KERNEL CONDITIONAL QUANTILE ESTIMATOR 
UNDER LEFT TRUNCATION 
FOR FUNCTIONAL REGRESSORS 

Naceré Helal and Elias Ould Said 

Communicated by Zbigniew Szabotnik 

Abstract. Let $Y$ be a random real response which is subject to left-truncation by another random variable $T$. In this paper, we study the kernel conditional quantile estimation when the covariable $X$ takes values in an infinite-dimensional space. A kernel conditional quantile estimator is given under some regularity conditions, among which in the small-ball probability, its strong uniform almost sure convergence rate is established. Some special cases have been studied to show how our work extends some results given in the literature. Simulations are drawn to lend further support to our theoretical results and assess the behavior of the estimator for finite samples with different rates of truncation and sizes.

Keywords: almost sure convergence, functional variables, kernel estimator, Lynden-Bell estimator, small-ball probability, truncated data.

Mathematics Subject Classification: 62G05, 62G07, 62G20, 62M09.

1. INTRODUCTION

Let $(X,Y)$ be a couple of random variables (r.v.’s) valued in $\mathcal{F} \times \mathbb{R}$, where $\mathcal{F}$ is a semi-metric space, $d$ denoting the semi-metric and $Y$ being with distribution function (d.f.) $F$. Our purpose is to study the co-variation between $X$ and $Y$ via the quantile regression estimation when the interest r.v. is subject to random left truncation and the regressors take values in an infinite dimensional space.

It is well-known that, in nonparametric modeling, quantile regression is a useful analysis tool since it is less sensitive to outliers compared to classical regression.

For complete data and when the regressors are of functional type, and to the best of our knowledge the first references are [9] and [13], where the authors established the strong consistency.
Many authors considered this problem when the explanatory variable is of functional type. Without pretending to be exhaustive, we quote the monographs by [4,11,14,20,34] and the recent one [3], where many new results are presented.

When the interest random variable \( Y \) is subject to random censorship and for functional covariates, we can cite the works on strong consistency and asymptotic normality for the independent and identically distributed (i.i.d.) case by [7] and [8] and for time series by [18] and [19].

This paper is devoted to extend the results of [9] to the case where the interest random variable is subject to truncation which definition is given hereafter.

Truncation is another type of incomplete data which is completely different from censorship (see [37]). In some practical situations (some examples are given below) the interest r.v. \( Y \) is inferred by another r.v. in the following sense:

Let \( T \) be another real r.v. with unknown d.f. \( G \). We consider a sample \((Y_1,T_1),(Y_2,T_2),\ldots,(Y_N,T_N)\), \( N \) copies of \((Y,T)\), where the sample size \( N \) is fixed but unknown. In this model \((Y_i,T_i)\) is observed only if \( Y_i \geq T_i \) no data is collected otherwise. Then the observed sample size \( n \) is random (but known) with \( n \leq N \). In practice, such models are considered in many applications. We quote two examples from the literature:

- **AIDS study** [22]: Let \( W \) be the infection time, where 1 corresponds to January 1978 and let \( T \) be the incubation time in months for people who were infected by contaminated blood transfusions and developed AIDS by 1 July 1986. Since the total study period is 102 months only individuals with \( W + T < 102 \) were included in the sample. Then, letting \( Y = 102 - W \) yields the model described: \((Y,T)\) is observed only if \( T < Y \).

- **Retirement House** [23]: In a retirement centre, subjects are observed only if they live long enough to enter the retirement house. The lifetime \( Y \) is then left truncated by the retirement house entry age, \( T \). People who enter the retirement house earlier may get better medical attention and therefore live longer. On the other hand, people with poor health and a shorter expected lifetime may retire earlier.

Left truncation in studies of developmental processes is not just of theoretical interest: It can cause substantial bias if ignored. Other examples in which a large fraction of potential observations are left truncated are the rate of spontaneous abortion [31] and the age at menopause transition stages [16].

Now let \( \{(X_i,Y_i,T_i),1 \leq i \leq N\} \) be a sequence of iid random vectors, where \( X_i \) takes values in some normed space \((\mathcal{S},\|\cdot\|)\), \( Y_i \) and \( T_i \) are as before.

Since \( N \) is unknown and \( n \) is known (although random), our results will not be stated with respect to the probability measure \( P \) (related to the \( N \)-sample). Without possible confusion, we still denote \((Y_i,T_i), i = 1,2,\ldots,n, (n \leq N)\) the observed pairs from the original \( N \)-sample. In all the remaining parts of this paper we suppose that \( T \) is independent of \( Y \).

Let \( \mathbf{P}_n(\cdot) = \mathbf{P}(\cdot|n) \) be the conditional probability. Since, independence is preserved, we can write \( \mathbf{P}_n = \mathbf{P}^\otimes n \), where \( \mathbf{P}(\cdot) = \mathbf{P}_1(\cdot) = \mathbb{P}(\cdot | Y \geq T) \). Estimation results are then established considering \( n \to \infty \) and so are expressed with respect to the probability \( \mathbf{P} \). Finally, let \( \mathbf{E}(\cdot) \) and \( \mathbb{E}(\cdot) \) denote the respective expectation operators of \( \mathbf{P}(\cdot) \) and \( \mathbb{P}(\cdot) \).
Now, for $x \in S$, we consider the conditional probability distribution of $Y_i$ given $X_i = x$ by
\[ F(y|x) = P(Y_i \leq y|X_i = x), \] (1.1)
where $F$ is supposed to be strictly monotone.

Let $p \in (0, 1)$. The conditional quantile is defined by
\[ \zeta_p(x) = \inf \{y : F(y|x) \geq p\}. \] (1.2)
It is clear that an estimator of $\zeta_p(x)$ can easily be deduced from an estimator of $F(\cdot|x)$.

We point out that $\zeta_p(x)$ satisfies
\[ F(\zeta_p(x)|x) = p. \] (1.3)

The rest of the paper is as follows: in Section 2, we recall some background for truncated data. Section 3 contains the definition of our estimators of the conditional distribution and conditional quantile. The assumptions and main results are given in Section 4. In Section 5 a discussion is given about the assumptions and examples for the most important assumption (A1). Section 6 considers some particular cases. A simulation study is detailed in Section 7. Finally, the proofs are relegated to Section 8.

2. BACKGROUND FOR TRUNCATION MODELS

In this section we give the main definitions and results related to truncation models. We refer the reader to [27] or [32] for more details.

Recall that our original sample is $(X_i, Y_i, T_i)_{1 \leq i \leq N}$. Taking into account the truncation effect we denote by $(X_1, Y_1, T_1), \ldots, (X_n, Y_n, T_n)$ the actually observed sample (i.e., $Y_i \geq T_i, 1 \leq i \leq n$) and suppose that $\alpha := P(Y_1 \geq T_1) > 0$. Conditionally on the value of $n$, these observed random vectors are still iid (see [26]). Note here that $n$ is a real random variable itself and that from the strong law of large numbers (SLLN) we have, as $N \to \infty$,
\[ \bar{\alpha}_n = \frac{n}{N} \to \alpha \quad \text{P-a.s.} \] (2.1)

For any real d.f. $L$ denote the left and right endpoints of its support by $a_L = \inf \{t : L(t) > 0\}$ and $b_L = \sup \{t : L(t) < 1\}$, respectively.

The conditional joint distribution function (see [37]) of $(Y_1, T_1)$ is
\[ J^*(y, t) = P(Y_1 \leq y, T_1 \leq t|Y_1 \geq T_1) = P(Y_1 \leq y, T_1 \leq t) = \alpha^{-1} \int_{-\infty}^{y} G(t \wedge u)dF(u), \]
where $t \wedge u = \min(t, u)$.

Following [37] the distribution functions of $Y$ and $T$ are
\[ F^*(y) = \alpha^{-1} \int_{-\infty}^{y} G(u)dF(u) \quad \text{and} \quad G^*(t) = \alpha^{-1} \int_{-\infty}^{+\infty} G(t \wedge u)dF(u), \]
respectively, and are estimated by

\[ F^*_n(y) = n^{-1} \sum_{i=1}^{n} 1 \{ Y_i \leq y \} \quad \text{and} \quad G^*_n(t) = n^{-1} \sum_{i=1}^{n} 1 \{ T_i \leq y \}, \]

respectively, where \( 1_A \) is the indicator of the set \( A \). Note that, in what follows, the star notation \((*)\) relates to any characteristic of the actually observed data (that is, conditionally on \( n \)).

Define

\[ C(y) = G^*(y) - F^*(y) \]

\[ = P(T_1 \leq y \leq Y_1 | Y_1 \geq T_1) = \alpha^{-1} G(y) \left( 1 - F(y) \right), \quad y \in [a_F, +\infty) \]

and consider its empirical estimate

\[ C_n(y) = n^{-1} \sum_{i=1}^{n} 1 \{ T_i \leq y \leq Y_i \} = G_n^*(y) - F_n^*(y^-). \]

It is well known that the respective nonparametric maximum likelihood of \( F \) and \( G \) are the product-limit estimators given by

\[ F_n(y) = 1 - \prod_{Y_i \leq y} \left[ \frac{nC_n(Y_i) - 1}{nC_n(Y_i)} \right] \quad \text{and} \quad G_n(y) = \prod_{T_i > y} \left[ \frac{nC_n(T_i) - 1}{nC_n(T_i)} \right] \]

which were obtained by [29]. Their asymptotic properties were studied by [38] who showed that

\[ \sup_{y} |F_n(y) - F(y)| \overset{\text{P}-a.s.}{\to} 0 \quad \text{and} \quad \sup_{y} |G_n(y) - G(y)| \overset{\text{P}-a.s.}{\to} 0, \quad (2.2) \]

provided \( a_G \leq a_F, b_G \leq b_F \) and \( \int dF/G < \infty \).

Consequently, \( \alpha \) is identifiable only if \( a_G \leq a_F \) and \( b_G \leq b_F \). Note that the estimator \( \tilde{\alpha}_n \) defined in (2.1) cannot be calculated (since \( N \) is unknown). Another estimator, namely

\[ \alpha_n = \frac{G_n(y) \{ 1 - F_n(y^-) \}}{C_n(y)} \quad (2.3) \]

is used. In [17] the authors proved that \( \alpha_n \) does not depend on \( y \) and its value can then be obtained for any \( y \) such that \( C_n(y) \neq 0 \). Furthermore, they showed (in their Corollary 2.5) its \( \text{P} - a.s. \) consistency.

3. QUANTILE AND DISTRIBUTION FUNCTIONS ESTIMATORS

In this section we recall some results and then define our quantile estimator. Our estimation of the conditional distribution function is based on the choice of weights. These are obtained in [32].
Recall that, in the case of complete data, a well-known kernel estimator of the regression function in infinite dimensional space is based on the Nadaraya-Watson weights

$$W_{i,N}(x) = \frac{K\left(\frac{\|x - X_i\|}{h_{N,K}}\right)}{\sum_{i=1}^{N} K\left(\frac{\|x - X_i\|}{h_{N,K}}\right)} = \frac{(N \phi(h))^{-1} K\left(\frac{\|x - X_i\|}{h_{N,K}}\right)}{g_{N}(x)}$$  

(3.1)

associated to the $N$-sample (with convention $0/0 = 0$) and $\phi(\cdot)$ is a function which will be described later.

As $N$ is unknown, we have to adapt the weights given in [27] which gives the following values

$$\tilde{W}_{i,n}(x) = \frac{\alpha_n^{-1} K\left(\frac{\|x - X_i\|}{h_{n,K}}\right)}{\sum_{i=1}^{n} G_n^{-1}(Y_i) K\left(\frac{\|x - X_i\|}{h_{n,K}}\right)}.$$  

(3.2)

Note that, in this formula and the forthcoming, the sum is taken only for $i$ such that $G_n(Y_i) \neq 0$. This in turn yields an estimator of the conditional distribution function $F(y|x)$ given by

$$F_n(y|x) = \alpha_n \sum_{i=1}^{n} \tilde{W}_{i,n}(x) \frac{1}{G_n(Y_i)} H\left(\frac{y - Y_i}{h_{n,H}}\right)$$

$$= \frac{\sum_{i=1}^{n} G_n^{-1}(Y_i) K\left(\frac{\|x - X_i\|}{h_{n,K}}\right) H\left(\frac{y - Y_i}{h_{n,H}}\right)}{\sum_{i=1}^{n} G_n^{-1}(Y_i) K\left(\frac{\|x - X_i\|}{h_{n,K}}\right)} = \psi_n(x, y)$$  

(3.3)

where

$$\psi_n(x, y) = \frac{\alpha_n}{n \phi(h_{n,K})} \sum_{i=1}^{n} \frac{1}{G_n(Y_i)} K\left(\frac{\|x - X_i\|}{h_{n,K}}\right) H\left(\frac{y - Y_i}{h_{n,H}}\right)$$  

(3.4)

and

$$g_n(x) = \frac{\alpha_n}{n \phi(h_{n,K})} \sum_{i=1}^{n} \frac{1}{G_n(Y_i)} K\left(\frac{\|x - X_i\|}{h_{n,K}}\right).$$  

(3.5)

Here $K$ is a real-valued kernel function, $H$ is a d.f. and $h_{n,K} =: h_K$ (resp. $h_{n,H} =: h_H$) is a sequence of positive real numbers which goes to zero as $n$ goes to infinity.

Let $p \in (0, 1)$. A natural estimator of $\zeta_p(\cdot)$ is given by

$$\zeta_{p,n}(x) = \inf\{y : F_n(y|x) \geq p\}$$  

(3.6)

which satisfies

$$F_n(\zeta_{p,n}(x)|x) = p.$$  

(3.7)

We consider the partial derivative of $\psi_n(x, y)$

$$\frac{\partial \psi_n(x, y)}{\partial y} = \psi'_n(x, y) = \frac{\alpha_n}{n h_H \phi(h_K)} \sum_{i=1}^{n} \frac{1}{G_n(Y_i)} K\left(\frac{\|x - X_i\|}{h_{n,K}}\right) H'(\frac{y - Y_i}{h_{n,H}}),$$

where $H'$ is derivative of $H$. 


Making use of (1.3) and (3.7), we get
\[
F_n(\zeta_p(x)|x) - F(\zeta_p(x)|x) = (\zeta_p(x) - \zeta_{p,n}(x))f_n(\zeta_{p,n}(x)|x),
\]
where
\[
f_n(y|x) = \frac{\partial F_n(y|x)}{\partial y} = \frac{\psi'_n(x, \cdot) - \phi_n(x)}{g_n(x)}
\]
is the estimate of the conditional density of \( Y \) (that is \( f(\cdot|x) \)) given \( X = x \), and \( \zeta_{p,n}(x) \) lies between \( \zeta_p(x) \) and \( \zeta_{p,n}(x) \). Equation (3.8) shows how from the behavior of \( (F_n(\zeta_p(x)|x) - F(\zeta_p(x)|x)) \), we can get asymptotic results for \( (\zeta_p(x) - \zeta_{p,n}(x)) \).

4. ASSUMPTIONS AND MAIN RESULTS

From now on we assume that \( 0 = a_G < a_F \) and \( b_G \leq b_F \) and suppose that \( T_i \) and \( (X_i, Y_i), 1 \leq i \leq N \) are independent. We consider two real numbers \( a \) and \( b \) such that \( a_F < a < b < b_F \).

Let \( B(x, h) \) be the ball of center \( x \) and radius \( h \) and let \( W_i = ||x - X_i|| \). Then \( P(X_i \in B(x, h)) = P(W_i \leq h) = F_x(h) \). We denote by \( \Xi \) some compact subset of \( S \) and \( \gamma(u) \) to be a neighborhood of \( u \).

Our assumptions are gathered here for easy reference.

(A1) There exist three functions \( g(\cdot), \phi(\cdot) \) (which are assumed to be increasing) and \( \zeta_0(\cdot) \) such that

(i) \( F_x(h_K) = g(x)\phi(h_K) + o(\phi(h_K)) \),

(ii) for all \( u \in [0, 1] \), \( \lim_{h_K \to 0} \frac{\phi(u h_K)}{\phi(h_K)} = \lim_{h_K \to 0} \zeta(h_K)(u) = \xi(u) \),

(iii) \( \lim_{h_K \to 0} \sup_{x \in \Xi} \left| \frac{F_x(h_K)}{\phi(h_K)} - a_1 g(x) \right| = 0 \), where \( a_1 = K(1) - \int_0^1 K'(u) \xi\phi(u)du \).

(A2) The kernel \( K \) is nonnegative, with compact support \([0, 1]\) of class \( C^1 \) on \((0, 1)\), \( K(0) > 0 \) and \( K(1) > 0 \) and its derivative \( K' \) is such that \( -\infty < C_1 < K'(t) < C_2 < 0 \) on \((0, 1)\).

(A3) The conditional probability satisfies a Hölder condition with respect to each variable, that is, there exist strictly positive constants \( \beta \) and \( \gamma \) such that:

(i) for all \( (y_1, y_2) \in \mathbb{R}^2 \) and \( (x_1, x_2) \in \vartheta(x) \times \vartheta(x) \),

\[
|F(y_1|x_1) - F(y_2|x_2)| \leq C_2(||x_1 - x_2||^\beta + |y_1 - y_2|^\gamma),
\]

where \( \vartheta(x) \) is a neighborhood of \( x \), and \( C_2 \) is a constant which depends on \( x \),

(ii) \( \int |z|^\gamma H'(z)dz < \infty \), \( H' \) is bounded.

(A4) The bandwidths \( h_K \) and \( h_H \) satisfy:

(i) \( \lim_{n \to \infty} h_K = \lim_{n \to \infty} h_H = 0 \),

(ii) \( \lim_{n \to \infty} \frac{\log n}{\sup h_K} = 0 \),
Now we are in position to state our main results:

**Theorem 4.1.** Under Assumptions (A1)–(A4) we have

\[
\sup_{x \in \Xi} \sup_{y \in [a,b]} \left| F_n(y|x) - F(y|x) \right| = O \left( h_K^{3/2} + h_H^{3/2} \right) a.s. \text{ as } n \to \infty.
\]

**Theorem 4.2.** Under the same assumptions as those of Theorem 4.1 and if \( f(y|x) > 0 \) for all \( y \) in a neighborhood of \( \xi_p(x) \) and \( x \) fixed, we have

\[
\sup_{x \in \Xi} |\xi_{p,n}(x) - \xi_p(x)| = O \left( h_K^{3/2} + h_H^{3/2} \right) a.s. \text{ as } n \to \infty.
\]

5. DISCUSSION OF THE ASSUMPTIONS

**Remark 5.1.** Assumption (A1) (i) plays an important role in our methodology. It is known as (for small \( h \)) the concentration hypothesis acting on the distribution of \( X \) in infinite-dimension spaces. In many examples, around zero the small ball probability \( P(W_i < h) \) can be written approximately as the product of two independent functions \( g(\cdot) \) and \( \phi(h) \). This idea has been adopted by [30] who reformulated it from [15], as a condition with respect to the functional distribution of \( W_1 \). The increasing assumption on \( \phi(\cdot) \) implies that \( \xi(\cdot) \) is bounded and then integrable (a fortiori \( \xi_0(\cdot) \) is integrable).

We point out that even if \( F_n(x) \) depends strongly on the ball center \( x \), but the function can be given explicitly and asymptotically for several well-known continuous time processes. Some examples can be found in [12]).

Assumption (A2) concerns the kernel and states conditions which are standard in functional nonparametric estimation. Assumption (A3) is the only condition involving the conditional probability density of \( Y \) given \( X \). It means that \( F(\cdot|\cdot) \) and its derivatives satisfy the Lipschitz condition with respect to each variable. It is sufficiently weak not to introduce the notion of the density for the functional random variable \( X \), and so, the concentration condition (A1) might play an important role. Here we point out that our assumptions are very usual in the functional estimation for functional regressors (see, e.g., [9]).

Assumption (A4) concerns the choice of the bandwidth which is closely linked to the small balls probability.

**Remark 5.2.** Here we give an example of a Gaussian process which satisfies Assumption (A1). Other examples can be found in [9] or in [11]. It is shown in Corollary 4.7.8 in [2, p. 186] that the expression of the Onsager-Machlup function of the couple \( (x,z) \), for the Gaussian measures on a semi-normed space \( (\mathcal{F}, || \cdot ||) \), is given by

\[
F(x,z) = \log \left( \lim_{h \to 0} \frac{P(X \in B(x,h))}{P(X \in B(z,h))} \right) = \frac{1}{2}||\pi(z)||_H^2 - \frac{1}{2}||\pi(x)||_H^2,
\]
where $\|\cdot\|_H$ is the Hilbert norm on the Cameron-Martin space of $\mathcal{F}$ associated to a Gaussian measure, denoted by $H$, and $\pi(\cdot)$ is the orthogonal projection onto the orthogonal complement of the set $\{a \in H, \text{ such that } \|a\| = 0\}$. So, in this case, the small ball probability function $F_x(h)$ can be written as $F_x(h) = g(x)\phi(h) + o(\phi(h))$ with

$$g(x) = \exp\left(-\frac{1}{2}\pi(x)^2_H\right) \quad \text{and} \quad \phi(h) = \mathbb{P}(X \in B(0,r)).$$

It is well known that the latter can be quantified for several continuous time processes such as Gaussian processes where the function $\phi(h)$ has the following general form

$$\phi(h) = h^\gamma \exp\left(-\frac{C}{h^p}\right) + o\left(h^\gamma \exp\left(-\frac{C}{h^p}\right)\right) \quad \text{for some } \gamma > 0 \text{ and } p > 0.$$

6. APPLICATIONS

6.1. BACK TO FINITE DIMENSIONAL SETTING AND/OR COMPLETE DATA

On the one hand, it is clear that our Theorem 4.2 extends the result existing in finite dimensional space. Indeed, taking $\mathcal{S} = \mathbb{R}^d$ it is clear that $g(x)$ is the probability density of $X$, and we get a similar result of strong consistency to that obtained in [27], where only the pointwise consistency is stated (see Theorem 1 therein).

On the other hand, in the case where there is no truncation, that is, $T = -\infty$, and for the infinite dimensional setting, our result reduces to those obtained in [9] (see Theorem 4.2 therein).

Finally, in the case of the finite-dimensional setting and no truncation, our result reduces to that in [35] (see Theorem 1 therein).

6.2. PREDICTION

In prediction problems, the main idea is to use the conditional median $\mu(x) = \xi_p(x)$ (for $p = 1/2$) which is a good alternative to standard methods based on the conditional mean to achieve robustness. Note that $\mu(x)$ is estimated by $\hat{\mu}_n(x) = \xi_{1/2,n}(x)$. For each $n \in \mathbb{N}$ and $t \in \mathbb{R}$, let $X_i(t), i \in \{1, \ldots, n\}$, be functional random variables. For each curve $X_i(t)$, we have a real response $Y_i$ which corresponds to some modality of our problem. The main task is: given a new curve $X_{n+1}(t) = x_{new}$, can we predict the corresponding response $y_{new}$? This is a prediction problem for infinite dimensional explanatory random variables. The predictor estimator is obtained by computing the quantity: $y_{new} := \hat{\mu}_n(x_{new}) = \xi_{1/2,n}(x_{new})$.

Now applying Theorem 4.2, we have the following corollary:

**Corollary 6.1.**

$$\sup_{x_{new} \in \Xi} \left|\xi_{1/2,n}(x_{new}) - \xi_{1/2}(x_{new})\right| = O\left(h^\beta_K + h^\gamma_H\right) + O\left(\frac{\log n}{n\phi(h_K)}\right)^{1/2} \quad \text{a.s. as } n \to +\infty.$$
7. SIMULATION STUDY

In this section we implement our methodology to assess the quality of predictor for a finite sample. To this purpose, we consider the classical nonparametric functional regression model

\[ Y = R(X) + \varepsilon, \]

where \( \varepsilon \) are normally distributed and generated as \( \mathcal{N}(0, 0.5) \). We proceed with the following algorithm:

We fix the random size \( n \) (recall that \( n \) is known).

- **Step 1.** We generate the random variables \( T_1, X_1(t), t \in [0, 1] \), in the following manner: \( T_1 \sim \mathcal{N}(\mu, 1) \), and we adapt \( \mu \) in the such way to get a different rate of truncation, \( X_1(t) \) is generated as follows:
  \[ X_1(t) = AS_1(t) + (1 - A)S_2(t) \]
  with \( A \sim \mathcal{B}(0.5) \), \( S_1(t) = 2 - \cos(\pi t W) \), \( S_2(t) = \cos(\pi t W) \) with \( W \sim \mathcal{U}(0, 1) \). Furthermore, we simulate \( \varepsilon_1 \sim \mathcal{N}(0, 0.5) \).

- **Step 2.** We calculate \( Y_1 = R(X_1(t)) + \varepsilon_1 \), where \( R(X_1(t)) = \int_0^1 (X_1(t))^2 dt \) and \( \varepsilon_1 \) is as indicated before.

- **Step 3.** Test:
  We begin by setting:
  \( N = 0 \), \( j = 0 \).
  While \( j \leq n \):
  We put \( N = N + 1 \). We test: if \( Y_1 < T_1 \) we reject the triplet \( (X_1(t), Y_1, T_1) \).
  Otherwise, we keep the triplet \( (X_1(t), Y_1, T_1) \). At the end of this count we get the deterministic \( N \), which permits to get the rate of truncation \( \tau = \frac{n}{N} \). More precisely, the rate of the observed triplet.
  We continue the process until \( n = 100 \).
  We get random vectors \( (X_i(t), Y_i, T_i) \), \( i = 1, \ldots, 100 \).

- Then we calculate the Lynden-Bell estimator with the observed couple \( (Y_i, T_i) \) for \( i = 1, \ldots, n \).

We choose the quadratic kernel:

\[ K(x) = \frac{3}{2}(1 - x^2)\mathbb{I}_{[0,1]} \]

and the distribution function \( H(\cdot) \) is defined by

\[ \int_{-\infty}^{x} \frac{3}{4}(1 - t^2)\mathbb{I}_{[1-1,1]}(t)dt. \]

Another important point for ensuring good behavior is to use a norm that is well-adapted to the kind of data we have to deal with. We used the norm defined by the \( L_1 - L_2 \)-distance between the second derivatives of the curves (for further
discussion see \[9\]). This choice is motivated by the regularity of the curves \(X(t)\). Concerning the semi-metric we have downloaded the file for choosing a different semi-metric on the web-site of STAPH team. In this simulation, the choice of optimal bandwidth is much more crucial than that of the kernel. Recall that here we compute the \(MSE(h)\) on all the nearest-neighbors. For a fixed observation of index \(i_0\) we minimize on all the distances \(d(X_{i_0}, X_i)\), this to avoid calculating all the distances. In other words, we minimize on \(H_{i_0} = h_k\) the \(k^{th}\) distance in increasing order, for some \(k\), as the kernel is \((0,1)\)-support. This \(k^{th}\) index allows us to calculate our estimate from the \(k\)-nearest observations of \(X_{i_0}\). This explains the name of \(k\)-NN method.

- **Step 4.**
  We split our data into randomly chosen subsets \( I \) and \( J \):

  \[(X_j, Y_j, \delta_j)_{j \in J}\] training sample \( (X_i, Y_i, \delta_i)_{i \in I}\) test sample.

- **Step 5.**
  For each \(X_i\) in the test sample, we set: \(i_* = \text{Argmin}_{j \in J} d(X_i, X_j)\).

- **Step 6.**
  We calculate our estimators in two steps. First we calculate the estimator of the conditional distribution function given by formula (3.3) and the second step is to calculate the conditional median by putting \(p = \frac{1}{2}\) which is described below.

**Remark 7.1.** Here we point out that the main idea of the downloaded subroutine, is to use the \(k\)-NN estimator which is well adapted to our problem as mentioned before. For a more thorough idea and for the interested reader, we cite the following references: [6] for real data and [25] for functional type. For the choice of the bandwidth, Rachdi and Vieu [33] recently described an automatic procedure. This choice is mainly based on the functional version of the cross-validation (CV) method. In other words they built some data-driven criterion and showed that the rule is asymptotically optimal (see their Theorem 1). We think that we can use the same idea for the conditional quantile operator to choose an optimal bandwidth. This work is beyond of the goal of this paper and can constitute a future challenge.

Our main purpose in the first part is to illustrate the consistency property of our estimator by examining its accuracy as a predictor in the following situations:

**First study.**
We fix the percentage of truncation by taking \(\mu = 2\) and we vary the sample-size \(n = 100, 300, 500\). For each case, we split our data into two subsets (learning sample and test sample). From the learning sample we compute the predicted values. We calculate the estimator \(\xi_{1/2, n}(X_j)\) for all \(j \in J\) by using the training sample and formula (3.6). Then we calculate the predictor values using the test sample.

In the following figures we plot the predicted values estimated by \(\xi_{1/2, n}(X_j)\) versus the true values. The continuous line represents what would be the perfect prediction. Typically, the efficiency of the prediction method is quantified by the closeness of the round point to this continuous line.
Remark 7.2. We point out that unlike for complete and censored data, the explanatory random function \( X(t) \) is not kept in each step but only if the interest random variable \( Y \) is bigger than the truncated random variable \( T \).

Fig. 1. The observed curves \( n = 100, \text{MSE} = 0.20 \)

Fig. 2. \( n = 300, \text{MSE} = 0.15 \), \( n = 500, \text{MSE} = 0.12 \)

Figures 1 and 2 show the curves and the prediction which are evaluated by the MSE. We see clearly that the quality of fit increase with the size \( n \).
The second study.
We fix the sample-size $n = 300$ and we vary the percentage of truncation by taking $\mu = -1, 1, 3, 7$ and 10 (see Figures 3–7).

![Figure 3. TR = 0%, Complete data $\mu = -1$, $MSE = 0.18$](image)

![Figure 4. TR = 12%, $\mu = 1$, $MSE = 0.20$](image)

We see clearly that the quality of fit deteriorates with the increase of the percentage of truncation, which is predictable.
Fig. 5. $TR = 32\%, \mu = 3, MSE = 0.26$

Fig. 6. $TR = 66\%, \mu = 7, MSE = 0.48$

Fig. 7. $TR = 79\%, \mu = 10, MSE = 0.58$
8. AUXILLIARY RESULTS AND PROOFS

In what follows, we denote by $C$ any generic constant which can vary from line to line. When there is a need for two constant on the same line, we denote them by $C$ and $C'$. First we need to introduce some notations. Similarly to formulae (3.4) and (3.5) we define the following pseudo-estimators:

$$
\tilde{\psi}_n(x,y) = \frac{\alpha}{n \phi(h_K)} \sum_{i=1}^{n} \frac{1}{G(Y_i)} K \left( \frac{\|x - X_i\|}{h_K} \right) H \left( \frac{y - Y_i}{h_H} \right) \tag{8.1}
$$

and

$$
\tilde{g}_n(x) = \frac{\alpha}{n \phi(h_K)} \sum_{i=1}^{n} \frac{1}{G(Y_i)} K \left( \frac{\|x - X_i\|}{h_K} \right). \tag{8.2}
$$

**Proof of Theorem 4.1.** The proof is based on the following decomposition

$$
F_n(y|x) - F(y|x) = \psi_n(x,y) - F(y|x)
$$

$$
= \frac{1}{g_n(x)} \left\{ \psi_n(x,y) - \tilde{\psi}_n(x,y) + \tilde{\psi}_n(x,y) - E[\tilde{\psi}_n(x,y)] \right\}
$$

$$
- F(y|x) [g_n(x) - \tilde{g}_n(x) + \tilde{g}_n(x) - E[\tilde{g}_n(x)]]
$$

$$
- \frac{1}{g_n(x)} \left\{ \alpha a_1 \psi(x,y) - E[\tilde{\psi}_n(x,y)] - F(y|x) [\alpha a_1 g(x) - E[\tilde{g}_n(x)]] \right\}.
$$

The first lemma will play the role of the classical Böchner lemma in finite dimension.

**Lemma 8.1.** Suppose that Assumptions (A1) and (A2) hold. For all fixed $x$, we have

$$
\sup_{x \in \Xi} \left| \frac{1}{\phi(h_K)} E \left[ K \left( \frac{\|x - X_i\|}{h_K} \right) \right] - a_1 g(x) \right| \to 0 \quad \text{as} \quad n \to +\infty. \tag{8.4}
$$

**Proof of Lemma 8.1.** Integrating by parts and by (A1) (i) and (A2), we have

$$
\frac{1}{\phi(h_K)} E \left[ K \left( \frac{\|x - X_i\|}{h_K} \right) \right] = \frac{1}{\phi(h_K)} \int_{0}^{h_K} K \left( \frac{u}{h_K} \right) dP_{\|x - X_i\|}(u)
$$

$$
= \frac{K(1) F_x(h_K)}{\phi(h_K)} - \frac{1}{\phi(h_K)} \int_{0}^{1} K'(u) F_x(h_K u) du \tag{8.5}
$$

$$
= \left( K(1) - \int_{0}^{1} K'(u) \xi_{h_K}(u) du \right) [g(x) + o(1)].
$$

From Assumption (A1) (ii) it follows that the term in the right-hand-side of (8.5) converges uniformly to $a_1 g(x)$ as $n \to +\infty$. □
Lemma 8.2. Under Assumptions (A3) (ii), (A4) (ii) and (iii) we have

\[
\sup_{x \in \Xi} |\widetilde{g}_n(x) - E[\widehat{g}_n(x)]| = O\left(\frac{\log n}{n\phi(h_K)}\right)^{1/2} \quad \text{a.s. as } n \to +\infty.
\]

Proof of Lemma 8.2. Let \(r_n\) be a real sequence which goes to zero at a rate given below. The compact subset \(\Xi\) can be covered by a finite number \(\ell_n\) of balls of radius \(r_n\) that is \(\Xi \subset \bigcup_{k=1}^{\ell_n} B(x_k, r_n)\). Furthermore, we suppose that there exists a positive finite constant \(C\) such that \(\ell_n r_n = C\). For any \(x \in \Xi\), we set \(k(x) = \arg\min_{1 \leq k \leq \ell_n} d(x_k, x)\).

We have

\[
\begin{aligned}
|\widetilde{g}_n(x) - E[\widehat{g}_n(x)]| &\leq |\widetilde{g}_n(x) - \widetilde{g}_n(x_k(x))| \\
&\quad + |\widetilde{g}_n(x_k(x)) - E[\widehat{g}_n(x_k(x))]| + |E[\widetilde{g}_n(x_k(x))] - E[\widetilde{g}_n(x)]|
\end{aligned}
\]

=: \(G_1(x) + G_2(x) + G_3(x)\).

Then for all \(\varepsilon > 0\),

\[
\begin{aligned}
P\left\{\sup_{x \in \Xi} |\widetilde{g}_n(x) - E[\widehat{g}_n(x)]| > 3\varepsilon\right\} &\leq P\left\{\sup_{x \in \Xi} |G_1(x)| \geq \varepsilon\right\} + P\left\{\sup_{x \in \Xi} |G_2(x)| \geq \varepsilon\right\} \\
&\quad + P\left\{\sup_{x \in \Xi} |G_3(x)| \geq \varepsilon\right\}
\end{aligned}
\]

(8.6)

Clearly, \(J_1\) and \(J_3\) can be treated in the same manner. Then thanks to (A3) (ii) and taking \(r_n = \left(\frac{h_K}{n\phi(h_K)}\right)^{1/2}\), we get

\[
\sup_{x \in \Xi} |G_1(x)| \leq \frac{\alpha \|K\|_{\infty}}{h_K} \left\|x - x_k(x)\right\| = \frac{\alpha \|K\|_{\infty} r_n}{h_K \phi(h_K)} = \frac{\alpha \|K\|_{\infty}}{(n h_K^2 \phi(h_K))^{1/2}}.
\]

For \(n\) large enough we have \(J_1 = J_3 = 0\).

For \(J_2\), set

\[
V_i(x_k) = \frac{\alpha}{\phi(h_K)G(Y_i)} \left[ K \left(\frac{\|x_k - X_i\|}{h_K}\right) - E \left[K \left(\frac{\|x_k - X_i\|}{h_K}\right)\right]\right]
\]

and then \(G_2(x_k) = \frac{1}{n} \sum_{i=1}^{n} V_i(x_k)\). So, we can get directly that

\[
|V_i(x_k)| \leq \frac{\alpha \|K\|_{\infty}}{\phi(h_K)G(a_F)} = d
\]

and by using conditional expectation property we have

\[
E[|V_i(x_k)|^2] \leq \frac{C\alpha \|K\|_{\infty}}{\phi(h_K)G(a_F)} = \delta^2.
\]
Thus, the use of the classical Bernstein inequality allows us to write

\[
J_2 = \mathbb{P}\left\{ \max_{1 \leq k \leq \ell_n} \sup_{x \in B(x_k, r_n)} |G_2(x_k)| > \varepsilon \right\} \leq \sum_{k=1}^{\ell_n} \mathbb{P}\left\{ \sup_{x \in B(x_k, r_n)} |G_2(x_k)| > \varepsilon \right\} \\
\leq \ell_n \max_{k=1, \ldots, \ell_n} \mathbb{P}\{ |G_2(x_k)| > \varepsilon \} \\
\leq 2 C \frac{\varepsilon}{\ell_n} \exp\left( -\frac{\ell_n^2 \phi(h_K) G(aF)}{\alpha \|K\|_{\infty}} \right).
\]

Taking \( r_n \) as before and \( \varepsilon = \varepsilon_0 \left( \frac{\log n}{n \phi(h_K)} \right)^{1/2} \) for \( \varepsilon_0 > 0 \), we get

\[
J_2 \leq 2C \left( \frac{\log n}{nh_K \phi(h_K)} \right)^{1/2} \times \frac{1}{(nh_K \log n)^{1/2}} \times n^{-1} \exp\left( -C' \frac{\log n}{\alpha \|K\|_{\infty}} \right).
\]

Thus, by choosing

\[
\frac{C' \varepsilon_0^2 G(aF)}{\alpha \|K\|_{\infty}} = 3 + \gamma \quad \text{with} \quad \gamma > 0,
\]

we have

\[
J_2 \leq 2C \left( \frac{\log n}{nh_K \phi(h_K)} \right)^{1/2} \times \frac{1}{(nh_K \log n)^{1/2}} \times n^{-1-\gamma}. \tag{8.7}
\]

Thanks to (A4) (ii) and (iii), the upper bound becomes a general term of a convergent Riemann series which in turn, by Borel-Cantelli’s lemma gives the result.

**Lemma 8.3.** Under Assumptions (A2) and (A4) we have

\[
\sup_{x \in \Xi} |g_n(x) - \tilde{g}_n(x)| = O(n^{-1/2}) \quad \text{as} \ n \to +\infty.
\]

**Proof of Lemma 8.3.** We have

\[
\sup_{x \in \Xi} |g_n(x) - \tilde{g}_n(x)|
\]

\[
= \sup_{x \in \Xi} \left| \alpha_n \sum_{i=1}^{n} \frac{1}{G_n(Y_i)K} \left( \frac{\|x - X_i\|}{h_K} \right) - \frac{\alpha}{n \phi(h_K)} \sum_{i=1}^{n} G(Y_i) \left( \frac{\|x - X_i\|}{h_K} \right) \right|
\]

\[
\leq \left( \frac{\|\alpha_n - \alpha\|}{G_n(aF)} + \frac{\alpha}{G_n(aF)G(aF)} \sup_{a \leq y \leq b} |G_n(y) - G(y)| \right) \sup_{x \in \Xi} |v_n(x)|
\]

with

\[
v_n(x) = \frac{1}{n \phi(h_K)} \sum_{i=1}^{n} \left( \frac{\|x - X_i\|}{h_K} \right).
\]

From Theorem 3.2 in [17] we have \( |\alpha_n - \alpha| = O(n^{-1/2}), \mathbb{P} \text{ - a.s.} \). Moreover, \( G_n(aF) \mathbb{P}^{-aF} G(aF) > 0 \). On the other hand, \( \sup_{a \leq y \leq b} |G_n(y) - G(y)| = O(n^{-1/2}), \mathbb{P} \text{ - a.s.} \) (see Remark 6 in [38]). Using Lemma 10 of [10] and under our assumptions, we get easily \( \sup_{x \in \Xi} |v_n(x)| = o(1) \) which permits us to conclude the proof.
The following lemma proves the asymptotic unbiasedness of the estimator $\tilde{g}_n(x)$.

**Lemma 8.4.** Under the same assumptions as those of Lemma 8.1 then we have
\[
\sup_{x \in \Xi} |\mathbb{E}[\tilde{g}_n(x)] - \alpha_1 g(x)| \rightarrow 0 \quad \text{as } n \rightarrow +\infty.
\]

**Proof of Lemma 8.4.** We have
\[
\mathbb{E}[\tilde{g}_n(x)] = \frac{\alpha}{\phi(h_K)} \mathbb{E} \left[ \frac{1}{G(Y_1)} K \left( \left\| \frac{x - X_1}{h_K} \right\| \right) \right] \\
= \frac{\alpha}{\phi(h_K)} \mathbb{E} \left[ \frac{1}{G(Y_1)} K \left( \left\| \frac{x - X_1}{h_K} \right\| \right) \mathbb{E} \left[ 1_{\{Y_1 \geq T_1\}} \left| X_1, Y_1 \right| \right] \right] \\
= \frac{\alpha}{\phi(h_K)} \mathbb{E} \left[ K \left( \left\| \frac{x - X_1}{h_K} \right\| \right) \right].
\]

Making use of Lemma 8.1 we get the result. \(\square\)

**Lemma 8.5.** Under the same assumptions as those of Lemma 8.1 we have
\[
\sup_{x \in \Xi} \sup_{a \leq y \leq b} |\psi_n(x, y) - \tilde{\psi}_n(x, y)| = O(n^{-1/2}) \quad \text{a.s. as } n \rightarrow +\infty.
\]

**Proof of Lemma 8.5.**
\[
|\psi_n(x, y) - \tilde{\psi}_n(x, y)| = \left| \frac{\alpha_n}{n\phi(h_K)} \sum_{i=1}^n \frac{1}{G(Y_i)} K \left( \left\| \frac{x - X_i}{h_K} \right\| \right) H \left( \frac{y - Y_i}{h_H} \right) \right| \\
- \frac{\alpha}{n\phi(h_K)} \sum_{i=1}^n \frac{1}{G(Y_i)} K \left( \left\| \frac{x - X_i}{h_K} \right\| \right) H \left( \frac{y - Y_i}{h_H} \right) \\
\leq \left\{ \frac{\alpha_n - \alpha}{G_n(a_F)} + \frac{\alpha}{G_n(a_F)G(a_F)} \sup_{a \leq y \leq b} |G_n(y) - G(y)| \right\} |\Psi_n(x, y)|
\]

with $\Psi_n(x, y) = \frac{1}{n\phi(h_K)} \sum_{i=1}^n K \left( \left\| \frac{x - X_i}{h_K} \right\| \right) H \left( \frac{y - X_i}{h_H} \right)$.

In the same way as in the Lemma 8.3, in conjunction with Lemma 3.3 in [9] for $j = 0$ we obtain
\[
\sup_{x \in \Xi} \sup_{a \leq y \leq b} |\psi_n(x, y) - \tilde{\psi}_n(x, y)| = O \left( n^{-1/2} \right). \quad \square
\]

**Lemma 8.6.** Under the same assumptions as those of Lemma 8.1 and (A4) (i), we have
\[
\sup_{x \in \Xi} \sup_{a \leq y \leq b} |\tilde{\psi}_n(x, y) - \mathbb{E}[\tilde{\psi}_n(x, y)]| = O \left( \frac{\log n}{n\phi(h_K)} \right)^{1/2} \quad \text{a.s. as } n \rightarrow +\infty.
\]
Proof of Lemma 8.6. For \( \Xi \), we use the same covering as in Lemma 8.2. Since \([a,b]\) is a fixed compact subset of \( \mathbb{R} \), it can be covered by a finite number \( s_n \) of intervals of length \( s_n \) of intervals \( [a,b] \subset \bigcup_{i=1}^{s_n} I_{t_i} \), where \( I_{t_i} = (y_t - u_n, y_t + u_n) \) and \( s_n = \frac{24}{u_n} \). Taking \( y_t = \arg\min_{t \in \{t_1, \ldots, t_{s_n}\}} |y - t| \) we have

\[
\sup_{x \in \Xi} \sup_{y \in [a,b]} |\tilde{\psi}_n(x, y) - E[\tilde{\psi}_n(x, y)]| \\
\leq \sup_{x \in \Xi} \sup_{y \in [a,b]} |\tilde{\psi}_n(x, y) - \tilde{\psi}_n(x_k, y)| \\
+ \sup_{x \in \Xi} \sup_{y \in [a,b]} |\tilde{\psi}_n(x_k, y) - \tilde{\psi}_n(x_k, y_t)| + \sup_{x \in \Xi} \sup_{y \in [a,b]} |\tilde{\psi}_n(x_k, y_t) - E[\tilde{\psi}_n(x_k, y_t)]| \\
+ \sup_{x \in \Xi} \sup_{y \in [a,b]} |E[\tilde{\psi}_n(x_k, y_t)] - E[\tilde{\psi}_n(x_k, t)]| + \sup_{x \in \Xi} \sup_{y \in [a,b]} |E[\tilde{\psi}_n(x_k, t)] - E[\tilde{\psi}_n(x, t)]| .
\]

(8.8)

Clearly, \( T_1 \) and \( T_5 \) can be treated in the same manner. We deal with \( T_1 \).

As the d.f. \( H \) is bounded by 1, then we come back directly to the Lemma 8.2. Then with the same choice of \( \varepsilon \), we have

\[
P \left\{ T_1 > \varepsilon_0 \left( \frac{\log n}{n\phi(h_K)} \right)^{1/2} \right\} = P \left\{ T_5 > \varepsilon_0 \left( \frac{\log n}{n\phi(h_K)} \right)^{1/2} \right\} = 0. \tag{8.9}
\]

Also, as \( T_2 \) and \( T_4 \) can be treated in the same manner, we deal only with \( T_2 \).

Using the Lipschitz argument we get

\[
\sup_{x \in \Xi} \sup_{y \in [a,b]} |\tilde{\psi}_n(x_k, y) - \tilde{\psi}_n(x_k, y_t)| \\
\leq \sup_{x \in \Xi} \sup_{y \in [a,b]} \frac{\alpha}{n\phi(h_K)} \sum_{i=1}^{n} G(Y_i) \left| H \left( \frac{y - Y_i}{h_H} \right) - H \left( \frac{y_t - Y_i}{h_H} \right) \right| K \left( \frac{\|x_k - X_i\|}{h_K} \right) \\
\leq \sup_{x \in \Xi} \sup_{y \in [a,b]} \frac{C|y - y_t|}{h_H} \left( \frac{\alpha}{n\phi(h_K)} \sum_{i=1}^{n} G(Y_i) K \left( \frac{\|x_k - X_i\|}{h_K} \right) \right) \\
\leq C \frac{\alpha u_n}{h_H G(a_F)} \sup_{x \in \Xi} v_n(x_k).
\]

By Lemma 8.1, the upper bound of the latter becomes

\[
O \left( \frac{Cu_n}{h_H G(a_F)} \alpha a_1 g(x) \right) = O \left( \frac{u_n}{h_H} \right).
\]
Choosing $u_n = n^{-\nu - \frac{1}{2}}$ with $\nu > 0$ and by (A4) (iii), we get $\frac{u_n H}{H_k} = O \left( \left( \frac{\log n}{n\phi(hK)} \right)^{1/2} \right)$.

Then for $n$ large enough and all $\varepsilon > 0$, we get

$$ P \left( T_2 > \varepsilon \left( \frac{\log n}{n\phi(hK)} \right)^{1/2} \right) = P \left( T_4 > \varepsilon \left( \frac{\log n}{n\phi(hK)} \right)^{1/2} \right) = 0. \quad (8.10) $$

Concerning $T_3$, clearly we have

$$ P \left( T_3 > \varepsilon \right) \leq \sum_{k=1}^{\ell_n} \sum_{y=1}^{\ell_n} \sup_{x \in B(x_k, r_n), \ y \in \{t_1, \ldots, t_n\}} \left| \psi_n(x_k, y_t) - E[\psi_n(x_k, y_t)] \right| \varepsilon \leq s_n \ell_n \left( \max_{n=1}^{\ell_n} \sup_{y \in \{t_1, \ldots, t_n\}} P \left( \left| \psi_n(x_k, y_t) - E[\psi_n(x_k, y_t)] \right| > \varepsilon \right) \right). $$

Let

$$ \lambda(x, y) = K \left( \frac{\|x - X_i\|}{h} \right) H \left( \frac{y_t - Y_i}{H} \right) - E \left( K \left( \frac{\|x - X_i\|}{h} \right) H \left( \frac{y_t - Y_i}{H} \right) \right). $$

By using similar arguments to those of the proof of Lemma 8.2 and by the fact that $H \leq 1$, we deduce that $E[\lambda] \leq 2 \|K\|_{\infty} |H| \leq C$ and $E[\lambda]^2 \leq \frac{C^2 n \|K\|_{\infty}}{\alpha \|F\|_{\infty}}.$

Now, we apply Bernstein’s exponential inequality to get

$$ P \left( T_3 > \varepsilon \right) \leq 2s_n \ell_n \exp \left\{ - \frac{-C' n \varepsilon^2 \phi(hK) G(a_F)}{2\alpha \|K\|_{\infty}} \right\} = \frac{2C}{r_n u_n} \exp \left\{ - \frac{-C' n \varepsilon^2 \phi(hK) G(a_F)}{\alpha \|K\|_{\infty}} \right\}. $$

The same choice of $r_n$, $u_n$ and $\varepsilon$ give that the upperbound of the latter becomes

$$ \frac{2C}{(n^2 h^2 \phi(hK))^{1/2}} \times n^{\nu + 2 - \frac{C' n \varepsilon^2 G(a_F)}{\alpha \|K\|_{\infty}}}. $$

Using the same choice of the power of $n$ as in the proof of (8.7), that is, $\frac{C' n \varepsilon^2 G(a_F)}{\alpha \|K\|_{\infty}} = 3 + \nu + \gamma$ with $\gamma > 0$, we have

$$ P \left( T_3 > \varepsilon \right) \leq 2C \left( \frac{\log n}{nH_k \phi(hK)} \right)^{1/2} \times \left( \frac{1}{nhK \log n} \right)^{1/2} \times n^{-1 - \gamma}. \quad (8.11) $$

Thanks to (A4) (ii) and (iii), the upper bound becomes a general term of a convergent Riemann series which in turn, by Borel-Cantelli’s lemma gives the result. Now, Lemma 8.6 can be easily deduced from (8.8)–(8.11).

**Lemma 8.7.** Under the same assumptions as for Lemma 8.2 there exists $\nu > 0$ such that

$$ \sum_{n=1}^{\infty} P \left( \inf_{x \in \Xi} \tilde{g}_n(x) \leq \nu \right) < \infty.$$
Proof of Lemma 8.7. From the inequality
\[ \inf_{x \in \Xi} \tilde{g}_n(x) \geq \inf_{x \in \Xi} E[\tilde{g}_n(x)] - \sup_{x \in \Xi} |\tilde{g}_n(x) - E[\tilde{g}_n(x)]| \]
and Lemma 8.2 we get the result. \qed

The following lemma states the asymptotic unbiasedness of the pseudo-estimator \( \tilde{\psi}_n(x) \).

**Lemma 8.8.** Under Assumptions (A1)–(A4) we have
\[ \sup_{a < y < b} |E[\tilde{\psi}_n(x, y)] - \alpha a_1 \psi(x, y)| = O \left( h_K^\beta + h_H^\gamma \right) \quad \text{as } n \to +\infty. \]

**Proof of Lemma 8.8.** We can write
\[
E[\tilde{\psi}_n(x, y)] = \frac{\alpha}{\phi(h_K)} E \left[ \frac{1}{G(Y_1)} K \left( \frac{||x - X_1||}{h_K} \right) H \left( \frac{y - Y_1}{h_H} \right) \right] 
= \frac{\alpha}{\phi(h_K)} E \left[ \frac{1}{G(Y_1)} K \left( \frac{||x - X_1||}{h_K} \right) E \left[ H \left( \frac{y - Y_1}{h_H} \right) \right] |X_1\right].
\]

Here and after, we suppose that \( X_1 \in B(x, h_K) \).

Moreover, we have by integration by parts and changing variables,
\[
E \left[ H \left( \frac{y - Y_1}{h_H} \right) \right] |X_1\right] = \int_{\mathbb{R}} H \left( \frac{y - u}{h_H} \right) f(u|X_1) du 
= \int_{\mathbb{R}} H'(z) F(y - zh_H|X_1) dz 
= \int_{\mathbb{R}} H'(z) |F(y - zh_H|X_1) - F(y|x)| dz + F(y|x).
\]

The last equality is due to the fact that \( H'(\cdot) \) is a probability density.

Thus, we have
\[
E[\tilde{\psi}_n(x, y)] = \frac{\alpha}{\phi(h_K)} E \left[ \frac{1}{G(Y_1)} K \left( \frac{||x - X_1||}{h_K} \right) \int_{\mathbb{R}} H'(z) |F(y - zh_H|X_1) - F(y|x)| dz \right] 
+ \frac{\alpha}{\phi(h_K)} F(y|x) E \left[ \frac{1}{G(Y_1)} K \left( \frac{||x - X_1||}{h_K} \right) \right] 
= L_1 + L_2.
\]

Making use of Lemma 8.4, \( L_2 \) tends to \( \alpha a_1 \psi(x, y) \) as \( n \) goes to infinity.

Now let us turn to \( L_1 \). By Assumption (A3) (i), we have
\[
\int_{\mathbb{R}} H'(z) |F(y - zh_H|X_1) - F(y|x)| dz \leq C_2 \int_{\mathbb{R}} H'(z) (h_K^\beta + |z|^\gamma h_H^\gamma) dz 
\leq C_2 h_K^\beta + C_2 h_H^\gamma \int_{\mathbb{R}} |z|^\gamma H'(z) dz.
\]
Making use of (A3) (ii) and Lemma 8.4, it is clear that, \( L_1 \) tends to zero as \( n \) goes to infinity, this completes the proof of Lemma 8.8.

Now to finish the proof of Theorem 4.1, we use the following inequality

\[
\sup_{x \in \Xi} \sup_{y \in [a, b]} |F_n(y|x) - F(y|x)| \\
\leq \frac{1}{\inf_{x \in \Xi} \mathbb{E}[g_n(x)]} \sup_{x \in \Xi} \sup_{y \in [a, b]} \left| \psi_n(x, y) - \tilde{\psi}_n(x, y) + \tilde{\psi}_n(x, y) - \mathbb{E}[\tilde{\psi}_n(x, y)] \right| \\
- F(y|x) \left| g_n(x) - \tilde{g}_n(x) + \tilde{g}_n(x) - \mathbb{E}[\tilde{g}_n(x)] \right| \\
+ \frac{1}{\inf_{x \in \Xi} \mathbb{E}[g_n(x)]} \sup_{x \in \Xi} \sup_{y \in [a, b]} \left| \alpha_1 \psi_n(x, y) - \mathbb{E}[\tilde{\psi}_n(x, y)] - F(y|x) [\alpha_1 g(x) - \mathbb{E}[\tilde{g}_n(x)]] \right|
\]

which together with Lemmas 8.1–8.8 concludes the proof of Theorem 4.1.

**Proof of Theorem 4.2.** The proof of this theorem is based in the following decomposition. As \( F(\cdot|x) \) is a distribution function with a unique quantile of order \( p \), then for any \( \epsilon > 0 \), let

\[
\eta(\epsilon) = \min \{ F(\zeta_p(x) + \epsilon|x) - F(\zeta_p(x)|x), F(\zeta_p(x)|x - F(\zeta_p(x) - \epsilon|x)) \}
\]

then

\[
\forall \epsilon > 0 \exists \forall y > 0 |\zeta_p(x) - y| \geq \epsilon \Rightarrow |F(\zeta_p(x)|x) - F(y|x)| \geq \eta(\epsilon).
\]

Now, using (1.3) and (3.7) we have

\[
\sup_{x \in \Xi} |F(\zeta_{p,n}(x)|x) - F(\zeta_p(x)|x)| = |F(\zeta_{p,n}(x)|x) - F_n(\zeta_{p,n}(x)|x)| \\
\leq \sup_{x \in \Xi} \sup_{y \in [a, b]} |F_n(y|x) - F(y|x)|.
\]

The consistency of \( \zeta_{p,n}(x) \) follows then immediately from Theorem 4.1, the continuity of \( F(\cdot|x) \) and the following inequality:

\[
\sum_{n \geq 1} \mathbb{P}(sup_{x \in \Xi} |\zeta_{p,n}(x) - \zeta_p(x)| \geq \epsilon) \leq \sum_{n \geq 1} \mathbb{P}(sup_{x \in \Xi} \sup_{y \in [a, b]} |F_n(y|x) - F(y|x)| \geq \epsilon).
\]

**Acknowledgments**

The authors would like to thank the referee for his/her remarks and comments which permitted us to improve the quality of the first and second version of the paper, in particular the study of the uniformity with respect to the regressors and add the simulation part to show how the methodology can be implemented.
REFERENCES


Kernel conditional quantile estimator under left truncation for functional regressors


Nacéra Helal

Département de Mathématiques
Université Djillali Liabès
BP 89, 22000, Sidi Bel Abbès, Algeria

Elias Ould Saïd

Université Lille Nord de France
F-59000 Lille, France
ULCO, LMPA, CS: 80699 Calais, France

Received: February 4, 2014.
Revised: February 26, 2015.
Accepted: March 21, 2015.