Afrika Statistika
Vol. 9, 2014, pages 671–683.
DOI: http://dx.doi.org/1016929.as.2014.671.61

Afrika Statistika

ISSN 2316-090X

On robust tail index estimation under random censorship

Abdallah Sayah*, **Djabrane Yahia and Brahim Brahimi**

Laboratory of Applied Mathematics, Mohamed Khider University, Biskra, Algeria

Received 07 May 2014; Accepted 13 October 2014

Abstract. In this paper, we propose a new robust tail index estimation procedure for Pareto-type distributions in the framework of randomly censored samples, based on the ideas of Kaplan-Meier integration using the huberized M-estimator of the tail index. We derive their asymptotic results. We illustrate the performance and the robustness of this estimator for small and large sample size in a simulation study.

Résumé. Dans cet article, nous proposons une nouvelle procédure de l'estimation robuste de l'indice de la queue pour les distributions de type Pareto dans le cas d'échantillons censurés, sur la base des idées de l'intgrale de Kaplan-Meier en utilisant le huberized M-estimateur de l'indice de la queue. Nous dérivons leurs résultats asymptotiques. Nous illustrons dans l'étude de la simulation la performance et la robustesse de cet estimateur pour un échantillon de petite et grande taille.

Key words: Heavy-tailed distributions; Hill estimator; Random censorship; Regular variation; Robust estimation; Tail index.
AMS 2000 Mathematics Subject Classification : 62G07; 62G20.

1. Introduction

Let $X_1, ..., X_n$ be n copies of independent and identically distributed random variable (rv) X, with common cumulative distribution function (cdf) F assumed to be heavy-tailed. In other words, the distribution tail $\overline{F} := 1 - F$ is regularly varying, with index $(-\alpha_1)$, notation: $\overline{F} \in \mathcal{RV}_{(-\alpha_1)}$. That is

$$\lim_{t\to\infty} \frac{\overline{F}(tx)}{\overline{F}(t)} = x^{-\alpha_1}, \text{ for any } x > 0,$$

*Corresponding author Abdallah Sayah: sayahabdel@yahoo.fr
Djabrane Yahia : yahia_dj@yahoo.fr
Brahim Brahimi: brah.brahim@gmail.com

where $\alpha_1 > 0$ is called shape parameter, tail index or extreme value index (EVI). It plays a very crucial role in the analysis of extremes as it governs the thickness of the distribution tails.

Suppose there exist a sequences of constants $(a_n > 0)$ and $(b_n \in \mathbb{R})$ such that the properly centered and normed sample maxima converge in distribution, as $n \to \infty$, to a non-degenerate limit distribution H, for all continuity points of H, i.e,

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{X_{n:n} - b_n}{a_n} \le x\right) = H_\alpha(x), \tag{1}$$

where $X_{1:n} \le ... \le X_{n:n}$ are the order statistics pertaining to the sample $(X_1, ..., X_n)$. The limit distribution H is necessarily of generalized extreme value type (Fisher and Tippett, 1928)

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

If (1) is satisfied, then F is said to belong to the maximum domain of attraction of H_α, denoted as $F \in D(H_\alpha)$.

The estimation of α_1 has a great interest for a complete data by many authors and common applications in a big variety of domains, as for example in economics, applied finance, insurance, business, industry, traffic, telecommunications, sociology and geology, as one might see the textbook Beirlant *et al.* (2007), Dekkers *et al.* (1989), Bacro and Brito (1995), Csörgő and Viharos (1998) and references therein.

The most celebrated estimator of α_1 is that proposed by Hill (1975)

$$\hat{\alpha}_1^H := \hat{\alpha}_1^H(k) = \left(\frac{1}{k}\sum_{i=1}^{n} \log(X_{n-i+1,n}) - \log(X_{n-k,n})\right)^{-1},$$

for $k = k_n$ is an integer sequence satisfying

$$1 < k < n, \quad k \to \infty \quad \text{and } k/n \to 0 \quad \text{as } n \to \infty. \tag{2}$$

The asymptotic properties of $\hat{\alpha}_1^H$ have been much studied. In the independent context, it is well known that, under some regularity conditions, $\hat{\alpha}_1^H$ is strongly consistent with asymptotic normal distribution when properly normalized Haeusler and Teugels (1985). The consistency of $\hat{\alpha}_1^H(k)$ was proved by Mason (1982) by only assuming the regular variation condition while its asymptotic normality was established under a suitable extra assumption, known as the second-order regular variation condition (see de Haan and Stadtmüller, 1996 and de Haan and Ferreira, 2006, page 117).

In many real applications, such as survival analysis, reliability theory or insurance..., the variable of interest X is not necessarily completely available. This is the case in the presence of random right censoring. The usual way to model this situation is to introduce a random variable Y called censoring rv, independent of X, and then to consider the rv $Z := \min(X, Y)$ and the indicator variable $\delta := \mathbf{1}(X \le Y)$, which determines whether or not X has been observed. The cdf's of Y and Z will be denoted by G and H respectively. Statistics of extremes of randomly censored data is a new research field. The topic was first mentioned in Reiss and

Thomas (1997), where an estimator of a positive extreme value index was introduced, but with no asymptotic results. Recently, Beirlant *et al.* (2007) proposed an estimators for the general extreme value index and for the extreme quantile with their asymptotic properties. Einmahl *et al.* (2008) adapted various extreme value index estimators to the case where the data are censored, by a random threshold and establish their asymptotic normality by imposing some assumptions that are rather unusual to the context of extreme value theory. More recently Brahimi *et al.* (2014), using the empirical process theory to approximate the adapted Hill estimator, for censored data, and derived its asymptotic normality.

The tail of the censoring distribution is assumed to be regularly varying too, that is $1 - G \in \mathcal{RV}_{(-\alpha_2)}$, for some $\alpha_2 > 0$. By virtue of the independence of X and Y, we have $1 - H(x) = (1 - F(x))(1 - G(x))$ and therefore $1 - H \in \mathcal{RV}_{(-\alpha)}$, with $\alpha := \alpha_1 + \alpha_2$. Let $\{(Z_i, \delta_i), \ 1 \leq i \leq n\}$ be a sample from the couple of rv's (Z, δ) and $Z_{1,n} \leq ... \leq Z_{n,n}$ represent the order statistics pertaining to $(Z_1, ..., Z_n)$. If we denote the concomitant of the ith order statistic by $\delta_{[i:n]}$ (i.e. $\delta_{[i:n]} = \delta_j$ if $Z_{i,n} = Z_j$), then the adapted Hill estimator of the tail index α_1 is defined by

$$\widehat{\alpha}_1^{(H,c)} := \frac{\widehat{\alpha}^H(k)}{\widehat{p}}, \tag{3}$$

where

$$\widehat{\alpha}^H(k) := \frac{1}{k}\sum_{i=1}^{k} \log Z_{n-i+1,n} - \log Z_{n-k:n} \tag{4}$$

and

$$\widehat{p} := \frac{1}{k}\sum_{i=1}^{k} \delta_{[n-i+1:n]}, \tag{5}$$

with $k := k_n$ satisfying (2). Roughly speaking, the adapted Hill estimator is equal to the quotient of the classical Hill estimator to the proportion of non censored data.

The rest of the paper is organized as follows. In Section 2, after a brief discussion on the huberized tail index M-functional and huberized M-estimator of the tail index, we derive our main result, namely the asymptotic normality of the robust tail index estimator in the framework of randomly censored samples. A small simulation study is cared to check the performance and the robustness of our estimator is given in Section 3. Concluding notes are given in Section 4. Proofs are relegated to Section 5.

2. Main results

The classical tail index estimators achieve consistency by relying on an asymptotically vanishing portion of high quantiles only. This results in a slow rate of convergence and applicability to relatively large samples only. The alternative approach is inspired by the theory of robust inference (Hampel *et al.*, 1986 and Huber, 1981) instead of exact consistency this theory aim at stability for small samples, possibly at the cost of a small asymptotic bias. This can be obtained by the definition of the following class of M-functional and M-estimators respectively that are defined as follows (see Beran and Shell, 2012).

Definition 1. Let $F_{Par}(x,\alpha) = 1 - x^{-\alpha}$ $(x \geq 1)$ and

$$\psi_v(x,\alpha) = [\alpha \log(x) - 1]_v - \int [\alpha \log(z) - 1]_v dF_{Par}(z,\alpha)$$
$$= [\alpha \log(x) - 1]_v - (v + \exp(-(v+1))),$$

where $[y]_v = \max(y,v)$, and denote by $\mathcal{F}$ a set of distributions with support in $\mathbb{R}_+$. Then the functional $T_{HUB} =: T_{HUB}^v$ defined on $\mathcal{F}$ as the solution t_0 of the equation

$$\beta_F(t) = \int \psi_v(x,t)\, dF(x) = 0, \quad (F \in \mathcal{F})$$

is called huberized tail index M-functional. The corresponding M-estimator $T_n =: T_n^v$, defined by

$$\sum_{j=1}^{n} \psi_v(X_j, T_n) = 0, \tag{6}$$

is called huberized M-estimator of the tail index.

The nonparametric maximum likelihood estimator of F in the case of censored data equals the famous estimator of Kaplan and Meire (1958) also called the product limit estimator, is given by

$$1 - \hat{F}_n(z) := \prod_{i:Z_{i,n} \leq z} \left(1 - \frac{\delta_{(i:n)}}{n-i+1}\right), \text{ for } z \in \mathbb{R}. \tag{7}$$

Stute and Wang (1993) and Stute (1995) studied the almost sure and distributional behavior of the so-called Kaplan-Meier integrals

$$I_n := \int \varphi(z)\, d\hat{F}_n(z),$$

where φ is an arbitrary integrable function. It is easily seen from (7) that

$$I_n = \sum_{i=n}^{n} W_{in} \varphi(Z_{i:,n}),$$

where for $1 \leq i \leq n$

$$W_{in} = \frac{\delta_{[i:n]}}{n-i+1} \prod_{j=1}^{i-1} \left[\frac{n-j}{n-j+1}\right]^{\delta_{[j:n]}}.$$

When there is no censorship, $W_{in} = 1/n$ so that I_n becomes the sample mean.

Stute (1995) obtained under random censoring and under some assumptions of central limit theorem for a general transformation φ, that

$$\int \varphi d\left(\hat{F}_n - F\right).$$

The functions defined below are crucial to our needs:

$$\tilde{H}_0(z) := P(Z \leq z, \delta = 0) = \int_{-\infty}^{z} (1 - F(t))\, G(dt),$$

$$\tilde{H}_1(z) := P(Z \le z, \delta = 1) = \int_{-\infty}^{z} (1 - G(t)) F(dt),$$

$$\lambda_0(z) := \exp\left(\int_{-\infty}^{z} \frac{\tilde{H}_0(dx)}{1 - H(x)} \right),$$

$$\lambda_1(z) := \frac{1}{1 - H(z)} \int \psi_v(x, \alpha_1) \lambda_0(x) \mathbf{1}_{\{z<x\}} \tilde{H}_1(dx),$$

and

$$\lambda_2(z) := \int \int \frac{\psi_v(x, \alpha_1) \lambda_0(x) \mathbf{1}_{\{y<z,\ y<x\}}}{1 - H(y)} \tilde{H}_0(dy) \tilde{H}_1(dx).$$

The following assumptions will be needed in theorem

$$\int \psi_v^2(x, \alpha_1) \lambda_0^2(x) \tilde{H}_1(dx) < \infty, \tag{8}$$

and

$$\int |\psi_v(x, \alpha_1)| A^{1/2}(x) F(dx) < \infty, \tag{9}$$

where

$$A(x) := \int_{-\infty}^{x} \frac{G(dy)}{(1 - H(y))(1 - G(y))}.$$

Theorem 1. *Let $X_i \sim F_{Par}(x, \alpha_1)$ and $Y_i \sim F_{Par}(y, \alpha_2)$, $x \ge 1$, $y \ge 1$ where $\alpha_1 > 0$ and $\alpha_2 > 0$, with $\alpha_2 < \alpha_1$. Moreover, let $\hat{F}_n$ be the Kaplan-Meier estimator of the df F and $(T_n)_{n \ge 0}$ a sequence of solutions of*

$$\lambda_{\hat{F}_n}(t) = \sum_{j=1}^{n} \psi_v(X_j, t) = 0, \qquad (n \in \mathbb{N}).$$

Then, under assumptions (8) and (9) we have

$$n^{1/2}(T_n - \alpha_1) \xrightarrow{D} \mathcal{N}\left(0, \sigma_{hub}^2\right),$$

where

$$\sigma_{hub}^2 = \frac{\alpha_1}{v+2} e^{v+1} \sigma^2,$$

and

$$\sigma^2 = Var\left\{ -\psi_v(z, \alpha_1) \lambda_0(z) \delta + \lambda_1(z)(1 - \delta) - \lambda_2(z) \right\}.$$

Remark 1. Condition (8) is the properly modified variance assumption on ψ_v and (9) only incorporates the first ψ_v−moment. It is mainly to control the bias of $\int \psi_v d\hat{F}_n$, which is a function of ψ_v rather than ψ_v^2. Stute (1994) and Stute (1995) gives a detailed account of this issue. In our case, this two assumptions are satisfied when $\alpha_2 < \alpha_1$. The expressions of λ_0, λ_1 and λ_2 are previously given.

3. Simulation study

3.1. Performance and comparative study

In this simulation study we examines the performance of our estimator given in Definition 1 and compare with the adapted Hill estimator given in (3) proposed by Einmahl *et al.* (2008). For this reason, we follows the steps below.

Step 1: We generate 1000 pseudorandoms samples X and C of size $n = 100, 200, 500$ and 1000 from Pareto cdf with $\alpha_1 = 0.6$ and $\alpha_2 = 0.25$ respectively. Here $v = 1$ and $p = 0.70$ that means the percentage of censorship is 30%.

Step 2: We obtained 1000 pseudorandom samples $Z = \min(X, C)$ and the indicator variable $\delta := \mathbf{1}(X \le C)$ of size $n = 100, 200, 500$ and 1000.

Step 3: We estimate the tail index parameter by the two estimators frame the observed data Z. We adopt the Reiss and Thomas algorithm (see Reiss and Thomas, 1997), for choosing the optimal numbers of upper extremes k in adapted Hill estimator. By this methodology, we define the optimal sample fraction of upper order statistics k by

$$k := \arg\min_j \frac{1}{j}\sum_{i=1}^{j} i^{\theta}\left|\widehat{\alpha}_1^H(i) - \text{median}\left\{\widehat{\alpha}_1^H(1), ..., \widehat{\alpha}_1^H(j)\right\}\right|.$$

On the light of our simulation study, we obtained reasonable results by choosing $\theta = 0.3$.

Step 4: We compute the bias and root mean squared error (RMSE) of the two estimators, the results are summarized in Table 1. we see that our estimator performs better.

	T_n		k	$\widehat{\alpha}_1^{(H,c)}(k)$	
n	bias	RMSE		bias	RMSE
100	0.0611	0.2511	17	−0.1143	0.2586
200	0.0431	0.1013	34	−0.0845	0.1821
500	0.0153	0.0684	86	−0.0245	0.1142
1000	0.0041	0.0356	169	−0.0070	0.0798

Table 1. Bias and RMSE of the two estimators based on 1000 samples of Pareto-distributed with tail index 0.6.

3.2. Comparative robustness study

We study the sensitivity to outliers of our estimator and compare with the adapted Hill estimator. We consider an ϵ-contaminated model known by mixture of Pareto distributions

$$F_{\alpha_1,\gamma_2,\epsilon}(z) = 1 - (1-\epsilon)\, z^{-1/\alpha_1} + \epsilon z^{-1/\gamma_2}, \tag{10}$$

where $\alpha_1, \gamma_2 > 0$ and $0 < \epsilon < 0.5$ is the fraction of contamination. Note that for $\epsilon = 0$, T_n and $\widehat{\alpha}_1^{(H,c)}(k)$ are asymptotically unbiased. Therefore, for $\epsilon > 0$, the effect of contamination becomes immediately apparent. If $\alpha_1 < \gamma_2$ and $\epsilon > 0$, (10) corresponds to a Pareto distribution contaminated by a longer tailed distribution.

For the implementation of mixtures models to the study outliers one refers, for instance, to (Barnett and Lewis, 1995, page 43). In this context, we proceed our study as follows.

We consider $\alpha_1 = 0.6$, $\gamma_2 = 2$ to have the contaminated model. Then we consider four contamination scenarios according to $\epsilon = 5\%, 10\%, 15\%, 25\%$.

For each value ϵ, we generate 1000 samples of size $n = 100$, 200 and 1000 from the model (10).

Finally, we compare the two estimators with this true value, by computing for each estimator, the appropriate bias and RMSE and summarize the results in Table 2.

		T_n		$\widehat{\alpha}_1^{(H,c)}(k)$	
n	% contamination	bias	RMSE	bias	RMSE
100	5	0.0685	0.2865	−0.1325	0.3251
	10	0.0754	0.3561	−0.6485	0.7546
	15	0.0791	0.3940	−1.0452	1.1256
	25	0.1245	0.5412	−1.1125	2.1951
200	5	0.0487	0.1965	−0.0911	0.2511
	10	0.0510	0.2213	−0.2496	0.3217
	15	0.0614	0.3889	−0.4518	0.9941
	25	0.1002	0.5001	−0.7120	1.4963
1000	5	0.0095	0.1002	−0.1194	0.1227
	10	0.0191	0.2249	−0.2162	0.3978
	15	0.0531	0.3449	−0.4101	0.4355
	25	0.0977	0.4250	−0.6788	0.9591

Table 2. Bias and RMSE of the two estimators based on 1000 samples of mixture of Pareto distributions with tail index 0.6, $\epsilon = 5\%, 10\%, 15\%, 25\%$.

The adapted Hill estimator is a pseudo-maximum likelihood estimator based on the exponential approximation of the normalized log-spacings $Y_j = j(\log Z_{j,n} - \log Z_{j+1,n})$ for $j = 1, ..., k$, So in practice, this estimator depends on the choice of k and is inherently not very robust to large values Y_j, which be sensitive to few particular observations, which constitutes a serious problem in terms of bias and RMSE. As expected, the adapted Hill estimator turn out to be more sensitive to this type of contaminations, for example, in 0% contamination for $n = 200$ the (bias, RMSE) of the adapted Hill estimator equals $(-0.0854, 0.1013)$, while for 25% contamination is $(-0.7120, 1.4963)$. We may conclude that the bias and RMSE of the adapted Hill estimator in more sensitive (or note robust) to outliers, however for 0% contamination the (bias, RMSE) of our estimator equals $(0.0431, 0.1821)$, while for 25% contamination is $(0.1002, 0.5001)$. Both the bias and the RMSE of our estimator are note sensitive to outliers, then we may conclude that is the better estimator.

4. Concluding notes

It has been shown that our estimator is more robust and perform better than the adapted Hill estimator proposed by Einmahl *et al.* (2008).

5. Proofs

Proof (Proof of Theorem 1). The proof is essentially based on Theorem 11 in (Stute and Wang, 1993, page 1594), Corollary 1.2 in (Stute, 1995, page 426) and (Beran and Shell, 2012, page 3432).

Let

$$\beta(\alpha) := \beta_{F_{Par}(x,\alpha_1)}(\alpha) = \int \psi_v(x,\alpha)\, dF_{Par}(x,\alpha_1).$$

Partial integration yields

$$\begin{aligned}\beta(\alpha) &= \int_1^\infty [\alpha \log(x) - 1]_v - \left(\int_1^\infty [\alpha \log(z) - 1]_v dF_{Par}(z,\alpha)\right) dF_{Par}(x,\alpha_1) \\ &= \frac{\alpha}{\alpha_1}\left(\exp(-\alpha_1 \frac{v+1}{\alpha})\right) - (\exp -(v+1)).\end{aligned}$$

Hence $\beta(\alpha_1) = 0$.

$$\begin{aligned}&\int \psi_v(z,T_n)\, d\hat{F}_n(z) - \int \psi_v(z,\alpha_1)\, dF(z) - \int \psi_v(z,\alpha_1)\, d\hat{F}_n(z) + \int \psi_v(z,\alpha_1)\, d\hat{F}_n(z) \\ &= \int (\psi_v(z,T_n) - \psi_v(z,\alpha_1))\, d\hat{F}_n(z) + \int \psi_v(z,\alpha_1)\, d\left(\hat{F}_n(z) - F(z)\right) = 0, \qquad (11)\end{aligned}$$

where

$$\psi_v(z,\alpha_1) = \begin{cases} v - \omega, & \text{if } 1 \le z \le \exp\left(\dfrac{v+1}{\alpha_1}\right) \\ \alpha_1 \ln z - 1 - \omega, & \text{if } \exp\left(\dfrac{v+1}{\alpha_1}\right) \le z < \infty \end{cases}$$

and

$$\omega = v + \exp(-(v+1)) \text{ with } -1 \le v < \infty.$$

By Taylors theorem (11)

$$n^{1/2}(T_n - \alpha_1) \int \psi_v^{'}(z,\alpha_1)\, d\hat{F}_n(z) = n^{1/2} \int (-\psi_v(z,\alpha_1))\, d\left(\hat{F}_n(z) - F(z)\right).$$

Then,

$$n^{1/2}(T_n - \alpha_1) = n^{1/2} \left(\int \left(\psi_v^{'}(z,\alpha_1)\right) d\hat{F}_n(z)\right)^{-1} \int (-\psi_v(z,\alpha_1))\, d\left(\hat{F}_n(z) - F(z)\right).$$

It was shown in theorem 1.1 in Stute and Wang (1993) that for any measurable real function φ, and under the condition $\int |\varphi|\, dF < \infty$, we get

$$\int \varphi d\hat{F}_n = \int \varphi dF + op(1). \qquad (12)$$

From Stute (1995) under assumptions (8) and (9) we have

$$n^{1/2} \int (-\psi_v(z,\alpha_1))\, d\left(\hat{F}_n(z) - F(z)\right) \xrightarrow{D} \mathcal{N}\left(0,\sigma^2\right),$$

where

$$\sigma^2 = Var\left(-\psi_v\left(z,\alpha_1\right)\lambda_0\left(z\right)\delta + \lambda_1\left(z\right)\left(1-\delta\right) - \lambda_2\left(z\right)\right). \tag{13}$$

Using (12) and (13) we get

$$n^{1/2}\left(T_n - \alpha_1\right) \overset{D}{\to} \mathcal{N}\left(0, \sigma_{hub}^2\right),$$

where

$$\sigma_{hub}^2 = \left(\int \left(\psi_v'\left(z,\alpha_1\right)\right) dF\left(z\right)\right)^{-1} \sigma^2.$$

A simple calculate implies that $\psi_v'\left(y,\alpha_1\right) = \ln y$, and

$$\begin{aligned}\int \psi_v'\left(y,\alpha_1\right) dF\left(y\right) &= \alpha_1 \int_{e^{\frac{v+1}{\alpha_1}}}^{\infty} y^{-\alpha_1-1} \ln y dy \\ &= -y^{-\alpha_1}\left(\ln y + \frac{1}{\alpha_1}\right)\Big|_{e^{\frac{v+1}{\alpha_1}}}^{\infty} \\ &= e^{-(v+1)}\left(\frac{v+2}{\alpha_1}\right).\end{aligned}$$

We showed that the conditions (8) and (9) are verified for $\alpha_2 < \alpha_1$. So that

$$\begin{aligned}\frac{1}{\alpha_1}\int_1^\infty \psi_v^2\left(x,\alpha_1\right)\lambda_0^2\left(x\right)\tilde{H}_1'\left(x\right)dx &= \int_1^{\exp\left(\frac{v+1}{\alpha_1}\right)} \left(v-\omega\right)^2 x^{\alpha_2-\alpha_1-1}dx \\ &+ \int_{\exp\left(\frac{v+1}{\alpha_1}\right)}^\infty \left(\alpha_1 \ln x - 1 - \omega\right)^2 x^{\alpha_2-\alpha_1-1} dx < \infty\end{aligned}$$

The condition (9) can be written in the form

$$c + \int_{\exp\left(\frac{v+1}{\alpha_1}\right)}^\infty \left|\alpha_1 \ln x - 1 - \omega\right| \left(x^{\alpha_2+\alpha_1} - 1\right)^{1/2} x^{-\alpha_1-1}dx$$

$$\sim c + \int_{\exp\left(\frac{v+1}{\alpha_1}\right)}^\infty \left|\alpha_1 \ln x - 1 - \omega\right| x^{\frac{\alpha_2-\alpha_1}{2}-1} dx,$$

where c is an arbitrary constant.

To complete the proof, it suffices to calculate the terms λ_0, λ_1 and λ_2 in (13).

Compute $\lambda_0\left(z\right)$:

We have

$$1 - H\left(z\right) = \left(1 - F\left(z\right)\right)\left(1 - G\left(z\right)\right)$$

where $F\left(z\right) = 1 - z^{-\alpha_1}$ and $G\left(z\right) = 1 - z^{-\alpha_2}$, therefore $H\left(z\right) = 1 - z^{-\alpha_2-\alpha_1}$.

The subdivisions functions are equal to

$$\begin{aligned}\tilde{H}_0\left(z\right) = P\left(Z \le z, \delta = 0\right) &= \int_1^z \left(1 - F\left(y\right)\right) dG\left(y\right) \\ &= \frac{\alpha_2}{\alpha_1+\alpha_2}\left(1 - z^{-\alpha_2-\alpha_1}\right),\end{aligned}$$

and

$$\tilde{H}_1(z) = \frac{\alpha_1}{\alpha_1+\alpha_2}\left(1 - z^{-\alpha_2-\alpha_1}\right),$$

then

$$\lambda_0(z) = \exp\left(\int_1^z \frac{d\tilde{H}_0(x)}{1-H(x)}\right) = z^{\alpha_2}. \tag{14}$$

To calculate $\lambda_1(z)$ there are two cases:

1°) If $\exp(\dfrac{v+1}{\alpha_1}) \le z < \infty$, then

$$\begin{aligned}\lambda_1(z) &= z^{\alpha_2}\left[\alpha_1 \ln z - \omega\right] \\ &=: \lambda_{11}(z),\end{aligned}$$

2°) If $1 \le z \le \exp(\dfrac{v+1}{\alpha_1})$, then

$$\begin{aligned}\lambda_1(z) = \frac{z^{\alpha_2}}{\alpha_1}(v-\omega)\left[1 - z^{\alpha_1}\exp\left(-(v+1)\right)\right] \\ + \lambda_{11}\left(\exp\left(\frac{v+1}{\alpha_1}\right)\right) =: \lambda_{12}(z)\end{aligned}$$

where

$$\lambda_{11}\left(\exp\left(\frac{v+1}{\alpha_1}\right)\right) = \left[v+1-\omega\right]\exp\left(\frac{v+1}{\alpha_1}\alpha_2\right).$$

Therefore

$$\lambda_1(z) = \begin{cases} \lambda_{11}(z), \text{ if } z \ge \exp\left(\dfrac{v+1}{\alpha_1}\right), \\ \lambda_{12}(z), \text{ if } 1 \le z \le \exp(\dfrac{v+1}{\alpha_1}). \end{cases} \tag{15}$$

Calculate $\lambda_2(z)$:

We have

$$\begin{aligned}\lambda_2(z) &= \int_1^z \left[\int_y^\infty \psi_v(x,\alpha_1)\lambda_0(x)\, d\tilde{H}_1(x)\right] \frac{d\tilde{H}_0(y)}{(1-H(y))^2} \\ &= \int_1^z \Phi(y)\frac{d\tilde{H}_0(y)}{(1-H(y))^2},\end{aligned}$$

where

$$\Phi(y) = \int_y^\infty \psi_v(x,\alpha_1)\lambda_0(x)\, d\tilde{H}_1(x).$$

Compute $\Phi(y)$:

$$\Phi(y)=\begin{cases}\Phi_1(y)=y^{-\alpha_1}(\alpha_1\ln y-\omega) & \text{if } \exp(\dfrac{v+1}{\alpha_1})\le z<\infty\\ \Phi_2(y)=\dfrac{v-\omega}{\alpha_1}(y^{-\alpha_1}-\exp-(v+1)) & \text{if } 1\le z\le \exp(\dfrac{v+1}{\alpha_1}).\\ +(v+1-\omega)\exp-((v+1)) & \end{cases}$$

1°) If $z\ge\exp\left(\dfrac{v+1}{\alpha_1}\right)$ then

$$\begin{aligned}\lambda_2(z)&=\int_1^{\exp\left(\frac{v+1}{\alpha_1}\right)}\Phi_2(y)\frac{d\tilde{H}_0(y)}{(1-H(y))^2}\\&+\int_{\exp\left(\frac{v+1}{\alpha_1}\right)}^{z}\Phi_1(y)\frac{d\tilde{H}_0(y)}{(1-H(y))^2}\\&:=\lambda_{21}+\lambda_{22}(z).\end{aligned}$$

Firstly

$$\begin{aligned}\lambda_{21}&=\int_1^{\exp\left(\frac{v+1}{\alpha_1}\right)}\Phi_2(y)\,\alpha_1 y^{\alpha_1+\alpha_2-1}dy\\&=\frac{v-\omega}{\alpha_1}\left(-1+\exp\left(\frac{v+1}{\alpha_1}\alpha_2\right)\right)+C_1,\end{aligned}$$

where

$$\begin{aligned}C_1=&\frac{\alpha_2}{\alpha_2+\alpha_1}\left(v-\omega+1-\frac{v-\omega}{\alpha_1}\right)\exp(-(v+1))\\&\times\left(-1+\exp\frac{v+1}{\alpha_1}(\alpha_1+\alpha_2)\right).\end{aligned}$$

Similarly

$$\begin{aligned}\lambda_{22}(z)&=\int_{\exp\left(\frac{v+1}{\alpha_1}\right)}^{z}\Phi_1(y)\frac{d\tilde{H}_0(y)}{(1-H(y))^2}\\&=\alpha_1\left[z^{\alpha_2}\left(\ln z-\frac{1}{\alpha_2}\right)-e^{\frac{v+1}{\alpha_1}\alpha_2}\left(\frac{v+1}{\alpha_1}-\frac{1}{\alpha_2}\right)\right]\\&-\omega\left(z^{\alpha_2}-\exp\left(\frac{v+1}{\alpha_1}\alpha_2\right)\right)\end{aligned}$$

2°) If $1\le z\le\exp\left(\dfrac{v+1}{\alpha_1}\right)$ then

$$\begin{aligned}\lambda_2(z)&=\int_1^z\Phi_2(y)\frac{d\tilde{H}_0(y)}{(1-H(y))^2}\\&=\frac{v-\omega}{\alpha_1}(z^{\alpha_2}-1)+\frac{\alpha_2}{\alpha_2+\alpha_1}\left(z^{\alpha_1+\alpha_2}-1\right)\\&\times\left[\left(v-\omega+1-\frac{v-\omega}{\alpha_1}\right)(\exp(-(v+1))\right]\\&=:\lambda_{23}(z).\end{aligned}$$

Then

$$\lambda_2(z) = \begin{cases} \lambda_{21} + \lambda_{22}(z), & \text{if } z \geq \exp\left(\dfrac{v+1}{\alpha_1}\right), \\ \lambda_{23}(z), & \text{if } 1 \leq z \leq \exp(\dfrac{v+1}{\alpha_1}). \end{cases} \tag{16}$$

We conclude, from (14), (15), (16) and (13) that

$$\sigma^2 = \begin{cases} Var((\omega - \alpha_1 \ln z + 1) z^{\alpha_2}\delta + \lambda_{11}(z)(1-\delta) \\ \qquad - (\lambda_{21} + \lambda_{22}(z))), & \text{if} \quad z \geq \exp\left(\dfrac{v+1}{\alpha_1}\right), \\ Var((\omega - v) z^{\alpha_2}\delta + \lambda_{12}(z)(1-\delta) - \lambda_{23}(z)), & \text{if} \quad 1 \leq z \leq \exp\left(\dfrac{v+1}{\alpha_1}\right). \end{cases}$$

Acknowledgements

The authors are grateful to anonymous referee and to Professor Necir Abdelhakim for their valuable comments and suggestions that helped to improve the quality of this paper.

References

Bacro, J.N. and Brito, M. 1995. Weak limiting behaviour of a simple tail Pareto-index estimator. *J. Statist. Plann. Inference* 45, 7-19.

Barnett, V. and Lewis, T., 1995. *Outliers in statistical data.* Third ed. John Willey & Sons. New York.

Beirlant, J., Guillou, A., Dierckx, G. and Fils-Villetard, A., 2007. Estimation of the extreme value index and extreme quantiles under random censoring. *Extremes.* **10**, no. 3, 151-174.

Beran, J. and Schell, D. 2012. On robust tail index estimation. *Comput. Statist. Data Anal.* **56**(11), 3430-3443.

Brahimi, B., Meraghni D. and Necir, A., 2014. Approximations to the tail index estimator of a heavy-tailed distribution under random censoring and application. *Extremes.* submitted.

Csörgő, S. and Viharos, L. 1998. Estimating the tail index. In Asymptotic Methods in Probability and Statistics, ed. B. Szyszkowicz. North Holland, Amesterdam, pp. 833-881.

Dekkers, A. L. M., Einmahl, J.H.J. and de Haan, L. 1989. A moment estimator for the index of an extreme-value distribution. Ann. Statist. 17, 1833-1855.

Einmahl, J.H.J., Fils-Villetard, A. and Guillou, A., 2008. Statistics of extremes under random censoring. *Bernoulli.* **14**, no.1, 207-227.

Fisher, R.A. and Tippett, L.H.C. 1928. Limiting forms of the frequency

Haeusler, E. and Teugels, J.L. 1985. On asymptotic normality of Hill's estimator for the exponent of regular variation. Ann. Statist. 13, 743-756.

de Haan, L. and Stadtmüller, U., 1996. Generalized regular variation of second order. *J. Australian Math. Soc.* (Series A) **61**, 381-395.

de Haan, L. and Ferreira, A., 2006. Extreme Value Theory: An Introduction. *Springer.*

Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J. and Stahel, W.A. 1986. *Robust statistics. The approach based on influence functions. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics.* John Wiley & Sons, Inc., New York.

Hill, B.M. 1975. A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3**, no.5. 1163-1174.

Huber, P.J. 1981. *Robust Statistics.* John Wiley and Sons, New York.

Kaplan, E.L. and Meier, P. 1958. Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* 53, 457-481.

Mason, D.M., 1982. Laws of large numbers for sums of extreme values. *Ann. Probab.* **10**, 756-764.

Reiss, R.D. and Thomas, M. 1997. *Statistical analysis of extreme values. From insurance, finance, hydrology and other fields.* With 1 Windows CD-ROM. Birkhäuser Verlag, Basel.

Stute, W. 1995. The central limit theorem under random censorship. *Ann. Statist.* **23**, 422-439.

Stute, W. 1994. The bias of Kaplan-Meir integrals. *Scand. J. Statist.* **21**, 475-484.

Stute, W. and Wang, J.L. 1993. The strong law under random censorchip. *Ann. Statist.* **21**, 1591-1607.