Electronic Journal of Statistics
Vol. 7 (2013) 2298–2343
ISSN: 1935-7524
DOI: 10.1214/13-EJS846

A moment estimator for the conditional extreme-value index

Gilles Stupfler

Aix Marseille Université, CERGAM, EA 4225, 15-19 allée Claude Forbin, 13628 Aix-en-Provence Cedex 1, France
e-mail: gilles.stupfler@univ-amu.fr

Abstract: In extreme value theory, the so-called extreme-value index is a parameter that controls the behavior of a distribution function in its right tail. Knowing this parameter is thus essential to solve many problems related to extreme events. In this paper, the estimation of the extreme-value index is considered in the presence of a random covariate, whether the conditional distribution of the variable of interest belongs to the Fréchet, Weibull or Gumbel max-domain of attraction. The pointwise weak consistency and asymptotic normality of the proposed estimator are established. We examine the finite sample performance of our estimator in a simulation study and we illustrate its behavior on a real set of fire insurance data.

MSC 2010 subject classifications: 62G05, 62G20, 62G30, 62G32.
Keywords and phrases: Extreme-value index, moment estimator, random covariate, consistency, asymptotic normality.

Received July 2013.

Contents

1. Introduction

The problem of studying extreme events arises in many fields of statistical applications. In hydrology, one could for instance be interested in forecasting the maximum level reached by the seawater along a coast over a given period, or studying extreme rainfall at a given location; in actuarial science, it is of primary interest for a company to estimate the probability that a claim which represents a threat to its solvency is filed. The pioneering result in extreme value theory, known as the Fisher-Tippett-Gnedenko theorem (see Fisher and Tippett [13] and Gnedenko [19]) states that if (Y_n) is an independent sequence of random copies of a random variable Y such that there exist normalizing non-random sequences of real numbers (a_n) and (b_n), with $a_n > 0$ and such that the sequence

$$\frac{1}{a_n}\left(\max_{1\leq i\leq n} Y_i - b_n\right)$$

converges in distribution to some nondegenerate limit, then the cumulative distribution function (cdf) of this limit can necessarily be written $y \mapsto G_\gamma(ay+b)$, with $a > 0$ and b, $\gamma \in \mathbb{R}$ where

$$G_\gamma(y) = \begin{cases} \exp\left(-(1+\gamma y)^{-1/\gamma}\right) & \text{if } \gamma \neq 0 \text{ and } 1+\gamma y > 0, \\ \exp\left(-\exp(-y)\right) & \text{if } \gamma = 0. \end{cases}$$

If the aforementioned convergence holds, we shall say that Y (or equivalently, its cdf F) belongs to the max-domain of attraction (MDA) of G_γ, with γ being the extreme-value index of Y, and we write $F \in \mathcal{D}(G_\gamma)$. Clearly, γ drives the behavior of F in its right tail:

- if $\gamma > 0$, namely Y belongs to the Fréchet MDA, then $1 - G_\gamma$ is heavy-tailed, *i.e.* it has a polynomial decay;
- if $\gamma < 0$, namely Y belongs to the Weibull MDA, then $1 - G_\gamma$ is short-tailed, *i.e.* it has a support bounded to the right;
- if $\gamma = 0$, namely Y belongs to the Gumbel MDA, then $1 - G_\gamma$ has an exponential decay.

The knowledge of γ is therefore necessary to tackle a number of problems in extreme value analysis, such as the estimation of extreme quantiles of Y, which made its estimation a central topic in the literature. Recent monographs on extreme value theory and especially univariate extreme-value index estimation include Beirlant *et al.* [3] and de Haan and Ferreira [21].

In practical applications, it is often the case that the variable of interest Y can be linked to a covariate X. In this situation, the extreme-value index of the conditional distribution of Y given $X = x$ may depend on x; the problem is then to estimate the conditional extreme-value index $x \mapsto \gamma(x)$. Motivating examples in the literature include the description of the right tail of the distribution of claim sizes in insurance or reinsurance (see [3]), the estimation of the maximal production level as a function of the quantity of labor (see Daouia *et al.* [6]), studying extreme temperatures as a function of various topological parameters (see Ferrez *et al.* [12]), the estimation of some quantitative physical characteristics of Martian soil (see Gardes *et al.* [17]), or analyzing extreme earthquakes as a function of the location (see Pisarenko and Sornette [26]).

In most recent works, this problem has been addressed in the "fixed design" case, namely when the covariates are nonrandom. For instance, Smith [27] and Davison and Smith [10] considered a regression model while Hall and Tajvidi [22] used a semiparametric approach in this context; a nonexhaustive list of fully nonparametric methods include Davison and Ramesh [9] for a local polynomial estimator, Chavez-Demoulin and Davison [5] for a method using splines, Gardes and Girard [14] for a moving window approach and Gardes and Girard [15] who used a nearest neighbor approach.

By contrast, the case when the covariate is random, which is very interesting as far as practical applications are concerned, has only been tackled in even newer works. In the actuarial science setting, one could for instance think of a situation in which an insurance firm covers damage done to policyholders by natural disasters: a typical covariate in this case is the location where a natural disaster happens. Another situation, which will be examined in this paper, is the case of an insurance firm covering damage done by fire accidents: a possible covariate of the claim size is the total sum insured by the firm. We refer to Wang and Tsai [28] for a maximum likelihood approach, Daouia *et al.* [7] who used a fixed number of nonparametric conditional quantile estimators to estimate the conditional extreme-value index, Gardes and Girard [16] who generalized the method of [7] to the case when the covariate space is infinite-dimensional, Goegebeur *et al.* [20] who studied a nonparametric regression estimator and Gardes and Stupfler [18] who introduced a smoothed local Hill estimator. Besides, the method of [7] was recently generalized in Daouia *et al.* [8] to a regression context with a response distribution belonging to the general max-domain of attraction: the latter study is the only one in this list which is not restricted to the case of the Fréchet MDA.

The aim of this paper is to introduce a moment estimator of the conditional extreme-value index, working in the three domains of attraction. In Section 2, we define our estimator of the conditional extreme-value index. The pointwise weak consistency and asymptotic normality of the estimator are stated in Section 3. The finite sample performance of the estimator is studied in Section 4. In Section 5, we illustrate the behavior of the proposed estimator on a real set of fire insurance data. Proofs of the main results are given in Section 6 and those of the auxiliary results are postponed to Section 7.

2. Estimation of the conditional extreme-value index

Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be n independent copies of a random pair (X, Y) taking its values in $E \times (0, \infty)$ where E is a metric space endowed with a metric d. For all $x \in E$, we assume that the conditional survival function (csf) $\overline{F}(\cdot|x) = 1 - F(\cdot|x)$ of Y given $X = x$ belongs to $\mathcal{D}(G_{\gamma(x)})$. Specifically, we shall work in the following setting:

(M_1) Y is a positive random variable and for every $x \in E$, there exist a real number $\gamma(x)$ and a positive function $a(\cdot|x)$ such that the left-continuous inverse $U(\cdot|x)$ of $1/\overline{F}(\cdot|x)$, defined by $U(z|x) = \inf\{y \in \mathbb{R} \,|\, 1/\overline{F}(y|x) \geq z\}$ for every

$z \geq 1$, satisfies

$$\forall z > 0,\ \lim_{t\to\infty} \frac{U(tz|x) - U(t|x)}{a(t|x)} = \begin{cases} \dfrac{z^{\gamma(x)} - 1}{\gamma(x)} & \text{if } \gamma(x) \neq 0 \\ \log z & \text{if } \gamma(x) = 0. \end{cases}$$

Model (M_1) is the conditional analogue of the classical extreme-value framework, see for instance [21], p. 19. In this model, for every $x \in E$, the function $U(\cdot|x)$ has a positive limit $U(\infty|x)$ at infinity; the function $U(\infty|\cdot)$ is called the conditional right endpoint of Y.

We now introduce our estimator, which is an adaptation of the moment estimator of Dekkers *et al.* [11]. To this end, we let, for an arbitrary $x \in E$ and $h = h(n) \to 0$ as $n \to \infty$, $N_n(x, h)$ be the total number of observations in the closed ball $B(x, h)$ having center x and radius h:

$$N_n(x, h) = \sum_{i=1}^{n} \mathbb{1}_{\{X_i \in B(x,h)\}} \quad \text{with} \quad B(x, h) = \{x' \in E \,|\, d(x, x') \leq h\},$$

where $\mathbb{1}_{\{\cdot\}}$ is the indicator function. The purpose of the bandwidth sequence $h(n)$ is to select those covariates which are close enough to x. Given $N_n(x, h) = p \geq 1$, we let, for $i = 1, \ldots, p$, $Z_i = Z_i(x, h)$ be the response variables whose associated covariates $W_i = W_i(x, h)$ belong to the ball $B(x, h)$. Let further $Z_{1,p} \leq \cdots \leq Z_{p,p}$ be the related order statistics (this way of denoting order statistics shall be used throughout the paper) and set for $j = 1, 2$

$$M_n^{(j)}(x, k_x, h) = \frac{1}{k_x} \sum_{i=1}^{k_x} [\log(Z_{p-i+1,p}) - \log(Z_{p-k_x,p})]^j$$

if $k_x \in \{1, \ldots, p-1\}$ and 0 otherwise. Given $N_n(x, h) = p$, the random variable $M_n^{(j)}(x, k_x, h)$ is then computed by using only the response variables whose values are greater than the random threshold $Z_{p-k_x,p}$ and whose associated covariates belong to a (small) neighborhood of x. For $j = 1$, this statistic is an analogue of Hill's estimator (see Hill [24]) in the presence of a random covariate; see also [15] for a nearest neighbor analogue of this quantity in the fixed design case. Our estimator, in the spirit of [11], is then

$$\widehat{\gamma}_n(x, k_x, h) = \widehat{\gamma}_{n,+}(x, k_x, h) + \widehat{\gamma}_{n,-}(x, k_x, h)$$

where

$$\begin{aligned} \widehat{\gamma}_{n,+}(x, k_x, h) &= M_n^{(1)}(x, k_x, h) \\ \text{and}\quad \widehat{\gamma}_{n,-}(x, k_x, h) &= 1 - \frac{1}{2}\left(1 - \frac{\left[M_n^{(1)}(x, k_x, h)\right]^2}{M_n^{(2)}(x, k_x, h)}\right)^{-1} \end{aligned}$$

if $[M_n^{(1)}(x, k_x, h)]^2 \neq M_n^{(2)}(x, k_x, h)$, with $\widehat{\gamma}_{n,-}(x, k_x, h) = 0$ otherwise.

The assumption that Y is a positive random variable makes the quantities $M_n^{(j)}(x, k_x, h)$ well-defined for every k_x. This simplifies somewhat a couple of technical results (see for instance Lemma 3). We point out that since we shall only compute our estimator using upper order statistics of the Z_i, this hypothesis may be replaced by the assumption $U(\infty|x) > 0$ for every $x \in E$, at the price of extra regularity conditions on the joint cumulative distribution function F of the pair (X, Y).

3. Main results

3.1. Weak consistency

We first wish to state the pointwise weak consistency of our estimator. To this end we let, for $x \in E$, $n_x = n_x(n, h) = n\mathbb{P}(X \in B(x, h))$ be the average total number of points in the ball $B(x, h)$ and we assume that $n_x(n, h) > 0$ for every n. Let $k_x = k_x(n)$ be a sequence of positive integers; furthermore, let $F_h(\cdot|x)$ be the conditional cdf of Y given $X \in B(x, h)$:

$$F_h(y|x) = \mathbb{P}(Y \leq y \,|\, X \in B(x, h))$$

and $U_h(\cdot|x)$ be the left-continuous inverse of $1/\overline{F}_h(\cdot|x)$. For u, $v \in (1, \infty)$ such that $u < v$, we introduce the quantity

$$\omega(u, v, x, h) = \sup_{z \in [u,v]} \left| \log \frac{U_h(z|x)}{U(z|x)} \right|.$$

Recall the notation $a \wedge b = \min(a, b)$ and $a \vee b = \max(a, b)$ for a, $b \in \mathbb{R}$. Our consistency result is then:

Theorem 1. *Assume that (M_1) holds. Pick $x \in E$. We assume that $n_x \to \infty$, $k_x \to \infty$, $k_x/n_x \to 0$ and for some $\delta > 0$*

$$\frac{U(n_x/k_x|x)}{a(n_x/k_x|x)} \omega\left(\frac{n_x}{(1+\delta)k_x}, n_x^{1+\delta}, x, h\right) \to 0 \quad \textit{as} \quad n \to \infty. \tag{1}$$

Then, setting $\gamma_+(x) = 0 \vee \gamma(x)$ and $\gamma_-(x) = 0 \wedge \gamma(x)$, it holds that

$$\widehat{\gamma}_{n,+}(x, k_x, h) \xrightarrow{\mathbb{P}} \gamma_+(x) \quad \textit{and} \quad \widehat{\gamma}_{n,-}(x, k_x, h) \xrightarrow{\mathbb{P}} \gamma_-(x) \quad \textit{as} \quad n \to \infty$$

and therefore $\widehat{\gamma}_n(x, k_x, h) \xrightarrow{\mathbb{P}} \gamma(x)$ as $n \to \infty$.

Theorem 1 is the conditional analogue of the consistency result proven in [11]; see also [21], Theorem 3.5.2. As far as the hypotheses of Theorem 1 are concerned, note that conditions $n_x \to \infty$, $k_x \to \infty$ and $k_x/n_x \to 0$ are standard hypotheses for the estimation of the conditional extreme-value index: they are the exact analogues of the conditions $n \to \infty$, $k = k(n) \to \infty$ and $k/n \to 0$ needed to ensure the convergence of Hill's estimator. Moreover, condition $n_x \to \infty$ is necessary to make sure that there are sufficiently many observations close to x, which is a standard assumption in the random covariate case.

Condition (1) is somewhat harder to grasp. To analyze this hypothesis further, we introduce the following conditions:

($\mathcal{E}$) E is a linear space.

(A_1) For every $x \in E$, $U(\cdot|x)$ is a continuous increasing function on $(1, \infty)$ and for every $y \in \mathbb{R}$, the function $\overline{F}(y|\cdot)$ is continuous on E.

We may now state the following result, which relates the behavior of the function $\log U_h(z|\cdot)$ around x to that of $\log U(z|\cdot)$:

Proposition 1. *Assume that* ($\mathcal{E}$) *and* (A_1) *hold. Assume further that* $x \in E$ *is such that*

$$\forall x' \in B(x, h),\ \forall r > 0,\ \mathbb{P}(X \in B(x', r)) > 0.$$

Then for every $x \in E$ *and for every* $z > 1$*, it holds that*

$$\left|\log \frac{U_h(z|x)}{U(z|x)}\right| \leq \sup_{x' \in B(x,h)} \left|\log \frac{U(z|x')}{U(z|x)}\right|.$$

Note that if $E = \mathbb{R}^d$ with X having a probability density function f on this space and with x being such that $f(x) > 0$ and f is continuous at x, the condition

$$\forall x' \in B(x, h),\ \forall r > 0,\ \mathbb{P}(X \in B(x', r)) > 0$$

appearing in Proposition 1 is satisfied when n is large enough.

With this result at hand, we define for u, $v \in (1, \infty)$ such that $u < v$:

$$\Omega\left(u, v, x, h\right) = \sup_{z \in [u,v]} \sup_{x' \in B(x,h)} \left|\log \frac{U(z|x')}{U(z|x)}\right|.$$

Proposition 1 entails that

$$\omega\left(\frac{n_x}{(1+\delta)k_x}, n_x^{1+\delta}, x, h\right) \leq \Omega\left(\frac{n_x}{(1+\delta)k_x}, n_x^{1+\delta}, x, h\right).$$

Consequently, if conditions ($\mathcal{E}$) and (A_1) are satisfied, a sufficient condition for (1) to hold is

$$\frac{U(n_x/k_x|x)}{a(n_x/k_x|x)}\Omega\left(\frac{n_x}{(1+\delta)k_x}, n_x^{1+\delta}, x, h\right) \to 0 \quad \text{as} \quad n \to \infty \tag{2}$$

which is a hypothesis on the uniform oscillation of $\log U$ in its second variable. To understand more about condition (2), we introduce an additional regularity assumption:

(A_2) The function γ is a continuous function on E.

If we omit the case $\gamma(x) = 0$ of the Gumbel MDA, then under (A_2), condition (2) can be made more explicit:

- If $\gamma(x) > 0$, namely $\overline{F}(\cdot|x)$ belongs to the Fréchet MDA, then Lemma 1.2.9 in [21] entails that $a(\cdot|x)/U(\cdot|x)$ converges to $\gamma(x)$ at infinity. Condition (2) then becomes

$$\Omega\left(\frac{n_x}{(1+\delta)k_x}, n_x^{1+\delta}, x, h\right) \to 0 \quad \text{as} \quad n \to \infty. \tag{3}$$

Since γ is continuous, one has $\gamma(x') > 0$ for x' close enough to x. Corollary 1.2.10 in [21] then yields for n large enough and every $x' \in B(x,h)$

$$\forall z \geq 1,\ U(z|x') = z^{\gamma(x')} L(z|x')$$

where for every $x' \in B(x,h)$, $L(\cdot|x')$ is a slowly varying function at infinity. Letting

$$\forall z \geq 1,\ L(z|x') = c(z|x') \exp\left(\int_1^z \frac{\Delta(v|x')}{v} dv\right) \tag{4}$$

be Karamata's representation of $L(\cdot|x')$ (see Theorem 1.3.1 in Bingham *et al.* [4]), where $c(\cdot|x')$ is a positive Borel measurable function converging to a positive constant at infinity and $\Delta(\cdot|x')$ is a Borel measurable function converging to 0 at infinity, condition (3) is thus a consequence of the convergences

$$\begin{aligned} \log n_x \sup_{x' \in B(x,h)} |\gamma(x') - \gamma(x)| &\to 0, \\ \sup_{z \in K_{x,\delta}} \sup_{x' \in B(x,h)} |\log c(z|x') - \log c(z|x)| &\to 0, \\ \text{and} \quad \log n_x \sup_{z \in K_{x,\delta}} \sup_{x' \in B(x,h)} |\Delta(z|x') - \Delta(z|x)| &\to 0 \end{aligned}$$

as $n \to \infty$, where $K_{x,\delta} = [n_x/[(1+\delta)k_x, n_x^{1+\delta}]$. Besides, if γ, $\log c$ and Δ satisfy some sort of Hölder condition, for instance

$$\sup_{x' \in B(x,h)} |\gamma(x') - \gamma(x)| = \mathrm{O}\left(h^\alpha\right), \tag{5}$$

$$\sup_{z \in K_{x,\delta}} \sup_{x' \in B(x,h)} |\log c(z|x') - \log c(z|x)| = \mathrm{O}\left(h^\alpha\right) \tag{6}$$

$$\text{and} \quad \sup_{z \in K_{x,\delta}} \sup_{x' \in B(x,h)} |\Delta(z|x') - \Delta(z|x)| = \mathrm{O}\left(h^\alpha\right) \tag{7}$$

for some $\alpha \in (0,1]$ as $n \to \infty$, then condition (3) becomes $h^\alpha \log n_x \to 0$ as $n \to \infty$. The regularity conditions above are fairly standard when estimating the conditional extreme-value index in the Fréchet MDA, see for instance [7].
- If $\gamma(x) < 0$, namely $\overline{F}(\cdot|x)$ belongs to the Weibull MDA, then according to Lemma 1.2.9 in [21], one has

$$\frac{U(\infty|x) - U(z|x)}{a(z|x)} \to -\frac{1}{\gamma(x)} \quad \text{as} \quad z \to \infty.$$

Furthermore, since one has $\gamma(x') < 0$ for x' close enough to x, Corollary 1.2.10 in [21] yields for n large enough and every $x' \in B(x,h)$ that

$$\forall z \geq 1,\ U(\infty|x') - U(z|x') = z^{\gamma(x')} L(z|x')$$

where for every $x' \in B(x,h)$, $L(\cdot|x')$ is a slowly varying function at infinity. Especially

$$\frac{U(z|x)}{a(z|x)} = -\frac{U(\infty|x)}{\gamma(x)} \frac{z^{-\gamma(x)}}{L(z|x)}(1 + \mathrm{o}(1)) \ \text{ as } \ z \to \infty.$$

Consequently, in this framework, condition (2) becomes

$$\frac{(n_x/k_x)^{-\gamma(x)}}{L(n_x/k_x|x)} \Omega\left(\frac{n_x}{(1+\delta)k_x}, n_x^{1+\delta}, x, h\right) \to 0 \ \text{ as } \ n \to \infty. \tag{8}$$

Write then for an arbitrary $z > 1$ and for $x' \in B(x,h)$

$$\begin{aligned} & \log U(z|x') - \log U(z|x) \\ = \ & \log \frac{U(\infty|x')}{U(\infty|x)} + \log\left(\frac{1 - [U(\infty|x')]^{-1} z^{\gamma(x')} L(z|x')}{1 - [U(\infty|x)]^{-1} z^{\gamma(x)} L(z|x)}\right). \end{aligned} \tag{9}$$

The first term on the right-hand side in (9) is readily controlled if the conditional right endpoint $x \mapsto U(\infty|x)$ is a positive Hölder continuous function on E:

$$\begin{aligned} \sup_{x' \in B(x,h)} \left|\frac{U(\infty|x')}{U(\infty|x)} - 1\right| &= \mathrm{O}\left(\sup_{x' \in B(x,h)} |U(\infty|x') - U(\infty|x)|\right) \\ &= \mathrm{O}(h^\beta) \end{aligned} \tag{10}$$

say, with $\beta \in (0,1]$. The second one can be bounded from above as follows: since $n_x/k_x \to \infty$ and $z^{\gamma(x)} L(z|x) \to 0$ as $z \to \infty$ (see Proposition 1.5.1 in [4]), we can write for n large enough

$$\begin{aligned} & \left|\frac{1 - [U(\infty|x')]^{-1} z^{\gamma(x')} L(z|x')}{1 - [U(\infty|x)]^{-1} z^{\gamma(x)} L(z|x)} - 1\right| \\ \leq \ & 2\left|[U(\infty|x)]^{-1} z^{\gamma(x)} L(z|x) - [U(\infty|x')]^{-1} z^{\gamma(x')} L(z|x')\right| \\ \leq \ & 2\frac{z^{\gamma(x)} L(z|x)}{U(\infty|x)} \left|1 - \left[\frac{U(\infty|x')}{U(\infty|x)}\right]^{-1} \frac{z^{\gamma(x')} L(z|x')}{z^{\gamma(x)} L(z|x)}\right| \end{aligned} \tag{11}$$

for every $z \geq n_x/[(1+\delta)k_x]$. Note now that for every $z \in K_{x,\delta}$ we have

$$\begin{aligned} \left|\log\left(\frac{z^{\gamma(x')} L(z|x')}{z^{\gamma(x)} L(z|x)}\right)\right| &\leq (1+\delta) \log n_x \sup_{x' \in B(x,h)} |\gamma(x') - \gamma(x)| \\ &\quad + \sup_{z \in K_{x,\delta}} \sup_{x' \in B(x,h)} \left|\log \frac{L(z|x')}{L(z|x)}\right|. \end{aligned}$$

Using Karamata's representation of $L(\cdot|x')$ (see (4)) and assuming that for some $\alpha \in (0,1]$

$$\sup_{x'\in B(x,h)} |\gamma(x')-\gamma(x)| = \mathrm{O}(h^\alpha), \tag{12}$$

$$\sup_{z\in K_{x,\delta}} \sup_{x'\in B(x,h)} |\log c(z|x') - \log c(z|x)| = \mathrm{O}(h^\alpha), \tag{13}$$

$$\text{and} \quad \sup_{z\in K_{x,\delta}} \sup_{x'\in B(x,h)} |\Delta(z|x') - \Delta(z|x)| = \mathrm{O}(h^\alpha) \tag{14}$$

as $n\to\infty$, then using the inequality

$$\forall t \in \mathbb{R},\ |e^t - 1| \le |t| e^{|t|} \tag{15}$$

it is readily seen that if $h^\alpha \log n_x \to 0$ as $n\to\infty$, we have

$$\sup_{z\in K_{x,\delta}} \sup_{x'\in B(x,h)} \left| \frac{z^{\gamma(x')}L(z|x')}{z^{\gamma(x)}L(z|x)} - 1 \right| = \mathrm{O}(h^\alpha \log n_x). \tag{16}$$

Note that conditions (12), (13) and (14) are exactly (5), (6) and (7). Equations (10), (11), (16) and inequality (15) now entail

$$\begin{aligned} &\sup_{z\in K_{x,\delta}} \sup_{x'\in B(x,h)} \left| \frac{1-[U(\infty|x')]^{-1} z^{\gamma(x')} L(z|x')}{1-[U(\infty|x)]^{-1} z^{\gamma(x)} L(z|x)} - 1 \right| \\ = \ & \mathrm{O}\left((h^\alpha \log n_x \vee h^\beta) \sup_{z\in K_{x,\delta}} z^{\gamma(x)} L(z|x) \right). \end{aligned}$$

Potter bounds for the regularly varying function $z \mapsto z^{\gamma(x)}L(z|x)$ (see Theorem 1.5.6 in [4]) yield

$$\limsup_{n\to\infty} \sup_{z\in K_{x,\delta}} \frac{z^{\gamma(x)}L(z|x)}{(n_x/k_x)^{\gamma(x)} L(n_x/k_x|x)} < \infty$$

so that

$$\begin{aligned} &\sup_{z\in K_{x,\delta}} \sup_{x'\in B(x,h)} \left| \frac{1-[U(\infty|x')]^{-1} z^{\gamma(x')} L(z|x')}{1-[U(\infty|x)]^{-1} z^{\gamma(x)} L(z|x)} - 1 \right| \\ = \ & \mathrm{O}\left((h^\alpha \log n_x \vee h^\beta) \frac{L(n_x/k_x|x)}{(n_x/k_x)^{-\gamma(x)}} \right). \end{aligned} \tag{17}$$

Finally, use together (9), (10) and (17) to get

$$\Omega\left(\frac{n_x}{(1+\delta)k_x}, n_x^{1+\delta}, x, h \right) = \mathrm{O}\left(h^\beta \vee \left[(h^\alpha \log n_x) \frac{L(n_x/k_x|x)}{(n_x/k_x)^{-\gamma(x)}} \right] \right). \tag{18}$$

Equation (18) makes it clear that in this case, condition (8) shall be satisfied provided it holds that $h^\alpha \log n_x \to 0$ (which was already required in the Fréchet MDA) and

$$\frac{(n_x/k_x)^{-\gamma(x)}}{L(n_x/k_x|x)} h^\beta \to 0 \quad \text{as} \quad n\to\infty.$$

> We can conclude that compared to the case of the Fréchet MDA, there is an additional condition for the pointwise consistency of our estimator to hold in the Weibull MDA. This condition compares the oscillation of the conditional right endpoint to the proportion of order statistics used in the expression of the estimator.

We end this paragraph by noting that Theorem 1 is only a pointwise result. In the case when $E = \mathbb{R}^d$ and X has a probability density function f whose support S has a nonempty interior, it may be possible to obtain a uniform consistency result on every compact subset Ω of the interior of S, using for instance a method introduced by Härdle and Marron [23]: since Ω is a compact subset of $\mathbb{R}^d$ we may, for all $n \in \mathbb{N} \setminus \{0\}$, find a finite subset Ω_n of Ω such that

$$\forall x \in \Omega,\ \exists \chi(x) \in \Omega_n,\ \|x - \chi(x)\| \le n^{-\eta} \ \text{ and } \ \exists c > 0,\ |\Omega_n| = \mathrm{O}\left(n^c\right)$$

as $n \to \infty$, where $|\Omega_n|$ stands for the cardinality of Ω_n and $\eta > 0$ is suitably chosen (*i.e.* large enough). In other words, we may cover for every n the set Ω by a finite number of balls having a common radius which converges to 0 at a polynomial rate; we may also require that the set Ω_n of the centers of these balls is such that the cardinality of Ω_n grows at a polynomial rate. If γ is continuous on S, it is then enough to prove that for every $\delta > 0$,

$$|\Omega_n| \sup_{\omega \in \Omega_n} \mathbb{P}(|\widehat{\gamma}_n(\omega, k_\omega, h) - \gamma(\omega)| > \delta) \quad \to \quad 0 \tag{19}$$

$$\text{and } \mathbb{P}\left(\sup_{x \in \Omega} \left|\widehat{\gamma}_n(x, k_x, h) - \widehat{\gamma}_n(\chi(x), k_{\chi(x)}, h)\right| > \delta\right) \quad \to \quad 0 \tag{20}$$

as $n \to \infty$. Showing (19) involves finding a uniform bound for the probabilities

$$\mathbb{P}(|\widehat{\gamma}_n(\omega, k_\omega, h) - \gamma(\omega)| > \delta),\ \omega \in \Omega_n$$

while the proof of (20) relies on a careful study of the oscillation of the random function $x \mapsto \widehat{\gamma}_n(x, k_x, h)$. This is of course a challenging task, which shall be part of future research on this estimator.

3.2. Asymptotic normality

To prove a pointwise asymptotic normality result for our estimator, we need to introduce a second-order condition on the function $U(\cdot|x)$:

(M_2) Condition (M_1) holds and for every $x \in E$, there exist a real number $\rho(x) \le 0$ and a function $A(\cdot|x)$ of constant sign converging to 0 at infinity such that the function $U(\cdot|x)$ satisfies

$$\forall z > 0,\ \lim_{t \to \infty} \frac{\frac{U(tz|x) - U(t|x)}{a(t|x)} - \frac{z^{\gamma(x)} - 1}{\gamma(x)}}{A(t|x)} = H_{\gamma(x),\rho(x)}(z)$$

where

$$H_{\gamma(x),\rho(x)}(z) = \int_1^z r^{\gamma(x)-1} \left[\int_1^r s^{\rho(x)-1} ds \right] dr.$$

Hypothesis (M_2) is the conditional analogue of the classical second-order condition on U, see for instance Definition 2.3.1 and Corollary 2.3.4 in [21]: the parameter $\rho(x)$ is the so-called second-order parameter of Y given $X = x$. Note that Theorem 2.3.3 in [21] shows that the function $|A(\cdot|x)|$ is regularly varying at infinity with index $\rho(x)$. Moreover, as shown in Lemma B.3.16 therein, if (M_2) holds with $\gamma(x) \neq \rho(x)$ and $\rho(x) < 0$ if $\gamma(x) > 0$, then defining $q(\cdot|x) = a(\cdot|x)/U(\cdot|x)$, a second-order condition also holds for the function $\log U(\cdot|x)$, namely:

$$\forall z > 0,\ \lim_{t\to\infty} \frac{\frac{\log U(tz|x) - \log U(t|x)}{q(t|x)} - \frac{z^{\gamma_-(x)}-1}{\gamma_-(x)}}{Q(t|x)} = H_{\gamma_-(x),\rho'(x)}(z) \tag{21}$$

with

$$\rho'(x) = \begin{cases} \rho(x) & \text{if } \gamma(x) < \rho(x) \leq 0 \\ \gamma(x) & \text{if } \rho(x) < \gamma(x) \leq 0 \\ -\gamma(x) & \text{if } 0 < \gamma(x) < -\rho(x) \text{ and } \ell(x) \neq 0 \\ \rho(x) & \text{if } (0 < \gamma(x) < -\rho(x) \text{ and } \ell(x) = 0) \text{ or } 0 < -\rho(x) \leq \gamma(x) \end{cases}$$

where we have defined

$$\ell(x) = \lim_{t\to\infty} \left(U(t|x) - \frac{a(t|x)}{\gamma(x)} \right)$$

and $Q(\cdot|x)$ has ultimately constant sign, converges to 0 at infinity and is such that $|Q(\cdot|x)|$ is regularly varying at infinity with index $\rho'(x)$; note that Lemma B.3.16 in [21] entails that one can choose

$$Q(t|x) = \begin{cases} A(t|x) & \text{if} \quad \gamma(x) < \rho(x) \leq 0 \\[2ex] \gamma_+(x) - \dfrac{a(t|x)}{U(t|x)} & \begin{array}{ll} \text{if} & \rho(x) < \gamma(x) \leq 0 \\ \text{or} & 0 < \gamma(x) < -\rho(x) \text{ and } \ell(x) \neq 0 \\ \text{or} & 0 < \gamma(x) = -\rho(x) \end{array} \\[2ex] \dfrac{\rho(x)}{\gamma(x) + \rho(x)} A(t|x) & \begin{array}{ll} \text{if} & 0 < \gamma(x) < -\rho(x) \text{ and } \ell(x) = 0 \\ \text{or} & 0 < -\rho(x) < \gamma(x). \end{array} \end{cases}$$

Besides, if $\gamma(x) > 0$ and $\rho(x) = 0$, then according to Lemma B.3.16 in [21], one has

$$\forall z > 0,\ \lim_{t\to\infty} \frac{\frac{\log U(tz|x) - \log U(t|x)}{q(t|x)} - \log z}{Q(t|x)} = 0 \tag{22}$$

for every $Q(\cdot|x)$ such that $A(t|x) = \mathrm{O}(Q(t|x))$ as $t \to \infty$; especially, we can and will take $Q(\cdot|x) = A(\cdot|x)$ in this case.

We can now state the asymptotic normality of our estimator.

Theorem 2. *Assume that* (M_2) *holds. Pick* $x \in E$. *We assume that* $n_x \to \infty$, $k_x \to \infty$, $k_x/n_x \to 0$, $\sqrt{k_x}\, Q(n_x/k_x|x) \to \lambda(x) \in \mathbb{R}$ *and for some* $\delta > 0$

$$\sqrt{k_x}\frac{U(n_x/k_x|x)}{a(n_x/k_x|x)}\omega\left(\frac{n_x}{(1+\delta)k_x}, n_x^{1+\delta}, x, h\right) \to 0 \quad as \quad n \to \infty. \tag{23}$$

Then if $\gamma(x) \neq \rho(x)$, *it holds that* $\sqrt{k_x}(\widehat{\gamma}_n(x, k_x, h) - \gamma(x))$ *is asymptotically Gaussian with mean* $\lambda(x)B(\gamma(x), \rho(x))$ *and variance* $V(\gamma(x))$ *where we have set*

$$B(\gamma(x),\rho(x)) = \begin{cases} \dfrac{(1-\gamma(x))(1-2\gamma(x))}{(1-\gamma(x)-\rho(x))(1-2\gamma(x)-\rho(x))} & \text{if} \quad \gamma(x) < \rho(x) \leq 0 \\ \dfrac{\gamma(x)(1+\gamma(x))}{(1-\gamma(x))(1-3\gamma(x))} & \text{if} \quad \rho(x) < \gamma(x) \leq 0 \\ -\dfrac{\gamma(x)}{(1+\gamma(x))^2} & \begin{array}{ll}\text{if} & 0 < \gamma(x) < -\rho(x) \\ \text{and} & \ell(x) \neq 0\end{array} \\ \dfrac{\gamma(x) - \gamma(x)\rho(x) + \rho(x)}{\rho(x)(1-\rho(x))^2} & \begin{array}{ll}\text{if} & (0 < \gamma(x) < -\rho(x) \\ \text{and} & \ell(x) = 0) \\ \text{or} & 0 < -\rho(x) \leq \gamma(x)\end{array} \\ 1 & \text{if} \quad 0 = \rho(x) < \gamma(x) \end{cases}$$

and

$$V(\gamma(x)) = \begin{cases} \gamma^2(x) + 1 & \text{if } \gamma(x) \geq 0 \\ \dfrac{(1-\gamma(x))^2(1-2\gamma(x))(1-\gamma(x)+6\gamma^2(x))}{(1-3\gamma(x))(1-4\gamma(x))} & \text{if } \gamma(x) < 0. \end{cases}$$

Theorem 2 is the conditional analogue of the asymptotic normality result stated in [11]; see also Theorem 3.5.4 in [21]. In particular, the asymptotic bias and variance of our estimator are similar to those obtained in the univariate setting. Note that in this result, contrary to the asymptotic normality result of [18], we do not condition on the value of $N_n(x, h)$. Besides, condition $\sqrt{k_x}\, Q(n_x/k_x|x) \to \lambda(x) \in \mathbb{R}$ as $n \to \infty$ in Theorem 2 is a standard condition needed to control the bias of the estimator. Finally, hypothesis (23) can be replaced by a hypothesis on the uniform relative oscillation of the function $\log U$ in its second argument, see Proposition 1, which in turn can be made explicit if suitable regularity conditions are satisfied, see Section 3.1.

To illustrate this last remark, we use Theorem 2 to obtain optimal rates of convergence for our estimator. For the sake of simplicity, we assume that $E = \mathbb{R}^d$, $d \geq 1$ is equipped with the standard Euclidean distance and that X has a probability density function f on $\mathbb{R}^d$ which is continuous on its support S, assumed to have nonempty interior. If x is a point lying in the interior of S which is such that $f(x) > 0$, it is straightforward to show that

$$n_x = n \int_{B(x,h)} f(u)du = nh^d \mathcal{V} f(x)(1 + \mathrm{o}(1)) \quad \text{as} \quad n \to \infty$$

with $\mathcal{V}$ being the volume of the unit ball in $\mathbb{R}^d$. Letting

$$k = \frac{k_x}{h^d \mathcal{V} f(x)}$$

it becomes clear that $k_x = kh^d \mathcal{V} f(x)$ and that hypotheses $n_x \to \infty$, $k_x \to \infty$ and $k_x/n_x \to 0$ as $n \to \infty$ are equivalent to $kh^d \to \infty$ and $k/n \to 0$ as $n \to \infty$. If k and h have respective order n^a and n^{-b}, with $a, b > 0$, the rate of convergence of the estimator $\widehat{\gamma}_n(x, k_x, h)$ to $\gamma(x)$ is then $n^{(a-bd)/2}$. Under the hypotheses of Theorem 2, provided that (A_1) and (A_2) hold, one can find the optimal values for a and b in the case $\gamma(x) \neq 0$:

- If $\gamma(x) > 0$, then under the Hölder conditions (5), (6) and (7), hypothesis (23) shall be satisfied if

$$\sqrt{kh^d} h^\alpha \log(nh^d) \to 0 \quad \text{as} \quad n \to \infty.$$

  Recalling the bias condition $\sqrt{kh^d}\, Q(n/k|x) \to \lambda(x)/\sqrt{\mathcal{V} f(x)} \in \mathbb{R}$ as $n \to \infty$ and letting

$$\rho''(x) = \begin{cases} \rho'(x) & \text{if } \rho(x) < 0 \\ 0 & \text{if } \rho(x) = 0 \end{cases}$$

  the problem is thus to maximize the quantity $a - bd$ under the constraints $a \in (0,1)$, $a - bd \geq 0$,

$$\begin{aligned} a - b(d + 2\alpha) &\leq 0 \\ \text{and} \quad a - bd + 2(1-a)\rho''(x) &\leq 0. \end{aligned}$$

  The solution of this problem is

$$a^* = \frac{-(d+2\alpha)\rho''(x)}{\alpha - (d+2\alpha)\rho''(x)} \quad \text{and} \quad b^* = \frac{-\rho''(x)}{\alpha - (d+2\alpha)\rho''(x)}$$

  for which

$$a^* - b^* d = \frac{-2\alpha\rho''(x)}{\alpha - (d+2\alpha)\rho''(x)}.$$

  The optimal convergence rate for our estimator in this case is therefore

$$n^{(a^*-b^*d)/2} = n^{-\alpha\rho''(x)/(\alpha-(d+2\alpha)\rho''(x))}.$$

- If $\gamma(x) < 0$, then under the Hölder conditions (10), (12), (13) and (14), hypothesis (23) shall be satisfied if

$$\sqrt{kh^d} h^\alpha \log(nh^d) \to 0 \quad \text{and} \quad \sqrt{kh^d} \frac{(n/k)^{-\gamma(x)}}{L(n/k|x)} h^\beta \to 0 \quad \text{as} \quad n \to \infty.$$

  Recalling the bias condition $\sqrt{kh^d}\, Q(n/k|x) \to \lambda(x)/\sqrt{\mathcal{V} f(x)} \in \mathbb{R}$ as $n \to \infty$, the problem thus consists in maximizing the quantity $a - bd$ under the constraints $a \in (0,1)$, $a - bd \geq 0$,

$$\begin{aligned} a - b(d+2\alpha) &\leq 0, \\ a - bd + 2(1-a)\rho'(x) &\leq 0 \\ \text{and} \quad a - 2(1-a)\gamma(x) - b(d+2\beta) &\leq 0. \end{aligned}$$

To make things easier, we shall assume that the conditional right endpoint $U(\infty|\cdot)$ is not more regular than γ, or in other words, that $\beta \leq \alpha$. In this case, since $\gamma(x) < 0$, the constraints reduce to $a \in (0,1)$, $a - bd \geq 0$,

$$\begin{aligned} a - bd + 2(1-a)\rho'(x) &\leq 0 \\ \text{and} \quad a - 2(1-a)\gamma(x) - b(d+2\beta) &\leq 0. \end{aligned}$$

The solution of this problem is

$$a^* = \frac{-(d+2\beta)\rho'(x) - d\gamma(x)}{\beta - (d+2\beta)\rho'(x) - d\gamma(x)} \quad \text{and} \quad b^* = \frac{-\rho'(x) - \gamma(x)}{\beta - (d+2\beta)\rho'(x) - d\gamma(x)}$$

for which

$$a^* - b^* d = \frac{-2\beta\rho'(x)}{\beta - (d+2\beta)\rho'(x) - d\gamma(x)}.$$

The optimal convergence rate for our estimator in this case is then

$$n^{(a^*-b^*d)/2} = n^{-\beta\rho'(x)/(\beta-(d+2\beta)\rho'(x)-d\gamma(x))}.$$

4. Simulation study

To have an idea of how our estimator behaves on a finite sample situation, we carried out a simulation study in the case $E = [0,1] \subset \mathbb{R}$ equipped with the standard Euclidean distance with a covariate X which is uniformly distributed on E. Furthermore, we let $\gamma : E \to \mathbb{R}$ be the positive function defined by

$$\forall x \in [0,1],\ \gamma(x) = \frac{2}{3} + \frac{1}{3}\sin(2\pi x).$$

We consider three different models for the csf of Y given $X = x$: the first one is

$$\forall y > 0,\ \overline{F}_1(y|x) = \left(1 + y^{-\tau}\right)^{1/\tau\gamma(x)},$$

where the parameter τ is chosen to be independent of x and its value is picked in the set $\{-1.2, -1, -0.8\}$. In other words, Y given $X = x$ is Burr type XII distributed; note that in this case the csf $\overline{F}_1(\cdot|x)$ belongs to the Fréchet MDA for every $x \in E$, the conditional extreme-value index is $\gamma(x)$ and the conditional second-order parameter is $\rho(x) = \tau\gamma(x)$ (see [3], p. 93). The second model is

$$\forall y \in [0, g(x)],\ \overline{F}_2(y|x) = \frac{\Gamma(2/\gamma(x))}{\Gamma^2(1/\gamma(x))} \int_{y/g(x)}^1 t^{1/\gamma(x)-1}(1-t)^{1/\gamma(x)-1} dt$$

where $\Gamma : (0,\infty) \to \mathbb{R}$ is Euler's Gamma function, defined by

$$\forall \alpha > 0,\ \Gamma(\alpha) = \int_0^\infty e^{-t} t^{\alpha-1} dt$$

and the frontier function g is defined by

$$\forall x \in [0,1],\ g(x) = 1 - c + 8cx(1-x)$$

with the constant $c > 0$ being picked in the set $\{0.1, 0.2, 0.3\}$. In this case, given $X = x$, $Y/g(x)$ is a $\mathrm{Beta}(1/\gamma(x), 1/\gamma(x))$ random variable: this conditional model is contained in the Weibull MDA with the conditional extreme-value index being $-\gamma(x)$. The final model is

$$\forall y > 0,\ \overline{F}_3(y|x) = \int_{\log y}^\infty \frac{1}{\sqrt{2\pi\sigma^2(x)}} \exp\left(-\frac{(t-\mu(x))^2}{2\sigma^2(x)}\right) dt$$

where μ and σ are the functions defined by

$$\forall x \in [0,1],\ \mu(x) = \frac{2}{3} + \frac{1}{3}\sin(2\pi x) \quad \text{and} \quad \sigma(x) = 0.7 + 2.4x(1-x).$$

In this model, Y given $X = x$ has a log-normal distribution with parameters $\mu(x)$ and $\sigma^2(x)$, which is an example of a conditional distribution belonging to the Gumbel MDA.

The aim of this simulation study is to estimate the conditional extreme-value index on a grid of points $\{x_1, \ldots, x_M\}$ of $[0,1]$. We need to choose two parameters: the bandwidth h and the number of upper order statistics k_x. We use a selection procedure that was introduced in [18], which we recall for the sake of completeness.

1) For every bandwidth h in a grid $\{h_1, \ldots, h_P\}$ of possible values of h, we first make a preliminary choice of k_x. Let $\widehat{\gamma}_{i,j}(k) = \widehat{\gamma}_n(x_i, k, h_j)$ and $\lfloor \cdot \rfloor$ denote the floor function: for each $i \in \{1, \ldots, M\}$, $j \in \{1, \ldots, P\}$ and $k \in \{q_{i,j}+1, \ldots, N_n(x_i, h_j) - q_{i,j}\}$, where $q_{i,j} = \lfloor N_n(x_i, h_j)/10 \rfloor \vee 1$, we introduce the set $E_{i,j,k} = \{\widehat{\gamma}_{i,j}(\ell),\ \ell \in \{k - q_{i,j}, \ldots, k + q_{i,j}\}\}$. We compute the variance of the set $E_{i,j,k}$ for every possible value of k and we record the number $K_{i,j}$ for which this variance is minimal. More precisely,

$$K_{i,j} = \arg\min_k \frac{1}{2q_{i,j}+1} \sum_{\ell=k-q_{i,j}}^{k+q_{i,j}} \left(\widehat{\gamma}_{i,j}(\ell) - \overline{\widehat{\gamma}}_{i,j}(k)\right)^2$$

$$\text{with}\ \ \overline{\widehat{\gamma}}_{i,j}(k) = \frac{1}{2q_{i,j}+1} \sum_{\ell=k-q_{i,j}}^{k+q_{i,j}} \widehat{\gamma}_{i,j}(\ell).$$

We record the value $k_{i,j}$ such that $\widehat{\gamma}_{i,j}(k_{i,j})$ is the median of the set $E_{i,j,K_{i,j}}$. For the sake of simplicity, the estimate $\widehat{\gamma}_{i,j}(k_{i,j})$ will be denoted by $\widetilde{\gamma}_{i,j}$.

2) We now select the bandwidth h: let q' be a positive integer such that $2q' + 1 < P$. For each $i \in \{1, \ldots, M\}$ and $j \in \{q'+1, \ldots, P - q'\}$, let

$F_{i,j} = \{\widetilde{\gamma}_{i,\ell},\ \ell \in \{j-q',\dots,j+q'\}\}$ and compute the standard deviation $\sigma_i(j)$ of $F_{i,j}$:

$$\sigma_i^2(j) = \frac{1}{2q'+1}\sum_{\ell=j-q'}^{j+q'}(\widetilde{\gamma}_{i,\ell} - \overline{\widetilde{\gamma}}_{i,j})^2 \quad \text{with} \quad \overline{\widetilde{\gamma}}_{i,j} = \frac{1}{2q'+1}\sum_{\ell=j-q'}^{j+q'}\widetilde{\gamma}_{i,\ell}.$$

Our stability criterion is then the average of these quantities over the grid $\{x_1,\dots,x_M\}$:

$$\overline{\sigma}(j) = \frac{1}{M}\sum_{i=1}^{M}\sigma_i(j).$$

We next record the integer j^* such that $\overline{\sigma}(j^*)$ is the first local minimum of the application $j \mapsto \overline{\sigma}(j)$ which is less than the average value of $\overline{\sigma}$. In other words, $j^* = q'+1$ if $\overline{\sigma}$ is increasing, $j^* = P-q'$ if $\overline{\sigma}$ is decreasing and

$$j^* \;=\; \min\left\{ j \text{ such that } \overline{\sigma}(j) \le \overline{\sigma}(j-1)\wedge\overline{\sigma}(j+1) \right.$$
$$\left. \text{and } \overline{\sigma}(j) \le \frac{1}{P-2q'}\sum_{\ell=q'+1}^{P-q'}\overline{\sigma}(\ell) \right\} \tag{24}$$

otherwise, where we extend $\overline{\sigma}$ by setting $\overline{\sigma}(q') = \overline{\sigma}(q'+1)$ and $\overline{\sigma}(P-q'+1) = \overline{\sigma}(P-q')$.

The selected bandwidth is then independent of x and is given by $h^* = h_{j^*}$ where j^* is defined in (24). The selected number of upper order statistics is given, for $x = x_i$, by $k^*_{x_i} = k_{i,j^*}$. The main idea of this procedure is that the bandwidth and the number of upper order statistics are selected in order to satisfy a stability criterion. This estimation procedure is carried out on $N = 100$ independent samples of size $n = 500$. The conditional extreme-value index is estimated on a grid of $M = 50$ evenly spaced points in $[0,1]$. Regarding the selection procedure, $P = 25$ values of h ranging from 0.05 to 0.3 are tested; the parameter q' is set to 1.

To have an idea of our estimator behaves compared to another estimator in the conditional extreme-value index estimation literature, we introduce the estimator $\widetilde{\gamma}_D = \widehat{\gamma}_n^{RP,1}$ of [8]. Let K be the triweight kernel:

$$K(t) = \frac{35}{32}(1-t^2)^3\mathbb{1}_{[-1,1]}(t).$$

Let $\widehat{\overline{F}}(\cdot,h|x)$ be the empirical kernel estimator of the csf:

$$\widehat{\overline{F}}(y,h|x) = \frac{\dfrac{1}{nh^d}\displaystyle\sum_{i=1}^{n}K\left(\frac{x-X_i}{h}\right)\mathbb{1}_{\{Y_i>y\}}}{\dfrac{1}{nh^d}\displaystyle\sum_{i=1}^{n}K\left(\frac{x-X_i}{h}\right)}$$

TABLE 1
MSEs associated to the estimators in all cases

Situation	Moment estimator $\widehat{\gamma}$	Estimator $\widetilde{\gamma}_D$ of Daouia *et al.*
Model 1		
$\tau = -0.8$	0.1496	0.1962
$\tau = -1$	0.0781	0.1616
$\tau = -1.2$	0.0553	0.1586
Model 2		
$c = 0.1$	0.0686	0.1329
$c = 0.2$	0.0689	0.1257
$c = 0.3$	0.0825	0.1313
Model 3	0.3384	0.2801

and let $\widehat{q}_n(\cdot, h|x)$ be the generalized inverse of $\widehat{\overline{F}}(\cdot, h|x)$: for $\alpha \in (0,1)$,

$$\widehat{q}_n(\alpha, h|x) = \inf\left\{y \in \mathbb{R},\ \widehat{\overline{F}}(y, h|x) \leq \alpha\right\}.$$

The quantity $\widehat{q}_n(\cdot, h|x)$ is the empirical estimator of the conditional quantile function. The estimator $\widetilde{\gamma}_D$ is then

$$\widetilde{\gamma}_D(x, \alpha_{n,x}, h) = \frac{1}{-\log 3} \log\left(\frac{\widehat{q}_n(\alpha_{n,x}, h|x) - \widehat{q}_n(\alpha_{n,x}/3, h|x)}{\widehat{q}_n(\alpha_{n,x}/3, h|x) - \widehat{q}_n(\alpha_{n,x}/9, h|x)}\right)$$

where $\alpha_{n,x} \to 0$ as $n \to \infty$ is a nonrandom sequence. This estimator is exactly the estimator $\widehat{\gamma}_n^{RP,1}$ of [8] with $J = 3$ and $r = 1/J$; it is a kernel version of the Pickands estimator, see Pickands [25]. To choose the parameters $\alpha_{n,x}$ and h for $\widetilde{\gamma}_D$, we restrict our search to a parameter $\alpha_{n,x}$ having the form $k_x/N_n(x_i, h_j)$, so that we are led to a choice of k_x and h just as for our estimator, and we use the procedure detailed above.

We give in Table 1 the empirical mean squared errors (MSEs) of each estimator, averaged over the M points of the grid. Table 1 shows that our estimator outperforms the estimator $\widetilde{\gamma}_D$ in terms of MSEs in every case except the Gumbel one. Besides, one can see that in the Fréchet MDA, the MSEs of both estimators increase as $|\rho(x)|$ gets closer to 0, which was expected since $\rho(x)$ controls the rate of convergence in (M_2): the closer $|\rho(x)|$ is to 0, the slower is this convergence and the harder is the estimation. Some illustrations are given in Figures 1–3, where the estimations corresponding to the median of the MSE are represented in each case for both estimators. One can see on these pictures that our estimator generally oscillates less than $\widetilde{\gamma}_D$; in the case when the conditional survival function belongs to the Fréchet or Weibull MDA, it also does a better job of mimicking the shape of the conditional extreme-value index.

5. Real data example

In this section, we introduce a real fire insurance data set, provided by the reinsurance broker Aon Re Belgium. The data set consists of $n = 1823$ observations (S_i, C_i) where C_i is the claim size related to the $i-$th fire accident and S_i is the

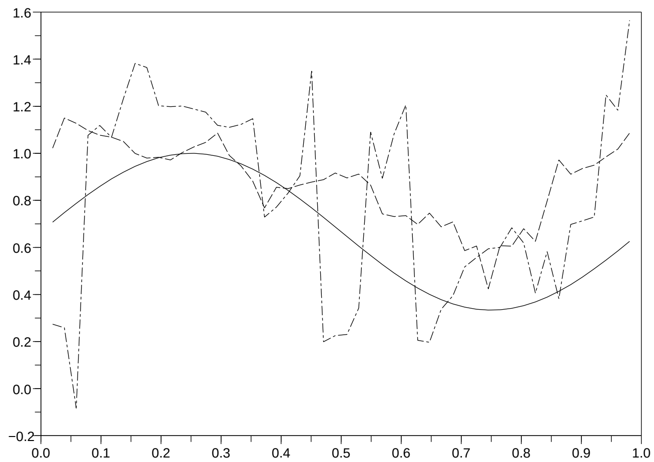

FIG 1. *Model 1, case $\tau = -1$: the true function γ (solid line), its estimators $\widehat{\gamma}$ (dashed line) and $\widetilde{\gamma}_D$ (dashed-dotted line), each corresponding to the median of the MSE.*

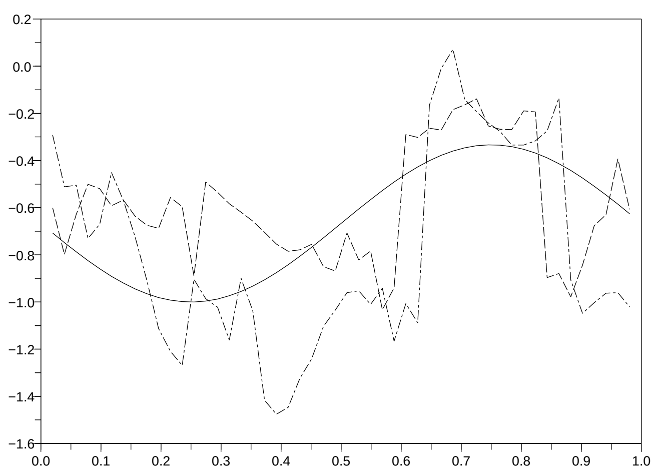

FIG 2. *Model 2, case $c = 0.3$: the true function γ (solid line) and its estimators $\widehat{\gamma}$ (dashed line) and $\widetilde{\gamma}_D$ (dashed-dotted line), each corresponding to the median of the MSE.*

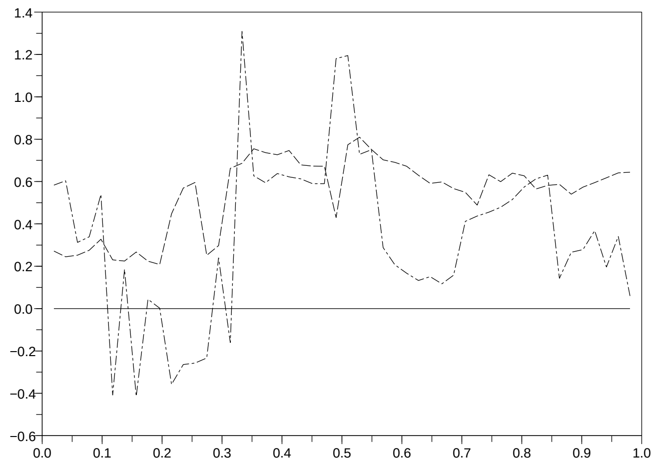

Fig 3. *Model 3: the true function γ (solid line) and its estimators $\widehat{\gamma}$ (dashed line) and $\widetilde{\gamma}_D$ (dashed-dotted line), each corresponding to the median of the MSE.*

associated total sum insured. It was, among others, considered by Beirlant and Goegebeur [2] and Beirlant *et al.* [1]; see also [3]. Our variable of interest is the ratio claim size/sum insured: in other words, we focus on the random variables $Y_i = C_i/S_i$. The covariate we consider is the total sum insured, which we can also consider as random; specifically, we let $X_i = \log S_i$. A scatterplot of the data is given in Figure 4.

In Section 7.6 of [3], the authors show that the distribution of the Y_i given $\log S_i$ can be approximated rather well by a General Pareto (GP) distribution. Our goal is then to provide an estimate of the conditional extreme-value index of the Y_i using our estimator. To this end, we use the selection procedure detailed in Section 4: the bandwidth h is selected among $h_1 \leq \cdots \leq h_{25}$ where the h_i are evenly spaced and

$$h_1 = 0.05(X_{n,n} - X_{1,n}) \quad \text{and} \quad h_{25} = 0.3(X_{n,n} - X_{1,n})$$

with $X_{1,n} \leq \ldots \leq X_{n,n}$ being the order statistics deduced from the X_i. This leads us to choose $h^* \approx 1.35$; a boxplot of the proportions $k^*_{x_i}/N_n(x_i, h^*)$ of order statistics used to compute the estimator is given in Figure 5. This allows us to give an estimate of the conditional extreme-value index, see Figure 6.

The first conclusion we draw from this study is that $\gamma(x) > 0$ for all x. This is somewhat surprising since one could expect the random variables Y_i to be bounded from above by 1. However, the GP fits discussed in [3] and the estimations carried out in [2] and [1] also lead to the same conclusion. One can think that in this case, modelling the distribution of the Y_i given $\log S_i$

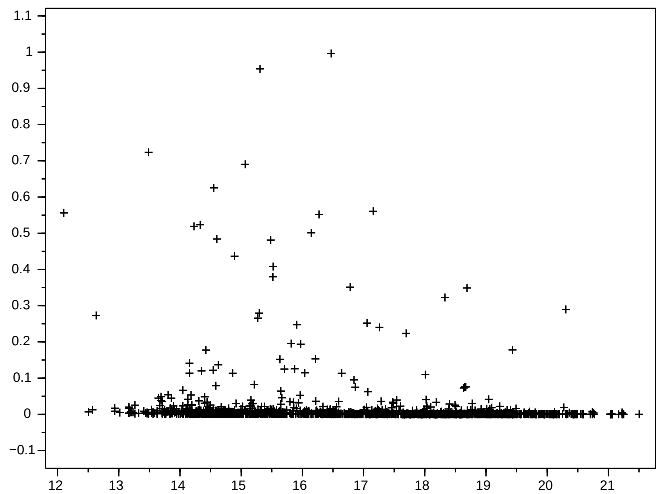

FIG 4. *Scatterplot of the Aon Re insurance data: $x-$axis: logarithm of the sum insured, $y-$axis: ratio claim/sum insured.*

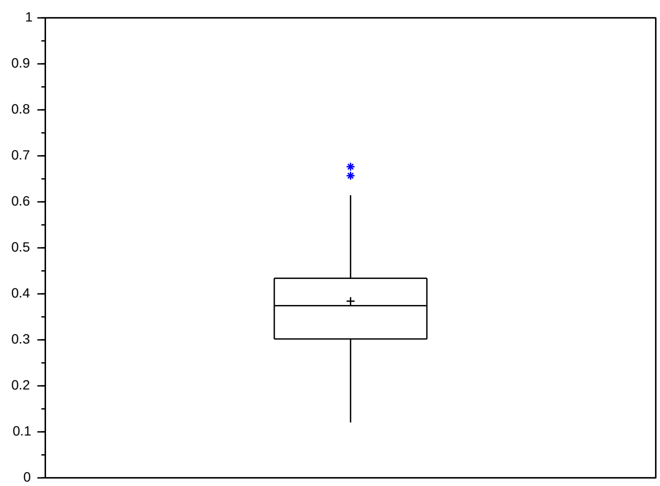

FIG 5. *Aon Re insurance data: boxplot of the proportions of order statistics used.*

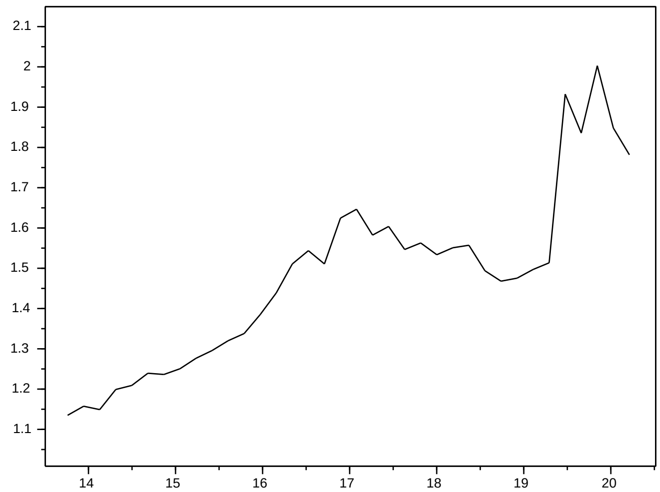

FIG 6. *Aon Re insurance data: $x-$axis: logarithm of the sum insured, full line: estimate $\widehat{\gamma}_n$.*

by a distribution belonging to the Fréchet domain of attraction is accurate in the "intermediate–upper" tail, namely, not too far into the upper tail of the distribution; an element backing this intuition is the exponential quantile plot given in Figure 7.17 of [3].

A second information is given by the shape of the estimated conditional extreme-value index. One can see that the estimator returns values that are greater than 1.1 for every considered value of the covariate. The study in [2], which considered the random variables C_i as variables of interest and splitted the random sample into three subgroups according to the type of buildings insured (which is an additional covariate information that we do not consider in this paper), provides estimations ranging from 1.027 to 1.413, while [1], which did not consider any covariate information at all, gives the estimate $\widehat{\gamma} = 1$. Our estimate can therefore be considered as a somewhat conservative one, especially when $\log S_i \geq 19.5$. Note that in this particular range, there are only very few (if any) high values of Y_i in the sample, which may be the cause for this phenomenon.

All in all, we can conclude that this study confirms previous findings about this data set, although the proposed estimator may at times give fairly conservative results. A possible direction for future research on this estimator is therefore to correct this behavior. One should keep in mind though that the essential advantage of the estimator studied in this paper is the fact that it works in every domain of attraction, making it superior to most others in this respect.

6. Proofs of the main results

6.1. Weak consistency

We start by proving the pointwise weak consistency of our estimator at a point x lying in E. To this end, since the $M_n^{(j)}(x, k_x, h)$ are defined conditionally on the value of the total number $N_n(x, h)$ of covariates belonging to $B(x, h)$, which is random, a natural idea is to condition on this value. A preliminary classical lemma is then required to control this random variable.

Lemma 1. *If $n_x \to \infty$ as $n \to \infty$ then for every $\delta > 0$*

$$\sqrt{n_x^{1-\delta}} \left| \frac{N_n(x,h)}{n_x} - 1 \right| \xrightarrow{\mathbb{P}} 0 \quad as \quad n \to \infty.$$

From Lemma 1, we deduce that if

$$I_x = \mathbb{N} \cap \left[\left(1 - n_x^{-1/4}\right) n_x, \left(1 + n_x^{-1/4}\right) n_x \right],$$

it holds that $N_n(x, h)$ lies in I_x with arbitrarily large probability as $n \to \infty$; in other words,

$$\sum_{p \in I_x} \mathbb{P}(N_n(x,h) = p) \to 1 \quad \text{as} \quad n \to \infty.$$

Furthermore, since $k_x/n_x \to 0$ as $n \to \infty$, we may and will, in the sequel, take n so large that $k_x < \inf I_x$.

The next step is to show that when n is large, studying the convergence in probability of the quantities $M_n^{(j)}(x, k_x, h)$ is equivalent to studying the behavior of analogous quantities defined in terms of upper order statistics of a sample of independent and identically distributed random variables having cdf $F(\cdot|x)$. To achieve that we begin by stating a lemma which gives the conditional distribution of the random variables Z_i.

Lemma 2. *Given $N_n(x, h) = p \geq 1$, the random variables Z_i, $1 \leq i \leq p$, are independent and identically distributed random variables having cdf $F_h(\cdot|x)$.*

Letting T_i, $i \geq 1$ be independent standard Pareto random variables, *i.e.* having cdf $t \mapsto 1 - 1/t$ on $(1, \infty)$, we deduce from this result that the distribution of the random vector $(Z_1, \dots, Z_p)$ given $N_n(x, h) = p \geq 1$ is the distribution of the random vector $(U_h(T_1|x), \dots, U_h(T_p|x))$. In other words, since $U_h(\cdot|x)$ is nondecreasing, we may focus on the behavior in probability of the quantities

$$\mathfrak{M}_{np}^{(j)}(x, k_x, h) = \frac{1}{k_x} \sum_{i=1}^{k_x} [\log U_h(T_{p-i+1,p}|x) - \log U_h(T_{p-k_x,p}|x)]^j$$

for $p > k_x$ and $j = 1, 2$. Lemma 3 below is the desired approximation of the statistics $\mathfrak{M}_{np}^{(j)}(x, k_x, h)$.

Lemma 3. *Given $N_n(x,h) = p \geq 1$, one has if $p > k_x$ that*

$$\left|\mathfrak{M}_{np}^{(1)}(x,k_x,h) - \mathcal{M}_{np}^{(1)}(x,k_x,h)\right| \leq 2\omega(T_{p-k_x,p}, T_{p,p}, x, h)$$

and

$$\begin{aligned} &\left|\mathfrak{M}_{np}^{(2)}(x,k_x,h) - \mathcal{M}_{np}^{(2)}(x,k_x,h)\right| \\ \leq \quad & 4\omega(T_{p-k_x,p}, T_{p,p}, x, h)\left[\omega(T_{p-k_x,p}, T_{p,p}, x, h) + \mathcal{M}_{np}^{(1)}(x,k_x,h)\right] \end{aligned}$$

where for $j = 1, 2$, we let

$$\mathcal{M}_{np}^{(j)}(x,k_x,h) = \frac{1}{k_x}\sum_{i=1}^{k_x}\left[\log U(T_{p-i+1,p}|x) - \log U(T_{p-k_x,p}|x)\right]^j.$$

The final lemmas are technical results. The first one is a simple result we shall repeatedly make use of.

Lemma 4. *Let (R_{ij}), (R'_{ij}), $1 \leq j \leq i$ be triangular arrays of random variables. Assume that there exist ℓ, $\ell' \in \mathbb{R}$ and a sequence of non-empty sets (I_n) contained in $\{1,\ldots,n\}$ such that for every $t > 0$*

$$\sup_{p\in I_n} \mathbb{P}(|R_{np} - \ell| > t) \to 0 \quad \text{and} \quad \sup_{p\in I_n} \mathbb{P}(|R'_{np} - \ell'| > t) \to 0 \quad \text{as} \quad n \to \infty.$$

Then for every Borel measurable function $h : \mathbb{R}^2 \to \mathbb{R}$ which is continuous at (ℓ,ℓ'), one has for every $t > 0$

$$\sup_{p\in I_n} \mathbb{P}(|h(R_{np}, R'_{np}) - h(\ell,\ell')| > t) \to 0 \quad \text{as} \quad n \to \infty.$$

The following result is the main technical tool we shall use to prove our asymptotic results. It is basically a conditional analogue of the additive version of Slutsky's lemma.

Lemma 5. *Let r be a positive integer, $(S_n) = (S_n^{(1)},\ldots,S_n^{(r)})$ be a sequence of random vectors and $S = (S^{(1)},\ldots,S^{(r)})$ be a random vector. Assume that there exist*

1. *a triangular array of events $(A_{ij})_{0\leq j\leq i}$ and a sequence of non-empty sets (I_n) contained in $\{1,\ldots,n\}$ such that*
   - *for every n the A_{np}, $0 \leq p \leq n$ have positive probability, are pairwise disjoint and*
   $$\sum_{p=0}^{n} \mathbb{P}(A_{np}) = 1;$$
   - *it holds that*
   $$\sum_{p\in I_n} \mathbb{P}(A_{np}) \to 1 \quad \text{as} \quad n \to \infty;$$

2. two triangular arrays of random vectors

$$(D_{ij} = (D_{ij}^{(1)}, \ldots, D_{ij}^{(r)}))_{1 \le j \le i} \quad \textit{and} \quad (R_{ij} = (R_{ij}^{(1)}, \ldots, R_{ij}^{(r)}))_{1 \le j \le i}$$

such that

- *for* $1 \le p \le n$*, the distribution of* S_n *given* A_{np} *is the distribution of* $D_{np} + R_{np}$*;*
- *it holds that for every* $t = (t_1, \ldots, t_r) \in \mathbb{R}^r$

$$\sup_{p \in I_n} \left| \mathbb{E}[\exp[it' D_{np}]] - \mathbb{E}[\exp[it' S]] \right| \to 0 \quad \textit{as} \quad n \to \infty$$

where t' *is the transpose vector of* t*;*

- *it holds that for every* $t > 0$ *and every* $j \in \{1, \ldots, r\}$

$$\sup_{p \in I_n} \mathbb{P}(|R_{np}^{(j)}| > t) \to 0 \quad \textit{as} \quad n \to \infty.$$

Then $S_n \xrightarrow{d} S$ *as* $n \to \infty$*. In particular, if* $D_{ij} = 0$ *for every* $1 \le j \le i$*, then* $S_n \xrightarrow{\mathbb{P}} 0$ *as* $n \to \infty$*.*

The next lemma, which specifies the asymptotic behavior in probability of the order statistic $T_{p-k_x,p}$ uniformly in $p \in I_x$, shall be used several times.

Lemma 6. *Assume that* $n_x \to \infty$*,* $k_x \to \infty$ *and* $k_x/n_x \to 0$ *as* $n \to \infty$*. Then for every* $t > 0$ *it holds that*

$$\sup_{p \in I_x} \mathbb{P}\left(\left| \frac{k_x}{p} T_{p-k_x,p} - 1 \right| > t \right) \to 0 \quad \textit{as} \quad n \to \infty.$$

Especially, for every $t > 0$

$$\sup_{p \in I_x} \mathbb{P}\left(T_{p-k_x,p} \le t \right) \to 0 \quad \textit{as} \quad n \to \infty$$

and for every function φ *which is regularly varying at infinity, we have both*

$$\sup_{p \in I_x} \mathbb{P}\left(\left| \frac{\varphi(T_{p-k_x,p})}{\varphi(p/k_x)} - 1 \right| > t \right) \to 0 \quad \textit{and} \quad \sup_{p \in I_x} \mathbb{P}\left(\left| \frac{\varphi(T_{p-k_x,p})}{\varphi(n_x/k_x)} - 1 \right| > t \right) \to 0$$

as $n \to \infty$*.*

Lemma 7 below is a useful corollary of Rényi's representation (see *e.g.* [21], p. 37).

Lemma 7. *For every Borel measurable functions* f *and* g*, every* $p \ge 2$ *and* $k \in \{1, \ldots, p-1\}$*, the random vectors*

$$\left(\frac{1}{k} \sum_{i=1}^{k} f\left(\log \frac{T_{p-i+1,p}}{T_{p-k,p}} \right), \ \frac{1}{k} \sum_{i=1}^{k} g\left(\log \frac{T_{p-i+1,p}}{T_{p-k,p}} \right) \right)$$

and

$$\left(\frac{1}{k}\sum_{i=1}^{k} f(\log T_i),\ \frac{1}{k}\sum_{i=1}^{k} g(\log T_i)\right)$$

have the same distribution.

The final lemma shows that the asymptotic behavior of the random variables $\mathcal{M}_{np}^{(j)}(x, k_x, h)$ is in some way uniform in $p \in I_x$. Before stating this result, we note that applying Theorem B.2.18 in [21], condition (M_1) entails that there exists a positive function $q_0(\cdot|x)$ which is equivalent to $a(\cdot|x)/U(\cdot|x)$ at infinity such that the following property holds: for each $\varepsilon > 0$, there exists $t_0 \geq 1$ such that for every $t \geq t_0$ and $z > 0$ with $tz \geq t_0$,

$$\left|\frac{\log U(tz|x) - \log U(t|x)}{q_0(t|x)} - \frac{z^{\gamma_-(x)} - 1}{\gamma_-(x)}\right| \leq \varepsilon z^{\gamma_-(x)}(z^{-\varepsilon} \vee z^{\varepsilon}). \tag{25}$$

Lemma 8. *Assume that (M_1) holds, and $n_x \to \infty$, $k_x \to \infty$ and $k_x/n_x \to 0$ as $n \to \infty$. Then for every $t > 0$ the convergences*

$$\sup_{p\in I_x} \mathbb{P}\left(\left|\frac{\mathcal{M}_{np}^{(1)}(x, k_x, h)}{q_0(p/k_x|x)} - \frac{1}{1-\gamma_-(x)}\right| > t\right) \to 0,$$

$$\textit{and} \quad \sup_{p\in I_x} \mathbb{P}\left(\left|\frac{\mathcal{M}_{np}^{(2)}(x, k_x, h)}{q_0^2(p/k_x|x)} - \frac{2}{(1-\gamma_-(x))(1-2\gamma_-(x))}\right| > t\right) \to 0$$

hold as $n \to \infty$.

We are now in position to examine the convergence in probability of the statistics $M_n^{(j)}(x, k_x, h)$, of which the consistency of our estimator is a simple corollary.

Proposition 2. *Assume that (M_1) holds, that $n_x \to \infty$, $k_x \to \infty$, $k_x/n_x \to 0$ and for some $\delta > 0$*

$$\frac{U(n_x/k_x|x)}{a(n_x/k_x|x)}\omega\left(\frac{n_x}{(1+\delta)k_x}, n_x^{1+\delta}, x, h\right) \to 0 \quad \textit{as} \quad n \to \infty.$$

Then it holds that

$$\frac{U(N_n(x,h)/k_x|x)}{a(N_n(x,h)/k_x|x)} M_n^{(1)}(x, k_x, h) \xrightarrow{\mathbb{P}} \frac{1}{1-\gamma_-(x)}$$

$$\textit{and} \quad \left[\frac{U(N_n(x,h)/k_x|x)}{a(N_n(x,h)/k_x|x)}\right]^2 M_n^{(2)}(x, k_x, h) \xrightarrow{\mathbb{P}} \frac{2}{(1-\gamma_-(x))(1-2\gamma_-(x))}$$

as $n \to \infty$.

This result is the analogue of Lemma 3.5.1 in [21] when there is a covariate: of course, a major difference here is that the total number of observations $N_n(x, h)$ is random.

Proof of Proposition 2. We start the proof by remarking that with the notation of (25), applying Lemma 1 yields

$$\frac{U(N_n(x,h)/k_x|x)}{a(N_n(x,h)/k_x|x)} q_0(N_n(x,h)/k_x|x) \xrightarrow{\mathbb{P}} 1 \quad \text{as} \quad n \to \infty. \tag{26}$$

Pick then an arbitrary $t > 0$ and introduce the two events

$$\begin{aligned} A_n^{(1)} &= \left\{ \left| \frac{M_n^{(1)}(x,k_x,h)}{q_0(N_n(x,h)/k_x|x)} - \frac{1}{1-\gamma_-(x)} \right| > t \right\} \\ \text{and} \quad A_n^{(2)} &= \left\{ \left| \frac{M_n^{(2)}(x,k_x,h)}{q_0^2(N_n(x,h)/k_x|x)} - \frac{2}{(1-\gamma_-(x))(1-2\gamma_-(x))} \right| > t \right\}. \end{aligned}$$

From (26), it is enough to prove that $\mathbb{P}(A_n^{(1)}) \to 0$ and $\mathbb{P}(A_n^{(2)}) \to 0$ as $n \to \infty$.

We start by controlling $\mathbb{P}(A_n^{(1)})$. Note that according to Lemma 2, one has

$$\mathbb{P}(A_n^{(1)}|N_n(x,h)=p) = \mathbb{P}\left(\left| \frac{\mathfrak{M}_{np}^{(1)}(x,k_x,h)}{q_0(p/k_x|x)} - \frac{1}{1-\gamma_-(x)} \right| > t \right).$$

Moreover, Lemma 3 entails

$$\begin{aligned} \left| \frac{\mathfrak{M}_{np}^{(1)}(x,k_x,h)}{q_0(p/k_x|x)} - \frac{1}{1-\gamma_-(x)} \right| &\leq \left| \frac{\mathcal{M}_{np}^{(1)}(x,k_x,h)}{q_0(p/k_x|x)} - \frac{1}{1-\gamma_-(x)} \right| \\ &+ \frac{2\omega(T_{p-k_x,p},T_{p,p},x,h)}{q_0(p/k_x|x)}. \end{aligned}$$

Introducing for an arbitrary $t' > 0$

$$\begin{aligned} u_{np}^{(1,1)} &= \mathbb{P}\left(\left| \frac{\mathcal{M}_{np}^{(1)}(x,k_x,h)}{q_0(p/k_x|x)} - \frac{1}{1-\gamma_-(x)} \right| > t' \right) \\ \text{and} \quad u_{np}^{(1,2)} &= \mathbb{P}\left(\frac{\omega(T_{p-k_x,p},T_{p,p},x,h)}{q_0(p/k_x|x)} > t' \right), \end{aligned}$$

Lemmas 4 and 5 with $A_{np} = \{N_n(x,h) = p\}$ make it enough to prove that $u_{np}^{(1,j)} \to 0$ as $n \to \infty$ uniformly in the integers $p \in I_x$ for every $j \in \{1,2\}$.

To control $u_{np}^{(1,1)}$ we apply Lemma 8 to obtain the convergence

$$\sup_{p\in I_x} u_{np}^{(1,1)} \to 0 \quad \text{as} \quad n \to \infty. \tag{27}$$

To control $u_{np}^{(1,2)}$ we recall that the function $q_0(\cdot|x)$ is regularly varying at infinity with index $\gamma_-(x)$ so that we can apply a uniform convergence result (see *e.g.* Theorem 1.5.2 in [4]) to get

$$\sup_{p\in I_x} \left| \frac{q_0(n_x/k_x|x)}{q_0(p/k_x|x)} - 1 \right| \to 0 \quad \text{as} \quad n \to \infty.$$

Especially, for n large enough, recalling that $q_0(\cdot|x)$ and $a(\cdot|x)/U(\cdot|x)$ are equivalent at infinity, we have

$$\sup_{p\in I_x} \frac{1}{q_0(p/k_x|x)} \leq 2\frac{U(n_x/k_x|x)}{a(n_x/k_x|x)}. \tag{28}$$

This inequality gives for n sufficiently large

$$\sup_{p\in I_x} u_{np}^{(1,2)} \leq \sup_{p\in I_x} \mathbb{P}\left(\frac{U(n_x/k_x|x)}{a(n_x/k_x|x)}\omega(T_{p-k_x,p}, T_{p,p}, x, h) > \frac{t'}{2}\right).$$

Using condition (1), we get for n large enough

$$\sup_{p\in I_x} u_{np}^{(1,2)} \leq \sup_{p\in I_x} \mathbb{P}\left(\left\{T_{p-k_x,p} < \frac{n_x}{(1+\delta)k_x}\right\} \cup \left\{T_{p,p} > n_x^{1+\delta}\right\}\right).$$

Because the random variables T_i are independent standard Pareto random variables, one has for n sufficiently large

$$\begin{aligned} \sup_{p\in I_x} u_{np}^{(1,2)} &\leq \sup_{p\in I_x} \mathbb{P}\left(T_{p-k_x,p} < \frac{n_x}{(1+\delta)k_x}\right) + \sup_{p\in I_x}\left[1-\left(1-n_x^{-1-\delta}\right)^p\right] \\ &\leq \sup_{p\in I_x} \mathbb{P}\left(T_{p-k_x,p} < \frac{n_x}{(1+\delta)k_x}\right) + \left[1-\left(1-n_x^{-1-\delta}\right)^{3n_x/2}\right] \end{aligned} \tag{29}$$

and the right-hand side in the last inequality converges to 0 as $n\to\infty$, by Lemma 6. Collecting (27) and (29) shows that $\mathbb{P}(A_n^{(1)})\to 0$ as $n\to\infty$.

Let us now consider $\mathbb{P}(A_n^{(2)})$. Applying Lemma 2, one has

$$\mathbb{P}(A_n^{(2)}|N_n(x,h)=p) = \mathbb{P}\left(\left|\frac{\mathfrak{M}_{np}^{(2)}(x,k_x,h)}{q_0^2(p/k_x|x)} - \frac{2}{(1-\gamma_-(x))(1-2\gamma_-(x))}\right| > t\right).$$

Lemma 3 yields

$$\begin{aligned} &\left|\frac{\mathfrak{M}_{np}^{(2)}(x,k_x,h)}{q_0^2(p/k_x|x)} - \frac{2}{(1-\gamma_-(x))(1-2\gamma_-(x))}\right| \\ \leq\ &\left|\frac{\mathcal{M}_{np}^{(2)}(x,k_x,h)}{q_0^2(p/k_x|x)} - \frac{2}{(1-\gamma_-(x))(1-2\gamma_-(x))}\right| \\ +\ &\left[\frac{2\omega(T_{p-k_x,p}, T_{p,p}, x, h)}{q_0(p/k_x|x)}\right]^2 + \frac{4\omega(T_{p-k_x,p}, T_{p,p}, x, h)}{q_0^2(p/k_x|x)}\mathcal{M}_{np}^{(1)}(x,k_x,h). \end{aligned}$$

Letting for an arbitrary $t'>0$

$$\begin{aligned} u_{np}^{(2,1)} &= \mathbb{P}\left(\left|\frac{\mathcal{M}_{np}^{(2)}(x,k_x,h)}{q_0^2(p/k_x|x)} - \frac{2}{(1-\gamma_-(x))(1-2\gamma_-(x))}\right| > t'\right), \\ u_{np}^{(2,2)} &= \mathbb{P}\left(\left[\frac{\omega(T_{p-k_x,p}, T_{p,p}, x, h)}{q_0(p/k_x|x)}\right]^2 > t'\right) \\ \text{and}\quad u_{np}^{(2,3)} &= \mathbb{P}\left(\frac{\omega(T_{p-k_x,p}, T_{p,p}, x, h)}{q_0^2(p/k_x|x)}\mathcal{M}_{np}^{(1)}(x,k_x,h) > t'\right), \end{aligned}$$

Lemmas 4 and 5 with $A_{np} = \{N_n(x,h) = p\}$ make it enough to prove that $u_{np}^{(2,j)} \to 0$ as $n \to \infty$ uniformly in the integers $p \in I_x$ for every $j \in \{1,2,3\}$. We start by noting that Lemma 8 leads to

$$\sup_{p \in I_x} u_{np}^{(2,1)} \to 0 \quad \text{as} \quad n \to \infty \tag{30}$$

and since

$$u_{np}^{(2,2)} = \mathbb{P}\left(\frac{\omega(T_{p-k_x,p}, T_{p,p}, x, h)}{q_0(p/k_x|x)} > \sqrt{t'}\right),$$

this term is similar to $u_{np}^{(1,2)}$ and therefore we obtain from (29) that

$$\sup_{p \in I_x} u_{np}^{(2,2)} \to 0 \quad \text{as} \quad n \to \infty. \tag{31}$$

Finally, the obvious inequality

$$\begin{aligned}
&\frac{\omega(T_{p-k_x,p}, T_{p,p}, x, h)}{q_0^2(p/k_x|x)} \mathcal{M}_{np}^{(1)}(x, k_x, h) \\
\leq \quad & \left[\frac{\omega(T_{p-k_x,p}, T_{p,p}, x, h)}{q_0(p/k_x|x)}\right] \left|\frac{\mathcal{M}_{np}^{(1)}(x, k_x, h)}{q_0(p/k_x|x)} - \frac{1}{1-\gamma_-(x)}\right| \\
+ \quad & \left[\frac{\omega(T_{p-k_x,p}, T_{p,p}, x, h)}{q_0(p/k_x|x)}\right] \frac{1}{1-\gamma_-(x)}
\end{aligned}$$

together with (27), (29) and Lemma 4 entails

$$\sup_{p \in I_x} u_{np}^{(2,3)} \to 0 \quad \text{as} \quad n \to \infty. \tag{32}$$

Collecting (30), (31) and (32) shows that $\mathbb{P}(A_n^{(2)}) \to 0$ as $n \to \infty$ which completes the proof. □

Proof of Theorem 1. Using Lemma 1.2.9 in [21] yields $a(t|x)/U(t|x) \to \gamma_+(x)$ as $t \to \infty$. Applying Proposition 2, we get

$$\widehat{\gamma}_{n,+}(x, k_x, h) \xrightarrow{\mathbb{P}} \gamma_+(x) \quad \text{and} \quad \widehat{\gamma}_{n,-}(x, k_x, h) \xrightarrow{\mathbb{P}} \gamma_-(x) \quad \text{as} \quad n \to \infty.$$

The result then follows from summing these two convergences. □

We conclude this section by proving Proposition 1. To this end, we state a couple of preliminary results. The first one of them links the behavior of a function having the form $1/\overline{F}$, where $\overline{F}$ is a csf on $\mathbb{R}$, to that of its left-continuous inverse.

Lemma 9. *Let $\overline{F}$ be a csf on $\mathbb{R}$ and U be the left-continuous inverse of $1/\overline{F}$.*

1. *If U is a continuous function on $(1, \infty)$ then*

$$\forall y \in \mathbb{R}, \; \overline{F}(y) \in (0,1) \Rightarrow \forall \delta > 0, \; \overline{F}(y+\delta) < \overline{F}(y).$$

2. *If U is an increasing function on $(1,\infty)$ then the function $\overline{F}$ is continuous on $\mathbb{R}$ and*

$$\forall z > 1,\ \frac{1}{\overline{F}(U(z))} = z.$$

The second lemma examines the properties of the csf $\overline{F}_h(\cdot|x)$.

Lemma 10. *Assume that $(\mathcal{E})$ and (A_1) hold. Assume further that $x \in E$ is such that*

$$\forall x' \in B(x,h),\ \forall r > 0,\ \mathbb{P}(X \in B(x',r)) > 0.$$

Then it holds that the function $\overline{F}_h(\cdot|x)$ is continuous on $\mathbb{R}$ and

$$\forall y \in \mathbb{R},\ \overline{F}_h(y|x) \in (0,1) \Rightarrow \forall \delta > 0,\ \overline{F}_h(y+\delta|x) < \overline{F}_h(y|x).$$

As a consequence

$$\forall z > 1,\ \frac{1}{\overline{F}_h(U_h(z|x)|x)} = z.$$

We may now show Proposition 1.

Proof of Proposition 1. We introduce the functions $\overline{F}_{\min}$ and $\overline{F}_{\max}$ defined by

$$\forall y \in \mathbb{R},\ \overline{F}_{\min}(y|x) = \inf_{x' \in B(x,h)} \overline{F}(y|x') \ \text{ and } \ \overline{F}_{\max}(y|x) = \sup_{x' \in B(x,h)} \overline{F}(y|x').$$

With this definition, we get

$$\overline{F}_{\min}(y|x) \le \overline{F}_h(y|x) = \frac{\mathbb{E}(\overline{F}(y|X)\mathbb{1}_{\{X \in B(x,h)\}})}{\mathbb{P}(X \in B(x,h))} \le \overline{F}_{\max}(y|x). \tag{33}$$

Applying Lemma 9, we obtain

$$\overline{F}_{\max}\left(\sup_{x' \in B(x,h)} U(z|x') \middle| x \right) \le \sup_{x' \in B(x,h)} \overline{F}(U(z|x')|x') = \frac{1}{z}$$

which, recalling (33), clearly entails

$$U_h(z|x) \le \sup_{x' \in B(x,h)} U(z|x'). \tag{34}$$

Likewise,

$$\overline{F}_{\min}\left(\inf_{x' \in B(x,h)} U(z|x') \middle| x \right) \ge \inf_{x' \in B(x,h)} \overline{F}(U(z|x')|x') = \frac{1}{z}.$$

Inequality (33) and Lemma 10 therefore entail

$$\overline{F}_h(U_h(z|x)|x) = \frac{1}{z} \le \overline{F}_h\left(\inf_{x' \in B(x,h)} U(z|x') \middle| x \right).$$

Applying Lemma 10 once again leads to the inequality

$$\inf_{x'\in B(x,h)} U(z|x') \le U_h(z|x). \tag{35}$$

From (34) and (35) we deduce that

$$\inf_{x'\in B(x,h)} \log\frac{U(z|x')}{U(z|x)} \le \log\frac{U_h(z|x)}{U(z|x)} \le \sup_{x'\in B(x,h)} \log\frac{U(z|x')}{U(z|x)}$$

because the logarithm function is increasing. This yields

$$\begin{aligned}\left|\log\frac{U_h(z|x)}{U(z|x)}\right| &= \left[\log\frac{U_h(z|x)}{U(z|x)}\right] \vee \left[-\log\frac{U_h(z|x)}{U(z|x)}\right] \\ &\le \left[\sup_{x'\in B(x,h)} \log\frac{U(z|x')}{U(z|x)}\right] \vee \left[-\inf_{x'\in B(x,h)} \log\frac{U(z|x')}{U(z|x)}\right]. \end{aligned} \tag{36}$$

The obvious inequality

$$-\left|\log\frac{U(z|x')}{U(z|x)}\right| \le \log\frac{U(z|x')}{U(z|x)} \le \left|\log\frac{U(z|x')}{U(z|x)}\right|$$

leads to

$$\sup_{x'\in B(x,h)} \log\frac{U(z|x')}{U(z|x)} \le \sup_{x'\in B(x,h)} \left|\log\frac{U(z|x')}{U(z|x)}\right| \tag{37}$$

$$\text{and} \quad -\inf_{x'\in B(x,h)} \log\frac{U(z|x')}{U(z|x)} \le \sup_{x'\in B(x,h)} \left|\log\frac{U(z|x')}{U(z|x)}\right|. \tag{38}$$

Collecting (36), (37) and (38) concludes the proof. □

6.2. Asymptotic normality

We proceed by proving the pointwise asymptotic normality of the estimator at a point $x \in E$ when condition (M_2) holds. We shall use the same ideas as in the proof of Proposition 2 to examine the asymptotic behavior of the statistics $M_n^{(j)}(x, k_x, h)$: if $\gamma(x) \neq \rho(x)$ and $\rho(x) < 0$ if $\gamma(x) > 0$, then from (21) and Theorem 2.3.6 in [21], there exist functions $q_0(\cdot|x)$ and $Q_0(\cdot|x)$ which are equivalent to $q(\cdot|x)$ and

$$\frac{1}{\rho'(x)} Q(\cdot|x) \mathbb{1}_{\{\rho'(x)<0\}} + Q(\cdot|x) \mathbb{1}_{\{\rho'(x)=0\}} \tag{39}$$

respectively at infinity such that for every $\varepsilon > 0$ there exists $t_0 \ge 1$ such that for every $t \ge t_0$ and $z > 0$ with $tz \ge t_0$,

$$\left| \frac{\frac{\log U(tz|x) - \log U(t|x)}{q_0(t|x)} - \frac{z^{\gamma_-(x)}-1}{\gamma_-(x)}}{Q_0(t|x)} - \psi_{\gamma_-(x),\rho'(x)}(z) \right| \le \varepsilon z^{\gamma_-(x)+\rho'(x)} (z^\varepsilon \vee z^{-\varepsilon}) \tag{40}$$

where

$$\psi_{\gamma_-(x),\rho'(x)}(z) = \begin{cases} \dfrac{z^{\gamma_-(x)+\rho'(x)} - 1}{\gamma_-(x) + \rho'(x)} & \text{if } \rho'(x) < 0, \\ \dfrac{z^{\gamma_-(x)} \log z}{\gamma_-(x)} & \text{if } \gamma_-(x) < \rho'(x) = 0, \\ \dfrac{1}{2}(\log x)^2 & \text{if } \gamma_-(x) = \rho'(x) = 0. \end{cases}$$

If now $\gamma(x) > 0$ and $\rho(x) = 0$, recalling the equality $q(\cdot|x) = a(\cdot|x)/U(\cdot|x)$, we get from Lemma B.3.16 in [21] that

$$q(t|x) - \gamma(x) = Q(t|x)(1 + \mathrm{o}(1)) \quad \text{as} \quad t \to \infty. \tag{41}$$

Equation (22) thus yields

$$\forall z > 0,\ \lim_{t\to\infty} \frac{\log U(tz|x) - \log U(t|x) - \gamma(x)\log z}{Q(t|x)} = \log z.$$

We may now apply Theorem B.2.18 in [21] to obtain that for every $\varepsilon > 0$ there exists $t_0 \geq 1$ such that for every $t \geq t_0$ and $z > 0$ with $tz \geq t_0$,

$$\left| \frac{\log U(tz|x) - \log U(t|x) - \gamma(x)\log z}{Q(t|x)} - \log z \right| \leq \varepsilon(z^{\varepsilon} \vee z^{-\varepsilon}). \tag{42}$$

Using together (41), (42) and the fact that the function $z \mapsto (z^{\varepsilon} \vee z^{-\varepsilon})^{-1} \log z$ is bounded on $(0, \infty)$, we get that for every $\varepsilon > 0$ there exists $t_0 \geq 1$ (possibly different) such that for every $t \geq t_0$ and $z > 0$ with $tz \geq t_0$,

$$\left| \frac{\frac{\log U(tz|x) - \log U(t|x)}{q(t|x)} - \log z}{Q(t|x)} \right| \leq \varepsilon(z^{\varepsilon} \vee z^{-\varepsilon}).$$

The following result is the analogue of Lemma 3.5.5 in [21] when there is a random covariate: let $\mathcal{V}(\gamma(x))$ be the matrix

$$\frac{1}{(1-\gamma_-(x))^2(1-2\gamma_-(x))} \begin{pmatrix} 1 & \dfrac{4}{1-3\gamma_-(x)} \\ \dfrac{4}{1-3\gamma_-(x)} & \dfrac{4(5-11\gamma_-(x))}{(1-2\gamma_-(x))(1-3\gamma_-(x))(1-4\gamma_-(x))} \end{pmatrix}$$

and note that if $\gamma(x) > 0$ then

$$\mathcal{V}(\gamma(x)) = \begin{pmatrix} 1 & 4 \\ 4 & 20 \end{pmatrix}.$$

Lemma 11. *Assume that* (M_2) *holds, that* $n_x \to \infty$, $k_x \to \infty$, $k_x/n_x \to 0$, $\sqrt{k_x}\, Q(n_x/k_x|x) \to \lambda(x) \in \mathbb{R}$ *and for some* $\delta > 0$

$$\sqrt{k_x} \frac{U(n_x/k_x|x)}{a(n_x/k_x|x)} \omega\left(\frac{n_x}{(1+\delta)k_x}, n_x^{1+\delta}, x, h \right) \to 0 \quad \text{as} \quad n \to \infty.$$

- *If* $\gamma(x) \neq \rho(x)$ *and* $\rho(x) < 0$ *if* $\gamma(x) > 0$*, it holds that the distribution of the random vector*

$$\sqrt{k_x}\left(\frac{\mathfrak{M}_{np}^{(1)}(x,k_x,h)}{q_0(T_{p-k_x,p}|x)}-\frac{1}{1-\gamma_-(x)},\frac{\mathfrak{M}_{np}^{(2)}(x,k_x,h)}{q_0^2(T_{p-k_x,p}|x)}-\frac{2}{(1-\gamma_-(x))(1-2\gamma_-(x))}\right)$$

is the distribution of a random vector $(D_{np}^{(1)}+R_{np}^{(1)},D_{np}^{(2)}+R_{np}^{(2)})$ *where*

  - *the triangular array* $(D_{ij}^{(1)},D_{ij}^{(2)})_{1\le j\le i}$ *is such that for every* $(t_1, t_2) \in \mathbb{R}^2$,

$$\sup_{p\in I_x}\left|\mathbb{E}[\exp[i(t_1D_{np}^{(1)}+t_2D_{np}^{(2)})]-\mathbb{E}[\exp[i(t_1P_1+t_2P_2)]]\right|\to 0 \quad as \quad n\to\infty$$

    where (P_1,P_2) *is a Gaussian random vector having mean* $(m^{(1)}(x), m^{(2)}(x))$ *with*

$$\begin{aligned} m^{(1)}(x) &= \lambda(x)\left(\frac{\mathbb{1}_{\{\rho'(x)<0\}}}{\rho'(x)}+\mathbb{1}_{\{\rho'(x)=0\}}\right)\mathbb{E}(\psi_{\gamma_-(x),\rho'(x)}(T)),\\ m^{(2)}(x) &= 2\lambda(x)\left(\frac{\mathbb{1}_{\{\rho'(x)<0\}}}{\rho'(x)}+\mathbb{1}_{\{\rho'(x)=0\}}\right)\\ &\qquad\times\mathbb{E}\left(\frac{T^{\gamma_-(x)}-1}{\gamma_-(x)}\psi_{\gamma_-(x),\rho'(x)}(T)\right)\end{aligned}$$

    and covariance matrix $\mathcal{V}(\gamma(x))$*;*

  - *the triangular arrays of random variables* $(R_{ij}^{(1)})_{1\le j\le i}$ *and* $(R_{ij}^{(2)})_{1\le j\le i}$ *are such that for every* $t>0$ *and* $j\in\{1,2\}$,

$$\sup_{p\in I_x}\mathbb{P}(|R_{np}^{(j)}|>t)\to 0 \quad as \quad n\to\infty.$$

- *If* $\gamma(x) > 0$ *and* $\rho(x) = 0$*, it holds that the distribution of the random vector*

$$\sqrt{k_x}\left(\frac{\mathfrak{M}_{np}^{(1)}(x,k_x,h)}{q(T_{p-k_x,p}|x)}-1,\frac{\mathfrak{M}_{np}^{(2)}(x,k_x,h)}{q^2(T_{p-k_x,p}|x)}-2\right)$$

is the distribution of a random vector $(D_{np}^{(1)}+R_{np}^{(1)},D_{np}^{(2)}+R_{np}^{(2)})$ *where*

  - *the triangular array* $(D_{ij}^{(1)},D_{ij}^{(2)})_{1\le j\le i}$ *is such that for every* $(t_1, t_2) \in \mathbb{R}^2$,

$$\sup_{p\in I_x}\left|\mathbb{E}[\exp[i(t_1D_{np}^{(1)}+t_2D_{np}^{(2)})]-\mathbb{E}[\exp[i(t_1P_1+t_2P_2)]]\right|\to 0 \quad as \quad n\to\infty$$

    where (P_1,P_2) *is a Gaussian centered random vector having covariance matrix* $\mathcal{V}(\gamma(x))$*;*

– *the triangular arrays of random variables* $(R_{ij}^{(1)})_{1\le j\le i}$ *and* $(R_{ij}^{(2)})_{1\le j\le i}$ *are such that for every* $t>0$ *and* $j\in\{1,2\}$,

$$\sup_{p\in I_x}\mathbb{P}(|R_{np}^{(j)}|>t)\to 0 \quad as \quad n\to\infty.$$

This result paves the way for a proof of Theorem 2.

Proof of Theorem 2. According to Lemma 2, the distribution of the random pair $(\widehat{\gamma}_{n,+}(x,k_x,h),\widehat{\gamma}_{n,-}(x,k_x,h))$ given $N_n(x,h)=p$ is that of

$$\left(\mathfrak{M}_{np}^{(1)}(x,k_x,h), 1-\frac{1}{2}\left(1-\frac{\left[\mathfrak{M}_{np}^{(1)}(x,k_x,h)\right]^2}{\mathfrak{M}_{np}^{(2)}(x,k_x,h)}\right)^{-1}\right).$$

Arguing along the first lines of the proof of Theorem 3.5.4 in [21] and applying Lemmas 4 and 5 with $A_{np}=\{N_n(x,h)=p\}$ together with Lemma 11 and the continuous mapping theorem, we then get that

$$\begin{pmatrix} \sqrt{k_x}\,(\widehat{\gamma}_{n,+}(x,k_x,h)-\gamma_+(x)) \\ \sqrt{k_x}\,(\widehat{\gamma}_{n,-}(x,k_x,h)-\gamma_-(x)) \end{pmatrix}$$
$$\stackrel{d}{\longrightarrow} \left(\begin{array}{c} \gamma_+(x)P_1+\dfrac{\lambda(x)}{1-\gamma_-(x)}\left(\mathbb{1}_{\{\gamma(x)>\rho(x)=0\}}-\mathbb{1}_{\{\rho(x)<\gamma(x)\le 0\}}\right) \\ (1-2\gamma_-(x))(1-\gamma_-(x))^2\left[\left(\dfrac{1}{2}-\gamma_-(x)\right)P_2-2P_1\right] \end{array}\right)$$

as $n\to\infty$, where (P_1,P_2) is the limit vector in Lemma 11. The result thus follows from Lemma 11 and some straightforward but lengthy computations. □

7. Proofs of the auxiliary results

Proof of Lemma 1. The proof is a straightforward consequence of the fact that $N_n(x,h)$ is a binomial random variable with parameters n and $\mathbb{P}(X\in B(x,h))$ and of Chebyshev's inequality. □

Proof of Lemma 2. If $(z_1,\ldots,z_p)\in(0,1)^p$, then since the random pairs (X_i,Y_i) have the same distribution, it holds that

$$\begin{aligned}\mathbb{P}\left(\bigcap_{i=1}^{p}\{Z_i\le z_i\}, N_n(x,h)=p\right) &= \binom{n}{p}\mathbb{P}\left(\bigcap_{i=1}^{p}\{Y_i\le z_i, X_i\in B(x,h)\}\right) \\ &\times \prod_{i=p+1}^{n}\mathbb{P}\left(X_i\notin B(x,h)\right).\end{aligned}$$

The definition of $F_h(\cdot|x)$ and the independence of the random pairs (X_i,Y_i), $i=1,\ldots,n$ entail that the above probability is

$$\prod_{i=1}^{p}F_h(z_i|x)\times\left[\binom{n}{p}\prod_{i=1}^{p}\mathbb{P}\left(X_i\in B(x,h)\right)\times\prod_{i=p+1}^{n}\mathbb{P}\left(X_i\notin B(x,h)\right)\right]$$

Since $N_n(x,h)$ is a binomial random variable with parameters n and $\mathbb{P}(X \in B(x,h))$, it becomes clear that given $N_n(x,h) = p$, the Z_i, $i = 1,\dots,p$ are independent and identically distributed random variables having cdf $F_h(\cdot|x)$, from which the result follows. □

Proof of Lemma 3. We start by writing the obvious inequality

$$|\log U_h(T_{p-i+1,p}|x) - \log U(T_{p-i+1,p}|x)| \le \omega(T_{p-k_x,p}, T_{p,p}, x, h) \qquad (43)$$

valid for every $i \in \{1,\dots,k_x+1\}$. The first part of the result is thus a straightforward consequence of (43) and the triangle inequality. To prove the second part, note that according to (43), for every $i = 1,\dots,k_x$

$$\begin{aligned}
&\left|\left[\log \frac{U_h(T_{p-i+1,p}|x)}{U_h(T_{p-k_x,p}|x)}\right]^2 - \left[\log \frac{U(T_{p-i+1,p}|x)}{U(T_{p-k_x,p}|x)}\right]^2\right| \\
\le\ & 2\omega(T_{p-k_x,p}, T_{p,p}, x, h)\left|\log \frac{U_h(T_{p-i+1,p}|x)}{U_h(T_{p-k_x,p}|x)} + \log \frac{U(T_{p-i+1,p}|x)}{U(T_{p-k_x,p}|x)}\right| \\
\le\ & 4\omega(T_{p-k_x,p}, T_{p,p}, x, h)\left[\omega(T_{p-k_x,p}, T_{p,p}, x, h) + \log \frac{U(T_{p-i+1,p}|x)}{U(T_{p-k_x,p}|x)}\right].
\end{aligned}$$

The result on $\mathfrak{M}^{(2)}_{np}(x,k_x,h)$ then follows from the triangle inequality and from summing the above inequalities for $i = 1,\dots,k_x$. □

Proof of Lemma 4. Since h is continuous at (ℓ,ℓ'), we can write

$$\forall t > 0,\ \exists \delta(t) > 0,\ |x-\ell| \vee |y-\ell'| \le \delta(t) \Rightarrow |h(x,y) - h(\ell,\ell')| \le t.$$

In other words, one has

$$|h(R_{np}, R'_{np}) - h(\ell,\ell')| > t \Rightarrow |R_{np} - \ell| \vee |R'_{np} - \ell'| > \delta(t).$$

This entails

$$\begin{aligned}
\sup_{p\in I_n} \mathbb{P}(|h(R_{np}, R'_{np}) - h(\ell,\ell')| > t) \quad &\le \quad \sup_{p\in I_n} \mathbb{P}(|R_{np} - \ell| > \delta(t)) \\
&+ \quad \sup_{p\in I_n} \mathbb{P}(|R'_{np} - \ell'| > \delta(t))
\end{aligned}$$

and the right-hand side converges to 0 as $n \to \infty$, which completes the proof. □

Proof of Lemma 5. Start by writing, for every $t = (t_1,\dots,t_r) \neq (0,\dots,0)$

$$\begin{aligned}
\mathbb{E}[\exp(it'S_n)] - \mathbb{E}[\exp(it'S)] \quad &= \quad \Big[\mathbb{E}[\exp(it'S_n)|A_{n0}] - \mathbb{E}[\exp(it'S)]\Big]\mathbb{P}(A_{n0}) \\
&+ \quad \sum_{p=1}^{n} \Big[\mathbb{E}[\exp(it'S_n)|A_{np}] - \mathbb{E}[\exp(it'S)]\Big]\mathbb{P}(A_{np}).
\end{aligned}$$

Pick an arbitrary $\delta > 0$: for n large enough, the triangle inequality yields

$$\begin{aligned} & \Big|\mathbb{E}[\exp(it'S_n)] - \mathbb{E}[\exp(it'S)]\Big| \\ & \leq \frac{\delta}{2} + \sup_{p\in I_n} \Big|\mathbb{E}[\exp[it'(D_{np}+R_{np})]] - \mathbb{E}[\exp(it'S)]\Big|. \end{aligned} \tag{44}$$

We now bound the term on the right-hand side of this inequality as

$$\begin{aligned} & \sup_{p\in I_n} \Big|\mathbb{E}[\exp[it'(D_{np}+R_{np})]] - \mathbb{E}[\exp(it'S)]\Big| \\ \leq & \sup_{p\in I_n} \Big|\mathbb{E}[\exp[it'(D_{np}+R_{np})]] - \mathbb{E}[\exp[it'D_{np}]]\Big| \\ + & \sup_{p\in I_n} \Big|\mathbb{E}[\exp[it'D_{np}]] - \mathbb{E}[\exp[it'S]]\Big|. \end{aligned} \tag{45}$$

The second term of the above inequality is controlled using the hypothesis on the array (D_{ij}): we have for n sufficiently large

$$\sup_{p\in I_n} \Big|\mathbb{E}[\exp[it'D_{np}]] - \mathbb{E}[\exp[it'S]]\Big| \leq \frac{\delta}{4}. \tag{46}$$

Besides, using once again the triangle inequality entails, if $||t||_\infty = \max_{1\leq j\leq r} |t_j|$,

$$\begin{aligned} & \Big|\mathbb{E}[\exp[it'(D_{np}+R_{np})]] - \mathbb{E}[\exp(it'D_{np})]\Big| \\ \leq \ & \mathbb{E}\left[|\exp(it'R_{np}) - 1| \mathbb{1}_{\{\max_{1\leq j\leq r} |R_{np}^{(j)}| \leq \delta/8r||t||_\infty\}}\right] \\ + \ & 2\mathbb{P}\left(\max_{1\leq j\leq r} |R_{np}^{(j)}| > \delta/8r||t||_\infty\right). \end{aligned}$$

Applying the mean value theorem to the function $z \mapsto e^{iz}$ and using the hypothesis on the array (R_{np}) leads to

$$\sup_{p\in I_n} \Big|\mathbb{E}[\exp[it'(D_{np}+R_{np})]] - \mathbb{E}[\exp(it'D_{np})]\Big| \leq \frac{\delta}{4} \tag{47}$$

for n large enough. Collecting (44), (45), (46) and (47) makes it clear that

$$\Big|\mathbb{E}[\exp(it'S_n)] - \mathbb{E}[\exp(it'S)]\Big| \leq \delta$$

for n large enough. Using the Cramér-Wold device concludes the proof. □

Proof of Lemma 6. Pick $t \in (0,1)$, $p \in I_x$ and write, if $t' = \log(1+t) \wedge (-\log(1-t))$,

$$\begin{aligned} \mathbb{P}\left(\left|\frac{k_x}{p} T_{p-k_x,p} - 1\right| > t\right) &= \mathbb{P}\left(\log(T_{p-k_x,p}) - \log\left(\frac{p}{k_x}\right) > \log(1+t)\right) \\ &+ \mathbb{P}\left(\log(T_{p-k_x,p}) - \log\left(\frac{p}{k_x}\right) < \log(1-t)\right) \\ &\leq 2\,\mathbb{P}\left(\left|\log(T_{p-k_x,p}) - \log\left(\frac{p}{k_x}\right)\right| > t'\right). \end{aligned}$$

It is therefore enough to prove that

$$u_{np} = \mathbb{P}\left(\left|\log(T_{p-k_x,p}) - \log\left(\frac{p}{k_x}\right)\right| > t'\right) \to 0$$

uniformly in $p \in I_x$ as $n \to \infty$. To this end, since the random variables $\log T_i$, $1 \leq i \leq p$ are independent standard exponential random variables, we get according to Rényi's representation

$$\log(T_{p-k_x,p}) \stackrel{d}{=} \sum_{j=1}^{p-k_x} \frac{\log T_j}{p-j+1}. \tag{48}$$

Besides, the inequalities

$$\log\left(\frac{p+1}{k_x+1}\right) = \int_{k_x+1}^{p+1} \frac{dv}{v} \leq \sum_{j=1}^{p-k_x} \frac{1}{p-j+1} \leq \int_{k_x}^{p} \frac{dv}{v} = \log\left(\frac{p}{k_x}\right)$$

yield

$$\left|\log\left(\frac{p}{k_x}\right) - \sum_{j=1}^{p-k_x} \frac{1}{p-j+1}\right| \leq \log\left(\frac{p}{k_x}\right) - \log\left(\frac{p+1}{k_x+1}\right) = \log\left(\frac{1+k_x^{-1}}{1+p^{-1}}\right).$$

Using the classical inequality $\log(1+s) \leq s$ valid for every $s > 0$, we get for n large enough the inequality

$$\sup_{p \in I_x} \left|\log\left(\frac{p}{k_x}\right) - \sum_{j=1}^{p-k_x} \frac{1}{p-j+1}\right| \leq \frac{1}{k_x} \leq \frac{t'}{2}. \tag{49}$$

Applying (48) and (49) then entails for n large enough

$$\sup_{p \in I_x} u_{np} \leq \sup_{p \in I_x} \mathbb{P}\left(\left|\sum_{j=1}^{p-k_x} \frac{\log(T_j) - 1}{p-j+1}\right| > \frac{t'}{2}\right). \tag{50}$$

Furthermore, Chebyshev's inequality and a comparison with an integral give

$$\begin{aligned} \sup_{p \in I_x} \mathbb{P}\left(\left|\sum_{j=1}^{p-k_x} \frac{\log(T_j) - 1}{p-j+1}\right| > \frac{t'}{2}\right) \quad &\leq \quad \frac{4}{t'^2} \sup_{p \in I_x} \sum_{j=1}^{p-k_x} \frac{1}{(p-j+1)^2} \\ &\leq \quad \frac{4}{t'^2 k_x} \to 0 \end{aligned} \tag{51}$$

as $n \to \infty$. Collecting (50) and (51) yields the first result. The second result is then a simple consequence of the first result and of the inequality

$$\sup_{p \in I_x} \mathbb{P}\left(T_{p-k_x,p} \leq t\right) \leq \sup_{p \in I_x} \mathbb{P}\left(\frac{k_x}{p} T_{p-k_x,p} - 1 < -\frac{1}{2}\right)$$

valid for n large enough. The third result is obtained by noting that since φ is regularly varying at infinity, writing

$$\frac{\varphi(T_{p-k_x,p})}{\varphi(p/k_x)} = \frac{1}{\varphi(p/k_x)}\varphi\left(\frac{p}{k_x}\left\{\frac{k_x}{p}T_{p-k_x,p}\right\}\right),$$

then Theorem 1.5.2 in [4] shows that there exists $t' > 0$ such that for n large enough

$$\left\{\left|\frac{k_x}{p}T_{p-k_x,p} - 1\right| \leq t'\right\} \subset \left\{\left|\frac{\varphi(T_{p-k_x,p})}{\varphi(p/k_x)} - 1\right| \leq t\right\}$$

for every $p \in I_x$; the first result then applies to yield

$$\sup_{p\in I_x} \mathbb{P}\left(\left|\frac{\varphi(T_{p-k_x,p})}{\varphi(p/k_x)} - 1\right| > t\right) \to 0 \quad \text{as} \quad n \to \infty.$$

Finally, since

$$\sup_{p\in I_x}\left|\frac{p}{n_x} - 1\right| \to 0 \quad \text{as } n \to \infty,$$

applying once again Theorem 1.5.2 in [4] gives

$$\sup_{p\in I_x}\left|\frac{\varphi(p/k_x)}{\varphi(n_x/k_x)} - 1\right| \to 0 \quad \text{as} \quad n \to \infty.$$

Using Lemma 4 completes the proof. □

Proof of Lemma 7. If T is a standard Pareto random variable, then $\log T$ is a standard exponential random variable. One can thus use the Cramér-Wold device and argue along the lines of the proof of Lemma 3.2.3 in [21]. □

Proof of Lemma 8. We start by proving the first statement. Pick δ, $t > 0$ and $\varepsilon \in (0, 1)$ such that

$$\frac{2\varepsilon}{t(1-(\gamma_-(x)+\varepsilon))} \leq \frac{\delta}{4}. \tag{52}$$

With the notation of (25), letting $B_{np} = \{T_{p-k_x,p} \leq t_0\}$, Lemma 6 shows that $\mathbb{P}(B_{np}) \to 0$ uniformly in $p \in I_x$ as $n \to \infty$. For every $p \in I_x$, on the complement B_{np}^c of B_{np}, one can apply (25) to write

$$\frac{\log U(T_{p-i+1,p}|x) - \log U(T_{p-k_x,p}|x)}{q_0(T_{p-k_x,p}|x)} \leq \frac{(T_i^\star(p))^{\gamma_-(x)} - 1}{\gamma_-(x)} + \varepsilon\,(T_i^\star(p))^{\gamma_-(x)+\varepsilon} \tag{53}$$

and

$$\frac{(T_i^\star(p))^{\gamma_-(x)} - 1}{\gamma_-(x)} - \varepsilon\,(T_i^\star(p))^{\gamma_-(x)+\varepsilon} \leq \frac{\log U(T_{p-i+1,p}|x) - \log U(T_{p-k_x,p}|x)}{q_0(T_{p-k_x,p}|x)} \tag{54}$$

where $T_i^\star(p) = T_{p-i+1,p}/T_{p-k_x,p} \geq 1$ for every $p \in I_x$ and $i = 1, \ldots, k_x$. Using (53) and (54), the probability of the event

$$C_{np} = \left\{ \left| \frac{\mathcal{M}_{np}^{(1)}(x, k_x, h)}{q_0(T_{p-k_x,p}|x)} - \frac{1}{1-\gamma_-(x)} \right| > t \right\}$$

is then bounded from above by $\mathbb{P}(B_{np}) + \mathbb{P}(C_{np}^{(1)}) + \mathbb{P}(C_{np}^{(2)}) \leq \delta/2 + \mathbb{P}(C_{np}^{(1)}) + \mathbb{P}(C_{np}^{(2)})$ uniformly in $p \in I_x$ for n large enough, where

$$\begin{aligned} C_{np}^{(1)} &= \left\{ \left| \frac{1}{k_x} \sum_{i=1}^{k_x} \frac{(T_i^\star(p))^{\gamma_-(x)} - 1}{\gamma_-(x)} - \frac{1}{1-\gamma_-(x)} \right| > \frac{t}{2} \right\} \\ \text{and} \quad C_{np}^{(2)} &= \left\{ \frac{1}{k_x} \sum_{i=1}^{k_x} (T_i^\star(p))^{\gamma_-(x)+\varepsilon} > \frac{t}{2\varepsilon} \right\}. \end{aligned}$$

Apply Lemma 7 to get for every $p \in I_x$

$$\begin{aligned} \mathbb{P}(C_{np}^{(1)}) &= \mathbb{P}\left(\left| \frac{1}{k_x} \sum_{i=1}^{k_x} \frac{T_i^{\gamma_-(x)} - 1}{\gamma_-(x)} - \frac{1}{1-\gamma_-(x)} \right| > \frac{t}{2} \right) \\ \text{and} \quad \mathbb{P}(C_{np}^{(2)}) &= \mathbb{P}\left(\frac{1}{k_x} \sum_{i=1}^{k_x} T_i^{\gamma_-(x)+\varepsilon} > \frac{t}{2\varepsilon} \right). \end{aligned}$$

Because

$$\mathbb{E}\left[\frac{T_i^{\gamma_-(x)} - 1}{\gamma_-(x)} \right] = \frac{1}{1-\gamma_-(x)},$$

Chebyshev's inequality leads to the inequality $\mathbb{P}(C_{np}^{(1)}) \leq \delta/4$ for n large enough, uniformly in $p \in I_x$. Furthermore, since $\varepsilon \in (0,1)$, using together (52) and Markov's inequality yields $\mathbb{P}(C_{np}^{(2)}) \leq \delta/4$ for every $p \in I_x$. Hence for n large enough the inequality

$$\sup_{p \in I_x} \mathbb{P}(C_{np}) \leq \delta.$$

In other words, it holds that for every $t > 0$

$$\sup_{p \in I_x} \mathbb{P}\left(\left| \frac{\mathcal{M}_{np}^{(1)}(x, k_x, h)}{q_0(T_{p-k_x,p}|x)} - \frac{1}{1-\gamma_-(x)} \right| > t \right) \to 0 \quad \text{as} \quad n \to \infty. \tag{55}$$

Recall that $q_0(\cdot|x)$ is regularly varying at infinity with index $\gamma_-(x)$ and apply Lemma 6 to get for every $t > 0$

$$\sup_{p \in I_x} \mathbb{P}\left(\left| \frac{q_0(T_{p-k_x,p}|x)}{q_0(p/k_x|x)} - 1 \right| > t \right) \to 0 \quad \text{as} \quad n \to \infty. \tag{56}$$

Finally, writing

$$\left|\frac{\mathcal{M}_{np}^{(1)}(x,k_x,h)}{q_0(p/k_x|x)}-\frac{1}{1-\gamma_-(x)}\right| \leq \frac{q_0(T_{p-k_x,p}|x)}{q_0(p/k_x|x)}\left|\frac{\mathcal{M}_{np}^{(1)}(x,k_x,h)}{q_0(T_{p-k_x,p}|x)}-\frac{1}{1-\gamma_-(x)}\right| + \frac{1}{1-\gamma_-(x)}\left|\frac{q_0(T_{p-k_x,p}|x)}{q_0(p/k_x|x)}-1\right|$$

and applying Lemma 4 together with (55) and (56) gives the first part of the result. To obtain the second part, square the inequalities (53) and (54) with $\varepsilon < 1/2$ small enough, use the equality

$$\mathbb{E}\left[\frac{T_i^{\gamma_-(x)}-1}{\gamma_-(x)}\right]^2 = \frac{2}{(1-\gamma_-(x))(1-2\gamma_-(x))}$$

and use the ideas developed for the proof of the first statement. □

Proof of Lemma 9. To prove the first statement, pick $y_0 \in \mathbb{R}$ such that $1/\alpha := \overline{F}(y_0) \in (0,1)$ and assume that for some $\delta > 0$, one has $\overline{F}(y_0+\delta) = \overline{F}(y_0)$. Then, since the function $\overline{F}$ is nonincreasing, it is constant equal to α on $[y_0, y_0+\delta]$. Thus

$$U(\alpha+\varepsilon) = \inf\{y \in \mathbb{R} \,|\, 1/\overline{F}(y) \geq \alpha+\varepsilon\} \geq y_0+\delta$$

for every $\varepsilon > 0$. Taking the limit $\varepsilon \downarrow 0$ yields $U(\alpha) \geq y_0+\delta$, which is a contradiction.

To show the second statement, assume that $\overline{F}$ is not continuous at $y_0 \in \mathbb{R}$: in other words, since $\overline{F}$ is right-continuous and nonincreasing,

$$\beta_- := \lim_{\substack{y\to y_0 \\ y<y_0}} \overline{F}(y) > \overline{F}(y_0) =: \beta_+.$$

It follows that for every $z \in (1/\beta_-, 1/\beta_+)$, one has $U(z) = y_0$, which is a contradiction. Finally, note that by the right-continuity of $\overline{F}$:

$$\forall z > 1,\ \eta := \frac{1}{\overline{F}(U(z))} - z \geq 0.$$

If one had $\eta > 0$, then it would hold that

$$U(z+\eta) = \inf\{y \in \mathbb{R} \,|\, 1/\overline{F}(y) \geq z+\eta\} \leq U(z)$$

which is a contradiction. □

Proof of Lemma 10. Write for every $y \in \mathbb{R}$

$$\overline{F}_h(y|x) = \frac{\mathbb{E}(\overline{F}(y|X)\mathbb{1}_{\{X\in B(x,h)\}})}{\mathbb{P}(X \in B(x,h))}.$$

The continuity assertion on $\overline{F}_h(\cdot|x)$ then follows from Lemma 9 and the dominated convergence theorem. Pick now $y \in \mathbb{R}$ such that $\overline{F}_h(y|x) \in (0,1)$ and $\delta > 0$. One has:

$$\overline{F}_h(y|x) - \overline{F}_h(y+\delta|x) = \frac{\mathbb{E}([\overline{F}(y|X) - \overline{F}(y+\delta|X)]\mathbb{1}_{\{X\in B(x,h)\}})}{\mathbb{P}(X \in B(x,h))}.$$

Assume that $\overline{F}_h(y|x) = \overline{F}_h(y+\delta|x)$. In this case, since for every $x' \in E$ the function $\overline{F}(\cdot|x')$ is nonincreasing, we get

$$[\overline{F}(y|X) - \overline{F}(y+\delta|X)]\mathbb{1}_{\{X\in B(x,h)\}} = 0 \text{ almost surely.} \tag{57}$$

Besides, since $\overline{F}_h(y|x) \in (0,1)$, there exist measurable sets A_0 and A_1 such that

$$\begin{aligned}
&\forall x' \in A_0 \cap B(x,h),\ \overline{F}(y|x') > 0 \text{ and } \mathbb{P}(X \in A_0 \cap B(x,h)) > 0,\\
&\forall x' \in A_1 \cap B(x,h),\ \overline{F}(y|x') < 1 \text{ and } \mathbb{P}(X \in A_1 \cap B(x,h)) > 0.
\end{aligned}$$

Since the ball $B(x,h)$ is a connected set in E because it is arc-connected, one may therefore apply the intermediate value theorem to the continuous map $x' \mapsto \overline{F}(y|x')$ to obtain that there exists $x' \in B(x,h)$ such that $\overline{F}(y|x') \in (0,1)$. Using once again the continuity of this map, we deduce that there exists $r > 0$, which we may choose in order to have $B(x',r) \subset B(x,h)$, such that

$$\forall x'' \in B(x',r),\ \overline{F}(y|x'') \in (0,1). \tag{58}$$

Because $\mathbb{P}(X \in B(x',r)) > 0$, using together (57) and (58) leads to the existence of some $x'' \in B(x,h)$ such that

$$\overline{F}(y|x'') = \overline{F}(y+\delta|x'') \text{ and } \overline{F}(y|x'') \in (0,1)$$

which, in view of Lemma 9, is a contradiction. Finally, note that together with the intermediate value theorem, the two properties of $\overline{F}_h(\cdot|x)$ we have shown here entail that for every $z \in (1,\infty)$, the inverse image of z under $1/\overline{F}_h$ consists of a unique point. Hence this point must be $U_h(z|x)$: consequently

$$\frac{1}{\overline{F}_h(U_h(z|x)|x)} = z$$

which completes the proof. □

Proof of Lemma 11. We only consider the case $\gamma(x) \neq \rho(x)$ and $\rho(x) < 0$ if $\gamma(x) > 0$, the proof being entirely similar in the case $\gamma(x) > \rho(x) = 0$. According to Lemma 3, the distribution of the random vector

$$\sqrt{k_x}\left(\frac{\mathfrak{M}^{(1)}_{np}(x,k_x,h)}{q_0(T_{p-k_x,p}|x)} - \frac{1}{1-\gamma_-(x)}, \frac{\mathfrak{M}^{(2)}_{np}(x,k_x,h)}{q_0^2(T_{p-k_x,p}|x)} - \frac{2}{(1-\gamma_-(x))(1-2\gamma_-(x))}\right)$$

is the distribution of the random vector

$$\begin{aligned}
&\sqrt{k_x}\left(\frac{\mathcal{M}^{(1)}_{np}(x,k_x,h)}{q_0(T_{p-k_x,p}|x)} - \frac{1}{1-\gamma_-(x)}, \frac{\mathcal{M}^{(2)}_{np}(x,k_x,h)}{q_0^2(T_{p-k_x,p}|x)} - \frac{2}{(1-\gamma_-(x))(1-2\gamma_-(x))}\right)\\
&+ \ (r^{(1)}_{np}, r^{(2)}_{np}),
\end{aligned} \tag{59}$$

where

$$\left|r_{np}^{(1)}\right| \leq 2\sqrt{k_x}\frac{\omega(T_{p-k_x,p}, T_{p,p}, x, h)}{q_0(T_{p-k_x,p}|x)},$$

$$\left|r_{np}^{(2)}\right| \leq 4\sqrt{k_x}\frac{\omega(T_{p-k_x,p}, T_{p,p}, x, h)\left[\omega(T_{p-k_x,p}, T_{p,p}, x, h) + \mathcal{M}_{np}^{(1)}(x, k_x, h)\right]}{q_0^2(T_{p-k_x,p}|x)}.$$

Recall that $q_0(\cdot|x)$ is equivalent to $q(\cdot|x)$ at infinity, which is itself regularly varying at infinity with index $\gamma_-(x)$. As a consequence, applying Lemma 6, we get for every $t > 0$ the convergence

$$\sup_{p\in I_x} \mathbb{P}\left(\left|\frac{q_0(T_{p-k_x,p}|x)}{q_0(p/k_x|x)} - 1\right| > t\right) \to 0 \text{ as } n \to \infty. \tag{60}$$

Besides, using (28) and (29) in the proof of Proposition 2, we get for every $t > 0$

$$\sup_{p\in I_x} \mathbb{P}\left(\sqrt{k_x}\frac{\omega(T_{p-k_x,p}, T_{p,p}, x, h)}{q_0(p/k_x|x)} > t\right) \leq \sup_{p\in I_x} \mathbb{P}\left(\sqrt{k_x}\frac{U(n_x/k_x|x)}{a(n_x/k_x|x)}\omega(T_{p-k_x,p}, T_{p,p}, x, h) > \frac{t}{2}\right) \tag{61}$$

for n large enough. Using condition (23), the right-hand side of the above inequality is bounded from above by

$$\sup_{p\in I_x} u_{np}^{(1,2)} \leq \sup_{p\in I_x} \mathbb{P}\left(\left\{T_{p-k_x,p} < \frac{n_x}{(1+\delta)k_x}\right\} \cup \left\{T_{p,p} > n_x^{1+\delta}\right\}\right)$$

for n sufficiently large; consequently (see (29))

$$\sup_{p\in I_x} \mathbb{P}\left(\sqrt{k_x}\frac{U(n_x/k_x|x)}{a(n_x/k_x|x)}\omega(T_{p-k_x,p}, T_{p,p}, x, h) > \frac{t}{2}\right) \to 0 \text{ as } n \to \infty.$$

Inequality (61) thus yields the convergence

$$\sup_{p\in I_x} \mathbb{P}\left(\sqrt{k_x}\frac{\omega(T_{p-k_x,p}, T_{p,p}, x, h)}{q_0(p/k_x|x)} > t\right) \to 0 \text{ as } n \to \infty. \tag{62}$$

Applying (27), (60), (62) and Lemma 4 we then obtain for every $t > 0$:

$$\sup_{p\in I_x} \mathbb{P}\left(\left|r_{np}^{(1)}\right| > t\right) \to 0 \text{ and } \sup_{p\in I_x} \mathbb{P}\left(\left|r_{np}^{(2)}\right| > t\right) \to 0 \text{ as } n \to \infty. \tag{63}$$

As in the proof of Lemma 8, letting $B_{np} = \{T_{p-k_x,p} \leq t_0\}$, Lemma 6 shows that $\mathbb{P}(B_{np}) \to 0$ uniformly in $p \in I_x$ as $n \to \infty$. Pick δ, $t > 0$ and choose $\varepsilon \in (0, 1)$ such that

$$\frac{12\varepsilon}{t(1 - (\gamma_-(x) + \rho'(x) + \varepsilon))}\left[1 + \left|\lambda(x)\left(\frac{\mathbb{1}_{\{\rho'(x)<0\}}}{\rho'(x)} + \mathbb{1}_{\{\rho'(x)=0\}}\right)\right|\right] \leq \frac{\delta}{4}. \tag{64}$$

For n large enough, one has $\mathbb{P}(B_{np}) \le \delta/4$ for every $p \in I_x$. Furthermore, on the complement B_{np}^c of B_{np}, one can apply (40) to write

$$\begin{aligned}
& \frac{(T_i^\star(p))^{\gamma_-(x)}-1}{\gamma_-(x)} + Q_0(T_{p-k_x,p}|x)\psi_{\gamma_-(x),\rho'(x)}(T_i^\star(p)) \\
- \quad & \varepsilon |Q_0(T_{p-k_x,p}|x)| (T_i^\star(p))^{\gamma_-(x)+\rho'(x)+\varepsilon} \\
\le \quad & \frac{\log U(T_{p-i+1,p}|x) - \log U(T_{p-k_x,p}|x)}{q_0(T_{p-k_x,p}|x)} \\
\le \quad & \frac{(T_i^\star(p))^{\gamma_-(x)}-1}{\gamma_-(x)} + Q_0(T_{p-k_x,p}|x)\psi_{\gamma_-(x),\rho'(x)}(T_i^\star(p)) \\
+ \quad & \varepsilon |Q_0(T_{p-k_x,p}|x)| (T_i^\star(p))^{\gamma_-(x)+\rho'(x)+\varepsilon}
\end{aligned} \tag{65}$$

where $T_i^\star(p) = T_{p-i+1,p}/T_{p-k_x,p} \ge 1$ for $p \in I_x$ and $i = 1, \ldots, k_x$. Note now that (39) implies that $|Q_0(\cdot|x)|$ is regularly varying at infinity, so that Lemma 6 entails for every $t' > 0$

$$\sup_{p \in I_x} \mathbb{P}\left(\left| \frac{Q_0(T_{p-k_x,p}|x)}{Q_0(n_x/k_x|x)} - 1 \right| > t' \right) \to 0 \quad \text{as} \quad n \to \infty. \tag{66}$$

Applying Lemma 7 and Chebyshev's inequality leads to

$$\sup_{p \in I_x} \mathbb{P}\left(\left| \frac{1}{k_x} \sum_{i=1}^{k_x} \psi_{\gamma_-(x),\rho'(x)}(T_i^\star(p)) - \mathbb{E}(\psi_{\gamma_-(x),\rho'(x)}(T)) \right| > \frac{t}{3} \right) \to 0 \tag{67}$$

as $n \to \infty$. Besides, (39) and the hypothesis $\sqrt{k_x}\, Q(n_x/k_x) \to \lambda(x)$ as $n \to \infty$ yield

$$\sqrt{k_x}\, Q_0(n_x/k_x) \to \lambda(x) \left(\frac{\mathbb{1}_{\{\rho'(x)<0\}}}{\rho'(x)} + \mathbb{1}_{\{\rho'(x)=0\}} \right) \quad \text{as} \quad n \to \infty \tag{68}$$

so that collecting (66), (67) and (68) and applying Lemma 4 shows that

$$\sup_{p \in I_x} \mathbb{P}\left(\left| \sqrt{k_x}\, Q_0(T_{p-k_x,p}|x) \frac{1}{k_x} \sum_{i=1}^{k_x} \psi_{\gamma_-(x),\rho'(x)}(T_i^\star(p)) - m^{(1)}(x) \right| > \frac{t}{3} \right) \to 0 \tag{69}$$

as $n \to \infty$. Meanwhile, letting

$$C_{np} = \left\{ \left| \frac{Q_0(T_{p-k_x,p}|x)}{Q_0(n_x/k_x|x)} - 1 \right| > \frac{1}{2} \right\}$$

then (66) entails that $\mathbb{P}(C_{np}) \to 0$ as $n \to \infty$, uniformly in $p \in I_x$; on C_{np}^c it holds that for n large enough

$$\begin{aligned}
\sqrt{k_x} |Q_0(T_{p-k_x,p}|x)| \quad & \le \quad \frac{3}{2} \sqrt{k_x} |Q_0(n_x/k_x|x)| \\
& \le \quad 2 \left[1 + \left| \lambda(x) \left(\frac{\mathbb{1}_{\{\rho'(x)<0\}}}{\rho'(x)} + \mathbb{1}_{\{\rho'(x)=0\}} \right) \right| \right].
\end{aligned} \tag{70}$$

Therefore, letting n be so large that $\mathbb{P}(C_{np}) \leq \delta/4$ for every $p \in I_x$, Lemma 7, Markov's inequality and (52) together imply that

$$\sup_{p\in I_x} \mathbb{P}\left(\varepsilon\sqrt{k_x}|Q_0(T_{p-k_x,p}|x)|\frac{1}{k_x}\sum_{i=1}^{k_x}(T_i^\star(p))^{\gamma_-(x)+\rho'(x)+\varepsilon} > \frac{t}{6}\right) \leq \frac{\delta}{2}. \quad (71)$$

Collecting (65), (69) and (71), we get for n large enough

$$\sup_{p\in I_x} \mathbb{P}\left(\left|\sqrt{k_x}\left[\frac{\mathcal{M}_{np}^{(1)}(x,k_x,h)}{q_0(T_{p-k_x,p}|x)} - \frac{1}{k_x}\sum_{i=1}^{k_x}\frac{(T_i^\star(p))^{\gamma_-(x)}-1}{\gamma_-(x)}\right] - m^{(1)}(x)\right| > t\right) \leq \delta.$$

Recalling (63) and applying Lemma 4, we obtain that

$$\begin{aligned} &\sqrt{k_x}\left(\frac{\mathcal{M}_{np}^{(1)}(x,k_x,h)}{q_0(T_{p-k_x,p}|x)} - \frac{1}{1-\gamma_-(x)}\right) + r_{np}^{(1)} \\ = \;& \sqrt{k_x}\left[\frac{1}{k_x}\sum_{i=1}^{k_x}\frac{(T_i^\star(p))^{\gamma_-(x)}-1}{\gamma_-(x)} - \frac{1}{1-\gamma_-(x)}\right] + m^{(1)}(x) + R_{np}^{(1)} \quad (72)\end{aligned}$$

with $R_{np}^{(1)}$ as in the statement of the result.

To obtain a similar result for $\mathcal{M}_{np}^{(2)}(x,k_x,h)$, we note that using (66), (68), Lemma 7 and Chebyshev's inequality, we have for every $t' > 0$

$$\begin{aligned}\sup_{p\in I_x} \mathbb{P}\Bigg(\Bigg|&2\sqrt{k_x}\,Q_0(T_{p-k_x,p}|x)\frac{1}{k_x}\sum_{i=1}^{k_x}\frac{(T_i^\star(p))^{\gamma_-(x)}-1}{\gamma_-(x)}\psi_{\gamma_-(x),\rho'(x)}(T_i^\star(p)) \\ &-m^{(2)}(x)\Bigg| > t'\Bigg) \to 0 \quad (73)\end{aligned}$$

as $n \to \infty$. Besides, picking $\delta > 0$, inequality (70), Lemma 7 and Markov's inequality yield for n large enough

$$\begin{aligned}&\sup_{p\in I_x} \mathbb{P}\left(\varepsilon\sqrt{k_x}|Q_0(T_{p-k_x,p}|x)|\frac{1}{k_x}\sum_{i=1}^{k_x}\frac{(T_i^\star(p))^{\gamma_-(x)}-1}{\gamma_-(x)}(T_i^\star(p))^{\gamma_-(x)+\rho'(x)+\varepsilon} > t'\right) \\ &\leq \frac{\delta}{4} \quad (74)\end{aligned}$$

if $\varepsilon > 0$ is chosen small enough. Using once again (66) and the convergence of $Q_0(\cdot|x)$ to 0, we get

$$\sup_{p\in I_x} \mathbb{P}\left(\sqrt{k_x}Q_0^2(T_{p-k_x,p}|x) > t'\right) \to 0 \text{ as } n \to \infty$$

which, together with Lemma 7 and Markov's inequality, entails

$$\sup_{p\in I_x} \mathbb{P}\left(\sqrt{k_x}Q_0^2(T_{p-k_x,p}|x)\frac{1}{k_x}\sum_{i=1}^{k_x}\psi^2_{\gamma_-(x),\rho'(x)}(T_i^\star(p)) > t'\right) \to 0 \quad (75)$$

and

$$\sup_{p\in I_x} \mathbb{P}\left(\sqrt{k_x} Q_0^2(T_{p-k_x,p}|x) \frac{1}{k_x} \sum_{i=1}^{k_x} \varepsilon^2 (T_i^\star(p))^{2(\gamma_-(x)+\rho'(x)+\varepsilon)} > t'\right) \to 0 \quad (76)$$

as $n \to \infty$. Square now the inequalities (65) and use (73), (74), (75) and (76) to obtain for n large enough

$$\sup_{p\in I_x} \mathbb{P}\left(\left|\sqrt{k_x}\left[\frac{\mathcal{M}_{np}^{(2)}(x,k_x,h)}{q_0^2(T_{p-k_x,p}|x)} - \frac{1}{k_x}\sum_{i=1}^{k_x}\left(\frac{(T_i^\star(p))^{\gamma_-(x)}-1}{\gamma_-(x)}\right)^2\right] - m^{(2)}(x)\right| > t\right) \le \delta.$$

Finally, recall (63) and apply Lemma 4 to get

$$\begin{aligned}
&\sqrt{k_x}\left(\frac{\mathcal{M}_{np}^{(2)}(x,k_x,h)}{q_0^2(T_{p-k_x,p}|x)} - \frac{2}{(1-\gamma_-(x))(1-2\gamma_-(x))}\right) + r_{np}^{(2)} \\
= &\sqrt{k_x}\left[\frac{1}{k_x}\sum_{i=1}^{k_x}\left(\frac{(T_i^\star(p))^{\gamma_-(x)}-1}{\gamma_-(x)}\right)^2 - \frac{2}{(1-\gamma_-(x))(1-2\gamma_-(x))}\right] \\
+ &m^{(2)}(x) + R_{np}^{(2)} \qquad (77)
\end{aligned}$$

with $R_{np}^{(2)}$ as in the statement of the result. Letting

$$\begin{aligned}
D_{np}^{(1)} &= \sqrt{k_x}\left[\frac{1}{k_x}\sum_{i=1}^{k_x}\frac{(T_i^\star(p))^{\gamma_-(x)}-1}{\gamma_-(x)} - \frac{1}{1-\gamma_-(x)}\right] + m^{(1)}(x), \\
D_{np}^{(2)} &= \sqrt{k_x}\left[\frac{1}{k_x}\sum_{i=1}^{k_x}\left(\frac{(T_i^\star(p))^{\gamma_-(x)}-1}{\gamma_-(x)}\right)^2 - \frac{2}{(1-\gamma_-(x))(1-2\gamma_-(x))}\right] + m^{(2)}(x)
\end{aligned}$$

and applying Lemma 7, it is obvious that for fixed n and every $p > k_x$, the random pair $(D_{np}^{(1)}, D_{np}^{(2)})$ has the same distribution as $(\mathcal{D}_n^{(1)}, \mathcal{D}_n^{(2)})$ where

$$\begin{aligned}
\mathcal{D}_n^{(1)} &= \sqrt{k_x}\left[\frac{1}{k_x}\sum_{i=1}^{k_x}\frac{T_i^{\gamma_-(x)}-1}{\gamma_-(x)} - \frac{1}{1-\gamma_-(x)}\right] + m^{(1)}(x), \\
\mathcal{D}_n^{(2)} &= \sqrt{k_x}\left[\frac{1}{k_x}\sum_{i=1}^{k_x}\left(\frac{T_i^{\gamma_-(x)}-1}{\gamma_-(x)}\right)^2 - \frac{2}{(1-\gamma_-(x))(1-2\gamma_-(x))}\right] + m^{(2)}(x).
\end{aligned}$$

The standard central limit theorem and some cumbersome computations show that the random vector $(\mathcal{D}_n^{(1)}, \mathcal{D}_n^{(2)})$ is asymptotically Gaussian with mean $(m^{(1)}(x), m^{(2)}(x))$ and covariance matrix $\mathcal{V}(\gamma(x))$, so that clearly

$$\begin{aligned}
&\sup_{p\in I_x}\left|\mathbb{E}[\exp[i(t_1 D_{np}^{(1)} + t_2 D_{np}^{(2)})] - \mathbb{E}[\exp[i(t_1 P_1 + t_2 P_2)]]\right| \\
= &\left|\mathbb{E}[\exp[i(t_1 \mathcal{D}_n^{(1)} + t_2 \mathcal{D}_n^{(2)})] - \mathbb{E}[\exp[i(t_1 P_1 + t_2 P_2)]]\right| \to 0 \ \text{ as } \ n \to \infty
\end{aligned}$$

for every t_1, $t_2 \in \mathbb{R}$. Now, according to (59), (72) and (77) the distribution of the random vector

$$\sqrt{k_x}\left(\frac{\mathfrak{M}_{np}^{(1)}(x,k_x,h)}{q_0(T_{p-k_x,p}|x)} - \frac{1}{1-\gamma_-(x)}, \frac{\mathfrak{M}_{np}^{(2)}(x,k_x,h)}{q_0^2(T_{p-k_x,p}|x)} - \frac{2}{(1-\gamma_-(x))(1-2\gamma_-(x))}\right)$$

is the distribution of $(D_{np}^{(1)} + R_{np}^{(1)}, D_{np}^{(2)} + R_{np}^{(2)})$, which completes the proof. □

Acknowledgements

The author acknowledges both the editor and an anonymous referee for their helpful comments which led to significant enhancements of this article.

References

[1] Beirlant, J., Boniphace, E., Dierckx, G. (2011). Generalized sum plots, *REVSTAT* **9**(2): 181–198. MR2851590

[2] Beirlant, J., Goegebeur, Y. (2004). Simultaneous tail index estimation, *REVSTAT* **2**(1): 15–39. MR2259344

[3] Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J. (2004). *Statistics of Extremes*, John Wiley and Sons. MR2108013

[4] Bingham, N.H., Goldie, C.M., Teugels, J.L. (1987). *Regular Variation*, Cambridge, U.K.: Cambridge University Press. MR0898871

[5] Chavez-Demoulin, V., Davison, A.C. (2005). Generalized additive modelling of sample extremes, *Journal of the Royal Statistical Society, Series C* **54**: 207–222. MR2134607

[6] Daouia, A., Florens, J.-P., Simar, L. (2010). Frontier estimation and extreme value theory, *Bernoulli* **16**(4): 1039–1063. MR2759168

[7] Daouia, A., Gardes, L., Girard, S., Lekina, A. (2011). Kernel estimators of extreme level curves, *Test* **20**(2): 311–333. MR2834049

[8] Daouia, A., Gardes, L., Girard, S. (2013). On kernel smoothing for extremal quantile regression, *Bernoulli*, to appear.

[9] Davison, A.C., Ramesh, N.I. (2000). Local likelihood smoothing of sample extremes, *Journal of the Royal Statistical Society, Series B* **62**: 191–208. MR1747404

[10] Davison, A.C., Smith, R.L. (1990). Models for exceedances over high thresholds, *Journal of the Royal Statistical Society, Series B* **52**: 393–442. MR1086795

[11] Dekkers, A.L.M., Einmahl, J.H.J., de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution, *Annals of Statistics* **17**(4): 1833–1855. MR1026315

[12] Ferrez, J., Davison, A.C., Rebetez, M. (2011). Extreme temperature analysis under forest cover compared to an open field, *Agricultural and Forest Meteorology* **151**: 992–1001.

[13] Fisher, R.A., 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.

[14] Gardes, L., Girard, S. (2008). A moving window approach for nonparametric estimation of the conditional tail index, *Journal of Multivariate Analysis* **99**: 2368–2388. MR2463396

[15] 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. MR2643556

[16] Gardes, L., Girard, S. (2012). Functional kernel estimators of large conditional quantiles, *Electronic Journal of Statistics* **6**: 1715–1744. MR2988462

[17] Gardes, L., Girard, S., Lekina, A. (2010). Functional nonparametric estimation of conditional extreme quantiles, *Journal of Multivariate Analysis* **101**(2): 419-433. MR2564351

[18] Gardes, L., Stupfler, G. (2013). Estimation of the conditional tail-index using a smoothed local Hill estimator, *Extremes*, to appear.

[19] Gnedenko, B.V. (1943). Sur la distribution limite du terme maximum d'une série aléatoire, *Annals of Mathematics* **44**: 423–453. MR0008655

[20] Goegebeur, Y., Guillou, A., Schorgen, A. (2013). Nonparametric regression estimation of conditional tails – the random covariate case, *Statistics*, to appear. MR3024038

[21] de Haan, L., Ferreira, A. (2006). *Extreme Value Theory: An Introduction*, New York: Springer. MR2234156

[22] Hall, P., Tajvidi, N. (2000). Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data, *Statistical Science* **15**: 153–167. MR1788730

[23] Härdle, W., Marron, J.S. (1985). Optimal bandwidth selection in nonparametric regression function estimation, *Annals of Statistics* **13**(4), 1465–1481. MR0811503

[24] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution, *Annals of Statistics* **3**: 1163–1174. MR0378204

[25] Pickands, J. (1975). Statistical inference using extreme order statistics, *Annals of Statistics* **3**: 119–131. MR0423667

[26] Pisarenko, V.F., Sornette, D. (2003). Characterization of the frequency of extreme earthquake events by the generalized Pareto distribution, *Pure and Applied Geophysics* **160**: 2343–2364.

[27] 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. MR1041763

[28] Wang, H., Tsai, C.L. (2009). Tail index regression, *Journal of the American Statistical Association* **104**(487): 1233–1240. MR2750246