrow{\mathbb{P}} \theta$ *as* $n \to \infty$.

Proof. We first show that, provided $n\,\mu_{p_n}/e^{p_n\theta} \to \infty$, the high order moment μ_{p_n} can be replaced by its empirical counterpart $\widehat{\mu}_{p_n}$ in (1). For all $\varepsilon > 0$, Chebyshev's inequality leads to

$$\mathbb{P}\left(\left|\frac{\widehat{\mu}_{p_n}}{\mu_{p_n}} - 1\right| > \varepsilon\right) \leq \frac{1}{\varepsilon^2}\,\frac{\mathrm{Var}\left(\widehat{\mu}_{p_n}\right)}{\mu_{p_n}^2} \leq \frac{1}{n\varepsilon^2}\,\frac{\mu_{2p_n}}{\mu_{p_n}^2},$$

where

$$\frac{\mu_{2p_n}}{\mu_{p_n}} = e^{p_n\theta}\,\frac{\mu_{2p_n}/e^{2p_n\theta}}{\mu_{p_n}/e^{p_n\theta}} \leq e^{p_n\theta} \tag{2}$$

and therefore

$$\mathbb{P}\left(\left|\frac{\widehat{\mu}_{p_n}}{\mu_{p_n}} - 1\right| > \varepsilon\right) \leq \frac{1}{\varepsilon^2}\,\frac{e^{p_n\theta}}{n\,\mu_{p_n}} \to 0$$

as $n \to \infty$. As a consequence, $\widehat{\mu}_{p_n}/\mu_{p_n} \xrightarrow{\mathbb{P}} 1$ as $n \to \infty$.

Since $\mu_{(a+1)p_n}/e^{(a+1)p_n\theta} \leq \mu_{p_n}/e^{p_n\theta}$, it follows that $n\,\mu_{p_n}/e^{p_n\theta} \to \infty$. Lemma 1 thus yields $n\,\mu_{p_n+1}/e^{(p_n+1)\theta} \to \infty$ and $n\,\mu_{(a+1)p_n+a+1}/e^{((a+1)p_n+a+1)\theta} \to \infty$ as $n \to \infty$. Consequently, from (1), $\widehat{\theta}_n = \Theta_n + \mathrm{o}_{\mathbb{P}}(1)$. Lemma 1 then entails $\widehat{\theta}_n \xrightarrow{\mathbb{P}} \theta$. ∎

Under a somewhat stronger condition of the rate of divergence of (p_n), a strong consistency result can be established for $\widehat{\theta}_n$:

Theorem 2. *If* (A_0) *holds and* $\dfrac{1}{\ln n}\,\dfrac{n\,\mu_{(a+1)p_n}}{e^{(a+1)p_n\theta}} \to \infty$ *as* $n \to \infty$, *then* $\widehat{\theta}_n \xrightarrow{a.s.} \theta$ *as* $n \to \infty$.

Proof. The result being obvious when $\mathbb{P}(X = \theta) = 1$, let us assume that $\mathbb{P}(X = \theta) < 1$. We start by showing that

$$\frac{1}{\ln n}\,\frac{n\,\mu_{p_n}}{e^{p_n\theta}} \to \infty \quad \Rightarrow \quad \frac{\widehat{\mu}_{p_n}}{\mu_{p_n}} \xrightarrow{\text{a.s.}} 1 \quad \text{as } n \to \infty. \tag{3}$$

To this end, let $Y_i = e^{X_i - \theta}$. Then $|Y_i^{p_n} - \mathbb{E}(Y_i^{p_n})| \leq 1$ a.s. and

$$\frac{\widehat{\mu}_{p_n} - \mu_{p_n}}{e^{p_n\theta}} = \frac{1}{n}\sum_{i=1}^{n}\left\{Y_i^{p_n} - \mathbb{E}(Y_i^{p_n})\right\}$$

is a mean of bounded, centered, independent and identically distributed random variables. Defining

$$\tau_n := \varepsilon\,\frac{n\,\mu_{p_n}}{e^{p_n\theta}} \quad \text{and} \quad \lambda_n := \varepsilon\,\frac{\mu_{p_n}}{e^{p_n\theta}}\,\frac{1}{\mathrm{Var}(Z_1^{p_n})} = \varepsilon\,\frac{\mu_{p_n}}{e^{p_n\theta}}\,\frac{e^{2p_n\theta}}{\mu_{2p_n} - \mu_{p_n}^2},$$

Bernstein's inequality (see Hoeffding, 1963) gives, for all $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\frac{\widehat{\mu}_{p_n}}{\mu_{p_n}} - 1\right| > \varepsilon\right) = \mathbb{P}\left(\left|\frac{\widehat{\mu}_{p_n} - \mu_{p_n}}{e^{p_n\theta}}\right| > \varepsilon\,\frac{\mu_{p_n}}{e^{p_n\theta}}\right) \leq \exp\left(-\frac{\tau_n\lambda_n}{2(1 + \lambda_n/3)}\right).$$

Note that since $e^{X-\theta} \in [0,\,1]$ a.s., we have $\mu_{p_n}/e^{p_n\theta} \to \mathbb{P}(X = \theta)$ as $n \to \infty$. Then, from (2),

$$\frac{1}{\lambda_n} \leq \frac{1}{\varepsilon}\left[1 - \frac{\mu_{p_n}}{e^{p_n\theta}}\right] \to \frac{1 - \mathbb{P}(X = \theta)}{\varepsilon} > 0$$

as $n \to \infty$, and therefore, for sufficiently large n, there exists a constant $C_\varepsilon > 0$ such that

$$\mathbb{P}\left(\left|\frac{\widehat{\mu}_{p_n}}{\mu_{p_n}} - 1\right| > \varepsilon\right) \leq \exp\left(-C_\varepsilon\,\frac{n\,\mu_{p_n}}{e^{p_n\theta}}\right).$$

Borel-Cantelli's lemma thus yields $\widehat{\mu}_{p_n}/\mu_{p_n} \xrightarrow{\text{a.s.}} 1$ as $n \to \infty$. Using once again the inequality $\mu_{(a+1)p_n}/e^{(a+1)p_n\theta} \leq \mu_{p_n}/e^{p_n\theta}$, Lemma 1 and (3), the result is now straightforward. ■

Let us now establish the asymptotic distribution of our estimator. To this end, additional assumptions are introduced on the survival function $\overline{F} = 1 - F$ of X:

(A_1) $\forall\, x < \theta$, $\overline{F}(x) = (\theta - x)^\alpha\, L((\theta - x)^{-1})$ where $\theta \in \mathbb{R}$, $\alpha > 0$ and L is a slowly varying function at infinity, *i.e.* such that $L(ty)/L(y) \to 1$ as $y \to \infty$ for all $t > 0$.

(A_2) $\forall\, x \geq 0$, $L(x) = c\,\exp\left(\int_1^x \eta(t)\,t^{-1}\,dt\right)$, where $c > 0$ and η is a bounded function tending to 0 at infinity, continuously differentiable on $(0,\,\infty)$, ultimately monotonic and non identically 0, such that $|\eta'|$ is regularly varying and there exists $\nu \leq 0$ with $x\,\eta'(x)/\eta(x) \to \nu$ as $x \to \infty$.

It is well-known that (A_1) holds if and only if F belongs to the Weibull max-domain of attraction, see Fisher and Tippett (1928) and Gnedenko (1943). (A_2) is the Karamata representation for the normalized slowly varying function L, see Bingham *et al.* (1987), p. 15. Under (A_2), the function $|\eta|$ is ultimately non-increasing and regularly varying at infinity with index ν, see Bingham *et al.*

(1987), paragraph 1.4.2, and the function $x \mapsto x\,|\eta'(x)|$ is regularly varying with index ν. In the extreme-value framework, ν is referred to as the second order parameter and (A_2) is a second order condition.

We can now state the asymptotic normality of $\widehat{\theta}_n$:

Theorem 3. *Assume (A_1) and (A_2) hold. If $n\,p_n^{-\alpha}\,L(p_n) \to \infty$ and $n\,p_n^{-\alpha}\,L(p_n)\,\eta^2(p_n) \to 0$, then*

$$v_n\left(\widehat{\theta}_n - \theta\right) \xrightarrow{d} \mathcal{N}\left(0,\, V(\alpha,\, a)\right) \quad \text{as } n \to \infty,$$

with $v_n = \sqrt{n\,L(p_n)}\,p_n^{-\alpha/2+1}$ and

$$V(\alpha,\, a) = \frac{\alpha+1}{a^2\,\Gamma(\alpha)}\left[2^{-\alpha-2} - 2\frac{(a+1)^{\alpha+1}}{(a+2)^{\alpha+2}} + 2^{-\alpha-2}(a+1)^{\alpha}\right].$$

Proof. Let us remark that $v_n\left(\widehat{\theta}_n - \theta\right) = v_n\left(\widehat{\theta}_n - \Theta_n\right) + v_n\left(\Theta_n - \theta\right)$ and focus on the random term. Our goal is to establish that $v_n\left(\widehat{\theta}_n - \Theta_n\right) \xrightarrow{d} \mathcal{N}\left(0,\, V(\alpha,\, a)\right)$ as $n \to \infty$. To this end, using the delta-method, it is enough to prove that the sequence of random variables

$$\xi_n := \frac{e^{-a\theta}}{a\,\sqrt{V(\alpha,\, a)}}\,v_n\left(e^{a\widehat{\theta}_n} - e^{a\Theta_n}\right)$$

converges in distribution to a standard Gaussian random variable. Notice that the change of variable $z = (\theta - x)^{-1}$ yields

$$\mu_p = p\int_{-\infty}^{\theta} e^{px}\,\overline{F}(x)\,dx = p^{-\alpha}\,e^{p\theta}\,L(p)\left[\Gamma(\alpha+1) + I_1\,\varepsilon_1(p) + I_2\,\varepsilon_2(p)\right] \tag{4}$$

where $I_1 = \displaystyle\int_1^{\infty} z^{\alpha}\,e^{-z}\,dz$, $I_2 = \displaystyle\int_0^{1} z^{\alpha}\,e^{-z}\,dz$ and

$$\begin{aligned}
\varepsilon_1(p) &= \frac{1}{I_1}\int_0^1 e^{-1/z}\,z^{-\alpha-3}\,\frac{L_1(pz)}{L_1(p)}\,dz - 1, \qquad L_1(z) = z\,L(z),\\
\varepsilon_2(p) &= \frac{1}{I_2}\int_1^{\infty} e^{-1/z}\,z^{-\alpha-1}\,\frac{L_2(pz)}{L_2(p)}\,dz - 1, \qquad L_2(z) = \frac{L(z)}{z}.
\end{aligned}$$

Using (4) together with Lemma 2i) entails

$$\xi_n = e^{-(a+1)\theta}\,\mu_{p_n+1}\,u_{n,\,a}\left(\Delta_n^{(1)} + \Delta_n^{(2)}\right)(1 + \mathrm{o}(1)) \tag{5}$$

with

$$\begin{aligned}
u_{n,\,a} &= \frac{1}{a\,\Gamma(\alpha+1)}\sqrt{\frac{1}{V(\alpha,\, a)}}\,\frac{p_n^{\alpha}\,v_n}{e^{p_n\theta}\,L(p_n)},\\
\Delta_n^{(1)} &= \left[\frac{\widehat{\mu}_{p_n}}{\widehat{\mu}_{p_n+1}} - \frac{\mu_{p_n}}{\mu_{p_n+1}}\right]\frac{\widehat{\mu}_{(a+1)p_n+a+1}}{\widehat{\mu}_{(a+1)p_n}},\\
\Delta_n^{(2)} &= \left[\frac{\widehat{\mu}_{(a+1)p_n+a+1}}{\widehat{\mu}_{(a+1)p_n}} - \frac{\mu_{(a+1)p_n+a+1}}{\mu_{(a+1)p_n}}\right]\frac{\mu_{p_n}}{\mu_{p_n+1}}.
\end{aligned}$$

Rewriting $\Delta_n^{(1)}$ and $\Delta_n^{(2)}$ yields

$$\xi_n = u_{n,\,a}\left[\zeta_n^{(1)} + \left(e^{-(a+1)\theta}\,\frac{\widehat{\mu}_{(a+1)p_n+a+1}}{\widehat{\mu}_{(a+1)p_n}}\cdot\frac{\mu_{p_n+1}}{\widehat{\mu}_{p_n+1}} - 1\right)\zeta_n^{(2)} + \left(\frac{\mu_{(a+1)p_n}}{\widehat{\mu}_{(a+1)p_n}} - 1\right)\zeta_n^{(3)}\right](1 + \mathrm{o}(1)),$$

where, setting $\nu_p = \widehat{\mu}_p - \mu_p$,

$$\begin{aligned}
\zeta_n^{(1)} &= \zeta_n^{(2)} + \zeta_n^{(3)}, \\
\text{with} \quad \zeta_n^{(2)} &= \nu_{p_n} - \frac{\mu_{p_n}}{\mu_{p_n+1}} \nu_{p_n+1}, \\
\text{and} \quad \zeta_n^{(3)} &= e^{-(a+1)\theta} \, \frac{\mu_{p_n} \, \mu_{(a+1)p_n+a+1}}{\mu_{(a+1)p_n}^2} \left[-\nu_{(a+1)p_n} + \frac{\mu_{(a+1)p_n}}{\mu_{(a+1)p_n+a+1}} \, \nu_{(a+1)p_n+a+1} \right].
\end{aligned}$$

In view of the above consistency results, it follows that

$$\xi_n = u_{n,\,a} \left[\zeta_n^{(1)} + \mathrm{o}_{\mathbb{P}} \left(\zeta_n^{(2)} \right) + \mathrm{o}_{\mathbb{P}} \left(\zeta_n^{(3)} \right) \right] (1 + \mathrm{o}(1)),$$

and it is therefore sufficient to show that

$$u_{n,\,a} \, \zeta_n^{(1)} \xrightarrow{d} \mathcal{N}(0, 1), \tag{6a}$$

$$u_{n,\,a} \, \zeta_n^{(2)} \xrightarrow{d} \mathcal{N}(0,\, C_2), \tag{6b}$$

$$u_{n,\,a} \, \zeta_n^{(3)} \xrightarrow{d} \mathcal{N}(0,\, C_3), \tag{6c}$$

where C_2 and C_3 are suitable constants. Let us then write $\zeta_n^{(1)} = \sum_{k=1}^{n} S_{n,\,k}$, where

$$\begin{aligned}
S_{n,\,k} &= \frac{1}{n} \left[e^{p_n X_k}, \; e^{(p_n+1) X_k}, \; e^{(a+1) p_n X_k}, \; e^{[(a+1)p_n+a+1] X_k} \right] A_n, \\
A_n &= [a_{n,\,0}, \; a_{n,\,1}, \; a_{n,\,2}, \; a_{n,\,3}]^t, \\
a_{n,\,0} &= 1, \\
a_{n,\,1} &= -\frac{\mu_{p_n}}{\mu_{p_n+1}}, \\
a_{n,\,2} &= -e^{-(a+1)\theta} \, \frac{\mu_{p_n} \, \mu_{(a+1)p_n+a+1}}{\mu_{(a+1)p_n}^2}, \\
a_{n,\,3} &= e^{-(a+1)\theta} \, \frac{\mu_{p_n}}{\mu_{(a+1)p_n}}.
\end{aligned}$$

Since the $S_{n,\,k}$, $1 \leq k \leq n$, are independent, identically distributed and centered random variables, we shall prove that

$$\frac{\mathbb{E} \left| S_{n,\,1} \right|^3}{\sqrt{n} \left[\mathrm{Var} \left(S_{n,\,1} \right) \right]^{3/2}} \to 0$$

as $n \to \infty$, and use Lyapounov's theorem (see e.g. Billingsley, 1979, p. 312) to obtain the asymptotic normality of $\zeta_n^{(1)}$.

An equivalent of $\mathrm{Var} \left(S_{n,\,1} \right)$ is obtained by using (4) and applying Lemma 2 to get

$$\mathrm{Var} \left(S_{n,\,1} \right) = a^2 \, \Gamma^2(\alpha + 1) \, V(\alpha,\, a) \, \frac{1}{n} \, e^{2 p_n \theta} \, p_n^{-\alpha-2} \, L(p_n) \, (1 + \mathrm{o}(1)).$$

To control $\mathbb{E} \left| S_{n,\,1} \right|^3$, introducing $Y = X - \theta$, Hölder's inequality yields

$$\begin{aligned}
\frac{\mathbb{E} \left| S_{n,\,1} \right|^3}{n^{-3} \, e^{3 p_n \theta}} &\leq 4 \, \mathbb{E} \left| e^{p_n Y} \left[a_{n,\,0} + a_{n,\,1} \, e^{\theta} \, e^{Y} \right] \right|^3 \\
&+ 4 \, \mathbb{E} \left| e^{(a+1) p_n Y} \left[a_{n,\,2} \, e^{a p_n \theta} + a_{n,\,3} \, e^{(a p_n + a + 1)\theta} \, e^{(a+1) Y} \right] \right|^3.
\end{aligned}$$

Let us remark that Y has survival function $\overline{G}$ defined by $\overline{G}(y) = (-y)^{\alpha}\, L\left((-y)^{-1}\right)$, for all $y \in (-\infty,\, 0)$. Setting

$$\begin{aligned} H_{n,\,0}(z) &= 1, \\ H_{n,\,1}(z) &= -\alpha z, \\ H_{n,\,2}(z) &= -e^{ap_n\theta}\frac{\mu_{p_n}}{\mu_{(a+1)p_n}}\,\frac{1-z^{a+1}}{1-z}, \\ H_{n,\,3}(z) &= \alpha\, e^{ap_n\theta}\frac{\mu_{p_n}}{\mu_{(a+1)p_n}}, \end{aligned}$$

some more easy computations show that there exist two sequences of Borel functions $(\chi_{n,\,1})$ and $(\chi_{n,\,2})$ uniformly converging to 0 on $[0,\, 1]$ such that for all $z \in [0,\, 1]$,

$$\begin{aligned} a_{n,\,0} + a_{n,\,1}\, e^{\theta}\, z &= H_{n,\,0}(z)(1-z) + \frac{H_{n,\,1}(z) + \chi_{n,\,1}(z)}{p_n}, \\ a_{n,\,2}\, e^{ap_n\theta} + a_{n,\,3}\, e^{(ap_n+a+1)\theta}\, z^{a+1} &= H_{n,\,2}(z)(1-z) + \frac{H_{n,\,3}(z) + \chi_{n,\,2}(z)}{p_n}. \end{aligned}$$

Applying Lemma 4 twice entails $\mathbb{E}\,|S_{n,\,1}|^3 = \mathrm{O}\left(n^{-3}\, e^{3p_n\theta}\, p_n^{-\alpha-3}\, L(p_n)\right)$. Lyapounov's theorem then gives (6a). Proofs of (6b) and (6c) are then similar.

Let us now focus on the nonrandom term $v_n(\Theta_n - \theta)$. Recalling (4) and letting

$$\tau(p,\, u) := \frac{I_1[\varepsilon_1(p) - \varepsilon_1(p+u)] + I_2[\varepsilon_2(p) - \varepsilon_2(p+u)]}{\Gamma(\alpha+1) + I_1\,\varepsilon_1(p+u) + I_2\,\varepsilon_2(p+u)},$$

one has

$$\forall\, u \geq 1, \quad \frac{\mu_p}{\mu_{p+u}} = e^{-u\theta}\left[1 + \frac{u}{p}\right]^{\alpha} \exp\left[-\int_p^{p+u} \frac{\eta(t)}{t}\, dt\right][1 + \tau(p,\, u)].$$

Let us note that $\displaystyle\int_p^{p+u} \frac{\eta(t)}{t}\, dt = \mathrm{O}\left(\frac{|\eta(p)|}{p}\right)$ and apply Lemma 2 to get

$$\forall\, u \geq 1, \quad \frac{\mu_p}{\mu_{p+u}} = e^{-u\theta}\left[1 + \frac{u}{p}\right]^{\alpha} + \mathrm{O}\left(\frac{|\eta(p)|}{p}\right).$$

It is then clear that

$$\Theta_n = \theta + \mathrm{O}\left(\frac{|\eta(p_n)|}{p_n}\right).$$

The result follows from Slutsky's lemma. ■

As far as the rate of convergence v_n of the estimator is concerned, note that up to a slowly varying factor, one has $v_n = \sqrt{n}\, p_n^{-\alpha/2+1}$, where (p_n) satisfies $n\, p_n^{-\alpha} \to \infty$ and $n\, p_n^{-\alpha}\, \eta^2(p_n) \to 0$. We shall consider the cases $\alpha \geq 2$ and $\alpha < 2$ separately:

1. If $\alpha \geq 2$, then the smaller p_n is, the higher v_n is. The constraint on (p_n) is therefore the condition $n\, p_n^{-\alpha}\, \eta^2(p_n) \to 0$. Since $|\eta|$ is regularly varying with index ν, this condition is essentially $n\, p_n^{2\nu-\alpha} \to 0$: the smallest possible sequence (p_n) satisfying this requirement has order $n^{1/(\alpha-2\nu)}$. Consequently, (v_n) has order $n^{(1-\nu)/(\alpha-2\nu)}$.

2. If now $\alpha < 2$, then the rate (v_n) increases as (p_n) increases: the constraint on (p_n) is the condition $n\, p_n^{-\alpha} \to \infty$. The largest possible sequence (p_n) satisfying this condition has order $n^{1/\alpha}$, which yields a rate (v_n) with order $n^{1/\alpha}$.

Hence, the estimator of Aarssen and de Haan (1994) and our estimator essentially have the same rate of convergence. Moreover, in the case $\alpha < 2$, the rate is the same as the one of the maximum estimator (see de Haan and Ferreira, 2006). The three above mentioned estimators are compared on finite sample situations in the next section.

3 Numerical illustration

Here, we examine the performances of our estimator by considering two different models. The first one has survival function

$$\forall\, x < 0, \quad \overline{F}(x) = \left[1 + (-x)^{-\tau_1}\right]^{-\tau_2} \tag{7}$$

with $\tau_1, \tau_2 > 0$, that is, $X = -1/Z$ where Z has a Burr$(1, \tau_1, \tau_2)$ type XII distribution as in Beirlant *et al.* (2004). Here (A_1) and (A_2) hold, with $\theta = 0$, $\alpha = \tau_1\, \tau_2$ and $\nu = -\tau_1$.

The second one has survival function

$$\forall\, x < 0, \quad \overline{F}(x) = \int_{\ln(1-1/x)}^{\infty} \lambda^2\, t\, e^{-\lambda t}\, dt \tag{8}$$

with $\lambda > 0$, which is tantamount to $X = -1/(e^Z - 1)$ where Z is Gamma$(2, \lambda)$ distributed. Some cumbersome computations show that (A_1) and (A_2) hold with $\theta = 0$, $\alpha = \lambda$ and $\nu = 0$.

Each of these models is considered with different sets of parameters, see the first column of Table 1. We choose $p_n = n^{1/\alpha}/\ln\ln n$, and a set $\mathcal{A} = \{0.1, 0.2, 0.3, \ldots, 25\}$ of different values of a is tested. In each situation, $N = 1000$ replications of a sample with size $n = 500$ are generated and the average L^1-error

$$E(a) = \frac{1}{N} \sum_{j=1}^{N} |\varepsilon(j,\, a)|\,, \quad \text{where} \quad \varepsilon(j,\, a) = \widehat{\theta}^{(j,\, a)} - \theta$$

is computed, with $\widehat{\theta}^{(j,\, a)}$ being the estimator computed on the jth replication with $a \in \mathcal{A}$ and an endpoint $\theta = 0$. Then, the "optimal" value of a is retained: $a^\star = \text{argmin}\{E(a),\, a \in \mathcal{A}\}$. The same procedure is applied to the extreme-value moment estimator of Aarssen and de Haan (1994), which depends on a parameter $k \in \{2, 3, \ldots, n-1\}$. The (naive) maximum estimator is also considered. Numerical results are summarized in Table 1, where $E(a^\star)$ is displayed. Let us notice that, in all the considered situations, our estimator yields slightly better (optimal) results than the maximum estimator and the extreme-value moment estimator.

To further compare the behavior of the estimators in the "optimal" case, boxplots of the associated errors $\varepsilon(j,\, a^*)$ are displayed on Figure 1–2. Clearly, the maximum as well as our estimator underestimate the endpoint. However, the error associated to our estimator is smaller than the error of

the maximum. Besides, the variance of our estimator is similar to the one of the maximum, and it is smaller than the one of the extreme-value moment estimator.

On Figure 3, we compare the functions E associated to the three estimators. On model (7), while the error associated to the extreme-value moment estimator appears to be very sensitive to the choice of k, the error associated to our estimator is stable for a large panel of a values. Results are similar in the other cases considered.

References

Aarssen, K., de Haan, L., 1994. On the maximal life span of humans. Math. Popul. Stud. 4(4), 259–281.

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., 2004. Statistics of Extremes, John Wiley and Sons.

Billingsley, P., 1979. Probability and measure, John Wiley and Sons.

Bingham, N.H., Goldie, C.M., Teugels, J.L., 1987. Regular Variation, Cambridge, U.K.: Cambridge University Press.

Fisher, R.A., Tippett, L.H.C., 1928. On estimating of the frequency distributions of the largest or smallest member of a sample. Proc. Camb. Phil. Soc. 24, 180–190.

Girard, S., Guillou, A., Stupfler, G., 2011. Estimating an endpoint with high order moments. Test, to appear. DOI: 10.1007/s11749-011-0277-8.

Gnedenko, B.V., 1943. Sur la distribution limite du maximum d'une série aléatoire. Ann. of Math. 44, 423–453.

de Haan, L., Ferreira, A., 2006. Extreme Value Theory, Springer.

Hall, P., 1982. On estimating the endpoint of a distribution. Ann. Statist. 10(2), 556–568.

Hall, P., Wang, J.Z., 2005. Bayesian likelihood methods for estimating the end point of a distribution. J. Roy. Statist. Soc. Ser. B 67(5), 717–729.

Hoeffding, W., 1963. Probability inequalities for sums of bounded random variables. J. Amer. Statist. Assoc. 58, 13–30.

Li, D., Peng, L., Xu, X., 2011a. Bias reduction for endpoint estimation. Extremes 14, 393–412.

Li, D., Peng, L, Qi, Y., 2011b. Empirical likelihood confidence intervals for the endpoint of a distribution function. Test 20, 353–366.

4 Appendix: Auxiliary results

The first two results are analogues of Lemmas 1 and 2 in Girard *et al.* (2011), as well as their proofs, which are omitted.

Lemma 1. *If* (A_0) *holds, then for all* $u \geq 1$*, one has* $\mu_p/\mu_{p+u} \to e^{-u\theta}$ *as* $p \to \infty$*.*

Lemma 2. *Assume that* (A_1) *holds and define* ε_1 *and* ε_2 *as in the proof of Theorem 3. Then, for all* $i = 1, 2$*,*

(i) $\varepsilon_i(p) \to 0$ *as* $p \to \infty$*.*

Moreover, if L *satisfies* (A_2)*, then for all* $i = 1, 2$ *and* $u, v \geq 1$*,*

(ii) $\varepsilon_i(p+u) - \varepsilon_i(p) = \mathrm{O}(|\eta(p)|/p)$*,*

(iii) $p^2(\varepsilon_i(p+u+v) - \varepsilon_i(p+v) - [\varepsilon_i(p+u) - \varepsilon_i(p)]) \to 0$ *as* $p \to \infty$*.*

The next lemma is a technical result which shall be useful in the proof of Lemma 4 below. It is a simple consequence of Lemma 2:

Lemma 3. *Assume that* (A_1) *holds. Then, as* $p \to \infty$*,*

$$\forall\, d \geq 0, \quad p \int_{-\infty}^{\theta} e^{px}\,(\theta - x)^d\,\overline{F}(x)\,dx = p^{-\alpha-d}\,e^{p\theta}\,L(p)\,\Gamma(\alpha+d+1)\,(1+\mathrm{o}(1)).$$

Proof of Lemma 3. Rewrite the left-hand side as in (4) and apply Lemma 2i). ∎

The final lemma of this section provides an asymptotic bound of the third-order moments appearing in the proof of Theorem 3.

Lemma 4. *Let* $m \in \mathbb{N}$*,* $(H_{n,\,j})$*,* $0 \leq j \leq m$ *be sequences of Borel uniformly bounded functions on* $(0,\,1)$ *and* (p_n) *be a real sequence tending to infinity. Introduce*

$$\forall\, z \in (0,\,1), \quad h_n(z) = \sum_{j=0}^{m} \frac{H_{n,\,j}(z)}{p_n^j}(1-z)^{m-j},$$

and let Y *be a random variable with survival function* $\overline{G}$ *defined by*

$$\forall\, y \in (-\infty,\,0), \quad \overline{G}(y) = (-y)^{\alpha}\,L\left((-y)^{-1}\right)$$

where $\alpha > 0$ *and* L *is a slowly varying function at infinity. Then*

$$\mathbb{E}\left|e^{p_n Y}\,h_n\left(e^{Y}\right)\right|^3 = \mathrm{O}\left(p_n^{-\alpha-3m}\,L(p_n)\right).$$

Proof of Lemma 4. Hölder's inequality yields

$$\mathbb{E}\left|e^{p_n Y}\,h_n\left(e^{Y}\right)\right|^3 \leq (m+1)^2 \left[\sum_{j=0}^{m} \frac{1}{p_n^{3j}} \sup_{\substack{[0,\,1]\\ n\in\mathbb{N}\setminus\{0\}}} |H_{n,\,j}|^3\,\mathbb{E}\left[e^{p_n Y}\,(1-e^{Y})^{m-j}\right]^3\right].$$

It is enough to show that $\forall\, j \in \{0, \ldots, m\},\ \mathbb{E}\left[e^{p_n Y}\,(1-e^Y)^{m-j}\right]^3 = \mathrm{O}\left(p_n^{-\alpha-(3m-3j)}\,L(p_n)\right)$. An integration by parts gives

$$\mathbb{E}\left[e^{p_n Y}\,(1-e^Y)^{m-j}\right]^3 = \int_{-\infty}^{0} \frac{d}{dy}\left[e^{3p_n y}\,(1-e^y)^{3m-3j}\right]\,\overline{G}(y)\,dy.$$

If (s_n) is a real sequence tending to infinity, φ is a positive bounded function on $(-\infty,\,0)$ and $\beta \geq 0$, by writing

$$\forall\,\delta > 0, \quad \int_{-\infty}^{0} e^{s_n y}\,(-y)^{\beta}\,\varphi(y)\,dy = \int_{-\delta}^{0} e^{s_n y}\,(-y)^{\beta}\,\varphi(y)\,dy\left[1 + \frac{\displaystyle\int_{-\infty}^{-\delta} e^{s_n y}\,(-y)^{\beta}\,\varphi(y)\,dy}{\displaystyle\int_{-\delta}^{0} e^{s_n y}\,(-y)^{\beta}\,\varphi(y)\,dy}\right],$$

it is readily shown that

$$\forall\,\delta > 0, \quad \int_{-\infty}^{0} e^{s_n y}\,(-y)^{\beta}\,\varphi(y)\,dy = \int_{-\delta}^{0} e^{s_n y}\,(-y)^{\beta}\,\varphi(y)\,dy\,(1+\mathrm{o}(1)). \tag{9}$$

Since $y/(1-e^y) \to -1$ as $y \to 0$, we get, for all $\varepsilon > 0$, choosing $\delta = \delta_0$ small enough,

$$1 - \frac{\varepsilon}{2} \leq \frac{\displaystyle\int_{-\delta_0}^{0} e^{s_n y}\,(1-e^y)^d\,\overline{G}(y)\,dy}{\displaystyle\int_{-\delta_0}^{0} e^{s_n y}\,(-y)^d\,\overline{G}(y)\,dy} \leq 1 + \frac{\varepsilon}{2}.$$

As a consequence, (9) yields, for all sufficiently large n,

$$1 - \varepsilon \leq \frac{\displaystyle\int_{-\delta_0}^{0} e^{s_n y}\,(1-e^y)^d\,\overline{G}(y)\,dy}{\displaystyle\int_{-\infty}^{0} e^{s_n y}\,(-y)^d\,\overline{G}(y)\,dy} \leq 1 + \varepsilon. \tag{10}$$

It only remains to use (9) once again and to apply Lemma 3 to obtain

$$\int_{-\infty}^{0} e^{s_n y}\,(1-e^y)^d\,\overline{G}(y)\,dy = s_n^{-\alpha-d-1}\,L(s_n)\,\Gamma(\alpha+d+1)\,(1+\mathrm{o}(1)).$$

Replacing in (10), it follows that $\mathbb{E}\left[e^{p_n Y}\,(1-e^Y)^{m-j}\right]^3 = \mathrm{O}\left(p_n^{-\alpha-(3m-3j)}\,L(p_n)\right)$, which establishes Lemma 4. ■

Distribution	Maximum	Estimator of Aarssen & de Haan	High order moments estimator
$-1/\text{Burr}(1, \tau_1, \tau_2)$			
$(\tau_1, \tau_2) = (1, 1)$ $\Rightarrow (\alpha, \nu) = (1, -1)$	$2.0 \cdot 10^{-3}$	$2.1 \cdot 10^{-3}$	$1.7 \cdot 10^{-3}$
$(\tau_1, \tau_2) = (5/6, 6/5)$ $\Rightarrow (\alpha, \nu) = (1, -5/6)$	$2.0 \cdot 10^{-3}$	$2.0 \cdot 10^{-3}$	$1.7 \cdot 10^{-3}$
$(\tau_1, \tau_2) = (2/3, 3/2)$ $\Rightarrow (\alpha, \nu) = (1, -2/3)$	$2.2 \cdot 10^{-3}$	$2.0 \cdot 10^{-3}$	$1.8 \cdot 10^{-3}$
$(\tau_1, \tau_2) = (1/2, 2)$ $\Rightarrow (\alpha, \nu) = (1, -1/2)$	$2.3 \cdot 10^{-3}$	$2.3 \cdot 10^{-3}$	$2.0 \cdot 10^{-3}$
$-1/(\exp(\text{Gamma}(2, \lambda)) - 1)$			
$\lambda = 1$ $\Rightarrow (\alpha, \nu) = (1, 0)$	$2.3 \cdot 10^{-4}$	$2.0 \cdot 10^{-4}$	$1.9 \cdot 10^{-4}$
$\lambda = 5/4$ $\Rightarrow (\alpha, \nu) = (5/4, 0)$	$1.1 \cdot 10^{-3}$	$9.2 \cdot 10^{-4}$	$8.5 \cdot 10^{-4}$
$\lambda = 5/3$ $\Rightarrow (\alpha, \nu) = (5/3, 0)$	$5.7 \cdot 10^{-3}$	$4.6 \cdot 10^{-3}$	$4.1 \cdot 10^{-3}$
$\lambda = 5/2$ $\Rightarrow (\alpha, \nu) = (5/2, 0)$	$3.1 \cdot 10^{-2}$	$2.4 \cdot 10^{-2}$	$2.3 \cdot 10^{-2}$

Table 1: Mean L^1-errors associated to the estimators in the eight situations.

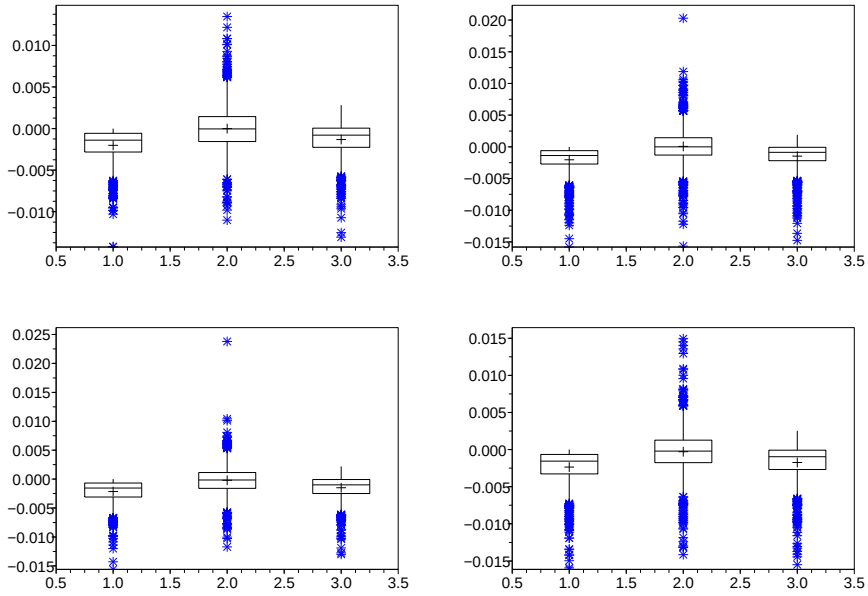

Figure 1: Boxplots of $\varepsilon(j, a^*)$ on model (7). Left: maximum estimator, middle: extreme-value moment estimator, right: high order moments estimator. Top left: $(\tau_1, \tau_2) = (1, 1)$; top right: $(\tau_1, \tau_2) = (5/6, 6/5)$; bottom left: $(\tau_1, \tau_2) = (2/3, 3/2)$; bottom right: $(\tau_1, \tau_2) = (1/2, 2)$.

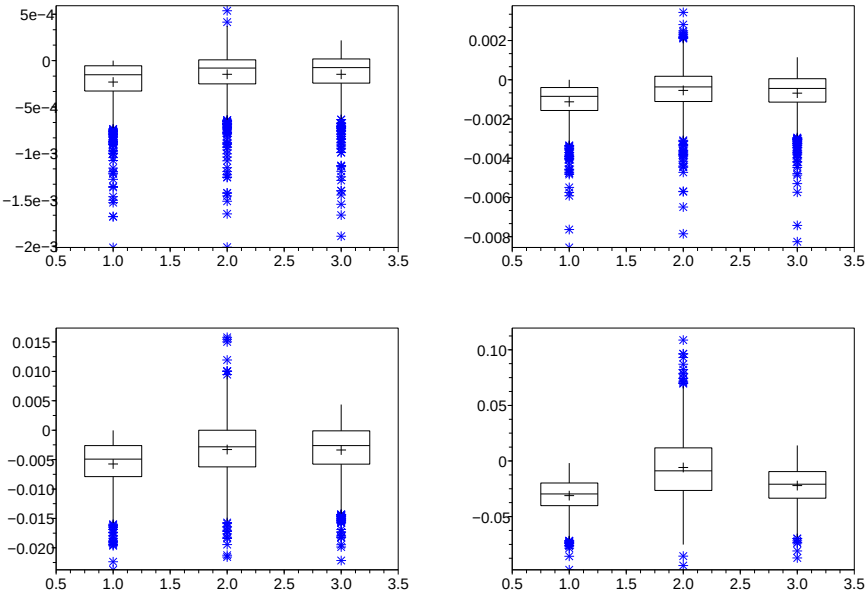

Figure 2: Boxplots of $\varepsilon(j, a^*)$ on model (8). Left: maximum estimator, middle: extreme-value moment estimator, right: high order moments estimator. Top left: $\lambda = 1$; top right: $\lambda = 5/4$; bottom left: $\lambda = 5/3$; bottom right: $\lambda = 5/2$.

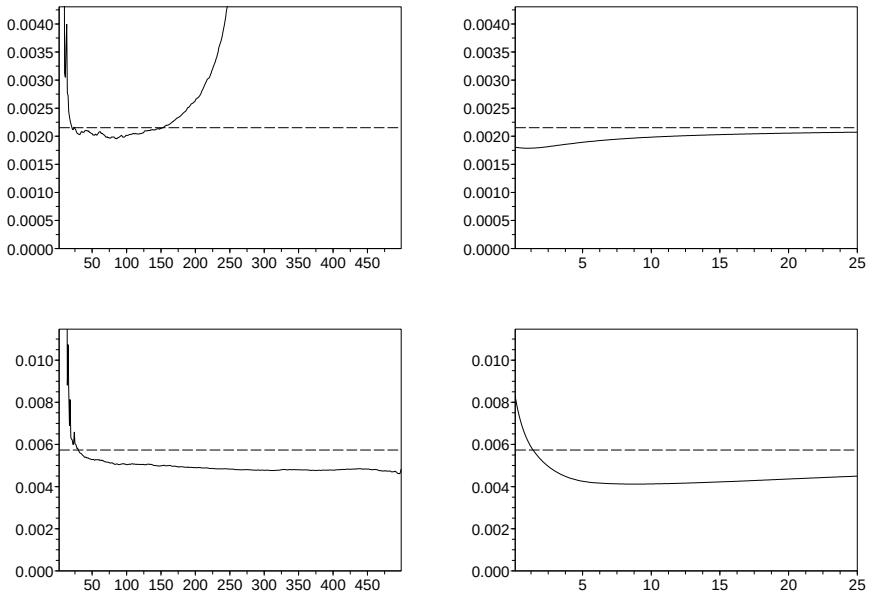

Figure 3: Comparison of the three estimators on models (7) (top) and (8) (bottom). Left: horizontally: threshold k, vertically: error E, dashed line: maximum estimator, solid line: extreme-value moment estimator. Right: horizontally: parameter a, vertically: error E, dashed line: maximum estimator, solid line: high order moments estimator. Top: model (7), $(\tau_1, \tau_2) = (2/3, 3/2)$. Bottom: model (8), $\lambda = 5/3$.