HAL
open science

Estimating the conditional tail index by integrating a kernel conditional quantile estimator

Laurent Gardes, Armelle Guillou, Antoine Schorgen

To cite this version:

Laurent Gardes, Armelle Guillou, Antoine Schorgen. Estimating the conditional tail index by integrating a kernel conditional quantile estimator. Journal of Statistical Planning and Inference, 2012, 142 (6), pp.1586-1598. 10.1016/j.jspi.2012.01.011 . inria-00578479

HAL Id: inria-00578479

https://inria.hal.science/inria-00578479

Submitted on 21 Mar 2011

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.

Estimating the conditional tail index by integrating a kernel conditional quantile estimator

L. Gardes[1], A. Guillou[2] and A. Schorgen[2]

[1] INRIA Rhône-Alpes, projet Mistis, Inovallée, 655 av. de l'Europe,
38334 Montbonnot, Saint-Ismier cedex, France
[2] Université de Strasbourg et CNRS, IRMA, UMR 7501, 7 rue René Descartes,
67084 Strasbourg cedex, France

Abstract

This paper deals with the estimation of an extreme value index of a heavy-tailed distribution in the presence of covariates. A class of estimators is proposed in this context and its asymptotic normality established under mild regularity conditions. These estimators are functions of a kernel conditional quantile estimator depending on some tuning parameters. The finite sample properties of our estimators are illustrated on a small simulation study.

Keywords: Heavy-tailed distribution; Covariates; Kernel estimator; Asymptotic normality.

AMS Subject Classifications: 62G32, 62G05, 62E20.

Corresponding author: Armelle Guillou. E-mail: armelle.guillou@math.unistra.fr; Tél: (+33) 3 68 85 01 99; Fax: (+33) 3 68 85 03 28.

1 Introduction

Tail index estimation has been extensively studied in the literature and several estimators proposed. The most popular estimator was proposed by Hill (1975) in the context of heavy-tailed distributions. Diebolt *et al.* (2008) considered the case of Weibull-tail distributions and the general case was studied by Dekkers *et al.* (1989). The aim of this paper is to extend this framework to the case where some covariate x is recorded simultaneously with the variable of interest Y. In this case, the tail index depends on the covariate x and is referred in the following to as the conditional tail index.

Such a context has been already considered by Smith (1989) or Davison and Smith (1990) who proposed parametric models, while semi-parametric approaches has been used in Hall and Tajvidi (2000) or Beirlant and Goegebeur (2003). Also, fully non-parametric methods can be used for instance based on splines (see Chavez-Demoulin and Davison, 2005) or local polynomials (see Davison and Ramesh, 2000). These latter estimators have been then extended by Beirlant and Goegebeur (2003) in case of multidimensional covariates. More recently, Gardes and Girard (2010) have addressed the problem of the estimation of a conditional extreme quantile via a nearest neighbour approach.

We propose here new conditional extreme value index estimators based on a class of functions satisfying some mild conditions. We only assume that the conditional distribution of Y given x is heavy-tailed whereas no parametric assumption on the covariates x is made. The conditional distribution of Y given x is then of polynomial-type, with a rate of convergence driven by the conditional tail index.

The remainder of this paper is organized as follows. Our class of estimators is given in Section 2. Their asymptotic normality is established in Section 3. In Section 4, a sub-class of estimators is studied and their finite sample properties are illustrated in Section 5. All the proofs are postponed to Section 6.

2 A class of conditional tail index estimators

Let E be a metric space associated to a distance d. For $x \in E$, denote by $F(.,x)$ (resp. $q(.,x)$) the conditional distribution function (resp. the conditional quantile function) of Y given x. We assume that for all $x \in E$,

$$\overline{F}(y,x) = y^{-1/\gamma(x)} L(y,x), \tag{1}$$

or, equivalently, for $\alpha \in (0,1)$,

$$q(\alpha,x) = \overline{F}^{\leftarrow}(\alpha,x) = \alpha^{-\gamma(x)} \ell(\alpha^{-1},x),$$

where $\overline{F}^{\leftarrow}(\alpha,x) = \inf\{t,\ \overline{F}(t,x) \leq \alpha\}$ denotes the generalized inverse of the conditional survival function and, for x fixed, $L(.,x)$ and $\ell(.,x)$ are slowly varying functions, that is for all $\lambda > 0$,

$$\lim_{y\to\infty} \frac{L(\lambda y,x)}{L(y,x)} = \lim_{y\to\infty} \frac{\ell(\lambda y,x)}{\ell(y,x)} = 1.$$

Here, $\gamma(.)$ is a unknown positive function of the covariate referred in the following to as the conditional tail index. For a given $x \in E$, our aim is to propose an estimator of $\gamma(x)$. Suppose that we have at our disposal pairs $(Y_1,x_1),...,(Y_n,x_n)$ of independent observations from model (1) where the design points $x_1,\ldots,x_n$ are assumed to be non random. We propose to estimate the conditional tail index as a function of the kernel conditional quantile estimator $\widehat{q}_n(\alpha,x) := \widehat{\overline{F}}_n^{\leftarrow}(\alpha,x)$, with

$$\widehat{\overline{F}}_n(y,x) = \sum_{i=1}^n H\left(\frac{d(x,x_i)}{h_{1,n}}\right) K\left(\frac{y-Y_i}{h_{2,n}}\right) \Bigg/ \sum_{i=1}^n H\left(\frac{d(x,x_i)}{h_{1,n}}\right), \tag{2}$$

where $h_{1,n}$ and $h_{2,n}$ are non random positive sequences. This kernel estimator was defined for instance in Ferraty and Vieu (2006). The function $H(.)$ is an asymmetrical kernel with support in $[0,1]$ and the function $K(.)$ is defined by

$$K(v) = \int_v^\infty g(s)ds,$$

where $g(.)$ is a bounded probability function with support included in $[-1,1]$. Our class of conditional tail index estimators is given by

$$\widehat{\gamma}(u_{n,x},x) = \int_0^{u_{n,x}} \Psi(\alpha,u_{n,x},x) \log \widehat{q}_n(\alpha,x) d\alpha, \tag{3}$$

where for all $u \in (0,1)$, $\Psi(.,u,x)$ is a non-null continuous function in $L_1(0,1)$ such that $\int_0^u \Psi(\alpha,u,x)d\alpha = 0$ and $u_{n,x}$ is a positive sequence. The sequence $u_{n,x}$ is introduced in order to select only the largest observation to estimate $\gamma(x)$. Its choice is thus very important in practice. Note that the estimator $\widehat{\gamma}(u_{n,x},x)$ depends only on the Y_i's for which the corresponding x_i's belong to the ball $B(x,h_{1,n}) = \{t \in E;\ d(t,x) \leq h_{1,n}\}$. The smoothness of the function $\widehat{\overline{F}}_n(.,x)$ is also controlled by the bandwidth $h_{2,n}$. Finally, note that $\widehat{\overline{F}}_n(y,x) \sim \overline{\mathbb{F}}_n(y,x)$ as $h_{2,n} \to 0$, where

$$\overline{\mathbb{F}}_n(y,x) = \sum_{i=1}^n H\left(\frac{d(x,x_i)}{h_{1,n}}\right) \mathbb{1}_{\{Y_i>y\}} \Bigg/ \sum_{i=1}^n H\left(\frac{d(x,x_i)}{h_{1,n}}\right),$$

is the empirical estimator of the conditional survival function. The same result is established for the associated conditional quantile in Lemma 1. Note also that in the special case where $H(.)$ is the indicator function, $\overline{\mathbb{F}}_n(.,x)$ leads to the estimator proposed in Gardes and Girard (2008).

3 Main result

In this section, we give the useful notations and assumptions in order to establish the asymptotic normality of our estimators.

(A.1) The slowly varying function $\ell(.,x)$ is normalized.

Assumption **(A.1)** is equivalent to suppose that for all $y > 1$,

$$\ell(y,x) = c(x)\exp\int_1^y \frac{\Delta(v,x)}{v}dv,$$

where $c(x) > 0$ and $\Delta(v,x) \to 0$ as $v \to \infty$.

(A.2) The function $|\Delta(.,x)|$ is ultimately decreasing.

The largest oscillation of the log-quantile function with respect to its second variable is defined for all $a \in (0,1/2)$ as:

$$\omega_n(a) = \sup\left\{\left|\log\frac{q(\alpha,x)}{q(\alpha,x')}\right|,\ \alpha \in (a,1-a),\ (x,x') \in B(x,h_{1,n}) \times B(x,h_{1,n})\right\}.$$

(A.3) The function $|\Delta(.,x)|$ is regularly varying with index $\rho(x) < 0$, that is for all $\lambda > 0$, $|\Delta(\lambda y,x)|/|\Delta(y,x)| \to \lambda^{\rho(x)}$ as $y \to \infty$. Moreover $\Delta(.,x)$ is assumed to have a constant sign at infinity.

Conditions **(A.1)** and **(A.3)** imply that for all $v > 0$,

$$\log\frac{\ell(vy,x)}{\ell(y,x)} = \Delta(y,x)\frac{v^{\rho(x)}-1}{\rho(x)}(1+o(1)) \text{ as } y \to \infty,$$

which is the so-called second order condition classically used to establish the asymptotic normality of tail index estimators. Note that the second order parameter $\rho(x)$ controls the rate of convergence of $\ell(\lambda y,x)/\ell(y,x)$ to 1. In particular if $\rho(x)$ is close to 0, this convergence is slow and thus the estimation of the conditional tail index is difficult. Let $m_x = m_{n,x}$ be the number of covariates in the ball $B(x,h_{1,n})$. The covariates belonging to the ball $B(x,h_{1,n})$ are denoted by $\{x_i^*, i = 1,\ldots,m_x\}$ and the associated observations by $\{Y_i^*, i = 1,\ldots,m_x\}$. The order statistics are denoted by $Y_{1,m_x}^* \leq \ldots \leq Y_{m_x,m_x}^*$ and the corresponding covariates by $\{x_{(1)}^*,\ldots,x_{(m_x)}^*\}$. We introduce the set $\{\beta_0,\ldots,\beta_{m_x}\}$ where $\beta_0 = 0$, $\beta_{m_x} = 1$ and, for $i = 1,\ldots,m_x-1$,

$$\beta_i = \sum_{j=m_x-i+1}^{m_x} H\left(\frac{d(x,x_{(j)}^*)}{h_{1,n}}\right) \Bigg/ \sum_{j=1}^{m_x} H\left(\frac{d(x,x_{(j)}^*)}{h_{1,n}}\right). \tag{4}$$

The next two conditions are related to the kernel function $H(.)$ and the function $\Psi(.)$. There are similar in spirit to the ones used by Beirlant *et al.* (2002). Let $k_x = k_{n,x}$ be a sequence of integers such that $1 < k_x < m_x$.

(B.1) There exists two positive functions $f_1(.,x)$ and $f_2(.,x)$ such that for all $i = 1,\ldots,k_x$,

$$f_1\left(\frac{i}{k_x},x\right) \leq \frac{k_x}{i}\int_0^{\beta_i}\Psi(\alpha,\beta_{k_x},x)d\alpha \leq f_2\left(\frac{i}{k_x},x\right),$$

satisfying

$$0 < \int_0^1 f_1^2(\alpha,x)d\alpha,$$

and for some $\xi > 0$,

$$\int_0^1 f_2^{2+\xi}(\alpha,x)d\alpha < \infty.$$

(B.2) There exists a positive function $f_3(.,x)$ such that for $i=1,\dots,k_x$

$$k_x\left|\int_{\beta_{i-1}}^{\beta_i}\Psi(\alpha,\beta_{k_x},x)d\alpha\right|\leq f_3\left(\frac{i}{k_x},x\right),$$

with

$$\int_0^1 f_3(\alpha,x)\max(\log(1/\alpha),1)d\alpha<\infty.$$

We are now in position to state our main result. We consider the case where the threshold $u_{n,x}$ in (3) is taken equal to β_{k_x}. Thus, the choice of the sequence $u_{n,x}\in(0,1)$ is replaced by the choice of a sequence of integers $k_x\in\{2,\dots,m_x-1\}$.

Theorem 1. *Suppose* **(A.1)**, **(A.2)**, **(A.3)**, **(B.1)** *and* **(B.2)** *hold. If* $k_x\to\infty$, $m_x/k_x\to\infty$,

$$k_x^{1/2}\Delta\left(\frac{m_x}{k_x},x\right)\to\lambda(x)\in\mathbb{R},\ \frac{k_x^{1/2}h_{2,n}}{q(k_x/m_x,x)}\to 0\ \textit{and}\ k_x^2\omega_n(m_x^{-(1+\delta)})\to 0,$$

for some $\delta>0$ *then*

$$\sigma_{n,x}\left(\widehat{\gamma}(\beta_{k_x},x)-\gamma(x)-\gamma(x)b_{1,n}(x)-\Delta\left(\frac{m_x}{k_x},x\right)b_{2,n}(x)\right),$$

converges in distribution to a $\mathcal{N}(0,\gamma^2(x))$ *random variable with*

$$\begin{aligned}\sigma_{n,x}^2 &= \left(\sum_{i=1}^{k_x}\left(\frac{1}{i}\int_0^{\beta_i}\Psi(\alpha,\beta_{k_x},x)d\alpha\right)^2\right)^{-1}\\ b_{1,n}(x) &= \sum_{i=1}^{k_x}\left(\frac{1}{i}\int_0^{\beta_i}\Psi(\alpha,\beta_{k_x},x)d\alpha\right)-1\\ b_{2,n}(x) &= \sum_{i=1}^{k_x}\left(\frac{1}{i}\int_0^{\beta_i}\Psi(\alpha,\beta_{k_x},x)d\alpha\right)\left(\frac{i}{k_x+1}\right)^{-\rho(x)}.\end{aligned}$$

The asymptotic bias is divided into two terms. The first one ($b_{1,n}(x)$) only depends on the functions $\Psi(.)$ and $H(.)$. The second term ($b_{2,n}(x)$) depends on the unknown second order parameter $\rho(x)$. Minimizing this second bias term requires the estimation of $\rho(x)$ (we refer to Cairo and Gomes (2002) for estimators of the second order parameter in the unconditional case). Note that, in order to cancel the first bias term $b_{1,n}(x)$, one can define the following estimator :

$$\widehat{\gamma}^*(\beta_{k_x},x)=\widehat{\gamma}(\beta_{k_x},x)\Bigg/\sum_{i=1}^{k_x}\left(\frac{1}{i}\int_0^{\beta_i}\Psi(\alpha,\beta_{k_x},x)d\alpha\right). \tag{5}$$

The asymptotic normality of this estimator is a straighforward consequence of Theorem 1.

Corollary 1. *Under the assumptions of Theorem 1,*

$$\sigma_{n,x}^*\left(\widehat{\gamma}^*(\beta_{k_x},x)-\gamma(x)-\Delta\left(\frac{m_x}{k_x},x\right)b_{2,n}^*(x)\right),$$

converges in distribution to a $\mathcal{N}(0,\gamma^2(x))$ *random variable with*

$$\begin{aligned}(\sigma_{n,x}^*)^2 &= \left(\sum_{i=1}^{k_x}\frac{1}{i}\int_0^{\beta_i}\Psi(\alpha,\beta_{k_x},x)d\alpha\right)^2\Bigg/\sum_{i=1}^{k_x}\left(\frac{1}{i}\int_0^{\beta_i}\Psi(\alpha,\beta_{k_x},x)d\alpha\right)^2\\ b_{2,n}^*(x) &= \sum_{i=1}^{k_x}\left(\frac{1}{i}\int_0^{\beta_i}\Psi(\alpha,\beta_{k_x},x)d\alpha\right)\left(\frac{i}{k_x+1}\right)^{-\rho(x)}\Bigg/\sum_{i=1}^{k_x}\left(\frac{1}{i}\int_0^{\beta_i}\Psi(\alpha,\beta_{k_x},x)d\alpha\right).\end{aligned}$$

4 A particular case

In order to simplify the assumptions of Theorem 1, we consider the case where the function $H(.)$ is given by

$$H(x) = \mathbb{1}_{\{x\in[0,1]\}}. \tag{6}$$

In this situation, $\beta_i = i/m_x$ for all $i = 0, \ldots, m_x$. Assumptions **(B.1)** and **(B.2)** are replaced by the following assumptions:

(C.1) Let $u \in (0,1)$. There exists a decreasing function $\Phi(.,x)$ such that for all $\beta \in (0,u]$ and for some $\xi > 0$:

$$\frac{u}{\beta}\int_0^\beta \Psi(\alpha,u,x)d\alpha = \Phi(\beta/u,x)$$

$$\text{with } \int_0^1 \Phi(\alpha,x)d\alpha = 1 \text{ and } \int_0^1 \Phi^{2+\xi}(\alpha,x)d\alpha < \infty.$$

(C.2) There exists a positive function $f_4(.,x)$ such that for $i = 1, \ldots, k_x$

$$\left| i\Phi\left(\frac{i}{k_x},x\right) - (i-1)\Phi\left(\frac{i-1}{k_x},x\right)\right| \leq f_4\left(\frac{i}{k_x},x\right),$$

with

$$\int_0^1 f_4(\alpha,x)\max(\log(1/\alpha),1)d\alpha < \infty.$$

We denote by $\widehat{\gamma}(k_x/m_x,x)$ the estimators obtained by taking the function $H(.)$ as in (6). The following result is a consequence of Theorem 1:

Corollary 2. *Suppose* **(A.1)**, **(A.2)**, **(A.3)**, **(C.1)** *and* **(C.2)** *hold. If* $k_x \to \infty$, $m_x/k_x \to \infty$,

$$k_x^{1/2}\int_0^{1/k_x}\Phi(\alpha,x)d\alpha \to 0,$$

$$k_x^{1/2}\Delta\left(\frac{m_x}{k_x},x\right) \to \lambda(x) \in \mathbb{R},\ \frac{k_x^{1/2}h_{2,n}}{q(k_x/m_x,x)} \to 0 \text{ and } k_x^2\omega_n(m_x^{-(1+\delta)}) \to 0,$$

for some $\delta > 0$ *then*

$$k_x^{1/2}\left(\widehat{\gamma}(k_x/m_x,x) - \gamma(x) - \Delta\left(\frac{m_x}{k_x},x\right)\mathcal{AB}(\Phi,\rho(x),x)\right)$$

converges to a $\mathcal{N}(0,\gamma^2(x)\mathcal{AV}(\Phi,x))$ *random variable with:*

$$\mathcal{AB}(\Phi,\rho(x),x) = \int_0^1 \Phi(\alpha,x)\alpha^{-\rho(x)}d\alpha \text{ and } \mathcal{AV}(\Phi,x) = \int_0^1 \Phi^2(\alpha,x)d\alpha.$$

As an example of functions satisfying the conditions of Corollary 2, one can mention for all $\alpha < u$:

$$\Psi_\theta(\alpha,u,x) := \begin{cases} \dfrac{(\theta+1)^2}{\theta u^{\theta+1}}\left(\dfrac{u^\theta}{\theta+1} - \alpha^\theta\right) & \text{if } \theta > 0 \\[2ex] \dfrac{1}{u}\left(\log\left(\dfrac{u}{\alpha}\right) - 1\right) & \text{if } \theta = 0 \\[2ex] \dfrac{1}{u} & \text{if } \theta = \infty. \end{cases}$$

Note that $\Psi_0(.,.,.)$ and $\Psi_\infty(.,.,.)$ correspond to the limiting cases of $\Psi_\theta(.,.,.)$.

Denote by $\check{\gamma}_\theta(x)$ the estimator obtained by taking the function $H(.)$ as in (6) and by using the function $\Psi_\theta(.,u,x)$. Its asymptotic normality is a direct consequence of Corollary 2.

Corollary 3. *Suppose* **(A.1)**, **(A.2)**, **(A.3)** *hold. If* $k_x \to \infty$, $m_x/k_x \to \infty$,

$$k_x^{1/2}\Delta\left(\frac{m_x}{k_x}, x\right) \to \lambda(x) \in \mathbb{R}, \ \frac{k_x^{1/2}h_{2,n}}{q(k_x/m_x, x)} \to 0 \text{ and } k_x^2\omega_n(m_x^{-(1+\delta)}) \to 0,$$

for some $\delta > 0$ *then*

$$k_x^{1/2}\left(\check{\gamma}_\theta(x) - \gamma(x) - \Delta\left(\frac{m_x}{k_x}, x\right)\frac{\theta+1}{(1-\rho(x))(1+\theta-\rho(x))}\right)$$

converges to a $\mathcal{N}(0, 2\gamma^2(x)(\theta+1)/(1+2\theta))$ *random variable.*

Note that when $\theta = 0$, $\mathcal{AB}(\Phi_0, \rho(x), x) = 1/(1-\rho(x))^2$ and $\mathcal{AV}(\Phi_0, x) = 2$ which are the same values than the ones obtain for the conditional Zipf estimator defined in Gardes and Girard (2008). This estimator is an extension in the conditional framework of the Zipf estimator introduced by Kratz and Resnick (1996) and Schultze and Steinebach (1996). When $\theta = \infty$, $\mathcal{AB}(\Phi_\infty, \rho(x), x) = 1/(1-\rho(x))$ and $\mathcal{AV}(\Phi_\infty, x) = 1$, which correspond to the asymptotic bias and variance of the conditional Hill estimator proposed again in Gardes and Girard (2008). To choose the value of θ, one can use for instance the criterion introduced in Gardes and Girard (2010). More precisely, we introduce a modified version of the Asymptotic Mean Squared Error (AMSE) as follows:

$$\begin{aligned} AMSE^* &:= \Delta^2\left(\frac{m_x}{k_x}, x\right)\int_{-\infty}^0 \mathcal{AB}^2(\Phi_\theta, \rho, x)d\rho + \frac{\gamma^2(x)\mathcal{AV}(\Phi_\theta, x)}{k_x} \\ &= \frac{\pi(\theta)}{k_x}\left\{\frac{k_x\Delta^2\left(\frac{m_x}{k_x}, x\right)}{\mathcal{AV}(\Phi_\theta, x)} + \frac{\gamma^2(x)}{\int_{-\infty}^0 \mathcal{AB}^2(\Phi_\theta, \rho, x)d\rho}\right\} \leq C(x)\frac{\pi(\theta)}{k_x} \end{aligned}$$

where $C(x)$ is a constant independent of θ and

$$\begin{aligned} \pi(\theta) &= \int_{-\infty}^0 \left(\frac{\theta+1}{(1-\rho)(1+\theta-\rho)}\right)^2 d\rho \times \frac{2(1+\theta)}{1+2\theta} \\ &= 2\left(\frac{\theta+1}{\theta}\right)^3 \frac{1}{1+2\theta}(\theta - 2\log(1+\theta) - 1/(1+\theta) + 1). \end{aligned}$$

Therefore, a good choice of θ could be to take θ_π minimizing the upper bound $\pi(\theta)$. Numerical optimization leads to $\theta_\pi \approx 0.6833$.

5 A small simulation study

In this section, we illustrate the efficiency of our estimator $\check{\gamma}_\theta(.)$ proposed in Section 4 on a small simulation study. Keeping the notations of Sections 3 and 4 in mind, we recall the expression of our estimator

$$\check{\gamma}_\theta(x) = \int_0^{k_x/m_x} \Psi_\theta(\alpha, k_x/m_x, x)\log\widehat{q}_n(\alpha, x)d\alpha. \tag{7}$$

The asymmetric kernel $H(.)$ is taken as the indicator function (6) and the integrated kernel $K(.)$ is chosen as the integrated biweight kernel:

$$K(x) = \int_x^\infty \frac{15}{16}(1-s^2)^2 \mathbb{1}_{\{s\in[-1,1]\}}ds.$$

Concerning the parameter θ, it will be fixed at the value $\theta_\pi = 0.6833$ as explained in Section 4. Moreover, we assume that the conditional survival function of Y given the covariate x is a Burr-type distribution, that is

$$\overline{F}(y, x) = \left(1 + y^{-\rho/\gamma(x)}\right)^{1/\rho}, \tag{8}$$

where $\rho < 0$ is the second order parameter defined in condition **(A.3)**. In this particular setting, the slowly varying function $\ell(.,x)$ and $\Delta(.,x)$ are given by

$$\ell(z,x) = (1-z^{\rho})^{-\gamma(x)/\rho},$$
$$\Delta(z,x) = \gamma(x)\,\frac{z^{\rho}}{1-z^{\rho}},$$

with $z > 1$.

Concerning the conditional tail index, as in Daouia *et al.* (2010), we choose

$$\gamma(x) = \frac{1}{2}\left(\frac{1}{10} + \sin(\pi x)\right)\left(\frac{11}{10} - \frac{1}{2}\exp\left(-64\left(x - \frac{1}{2}\right)^2\right)\right). \tag{9}$$

In order to proceed to our simulations, the next step is to select k_x and the two bandwidths $h_{1,n}$ and $h_{2,n}$ of kernels $H(.)$ and $K(.)$, i.e to find their optimal values according to some criterion.

5.1 Optimization of (m_x, k_x) - AMSE minimization

In our case, $H(.)$ is the indicator function. Therefore the estimator only depends on the number of points in the ball $B(x, h_{1,n})$. Thus to find the optimal numbers (m_x, k_x), we minimize:

$$AMSE(m_x,k_x) = \frac{2(\theta_\pi+1)}{2\theta_\pi+1}\frac{\gamma^2(x)}{k_x} + \left[\frac{\theta_\pi+1}{(1-\rho)(1+\theta_\pi-\rho)}\Delta\left(\frac{m_x}{k_x},x\right)\right]^2$$

with respect to these parameters.

Now, concerning the two unknown quantities $\gamma(x)$ and $\Delta(m_x/k_x, x)$, they will be estimated using an exponential regression model (see (18) in the proof of Theorem 1), taking $\rho(x) = \rho := -1$ as a canonical value. This choice is usual (see for instance Feuerverger and Hall (1999) or Beirlant *et al.* (1999)) and leads to a least squares estimator for $\gamma(x)$ and $\Delta(\frac{m_x}{k_x}, x)$, denoted respectively by $\widehat{\gamma}_{LS}(x)$ and $\widehat{\Delta}_{LS}(\frac{m_x}{k_x}, x)$. Therefore, the AMSE can be estimated by

$$\widehat{AMSE}(m_x,k_x) = \frac{2(\theta_\pi+1)}{2\theta_\pi+1}\frac{\widehat{\gamma}^2_{LS}(x)}{k_x} + \left[\frac{\theta_\pi+1}{2(\theta_\pi+2)}\widehat{\Delta}_{LS}\left(\frac{m_x}{k_x},x\right)\right]^2$$

from which we deduce our criterion of selection for (m_x, k_x) :

$$(\widehat{m}_x, \widehat{k}_x) = \underset{m_x,k_x}{\arg\min}\,\widehat{AMSE}(m_x,k_x).$$

Note that the AMSE does not depend on the bandwidth parameter $h_{2,n}$. Therefore we need another criterion to find its optimal value.

5.2 Optimization of $h_{2,n}$ - Subsample division

The task is now to find an optimal value for the bandwidth $h_{2,n}$ of the integrated kernel $K(.)$. To this end, in the spirit of Ferraty and Vieu (2006), we propose to split our sample into two random subsamples I_1 and I_2 of the same sample size. Then, we fix the covariate on x_0 (for instance taken in I_1) and find the nearest neighbour x_0^* of x_0 in I_2. Finally, we calculate the estimation $\check{\gamma}^{(1)}_{\theta_\pi}(x_0)$ of $\gamma(x_0)$ from I_1 and $\check{\gamma}^{(2)}_{\theta_\pi}(x_0^*)$ from I_2, using the optimal value for (m_x, k_x) as described in Section 5.2. The resulting optimization criterion for $h_{2,n}$ can then be proposed :

$$\widehat{h}_{2,n} = \underset{h_{2,n}}{\arg\min}\,|\check{\gamma}^{(1)}_{\theta_\pi}(x_0) - \check{\gamma}^{(2)}_{\theta_\pi}(x_0^*)|.$$

5.3 Graphs and computational issues

We simulate 100 replications of $n = 5\,000$ pairs $(Y_i, x_i), i = 1, ..., n$ where $x_i = i/n$ and the distribution of Y given x is the Burr-type distribution (8). The conditional quantile estimator $\widehat{q}_n(\alpha, x)$ is obtained by dichotomous inversion of the conditional empirical survival function defined in (2).
For a fixed value of the covariate x, we perform the optimizations of (m_x, k_x) and $h_{2,n}$ using grids:

- the number m_x of neighbours of x belongs to a grid from 2% to 10% of the sample size, i.e from $m_x = 100$ to 500 in our case; to avoid too small or too large values for k_x, this parameter has been taken in the interval $(0.15 \times m_x; 0.85 \times m_x)$;
- $h_{2,n}$ goes from 0.05 to 1 by step of 0.05.

Finally, two values of ρ were used: $\rho = -1$ and -0.8. Note that as usual in the absence of covariates, the closer ρ is to 0, the more difficult is the estimation of $\gamma(.)$.

Figure 1 below illustrates the procedure presented in Section 5.1 for $x = 0.4$. In panel (a), we plot the estimated and theoretical AMSE for one realization. The boxplot of $\widehat{k}_x$ for the 100 replications is given in panel (b). As observed, $\widehat{k}_{0.4}$ is very close with small variability to the optimal value $(k_{0.4}^{opt})$ which minimizes the AMSE.
Now, we come back to our original problem, that is the estimation of $\gamma(.)$. Again we simulate 100 replications from the Burr-type distribution (8). The L^2-errors between $\widehat{\gamma}(.)$ and $\gamma(.)$ are computed. The median, 10% and 90% quantiles of these errors are represented in Figure 2. From these plots, we observe that our estimator performs quite well with a deterioration in the estimation of $\gamma(.)$ if $|\rho|$ decreases.

6 Proofs

6.1 Preliminary results

The difference between the estimators $\widehat{q}_n(., x)$ and $\overline{\mathbb{F}}_n^{\leftarrow}(., x)$ of the conditional quantile is controlled by the following lemma:

Lemma 1. *For all $x \in E$,*

$$\sup_{\alpha \in [0,1)} \left| \widehat{q}_n(\alpha, x) - \overline{\mathbb{F}}_n^{\leftarrow}(\alpha, x) \right| \leq h_{2,n}.$$

Proof − The empirical estimator of the survival function can be rewritten as:

$$\overline{\mathbb{F}}_n(y, x) = \begin{cases} \beta_0 = 0 & \text{if } y \geq Y^*_{m_x, m_x}, \\ \beta_j & \text{if } y \in [Y^*_{m_x - j, m_x}, Y^*_{m_x - j + 1, m_x}), \\ \beta_{m_x} = 1 & \text{if } y < Y^*_{1, m_x}, \end{cases}$$

where β_j, $j = 0, \ldots, m_x$ are defined in (4). It is then easy to check that:

$$\overline{\mathbb{F}}_n^{\leftarrow}(\alpha, x) = \sum_{i=0}^{m_x - 1} Y^*_{m_x - i, m_x} \mathbb{1}_{\{\alpha \in [\beta_i, \beta_{i+1})\}}. \tag{10}$$

Let $j \in \{0, \ldots, m_x - 1\}$. Since $K(u) = 1$ for all $u \leq -1$ and $K(u) = 0$ for all $u \geq 1$, we have for all $y \geq Y^*_{m_x - j, m_x} + h_{2,n}$ that

$$K\left(\frac{y - Y^*_{i, m_x}}{h_{2,n}}\right) = 0 \text{ for all } i \in \{1, \ldots, m_x - j\}.$$

Furthermore, for all $y \leq Y^*_{m_x - j, m_x} - h_{2,n}$,

$$K\left(\frac{y - Y^*_{i, m_x}}{h_{2,n}}\right) = 1 \text{ for all } i \in \{m_x - j, \ldots, m_x\}.$$

Thus, for all $y \geq Y^*_{m_x-j,m_x} + h_{2,n}$, $\widehat{\overline{F}}_n(y,x) \leq \beta_j$ and for all $y \leq Y^*_{m_x-j,m_x} - h_{2,n}$, $\widehat{\overline{F}}_n(y,x) \geq \beta_{j+1}$. Hence, for all $\alpha \in [\beta_j, \beta_{j+1})$,

$$\widehat{\overline{F}}_n^{\leftarrow}(\alpha, x) \in (Y^*_{m_x-j,m_x} - h_{2,n}, Y^*_{m_x-j,m_x} + h_{2,n}). \tag{11}$$

Equations (10) and (11) conclude the proof. ■

The following lemma establish the rate of convergence of Riemann sums.

Lemma 2. *Let $\varphi(.)$ be an integrable function defined on $(0,1]$.*

i)

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \varphi\left(\frac{i}{N}\right) = \int_0^1 \varphi(x)dx.$$

ii) Furthermore, if $\varphi(.)$ is a decreasing function, we have for all $N \in \mathbb{N}^*$,

$$0 \leq \int_0^1 \varphi(x)dx - \frac{1}{N}\sum_{i=1}^{N} \varphi\left(\frac{i}{N}\right) \leq \int_0^{1/N} \varphi(x)dx - \frac{\varphi(1)}{N}.$$

Proof – The proof of part i) is straightforward. Concerning ii), remark that, since φ is decreasing,

$$\int_0^1 \varphi(x)dx = \sum_{i=1}^{N} \int_{(i-1)/N}^{i/N} \varphi(x)dx \geq \frac{1}{N}\sum_{i=1}^{N} \varphi\left(\frac{i}{N}\right). \tag{12}$$

Next,

$$\int_{1/N}^1 \varphi(x)dx = \sum_{i=1}^{N-1} \int_{i/N}^{(i+1)/N} \varphi(x)dx \leq \frac{1}{N}\sum_{i=1}^{N-1} \varphi\left(\frac{i}{N}\right). \tag{13}$$

Equations (12) and (13) conclude the proof. ■

Let us introduce the rescaled log-spacings:

$$Z_{i,m_x}(x) = i\left\{\log(Y^*_{m_x-i+1,m_x}) - \log(Y^*_{m_x-i,m_x})\right\},\ i = 1,\ldots,m_x-1.$$

The next result provides an expansion of the conditional tail index estimator $\widehat{\gamma}(\beta_{k_x}, x)$ where β_{k_x} is defined in (4).

Lemma 3. *Assume that conditions* **(A.1)**, **(A.2)** *hold. If $k_x \to \infty$, $m_x/k_x \to \infty$ and $k_x^2\omega_n(m_x^{-(1+\delta)}) \to 0$ for some $\delta > 0$ then,*

$$\widehat{\gamma}(\beta_{k_x}, x) = \sum_{i=1}^{k_x}\left(\frac{1}{i}\int_0^{\beta_i} \Psi(\alpha, \beta_{k_x}, x)d\alpha\right) Z_{i,m_x}(x) + O_{\mathbb{P}}\left(\frac{h_{2,n}}{q(k_x/m_x, x)}\right).$$

Proof of Lemma 3 – From Lemma 1,

$$\sup_{\alpha\in[0,\beta_{k_x})} \left|\frac{\widehat{q}_n(\alpha,x)}{\overline{\mathbb{F}}_n^{\leftarrow}(\alpha,x)} - 1\right| \leq \frac{h_{2,n}}{Y^*_{m_x-k_x,m_x}}, \tag{14}$$

since, for all $\alpha \in [0,\beta_{k_x})$, $\overline{\mathbb{F}}_n^{\leftarrow}(\alpha,x) \geq \overline{\mathbb{F}}_n^{\leftarrow}(\beta_{k_x},x) = Y^*_{m_x-k_x,m_x}$. From Gardes and Girard (2010, Lemma 1), there exists a sequence of event $\mathcal{A}_n$ with $\mathbb{P}(\mathcal{A}_n) \to 1$ as $n \to \infty$ such that for some $\delta > 0$,

$$\log Y^*_{m_x-k_x,m_x}|\mathcal{A}_n \stackrel{d}{=} \log q(U_{k_x+1,m_x}, x) + O_{\mathbb{P}}\left(\omega_n(m_x^{-(1+\delta)})\right)\Big|\mathcal{A}_n, \tag{15}$$

where U_i, $i = 1,\ldots,m_x$ are independent standard uniform variables, $U_{1,m_x} \leq \ldots \leq U_{m_x,m_x}$ are the associated order statistics. Hence, for all $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\log Y^*_{m_x-k_x,m_x} - \log q\left(\frac{k_x}{m_x}, x\right)\right| > \varepsilon\right) =: T_{1,n} + T_{2,n},$$

where

$$T_{1,n} := \mathbb{P}\left(\left|\log Y^*_{m_x-k_x,m_x} - \log q\left(\frac{k_x}{m_x},x\right)\right| > \varepsilon \middle| \mathcal{A}_n\right)\mathbb{P}(\mathcal{A}_n)$$

and

$$T_{2,n} := \mathbb{P}\left(\left|\log Y^*_{m_x-k_x,m_x} - \log q\left(\frac{k_x}{m_x},x\right)\right| > \varepsilon \cap \mathcal{A}_n^C\right).$$

Clearly from (15),

$$T_{1,n} \stackrel{d}{=} \mathbb{P}\left(\left|\log\frac{q(U_{k_x+1,m_x},x)}{q\left(k_x/m_x,x\right)} + O_{\mathbb{P}}\left(\omega_n(m_x^{-(1+\delta)})\right)\right| > \varepsilon \middle| \mathcal{A}_n\right)\mathbb{P}(\mathcal{A}_n).$$

Using the fact that for a positive sequence k_x such that $k_x \to \infty$ and $m_x/k_x \to \infty$, $U_{k_x+1,m_x} \stackrel{\mathbb{P}}{\sim} k_x/m_x$ and since $q(.,x)$ is a regularly varying function we have $T_{1,n} \to 0$ using the fact that $\omega_n(m_x^{-(1+\delta)}) \to 0$. Furthermore, $T_{2,n} \leq \mathbb{P}(\mathcal{A}_n^C) \to 0$ since $\mathbb{P}(\mathcal{A}_n) \to 1$. Hence, for all $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\log Y^*_{m_x-k_x,m_x} - \log q\left(\frac{k_x}{m_x},x\right)\right| > \varepsilon\right) \to 0,$$

and thus (14) implies that

$$\sup_{\alpha\in[0,\beta_{k_x})}\left|\frac{\widehat{q}_n(\alpha,x)}{\overline{\mathbb{F}}_n^{\leftarrow}(\alpha,x)} - 1\right| = O_{\mathbb{P}}\left(\frac{h_{2,n}}{q(k_x/m_x,x)}\right). \tag{16}$$

Now remark that

$$\widehat{\gamma}(\beta_{k_x},x) = \int_0^{\beta_{k_x}} \Psi(\alpha,\beta_{k_x},x)\log\overline{\mathbb{F}}_n^{\leftarrow}(\alpha,x)d\alpha + \int_0^{\beta_{k_x}} \Psi(\alpha,\beta_{k_x},x)\log\frac{\widehat{q}_n(\alpha,x)}{\overline{\mathbb{F}}_n^{\leftarrow}(\alpha,x)}d\alpha.$$

From (16) and since $\Psi(.,u,x) \in L_1(0,1)$,

$$\left|\int_0^{\beta_{k_x}} \Psi(\alpha,\beta_{k_x},x)\log\frac{\widehat{q}_n(\alpha,x)}{\overline{\mathbb{F}}_n^{\leftarrow}(\alpha,x)}d\alpha\right| = O_{\mathbb{P}}\left(\frac{h_{2,n}}{q(k_x/m_x,x)}\right). \tag{17}$$

Furthermore, from (10) and since $\int_0^{\beta_{k_x}} \Psi(\alpha,\beta_{k_x},x)d\alpha = 0$, we have

$$\begin{aligned}\int_0^{\beta_{k_x}} \Psi(\alpha,\beta_{k_x},x)\log\overline{\mathbb{F}}_n^{\leftarrow}(\alpha,x)d\alpha &= \sum_{i=1}^{k_x}\left(\int_{\beta_{i-1}}^{\beta_i}\Psi(\alpha,\beta_{k_x},x)d\alpha\right)\log Y^*_{m_x-i+1,m_x}\\ &= \sum_{i=1}^{k_x}\left(\frac{1}{i}\int_0^{\beta_i}\Psi(\alpha,\beta_{k_x},x)d\alpha\right)Z_{i,m_x}(x)\end{aligned}$$

which achieves the proof. ■

6.2 Proofs of the main results

Proof of Theorem 1 – From Lemma 1 in Gardes and Girard (2010), there exists a sequence of event $\mathcal{A}_n$ with $\mathbb{P}(\mathcal{A}_n) \to 1$ as $n \to \infty$ such that $\{Z_{i,m_x}(x),\ i = 1,\ldots,k_x|\mathcal{A}_n\}$ has the same distribution as

$$\left\{i\log\frac{q(U_{i,m_x},x)}{q(U_{i+1,m_x},x)} + iO_{\mathbb{P}}\left(\omega_n(m_x^{-(1+\delta)})\right),\ i = 1,\ldots,k_x \middle| \mathcal{A}_n\right\}.$$

Denoting by

$$\widetilde{\gamma}(\beta_{k_x},x) = \sum_{i=1}^{k_x}\left(\frac{1}{i}\int_0^{\beta_i}\Psi(\alpha,\beta_{k_x},x)d\alpha\right)Z_{i,m_x}(x),$$

we thus have that $\widetilde{\gamma}(\beta_{k_x},x)|\mathcal{A}_n$ has the same distribution as

$$\sum_{i=1}^{k_x}\left(\frac{1}{i}\int_0^{\beta_i}\Psi(\alpha,\beta_{k_x},x)d\alpha\right)i\log\frac{q(U_{i,m_x},x)}{q(U_{i+1,m_x},x)} + O_{\mathbb{P}}\left(k_x\omega_n(m_x^{-(1+\delta)})\right)(b_{1,n}(x)+1)\Bigg|\mathcal{A}_n.$$

Let us introduce the following notation $\Delta_x = \Delta\left(m_x/k_x, x\right)$. From Gardes and Girard (2010, Lemma 4), to show that $\sigma_{n,x}(\widetilde{\gamma}(\beta_{k_x}, x) - \gamma(x)(b_{1,n}(x)+1) - \Delta_x b_{2,n}(x))$ converges to a $\mathcal{N}(0, \gamma^2(x))$ random variable, it is sufficient to prove that $\sigma_{n,x}(T_{1,n} - \gamma(x)(b_{1,n}(x)+1) - \Delta_x b_{2,n}(x)) + \sigma_{n,x} T_{2,n}$ converges to a $\mathcal{N}(0, \gamma^2(x))$ random variable, where

$$T_{1,n} = \sum_{i=1}^{k_x} \left(\frac{1}{i} \int_0^{\beta_i} \Psi(\alpha, \beta_{k_x}, x) d\alpha \right) i \log \frac{q(U_{i,m_x}, x)}{q(U_{i+1,m_x}, x)},$$

and

$$T_{2,n} = O_{\mathbb{P}}\left(k_x \omega_n (m_x^{-(1+\delta)}) \right) (b_{1,n}(x) + 1).$$

From Beirlant *et al.* (2002, Theorem 2.1), $\{i(\log q(U_{i,m_x}, x) - \log q(U_{i+1,m_x}, x)),\ i = 1, \ldots, k_x\}$ has the same distribution as

$$\left\{ \left(\gamma(x) + \Delta_x \left(\frac{i}{k_x + 1} \right)^{-\rho(x)} \right) F_i \ + \nu_{i,n}(x) + o_{\mathbb{P}}(\Delta_x),\ i = 1, \ldots, k_x \right\}, \tag{18}$$

where $F_1, \ldots, F_{k_x}$ are independent standard exponential random variables and, uniformly in $i = 1, \ldots, k_x$,

$$\sum_{j=i}^{k_x} \frac{|\nu_{j,n}(x)|}{j} = o_{\mathbb{P}}\left(\Delta_x \max\left(\log \frac{k_x + 1}{i}, 1 \right) \right). \tag{19}$$

Thus, $\sigma_{n,x}(T_{1,n} - \gamma(x)(b_{1,n}(x) + 1) - \Delta_x b_{2,n}(x))$ has the same distribution as

$$\sigma_{n,x} \left(\gamma(x) T_{1,1,n} + \Delta_x T_{1,2,n} + T_{1,3,n} + (b_{1,n}(x) + 1) o_{\mathbb{P}}(\Delta_x) \right),$$

with

$$\begin{aligned}
T_{1,1,n} &= \sum_{i=1}^{k_x} \left(\frac{1}{i} \int_0^{\beta_i} \Psi(\alpha, \beta_{k_x}, x) d\alpha \right) (F_i - 1) \\
T_{1,2,n} &= \sum_{i=1}^{k_x} \left(\frac{1}{i} \int_0^{\beta_i} \Psi(\alpha, \beta_{k_x}, x) d\alpha \right) \left(\frac{i}{k_x + 1} \right)^{-\rho(x)} (F_i - 1) \\
T_{1,3,n} &= \sum_{i=1}^{k_x} \left(\frac{1}{i} \int_0^{\beta_i} \Psi(\alpha, \beta_{k_x}, x) d\alpha \right) \nu_{i,n}(x).
\end{aligned}$$

Let us first consider the term $T_{1,1,n}$. From Lyapunov's Theorem, if for some $\xi > 0$,

$$\left(\sum_{i=1}^{k_x} \left(\frac{1}{i} \int_0^{\beta_i} \Psi(\alpha, \beta_{k_x}, x) d\alpha \right)^2 \right)^{-(1+\xi/2)} \sum_{i=1}^{k_x} \left| \frac{1}{i} \int_0^{\beta_i} \Psi(\alpha, \beta_{k_x}, x) d\alpha \right|^{2+\xi} \to 0, \tag{20}$$

then,

$$\sigma_{n,x} T_{1,1,n} \xrightarrow{d} \mathcal{N}(0, 1). \tag{21}$$

From condition **(B.1)**,

$$\begin{aligned}
&\left(\sum_{i=1}^{k_x} \left(\frac{1}{i} \int_0^{\beta_i} \Psi(\alpha, \beta_{k_x}, x) d\alpha \right)^2 \right)^{-(1+\xi/2)} \sum_{i=1}^{k_x} \left| \frac{1}{i} \int_0^{\beta_i} \Psi(\alpha, \beta_{k_x}, x) d\alpha \right|^{2+\xi} \\
&\qquad \leq k_x^{-\xi/2} \left(\frac{1}{k_x} \sum_{i=1}^{k_x} f_1^2 \left(\frac{i}{k_x}, x \right) \right)^{-(1+\xi/2)} \left(\frac{1}{k_x} \sum_{i=1}^{k_x} f_2^{2+\xi} \left(\frac{i}{k_x}, x \right) \right).
\end{aligned}$$

Using Lemma 2 i) and condition **(B.1)**, it then easy to check that condition (20) is satisfied. Now, let us focus on the term $T_{1,2,n}$. We have $\mathbb{E}(T_{1,2,n}) = 0$ and

$$\mathrm{Var}(T_{1,2,n}) = \frac{1}{k_x^2} \sum_{i=1}^{k_x} \left(\frac{k_x}{i} \int_0^{\beta_i} \Psi(\alpha, \beta_{k_x}, x) d\alpha \right)^2 \left(\frac{i}{k_x + 1} \right)^{-2\rho(x)} = O(k_x^{-1}),$$

from Lemma 2 i) and condition **(B.1)**. Thus,

$$T_{1,2,n} = o_{\mathbb{P}}(1). \tag{22}$$

Furthermore,

$$\begin{aligned} T_{1,3,n} &= \sum_{i=1}^{k_x} \left(\frac{1}{i} \int_0^{\beta_i} \Psi(\alpha, \beta_{k_x}, x) d\alpha \right) \nu_{i,n}(x) \\ &= \sum_{i=1}^{k_x} \left(\sum_{j=1}^{i} \int_{\beta_{j-1}}^{\beta_j} \Psi(\alpha, \beta_{k_x}, x) d\alpha \right) \frac{\nu_{i,n}(x)}{i} \\ &= \sum_{j=1}^{k_x} \int_{\beta_{j-1}}^{\beta_j} \Psi(\alpha, \beta_{k_x}, x) d\alpha \sum_{i=j}^{k_x} \frac{\nu_{i,n}(x)}{i}. \end{aligned}$$

Thus, from **(B.2)**, (19) and using Lemma 2 i)

$$\begin{aligned} |T_{1,3,n}| &= \Delta_x o_{\mathbb{P}} \left(\sum_{j=1}^{k_x} \left| \int_{\beta_{j-1}}^{\beta_j} \Psi(\alpha, \beta_{k_x}, x) d\alpha \right| \max \left(\log \frac{k_x+1}{j}, 1 \right) \right) \\ &= \Delta_x o_{\mathbb{P}} \left(\frac{1}{k_x} \sum_{j=1}^{k_x} f_3 \left(\frac{j}{k_x}, x \right) \max \left(\log \frac{k_x+1}{j}, 1 \right) \right) \\ &= o_{\mathbb{P}}(\Delta_x). \end{aligned} \tag{23}$$

Finally, using **(B.1)** and Lemma 2 i), $b_{1,n}(x) + 1 = O(1)$ and thus

$$\sigma_{n,x} T_{2,n} = O_{\mathbb{P}} \left(k_x^{3/2} \omega_n (m_x^{-(1+\delta)} \right) = o_{\mathbb{P}}(1). \tag{24}$$

As a first conclusion, (21) to (24) entail that $\sigma_{n,x}(\widetilde{\gamma}(\beta_{k_x}, x) - \gamma(x)(b_{1,n}(x) + 1) - \Delta_x b_{2,n}(x))$ converges to a $\mathcal{N}(0, \gamma^2(x))$. The conclusion of the proof is then straighforward using Lemma 3 and the fact that under **(B.1)** and Lemma 2 i),

$$\sigma_{n,x} O_{\mathbb{P}} \left(\frac{h_{2,n}}{q(k_x/m_x, x)} \right) = O_{\mathbb{P}} \left(\frac{k_x^{1/2} h_{2,n}}{q(k_x/m_x, x)} \right) = o_{\mathbb{P}}(1). \tag{25}$$

■

Proof of Corollary 2 – It is quite easy to check that the assumptions of Corollary 2 imply the ones of Theorem 1. Thus, it is sufficient to prove the three following statements:

$$\sigma_{n,x} \sim k_x^{1/2} \left(\int_0^1 \Phi^2(\alpha, x) d\alpha \right)^{-1/2} \tag{26}$$

$$\sigma_{n,x} b_{1,n}(x) \to 0 \tag{27}$$

$$b_{2,n}(x) \to \int_0^1 \Phi(\alpha, x) \alpha^{-\rho(x)} d\alpha. \tag{28}$$

Equivalence (26) is a direct consequence of Lemma 2 i) and condition **(C.1)**. Using Lemma 2 ii) we have that $b_{1,n}(x) = o(k_x^{-1/2})$ since

$$\int_0^1 \Phi(\alpha, x) d\alpha = 1.$$

Hence, from (26), $\sigma_{n,x} b_{1,n}(x) = o(1)$. Finally, (28) is straightforward using Lemma 2 i). This achieves the proof of Corollary 2. ■

Proof of Corollary 3 – It is sufficient to prove that the function $\Psi_\theta(.,u,x)$ satisfies conditions **(C.1)** and **(C.2)**. Clearly, **(C.1)** is satisfied with

$$\Phi_\theta(\alpha,x)=\begin{cases}\dfrac{\theta+1}{\theta}(1-\alpha^\theta) & \text{if } \theta>0\\ -\log(\alpha) & \text{if } \theta=0\\ 1 & \text{if } \theta=\infty,\end{cases}$$

for $\alpha<1$. Note that $\Phi_0(.,.)$ and $\Phi_\infty(.,.)$ correspond to the limiting cases of $\Phi_\theta(.,.)$. Concerning the condition **(C.2)**, the mean value theorem leads to :

$$\begin{aligned}\left|i\Phi_\theta\left(\frac{i}{k_x},x\right)-(i-1)\Phi_\theta\left(\frac{i-1}{k_x},x\right)\right| &\leq \sup_{\alpha\in I_i}\left|\frac{\theta+1}{\theta}(1-\alpha^\theta)-\alpha^\theta(1+\theta)\right|\\ &\leq \sup_{\alpha\in I_i}\frac{\theta+1}{\theta}(1+\alpha^\theta(\theta-1)),\end{aligned}$$

where $I_i=[(i-1)/k_x,i/k_x]$. Clearly, if $\theta\geq 1$, for all $i=1,\ldots,k_x$,

$$\left|i\Phi_\theta\left(\frac{i}{k_x},x\right)-(i-1)\Phi_\theta\left(\frac{i-1}{k_x},x\right)\right|\leq\frac{\theta+1}{\theta}\left(1+\left(\frac{i}{k_x}\right)^\theta(\theta-1)\right). \tag{29}$$

If $\theta<1$, then for $i=1$,

$$\left|i\Phi_\theta\left(\frac{i}{k_x},x\right)-(i-1)\Phi_\theta\left(\frac{i-1}{k_x},x\right)\right|=\Phi_\theta\left(\frac{1}{k_x},x\right), \tag{30}$$

and for $i=2,\ldots,k_x$,

$$\begin{aligned}\left|i\Phi_\theta\left(\frac{i}{k_x},x\right)-(i-1)\Phi_\theta\left(\frac{i-1}{k_x},x\right)\right| &\leq \frac{\theta+1}{\theta}\left(1+\left(\frac{i-1}{k_x}\right)^\theta(\theta-1)\right)\\ &\leq \frac{\theta+1}{\theta}\left(1+\left(\frac{i}{2k_x}\right)^\theta(\theta-1)\right).\end{aligned} \tag{31}$$

Collecting (29)-(31), leads to

$$\left|i\Phi_\theta\left(\frac{i}{k_x},x\right)-(i-1)\Phi_\theta\left(\frac{i-1}{k_x},x\right)\right|\leq f_4\left(\frac{i}{k_x},x\right),$$

where

$$f_4(\alpha,x)=\begin{cases}(\theta+1)(1+(\theta-1)\alpha^\theta)/\theta & \text{if } \theta\geq 1,\\ (\theta+1)(1+(\theta-1)(\alpha/2)^\theta)/\theta & \text{if } \theta<1.\end{cases}$$

Since $f_4(.,x)$ satisfies the condition in **(C.2)**, the proof is complete. ■

References

Beirlant, J., Dierckx, G., Goegebeur, Y., Matthys, G. 1999. Tail index estimation and an exponential regression model, Extremes, 2, 177–200.

Beirlant, J., Dierckx, G., Guillou, A., Stărică, C. 2002. On exponential representations of log-spacings of extreme order statistics, Extremes, 5, 157–180.

Beirlant, J., Goegebeur, Y. 2003. Regression with response distributions of Pareto-type, Computational Statistics and Data Analysis, 42, 595–619.

Caeiro, F., Gomes, M.I. 2002. A class of asymptotically unbiased semi-parametric estimators of the tail index, Test, 11, 345–364.

Chavez-Demoulin, V., Davison, A.C. 2005. Generalized additive modelling of sample extremes, Journal of the Royal Statistical Society, series C, 54, 207–222.

Daouia, A., Gardes, L., Girard, S., Lekina, A. 2010. Kernel estimators of extreme level curves, Test, to appear.
Davison, A.C., Ramesh, N.I. 2000. Local likelihood smoothing of sample extremes, Journal of the Royal Statistical Society, series B, 62, 191–208.
Davison, A.C., Smith, R.L. 1990. Models for exceedances over high thresholds, Journal of the Royal Statistical Society, series B, 52, 393–442.
Dekkers, A., Einmahl, J., de Haan, L. 1989. A moment estimation for the index of an extreme-value distribution, Annals of Statistics, 17, 1833–1855.
Diebolt, J., Gardes, L., Girard, S., Guillou, A. 2008. Bias-reduced estimators of the Weibull tail-coefficient, Test, 17, 311–331.
Ferraty, F., Vieu, P. 2006. Nonparametric Functional Data Analysis: Theory and Practice, Springer Series in Statistics, Springer.
Feuerverger, A., Hall, P. 1999. Estimating a tail exponent by modelling departure from a Pareto distribution, Annals of Statistics, 27, 760–781.
Gardes, L., Girard, S. 2008. A moving window approach for nonparametric estimation of the conditional tail index, Journal of Multivariate Analysis, 99, 2368–2388.
Gardes, L., Girard, S. 2010. Conditional extremes from heavy-tailed distributions: an application to the estimation of extreme rainfall return levels, Extremes, 13, 177–204.
Hall, P., Tajvidi, N. 2000. Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data, Statistical Science, 15, 153–167.
Hill, B.M. 1975. A simple general approach to inference about the tail of a distribution, Annals of Statistics, 3, 1163–1174.
Kratz, M., Resnick, S. 1996. The QQ-estimator and heavy tails, Stochastic Models, 12, 699–724.
Schultze, J., Steinebach, J. 1996. On least squares estimates of an exponential tail coefficient, Statistics and Decisions, 14, 353–372.
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.

Figure 1: (a) Theoretical (—) and estimated $(\cdots)$ AMSE for one realization of the Burr-type distribution (8) with $\rho = -1$ and $\gamma(0.4)$ defined in (9). The vertical lines correspond to the optimal values $k_{0.4}^{opt}$ and $\widehat{k}_{0.4}$; (b) Boxplot of $\widehat{k}_{0.4}$ for the 100 replications. The horizontal line is the optimal value $k_{0.4}^{opt}$.

Figure 2: Estimation of $\gamma(.)$: Median (full line), 10% (dashed line) and 90% quantiles (dotted line) of the L^2-errors and true function (—) based on 100 replications from the Burr-type distribution (8) with $\gamma(.)$ defined in (9) and (a) $\rho = -1$ or (b) $\rho = -0.8$.