



**HAL**  
open science

# Moment estimators of the extreme value index for randomly censored data in the Weibull domain of attraction

Julien Worms, Rym Worms

► **To cite this version:**

Julien Worms, Rym Worms. Moment estimators of the extreme value index for randomly censored data in the Weibull domain of attraction. 2014. hal-01162967

**HAL Id: hal-01162967**

**<https://hal.science/hal-01162967v1>**

Preprint submitted on 11 Jun 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Moment estimators of the extreme value index for randomly censored data in the Weibull domain of attraction

Julien Worms<sup>1</sup> , Rym Worms<sup>2</sup>

21 october 2014

## *Abstract*

This paper addresses the problem of estimating the extreme value index in presence of random censoring for distributions in the Weibull domain of attraction. The methodologies introduced in [Worms (2014)], in the heavy-tailed case, are adapted here to the negative extreme value index framework, leading to the definition of weighted versions of the popular moments of relative excesses with arbitrary exponent . This leads to the definition of two families of estimators (with an adaptation of the so called Moment estimator as a particular case), for which the consistency is proved under a first order condition. Illustration of their performance, issued from an extensive simulation study, are provided.

*Keywords* : Extreme value index, Tail inference , Random censoring , Kaplan-Meier integration

*AMS Classification* : 62G32 (Extreme value statistics) , 62N02 (Estimation for censored data)

---

<sup>1</sup>*Université de Versailles-Saint-Quentin-en-Yvelines, Laboratoire de Mathématiques de Versailles (CNRS UMR 8100), F-78035 Versailles Cedex, France, julien.worms@uvsq.fr*

<sup>2</sup>*Université Paris-Est, Laboratoire d'Analyse et de Mathématiques Appliquées (CNRS UMR 8050), UPEMLV, UPEC, F-94010, Créteil, France, rym.worms@u-pec.fr*

# 1 Introduction

Extreme value statistics is an active domain of research, with numerous fields of application, and which benefits from an important literature in the context of i.i.d. data, dependent data, and (more recently) multivariate or spatial data. By contrast, methodological articles in the case of randomly censored data are quite recent and few : [Einmahl et al. (2008)] presents a general method for adapting estimators of the extreme value index in a censorship framework (a methodology based on a previous work [Beirlant et al. (2007)]), [Diop et al. (2014)] extends the framework to data with covariate information, and [Worms (2014)] proposes a more survival analysis-oriented approach restricted to the heavy tail case. Other existing works on the topic of extremes for censored data are [Brahimi et al. (2013)] and the review paper [Gomes and Neves (2011)].

In this paper, the topic of extreme value statistics for randomly censored data with negative extreme value index is addressed. Our initial purpose was to rely on the ideas of [Worms (2014)] in order to define a more "natural" version (with respect to that proposed in [Einmahl et al. (2008)]) of the moment estimator in the context of censored observations. We finally came out to propose weighted versions of the popular moments of the relative excesses (with arbitrary exponent), and therefore define competitive estimators of the extreme value index in this censoring situation, for distributions in the Weibull maximum domain of attraction.

Let us first define more precisely the framework, the data, and the notations.

In the classical univariate framework of i.i.d. data, a central task is to estimate the extreme value index  $\gamma$ , which captures the main information about the behavior of the tail distribution of the data. More precisely, a distribution function (d.f.)  $F$  is said to be in the maximum domain of attraction of  $H_\gamma$  (noted  $F \in D(H_\gamma)$ ) with

$$H_\gamma(x) := \begin{cases} \exp(-(1 + \gamma x)^{-1/\gamma}) & \text{for } \gamma \neq 0 \text{ and } 1 + \gamma x > 0 \\ \exp(-\exp(-x)) & \text{for } \gamma = 0 \text{ and } x \in \mathbb{R} \end{cases},$$

if there exist two normalizing sequences  $(a_n) \subset \mathbb{R}^+$  and  $(b_n) \subset \mathbb{R}$  such that

$$F^n(a_n x + b_n) \xrightarrow{n \rightarrow \infty} H_\gamma(x) \quad (\forall x \in \mathbb{R}).$$

We consider in this paper two independent i.i.d. non-negative samples  $(X_i)_{i \leq n}$  and  $(C_i)_{i \leq n}$  with respective continuous distribution functions  $F$  and  $G$  (with end-points  $\tau_F$  and  $\tau_G$ , where  $\tau_F := \sup\{x, F(x) < 1\}$ ). In the context of randomly right-censored observations, one only observes, for  $1 \leq i \leq n$ ,

$$Z_i = X_i \wedge C_i \quad \text{and} \quad \delta_i = \mathbb{I}_{X_i \leq C_i}.$$

We denote by  $H$  the distribution function of the  $Z$ -sample, satisfying

$$1 - H = (1 - F)(1 - G)$$

and by  $Z_{1,n} \leq \dots \leq Z_{n,n}$  the associated order statistics. In the whole paper,  $\delta_{1,n}, \dots, \delta_{n,n}$  denote the  $\delta$ 's corresponding to  $Z_{1,n}, \dots, Z_{n,n}$ , respectively.  $F$  and  $G$

are assumed to belong to the maximum domains of attraction  $D(H_{\gamma_X})$  and  $D(H_{\gamma_C})$  respectively, where  $\gamma_X$  and  $\gamma_C$  are real numbers, which implies that  $H \in D(H_\gamma)$ , for some  $\gamma \in \mathbb{R}$ .

Our goal is to estimate the extreme value index  $\gamma_X$  in this context of right censorship. The most interesting cases, described in [Einmahl et al. (2008)], are the following :

$$\begin{array}{lll}
\text{case 1:} & \gamma_X > 0, \gamma_C > 0 & \text{in this case } \gamma = \frac{\gamma_X \gamma_C}{\gamma_X + \gamma_C} \\
\text{case 2:} & \gamma_X < 0, \gamma_C < 0, \tau_F = \tau_G & \text{in this case } \gamma = \frac{\gamma_X \gamma_C}{\gamma_X + \gamma_C} \\
\text{case 3:} & \gamma_X = \gamma_C = 0, \tau_F = \tau_G = +\infty & \text{in this case } \gamma = 0.
\end{array}$$

In [Worms (2014)], case 1 above was considered and an adaptation of the so-called Hill estimator to the right censoring framework was proposed. In this paper, our aim is to consider case 2 above and adapt the approach leading to the so-called Moment Estimator to this censored situation. An adaptation of this estimator was already proposed in [Einmahl et al. (2008)] : it consists in dividing the classical Moment Estimator  $\hat{\gamma}_n^Z$  of  $\gamma$  (calculated from the  $Z$ -sample) by the proportion

$$\hat{p} := k_n^{-1} \sum_{i=1}^{k_n} \delta_{n-i+1,n}$$

of uncensored data in the tail, where  $k_n$  is the number of upper order statistics retained. Note that  $\hat{\gamma}_n^Z$  is an appropriate combination of the following moments

$$\mathcal{M}_{n,k_n}^{(\alpha)} := \frac{1}{k_n} \sum_{i=1}^{k_n} \log^\alpha \left( \frac{Z_{n-i+1,n}}{Z_{n-k_n,n}} \right),$$

for  $\alpha = 1$  or  $2$  (where  $\log^\alpha(x)$  stands for  $(\log(x))^\alpha$ ), and that  $\hat{p}$  estimates the ultimate proportion  $p$  of uncensored observations in the tail, which turns out to be equal to

$$p := \frac{\gamma}{\gamma_X} = \frac{\gamma_C}{\gamma_X + \gamma_C}.$$

Our goal is to show that relying on usual strategies in the survival analysis literature leads to estimators of  $\gamma_X$  which are often sharper than those obtained by simply dividing an estimator of  $\gamma$  by the proportion of uncensored observations. By “usual” strategy we mean using “Kaplan-Meier”-like random weights : we refer to [Worms (2014)] for more detailed informations concerning the origin of the two kinds of random weights appearing in the formulas below. As a matter of fact, we define, for any given  $\alpha \geq 1$ , the following two versions of randomly weighted moments of the log relative excesses :

$$M_{n,k_n}^{(\alpha)} := \frac{1}{n(1 - \hat{F}_n(Z_{n-k_n,n}))} \sum_{i=1}^{k_n} \frac{\delta_{n-i+1,n}}{1 - \hat{G}_n(Z_{n-i+1,n}^-)} \left( \log^\alpha \left( \frac{Z_{n-i+1,n}}{Z_{n-k_n,n}} \right) \right) \quad (1)$$

and

$$\widetilde{M}_{n,k_n}^{(\alpha)} := \frac{1}{n(1 - \widehat{F}_n(Z_{n-k_n,n}))} \sum_{i=1}^{k_n} \frac{1}{1 - \widehat{G}_n(Z_{n-i+1,n}^-)} \xi_{i,n} \quad (2)$$

where

$$\xi_{i,n} := i \left( \log^\alpha \left( \frac{Z_{n-i+1,n}}{Z_{n-k_n,n}} \right) - \log^\alpha \left( \frac{Z_{n-i,n}}{Z_{n-k_n,n}} \right) \right) \quad (3)$$

and  $(k_n)$  is a sequence of integers satisfying, as  $n$  tends to  $+\infty$ ,

$$k_n \rightarrow +\infty \quad \text{and} \quad k_n = o(n). \quad (4)$$

Above,  $\widehat{F}_n$  and  $\widehat{G}_n$  naturally denote the Kaplan-Meier estimators of  $F$  and  $G$ , respectively, defined as follows : for  $t < Z_{n,n}$ ,

$$1 - \widehat{F}_n(t) = \prod_{Z_{i,n} \leq t} \left( \frac{n-i}{n-i+1} \right)^{\delta_{i,n}} \quad \text{and} \quad 1 - \widehat{G}_n(t) = \prod_{Z_{i,n} \leq t} \left( \frac{n-i}{n-i+1} \right)^{1-\delta_{i,n}}.$$

It should be noted that these 2 weighted versions of the moments of the log-excesses defined in (1) and (2) are in fact closely related : as a matter of fact, they differ only when the maximum observation  $Z_{n,n}$  is censored (when  $\delta_{n,n} = 1$ , we have indeed  $M_{n,k_n}^{(\alpha)} = \widetilde{M}_{n,k_n}^{(\alpha)}$ , see Proposition 1 in Section 5). However, both versions deserve attention : firstly because in practice the last observation is often a censored one, and secondly because when they do differ, the difference is the only term involving the information contained in the maximum observation  $Z_{n,n}$  (this difference is therefore non-asymptotically not negligible, although it tends to 0 in probability, as stated in Proposition 1 in Section 5).

In section 2 below, assumptions are presented and discussed, convergence results for the weighted moments  $M_{n,k_n}^{(\alpha)}$  and  $\widetilde{M}_{n,k_n}^{(\alpha)}$  are stated, and we describe how classes of estimators of  $\gamma_X$  can be deduced by combining these moments for different values of  $\alpha$ . In Section 3, performance of these estimators will be presented on the basis of simulations. Section 4 provides some words of conclusion, Section 5 is devoted to the proof of Theorem 1 below, and finally the Appendix includes standard (but central to our proofs) results on regularly varying functions, as well as the proofs of the different lemmas which were used in Section 5.

## 2 Results

### 2.1 Assumptions

In addition to (4), our results need the following minimal assumption :

$$(A) \quad F \in D(H_{\gamma_X}), \quad G \in D(H_{\gamma_C}) \quad \text{with} \quad \gamma_X < 0, \gamma_C < 0 \quad \text{and} \quad x^* := \tau_F = \tau_G.$$

As noted earlier, this assumption implies that  $H \in D(H_\gamma)$  with  $\tau_H = x^*$  and

$$\gamma = \frac{\gamma_X \gamma_C}{\gamma_X + \gamma_C} < 0.$$

If we note  $U(t) = H^{\leftarrow}(1 - 1/t)$  the quantile function associated to  $H$ , then  $x^* = U(\infty)$  and  $H \in D(H_\gamma)$  is equivalent to the existence of some positive function  $a$  such that

$$\lim_{t \rightarrow +\infty} \frac{\log U(tx) - \log U(t)}{a(t)/U(t)} = \frac{x^\gamma - 1}{\gamma}, \quad \forall x > 0, \quad (5)$$

which, since  $\gamma < 0$ , is itself equivalent to

$$\lim_{t \rightarrow +\infty} \frac{U(\infty) - U(tx)}{U(\infty) - U(t)} = x^\gamma, \quad \forall x > 0. \quad (6)$$

This means that the function  $U(\infty) - U$  is regularly varying (at  $+\infty$ ) with index  $\gamma$  (see the appendix for the definition of regular variation at  $+\infty$ ). A reference for the equivalence of conditions (5) and (6) to (A) is [Haan and Ferreira (2006)] (respectively relation (3.5.4) and Corollary 1.2.10 there).

Finally, we will need some very mild additional assumption on  $(k_n)$

(K) there exists some  $\delta > 0$ , or some  $\delta \geq \frac{\gamma_X - \gamma_C}{\gamma_X + \gamma_C}$  if  $\gamma_C \geq \gamma_X$ ,  
such that

$$-\log(k_n/n)/k_n = O(n^{-\delta}). \quad (7)$$

## 2.2 Asymptotic results

Let us introduce the notation  $a_{n,k} := a(n/k_n)/U(n/k_n)$  (see the previous paragraph for the definition of functions  $U$  and  $a$ ), where  $a_{n,k} \rightarrow 0$  (cf equation (3.5.5) in [Haan and Ferreira (2006)]). In the paper,  $Beta(\cdot, \cdot)$  denotes the usual Beta function,  $Beta(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt$  ( $a > 0, b > 0$ ).

**Theorem 1** *Under assumption (A) and conditions (4) and (K), for any given  $\alpha \geq 1$ , both  $\frac{M_{n,k_n}^{(\alpha)}}{(a_{n,k})^\alpha}$  and  $\frac{\widetilde{M}_{n,k_n}^{(\alpha)}}{(\widetilde{a}_{n,k})^\alpha}$  converge in probability, as  $n$  tends to  $\infty$ , to*

$$|\gamma_X|^{-1} |\gamma|^{-\alpha} Beta(|\gamma_X|^{-1}; \alpha + 1).$$

The following corollary states the consistency of our two different adaptations of the Moment estimator to this censored framework.

**Corollary 1** *Under conditions of Theorem 1, as  $n \rightarrow \infty$ ,*

$$\widehat{\gamma}_{n,Mom} := M_{n,k_n}^{(1)} + 1 - \frac{1}{2} \left( 1 - \frac{(M_{n,k_n}^{(1)})^2}{M_{n,k_n}^{(2)}} \right)^{-1} \xrightarrow{\mathbb{P}} \gamma_X$$

and

$$\widetilde{\gamma}_{n,Mom} := \widetilde{M}_{n,k_n}^{(1)} + 1 - \frac{1}{2} \left( 1 - \frac{(\widetilde{M}_{n,k_n}^{(1)})^2}{\widetilde{M}_{n,k_n}^{(2)}} \right)^{-1} \xrightarrow{\mathbb{P}} \gamma_X.$$

In fact, by using the elementary properties of the Beta function, the weighted moments  $M_n^{(\alpha)}$  or  $\widetilde{M}_n^{(\alpha)}$  can be combined in different ways, leading to the definition of two different classes of consistent estimators of  $\gamma_X$ , parametrized by  $\alpha \geq 1$  (proofs of the 3 corollaries are easy and omitted). In the next section, we study their finite sample performance.

**Corollary 2** *Under conditions of Theorem 1, as  $n \rightarrow \infty$ ,*

$$\widehat{\gamma}_{n,1}^{(\alpha)} := (V_{n,\alpha}^{-1} + \alpha + 1)^{-1} \xrightarrow{\mathbb{P}} \gamma_X$$

and

$$\widetilde{\gamma}_{n,1}^{(\alpha)} := (\widetilde{V}_{n,\alpha}^{-1} + \alpha + 1)^{-1} \xrightarrow{\mathbb{P}} \gamma_X$$

where

$$V_{n,\alpha} := 1 - \frac{\alpha + 2}{\alpha + 1} \frac{(M_n^{(\alpha+1)})^2}{M_n^{(\alpha)} M_n^{(\alpha+2)}} \quad \text{and} \quad \widetilde{V}_{n,\alpha} := 1 - \frac{\alpha + 2}{\alpha + 1} \frac{(\widetilde{M}_n^{(\alpha+1)})^2}{\widetilde{M}_n^{(\alpha)} \widetilde{M}_n^{(\alpha+2)}}.$$

**Corollary 3** *Under conditions of Theorem 1, as  $n \rightarrow \infty$ ,*

$$\widehat{\gamma}_{n,2}^{(\alpha)} := \frac{1 - (\alpha + 1)R_{n,\alpha}}{(\alpha + 1)(1 - R_{n,\alpha})} \xrightarrow{\mathbb{P}} \gamma_X$$

and

$$\widetilde{\gamma}_{n,2}^{(\alpha)} := \frac{1 - (\alpha + 1)\widetilde{R}_{n,\alpha}}{(\alpha + 1)(1 - \widetilde{R}_{n,\alpha})} \xrightarrow{\mathbb{P}} \gamma_X,$$

where

$$R_{n,\alpha} := \frac{M_n^{(1)} M_n^{(\alpha)}}{M_n^{(\alpha+1)}} \quad \text{and} \quad \widetilde{R}_{n,\alpha} := \frac{\widetilde{M}_n^{(1)} \widetilde{M}_n^{(\alpha)}}{\widetilde{M}_n^{(\alpha+1)}}.$$

**Remark 1** *It is straightforward to see that  $\widehat{\gamma}_{n,2}^{(\alpha)}$  with  $\alpha = 1$  equals  $1 - \frac{1}{2}(1 - R_{n,1})^{-1}$ , which is very close to  $\widehat{\gamma}_{n,Mom}$ , since  $M_{n,k_n}^{(1)} \rightarrow 0$  in our finite endpoint framework.*

**Remark 2** *If  $\mathcal{M}_{n,k_n}^{(\alpha)}$  denotes the unweighted moments defined in the introduction, it can be proved that under (A) and (4), for  $\alpha \geq 1$ ,*

$$\frac{\mathcal{M}_{n,k_n}^{(\alpha)}}{(a_{n,k})^\alpha} \xrightarrow{\mathbb{P}} |\gamma|^{-\alpha-1} \text{Beta}(|\gamma|^{-1}; \alpha + 1).$$

Therefore, it is easy to check that combining those moments as described in Corollaries 2 and 3 leads to consistent estimators of  $\gamma$ , and thus dividing the latter by  $\hat{p}$  (defined in the introduction) leads to 2 classes of consistent estimators  $\check{\gamma}_{n,1}^{(\alpha)}$  and  $\check{\gamma}_{n,2}^{(\alpha)}$  of  $\gamma_X$ . We also define  $\check{\gamma}_{n,Mom}$  as the estimator of  $\gamma_X$  obtained by dividing the classical Moment estimator of  $\gamma$  by the proportion  $\hat{p}$ . A finite-sample comparison of those estimators with our new competitors is presented in the following section.

**Remark 3** Note that the combination of moments proposed in Corollaries 2 and 3 become inadequate in the framework of a positive extreme value index: it can be indeed proved that, in this framework, the combinations  $\hat{\gamma}_{n,j}^{(\alpha)}$  and  $\tilde{\gamma}_{n,j}^{(\alpha)}$ , for  $j = 1$  or  $2$ , converge in probability to zero, by proving that  $M_{n,k_n}^{(\alpha)}$  and  $\widetilde{M}_{n,k_n}^{(\alpha)}$  converge in probability to  $\gamma_X^\alpha \Gamma(\alpha + 1)$  (in the complete data case, this result is known for  $\mathcal{M}_{n,k_n}^{(\alpha)}$ , see [Segers (2001)]). This could suggest, in the positive index case, the definition of estimators of  $\gamma_X$  which would be equal to  $\hat{\gamma}_{n,j}^{(\alpha)}$  or  $\tilde{\gamma}_{n,j}^{(\alpha)}$  (for  $j = 1$  or  $2$ ) plus a "censored version" of the Hill estimator (in the same spirit as the definition of the Moment estimator, which equals the Hill estimator plus a term converging to 0 in the positive index case).

### 3 Finite sample behavior

The goal of this Section is to present our results concerning the finite sample performances of our new estimators of the extreme value index in presence of random censorship, presented in Corollaries 1, 2 and 3. In each case considered, 2000 random samples of size  $n = 500$  were generated, and the median bias and mean squared error (MSE) of the different estimators of  $\gamma$  were plotted against the number  $k_n$  of excesses used.

A great variety of situations can be (and has been) considered in our simulation study : various values of  $\gamma_X$  and  $\gamma_C$  (and therefore various censoring rates in the tail), various families of underlying distributions (Reverse Burr, generalized Pareto, Beta), and choice of the value of  $\alpha$ . It is impossible to illustrate here the different possible combinations of these features : we will therefore try to draw some general conclusions from the many different situations we have observed, and provide a partial illustration with 3 particular cases.

Concerning the choice of the tuning parameter  $\alpha$ , we did not find a value which seemed preferable in every situation : nonetheless, in general, for small values of  $k_n$ , a value of  $\alpha$  around 1 or 2 yields better MSE, whereas for high values of  $k_n$ , the MSE is lower for values of  $\alpha$  greater than 2. We decided not to include this preliminary study in this article, and chose (almost arbitrarily) the value  $\alpha = 2$  in all our subsequent simulations.

Let us now settle the vocabulary used in this section. We will call *Moment estimators* the estimators  $\hat{\gamma}_{n,Mom}$  and  $\tilde{\gamma}_{n,Mom}$  appearing in Corollary 1, as well as the estimator  $\check{\gamma}_{n,Mom}$  introduced in Remark 2 above. We will call *type 1 (resp. type 2) estimators* the estimators  $\hat{\gamma}_{n,1}^{(\alpha)}$  and  $\tilde{\gamma}_{n,1}^{(\alpha)}$  (resp.  $\hat{\gamma}_{n,2}^{(\alpha)}$  and  $\tilde{\gamma}_{n,2}^{(\alpha)}$ ) appearing in Corollary 2 (resp. 3), as well as the estimator  $\check{\gamma}_{n,1}^{(\alpha)}$  (resp.  $\check{\gamma}_{n,2}^{(\alpha)}$ ) introduced in Remark 2.

We will also consider names for the different methods : the KM method (for Kaplan-Meier-like weights, appearing in the definition of  $M_{n,k_n}^{(\alpha)}$ ), leading to  $\hat{\gamma}$  estimators, the L method (for Leurgans-like weights) leading to  $\tilde{\gamma}$  estimators (the name comes from the mathematician Sue Leurgans who inspired the weights, see [Worms (2014)] for details and a reference), and the EFG method (for constant



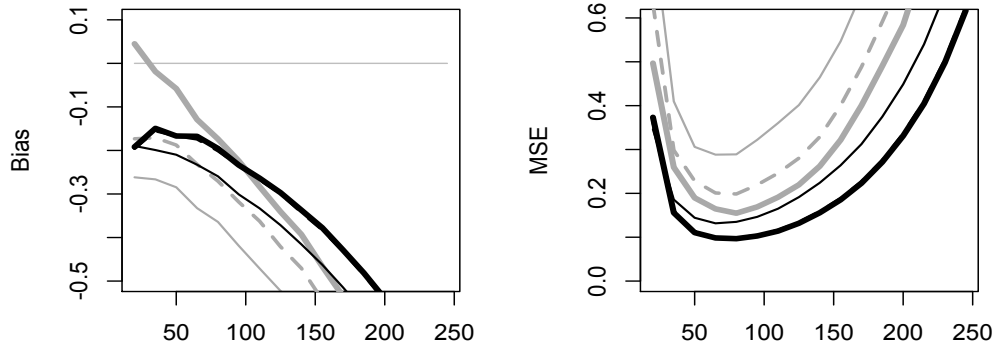


Figure 1: Comparison between  $\hat{\gamma}_{n,1}^{(2)}$  (thick black),  $\tilde{\gamma}_{n,1}^{(2)}$  (dashed black),  $\check{\gamma}_{n,1}^{(2)}$  (thin black),  $\hat{\gamma}_{n,Mom}$  (thick grey),  $\tilde{\gamma}_{n,Mom}$  (dashed grey) and  $\check{\gamma}_{n,Mom}$  (thin grey) for a RevBurr(1, 1, 1, 10) censored by a RevBurr(10, 2/3, 1, 10) ( $\gamma_X = -1 > \gamma_C = -3/2$ ,  $p=2/5$ , weak censoring)

weighting by  $\hat{p}$ ), leading to  $\check{\gamma}$  estimators (the names comes from the initials of the authors of [Einmahl et al. (2008)]).

There are two main questions addressed in this empirical study : is one of the 3 methods preferable to the others (and in which conditions) and is there a better choice for the type of estimator (type 1 , type 2, or classical Moment estimator) ? Unsurprisingly, after our intensive simulation study, we may say that the answer is no for the 2 questions, if an overall superiority is looked for. However, we can make some partial comments concerning the choice of the method and of the estimator type, whether the censoring is strong or weak, or the value of  $|\gamma|$  is small or not.

Note first that, if the censoring rate  $1 - p$  in the tail is very low (say lower than 10%), we observed that there was not much difference between the 3 methods (KM, L, EFG), and that it was just a question of choosing between type 1, type 2, and moment estimator. This is why, in the following, we only consider cases where the censoring rate  $1 - p = \frac{\gamma_X}{\gamma_X + \gamma_C}$  is larger than 1/4, and talk about strong censoring in the tail when this rate is greater than 1/2 (*i.e.*  $\gamma_X \leq \gamma_C$ ), and weak censoring otherwise (when  $\gamma_X > \gamma_C$ ).

For “high” values of  $\gamma_X$ , *i.e.* lower than  $-1/2$ , we have most of the time observed better performance of the KM and L methods with respect to the EFG method, in strong or weak censoring frameworks. In this context, the type 1 estimators are generally preferable to the type 2 estimators, and comparable or preferable to the moment estimator.

For values of  $\gamma_X$  between  $-1/2$  and 0 (sometimes called the “regular” case, and which is the most frequently encountered in practice), there exists a great variety

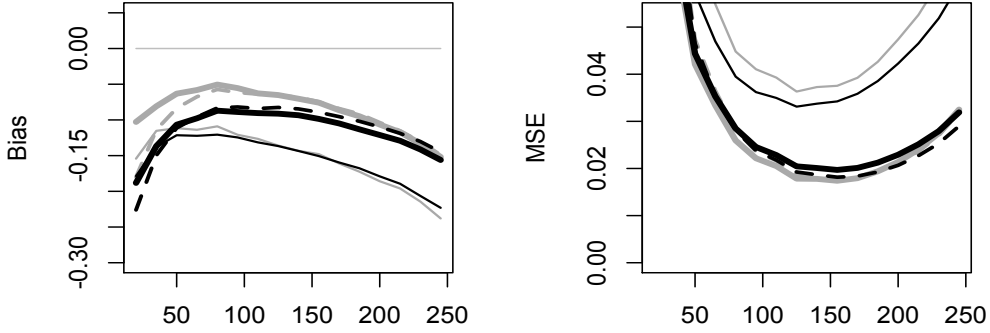


Figure 2: Comparison between  $\hat{\gamma}_{n,2}^{(2)}$  (thick black),  $\tilde{\gamma}_{n,2}^{(2)}$  (dashed black),  $\check{\gamma}_{n,2}^{(2)}$  (thin black),  $\hat{\gamma}_{n,Mom}$  (thick grey),  $\tilde{\gamma}_{n,Mom}$  (dashed grey) and  $\check{\gamma}_{n,Mom}$  (thin grey) for a  $RevBurr(1, 8, 1/2, 10)$  censored by a  $RevBurr(10, 4, 1/2, 10)$  ( $\gamma_X = -1/4 > \gamma_C = -1/2$ ,  $p=1/3$ , weak censoring)

of situations. We observed that the moment estimators were generally better than the type 2 estimators, which were themselves generally better than the type 1 ones. Concerning the choice of the method, for the moment estimator, it seems difficult to suggest a particular one, between the KM, L, and EFG methods (even though in many cases, at least one among the KM and L methods was better than the EFG method). Concerning the inferiority of types 1 and 2 versus the moment estimator, it should be noted that it is mainly due to the bias, which contributes the most to the MSE (in fact, we clearly noticed that the variances of the types 1 and 2, for  $\alpha = 2$ , are almost always lower than the variance of the moment estimator).

The 3 particular situations we chose as illustrations of the comments above involve the Reverse Burr class of distributions  $RevBurr(\beta, \tau, \lambda, x^*)$  (with  $\beta, \tau, \lambda > 0$ ) : its survival function is

$$\mathbb{P}(X > x) = (1 + \beta^{-1}(x^* - x)^{-\tau})^{-\lambda},$$

and its extreme value index is  $-1/(\lambda\tau)$ .

In Figure 1, the value of  $\gamma_X$  is lower than  $-1/2$ , and therefore, as motivated above, for readability purposes we only kept the type 1 estimators on the graph, whereas for the other two figures, the value of  $\gamma_X$  is between  $-1/2$  and 0 and we therefore only kept the type 2 estimator illustrated. Remind here that these 3 examples are only 3 particular cases of the numerous combinations of features we have considered in our simulation study.

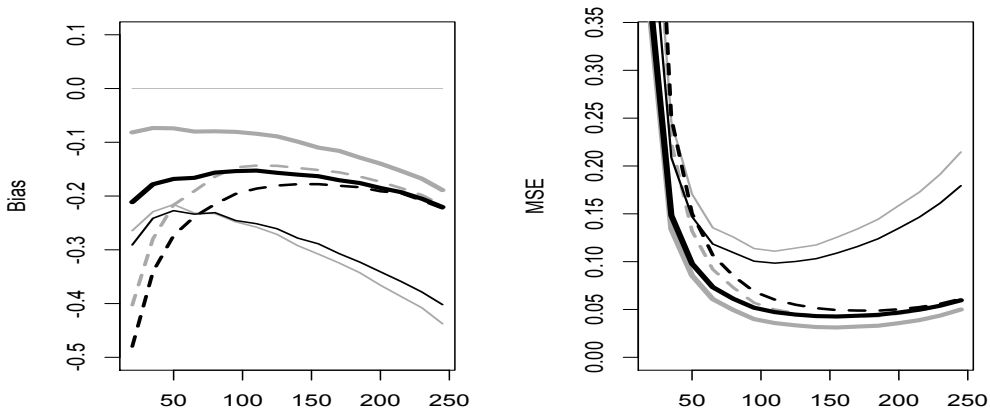


Figure 3: Comparison between  $\hat{\gamma}_{n,2}^{(2)}$  (thick black),  $\tilde{\gamma}_{n,2}^{(2)}$  (dashed black),  $\check{\gamma}_{n,2}^{(2)}$  (thin black),  $\hat{\gamma}_{n,Mom}$  (thick grey),  $\tilde{\gamma}_{n,Mom}$  (dashed grey) and  $\check{\gamma}_{n,Mom}$  (thin grey) for a RevBurr(10, 8, 1/2, 10) censored by a RevBurr(10, 5, 1, 10) ( $\gamma_X = -1/4 < \gamma_C = -1/5$ ,  $p=5/9$ , strong censoring)

## 4 Conclusion

In this paper, we applied the methodology introduced in [Worms (2014)] to define weighted versions of the moments of relative excesses, and consequently construct new estimators of the extreme value index for randomly-censored data with distributions in the Weibull domain of attraction. We proposed, in particular, a new adaptation of the famous Moment estimator. Our intensive simulation study shows that the proposed estimators are competitive even if, in many cases, the bias would need to be reduced. A future possible work would be to exploit our weighting methodology in order to estimate other parameters of the tail (for reducing the bias, for example) as well as extreme quantiles. The asymptotic normality remains a question to be addressed (difficulties come from the control of the Kaplan-Meier estimates in the tail).

## 5 Proof of Theorem 1

Before proceeding to the proof of Theorem 1, we state the following Proposition which explains the link between our two proposals of weighted moments.

**Proposition 1** (i) For any  $\alpha \geq 1$ ,  $\widetilde{M}_{n,k_n}^{(\alpha)} = M_{n,k_n}^{(\alpha)} + (1 - \delta_{n,n})D_{n,k_n}^{(\alpha)}$ , where

$$D_{n,k_n}^{(\alpha)} = \frac{1}{n(1 - \widehat{F}_n(Z_{n-k_n,n}))(1 - \widehat{G}_n(Z_{n,n}^-))} \log^\alpha \left( \frac{Z_{n,n}}{Z_{n-k_n,n}} \right).$$

(ii) Under the same assumptions as Theorem 1, for any  $\alpha \geq 1$ , we have  $D_{n,k_n}^{(\alpha)} = o_{\mathbb{P}}(a_{n,k}^\alpha)$ .

According to this proposition, the validity of Theorem 1 for  $M_{n,k_n}^{(\alpha)}$  is a consequence of Theorem 1 for  $\widetilde{M}_{n,k_n}^{(\alpha)}$ . Proposition 1 will be proved at the end of the appendix.

We now need to state the following technical Lemmas, which will be proved in the appendix. The notation  $\xi_{i,n}$  was introduced in (3) and we note

$$\widetilde{Z}_{i,n} := \frac{x^* - Z_{n-i+1,n}}{x^* - Z_{n-k_n,n}}.$$

**Lemma 1** Let  $\alpha \geq 1$  and  $u_i = u_{i,n} = \frac{i}{k_n+1}$ . Under assumptions (A) and (4):

(i) if  $0 < a < 1$  then

$$\frac{1}{k_n} \sum_{i=1}^{k_n} u_i^{-a} \frac{\xi_{i,n}}{a_{n,k}^\alpha} \xrightarrow{\mathbb{P}} (1-a)|\gamma|^{-\alpha-1} \text{Beta} \left( \frac{1-a}{|\gamma|}; \alpha+1 \right),$$

(ii) if  $a > 1$ , then for any  $\delta' > 0$ ,

$$\frac{1}{k_n^{a+\delta'}} \sum_{i=1}^{k_n} u_i^{-a} \frac{\xi_{i,n}}{a_{n,k}^\alpha} \xrightarrow{\mathbb{P}} 0.$$

**Lemma 2** For any given positive exponents  $\theta$  and  $\theta' > 0$ , there exist constants  $c > 1$ ,  $c' < 1$  both arbitrarily close to 1, and  $a_+ > 0$ ,  $a_- > 0$ , arbitrarily close to  $\gamma\theta$  and to  $\gamma\theta'$  respectively, such that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( \max_{i \leq k_n} \frac{\widetilde{Z}_{i,n}^\theta}{c u_i^{-a_+}} > 1 \right) = \lim_{n \rightarrow \infty} \mathbb{P} \left( \min_{i \leq k_n} \frac{\widetilde{Z}_{i,n}^{\theta'}}{c' u_i^{-a_-}} < 1 \right) = 0.$$

We now proceed to the proof of Theorem 1 (for  $\widetilde{M}_{n,k_n}^{(\alpha)}$ ), which has structural similarities with the proof of Theorem 2 in [Worms (2014)]. We shall refer to the latter when necessary. We have the decomposition

$$\widetilde{M}_{n,k_n}^{(\alpha)} = A_n(\overline{W}_n + R_n),$$

where

$$\begin{aligned} A_n &:= \frac{1 - \hat{G}_n(Z_{n-k_n, n})}{1 - G(Z_{n-k_n, n})} & , & \quad \bar{W}_n := \frac{1}{k_n} \sum_{i=1}^{k_n} W_{i, n}, \\ R_n &:= \frac{1}{k_n} \sum_{i=1}^{k_n} (C_{i, n} - 1) W_{i, n} & , & \quad C_{i, n} := \frac{1 - G(Z_{n-i+1, n})}{1 - \hat{G}_n(Z_{n-i+1, n})}, \end{aligned}$$

and

$$W_{i, n} := \frac{1 - G(Z_{n-k_n, n})}{1 - G(Z_{n-i+1, n})} \xi_{i, n}.$$

Since  $A_n \xrightarrow{\mathbb{P}} 1$  as  $n \rightarrow \infty$  (see Theorem 2 in [Csörgö (1996)]), we need to prove that  $R_n = o_{\mathbb{P}}(a_{n, k}^\alpha)$  and

$$\frac{\bar{W}_n}{a_{n, k}^\alpha} \xrightarrow{\mathbb{P}} l_\alpha, \quad (8)$$

where  $l_\alpha$  denote the limit in the statement of Theorem 1.

### 5.1 Proof of $\bar{W}_n/a_{n, k}^\alpha \xrightarrow{\mathbb{P}} l_\alpha$

Since  $G \in D(H_{\gamma_C})$ , with  $\gamma_C < 0$ , is equivalent to  $t \rightarrow 1 - G(x^* - t)$  being regularly varying at 0 with index  $-1/\gamma_C$  (see the appendix for the definition of regular variation at 0), the bounds (12) in Corollary 4 (in the appendix) applied to  $f = G$ ,  $x = (x^* - Z_{n-i+1, n})/(x^* - Z_{n-k_n, n})$  and  $t = x^* - Z_{n-k_n, n}$ , yield, for  $\epsilon > 0$ ,  $n$  sufficiently large and every  $1 \leq i \leq k_n$ ,

$$(1 - \epsilon) \xi_{i, n} \tilde{Z}_{i, n}^{\gamma_C^{-1} + \epsilon} \leq W_{i, n} \leq (1 + \epsilon) \xi_{i, n} \tilde{Z}_{i, n}^{\gamma_C^{-1} - \epsilon}. \quad (9)$$

Let  $\eta > 0$ . We first write, for  $\epsilon$  sufficiently small,

$$\mathbb{P}(\bar{W}_n/a_{n, k}^\alpha - l_\alpha > \eta) \leq \mathbb{P}(k_n^{-1} \sum_{i=1}^{k_n} \tilde{Z}_{i, n}^{\gamma_C^{-1} - \epsilon} \frac{\xi_{i, n}}{a_{n, k}^\alpha} - l_\alpha > \frac{\eta}{2}),$$

$$\mathbb{P}(l_\alpha - \bar{W}_n/a_{n, k}^\alpha > \eta) \leq \mathbb{P}(l_\alpha - k_n^{-1} \sum_{i=1}^{k_n} \tilde{Z}_{i, n}^{\gamma_C^{-1} + \epsilon} \frac{\xi_{i, n}}{a_{n, k}^\alpha} > \frac{\eta}{2}).$$

Let us now consider constants  $c > 1$  and  $c' < 1$ , both arbitrary close to 1, and  $a_+ > 0$  and  $a_- > 0$  both arbitrary close to  $\gamma/\gamma_C$ . These constants come from the application of Lemma 2 above with  $\theta = \gamma_C^{-1} - \epsilon$  and  $\theta' = \gamma_C^{-1} + \epsilon$ . Using positivity of  $\xi_{i, n}$ , it comes

$$\begin{aligned} \mathbb{P}\left(\frac{\bar{W}_n}{a_{n, k}^\alpha} - l_\alpha > \eta\right) &\leq \mathbb{P}\left(\max_{i \leq k_n} \frac{\tilde{Z}_{i, n}^{\gamma_C^{-1} - \epsilon}}{c u_i^{-a_+}} > 1\right) \\ &\quad + \mathbb{P}\left(c k_n^{-1} \sum_{i=1}^{k_n} u_i^{-a_+} \frac{\xi_{i, n}}{a_{n, k}^\alpha} - l_\alpha > \frac{\eta}{2}\right) \\ \mathbb{P}\left(l_\alpha - \frac{\bar{W}_n}{a_{n, k}^\alpha} > \eta\right) &\leq \mathbb{P}\left(\min_{i \leq k_n} \frac{\tilde{Z}_{i, n}^{\gamma_C^{-1} + \epsilon}}{c' u_i^{-a_-}} < 1\right) \\ &\quad + \mathbb{P}\left(l_\alpha - c' k_n^{-1} \sum_{i=1}^{k_n} u_i^{-a_-} \frac{\xi_{i, n}}{a_{n, k}^\alpha} > \frac{\eta}{2}\right) \end{aligned}$$

where  $u_i = u_{i,n} = i/(k_n + 1)$  for  $1 \leq i \leq k_n$ . If we call  $l_{\alpha,a}$  the limit in the statement of Lemma 1 (i), and if we apply Lemma 2 as indicated previously, we have

$$\begin{aligned} \limsup_{n \rightarrow \infty} \mathbb{P} \left( \frac{\bar{W}_n}{a_{n,k}^\alpha} - l_\alpha > \eta \right) &\leq \limsup_{n \rightarrow \infty} \mathbb{P} \left( k_n^{-1} \sum_{i=1}^{k_n} u_i^{-a_+} \frac{\xi_{in}}{a_{n,k}^\alpha} - l_{\alpha,a_+} > \frac{1}{c} \left( \frac{\eta}{2} + l_\alpha \right) - l_{\alpha,a_+} \right) \\ \limsup_{n \rightarrow \infty} \mathbb{P} \left( l_\alpha - \frac{\bar{W}_n}{a_{n,k}^\alpha} > \eta \right) &\leq \limsup_{n \rightarrow \infty} \mathbb{P} \left( l_{\alpha,a_-} - k_n^{-1} \sum_{i=1}^{k_n} u_i^{-a_-} \frac{\xi_{in}}{a_{n,k}^\alpha} > \frac{1}{c'} \left( \frac{\eta}{2} - l_\alpha \right) + l_{\alpha,a_-} \right) \end{aligned}$$

Since  $l_\alpha = l_{\alpha,\gamma/\gamma_C}$ , and both  $a_+$  and  $a_-$  are arbitrary close to  $\gamma/\gamma_C < 1$ , it is easy to see that (8) comes from the application of Lemma 1 (i) to  $a = a_+$  and  $a = a_-$ .

## 5.2 Proof of $R_n = o_{\mathbb{P}}(a_{n,k}^\alpha)$

Let us use the same decomposition as in the proof of the negligibility of the term  $R_n$  in [Worms (2014)] (see subsection 5.1.2 there). In other words, we define, for some  $\delta' > 0$ ,

$$\tilde{C}(t) := \int_0^t \frac{dG(x)}{(1-G(x))^2(1-F(x))} \quad \text{and} \quad h_{i,n} := (\tilde{C}(Z_{n-i+1,n}))^{-\frac{1}{2}-\delta'},$$

and we readily have  $|R_n| \leq T_n^1 T_n^2$ , where

$$T_n^1 := \sup_{1 \leq i \leq k_n} \sqrt{n} |h_{i,n}(C_{i,n} - 1)| \quad \text{and} \quad T_n^2 := \frac{1}{k_n} \sum_{i=1}^{k_n} W_{i,n} h_{i,n}^{-1} n^{-\frac{1}{2}}.$$

Using sharp results of the survival analysis literature, we have already proved in [Worms (2014)] that  $T_n^1 = O_{\mathbb{P}}(1)$ . It remains to prove that

$$T_n^2/a_{n,k}^\alpha = o_{\mathbb{P}}(1).$$

First, from the definition of  $h_{in}$  and  $\tilde{C}$ , since  $(1-H) = (1-F)(1-G)$  we clearly have

$$h_{in}^{-1} < \left( \frac{-\log(1-G(Z_{n-i+1,n}))}{1-H(Z_{n-i+1,n})} \right)^{\frac{1}{2}+\delta'}.$$

Moreover, under assumption (A),  $1-H(x^*-\cdot)$  is regularly varying at zero with index  $-1/\gamma$  and  $-\log(1-G(x^*-\cdot))$  is slowly varying at 0 : therefore, the application of (9), as well as bound (11) to  $-\log(1-G(x^*-\cdot))$  and bound (12) to  $f = G$  and  $f = H$ , implies that , for  $n$  sufficiently larger,  $T_n^2 \leq 4P_n Q_n$  where

$$\begin{aligned} P_n &:= n^{-\frac{1}{2}} \left( \frac{-\log(1-G(Z_{n-k_n,n}))}{1-H(Z_{n-k_n,n})} \right)^{\frac{1}{2}+\delta'} \\ Q_n &:= \frac{1}{k_n} \sum_{i=1}^{k_n} \xi_{in} \tilde{Z}_{i,n}^\beta, \end{aligned}$$

where  $\beta = (2\gamma)^{-1} + \gamma_C^{-1} - \epsilon'$ , for some  $\epsilon' > 0$ .

We thus need to prove that  $P_n Q_n / a_{n,k}^\alpha = o_{\mathbb{P}}(1)$ .

Let  $\eta > 0$  and consider constants  $c > 1$  arbitrarily close to 1 and  $a_+ > 0$  arbitrarily close to  $\gamma\beta = \frac{1}{2} + \frac{\gamma}{\gamma_C} - \gamma\epsilon' > \frac{1}{2} + \frac{\gamma}{\gamma_C}$ . We have,

$$\mathbb{P} \left( \frac{P_n Q_n}{a_{n,k}^\alpha} > \eta \right) \leq \mathbb{P} \left( \max_{i \leq k_n} \frac{\tilde{Z}_{i,n}^\beta}{c u_i^{-a_+}} > 1 \right) + \mathbb{P} \left( P_n k_n^{-1} \sum_{i=1}^{k_n} u_i^{-a_+} \frac{\xi_{in}}{a_{n,k}^\alpha} > \frac{\eta}{c} \right). \quad (10)$$

First, Lemma 2 is applied with  $\theta = \beta$  and thus the first term of the right-hand side of (10) tends to 0. Next, Lemma 1 is applied with  $a = a_+$ : we thus need to distinguish the case  $\gamma_C < \gamma_X$  (for which  $\gamma\beta < 1$ ) from the case  $\gamma_C \geq \gamma_X$  (for which  $\gamma\beta > 1$  when  $\epsilon' > 0$  gets small).

(i) Case  $\gamma_C < \gamma_X$

First of all, assumption (K) implies that  $P_n = o_{\mathbb{P}}(1)$  (see relation (20) in [Worms (2014)]). Since  $a_+ < 1$  in this case, Lemma 1 (ii) implies that  $\frac{1}{k_n} \sum_{i=1}^{k_n} u_i^{-a_+} \frac{\xi_{in}}{a_{n,k}^\alpha} = O_{\mathbb{P}}(1)$  and consequently the second term of the right hand-side of (10) tends to 0.

(ii) Case  $\gamma_C \geq \gamma_X$

In this case  $a_+ > 1$ , therefore Lemma 1 (ii) implies that, for any given  $\delta' > 0$ ,  $k_n^{-(a_+\delta')} \sum_{i=1}^{k_n} u_i^{-a_+} \frac{\xi_{in}}{a_{n,k}^\alpha} = o_{\mathbb{P}}(1)$ . Moreover, assumption (K) implies that, for  $\delta' > 0$  small enough and  $a_+$  sufficiently close to  $1/2 + \gamma/\gamma_C$ , we have  $k_n^{a_+\delta'-1} P_n = O_{\mathbb{P}}(1)$ . Hence, the second term of the right hand-side of (10) also tends to 0.  $\diamond$

## 6 Appendix

### 6.1 Regular variation and Potter-type bounds

**Definition 1** *An ultimately positive function  $f : \mathbb{R}^+ \rightarrow \mathbb{R}$  is regularly varying (at infinity) with index  $\alpha \in \mathbb{R}$ , if*

$$\lim_{t \rightarrow +\infty} \frac{f(tx)}{f(t)} = x^\alpha \quad (\forall x > 0).$$

*This is noted  $f \in RV_\alpha$ . If  $\alpha = 0$ ,  $f$  is said to be slowly varying.*

**Remark 4** *Regular variation (and slow variation) can be defined at zero as well. A function  $f$  is said to be regularly varying at zero with index  $\alpha$  if the function  $x \rightarrow f(1/x)$  is regularly varying at infinity, with index  $-\alpha$ .*

**Proposition 2** *(See [Haan and Ferreira (2006)] Proposition B.1.9)*

*Suppose  $f \in RV_\alpha$ . If  $x > 0$  and  $\delta_1, \delta_2 > 0$  are given, then there exists  $t_0 = t_0(\delta_1, \delta_2)$  such that for any  $t \geq t_0$  satisfying  $tx \geq t_0$ , we have*

$$(1 - \delta_1)x^\alpha \min(x^{\delta_2}, x^{-\delta_2}) < \frac{f(tx)}{f(t)} < (1 + \delta_1)x^\alpha \max(x^{\delta_2}, x^{-\delta_2}).$$

If  $x < 1$  and  $\epsilon > 0$ , then there exists  $t_0 = t_0(\epsilon)$  such that for every  $t \geq t_0$ ,

$$(1 - \epsilon)x^{\alpha+\epsilon} < \frac{f(tx)}{f(t)} < (1 + \epsilon)x^{\alpha-\epsilon}$$

and if  $x \geq 1$ ,

$$(1 - \epsilon)x^{\alpha-\epsilon} < \frac{f(tx)}{f(t)} < (1 + \epsilon)x^{\alpha+\epsilon}.$$

**Corollary 4** *If  $f$  is a positive function with end-point  $x^*$ , such that  $t \rightarrow 1 - f(x^* - t)$  is regularly varying at 0 with index  $\alpha$ , i.e.*

$$\frac{1 - f(x^* - tx)}{1 - f(x^* - t)} \rightarrow x^\alpha, \text{ as } t \rightarrow 0,$$

for some  $\alpha \in \mathbb{R}$ , then for every  $\epsilon > 0$ , there exists  $t_0 > 0$  such that,  $\forall 0 < t < t_0$ ,  $\forall 0 < x < 1$ ,

$$(1 - \epsilon)x^{\alpha+\epsilon} \leq \frac{1 - f(x^* - tx)}{1 - f(x^* - t)} \leq (1 + \epsilon)x^{\alpha-\epsilon} \quad (11)$$

and

$$(1 - \epsilon)x^{-\alpha+\epsilon} \leq \frac{1 - f(x^* - t)}{1 - f(x^* - tx)} \leq (1 + \epsilon)x^{-\alpha-\epsilon} \quad (12)$$

Below,  $U$  corresponds to the quantile function associated to  $H$  introduced in paragraph 2.1.

**Corollary 5** *If  $U$  satisfies condition (6), then for every  $\epsilon > 0$ , there exists  $t_0 > 0$  such that,  $\forall 0 < t < t_0$ ,  $\forall x \geq 1$ ,*

$$(1 - \epsilon)x^{\gamma-\epsilon} \leq \frac{U(\infty) - U(tx)}{U(\infty) - U(t)} \leq (1 + \epsilon)x^{\gamma+\epsilon}. \quad (13)$$

**Proposition 3** (see [Haan and Ferreira (2006)] Theorem B.2.18)

*If  $U$  satisfies condition (5) with the positive function  $a$ , then there exists a function  $q_0$  equivalent to  $a/U$  at infinity such that  $\forall \epsilon > 0$ ,  $\exists t_0 > 0$ ,  $\forall t \geq t_0$ ,  $\forall x \geq 1$ ,*

$$\frac{x^\gamma - 1}{\gamma} - \epsilon x^{\gamma+\epsilon} \leq \frac{\log U(tx) - \log U(t)}{q_0(t)} \leq \frac{x^\gamma - 1}{\gamma} + \epsilon x^{\gamma+\epsilon}. \quad (14)$$

We now proceed to the proofs of the different lemmas stated previously and finally of Proposition 1 .



## 6.2 Proof of Lemma 1

Let  $c_n := (q_0(n/k_n)/a_{n,k})^\alpha$  (which tends to 1 as  $n \rightarrow \infty$ ) and  $(Y_i)$  be a sequence of i.i.d standard Pareto random variables. Let

$$LL_{i,k} := \frac{\log(U(Y_{n-i+1,n})) - \log(U(Y_{n-k_n,n}))}{q_0(Y_{n-k_n,n})} \text{ and } QQ_{i,k} := \frac{\left(\frac{Y_{n-i+1,n}}{Y_{n-k_n,n}}\right)^\gamma - 1}{\gamma}.$$

For every  $1 \leq i \leq k_n$ , we thus have

$$\begin{aligned} c_n^{-1} \frac{\xi_{i,n}}{a_{n,k}^\alpha} &\stackrel{d}{=} i((LL_{i,k})^\alpha - (LL_{i+1,k})^\alpha) \\ &= \alpha_{i,n} + \beta_{i,n}, \end{aligned}$$

where

$$\begin{aligned} \alpha_{i,n} &:= i((QQ_{i,k})^\alpha - (QQ_{i+1,k})^\alpha) \\ \beta_{i,n} &:= i(B_{k,n}(i) - B_{k,n}(i+1)) \\ B_{k,n}(i) &:= (LL_{i,k})^\alpha - (QQ_{i,k})^\alpha, \end{aligned}$$

with  $B_{k,n}(k_n + 1) = 0$ .

Our first step will be to prove that (with  $\delta' > 0$ )

$$\frac{1}{k_n} \sum_{i=1}^{k_n} u_i^{-a} \beta_{i,n} \text{ or } \frac{1}{k_n^{\alpha+\delta'}} \sum_{i=1}^{k_n} u_i^{-a} \beta_{i,n} \text{ is } o_{\mathbb{P}}(1) \text{ whether } 0 < a < 1 \text{ or } a > 1.$$

Using bounds (14) for some  $\epsilon' > 0$ , with  $t = Y_{n-k_n,n}$  and  $x = Y_{n-i+1,n}/Y_{n-k_n,n} > 1$ , and relying on the mean value theorem, we easily prove that, since  $\alpha \geq 1$  and  $\gamma < 0$ ,

$$|(LL_{i,k})^\alpha - (QQ_{i,k})^\alpha| \leq c \epsilon' \left(\frac{Y_{n-i+1,n}}{Y_{n-k_n,n}}\right)^{\gamma+\epsilon'} \leq c \epsilon' \quad (15)$$

for some constant  $c$  (close to  $\alpha|\gamma|^{1-\alpha}$ ). Therefore  $|B_{k,n}(i)| = o_{\mathbb{P}}(1)$ , uniformly on  $1 \leq i \leq k_n$ . Since  $|B_{k,n}(i)| = \left|\sum_{j=i}^{k_n} \frac{\beta_{j,n}}{j}\right|$ , we thus have, when  $0 < a < 1$ ,

$$\begin{aligned} \left|\frac{1}{k_n+1} \sum_{i=1}^{k_n} u_i^{-a} \beta_{i,n}\right| &= \left|\sum_{i=1}^{k_n} \frac{\beta_{i,n}}{i} u_i^{1-a} ds\right| \\ &= (1-a) \left|\sum_{i=1}^{k_n} \frac{\beta_{i,n}}{i} \int_0^{u_i} s^{-a} ds\right| \\ &\leq (1-a) \sum_{i=1}^{k_n} \left|\sum_{j=i}^{k_n} \frac{\beta_{j,n}}{j}\right| \int_{u_{i-1}}^{u_i} s^{-a} ds \\ &= o_{\mathbb{P}}(1) \sum_{i=1}^{k_n} (u_i^{1-a} - u_{i-1}^{1-a}) \\ &\leq o_{\mathbb{P}}(1) \frac{1}{k_n} \sum_{i=1}^{k_n} u_i^{-a} \\ &= o_{\mathbb{P}}(1). \end{aligned}$$

The proof for  $a > 1$  is similar (see end of subsection 5.2.1 in [Worms (2014)] for more details, with the difference that now  $|B_{k,n}(i)| = o_{\mathbb{P}}(1)$  uniformly in  $i$ ).

Since we have dealt with the  $\beta_{i,n}$  part, the lemma will be proved as soon as we obtain that, when  $0 < a < 1$ ,

$$\frac{1}{k_n} \sum_{i=1}^{k_n} u_i^{-a} \alpha_{i,n} \xrightarrow{\mathbb{P}} (1-a) |\gamma|^{-\alpha-1} \text{Beta} \left( \frac{1-a}{|\gamma|}; \alpha+1 \right) \quad (16)$$

and, when  $a > 1$ ,

$$\frac{1}{k_n^{a+\delta'}} \sum_{i=1}^{k_n} u_i^{-a} \alpha_{i,n} \xrightarrow{\mathbb{P}} 0. \quad (17)$$

From now on we will sometimes write  $k$  instead of  $k_n$ . Let  $(E_i)$  be a sequence of i.i.d standard exponential random variables. According to (23), and by applying the mean value theorem, there exist some random variables  $E_{i,k}^* \in [E_{k-i,k}, E_{k-i+1,k}]$  such that (remind below that  $\gamma$  is  $< 0$  and  $\alpha \geq 1$ )

$$\begin{aligned} \alpha_{i,n} &\stackrel{d}{=} i \left( \left( \frac{e^{\gamma E_{k-i+1,k-1}}}{\gamma} \right)^\alpha - \left( \frac{e^{\gamma E_{k-i,k-1}}}{\gamma} \right)^\alpha \right) \\ &= \alpha i (E_{k-i+1,k} - E_{k-i,k}) e^{\gamma E_{i,k}^*} \left( \frac{1 - e^{\gamma E_{i,k}^*}}{|\gamma|} \right)^{\alpha-1} \\ &= \alpha |\gamma|^{1-\alpha} \times i (E_{k-i+1,k} - E_{k-i,k}) \left\{ u_i^{|\gamma|} (1 - u_i^{|\gamma|})^{\alpha-1} + \Delta_{i,n} \right\} \end{aligned} \quad (18)$$

where

$$\Delta_{i,n} := e^{\gamma E_{i,k}^*} (1 - e^{\gamma E_{i,k}^*})^{\alpha-1} - u_i^{|\gamma|} (1 - u_i^{|\gamma|})^{\alpha-1}.$$

We will prove later that

$$\max_{i \leq k_n} |\Delta_{i,n}| = o_{\mathbb{P}}(1). \quad (19)$$

For the moment, note that  $(i (E_{k-i+1,k} - E_{k-i,k}))_{i \leq k_n} \stackrel{d}{=} (f_i)_{i \leq k_n}$  due to the Renyi representation, where  $(f_i)$  denotes a sequence of i.i.d standard exponential random variables. Moreover, application of the law of large numbers for triangular arrays of independent random variables (cf [Chow and Teicher (1997)] ; details are omitted) implies that, when  $0 < a < 1$ ,

$$\frac{1}{k_n} \sum_{i=1}^{k_n} u_i^{-a} f_i = O_{\mathbb{P}}(1) \quad \text{and} \quad \frac{1}{k_n} \sum_{i=1}^{k_n} u_i^{-a} f_i u_i^{|\gamma|} (1 - u_i^{|\gamma|})^{\alpha-1} \xrightarrow{\mathbb{P}} \int_0^1 x^{|\gamma|-a} (1 - x^{|\gamma|})^{\alpha-1} dx \quad (20)$$

and, when  $a > 1$  (and  $\delta' > 0$  is given),

$$\frac{1}{k_n^{a+\delta'}} \sum_{i=1}^{k_n} u_i^{-a} f_i = o_{\mathbb{P}}(1). \quad (21)$$

Considering first the situation  $0 < a < 1$ , combining (18), (19) and (20) shows that relation (16) will hold as soon as

$$\int_0^1 x^{|\gamma|-a} (1 - x^{|\gamma|})^{\alpha-1} dx = \alpha^{-1} (1-a) |\gamma|^{-2} \text{Beta} \left( \frac{1-a}{|\gamma|}; \alpha+1 \right).$$

Use of the formulas  $Beta(u, v) = \Gamma(u)\Gamma(v)/\Gamma(u+v)$  and  $u\Gamma(u) = \Gamma(u+1)$  proves the latter relation. When  $a > 1$ , relation (17) is a consequence of (18), (19) and (21).

It remains to prove relation (19). For this purpose, we introduce the sequence  $(V_i)$  of i.i.d. standard uniform random variables such that  $V_{i,k} = e^{-E_{k-i+1,k}}$ , and we note  $V_{i,k}^* := e^{-E_{i,k}^*}$ . If we set  $W_i := V_i^{|\gamma|}$ ,  $W_i^* := (V_{i,k}^*)^{|\gamma|}$  and  $v_i := u_i^{|\gamma|}$ , then relation (19) is now

$$M_n := \max_{1 \leq i \leq k_n} |W_i^*(1 - W_i^*)^{\alpha-1} - v_i(1 - v_i)^{\alpha-1}| = o_{\mathbb{P}}(1),$$

with  $0 < W_i \leq W_i^* \leq W_{i+1} < 1$ . Unfortunately, the function  $x \mapsto x(1-x)^{\alpha-1}$  is not uniformly continuous on  $]0, 1[$  if  $\alpha$  is smaller than 2. Until the end of the proof we will note  $\max_i$  instead of  $\max_{1 \leq i \leq k_n}$ . Since  $|x - a| \leq \max\{|y - a|, |z - a|\}$  whenever  $|x - y| \leq |z - y|$  (this yields the first inequality below), we have

$$\begin{aligned} M_n &\leq \max \left\{ \max_i |W_i(1 - W_{i+1})^{\alpha-1} - v_i(1 - v_i)^{\alpha-1}| ; \right. \\ &\quad \left. \max_i |W_{i+1}(1 - W_i)^{\alpha-1} - v_i(1 - v_i)^{\alpha-1}| \right\} \\ &\leq \max \left\{ \max_i |W_i(1 - W_{i+1})^{\alpha-1} - W_i(1 - W_i)^{\alpha-1}| ; \right. \\ &\quad \left. \max_i |(W_{i+1} - W_i)(1 - W_i)^{\alpha-1}| ; \max_i |W_i(1 - W_i)^{\alpha-1} - v_i(1 - v_i)^{\alpha-1}| \right\} \\ &\leq \max \left\{ \max_i |(1 - W_{i+1})^{\alpha-1} - (1 - W_i)^{\alpha-1}| ; \max_i |W_{i+1} - W_i| ; \right. \\ &\quad \left. \max_i |W_i - v_i| ; \max_i |(1 - W_i)^{\alpha-1} - (1 - v_i)^{\alpha-1}| \right\} \\ &=: \max \{ M_{n,1} ; M_{n,2} ; M_{n,3} ; M_{n,4} \}. \end{aligned}$$

Now, since  $|x^a - y^a| \leq (a \wedge 1)|x - y|^{a \vee 1}$  whenever  $x$  and  $y$  belong to  $[0, 1]$  and  $a > 0$ , we have  $M_{n,1} \leq cM_{n,2}^{c'}$  and  $M_{n,4} \leq cM_{n,3}^{c'}$  for some positive constants  $c$  and  $c'$ . On the other hand,  $M_{n,2}$  is bounded by  $2M_{n,3} + \max_i |v_{i+1} - v_i| = 2M_{n,3} + o(1)$ . Therefore, the negligibility of  $M_n$  amounts to

$$M_{n,3} = \max_{1 \leq i \leq k_n} |V_i^{|\gamma|} - u_i^{|\gamma|}| = o_{\mathbb{P}}(1).$$

This property is proved in details in [Beirlant et al. (2002)] (page 164, with  $-\rho$  instead of  $|\gamma|$ ), so we do not reproduce it here.  $\diamond$

### 6.3 Proof of Lemma 2

If  $Y_{1,n}, \dots, Y_{n,n}$  denote the ascending order statistics of  $n$  i.i.d standard Pareto random variables, we have

$$\tilde{Z}_{i,n} \stackrel{d}{=} \frac{U(\infty) - U(Y_{n-i+1,n})}{U(\infty) - U(Y_{n-k_n,n})}.$$

Applying bounds (13), it comes, for some given  $\epsilon' > 0$  and  $n$  sufficiently large,

$$\begin{aligned} \left( \frac{U(\infty) - U(Y_{n-i+1,n})}{U(\infty) - U(Y_{n-k_n,n})} \right)^\theta &\leq (1 + \epsilon')^\theta \left( \frac{Y_{n-i+1,n}}{Y_{n-k_n,n}} \right)^{a_+} \\ \left( \frac{U(\infty) - U(Y_{n-i+1,n})}{U(\infty) - U(Y_{n-k_n,n})} \right)^{\theta'} &\geq (1 - \epsilon')^{\theta'} \left( \frac{Y_{n-i+1,n}}{Y_{n-k_n,n}} \right)^{a_-} \end{aligned}$$

where

$$a_+ = (\gamma + \epsilon')\theta \quad \text{and} \quad a_- = (\gamma - \epsilon')\theta'.$$

We finish the proof as for Lemma 1 of [Worms (2014)].  $\diamond$

## 6.4 Proof of Proposition 1

(i) First, note that for any sequences  $(a_n)$  and  $(b_n)$  such that  $a_0 = 0$  and  $b_{k+1} = 0$ , we have  $\sum_{i=1}^k a_i(b_i - b_{i+1}) = \sum_{i=1}^k (a_i - a_{i-1})b_i$ . By letting  $a_i = \frac{i}{1 - \hat{G}_n(Z_{n-i+1,n}^-)}$  and  $b_i = \log^\alpha \left( \frac{Z_{n-i+1,n}}{Z_{n-k_n,n}} \right)$ , we have  $b_{k+1} = 0$  but  $a_0$  is undefined, so we set  $a_0 = 0$ . Therefore, by the definition (3) of  $\xi_{i,n}$ , this implies that

$$\begin{aligned} \sum_{i=1}^{k_n} \frac{\xi_{i,n}}{1 - \hat{G}_n(Z_{n-i+1,n}^-)} &= \sum_{i=2}^{k_n} \left( \frac{i}{1 - \hat{G}_n(Z_{n-i+1,n}^-)} - \frac{i-1}{1 - \hat{G}_n(Z_{n-i+2,n}^-)} \right) \log^\alpha \left( \frac{Z_{n-i+1,n}}{Z_{n-k_n,n}} \right) \\ &\quad + \frac{1}{1 - \hat{G}_n(Z_{n,n})} \log^\alpha \left( \frac{Z_{n,n}}{Z_{n-k_n,n}} \right). \end{aligned}$$

By definition of  $M_{n,k_n}^{(\alpha)}$  and  $D_{n,k_n}^{(\alpha)}$ , part (i) of Proposition 1 will be proved as soon as we show that, for  $2 \leq i \leq k_n$ ,

$$\frac{i}{1 - \hat{G}_n(Z_{n-i+1,n}^-)} - \frac{i-1}{1 - \hat{G}_n(Z_{n-i+2,n}^-)} = \frac{\delta_{n-i+1,n}}{1 - \hat{G}_n(Z_{n-i+1,n}^-)}.$$

Note that

$$\begin{aligned} &\frac{i}{1 - \hat{G}_n(Z_{n-i+1,n}^-)} - \frac{i-1}{1 - \hat{G}_n(Z_{n-i+2,n}^-)} \\ &= \frac{1}{1 - \hat{G}_n(Z_{n-i+1,n}^-)} \left( i - (i-1) \frac{1 - \hat{G}_n(Z_{n-i+1,n}^-)}{1 - \hat{G}_n(Z_{n-i+2,n}^-)} \right) \end{aligned} \quad (22)$$

where

$$\frac{1 - \hat{G}_n(Z_{n-i+1,n}^-)}{1 - \hat{G}_n(Z_{n-i+2,n}^-)} = \frac{\prod_{j=1}^{n-i} \left( \frac{n-j}{n-j+1} \right)^{1-\delta_{j,n}}}{\prod_{j=1}^{n-i+1} \left( \frac{n-j}{n-j+1} \right)^{1-\delta_{j,n}}}.$$

The right-hand side is clearly equal to  $\frac{i}{i-1} \left( \frac{i-1}{i} \right)^{\delta_{n-i+1,n}}$ , which may be rewritten as  $\frac{i}{i-1} \left( 1 - \frac{\delta_{n-i+1,n}}{i} \right)$ . Replacing in (22) concludes the proof of (i).

(ii) Note first that

$$D_{n,k_n}^{(\alpha)} = \frac{1}{k_n} A_n C_{1,n} W_{1,n},$$

where  $A_n$  and  $C_{i,n}$  are defined in the beginning of the proof of Theorem 1 and

$$W_{1,n} = \frac{1 - G(Z_{n-k_n,n})}{1 - G(Z_{n,n})} \log^\alpha \left( \frac{Z_{n,n}}{Z_{n-k_n,n}} \right).$$

We know that  $A_n = o_{\mathbb{P}}(1)$ . Since  $C_{1,n} = O_{\mathbb{P}}(1)$  (see Theorem 2.2 in [Zhou (1991)]), we only have to prove that  $\frac{1}{k_n} \frac{W_{1,n}}{a_{n,k}^\alpha} = o_{\mathbb{P}}(1)$ . For the same reasons that led to (9), for any  $\epsilon > 0$  and  $n$  sufficiently large, we get  $W_{1,n} \leq W_{1,n}^+$  with

$$W_{1,n}^+ = (1 + \epsilon) \log^\alpha \left( \frac{Z_{n,n}}{Z_{n-k_n,n}} \right) \tilde{Z}_{1,n}^{\gamma_C^{-1} - \epsilon}.$$

If  $(Y_i)_{1 \leq i \leq n}$  is an i.i.d. sequence of standard Pareto random variables, then

$$c_n^{-1} \frac{W_{1,n}^+}{a_{n,k}^\alpha} \stackrel{d}{=} (1 + \epsilon) (LL_{1,k})^\alpha (U_{1,k})^{\gamma_C^{-1} - \epsilon},$$

where  $c_n \rightarrow 1$  and  $LL_{1,k}$  were defined in the proof of Lemma 1 and  $U_{1,k} = \frac{U(\infty) - U(Y_{n,n})}{U(\infty) - (Y_{n-k_n,n})}$ . On one hand, Potter bounds (13) yield for any  $\epsilon' > 0$  and  $n$  sufficiently large,

$$(U_{1,k})^{\gamma_C^{-1} - \epsilon} \leq (1 + \epsilon') \left( \frac{Y_{n,n}}{Y_{n-k_n,n}} \right)^{a^+},$$

with  $a^+ = (\gamma + \epsilon')(\gamma_C^{-1} - \epsilon)$ . On the other hand, (15) implies that for  $\epsilon'' > 0$ ,  $|(LL_{1,k})^\alpha - (QQ_{1,k})^\alpha| \leq c\epsilon''$ , for some constant  $c$ , where  $QQ_{i,k}$  is also defined in the proof of Lemma 1. Moreover, it is known that

$$(Y_{n-i+1,n}/Y_{n-k_n,n})_{1 \leq i \leq k_n} \stackrel{d}{=} (\tilde{Y}_{k_n-i+1,k_n})_{1 \leq i \leq k_n}, \quad (23)$$

where  $\tilde{Y}_{1,k_n}, \dots, \tilde{Y}_{k_n,k_n}$  are the ascending order statistics of  $k_n$  i.i.d random variables  $\tilde{Y}_1, \dots, \tilde{Y}_{k_n}$  with standard Pareto distribution.

Consequently,  $(LL_{1,k})^\alpha (U_{1,k})^{\gamma_C^{-1} - \epsilon} \leq Q_{1,n}$ , with

$$Q_{1,n} \stackrel{d}{=} (1 + \epsilon') \left( \frac{\tilde{Y}_{k_n,k_n} - 1}{\gamma} \right)^\alpha (\tilde{Y}_{k_n,k_n})^{a^+} + c\epsilon''(1 + \epsilon')(\tilde{Y}_{k_n,k_n})^{a^+}.$$

Since  $\tilde{Y}_{k_n,k_n} \geq 1$  and  $\gamma < 0$ , the right-hand side of the relation above is lower than  $c'(\tilde{Y}_{k_n,k_n})^{a^+}$ , for some constant  $c'$ .

Now, Standard Pareto distributions having moments of order less than 1,

$$\tilde{Y}_{k_n,k_n} = \max_{1 \leq i \leq k_n} \tilde{Y}_{i,k_n} = o_{\mathbb{P}}(k_n^p),$$

for any  $p > 1$ , therefore,  $\frac{1}{k_n} (\tilde{Y}_{k_n,k_n})^{a^+} = o_{\mathbb{P}}(k_n^{pa^+ - 1})$ , which is clearly  $o_{\mathbb{P}}(1)$ . This concludes the proof.  $\diamond$

## References

- [Beirlant et al. (2002)] J. Beirlant, G. Dierckx, A. Guillou and C. Stărică . On exponential representations of log spacings of order statistics. In *Extremes* **5**, pages 157-180 (2002)
- [Beirlant et al. (2007)] J. Beirlant, G. Dierckx, A. Guillou and A. Fils-Villetard . Estimation of the extreme value index and extreme quantiles under random censoring. In *Extremes* **10**, pages 151-174 (2007)
- [Beirlant et al. (2010)] J. Beirlant, A. Guillou and G. Toulemonde . Peaks-Over-Threshold modeling under random censoring. In *Comm. Stat. : Theory and Methods* **39**, pages 1158-1179 (2010)
- [Brahimi et al. (2013)] B. Brahimi, D. Meraghni and A. Necir . Approximations to the tail index estimator of a heavy-tailed distribution under random censoring and application. working paper, arXiv:1302.1666 (2013)
- [Chow and Teicher (1997)] Y.S. Chow and H. Teicher . Probability theory. Independence, interchangeability, martingales. *Springer* (1997)
- [Csörgő (1996)] S. Csörgő . Universal Gaussian approximations under random censorship. In *Annals of statistics* **24 (6)**, pages 2744-2778 (1996)
- [Diop et al. (2014)] A. Diop, J-F. Dupuy and P. Ndao. Nonparametric estimation of the conditional tail index and extreme quantiles under random censoring. In *Computational Statistics & Data Analysis* **79**, pages 63-79 (2014)
- [Einmahl et al. (2008)] J. Einmahl, A. Fils-Villetard and A. Guillou . Statistics of extremes under random censoring. In *Bernoulli* **14**, pages 207-227 (2008)
- [Gomes and Neves (2011)] M.I. Gomes and M.M. Neves . Estimation of the extreme value index for randomly censored data. In *Biometrical Letters* **48 (1)**, pages 1-22 (2011)
- [Haan and Ferreira (2006)] L. de Haan and A. Ferreira . Extreme Value Theory : an Introduction. *Springer Science + Business Media* (2006)
- [Segers (2001)] J. Segers. Residual estimators. In *Journal of Statistical Planning and Inference* **98**, pages 15-27 (2001)
- [Worms (2014)] J. Worms and R. Worms New estimators of the extreme value index under random right censoring, for heavy-tailed distributions. In *Extremes* **17 (2)**, pages 337-358 (2014)
- [Zhou (1991)] M. Zhou. Some Properties of the Kaplan-Meier Estimator for Independent Nonidentically Distributed Random Variables. In *Annals of statistics* **19 (4)**, pages 2266-2274 (1991)