



Estimation of the extreme value index for non necessarily identically distributed observations

Laurent Gardes

► To cite this version:

Laurent Gardes. Estimation of the extreme value index for non necessarily identically distributed observations. 2014. hal-01075824v1

HAL Id: hal-01075824

<https://hal.science/hal-01075824v1>

Preprint submitted on 20 Oct 2014 (v1), last revised 3 Jun 2015 (v3)

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.

Estimation of the extreme value index for non necessarily identically distributed observations

Laurent Gardes

Université de Strasbourg & CNRS, IRMA, UMR 7501, 7 rue René Descartes,
67084 Strasbourg cedex, France.

Abstract

It is well known that the tail behavior of a survival function S is controlled by the so-called extreme value index. The aim of this paper is the estimation of this extreme value index in the case where the observations are independent but not necessarily distributed from the same distribution S . The proposed estimator only depends on a nonparametric estimator of the survival function S and can be applied to many different situations such as conditional extremes and heteroscedastic extremes. Its consistency is established under general conditions and its finite sample behavior is investigated through a simulation study.

Keywords. Extreme value index, non-identically distributed, conditional extremes, heteroscedastic extremes.

AMS Subject Classifications. 62G05; 62G20; 62G30; 62G32.

1 Introduction

In various applications, the behavior of large values (instead of central values) of a random variable can be of high interest. For instance, in climatology, the study of high temperatures or important precipitations is a key point to understand the effect of global warming. To make inference on the tail of a distribution characterized by its survival function S , the common departure point is to assume the existence of two sequences $\alpha_n > 0$ and β_n and of a non degenerate cumulative distribution function H such that

$$\lim_{n \rightarrow \infty} nS(\alpha_n y + \beta_n) = -\log H(y). \quad (1)$$

A result established by Fisher and Tippet [11] and Gnedenko [20] ensures that H is necessarily of the form $H_\gamma(y) := \exp[-(1 + \gamma y)_+^{-1/\gamma}]$ where $(\cdot)_+$ and $(\cdot)_-$ are respectively the positive and negative part functions. The parameter $\gamma \in \mathbb{R}$ is called the extreme value index of S . If limit (1) holds, we say that S belongs to the maximum domain of attraction (*MDA*) of the extreme value distribution H_γ . The class of survival functions that belongs to the *MDA* of H_γ is denoted $\mathcal{DA}(H_\gamma)$. Clearly the knowledge of γ brings us important information on the tail distribution. For this reason, the estimation of the extreme value index has been widely studied essentially in the situation where observations of n independent copies of a random value with survival function S are recorded. The most notable estimator is probably the Hill's estimator [24] dedicated to heavy tailed distributions (*i.e.* when $\gamma > 0$). For the general case $\gamma \in \mathbb{R}$, one can cite the estimator proposed by Pickands [27] and the moment estimator defined by Dekkers *et al.* [6].

In this paper, the estimation of the extreme value index is addressed in a more general setting. More precisely, we consider the situation where we have at our disposal observations of independent positive random variables Z_1, \dots, Z_n but not necessarily with the same distribution S . We only assume the existence of a stochastic process $\{\hat{S}_n(y) := T_n(y, Z_1, \dots, Z_n), y \geq 0\}$ (where T_n is a known function) which is a consistent estimator (in a sense that will be specified later) of the survival function S . The aim of this paper is to estimate the extreme value index by a function on this stochastic process. There are many practical situations where the estimation of the extreme value index must be done without having observations of a random sample from S . We list some examples below that will be more detailed in Section 3.

1) The first example is the situation of heteroscedastic extreme as introduced by Einmahl *et al.* [10] and de Haan *et al.* [23]. In numerous applications, the observed data are not generated from the same distribution. This is the case for instance in hydrology where the distribution of the amount of rain can differ from a season to another. To handle this kind of data, a model called *heteroscedastic extremes* was introduced in [10] and [23]. This model is defined as follows. Let $-\infty < b_1 \leq x_{n,1} < \dots < x_{n,n} \leq b_2 < \infty$ be deterministic points (for instance, one can see these points as time points). For every point $x_{n,i}$, we observe the realization of a positive random variable $Z_{n,i}$ with survival function $S_{x_{n,i}}$. It is assumed that for all $i = 1, \dots, n$, the survival function $S_{x_{n,i}}$ is asymptotically proportional to a given survival function $S \in \mathcal{DA}(H_\gamma)$, $\gamma \in \mathbb{R}$ (more details are given in Section 3). Hence, all the random variables share the same extreme value index but have different scales in the tails and the estimation of γ must be done by using the independent but not identically distributed random variables $Z_{n,1}, \dots, Z_{n,n}$.

The two following examples deal with the estimation of the extreme value index in presence of a covariate (random or deterministic).

2) In a regression context, a positive random variable Y of interest is linked with a covariate $X \in \mathbb{R}^p$ bringing an additional information on Y . This context can be encountered in various situations like insure, finance, climatology to name but a few. For a fixed value $x_0 \in \mathbb{R}^p$, we focus on the conditional distribution of Y given $X = x_0$. The survival function of interest is then given for all $y \geq 0$ by $S(y|x_0) = \mathbb{P}(Y > y|X = x_0)$ and is supposed to belong to the maximum domain of attraction of $H_{\gamma(x_0)}$. We are interested in the estimation of $\gamma(x_0)$ using observations of n independent copies $Z_1 = (X_1, Y_1), \dots, Z_n = (X_n, Y_n)$ of Z . The difficulty here is that we do not have at our disposal independent and identically distributed observations from $S(\cdot|x_0)$ to estimate $\gamma(x_0)$. In the literature, the function $\gamma(\cdot)$ is referred to as the conditional extreme value index. Its estimation has already been considered in [3] and [15] with an estimator inspired from the Refined Pickands estimator proposed by Drees [8], in [21] and [30] with an adaption of the moment estimator [6] and a maximum likelihood approach was considered by Wang and Tsai [32]. In the particular case of a positive conditional extreme value index, Gardes and Stupfler [16] propose an adaption of the Hill estimator.

3) In some applications, the covariate is deterministic and belongs to a set $E \subset \mathbb{R}^p$. This is the case for instance in hydrology where the covariate is the geographical position of a rain-gauge station and the variable of interest is the amount of rain. In this situation, one observe positive independent random variables $Z_{n,1}, \dots, Z_{n,n}$ with respective survival function $S_{x_{n,i}}$ for $i \in \{1, \dots, n\}$ where $x_{n,i} \in E$ is the value of the covariate associated to the i -th observation. The survival function of interest is, for a fixed $x_0 \in E$, given by $S_{x_0} \in \mathcal{DA}(H_{\gamma(x_0)})$ and our goal is the estimation of $\gamma(x_0)$. This problem has already been addressed by Gardes and

Girard using a moving window method [13] and a nearest neighbor approach [14] but only for conditional heavy-tailed distributions (*i.e.* for a positive function γ). One can also cite the papers of [5] and [28] using a regression model, Chavez-Demoulin and Davison [2] for a method using splines and Davison and Ramesh [4] for a local polynomial estimator.

These three situations are clearly not the only ones where the extreme value index should be estimated without a sample of independent and identically distributed random variables. One can also think for example on the estimation of γ under random censoring (see Einmahl *et al.* [9]) or for truncated data (see Gardes and Stupfler [17]). The rest of the paper is organized as follows. In Section 2, the definition of our extreme value index estimator is given and its consistency is established under a general setting. The three specific models introduced in this introduction are treated in details in Section 3. The finite sample behavior of the proposed estimator is illustrated through a simulation study in Section 4 and a short conclusion is given in Section 5. Section 6 is devoted to the proofs.

2 Estimation of the extreme value index

Consider a survival function $S \in \mathcal{DA}(H_\gamma)$, $\gamma \in \mathbb{R}$. The aim of this section is to estimate the extreme value index γ by a function of a given nonparametric estimator of S . As we will see in Section 3, this procedure of estimation allows us to consider situation where non necessarily identically distributed random variables are observed. In the rest of the paper, it is assumed that $S(y) = 0$ for all $y < 0$ (*i.e.* S is associated to a positive random variable). Denoting by $S^\leftarrow(u) = \inf\{t, S(t) \leq u\}$ the right-continuous inverse of S , it is well known (see [22, Theorem 1.1.6]) that the property $S \in \mathcal{DA}(H_\gamma)$ is equivalent to

$$\lim_{\alpha \rightarrow 0} \frac{S^\leftarrow(u\alpha) - S^\leftarrow(\alpha)}{a(\alpha^{-1})} - L_\gamma(1/u) \rightarrow 0, \quad (2)$$

for all $u \in (0, 1]$ where a is a positive function and for all $v \geq 1$ and $s \in \mathbb{R}$, $L_s(v) = \int_1^v u^{s-1} du$. In the rest of the paper, a survival function S satisfying (2) is said to be of extended regular variation with index $\gamma \in \mathbb{R}$ and positive auxiliary function a (we write $S \in \mathcal{ERV}(\gamma, a)$). The class of extended regular variation functions was also introduced in [22, Definition B.2.3] for instance. Note that the auxiliary function a satisfies $a(\alpha^{-1})/S^\leftarrow(\alpha) \rightarrow \gamma_+$ as $\alpha \rightarrow 0$ (see for instance [12, Lemma 3.1]). Furthermore, it is shown in Lemma 2 that convergence (2) is uniform on $u \in [\eta, 1]$ for all $\eta \in (0, 1)$. In paragraph 2.1, our estimator of γ is defined and its consistency is established in paragraph 2.2.

2.1 Definition of the estimator

Let S be a survival function of class $\mathcal{ERV}(\gamma, a)$. Assume that we have at our disposal an estimator $\{\tilde{S}_n(y), y \geq 0\}$ of the survival function S where \tilde{S}_n is a non-increasing and right-continuous function. The estimator of γ proposed in this paper is based on the following statistics:

$$H_{\eta, \varphi}^{(\delta)}(\tilde{S}_n) := \int_\eta^1 \varphi(u) \left(\log \frac{\tilde{S}_n^\leftarrow(u\alpha_n)}{\tilde{S}_n^\leftarrow(\alpha_n)} \right)^\delta du \Bigg/ \left(\int_\eta^1 \varphi(u) L_0(1/u) du \right)^\delta, \quad (3)$$

where $\eta \in (0, 1)$, $\delta \in \mathbb{N}$ and φ is a positive and bounded function on $(\eta, 1)$. In order to define the estimator of γ , the following function is introduced: for $s \leq 0$ let

$$\Psi_{\eta, \varphi}(s) := \left(\int_\eta^1 \varphi(u) L_s(1/u) du \right)^2 \Bigg/ \int_\eta^1 \varphi(u) L_s^2(1/u) du. \quad (4)$$

It is shown in Lemma 1 that $\Psi_{\eta,\varphi}$ is a decreasing function on $(-\infty, 0]$ and thus its inverse $\Psi_{\eta,\varphi}^{\leftarrow}$ is well defined. We propose to estimate the extreme value index $\gamma \in \mathbb{R}$ by the function of $\{\tilde{S}_n(y), y \geq 0\}$ defined below:

$$\gamma_{\eta,\varphi}(\tilde{S}_n; \alpha_n) := H_{\eta,\varphi}^{(1)}(\tilde{S}_n) + \Psi_{\eta,\varphi}^{\leftarrow} \left(\max \left\{ \frac{[H_{\eta,\varphi}^{(1)}(\tilde{S}_n)]^2}{H_{\eta,\varphi}^{(2)}(\tilde{S}_n)}, \Psi_{\eta,\varphi}(0) \right\} \right). \quad (5)$$

Let us highlight that in the simplest situation where independent random variables Z_1, \dots, Z_n with common survival function $S \in \mathcal{DA}(H_\gamma)$ are recorded, the natural choice for \tilde{S}_n is the empirical distribution function for which, $\tilde{S}_n^{\leftarrow}(\alpha) = Z_{(n-\lfloor n\alpha \rfloor)}$ for $\alpha \in (0, 1)$, where $\lfloor x \rfloor$ denotes the integer part of x and $Z_{(1)} \leq \dots \leq Z_{(n)}$ are the ordered statistics associated to the sample Z_1, \dots, Z_n . Hence, in the limit case $\eta = 0$ (not allowed here) and if $\varphi = \text{cst}$ is a constant function, $H_{0,\text{cst}}^{(\delta)}(\tilde{S}_n) = \lfloor n\alpha_n \rfloor / (n\alpha_n) \mathcal{H}_n^{(\delta)} = \mathcal{H}_n^{(\delta)}(1 + o(1))$ as $n \rightarrow \infty$ where

$$\mathcal{H}_n^{(\delta)} := \frac{1}{\lfloor n\alpha_n \rfloor} \sum_{i=1}^{\lfloor n\alpha_n \rfloor} \left(\log \frac{Z_{(n-i+1)}}{Z_{(n-\lfloor n\alpha_n \rfloor)}} \right)^\delta.$$

Since $\Psi_{0,\text{cst}}^{\leftarrow}(s) = 1 - 1/2(1-s)^{-1}$, the estimator $\gamma_{0,\text{tilde}}(\tilde{S}_n; \alpha_n)$ is thus, in presence of independent and identically distributed random variables, asymptotically equivalent to the well known moment estimator introduced by Dekkers *et al.* [6] and given by

$$\hat{\gamma}_M(\alpha_n) := \mathcal{H}_n^{(1)} - 1 - \frac{1}{2} \left(1 - \frac{(\mathcal{H}_n^{(1)})^2}{\mathcal{H}_n^{(2)}} \right)^{-1}.$$

Our estimator (5) can thus be seen as a weighted version of the moment estimator adapted to more general settings.

2.2 Consistency

We are now interested in proving the consistency of the estimator defined in (5). As usual in extreme value theory, second-order condition is required in order to precise the rate of convergence in (2). In what follows, for sequences τ_n and α_n converging to 0, a survival function S is said to be of class $\mathcal{ERV}_2(\gamma, a, \tau_n, \alpha_n)$ if $S \in \mathcal{ERV}(\gamma, a)$ with

$$\lim_{n \rightarrow \infty} \tau_n^{-1} \max \left\{ \sup_{u \in (\eta, 1)} \left| \frac{S^{\leftarrow}(u\alpha_n) - S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} - L_\gamma(1/u) \right|, \left| \frac{a(\alpha_n^{-1})}{S^{\leftarrow}(\alpha_n)} - \gamma_+ \right| \right\} = 0.$$

Note that the class $\mathcal{ERV}_2(\gamma, a, \tau_n, \alpha_n)$ of second-order extended regular varying functions is considered in most of the papers dealing with the estimation of $\gamma \in \mathbb{R}$ (see for instance [6], [21], [30]). Our main result is given below.

Theorem 1. *Let (τ_n) and (α_n) be sequences converging to 0 as $n \rightarrow \infty$. Let $\{\tilde{S}_n(y), y \geq 0\}$ be a stochastic process such that*

$$\frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \tau_n^{-1} \sup_{u \in (\eta, 1)} \left| \frac{\tilde{S}_n^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(u\alpha_n)} - 1 \right| = \mathcal{O}_{\mathbb{P}}(1). \quad (6)$$

If $S \in \mathcal{ERV}_2(\gamma, a, \tau_n, \alpha_n)$ then $|\gamma_{\eta,\varphi}(\tilde{S}_n; \alpha_n) - \gamma| = \mathcal{O}_{\mathbb{P}}(\tau_n)$.

The previous result ensures that for every uniformly consistent estimator of S , the extreme value index estimator defined by (5) is consistent. We now focus on a particular family of estimators of the survival function S .

$$\hat{S}_n(y) = \sum_{i=1}^n R_{n,i}(y) \bigg/ \sum_{i=1}^n R_{n,i}(0), \quad (7)$$

where $\{R_{n,i}(y), y \geq 0\}$, $i = 1, \dots, n$ are n independent stochastic processes with $R_{n,i}$ positive, non-increasing and right-continuous and such that $\sup\{R_{n,i}(y), y \geq 0\} \leq C_R$ almost surely for some positive constant C_R . Note that, as a survival function, \hat{S}_n is a $[0, 1]$ -valued function which is right-continuous and non-increasing. Clearly this family of estimators encompasses the empirical distribution function and all the estimators that are considered in Section 3 can be written as in (7). Before establishing the consistency of the associated estimator of γ , the following notations are introduced: for a sequence (α_n) converging to 0,

$$y_{\kappa_n}(u; \alpha_n) := S^{\leftarrow}(u\alpha_n) + \kappa_n a(\alpha_n^{-1}) \text{ and } \mu_{\kappa_n}(u; \alpha_n) := \mathbb{E} \left(\sum_{i=1}^n R_{n,i}(y_{\kappa_n}(u; \alpha_n)) \right),$$

where (κ_n) is a sequence converging to 0. The consistency of $\gamma_{\eta, \varphi}(\hat{S}_n; \alpha_n)$ is established in the following result.

Theorem 2. *Let (σ_n) and (α_n) be sequences converging to 0 as $n \rightarrow \infty$. Let $\{\hat{S}_n(y), y \geq 0\}$ be a stochastic process defined as in (7). For $\tau_n = \sigma_n(\log(\sigma_n^{-1}))^{1/2}$, if $S \in \mathcal{ERV}_2(\gamma, a, \tau_n, \alpha_n)$ and if there exists a constant $r_0 > 0$ such that for every sequence $\kappa_n \rightarrow 0$,*

$$\sup_{u \in (\eta, 1)} \left| \frac{r_0 \alpha_n \sigma_n^2 \mu_{\kappa_n}(u; \alpha_n)}{S(y_{\kappa_n}(u; \alpha_n))} - 1 \right| = o(\tau_n) \text{ and } \sup_{u \in (\eta, 1)} \left| \alpha_n \sigma_n^2 \sum_{i=1}^n R_{n,i}(0) - r_0 \right| = o_{\mathbb{P}}(\tau_n), \quad (8)$$

then $|\gamma_{\eta, \varphi}(\hat{S}_n; \alpha_n) - \gamma| = \mathcal{O}_{\mathbb{P}}(\tau_n)$.

To prove this result, it suffices to show that under the assumptions of Theorem 2, the right-continuous inverse of the estimator \hat{S}_n given in (7) is a uniform consistent estimator of the quantile function S^{\leftarrow} i.e. such that (6) holds. In the next section, some examples where an estimator \hat{S}_n satisfying the assumptions of Theorem 2 can be constructed are presented.

3 Applications

3.1 Heteroscedastic extremes

In various applications, the observed data do not generally satisfy the independence assumption and/or are not identically distributed. This is obviously the case for instance when we are interested in the study of daily rainfall over time. To handle the case when observations are independent but drawn from different distributions, Einmahl *et al.* [10] and de Haan *et al.* [23] assume that the distribution tails are proportional, situation which is called by authors *heteroscedastic extremes*. The framework considered here is mostly the same as the one considered in [10]. For $-\infty < b_1 < b_2 < \infty$, let $\{S_x, x \in [b_1, b_2]\}$ be a family of survival functions satisfying the heteroscedasticity condition given by:

(M) For all $x \in [b_1, b_2]$, the survival functions S_x share the same right endpoint z^* . Furthermore, there exist a survival function S in the maximum domain of attraction of H_γ , $\gamma \in \mathbb{R}$ with $S^{\leftarrow}(0) = z^*$ and a continuous positive function c defined on $[b_1, b_2]$ such that:

$$\lim_{z \rightarrow z^*} \sup_{x \in [b_1, b_2]} \left| \frac{S_x(z)}{S(z)} - c(x) \right| = 0 \text{ with } \int_a^b c(s) ds < \infty.$$

Note that the survival function S and the function c in (M) are not uniquely defined. To ensure unicity, we impose from now on that $\int_{b_1}^{b_2} c(x) dx = b_2 - b_1$.

For every deterministic points $-\infty \leq b_1 < x_{n,1} < \dots < x_{n,n} \leq b_2 < \infty$, we observe a realization of a positive random variable $Z_{n,i}$ with survival function $S_{x_{n,i}}$. The random variables $Z_{n,1}, \dots, Z_{n,n}$ are supposed positive and independent. Let us highlight that in [10], the points

$\{x_{n,i}, i = 1, \dots, n\}$ are assumed to be regularly distributed on $[0, 1]$ (i.e. $x_{n,i} = i/n$). This assumption can be too restrictive for an application purpose like for instance in hydrology since the times for which a certain non null amount of rain is observed are clearly not regularly distributed. In this paper, letting $x_{n,0} = b_1$ and $x_{n,n+1} = b_2$, it is only assumed that

(A.1) there exists $d > 0$ such that $\max\{x_{n,i} - x_{n,i-1}, i = 1, \dots, n+1\} \leq d/n$.

Under model **(M)**, it is easy to see that for all $i \in \{1, \dots, n\}$, $S_{x_{n,i}}$ is in the maximum domain of attraction of H_γ . Thus, the random variables $Z_{n,1}, \dots, Z_{n,n}$ share the same extreme value index but differ asymptotically in scale. As a consequence, the associated quantiles depend on the value of the function c and thus, the model of heteroscedastic extremes permits us to take into account the effect of the point of observation (for example the time) on extremal events. As mentioned in the introduction we focus in this paper on the estimation of γ . To reach this goal, only an estimator of the survival function S is required. We propose here to use the estimator

$$\hat{S}_n^{(1)}(z) = \sum_{i=1}^n \frac{x_{n,i} - x_{n,i-1}}{x_{n,n} - b_1} \mathbb{I}_{\{Z_{n,i} > z\}}, \quad (9)$$

where $\mathbb{I}_{\{\cdot\}}$ is the indicator function. Clearly, $\hat{S}_n^{(1)}$ is of the form (7) with for $i \in \{1, \dots, n\}$

$$R_{n,i}(z) =: R_{n,i}^{(1)}(z) = n \frac{x_{n,i} - x_{n,i-1}}{x_{n,n} - b_1} \mathbb{I}_{\{Z_{n,i} > z\}} \leq \frac{2d}{b_2 - b_1}, \quad (10)$$

under condition **(A.1)**. An asymptotic property of the associated extreme value index estimator $\gamma_{\eta,\varphi}(\hat{S}_n^{(1)}; \alpha_n)$ is established in the following result.

Corollary 1. *For a sequence (α_n) converging to 0, let $\sigma_n = (n\alpha_n)^{-1/2}$ and $\tau_n = \sigma_n(\log(\sigma_n^{-1}))^{1/2}$. Under model **(M)**, suppose that condition **(A.1)** holds and that*

$$\sup \{|c(u) - c(u')|, (u, u') \in [b_1, b_2]^2 \text{ with } |u - u'| \leq d/n\} = o(\tau_n). \quad (11)$$

If $\sigma_n \rightarrow 0$, if there exist a positive function a_1 such that $S \in \mathcal{ERV}_2(\gamma, a_1, \tau_n, \alpha_n)$ and if there exists $\delta > 1$ such that

$$\sup \left\{ \left| \frac{S_x(z)}{S(z)} - c(x) \right|, x \in [b_1, b_2], \frac{z}{S^{\leftarrow}(\alpha_n)} \in [\delta^{-1}, \delta] \right\} = o(\tau_n), \quad (12)$$

then, $|\gamma_{\eta,\varphi}(\hat{S}_n^{(1)}; \alpha_n) - \gamma| = \mathcal{O}_{\mathbb{P}}(\tau_n)$.

Conditions (11) and (12) are very similar to the ones used in [10]. Condition (11) is a regularity condition on the function c involved in model **(M)**. It is satisfied for instance if the function c is Lipschitz continuous of order at least $1/2$. Condition (11) is also a regularity condition but on the function $S_x(z)$ considered as a function of x . The sequence (α_n) represents the proportion of largest observations used in the estimation procedure and thus $k_n := n\alpha_n$ is the number of kept observations. Conditions $\alpha_n \rightarrow 0$ and $n\alpha_n \rightarrow \infty$ (or equivalently $k_n/n \rightarrow 0$ and $k_n \rightarrow \infty$) are standard hypothesis for the estimation of the extreme value index. Finally note that in Corollary 1 (and also in the two following corollaries), we do not find the classical rate of convergence in extreme value theory (namely $(n\alpha_n)^{1/2}$). This is mainly due to the fact that Corollary 1 requires the uniform consistency result (6). Using others theoretical tools in the proof of Corollary will probably permit us to obtain the classical rate of convergence but this is beyond the scope of this paper.

3.2 Conditional extreme value index estimation for a random covariate

This example takes place in a regression context where a positive response variable Y is measured with a random explanatory variable $X \in \mathbb{R}^p$. More specifically, we are interested in

the influence of X on the large values of the response variable Y or, in other words, on the conditional tail distribution of Y given X . This can be of high interest in many applications like insurance [1], finance [31], climatology [28] to name a few. In what follows, we consider n independent copies $Z_1 = (X_1, Y_1), \dots, Z_n = (X_n, Y_n)$ of a random vector $Z = (X, Y) \in \mathbb{R}^p \times [0, \infty)$. The probability density function of X is denoted by g and it is assumed that

(A.2) there exists a constant $c_g > 0$ such that $|g(x) - g(x')| \leq c_g \|x - x'\|_\infty$.

Note that the uniform norm was used in condition **(A.2)** but obviously, any norm on \mathbb{R}^p can be also considered. For a given $x_0 \in \mathbb{R}^p$ such that $g(x_0) > 0$, we suppose that the conditional distribution of Y given $X = x_0$ belongs to the maximum domain of attraction of $H_{\gamma(x_0)}$, $\gamma(x_0) \in \mathbb{R}$ and we are interested in the estimation of this extreme value index. To reach this goal, we consider the estimator of the survival function $S(\cdot|x_0) := \mathbb{P}(Y > y|X = x_0)$ defined by:

$$\hat{S}_n^{(2)}(y|x_0) = \sum_{i=1}^n \mathbb{I}_{\{Y_i > y\}} K_{H_n}(x_0 - X_i) \bigg/ \sum_{i=1}^n K_{H_n}(x_0 - X_i), \quad (13)$$

which is the kernel estimator introduced by Nadaraya [25] and Watson [33] where H_n is a positive-definite matrix. For the sake of shortness, we have introduced the notation $K_{H_n}(t) := |H_n|^{-1} K(H_n^{-1}t)$, $t \in \mathbb{R}^p$ where $K(\cdot)$ is called the kernel function and, for all square matrix M , $|M|$ denotes the determinant of M . In what follow, we consider kernel functions satisfying:

(A.3) K is a bounded density on \mathbb{R}^p with support \mathcal{U}_p , the unit ball of \mathbb{R}^p .

This condition is classical in local estimation (see for instance [3, 21]). It is easy to check that $\hat{S}_n^{(2)}(\cdot|x_0)$ is of the form (7) with

$$R_{n,i}(y) =: R_{n,i}^{(2)}(y|x_0) = |H_n| K_{H_n}(x_0 - X_i) \mathbb{I}_{\{Y_i > y\}} \leq \|K\|_\infty, \quad (14)$$

under **(A.3)**. We propose to estimate the extreme value index $\gamma(x_0)$ by the associated estimator $\gamma_{n,\varphi}(\hat{S}_n^{(2)}(\cdot|x_0); \alpha_n)$ (see equation (5)). Its asymptotic behavior is established in the next result. Before, the following notation is required: for a positive-definite matrix M of size p and $c \in \mathbb{R}^p$, $B(c, M) := \{x \in \mathbb{R}^p \mid \|M^{-1}(c - x)\|_\infty \leq 1\}$ denotes the ball of center c and radius M .

Corollary 2. *For a sequence (α_n) converging to 0, let $\sigma_n = (n|H_n|\alpha_n)^{-1/2}$ and $\tau_n = \sigma_n(\log(\sigma_n^{-1}))^{1/2}$. For $x_0 \in \mathbb{R}^p$ such that $g(x_0) > 0$, assume that **(A.2)** and **(A.3)** hold. If there exist of a positive function $a_2(\cdot|x_0)$ such that $S(\cdot|x_0) \in \mathcal{ERV}_2(\gamma(x_0), a_2(\cdot|x_0), \tau_n, \alpha_n)$, if $\sigma_n \rightarrow 0$, $\tau_n^{-1}\|H_n\|_\infty \rightarrow 0$ and if there exists $\delta > 1$ such that*

$$\sup \left\{ \left| \frac{S(y|x)}{S(y|x_0)} - 1 \right|, x \in B(x_0, H_n), \frac{y}{S^{\leftarrow}(\alpha_n|x_0)} \in [\delta^{-1}, \delta] \right\} = o(\tau_n), \quad (15)$$

then $|\gamma_{n,\varphi}(\hat{S}_n^{(2)}(\cdot|x_0); \alpha_n) - \gamma(x_0)| = \mathcal{O}_{\mathbb{P}}(\tau_n)$.

Note that the conditions used in the previous corollary are similar to the ones considered in [3] where the asymptotic normality of $(\hat{S}_n^{(2)})^{\leftarrow}(\cdot|x_0)$ is established. Recall that under the assumptions of Corollary 2, it can be shown that $(\hat{S}_n^{(2)})^{\leftarrow}(\cdot|x_0)$ is a uniform consistent estimator in the sense given by (6). The expected number of points kept for the estimation is given by $n|H_n|\alpha_n$ and thus conditions $\alpha_n \rightarrow 0$ and $n|H_n|\alpha_n \rightarrow \infty$ are classical in extreme value theory. Finally, condition (15) controls the oscillations of the conditional survival function $S(\cdot|x_0)$. An interesting discussion on this condition can be found in [30].

3.3 Conditional extreme value index estimation for a fixed design

In this last example, we are interested in large values of a response variable recorded with a nonrandom explanatory variable. A motivating example is the study of extreme rainfalls

at different geographical positions. In this paragraph, the following framework is considered: let E be a compact subset of \mathbb{R}^p , $p \in \mathbb{N}^*$ and $\{S_x, x \in E\}$ a family of survival functions. Associated to nonrandom covariates $x_{n,1} \in E, \dots, x_{n,n} \in E$, we have positive and independent random variables $Z_{n,1}, \dots, Z_{n,n}$, where for all $i \in \{1, \dots, n\}$, the survival function of $Z_{n,i}$ is given by $S_{x_{n,i}}$. For a given $x_0 \in \mathring{E}$ where \mathring{E} denotes the interior of E , it is assumed that $S_{x_0} \in \mathcal{DA}(H_{\gamma(x_0)})$ and our goal is the estimation of the extreme value index $\gamma(x_0) \in \mathbb{R}$ using estimator (5). To this end, only the estimation of the survival function S_{x_0} is required. We propose here to use the estimator introduced by Stone [29] with the Gasser and Müller [18] weights. For a positive-definite matrix H_n it is given by:

$$\hat{S}_{n,x_0}^{(3)}(z) = \sum_{i=1}^n \mathbb{I}_{\{Z_{n,i} > z\}} \int_{A_{n,i}} K_{H_n}(x_0 - t) dt, \quad (16)$$

where $A_{n,i}$ are sets that partition the subset E with $x_{n,i} \in A_{n,i}$ and where K_{H_n} is defined as in paragraph 3.2. The following notation will be used: for $i \in \{1, \dots, n\}$, let $\mathcal{V}_{n,i}$ the volume of $A_{n,i}$ and let $\bar{\mathcal{V}}_n := \max(\mathcal{V}_{n,1}, \dots, \mathcal{V}_{n,n})$. We assume that

(A.4) There exists a positive constant C_V such that $n\bar{\mathcal{V}}_n \leq C_V$.

Since E is a compact subset, this condition is reasonable and classical in nonparametric regression for the fixed design case (see for instance [19]). Now let

$$R_{n,i}(y) =: R_{n,i,x_0}^{(3)}(y) = n|H_n| \mathbb{I}_{\{Z_{n,i} > y\}} \int_{A_{n,i}} K_{H_n}(x_0 - t) dt. \quad (17)$$

Under **(A.3)** and **(A.4)**, $R_{n,i,x_0}^{(3)}(y) \leq n\|K\|_\infty \mathcal{V}_{n,i} \leq C_V\|K\|_\infty$ and, since $x_0 \in \mathring{E}$,

$$\int_E K_{H_n}(x_0 - t) dt = 1 \text{ and thus } \sum_{i=1}^n R_{n,i,x_0}^{(3)}(0) = n|H_n|.$$

Thus, estimator $\hat{S}_{n,x_0}^{(3)}$ admits the same form as (7). We now give an asymptotic property of the extreme value index estimator $\gamma_{\eta,\varphi}(\hat{S}_{n,x_0}^{(3)}; \alpha_n)$.

Corollary 3. *For a sequence (α_n) converging to 0, let $\sigma_n = (n|H_n|\alpha_n)^{-1/2}$ and $\tau_n = \sigma_n(\log(\sigma_n^{-1}))^{1/2}$. Under **(A.3)** and **(A.4)**, if there exist a positive function a_{3,x_0} such that $S_{x_0} \in \mathcal{ERV}_2(\gamma(x_0), a_{3,x_0}, \tau_n, \alpha_n)$, if H_n converges to the zero matrix, if $\sigma_n \rightarrow 0$ and if there exist $\delta > 1$ such that*

$$\sup \left\{ \left| \frac{S_x(z)}{S_{x_0}(z)} - 1 \right|, x \in B(x_0, 2H_n), \frac{z}{S_{x_0}^{-1}(\alpha_n)} \in [\delta^{-1}, \delta] \right\} = o(\tau_n), \quad (18)$$

then $|\gamma_{\eta,\varphi}(\hat{S}_{n,x_0}^{(3)}; \alpha_n) - \gamma(x_0)| = \mathcal{O}_{\mathbb{P}}(\tau_n)$.

The conditions in Corollary 3 are very similar to the ones of Corollary 2. The only difference is that condition $\tau_n^{-1}\|H_n\|_\infty$ is not required here since this condition was used in Corollary 2 to ensure the consistency of the probability density estimator.

Link with heteroscedastic extremes In the special case where $E = [b_1, b_2]$ and where the family of survival functions $\{S_x, x \in E\}$ satisfy the heteroscedasticity condition **(M)** (see Section 3.1) with an extreme value index γ then, under the assumptions of Corollary 3 and for all $x_0 \in [b_1, b_2]$, $\gamma_{\eta,\varphi}(\hat{S}_{n,x_0}^{(3)}; \alpha_n)$ and $\gamma_{\eta,\varphi}(\hat{S}_n^{(1)}; \alpha_n)$ are consistent estimators of γ . Hence, given a sample $Z_{n,1}, \dots, Z_{n,n}$ associated to nonrandom covariates $b_1 \leq x_{n,1} < \dots < x_{n,n} \leq b_2$, one way to check the validity of model **(M)** is to plot the ratio $\gamma_{\eta,\varphi}(\hat{S}_{n,x_0}^{(3)}; \alpha_n) / \gamma_{\eta,\varphi}(\hat{S}_n^{(1)}; \alpha_n)$ as a function of $x_0 \in [b_1, b_2]$. If this ratio seems to be very different from 1, it is unreasonable to accept the validity of model **(M)**. In presence of heavy-tailed distribution (*i.e.* $\gamma > 0$), a similar idea can be found in [10] where a statistical test for the validity of **(M)** is proposed.

This test is based on the ratio of two different Hill's estimators. Establishing a statistical test is beyond the scope of the present paper and will be part of a future work. If model **(M)** is assumed, the function c can be estimated using the statistics

$$\hat{c}_n(x_0) := \hat{S}_{n,x_0}^{(3)}(Z_{(n,n-\lfloor n\alpha_n \rfloor)}) \Big/ \hat{S}_n^{(1)}(Z_{(n,n-\lfloor n\alpha_n \rfloor)}) ,$$

where (α_n) is a sequence on $(0,1)$ converging to 0 and $Z_{(n,1)} \leq \dots \leq Z_{(n,n)}$ are the order statistics deduced from the sample $Z_{n,1}, \dots, Z_{n,n}$. Note that taking $A_{n,i} = (x_{n,i-1}, x_{n,i}]$ in the expression of $\hat{S}_{n,x_0}^{(3)}$, it is easy to check that

$$\int_{b_1}^{b_2} \hat{c}_n(x) dx = x_{n,n} - b_1 = b_2 - b_1 + \mathcal{O}(1/n).$$

In [10], a slightly different estimator of function c is proposed. The behavior of the estimator \hat{c}_n is briefly illustrated in the simulation study.

4 Simulations

The behavior of the estimators presented in this paper is examined through Monte Carlo simulations. The data are generated using one of the two following processes:

P1 - For $i = 1, \dots, n$, let $x_{n,i} = G(i/n)$ where G is the distribution function of a beta distribution with parameters $a = b = 2$. For a given function $\tilde{c} : [0, 1] \mapsto [1, \infty)$, we generate n independent random variables $Z_{n,1}, \dots, Z_{n,n}$ where the survival function of $Z_{n,i}$ is one of the three following:

P1-1: Heteroscedastic Fréchet distribution: for $z \geq 0$,

$$S_{x_{n,i}}(z) = 1 - \exp\left(-\frac{\tilde{c}(x_{n,i})}{z}\right).$$

P1-2: Heteroscedastic uniform distribution: for $z \in [1 - 1/\tilde{c}(x_{n,i}), 1]$,

$$S_{x_{n,i}}(z) = (1 - z)\tilde{c}(x_{n,i}).$$

P1-3: Heteroscedastic Weibull distribution: for $\theta \in (0, 1]$ and $z \geq \log \tilde{c}(x_{n,i})$,

$$S_{x_{n,i}}(z) = \exp\left[-(z - \log \tilde{c}(x_{n,i}))^\theta\right].$$

P2 - Let $(X_1, Y_1), \dots, (X_n, Y_n)$ be n independent copies of a random vector (X, Y) where X is uniformly distributed on $[0, 1]$ and, for a positive function γ defined on $[0, 1]$ and a function $\tilde{c} : [0, 1] \mapsto [1, \infty)$, the conditional survival function of Y given $X = x$ is one of the three following:

P2-1 Conditional Fréchet distribution: for $y \geq 0$,

$$S(y|x) = 1 - \exp(-\tilde{c}(x)y^{-1/\gamma(x)})$$

P2-2: Conditional beta distribution: let $G(\cdot|x)$ be the distribution function of a beta distribution with parameters $a = b = -1/\gamma(x)$, for $y \in [G^{\leftarrow}(1 - 1/\tilde{c}(x)|x), 1]$,

$$S(y|x) = (1 - G(y|x))\tilde{c}(x).$$

P2-3: Conditional Weibull distribution: for a function $\theta : [0, 1] \mapsto (0, 1]$ and $y \geq \log c(x)$,

$$S(y|x) = \exp\left[-(y - \log \tilde{c}(x))^{\theta(x)}\right].$$

Note that process **P1** corresponds to the situation of heteroscedastic extremes as described in Section 3.1. Indeed, the three survival functions given in **P1-1**, **P1-2** and **P1-3** satisfy

model **(M)** with $b_1 = 0$ and $b_2 = 1$ and where the common extreme value index and right endpoint are given by $\gamma = 1$ and $z^* = +\infty$ for **P1-1**, $\gamma = -1$ and $z^* = 1$ for **P1-2** and $\gamma = 0$ and $z^* = +\infty$ for **P1-3**. The function c is given by $\tilde{c}/\int_0^1 c(s)ds$ for processes **P1-1** and **P1-2**. For process **P1-3**, it is easy to check that if $\theta \in (0, 1)$, $c = 1$ and if $\theta = 1$, $c = \tilde{c}/\int_0^1 c(s)ds$.

Process **P2** corresponds to a conditional extreme value framework with a random covariate (see Section 3.2) where the conditional extreme value index is given by the function γ .

The rest of this section is organized as follows: first, the role of parameter η and of function φ in the estimator $\gamma_{n,\varphi}(\hat{S}_n; \alpha_n)$ are illustrated on data generated by process **P1**. Next, always for data generated by process **P1**, our estimator $\gamma_{n,\varphi}(\hat{S}_n^{(1)}; \alpha_n)$ defined in section 3.1 is compared to the classical moment estimator proposed by Dekkers *et al.* [6]. Recall that the consistency of moment estimator has been proved only in presence of independent and identically distributed random variables but, using similar techniques as in [10], the consistency must be also true for heteroscedastic extremes. Finally, with data generated by the process **P2**, we compare our estimator $\gamma_{n,\varphi}(\hat{S}_n^{(2)}; \alpha_n)$ defined in section 3.2 with the estimator proposed by Stupfler [30].

4.1 Influence of parameter η and of function φ

The estimator proposed in this paper mainly depends on a parameter $\eta \in (0, 1)$, a bounded function φ and a sequence (α_n) . Here, α_n is fixed to $n^{-1/3}$ and we focus on the influence that both η and φ have on the behavior of $\gamma_{n,\varphi}(\hat{S}_n^{(1)}; \alpha_n)$. We generate $N = 500$ samples of size $n \in \{50, 100, 200, 400\}$ using the generating processes **P1-1**, **P1-2** and **P1-3** with $\tilde{c}(\cdot) = 1 + \log(1/\cdot)$ and, for **P1-3**, with $\theta = 1$. For each values of $\eta \in \{0.005 \times 2^j, j = 0, 1, \dots, 5\}$ and for the two functions $\varphi = 1$ and $\varphi(\cdot) = \log(1/\cdot)$, the empirical mean squared error of the estimator $\gamma_{n,\varphi}(\hat{S}_n^{(1)}; \alpha_n)$ is computed. Recall that for any estimator $\hat{\gamma}_n$ of γ , denoting by $\hat{\gamma}_n^{(r)}$, $r = 1, \dots, N$ the values of $\hat{\gamma}_n$ observed on each replications, the empirical mean squared error of $\hat{\gamma}_n$ is given by:

$$\text{EMSE}(\hat{\gamma}_n) = \frac{1}{N} \sum_{r=1}^N (\hat{\gamma}_n^{(r)} - \gamma)^2.$$

The results are collected in Table 1 (for $\varphi = 1$) and Table 2 (for $\varphi(\cdot) = \log(1/\cdot)$). The bold numbers are the best results obtained for each values of n . Concerning the function φ , one can see that the choice $\varphi = 1$ provides slightly better results for the Fréchet distribution (*i.e* process **P1-1**) but, for the two other generating processes, the choice $\varphi(\cdot) = \log(1/\cdot)$ is clearly better (especially for the uniform distribution (**P1-2**)). For the choice of η , it seems that taking η small provides better results (except for the Fréchet distribution with $\varphi(\cdot) = \log(1/\cdot)$). Let us also highlight that taking $\eta = 0.02$ for each values of n and each generating process leads to mean squared errors quite close to the bold numbers. In conclusion, the choice $\eta = 0.02$ and $\varphi(\cdot) = \log(1/\cdot)$ seem reasonable and will be used in the rest of this simulation study.

4.2 Behavior of the estimator for heteroscedastic data

We generate $N = 500$ samples of size $n = 100$ using the generating processes **P1-1**, **P1-2** and **P1-3** with $\tilde{c} = 1$ and $\tilde{c}(\cdot) = 1 + \log(1/\cdot)$ and, for **P1-3**, with $\theta = 1$. Note that when $\tilde{c} = 1$, the observations are independent and identically distributed. In order to appreciate the effect of the sequence α_n , the estimator is computed for $\alpha_n \in \{n^{-1/a}, a = 2, \dots, 6\}$ while η is fixed to 0.02 and $\varphi = \log(1/\cdot)$. Using the empirical mean squared error, our estimator $\gamma_{n,\varphi}(\hat{S}_n^{(1)}; \alpha_n)$ is compared to the moment estimator $\hat{\gamma}_M(\alpha_n)$. More precisely, the ratio

$$R_1(\alpha_n) := \text{EMSE}\left(\gamma_{n,\varphi}(\hat{S}_n^{(1)}; \alpha_n)\right) / \text{EMSE}(\hat{\gamma}_M(\alpha_n)),$$

is computed for each values of α_n and each functions \tilde{c} . Clearly, a ratio lower than 1 means that our estimator provides better results (in term of mean squared error) than the moment

estimator. The results are presented in Table 3. It appears that, for the Fréchet distribution (process **P1-1**), the moment estimator is slightly better than our estimator but for the two other generating processes, our method provides better results in term of mean squared error. One can also notice that the function \tilde{c} has not a strong influence on the estimation of γ . As mentioned at the end of Section 3.3, the function $c = \tilde{c} / \int_0^1 c(s)ds$ can be estimated for every $x_0 \in [0, 1]$ by

$$\hat{c}_n(x_0) := \hat{S}_{n,x_0}^{(3)}(Z_{(n,n-\lfloor n\alpha_n \rfloor)}) / \hat{S}_n^{(1)}(Z_{(n,n-\lfloor n\alpha_n \rfloor)}) .$$

The finite sample behavior of this estimator is illustrated for data generated by **P1-3** with $\tilde{c}(\cdot) = 1 + \log(1/\cdot)$, $\theta = 1/2$ (and thus $c = \tilde{c}/2$) and $\theta = 1$ (and thus $\tilde{c} = 1$). To compute $\hat{S}_{n,x_0}^{(3)}$, we take $A_{n,i} = (x_{n,i-1}, x_{n,i}]$ and a bandwidth $H_n = h(x_0)$ with:

$$h(x_0) := 2n(h^*)^2 / \sum_{i=1}^n \mathbb{I}_{\{|x_{n,i} - x_0| \leq h^*\}} ,$$

where $h^* = 0.05$. Note that the bandwidth depends on x_0 and takes into account the repartition of the $x_{n,i}$'s on $[0, 1]$. The sequence α_n is fixed to $n^{-1/3}$. The estimator $\hat{c}_n(x_0)$ is computed for $N = 500$ samples of size $n = 1000$ and for $x_0 \in \{10(i-1)/n, i = 1, \dots, \lfloor n/10 \rfloor + 1\}$. For each x_0 and each θ , we represent on Figure 1 the quantiles of order 0.1, 0.5 (median) and 0.9.

4.3 Behavior of the estimator in presence of random covariates

We are now interested in the behavior of the estimator $\gamma_{n,\varphi}(\hat{S}_n^{(2)}(\cdot|x_0); \alpha_n)$ in the situation where a variable of interest Y is recorded with a random covariate $X \in [0, 1]$. We generate $N = 500$ samples of size $n = 500$ using process **P2**. The function \tilde{c} is taken equal to $\log(1/\cdot) + 1$ and the conditional extreme value index is given by $\gamma(x) = 2/3 + 1/3 \sin(2\pi x)$. For the process **P2-3**, we choose $\theta(x) = \gamma(x)$ for all $x \in [0, 1]$. As before, the value of η is fixed to 0.02 and $\varphi(\cdot) = \log(1/\cdot)$. The couple of sequences (α_n, H_n) required to compute our estimator are picked in the set $\{(n^{-1/i}, n^{(1-i)/(ij)}), i, j = 2, \dots, 6\}$. Note that taking (α_n, H_n) in this set ensures that $\log(nH_n\alpha_n) = (i-1)(j-1)/(ij) \log(n) \rightarrow \infty$ as $n \rightarrow \infty$. The value of x_0 is fixed to $1/4$ for which the maximum of the function γ is reached ($\gamma(1/4) = 1$).

We compare the estimator $\gamma_{n,\varphi}(\hat{S}_n^{(2)}(\cdot|x_0); \alpha_n)$ to the estimator proposed by Stupfler [30] and defined as follows: for $\delta \in \mathbb{N}$, let $k_n(x_0) := \lfloor \alpha_n M(x_0, H_n) \rfloor$ and

$$\mathcal{H}_{n,S}^{(\delta)}(x_0) := \frac{1}{k_n(x_0)} \sum_{i=1}^{k_n(x_0)} \left(\log \frac{Y_{(M(x_0, H_n) - i + 1)}^*(x_0)}{Y_{(M(x_0, H_n) - k_n(x_0))}^*(x_0)} \right)^\delta ,$$

where

$$M(x_0, H_n) := \sum_{i=1}^n \mathbb{I}_{\{|X_i - x_0| \leq H_n\}} ,$$

and, given $M(x_0, H_n) = p$, $Y_1^*(x_0), \dots, Y_p^*(x_0)$ are the response variables whose associated covariate X_i is such that $|X_i - x_0| \leq H_n$. As previously mentioned, $Y_{(1)}^*(x_0) \leq \dots \leq Y_{(p)}^*(x_0)$ are the associated ordered statistics. The estimator introduced in [30] is then given by:

$$\hat{\gamma}_S(\alpha_n, H_n|x_0) := \mathcal{H}_{n,S}^{(1)} - 1 - \frac{1}{2} \left(1 - \frac{(\mathcal{H}_{n,S}^{(1)})^2}{\mathcal{H}_{n,S}^{(2)}} \right)^{-1} .$$

To make the comparison, we compute for each values of α_n and H_n the ratio

$$R_2(\alpha_n, H_n) := \text{EMSE} \left(\gamma_{n,\varphi}(\hat{S}_n^{(2)}(\cdot|x_0); \alpha_n) \right) / \text{EMSE}(\hat{\gamma}_S(\alpha_n, H_n|x_0)) .$$

The results are gathered in Table 4. For the conditional Fréchet and Weibull distributions (processes **P2-1** and **P2-3**), the estimator $\hat{\gamma}_S(\alpha_n, H_n|x_0)$ provides, for some couples (α_n, H_n) slightly better results than our estimator. For the conditional beta distribution, our estimator is clearly better for all couples (α_n, H_n) .

5 Conclusion

In this paper, an estimator of the extreme value index γ associated to a survival function S was proposed. It is defined as a deterministic function of a nonparametric estimator of S and can be used even in the situation where the observations are not identically distributed from S . Its consistency was established under mild conditions and, through a simulation study, it appears that our estimator performs quite well in different situations. For a future work, we are interested in establishing the asymptotic normality and also treat the problem of extreme quantile estimation.

6 Proofs

6.1 Preliminary results

The first lemma is dedicated to the function Ψ_η defined in (4) and its derivative. We first introduce the integral,

$$I_\eta^{(\delta)}(s, t) := \int_\eta^1 \varphi(u) L_s^\delta(1/u) u^{-t} du < \infty,$$

where $(s, t) \in [0, \infty)^2$ and $\delta \in \mathbb{N}$. We thus have $\Psi_\eta(s) = [I_\eta^{(1)}(s, 0)]^2 / I_\eta^{(2)}(s, 0)$.

Lemma 1. *For all $\eta \in (0, 1)$, the function Ψ_η is decreasing on $(-\infty, 0]$ with $\Psi_\eta(s) \rightarrow I_\eta^{(0)}(0, 0)$ as $s \rightarrow -\infty$. Furthermore, the function Ψ_η is continuously differentiable on $(-\infty, 0]$ with $\Psi'_\eta(s) \rightarrow 0^-$ as $s \rightarrow -\infty$ and $\Psi'_\eta(s) \rightarrow \{[I_\eta^{(2)}(0, 0)]^2 I_\eta^{(1)}(0, 0) - [I_\eta^{(1)}(0, 0)]^2 I_\eta^{(3)}(0, 0)\} / [I_\eta^{(2)}(0, 0)]^2$ as $s \rightarrow 0^-$.*

Proof – First, remark that

$$\Psi_\eta(s) = \left(\int_\eta^1 \varphi(u) s L_s(1/u) du \right)^2 \Big/ \int_\eta^1 \varphi(u) (s L_s(1/u))^2 du.$$

Since φ is a positive bounded function, $0 \leq -s L_s(1/u) \leq 1$ for all $s \geq 0$ and $s L_s(1/u) \rightarrow -1$ as $s \rightarrow -\infty$, the dominated convergence entails that

$$I_\eta^{(\delta)}(s, 0) = -\frac{1}{s^\delta} \int_\eta^1 \varphi(u) du (1 + o(1)), \quad (19)$$

as $s \rightarrow 0$ and hence, $\Psi_\eta(s) \rightarrow I_\eta^{(0)}(0, 0)$ as s goes to zero.

We now compute the derivative of the function Ψ_η . Since φ is a positive bounded function, $0 \leq L_s(1/u) \leq \log(1/\eta)$ and $0 \leq u^{-s} \leq 1$ for all $s \geq 0$, the functions $I_\eta^{(\delta)}(s, 0)$ and $I_\eta^{(\delta)}(0, s)$ are continuous on $s \in (-\infty, 0]$ and one can interchange the derivative and the integral sign. Hence, for $s < 0$,

$$A'(s) := \frac{d}{ds} I_\eta^{(1)}(s, 0) = \frac{1}{s} \left(I_\eta^{(1)}(0, s) - I_\eta^{(1)}(s, 0) \right) = \int_\eta^1 \frac{\varphi(u)}{s} (\log(1/u) u^{-s} - L_s(1/u)) du. \quad (20)$$

Furthermore, remarking that

$$I_\eta^{(2)}(s, 0) = \frac{2}{s} \left(I_\eta^{(1)}(2s, 0) - I_\eta^{(1)}(s, 0) \right),$$

one has for $s < 0$;

$$\begin{aligned} B'(s) &= \frac{d}{ds} I_\eta^{(2)}(s, 0) = \frac{2}{s^2} \left(I_\eta^{(1)}(0, 2s) - I_\eta^{(1)}(0, s) + 2 \left(I_\eta^{(1)}(s, 0) - I_\eta^{(1)}(2s, 0) \right) \right) \\ &= \int_\eta^1 \frac{2\varphi(u)}{s} L_s(1/u) [\log(1/u) u^{-s} - L_s(1/u)] du. \end{aligned} \quad (21)$$

Since the functions A' and B' are continuous, the derivative

$$\Psi'_\eta(s) = \frac{2A'(s)I_\eta^{(1)}(s,0)I_\eta^{(2)}(s,0) - [I_\eta^{(1)}(s,0)]^2 B'(s)}{[I_\eta^{(2)}(s,0)]^2},$$

is also a continuous function on $s \in (-\infty, 0)$.

We now compute the limit of $\Psi'_\eta(s)$ as $s \rightarrow 0$. First, we focus on the function A' . Since for all $u \in (0, 1]$, $L_s(1/u)$ is a non-decreasing function in $s \in \mathbb{R}$, the following inequalities hold:

$$L_s(1/u) \leq \log(1/u) \text{ if } s \leq 0 \text{ and } L_s(1/u) \geq \log(1/u) \text{ if } s \geq 0. \quad (22)$$

A straightforward consequence is that for all $s < 0$, $0 \leq [\log(1/u)u^{-s} - L_s(1/u)]/s \leq \log^2(1/u)$. Remarking that $\log(1/u)u^{-s} - L_s(1/u) \rightarrow \log^2(1/u)/2$ as $s \rightarrow 0$, the dominated convergence theorem lead to

$$\lim_{s \rightarrow 0} A'(s) = \frac{1}{2} \int_\eta^1 \varphi(u) \log^2(1/u) du = \frac{1}{2} I_\eta^{(2)}(0, 0). \quad (23)$$

We now focus on the function B' . Using the inequalities (22) leads to

$$0 \leq \frac{1}{s} L_s(1/u) [\log(1/u)u^{-s} - L_s(1/u)] \leq \log^3(1/u),$$

for all $s < 0$. Remarking that $L_s(1/u) [\log(1/u)u^{-s} - L_s(1/u)]/s \rightarrow \log^3(1/u)/2$ as $s \rightarrow 0$ and using the dominated convergence theorem leads to

$$\lim_{s \rightarrow 0} B'(s) = \int_\eta^1 \varphi(u) \log^3(1/u) du = I_\eta^{(3)}(0, 0). \quad (24)$$

Collecting (23) and (24) and since $I_\eta^{(\delta)}(s, 0)$ and $I_\eta^{(\delta)}(0, s)$ are continuous on $(-\infty, 0]$, the limit of Ψ'_η as $s \rightarrow 0$ is given by $\{[I_\eta^{(2)}(0, 0)]^2 I_\eta^{(1)}(0, 0) - [I_\eta^{(1)}(0, 0)]^2 I_\eta^{(3)}(0, 0)\} / [I_\eta^{(2)}(0, 0)]^2$.

We are now interested in the limit of the derivative of Ψ_η as $s \rightarrow -\infty$. Collecting (20) and (21),

$$\Psi'_\eta(s) = \frac{I_\eta^{(1)}(s, 0) \left(2I_\eta^{(1)}(0, s)I_\eta^{(1)}(2s, 0) - I_\eta^{(1)}(0, s)I_\eta^{(1)}(s, 0) - I_\eta^{(1)}(0, 2s)I_\eta^{(1)}(s, 0) \right)}{2 \left(I_\eta^{(1)}(2s, 0) - I_\eta^{(1)}(s, 0) \right)^2}, \quad (25)$$

for $s < 0$. Using (19) and remarking that $I_\eta^{(1)}(0, s) \rightarrow 0$ as $s \rightarrow -\infty$ shows that $\Psi'_\eta(s) \rightarrow 0$ as $s \rightarrow -\infty$.

It remains to prove that $\psi'_\eta(s) < 0$ for all $s < 0$ or equivalently, in view of (25) that

$$\Delta_\eta(s) := 2I_\eta^{(1)}(0, s)I_\eta^{(1)}(2s, 0) - I_\eta^{(1)}(0, s)I_\eta^{(1)}(s, 0) - I_\eta^{(1)}(0, 2s)I_\eta^{(1)}(s, 0) < 0,$$

since $I_\eta^{(1)}(s, 0) > 0$ for $s \leq 0$. Furthermore, using the fact that $2L_{2s}(1/u) = L_s(1/u)(u^{-s} + 1)$, one has

$$\frac{I_\eta^{(1)}(2s, 0)}{I_\eta^{(1)}(s, 0)} = \frac{1}{2} \left(1 + \frac{I_\eta^{(1)}(s, s)}{I_\eta^{(1)}(s, 0)} \right),$$

Hence,

$$\begin{aligned} \Delta_\eta(s) &= I_\eta^{(1)}(s, 0)I_\eta^{(1)}(0, s) \left(2 \frac{I_\eta^{(1)}(2s, 0)}{I_\eta^{(1)}(s, 0)} - 1 - \frac{I_\eta^{(1)}(0, 2s)}{I_\eta^{(1)}(0, s)} \right) \\ &= I_\eta^{(1)}(s, s)I_\eta^{(1)}(0, s) - I_\eta^{(1)}(s, 0)I_\eta^{(1)}(0, 2s) \\ &= \int_\eta^1 \int_\eta^1 \varphi(u)\varphi(v)L_s(1/u)v^{-s} \log(1/v)(u^{-s} - v^{-s}) dv du \end{aligned}$$

Next, using the decomposition $\int_\eta^1 = \int_\eta^u + \int_u^1$ and Fubini's Theorem (which can be applied because the function involved is of constant sign),

$$\Delta_\eta(s) = \int_\eta^1 \int_\eta^u \varphi(u)\varphi(v)(u^{-s} - v^{-s})\kappa(u, v) dv du,$$

where $\kappa(u, v) := L_s(1/u)v^{-s} \log(1/v) - L_s(1/v)u^{-s} \log(1/u)$. Remark now that for $v \in (\eta, u)$ and $s < 0$,

$$\begin{aligned} \frac{\partial}{\partial v} \kappa(u, v) &= v^{-(s+1)} (u^{-s} \log(1/u) - L_s(1/u)(s \log(1/v) + 1)) \\ &\geq v^{-(s+1)} (\log(1/u) - L_s(1/u)) > 0. \end{aligned}$$

Hence, for all $u \in (\eta, 1)$ and $v \in (\eta, u)$, $\kappa(u, v) < \kappa(u, u) = 0$ and the proof is complete since $\varphi(u)\varphi(v)(u^{-s} - v^{-s}) > 0$ when $u \in (\eta, 1)$ and $v \in (\eta, u)$. \blacksquare

The next two lemmas are general results on extended regular varying functions. The first result shows that the convergence characterizing a function of extended regular variation is locally uniform.

Lemma 2. *Let S a survival function of class $\mathcal{ERV}(\gamma, \mathbf{a})$ then for all $\eta > 0$,*

$$\lim_{\alpha \rightarrow 0} \sup_{u \in [\eta, 1]} \left| \frac{S^\leftarrow(u\alpha) - S^\leftarrow(\alpha)}{a(\alpha^{-1})} - L_\gamma(1/u) \right| = 0.$$

Proof of Lemma 2 – From [22, Theorem B.2.18], for all $\varepsilon \in (0, 1)$, there exists $\alpha_0(\varepsilon)$ such that for all $\alpha < \alpha_0(\varepsilon)$ and all $u \in [\eta, 1]$,

$$\Delta_{S^\leftarrow}(\alpha, u) := \left| \frac{S^\leftarrow(u\alpha) - S^\leftarrow(\alpha)}{a(\alpha^{-1})} - L_\gamma(1/u) \right| \leq \frac{\varepsilon}{\eta} \max\{1, \eta^{-\gamma}\},$$

where for $y \geq 0$,

$$a_0(y) = \begin{cases} \gamma S^\leftarrow(1/y) & \text{if } \gamma > 0, \\ -\gamma(S^\leftarrow(0) - S^\leftarrow(1/y)) & \text{if } \gamma < 0, \\ S^\leftarrow(1/y) - y^{-1} \int_0^y S^\leftarrow(1/s) ds & \text{if } \gamma = 0. \end{cases}$$

Clearly, $a(y)/a_0(y) \rightarrow 1$ as y goes to infinity. Hence, there exists $\alpha_1(\varepsilon)$ such that for $\alpha < \alpha_1(\varepsilon)$, $|1 - a(\alpha^{-1})/a_0(\alpha^{-1})| \leq \varepsilon$. For $\alpha < \alpha_0(\varepsilon) \wedge \alpha_1(\varepsilon)$, we thus have the inequality

$$\begin{aligned} \Delta_{S^\leftarrow}(\alpha, u) &\leq \frac{a_0(\alpha^{-1})}{a(\alpha^{-1})} \left| \frac{S^\leftarrow(u\alpha) - S^\leftarrow(\alpha)}{a_0(\alpha^{-1})} - L_\gamma(1/u) \right| + |L_\gamma(1/u)| \left| 1 - \frac{a(\alpha^{-1})}{a_0(\alpha^{-1})} \right| \\ &\leq (1 + \varepsilon) \frac{\varepsilon}{\eta} \max\{1, \eta^{-\gamma}\} + L_\gamma(1/\eta) \varepsilon \leq \left(\frac{2 \max\{1, \eta^{-\gamma}\}}{\eta} + L_\gamma(1/\eta) \right) \varepsilon, \end{aligned}$$

which concludes the proof. \blacksquare

Lemma 3. *If there exist positive sequences (α_n) and (τ_n) converging to 0 as $n \rightarrow \infty$ such that the survival function S is of class $\mathcal{ERV}_2(\gamma, \mathbf{a}, \tau_n, \alpha_n)$ then for all $\eta \in (0, 1)$,*

$$\lim_{n \rightarrow \infty} \tau_n^{-1} \sup_{u \in [\eta, 1]} \left| \frac{S^\leftarrow(\alpha_n)}{a(\alpha_n^{-1})} \log \frac{S^\leftarrow(u\alpha_n)}{S^\leftarrow(\alpha_n)} - L_{\gamma_-}(1/u) \right| = 0, \quad (26)$$

and

$$\lim_{n \rightarrow \infty} \tau_n^{-1} \sup_{v \in [0, L_\gamma(1/\eta)]} \left| \frac{\alpha_n}{S(S^\leftarrow(\alpha_n) + v a(\alpha_n^{-1}))} - L_\gamma^\leftarrow(v) \right| = 0. \quad (27)$$

Proof of Lemma 3 – We first focus on (26). Let us introduce the notations

$$\Delta_n(u) := \frac{S^\leftarrow(u\alpha_n) - S^\leftarrow(\alpha_n)}{a(\alpha_n^{-1})} - L_\gamma(1/u) \text{ and } \bar{\Delta}_n := \sup_{u \in [\eta, 1]} |\Delta_n(u)|.$$

We start with the following equality:

$$\frac{S^\leftarrow(\alpha_n)}{a(\alpha_n^{-1})} \log \frac{S^\leftarrow(u\alpha_n)}{S^\leftarrow(\alpha_n)} = \frac{S^\leftarrow(\alpha_n)}{a(\alpha_n^{-1})} \log \left[1 + \frac{a(\alpha_n^{-1})}{S^\leftarrow(\alpha_n)} (L_\gamma(1/u) + \Delta_n(u)) \right].$$

We consider the case $\gamma > 0$. A straightforward calculus leads to

$$\log \left[1 + \frac{a(\alpha_n^{-1})}{S^{\leftarrow}(\alpha_n)} (L_\gamma(1/u) + \Delta_n(u)) \right] - \log(u^{-\gamma}) = \log(1 + D_{n,1}(u)),$$

where

$$\begin{aligned} |D_{n,1}(u)| &= u^\gamma \left| \gamma \Delta_n(u) + \left(\frac{a(\alpha_n^{-1})}{S^{\leftarrow}(\alpha_n)} - \gamma \right) (L_\gamma(1/u) + \Delta_n(u)) \right| \\ &\leq \gamma \bar{\Delta}_n + \left| \frac{a(\alpha_n^{-1})}{S^{\leftarrow}(\alpha_n)} - \gamma \right| (L_\gamma(1/\eta) + \bar{\Delta}_n) \rightarrow 0. \end{aligned}$$

Hence, using inequality $|\log(1+x)| \leq 3|x|/2$ for $x \in [-1/2, 1/2]$, one has for n large enough and $\gamma > 0$

$$\begin{aligned} \tau_n^{-1} \sup_{u \in [\eta, 1]} \left| \frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \log \frac{S^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(\alpha_n)} - \log(1/u) \right| &\leq \frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \frac{3\tau_n^{-1}}{2} \sup_{u \in [\eta, 1]} |D_{n,1}(u)| \\ &\quad + \log(1/\eta) \tau_n^{-1} \left| \frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \gamma - 1 \right|, \quad (28) \end{aligned}$$

which converges to 0 by assumption. Now, assume that $\gamma \leq 0$, since

$$|D_{n,2}(u)| := \left| \frac{a(\alpha_n^{-1})}{S^{\leftarrow}(\alpha_n)} (L_\gamma(1/u) + \Delta_n(u)) \right| \leq \frac{a(\alpha_n^{-1})}{S^{\leftarrow}(\alpha_n)} (L_\gamma(1/\eta) + \bar{\Delta}_n) \rightarrow 0$$

and using the inequality $x(1-x) \leq \log(1+x) \leq x$ for $x \in [-1/2, 1/2]$, one has:

$$-L_\gamma(1/u) D_{n,2}(u) + \Delta_n(u)(1 - D_{n,2}(u)) \leq \frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \log \frac{S^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(\alpha_n)} - L_\gamma(1/u) \leq \Delta_n(u).$$

Since $\tau_n^{-1} \sup |D_{n,2}(u)| \rightarrow 0$ and $\tau_n^{-1} \bar{\Delta}_n \rightarrow 0$, it is clear that

$$\tau_n^{-1} \sup_{u \in [\eta, 1]} \left| \frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \log \frac{S^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(\alpha_n)} - L_\gamma(1/u) \right| \rightarrow 0. \quad (29)$$

Collecting (28) and (29) conclude the proof of (26). The proof of equation (27) is a direct consequence of Vervaat's Lemma (see [22, Lemma A.0.2]). \blacksquare

The following lemma is a technical result that will be useful in the proof of Lemma 5. A proof of this result can be found in [17, Lemma 6].

Lemma 4. *Let (X_n) be a sequence of positive real-valued random variables such that for every positive nonrandom sequence δ_n converging to 0, the random sequence $\delta_n X_n$ converges to 0 in probability. Then $X_n = \mathcal{O}_{\mathbb{P}}(1)$.*

Let $\{\hat{S}_n(y), y \geq 0\}$ be a $[0, 1]$ -valued stochastic process where \hat{S}_n is a non-increasing and right-continuous function. In the next result it is shown that if for large values of y , $\hat{S}_n(y)$ is a consistent estimator of $S(y)$ then $\hat{S}_n^{\leftarrow}(\alpha)$ is also a consistent estimator of $S^{\leftarrow}(\alpha)$ for small values of α . This result is a cornerstone in the proof of Theorem 2.

Lemma 5. *Let (α_n) and (τ_n) be sequences converging to 0 as $n \rightarrow \infty$ and such that the survival function S is of class $\mathcal{ERV}_2(\gamma, \mathbf{a}, \boldsymbol{\tau}_n, \boldsymbol{\alpha}_n)$. If for every sequence $\kappa_n \rightarrow 0$, $\{\hat{S}_n(y), y \geq 0\}$ is such that for all $\eta \in (0, 1)$,*

$$\tau_n^{-1} \sup_{u \in (\eta, 1)} \left| \frac{\hat{S}_n(S^{\leftarrow}(u\alpha_n) + \kappa_n a(\alpha_n^{-1}))}{S(S^{\leftarrow}(u\alpha_n) + \kappa_n a(\alpha_n^{-1}))} - 1 \right| = \mathcal{O}_{\mathbb{P}}(1),$$

then

$$\frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \tau_n^{-1} \sup_{u \in (\eta, 1)} \left| \frac{\hat{S}_n^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(u\alpha_n)} - 1 \right| = \mathcal{O}_{\mathbb{P}}(1). \quad (30)$$

From equation (27) in Lemma 3, it is easy to see that uniformly on $u \in (\eta, 1)$, $S(S^{\leftarrow}(u\alpha_n) + \kappa_n a(\alpha_n^{-1})) = u\alpha_n(1 + o(1)) \geq \eta\alpha_n/2 > 0$ for n large enough. Hence, the division by $S(S^{\leftarrow}(u\alpha_n) + \kappa_n a(\alpha_n^{-1}))$ is allowed for every sequence $\kappa_n \rightarrow 0$ at least for n large enough.

Proof of Lemma 5 – Let us introduce the sequence $m_n := \lfloor \tau_n^{-1} \rfloor$. For $j = 1, \dots, m_n$, let $\theta_n(j) := \eta + j(1 - \eta)/(m_n + 1) \in (\eta, 1)$. Clearly, for all $u \in (\eta, 1)$, there exists $j_u \in \{1, \dots, m_n - 1\}$ such that $\theta_n(j_u) \leq u < \theta_n(j_u + 1)$. Since S^{\leftarrow} and \hat{S}_n^{\leftarrow} are non-increasing and right-continuous functions, it is easy to check that for all $u \in (\eta, 1)$,

$$\begin{aligned} \left| \hat{S}_n^{\leftarrow}(u\alpha_n) - S^{\leftarrow}(u\alpha_n) \right| &\leq 2 \left| \hat{S}_n^{\leftarrow}(\theta_n(j_u)\alpha_n) - S^{\leftarrow}(\theta_n(j_u)\alpha_n) \right| \\ &\quad + \left| \hat{S}_n^{\leftarrow}(\theta_n(j_u + 1)\alpha_n) - S^{\leftarrow}(\theta_n(j_u + 1)\alpha_n) \right| \\ &\quad + 2(S^{\leftarrow}(\theta_n(j_u)\alpha_n) - S^{\leftarrow}(\theta_n(j_u + 1)\alpha_n)). \end{aligned}$$

Hence,

$$\sup_{u \in (\eta, 1)} \left| \hat{S}_n^{\leftarrow}(u\alpha_n) - S^{\leftarrow}(u\alpha_n) \right| \leq 3(T_{n,1} + T_{n,2}),$$

with

$$T_{n,1} := \max_{j=1, \dots, m_n-1} [S^{\leftarrow}(\theta_n(j)\alpha_n) - S^{\leftarrow}(\theta_n(j+1)\alpha_n)],$$

and

$$T_{n,2} := \max_{j=1, \dots, m_n} \left| \hat{S}_n^{\leftarrow}(\theta_n(j)\alpha_n) - S^{\leftarrow}(\theta_n(j)\alpha_n) \right|.$$

Let us first focus on the term $T_{n,1}$. Since S is of class $\mathcal{ERV}_2(\gamma, a, \tau_n, \alpha_n)$, a straightforward calculus entails that

$$S^{\leftarrow}(\theta_n(j)\alpha_n) - S^{\leftarrow}(\theta_n(j+1)\alpha_n) = a(\alpha_n^{-1}) [L_\gamma(1/\theta_n(j)) - L_\gamma(1/\theta_n(j+1)) + o(\tau_n)],$$

where the term $o(\tau_n)$ converges to 0 uniformly on $u \in (\eta, 1)$.

Since the derivative of the function $L_\gamma(1/\cdot)$ is bounded on $(\eta, 1)$, a first order Taylor expansion leads to $S^{\leftarrow}(\theta_n(j)\alpha_n) - S^{\leftarrow}(\theta_n(j+1)\alpha_n) = a(\alpha_n^{-1}) (\mathcal{O}(m_n^{-1}) + o(\tau_n))$, uniformly on $u \in (\eta, 1)$ and thus

$$T_{n,1} = a(\alpha_n^{-1}) \mathcal{O}(\tau_n). \quad (31)$$

Let us now consider the term $T_{n,2}$. Our goal is to show that

$$\frac{\tau_n^{-1}}{a(\alpha_n^{-1})} T_{n,2} = \mathcal{O}_{\mathbb{P}}(1). \quad (32)$$

To this end, it suffices, from Lemma 4, to show that for every sequence $\delta_n \rightarrow 0$ and for every $\varepsilon > 0$,

$$p_n(\varepsilon) := \mathbb{P} \left[\frac{\delta_n \tau_n^{-1}}{a(\alpha_n^{-1})} \max_{j=1, \dots, m_n} \left| \hat{S}_n^{\leftarrow}(\theta_n(j)\alpha_n) - S^{\leftarrow}(\theta_n(j)\alpha_n) \right| > \varepsilon \right] \rightarrow 0.$$

Note that it is sufficient to consider sequences $\delta_n \rightarrow 0$ such that $\delta_n^{-1} \tau_n \rightarrow 0$ as n goes to infinity. Introducing the sequence

$$y_n^\pm(\theta_n(j)) := S^{\leftarrow}(\theta_n(j)\alpha_n) \pm \varepsilon \tau_n \frac{a(\alpha_n^{-1})}{\delta_n},$$

one has

$$\begin{aligned} p_n(\varepsilon) &\leq \mathbb{P} \left[\bigcup_{j=1}^{m_n} \left\{ \hat{S}_n^{\leftarrow}(\theta_n(j)\alpha_n) > y_{j,n}^+ \right\} \cup \left\{ \hat{S}_n^{\leftarrow}(\theta_n(j)\alpha_n) \leq y_{j,n}^- \right\} \right] \\ &\leq 1 - \mathbb{P} \left[\bigcap_{j=1}^{m_n} \left\{ \hat{S}_n^{\leftarrow}(\theta_n(j)\alpha_n) \leq y_{j,n}^+ \right\} \right] + \mathbb{P} \left[\bigcup_{j=1}^{m_n} \left\{ \hat{S}_n^{\leftarrow}(\theta_n(j)\alpha_n) \leq y_{j,n}^- \right\} \right]. \end{aligned}$$

Since \hat{S}_n^{\leftarrow} is a non-increasing and right-continuous function, it is easy to see that

$$p_n(\varepsilon) \leq \mathbb{P} \left[\bigcup_{j=1}^{m_n} \{W_n^+(\theta_n(j)) > b_{n,j}^+(\varepsilon)\} \right] + \mathbb{P} \left[\bigcup_{j=1}^{m_n} \{W_n^-(\theta_n(j)) \leq b_{n,j}^-(\varepsilon)\} \right],$$

where

$$W_n^\pm(\theta_n(j)) := \tau_n^{-1} \left(\frac{\hat{S}_n(y_n^\pm(\theta_n(j)))}{S(y_n^\pm(\theta_n(j)))} - 1 \right) \text{ and } b_{n,j}^\pm(\varepsilon) := \tau_n^{-1} \left(\frac{\theta_n(j)\alpha_n}{S(y_n^\pm(\theta_n(j)))} - 1 \right).$$

By assumption, $W_n^\pm(u) = \mathcal{O}_{\mathbb{P}}(1)$ uniformly on $u \in [\eta, 1]$. Since S is of class $\mathcal{ERV}_2(\gamma, a, \tau_n, \alpha_n)$,

$$y_n^\pm(\theta_n(j))(\varepsilon) = S^{\leftarrow}(\alpha_n) + a(\alpha_n^{-1}) [L_\gamma(1/\theta_n(j)) + \zeta_n^\pm(\varepsilon)],$$

where $\zeta_n^\pm(\varepsilon) = \tau_n(\pm\varepsilon\delta_n^{-1} + o(1))$, the term $o(1)$ converging to 0 uniformly on $u \in (\eta, 1)$. Thus, from Lemma 3, equation (27), one has for all $j \in \{1, \dots, m_n\}$,

$$b_{n,j}^+(\varepsilon) = \tau_n^{-1} [\theta_n(j)L_\gamma^{\leftarrow}(L_\gamma(1/\theta_n(j)) + \zeta_n^+(\varepsilon)) - 1 + o(\tau_n)].$$

Since for $v \in [0, L_\gamma(1/\eta)]$ the derivative of L_γ^{\leftarrow} is larger than $1 \wedge \eta^{\gamma-1}$, a Taylor expansion leads to, for n large enough:

$$b_{n,j}^+ \geq \frac{\eta\varepsilon(1 \wedge \eta^{\gamma-1})}{2\delta_n}.$$

Similarly, for n large enough, one can show that

$$b_{n,j}^- \leq -\frac{\eta\varepsilon(1 \wedge \eta^{\gamma-1})}{2\delta_n}.$$

Hence,

$$\begin{aligned} p_n(\varepsilon) &\leq \mathbb{P} \left[\bigcup_{j=1}^{m_n} \left\{ \delta_n W_n^+(\theta_n(j)) > \frac{\eta\varepsilon(1 \wedge \eta^{\gamma-1})}{2} \right\} \right] + \mathbb{P} \left[\bigcup_{j=1}^{m_n} \left\{ \delta_n W_n^-(\theta_n(j)) \leq -\frac{\eta\varepsilon(1 \wedge \eta^{\gamma-1})}{2} \right\} \right] \\ &\leq \mathbb{P} \left[\sup_{u \in [\eta, 1]} \delta_n |W_n^+(u)| > \frac{\eta\varepsilon(1 \wedge \eta^{\gamma-1})}{2} \right] + \mathbb{P} \left[\sup_{u \in [\eta, 1]} \delta_n |W_n^-(u)| \geq -\frac{\eta\varepsilon(1 \wedge \eta^{\gamma-1})}{2} \right] \rightarrow 0, \end{aligned}$$

proving (32). Finally, since S^{\leftarrow} is a non-increasing function and collecting (31) and (32),

$$\frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \tau_n^{-1} \sup_{u \in [\eta, 1]} \left| \frac{\hat{S}_n^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(\alpha_n)} - 1 \right| \leq \frac{3\tau_n^{-1}}{a(\alpha_n^{-1})} (T_{n,1} + T_{n,2}) = \mathcal{O}_{\mathbb{P}}(1),$$

which conclude the proof. ■

The next lemma establishes a uniform convergence result on processes of the form:

$$\hat{\Phi}_n(u) := \sum_{i=1}^n X_{n,i}(u),$$

where, for $\eta \in (0, 1)$, $\{X_{n,i}(u), u \in (\eta, 1)\}$, $i = 1, \dots, n$ are n independent stochastic processes with $X_{n,i}$ non-decreasing and positive. The expectation of $\hat{\Phi}_n(u)$ is denoted $\mu_n(u)$. This lemma will be used to study the numerator of (7).

Lemma 6. *Let $\tau_n = (\log(\mu_n(1))/\mu_n(1))^{1/2}$. If $\mu_n(\eta) \rightarrow \infty$ as n goes to infinity, if there exist positive constants C_X and C_μ such that for all $i \in \{1, \dots, n\}$ and $u \in [\eta, 1]$, $X_{n,i}(u) \leq C_X$, $\mu_n(\eta)/\mu_n(1) \geq C_\mu$ for n large enough and*

$$\sup \left\{ \left| \frac{\mu_n(u)}{\mu_n(u')} - 1 \right|, u \in (\eta, 1) \text{ with } |u - u'| \leq (\mu_n(1))^{-1/2} \right\} = \mathcal{O}(\tau_n), \quad (33)$$

then,

$$\sup_{u \in (\eta, 1)} \left| \frac{\hat{\Phi}_n(u)}{\mu_n(u)} - 1 \right| = \mathcal{O}_{\mathbb{P}}(\tau_n).$$

Proof of Lemma 6 – Let $C_\varepsilon := (3C_X/C_\mu)^{1/2}$. Using a multiplicative form of the Chernoff's inequality for bounded variables (see for instance [7, Theorem 1.1]), one has for all $u \in (\eta, 1)$

$$\mathbb{P} \left[\tau_n^{-1} \left| \frac{\hat{\Phi}_n(u)}{\mu_n(u)} - 1 \right| > C_\varepsilon \right] \leq 2 \exp \left(- \frac{C_\varepsilon^2}{3C_X} \frac{\mu_n(u)}{\mu_n(1)} \log(\mu_n(1)) \right) \leq \frac{2}{\mu_n(1)}. \quad (34)$$

Let us now introduce the sequence $m_n := \lceil (\mu_n(1))^{1/2} \rceil + 1 \rightarrow \infty$, and, for $j = 1, \dots, m_n$, let $\theta_n(j) := \eta + (j-1)(1-\eta)/(m_n-1) \in [\eta, 1]$. Here $\lceil \cdot \rceil$ is the notation for the ceiling function. Clearly, for all $u \in (\eta, 1)$, there exists $j_u \in \{1, \dots, m_n-1\}$ such that $\theta_n(j_u) \leq u < \theta_n(j_u+1)$ and then, since $X_{n,i}$, $i = 1, \dots, n$ are non-increasing,

$$\begin{aligned} \left| \frac{\hat{\Phi}_n(u)}{\mu_n(u)} - 1 \right| &\leq \frac{1}{\mu_n(\eta)} \left[\left| \hat{\Phi}_n(\theta_n(j_u+1)) - \mu_n(\theta_n(j_u+1)) \right| + 2 \left| \hat{\Phi}_n(\theta_n(j_u)) - \mu_n(\theta_n(j_u)) \right| \right. \\ &\quad \left. + 2(\mu_n(\theta_n(j_u+1)) - \mu_n(\theta_n(j_u))) \right], \end{aligned}$$

leading to

$$\sup_{u \in (\eta, 1)} \left| \frac{\hat{\Phi}_n(u)}{\mu_n(u)} - 1 \right| \leq \frac{3}{\mu_n(\eta)} (T_{n,1} + T_{n,2}),$$

with

$$T_{n,1} := \max_{j=1, \dots, m_n-1} (\mu_n(\theta_n(j+1)) - \mu_n(\theta_n(j))) \text{ and } T_{n,2} := \max_{j=1, \dots, m_n} \left| \hat{\Phi}_n(\theta_n(j)) - \mu_n(\theta_n(j)) \right|.$$

Under (33), since for all $j \in \{1, \dots, m_n-1\}$, $\theta_n(j) - \theta_n(j+1) = (m_n-1)^{-1} \leq (\mu_n(1))^{-1/2}$, the following holds for n large enough:

$$T_{n,1} = \mathcal{O} \left((\mu_n(1) \log(\mu_n(1)))^{1/2} \right). \quad (35)$$

Furthermore, using (34)

$$\begin{aligned} \mathbb{P} \left[(\mu_n(1) \log(\mu_n(1)))^{-1/2} T_{n,2} > C_\varepsilon \right] &\leq \mathbb{P} \left[\left(\frac{\mu_n(1)}{\log(\mu_n(1))} \right)^{1/2} \bigcup_{j=1}^{m_n} \left\{ \left| \frac{\hat{\Phi}_n(\theta_n(j))}{\mu_n(\theta_n(j))} - 1 \right| > C_\varepsilon \right\} \right] \\ &\leq 2(\mu_n(1))^{-1/2} \rightarrow 0, \end{aligned}$$

since $\mu_n(b) \rightarrow \infty$. Hence,

$$T_{n,2} = \mathcal{O}_{\mathbb{P}} \left((\mu_n(1) \log(\mu_n(1)))^{1/2} \right). \quad (36)$$

From (35) and (36), since $\mu_n(1)/\mu_n(\eta) \leq C_\mu^{-1}$, $\tau_n^{-1} T_{n,1}/\mu_n(\eta) = \mathcal{O}(1)$ and $\tau_n^{-1} T_{n,2}/\mu_n(\eta) = \mathcal{O}_{\mathbb{P}}(1)$, which conclude the proof. \blacksquare

6.2 Proofs of main results

Proof of Theorem 1 – We start by showing the following equation:

$$\left| \left(\frac{S_n^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \right)^\delta H_{\eta, \varphi}^{(\delta)}(\tilde{S}_n) - \int_\eta^1 \varphi(u) L_{\gamma_-}^\delta(1/u) du \right| \bigg/ \left(\int_\eta^1 \varphi(u) L_0(1/u) du \right)^\delta = \mathcal{O}_{\mathbb{P}}(\tau_n), \quad (37)$$

Let us first introduce the following notations: for $u \in [\eta, 1]$,

$$\Delta_n(u) = \left(\frac{\tilde{S}_n^{\leftarrow}(u\alpha_n)}{S_n^{\leftarrow}(u\alpha_n)} - 1 \right), \quad R_n(u) = \log \frac{1 + \Delta_n(u)}{1 + \Delta_n(1)},$$

$$\text{and } b_n(u) = \frac{S_n^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \log \frac{S_n^{\leftarrow}(u\alpha_n)}{S_n^{\leftarrow}(\alpha_n)} - L_{\gamma_-}(1/u).$$

One has

$$\begin{aligned}
\left(\frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \log \frac{\tilde{S}_n^{\leftarrow}(u\alpha_n)}{\tilde{S}_n^{\leftarrow}(\alpha_n)} \right)^{\delta} &= \left(\frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \right)^{\delta} \left(\log \frac{S^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(\alpha_n)} + R_n(u) \right)^{\delta} \\
&= \left(\frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \log \frac{S^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(\alpha_n)} \right)^{\delta} \\
&\quad + \sum_{j=0}^{\delta-1} C_{\delta}^j \left(\frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \log \frac{S^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(\alpha_n)} \right)^j \left(\frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} R_n(u) \right)^{\delta-j}
\end{aligned}$$

Since \tilde{S}_n satisfies (30),

$$\bar{\Delta}_n := \sup_{u \in (\eta, 1)} |\Delta_n(u)| = \mathcal{O}_{\mathbb{P}} \left(\frac{a(\alpha_n^{-1})}{S^{\leftarrow}(\alpha_n)} \tau_n \right) = o_{\mathbb{P}}(1),$$

and thus, $|R_n(u)| = \mathcal{O}_{\mathbb{P}}(\bar{\Delta}_n) = o_{\mathbb{P}}(1)$. Remark that

$$\sum_{j=0}^{\delta-1} C_{\delta}^j \left(\frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \log \frac{S^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(\alpha_n)} \right)^j \leq \sum_{j=0}^{\delta} C_{\delta}^j (L_{\gamma_-}(1/u) + b_n(u))^j \leq (1 + L_{\gamma_-}(1/\eta) + \bar{b}_n)^{\delta},$$

with $\bar{b}_n = \sup\{|b_n(u)|, u \in [\eta, 1]\}$ which converges to 0 from the first part of Lemma 3. Hence,

$$\sum_{j=0}^{\delta-1} C_{\delta}^j \left(\frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \log \frac{S^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(\alpha_n)} \right)^j = \mathcal{O}(1),$$

and thus,

$$\left(\frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \right)^{\delta} \left[\left(\log \frac{\tilde{S}_n^{\leftarrow}(u\alpha_n)}{\tilde{S}_n^{\leftarrow}(\alpha_n)} \right)^{\delta} - \left(\log \frac{S^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(\alpha_n)} \right)^{\delta} \right] = \mathcal{O}_{\mathbb{P}}(\tau_n),$$

uniformly on $u \in (\eta, 1)$. Since

$$\left(\frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \log \frac{S^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(\alpha_n)} \right)^{\delta} - L_{\gamma_-}^{\delta}(1/u) = (L_{\gamma_-}(1/u) + b_n(u))^{\delta} - L_{\gamma_-}^{\delta}(1/u) = \mathcal{O}(\bar{b}_n),$$

we have, as a first conclusion that, uniformly on $u \in (\eta, 1)$,

$$\left(\frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \log \frac{\tilde{S}_n^{\leftarrow}(u\alpha_n)}{\tilde{S}_n^{\leftarrow}(\alpha_n)} \right)^{\delta} = L_{\gamma_-}^{\delta}(1/u) + \mathcal{O}_{\mathbb{P}}(\tau_n), \quad (38)$$

since, from Lemma 3, $\tau_n^{-1}\bar{b}_n \rightarrow 0$. Multiplying equation (38) by $\varphi(u)$ and integrating between η and 1 lead to (37).

The rest of the proof is based on the decomposition

$$\gamma_{n,\varphi}(\tilde{S}_n) - \gamma = \left(H_{\eta,\varphi}^{(1)}(\tilde{S}_n) - \gamma_+ \right) + \left(\Psi_{\eta,\varphi}^{\leftarrow} \left(\max \left\{ \frac{[H_{\eta,\varphi}^{(1)}(\tilde{S}_n)]^2}{H_{\eta,\varphi}^{(2)}(\tilde{S}_n)}, \Psi_{\eta,\varphi}(0) \right\} \right) - \gamma_- \right) =: D_n^+ + D_n^-.$$

Let us first consider the term D_n^+ . From (37), one has

$$D_n^+ = \frac{a(\alpha_n^{-1})}{S^{\leftarrow}(\alpha_n)} \mathcal{O}_{\mathbb{P}}(\tau_n) + \frac{a(\alpha_n^{-1})}{S^{\leftarrow}(\alpha_n)} \int_{\eta}^1 \varphi(u) L_{\gamma_-}(1/u) du \Big/ \int_{\eta}^1 \varphi(u) L_0(1/u) du - \gamma_+.$$

Since S is of class $\mathcal{ERV}_2(\gamma, a, \tau_n, \alpha_n)$ and remarking that if $\gamma \geq 0$,

$$\int_{\eta}^1 \varphi(u) L_{\gamma_-}(1/u) du \Big/ \int_{\eta}^1 \varphi(u) L_0(1/u) du = 1,$$

it is clear that $D_n^+ = \mathcal{O}_{\mathbb{P}}(\tau_n)$. Now, using again (37), it is easy to check that

$$\frac{[H_{\eta,\varphi}^{(1)}(\tilde{S}_n)]^2}{H_{\eta,\varphi}^{(2)}(\tilde{S}_n)} = \Psi_{\eta,\varphi}(\gamma_-) (1 + \mathcal{O}_{\mathbb{P}}(\tau_n)).$$

Since $\Psi_{\eta,\varphi}$ is a decreasing function (see Lemma 1) and $\tau_n \rightarrow 0$,

$$\max \left\{ \frac{[H_{\eta,\varphi}^{(1)}(\tilde{S}_n)]^2}{H_{\eta,\varphi}^{(2)}(\tilde{S}_n)}, \Psi_{\eta,\varphi}(0) \right\} = \Psi_{\eta,\varphi}(\gamma_-) (1 + \mathcal{O}_{\mathbb{P}}(\tau_n)).$$

Finally, since from Lemma 1, the derivative of $\Psi_{\eta,\varphi}^{\leftarrow}$ is bounded in a neighborhood of $\Psi_{\eta,\varphi}(\gamma_-)$,

$$D_n^- = \Psi_{\eta,\varphi}^{\leftarrow} \left(\max \left\{ \frac{[H_{\eta,\varphi}^{(1)}(\tilde{S}_n)]^2}{H_{\eta,\varphi}^{(2)}(\tilde{S}_n)}, \Psi_{\eta,\varphi}(0) \right\} \right) - \Psi_{\eta,\varphi}^{\leftarrow}(\Psi_{\eta,\varphi}(\gamma_-)) = \mathcal{O}_{\mathbb{P}}(\tau_n),$$

which concludes the proof. \blacksquare

Proof of Theorem 2 – From Theorem 1, it suffices to show that

$$\frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \tau_n^{-1} \sup_{u \in (\eta, 1)} \left| \frac{\hat{S}_n^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(u\alpha_n)} - 1 \right| = \mathcal{O}_{\mathbb{P}}(1), \quad (39)$$

where \hat{S}_n is given by (7). From Lemma 5, letting $y_{\kappa_n}(u; \alpha_n) := S^{\leftarrow}(u\alpha_n) + \kappa_n a(\alpha_n^{-1})$, it suffices to prove that for every sequence $\kappa_n \rightarrow 0$,

$$\tau_n^{-1} \sup_{u \in (\eta, 1)} \left| \frac{\hat{S}_n(y_{\kappa_n}(u; \alpha_n))}{S(y_{\kappa_n}(u; \alpha_n))} - 1 \right| = \mathcal{O}_{\mathbb{P}}(1). \quad (40)$$

First, let us consider the numerator of $\hat{S}_n(y_{\kappa_n}(u; \alpha_n))$ given by:

$$\hat{\Phi}_n(u) := \sum_{i=1}^n R_{n,i}(y_{\kappa_n}(u; \alpha_n)).$$

Recall that, by assumption,

$$\mu_{\kappa_n}(u; \alpha_n) = \mathbb{E}(\hat{\Phi}_n(u)) = r_0^{-1} \frac{S(y_{\kappa_n}(u; \alpha_n))}{\alpha_n \sigma_n^2} (1 + o(\tau_n)),$$

uniformly on $u \in (\eta, 1)$. Since the survival function S is of class $\mathcal{ERV}_2(\gamma, a, \tau_n, \alpha_n)$, Lemma 3, equation (27) entails that, uniformly on $u \in (\eta, 1)$

$$S(y_{\kappa_n}(u; \alpha_n)) = \alpha_n [1/L_{\gamma}^{\leftarrow}(L_{\gamma}(1/u) + \kappa_n + o(\tau_n)) + o(\tau_n)] = u\alpha_n(1 + o(1)),$$

and thus, uniformly on $u \in (\eta, 1)$,

$$\mu_{\kappa_n}(u; \alpha_n) = r_0^{-1} \sigma_n^{-2} [1/L_{\gamma}^{\leftarrow}(L_{\gamma}(1/u) + \kappa_n + o(\tau_n)) + o(\tau_n)].$$

As a first conclusion, $(\mu_{\kappa_n}(1; \alpha_n)/\log(\mu_{\kappa_n}(1; \alpha_n)))^{-1/2}$ is asymptotically proportional to τ_n and, for n large enough,

$$\frac{\mu_{\kappa_n}(\eta; \alpha_n)}{\mu_{\kappa_n}(1; \alpha_n)} \leq 2\eta. \quad (41)$$

We now show that $\mu_{\kappa_n}(u; \alpha_n)$ satisfies condition (33) of Lemma 6. For $(u, u') \in (\eta, 1)^2$ such that $|u - u'| \leq (\mu_{\kappa_n}(1; \alpha_n))^{-1/2}$, since the derivatives of $1/L_{\gamma}^{\leftarrow}$ and $L_{\gamma}(1/\cdot)$ are bounded on $(\eta, 1)$, a Taylor expansion leads to $|\mu_{\kappa_n}(u; \alpha_n) - \mu_{\kappa_n}(u'; \alpha_n)| = \mathcal{O}(\sigma_n^{-2} \tau_n)$. Thus,

$$\left| \frac{\mu_n(u)}{\mu_n(u')} - 1 \right| \leq \frac{1}{\mu_n(\eta)} |\mu_n(u) - \mu_n(u')| = \mathcal{O}(\tau_n). \quad (42)$$

Hence, collecting (41) and (42), Lemma 6 entails that

$$\sup_{u \in (\eta, 1)} \left| \frac{\hat{\Phi}_n(u) r_0 \alpha_n \sigma_n^2}{S(y_{\kappa_n}(u; \alpha_n))} (1 + o(\tau_n)) - 1 \right| = \mathcal{O}_{\mathbb{P}}(\tau_n).$$

Finally, since by assumption,

$$\alpha_n \sigma_n^2 \sum_{i=1}^n R_{n,i}(0) = r_0 + o_{\mathbb{P}}(\tau_n),$$

uniformly on $u \in (\eta, 1)$, one has

$$\sup_{u \in (\eta, 1)} \left| \frac{\hat{S}_n(y_{\kappa_n}(u; \alpha_n))}{S(y_{\kappa_n}(u; \alpha_n))} (1 + o_{\mathbb{P}}(\tau_n)) - 1 \right| = \mathcal{O}_{\mathbb{P}}(\tau_n).$$

It is easy to check that the factor $1 + o_{\mathbb{P}}(\tau_n)$ can be removed proving (40) for every sequence $\kappa_n \rightarrow 0$ and consequently (39) is established. This concludes the proof. \blacksquare

Proof of Corollary 1 – Let $z_{\kappa_n}(u; \alpha_n) = S^{\leftarrow}(u\alpha_n) + \kappa_n a_1(\alpha_n^{-1})$, where κ_n is a sequence converging to zero as $n \rightarrow \infty$. Keeping in mind the notation of equation (10), it is easy to check that

$$\sum_{i=1}^n R_{n,i}^{(1)}(0) = n,$$

and thus the second part of condition (8) of Theorem 2 is clearly satisfied with $r_0 = 1$. Now, let

$$\hat{\Phi}_n^{(1)}(u) := \sum_{i=1}^n R_{n,i}^{(1)}(z_{\kappa_n}(u; \alpha_n)).$$

Our goal is to provide an expansion of $\mu_{\kappa_n}^{(1)}(u; \alpha_n) := \mathbb{E}(\hat{\Phi}_n^{(1)}(u))$ in order to check the validity of the first part of condition (8) in Theorem 2. We start with

$$\frac{\mu_{\kappa_n}(u; \alpha_n)}{nS(z_{\kappa_n}(u; \alpha_n))} = \sum_{i=1}^n \frac{x_{n,i} - x_{n,i-1}}{x_{n,n} - b_1} \left(\frac{S_{x_{n,i}}(z_{\kappa_n}(u; \alpha_n))}{S(z_{\kappa_n}(u; \alpha_n))} - c(x_{n,i}) \right) \quad (43)$$

$$+ \sum_{i=1}^n \frac{x_{n,i} - x_{n,i-1}}{x_{n,n} - b_1} c(x_{n,i}). \quad (44)$$

Since S is of class $\mathcal{ERV}_2(\gamma, \mathbf{a}_1, \tau_n, \alpha_n)$, it is easy to check that there exists $\delta > 1$ such that for all $u \in (\eta, 1)$, $z_{\kappa_n}(u; \alpha_n)/S^{\leftarrow}(\alpha_n) \in [\delta^{-1}, \delta]$. Hence, from condition (12), the term (43) is a $o(\tau_n)$ uniformly on $u \in (\eta, 1)$. Furthermore,

$$\begin{aligned} 1 = \frac{1}{b_2 - b_1} \int_{b_1}^{b_2} c(s) ds &= \frac{x_{n,n} - b_1}{b_2 - b_1} \sum_{i=1}^n \frac{x_{n,i} - x_{n,i-1}}{x_{n,n} - b_1} c(x_{n,i}) + c(b_2) \frac{x_{n,n} - b_1}{b_2 - b_1} (b_2 - x_{n,n}) \\ &+ \frac{1}{b_2 - b_1} \sum_{i=1}^{n+1} \int_{x_{n,i-1}}^{x_{n,i}} (c(s) - c(x_{n,i})) ds \end{aligned}$$

From conditions **(A.3)** and (11), since $n\sigma_n \rightarrow \infty$, we deduce from the previous equation that the term (44) is equal to $1 + o(\tau_n)$ uniformly on $u \in (\eta, 1)$. Hence, condition (8) is satisfied and the conclusion follows applying Theorem 2. \blacksquare

Proof of Corollary 2 – Let $y_{\kappa_n}(u; \alpha_n) = S^{\leftarrow}(u\alpha_n|x_0) + \kappa_n a_2(\alpha_n^{-1}|x_0)$, where κ_n is a sequence converging to zero as $n \rightarrow \infty$. Let us first focus on the denominator of $\hat{S}_n^{(2)}(\cdot|x_0)$. Let

$$\hat{g}_n(x_0) = \frac{1}{n} \sum_{i=1}^n K_{H_n}(x_0 - X_i) = (n|H_n|)^{-1} \sum_{i=1}^n R_{n,i}^{(2)}(0|x_0),$$

where $R_{n,i}^{(2)}$ is defined in (14). Under **(A.2)** and **(A.3)**, since $n|H_n| \rightarrow \infty$, it is well known that

$$\hat{g}_n(x_0) = g(x_0) \left(1 + \mathcal{O}_{\mathbb{P}}(\|H_n\|_{\infty}) + \mathcal{O}_{\mathbb{P}}\left((n|H_n|)^{-1/2}\right) \right),$$

(see Parzen [26] for a proof). Then, since by assumption $\tau_n^{-1}\|H_n\|_{\infty}$ and $\tau_n^{-1}(n|H_n|)^{-1/2}$ converge to 0, one has that $\hat{g}_n(x_0)/g(x_0) = 1 + o_{\mathbb{P}}(\tau_n)$. Thus the second part of condition (8) in Theorem 2 is satisfied. Now, let

$$\hat{\Phi}_n^{(2)}(u|x_0) := \sum_{i=1}^n R_{n,i}^{(2)}(z_{\kappa_n}(u; \alpha_n)|x_0).$$

Let $\mu_{\kappa_n}(u; \alpha_n | x_0) = \mathbb{E}(\hat{\Phi}_n^{(2)}(u | x_0))$. Straightforward calculus leads to:

$$\begin{aligned} \frac{\mu_{\kappa_n}(u; \alpha_n | x_0)}{n|H_n|g(x_0)S(y_{\kappa_n}(u; \alpha_n) | x_0)} &= \int_{\mathcal{U}_p} \left(\frac{S(y_{\kappa_n}(u; \alpha_n) | x_0 - H_n t)}{S(y_{\kappa_n}(u; \alpha_n) | x_0)} - 1 \right) K(t) \frac{g(x_0 - H_n t)}{g(x_0)} dt \\ &+ \int_{\mathcal{U}_p} K(t) \frac{g(x_0 - H_n t)}{g(x_0)} dt. \end{aligned}$$

Let us first focus on the second term. Under condition **(A.2)**,

$$\int_{\mathcal{U}_p} K(t) \frac{g(x_0 - H_n t)}{g(x_0)} dt - 1 \leq \frac{c_g \|H_n\|_\infty}{g(x_0)} \int_{\mathcal{U}_p} K(t) \|t\|_\infty dt = \mathcal{O}(\|H_n\|_\infty) = o(\tau_n), \quad (45)$$

since $\tau_n^{-1} \|H_n\|_\infty \rightarrow 0$. Now, since $S(\cdot | x_0)$ is of class $\mathcal{ERV}_2(\gamma, \mathbf{a}_2(\cdot | x_0), \tau_n, \alpha_n)$, it is easy to check that there exists $\delta > 1$ such that for all $u \in (\eta, 1)$, $y_{\kappa_n}(u; \alpha_n) / S^{\leftarrow}(\alpha_n | x_0) \in [\delta^{-1}, \delta]$. Hence, from condition (15) and using (45)

$$\int_{\mathcal{U}_p} \left(\frac{S(y_{\kappa_n}(u; \alpha_n) | x_0 - H_n t)}{S(y_{\kappa_n}(u; \alpha_n) | x_0)} - 1 \right) K(t) \frac{g(x_0 - H_n t)}{g(x_0)} dt = o(\tau_n). \quad (46)$$

Collecting (45) and (46), the first part of condition (8) is satisfied. Theorem 2 concludes the proof. \blacksquare

Proof of Corollary 3 – Let $z_{\kappa_n}(u; \alpha_n) = S_{x_0}^{\leftarrow}(u \alpha_n) + \kappa_n a_{3, x_0}(\alpha_n^{-1})$, where κ_n is a sequence converging to zero as $n \rightarrow \infty$. First, recall that

$$\sum_{i=1}^n R_{n, i, x_0}^{(3)}(0) = n|H_n|,$$

and thus second part of condition (8) is obviously satisfied. Now let

$$\hat{\Phi}_{n, x_0}^{(3)}(u) = \sum_{i=1}^n R_{n, i, x_0}^{(3)}(z_{\kappa_n}(u; \alpha_n)),$$

and denote by $\mu_{\kappa_n, x_0}(u; \alpha_n)$ its expectation. Recalling that

$$\sum_{i=1}^n \int_{A_i} K_{H_n}(x_0 - t) dt = 1,$$

it is easy to check that,

$$\frac{\mu_{\kappa_n, x_0}(u; \alpha_n)}{n|H_n|S_{x_0}(z_{\kappa_n}(u; \alpha_n))} = 1 + \sum_{i=1}^n \int \left(\frac{S_{x_{n, i}}(z_{\kappa_n}(u; \alpha_n))}{S_{x_0}(z_{\kappa_n}(u; \alpha_n))} - 1 \right) \mathbb{I}_{\{t \in A_i \cap B(x_0, H_n)\}} K_{H_n}(x_0 - t) dt.$$

Since $n\bar{\nu}_n \leq C_\nu$, for all $i \in \{1, \dots, n\}$, $A_i \subset B(x_i, r_{n, p} I_p)$ with $2r_{n, p} := (C_\nu/n)^{1/p}$. Thus if $A_i \cap B(x_0, H_n) \neq \emptyset$, $x_i \in B(x_0, 2H_n)$. Indeed, if there exists $t \in A_i \cap B(x_0, H_n)$,

$$\|H_n^{-1}(x_i - x_0)\|_\infty \leq \|H_n^{-1}(x_i - t)\|_\infty + \|H_n^{-1}(t - x_0)\|_\infty \leq \|H_n^{-1}\|_\infty r_{n, p} + 1.$$

By assumption $n|H_n| \rightarrow \infty$ and since for all positive-definite matrix M of size p , $|M|^{1/p} \leq \|M\|_\infty$, one has for n large enough that $\|H_n^{-1}\|_\infty r_{n, p} \leq 1$ and thus, $\|H_n^{-1}(x_i - x_0)\|_\infty \leq 2$ i.e. $x_{n, i} \in B(x_0, 2H_n)$. Furthermore, since S_{x_0} is of class $\mathcal{ERV}_2(\gamma, \mathbf{a}_{3, x_0}, \tau_n, \alpha_n)$, it is easy to check that there exists $\delta > 1$ such that for all $u \in (\eta, 1)$ $z_{\kappa_n}(u; \alpha_n) / S_{x_0}^{\leftarrow}(\alpha_n) \in [\delta^{-1}, \delta]$. Hence, from condition (18),

$$\mu_{\kappa_n, x_0}(u; \alpha_n) = n|H_n|S_{x_0}(z_{\kappa_n}(u; \alpha_n)) (1 + o(\tau_n)),$$

proving the first part of condition (8). Theorem 2 concludes the proof. \blacksquare

References

- [1] Beirlant, J., Goegebeur, Y., Segers, J. and Teugels J. (2004). *Statistics of extremes: Theory and applications*, Wiley.
- [2] Chavez-Demoulin, V. and Davison, A.C. (2005). Generalized additive modelling of sample extremes, *Journal of the Royal Statistical Society, Series C*, **54**, 207–222.
- [3] Daouia, A., Gardes, L. and Girard, S. (2013). On kernel smoothing for extremal quantile regression, *Bernoulli*, **19(5B)**, 2557–2589.
- [4] Davison, A.C. and Ramesh, N.I. (2000). Local likelihood smoothing of sample extremes, *Journal of the Royal Statistical Society, Series B*, **62**, 191–208.
- [5] Davison, A.C. and Smith, R.L. (1990). Models for exceedances over high thresholds, *Journal of the Royal Statistical Society, Series B*, **52**, 393–442. MR1086795
- [6] Dekkers, A. L. M., Einmahl, J. H. J. and de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution, *The Annals of Statistics*, **17(4)**, 1833–1855.
- [7] Dubhashi, D. and Parconesi, A. (2009). *Concentration of measure for the analysis of randomized algorithms*, Cambridge University Press, New York.
- [8] Drees, H. (1995). Refined Pickands estimators of the extreme value index, *The Annals of Statistics*, **23**, 2059–2080.
- [9] Einmahl, J.H.J., Fils-Villetard, A. and Guillou, A. (2008). Statistics of extremes under random censoring, *Bernoulli*, **14(1)**, 207–227.
- [10] Einmahl, J. H. J., de Haan, L. and Zhou, C. (2014). Statistics of Heteroscedastic Extremes, *CentER Discussion Paper Series*, **No. 2014-015**.
- [11] Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample, *Proceedings of the Cambridge Philosophical Society*, **24**, 180–190.
- [12] Fraga Alves, M. I., Gomes, M. I., de Haan, L. and Neves, C. (2007). A note on second order conditions in extreme value theory: linking general and heavy tail conditions, *REVSTAT - Statistical Journal*, **5(3)**, 285–304.
- [13] Gardes, L. and Girard, S. (2008). A moving window approach for nonparametric estimation of the conditional tail index, *Journal of Multivariate Analysis*, **99**, 2368–2388.
- [14] Gardes, L. and Girard, S. (2010). Conditional extremes from heavy-tailed distributions: an application to the estimation of extreme rainfall return levels, *Extremes*, **13(2)**, 177–204.
- [15] Gardes, L. and Girard, S. (2012). Functional kernel estimators of large conditional quantiles, *Electronic Journal of Statistics*, **6**, 1715–1744.
- [16] Gardes, L. and Stupfler, G. (2014). Estimation of the conditional tail index using a smoothed local Hill estimator, *Extremes*, **17(1)**, 45–75.
- [17] Gardes, L. and Stupfler, G. (2014). Estimating extreme quantiles under random truncation, *TEST*, DOI: 10.1007/s11749-014-0403-5.
- [18] Gasser, T. and Müller, H. (1979). Kernel estimation of regression functions, *Curve Estimation*, Springer-Verlag, Heidelberg.
- [19] Georgiev, A.A. (1988). Consistent nonparametric multiple regression: the fixed design case, *Journal of Multivariate Analysis*, **25**, 100–110.
- [20] Gnedenko, B. V. (1943). Sur la distribution limite du terme maximum d’une série aléatoire, *Annals of Mathematics*, **44**, 423–453.

- [21] Goegebeur, Y., Guillou, A. and Osmann, M. (2014). A local moment type estimator for the extreme value index in regression with random covariates, *Canadian Journal of Statistics*, to appear.
- [22] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*, Springer.
- [23] de Haan, L., Klein Tank, A. and Neves, C. (2011). On tail trend detection: modeling relative risk, *arXiv preprint*, *arXiv:1106.4149*.
- [24] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution, *The Annals of Statistics*, **3(5)**, 1163–1174.
- [25] Nadaraya, E.A. (1964). On estimating regression, *Theory of Probability and its Application*, **9(1)**, 141–142.
- [26] Parzen, E. (1962). On the estimation of a probability density function and mode, *Annals of Mathematical Statistics*, **33**, 1065–1076.
- [27] Pickands, III J. (1975). Statistical inference using extreme order statistics, *The Annals of Statistics*, **3(1)**, 119–131.
- [28] Smith, R.L. (1989). Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone (with discussion), *Statistical Science*, **4**, 367–393.
- [29] Stone, C. J. (1977). Consistent nonparametric regression, *Annals of Statistics*, **5**, 595–645.
- [30] Stupfler, G. (2013). A moment estimator for the conditional extreme-value index, *Electronic Journal of Statistics*, **7**, 2298–2343.
- [31] Tsay, R.S. (2002). *Analysis of financial time series*. Wiley, New York.
- [32] Wang, H. and Tsai, C.L. (2009). Tail index regression, *Journal of the American Statistical Association*, **104(487)**, 1233–1240.
- [33] Watson, G. S. (1964). Smooth regression analysis, *Sankhya*, **26(15)**, 175–184.

η	P1-1				P1-2			
	$n = 50$	$n = 100$	$n = 200$	$n = 400$	$n = 50$	$n = 100$	$n = 200$	$n = 400$
0.005	0.194	0.098	0.056	0.035	0.682	0.320	0.176	0.112
0.01	0.195	0.100	0.057	0.035	0.700	0.330	0.182	0.116
0.02	0.210	0.102	0.059	0.037	0.735	0.349	0.193	0.123
0.04	0.229	0.112	0.067	0.044	0.804	0.390	0.218	0.139
0.08	0.303	0.133	0.093	0.056	0.960	0.478	0.276	0.172
0.16	0.438	0.200	0.158	0.088	1.338	0.699	0.419	0.259

η	P1-3			
	$n = 50$	$n = 100$	$n = 200$	$n = 400$
0.005	0.108	0.077	0.044	0.030
0.01	0.112	0.082	0.048	0.034
0.02	0.121	0.091	0.056	0.040
0.04	0.143	0.110	0.068	0.050
0.08	0.201	0.166	0.096	0.069
0.16	0.360	0.292	0.159	0.112

Table 1: Values for different values of η and n of $\text{EMSE}(\gamma_{n,\varphi}(\hat{S}_n^{(1)}; n^{-1/3}))$ in the case $\varphi = 1$ using the generating processes **P1-1**, **P1-2** and **P1-3** with $\tilde{c}(\cdot) = 1 + \log(1/\cdot)$.

η	P1-1				P1-2			
	$n = 50$	$n = 100$	$n = 200$	$n = 400$	$n = 50$	$n = 100$	$n = 200$	$n = 400$
0.005	0.242	0.129	0.074	0.047	0.268	0.151	0.085	0.054
0.01	0.228	0.125	0.073	0.046	0.278	0.158	0.089	0.057
0.02	0.224	0.122	0.070	0.045	0.299	0.171	0.096	0.063
0.04	0.221	0.124	0.069	0.050	0.341	0.197	0.113	0.075
0.08	0.265	0.140	0.090	0.057	0.440	0.256	0.148	0.100
0.16	0.373	0.193	0.134	0.080	0.634	0.404	0.241	0.166

η	P1-3			
	$n = 50$	$n = 100$	$n = 200$	$n = 400$
0.005	0.094	0.063	0.034	0.024
0.01	0.095	0.066	0.038	0.027
0.02	0.100	0.072	0.044	0.031
0.04	0.112	0.082	0.053	0.039
0.08	0.149	0.124	0.076	0.056
0.16	0.273	0.221	0.122	0.097

Table 2: Values for different values of η and n of $\text{EMSE}(\gamma_{n,\varphi}(\hat{S}_n^{(1)}; n^{-1/3}))$ in the case $\varphi = \log(1/\cdot)$ using the generating processes **P1-1**, **P1-2** and **P1-3** with $\tilde{c}(\cdot) = 1 + \log(1/\cdot)$.

α_n	P1-1		P1-2		P1-3	
	$\tilde{c} = 1$	$\tilde{c}(\cdot) = 1 + \log(1/\cdot)$	$\tilde{c} = 1$	$\tilde{c}(\cdot) = 1 + \log(1/\cdot)$	$\tilde{c} = 1$	$\tilde{c}(\cdot) = 1 + \log(1/\cdot)$
$n^{-1/2}$	1.018	0.725	0.195	0.141	0.484	0.534
$n^{-1/3}$	1.211	1.070	0.402	0.352	0.809	0.735
$n^{-1/4}$	1.232	1.151	0.428	0.380	0.835	0.930
$n^{-1/5}$	1.210	1.177	0.400	0.252	0.933	0.865
$n^{-1/6}$	1.274	1.172	0.433	0.284	1.000	0.811

Table 3: Comparison of $\gamma_{n,\varphi}(\hat{S}_n^{(1)}; \alpha_n)$ and the moment estimator: values of the ratio $R_1(\alpha_n)$ with $\varphi(\cdot) = \log(1/\cdot)$ and $\tau = 0.02$ using the generating processes **P1-1**, **P1-2** and **P1-3** with different values of α_n and different functions \tilde{c} .

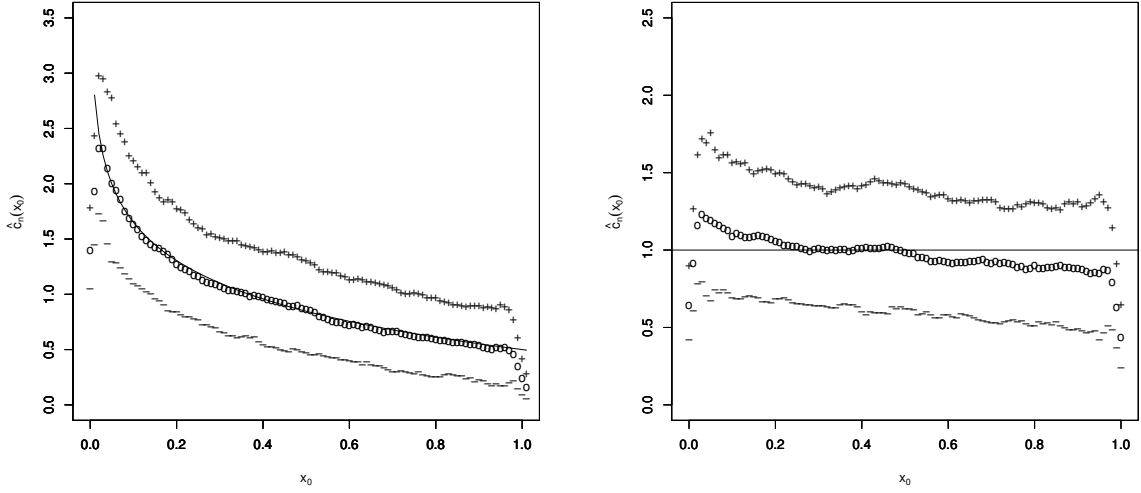


Figure 1: Estimation of $c(x_0)$ for different values of x_0 and for data generated by **P1-3** with $\theta = 1$ (left) and $\theta = 1/2$ (right). The points \circ represent the median over the $N = 500$ replications while the points $-$ and $+$ represent respectively the quantiles of order 0.1 and 0.9. The full line is the true function c .

$(\alpha_{n,i}, H_{n,i,j})$	P2-1					P2-2				
	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$
$i = 2$	0.390	0.713	0.653	1.051	1.066	0.055	0.139	0.179	0.312	0.327
$i = 3$	0.666	1.089	1.130	1.078	1.036	0.109	0.288	0.571	0.600	0.698
$i = 4$	0.789	1.079	1.268	1.136	1.132	0.208	0.435	0.617	0.692	0.833
$i = 5$	0.936	1.121	1.352	1.286	1.297	0.295	0.454	0.557	0.596	0.642
$i = 6$	0.885	1.256	1.400	1.371	1.355	0.223	0.470	0.580	0.622	0.686

$(\alpha_{n,i}, H_{n,i,j})$	P2-3				
	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$
$i = 2$	0.018	0.315	0.423	0.407	0.494
$i = 3$	0.214	0.586	0.802	1.058	0.882
$i = 4$	0.457	0.842	1.065	1.167	1.118
$i = 5$	0.354	1.062	1.355	1.172	1.103
$i = 6$	0.532	1.148	1.300	1.148	1.077

Table 4: Comparison of $\gamma_{n,\varphi}(\hat{S}_n^{(3)}(\cdot|x_0); \alpha_n)$ and the estimator $\hat{\gamma}_S(\alpha_n, H_n|x_0)$: values of the ratio $R_2(\alpha_n, H_n)$ with $\varphi(\cdot) = \log(1/\cdot)$ and $\tau = 0.02$ using the generating processes **P2-1**, **P2-2** and **P2-3** for $(\alpha_{n,i} = n^{-1/i}, H_{n,i,j} = n^{(1-i)/(ij)})$, $i, j = 2, \dots, 6$.