Acta Mathematica Sinica, English Series
Published online: June 12, 2018
https://doi.org/10.1007/s10114-018-7095-9
http://www.ActaMath.com

Acta Mathematica Sinica, English Series

Nonparametric Estimation of Extreme Conditional Quantiles with Functional Covariate

Feng Yang HE
Department of Statistics, Ji'nan University, Guangzhou 510632, *P. R. China*
E-mail: 198578*hfy@*163*.com*

Ye Bin CHENG
Glorious Sun School of Business and Management, Donghua University, Shanghai 250001, *P. R. China*
E-mail: *chengyebin@hotmail.com*

Tie Jun TONG
Department of Mathematics, Hong Kong Baptist University, Hong Kong, P. R. China
E-mail: *tongt@hkbu.edu.hk*

Abstract Estimation of the extreme conditional quantiles with functional covariate is an important problem in quantile regression. The existing methods, however, are only applicable for heavy-tailed distributions with a positive conditional tail index. In this paper, we propose a new framework for estimating the extreme conditional quantiles with functional covariate that combines the nonparametric modeling techniques and extreme value theory systematically. Our proposed method is widely applicable, no matter whether the conditional distribution of a response variable Y given a vector of functional covariates X is short, light or heavy-tailed. It thus enriches the existing literature.

Keywords Extreme conditional quantile, extreme value theory, nonparametric modeling, functional covariate

MR(2010) Subject Classification 62G32

1 Introduction

Regression quantiles involving various extremal phenomena have been attracting a lot of attention with real applications. In these models, the covariates may appear to be curves in many disciplines such as biometrics, econometrics and statistical medicine. One way to deal with the curve data is to treat them as the observations of functional variables. As a consequence, we encounter the problem of estimating the extreme conditional quantiles with functional covariate, which has not been extensively studied in the literature. To solve the problem, nonparametric modeling techniques adapted to functional data are usually required to deal with the covariates, see, for example, [6] and the references therein. On the other hand, extreme value theory is

Received February 26, 2017, revised August 31, 2017, accepted January 8, 2018
Supported by the National Natural Science Foundation of China (Grant No. 11671338) and the Hong Kong Baptist University (Grant Nos. FRG1/16-17/018 and FRG2/16-17/074)

involved in estimating the extreme conditional quantile, for which one may refer to [5] or [11] for existing extreme value methods.

Note that parametric models are proposed in [4, 12] and semi-parametric methods are considered in [10]. Nonparametric methods have been first used for modeling trends in sample extremes for univariate case in [3]. Under regression setting, nonparametric estimators of extreme conditional quantiles with finite dimentional covariate have been discussed in [1, 2].

Recently, [7] and [8] have devoted to the nonparametric estimation of conditional quantiles when the functional covariate information is available and when the order of the quantiles converges to one as the sample size increases. [9] presented a nonparametric family of estimators for the tail index of a Weibull tail-distribution with functional covariate. We note, however, that the estimation methods in [7] and [8] are only applicable for some special heavy-tailed distributions with a positive conditional tail index. Also, [9] focused only on the estimation of the tail index for the special Weibull tail-distributions which encompass a variety of light-tailed distributions, such as Weibull, Gaussian, Gamma and Logistic distributions, and their estimators were based on a kernel estimator of extreme conditional quantile $q(\alpha_n|x) = \inf\{z : \bar{F}_Y(z|x) \leq \alpha_n\}$ when $\alpha_n \to 0$ as $n \to \infty$, where $\bar{F}_Y(y|x) = \mathbb{P}(Y > y|X = x)$ is the conditional survival function of Y given $X = x$. In such a case, it remains a question how to estimate a higher order quantile $q(\varphi_n|x)$ than $q(\alpha_n|x)$ where $\alpha_n \to 0$ and $\varphi_n/\alpha_n \to 0$ as $n \to \infty$. To answer this question, we propose a general framework for estimating the extreme conditional quantiles with functional covariate that combines the nonparametric modeling techniques and extreme value theory systematically. Our proposed method is widely applicable, no matter whether the conditional distribution of Y given $X = x$ is short, light or heavy-tailed, which correspond to the conditional tail index $\gamma(x) < 0$, $\gamma(x) = 0$ or $\gamma(x) > 0$, respectively.

2 Main Results

2.1 Notation and Assumptions

Let $\{(X_i, Y_i) : i = 1, \dots, n\}$ be an independent sample of the random pair (X, Y) in $\mathcal{F} \times \mathbb{R}$ where $\mathcal{F}$ is a functional space associated to a semi-metric $d : \mathcal{F} \times \mathcal{F} \to [0, +\infty)$, and a semi-metric d is not a metric because $d(x, x') = 0$ can not derive $x = x'$ for any $(x, x') \in \mathcal{F}^2$. The conditional density function of Y given $X = x$ is denoted by $f(y|x)$, and the functional nonparametric estimation of $\bar{F}_Y(y|x)$ is given for all $(x, y) \in (\mathcal{F}, \mathbb{R})$ by

$$\hat{\bar{F}}_n(y|x) = \sum_{i=1}^{n} \mathcal{K}(h^{-1}d(x, X_i))H(g^{-1}(Y_i - y)) / \sum_{i=1}^{n} \mathcal{K}(h^{-1}d(x, X_i)), \tag{2.1}$$

with $H(u) = \int_{-\infty}^{u} \mathcal{K}_0(v)dv$ $(\forall u \in \mathbb{R})$ where $\mathcal{K} : \mathbb{R}^+ \to \mathbb{R}^+$ and $\mathcal{K}_0 : \mathbb{R} \to \mathbb{R}^+$ are two kernel functions, and h and g are two real numbers (depending on n) such that $h \to 0$ and $g \to 0$ as $n \to \infty$.

The functional estimator of the conditional quantile $q(\tau|x)$ is defined as follows:

$$\hat{q}_n(\tau|x) = \hat{\bar{F}}_n^{-1}(\tau|x) = \inf\{y|\hat{\bar{F}}_n(y|x) \leq \tau\}, \tag{2.2}$$

where $\tau \in [\tau_L, 1)$ with τ_L being a small positive number. To establish the asymptotic results, we need some regularity assumptions, in which (F2) and (F3) are the same as conditions $(A.3)$ and $(A.4)$ in [8], respectively.

(F1) Suppose $F''(\cdot|x)$ exists and

$$\lim_{y \uparrow y_F^*(x)} \left(\frac{\bar{F}(y|x)}{\bar{F}'(y|x)} \right)' = -\gamma(x), \tag{2.3}$$

where $\gamma(x)$ is the so-called conditional tail index or conditional extreme value index and $\bar{F}(\cdot|x)$ has a right endpoint $y_F^*(x)$, which may be infinite.

(F2) $\mathcal{K}(\cdot)$ is a kernel function with support [0,1] and there exist two positive constants $0 < C_1 < C_2 < \infty$ such that $C_1 \leq \mathcal{K}(t) \leq C_2$ for any $t \in [0, 1]$.

(F3) $\mathcal{K}_0(\cdot)$ is a probability density function with support [−1,1].

(F1) is the conditional version of the famous von Mises' condition in extreme value theory. For ease of notation, we define the U_z function by $U_z(v) = \int_1^v u^{z-1} du$ for all $v \in \mathbb{R}$ and $z \in \mathbb{R}$. Under (F2), we introduce the l-th moment $m_x^{(l)}(h) = \mathbb{E}\{\mathcal{K}^l(h^{-1}d(x, X))\}$ for all $l > 0$. By Lemma 5.4, it can be shown that $m_x^{(l)}(h)$ is of the same asymptotic order as $\rho_x(h)$. Let also U_z^{-1} be the inverse function of U_z. By Theorem 1.1.6 in [5], (F1) implies that there exists a positive auxiliary function $a(\cdot|x)$ such that

$$\lim_{y \uparrow y_F^*(x)} \frac{\bar{F}(y + u(x)a(y|x)|x)}{\bar{F}(y|x)} = \frac{1}{U_{\gamma(x)}^{-1}(u(x))}, \tag{2.4}$$

for all $u(x)$ satisfying $1 + u(x)\gamma(x) > 0$. Furthermore, by Lemma 1.1.1 in [5], (2.4) is also equivalent to

$$d(t, \tau|x) := \frac{q(t\tau|x) - q(\tau|x)}{a(q(\tau|x)|x)} - U_{\gamma(x)}(1/t) \to 0, \tag{2.5}$$

for all $t > 0$ as $\tau \to 0$.

2.2 Estimation

Let $\tau_n \to 0$ and $\varphi_n/\tau_n \to 0$ as $n \to \infty$. Note that the conventional kernel estimator $\tilde{q}(\varphi_n|x)$ by (2.2) for $q(\varphi_n|x)$ lacks reliability due to the sparsity of sample data at the extreme tails. Therefore, the main purpose of this paper is to estimate $q(\varphi_n|x)$ precisely, which is a higher order quantile than $q(\tau_n|x)$ and therefore is called extreme conditional quantile. Following (2.5), we construct the estimator of $q(\varphi_n|x)$ as

$$\hat{q}_n(\varphi_n|x) = \hat{q}_n(\tau_n|x) + U_{\hat{\gamma}_n(x)}(\tau_n/\varphi_n)\hat{a}_n(x),$$

where $\hat{\gamma}_n(x)$ and $\hat{a}_n(x)$ are the estimators of $\gamma(x)$ and $a(q(\tau_n|x)|x)$, respectively. Inspired by [1], we introduce the following estimators:

$$\hat{\gamma}_n^P(x) = \frac{1}{\log(r)} \sum_{k=1}^{K-2} w_k \log \left(\frac{\hat{q}_n(l_k\tau_n|x) - \hat{q}_n(l_{k+1}\tau_n|x)}{\hat{q}_n(l_{k+1}\tau_n|x) - \hat{q}_n(l_{k+2}\tau_n|x)} \right),$$

$$\hat{a}_n^P(x) = \frac{1}{U_{\hat{\gamma}_n^P(x)}(r)} \sum_{k=1}^{K-2} w_k r^{\hat{\gamma}_n^P(x)k} (\hat{q}_n(l_k\tau_n|x) - \hat{q}_n(l_{k+1}\tau_n|x)),$$

where $l_k = r^{k-1}$, $k = 1, \ldots, K$, for all $r \in (0, 1)$, and $\{w_k\}$ are the weights summing to one. When $w_1 = \cdots = w_{K-2} = 1/(K-2)$, the estimator of $\gamma(x)$ leads to

$$\hat{\gamma}_n^{P,1}(x) = \frac{1}{(K-2)\log(r)} \left(\frac{\hat{q}_n(l_1\tau_n|x) - \hat{q}_n(l_2\tau_n|x)}{\hat{q}_n(l_{K-1}\tau_n|x) - \hat{q}_n(l_K\tau_n|x)} \right).$$

Letting $K=3$ and $r=1/2$, $\hat{\gamma}_n^{P,1}(x)$ corresponds to the well-known Pickands estimator. Additionally, if we consider the linear weights $w_k=2k/[(K-1)(K-2)]$ for $k=1,\ldots,K-2$, the estimator is given as

$$\hat{\gamma}_n^{P,2}(x)=\frac{2}{(K-1)(K-2)\log(r)}\sum_{k=1}^{K-2}\log\left(\frac{\hat{q}_n(l_k\tau_n|x)-\hat{q}_n(l_{k+1}\tau_n|x)}{\hat{q}_n(l_{K-1}\tau_n|x)-\hat{q}_n(l_K\tau_n|x)}\right).$$

2.3 Asymptotic Results

To investigate the asymptotic results of the proposed estimators, we further let $y_n(x)$ be a sequence such that $y_n(x)\uparrow y_F^*(x)$ and $y_{n,k}(x)=y_n(x)+U_{\gamma(x)}(1/l_k)a(y_n(x)|x)(1+o(1))$ for all $k=1,\ldots,K$. Additionally, the oscillations of the conditional survival function are controlled by

$$V_n(x)=\max_{k=1,\ldots,K}\sup_{|t-y_{n,k}(x)|\le g}\sup_{u\in B(x,h)}\left|\frac{\bar{F}(t|u)}{\bar{F}(y_{n,k}(x)|x)}-1\right|,$$

where $h\to 0$ and $g\to 0$ as $n\to\infty$. For convenience, the following function is introduced for all $v\in\mathbb{R}$ and $z\in\mathbb{R}$ by

$$U_z'(v)=\frac{\partial U_z(v)}{\partial z}=\int_1^v u^{z-1}\log(u)du.$$

For ease of notation, let also $\phi_n(x)=(n\bar{F}(y_n(x)|x)(m_x^{(1)}(h))^2/m_x^{(2)}(h))^{-1/2}$ and $\delta_n(x)=(n\tau_n(m_x^{(1)}(h))^2/m_x^{(2)}(h))^{-1/2}$. In addition, we note that the derived theoretical results in this paper is an adaptation of [1] to the situation of functional covariate.

Theorem 2.1 *Assume* (F1)–(F3) *hold. Let $x\in\mathcal{F}$ such that $\rho_x(h)>0$ for any $h>0$ and consider a sequence $0<l_K<\cdots<l_2<l_1\le 1$, where K is a positive integer. If $y_n(x)\uparrow y_F^*(x)$ such that $n\rho_x(h)\bar{F}(y_n(x)|x)\to\infty$ and $n\rho_x(h)\bar{F}(y_n(x)|x)V_n^2(x)\to 0$ as $n\to\infty$, then*

$$\left\{\phi_n^{-1}(x)\left(\frac{\hat{\bar{F}}(y_{n,k}(x)|x)}{\bar{F}(y_{n,k}(x)|x)}-1\right)\right\}_{k=1,\ldots,K}$$

is asymptotically normal distributed with mean zero and covariance matrix $V(x)$, where $V_{k,k'}(x)=l_{k\wedge k'}^{-1}$ for $(k,k')\in\{1,\ldots,K\}^2$.

Theorem 2.2 *Assume* (F1)–(F3) *hold. Let $x\in\mathcal{F}$ such that $\rho_x(h)>0$ for any $h>0$ and consider a sequence $0<l_K<\cdots<l_2<l_1\le 1$, where K is a positive integer. If $\tau_n\to 0$ such that $\delta_n(x)\to 0$ and $\delta_n^{-1}(x)V_n(x)\to 0$ as $n\to\infty$, then*

$$f(q(\tau_n|x)|x)\sqrt{n\tau_n^{-1}(m_x^{(1)}(h))^2/m_x^{(2)}(h)}\{\hat{q}_n(l_k\tau_n|x)-q(l_k\tau_n|x)\}_{k=1,\ldots,K} \tag{2.6}$$

is asymptotically normal distributed with mean zero and covariance matrix $C(x)$, where $C_{k,k'}(x)=(l_kl_{k'})^{-\gamma(x)}l_{k\wedge k'}^{-1}$ for $(k,k')\in\{1,\ldots,K\}^2$.

Note that (2.6) in Theorem 2.2 can be represented as

$$\left\{\delta_n^{-1}(x)\frac{q(\tau_n|x)(1+o(1))}{a(q(\tau_n|x)|x)}\frac{q(l_k\tau_n|x)}{q(\tau_n|x)}\left(\frac{\hat{q}_n(l_k\tau_n|x)}{q(l_k\tau_n|x)}-1\right)\right\}_{k=1,\ldots,K}.$$

Then together with the fact that $\lim_{n\to\infty}q(l_k\tau_n|x)/q(\tau_n|x)=l_k^{-(\gamma(x)\vee 0)}$ by Lemma 5.1, we have the following corollary.

Corollary 2.3 *Assume the conditions of Theorem 2.2 hold. The random vector*

$$\left\{\delta_n^{-1}(x)\frac{q(\tau_n|x)}{a(q(\tau_n|x)|x)}\left(\frac{\hat{q}_n(l_k\tau_n|x)}{q(l_k\tau_n|x)}-1\right)\right\}_{k=1,\ldots,K}$$

is asymptotically normal distributed with mean zero and covariance matrix $\tilde{C}(x)$*, where* $\tilde{C}_{k,k'}(x) = (l_k l_{k'})^{-(\gamma(x)\wedge 0)} l_{k\wedge k'}^{-1}$ *for* $(k,k')\in\{1,\ldots,K\}^2$*.*

Theorem 2.4 *Assume* (F1) *holds and* $\hat{q}_n(\tau_n|x)$ *is a functional kernel estimator of* $q(\tau_n|x)$ *defined in* (2.2)*. Let* $\hat{\gamma}_n(x)$ *and* $\hat{a}_n(x)$ *be two estimators of* $\gamma(x)$ *and* $a(q(\tau_n|x)|x)$*, respectively, such that*

$$\delta_n^{-1}(x)\left(\hat{\gamma}_n(x)-\gamma(x),\frac{\hat{a}_n(x)}{a(q(\tau_n|x)|x)}-1,\frac{\hat{q}_n(\tau_n|x)-q(\tau_n|x)}{a(q(\tau_n|x)|x)}\right)'\xrightarrow{d}\vartheta(x),$$

where $\xrightarrow{d}$ *denotes "convergence in distribution",* $\vartheta(x)$ *is a non-degenerate random vector in* $\mathbb{R}^3$ *and*

$$\delta_n(x)\log(\tau_n/\varphi_n)\to 0 \quad \text{and} \quad \delta_n^{-1}(x)\frac{d(\varphi_n/\tau_n,\tau_n|x)}{U'_{\gamma(x)}(\tau_n/\varphi_n)}\to 0 \text{ as } n\to\infty.$$

Then with $d(x)'=(1,-(\gamma(x)\wedge 0),(\gamma(x)\wedge 0)^2)$*, we have*

$$\delta_n^{-1}(x)\left(\frac{\hat{q}_n(\varphi_n|x)-q(\varphi_n|x)}{a(q(\tau_n|x)|x)U'_{\gamma(x)}(\tau_n/\varphi_n)}\right)\xrightarrow{d} d(x)'\vartheta(x).$$

Theorem 2.5 *Assume* (F1)–(F3) *hold. Denote* $\eta_n(x)=\max_{k=1,\ldots,K}|d(l_k,\tau_n|x)|$*, and let* $\tau_n\to 0$ *such that*

$$n\tau_n\rho_x(h)\to\infty \quad \text{and} \quad \delta_n^{-1}(x)(V_n(x)\vee\eta_n(x))\to 0 \text{ as } n\to\infty.$$

Then,

$$\delta_n^{-1}(x)\left(\hat{\gamma}_n^P(x)-\gamma(x),\frac{\hat{a}_n^P(x)}{a(q(\tau_n|x)|x)}-1,\frac{\hat{q}_n(\tau_n|x)-q(\tau_n|x)}{a(q(\tau_n|x)|x)}\right)'\xrightarrow{d} N(\boldsymbol{0},\mathcal{C}(x)),$$

where $\mathcal{C}(x)=A(x)C(x)A(x)'/(U_{\gamma(x)}(r))^2$ *and* $A(x)$ *is a* $3\times K$ *matrix defined by*

$$\begin{aligned}
A_{1,k}(x)&=r^{\gamma(x)k}(\beta_0^{(\gamma)}\omega_k+\beta_1^{(\gamma)}\omega_{k-1}+\beta_2^{(\gamma)}\omega_{k-2}),\\
A_{2,k}(x)&=r^{\gamma(x)k}(\beta_0^{(a)}\omega_k+\beta_1^{(a)}\omega_{k-1}+\beta_2^{(a)}\omega_{k-2}),\\
A_{3,k}(x)&=U_{\gamma(x)}(r)I\{k=1\},
\end{aligned}$$

for all $k=1,\ldots,K$*.*

Theorem 2.5 establishes the joint asymptotic normality of $(\hat{\gamma}_n^P(x),\hat{a}_n^P(x),\hat{q}_n(\tau_n|x))$. Following the above theorems, we can obtain the asymptotic normality of the extreme conditional quantile estimator, denoted by $\hat{q}_n^{\mathrm{P}}(\varphi_n|x)=\hat{q}_n(\tau_n|x)+U_{\hat{\gamma}_n^{\mathrm{P}}(x)}(\tau_n/\varphi_n)\hat{a}_n^{\mathrm{P}}(x)$.

Corollary 2.6 *Assume the conditions of Theorem 2.5 hold. Let* $\hat{q}_n(\tau_n|x)$ *be a functional kernel estimator of* $q(\tau_n|x)$ *defined in* (2.2)*. If* $\tau_n\to 0$ *such that* $\delta_n(x)\log(\tau_n/\varphi_n)\to 0$ *and* $[\delta_n(x)U'_{\gamma(x)}(\tau_n/\varphi_n)]^{-1}d(\varphi_n/\tau_n,\tau_n|x)\to 0$ *as* $n\to\infty$*, then*

$$\delta_n^{-1}(x)\left(\frac{\hat{q}_n^P(\varphi_n|x)-q(\varphi_n|x)}{a(q(\tau_n|x)|x)U'_{\gamma(x)}(\tau_n/\varphi_n)}\right)\xrightarrow{d} N(0,\mathcal{C}'(x)),$$

where $\mathcal{C}'(x)=d(x)'\mathcal{C}(x)d(x)$ *with* $d(x)'=(1,-(\gamma(x)\wedge 0),(\gamma(x)\wedge 0)^2)$*.*

3 Simulation Study

In this section, we conduct simulation studies to evaluate the finite sample performance of the proposed conditional tail index and the extreme conditional quantile estimators. The functional covariate $X \in \mathcal{F} = C[0,1]$ is defined by $X(t) = \sin(2\pi St)$ for all $t \in [0,1]$, where S is uniformly distributed on $[1/4, 3/4]$ and

$$\|X(t)\|_2^2 = \int_0^1 X^2(t)dt = \frac{1}{2}\left(1 - \frac{\sin(4\pi S)}{4\pi S}\right).$$

Moreover, we consider three cases for the conditional distribution of Y given X:

Case 1 The Beta distribution $\mathrm{Beta}(v(X), v(X))$ with $v(X) = 1 + 2\|X(t)\|_2^2$;

Case 2 The normal distribution $\mathbb{N}(0, 4\|X(t)\|_2^2)$;

Case 3 The Fréchet distribution with $\bar{F}_Y(y|X) = 1 - \exp(-y^{-(2+4\|X(t)\|_2^2)})$.

Note that the Beta distribution corresponds to $\gamma(X) = -1/v(X)$, the normal distribution corresponds to $\gamma(X) = 0$, and the Fréchet distribution corresponds to $\gamma(X) = 1/(2+4\|X(t)\|_2^2))$.

We focus on the estimation of $q(\varphi_n|x)$ with $\varphi_n = 1/n$ where $n = 200$ stands for the sample size, and $\{S_i, 1 \leq i \leq n\}$ are randomly drawn from the uniform distribution $U[1/4, 3/4]$ to generate the sample curves. The choice of the semi-metric d is $d(X, X') = \|X - X'\|_2^2$ for all $(X, X') \in \mathcal{F}^2$. In the functional nonparametric estimation of $\bar{F}(y|x)$, two kernel functions $\mathcal{K}(u) = (1-u^2)\mathbb{I}\{0 \leq u \leq 1\}$ and $\mathcal{K}_0(u) = \frac{3}{4}(1-u^2)\mathbb{I}\{-1 \leq u \leq 1\}$ are used. To assess the accuracy of the tail index estimators, we can compute the mean squared error (MSE) and the bias term (Bias) for both $\hat{\gamma}_n^{\mathrm{P},1}(x)$ and $\hat{\gamma}_n^{\mathrm{P},2}(x)$ in each simulation:

$$\mathrm{MSE}\{\hat{\gamma}_n(\cdot)\} = \frac{1}{N}\sum_{j=1}^{N}\{\hat{\gamma}_n(x_j) - \gamma(x_j)\}^2 \quad \text{and} \quad \mathrm{Bias}\{\hat{\gamma}_n(\cdot)\} = \frac{1}{N}\sum_{j=1}^{N}|\hat{\gamma}_n(x_j) - \gamma(x_j)|,$$

where $N = 100$ and $\{x_j, 1 \leq j \leq N\}$ represent random curves $\{x_j(t) = \sin(2\pi P_j t), 1 \leq j \leq N\}$ with $\{P_j, 1 \leq j \leq N\}$ being randomly drawn from the uniform distribution $U[1/4, 3/4]$.

	MSE			Bias		
	$\hat{\gamma}_n^{P,1}(x)$	$\hat{\gamma}_n^{P,2}(x)$	$\hat{\gamma}_n^{\phi_1}(x)$	$\hat{\gamma}_n^{P,1}(x)$	$\hat{\gamma}_n^{P,2}(x)$	$\hat{\gamma}_n^{\phi_1}(x)$
Beta	0.0819	0.0819		0.2773	0.2773	
Gaussian	0.0529	0.0529		0.2110	0.2110	
Fréchet	0.0672	0.0676	0.0041	0.2247	0.2264	0.0470

Table 1 Performance of the three tail index estimators with $n = 200$

In each simulation, for the performance of the extreme conditional quantile estimators, we consider the following two measures for evaluating the performance of $\tilde{q}_n(\varphi_n|x) = \hat{q}_n^{\mathrm{P},1}(\varphi_n|x)$, $\hat{q}_n^{\mathrm{P},2}(\varphi_n|x)$:

$$\mathrm{MSE}\{\tilde{q}_n(\varphi_n|\cdot)\} = \frac{1}{N}\sum_{j=1}^{N}\{\tilde{q}_n(\varphi_n|x_j) - q(\varphi_n|x_j)\}^2,$$

$$\mathrm{Bias}\{\tilde{q}_n(\varphi_n|\cdot)\} = \frac{1}{N}\sum_{j=1}^{N}|\tilde{q}_n(\varphi_n|x_j) - q(\varphi_n|x_j)|.$$

Note that the tail index estimators $\hat{\gamma}_n^{\mathrm{P},1}(x)$ and $\hat{\gamma}_n^{\mathrm{P},2}(x)$ depend on the unknown tuning parameters K and r. In our simulations, we use a grid search method to suggest the optimal tuning parameters. By Theorem 2.5 and Corollary 2.6, $\hat{q}_n^{\mathrm{P}}(\varphi_n|x)$ has an equal asymptotic variance as $\hat{\gamma}_n^{\mathrm{P}}(x)$ when $\gamma(x) \geq 0$. Hence, for consistency, we can choose the parameters (K, r) in our simulations by minimizing the asymptotic variance of $\hat{q}_n^{\mathrm{P}}(\varphi_n|x)$ with K ranging from $\mathbf{K} = \{3, 4, \ldots, 30\}$, r ranging from $\mathbf{R} = \{0.01, 0.02, \ldots, 0.99\}$ and S ranging from $\mathbf{S} = \{0.25, 0.26, \ldots, 0.75\}$. In Case 1, both $\hat{\gamma}_n^{\mathrm{P},1}(x)$ and $\hat{\gamma}_n^{\mathrm{P},2}(x)$ suggest the optimal values at $K = 3$ and $r = 0.01$. In Case 2, as $\gamma(x) = 0$, both $\hat{\gamma}_n^{\mathrm{P},1}(x)$ and $\hat{\gamma}_n^{\mathrm{P},2}(x)$ suggest the optimal values at $K = 3$ and $r = 0.14$. In Case 3, $\hat{\gamma}_n^{\mathrm{P},1}(x)$ suggests $K = 4$ and $r = 0.33$ whereas $\hat{\gamma}_n^{\mathrm{P},2}(x)$ suggests $K = 3$ and $r = 0.18$.

Also for the intermediate quantile level τ_n, in the above three cases, we consider to select it by minimizing the averaged MSE of $\hat{\gamma}_n^{\mathrm{P}}(x)$ when evaluating the performance of $\hat{\gamma}_n^{\mathrm{P}}(x)$, and minimizing the averaged MSE of $\hat{q}_n^P(\varphi_n|x)$ when evaluating the performance of $\hat{q}_n^P(\varphi_n|x)$, with τ_n ranging from $\mathcal{T} = \{0.10, 0.12, \ldots, 0.26\}$.

We use the cross-validation method to choose the smoothing parameters g and h. Specifically, we define

$$(g^{\mathrm{opt}}, h^{\mathrm{opt}}) = \arg\min_{(g,h)} \left\{ \sum_{i=1}^{n} \sum_{i=1}^{n} (H(g^{-1}(Y_i - Y_j)) - \hat{\hat{F}}_{n,-i}(Y_j|X_i))^2 : \ g \in \mathcal{G}, h \in \mathcal{H} \right\},$$

where $\hat{\hat{F}}_{n,-i}$ is the estimator given in (2.1) computed from the sample $\{(X_m, Y_m) : 1 \leq m \leq n, m \neq i\}$. Here, $\mathcal{G}$ and $\mathcal{H}$ form a regular grid. In Case 1, we suggest $\mathcal{G} = \{0.001, 0.002, \ldots, 0.01\}$ and $\mathcal{H} = \{1.50, 1.52, \ldots, 2.00\}$, which result in stable results. In Cases 2 and 3, $\mathcal{G} = \{0.02, 0.06, \ldots, 1.98\}$ and $\mathcal{H} = \{0.02, 0.04, \ldots, 1.00\}$.

The simulated MSE and Bias for evaluating both $\hat{\gamma}_n^{\mathrm{P}}(x)$ and $\hat{q}_n^P(\varphi_n|x)$ are averaged over 100 Monte Carlo simulations and reported in Tables 1 and 2. We note that $\hat{\gamma}_n^{\mathrm{P},2}(x)$ is either equally or more efficient than $\hat{\gamma}_n^{\mathrm{P},1}(x)$ in all settings, $\hat{q}_n^{P,2}(\varphi_n|x)$ outperforms $\hat{q}_n^{P,1}(\varphi_n|x)$ in Case 3 and they have the same performance in Cases 1 and 2.

	MSE				Bias			
	$\hat{q}_n^{P,1}$	$\hat{q}_n^{P,2}$	$\hat{q}_n^{w}$	$\hat{q}_n$	$\hat{q}_n^{P,1}$	$\hat{q}_n^{P,2}$	$\hat{q}_n^{w}$	$\hat{q}_n$
Beta	2.88E-4	2.88E-4		3.68E-4	0.0141	0.0141		0.0156
Gaussian	0.5588	0.5588		0.5770	0.6321	0.6321		0.6375
Fréchet	0.4807	0.4172	0.3274	0.4828	0.5070	0.4848	0.4279	0.5117

Table 2 Performance of the extreme conditional quantile estimators with $\varphi_n = 0.005$

In what follows, we first compare our estimators with the conventional quantile estimator for $q(\varphi_n|x)$, where the conventional estimator $\hat{q}_n(\varphi_n|x)$ can be obtained by (2.2). The averaged results of MSE and Bias of $\hat{q}_n(\varphi_n|x)$ over 100 Monte Carlo simulations are given in Table 2. It is clear that our extrapolated estimators for extreme conditional quantiles are more efficient than the conventional estimator $\hat{q}_n(\varphi_n|x)$ in terms of both the averaged MSE and the averaged Bias.

For further comparison, we also use the estimation method in [8] to estimate both the conditional tail index $\gamma(x)$ and the extreme conditional quantile $q(\varphi_n|x)$. Because their estimation method is suitable only for $\gamma(x) > 0$, we thus only need to consider the data generating process of Case 3. According to [8], the functional Weissman estimator of $q(\varphi_n|x)$ is given by $\hat{q}_n^w(\varphi_n|x) = \hat{q}_n(\tau_n|x)(\tau_n/\varphi_n)^{\hat{\gamma}_n^{\phi_1}(x)}$, where the functional estimator $\hat{q}_n(\tau_n|x)$ is obtained by (2.2), and $\hat{\gamma}_n^{\phi_1}(x) = \sum_{j=1}^{J}[\log \hat{q}_n(l_j\tau_n|x) - \log \hat{q}_n(\tau_n|x)]/\sum_{j=1}^{J}\log(1/l_j)$ is the functional Hill estimator for $\gamma(x)$. [8] advised to investigate different values of τ_n and l_j: $\tau_n = c\log(n)/n$ with $c \in \{5, 10, 15, 20\}$ and $l_j = (1/j)^s$ with $s \in \{1, 2, 3, 10\}$. Hence, when $l_j = (1/j)^s$, we can choose the parameters (J, s) by minimizing the asymptotic variance of $\hat{q}_n^w(\varphi_n|x)$, which has an equal asymptotic variance as $\hat{\gamma}_n^{\phi_1}(x)$, with J ranging from $\mathbf{J} = \{2, 3, \ldots, 30\}$ and s ranging from $\mathbf{S} = \{1, 2, 3, 10\}$. As a result, $\hat{\gamma}_n^{\phi_1}(x)$ suggests the optimal values at $J = 9$ and $s = 1$. For the choice of τ_n, we select it by minimizing the averaged MSE of $\hat{\gamma}_n^{\phi_1}(x)$ when evaluating the performance of $\hat{\gamma}_n^{\phi_1}(x)$, and minimizing the averaged MSE of $\hat{q}_n^w(\varphi_n|x)$ when evaluating the performance of $\hat{q}_n^w(\varphi_n|x)$, with τ_n ranging from $\mathcal{T} = \{\tau_n = cn^{-1}\log(n)\colon c = 5, 10, 15, 20\}$. Following [8], the smoothing parameter h is selected using the cross-validation method with h ranging from a regular grid $\mathcal{H} = \{h_1 \leq h_2 \leq \cdots \leq h_M\}$, where $h_1 = 0.01$, $h_M = 0.1$ and $M = 20$, and the choice of g is fixed at $g = 0.1$. The averaged MSE and the averaged Bias of $\hat{\gamma}_n^{\phi_1}(x)$ and $\hat{q}_n^w(\varphi_n|x)$ over 100 Monte Carlo simulations are reported in Tables 1 and 2, respectively. The results display that $\hat{\gamma}_n^{\phi_1}(x)$ and $\hat{q}_n^w(\varphi_n|x)$ have better performance than our estimators in Case 3. However, the restriction that the estimation method in [8] is only applicable for the case of $\gamma(x) > 0$ induces that it does not have a wide range of application. This limitation can also be seen in real data analysis in [8], where they have to consider the inverse of the original response variable so that their estimation method can be applied. As the obtained results from the inverse response variable may lack practical explanation, this suggests that our proposed method has its own merits because it also works for $\gamma(x) < 0$ and $\gamma(x) = 0$.

4 Real Data Analysis

We apply the proposed method to a real data example and estimate the extreme conditional quantiles of fat content with functional covariate being the Spectrometric Curves. The data set is described in [6, Section 2.1] and it includes $n = 215$ pieces of finely chopped meat. As mentioned earlier, [8] has applied the same real data to illustrate the behaviour of their large conditional quantiles estimators by considering the inverse of the original variable $\tilde{Y}$ (fat content) defined as $Y = 100/\tilde{Y}$. However, such an inverse transformation may result in a lack of practical explanation for the obtained results.

For each unit i, there is one spectrometric discretized curve x_i which corresponds to the absorbance measured at a grid of 100 wavelengths $x_i(\lambda_1), \ldots, x_i(\lambda_{100})$. As pointed out in [6], we view each unit as a continuous curve because of the fineness of the grid. As we can see from Figure 1, the shapes of the spectrometric curves are very smooth and looks very similar. Moreover, for each curve x_i, its fat content y_i is obtained by analytical chemical processing. In this study, we are interested in estimating the high conditional quantile of fat content with order φ_n. To achieve this goal, given two observed curves x_i and x_j, we apply the following semi-metric:

$$d^2(x_i, x_j) = \int (x_i^{(1)}(t) - x_j^{(1)}(t))^2 dt,$$

where $x^{(1)}$ denotes the first derivative of x. In practice, one can use a B-spline approximation for the curves as proposed in [6] to compute the above semi-metric.

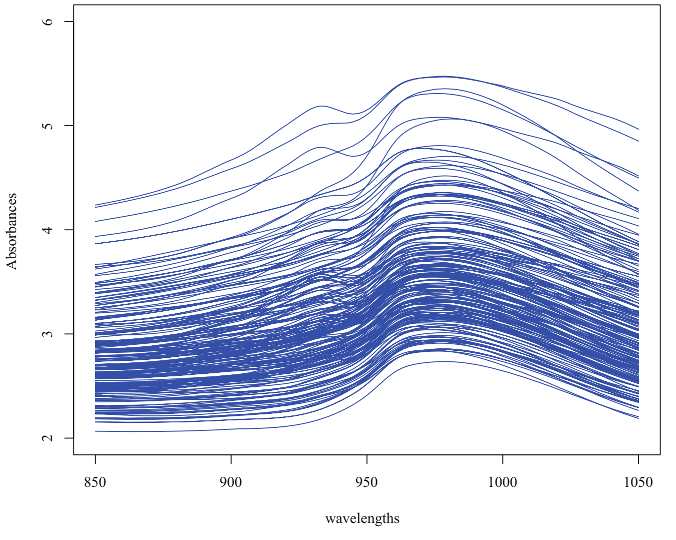


Figure 1 The spectrometric curves

To estimate the conditional tail index $\gamma(x)$ given $X = x$, we need to select the optimal parameters (K, r) and the intermediate quantile level τ_n. For this purpose, we first introduce the sample mean squared error $\mathrm{SMSE}\{\hat{\gamma}_n(\cdot)\} = n^{-1}\sum_{i=1}^n\{\hat{\gamma}_n(x_i) - \bar{\hat{\gamma}}\}^2$ for $\hat{\gamma}_n(x) = \hat{\gamma}_n^{\mathrm{P},1}(x)$, $\hat{\gamma}_n^{\mathrm{P},2}(x)$, and $\bar{\hat{\gamma}} = n^{-1}\sum_{i=1}^n \hat{\gamma}_n(x_i)$. We then select the tuning parameters (K, r) and τ_n for $\hat{\gamma}_n(x)$ on the principle of minimizing the SMSE of $\hat{\gamma}_n(x)$, with K ranging from $\mathbf{K} = \{3, 4\}$, r ranging from $\mathbf{R} = \{0.01, 0.14, 0.18, 0.33\}$, and τ_n ranging from $\mathcal{A} = \{0.1, 0.12, \ldots, 0.26\}$. Consequently, we yield by calculations that both $\hat{\gamma}_n^{\mathrm{P},1}(x)$ and $\hat{\gamma}_n^{\mathrm{P},2}(x)$ suggest the optimal values at $(K, r) = (3, 0.01)$ and $\tau_n = 0.22$. Moreover, the smoothing parameters g and h are selected using the cross-validation method as proposed in Section 3 with $\mathcal{G} = \{0.02, 0.06, \ldots, 1.98\}$ and $\mathcal{H} = \{0.02, 0.04, \ldots, 1\}$. Note that $\hat{\gamma}_n^{\mathrm{P},1}(x) = \hat{\gamma}_n^{\mathrm{P},2}(x)$ when $J = 3$, and then the two estimators $\hat{q}_n^{\mathrm{P},1}(\varphi_n|x)$ and $\hat{q}_n^{\mathrm{P},2}(\varphi_n|x)$ have the same performance in this case. Hence, we only need to adopt $\hat{\gamma}_n^{\mathrm{P},1}(x)$ and $\hat{q}_n^{\mathrm{P},1}(\varphi_n|x)$ to estimate the conditional tail index $\gamma(x)$ and the extreme conditional quantile $q(\varphi_n|x)$, respectively. For comparison, we also apply the conventional method to estimate $q(\varphi_n|x)$, where the conventional quantile estimator for $q(\varphi_n|x)$ is denoted by $\hat{q}_n(\varphi_n|x)$. Figure 2 gives the conditional tail index estimator $\hat{\gamma}_n^{\mathrm{P},1}(x_i)$ versus i with $i \in \{1, 2, \ldots, 215\}$ when $(K, r) = (3, 0.01)$ and $\tau_n = 0.22$. It appears that the estimated value of each $\gamma(x_i)$ is negative.

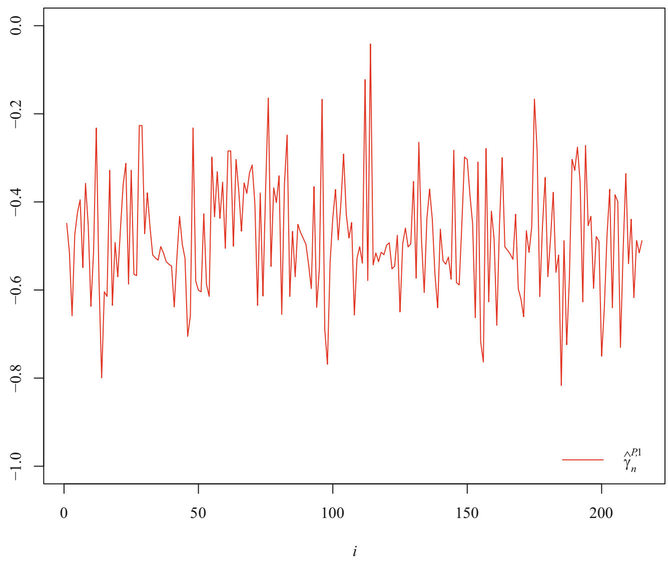


Figure 2 The estimated results of the conditional tail index

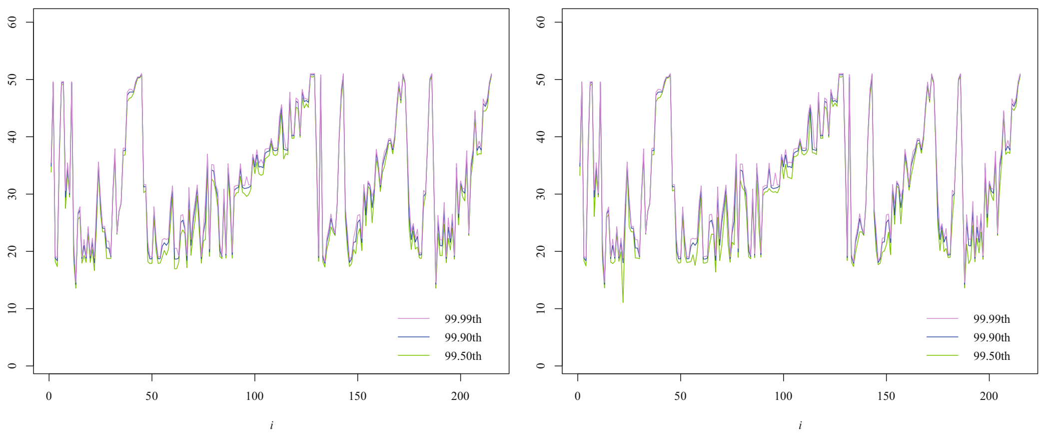
For the performance of $\hat{q}_n^{\mathrm{P},1}(\varphi_n|x)$ and $\hat{q}_n(\varphi_n|x)$, we consider the following measure:

$$\mathrm{SMSE}\{\tilde{q}_n(\varphi_n|\cdot)\} = \frac{1}{n}\sum_{i=1}^{n}\{\tilde{q}_n(\varphi_n|x_i) - \bar{\tilde{q}}_n(\varphi_n)\}^2,$$

where $\tilde{q}_n(\varphi_n|\cdot) = \hat{q}_n^{P,1}(\varphi_n|\cdot)$, $\hat{q}_n(\varphi_n|\cdot)$ and

$$\bar{\tilde{q}}_n(\varphi_n) = n^{-1}\sum_{i=1}^{n}\tilde{q}_n(\varphi_n|x_i).$$

For consistency, we select the same tuning parameters $(K, r) = (3, 0.01)$ as those suggested for evaluating $\hat{\gamma}_n^{\mathrm{P},1}(x)$ and $\hat{\gamma}_n^{\mathrm{P},2}(x)$. As for τ_n, we choose it by minimizing the SMSE of $\hat{q}_n^{\mathrm{P},1}(\varphi_n|\cdot)$ with τ_n ranging from $\mathcal{A} = \{0.1, 0.12, \ldots, 0.26\}$. The smoothing parameters g and h are selected using the cross-validation method as proposed in Section 3 with $\mathcal{G} = \{0.02, 0.06, \ldots, 1.98\}$ and $\mathcal{H} = \{0.02, 0.04, \ldots, 1\}$. Figure 3 shows the comparison between $\hat{q}_n^{P,1}(\varphi_n|x)$ and $\hat{q}_n(\varphi_n|x)$ for $\varphi_n \in \{0.005, 0.001, 0.0001\}$. Figure 3 graphically displays that the estimated quantiles regression curves for the conventional method and the proposed method, respectively, and they look the same. However, when $\varphi_n \in \{0.005, 0.001, 0.0001\}$, we obtain by calculations that the respective value of $\mathrm{SMSE}\{\hat{q}_n^{\mathrm{P},1}(\varphi_n|\cdot)\}$ is smaller than or equal to that of $\mathrm{SMSE}\{\hat{q}_n(\varphi_n|\cdot)\}$. Hence, the estimation results of $\hat{q}_n^{\mathrm{P},1}(\varphi_n|x)$ are either equally or more stable than that of $\hat{q}_n(\varphi_n|x)$ for each $\varphi_n \in \{0.005, 0.001, 0.0001\}$. In addition, the ranges of $\hat{q}_n^{\mathrm{P},1}(0.001|x_i)$-$\hat{q}_n^{\mathrm{P},1}(0.005|x_i)$ and $\hat{q}_n^{\mathrm{P},1}(0.0001|x_i)$-$\hat{q}_n^{\mathrm{P},1}(0.001|x_i)$ with $i = 1, 2, \ldots, 215$ are $[0.1036, 3.2395]$ and $[0.0468, 3.4397]$, respectively. And the ranges of $\hat{q}_n(0.001|x_i)$-$\hat{q}_n(0.005|x_i)$ and $\hat{q}_n(0.0001|x_i)$-$\hat{q}_n(0.001|x_i)$ with $i = 1, 2, \ldots, 215$ are $[0.1, 6.94]$ and $[0.04, 5.34]$, respectively. In summary, the estimated quantile curves from the proposed method and the conventional method are consistent and they do not raise any conflict with each other.



(a) Estimates by the extrapolation method (b) Estimates by the conventional method

Figure 3 The estimated quantile curves with $\varphi_n \in \{0.005, 0.001, 0.0001\}$

5 Proofs

To prove the theorems in Section 2.3, we first present some preliminary results. Lemmas 5.1 through 5.4 are similar as those in [1] and [8], respectively. Let $B(x, r)$ denote the ball of radius $r > 0$ with center at $x \in \mathcal{F}$, defined by $B(x, r) = \{t \in \mathcal{F} : d(t, x) \le r\}$. When n is taken very large, h is close to zero and then $B(x, h)$ is considered as a small ball and $\rho_x(h) = \mathbb{P}(X \in B(x, h))$ as the small ball probability of X.

Lemma 5.1 *Assume* (F1) *holds, and let* $\tau_n \to 0$ *as* $n \to \infty$*. Then for any* $m > 0$*,*

$$\lim_{n\to\infty} \frac{q(m\tau_n|x)}{q(\tau_n|x)} = m^{-(\gamma(x)\vee 0)}.$$

Lemma 5.2 *Assume* (F1) *holds. Let* $u_n(x) \uparrow y_F^*(x)$ *and an arbitrary sequence of functions* $\{t_n(x)\}$ *satisfy* $t_n(x) \to t_0(x)$ *as* $n \to \infty$ *where* $t_0(x)$ *is such that there exists* $\xi > 0$ *for which* $1 + \gamma(x)t_0(x) \le \xi$*. Then*

$$\lim_{n\to\infty} \frac{\bar{F}(u_n(x) + t_n(x)a(u_n(x)|x))}{\bar{F}(u_n(x)|x)} = \frac{1}{U^{-1}_{\gamma(x)}(t_0(x))}.$$

Lemma 5.3 *Assume* $\xi_n^{(\gamma)}(x) = \Delta_n^{-1}(\hat{\gamma}_n(x) - \gamma(x)) = O_{I\!P}(1)$ *where* $\Delta_n \to 0$*. Let* $c_n \ge 1$ *or* $c_n \le 1$ *such that* $\Delta_n \log(c_n) \to 0$*. Then*

$$\Delta_n^{-1}\left(\frac{U_{\hat{\gamma}_n(x)}(c_n) - U_{\gamma(x)}(c_n)}{U'_{\gamma_n(x)}}\right) = \xi_n^{(\gamma)}(x)(1 + o_{I\!P}(1)).$$

Lemma 5.4 *Assume* (F2) *holds. For all* $l > 0$ *and* $x \in \mathcal{F}$*, we have* $0 < C_1^l \rho_x(h) \le m_x^{(l)}(h) \le C_2^l \rho_x(h)$*.*

Lemma 5.4 provides a control on the moments $m_x^{(l)}(h)$ for all $l > 0$. Recall that the functional kernel estimator of $\bar{F}_Y(y|x)$ defined in (2.1) can be rewritten as $\hat{\bar{F}}_n(y|x) = \hat{\psi}_n(y, x)/\hat{g}_n(x)$, where

$$\hat{\psi}_n(y, x) = \frac{1}{m_x^{(1)}(h)} \sum_{i=1}^n \mathcal{K}(h^{-1}d(x, X_i))H(g^{-1}(Y_i - y)),$$

$$\hat{g}_n(x) = \frac{1}{m_x^{(1)}(h)} \sum_{i=1}^{n} \mathcal{K}(h^{-1}d(x, X_i)).$$

Lemma 5.5 *Assume* (F2) *holds, and let $x \in \mathcal{F}$ such that $\rho_x(h) > 0$. We have*

(1) $E(\hat{g}_n(x)) = 1$.

(2) *If $\rho_x(h) \to 0$ as $h \to 0$, then*

$$0 < \lim_{n\to\infty} n\rho_x(h)\mathrm{var}(\hat{g}_n(x)) < \infty.$$

Therefore, under (F2), if $\rho_x(h) \to 0$ and $n\rho_x(h) \to \infty$, then $\hat{g}_n(x) \xrightarrow{p} 1$.

Lemma 5.6 *Assume* (F3) *holds, and let g be an arbitrary small positive number. Then for any $y, z \in \mathbb{R}$,*

$$\mathbb{E}(H((Y-y)/g)H((Y-z)/g)|X) = \begin{cases} \mathbb{E}(H((Y-y\vee z)/g)|X), & y \neq z, \\ \mathbb{E}(H^2((Y-y)/g)|X), & y = z. \end{cases}$$

Proof When $y \neq z$, without loss of generality, we consider the case of $y < z$. Because

$$H((t-y)/g) = \begin{cases} 0, & t < y-g, \\ \displaystyle\int_{-1}^{(t-y)/g} \mathcal{K}_0(u)du, & y-g \le t < y+g, \\ 1, & t \ge y+g, \end{cases}$$

$$H((t-z)/g) = \begin{cases} 0, & t < z-g, \\ \displaystyle\int_{-1}^{(t-z)/g} \mathcal{K}_0(u)du, & z-g \le t < z+g, \\ 1, & t \ge z+g, \end{cases}$$

and $y + g < z - g$ in light of g being an arbitrary small positive-real number, we have

$$H((t-y)/g)H((t-z)/g) = \begin{cases} 0, & t < z-g, \\ \displaystyle\int_{-1}^{(t-z)/g} \mathcal{K}_0(u)du, & z-g \le t < z+g, \\ 1, & t \ge z+g. \end{cases}$$

It is clear that $H((t-y)/g)H((t-z)/g) = H((t-z)/g)$ for $y < z$ and $t \in \mathbb{R}$. Therefore, $\mathbb{E}(H((Y-y)/g)H((Y-z)/g)|X) = \mathbb{E}(H((Y-z)/g)|X)$. For symmetry, it follows that, for all $y \neq z$, $\mathbb{E}(H((Y-y)/g)H((Y-z)/g)|X) = \mathbb{E}(H((Y-y\vee z)/g)|X)$. Finally, when $y = z$, it is apparent that

$$\mathbb{E}(H((Y-y)/g)H((Y-z)/g)|X) = \mathbb{E}(H^2((Y-y)/g)|X). \qquad \square$$

Lemma 5.7 *Assume* (F1)–(F3) *hold, and let $x \in \mathcal{F}$ such that $\rho_x(h) > 0$. Let also $y_{n,k}(x) = y_n(x) + U_{\gamma(x)}(1/l_k)\ a(y_n(x)|x)(1 + o(1))$ with $0 < l_K < \cdots < l_2 < l_1 \le 1$, where K is a positive integer. If $y_n(x) \uparrow y_F^*(x)$ and $g \to 0$ such that $V_n(x) \to 0$, $a(y_n(x)|x)/g \to \infty$ and $n\bar{F}(y_n(x)|x)\rho_x(h) \to \infty$ as $n \to \infty$, then*

(1) $\mathbb{E}(\hat{\psi}_n(y_{n,k}(x), x)) = \bar{F}(y_{n,k}(x)|x)(1 + O(V_n(x)))$, *and*

(2) *the random vector*

$$\left\{\phi_n^{-1}(x)\left(\frac{\hat{\psi}_n(y_{n,k}(x),x)-\mathbb{E}(\hat{\psi}_n(y_{n,k}(x),x))}{\bar{F}(y_{n,k}(x)|x)}\right)\right\}_{k=1,\dots,K}$$

is asymptotically normal distributed with mean zero and covariance matrix $V(x)$, *where* $V_{k,k'}(x) = l_{k\wedge k'}^{-1}$ *for* $(k,k')\in\{1,\dots,K\}^2$.

Proof (1) Note that $\{(X_i,Y_i)\}_{i=1,\dots,n}$ are identically distributed. It follows that

$$\begin{aligned}\mathbb{E}(\hat{\psi}_n(y_{n,k}(x),x)) &= \frac{1}{m_x^{(1)}(h)}\mathbb{E}\{\mathcal{K}(h^{-1}d(x,X))H(g^{-1}(Y-y_{n,k}(x)))\}\\ &= \frac{1}{m_x^{(1)}(h)}\mathbb{E}\{\mathcal{K}(h^{-1}d(x,X))\mathbb{E}(H(g^{-1}(Y-y_{n,k}(x)))|X)\}.\end{aligned}$$

According to (F3), if $\tilde{F}_{n,k}(x,X)=\bar{F}(y_{n,k}(x)+gu|X)-\bar{F}(y_{n,k}(x)|x)$, we have

$$\begin{aligned}\mathbb{E}(H(g^{-1}(Y-y_{n,k}(x)))|X) &= \int_{\mathbb{R}} H(g^{-1}(y-y_{n,k}(x)))dF(y|X)\\ &= \int_{y_{n,k}(x)-g}^{y_{n,k}(x)+g}\int_{-1}^{g^{-1}(y-y_{n,k}(x))}\mathcal{K}_0(u)du dF(y|X)\\ &\quad+\int_{y_{n,k}(x)+g}^{\infty} dF(y|X)\\ &= \int_{-1}^{1}\mathcal{K}_0(u)du\int_{y_{n,k}(x)+gu}^{\infty} dF(y|X)\\ &= \int_{-1}^{1}\bar{F}(y_{n,k}(x)+gu|X)\mathcal{K}_0(u)du\\ &= \bar{F}(y_{n,k}(x)|x)+\int_{-1}^{1}\mathcal{K}_0(u)\tilde{F}_{n,k}(x,X)du.\end{aligned}$$

Consequently, $\mathbb{E}(\hat{\psi}_n(y_{n,k}(x),x))=\bar{F}(y_{n,k}(x)|x)+I_{1,n}+I_{2,n}$, where

$$\begin{aligned}I_{1,n} &= \frac{1}{m_x^{(1)}(h)}\mathbb{E}\{\mathcal{K}(h^{-1}d(x,X))(\bar{F}(y_{n,k}(x)|X)-\bar{F}(y_{n,k}(x)|x))\},\\ I_{2,n} &= \frac{1}{m_x^{(1)}(h)}\mathbb{E}\left\{\mathcal{K}(h^{-1}d(x,X))\bar{F}(y_{n,k}(x)|X)\int_{-1}^{1}\mathcal{K}_0(u)\left(\frac{\bar{F}(y_{n,k}(x)+gu|X)}{\bar{F}(y_{n,k}(x)|X)}-1\right)du\right\}.\end{aligned}$$

For $I_{1,n}$, by (F2) and Jensen's inequality, we have

$$\begin{aligned}|I_{1,n}| &= \frac{1}{m_x^{(1)}(h)}|\mathbb{E}\{\mathcal{K}(h^{-1}d(x,X))(\bar{F}(y_{n,k}(x)|X)-\bar{F}(y_{n,k}(x)|x))\mathbb{I}\{d(x,X)\le h\}\}|\\ &\le \frac{\bar{F}(y_{n,k}(x)|x)}{m_x^{(1)}(h)}\mathbb{E}\left\{\mathcal{K}(h^{-1}d(x,X))\left|\frac{\bar{F}(y_{n,k}(x)|X)}{\bar{F}(y_{n,k}(x)|x)}-1\right|\mathbb{I}\{d(x,X)\le h\}\right\}\\ &= \bar{F}(y_{n,k}(x)|x)O(V_n(x)). \end{aligned} \tag{5.1}$$

For $I_{2,n}$, let

$$\left|\frac{\bar{F}(y_{n,k}(x)+gu|X)}{\bar{F}(y_{n,k}(x)|X)}-1\right|\mathbb{I}\{d(x,X)\le h\}=\left|\frac{A_{n,k}(x,u|X)}{B_{n,k}(x|X)}-1\right|\mathbb{I}\{d(x,X)\le h\},$$

where

$$A_{n,k}(x,u|X)=\frac{\bar{F}(y_{n,k}(x)+gu|X)}{\bar{F}(y_{n,k}(x)|x)}\quad\text{and}\quad B_{n,k}(x|X)=\frac{\bar{F}(y_{n,k}(x)|X)}{\bar{F}(y_{n,k}(x)|x)}.$$

Noting that $|A_{n,k}(x,u|X)-1| \le V_n(x)$ and $|B_{n,k}(x|X)-1| \le V_n(x)$ for $d(x,X) \le h$ and $\forall u \in [-1,1]$, we have

$$\frac{1-V_n(x)}{1+V_n(x)} \le \frac{A_{n,k}(x,u|X)}{B_{n,k}(x|X)} \le \frac{1+V_n(x)}{1-V_n(x)},$$

for $d(x,X) \le h$ and $\forall u \in [-1,1]$. Furthermore, for $\forall u \in [-1,1]$,

$$\frac{-2V_n(x)}{1+V_n(x)} \le \left(\frac{A_{n,k}(x,u|X)}{B_{n,k}(x|X)} - 1\right)\mathbb{I}\{d(x,X) \le h\} \le \frac{2V_n(x)}{1-V_n(x)}.$$

This implies that

$$\left|\left(\frac{A_{n,k}(x,u|X)}{B_{n,k}(x|X)} - 1\right)\right|\mathbb{I}\{d(x,X) \le h\} \le \frac{2V_n(x)}{1-V_n(x)}.$$

Finally, by (5.1) and the fact that $V_n(x) \to 0$ as $n \to \infty$,

$$\begin{aligned}
|I_{2,n}| &\le \frac{2V_n(x)}{(1-V_n(x))m_x^{(1)}(h)}\mathbb{E}(\mathcal{K}(h^{-1}d(x,X))\bar{F}(y_{n,k}(x)|x)|X) \\
&= \frac{2V_n(x)}{(1-V_n(x))m_x^{(1)}(h)}(m_x^{(1)}(h)I_{1,n} + m_x^{(1)}(h)\bar{F}(y_{n,k}(x)|x)) \\
&= \frac{2V_n(x)}{1-V_n(x)}(I_{1,n} + \bar{F}(y_{n,k}(x)|x)) \\
&\le \frac{2V_n(x)}{1-V_n(x)}\bar{F}(y_{n,k}(x)|x)(1+O(V_n(x))) \\
&= \bar{F}(y_{n,k}(x)|x)O(V_n(x)).
\end{aligned}$$

This proves part (1).

(2) Let $\alpha = (\alpha_1, \ldots, \alpha_K)' \in \mathbb{R}$ and $\alpha \neq 0$,

$$\Psi_n = \sum_{k=1}^{K} \alpha_k \left(\frac{\hat{\psi}_n(y_{n,k}(x),x) - \mathbb{E}(\hat{\psi}_n(y_{n,k}(x),x))}{\phi_n(x)\bar{F}(y_{n,k}(x)|x)}\right) = \sum_{i=1}^{n} Z_{i,n},$$

where the random variable $Z_{i,n}$ is defined by

$$\begin{aligned}
n\phi_n(x)m_x^{(1)}(h)Z_{i,n} = &\sum_{k=1}^{K} \frac{\alpha_k \mathcal{K}(h^{-1}d(x,X_i))H(g^{-1}(Y_i - y_{n,k}(x)))}{\bar{F}(y_{n,k}(x)|x)} \\
&- \mathbb{E}\left(\sum_{k=1}^{K} \frac{\alpha_k \mathcal{K}(h^{-1}d(x,X_i))H(g^{-1}(Y_i - y_{n,k}(x)))}{\bar{F}(y_{n,k}(x)|x)}\right).
\end{aligned}$$

Obviously, $\{Z_{i,n}, i = 1, \ldots, n\}$ is a set of independent and identically distributed random variables with mean zero. In what follows, we study their asymptotic variance.

$$\begin{aligned}
\mathrm{var}(Z_{i,n}) &= \frac{1}{n^2(m_x^{(1)}(h))^2\phi_n^2(x)}\mathrm{var}\left(\sum_{k=1}^{K} \alpha_k \mathcal{K}(h^{-1}d(x,X_i))\frac{H(g^{-1}(Y_i - y_{n,k}(x)))}{\bar{F}(y_{n,k}(x)|x)}\right) \\
&= \frac{\alpha' B(x)\alpha}{n^2(m_x^{(1)}(h))^2\phi_n^2(x)} \\
&= \frac{\bar{F}(y_n(x)|x)}{nm_x^{(2)}(h)}\alpha' B(x)\alpha,
\end{aligned}$$

where $B(x)$ is a $K \times K$ covariance matrix with coefficients defined for $(k, k') \in \{1, \dots, K\}^2$ by

$$
\begin{aligned}
B_{k,k'}(x) &= \frac{A_{k,k'}(x)}{\bar{F}(y_{n,k}(x)|x)\bar{F}(y_{n,k'}(x)|x)}, \\
A_{k,k'}(x) &= \mathrm{cov}\{\mathcal{K}(h^{-1}d(x,X))H(g^{-1}(Y - y_{n,k}(x))), \mathcal{K}(h^{-1}d(x,X))H(g^{-1}(Y - y_{n,k'}(x)))\} \\
&= I_{3,n} - I_{4,n},
\end{aligned}
$$

where

$$
\begin{aligned}
I_{3,n} &= \mathbb{E}\{\mathcal{K}^2(h^{-1}d(x,X))H(g^{-1}(Y - y_{n,k}(x)))H(g^{-1}(Y - y_{n,k'}(x)))\}, \\
I_{4,n} &= \mathbb{E}\{\mathcal{K}(h^{-1}d(x,X))H(g^{-1}(Y - y_{n,k}(x)))\}\mathbb{E}\{\mathcal{K}(h^{-1}d(x,X))H(g^{-1}(Y - y_{n,k'}(x)))\}.
\end{aligned}
$$

For $I_{3,n}$ with $k \neq k'$, by Lemma 5.6 we have

$$
\begin{aligned}
I_{3,n} &= \mathbb{E}\{\mathcal{K}^2(h^{-1}d(x,X))\mathbb{E}(H(g^{-1}(Y - y_{n,k\vee k'}(x)))|X)\} \\
&= \mathbb{E}\{\mathcal{K}^2(h^{-1}d(x,X))H(g^{-1}(Y - y_{n,k\vee k'}(x)))\}.
\end{aligned}
$$

Noting also that $\mathcal{K}^2(\cdot)$ is a kernel function satisfying the assumption (F2), the proof of part (1) implies that

$$
I_{3,n} = m_x^{(2)}(h)\bar{F}(y_{n,k\vee k'}(x)|x)(1 + O(V_n(x))), \tag{5.2}
$$

for all $k \neq k'$. For the case of $k = k'$, by definition

$$
I_{3,n} = \mathbb{E}\{\mathcal{K}^2(h^{-1}d(x,X))H^2(g^{-1}(Y - y_{n,k}(x)))\},
$$

where $\mathcal{K}^2(\cdot)$ is a kernel function satisfying the assumption (F2) and the probability density function of $H^2(\cdot)$ satisfies (F3). Therefore, (5.2) also holds for $k = k'$. Second, the proof of part (1) implies that

$$
I_{4,n} = (m_x^{(1)}(h))^2\bar{F}(y_{n,k}(x)|x)\bar{F}(y_{n,k'}(x)|x)(1 + O(V_n(x))).
$$

As a result,

$$
\begin{aligned}
A_{k,k'}(x) &= m_x^{(2)}(h)\bar{F}(y_{n,k\vee k'}(x)|x)(1 + O(V_n(x))) \\
&\quad - (m_x^{(1)}(h))^2\bar{F}(y_{n,k}(x)|x)\bar{F}(y_{n,k'}(x)|x)(1 + O(V_n(x))).
\end{aligned}
$$

This leads to

$$
\begin{aligned}
B_{k,k'}(x) &= \frac{m_x^{(2)}(h)}{\bar{F}(y_{n,k\wedge k'}(x)|x)}(1 + O(V_n(x))) - (m_x^{(1)}(h))^2(1 + O(V_n(x))) \\
&= \frac{m_x^{(2)}(h)}{\bar{F}(y_{n,k\wedge k'}(x)|x)}\left(1 + O(V_n(x)) - \frac{(m_x^{(1)}(h))^2}{m_x^{(2)}(h)}\bar{F}(y_{n,k\wedge k'}(x)|x)(1 + O(V_n(x)))\right).
\end{aligned}
$$

By Lemma 5.4, $(m_x^{(1)}(h))^2/m_x^{(2)}(h)$ is bounded, and also by $\bar{F}(y_{n,k\wedge k'}(x)|x) \to 0$ as $n \to \infty$, we get

$$
B_{k,k'}(x) = \frac{m_x^{(2)}(h)}{\bar{F}(y_{n,k\wedge k'}(x)|x)}(1 + o(1)).
$$

Additionally, in light of Lemma 5.2,

$$
\lim_{n\to\infty} \frac{\bar{F}(y_{n,k\wedge k'}(x)|x)}{\bar{F}(y_n(x)|x)} = \frac{1}{U_{\gamma(x)}^{-1}(U_{\gamma(x)}(1/l_{k\wedge k'}))} = l_{k\wedge k'}
$$

entails

$$B_{k,k'}(x) = \frac{m_x^{(2)}(h)}{\bar{F}(y_n(x)|x)} l_{k\wedge k'}^{-1}(1+o(1)) = \frac{m_x^{(2)}(h)V_{k,k'}}{\bar{F}(y_n(x)|x)}(1+o(1)).$$

This results in $\mathrm{var}(Z_{i,n}) = \alpha' V(x)\alpha(1+o(1))/n$ for all $i = 1,\dots,n$. As a result, $\mathrm{var}(\Psi_n) \to \alpha' V\alpha$ as $n \to \infty$. Hence, Lyapunov criteria for the asymptotic normality of sums of Ψ_n reduces to $\sum_{i=1}^n \mathbb{E}|Z_{i,n}|^3 = n\mathbb{E}|Z_{1,n}|^3 \to 0$ as $n \to \infty$. Note that $Z_{1,n}$ is a bounded random variable, because

$$\begin{aligned}
|Z_{1,n}| &\le \frac{2C_2\sum_{k=1}^K |\alpha_k|}{n\phi_n(x)m_x^{(1)}(h)\bar{F}(y_{n,K}(x)|x)} \\
&= \frac{2C_2\sum_{k=1}^K |\alpha_k|(1+o(1))}{nl_K\phi_n(x)m_x^{(1)}(h)\bar{F}(y_{n,K}(x)|x)} \\
&= 2C_2 l_K^{-1}\frac{m_x^{(1)}(h)}{m_x^{(2)}(h)}\sum_{k=1}^K |\alpha_k|\phi_n(x)(1+o(1)) \\
&\le 2(C_2/C_1)^2 l_K^{-1}\sum_{k=1}^K |\alpha_k|\phi_n(x)(1+o(1)),
\end{aligned}$$

in light of Lemma 5.4. Thus,

$$\begin{aligned}
n\mathbb{E}|Z_{1,n}|^3 &\le 2(C_2/C_1)^2 l_K^{-1}\sum_{k=1}^K |\alpha_k|\phi_n(x)n\mathrm{var}(Z_{1,n})(1+o(1)) \\
&= 2(C_2/C_1)^2 l_K^{-1}\sum_{k=1}^K |\alpha_k|\alpha' V(x)\alpha\phi_n(x)(1+o(1)) \to 0
\end{aligned}$$

as $n \to \infty$. In addition, $V(x)$ is a positive definitive matrix, and therefore

$$B_n^3 = \left[\sum_{i=1}^n \mathrm{var}(Z_{i,n})\right]^{3/2} = [\mathrm{var}(\Psi_n)]^{3/2} = (\alpha' V\alpha(1+o(1)))^{3/2} \to (\alpha' V\alpha)^{3/2} > 0.$$

Accordingly, we can conclude that Ψ_n converges in distribution to a normal random variable with mean zero and variance $\alpha' V(x)\alpha$ for all $\alpha \neq 0$ in $\mathbb{R}^K$. □

5.1 Proof of Theorem 2.1

Following Lemma 5.7, we have the following expansion:

$$\phi_n^{-1}(x)\sum_{k=1}^K \beta_k\left(\frac{\hat{\bar{F}}_n(y_{n,k}(x)|x)}{\bar{F}(y_{n,k}(x)|x)} - 1\right) = \frac{\Lambda_{1,n}+\Lambda_{2,n}+\Lambda_{3,n}}{\hat{g}_n(x)},$$

where

$$\begin{aligned}
\Lambda_{1,n} &= \phi_n^{-1}(x)\sum_{k=1}^K \beta_k\left(\frac{\hat{\psi}_n(y_{n,k}(x),x) - \mathbb{E}(\hat{\psi}_n(y_{n,k}(x),x))}{\bar{F}(y_{n,k}(x)|x)}\right), \\
\Lambda_{2,n} &= \phi_n^{-1}(x)\sum_{k=1}^K \beta_k\left(\frac{\mathbb{E}(\hat{\psi}_n(y_{n,k}(x),x)) - \bar{F}(y_{n,k}(x)|x)}{\bar{F}(y_{n,k}(x)|x)}\right), \\
\Lambda_{3,n} &= \phi_n^{-1}(x)\left(\sum_{k=1}^K \beta_k\right)(\hat{g}_n(x) - 1).
\end{aligned}$$

$\Lambda_{1,n}$ can be rewritten as $\Lambda_{1,n} = \sqrt{\alpha' V \alpha}\eta_n$, where η_n converges in distribution to a standard random variable. Also, $\Lambda_{2,n} = O(\phi_n^{-1}(x)V_n(x)) = o(1)$. For $\Lambda_{3,n}$, using Chebyshev's inequality, for any $\epsilon > 0$ we have

$$\begin{aligned}\mathbb{P}\{\phi_n^{-1}(x)|\hat{g}_n(x) - 1| \geq \epsilon\} &\leq \frac{\phi_n^{-2}(x)\mathrm{var}(\hat{g}_n(x))}{\epsilon^2} \\ &= \frac{n\varrho_x(h)\mathrm{var}(\hat{g}_n(x))}{\epsilon^2}[(m_x^{(1)}(h))^2/(\varrho_x(h)m_x^{(2)}(h))]\bar{F}(y_n(x)|x) \\ &\to 0,\end{aligned}$$

as $y_n(x) \uparrow y_F^*(x)$ and where $0 < (C_1/C_2)^2 \leq (m_x^{(1)}(h))^2/(\varrho_x(h)m_x^{(2)}(h)) \leq (C_2/C_1)^2$. Therefore, $\Lambda_{3,n} = o_{\mathbb{P}}(1)$. We can conclude that

$$\hat{g}_n(x)\phi_n^{-1}(x)\sum_{k=1}^{K}\beta_k\left(\frac{\hat{\bar{F}}_n(y_{n,k}(x)|x)}{\bar{F}(y_{n,k}(x)|x)} - 1\right) = \sqrt{\alpha' V \alpha}\eta_n + o_{\mathbb{P}}(1).$$

Finally, as $\hat{g}_n(x) \xrightarrow{p} 1$, we have Theorem 2.1.

5.2 Proof of Theorem 2.2

For $z_k \in \mathbb{R}$ and $k = 1, 2, \ldots, K$, we define

$$\begin{aligned}&v_n = (n\tau_n^{-1}(m_x^{(1)}(h))^2/m_x^{(2)}(h))^{1/2}, \\ &\sigma_n(x) = (v_n f(q(\tau_n|x)|x))^{-1}, \\ &W_{n,k}(x) = v_n(\hat{\bar{F}}_n(q(l_k\tau_n|x) + \sigma_n(x)|x) - \bar{F}(q(l_k\tau_n|x) + \sigma_n(x)|x)), \\ &a_{n,k}(x) = v_n(l_k\tau_n - \bar{F}(q(l_k\tau_n|x) + \sigma_n(x)|x)).\end{aligned}$$

We study the asymptotic behavior of K-variate function defined by

$$\begin{aligned}\mathbb{F}_n(z_1, \ldots, z_K) &= \mathbb{P}\bigg(\bigcap_{k=1}^{K}\{\sigma_n^{-1}(x)(\hat{q}_n(l_k\tau_n|x) - q(l_k\tau_n|x)) \leq z_k\}\bigg) \\ &= \mathbb{P}\bigg(\bigcap_{k=1}^{K}\{W_{n,k}(x) \leq a_{n,k}(x)\}\bigg).\end{aligned}$$

For the non-random term $a_{n,k}(x)$ with $k \in \{1, \ldots, K\}$, there exists $\theta_{n,k} \in (0, 1)$ such that

$$a_{n,k}(x) = v_n\sigma_n(x)z_k f(q(\tau_n|x)|x) = z_k\frac{f(q_{n,k}(x)|x)}{f(q(\tau_n|x)|x)},$$

where

$$\begin{aligned}q_{n,k}(x) &= q(l_k\tau_n|x) + \theta_{n,k}\sigma_n(x)z_k \\ &= q(l_k\tau_n|x) + \theta_{n,k}\frac{z_k}{l_k}(n\tau_n(m_x^{(1)}(h))^2/m_x^{(2)}(h))^{-1/2}\frac{l_k\tau_n}{f(q(l_k\tau_n|x)|x)} \cdot \frac{f(q(l_k\tau_n|x)|x)}{f(q(\tau_n|x)|x)} \\ &= q(l_k\tau_n|x) + \theta_{n,k}z_k l_k^{\gamma(x)}(n\tau_n(m_x^{(1)}(h))^2/m_x^{(2)}(h))^{-1/2}\frac{l_k\tau_n}{f(q(l_k\tau_n|x)|x)}(1 + o(1)),\end{aligned}$$

since $f(q(\cdot|x)|x)$ is regularly varying at 0 with index $1 + \gamma(x)$; see (1.1.33) of Corollary 1.1.10 in [5] for detail. Moreover, in light of Theorem 1.2.6 and Remark 1.2.7 in [5], a possible choice of the auxiliary function $a(\cdot|x)$ is

$$a(t|x) = \frac{\bar{F}(t|x)}{f(t|x)}(1 + o(1)),$$

conducing to

$$q_{n,k}(x)=q(l_k\tau_n|x)+\theta_{n,k}z_kl_k^{\gamma(x)}(n\tau_n(m_x^{(1)}(h))^2/m_x^{(2)}(h))^{-1/2}a(q(l_k\tau_n|x)|x)(1+o(1)).$$

Using Lemma 5.2 with $t_n(x)=\theta_{n,k}z_kl_k^{\gamma(x)}(n\tau_n(m_x^{(1)}(h))^2/m_x^{(2)}(h))^{-1/2}(1+o(1))$, $u_n(x)=q(l_k\tau_n|x)$ and $t_0(x)=0$, we get

$$\frac{\bar{F}(q_{n,k}(x)|x)}{l_k\tau_n}\to\frac{1}{U_{\gamma(x)}^{-1}(0)}=1\quad\text{ as } n\to\infty.$$

Recalling that $f(q(\cdot|x)|x)$ is regularly varying, we have

$$\frac{f(q_{n,k}(x)|x)}{f(q(\tau_n|x)|x)}\to l_k^{1+\gamma(x)}\quad\text{ as } n\to\infty.$$

Consequently,

$$a_{n,k}(x)=z_kl_k^{1+\gamma(x)}(1+o(1)),\quad k=1,\ldots,K. \tag{5.3}$$

In what follows, we examine the random term $W_{n,k}(x)$. Let $z_{n,k}(x)=q(l_k\tau_n|x)+\sigma_n(x)z_k$ for $k=1,\ldots,K$, $y_n(x)=q(\tau_n|x)$, and consider the expansion

$$\frac{z_{n,k}(x)-y_n(x)}{a(y_n(x)|x)}=\frac{q(l_k\tau_n|x)-q(\tau_n|x)}{a(q(\tau_n|x)|x)}+\frac{\sigma_n(x)z_k}{a(q(\tau_n|x)|x)}.$$

According to (2.5), we have

$$\lim_{n\to\infty}\frac{q(l_k\tau_n|x)-q(\tau_n|x)}{a(q(\tau_n|x)|x)}=U_{\gamma(x)}(1/l_k),$$

and

$$\begin{aligned}\lim_{n\to\infty}\frac{\sigma_n(x)z_k}{a(q(\tau_n|x)|x)}&=\lim_{n\to\infty}(n\tau_n(m_x^{(1)}(h))^2/m_x^{(2)}(h))^{-1/2}\\&=0,\end{aligned}$$

giving rise to $z_{n,k}(x)=y_n(x)+U_{\gamma(x)}(1/l_k)a(y_n(x)|x)|x)(1+o(1))$. By Theorem 2.1 and Lemma 5.2, it yields

$$\begin{aligned}W_{n,k}(x)&=\frac{\bar{F}(y_n(x)+U_{\gamma(x)}(1/l_k)a(y_n(x)|x)|x)(1+o(1))|x)}{\tau_n}\xi_{n,k}\\&=l_k\xi_{n,k}(1+o(1)),\end{aligned}$$

where $\xi_n=(\xi_{n,1},\ldots,\xi_{n,K})'$ converges to a normal random vector with mean zero and covariance matrix $V(x)$. Finally, by (5.3), we conclude Theorem 2.2.

5.3 Proof of Theorem 2.4

By definition, we have

$$\begin{aligned}q(\varphi_n|x)&=q(\tau_n|x)+(U_{\gamma(x)}(\tau_n/\varphi_n)+d(\varphi_n/\tau_n,\tau_n|x))a(q(\tau_n|x)|x),\\\hat{q}_n(\varphi_n|x)&=\hat{q}(\tau_n|x)+U_{\hat{\gamma}_n(x)}(\tau_n/\varphi_n)\hat{a}_n(x),\end{aligned}$$

where $\hat{\gamma}_n(x)$ and $\hat{a}_n(x)$ are two estimators of $\gamma(x)$ and $a(q(\tau_n|x)|x)$, respectively. We have the following expansion:

$$\delta_n^{-1}(x)\left(\frac{\hat{q}_n(\varphi_n|x)-q(\varphi_n|x)}{a(q(\tau_n|x)|x)U'_{\gamma(x)}(\tau_n/\varphi_n)}\right)=T_{n,1}+T_{n,2}+T_{n,3}+T_{n,4},$$

where

$$T_{n,1} = \delta_n^{-1}(x)\Big(\frac{\hat{q}_n(\tau_n|x) - q(\tau_n|x)}{a(q(\tau_n|x)|x)U'_{\gamma(x)}(\tau_n/\varphi_n)}\Big),$$
$$T_{n,2} = \delta_n^{-1}(x)\Big(\frac{U_{\hat{\gamma}_n(x)}(\tau_n/\varphi_n) - U_{\gamma(x)}(\tau_n/\varphi_n)}{U'_{\gamma(x)}(\tau_n/\varphi_n)}\Big)\frac{\hat{a}_n(x)}{a(q(\tau_n|x)|x)},$$
$$T_{n,3} = \delta_n^{-1}(x)\frac{U_{\gamma(x)}(\tau_n/\varphi_n)}{U'_{\gamma(x)}(\tau_n/\varphi_n)}\Big(\frac{\hat{a}_n(x)}{a(q(\tau_n|x)|x)} - 1\Big),$$
$$T_{n,3} = -\delta_n^{-1}(x)\frac{d(\varphi_n/\tau_n, \tau_n|x)}{U'_{\gamma(x)}(\tau_n/\varphi_n)}.$$

For ease of notation, let

$$(\xi_n^{(\gamma)}(x), \xi_n^{(a)}(x), \xi_n^{(q)}(x)) = \delta_n^{-1}(x)\Big(\hat{\gamma}_n(x) - \gamma(x), \frac{\hat{a}_n(x)}{a(q(\tau_n|x)|x)} - 1, \frac{\hat{q}_n(\tau_n|x) - q(\tau_n|x)}{a(q(\tau_n|x)|x)}\Big).$$

Note also that, when $v \to \infty$,

$$U_z(v) = \int_1^v u^{z-1}du = (1 + o(1))\begin{cases} -\frac{1}{z} & \text{if } z < 0, \\ \log(v) & \text{if } z = 0, \\ \frac{v^z}{z} & \text{if } z > 0, \end{cases}$$

and

$$U'_z(v) = \int_1^v u^{z-1}\log(u)du = (1 + o(1))\begin{cases} \frac{1}{z^2} & \text{if } z < 0, \\ \frac{\log^2(v)}{2} & \text{if } z = 0, \\ \frac{u^z\log(v)}{z} & \text{if } z > 0. \end{cases}$$

By Lemma 5.3 and the assumptions, we conclude Theorem 2.4 by noting that

$$T_{n,1} = \frac{\xi_n^{(q)}(x)}{U'_{\gamma(x)}(\frac{\tau_n}{\varphi_n})} = (\gamma(x) \wedge 0)^2\xi_n^{(q)}(x)(1 + o_{I\!P}(1)),$$
$$T_{n,2} = \delta_n^{-1}(x)\Big(\frac{U_{\hat{\gamma}_n(x)}(\tau_n/\varphi_n) - U_{\gamma(x)}(\tau_n/\varphi_n)}{U'_{\gamma(x)}(\tau_n/\varphi_n)}\Big)(1 + o_{I\!P}(1)) = \xi_n^{(\gamma)}(x)(1 + o_{I\!P}(1)),$$
$$T_{n,3} = \frac{U_\gamma(\tau_n/\varphi_n)}{U'_\gamma(\tau_n/\varphi_n)}\xi_n^{(a)}(x) = -(\gamma(x) \wedge 0)\xi_n^{(a)}(x)(1 + o_{I\!P}(1)),$$
$$T_{n,4} = o_{I\!P}(1).$$

5.4 Proof of Theorem 2.5

(i) For ease of notation, we define

$$\gamma_{n,k}(x) = \frac{1}{\log(r)}\log\Big(\frac{\hat{q}_n(l_k\tau_n|x) - \hat{q}_n(l_{k+1}\tau_n|x)}{\hat{q}_n(l_{k+1}\tau_n|x) - \hat{q}_n(l_{k+2}\tau_n|x)}\Big). \tag{5.4}$$

Then, $\hat{\gamma}_n^P(x) = \sum_{k=1}^{K-2}\omega_k\gamma_{n,k}(x)$. By Theorem 2.2 and (2.5), for all $k = 1, \ldots, K$,

$$\hat{q}_n(l_k\tau_n|x) = q(\tau_n|x) + a(q(\tau_n|x)|x)(U_{\gamma(x)}(1/l_k) + d(l_k, \tau_n|x)) + \sigma_n(x)\xi_{k,n}, \tag{5.5}$$

where $\sigma_n^{-1}(x) = f(q(\tau_n|x)|x)\sqrt{n\tau_n^{-1}(m_x^{(1)}(h))^2/m_x^{(2)}(h)}$ and $\xi_n = \{\xi_{k,n}\}_{k=1,\ldots,K}$ is asymptotically normal distributed with mean zero and covariance matrix $C(x)$. Let

$$\eta_n(x) = \max_{k=1,\ldots,K} |d(l_k, \tau_n|x)| \quad \text{and} \quad \epsilon_n = \frac{\sigma_n(x)}{a(q(\tau_n|x)|x)} = \delta_n(x)(1+o(1)).$$

We have

$$\begin{aligned}\frac{\hat{q}_n(l_k\tau_n|x) - \hat{q}_n(l_{k+1}\tau_n|x)}{a(q(\tau_n|x)|x)} &= U_{\gamma(x)}(1/l_k) + d(l_k, \tau_n) + \frac{\sigma_n(x)\xi_{k,n}}{a(q(\tau_n|x)|x)} - U_{\gamma(x)}(1/l_{k+1}) \\ &\quad - d(l_{k+1}, \tau_n) - \frac{\sigma_n(x)\xi_{k+1,n}}{a(q(\tau_n|x)|x)} \\ &= \epsilon_n(\xi_{k,n} - \xi_{k+1,n}) + U_{\gamma(x)}(r)r^{-k\gamma(x)} + O(\eta_n(x)). \end{aligned} \tag{5.6}$$

By (5.4), we can obtain that

$$\begin{aligned}\log(r)(\gamma_{n,k}(x) - \gamma(x)) &= \log\left(1 + \frac{\epsilon_n(\xi_{k,n} - \xi_{k+1,n})r^{\gamma(x)k}}{U_{\gamma(x)}(r)} + O(\eta_n(x))\right) \\ &\quad - \log\left(1 + \frac{\epsilon_n(\xi_{k+1,n} - \xi_{k+2,n})r^{\gamma(x)(k+1)}}{U_{\gamma(x)}(r)} + O(\eta_n(x))\right).\end{aligned}$$

By the first order Taylor expansion, we have

$$\begin{aligned}\log(r)\epsilon_n^{-1}(\gamma_{n,k}(x) - \gamma(x)) &= \frac{r^{\gamma(x)k}}{U_{\gamma(x)}(r)}(\xi_{k,n} - (1 + r^{\gamma(x)})\xi_{k+1,n} + r^{\gamma(x)}\xi_{k+2,n}) + O(\epsilon_n^{-1}\eta_n(x)) \\ &\quad + o_{I\!P}(1),\end{aligned}$$

and thus, under the assumption $\delta_n^{-1}(x)\eta_n(x) \to 0$ as $n \to \infty$,

$$\begin{aligned}\delta_n^{-1}(x)(\hat{\gamma}_n^{\mathrm{P}}(x) - \gamma(x)) &= \frac{1}{\log(r)U_{\gamma(x)}(r)} \sum_{k=1}^{K-2} \omega_k r^{\gamma(x)k}(\xi_{k,n} - (1 + r^{\gamma(x)})\xi_{k+1,n} + r^{\gamma(x)}\xi_{k+2,n}) \\ &\quad + o_{I\!P}(1).\end{aligned}$$

Let $\omega_{-1} = \omega_0 = \omega_{K-1} = \omega_K = 0$, $\beta_0^{(\gamma)} = 1/\log(r)$, $\beta_1^{(\gamma)} = -(1 + r^{-\gamma(x)})/\log(r)$, and $\beta_2^{(\gamma)} = r^{-\gamma(x)}/\log(r)$. Then,

$$\xi_n^{(\gamma)}(x) = \frac{1}{U_{\gamma(x)}(r)} \sum_{k=1}^{K} r^{\gamma(x)k}(\beta_0^{(\gamma)}\omega_k + \beta_1^{(\gamma)}\omega_{k-1} + \beta_2^{(\gamma)}\omega_{k-2})\xi_{k,n} + o_{I\!P}(1). \tag{5.7}$$

(ii) We define

$$a_{n,k}(x) = \frac{r^{\hat{\gamma}_n^{\mathrm{P}}(x)k}(\hat{q}_n(l_k\tau_n|x) - \hat{q}_n(l_{k+1}\tau_n|x))}{U_{\hat{\gamma}_n^{\mathrm{P}}(x)}(r)}.$$

Then, $\hat{a}_n^{\mathrm{P}}(x) = \sum_{k=1}^{K-2} \omega_k a_{n,k}(x)$. By (5.6), for all $j = 1, \ldots, K-2$,

$$\frac{a_{n,k}(x)}{a(q(\tau_n|x)|x)} = \frac{r^{\hat{\gamma}_n^P(x)k}}{U_{\hat{\gamma}_n^P(x)}(r)}[\epsilon_n(\xi_{k,n} - \xi_{k+1,n}) + U_{\gamma(x)}(r)r^{-k\gamma(x)} + O(\eta_n(x))].$$

Also by $\hat{\gamma}_n^{\mathrm{P}}(x) = \gamma(x) + \delta_n(x)\xi_n^{\gamma}(x)$ and Lemma 5.3, we have

$$\delta_n^{-1}(x)(U_{\hat{\gamma}_n^P(x)}(r) - U_{\gamma(x)}(r))/U'_{\gamma(x)}(r) = \xi_n^{\gamma}(x)(1 + o_{I\!P}(1)).$$

Accordingly,

$$\frac{U_{\hat{\gamma}_n^{\mathrm{P}}(x)}(r)}{U_{\gamma(x)}(r)} = 1 + \frac{U'_{\gamma(x)}(r)}{U_{\gamma(x)}(r)}\delta_n(x)\xi_n^{(\gamma)}(x)(1 + o_{I\!P}(1)),$$

and

$$r^{(\hat{\gamma}_n^{\mathrm{P}}(x)-\gamma(x))k} = r^{\delta_n(x)k\xi_n^{(\gamma)}(x)} = \exp[\delta_n(x)k\xi_n^{(\gamma)}(x)\log(r)].$$

It then yields that

$$\frac{a_{n,k}(x)}{a(q(\tau_n|x)|x)} = \exp[\delta_n(x)k\xi_n^{(\gamma)}(x)\log(r)]\frac{1 + (r^{k\gamma(x)}/U_{\gamma(x)}(x))(\xi_{k,n} - \xi_{k+1,n})\epsilon_n + O(\eta_n(x))}{1 + (U'_{\gamma(x)}(r)/U_{\gamma(x)}(r))\delta_n(x)\xi_n^{(\gamma)}(x)(1 + o_{I\!P}(1))}.$$

Furthermore, by the first order Taylor expansion, we have

$$\begin{aligned}\delta_n^{-1}(x)\left(\frac{a_{n,k}(x)}{a(q(\tau_n|x)|x)} - 1\right) &= \frac{r^{\gamma(x)k}}{U_{\gamma(x)}(r)}(\xi_{k,n} - \xi_{k+1,n}) + \xi_n^{(\gamma)}(x)\left(k\log(r) - \frac{U'_{\gamma(x)}(r)}{U_{\gamma(x)}(r)}\right)\\ &\quad + O(\delta_n^{-1}(x)\eta_n(x)) + o_{I\!P}(1).\end{aligned}$$

Let $E(\omega) = \sum_{k=1}^{K} k\omega_k$, $c_0 = \log(r)E(\omega) - U'_{\gamma(x)}(r)/U_{\gamma(x)}(r)$, $\beta_0^{(a)} = c_0\beta_0^{(\gamma)} + 1$, $\beta_1^{(a)} = c_0\beta_1^{(\gamma)} - r^{-\gamma(x)}$, and $\beta_2^{(a)} = c_0\beta_2^{(\gamma)}$. Noting that $\omega_{-1} = \omega_0 = \omega_{K-1} = \omega_K = 0$, we have

$$\begin{aligned}\xi_n^{(a)}(x) &= \delta_n^{-1}(x)\left(\frac{\hat{a}_n^P(x)}{a(q(\tau_n|x)|x)} - 1\right)\\ &= \frac{1}{U_{\gamma(x)}(r)}\sum_{k=1}^{k} r^{\gamma(x)K}[(c_0\beta_0^{(\gamma)} + 1)\omega_k + (c_0\beta_1^{(\gamma)} - r^{-\gamma})\omega_{k-1} + c_0\beta_2^{(\gamma)}\omega_{k-2}]\xi_{k,n} + o_{I\!P}(1)\\ &= \frac{1}{U_{\gamma(x)}(r)}\sum_{k=1}^{K} r^{\gamma(x)k}(\beta_0^{(a)}\omega_k + \beta_1^{(a)}\omega_{k-1} + \beta_2^{(a)}\omega_{k-2})\xi_{k,n} + o_{I\!P}(1). \qquad (5.8)\end{aligned}$$

(iii) Letting $k = 1$ in (5.5), we have $l_1 = 1$ and $\xi_{1,n} = \sigma_n^{-1}(x)(\hat{q}_n(\tau_n|x) - q(\tau_n|x))$. It follows hereby that

$$(1 + o(1))\xi_{1,n} = \delta_n^{-1}(x)\frac{(\hat{q}_n(\tau_n|x) - q(\tau_n|x))}{a(q(\tau_n|x)|x)}. \qquad (5.9)$$

Finally, by (5.7), (5.8) and (5.9), we have

$$(\xi_n^{\gamma}(x), \xi_n^{(a)}(x), \xi_n^{(q)}(x))' = \frac{1}{U_{\gamma(x)}(r)}A(x)\xi_n + (o_{I\!P}(1), o_{I\!P}(1), o_{I\!P}(1))',$$

where $A(x)$ is a $3 \times K$ matrix defined in Theorem 2.5. Consequently, the random vector $(\xi_n^{\gamma}(x), \xi_n^{(a)}(x), \xi_n^{(q)}(x))'$ is asymptotically normal with mean zero and covariance matrix $\mathcal{C}(x) = A(x)C(x)A'(x)/(U_{\gamma(x)}(r))^2$.

Acknowledgements The authors thank the editor, the associate editor and two referees for their constructive comments that have led to a substantial improvement of the paper.

References

[1] Daouia, A., Gardes, L., Girard, S.: On kernel smoothing for extremal quantile regression. *Bernoulli*, **19**, 2557–2589 (2013)

[2] Daouia, A., Gardes, L., Girard, S., et al.: Kernel estimators of extreme level curves. *TEST*, **20**, 311–333 (2011)

[3] Davison, A. C., Ramesh, N. I.: Local likelihood smoothing of sample extremes. *J. R. Stat. Soc. Ser. B*, **62**, 191–208 (2000)

[4] Davison, A. C., Smith, R. L.: Models for exceedances over high thresholds. *J. R. Stat. Soc. Ser. B*, **52**, 393–442 (1990)

[5] De Haan, L., Ferreira, A.: Extreme Value Theory: An Introduction, Springer-Verlag, New York, 2006

[6] Ferraty, F., Vieu, P.: Nonparametric Functional Data Analysis, Springer-Verlag, New York, 2006

[7] Gardes, L., Girard, S., Lekina, A.: Functional nonparametric estimation of conditional extreme quantiles. *J. Mult. Anal.*, **101**, 419–433 (2010)

[8] Gardes, L., Girard, S.: Functional kernel estimators of large conditional quantiles. *Elec. J. Stat.*, **6**, 1715–1744 (2012)

[9] Gardes, L., Girard, S.: On the estimation of the functional Weibull tail-coefficient. *J. Mult. Anal.*, **146**, 29–45 (2016)

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

[11] Resnick, S. I.: Extreme Values, Regular Variation and Point Processes, Springer-Verlag, New York, 1987

[12] Smith, R. L.: Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. *Statist. Sci.*, **4**, 367–393 (1989)