On the strong consistency of the kernel estimator of extreme conditional quantiles

Stéphane Girard[1,2] and Sana Louhichi[2]

[1] Mistis, Inria Grenoble Rhône-Alpes, France.
[2] Laboratoire Jean Kuntzmann, France.

Abstract

Nonparametric regression quantiles can be obtained by inverting a kernel estimator of the conditional distribution. The asymptotic properties of this estimator are well-known in the case of ordinary quantiles of fixed order. The goal of this paper is to establish the strong consistency of the estimator in case of extreme conditional quantiles. In such a case, the probability of exceeding the quantile tends to zero as the sample size increases, and the extreme conditional quantile is thus located in the distribution tails.

1 Introduction

Quantile regression plays a central role in various statistical studies. In particular, nonparametric regression quantiles obtained by inverting a kernel estimator of the conditional distribution function are extensively investigated in the sample case [3, 27, 29, 30]. Extensions to random fields [1], time series [14], functional data [11] and truncated data [25] are also available. However, all these papers are restricted to conditional quantiles having a fixed order $\alpha \in (0,1)$. In the following, α denotes the conditional probability to be larger than the conditional quantile. Consequently, the above mentioned asymptotic theories do not apply in the distribution tails, *i.e* when $\alpha = \alpha_n \to 0$ or $\alpha_n \to 1$ as the sample size n goes to infinity. Motivating applications include for instance environmental studies [16, 28], finance [31], assurance [4] and image analysis [26].

The asymptotics of extreme conditional quantile estimators have been established in a number of regression models. Chernozhukov [6] and Jurecková [22] considered the extreme quantiles in the linear regression model and derived their asymptotic distributions under various error distributions. Other parametric models are considered in [10, 28]. A semi-parametric approach to modeling trends in extremes has been introduced in [9] basing on local polynomial fitting of the Generalized extreme-value distribution. Hall and Tajvidi [21] suggested a nonparametric estimation of the temporal trend when fitting parametric models to extreme-values. Another semi-parametric method has been developed in [2] using a conditional Pareto-type distribution for the response. Fully nonparametric estimators of extreme conditional quantiles have been discussed in [2, 5] including

local polynomial maximum likelihood estimation, and spline fitting via maximum penalized likelihood. Recently, [15, 18] proposed, respectively, a moving-window based estimator for the tail index and extreme quantiles of heavy-tailed conditional distributions, and they established their asymptotic properties.

In the kernel-smoothing case, the asymptotic theory for quantile regression in the tails is still in full development. [19, 20] have analyzed the case $\alpha_n = 1/n$ in the particular situation where the response Y given $X = x$ is uniformly distributed. The asymptotic distribution of the kernel estimator of extreme conditional quantile is established by [7, 17] for heavy-tailed conditional distributions. This result is extended to all types of tails in [8].

Here, we focus on the strong consistency of the kernel estimator for extreme conditional quantiles. Our main result is established in Section 2. Some illustrative examples are provided in Section 3. The proofs of the main results are given in Section 4 and the proofs of the auxiliary results are postponed to the Appendix.

2 Main results

Let $(X_i, Y_i)_{1\leq i\leq n}$ be independent copies of a random pair $(X,Y) \in \mathbb{R}^d \times \mathbb{R}^+$ with density $f_{(X,Y)}$. Let g be the density of X that we suppose strictly positive. The conditional survival function of Y given $X = x$ is denoted by,

$$\bar{F}(y|x) = \mathbb{P}(Y > y|X = x) = \frac{1}{g(x)} \int_y^{+\infty} f_{(X,Y)}(x,z)dz$$

The kernel estimator of $\bar{F}(y|x)$ is, for x such that $\sum_{i=1}^n K_h(x - X_i) \neq 0$,

$$\bar{F}_n(y|x) = \frac{\sum_{i=1}^n K_h(x - X_i)\mathbb{1}_{Y_i>y}}{\sum_{i=1}^n K_h(x - X_i)},$$

where $h = h_n \to 0$ as $n \to \infty$ and $K_h(u) = \frac{1}{h^d} K(\frac{u}{h})$, the kernel K is a measurable function which satisfies the conditions:

$(\mathscr{K}_1)$$K$ is a continuous probability density.

$(\mathscr{K}_2)$$K$ is with compact support: $\exists R > 0$ such that $K(u) = 0$ for any $\|u\| \leq R$. Define $\kappa := \|K\|_\infty = \sup_{x\in\mathbb{R}^d} K(x) < \infty$.

Recall that, for a class of function $\mathscr{G}$, $\mathscr{N}(\varepsilon, \mathscr{G}, d_Q)$ denotes the minimal number of balls $\{g, d_Q(g,g') < \varepsilon\}$ of d_Q-radius ε needed to cover $\mathscr{G}$ and d_Q is the $L_2(Q)$-metric. Let $\mathscr{K}$ be the set of functions $\mathscr{K} = \{K((x-\cdot)/h),\ h > 0, x \in \mathbb{R}^d\}$ and $\mathscr{N}(\varepsilon, \mathscr{K}) = \sup_Q \mathscr{N}(\varepsilon\|K\|_\infty, \mathscr{K}, d_Q)$ where the supremum is taken over all the probability measure Q on $\mathbb{R}^d \times \mathbb{R}$. Suppose that,

$(\mathscr{K}_3)$for some $C, \nu > 1$, $\mathscr{N}(\varepsilon, \mathscr{K}) \leq C\varepsilon^{-\nu}$ for any $\varepsilon \in]0,1[$.

A number of sufficient conditions for which $(\mathscr{K}_3)$ holds are discussed in [13] and the references therein. Finally suppose that

$(\mathscr{A}_1)$for all $\alpha \in (0,1)$ there exists an unique $q(\alpha|x) \in \mathbb{R}$ such that $\bar{F}(q(\alpha|x)|x) = \alpha$.

The conditional quantile $q(\alpha|x)$ is then the inverse of the function $\bar{F}(\cdot|x)$ at the point α. Let (α_n) be a fixed sequence of levels with values in $[0,1]$. For any $x \in \mathbb{R}^d$, define $q(\alpha_n|x)$ as the unique solution of the equation,

$$\alpha_n = \bar{F}(q(\alpha_n|x)|x), \tag{1}$$

whose existence is guaranteed by Assumption $(\mathscr{A}_1)$. The kernel estimator of the conditional quantiles $q(\alpha|x)$ is:

$$\hat{q}_n(\alpha|x) = \inf\{y \in \mathbb{R}, \bar{F}_n(y|x) \leq \alpha\}. \tag{2}$$

Finally, denote by $\hat{g}_n(x)$ the kernel density estimator of the probability density g i.e.

$$\hat{g}_n(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - X_i).$$

Our main result is the following.

Theorem 1 *Let $(X_i, Y_i)_{1\leq i\leq n}$ be independent copies of a random pair $(X,Y) \in \mathbb{R}^d \times \mathbb{R}^+$ with density $f_{(X,Y)}$. Let g be the density of X that we suppose bounded and strictly positive. Suppose that assumption $(\mathscr{A}_1)$ is satisfied. Let (α_n) be a sequence of levels in $[0,1]$ for which*

$$\limsup_{n\to\infty} \sup_{x\in\mathbb{R}^d} \frac{\hat{q}_n(\alpha_n|x)}{q(\alpha_n|x)} \leq Cst, \tag{3}$$

almost surely. Suppose that the kernel K satisfies Conditions $(\mathscr{K}_1), (\mathscr{K}_2), (\mathscr{K}_3)$. Define

$$A(y,z,x,h_n) = \sup_{u:d(u,x)\leq Rh_n} \left| \frac{\bar{F}(y|u)}{\bar{F}(z|x)} - 1 \right|.$$

Suppose that for some fixed positive ε_0 and for $z \in \{q(\alpha_n|x), (1+\varepsilon)q(\alpha_n|x)\}$

$$\limsup_{n\to\infty} \sup_{x\in\mathbb{R}^d, |\varepsilon|\leq\varepsilon_0} A((1+\varepsilon)q(\alpha_n|x), z, x, h_n) \leq C < \infty. \tag{4}$$

If, moreover,

$$\lim_{n\to\infty} nh_n^d\alpha_n = \infty \text{ and } \lim_{n\to\infty} \frac{\ln(\alpha_n h_n^d \wedge \alpha_n^2)}{nh_n^d\alpha_n} = 0, \tag{5}$$

then there exists a positive constant C, not dependent on x, such that one has for n sufficiently large

$$\begin{aligned}\left|1 - \frac{\bar{F}(\hat{q}_n(\alpha_n|x)|x)}{\bar{F}(q(\alpha_n|x)|x)}\right| &\leq C \sup_{|\varepsilon|\leq\varepsilon_0} A((1+\varepsilon)q(\alpha_n|x), (1+\varepsilon)q(\alpha_n|x), x, h_n) \\ &+ \frac{C}{\hat{g}_n(x)}\sqrt{\frac{\ln(\alpha_n^{-1}h_n^{-d}) \vee \ln\ln n}{n\alpha_n h_n^d}}, \text{ almost surely.}\end{aligned}$$

The first term of the bound can be interpreted as a bias term due to the kernel smoothing. The second term can be seen as a variance term, $n\alpha_n h_n^d$ being the effective number of points used in the estimation. The following proposition gives conditions under which (3) is satisfied.

Proposition 1 *Suppose that g is Lipschitzian and bounded above by $g_{\max}$. Let $v_d = \int_{\|v\|\leq 1} dv$ be the volume of the unit sphere. If $A(q(\alpha_n|x), q(\alpha_n|x), x, 0, h) \to 0$ and there exist $\varepsilon > 0$ such that*

$$\sum_{n=1}^{\infty} nh^d\alpha_n \exp\{-v_d g_{\max} nh_d\alpha_n(1+\varepsilon)\} = \infty, \tag{6}$$

then,

$$\limsup_{n\to\infty} \sup_{x\in\mathbb{R}^d} \frac{\hat{q}_n(\alpha_n|x)}{q(\alpha_n|x)} \leq 1, \text{ almost surely.}$$

Some examples of distributions satisfying condition (4) are provided in the next section.

3 Examples

Let us first focus on a conditional Pareto distribution defined as

$$\bar{F}(y|x) = y^{-\theta(x)}, \text{ for all } y > 0. \tag{7}$$

Here, $\theta(x) > 0$ can be read as the inverse of the conditional extreme-value index. The above distribution belongs to the so-called Fréchet maximum domain of attraction which encompasses all distributions with heavy tails. As a consequence of Theorem 1, we have:

Corollary 1 *Let us consider a conditional Pareto distribution (7) such that* $0 < \theta_{\min} \leq \theta(x) \leq \theta_{\max}$ *for all* $x \in \mathbb{R}^d$. *Assume that* θ *is Lipschitzian. If the sequences* (α_n) *and* (h_n) *are such that* $h_n \log \alpha_n \to 0$ *as* $n \to \infty$ *and (5), (6) hold, then* $\hat{q}_n(\alpha_n|x)/q(\alpha_n|x) \to 1$ *almost surely as* $n \to \infty$.

Let us now consider a conditional exponential distribution defined as

$$\bar{F}(y|x) = \exp(-\theta(x)y), \text{ for all } y > 0, \tag{8}$$

where $\theta(x) > 0$ is the inverse of the conditional expectation of Y given $X = x$. This distribution belongs to the Gumbel maximum domain of attraction which collects all distributions with a null conditional extreme-value index. These distributions are often referred to as light-tailed distributions. In such a case, Theorem 1 yields a stronger convergence result than in the heavy-tail framework:

Corollary 2 *Let us consider a conditional exponential distribution (8) with* $0 < \theta_{\min} \leq \theta(x) \leq \theta_{\max}$ *for all* $x \in \mathbb{R}^d$. *Assume that* θ *is Lipschitzian. If the sequences* (α_n) *and* (h_n) *are such that* $h_n \log \alpha_n \to 0$ *as* $n \to \infty$ *and (5), (6) hold, then* $(\hat{q}_n(\alpha_n|x) - q(\alpha_n|x)) \to 0$ *almost surely as* $n \to \infty$.

4 Proofs

4.1 Proof of Theorem 1

Clearly, by (1),

$$\begin{aligned}\frac{|\bar{F}(q(\alpha_n|x)|x) - \bar{F}(\hat{q}_n(\alpha_n|x)|x)|}{\bar{F}(q(\alpha_n|x)|x)} &\leq \frac{|\alpha_n - \bar{F}_n(\hat{q}_n(\alpha_n|x)|x)|}{\bar{F}(q(\alpha_n|x)|x)} \\ &+ \frac{|\bar{F}_n(\hat{q}_n(\alpha_n|x)|x) - \bar{F}(\hat{q}_n(\alpha_n|x)|x)|}{\bar{F}(q(\alpha_n|x)|x)}.\end{aligned}$$

First, from (2), $|\alpha_n - \bar{F}_n(\hat{q}_n(\alpha_n|x)|x)|$ is bounded above by the maximal jump of $\bar{F}_n(y|x)$ at some observation point (X_j, Y_j):

$$|\alpha_n - \bar{F}_n(\hat{q}_n(\alpha_n|x)|x)| \leq \frac{\max_{j=1,\ldots,n} K_h(x - X_j)}{\sum_{i=1}^n K_h(x - X_i)}.$$

It follows from $(\mathscr{K}_2)$ that

$$\frac{|\alpha_n - \bar{F}_n(\hat{q}_n(\alpha_n|x)|x)|}{\bar{F}(q(\alpha_n|x)|x)} \leq \frac{\kappa}{nh^d \alpha_n \hat{g}_n(x)}.$$

Let us then focus on the second term:

$$\frac{|\bar{F}_n(\hat{q}_n(\alpha_n|x)|x)-\bar{F}(\hat{q}_n(\alpha_n|x)|x)|}{\bar{F}(q(\alpha_n|x)|x)}=\frac{\bar{F}(\hat{q}_n(\alpha_n|x)|x)}{\bar{F}(q(\alpha_n|x)|x)}\left|\frac{\bar{F}_n(\hat{q}_n(\alpha_n|x)|x)}{\bar{F}(\hat{q}_n(\alpha_n|x)|x)}-1\right|. \quad (9)$$

We write $\hat{q}_n(\alpha_n|x)=(1+\varepsilon)\,q(\alpha_n|x)$, with $\varepsilon=\frac{\hat{q}_n(\alpha_n|x)}{q(\alpha_n|x)}-1$. Condition (3) allows to deduce that there exists, for n sufficiently large, ε_0 not dependent on x such that $|\varepsilon|\leq\varepsilon_0$. Consequently and taking into account (9) there exists a positive constant ε_0 not dependent on x and n such that (for the sake of simplicity, we write $q=q(\alpha_n|x)$)

$$\begin{aligned}\left|\frac{\bar{F}(\hat{q}_n(\alpha_n|x)|x)}{\bar{F}(q(\alpha_n|x)|x)}-1\right| &\leq \frac{\kappa}{nh^d\alpha_n\hat{g}_n(x)}\\ &+ \sup_{|\varepsilon|\leq\varepsilon_0}\left(\frac{\bar{F}(q(1+\varepsilon))|x)}{\bar{F}(q|x)}\left|\frac{\bar{F}_n(q(1+\varepsilon)|x)}{\bar{F}(q(1+\varepsilon))|x)}-1\right|\right) \quad (10)\end{aligned}$$

Our purpose now is to control the term $\left|\frac{\bar{F}_n(q(1+\varepsilon)|x)}{\bar{F}(q(1+\varepsilon)|x)}-1\right|$ of (10). For this, write

$$\bar{F}_n(y|x)=\frac{\sum_{i=1}^n K_h(x-X_i)\mathbf{1}_{Y_i>y}}{\sum_{i=1}^n K_h(x-X_i)}=:\frac{\hat{\psi}_n(y,x)}{\hat{g}_n(x)},$$

with

$$\hat{\psi}_n(y,x)=\frac{1}{n}\sum_{i=1}^n K_h(x-X_i)\mathbf{1}_{Y_i>y},\ \ \hat{g}_n(x)=\frac{1}{n}\sum_{i=1}^n K_h(x-X_i).$$

We need the following lemma.

Lemma 1 *Suppose that Condition ($\mathscr{H}_2$) holds. Then, for each n, one has for any $y\in\mathbb{R}$ and $x\in\mathbb{R}^d$ for which $\bar{F}(y|x)\hat{g}_n(x)\neq 0$,*

$$\begin{aligned}&\left|\frac{\bar{F}_n(y|x)}{\bar{F}(y|x)}-1\right|\\ &\leq A(y,y,x,h_n)+(1+A(y,y,x,h_n))\frac{|\hat{g}_n(x)-\mathbb{E}\hat{g}_n(x)|}{\hat{g}_n(x)}+\frac{|\hat{\psi}_n(y,x)-\mathbb{E}\,(\hat{\psi}_n(y,x))\,|}{\bar{F}(y|x)\hat{g}_n(x)}.\end{aligned}$$

The proof is postponed to the Appendix. According to Lemma 1, we have to control the two quantities $\mathbb{E}\,|\hat{g}_n(x)-\mathbb{E}\hat{g}_n(x)|$ and $\frac{|\hat{\psi}_n(y,x)-\mathbb{E}(\hat{\psi}_n(y,x))|}{\bar{F}(y|x)}$. This is the purpose of Propositions 2 and 3 below.

Proposition 2 *Einmahl-Mason (2005).*
Suppose that g is a bounded density on $\mathbb{R}^d$, and that the assumptions $(K.i),\cdots,(K.iv)$ of [13] are all satisfied. Then, for any $c>0$,

$$\limsup_{n\to\infty}\sup_{\{c\ln n/n\leq h_n^d\leq 1\}}\sqrt{\frac{nh_n^d}{\ln(h_n^{-d})\vee\ln\ln n}}\sup_{x\in\mathbb{R}^d}|\hat{g}_n(x)-\mathbb{E}\hat{g}_n(x)|=:K(c)<\infty,$$

almost surely.

Our task now is to control $\frac{|\hat{\psi}_n(y,x)-\mathbb{E}(\hat{\psi}_n(y,x))|}{\bar{F}(y|x)}$. Let ε be a fixed real in $[-\varepsilon_0,\varepsilon_0]$ for some arbitrary positive ε_0. The following proposition evaluates the almost sure asymptotic behaviour of

$$\frac{|\hat{\psi}_n(q(1+\varepsilon),x)-\mathbb{E}(\hat{\psi}_n(q(1+\varepsilon),x))|}{\bar{F}(q|x)}.$$

Proposition 3 *Let (α_n) be a sequence in $[0,1]$ and for $x \in \mathbb{R}^d$, $q = q(\alpha_n|x)$ be the conditional quantile as defined by (1). Define the set of functions $\mathscr{F}$ by,*

$$\mathscr{F} = \left\{ (u,v) \longmapsto K\left(\frac{x-u}{h}\right) I_{v>q(1+\varepsilon)},\ n \in \mathbb{N},\ h>0,\ x \in \mathbb{R}^d,\ |\varepsilon| \le \varepsilon_0 \right\} \quad (11)$$

and suppose that $\mathscr{N}(\varepsilon,\mathscr{F}) \le C\varepsilon^{-\nu}$, for some $C,\nu > 1$ and all $\varepsilon \in]0,1[$. Suppose also that Condition (4) is satisfied. If $nh_n^d\alpha_n \to \infty$ and $\frac{\ln(\alpha_n h_n^d \wedge \alpha_n^2)}{nh_n^d\alpha_n} \to 0$ as $n \to \infty$, then there exists a positive constant C_1 such that

$$\limsup_{n\to\infty} \sqrt{\frac{n\alpha_n h_n^d}{\ln(\alpha_n^{-1}h_n^{-d}) \vee \ln\ln n}} \sup_{x\in\mathbb{R}^d, |\varepsilon|\le\varepsilon_0} \frac{|\hat{\psi}_n(q(1+\varepsilon),x) - \mathbb{E}(\hat{\psi}_n(q(1+\varepsilon),x))|}{F(q|x)} \le C_1,$$

almost surely.

Proof of Proposition 3. We have,

$$\begin{aligned}
&\frac{1}{F(q|x)}(\hat{\psi}_n(q(1+\varepsilon),x) - \mathbb{E}(\hat{\psi}_n(q(1+\varepsilon),x))) \\
&= \frac{1}{nF(q|x)} \sum_{i=1}^n [K_h(x-X_i)\mathbf{1}_{Y_i>q(1+\varepsilon)} - \mathbb{E}(K_h(x-X_i)\mathbf{1}_{Y_i>q(1+\varepsilon)})] \\
&= \frac{1}{nh_n^d} \sum_{i=1}^n (v_{h,x,\varepsilon}(X_i,Y_i) - \mathbb{E}(v_{h,x,\varepsilon}(X_i,Y_i))) = \frac{1}{h_n^d\sqrt{n}}\beta_n(v_{h,x,\varepsilon}),
\end{aligned}$$

where,

$$v_{h,x,\varepsilon}(u,v) = K(\frac{x-u}{h})\frac{\mathbf{1}_{v>q(1+\varepsilon)}}{F(q|x)},\ \beta_n(g) = \frac{1}{\sqrt{n}}\sum_{i=1}^n (g(X_i,Y_i) - \mathbb{E}(g(X_i,Y_i)))$$

Define the class of functions:

$$\mathscr{G} := \mathscr{G}_{n,h} = \{v_{h,x,\varepsilon},\ x \in \mathbb{R}^d, \varepsilon \in [-\varepsilon_0,\varepsilon_0]\} \quad (12)$$

and let $\|\beta_n\|_{\mathscr{G}} = \sup_{g\in\mathscr{G}} |\beta_n(g)|$ and

$$\Theta_n = \sup_{x\in\mathbb{R}^d, \varepsilon\in[-\varepsilon_0,\varepsilon_0]} \frac{|\hat{\psi}_n(q(1+\varepsilon),x) - \mathbb{E}(\hat{\psi}_n(q(1+\varepsilon),x))|}{F(q|x)}.$$

Consequently, for any $\gamma > 0$,

$$\mathbb{P}(\Theta_n > \gamma) \le \mathbb{P}\left(\sqrt{n}\|\beta_n\|_{\mathscr{G}} > \gamma nh^d\right) \le \mathbb{P}\left(\max_{1\le m\le n}\sqrt{m}\|\beta_m\|_{\mathscr{G}} > \gamma nh^d\right) \quad (13)$$

We have then to evaluate $\max_{1\le m\le n}\sqrt{m}\|\beta_m\|_{\mathscr{G}}$. By Talagrand Inequality, (see A.1. in [12]), we have for any $t>0$ and suitable finite constants $A_1, A_2 > 0$,

$$\begin{aligned}
&\mathbb{P}\left(\max_{1\le m\le n}\sqrt{m}\|\beta_m\|_{\mathscr{G}} > A_1\left(\mathbb{E}\|\sum_{i=1}^n \varepsilon_i g(X_i,Y_i)\|_{\mathscr{G}} + t\right)\right) \\
&\le 2\exp(-A_2t^2/n\sigma^2) + 2\exp(-A_2t/M),
\end{aligned}$$

where $(\varepsilon_i)_i$ is a sequence of independent Rademacher random variables independent of the random vectors $(X_i,Y_i)_{1\le i\le n}$ and

$$\sup_{g\in\mathscr{G}}\|g\|_\infty \le M,\ \sup_{g\in\mathscr{G}}\mathrm{Var}(g(X,Y)) \le \sigma^2.$$

Here $\|g\|_\infty \leq \frac{\|K\|_\infty}{\alpha_n} =: M$, and

$$\mathrm{Var}(v^2_{h,x,\varepsilon}(X,Y)) \leq \mathbb{E}(v^2_{h,x,\varepsilon}(X,Y)) = \frac{1}{\bar{F}^2(q|x)}\mathbb{E}(K^2(\frac{x-X}{h})\mathbf{1}_{Y>q(1+\varepsilon)})$$
$$= \frac{1}{F(q|x)}\int K^2(\frac{x-u}{h})\frac{\bar{F}(q(1+\varepsilon)|u)}{F(q|x)}g(u)du$$
$$\leq \frac{h_n^d}{\alpha_n}\sup_{x\in\mathbb{R}^d,\varepsilon\in[-\varepsilon_0,\varepsilon_0]}(1+A(q(1+\varepsilon),q,x,h))\|K\|_2^2\|g\|_\infty = \frac{h_n^d}{\alpha_n}L =: \sigma^2, \quad (14)$$

for some positive constant L, since for n sufficiently large

$$\sup_{x\in\mathbb{R}^d,\varepsilon\in[-\varepsilon_0,\varepsilon_0]} A(q(1+\varepsilon),q,x,h) \leq cst.$$

We obtain combining this with the above Talagrand's Inequality,

$$\mathbb{P}\left(\max_{1\leq m\leq n}\sqrt{m}\|\beta_m\|_\infty > A_1\left(\mathbb{E}\|\sum_{i=1}^n \varepsilon_i g(X_i,Y_i)\|_{\mathscr{G}} + t\right)\right)$$
$$\leq 2\exp(-A_2t^2\frac{\alpha_n}{nh_n^dL}) + 2\exp(-A_2t\frac{\alpha_n}{\|K\|_\infty}). \quad (15)$$

The last bound together with (13) gives

$$\mathbb{P}\left(\Theta_n > A_1n^{-1}h_n^{-d}\left(\mathbb{E}\|\sum_{i=1}^n \varepsilon_i g(X_i,Y_i)\|_{\mathscr{G}} + t\right)\right)$$
$$\leq 2\exp(-A_2t^2\frac{\alpha_n}{nh_n^dL}) + 2\exp(-A_2t\frac{\alpha_n}{\|K\|_\infty}). \quad (16)$$

We have now to evaluate $\mathbb{E}\|\sum_{i=1}^n \varepsilon_i g(X_i,Y_i)\|_{\mathscr{G}}$. We will argue as for the proof of Proposition A.1. in [12]. We have by (6.9) of Proposition 6.8 in [24],

$$\mathbb{E}\|\sum_{i=1}^n \varepsilon_i g(X_i,Y_i)\|_{\mathscr{G}} \leq 6t_n + 6\mathbb{E}(\max_{i\leq n}\|\varepsilon_i g(X_i,Y_i)\|_{\mathscr{G}}) \leq 6t_n + 6\frac{\|K\|_\infty}{\alpha_n}, \quad (17)$$

where we have defined

$$t_n = \inf\left\{t>0, \mathbb{P}\left(\|\sum_{i=1}^n \varepsilon_i g(X_i,Y_i)\|_{\mathscr{G}} > t\right) \leq \frac{1}{24}\right\}.$$

Our purpose is then to control t_n. Define the event

$$F_n = \left\{n^{-1}\sup_{g\in\mathscr{G}}\sum_{j=1}^n g^2(X_j,Y_j) \leq 64\sigma^2\right\},$$

where σ^2 is as in (14). Let g_0 be a fixed element of $\mathscr{G}$. We have,

$$\mathbb{E}|\sum_{i=1}^n \varepsilon_i g_0(X_i,Y_i)\mathbf{1}_{F_n}| \leq 8\sigma\sqrt{n}.$$

By (A8) of [12], we have now to control $\mathscr{N}(\varepsilon,\mathscr{G},d_{n,2})$. Recall that $\mathscr{N}(\varepsilon,\mathscr{G},d_Q)$ is the minimal number of balls $\{g, d_Q(g,g') < \varepsilon\}$ of d_Q-radius ε needed to cover $\mathscr{G}$, d_Q is the $L_2(Q)$-metric and

$$d_{n,2}(f,g) = d_{Q_n}(f,g) = \int (f(x)-g(x))^2 dQ_n(x),$$

with $Q_n = \frac{1}{n}\sum_{i=1}^n \delta_{(X_i,Y_i)}$. In other words

$$d_{n,2}(f,g) = \frac{1}{n}\sum_{i=1}^n (f(X_i,Y_i) - g(X_i,Y_i))^2.$$

We note first that, on the event F_n,

$$\mathscr{N}(\varepsilon, \mathscr{G}, d_{n,2}) = 1, \text{ whenever } \varepsilon > 16\sigma.$$

We will suppose then $\varepsilon \leq 16\sigma$. We have

$$\mathscr{N}(\varepsilon, \mathscr{G}, d_{n,2}) = \mathscr{N}(\varepsilon, \mathscr{G}, d_{Q_n}) \leq \mathscr{N}(\varepsilon\alpha_n, \mathscr{F}, d_{Q_n}),$$

where $\mathscr{F}$ is as defined by (11). Recall that $\mathscr{N}(\varepsilon, \mathscr{F}) = \sup_Q \mathscr{N}(\varepsilon\|K\|_\infty, \mathscr{F}, d_Q)$ where the supremum is taken over all the probability measure Q on $\mathbb{R}^d \times \mathbb{R}$. We have supposed that,

$$\mathscr{N}(\varepsilon, \mathscr{F}) \leq C\varepsilon^{-\nu}.$$

for some $C, \nu > 1$ and all $\varepsilon \in]0,1[$. Consequently,

$$\mathscr{N}(\varepsilon, \mathscr{G}, d_{n,2}) \leq C(\frac{\alpha_n\varepsilon}{\|K\|_\infty})^{-\nu}, \tag{18}$$

as soon as $\alpha_n\varepsilon < \|K\|_\infty$. Hence, we have almost surely on the event F_n,

$$\begin{aligned}
\int_0^\infty \sqrt{\ln(\mathscr{N}(\varepsilon, \mathscr{G}, d_{n,2}))}d\varepsilon &= \int_0^{16\sigma} \sqrt{\ln(\mathscr{N}(\varepsilon, \mathscr{G}, d_{n,2}))}d\varepsilon \\
&\leq \sum_{i=0}^\infty \int_{2^{-i-1}16\sigma}^{2^{-i}16\sigma} \sqrt{\ln(C) + \nu\ln(\frac{\|K\|_\infty}{\alpha_n\varepsilon})}d\varepsilon \\
&\leq \sum_{i=0}^\infty 2^{-i-1}16\sigma\sqrt{\ln(C) + \nu\ln(\frac{2^{i+1}\|K\|_\infty}{\alpha_n 16\sigma})} \\
&\leq C_2 \sum_{i=0}^\infty \frac{\sqrt{i+1}}{2^{i+1}}\sqrt{\sigma^2\max(\ln(C_1), \ln(\frac{1}{\alpha_n^2\sigma^2}))},
\end{aligned}$$

for some positive constants C_1, C_2 that depend only on C, ν and $\|K\|_\infty$. We conclude, using (A8) of [12],

$$\mathbb{E}\left(\|\sum_{i=1}^n \varepsilon_i g(X_i,Y_i)\|_{\mathscr{G}} \mathbf{1}_{F_n}\right) \leq C_2'\sqrt{n\sigma^2\max(\ln(C_1), \ln(\frac{1}{\alpha_n^2\sigma^2}))}. \tag{19}$$

We now use Inequality A2 in [12] (which is due to Giné and Zinn), with $t = 64\sqrt{n}\sigma^2$. We obtain, for $m \geq 1$, since for any $g \in \mathscr{G}$, $\|g\|_\infty \leq \frac{\|K\|_\infty}{\alpha_n}$,

$$\begin{aligned}
\mathbb{P}(F_n^c) &= \mathbb{P}\left(n^{-1}\sup_{g\in\mathscr{G}}\sum_{j=1}^n g^2(X_i,Y_i) > 64\sigma^2\right) \\
&\leq 4\mathbb{P}(\mathscr{N}(\rho n^{-1/4}, \mathscr{G}, d_{n,2}) \geq m) + 8m\exp(-n\sigma^2\alpha_n^2/\|K\|_\infty^2)
\end{aligned}$$

where $n^{-1/4}\rho = n^{-1/4}\min(\sigma n^{1/4}, n^{1/4}) = \min(\sigma, 1)$. Hence by (18),

$$\mathscr{N}(\rho n^{-1/4}, \mathscr{G}, d_{n,2}) \leq C(\frac{\alpha_n\min(\sigma,1)}{\|K\|_\infty})^{-\nu},$$

Consequently, we have for $m = [2C(\frac{\alpha_n \min(\sigma,1)}{\|K\|_\infty})^{-\nu}]$,

$$\mathbb{P}(F_n^c) \leq 16C(\frac{\alpha_n \min(\sigma,1)}{\|K\|_\infty})^{-\nu} \exp(-n\sigma^2\alpha_n^2/\|K\|_\infty^2)$$

The last bound together with (19) gives,

$$\begin{aligned}\mathbb{P}\left(\|\sum_{i=1}^n \varepsilon_i g(X_i,Y_i)\|_{\mathscr{G}} > t\right) &\leq \mathbb{P}(F_n^c) + \frac{1}{t}\mathbb{E}\left(\|\sum_{i=1}^n \varepsilon_i g(X_i,Y_i)\|_{\mathscr{G}} \mathbb{1}_{F_n}\right) \\ &\leq 16C(\frac{\|K\|_\infty}{\alpha_n \min(\sigma,1)})^{\nu} \exp(-n\sigma^2\alpha_n^2/\|K\|_\infty^2) \\ &+ \frac{C_2'}{t}\sqrt{n\sigma^2 \max(\ln(C_1), \ln(\frac{1}{\alpha_n^2\sigma^2}))}. \qquad (20)\end{aligned}$$

Let us control the second term in (20). Recall that by (14), $\sigma^2 = L\frac{h_n^d}{\alpha_n}$ and thus $\alpha_n^2\sigma^2 = Lh_n^d\alpha_n$ for some positive constant L. This fact together with $h_n^d\alpha_n \to 0$ as $n \to \infty$ allows to deduce that for n sufficiently large,

$$\frac{C_2'}{t}\sqrt{n\sigma^2 \max(\ln(C_1), \ln(\frac{1}{\alpha_n^2\sigma^2}))} \leq \frac{Cst}{t}\sqrt{\frac{nh_n^d}{\alpha_n}\ln(\frac{1}{\alpha_n h_n^d})}. \qquad (21)$$

Our task now is to control the first term in (20). We have,

$$\begin{aligned}&16C(\frac{\|K\|_\infty}{\alpha_n \min(\sigma,1)})^{\nu} \exp(-n\sigma^2\alpha_n^2/\|K\|_\infty^2) \\ &\leq L_0^{\nu/2} \exp\left(-nh_n^d\alpha_n(L_1 + \frac{\nu}{2nh_n^d\alpha_n}\ln(\alpha_n h_n^d \wedge \alpha_n^2))\right),\end{aligned}$$

which tends to 0 as $n \to \infty$ as soon as $nh_n^d\alpha_n \to 0$ and $\frac{\ln(\alpha_n h_n^d \wedge \alpha_n^2)}{nh_n^d\alpha_n} \to 0$. Hence for $t \geq 48Cst\sqrt{\frac{nh_n^d}{\alpha_n}\ln(\frac{1}{\alpha_n h_n^d})} =: t_n$, Inequalities (20) and (21) give for n sufficiently large and for $t \geq t_n$

$$\mathbb{P}\left(\|\sum_{i=1}^n \varepsilon_i g(X_i,Y_i)\|_{\mathscr{G}} > t\right) \leq \frac{1}{24}.$$

We conclude using this fact together with Inequality (17),

$$\mathbb{E}\|\sum_{i=1}^n \varepsilon_i g(X_i,Y_i)\|_{\mathscr{G}} \leq c\left(\frac{1}{\alpha_n} + \sqrt{\frac{nh_n^d}{\alpha_n}\ln(\frac{1}{\alpha_n h_n^d})}\right) = \mathscr{O}\left(\sqrt{\frac{nh_n^d}{\alpha_n}\ln(\frac{1}{\alpha_n h_n^d})}\right). \quad (22)$$

Recalling that

$$\Theta_n = \sup_{x\in\mathbb{R}^d, \varepsilon\in[-\varepsilon_0,\varepsilon_0]} \frac{|\hat{\psi}_n(q(1+\varepsilon),x) - \mathbb{E}(\hat{\psi}_n(q(1+\varepsilon),x))|}{F(q|x)},$$

and collecting (22), (16), yields

$$\begin{aligned}&\mathbb{P}\left(\Theta_n > A_1 n^{-1} h_n^{-d}\left(\mathbb{E}\|\sum_{i=1}^n \varepsilon_i g(X_i,Y_i)\|_{\mathscr{G}} + t\right)\right) \\ &\leq 2\left[\exp\left(-A_2\tilde{L}^2\frac{\ln(\ln(n))}{L}\right) + \exp\left(-\frac{A_2\tilde{L}}{\|K\|_\infty}\sqrt{nh_n^d\alpha_n\ln(\ln n)}\right)\right], (23)\end{aligned}$$

for any $t \geq \left(Cst \sqrt{\frac{nh_n^d}{\alpha_n} \ln(\frac{1}{\alpha_n h_n^d})} \right) \vee \tilde{L} \sqrt{\frac{nh_n^d}{\alpha_n} \ln(\ln n)}$ and some $\tilde{L} > 0$. Now,

$$\frac{\ln n}{n\alpha_n h_n^d} + \frac{\ln(\alpha_n h_n^d \wedge \alpha_n^2)}{n\alpha_n h_n^d} \leq \frac{\ln(n\alpha_n h_n^d)}{n\alpha_n h_n^d},$$

which tends to 0 as n tends to infinity, since $\lim_{n\to\infty} n\alpha_n h_n^d = \infty$. Hence,

$$\lim_{n\to\infty} \frac{\ln n}{n\alpha_n h_n^d} = 0,$$

which proves that for n sufficiently large $nh_n^d \alpha_n \geq \ln n$. We conclude then from (23) that,

$$\mathbb{P}\left(\Theta_n > A_1 n^{-1} h_n^{-d} \left(\mathbb{E}\| \sum_{i=1}^n \varepsilon_i g(X_i, Y_i) \|_{\mathscr{G}} + t \right) \right) \leq 4(\ln n)^{-\rho},$$

since we have $nh_n^d \alpha_n \geq \ln n$. We choose $\tilde{L}$ in such a way that $\rho > 1$. Proposition 3 is proved thanks to Borel-Cantelli lemma. ■

We continue the proof of Theorem 1. Inequality (10), together with Lemma 1 and Condition (4), gives for some universal positive constant C

$$\begin{aligned}
&\left| \frac{\bar{F}(\hat{q}_n(\alpha_n|x)|x)}{F(q(\alpha_n|x)|x)} - 1 \right| \leq \frac{\kappa}{nh^d \alpha_n \hat{g}_n(x)} \\
&+C \sup_{|\varepsilon|\leq\varepsilon_0} A(q(1+\varepsilon), q(1+\varepsilon), x, h_n) \\
&+C \sup_{|\varepsilon|\leq\varepsilon_0} (1 + A(q(1+\varepsilon), q(1+\varepsilon), x, h_n)) \frac{|\hat{g}_n(x) - \mathbb{E}\hat{g}_n(x)|}{\hat{g}_n(x)} \\
&+C \sup_{|\varepsilon|\leq\varepsilon_0} \frac{|\hat{\psi}_n(q(1+\varepsilon), x) - \mathbb{E}(\hat{\psi}_n(q(1+\varepsilon), x))|}{F(q|x)\hat{g}_n(x)}.
\end{aligned} \tag{24}$$

We first use Einmahl and Mason's result (cf. Proposition 2 above). All the requirements of Einmahl and Mason result are satisfied from that of Theorem 1. This gives that, for $c \ln n/n \leq h_n^d \leq 1$,

$$\sqrt{\frac{nh_n^d}{\ln(h_n^{-d}) \vee \ln\ln n}} \sup_{x\in\mathbb{R}^d} |\hat{g}_n(x) - \mathbb{E}\hat{g}_n(x)| < C, \tag{25}$$

almost surely. Our task now is to apply Proposition 3. We first claim that

Lemma 2 *Under Condition* $(\mathscr{K}_3)$*, the class of function* $\mathscr{F}$ *defined by* (11) *satisfies* $\mathscr{N}(\varepsilon, \mathscr{F}) \leq C\varepsilon^{-\nu}$*, for* $C > 0, \varepsilon > 0, \nu > 1$.

Proof of Lemma 2. Define the set of function $\mathscr{F} = \mathscr{K}\mathscr{I}$, where the set of functions $\mathscr{K}$ is $\mathscr{K} = \left\{ u \longmapsto K\left(\frac{x-u}{h}\right), x \in \mathbb{R}^d, h > 0 \right\}$, and $\mathscr{I} = \{ v \longmapsto \mathbf{1}_{v>q(1+\varepsilon)}, x \in \mathbb{R}^d, n \in \mathbb{N}, |\varepsilon| \leq \varepsilon_0 \}$. The proof of Lemma 2 follows from Lemma A.1 of [12] since $\mathscr{N}(\varepsilon, \{ v \longmapsto \mathbf{1}_{v>y}, y \in \mathbb{R} \}) \leq C\varepsilon^{-\tilde{\nu}}$ with $\tilde{\nu} > 0$ and $C > 0$. ■

Consequently, all the requirements of Proposition 3 are satisfied from that of Theorem 1. The conclusion of Proposition 3 together with (25), (24) and the facts that

$$\begin{aligned}\sqrt{\frac{\ln(h_n^{-d})\vee\ln\ln n}{nh_n^d}} &\leq \sqrt{\frac{\ln(\alpha_n^{-1}h_n^{-d})\vee\ln\ln n}{nh_n^d\alpha_n}},\\ \frac{1}{nh_n^d\alpha_n} &\leq \sqrt{\frac{\ln(\alpha_n^{-1}h_n^{-d})\vee\ln\ln n}{n\alpha_n h_n^d}}\end{aligned}$$

complete the proof of Theorem 1. ■

4.2 Proof of Proposition 1

Let us introduce $Z_i^{(n)}(x)$ for $i=1,\ldots,n$ a triangular array of i.i.d. random variables defined by $Z_i^{(n)}(x)=Y_iI_{\|x-X_i\|\leq h}$. Their common survival distribution function can be expanded as:

$$\begin{aligned}\bar{\Psi}_n(t,x) &= \mathbb{P}(Z_1^{(n)}(x)>t)=\int_{\|x-u\|\leq h}\bar{F}(t|u)g(u)du\\ &= h^d\int_{\|v\|\leq 1}\bar{F}(t|x-hv)g(x-hv)dv,\end{aligned}$$

or equivalently,

$$\begin{aligned}\frac{\bar{\Psi}_n(t,x)}{h^d\bar{F}(t|x)g(x)} &= \int_{\|v\|\leq 1}dv\\ &+ \int_{\|v\|\leq 1}\left(\frac{\bar{F}(t|x-hv)}{F(t|x)}-1\right)\frac{g(x-hv)}{g(x)}dv\\ &+ \int_{\|v\|\leq 1}\left(\frac{g(x-hv)}{g(x)}-1\right)dv.\end{aligned}$$

Letting $v_d=\int_{\|v\|\leq 1}dv$ the volume of the unit sphere, and assuming that g is Lipschitzian, it follows,

$$\frac{\bar{\Psi}_n(t,x)}{h^d\bar{F}(t|x)g(x)}=v_d+o(1)+O(A(t,t,x,0,h))$$

and introducing $\beta_n(x)=n\bar{\Psi}_n(q(\alpha_n|x)|x)$, we obtain

$$\beta_n(x)=v_dg(x)nh_d\alpha_n(1+o(1))$$

under condition $A(q(\alpha_n|x),q(\alpha_n|x),x,0,h)\to 0$ as $n\to\infty$. We now need the following lemma (also available to triangular arrays).

Lemma 3 *(Klass, 1985) [23] Let $Z,Z_1,Z_2,\ldots$ be a sequence of i.i.d. random vectors and define $M_n=\max\{Z_1,\ldots,Z_n\}$. Suppose that (b_n) is nondecreasing, $\mathbb{P}(Z>b_n)\to 0$ and $n\mathbb{P}(Z>b_n)\to\infty$ as $n\to\infty$. If, moreover,*

$$\sum_{n=1}^{\infty}\mathbb{P}(Z>b_n)\exp\{-n\mathbb{P}(Z>b_n)\}=\infty,$$

then $\limsup_{n\to\infty}M_n/b_n<1$ a.s.

From Lemma 3, a sufficient condition for

$$\limsup_{n\to\infty} \frac{\max_{1\leq i\leq n} Z_i^{(n)}(x)}{q(\alpha_n|x)} < 1 \; a.s. \tag{26}$$

is

$$\sum_{n=1}^{\infty} \frac{\beta_n(x)}{n} \exp\{-\beta_n(x)\} = \infty,$$

which is fulfilled under (6). Finally,

$$\frac{\hat{q}_n(\alpha_n|x)}{q(\alpha_n|x)} \leq \frac{\max_{1\leq i\leq n} Z_i^{(n)}(x)}{q(\alpha_n|x)}$$

and the conclusion follows from (26). ■

4.3 Proofs of corollaries

Proof of Corollary 1. For all $\tau \in [0,1]$ and (u,x) such that $d(u,x) \leq Rh_n$, we have

$$\begin{aligned}
\frac{\bar{F}(q(\alpha_n|x)(1+\varepsilon)|u)}{\bar{F}(q(\alpha_n|x)(1+\varepsilon)^{\tau}|x)} &= q(\alpha_n|x)^{\theta(x)-\theta(u)}(1+\varepsilon)^{\tau\theta(x)-\theta(u)} \\
&= \exp\left\{\frac{\theta(u)-\theta(x)}{\theta(x)}\log\alpha_n + (\tau\theta(x)-\theta(u))\log(1+\varepsilon)\right\} \\
&= \exp\{O(h_n\log\alpha_n) + O(\tau\theta(x)-\theta(u))\} \\
&= (1+O(h_n\log\alpha_n))\exp\{O(\tau\theta(x)-\theta(u))\}. \qquad (27)
\end{aligned}$$

Assuming that $h_n\log\alpha_n \to 0$ as $n\to\infty$, it follows that (27) is bounded above for all $\tau\in[0,1]$ and (u,x) such that $d(u,x)\leq Rh_n$ and therefore condition (4) is fulfilled. If, moreover, $\tau = 1$ then $O(\theta(x)-\theta(u)) = O(h_n)$ and thus (27) tends to zero as n goes to infinity. Proposition 1 then entails that assumption (3) holds. Theorem 1 implies that

$$\left|1 - \frac{\bar{F}(\hat{q}_n(\alpha_n|x)|x)}{\bar{F}(q(\alpha_n|x)|x)}\right| = \left|1 - \left(\frac{\hat{q}_n(\alpha_n|x)}{q(\alpha_n|x)}\right)^{-\theta(x)}\right| \to 0$$

almost surely as $n\to\infty$. The conclusion follows. ■

Proof of Corollary 2. For all $\tau\in[0,1]$ and (u,x) such that $d(u,x)\leq Rh_n$, we have

$$\begin{aligned}
\frac{\bar{F}(q(\alpha_n|x)(1+\varepsilon)|u)}{\bar{F}(q(\alpha_n|x)(1+\varepsilon)^{\tau}|x)} &= \exp\left\{(1+\varepsilon)\log(\alpha_n)\left(\frac{\theta(u)-\theta(x)}{\theta(x)} + 1 - (1+\varepsilon)^{\tau-1}\right)\right\} \\
&= \exp\{O(h_n\log\alpha_n)\}\exp\left\{(1+\varepsilon)\log(\alpha_n)\left(1-(1+\varepsilon)^{\tau-1}\right)\right\} \\
&= (1+O(h_n\log\alpha_n)o(1) = o(1). \qquad (28)
\end{aligned}$$

Assuming that $h_n\log\alpha_n\to 0$ as $n\to\infty$, it follows that (28) tends to zero as n goes to infinity. Assumptions (3) and (4) both hold. Theorem 1 implies that

$$\left|1 - \frac{\bar{F}(\hat{q}_n(\alpha_n|x)|x)}{\bar{F}(q(\alpha_n|x)|x)}\right| = |1-\exp\left((q(\alpha_n|x)-\hat{q}_n(\alpha_n|x))\theta(x)\right)| \to 0$$

almost surely as $n\to\infty$. The conclusion follows. ■

Appendix: proof of auxiliary results

Proof of Lemma 1. Clearly,

$$\left|\frac{\bar{F}_n(y|x)}{F(y|x)}-1\right| \leq \left|\frac{\bar{F}_n(y|x)}{F(y|x)}-\frac{\mathbb{E}\left(\hat{\psi}_n(y,x)\right)}{F(y|x)\mathbb{E}\left(\hat{g}_n(x)\right)}\right| + \left|\frac{\mathbb{E}\left(\hat{\psi}_n(y,x)\right)}{F(y|x)\mathbb{E}\left(\hat{g}_n(x)\right)}-1\right| \tag{29}$$

We have,

$$\begin{aligned}
&\mathbb{E}\left(\hat{\psi}_n(y,x)\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(K_h(x-X_i)\mathbf{1}_{Y_i>y}) = \mathbb{E}(K_h(x-X_1)\mathbf{1}_{Y_1>y}) \\
&= \mathbb{E}(K_h(x-X_1)\mathbb{P}(Y_1>y|X_1)) = \int K_h(x-z)\mathbb{P}(Y_1>y|X_1=z)g(z)dz \\
&= \int K_h(x-z)\bar{F}(y|z)g(z)dz,
\end{aligned}$$

and $\mathbb{E}\left(\hat{g}_n(x)\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(K_h(x-X_i)) = \mathbb{E}(K_h(x-X_1))$. Consequently,

$$\begin{aligned}
&\frac{\mathbb{E}\left(\hat{\psi}_n(y,x)\right)}{F(y|x)\mathbb{E}\left(\hat{g}_n(x)\right)}-1 = \\
&= \frac{1}{\mathbb{E}(K_h(x-X_1))}\left(\int K_h(x-z)\left[\frac{\bar{F}(y|z)}{F(y|x)}-1\right]g(z)dz\right) \\
&= \frac{1}{\mathbb{E}(K_h(x-X_1))}\left(\int K_h(u)\left[\frac{\bar{F}(y|x-u)}{F(y|x)}-1\right]g(x-u)du\right).
\end{aligned}$$

We conclude, since the kernel K is compactly supported, that for some $R>0$,

$$\left|\frac{\mathbb{E}\left(\hat{\psi}_n(y,x)\right)}{F(y|x)\mathbb{E}\left(\hat{g}_n(x)\right)}-1\right| \leq \sup_{\{x',d(x,x')\leq hR\}}\left|\frac{\bar{F}(y|x')}{F(y|x)}-1\right| = A(y,y,x,h). \tag{30}$$

Now,

$$\begin{aligned}
&\left|\frac{\bar{F}_n(y|x)}{F(y|x)}-\frac{\mathbb{E}\left(\hat{\psi}_n(y,x)\right)}{F(y|x)\mathbb{E}\left(\hat{g}_n(x)\right)}\right| \\
&\leq \frac{\left|\hat{\psi}_n(y,x)-\mathbb{E}\left(\hat{\psi}_n(y,x)\right)\right|}{F(y|x)\hat{g}_n(x)} + \frac{\mathbb{E}\left(\hat{\psi}_n(y,x)\right)\left|\hat{g}_n(x)-\mathbb{E}\hat{g}_n(x)\right|}{F(y|x)\hat{g}_n(x)\mathbb{E}\hat{g}_n(x)} \\
&\leq \frac{\left|\hat{\psi}_n(y,x)-\mathbb{E}\left(\hat{\psi}_n(y,x)\right)\right|}{F(y|x)\hat{g}_n(x)} + (1+A(y,y,x,h))\frac{\mathbb{E}\left|\hat{g}_n(x)-\mathbb{E}\hat{g}_n(x)\right|}{\hat{g}_n(x)},
\end{aligned}$$

by (30). The last bound together with (30) and (29) prove Lemma 1. ■

References

[1] Abdi, S, Abdi, A, Dabo-Niang, S. and Diop, A. (2010). Consistency of a nonparametric conditional quantile estimator for random fields. Mathematical Methods and Statistics, 2010, Vol. 19, No. 1, 1-21.

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

[3] Berlinet, A., Gannoun, A. and Matzner-Lober, E. (2001). Asymptotic normality of convergent estimates of conditional quantiles. Statistics, 35, 139–169.

[4] Beirlant, J., Goegebeur, Y., Segers, J., and Teugels J. (2004). Statistics of extremes: Theory and applications, Wiley.

[5] 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.

[6] Chernozhukov, V. (2005). Extremal quantile regression, The Annals of Statistics, 33(2), 806–839.

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

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

[9] 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.

[10] 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.

[11] Dabo-Niang, S. and Laksaci, A. (2012). Nonparametric Quantile Regression Estimation for Functional Dependent Data. Comm. Statist. Theory Methods, 41-7: 1254-1268.

[12] Einmahl, U and Mason, D. M. (2000). An empirical process approach to the uniform consistency of kernel type function estimators. J. Theor. Probab. 13, 1-37.

[13] Einmahl, U and Mason, D. M. (2005). Uniform in bandwidth consistency of kernel type function estimators. Ann. Statist. Volume 33, Number 3, 1380-1403.

[14] Ezzahrioui, M. and Ould-Saïd, E. (2008). Asymptotic results of a nonparametric conditional quantile estimator for functional time series. Comm. Statist. Theory Methods, 37(16-17):2735-2759.

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

[16] Gardes, L., and Girard, S. (2010). Conditional extremes from heavy-tailed distributions: An application to the estimation of extreme rainfall return levels. Extremes, 13, 177–204.

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

[18] Gardes, L., Girard, S., and Lekina, A. (2010). Functional nonparametric estimation of conditional extreme quantiles. Journal of Multivariate Analysis, 101, 419–433.

[19] Girard, S., and Jacob, P. (2008). Frontier estimation via kernel regression on high power-transformed data. Journal of Multivariate Analysis, 99, 403–420.

[20] Girard, S., and Menneteau, L. (2005). Central limit theorems for smoothed extreme value estimates of Poisson point processes boundaries. Journal of Statistical Planning and Inference, 135, 433–460.

[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] Jurecková, J. (2007). Remark on extreme regression quantile. Sankhya, 69, Part 1, 87–100.

[23] Klass, M. (1985). The Robbins-Siegmund series criterion for partial maxima. Ann. Probab. Volume 13, Number 4, 1369-1370.

[24] Ledoux, M., Talagrand, M. (1991) Probability in Banach spaces Ergebnisse der Mathematik. Springer.

[25] Ould-Saïd, E., Yahia, D., and Necir, A. (2009). A strong uniform convergence rate of a kernel conditional quantile estimator under random left-truncation and dependent data. Electron. J. Statist. Volume 3, 426-445.

[26] Park, B.U. (2001). On nonparametric estimation of data edges. Journal of the Korean Statistical Society, 30, 265–280.

[27] Samanta, T. (1989). Non-parametric estimation of conditional quantiles. Statistics and Probability Letters, 7, 407–412.

[28] Smith, R.L. (1989). Extreme value analysis of environmental time series: An application to trend detection in ground-level ozone (with discussion). Statistical Science, 4, 367–393.

[29] Stone, C.J. (1977). Consistent nonparametric regression (with discussion). The Annals of Statistics, 5, 595–645.

[30] Stute, W. (1986). Conditional empirical processes. The Annals of Statistics, 14, 638–647.

[31] Tsay, R.S. (2002). Analysis of financial time series. Wiley, New York.