The current issue and full text archive of this journal is available on Emerald Insight at:
https://www.emerald.com/insight/1319-5166.htm

Semiparametric tail-index estimation for randomly right-truncated heavy-tailed data

Saida Mancer
Universite Mohamed Khider de Biskra, Biskra, Algeria
Abdelhakim Necir
Department of Mathematics, University of Biskra, Biskra, Algeria, and
Souad Benchaira
Universite Mohamed Khider de Biskra, Biskra, Algeria

Received 8 February 2022
Revised 20 May 2022
Accepted 8 June 2022

Abstract

Purpose – The purpose of this paper is to propose a semiparametric estimator for the tail index of Pareto-type random truncated data that improves the existing ones in terms of mean square error. Moreover, we establish its consistency and asymptotic normality.

Design/methodology/approach – To construct a root mean squared error (RMSE)-reduced estimator of the tail index, the authors used the semiparametric estimator of the underlying distribution function given by Wang (1989). This allows us to define the corresponding tail process and provide a weak approximation to this one. By means of a functional representation of the given estimator of the tail index and by using this weak approximation, the authors establish the asymptotic normality of the aforementioned RMSE-reduced estimator.

Findings – In basis on a semiparametric estimator of the underlying distribution function, the authors proposed a new estimation method to the tail index of Pareto-type distributions for randomly right-truncated data. Compared with the existing ones, this estimator behaves well both in terms of bias and RMSE. A useful weak approximation of the corresponding tail empirical process allowed us to establish both the consistency and asymptotic normality of the proposed estimator.

Originality/value – A new tail semiparametric (empirical) process for truncated data is introduced, a new estimator for the tail index of Pareto-type truncated data is introduced and asymptotic normality of the proposed estimator is established.

Keywords Extreme value index, Product-limit estimator, Semiparametric, Tail-empirical process, Truncated data

Paper type Research paper

1. Introduction

Let $(\mathbf{X}_i, \mathbf{Y}_i), i = 1, \ldots, N \geq 1$ be a sample from a couple $(\mathbf{X}, \mathbf{Y})$ of independent positive random variables (rv's) defined over a probability space $(\Omega, \mathcal{A}, \mathbf{P})$, with continuous distribution functions (df's) $\mathbf{F}$ and $\mathbf{G}$, respectively. Suppose that $\mathbf{X}$ is right-truncated by $\mathbf{Y}$, in the sense that $\mathbf{X}_i$ is only observed when $\mathbf{X}_i \leq \mathbf{Y}_i$. Thus, let us denote (X_i, Y_i), $i = 1, \ldots, n$ to be the observed data, as copies of a couple of dependent rv's (X, Y) corresponding to the truncated sample $(\mathbf{X}_i, \mathbf{Y}_i), i = 1, \ldots, N$, where $n = n_N$ is a random sequence of discrete rv's. By the weak law of large numbers, we have

The authors are indebted to the reviewers for their pertinent remarks and valuable suggestions that led to a real improvement of the paper.

Arab Journal of Mathematical Sciences
Emerald Publishing Limited
e-ISSN: 2588-9214
p-ISSN: 1319-5166
DOI 10.1108/AJMS-02-2022-0033

$$n/N \xrightarrow{\mathbf{P}} p := \mathbf{P}(\mathbf{X} \leq \mathbf{Y}) = \int_0^\infty \mathbf{F}(w)d\mathbf{G}(w), \text{ as } N \to \infty, \tag{1.1}$$

where the notation $\xrightarrow{\mathbf{P}}$ stands for the convergence in probability. The constant p corresponds to the probability of observed sample which is supposed to be non-null, otherwise nothing is observed. The truncation phenomena frequently occurs in medical studies, when one wants to study the length of survival after the start of the disease: if $\mathbf{Y}$ denotes the elapsed time between the onset of the disease and death, and if the follow-up period starts $\mathbf{X}$ units of time after the onset of the disease then, clearly, $\mathbf{X}$ is right-truncated by $\mathbf{Y}$. For concrete examples of truncated data in medical treatments one refers, among others, to Refs. [1, 2]. Truncated data schemes may also occur in many other fields, namely actuarial sciences, astronomy, demography and epidemiology, see for instance the textbook of [3].

From [4] the marginal df's F^* and G^* corresponding to the joint df of (X, Y) are given by

$$F^*(x) := p^{-1} \int_0^x \overline{\mathbf{G}}(w)d\mathbf{F}(w) \text{ and } G^*(x) := p^{-1} \int_0^x \mathbf{F}(w)d\mathbf{G}(w).$$

By the previous first equation, we derive a representation of the underlying df $\mathbf{F}$ as follows:

$$\mathbf{F}(x) = p \int_0^x \frac{dF^*(w)}{\overline{\mathbf{G}}(w)}, \tag{1.2}$$

which will be for a great interest thereafter. In the sequel, we are dealing with the concept of regular variation. A function φ is said to be regularly varying at infinity with negative index $-1/\eta$, notation $\varphi \in \mathcal{RV}(-1/\eta)$, if

$$\varphi(st)/\varphi(t) \to s^{-1/\eta}, \text{ as } t \to \infty, \tag{1.3}$$

for $s > 0$. This relation is known as the first-order condition of regular variation and the corresponding uniform convergence is formulated in terms of "Potter's inequalities" as follows: for any small $\epsilon > 0$, there exists $t_0 > 0$ such that for any $t \geq t_0$ and $s \geq 1$, we have

$$(1-\epsilon)s^{-1/\eta-\epsilon} < \varphi(st)/\varphi(t) < (1+\epsilon)s^{-1/\eta+\epsilon}. \tag{1.4}$$

See for instance Proposition B.1.9 (assertion 5, page 367) in Ref. [5]. The second-order condition (see Ref. [6] expresses the rate of the convergence (1.3) above. For any $x > 0$, we have

$$\frac{\varphi(tx)/\varphi(t) - x^{-1/\eta}}{A(t)} \to x^{-1/\eta} \frac{x^{\tau/\eta}-1}{\tau\eta}, \text{ as } t \to \infty, \tag{1.5}$$

where $\tau < 0$ denotes the second-order parameter and A is a function tending to zero and not changing signs near infinity with regularly varying absolute value with positive index τ/η. A function φ that satisfies assumption (1.5) is denoted $\varphi \in \mathcal{RV}_2(-1/\eta; \tau, A)$. We now have enough material to tackle the main goal of the paper. To begin, let us assume that the tails of both df's $\mathbf{F}$ and $\mathbf{G}$ are regularly varying. That is

$$\overline{\mathbf{F}} \in \mathcal{RV}(-1/\gamma_1) \text{ and } \overline{\mathbf{G}} \in \mathcal{RV}(-1/\gamma_2), \text{ with } \gamma_1, \gamma_2 > 0. \tag{1.6}$$

Under this assumption, [4] showed that

$$\overline{F}^* \in \mathcal{RV}(-1/\gamma_1) \text{ and } \overline{G}^* \in \mathcal{RV}(-1/\gamma), \tag{1.7}$$

where

$$\gamma := \frac{\gamma_1 \gamma_2}{\gamma_1 + \gamma_2}. \tag{1.8}$$

For further details on the proof of this statement one refers to Ref. [7] (Lemma A1). The estimation of the tail index γ_1 was recently addressed for the first time in Ref. [4] where the authors used equation (1.8) to propose an estimator to γ_1 as a ratio of Hill estimators [8] of the tail indices γ and γ_2. These estimators are based on the top order statistics $X_{n-k:n} \leq \ldots \leq X_{n:n}$ and $Y_{n-k:n} \leq \ldots \leq Y_{n:n}$ pertaining to the samples $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ respectively. The sample fraction $k = k_n$ being a sequence of integers such that, $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$. The asymptotic normality of the given estimator is established in Ref. [9]. By using a Lynden-Bell integral, [10] proposed the following estimator for the tail index γ_1:

$$\widehat{\gamma}_1^{(\mathbf{W})}(u) := \frac{1}{\overline{\mathbf{F}}_n^{(1)}(u)} \sum_{i=1}^{n} \mathbf{1}(X_i > u) \frac{\mathbf{F}_n^{(1)}(X_i)}{C_n(X_i)} \log \frac{X_i}{u},$$

for a given deterministic threshold $u > 0$, where

$$\mathbf{F}_n^{(1)}(x) := \prod_{X_i > x} \left[1 - \frac{1}{nC_n(X_i)}\right],$$

is the popular nonparametric maximum likelihood estimator of cdf $\mathbf{F}$ introduced in the well-known work [11]; with

$$C_n(x) := \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(X_i \leq x \leq Y_i).$$

Independently, [7] used a Woodroofe integral with a random threshold, to derive the following estimator

$$\widehat{\gamma}_1^{(\mathbf{BMN})} := \frac{1}{\overline{\mathbf{F}}_n^{(2)}(X_{n-k:n})} \sum_{i=1}^{k} \frac{\mathbf{F}_n^{(2)}(X_{n-i+1:n})}{C_n(X_{n-i+1:n})} \log \frac{X_{n-i+1:n}}{X_{n-k:n}}, \tag{1.9}$$

where

$$\mathbf{F}_n^{(2)}(x) := \prod_{X_i > x} \exp\left\{-\frac{1}{nC_n(X_i)}\right\},$$

is the so-called Woodroofe's nonparametric estimator [12] of df $\mathbf{F}$. To improve the performance of $\widehat{\gamma}_1^{(\mathbf{BMN})}$, [13, 14], respectively, proposed a Kernel-smoothed and a reduced-bias versions of this estimator and established their consistency and asymptotic normality. It is worth mentioning that Lynden-Bell integral estimator $\widehat{\gamma}_1^{(\mathbf{W})}(u)$ with a random threshold $u = X_{n-k:n}$ becomes

$$\widehat{\gamma}_1^{(\mathbf{W})} := \frac{1}{\overline{\mathbf{F}}_n^{(1)}(X_{n-k:n})} \sum_{i=1}^{k} \frac{\mathbf{F}_n^{(1)}(X_{n-i+1:n})}{C_n(X_{n-i+1:n})} \log \frac{X_{n-i+1:n}}{X_{n-k:n}}. \tag{1.10}$$

In a simulation study, [15] compared this estimator with $\widehat{\gamma}_1^{(\mathbf{BMN})}$. They pointed out that both estimators have similar behaviors in terms of biases and mean squared errors.

Recall that the nonparametric Lynden-Bell estimator $\mathbf{F}_n^{(1)}$ was constructed on the basis of the fact that $\mathbf{F}$ and $\mathbf{G}$ are both unknown. In this paper, we are dealing with the situation when

$\mathbf{F}$ is unknown but $\mathbf{G}$ is parametrized by a known model $\mathbf{G}_\theta$, $\theta \in \Theta \subset \mathbb{R}^d$, $d \geq 1$ having a density $\mathbf{g}_\theta$ with respect to Lebesgue measure. [2] considered this assumption and introduced a semiparametric estimator for df $\mathbf{F}$ defined by

$$\mathbf{F}_n\left(x; \widehat{\theta}_n\right) := P_n\left(\widehat{\theta}_n\right) \frac{1}{n} \sum_{i=1}^{n} \frac{\mathbf{1}(X_i \leq x)}{\overline{\mathbf{G}}_{\widehat{\theta}_n}(X_i)}, \tag{1.11}$$

where $1/P_n\left(\widehat{\theta}_n\right) := n^{-1}\sum_{i=1}^{n} 1/\overline{\mathbf{G}}_{\widehat{\theta}_n}(X_i)$ and

$$\widehat{\theta}_n := \arg \max_{\theta\in\Theta} \prod_{i=1}^{n} g_\theta(Y_i)/\overline{\mathbf{G}}_\theta(X_i), \tag{1.12}$$

denoting the conditional maximum likelihood estimator (CMLE) of θ, which is consistent and asymptotically normal, see for instance Ref. [16]. On the other hand, [2] showed that $\mathbf{F}_n\left(x; \widehat{\theta}_n\right)$ is an uniformly consistent estimator over the x-axis and established, under suitable regularity assumptions, its asymptotic normality. [2, 17] pointed out that the semiparametric estimate has greater efficiency uniformly over the x-axis. In the light of a simulation study, the authors suggest that the semiparametric estimate is a better choice when parametric information of the truncation distribution is available. Since the apparition of this estimation method many papers are devoted to the statistical inference with truncation data, see for instance Refs. [18–22] and [23].

Motivated by the features of the semiparametric estimation, we next propose a new estimator for γ_1 by means of a suitable functional of $\mathbf{F}_n\left(x; \widehat{\theta}_n\right)$. We start our construction by noting that from Theorem 1.2.2 in de [5]; the first-order condition (1.6) (for $\mathbf{F}$) implies that

$$\lim_{t\to\infty} \frac{1}{\overline{\mathbf{F}}(t)} \int_t^\infty \log(x/t) d\mathbf{F}(x) = \gamma_1. \tag{1.13}$$

In other words, γ_1 may viewed as a functional $\psi_t(\mathbf{F})$, for a large t, where

$$\psi_t(\mathbf{F}) := \frac{1}{\overline{\mathbf{F}}(t)} \int_t^\infty \log(x/t) d\mathbf{F}(x).$$

Replacing $\mathbf{F}$ by $\mathbf{F}_n\left(\cdot; \widehat{\theta}_n\right)$ and letting $t = X_{n-k:n}$ yield

$$\begin{aligned} \widehat{\gamma}_1 &= \psi_{X_{n-k:n}}\left(\mathbf{F}_n\left(\cdot; \widehat{\theta}_n\right)\right) \\ &= \frac{1}{\overline{\mathbf{F}}_n\left(X_{n-k:n}; \widehat{\theta}_n\right)} \int_{X_{n-k:n}}^\infty \log(x/X_{n-k:n}) d\mathbf{F}_n\left(x; \widehat{\theta}_n\right), \end{aligned} \tag{1.14}$$

as new estimator for γ_1. Observe that

$$\begin{aligned} &\int_t^\infty \log(x/t) d\mathbf{F}_n\left(x; \widehat{\theta}_n\right) \\ &= P_n\left(\widehat{\theta}\right) \int_{X_{n-k:n}}^\infty \log(x/X_{n-k:n}) \mathbf{1}(x \geq X_{n-k}) d\mathbf{F}_n\left(x; \widehat{\theta}_n\right), \end{aligned}$$

which may be rewritten into

$$\frac{P_n\left(\widehat{\theta}_n\right)1}{n}\sum_{i=1}^{n}\int_{X_{n-k:n}}^{\infty}\frac{\log(x/X_{n-k:n})\mathbf{1}(x\geq X_{n-k})}{\overline{\mathbf{G}}_{\widehat{\theta}_n}(X_i)}d\mathbf{1}(X_i\leq x)$$

$$=P_n\left(\widehat{\theta}_n\right)\frac{1}{n}\sum_{i=1}^{k}\frac{\log(X_{n-i+1}/X_{n-k:n})}{\overline{\mathbf{G}}_{\widehat{\theta}_n}(X_{n-i+1:n})}.$$

On the other hand, $\mathbf{F}\left(X_{n-k:n};\widehat{\theta}_n\right)$ equals

$$P_n\left(\widehat{\theta}_n\right)\frac{1}{n}\sum_{i=1}^{n}\frac{\mathbf{1}(X_{i:n}\leq X_{n-k:n})}{\overline{\mathbf{G}}_{\widehat{\theta}_n}(X_{i:n})}=P_n\left(\widehat{\theta}_n\right)\frac{1}{n}\sum_{i=1}^{n-k}1/\overline{\mathbf{G}}_{\widehat{\theta}_n}(X_{i:n}).$$

Hence,

$$\overline{\mathbf{F}}\left(X_{n-k:n};\widehat{\theta}_n\right)=\frac{\frac{1}{n}\sum_{i=1}^{n}1/\overline{\mathbf{G}}_{\widehat{\theta}_n}(X_{i:n})-\frac{1}{n}\sum_{i=1}^{n-k}1/\overline{\mathbf{G}}_{\widehat{\theta}_n}(X_{i:n})}{1n\sum_{i=1}^{n}1/\overline{\mathbf{G}}_{\widehat{\theta}_n}(X_{i:n})}$$

$$=P_n\left(\widehat{\theta}_n\right)\frac{1}{n}\sum_{i=1}^{k}1/\overline{\mathbf{G}}_{\widehat{\theta}_n}(X_{n-i+1:n}).$$

Thereby, the form of our new estimator is

$$\widehat{\gamma}_1=\frac{\sum_{i=1}^{k}\left(\overline{\mathbf{G}}_{\widehat{\theta}_n}(X_{n-i+1:n})\right)^{-1}\log(X_{n-i+1}/X_{n-k:n})}{\sum_{i=1}^{k}\left(\overline{\mathbf{G}}_{\widehat{\theta}_n}(X_{n-i+1:n})\right)^{-1}}. \tag{1.15}$$

The asymptotic behavior of $\widehat{\gamma}_1$ will be established by means of the following tail empirical process

$$\mathbf{D}_n\left(x;\widehat{\theta}_n;\gamma_1\right):=\sqrt{k}\left(\frac{\overline{\mathbf{F}}_n\left(xX_{n-k:n};\widehat{\theta}_n\right)}{\overline{\mathbf{F}}_n\left(X_{n-k:n};\widehat{\theta}_n\right)}-x^{-1/\gamma_1}\right),\text{for } x>1.$$

This method was already used to establish the asymptotic behavior of Hill's estimator for complete data [5]; page 162) that we will adapt to the truncation case. Indeed, by using an integration by parts and a change of variables of the integral (1.14), one gets

$$\widehat{\gamma}_1=\int_1^{\infty}x^{-1}\frac{\overline{\mathbf{F}}_n\left(xX_{n-k:n};\widehat{\theta}_n\right)}{\overline{\mathbf{F}}_n\left(X_{n-k:n};\widehat{\theta}_n\right)}dx,$$

and therefore

$$\sqrt{k}(\widehat{\gamma}_1-\gamma_1)=\int_1^{\infty}x^{-1}\mathbf{D}_n\left(x;\widehat{\theta}_n;\gamma_1\right)dx. \tag{1.16}$$

Thus, for a suitable weighted weak approximation to $\mathbf{D}_n\left(\cdot;\widehat{\theta}_n;\gamma_1\right)$, we may easily deduce the consistency and asymptotic normality of $\widehat{\gamma}_1$. This process may also contribute to the goodness-of-fit test to fitting heavy-tailed distributions via, among others, the Kolmogorov–Smirnov and Cramér–von Mises type statistics

$$\sup_{x>1}\left|\mathbf{D}_n\left(x;\widehat{\theta}_n,\widehat{\gamma}_1\right)\right| \text{ and } \int_1^\infty \mathbf{D}_n^2\left(x;\widehat{\theta}_n,\widehat{\gamma}_1\right)dx^{-1/\widehat{\gamma}_1}.$$

More precisely, these statistics are used when testing the null hypothesis H_0: "both $\mathbf{F}$ and $\mathbf{G}$ are heavy-tailed" versus the alternative one H_1: "at least one of $\mathbf{F}$ and $\mathbf{G}$ is not heavy-tailed", that is H_0: "(1.6) holds" versus H_1: "(1.6) does not hold". This problem has been already addressed by Refs. [24, 25] in the case of complete data. The (uniform) weighted weak convergence of $\mathbf{D}_n\left(x;\widehat{\theta}_n,\gamma_1\right)$ and the asymptotic normality of $\widehat{\gamma}_1$, stated below, will be of great interest to establish the limit distributions of the aforementioned test statistics. This is out of the scope of this paper whose remainder is structured as follows. In Section 2, we present our main results which consist in the consistency and asymptotic normality of estimator $\widehat{\gamma}_1$. The performance of the proposed estimator is checked by simulation in Section 3. An application to a real dataset composed of induction times of AIDS diseases is given in Section 4. The proofs are gathered in Section 5. A useful lemma and its proof are postponed to Appendix.

2. Main results

The regularity assumptions, denoted $[A0]$, concerning the existence, consistency and asymptotic normality of the CLME estimator $\widehat{\theta}_n$, given in (1.12), are discussed in Ref. [16]. Here, we only state additional conditions on df $\mathbf{G}_\theta$ corresponding to Pareto-type models which are required to establish the asymptotic behavior of our newly estimator $\widehat{\gamma}_1$.

(1) $[A1]$ For each fixed y, the function $\theta \to \mathbf{G}_\theta(y)$ is continuously differentiable of partial derivatives $\mathbf{G}_\theta^{(j)} =: \partial\mathbf{G}_\theta/\partial\theta_j, j = 1, \ldots, d$.

(2) $[A2]$ $\overline{\mathbf{G}}_\theta^{(j)} \in \mathcal{RV}(-1/\gamma_2)$.

(3) $[A3]$ $y^{-\epsilon}\overline{\mathbf{G}}_\theta^{(j)}(y)/\overline{\mathbf{G}}_\theta(y) \to 0$, as $y \to \infty$, for any $\epsilon > 0$.

For common Pareto-type models, one may easily check that there exist some constants $a_j \geq 0$, c_j and d_j, such that $\overline{\mathbf{G}}_\theta^{(j)}(y) \sim c_j\left(y^{-1/\gamma_2} + d_j\right)\log y$, for all large x. Then one may consider that the assumptions $[A1]-[A3]$ are not very restrictive and they may be acceptable in the extreme value theory.

Theorem 2.1. *Assume that* $\overline{\mathbf{F}} \in \mathcal{RV}_2(-1/\gamma_1;\rho_1,\mathbf{A})$ *and* $\mathbf{G}_\theta \in \mathcal{RV}(-1/\gamma_2)$ *satisfying the assumptions* $[A0]-[A3]$, *and suppose that* $\gamma_1 < \gamma_2$. *Then on the probability space* $(\Omega,\mathcal{A},\mathbf{P})$, *there exists a standard Wiener process* $\{W(s), 0 \leq s \leq 1\}$ *such that, for any small* $0 < \epsilon < 1/2$, *we have*

$$\sup_{x>1} x^\epsilon \left|\mathbf{D}_n\left(x;\widehat{\theta}_n,\gamma_1\right) - \Gamma(x;W) - x^{-1/\gamma_1}\frac{x^{\rho_1/\gamma_1}-1}{\rho_1\gamma_1}\sqrt{k}\,\mathbf{A}(a_k)\right| \overset{\mathbf{P}}{\to} 0,$$

provided that $\sqrt{k}\mathbf{A}(a_k) = O(1)$, *where*

$$\Gamma(x;W) := \frac{\gamma}{\gamma_1}x^{-1/\gamma_1}\left\{x^{1/\gamma}W\left(x^{-1/\gamma}\right) - W(1)\right\}$$
$$+\frac{\gamma}{\gamma_1+\gamma_2}x^{-1/\gamma_1}\int_0^1 s^{-\gamma/\gamma_2-1}\left\{x^{1/\gamma}W\left(x^{-1/\gamma}s\right) - W(s)\right\}ds,$$

is a centered Gaussian process and $a_k := F^{*\leftarrow}(1-k/n)$, *where*

$$F^{*\leftarrow}(s) := \inf\{x : F^*(x) \geq s\},\ 0 < s < 1,$$

denotes the quantile (or the generalized inverse) function pertaining to df F*.

Applying this weak approximation, we establish both consistency and asymptotic normality of our new estimator $\widehat{\gamma}_1$, that we state in the following Theorem.

Theorem 2.2. *Under the assumptions of Theorem 2.1, we have*

$$\begin{aligned}&\widehat{\gamma}_1 - \gamma_1 \\ &= k^{-1/2}\int_1^\infty x^{-1}\Gamma(x;W)dx + \mathbf{A}(a_k)\int_1^\infty x^{-1/\gamma_1-1}\frac{x^{\rho_1/\gamma_1}-1}{\rho_1\gamma_1}dx + o_{\mathbf{P}}\left(k^{-1/2}\right),\end{aligned}$$

this implies that $\widehat{\gamma}_1 \xrightarrow{\mathbf{P}} \gamma_1$. *Whenever* $\sqrt{k}\mathbf{A}(a_k) \to \lambda < \infty$, *we get*

$$\sqrt{k}(\widehat{\gamma}_1 - \gamma_1) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\lambda}{1-\rho_1}, \sigma^2\right),$$

where $\sigma^2 := \gamma^2(1+\gamma_1/\gamma_2)\left(1+(\gamma_1/\gamma_2)^2\right)(1-\gamma_1/\gamma_2)^3$, *and* $\mathbf{1}(\mathcal{A})$ *stands for the indicator function pertaining to a set* $\mathcal{A}$.

3. Simulation study

In this section, we will perform a simulation study in order to compare the finite sample behavior of our new semiparametric estimator $\widehat{\gamma}_1$, given in (1.15), with Woodrofee and Lynden-Bell integral estimators $\widehat{\gamma}_1^{(\mathbf{BMN})}$ and $\widehat{\gamma}_1^{(\mathbf{W})}$, given respectively in (1.9) and (1.10). The truncation and truncated distributions functions **F** and **G** will be chosen among the following two models:

(1) Burr (γ, δ) distribution with right-tail function:

$$\overline{H}(x) = \left(1 + x^{1/\delta}\right)^{-\delta/\gamma},\ x \geq 0,\ \delta > 0,\ \gamma > 0;$$

(2) Fréchet (γ) distribution with right-tail function:

$$\overline{H}(x) = 1 - \exp\left(-x^{-1/\gamma}\right), x > 0, \gamma > 0.$$

The simulation study is being made in fours scenarios following to the choice of the underlying df's **F** and $\mathbf{G}_\theta$:

(3) $[S1]$ Burr (γ_1, δ) truncated by Burr (γ_2, δ); with $\theta = (\gamma_2, \delta)$

(4) $[S2]$ Fréchet (γ_1) truncated by Fréchet (γ_2); with $\theta = \gamma_2$

(5) $[S3]$ Fréchet (γ_1) truncated by Burr (γ_2, δ); with $\theta = (\gamma_2, \delta)$

(6) $[S4]$ Burr (γ_1, δ) truncated by Fréchet (γ_2); with $\theta = \gamma_2$

To this end, we fix $\delta = 1/4$ and choose the values 0.6 and 0.8 for γ_1 and 55% and 90% for the portions of observed truncated data given in (1.1) so that the assumption $\gamma_1 < \gamma_2$ stated in Theorem 2.1 holds. In other words, the values of p have to be greater than 50%. For each couple (γ_1, p), we solve the equation (1.1) to get the pertaining γ_2-value, which we summarize as follows:

$$(p, \gamma_1, \gamma_2) = (55\%, 0.6, 1.4), (90\%, 0.6, 5.4), (55\%, 0.8, 1.9), (90\%, 0.8, 7.2). \quad (3.17)$$

For each scenario, we simulate 1000 random samples of size $N = 300$ and compute the root mean squared error (RMSE) and the absolute bias (ABIAS) corresponding to each estimator $\widehat{\gamma}_1, \widehat{\gamma}_1^{(\mathbf{BMN})}$ and $\widehat{\gamma}_1^{(\mathbf{W})}$. The comparison is done by plotting the ABIAS and RMSE as functions of the sample fraction k which varies from 2 to 120. This range is chosen so that it contains the optimal number of upper extremes k^* used in the computation of the tail index estimate. There are many heuristic methods to select k^*, see for instance Ref. [26]; here we use the algorithm proposed by Ref. [27] in page 137, which is incorporated in the R software "Xtremes" package. Note that the computation the CMLE of θ is made by means of the syntax "maxLik" of the MaxLik R software package. The optimal sample fraction k^* is defined, in this procedure, by

$$k^* := \arg\min_{1<k<n} \frac{1}{k} \sum_{i=1}^{k} i^{\omega} |\widehat{\gamma}(i) - \text{median}\{\widehat{\gamma}(1), \ldots, \widehat{\gamma}(k)\}|,$$

for suitable constant $0 \le \omega \le 1/2$, where $\widehat{\gamma}(i)$ corresponds to an estimator of tail index γ, based on the i upper order statistics, of a Pareto-type model. We observed, in our simulation study, that $\omega = 0.3$ allows better results both in terms of bias and RMSE. It is worth mentioning that making N vary did not provide notable findings; therefore, we kept the size N fixed. The finite sample behaviors of the above-mentioned estimators are illustrated in Figures 1–8. The overall conclusion is that the biases of three estimators are almost equal, however, in the case of medium truncation ($p \approx 50\%$), the RMSE of our new semiparametric $\widehat{\gamma}_1$ is clearly the smallest compared that of $\widehat{\gamma}_1^{(\mathbf{BMN})}$ and $\widehat{\gamma}_1^{(\mathbf{W})}$. Actually, the medium truncation situation is the most frequently encountered in real data, while the strong truncation ($p \gg 50\%$) remains, up

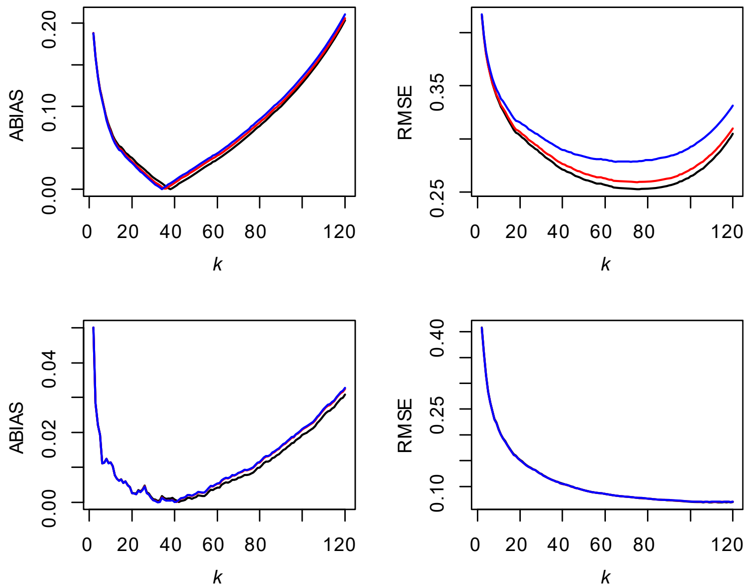

Figure 1.
Absolute bias (left two panels) and RMSE (right two panels) of $\widehat{\gamma}_1$ (black) and $\widehat{\gamma}_1^{(\mathbf{BMN})}$ (red) and $\widehat{\gamma}_1^{(\mathbf{W})}$(blue), corresponding to two situations of scenario $S_1 : (\gamma_1 = 0.6, p = 55\%)$ (top two panels) and $(\gamma_1 = 0.6, p = 90\%)$ (bottom two panels) based on 1000 samples of size 300

to our knowledge, theoretical. In this sense, we may consider that the semiparametric estimator is more efficient than the two other ones. We point out that the two estimators $\widehat{\gamma}_1^{(\mathbf{BMN})}$ and $\widehat{\gamma}_1^{(\mathbf{W})}$ have almost the same behavior which actually was noticed before by Ref. [15]. The optimal sample fractions and estimate values of the tail index obtained through the three estimators are given in Tables 1–4.

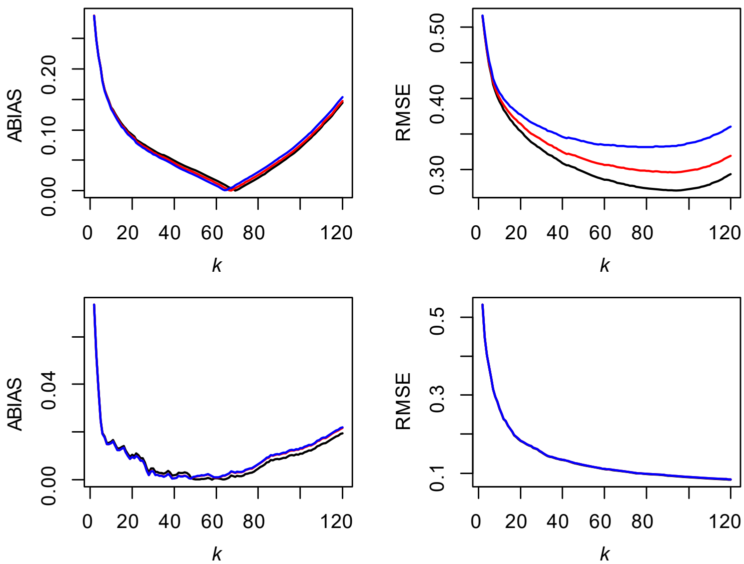

Figure 2.
Absolute bias (left two panels) and RMSE (right two panels) of $\widehat{\gamma}_1$ (black) and $\widehat{\gamma}_1^{(\mathbf{BMN})}$ (red) and $\widehat{\gamma}_1^{(\mathbf{W})}$(blue), corresponding to two situations of scenario $S_1 : (\gamma_1 = 0.8, p = 55\%)$ (top two panels) and $(\gamma_1 = 0.8, p = 90\%)$ (bottom two panels) based on 1000 samples of size 300

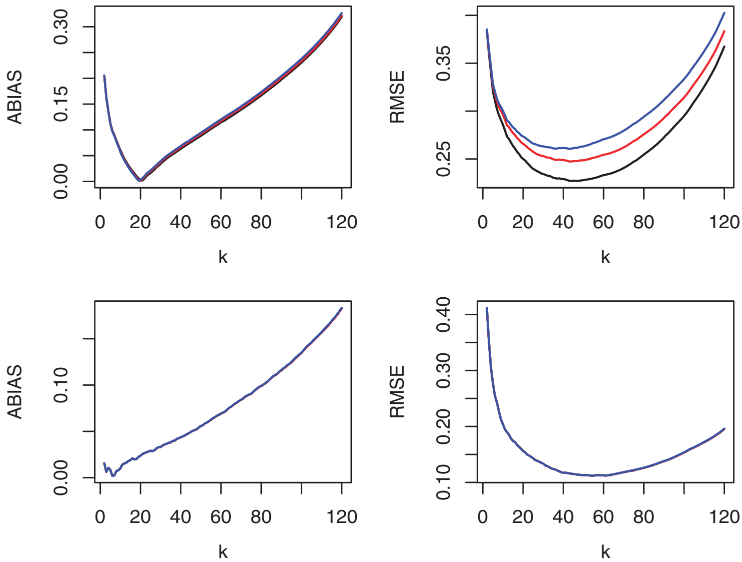

Figure 3.
Absolute bias (left two panels) and RMSE (right two panels) of $\widehat{\gamma}_1$ (black) and $\widehat{\gamma}_1^{(\mathbf{BMN})}$ (red) and $\widehat{\gamma}_1^{(\mathbf{W})}$(blue), corresponding to two situations of scenario $S_2 : (\gamma_1 = 0.6, p = 55\%)$ (top two panels) and $(\gamma_1 = 0.6, p = 90\%)$ (bottom two panels) based on 1000 samples of size 300

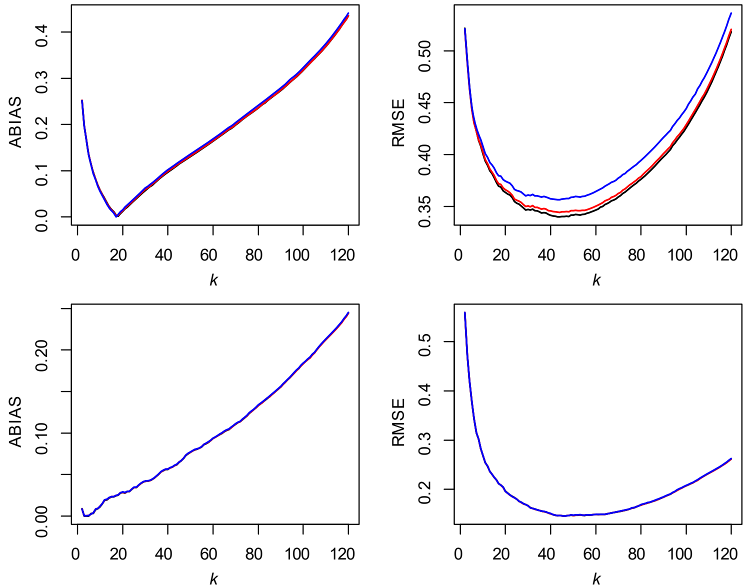

Figure 4.
Absolute bias (left two panels) and RMSE (right two panels) of $\widehat{\gamma}_1$ (black) and $\widehat{\gamma}_1^{(\mathbf{BMN})}$ (red) and $\widehat{\gamma}_1^{(\mathbf{W})}$(blue), corresponding to two situations of scenario $S_2 : (\gamma_1 = 0.8, p = 55\%)$ (top two panels) and $(\gamma_1 = 0.8, p = 90\%)$ (bottom two panels) based on 1000 samples of size 300

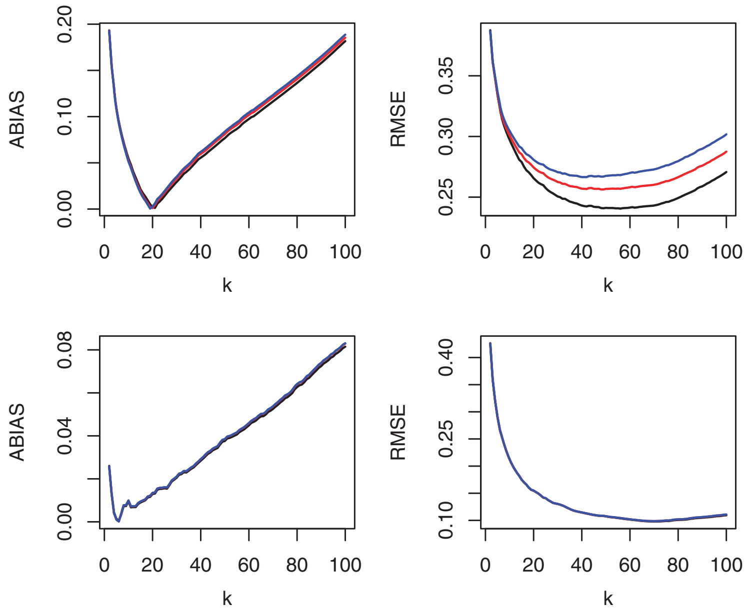

Figure 5.
Absolute bias (left two panels) and RMSE (right two panels) of $\widehat{\gamma}_1$ (black) and $\widehat{\gamma}_1^{(\mathbf{MBN})}$ (red) and $\widehat{\gamma}_1^{(\mathbf{W})}$(blue), corresponding to two situations of scenario $S_3 : (\gamma_1 = 0.6, p = 55\%)$ (top two panels) and $(\gamma_1 = 0.6, p = 90\%)$ (bottom two panels) based on 1000 samples of size 300

4. Real data example

In this section, we give an application to the AIDS data set, available in the "DTDA" *R* package and the textbook of [28] (page 19) and already used by Ref. [1]. The data present the infection and induction times for $n = 258$ adults who were infected with HIV virus and developed AIDS by June 30, 1986. The variable of interest here is the time of induction $\mathbf{T}$ of the

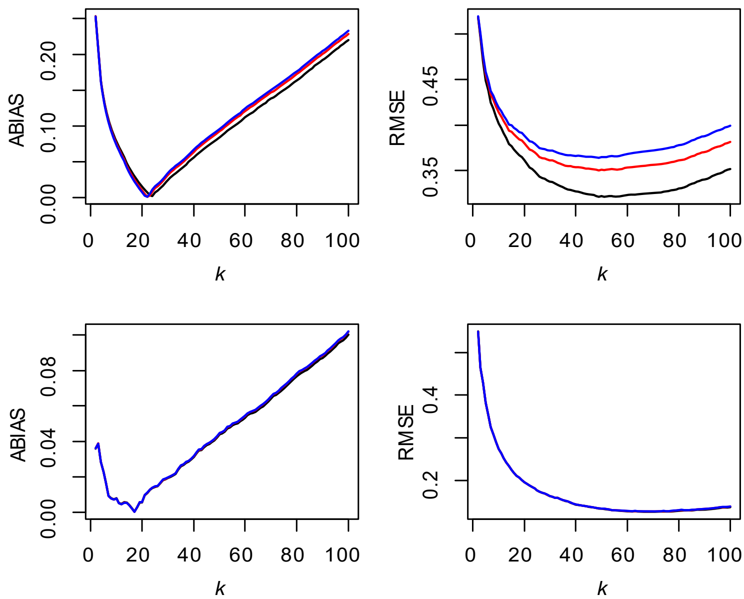

Figure 6. Absolute bias (left two panels) and RMSE (right two panels) of $\widehat{\gamma}_1$ (black) and $\widehat{\gamma}_1^{(\mathbf{BMN})}$ (red) and $\widehat{\gamma}_1^{(\mathbf{W})}$(blue), corresponding to two situations of scenario $S_3 : (\gamma_1 = 0.8,$ $p = 55\%)$ (top two panels) and $(\gamma_1 = 0.8,$ $p = 90\%)$ (bottom two panels) based on 1000 samples of size 300

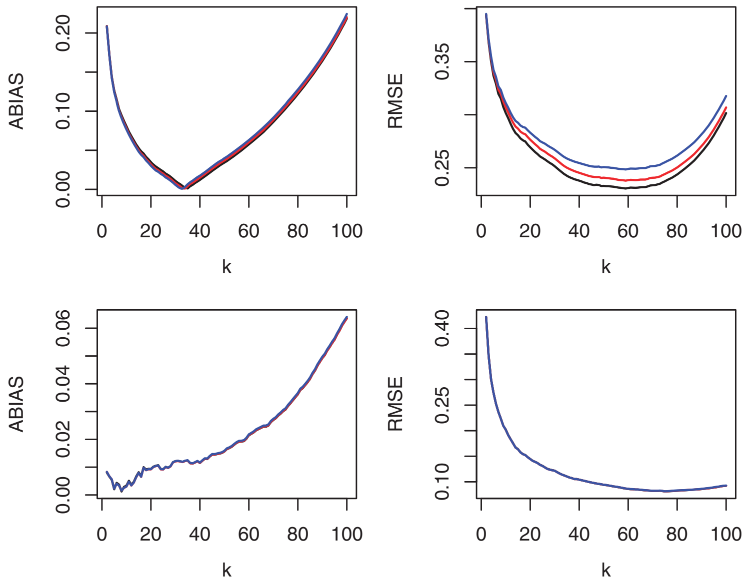

Figure 7. Absolute bias (left two panels) and RMSE (right two panels) of $\widehat{\gamma}_1$ (black) and $\widehat{\gamma}_1^{(\mathbf{BMN})}$ (red) and $\widehat{\gamma}_1^{(\mathbf{W})}$(blue), corresponding to two situations of scenario $S_4 : (\gamma_1 = 0.6,$ $p = 55\%)$ (top two panels) and $(\gamma_1 = 0.6,$ $p = 90\%)$ (bottom two panels) based on 1000 samples of size 300

disease duration which elapses between the date of infection M and the date $M + T$ of the declaration of the disease. The sample $(T_1, M_1), \ldots, (T_n, M_n)$ are taken between two fixed dates: "0" and "8", i.e. between April 1, 1978, and June 30, 1986. The initial date "0" denotes an infection occurring in the three months: from April 1, 1978, to June 30, 1978. Let us assume that M and T are the observed rv's, corresponding to the underlying rv's $\mathbf{M}$ and $\mathbf{T}$, given by the truncation scheme $0 \leq M + T \leq 8$, which in turn may be rewritten into

$$0 \leq M \leq S, \tag{4.18}$$

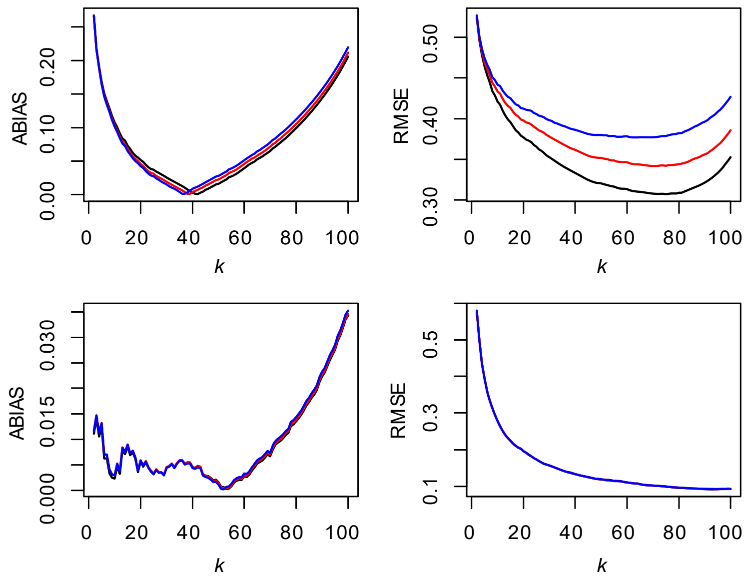

Figure 8. Absolute bias (left two panels) and RMSE (right two panels) of $\widehat{\gamma}_1$ (black) and $\widehat{\gamma}_1^{(\mathbf{BMN})}$ (red) and $\widehat{\gamma}_1^{(\mathbf{W})}$ (blue), corresponding to two situations of scenario $S_4 : (\gamma_1 = 0.8, p = 55\%)$ (top two panels) and $(\gamma_1 = 0.8, p = 90\%)$ (bottom two panels) based on 1000 samples of size 300

Table 1. Optimal sample fractions and estimate values of the tail index $\gamma_1 = 0.6$ based on 1,000 samples of size 300 for the four scenarios with $p = 0.55$

	k^*	$\widehat{\gamma}_1$	k^*	$\widehat{\gamma}_1^{(\mathbf{BMN})}$	k^*	$\widehat{\gamma}_1^{(\mathbf{W})}$
$S1$	44	0.600	41	0.599	40	0.600
$S2$	18	0.601	17	0.600	16	0.597
$S3$	21	0.601	20	0.601	19	0.599
$S4$	30	0.603	27	0.600	25	0.598

Table 2. Optimal sample fractions and estimate values of the tail index $\gamma_1 = 0.6$ based on 1,000 samples of size 300 for the four scenarios with $p = 0.9$

	k^*	$\widehat{\gamma}_1$	k^*	$\widehat{\gamma}_1^{(\mathbf{BMN})}$	k^*	$\widehat{\gamma}_1^{(\mathbf{W})}$
$S1$	82	0.610	82	0.611	82	0.611
$S2$	37	0.640	37	0.640	37	0.640
$S3$	46	0.633	37	0.625	37	0.625
$S4$	52	0.610	52	0.610	52	0.610

Table 3. Optimal sample fractions and estimate values of the tail index $\gamma_1 = 0.8$ based on 1,000 samples of size 300 for the four scenarios with $p = 0.55$

	k^*	$\widehat{\gamma}_1$	k^*	$\widehat{\gamma}_1^{(\mathbf{BMN})}$	k^*	$\widehat{\gamma}_1^{(\mathbf{W})}$
$S1$	59	0.799	57	0.800	54	0.799
$S2$	21	0.803	21	0.803	20	0.799
$S3$	24	0.802	22	0.798	22	0.801
$S4$	51	0.799	52	0.800	50	0.801

where $S:=8-T$. To work within the framework of the present paper, let us make the following transformations:

$$X:=\frac{1}{S+\epsilon} \text{ and } Y:=\frac{1}{M+\epsilon}, \tag{4.19}$$

where $\epsilon=0.05$ so that the two denominators be non-null. Thus, in view of (4.18), we have $X\leq Y$, which means that X is randomly right-truncated by Y. Thereby, for the given sample $(T_1,M_1),\ldots,(T_n,M_n)$, from (T,M), the previous transformations produce a new one $(X_1,Y_1),\ldots,(X_n,Y_n)$ from (X,Y).

Let us now denote by $\mathbf{F}$ and $\mathbf{G}$ the df's of the underling rv's $\mathbf{X}$ and $\mathbf{Y}$ corresponding to the truncated rv's X and Y, respectively. By using parametric likelihood methods, [29] fits both df's of $\mathbf{M}$ and $\mathbf{S}$ by the two-parameter Weibull model, this implies that the df's of $\mathbf{F}$ and $\mathbf{G}$ by may be fitted by two-parameter Fréchet model, namely $\mathbf{H}_{(a,r)}(x)=\exp(-a^r x^{-r}),x>0,a>0$, $r>0$, hence both $\mathbf{F}$ and $\mathbf{G}$ are heavy-tailed. The estimated parameters corresponding to the fitting of df $\mathbf{G}$ are $a_0=0.004$ and $r_0=2.1$, see also [1] page 520. Thus, one may consider that df $\mathbf{G}$ is known and equals $\mathbf{G}_\theta=\mathbf{H}_{(a_0,r_0)}$, where $\theta=(a_0,r_0)$. By using the Thomas and Reiss algorithm, given above, we compute the optimal sample fraction k^* corresponds to the tail index estimator $\widehat{\gamma}_1$ of df $\mathbf{F}$ is γ_1. We find

$$k^*=19,\ X_{n-k:n}=0.356 \text{ and } \widehat{\gamma}_1=0.917. \tag{4.20}$$

The well-known Weissman estimator [30] of the high quantile, $q_v:=\mathbf{F}^{-1}(1-v_n)$, corresponding to the underling df $\mathbf{F}$ is given by

$$\widehat{q}_v:=X_{n-k:n}\left(\frac{v}{\overline{\mathbf{F}}_n(X_{n-k:n})}\right)^{-\widehat{\gamma}_1},$$

where $v=1/(2n)$ and $\mathbf{F}_n$ is the semiparametric estimator of df $\mathbf{F}$ of $\mathbf{X}$ given in (1.11). From the values (4.20), we get $\widehat{q}_v=0.061$. Let us now compute the high quantile of $\mathbf{T}$ based on the original data, $T_1,\ldots,T_n$. Recall that $\mathbf{P}(\mathbf{X}\geq q_v)=v$ and $\mathbf{X}=\mathbf{1}/(8-\mathbf{T}+\epsilon)$, this implies that $\mathbf{P}(\mathbf{T}\geq 1/q_v-8+\epsilon)=v$, this means that $1/q_v-8+\epsilon$ is the high quantile of $\mathbf{T}$, which corresponds to the end-time t_{end} that we want to estimate. Thereby $\widehat{t}_{end}=1/\widehat{q}_v-8+10^{-2}=1/0.061-8+10^{-2}=8.40$, the value the end time of induction of AIDS is: 8 years, 4 months and 24 days.

5. Proofs

5.1 Proof of Theorem 2.1

Let us first notice that the semiparametric estimator of df $\mathbf{F}$ given in (1.12) may be rewritten into

$$\mathbf{F}_n\left(x;\widehat{\theta}_n\right)=P_n\left(\widehat{\theta}_n\right)\int_0^x\frac{dF_n^*(w)}{\overline{\mathbf{G}}_{\widehat{\theta}_n}(w)}, \tag{5.21}$$

Table 4. Optimal sample fractions and estimate values of the tail index $\gamma_1=0.8$ based on 1,000 samples of size 300 for the four scenarios with $p=0.9$

	k^*	$\widehat{\gamma}_1$	k^*	$\widehat{\gamma}_1^{(\mathbf{BMN})}$	k^*	$\widehat{\gamma}_1^{(\mathbf{W})}$
S1	90	0.804	90	0.806	90	0.807
S2	34	0.845	34	0.846	34	0.846
S3	40	0.831	40	0.831	40	0.831
S4	71	0.814	71	0.814	71	0.815

and $1/P_n\left(\widehat{\theta}\right) = \int_0^\infty dF_n^*(w)/\overline{\mathbf{G}}_{\widehat{\theta}_n}(w)$, where $F_n^*(w) := n^{-1}\sum_{i=1}^n \mathbf{1}(X_i \leq w)$ denotes the usual empirical df pertaining to the observed sample $X_1, \ldots, X_n$. It is worth mentioning that by using the strong law of large numbers $P_n\left(\widehat{\theta}_n\right) \to P(\theta)$ (almost surely) as $n \to \infty$, where $P(\theta) = 1/\int_0^\infty dF^*(w)/\overline{\mathbf{G}}_\theta(w)$ (see e.g. Lemma 3.2 in Ref. [2]. On the other hand from equation (1.2), we deduce that $p = 1/\int_0^\infty dF^*(w)/\overline{\mathbf{G}}(w)$, it follows that $p \equiv P(\theta)$ because we already assumed that $\mathbf{G} \equiv \mathbf{G}_\theta$. Next we use the distribution tail

$$\overline{\mathbf{F}}(x) = P(\theta) \int_x^\infty \frac{dF^*(w)}{\overline{\mathbf{G}}_\theta(w)}, \tag{5.22}$$

and its empirical counterpart

$$\overline{\mathbf{F}}_n\left(x; \widehat{\theta}_n\right) = P_n\left(\widehat{\theta}_n\right) \int_x^\infty \frac{dF_n^*(w)}{\overline{\mathbf{G}}_{\widehat{\theta}_n}(w)}.$$

We begin by decomposing $k^{-1/2}\mathbf{D}_n\left(x; \widehat{\theta}_n\right)$, for $x > 1$, into the sum of

$$\mathbf{M}_{n1}(x) := x^{-1/\gamma_1} \frac{\overline{\mathbf{F}}_n\left(xX_{n-k:n}; \widehat{\theta}_n\right) - \overline{\mathbf{F}}_n(xX_{n-k:n}; \theta)}{\overline{\mathbf{F}}(xX_{n-k:n})},$$

$$\mathbf{M}_{n2}(x) := x^{-1/\gamma_1} \frac{\overline{\mathbf{F}}_n(xX_{n-k:n}; \theta) - \overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}(xX_{n-k:n})},$$

$$\mathbf{M}_{n3}(x) := -\frac{\overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}_n(X_{n-k:n}; \theta)} \frac{\overline{\mathbf{F}}_n(X_{n-k:n}; \theta) - \overline{\mathbf{F}}(X_{n-k:n})}{\overline{\mathbf{F}}(X_{n-k:n})},$$

$$\mathbf{M}_{n4}(x) := \left(\frac{\overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}_n(X_{n-k:n}; \theta)} - x^{-1/\gamma_1}\right) \frac{\overline{\mathbf{F}}_n(xX_{n-k:n}; \theta) - \overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}(xX_{n-k:n})}$$

and

$$\mathbf{M}_{n5}(x) := \frac{\overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}(X_{n-k:n})} - x^{-1/\gamma_1}.$$

Our goal is to provide a weighted weak approximation to the tail empirical process $\mathbf{D}_n\left(x; \widehat{\theta}_n; \gamma_1\right)$. Let $\xi_i := \overline{F}^*(X_i)$, $i = 1, \ldots, n$ be a sequence of independent and identically distributed rv's. Recall that both df's $\mathbf{F}$ and $\mathbf{G}_\theta$ are assumed to be continuous, this implies that F^* is continuous as well, therefore $\mathbf{P}(\xi_i \leq u) = u$, this means that $(\xi_i)_{i=1,n}$ are uniformly distributed on $(0, 1)$. Let us now define the corresponding uniform tail empirical process

$$\alpha_n(s) := \sqrt{k}(\mathbf{U}_n(s) - s), \text{ for } 0 \leq s \leq 1, \tag{5.23}$$

where

$$\mathbf{U}_n(s) := k^{-1} \sum_{i=1}^n \mathbf{1}(\xi_i < ks/n), \tag{5.24}$$

denotes the tail empirical df pertaining to the sample $(\xi_i)_{i=1,n}$. In view of Proposition 3.1 of [31], there exists a Wiener process W such that for every $0 \leq \epsilon < 1/2$,

$$\sup_{0 \le s < 1} s^{-\epsilon} |\alpha_n(s) - W(s)| \xrightarrow{\mathbf{P}} 0, \text{ as } n \to \infty. \tag{5.25}$$

Let us fix a sufficiently small $0 < \epsilon < 1/2$. We will successively show that, under the first-order conditions of regular variation (1.6), we have, uniformly on $x \geq 1$, for all large n:

$$\sqrt{k}\,\mathbf{M}_{n2}(x) = \frac{\gamma}{\gamma_1} x^{1/\gamma_2} W\left(t^{-1/\gamma}\right) + \frac{\gamma}{\gamma_1} \int_{x^{1/\gamma_2}}^{\infty} W\left(t^{-\gamma_2/\gamma}\right) dt + o_{\mathbf{P}}\left(x^{\frac{1}{2}\left(\frac{1}{\gamma_2} - \frac{1}{\gamma_1}\right) + \epsilon}\right) \tag{5.26}$$

and

$$\sqrt{k}\,\mathbf{M}_{n3}(x) = -x^{-1/\gamma_1}\left(\frac{\gamma}{\gamma_1} W(1) + \frac{\gamma}{\gamma_1} \int_{1}^{\infty} W\left(t^{-\gamma_2/\gamma}\right) dt\right) + o_{\mathbf{P}}\left(x^{-1/\gamma_1 + \epsilon}\right), \tag{5.27}$$

while

$$\sqrt{k}\,\mathbf{M}_{n1}(x) = o_{\mathbf{P}}\left(x^{-1/\gamma_1 + \epsilon}\right), \ \sqrt{k}\,\mathbf{M}_{n4}(x) = o_{\mathbf{P}}\left(x^{\frac{1}{2}\left(\frac{1}{\gamma_2} - \frac{1}{\gamma_1}\right) + \epsilon}\right), \tag{5.28}$$

and

$$\sqrt{k}\,\mathbf{M}_{n5}(x) = x^{-1/\gamma_1} \frac{x^{\rho_1/\gamma_1} - 1}{\rho_1 \gamma_1} \sqrt{k}\,\mathbf{A}(a_k) + o_{\mathbf{P}}\left(x^{-1/\gamma_1}\right). \tag{5.29}$$

Throughout the proof, without loss of generality, we assume that $a\epsilon \equiv \epsilon$, for any constant $a > 0$. We point out that all the rest terms of the previous approximations are negligible in probability, uniformly on $x > 1$. Let us begin by the term $\mathbf{M}_{n1}(x)$ which may be made into

$$\frac{x^{-1/\gamma_1}}{\overline{\mathbf{F}}(xX_{n-k:n})} P_n\left(\widehat{\theta}_n\right)\left(\int_x^{\infty} \frac{dF_n^*(X_{n-k:n}w)}{\overline{\mathbf{G}}_{\widehat{\theta}}(X_{n-k:n}w)} - \int_x^{\infty} \frac{dF_n^*(X_{n-k:n}w)}{\overline{\mathbf{G}}_{\theta}(X_{n-k:n}w)}\right)$$

$$= \frac{x^{-1/\gamma_1}}{\overline{\mathbf{F}}(xX_{n-k:n})} P_n\left(\widehat{\theta}_n\right) \int_x^{\infty} \left(\frac{1}{\overline{\mathbf{G}}_{\widehat{\theta}}(X_{n-k:n}w)} - \frac{1}{\overline{\mathbf{G}}_{\theta}(X_{n-k:n}w)}\right) dF_n^*(X_{n-k:n}w).$$

Applying the mean value theorem (for several variables) to function $\theta \to 1/\overline{\mathbf{G}}_{\theta}(\cdot)$, yields

$$\frac{1}{\overline{\mathbf{G}}_{\widehat{\theta}}(z)} - \frac{1}{\overline{\mathbf{G}}_{\theta}(z)} = \sum_{i=1}^{d} \left(\widehat{\theta}_{i,n} - \theta_i\right) \frac{\overline{\mathbf{G}}_{\tilde{\theta}}^{(i)}(z)}{\overline{\mathbf{G}}_{\tilde{\theta}}^{2}(z)}, \text{ for any } z > 1,$$

where $\tilde{\theta}_n$ is such that $\tilde{\theta}_{i,n}$ is between θ_i and $\widehat{\theta}_{i,n}$, for $i = 1, \ldots, d$, therefore

$$\mathbf{M}_{n1}(x) = \frac{x^{-1/\gamma_1}}{\overline{\mathbf{F}}(xX_{n-k:n})} P_n\left(\widehat{\theta}_n\right) \sum_{i=1}^{d} \left(\widehat{\theta}_i - \theta_i\right) \int_x^{\infty} \frac{\overline{\mathbf{G}}_{\tilde{\theta}}^{(i)}(X_{n-k:n}w)}{\overline{\mathbf{G}}_{\tilde{\theta}}^{2}(X_{n-k:n}w)} dF_n^*(X_{n-k:n}w).$$

Recall that by assumptions (1.6) and $[A2]$ both $\overline{\mathbf{G}}_{\theta}$ and $\overline{\mathbf{G}}_{\theta}^{(i)}$ are regularly varying with the same index $(-1/\gamma_2)$ and, on the other hand, $X_{n-k:n} \xrightarrow{\mathbf{P}} \infty$ and $w > 1$ imply that $X_{n-k:n}w \xrightarrow{\mathbf{P}} \infty$. Applying Pooter's inequalities (1.4), we get

$$\frac{\overline{\mathbf{G}}_{\tilde{\theta}}(X_{n-k:n}w)}{\overline{\mathbf{G}}_{\tilde{\theta}}(X_{n-k:n})} = (1+o_{\mathbf{P}}(1))w^{-1/\gamma_2+\epsilon} = \frac{\overline{\mathbf{G}}^{(i)}_{\tilde{\theta}}(X_{n-k:n}w)}{\overline{\mathbf{G}}^{(i)}_{\tilde{\theta}}(X_{n-k:n})},$$

it follows that

$$\begin{aligned}\mathbf{M}_{n1}(x) &= (1+o_{\mathbf{P}}(1))P_n\left(\widehat{\theta}_n\right)\frac{x^{-1/\gamma_1}}{\overline{\mathbf{G}}_{\tilde{\theta}}(X_{n-k:n})\overline{\mathbf{F}}(xX_{n-k:n})} \\ &\quad \times \sum_{i=1}^{d}\frac{\overline{\mathbf{G}}^{(i)}_{\tilde{\theta}}(X_{n-k:n})}{\overline{\mathbf{G}}_{\tilde{\theta}}(X_{n-k:n})}\left|\widehat{\theta}_{i,n}-\theta_i\right|\int_x^{\infty} w^{1/\gamma_2-\epsilon}dF_n^*(X_{n-k:n}w).\end{aligned}$$

Under some regularity assumptions, [16] stated that $\sqrt{n}\left(\widehat{\theta}_n-\theta\right)$ is asymptotically a centered multivariate normal rv, which implies that $\widehat{\theta}_{i,n}-\theta_i = O_{\mathbf{P}}\left(n^{-1/2}\right)$ and thus $\widehat{\theta}_n \xrightarrow{\mathbf{P}} \theta$. On the other hand, by the law of large numbers $P_n(\theta) \xrightarrow{\mathbf{P}} P(\theta)$ as $n\to\infty$, then we may readily show that $P_n\left(\widehat{\theta}_n\right) \xrightarrow{\mathbf{P}} P(\theta)$ as $n\to\infty$ as well. Note that since $\widehat{\theta}_n$ is a consistent estimator of θ then $\tilde{\theta}_n$ is too. Then by using the fact that $X_{n-k:n} \xrightarrow{\mathbf{P}} \infty$ and both conditions $[A1]$ and $[A3]$, we show readily that

$$(X_{n-k:n})^{-\epsilon}\frac{\overline{\mathbf{G}}^{(i)}_{\tilde{\theta}_n}(X_{n-k:n})}{\overline{\mathbf{G}}_{\tilde{\theta}_n}(X_{n-k:n})} \xrightarrow{\mathbf{P}} 0, \text{ as } n\to\infty,$$

and $\overline{\mathbf{G}}_{\theta}(X_{n-k:n})/\overline{\mathbf{G}}_{\tilde{\theta}_n}(X_{n-k:n}) \xrightarrow{\mathbf{P}} 1$. In view of Lemma A1 in Ref. [7], we infer that $X_{n-k:n} = (1+o_{\mathbf{P}}(1))(k/n)^{-\gamma}$, thus

$$\mathbf{M}_{n1}(x) = (k/n)^{-\epsilon\gamma}o_{\mathbf{P}}\left(n^{-1/2}\right)\tilde{\mathbf{M}}_{n1}(x),$$

where

$$\tilde{\mathbf{M}}_{n1}(x) := \frac{x^{-1/\gamma_1}P(\theta)}{\overline{\mathbf{G}}_{\theta}(X_{n-k:n})\overline{\mathbf{F}}(xX_{n-k:n})}\int_x^{\infty} w^{1/\gamma_2-\epsilon}dF_n^*(X_{n-k:n}w).$$

Making use of representation (5.22), we write

$$\begin{aligned}\tilde{\mathbf{M}}_{n1}(x) &= x^{-1/\gamma_1}\left(\int_x^{\infty}\frac{\overline{\mathbf{G}}_{\theta}(X_{n-k:n})}{\overline{\mathbf{G}}_{\theta}(X_{n-k:n}w)}d\frac{F^*(X_{n-k:n}w)}{\overline{F}^*(X_{n-k:n})}\right)^{-1} \\ &\quad \times\left(\int_x^{\infty} w^{1/\gamma_2-\epsilon}d\frac{F_n^*(X_{n-k:n}w)}{\overline{F}^*(X_{n-k:n})}\right).\end{aligned} \tag{5.30}$$

Once again by using the routine manipulations of Potter's inequalities, we show that the first integral in (5.30) is equal to

$$(1+o_{\mathbf{P}}(1))\int_x^{\infty} w^{1/\gamma_2+\epsilon/2}d\frac{F^*(X_{n-k:n}w)}{\overline{F}^*(X_{n-k:n})}.$$

An integration by parts to the previous integral yields

$$x^{1/\gamma_2+\epsilon/2}\frac{\overline{F}^*(X_{n-k:n}x)}{\overline{F}^*(X_{n-k:n})}+(1/\gamma_2+\epsilon/2)\int_x^\infty w^{1/\gamma_2+\epsilon/2-1}\frac{\overline{F}^*(X_{n-k:n}w)}{\overline{F}^*(X_{n-k:n})}dw.$$

Recall that from (1.7),we have $\overline{F}^*\in\mathcal{RV}_{(-1/\gamma)}$, then

$$\frac{\overline{F}^*(X_{n-k:n}w)}{\overline{F}^*(X_{n-k:n})}=(1+o_{\mathbf{P}}(1))w^{-1/\gamma+\epsilon/2},$$

uniformly on $w>1$. Therefore, the previous quantity reduces into

$$(1+o_{\mathbf{P}}(1))\left(1+\frac{1/\gamma_2+\epsilon/2}{-1/\gamma_1+\epsilon}\right)x^{-1/\gamma_1+\epsilon}.$$

Thereby the first expression between two brackets in (5.30) equals $O_{\mathbf{P}}\left(x^{1/\gamma_1-\epsilon}\right)$. Let us consider the second factor in (5.30). By similar arguments as used for the first factor, we show that

$$x^{1/\gamma_2+\epsilon/2}\frac{\overline{F}_n^*(X_{n-k:n}x)}{\overline{F}^*(X_{n-k:n})}+(1/\gamma_2+\epsilon/2)\int_x^\infty w^{1/\gamma_2+\epsilon/2}\frac{\overline{F}_n^*(X_{n-k:n}w)}{\overline{F}^*(X_{n-k:n})}dw,$$

multiplied by $(1+o_{\mathbf{P}}(1))$, uniformly on $x>1$. From Lemma 7.1, we have

$$\frac{\overline{F}_n^*(X_{n-k:n}w)}{\overline{F}^*(X_{n-k:n})}=O_{\mathbf{P}}\left(w^{-1/\gamma+\epsilon/2}\right),$$

which implies that the previous expression equals $O_{\mathbf{P}}\left(x^{-1/\gamma_1+\epsilon}\right)$, thus $\tilde{\mathbf{M}}_{n1}(x)=O_{\mathbf{P}}\left(x^{-1/\gamma+\epsilon}\right)$ and therefore

$$\sqrt{k}\mathbf{M}_{n1}(x)=(k/n)^{1/2-\epsilon\gamma}O_{\mathbf{P}}\left(x^{-1/\gamma_1+\epsilon}\right).$$

By assumption $k/n\to 0$, it follows that $\sqrt{k}\mathbf{M}_{n1}(x)=o_{\mathbf{P}}\left(x^{-1/\gamma_1+\epsilon}\right)$ which meets the result of (5.30). Let now consider the second term $\mathbf{M}_{n2}(x)$ which may be rewritten into

$$-x^{-1/\gamma_1}\frac{k/n}{\overline{F}^*(X_{n-k:n})}\frac{\overline{\mathbf{F}}(X_{n-k:n})}{\overline{\mathbf{F}}(xX_{n-k:n})}\frac{\overline{\mathbf{G}}_\theta(X_{n-k:n})/\overline{F}^*(X_{n-k:n})}{\overline{\mathbf{F}}(X_{n-k:n})}$$
$$\times\int_x^\infty\frac{\overline{\mathbf{G}}_\theta(X_{n-k:n})}{\overline{\mathbf{G}}_\theta(X_{n-k:n}w)}d\frac{\overline{F}_n^*(X_{n-k:n}w)-\overline{F}^*(X_{n-k:n}w)}{k/n}.$$

In view of Potter's inequalities, it is clear that

$$\frac{\overline{\mathbf{F}}(X_{n-k:n})}{\overline{F}^*(X_{n-k:n})/\overline{\mathbf{G}}_\theta(X_{n-k:n})}\xrightarrow{\mathbf{P}}\frac{\gamma_1}{\gamma}P(\theta)$$

and

$$\frac{\overline{\mathbf{F}}(X_{n-k:n})}{\overline{\mathbf{F}}(xX_{n-k:n})}\xrightarrow{\mathbf{P}}x^{1/\gamma_1}.$$

Smirnov's lemma (see, e.g. Lemma 2.2.3 in Ref. [5] with the fact that $\overline{F}^{*}(X_{n-k:n})\overset{d}{=}\xi_{k+1:n}$ imply that $\frac{n}{k}\,\xi_{k+1:n}\xrightarrow{\mathbf{P}}1$, hence $\frac{n}{k}\,\overline{F}^{*}(X_{n-k:n})=1+o_{\mathbf{P}}(1)$. Therefore,

$$\mathbf{M}_{n2}(x)=-(1+o_{\mathbf{P}}(1))\frac{\gamma}{\gamma_1}\int_x^\infty\frac{\overline{\mathbf{G}}_\theta(X_{n-k:n})}{\overline{\mathbf{G}}_\theta(X_{n-k:n}w)}d\frac{\overline{F}_n^{*}(X_{n-k:n}w)-\overline{F}^{*}(X_{n-k:n}w)}{k/n}.$$

On the other hand, using an integration by parts yields

$$\mathbf{M}_{n2}(x)=(1+o_{\mathbf{P}}(1))\frac{\gamma_1}{\gamma}\left(\mathbf{M}_{n2}^{(1)}(x)+\mathbf{M}_{n2}^{(2)}(x)\right),$$

where

$$\mathbf{M}_{n2}^{(1)}(x):=\int_x^\infty\frac{\overline{F}_n^{*}(X_{n-k:n}w)-\overline{F}^{*}(X_{n-k:n}w)}{k/n}d\frac{\overline{\mathbf{G}}_\theta(X_{n-k:n})}{\overline{\mathbf{G}}_\theta(X_{n-k:n}w)}$$

and

$$\mathbf{M}_{n2}^{(2)}(x):=\frac{\overline{\mathbf{G}}_\theta(X_{n-k:n})}{\overline{\mathbf{G}}_\theta(X_{n-k:n}x)}\frac{\overline{F}_n^{*}(X_{n-k:n}x)-\overline{F}^{*}(xX_{n-k:n})}{k/n}.$$

By using the change of variables $t=\overline{\mathbf{G}}_\theta(X_{n-k:n})/\overline{\mathbf{G}}_\theta(X_{n-k:n}w)$, it is easy to verify that

$$\mathbf{M}_{n2}^{(1)}(x)=\int_{\frac{\overline{\mathbf{G}}_\theta(X_{n-k:n})}{\overline{\mathbf{G}}_\theta(X_{n-k:n}x)}}^\infty\frac{n}{k}\left\{\overline{F}_n^{*}\left(\mathbf{G}_\theta^{\leftarrow}\left(1-\overline{\mathbf{G}}_\theta(X_{n-k:n})t^{-1}\right)\right)-\overline{F}^{*}\left(\mathbf{G}_\theta^{\leftarrow}\left(1-\overline{\mathbf{G}}_\theta(X_{n-k:n})t^{-1}\right)\right)\right\}dt.$$

Observe that

$$\mathbf{M}_{n2}^{(1)}(x)=\int_{\frac{\overline{\mathbf{G}}_\theta(X_{n-k:n})}{\overline{\mathbf{G}}_\theta(X_{n-k:n}x)}}^\infty(\mathbf{U}_n(\vartheta_n(t;\theta))-\vartheta_n(t;\theta))dt,$$

where $\vartheta_n(t;\theta):=\frac{n}{k}\overline{F}^{*}\left(\mathbf{G}_\theta^{\leftarrow}\left(1-\overline{\mathbf{G}}_\theta(X_{n-k:n})t^{-1}\right)\right)$ and $\mathbf{U}_n$ are the tail empirical df given in (5.24). Thereby,

$$\sqrt{k}\,\mathbf{M}_{n2}^{(1)}(x)=\int_{\frac{\overline{\mathbf{G}}_\theta(X_{n-k:n})}{\overline{\mathbf{G}}_\theta(X_{n-k:n}x)}}^\infty\alpha_n(\vartheta_n(t;\theta))dt,$$

with α_n being the tail empirical process defined in (5.23). Let us decompose the previous integral into

$$\begin{aligned}&\int_{\frac{\overline{\mathbf{G}}_\theta(X_{n-k:n})}{\overline{\mathbf{G}}_\theta(X_{n-k:n}x)}}^\infty(\alpha_n(\vartheta_n(t;\theta))-W(\vartheta_n(t;\theta)))dt+\int_{\frac{\overline{\mathbf{G}}_\theta(X_{n-k:n})}{\overline{\mathbf{G}}_\theta(X_{n-k:n}x)}}^\infty W(\vartheta_n(t;\theta))dt\\&=S_n+R_n.\end{aligned}$$

By applying weak approximation (5.25), we get

$$S_n=o_{\mathbf{P}}(1)\int_{\frac{\overline{\mathbf{G}}_\theta(X_{n-k:n})}{\overline{\mathbf{G}}_\theta(X_{n-k:n}x)}}^\infty(\vartheta_n(t;\theta))^{1/2-\epsilon}dt.$$

Observe that $\overline{F}^*\left(\mathbf{G}_\theta^{\leftarrow}\left(1-\overline{\mathbf{G}}_\theta(X_{n-k:n})\right)\right)=\overline{F}^*(X_{n-k:n})$, thereby

$$\vartheta_n(t;\theta)=\frac{n}{k}\,\overline{F}^*(X_{n-k:n})\frac{\overline{F}^*\left(\mathbf{G}_\theta^{\leftarrow}\left(1-\overline{\mathbf{G}}_\theta(X_{n-k:n})t^{-1}\right)\right)}{\overline{F}^*\left(\mathbf{G}_\theta^{\leftarrow}\left(1-\overline{\mathbf{G}}_\theta(X_{n-k:n})\right)\right)}.$$

It is easy to check that $\overline{F}^*\left(\mathbf{G}_\theta^{\leftarrow}(1-\cdot)\right)\in\mathcal{RV}(\gamma_2/\gamma)$, then once again by means of Pooter's inequality, we show that $\vartheta_n(t;\theta)=(1+o_{\mathbf{P}}(1))t^{-\gamma_2/\gamma+\epsilon}$, therefore

$$S_n=o_{\mathbf{P}}(1)\int_{\frac{\overline{\mathbf{G}}_\theta(X_{n-k:n})}{\overline{\mathbf{G}}_\theta(X_{n-k:n}x)}}^{\infty}\left(t^{-\gamma_2/\gamma+\epsilon}\right)^{1/2-\epsilon}dt.$$

By using an elementary integration, we get

$$S_n=o_{\mathbf{P}}(1)\left(\frac{\overline{\mathbf{G}}_\theta(X_{n-k:n})}{\overline{\mathbf{G}}_\theta(X_{n-k:n}x)}\right)^{(-\gamma_2/\gamma+\epsilon)(1/2-\epsilon)+1}=o_{\mathbf{P}}\left(x^{\frac{1}{\gamma_2}-\frac{1}{2\gamma}+\epsilon}\right).$$

By replacing γ by its by its expression given in (1.8), we end up with

$$S_n=o_{\mathbf{P}}\left(x^{\frac{1}{2}\left(\frac{1}{\gamma_2}-\frac{1}{\gamma_1}\right)+\epsilon}\right).$$

The term R_n may be decomposed into

$$\int_{\frac{\overline{\mathbf{G}}_\theta(X_{n-k:n})}{\overline{\mathbf{G}}_\theta(X_{n-k:n}x)}}^{x^{1/\gamma_2}}W(\vartheta_n(t;\theta))dt+\int_{x^{1/\gamma_2}}^{\infty}W(\vartheta_n(t;\theta))dt=R_{n1}+R_{n2}.$$

It is clear that

$$|R_{n1}|<\left\{\sup_{t>\frac{\overline{\mathbf{G}}_\theta(X_{n-k:n})}{\overline{\mathbf{G}}_\theta(X_{n-k:n}x)}}\frac{|W(\vartheta_n(t;\theta))|}{(\vartheta_n(t;\theta))^\epsilon}\right\}\int_{\frac{\overline{\mathbf{G}}_\theta(X_{n-k:n})}{\overline{\mathbf{G}}_\theta(X_{n-k:n}x)}}^{x^{1/\gamma_2}}(\vartheta_n(t;\theta))^\epsilon dt.$$

It is ready to check, by using the change of variables $\vartheta_n(t;\theta)=s$, that the previous first factor between the curly brackets equals

$$\sup_{0<s<\frac{n}{k}\overline{F}^*(X_{n-k:n}x;\theta)}\frac{|W(s)|}{s^\epsilon}<\sup_{0<s<\frac{n}{k}\overline{F}^*(X_{n-k:n};\theta)}\frac{|W(s)|}{s^\epsilon}.$$

From Lemma 3.2 in Ref. [31] $\sup_{0<s\leq 1}s^{-\delta}|W(s)|=O_{\mathbf{P}}(1)$, for any $0<\delta<1/2$, then since $n\overline{F}^*(X_{n-k:n};\theta)/k\xrightarrow{\mathbf{P}}1$, as $n\to\infty$, we infer that

$$\sup_{0<s<\frac{n}{k}\overline{F}^*(X_{n-k:n};\theta)}s^{-\epsilon}|W(s)|=O_{\mathbf{P}}(1).$$

for all large n. On the other hand, we already pointed out above that

$$\vartheta_n(t;\theta)=(1+o_{\mathbf{P}}(1))t^{-\gamma_2/\gamma+\epsilon},$$

which implies that the second factor is equal to

$$O_{\mathbf{P}}(1)\int_{\frac{\overline{\mathbf{G}}_\theta(X_{n-k:n})}{\overline{\mathbf{G}}_\theta(X_{n-k:n}x)}}^{x^{1/\gamma_2}}\left(t^{-\gamma_2/\gamma+\epsilon}\right)^{\epsilon}dt = O_{\mathbf{P}}(1)\int_{\frac{\overline{\mathbf{G}}_\theta(X_{n-k:n})}{\overline{\mathbf{G}}_\theta(X_{n-k:n}x)}}^{x^{1/\gamma_2}} t^{-\epsilon\gamma_2/\gamma+\epsilon}dt,$$

which after integration yields

$$O_{\mathbf{P}}(1)\left\{\left(\frac{\overline{\mathbf{G}}_\theta(X_{n-k:n})}{\overline{\mathbf{G}}_\theta(X_{n-k:n}x)}\right)^{-\epsilon\gamma_2/\gamma+\epsilon+1}-\left(x^{-1/\gamma}\right)^{-\epsilon\gamma_2/\gamma+\epsilon+1}\right\}.$$

Recall that from formula (1.8), we have $\gamma_2/\gamma>1$, then by using the mean value theorem and Pooter's inequalities, we get $R_{n1}=o_{\mathbf{P}}(x^{-\epsilon})$. The second term R_{n2} may be decomposed into

$$R_{n2}=\int_{x^{1/\gamma_2}}^{\infty}\left(W(\vartheta_n(t;\theta))-W\left(t^{-\gamma_2/\gamma}\right)\right)dt+\int_{x^{1/\gamma_2}}^{\infty}W\left(t^{-\gamma_2/\gamma}\right)dt.$$

From Proposition B.1.10 in Ref. [5], we have with high probability,

$$c_n(t;\theta):=\left|\vartheta_n(t;\theta)-t^{-\gamma_2/\gamma}\right|\leq\epsilon t^{-\gamma_2/\gamma-\epsilon},\ \text{as}\ n\to\infty, \tag{5.31}$$

this means that $\sup_{x>1}\sup_{t>x^{1/\gamma_2}}c_n(t;\theta)\xrightarrow{\mathbf{P}}0$, as $n\to\infty$. This implies by using Levy's modulus of continuity of the Wiener process (see, e.g. Theorem 1.1.1 in Ref. [32]) that

$$\left|W(\vartheta_n(t;\theta))-W\left(t^{-\gamma_2/\gamma}\right)\right|\leq 2\sqrt{c_n(t;\theta)\log(1/c_n(t;\theta))},$$

with high probability. By using the fact that $\log\ s<\epsilon s^{-\epsilon}$, for $s\downarrow 0$ together with inequality (5.31), we show that

$$\left|W(\vartheta_n(t;\theta))-W\left(t^{-\gamma_2/\gamma}\right)\right|<2\epsilon t^{-(\gamma_2/\gamma-\epsilon)/2},$$

uniformly on $t>x^{1/\gamma_2}$, it follows that

$$\left|\int_{x^{1/\gamma_2}}^{\infty}\left(W(\vartheta_n(t;\theta))-W\left(t^{-\gamma_2/\gamma}\right)\right)dt\right|=o_{\mathbf{P}}(1)\left|\int_{x^{1/\gamma_2}}^{\infty}t^{-(\gamma_2/\gamma-\epsilon)/2}dt\right|.$$

Recall that the assumption $\gamma_1<\gamma_2$ together with equation $1/\gamma=1/\gamma_1+1/\gamma_2$, imply that $\gamma_2/(2\gamma)>1$, thus $-(\gamma_2/\gamma-\epsilon)/2+1<0$, therefore $\left|\int_{x^{1/\gamma_2}}^{\infty}t^{-(\gamma_2/\gamma-\epsilon)/2}dt\right|=o_{\mathbf{P}}\left(x^{-1/\gamma_1-\epsilon}\right)$. Then we showed that

$$R_{n1}=o_{\mathbf{P}}(x^{-\epsilon})\ \text{and}\ R_{n2}=\int_{x^{1/\gamma_2}}^{\infty}W\left(t^{-\gamma_2/\gamma}\right)dt+o_{\mathbf{P}}\left(x^{-1/\gamma_1-\epsilon}\right),$$

hence

$$\sqrt{k}\mathbf{M}_{n2}^{(1)}(x)=R_n+S_n=\int_{x^{1/\gamma_2}}^{\infty}W\left(t^{-\gamma_2/\gamma}\right)dt+o_{\mathbf{P}}\left(x^{-1/\gamma_1-\epsilon}\right)+o_{\mathbf{P}}\left(x^{\frac{1}{2}\left(\frac{1}{\gamma_2}-\frac{1}{\gamma_1}\right)+\epsilon}\right).$$

It is clear that

$$\left(-\frac{1}{\gamma_1}-\epsilon\right)-\left(\frac{1}{2}\left(\frac{1}{\gamma_2}-\frac{1}{\gamma_1}\right)+\epsilon\right)=-\frac{\gamma_1+\gamma_2+4\epsilon\gamma_1\gamma_2}{2\gamma_1\gamma_2}<0.$$

then

$$\sqrt{k}\mathbf{M}_{n2}^{(1)}(x) = \int_{x^{1/\gamma_2}}^{\infty} W\left(t^{-\gamma_2/\gamma}\right)dt + o_{\mathbf{P}}\left(x^{\frac{1}{2}\left(\frac{1}{\gamma_2}-\frac{1}{\gamma_1}\right)+\epsilon}\right).$$

By using similar arguments, we end up with

$$\sqrt{k}\mathbf{M}_{n2}^{(2)}(x) = x^{1/\gamma_2} W\left(t^{-1/\gamma}\right) + o_{\mathbf{P}}\left(x^{-\frac{1}{\gamma_1}+\epsilon}\right),$$

therefore, we omit further details. Finally, we have

$$\sqrt{k}\mathbf{M}_{n2}(x) = \frac{\gamma}{\gamma_1} x^{1/\gamma_2} W\left(t^{-1/\gamma}\right) + \frac{\gamma}{\gamma_1}\int_{x^{1/\gamma_2}}^{\infty} W\left(t^{-\gamma_2/\gamma}\right)dt + o_{\mathbf{P}}\left(x^{\frac{1}{2}\left(\frac{1}{\gamma_2}-\frac{1}{\gamma_1}\right)+\epsilon}\right).$$

Let us now focus on the term $\mathbf{M}_{n3}(x)$. From the latter approximation, we infer that

$$\begin{aligned}\sqrt{k}\mathbf{M}_{n2}(1) &= \sqrt{k}\,\frac{\overline{\mathbf{F}}_n(X_{n-k:n};\theta) - \overline{\mathbf{F}}(X_{n-k:n})}{\overline{\mathbf{F}}(X_{n-k:n})} \\ &= \frac{\gamma}{\gamma_1} W(1) + \frac{\gamma}{\gamma_1}\int_1^{\infty} W\left(t^{-\gamma_2/\gamma}\right)dt + o_{\mathbf{P}}(1),\end{aligned} \tag{5.32}$$

which implies that

$$\sqrt{k}\,\frac{\overline{\mathbf{F}}_n(X_{n-k:n};\theta) - \overline{\mathbf{F}}(X_{n-k:n})}{\overline{\mathbf{F}}(X_{n-k:n})} = O_{\mathbf{P}}(1).$$

In other words, we have

$$\frac{\overline{\mathbf{F}}_n(X_{n-k:n};\theta)}{\overline{\mathbf{F}}(X_{n-k:n})} = 1 + O_{\mathbf{P}}\left(k^{-1/2}\right). \tag{5.33}$$

The regular variation of $\overline{\mathbf{F}}(\cdot)$ and (5.33) together imply that

$$\frac{\overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}_n(X_{n-k:n};\theta)} = x^{-1/\gamma_1} + o_{\mathbf{P}}\left(x^{-1/\gamma_1+\epsilon}\right). \tag{5.34}$$

By combining the results (5.32) and (5.34), we get

$$\sqrt{k}\mathbf{M}_{n3}(x) = -x^{-1/\gamma_2}\left(\frac{\gamma}{\gamma_1} W(1) + \frac{\gamma}{\gamma_1}\int_1^{\infty} W\left(t^{-\gamma_2/\gamma}\right)dt\right) + o_{\mathbf{P}}\left(x^{-1/\gamma_1+\epsilon}\right).$$

For the fourth term $\mathbf{M}_{n4}(x)$, we write

$$\sqrt{k}\mathbf{M}_{n4}(x) = \left(\frac{\overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}_n(X_{n-k:n};\theta)} - x^{-1/\gamma_1}\right)\left(\sqrt{k}\,\frac{\overline{\mathbf{F}}_n(xX_{n-k:n};\theta) - \overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}(xX_{n-k:n})}\right).$$

From (5.34) the first factor of the previous equation equals $o_{\mathbf{P}}\left(x^{-1/\gamma_1+\epsilon}\right)$. On the other hand, the change of variables $s = t^{-\gamma_2/\gamma}$ yields

$$\int_{x^{1/\gamma_2}}^{\infty} W\left(t^{-\gamma_2/\gamma}\right)dt = \frac{\gamma}{\gamma_2}\int_0^{x^{-1/\gamma}} s^{-\gamma/\gamma_2-1} W(s)ds.$$

Since $\sup_{0<s<1} s^{-1/2+\epsilon}|W(s)| = O_{\mathbf{P}}(1)$, then we easily show that

$$\int_{x^{1/\gamma_2}}^{\infty} W\left(t^{-\gamma_2/\gamma}\right)dt = O_{\mathbf{P}}\left(x^{\frac{1}{2}\left(\frac{1}{\gamma_2}-\frac{1}{\gamma_1}\right)+\epsilon}\right),$$

it follows that $\sqrt{k}M_{n2}(x) = O_{\mathbf{P}}\left(x^{\frac{1}{2}\left(\frac{1}{\gamma_2}-\frac{1}{\gamma_1}\right)+\epsilon}\right)$ as well. Therefore,

$$\sqrt{k}\,\frac{\overline{\mathbf{F}}_n(xX_{n-k:n};\theta) - \overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}(xX_{n-k:n})} = x^{1/\gamma_1}O_{\mathbf{P}}\left(x^{\frac{1}{2}\left(\frac{1}{\gamma_2}-\frac{1}{\gamma_1}\right)+\epsilon}\right) = O_{\mathbf{P}}\left(x^{\frac{1}{2\gamma}+\epsilon}\right).$$

Hence, we have

$$\sqrt{k}\mathbf{M}_{n4}(x) = o_{\mathbf{P}}\left(x^{-1/\gamma_1+\epsilon}\right)O_{\mathbf{P}}\left(x^{\frac{1}{2\gamma}+\epsilon}\right) = o_{\mathbf{P}}\left(x^{\frac{1}{2}\left(\frac{1}{\gamma_2}-\frac{1}{\gamma_1}\right)+\epsilon}\right).$$

By assumption, $\overline{\mathbf{F}}$ satisfies the second-order condition of regular variation (1.5), this means that for

$$\lim_{t\to\infty}\frac{\overline{\mathbf{F}}(tx)/\overline{\mathbf{F}}(t) - x^{-1/\gamma_1}}{\mathbf{A}(t)} = x^{-1/\gamma_1}\frac{x^{\rho_1/\gamma_1}-1}{\rho_1\gamma_1}, \tag{5.35}$$

for any $x > 0$, where $\rho_1 < 0$ is the second-order parameter and $\mathbf{A}$ is $\mathcal{RV}(\rho_1/\gamma_1)$. The uniform inequality corresponding to (5.35) says: there exist $t_0 > 0$, such that for any $t > t_0$, we have

$$\left|\frac{\overline{\mathbf{F}}(tx)/\overline{\mathbf{F}}(t) - x^{-1/\gamma_1}}{\mathbf{A}(t)} - x^{-1/\gamma_1}\frac{x^{\rho_1/\gamma_1}-1}{\rho_1\gamma_1}\right| < \epsilon x^{-1/\gamma_1+\rho_1/\gamma_1+\epsilon},$$

see for instance assertion (2.3.23) of Theorem 2.3.9 in Ref. [5]. It is easy to check that the latter inequality implies that

$$\begin{aligned}\sqrt{k}\mathbf{M}_{n5}(x) &= \sqrt{k}\left(\frac{\overline{\mathbf{F}}(xX_{n-k:n})}{\overline{\mathbf{F}}(X_{n-k:n})} - x^{-1/\gamma_1}\right)\\ &= x^{-1/\gamma_1}\frac{x^{\rho_1/\gamma_1}-1}{\rho_1\gamma_1}\sqrt{k}\mathbf{A}(X_{n-k:n}) + o_{\mathbf{P}}\left(x^{-1/\gamma_1}\frac{x^{\rho_1/\gamma_1}-1}{\rho_1\gamma_1}\sqrt{k}\,\mathbf{A}(X_{n-k:n})\right).\end{aligned}$$

Recall that $a_k = F^{*} \leftarrow (1-k/n)$ and notice that $X_{n-k:n}/a_k \xrightarrow{\mathbf{P}} 1$ as $n \to \infty$, then in view of the regular variation of $\mathbf{A}$, we infer that $\mathbf{A}(X_{n-k:n}) = (1 + o_{\mathbf{P}}(1))\mathbf{A}(a_k)$. On the other hand, by assumption $\sqrt{k}\mathbf{A}(a_k)$ is asymptotically bounded, therefore

$$\sqrt{k}\mathbf{M}_{n5}(x) = x^{-1/\gamma_1}\frac{x^{\rho_1/\gamma_1}-1}{\rho_1\gamma_1}\sqrt{k}\mathbf{A}(a_k) + o_{\mathbf{P}}\left(x^{-1/\gamma_1}\right).$$

To summarize, at this stage, we showed that

$$\begin{aligned}\mathbf{D}_n\left(x;\widehat{\theta}\right) &= \frac{\gamma}{\gamma_1}x^{1/\gamma_2}W\left(t^{-1/\gamma}\right)+\frac{\gamma}{\gamma_1}\int_{x^{1/\gamma_2}}^{\infty}W\left(t^{-\gamma_2/\gamma}\right)dt\\ &-x^{-1/\gamma_2}\left(\frac{\gamma}{\gamma_1}W(1)+\frac{\gamma}{\gamma_1}\int_1^{\infty}W\left(t^{-\gamma_2/\gamma}\right)dt\right)\\ &+x^{-1/\gamma_1}\frac{x^{\rho_1/\gamma_1}-1}{\rho_1\gamma_1}\sqrt{k}\mathbf{A}(a_k)+\varsigma(x),\end{aligned}$$

where $\varsigma(x):=o_{\mathbf{P}}\left(x^{-1/\gamma_1+\epsilon}\right)+o_{\mathbf{P}}\left(x^{-1/\gamma_1}\right)+o_{\mathbf{P}}\left(x^{\frac{1}{2}\left(\frac{1}{\gamma_2}-\frac{1}{\gamma_1}\right)+\epsilon}\right)$. By using a change of variables, we show that sum of the first three terms equals the Gaussian process $\Gamma(x;W)$ stated in Theorem 2.1. Recall that $\gamma_1<\gamma_2$ and

$$\frac{1}{2}\left(\frac{1}{\gamma_2}-\frac{1}{\gamma_1}\right)+\epsilon<0,$$

then it is easy to verify that $\varsigma(x)=o_{\mathbf{P}}\left(x^{\frac{1}{2}\left(\frac{1}{\gamma_2}-\frac{1}{\gamma_1}\right)+\epsilon}\right)$. It follows that

$$\begin{aligned}&x^{\epsilon}\left\{\mathbf{D}_n\left(x;\widehat{\theta}\right)-\Gamma(x;W)-x^{-1/\gamma_1}\frac{x^{\rho_1/\gamma_1}-1}{\rho_1\gamma_1}\sqrt{k}\,\mathbf{A}(a_k)\right\}\\ &=o_{\mathbf{P}}\left(x^{\frac{1}{2}\left(\frac{1}{\gamma_2}-\frac{1}{\gamma_1}\right)+2\epsilon}\right)=o_{\mathbf{P}}(1),\end{aligned}$$

uniformly on $x>1$, therefore

$$\sup_{x>1}x^{\epsilon}\left|\mathbf{D}_n\left(x;\widehat{\theta}\right)-\Gamma(x;W)-x^{-1/\gamma_1}\frac{x^{\rho_1/\gamma_1}-1}{\rho_1\gamma_1}\sqrt{k}\,\mathbf{A}(a_k)\right|=o_{\mathbf{P}}(1),$$

for any sample $0<\epsilon<1/2$, which completes the proof of Theorem 2.1.

5.2 Proof of Theorem 2.2

From the representation (1.16), we write

$$\widehat{\gamma}_1-\gamma_1=T_{n1}+T_{n2}+T_{n3},$$

where

$$T_{n1}:=k^{-1/2}\int_1^{\infty}x^{-1}\left\{\mathbf{D}_n\left(x;\widehat{\theta};\gamma_1\right)-\Gamma(x;W)-x^{-1/\gamma_1}\frac{x^{\rho_1/\gamma_1}-1}{\rho_1\gamma_1}\sqrt{k}\,\mathbf{A}(a_k)\right\}dx$$

$$T_{n2}:=k^{-1/2}\int_1^{\infty}x^{-1}\Gamma(x;W)dx$$

and

$$T_{n3}:=-\mathbf{A}(a_k)\int_1^{\infty}x^{-1/\gamma_1-1}\frac{x^{\rho_1/\gamma_1}-1}{\rho_1\gamma_1}dx.$$

Using Theorem 2.1 yields $T_{n1} = o_{\mathbf{P}}\left(k^{-1/2}\right)\int_1^\infty x^{-1+\epsilon}dx = o_{\mathbf{P}}\left(k^{-1/2}\right) = o_{\mathbf{P}}(1)$. Since $\mathbf{E}|W(s)| \leq s^{1/2}$, then it is easy to show that $\int_1^\infty x^{-1}\Gamma(x;W)dx = O_{\mathbf{P}}(1)$, it follows that $T_{n2} = O_{\mathbf{P}}\left(k^{-1/2}\right) = o_{\mathbf{P}}(1)$. Using an elementary integration, we get $T_{n3} = \mathbf{A}(a_k)/(1-\rho_1)$ which tends to zero as $n \to \infty$, because $a_k \to \infty$ and $|\mathbf{A}|$ is regularly varying with negative index. Therefore, $\widehat{\gamma}_1 \xrightarrow{\mathbf{P}} \gamma_1$, as $n \to \infty$ which gives the first result of Theorem. To establish the asymptotic normality, we write

$$\sqrt{k}(\widehat{\gamma}_1 - \gamma_1) = \sqrt{k}T_{n1} + \sqrt{k}T_{n2} + \sqrt{k}T_{n3},$$

where

$$\sqrt{k}T_{n1} = o_{\mathbf{P}}(1), \sqrt{k}T_{n2} = \int_1^\infty x^{-1}\Gamma(x;W)dx$$

and

$$\sqrt{k}T_{n3} = \frac{\sqrt{k}\mathbf{A}(a_k)}{1-\rho_1}.$$

Note that $\Gamma(x;W)$ is a centered Gaussian process and by using the assumption $\sqrt{k}\mathbf{A}(a_k) \to \lambda < \infty$, we end up with

$$\sqrt{k}(\widehat{\gamma}_1 - \gamma_1) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\lambda}{1-\rho_1}, \mathbf{E}\left[\int_1^\infty x^{-1}\Gamma(x;W)dx\right]^2\right).$$

By elementary calculations (we omit the details), we show that

$$\mathbf{E}\left[\int_1^\infty x^{-1}\Gamma(x;W)dx\right]^2 = \sigma^2.$$

6. Conclusion

On the basis of a semiparametric estimator of the underlying distribution function, we proposed a new estimation method to the tail index of Pareto-type distributions for randomly right-truncated data. Compared with the existing ones, this estimator behaves well both in terms of bias and RMSE. A useful weak approximation of the corresponding tail empirical process allowed us to establish both the consistency and asymptotic normality of the proposed estimator.

References

1. Lagakos SW, Barraj LM, De Gruttola V. Nonparametric analysis of truncated survival data, with applications to AIDS. Biometrika. 1988; 75: 515-23.
2. Wang MC. A semiparametric model for randomly truncated data. J Amer Statist Assoc. 1989; 84: 742-8.
3. Lawless JF. Statistical models and methods for lifetime data. 2nd ed. Wiley Series in Probability and Statistics; 2002.
4. Gardes L, Stupfler G. Estimating extreme quantiles under random truncation. TEST. 2015; 24: 207-27.
5. de Haan L, Ferreira A. Extreme value theory: an introduction. Springer; 2006.
6. de Haan L, Stadtmüller U. Generalized regular variation of second order. J Aust Math. Soc. (Series A). 1996; 61: 381-95.

7. Benchaira S, Meraghni D, Necir A. Tail product-limit process for truncated data with application to extreme value index estimation. Extremes. 2016a; 19: 219-51.
8. Hill BM. A simple general approach to inference about the tail of a distribution. Ann Statist. 1975; 3: 1163-74.
9. Benchaira S, Meraghni D, Necir A. On the asymptotic normality of the extreme value index for right-truncated data. Statist Probab Lett. 2015; 107: 378-84.
10. Worms J, Worms R. A Lynden-Bell integral estimator for extremes of randomly truncated data. Statist Probab Lett. 2016; 109: 106-17.
11. Lynden-Bell D. A method of allowing for known observational selection in small samples applied to 3CR quasars. Monthly Notices Roy Astron Soc. 1971; 155: 95-118.
12. Woodroofe M. Estimating a distribution function with truncated data. Ann Statist. 1985; 13: 163-77.
13. Benchaira S, Meraghni D, Necir A. Kernel estimation of the tail index of a right-truncated Pareto-type distribution. Statist Probab Lett. 2016b; 119: 186-93.
14. Haouas N, Necir A, Brahimi B. Estimating the second-order parameter of regular variation and bias reduction in tail index estimation under random truncation. J Stat Theor Pract. 2019; 13: 110-144.
15. Haouas N, Necir A, Meraghni D, Brahimi B. A Lynden-Bell integral estimator for the tail index of right-truncated data with a random threshold. Afr Stat. 2018; 12: 1159-70.
16. Andersen EB. Asymptotic properties of conditional maximum-likelihood estimators. J Roy Statist Soc Ser B. 1970; 32: 283-301.
17. Moreira C, de Uña-Álvarez J. A semiparametric estimator of survival for doubly truncated data. Stat Med. 2010; 29: 3147-59.
18. Bilker WB, Wang MC. Asemiparametric extension of the Mann–Whitney test for randomly truncated data. Biometrics. 1996; 52: 10-20.
19. Li G, Qin J, Tiwari RC. Semiparametric likelihood ratio-based inferences for truncated data. J Amer Statist Assoc. 1997; 92: 236-45.
20. Moreira C, de Uña-Álvarez J, Van Keilegom I. Goodness-of-fit tests for a semiparametric model under random double truncation. Comput Statist. 2014; 29: 1365-137.
21. Qin J, Wang MC. Semiparametric analysis of truncated data. Lifetime Data Anal. 2001; 7(3): 225-42.
22. Shen PS. Semiparametric analysis of doubly truncated data. Comm Statist Theor Methods. 2010; 39: 3178-90.
23. Shen PS, Hsu H. Conditional maximum likelihood estimation for semiparametric transformation models with doubly truncated data. Comput Statist Data Anal. 2020; 144: 15: 106862.
24. Drees H, de Haan L, Li D. Approximations to the tail empirical distribution function with application to testing extreme value conditions. J Statist Plann Inference. 2006; 136: 3498-538.
25. Koning AJ, Peng L. Goodness-of-fit tests for a heavy tailed distribution. J Statist Plann Inference. 2008; 138: 3960-81.
26. Caeiro F, Gomes MI. Threshold selection in extreme value analysis. Chapter in. In: Dey D, Yan J, (Eds). Extreme value modeling and risk analysis: methods and applications. 9781498701297. Chapman-Hall/CRC; 2015. p. 69-87.
27. Reiss RD, Thomas M. Statistical analysis of extreme values with applications to insurance, finance, hydrology and other fields. 3rd ed. Basel, Boston, Berlin: Birkhäuser Verlag; 2007.
28. Klein JP, Moeschberger S. Survival analysis: techniques for censored and truncated data. Berlin: Springer; 1997. doi: 10.1007/978-1-4757-2728-9.
29. Lui KJ, Lawrence DN, Morgan WM, Peterman TA, Haverkos HH, Breakman DJ. A model-based approach for estimating the mean incubation period of transfusion-associated acquired immunodeficiency syndrome. Proc Nat Acad Sc. 1986; 83: 2913-7.

30. Weissman I. Estimation of parameters and large quantiles based on the k largest observations. J Am Statist Assoc. 1978; 73: 812-15.

31. Einmahl JHJ, de Haan L, Li D. Weighted approximations of tail copula processes with application to testing the bivariate extreme value condition. Ann Statist. 2006; 34: 1987-2014.

32. Csörgő M, Révész P. Strong approximations in probability and statistics. Probability andMathematical statistics. New York, London: Academic Press [Harcourt Brace Jovanovich, Publishers]; 1981.

33. Shorack GR, Wellner JA. Empirical processes with applications to statistics. New York: Wiley; 1986.

Further reading

34. He S, Yang GL. Estimation of the truncation probability in the random truncation model. Ann Statist. 1998; 26: 1011-27.

35. Hua L, Joe H. Second order regular variation and conditional tail expectation of multiple risks. Insurance Math Econom. 2011; 49: 537-46.

36. Neves C, Fraga Alves MI. Reiss and Thomas' automatic selection of the number of extremes. Comput Statist Data Anal. 2004; 47: 689-704.

0, we haveF¯n*Xn−k:nwF¯*Xn−k:n=OPw−1/⊠+⊠/2,uniformly on w≥1.Proof. Let Vnt⊠n−1∑i=1n1⊠i ...",5,1,1,0>Appendix

Lemma 7.1. *For any small* $\epsilon > 0$*, we have*

$$\frac{\overline{F}_n^*(X_{n-k:n}w)}{\overline{F}^*(X_{n-k:n})} = O_{\mathbf{P}}\left(w^{-1/\gamma+\epsilon/2}\right), \text{ uniformly on } w \geq 1.$$

Proof. Let $V_n(t) := n^{-1}\sum_{i=1}^n \mathbf{1}(\xi_i \leq t)$ be the uniform empirical df pertaining to the sample $\xi_i := \overline{F}^*(X_i)$, $i = 1, \ldots, n$, of independent and identically distributed uniform$(0,1)$ rv's. It is clear that, for an arbitrary x, we have $V_n\left(\overline{F}^*(x)\right) = \overline{F}_n^*(x)$ almost surely. From Assertion 7 in Ref. [33] (page 415), $V_n(t)/t = O_{\mathbf{P}}(1)$ uniformly on $1/n \leq t \leq 1$, this implies that

$$\frac{\overline{F}_n^*(X_{n-k:n}w)}{\overline{F}^*(X_{n-k:n}w)} = O_{\mathbf{P}}(1), \text{ uniformly on } w \geq 1. \tag{7.36}$$

On the other hand, by applying Potter's inequalities (1.4) to $\overline{F}^*$, we get

$$\frac{\overline{F}^*(X_{n-k:n}w)}{\overline{F}^*(X_{n-k:n})} = O_{\mathbf{P}}\left(w^{-1/\gamma+\epsilon/2}\right), \text{ uniformly on } w \geq 1. \tag{7.37}$$

Combining the two statements, (7.36) and (7.37), gives the desired result. □

Corresponding author

Abdelhakim Necir can be contacted at: ah.necir@univ-biskra.dz