FUNCTIONAL NONPARAMETRIC ESTIMATION OF CONDITIONAL EXTREME QUANTILES

Laurent Gardes, Stéphane Girard, Alexandre Lekina

FUNCTIONAL NONPARAMETRIC ESTIMATION OF CONDITIONAL EXTREME QUANTILES

Laurent GARDES[1], Stéphane GIRARD and Alexandre LEKINA

Team Mistis, INRIA Rhône-Alpes and LJK,
655, avenue de l'Europe, Montbonnot
38334 Saint-Ismier cedex, France.
Laurent.Gardes@inrialpes.fr

Abstract – We address the estimation of quantiles from heavy-tailed distributions when functional covariate information is available and in the case where the order of the quantile converges to one as the sample size increases. Such "extreme" quantiles can be located in the range of the data or near and even beyond the boundary of the sample, depending on the convergence rate of their order to one. Nonparametric estimators of these functional extreme quantiles are introduced, their asymptotic distributions are established and an illustration on a real data set is presented.

Keywords – Conditional quantile, extreme-values, nonparametric estimation, functional data.

AMS Subject classifications – 62G32, 62G05, 62E20.

1 Introduction

An important literature is dedicated to the estimation of extreme quantiles, *i.e.* quantiles of order $1 - \alpha$ with α tending to zero. The most popular estimator was proposed by Weissman [27], in the context of heavy-tailed distributions, and adapted to Weibull-tail distributions in [9, 18]. We also refer to [10] for the general case.
In a lot of applications, some covariate information is recorded simultaneously with the quantity of interest. For instance, in climatology one may be interested in the estimation of return periods associated to extreme rainfall as a function of the geographical location. The extreme quantile thus depends on the covariate and is referred in the sequel to as the conditional extreme quantile. Parametric models for conditional extremes are proposed in [8, 26] whereas semi-parametric methods are considered in [1, 21]. Fully

[1]Corresponding author

non-parametric estimators have been first introduced in [7], where a local polynomial modelling of the extreme observations is used. Similarly, spline estimators are fitted in [6] through a penalized maximum likelihood method. In both cases, the authors focus on univariate covariates and on the finite sample properties of the estimators. These results are extended in [2] where local polynomials estimators are proposed for multivariate covariates and where their asymptotic properties are established.

Besides, covariates may be curves in many situations coming from applied sciences such as chemometrics (see Section 5 for an illustration) or astrophysics [3]. However, the estimation of conditional extreme quantiles with functional covariates has not been addressed yet. Two statistical fields are involved in this study. In the one hand, nonparametric smoothing techniques adapted to functional data are required in order to deal with the covariate. We refer to [5, 16, 23, 24] for overviews on this literature. We propose here to select the observations to be used in the conditional quantile estimator by a moving window approach. In the second hand, once this selection is achieved, extreme-value methods are used to estimate the conditional quantile, see [12] for a comprehensive treatment of extreme-value methodology in various frameworks.

Whereas no parametric assumption is made on the functional covariate, we assume that the conditional distribution is heavy-tailed. This semi-parametric assumption amounts to supposing that the conditional survival function decreases at a polynomial rate. To estimate the conditional quantile, we focus on three different situations. In the first one, the convergence of α to zero is slow enough so that the quantile is located in the range of the data. In the second situation, the quantile is located near the boundary of the sample. Finally, in the third situation, the convergence of α to zero is sufficiently fast so that the quantile may be beyond the boundary of the sample. This situation is clearly the most difficult one since an extrapolation outside the range of the sample is needed to achieve the estimation.

Nonparametric estimators are defined in Section 2 for each situation. Their asymptotic distributions are derived in Section 3. Some examples are provided in Section 4 and an illustration on chemometric data is given in Section 5. Proofs are postponed to Section 6.

2 Estimators of conditional extreme quantiles

Let E be a (finite or infinite dimensional) metric space associated to a metric d. Let us denote by $F(.,x)$ the conditional cumulative distribution function of a real random variable Y given $x \in E$ and by $q(\alpha,x)$ the associated conditional quantile of order $1-\alpha$ defined by

$$F(q(\alpha,x),x) = 1-\alpha,$$

for all $x \in E$ and $\alpha \in (0,1)$. In this paper, we focus on the case where, for all $x \in E$, $F(.,x)$ is the cumulative distribution function of a heavy-tailed distribution. In such a situation, the conditional quantile $q(.,x)$ satisfies, for all $\lambda > 0$,

$$\lim_{\alpha \to 0} \frac{q(\lambda\alpha, x)}{q(\alpha, x)} = \lambda^{-\gamma(x)}, \tag{1}$$

where $\gamma(.)$ is an unknown positive function of the covariate x referred to as the conditional tail index. Loosely speaking, the conditional quantile $q(.,x)$ decreases towards 0 at a polynomial rate driven by $\gamma(x)$. The conditional quantile is said to be regularly varying at 0 with index $-\gamma(x)$, and this property characterizes heavy-tailed distributions. We refer to [4] for a general account on regular variation theory and to paragraph 4.2 for some examples of distributions satisfying (1).

Given a sample $(Y_1, x_1), \ldots, (Y_n, x_n)$ of independent observations, our aim is to build point-wise estimators of conditional quantiles. More precisely, for a given $t \in E$, we want to estimate $q(\alpha, t)$, focusing on the case where the design points $x_1, \ldots, x_n$ are non random. To this end, for all $r > 0$, let us denote by $B(t, r)$ the ball centered at point t and with radius r defined by

$$B(t,r) = \{x \in E,\ d(x,t) \leq r\}$$

and let $h_{n,t} = h_t$ be a positive sequence tending to zero as n goes to infinity. The proposed estimator uses a moving window approach since it is based on the response variables $Y_i's$ for which the associated covariates $x_i's$ belong to the ball $B(t, h_t)$. The proportion of such design points is thus defined by

$$\varphi(h_t) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}\{x_i \in B(t, h_t)\}$$

and plays an important role in this study. It describes how the design points concentrate in the neighborhood of t when h_t goes to zero, similarly to the small ball probability does, see for instance the monograph on functional data analysis [16]. Thus, the nonrandom number of observations in the slice $S_t = (0, \infty) \times B(t, h_t)$ is given by $m_{n,t} = m_t = n\varphi(h_t)$. Let $\{Z_i(t),\ i = 1, \ldots, m_t\}$ be the response variables $Y_i's$ for which the associated covariates $x_i's$ belong to the ball $B(t, h_t)$ and let $Z_{1,m_t}(t) \leq \ldots \leq Z_{m_t,m_t}(t)$ be the corresponding order statistics.

In this paper, we focus on the estimation of conditional "extreme" quantile of order $1 - \alpha_{m_t}$. Here, the word "extreme" means that α_{m_t} tends to zero as n goes to infinity, making kernel based estimators [14] non adapted. In the sequel, three situations are considered:

(S.1) $\alpha_{m_t} \to 0$ and $m_t \alpha_{m_t} \to \infty$,

(S.2) $\alpha_{m_t} \to 0$, $m_t \alpha_{m_t} \to c \in [1, \infty)$ and $\lfloor m_t \alpha_{m_t} \rfloor \to \lfloor c \rfloor$.

(S.3) $\alpha_{m_t} \to 0$ and $m_t \alpha_{m_t} \to c \in [0, 1)$,

where $\lfloor x \rfloor$ denotes the largest integer smaller than x. Let us highlight that, in the unconditional case, situations **(S.1)** and **(S.3)** with $c \neq 0$ have already been examined by Dekkers and de Haan [10], the extreme case $c = 0$ being considered in [20], Theorem 5.1. A summary of their results can be found in [12], Theorem 6.4.14 and Theorem 6.4.15. In situation **(S.1)**, α_{m_t} goes to 0 slower than $1/m_t$ and the point-wise estimation of the conditional extreme quantile relies on an interpolation inside the sample, since, from Proposition 2 below, $q(\alpha_{m_t}, t)$ is eventually almost surely smaller that the maximal observation $Z_{m_t,m_t}(t)$ in the slice S_t. In such a situation, we propose to estimate $q(\alpha_{m_t}, t)$ by:

$$\hat{q}_1(\alpha_{m_t}, t) = Z_{m_t - \lfloor m_t \alpha_{m_t} \rfloor + 1, m_t}(t). \tag{2}$$

In the intermediate situation **(S.2)**, estimator (2) can still be used, since for n large enough, $\lfloor m_t \alpha_{m_t} \rfloor = \lfloor c \rfloor > 0$ and thus the estimation relies on a conditional extreme value of the sample. Let us note that, if c is not an integer, then $m_t \alpha_{m_t} \to c$ implies $\lfloor m_t \alpha_{m_t} \rfloor \to \lfloor c \rfloor$. Otherwise, if c is an integer, then condition $\lfloor m_t \alpha_{m_t} \rfloor \to \lfloor c \rfloor$ is necessary to prevent the sequence $\lfloor m_t \alpha_{m_t} \rfloor$ from having two adherence values and $\hat{q}_1(\alpha_{m_t}, t)$ from oscillating. In situation **(S.3)**, α_{m_t} goes to 0 at the same speed or faster than $1/m_t$ and the conditional extreme quantile is eventually larger than $Z_{m_t,m_t}(t)$ with positive probability $\mathrm{e}^{-c} \geq \mathrm{e}^{-1}$. Thus, its estimation is more difficult since it requires an estimation outside the sample. We propose in this case to estimate $q(\alpha_{m_t}, t)$ by:

$$\begin{aligned} \hat{q}_2(\alpha_{m_t}, t) &= \hat{q}_1(\beta_{m_t}, t) \, (\beta_{m_t}/\alpha_{m_t})^{\hat{\gamma}_n(t)} \\ &= Z_{m_t - \lfloor m_t \beta_{m_t} \rfloor + 1, m_t}(t) \, (\beta_{m_t}/\alpha_{m_t})^{\hat{\gamma}_n(t)}, \end{aligned} \tag{3}$$

where β_{m_t} satisfies **(S.1)** and $\hat{\gamma}_n(t)$ is a point-wise estimator of the conditional tail index $\gamma(t)$. Such estimators have been proposed both in the finite dimensional setting [2] and in the general case [19], see also paragraph 4.1 for some examples. Note that (3) is an adaptation of Weissman estimator [27] in the case where covariate information is available. The extrapolation is achieved thanks to the multiplicative term $(\beta_{m_t}/\alpha_{m_t})^{\hat{\gamma}_n(t)}$ which magnitude is driven by the estimated tail index $\hat{\gamma}_n(t)$. As expected, the extrapolation is all the more important as the tail is heavy.

3 Main results

We first give some notations and conditions useful to establish the asymptotic distributions of our estimators. In the sequel, we fix $t \in E$ and we assume:

(A) The conditional quantile function

$$\alpha \in (0,1) \mapsto q(\alpha,t) \in (0,+\infty)$$

is differentiable, the function defined by

$$\alpha \in (0,1) \mapsto \Delta(\alpha,t) = \gamma(t) + \alpha \frac{\partial \log q}{\partial \alpha}(\alpha,t) \in (0,+\infty)$$

is continuous and such that $\lim_{\alpha \to 0} \Delta(\alpha,t) = 0$.

Assumption **(A)** controls the behavior of the log-quantile function with respect to its first variable. It is a sufficient condition to obtain the heavy-tail property (1), see for instance [4], Chapter 1. For all $a \in (0,1)$, let us introduce

$$\bar{\Delta}(a,t) = \sup_{\alpha \in (0,a)} |\Delta(\alpha,t)|.$$

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

$$\omega_n(a) = \sup \left\{ \left| \log \frac{q(\alpha,x)}{q(\alpha,x')} \right| \ , \ \alpha \in (a,1-a) \ , \ (x,x') \in B(t,h_t)^2 \right\}.$$

Finally, let $k_t \in \{1,\dots,m_t\}$ and $J_{k_t} = \{1,\dots,k_t\}$. Our first result establishes a representation in distribution of the largest random variables of the sample $Z_i(t)$, $i \in \{1,\dots,m_t\}$.

Proposition 1 *If $k_t/m_t \to 0$ and $k_t^2 \omega_n(m_t^{-(1+\delta)}) \to 0$ for some $\delta > 0$, then, there exists an event $\mathcal{A}_n$ with $\mathbb{P}(\mathcal{A}_n) \to 1$ as $n \to \infty$ such that*

$$\{(\log Z_{m_t-i+1,m_t}, i \in J_{k_t}) \,|\mathcal{A}_n\} \stackrel{d}{=} \{(\log q(V_{i,m_t},T_i), i \in J_{k_t}) \,|\mathcal{A}_n\},$$

where $V_{1,m_t} \le \ldots \le V_{m_t,m_t}$ are the order statistics associated to the sample $\{V_1,\dots,V_{m_t}\}$ of independent uniform variables and $\{T_1,\dots,T_{k_t}\}$ are random variables in the ball $B(t,h_t)$.

Note that this result is implicitly used in [19], proof of Theorem 1. We also refer to [13], Theorem 3.5.2, for the approximation of the nearest neighbors distribution using the Hellinger distance and to [17] for the study of their asymptotic distribution. Here, condition $k_t^2 \omega_n(m_t^{-(1+\delta)}) \to 0$ shows that, the smoother the quantile function is on the slice S_t, *i.e.* the smaller its oscillation is, the easier the control of the uppest observations is, *i.e* the larger k_t can be.

The next proposition is dedicated to the study of the position of the conditional extreme quantile $q(\alpha,t)$ with respect to the largest observation in the slice S_t.

Proposition 2 *If $\omega_n(m_t^{-(1+\delta)}) \to 0$ for some $\delta > 0$, then*

- *under* **(S.1)**, $\mathbb{P}(Z_{m_t,m_t} < q(\alpha_{m_t}, t)) \to 0$,
- *under* **(S.2)** *or* **(S.3)**, $\mathbb{P}(Z_{m_t,m_t} < q(\alpha_{m_t}, t)) \to \mathrm{e}^{-c}$.

Let us first focus on situation **(S.1)** where the estimation of the conditional extreme quantile is addressed using $\hat{q}_1(\alpha_{m_t}, t)$, an upper order statistic chosen in the considered slice.

Theorem 1 *Let (α_{m_t}) be a sequence satisfying* **(S.1)**.
If $(m_t\alpha_{m_t})^2\omega_n(m_t^{-(1+\delta)}) \to 0$ for some $\delta > 0$ then,

$$(m_t\alpha_{m_t})^{1/2}\left(\frac{\hat{q}_1(\alpha_{m_t}, t)}{q(\alpha_{m_t}, t)} - 1\right) \xrightarrow{d} \mathcal{N}(0, \gamma^2(t)).$$

It appears that the estimator is asymptotically Gaussian, with asymptotic variance proportional to $\gamma^2(t)/(m_t\alpha_{m_t})$. Thus, the heavier is the tail, the larger is $\gamma(t)$, and the larger is the variance. Besides, the asymptotic variance being inversely proportional to α_{m_t}, the estimation remains more stable when the extreme quantile is far from the boundary of the sample. Considering now situation **(S.2)**, an asymptotically Gaussian behavior cannot be expected since, in this case, the estimator is based on the $\lfloor c \rfloor$th uppest order statistic in the considered slice.

Theorem 2 *Let (α_{m_t}) be a sequence satisfying* **(S.2)**.
If $\omega_n(m_t^{-(1+\delta)}) \to 0$ for some $\delta > 0$ then,

$$\left(\frac{\hat{q}_1(\alpha_{m_t}, t)}{q(\alpha_{m_t}, t)} - 1\right) \xrightarrow{d} \mathcal{E}(c, \gamma(t)),$$

where $\mathcal{E}(c, \gamma(t))$ is a non-degenerated distribution.

The asymptotic distribution $\mathcal{E}(c, \gamma(t))$ could be explicitly deduced from the proof of the result. It is omitted here for the sake of simplicity. Situation **(S.3)** is more complex since the asymptotic distribution of $\hat{q}_2$ may depend both on the behavior of $\hat{q}_1$ and $\hat{\gamma}_n$. In the next theorem, two cases are investigated. In situation (i), the asymptotic distribution of $\hat{q}_2$ is driven by $\hat{q}_1$. At the opposite, in situation (ii), $\hat{q}_2$ inherits its asymptotic distribution from $\hat{\gamma}_n$.

Theorem 3 *Let (β_{m_t}) be a sequence satisfying* **(S.1)** *and let (α_{m_t}) be a sequence eventually smaller than (β_{m_t}). Define $\zeta_{m_t} = (m_t\beta_{m_t})^{1/2}\log(\beta_{m_t}/\alpha_{m_t})$. If $(m_t\beta_{m_t})^2\omega_n(m_t^{-(1+\delta)}) \to 0$ for some $\delta > 0$ and there exists a positive sequence $\upsilon_n(t)$ and a distribution $\mathcal{D}$ such that*

$$\upsilon_n(t)(\hat{\gamma}_n(t) - \gamma(t)) \xrightarrow{d} \mathcal{D}, \tag{4}$$

then, two situations arise:

(i) Under the additional condition

$$\zeta_{m_t} \max\left\{\upsilon_n^{-1}(t), \bar{\Delta}(\beta_{m_t}, t)\right\} \to 0, \tag{5}$$

we have

$$(m_t \beta_{m_t})^{1/2} \left(\frac{\hat{q}_2(\alpha_{m_t}, t)}{q(\alpha_{m_t}, t)} - 1\right) \xrightarrow{d} \mathcal{N}(0, \gamma^2(t)). \tag{6}$$

(ii) Otherwise, under the additional condition

$$\upsilon_n(t) \max\left\{\zeta_{m_t}^{-1}, \bar{\Delta}(\beta_{m_t}, t)\right\} \to 0, \tag{7}$$

we have

$$\frac{\upsilon_n(t)}{\log\left(\beta_{m_t}/\alpha_{m_t}\right)} \left(\frac{\hat{q}_2(\alpha_{m_t}, t)}{q(\alpha_{m_t}, t)} - 1\right) \xrightarrow{d} \mathcal{D}. \tag{8}$$

Note that, even though the main interest of this result is to tackle the case where (α_{m_t}) is a sequence satisfying **(S.3)**, it can also be applied in the more general situation where α_{m_t} is eventually smaller than β_{m_t}. For instance, it appears that, in situation **(S.2)**, $\hat{q}_2(\alpha_{m_t}, t)$ is a consistent estimator of $q(\alpha_{m_t}, t)$ in the sense that the ratio converges to one in probability whereas, in view of Theorem 2, $\hat{q}_1(\alpha_{m_t}, t)$ is not consistent. Some applications of Theorem 3 are provided in the next section.

4 Examples

In paragraph 4.1, the above theorem is illustrated with a particular family of conditional tail index estimators. The corresponding assumptions are simplified in paragraph 4.2 for some classical heavy-tailed distributions.

4.1 Some conditional tail-index estimators

In [19], a family of conditional tail index estimators is introduced. They are based on a weighted sum of the log-spacings between the k_t largest order statistics $Z_{m_t-k_t+1,m_t}, \ldots, Z_{m_t,m_t}$. The family is defined by

$$\hat{\gamma}_n(t, W) = \sum_{i=1}^{k_t} i \log\left(\frac{Z_{m_t-i+1,m_t}(t)}{Z_{m_t-i,m_t}(t)}\right) W\left(i/k_t, t\right) \Bigg/ \sum_{i=1}^{k_t} W\left(i/k_t, t\right), \tag{9}$$

where $W(., t)$ is a weight function defined on $(0, 1)$ and integrating to one. Basing on (9) and considering $\beta_{m,t} = k_t/m_t$, the conditional extreme quantile estimator (3) can be written as

$$\hat{q}_2(\alpha_{m_t}, t, W) = Z_{m_t-k_t+1,m_t}(t) \left(\frac{k_t}{m_t \alpha_{m_t}}\right)^{\hat{\gamma}_n(t,W)}.$$

From [19], Theorem 2, under some conditions on the weight function, $\hat{\gamma}_n(t, W)$ is asymptotically Gaussian:

$$k_t^{1/2}(\hat{\gamma}_n(t, W) - \gamma(t)) \xrightarrow{d} \mathcal{N}(0, \gamma^2(t)\mathcal{AV}(t, W)),$$

where $\mathcal{AV}(t, W) = \int_0^1 W^2(s, t)ds$. Letting $\upsilon_n(t) = k_t^{1/2}$, we obtain

$$\zeta_{m_t}\upsilon_n^{-1}(t) = \log\left(\frac{k_t}{m_t\alpha_{m_t}}\right) \to \infty,$$

in situation **(S.2)** or **(S.3)**, which means that condition (5) cannot be satisfied. Thus, only situation (ii) of Theorem 3 may arise leading to the following corollary.

Corollary 1 *Suppose the assumptions of [19], Theorem 2 hold. Let $k_t \to \infty$ such that*

$$k_t^{1/2}\bar{\Delta}(k_t/m_t, t) \to 0 \quad \textit{and} \tag{10}$$

$$k_t^2\omega_n(m_t^{-(1+\delta)}) \to 0 \quad \textit{for some } \delta > 0. \tag{11}$$

Let (α_{m_t}) be a sequence satisfying **(S.2)** *or* **(S.3)**. *Then,*

$$\frac{k_t^{1/2}}{\log(k_t/(m_t\alpha_{m_t}))}\left(\frac{\hat{q}_2(\alpha_{m_t}, t, W)}{q(\alpha_{m_t}, t)} - 1\right) \xrightarrow{d} \mathcal{N}(0, \gamma^2(t)\mathcal{AV}(t, W)).$$

As an example, one can use constant weights $W^{\mathrm{H}}(s, t) = 1$ to obtain the so-called conditional Hill estimator with $\mathcal{AV}(t, W^{\mathrm{H}}) = 1$ or logarithmic weights $W^{\mathrm{Z}}(s, t) = -\log(s)$ leading to the conditional Zipf estimator with $\mathcal{AV}(t, W^{\mathrm{Z}}) = 2$. We refer to [19], Section 4, for further details.

4.2 Illustration on some heavy-tailed distributions

Standard Pareto distribution is the simplest example of heavy-tailed distribution. Its conditional quantile of order $1 - \alpha$ decreases as a power function of α since, in this case, $q(\alpha, t) = \alpha^{-\gamma(t)}$. Therefore $\Delta(\alpha, t) = 0$ for all $\alpha \in (0, 1)$ and condition (10) of Corollary 1 vanishes. Another example is Fréchet distribution for which

$$q(\alpha, t) = \alpha^{-\gamma(t)}\left\{\frac{1}{\alpha}\log\left(\frac{1}{1-\alpha}\right)\right\}^{-\gamma(t)}.$$

Here, the conditional quantile approximatively decreases as a power function of α since, in this case, $q(\alpha, t) \sim \alpha^{-\gamma(t)}$, the quality of this approximation being controlled by

$$\Delta(\alpha, t) = -\frac{\gamma(t)}{2}\alpha(1 + O(\alpha)) \text{ as } \alpha \to 0.$$

A similar example is given by Burr distributions for which

$$q(\alpha,t) = \alpha^{-\gamma(t)}\left(1-\alpha^{-\rho(t)}\right)^{-\gamma(t)/\rho(t)}$$

and

$$\Delta(\alpha,t) = -\gamma(t)\alpha^{-\rho(t)}(1+O(\alpha^{-\rho(t)})),$$

with $\rho(t) < 0$. These results are collected in Table 1. In both Fréchet and Burr cases, $\Delta(\alpha,t)$ is asymptotically proportional to $\alpha^{-\rho(t)}$ as $\alpha \to 0$ with the convention $\rho(t) = -1$ for the Fréchet distribution. Note that $\rho(t)$ is known as the second-order parameter in the extreme-value theory. It drives the quality of the approximation of the conditional quantile $q(\alpha,t)$ by the power function $\alpha^{-\gamma(t)}$. Furthermore, it is easily seen, that for these two distributions, the function $|\Delta(.,t)|$ is increasing. Thus, condition (10) of Corollary 1 can be simplified as $m_t^{2\rho(t)}k_t^{1-2\rho(t)} \to 0$ which shows that, the smaller $\rho(t)$ is, the larger k_t can be. Finally, if γ and ρ are Lipschitzian, *i.e.* if there exist constants $c_\gamma > 0$ and $c_\rho > 0$ such that

$$|\gamma(x)-\gamma(x')| \leq c_\gamma d(x,x') \text{ and } |\rho(x)-\rho(x')| \leq c_\rho d(x,x')$$

for all $(x,x') \in B(t,h_t)^2$, then the oscillation can be bounded by $\omega_n(a) = O(h_t \log(1/a))$ as $a \to 0$ and thus condition (11) of Corollary 1 can be simplified as $k_t^2 h_t \log m_t \to 0$.

5 Illustration on real data

In this section, we propose to illustrate the behaviour of our conditional extreme quantiles estimators on functional chemometric data. The dataset can be found at `http://lib.stat.cmu.edu/datasets/tecator`. It consists of $n = 215$ samples of finely chopped meat (see for example [15] for more details). For each unit i taken among this sample, we observe one spectrometric curve χ_i discretized at 100 wavelengths $\lambda_1,\dots,\lambda_{100}$. The covariate x_i is thus defined by $x_i = (x_{i,1},\dots,x_{i,100})^t$ with $x_{i,j} = \chi_i(\lambda_j)$ for all $j = 1,\dots,100$. Each variable $x_{i,j}$ is the $-\log_{10}$ of the transmittance, that is the absorbance, recorded by the Tecator Infratec Food and Feed Analyzer spectrometer. Clearly, the covariate x_i is in fact a discretized curve but, as mentioned in [22], the fineness of the grid spanning the discretization allows us to consider each subject as a continuous curve. Hence, the covariate can be considered as belonging to an infinite dimensional space E. For each spectrometric curve χ_i, the fat content $\tilde{Y}_i \in [0,100]$ (in percentage) is given. Since these values are upper bounded, they cannot satisfy model (1) and we propose to use as variable of interest $Y_i = 100/\tilde{Y}_i \in [1,\infty)$, $i = 1,\dots,n$. The adequation of our model to the new sample (Y_i, x_i), $i = 1,\dots,n$ will be graphically checked below.

In the following, the semi-metric distance based on the second derivative is adopted, as advised in [16], Chapter 9:

$$d^2(\chi_i, \chi_j) = \int \left(\chi_i^{(2)}(t) - \chi_j^{(2)}(t)\right)^2 dt,$$

where $\chi^{(2)}$ denotes the second derivative of χ. To compute this semi-metric, one can use an approximation of the functions χ_i and χ_j based on B-splines as proposed in [16], Chapter 3. Here, we limit ourselves to a discretized version $\tilde{d}$ of d:

$$\tilde{d}^2(x_i, x_j) = \sum_{l=2}^{99} \{(x_{i,l+1} - x_{j,l+1}) + (x_{i,l-1} - x_{j,l-1}) - 2(x_{i,l} - x_{j,l})\}^2 .$$

We propose to estimate the conditional extreme quantiles in situation **(S.3)** in a given direction of the space E. More precisely, we focus on the segment $[\chi_{i_0}, \chi_{i_1}]$ where χ_{i_0} and χ_{i_1} denote the most different curves in the sample, *i.e.*

$$(i_0, i_1) = \arg\max_{1 \le i < j \le n} \tilde{d}(x_i, x_j).$$

The selected curves χ_{i_0} and χ_{i_1} are plotted in Figure 1. The conditional extreme quantile to estimate is $q(\alpha, t(r))$ where $t(r) = r\chi_{i_1} + (1-r)\chi_{i_0}$ for $r \in [0,1]$. To this end, the estimator $\hat{q}_2(\alpha, t(r), W^{\mathrm{Z}})$ defined in paragraph 4.1 is considered. Parameters $h_{t(r)}$ and $k_{t(r)}$ are selected thanks to the heuristics proposed in [19] which consists in minimizing the distance between two different estimators of the conditional extreme quantile:

$$(\hat{h}_{t(r)}, \hat{k}_{t(r)}) = \arg\min_{h,k} |\hat{q}_2(\alpha, t(r), W^{\mathrm{H}}) - \hat{q}_2(\alpha, t(r), W^{\mathrm{Z}})|.$$

The estimated quantiles $\hat{q}_2(1/300, t(r), W^{\mathrm{Z}})$ and $\hat{q}_2(1/500, t(r), W^{\mathrm{Z}})$ are plotted as functions of r in Figure 2. As a comparison, the maximal observation $Z_{m_{t(r)}, m_{t(r)}}$ in the ball $B(t(r), \hat{h}_{t(r)})$ is also represented as a function of r. It appears that, for most values of r, $\hat{q}_2(1/300, t(r), W^{\mathrm{Z}})$ is larger than the maximum observation. Unsurprisingly, we can also observe that $\hat{q}_2(1/500, t(r), W^{\mathrm{Z}}) > \hat{q}_2(1/300, t(r), W^{\mathrm{Z}})$ for all $r \in [0,1]$ since this property is ensured by the definition of the estimator itself (3). The globally decreasing shape of the curves indicates that heaviest tails (*i.e.* largest values of $\gamma(t(r))$) are found in the neighbourhood of the curve χ_{i_0} (*i.e.* for small values of r). At the opposite, lightest tails are found in the neighbourhood of the curve χ_{i_1}. This observation can be confirmed on the QQ-plots obtained by drawing some log-spacings versus standard exponential quantiles:

$$\left(\log \frac{\hat{k}_{t(r)}}{j}, \log \frac{Z_{m_{t(r)}-j+1, m_{t(r)}}}{Z_{m_{t(r)}-\hat{k}_{t(r)}+1, m_{t(r)}}},\ j = 1, \dots, \hat{k}_{t(r)}\right)$$

for $r \in \{0, 1\}$. These QQ-plots rely on the property that, under model (1), the log-spacings $\log(Z_{m_{t(r)}-j+1,m_{t(r)}}/Z_{m_{t(r)}-\hat{k}_{t(r)}+1,m_{t(r)}})$ computed in the ball $B(t(r), \hat{h}_{t(r)})$ are approximatively distributed from an exponential distribution with scale parameter $\gamma(t(r))$. See [12], Section 6.2, for a review on exploratory data analysis methods for extremes. The obtained QQ-plots for $r \in \{0, 1\}$ are presented on Figure 3. Let us note that the plots are approximatively linear, confirming the good adequation of the heavy-tail model (1) to the dataset. Two lines with slopes $\hat{\gamma}(t(0)) \simeq 0.36$ and $\hat{\gamma}(t(1)) \simeq 0.09$ have been superimposed. These very different slopes confirm a strong heterogeneity of the sample in terms of tail behaviour.

6 Proofs

6.1 Preliminary results

Our first auxiliary lemma is a simple unconditioning tool for determining the asymptotic distribution of a random variable.

Lemma 1 *Let (X_n) and (Y_n) be two sequences of real random variables. Suppose there exists an event $\mathcal{A}_n$ such that $(X_n|\mathcal{A}_n) \stackrel{d}{=} (Y_n|\mathcal{A}_n)$ with $\mathbb{P}(\mathcal{A}_n) \to 1$. Then, $Y_n \stackrel{d}{\to} Y$ implies $X_n \stackrel{d}{\to} Y$.*

Proof of Lemma 1 − For all $x \in \mathbb{R}$, the well-known expansion

$$\mathbb{P}(X_n \le x) = \mathbb{P}(\{X_n \le x\}|\mathcal{A}_n)\mathbb{P}(\mathcal{A}_n) + \mathbb{P}(\{X_n \le x\}|\mathcal{A}_n^C)\mathbb{P}(\mathcal{A}_n^C),$$

where $\mathcal{A}_n^C$ is the complementary event associated to $\mathcal{A}_n$, leads to the following inequalities:

$$\mathbb{P}(\{X_n \le x\}|\mathcal{A}_n)\mathbb{P}(\mathcal{A}_n) \le \mathbb{P}(X_n \le x) \le \mathbb{P}(\{X_n \le x\}|\mathcal{A}_n)\mathbb{P}(\mathcal{A}_n) + \mathbb{P}(\mathcal{A}_n^C).$$

Since $(X_n|\mathcal{A}_n) \stackrel{d}{=} (Y_n|\mathcal{A}_n)$, it follows that:

$$\mathbb{P}(\{Y_n \le x\} \cap \mathcal{A}_n) \le \mathbb{P}(X_n \le x) \le \mathbb{P}(\{Y_n \le x\} \cap \mathcal{A}_n) + \mathbb{P}(\mathcal{A}_n^C).$$

Taking into account of

$$\mathbb{P}(Y_n \le x) - \mathbb{P}(\mathcal{A}_n^C) \le \mathbb{P}(\{Y_n \le x\} \cap \mathcal{A}_n) \le \mathbb{P}(Y_n \le x)$$

leads to:

$$\mathbb{P}(Y_n \le x) - \mathbb{P}(\mathcal{A}_n^C) \le \mathbb{P}(X_n \le x) \le \mathbb{P}(Y_n \le x) + \mathbb{P}(\mathcal{A}_n^C).$$

The conclusion is then straightforward since $\mathbb{P}(Y_n \le x) \to \mathbb{P}(Y \le x)$ and $\mathbb{P}(\mathcal{A}_n^C) \to 0$. ■

The next lemma provides the asymptotic distribution of extreme quantile estimators from an uniform distribution in a situation analogous to **(S.1)** in the unconditional case.

Lemma 2 *Let $V_1, \dots, V_M$ be independent uniform random variables. For any sequence $(\theta_M) \subset (0,1)$ such that $\theta_M \to 0$ and $M\theta_M \to \infty$,*

$$\left(\frac{M}{\theta_M}\right)^{1/2} (V_{\lfloor M\theta_M \rfloor, M} - \theta_M) \xrightarrow{d} \mathcal{N}(0,1).$$

Proof of Lemma 2 – For the sake of simplicity, let us introduce $k_M = \lfloor M\theta_M \rfloor$. From Rényi's representation theorem,

$$V_{k_M,M} \stackrel{d}{=} \sum_{i=1}^{k_M} E_i \Big/ \sum_{i=1}^{M+1} E_i$$

where $E_1, \dots, E_{M+1}$ are independent random variables from a standard exponential distribution. Thus,

$$\begin{aligned}
\xi_M &\stackrel{def}{=} \left(\frac{M}{\theta_M}\right)^{1/2} (V_{k_M,M} - \theta_M) \stackrel{d}{=} \left(\frac{1}{M}\sum_{i=1}^{M+1} E_i\right)^{-1} \left(\frac{M}{\theta_M}\right)^{1/2} \\
&\times \left[\frac{1}{k_M}\sum_{i=1}^{k_M} E_i \left(\frac{k_M}{M} - \theta_M\right) + \theta_M \left(\frac{1}{k_M}\sum_{i=1}^{k_M} E_i - 1\right)\right. \\
&- \left.\theta_M \left(\frac{1}{M}\sum_{i=1}^{M+1} E_i - 1\right)\right],
\end{aligned}$$

and, in view of the law of large numbers, we have

$$\begin{aligned}
\xi_M &\stackrel{\mathrm{P}}{\sim} \left(\frac{M}{\theta_M}\right)^{1/2} \left(\frac{k_M}{M} - \theta_M\right)(1 + o_P(1)) + (M\theta_M)^{1/2} \left(\frac{1}{k_M}\sum_{i=1}^{k_M} E_i - 1\right) \\
&- (M\theta_M)^{1/2} \left(\frac{1}{M}\sum_{i=1}^{M+1} E_i - 1\right) \stackrel{def}{=} \xi_{1,M} + \xi_{2,M} - \xi_{3,M}.
\end{aligned}$$

Let us consider the three terms separately. First, writing $k_M = M\theta_M - \tau_M$ with $\tau_M \in [0,1)$, we have

$$\xi_{1,M} \stackrel{\mathrm{P}}{\sim} \left(\frac{M}{\theta_M}\right)^{1/2} \frac{\tau_M}{M} = \frac{\tau_M}{(M\theta_M)^{1/2}} \to 0, \tag{12}$$

since $M\theta_M \to \infty$. Second, since $k_M \sim M\theta_M$, the central limit theorem entails

$$\xi_{2,M} \sim k_M^{1/2} \left(\frac{1}{k_M}\sum_{i=1}^{k_M} E_i - 1\right) \xrightarrow{d} \mathcal{N}(0,1). \tag{13}$$

Similarly, it is easy to check that

$$\xi_{3,M} = O_{\mathrm{P}}(\theta_M^{1/2}) = o_{\mathrm{P}}(1), \tag{14}$$

since $\theta_M \to 0$. Collecting (12), (13) and (14) concludes the proof. ■

6.2 Proofs of main results

Proof of Proposition 1 – Under **(A)** and since the random values $\{Z_i(t),\ i = 1, \dots, m_t\}$ are independent, we have:

$$\{\log Z_i(t),\ i = 1, \dots, m_t\} \stackrel{d}{=} \{\log q(V_i, x_i)\ i = 1, \dots, m_t\},$$

where x_i is the covariate associated to $Z_i(t)$. Denoting by $\psi(i)$ the random index of the covariate associated to the observation $Z_{m_t-i+1,m_t}(t)$, we obtain

$$\{\log Z_{m_t-i+1,m_t}(t),\ i = 1, \dots, m_t\} \stackrel{d}{=} \{\log q(V_{\psi(i)}, x_{\psi(i)})\ i = 1, \dots, m_t\}.$$

Let us consider the event $\mathcal{A}_n = \mathcal{A}_{1,n} \cap \mathcal{A}_{2,n}$ where

$$\begin{aligned}
\mathcal{A}_{1,n} &= \left\{ \min_{i=1,\dots,k_t-1} \log \frac{q(V_{i,m_t}, u_i)}{q(V_{i+1,m_t}, u_{i+1})} > 0, \forall (u_1, \dots, u_{k_t}) \subset B(t, h_t) \right\} \text{ and} \\
\mathcal{A}_{2,n} &= \left\{ \min_{i=k_t+1,\dots,m_t} \log \frac{q(V_{k_t,m_t}, u_{k_t})}{q(V_{i,m_t}, u_i)} > 0, \forall (u_{k_t+1}, \dots, u_{m_t}) \subset B(t, h_t) \right\}.
\end{aligned}$$

Conditionally to $\mathcal{A}_{1,n}$, the random variables $q(V_{i,m_t}, u_i)$, $i = 1, \dots, k_t$ are ordered as

$$q(V_{k_t,m_t}, u_{k_t}) \le q(V_{k_t-1,m_t}, u_{k_t-1}) \le \dots \le q(V_{1,m_t}, u_1),$$

and, conditionally to $\mathcal{A}_{2,n}$, the remaining random variables $q(V_{i,m_t}, u_i)$, $i = k_t + 1, \dots, m_t$ are smaller since

$$\max_{i=k_t+1,\dots,m_t} q(V_{i,m_t}, u_i) \le q(V_{k_t,m_t}, u_{k_t}).$$

Thus, conditionally to $\mathcal{A}_n$, the k_t largest random values taken from the set $\{\log q(V_{\psi(i)}, x_{\psi(i)}),\ i = 1, \dots, m_t\}$ are $\{\log q(V_{i,m_t}, x_{\psi(i)}),\ i = 1, \dots, k_t\}$. Consequently, for $J_{k_t} = \{1, \dots, k_t\}$ and letting $T_i \stackrel{def}{=} x_{\psi(i)}$, we have:

$$\{\log Z_{m_t-i+1,m_t}(t), i \in J_{k_t} | \mathcal{A}_n\} \stackrel{d}{=} \{\log q(V_{i,m_t}, T_i), i \in J_{k_t} | \mathcal{A}_n\}.$$

To conclude the proof, it remains to show that $\mathbb{P}(\mathcal{A}_n) \to 1$ as $n \to \infty$. Let us define $\delta_{m_t} = m_t^{-(1+\delta)}$ and consider the events

$$\begin{aligned}
\mathcal{A}_{3,n} &= \{V_{1,m_t} > \delta_{m_t}\} \cap \{V_{m_t,m_t} < 1 - \delta_{m_t}\} \\
\mathcal{A}_{4,n} &= \left\{ \min_{i=1,\dots,k_t} \log \frac{q(V_{i,m_t}, t)}{q(V_{i+1,m_t}, t)} > 2\omega_n(\delta_{m_t}) \right\}.
\end{aligned}$$

Under $\mathcal{A}_{3,n}$, we have $\delta_{m_t} < V_{i,m_t} < 1-\delta_{m_t}$ for all $i = 1,\dots,m_t$. Hence, for all $(u_i,u_j) \in B(t,h_t)^2$, it follows that, on the one hand

$$\begin{aligned}\log \frac{q(V_{j,m_t},u_j)}{q(V_{i,m_t},u_i)} &= \log \frac{q(V_{j,m_t},t)}{q(V_{i,m_t},t)} + \log \frac{q(V_{j,m_t},u_j)}{q(V_{j,m_t},t)} + \log \frac{q(V_{i,m_t},t)}{q(V_{i,m_t},u_i)} \\ &\geq \log \frac{q(V_{j,m_t},t)}{q(V_{i,m_t},t)} - 2\omega_n(\delta_{m_t}),\end{aligned}$$

and on the other hand,

$$\begin{aligned}\min_{i=k_t+1,\dots,m_t} \log \frac{q(V_{k_t,m_t},u_{k_t})}{q(V_{i,m_t},u_i)} &\geq \min_{i=k_t+1,\dots,m_t} \log \frac{q(V_{k_t,m_t},t)}{q(V_{i,m_t},t)} - 2\omega_n(\delta_{m_t}) \\ &\geq \log \frac{q(V_{k_t,m_t},t)}{q(V_{k_t+1,m_t},t)} - 2\omega_n(\delta_{m_t}).\end{aligned}$$

Consequently $\mathcal{A}_{3,n} \cap \mathcal{A}_{4,n} \subset \mathcal{A}_n$. Remarking that

$$\mathbb{P}(\mathcal{A}_{3,n}) \geq \mathbb{P}(V_{1,m_t} > \delta_{m_t}) + \mathbb{P}(V_{m_t,m_t} < 1-\delta_{m_t}) - 1 = 2\mathbb{P}(V_{1,m_t} > \delta_{m_t}) - 1 \to 1,$$

since $V_{m_t,m_t} \stackrel{d}{=} 1 - V_{1,m_t}$ and $\mathbb{P}(V_{1,m_t} > \delta_{m_t}) = (1-\delta_{m_t})^{m_t} \to 1$, it thus remains to prove that $\mathbb{P}(\mathcal{A}_{4,n}) \to 1$. From [4], paragraph 1.3.1, condition **(A)** implies that there exists $c(t) > 0$, depending only on t such that, for all $\alpha \in (0,1)$,

$$q(\alpha,t) = c(t) \exp\left\{ \int_\alpha^1 \frac{\gamma(t) + \Delta(u,t)}{u} du \right\},$$

which is the so-called Karamata representation for normalised regularly varying functions. Hence, for all $i \in J_{k_t}$,

$$\log \frac{q(V_{i,m_t},t)}{q(V_{i+1,m_t},t)} = \int_{V_{i,m_t}}^{V_{i+1,m_t}} \frac{\gamma(t) + \Delta(u,t)}{u} du,$$

and it follows that

$$\log \frac{q(V_{i,m_t},t)}{q(V_{i+1,m_t},t)} \geq (\gamma(t) - \bar{\Delta}(V_{k_t+1,m_t},t)) \log \frac{V_{i+1,m_t}}{V_{i,m_t}},$$

leading to

$$\begin{aligned}\mathbb{P}(\mathcal{A}_{4,n}) &\geq \mathbb{P}\left((\gamma(t) - \bar{\Delta}(V_{k_t+1,m_t},t)) \min_{i=1,\dots,k_t} \log \frac{V_{i+1,m_t}}{V_{i,m_t}} > 2\omega_n(\delta_{m_t}) \right) \\ &\geq \mathbb{P}\left(\left\{ \min_{i=1,\dots,k_t} \log \frac{V_{i+1,m_t}}{V_{i,m_t}} \geq \frac{4\omega_n(\delta_{m_t})}{\gamma(t)} \right\} \cap \left\{ \bar{\Delta}(V_{k_t+1,m_t},t) < \gamma(t)/2 \right\} \right) \\ &\geq \mathbb{P}\left(\min_{i=1,\dots,k_t} \log \frac{V_{i+1,m_t}}{V_{i,m_t}} \geq \frac{4\omega_n(\delta_{m_t})}{\gamma(t)} \right) + \mathbb{P}\left(\bar{\Delta}(V_{k_t+1,m_t},t) < \gamma(t)/2 \right) - 1 \\ &\stackrel{def}{=} P_{1,m_t} + P_{2,m_t} - 1.\end{aligned}$$

In view of Rényi representation for uniform ordered random variables,

$$\{i\log(V_{i,m_t}^{-1}/V_{i+1,m_t}^{-1}),\ i\in J_{k_t}\}\stackrel{d}{=}\{F_i,\ i\in J_{k_t}\},$$

where $F_1,\dots,F_{k_t}$ are independent random variables from a standard exponential distribution, we have

$$\begin{aligned}P_{1,m_t} &= \mathbb{P}\left(\min_{i=1,\dots,k_t}\frac{F_i}{i}\geq\frac{4\omega_n(\delta_{m_t})}{\gamma(t)}\right)=\prod_{i=1}^{k_t}\exp\left(-\frac{4i\omega_n(\delta_{m_t})}{\gamma(t)}\right)\\ &= \exp\left(-\frac{2}{\gamma(t)}k_t(k_t+1)\omega_n(\delta_{m_t})\right)\to 1,\end{aligned}$$

since $k_t^2\omega_n(\delta_{m_t})\to 0$. Furthermore, $V_{k_t+1,m_t}=(k_t/m_t)(1+o_{\mathrm{P}}(1))\stackrel{\mathrm{P}}{\to}0$ and $\Delta(\alpha,t)\to 0$ as $\alpha\to 0$ entail $P_{2,m_t}\to 1$. The conclusion follows. ■

Proof of Proposition 2 – From Proposition 1, there exists an event $\mathcal{A}_n$ with $\mathbb{P}(\mathcal{A}_n)\to 1$ such that $(Z_{m_t,m_t}(t)|\mathcal{A}_n)\stackrel{d}{=}(q(V_{1,m_t},T_1)|\mathcal{A}_n)$ and thus,

$$\begin{aligned}\mathbb{P}(Z_{m_t,m_t}(t)<q(\alpha_{m_t},t)) &= \mathbb{P}\left(\left\{\log\frac{q(V_{1,m_t},T_1)}{q(\alpha_{m_t},t)}<0\right\}\cap\mathcal{A}_n\right)\\ &+ \mathbb{P}\left(\left\{\log\frac{Z_{m_t,m_t}(t)}{q(\alpha_{m_t},t)}<0\right\}\cap\mathcal{A}_n^C\right)\\ &\stackrel{def}{=} P_{3,m_t}+P_{4,m_t}.\end{aligned}\tag{15}$$

Clearly, $P_{4,m_t}\leq\mathbb{P}(\mathcal{A}_n^C)\to 0$. Let us now consider the term P_{3,m_t}. Introducing $\delta_{m_t}=m_t^{-(1+\delta)}$ and $\mathcal{A}_{5,n}=\{V_{1,m_t}\in[\delta_{m_t},1-\delta_{m_t}]\}$, we have

$$\begin{aligned}P_{3,m_t} &= \mathbb{P}\left(\left\{\log\frac{q(V_{1,m_t},T_1)}{q(\alpha_{m_t},t)}<0\right\}\cap\mathcal{A}_n\cap\mathcal{A}_{5,n}\right)\\ &+ \mathbb{P}\left(\left\{\log\frac{q(V_{1,m_t},T_1)}{q(\alpha_{m_t},t)}<0\right\}\cap\mathcal{A}_n\cap\mathcal{A}_{5,n}^C\right)\end{aligned}$$

and standard calculations lead to:

$$\begin{aligned}&\mathbb{P}\left(\left\{\log\frac{q(V_{1,m_t},T_1)}{q(\alpha_{m_t},t)}<0\right\}\cap\mathcal{A}_{5,n}\right)+\mathbb{P}(\mathcal{A}_n)-1\leq P_{3,m_t}\\ \leq\ &\mathbb{P}\left(\left\{\log\frac{q(V_{1,m_t},T_1)}{q(\alpha_{m_t},t)}<0\right\}\cap\mathcal{A}_{5,n}\right)+\mathbb{P}(\mathcal{A}_{5,n}^C).\end{aligned}$$

Furthermore, $\mathcal{A}_{5,n}$ implies

$$\left|\log\frac{q(V_{1,m_t},T_1)}{q(V_{1,m_t},t)}\right|\leq\omega_n(\delta_{m_t}),$$

and thus

$$\begin{aligned}
& \mathbb{P}\left(\left\{\log\frac{q(V_{1,m_t},t)}{q(\alpha_{m_t},t)} < -\omega_n(\delta_{m_t})\right\}\cap\mathcal{A}_{5,n}\right)+\mathbb{P}(\mathcal{A}_n)-1 \le P_{3,m_t} \\
\le \quad & \mathbb{P}\left(\left\{\log\frac{q(V_{1,m_t},t)}{q(\alpha_{m_t},t)} < \omega_n(\delta_{m_t})\right\}\cap\mathcal{A}_{5,n}\right)+\mathbb{P}(\mathcal{A}_{5,n}^C),
\end{aligned}$$

which entails

$$\begin{aligned}
& \mathbb{P}\left(\log\frac{q(V_{1,m_t},t)}{q(\alpha_{m_t},t)} < -\omega_n(\delta_{m_t})\right)+\mathbb{P}(\mathcal{A}_{5,n})+\mathbb{P}(\mathcal{A}_n)-2 \le P_{1,m_t} \\
\le \quad & \mathbb{P}\left(\log\frac{q(V_{1,m_t},t)}{q(\alpha_{m_t},t)} < \omega_n(\delta_{m_t})\right)+\mathbb{P}(\mathcal{A}_{5,n}^C). \qquad (16)
\end{aligned}$$

Let us now focus on the quantity

$$\begin{aligned}
P_{5,m_t} & \stackrel{def}{=} \mathbb{P}\left(\log\frac{q(V_{1,m_t},t)}{q(\alpha_{m_t},t)} < \pm\omega_n(\delta_{m_t})\right) \\
& = \left[\mathbb{P}\left(\log\frac{q(V_1,t)}{q(\alpha_{m_t},t)} < \pm\omega_n(\delta_{m_t})\right)\right]^{m_t} \\
& = \left[\mathbb{P}\left(q(V_1,t) < \mathrm{e}^{\pm\omega_n(\delta_{m_t})}q(\alpha_{m_t},t)\right)\right]^{m_t} \\
& = \left[\mathbb{P}\left(1-V_1 < F\left(\mathrm{e}^{\pm\omega_n(\delta_{m_t})}q(\alpha_{m_t},t),t\right)\right)\right]^{m_t} \\
& = \exp\left[m_t\log F\left(\mathrm{e}^{\pm\omega_n(\delta_{m_t})}q(\alpha_{m_t},t),t\right)\right].
\end{aligned}$$

Since $\mathrm{e}^{\pm\omega_n(\delta_{m_t})}q(\alpha_{m_t},t)\to\infty$ and introducing the conditional survival function $\bar{F}(.,t)=1-F(.,t)$, we have

$$\begin{aligned}
m_t\log F\left(\mathrm{e}^{\pm\omega_n(\delta_{m_t})}q(\alpha_{m_t},t),t\right) & = -m_t\bar{F}\left(\mathrm{e}^{\pm\omega_n(\delta_{m_t})}q(\alpha_{m_t},t),t\right)(1+o(1)) \\
& = -m_t\alpha_{m_t}\frac{\bar{F}\left(\mathrm{e}^{\pm\omega_n(\delta_{m_t})}q(\alpha_{m_t},t),t\right)}{\bar{F}\left(q(\alpha_{m_t},t),t\right)}(1+o(1)).
\end{aligned}$$

As already mentioned, **(A)** implies (1) which, in turn, shows that $\bar{F}(.,t)$ is a regularly function at infinity with index $-1/\gamma(t)$. Hence, since $\mathrm{e}^{\pm\omega_n(\delta_{m_t})}\to 1$, we thus have (see [4], Theorem 1.5.2),

$$\frac{\bar{F}\left(\mathrm{e}^{\pm\omega_n(\delta_{m_t})}q(\alpha_{m_t},t),t\right)}{\bar{F}\left(q(\alpha_{m_t},t),t\right)}\to 1.$$

As a conclusion,

$$P_{5,m_t}=\left[1-\alpha_{m_t}(1+o(1))\right]^{m_t}, \qquad (17)$$

and collecting (16) and (17) leads to:

$$\begin{aligned}
& \left[1-\alpha_{m_t}(1+o(1))\right]^{m_t}+\mathbb{P}(\mathcal{A}_{5,n})+\mathbb{P}(\mathcal{A}_n)-2 \\
\le \quad & P_{3,m_t}\le\left[1-\alpha_{m_t}(1+o(1))\right]^{m_t}+\mathbb{P}(\mathcal{A}_{5,n}^C).
\end{aligned}$$

Since $\mathbb{P}(\mathcal{A}_{5,n})\to 1$ and $\mathbb{P}(\mathcal{A}_n)\to 1$, it is then straightforward that $P_{3,m_t}\to 0$ under **(S.1)** and $P_{3,m_t}\to\mathrm{e}^{-c}$ under **(S.2)** or **(S.3)**. Equation (15) concludes the proof. ■

Proof of Theorem 1 – Let us introduce, for the sake of simplicity, $k_t = \lfloor m_t \alpha_{m_t} \rfloor$. From Proposition 1, there exists an event $\mathcal{A}_n$ such that:

$$\left((m_t \alpha_{m_t})^{1/2} \log \frac{\hat{q}_1(\alpha_{m_t}, t)}{q(\alpha_{m_t}, t)} \middle| \mathcal{A}_n \right) \stackrel{d}{=} \left((m_t \alpha_{m_t})^{1/2} \log \frac{q(V_{k_t, m_t}, T_{k_t})}{q(\alpha_{m_t}, t)} \middle| \mathcal{A}_n \right),$$

where $\mathbb{P}(\mathcal{A}_n) \to 1$. From Lemma 1, the convergence in distribution

$$(m_t \alpha_{m_t})^{1/2} \log \frac{q(V_{k_t, m_t}, T_{k_t})}{q(\alpha_{m_t}, t)} \xrightarrow{d} \mathcal{N}(0, \gamma^2(t)), \tag{18}$$

is a sufficient condition to obtain

$$(m_t \alpha_{m_t})^{1/2} \log \frac{\hat{q}_1(\alpha_{m_t}, t)}{q(\alpha_{m_t}, t)} \xrightarrow{d} \mathcal{N}(0, \gamma^2(t)).$$

A straightforward application of the δ-method will then conclude the proof. Let us prove the convergence in distribution (18). To this end, consider

$$R_n = \left| \log \frac{q(V_{k_t, m_t}, T_{k_t})}{q(V_{k_t, m_t}, t)} \right|$$

and let $\delta_{m_t} = m_t^{-(1+\delta)}$. Remark that, under **(S.1)**,

$$\mathbb{P}(R_n \leq \omega_n(\delta_{m_t})) \geq \mathbb{P}(V_{k_t, m_t} \in [\delta_{m_t}, 1 - \delta_{m_t}]) \to 1.$$

Thus, $R_n = O_{\mathrm{P}}(\omega_n(\delta_{m_t}))$ and we have

$$\log \frac{q(V_{k_t, m_t}, T_{k_t})}{q(\alpha_{m_t}, t)} = \log \frac{q(V_{k_t, m_t}, t)}{q(\alpha_{m_t}, t)} + O_{\mathrm{P}}(\omega_n(\delta_{m_t})). \tag{19}$$

Let us introduce the log-quantile function $g(.) = \log q(., t)$. Clearly, for all $\alpha \in (0, 1)$,

$$g'(\alpha) = \frac{\Delta(\alpha, t) - \gamma(t)}{\alpha}$$

and a first-order Taylor expansion leads to:

$$\begin{aligned} (m_t \alpha_{m_t})^{1/2} \log \frac{q(V_{k_t, m_t}, t)}{q(\alpha_{m_t}, t)} &= (m_t \alpha_{m_t})^{1/2} g'(\theta_{m_t})(V_{k_t, m_t} - \alpha_{m_t}) \\ &= \alpha_{m_t} g'(\theta_{m_t}) \left(\frac{m_t}{\alpha_{m_t}} \right)^{1/2} (V_{k_t, m_t} - \alpha_{m_t}), \end{aligned}$$

where $\theta_{m_t} \in [\min(\alpha_{m_t}, V_{k_t, m_t}), \max(\alpha_{m_t}, V_{k_t, m_t})]$. Now, $V_{k_t, m_t} \stackrel{\mathrm{P}}{\sim} \alpha_{m_t}$ entails $\theta_{m_t} \stackrel{\mathrm{P}}{\sim} \alpha_{m_t} \to 0$ and, from **(A)**,

$$\alpha_{m_t} g'(\theta_{m_t}) \stackrel{\mathrm{P}}{\sim} \theta_{m_t} g'(\theta_{m_t}) = \Delta(\theta_{m_t}, t) - \gamma(t) \xrightarrow{\mathrm{P}} -\gamma(t).$$

Then, Lemma 2 implies that

$$(m_t \alpha_{m_t})^{1/2} \log \frac{q(V_{k_t, m_t}, t)}{q(\alpha_{m_t}, t)} \xrightarrow{d} \mathcal{N}(0, \gamma(t)^2). \tag{20}$$

Collecting (19) and (20) concludes the proof after remarking that condition $(m_t \alpha_{m_t})^2 \omega_n(\delta_{m_t}) \to 0$ implies $(m_t \alpha_{m_t})^{1/2} \omega_n(\delta_{m_t}) \to 0$. ■

Proof of Theorem 2 – Since $q(.,t)$ is regularly varying with index $-\gamma(t)$, we have under **(S.2)** that $q(1/m_t,t)/q(\alpha_{m_t},t) \sim (m_t\alpha_{m_t})^{\gamma(t)} \to c^{\gamma(t)}$ and the following asymptotic expansion holds

$$\begin{aligned}\log\frac{\hat{q}_1(\alpha_{m_t},t)}{q(\alpha_{m_t},t)} &= \log\frac{\hat{q}_1(\alpha_{m_t},t)}{q(1/m_t,t)}+\frac{q(1/m_t,t)}{q(\alpha_{m_t},t)}\\ &= \log\frac{\hat{q}_1(\alpha_{m_t},t)}{q(1/m_t,t)}+\gamma(t)\log(c)+o(1).\end{aligned}$$

Now, recall that in situation **(S.2)**, for n large enough, $\lfloor m_t\alpha_{m_t}\rfloor = \lfloor c\rfloor$. Thus, from Proposition 1, there exists an event $\mathcal{A}_n$ such that $\mathbb{P}(\mathcal{A}_n)\to 1$ and

$$\left(\log\frac{\hat{q}_1(\alpha_{m_t},t)}{q(1/m_t,t)}\middle|\mathcal{A}_n\right)\stackrel{d}{=}\left(\log\frac{q(V_{\lfloor c\rfloor,m_t},T_{\lfloor c\rfloor})}{q(1/m_t,t)}\middle|\mathcal{A}_n\right).$$

Mimicking the proof of Theorem 1, we obtain

$$\log\frac{q(V_{\lfloor c\rfloor,m_t},T_{\lfloor c\rfloor})}{q(1/m_t,t)}=\log\frac{q(V_{\lfloor c\rfloor,m_t},t)}{q(1/m_t,t)}+O_{\mathrm{P}}(\omega_n(\delta_{m_t})).$$

To conclude, one can remark that $q(V_{\lfloor c\rfloor,m_t},t)$ is the $\lfloor c\rfloor$th uppest order statistics associated to a heavy-tailed distribution. In such a case, Corollary 4.2.4 of [12] states that $q(V_{\lfloor c\rfloor,m_t},t)/q(1/m_t,t)$ converges to a non-degenerated distribution. This asymptotic distribution is explicit even though it is not reproduced here. ■

Proof of Theorem 3 – Observing that

$$\log\hat{q}_2(\alpha_{m_t},t)=\log\hat{q}_1(\beta_{m_t},t)+\hat{\gamma}_n(t)\log\left(\frac{\beta_{m_t}}{\alpha_{m_t}}\right)$$

leads to the following expansion

$$\begin{aligned}\log\frac{\hat{q}_2(\alpha_{m_t},t)}{q(\alpha_{m_t},t)}-1 &= \log\frac{\hat{q}_1(\beta_{m_t},t)}{q(\beta_{m_t},t)}-1\\ &+ \log\left(\frac{\beta_{m_t}}{\alpha_{m_t}}\right)(\hat{\gamma}_n(t)-\gamma(t))\\ &- \log\frac{q(\alpha_{m_t},t)}{q(\beta_{m_t},t)}-\gamma(t)\log\left(\frac{\beta_{m_t}}{\alpha_{m_t}}\right)\\ &\stackrel{def}{=} \xi_{4,m_t}+\xi_{5,m_t}-\xi_{6,m_t}.\end{aligned}$$

First remark that, under **(A)**, as already shown in the proof of Proposition 1,

$$\log\frac{q(\alpha_{m_t},t)}{q(\beta_{m_t},t)}=\int_{\alpha_{m_t}}^{\beta_{m_t}}\frac{\gamma(t)+\Delta(u,t)}{u}du,$$

and thus, ξ_{6,m_t} can be simplified as

$$\xi_{6,m_t} = \int_{\alpha_{m_t}}^{\beta_{m_t}} \frac{\Delta(u,t)}{u} du$$

which leads to the bound:

$$|\xi_{6,m_t}| \leq \bar{\Delta}(\beta_{m_t}, t) \log\left(\frac{\beta_{m_t}}{\alpha_{m_t}}\right).$$

The two additional conditions are now treated separately since, under condition (5), the asymptotic distribution is imposed by ξ_{4,m_t} whereas, under (7), the asymptotic distribution is imposed by ξ_{5,m_t}.
(i) Under (5), Theorem 1 entails that

$$(m_t \beta_{m_t})^{1/2} \xi_{4,m_t} \xrightarrow{d} \mathcal{N}(0, \gamma^2(t)) \tag{21}$$

and

$$(m_t \beta_{m_t})^{1/2} \xi_{5,m_t} = \zeta_{m_t} \upsilon_n^{-1}(t) \upsilon_n(t)(\hat{\gamma}_n(t) - \gamma(t)) \xrightarrow{\mathrm{P}} 0, \tag{22}$$

from (4) and (5). Finally,

$$(m_t \beta_{m_t})^{1/2} |\xi_{6,m_t}| \leq \zeta_{m_t} \bar{\Delta}(\beta_{m_t}, t) \to 0, \tag{23}$$

from (5). Collecting (21), (22) and (23) concludes the proof of (6).

(ii) Under (7), Theorem 1 implies

$$\frac{\upsilon_n(t)}{\log(\beta_{m_t}/\alpha_{m_t})} \xi_{4,m_t} = \upsilon_n(t) \zeta_{m_t}^{-1} (m_t \beta_{m_t})^{1/2} \xi_{4,m_t} \xrightarrow{\mathrm{P}} 0. \tag{24}$$

Moreover, from (4),

$$\frac{\upsilon_n(t)}{\log(\beta_{m_t}/\alpha_{m_t})} \xi_{5,m_t} = \upsilon_n(t)(\hat{\gamma}_n(t) - \gamma(t)) \xrightarrow{d} \mathcal{D} \tag{25}$$

and finally,

$$\frac{\upsilon_n(t)}{\log(\beta_{m_t}/\alpha_{m_t})} |\xi_{6,m_t}| \leq \bar{\Delta}(\beta_{m_t}, t) \upsilon_n(t) \to 0, \tag{26}$$

under (7). Collecting (24), (25) and (26) concludes the proof of (8). ■

References

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

[2] Beirlant, J. and Goegebeur, Y. (2004). Local polynomial maximum likelihood estimation for Pareto-type distributions, *Journal of Multivariate Analysis*, **89**, 97–118.

[3] Bernard-Michel, C., Gardes, L. and Girard, S. (2008). Gaussian Regularized Sliced Inverse Regression, *Statistics and Computing*, to appear.

[4] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular variation*, Encyclopedia of Mathematics and its Applications, **27**, Cambridge University Press.

[5] Bosq, D. (2000). *Linear processes in function spaces*, Lecture Notes in Statistics, **149**, Springer-Verlag.

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

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

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

[9] Diebolt, J., Gardes, L., Girard, S. and Guillou, A. (2008). Bias-reduced extreme quantiles estimators of Weibull distributions, *Journal of Statistical Planning and Inference*, **138**, 1389–1401.

[10] Dekkers, A. and de Haan, L. (1989). On the estimation of the extreme-value index and large quantile estimation, *Annals of Statistics*, **17**, 1795–1832.

[11] Drees, H., de Haan, L. and Resnick, S. (2000). How to make a Hill plot, *Annals of Statistics*, **28**, 254–274.

[12] Embrechts, P., Klüppelberg, C., Mikosch, T. (1997). *Modelling extremal events*, Springer.

[13] Falk, M., Hüsler, J. and Reiss, R.D. (2004). *Laws of small numbers: Extremes and rare events*, 2nd edition, Birkhäuser.

[14] Ferraty, F., Laksaci, A. and Vieu, P. (2006). Estimating some characteristics of the conditional distribution in nonparametric functional models, *Statistical Inference for Stochastic Processes*, **9**, 47–76.

[15] Ferraty, F. and Vieu, P. (2002). The functional nonparametric model and application to spectrometric data, *Computational Statistics*, **17**, 545–564.

[16] Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis: Theory and Practice*, Springer Series in Statistics, Springer.

[17] Gangopadhyay, A.K. (1995). A note on the asymptotic behavior of conditional extremes, *Statistics and Probability Letters*, **25**, 163–170.

[18] Gardes, L. and Girard, S. (2005). Estimating extreme quantiles of Weibull tail-distributions, *Communication in Statistics - Theory and Methods*, **34**, 1065–1080.

[19] Gardes, L. and Girard, S. (2008). A moving window approach for nonparametric estimation of the conditional tail index, *Journal of Multivariate Analysis*, to appear.

[20] de Haan, L. (1984). Slow variation and the characterization of domains of attraction. In: Tiago de Oliveira, J. (Ed.) *Statistical extremes and applications*, pp. 31–48, Reidel, Dordrecht.

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

[22] Leurgans, S.E., Moyeed, R.A. and Silverman, B.W. (1993). Canonical correlation analysis when the data are curves, *Journal of the Royal Statistical Society Series B*, **55**, 725–740.

[23] Ramsay, J. and Silverman, B. (1997). *Functional Data Analysis*, Springer-Verlag.

[24] Ramsay, J. and Silverman, B. (2002). *Applied functional Data Analysis*, Springer-Verlag.

[25] Reiss, R.D. and Thomas, M. (2001). *Statistical analysis of extreme values*, Birkhäuser, Basel.

[26] 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.

[27] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations, *Journal of the American Statistical Association*, **73**, 812–815.

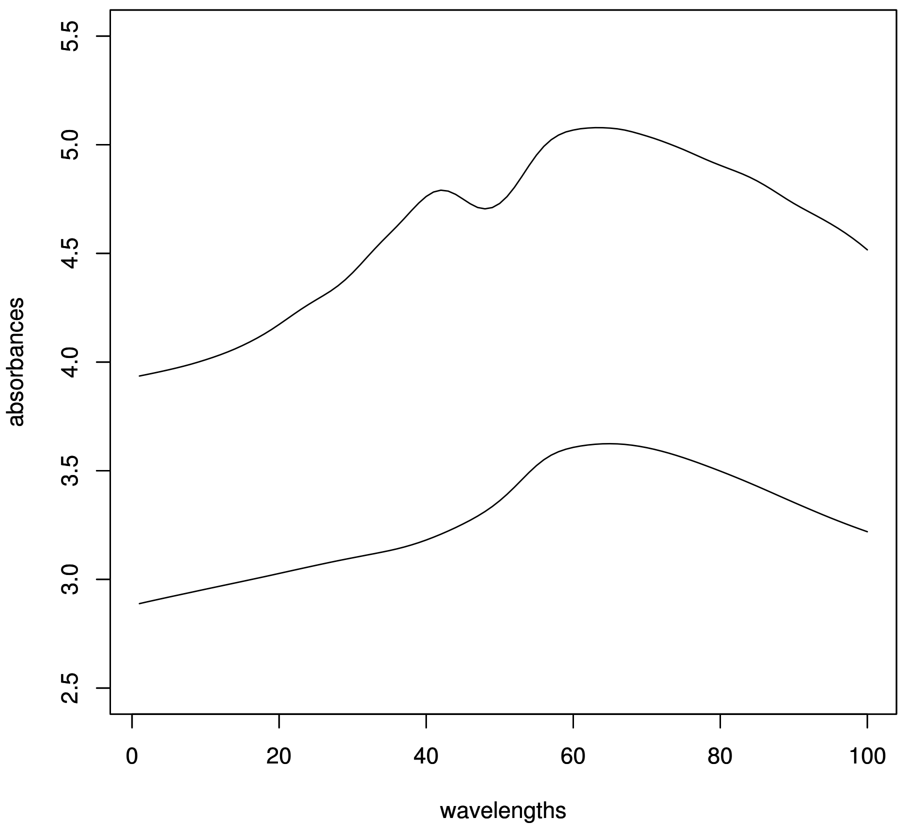

Figure 1: Selected spectrometric curves χ_{i_0} and χ_{i_1}.

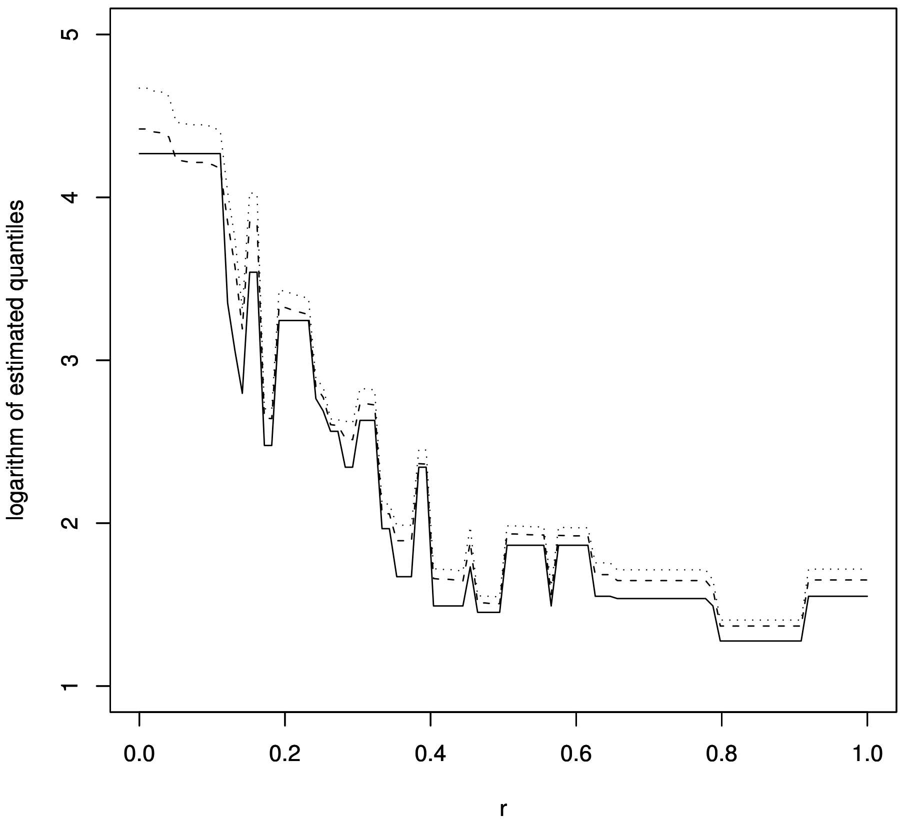

Figure 2: Estimated conditional extreme quantiles $\hat{q}_2(\alpha, t(r), W^{\mathrm{z}})$ in the selected direction as a function of r. Continuous line: maximum, dashed line: estimated quantile of order $\alpha = 1/300$, dotted line: estimated quantile of order $\alpha = 1/500$.

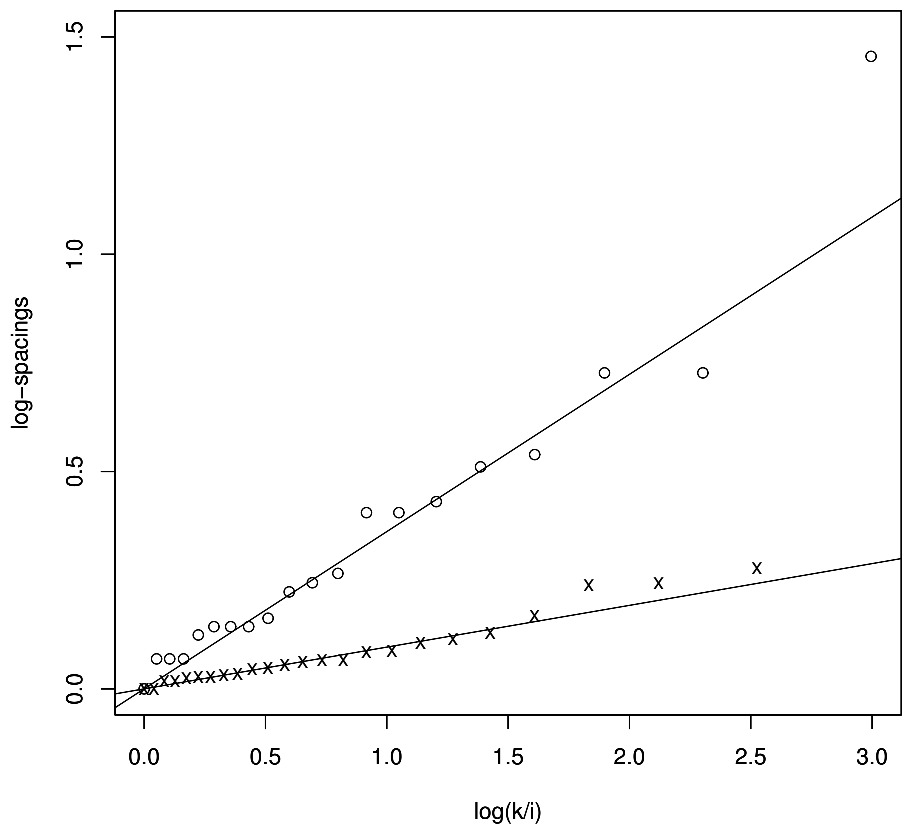

Figure 3: QQ-plots obtained for the two selected spectrometric curves ($\circ$: χ_{i_0} and $\times$: χ_{i_1}). The straight lines have slope $\hat{\gamma}_n(W^{\mathrm{H}}, t(r))$, $r \in \{0, 1\}$.

	$q(\alpha,t)$	$\Delta(\alpha,t)$
Pareto	$\alpha^{-\gamma(t)}$	0
Fréchet	$\alpha^{-\gamma(t)}\left\{\frac{1}{\alpha}\log\left(\frac{1}{1-\alpha}\right)\right\}^{-\gamma(t)}$	$-\frac{\gamma(t)}{2}\alpha(1+O(\alpha))$
Burr	$\alpha^{-\gamma(t)}\left(1-\alpha^{-\rho(t)}\right)^{-\gamma(t)/\rho(t)}$	$-\gamma(t)\alpha^{-\rho(t)}(1+O(\alpha^{-\rho(t)}))$

Table 1: Some examples of heavy-tailed distributions. For all distributions, $\gamma(t) > 0$ is the tail-index and $\rho(t) < 0$ is referred to as the second-order parameter in extreme-value theory.