KYBERNETIKA — VOLUME *59* (2023), NUMBER 4, PAGES 548–574

LOCAL LINEAR ESTIMATION OF THE CONDITIONAL MODE UNDER LEFT TRUNCATION FOR FUNCTIONAL REGRESSORS

Halima Boudada and Sarra Leulmi

In this work, we introduce a local linear estimator of the conditional mode for a random real response variable which is subject to left-truncation by another random variable where the covariate takes values in an infinite dimensional space. We first establish both of pointwise and uniform almost sure convergences, with rates, of the conditional density estimator. Then, we deduce the strong consistency of the obtained conditional mode estimator. We finally illustrate the outperformance of our method with respect to the kernel one through a simulation study for a finite sample with different rates of truncation and sizes.

Keywords: functional regressors, left truncation model, conditional mode, almost sure convergence, local linear estimator

Classification: 62G07, 62G20, 62R10, 62N99

1. INTRODUCTION

The functional data is widely used in practice. Hence, many statisticians are motivated to study the relationship between a real response variable (r.v.) and a covariate that takes its values in an infinite dimensional space. The classical way to investigate this link is the regression method which is based on the conditional expectation. However, two other techniques are used: the conditional quantile based on the conditional distribution function and the conditional mode based on the conditional density.

One of the pioneer works in this research field is the monograph of [11], where the authors established the pointwise almost-complete (a.co.) convergence for different non-parametric kernel type estimators. However, lot of works show that the local linear method not only generalizes the kernel's one but it is also in possession of superior bias properties. We refer to [9] for more comparison between the two methods.

In the functional data case and when the response random variable is complete, the local linear estimation is studied by [18, 19] for conditional regression function, [21] for conditional distribution function and[6] for conditional density function. Nevertheless, in diverse fields, such as medicine, biology, public health, epidemiology, engineering, economics and demography, the response random variable may be incomplete and subject to random censorship or truncation model.

DOI: 10.14736/kyb-2023-4-0548

Using the non-parametric kernel method, [14] studied the strong consistency of the conditional quantile estimator for functional censored dependent data. Furthermore, based on the so-called synthetic data, [1] defined an M-estimator for the functional regression and established the strong consistency and the asymptotic normality for censored independent data. [20] examined, under mild conditions, the almost complete consistency and the asymptotic normality of the estimator of the relative error in functional regression and censored data. Whereas, for functional truncated data, [13] established the strong uniform almost sure (a.s.) convergence rate of a conditional quantile estimator. [7] studied the almost complete convergence rate of estimators for the Ψ-regression model. The almost complete convergence rate of the M-estimator of the regression is established by [8], when the sample is an α-mixing sequence.

Based on the local linear approach, we can refer to [16, 17], where the rates of the pointwise and the uniform almost complete convergences of the conditional quantile and the regression estimators are obtained for functional censored independent data. [3] constructed a conditional density's estimator and established its pointwise almost sure convergence, for functional censored data under dependence condition. While for functional truncated data, the estimation of the generalized regression is introduced in [5] where the author is studied its pointwise and uniform almost sure convergences.

Unfortunately, the estimation of the regression is sensitive to outliers and even inappropriate especially when the distribution is strongly asymmetric. The estimation of the conditional mode then constitutes an interesting alternative. It is more robust and useful to better understand the relationship between the response variable and the set of covariates in comparison with the regression estimation methods. To our knowledge, the local linear estimation of the conditional mode under left truncation for functional regressors has not been studied in statistical literature, what prompted us to study this topic and organize our work as follows. in Section 2, we recall some background for the truncated data and we present the new local linear estimator of the conditional mode. Then, we establish in section 3 the pointwise almost sure convergence of the conditional density estimator. Subsection 4.1 is devoted to its uniform version. Moreover, we apply the previous results on the conditional mode estimator, in Subsection 4.2. A simulation study and a real data application are made to illustrate the good accuracy of the proposed estimator in Section 5. Finally, proofs of our theoretical results are relegated to the Appendix.

2. MODEL AND ESTIMATION

Let us consider N independent pairs of random variables $(X_i, Y_i)_{i=1,\dots,N}$ which are assumed drawn from the pair (X, Y). This later is valued in $\mathcal{F} \times \mathbb{R}$, where $\mathcal{F}$ is a semi metric space equipped with a semi metric d and Y being with unknown distribution function (d.f.) F. In the complete case, the local linear estimator of the conditional density function $f(y|x)$ (see [6]) is given by

$$\tilde{f}_n(y|x) = \frac{\sum_{i,j=1}^{n} \Delta_{ij}(x) H(h_H^{-1}(y - Y_j))}{h_H \sum_{i,j=1}^{n} \Delta_{ij}(x)}, \qquad \left(\frac{0}{0} := 0\right), \tag{1}$$

with

$$\Delta_{ij}(x) := \beta(X_i, x)\left(\beta(X_i, x) - \beta(X_j, x)\right) K(h_K^{-1} d(X_i, x)) K(h_K^{-1} d(X_j, x)), \qquad (2)$$

where K and H are kernels, the bandwidth $h_K := h_{K,n}$ (resp. $h_H := h_{H,n}$) is a sequence of strictly positive real numbers which plays a smoothing parameter role and $\beta(\cdot,\cdot)$ is a known operator from $\mathcal{F} \times \mathcal{F}$ into $\mathbb{R}$ such that, $\forall x \in \mathcal{F}$, $\beta(x, x) = 0$.

Let now consider $(T_i)_{i=1,\ldots,N}$ a sample of independent and identically distributed (i.i.d.) random variables that are distributed as T which has unknown d.f G. T is supposed independent of (X, Y). N is unknown but deterministic. In the left truncation model, the lifetime Y_i and the truncation r.v. T_i are both observable only when $Y_i \geq T_i$. We shall denote $(Y_i, T_i)_{i=1,\ldots,n}$; $(n \leq N)$ the actual observed sample which its size n, as a consequence of truncation, is a binomial r.v. with parameters N and $\mu = \mathbb{P}(Y \geq T)$. It is clear that if $\mu = 0$, no data can be observed, and therefore, we suppose throughout this article that $\mu > 0$.

By the strong law of large numbers, we have

$$\widehat{\mu}_n := \frac{n}{N} \to \mu, \mathbb{P} - a.s.$$

We point out that if the original data $(Y_i, T_i)_{i=1,\ldots,N}$ are i.i.d., the observed data (Y_i, T_i), $i = 1, 2, \ldots, n$ are still i.i.d. ([15]). Under random left truncation model, following [22], the d.f.s of Y and T are expressed respectively as

$$F^*(y) = \mu^{-1} \int_{-\infty}^{y} G(u)\,\mathrm{d}F(u) \qquad \text{and} \qquad G^*(t) = \mu^{-1} \int_{-\infty}^{\infty} G(t \wedge u)\,\mathrm{d}F(u),$$

where $t \wedge u = \min(t, u)$ and are estimated by their empirical estimators

$$F_n^*(y) = n^{-1} \sum_{i=1}^{n} 1_{\{Y_i \leq y\}} \qquad \text{and} \qquad G_n^*(t) = n^{-1} \sum_{i=1}^{n} 1_{\{T_i \leq t\}},$$

where 1_A denotes the indicator function of the set A.

Define

$$C(y) := G^*(y) - F^*(y) = \mu^{-1} G(y)(1 - F(y)),$$

the empirical estimator of $C(y)$ is defined by

$$C_n(y) = n^{-1} \sum_{i=1}^{n} 1_{\{T_i \leq y \leq Y_i\}}.$$

The non-parametric maximum likelihood estimators of F and G are given respectively by

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

According to [12], μ can be estimated by

$$\mu_n = C_n^{-1}(y) G_n(y)(1 - F_n(y))$$

which is independent of y.

Our results will be stated with respect to the conditional probability $\mathbf{P}(\cdot)$ related to the n-sample instead of the probability measure $\mathbb{P}(\cdot)$ related to the N-sample. We donate by $\mathbf{E}$ and $\mathbb{E}$ the respective expectation operators of $\mathbf{P}(.)$ and $\mathbb{P}(\cdot)$.

For any d.f. L, let $a_L = \inf\{y : L(y) > 0\}$ and $b_L = \sup\{y : L(y) < 1\}$ be its two endpoints. The asymptotic properties of F_n, G_n and μ_n are obtained only if $a_G \leq a_F$ and $b_G \leq b_F$, that's why we consider this condition an important one in truncation model. We take two real numbers c and d such that $[c, d] \subset [a_F, b_F]$, we are going to use this inclusion in the uniform consistency of the distribution law $G(\cdot)$ of the truncated r.v. T which is stated over a compact set (see Remark 6 in [23]).

Combining the ideas in [6] and [13], the local linear estimator of $f(y|x)$ in the case of truncated data is the coefficient $\hat{a}$ obtained by minimizing the following quantity

$$\min_{(a,b)\in\mathbb{R}^2} \sum_{i=1}^{n} \left[h_H^{-1} H(h_H^{-1}(y - Y_i)) - a - b\beta(X_i, x)\right]^2 K(h_K^{-1} d(X_i, x)) G_n^{-1}(Y_i)$$

and we have

$$f_n(y|x) = e_1' \left(Q'DQ\right)^{-1} Q'DZ,$$

where

$$Q' = \begin{bmatrix} 1 & \dots & 1 \\ \beta(X_1, x) & \dots & \beta(X_n, x) \end{bmatrix} \quad , \quad Z = \begin{bmatrix} h_H^{-1} H(h_H^{-1}(y - Y_1)) \\ \vdots \\ h_H^{-1} H(h_H^{-1}(y - Y_n)) \end{bmatrix},$$

$D = diag(K(h^{-1}d(X_1, x))G_n^{-1}(Y_1), \cdots, K(h^{-1}d(X_n, x))G_n^{-1}(Y_n))$ and $e_1' = (1, 0) \in \mathbb{R}^2$. By a simple calculus, one can derive the following explicit estimator

$$f_n(y|x) = \frac{\sum_{i,j=1}^{n} W_{ij}(x) H(h_H^{-1}(y - Y_j))}{h_H \sum_{i,j=1}^{n} W_{ij}(x)}, \qquad \left(\frac{0}{0} := 0\right), \tag{3}$$

where

$$W_{ij}(x) = \Delta_{ij}(x) G_n^{-1}(Y_i) G_n^{-1}(Y_j),$$

with $\Delta_{ij}(x)$ are defined by (2).

Remark 2.1. Notice that $f(y|x)$ is the derivative of the conditional distribution function $F(y|x) = P(Y \leq y|X = x)$. So, the feasible estimator $f_n(y|x)$ is also given by $f_n(y|x) = \frac{\partial F_n(y|x)}{\partial y}$, where $F_n(y|x)$ is the local linear estimator of $F(y|x)$. It was introduced by [4] who studied its pointwise and uniform almost sure convergences.

It is very well known that the conditional mode $\theta(x)$, on a set $[c, d]$, is given by

$$\theta(x) = \arg \sup_{y\in[c,d]} f(y|x),$$

this definition assumes implicitly that $\theta(x)$ exists on $[c, d]$. Therefore a natural estimator of $\theta(x)$ is defined by

$$\theta_n(x) = \arg \sup_{y\in[c,d]} f_n(y|x),$$

where $f_n(y|x)$ is defined in (3).

3. THE POINTWISE ALMOST SURE CONVERGENCE

Let x be a fixed point in $\mathcal{F}$, for any positive real r, $B(x,r) := \{y \in \mathcal{F}; d(x,y) \leq r\}$ denotes a closed ball in $\mathcal{F}$ of center x and radius r. We also define $\Phi_x(r_1, r_2) := \mathbf{P}(r_1 \leq d(x,X) \leq r_2)$, where r_1 and r_2 are two real numbers.

The following assumptions are necessary to study the asymptotic behaviour of our estimator $f_n(y|x)$.

(H1) For any $r > 0; \Phi_x(r) := \Phi_x(0,r) > 0$.

(H2) The conditional density $f(y|x)$ satisfies for some strictly positive constants b_1, b_2 and for all $(y_1, y_2) \in [c,d] \times [y_1 - h_H, y_1 + h_H]$ and $x_1, x_2 \in B(x, h_K)$,

$$|f(y_1|x_1) - f(y_2|x_2)| \leq C_x \left(d^{b_1}(x_1,x_2) + |y_1 - y_2|^{b_2}\right),$$

where C_x is a positive constant depending on x.

(H3) The function $\beta(\cdot,\cdot)$ is such that

$$\exists 0 < M_1 < M_2, \forall x' \in \mathcal{F}, M_1 d(x,x') \leq |\beta(x,x')| \leq M_2 d(x,x').$$

(H4) The kernel K is a positive and differentiable function on its support $[-1,1]$.

(H5) The kernel H is a positive, bounded and Lipschitzian continuous function, satisfying

$$\int |t|^{b_2} H(t)\,\mathrm{d}t < \infty \qquad \text{and} \qquad \int H^2(t)\,\mathrm{d}t < \infty.$$

(H6) The bandwidths h_K and h_H satisfy

$$\lim_{n\to\infty} h_K = 0, \qquad \lim_{n\to\infty} \left(\frac{\ln n}{n h_H \Phi_x(h_K)}\right) = 0,$$
$$\lim_{n\to\infty} h_H = 0 \quad \text{and} \quad \lim_{n\to\infty} n^{\gamma} h_H = \infty \quad \text{for some} \quad \gamma > 0.$$

(H7) There exists an integer n_0, such that

$$\forall n > n_0, \frac{1}{\Phi_x(h_K)} \int_0^1 \Phi_x(z h_K, h_K) \frac{\mathrm{d}}{\mathrm{d}z}\left(z^2 K(z)\right)\,\mathrm{d}z > 0$$

and

$$h_K \int_{B(x,h_K)} \beta(u,x)\,\mathrm{d}\mathbf{P}_X(u) = o\left(\int_{B(x,h_K)} \beta^2(u,x)\,\mathrm{d}\mathbf{P}_X(u)\right),$$

where $\mathrm{d}\mathbf{P}_X$ is the distribution of X.

Remark 3.1. Notice that these hypotheses are standard in this context and they are very similar to those used in [6].

We are now in position to state the pointwise a.s. convergence of $f_n(y|x)$.

Theorem 3.2. Assume that assumptions (H1) – (H7) are satisfied, we have

$$\sup_{y\in[c,d]} |f_n(y|x) - f(y|x)| = O(h_K^{b_1} + h_H^{b_2}) + O_{a.s.}\left(\sqrt{\frac{\ln n}{nh_H\Phi_x(h_K)}}\right).$$

Proof. Let us set the following pseudo-estimator of $f(y|x)$ defined by

$$\begin{aligned}
\widetilde{f}_n(y|x) &= \frac{\sum_{i,j=1}^n G^{-1}(Y_i)G^{-1}(Y_j)\Delta_{ij}(x)H(h_H^{-1}(y-Y_j))}{h_H\sum_{i,j=1}^n G^{-1}(Y_i)G^{-1}(Y_j)\Delta_{ij}(x)} \\
&= \frac{\frac{\mu_n^2}{n(n-1)h_H\mathbf{E}\left(\Delta_{12}(x)\right)}\sum_{i,j=1}^n G^{-1}(Y_i)G^{-1}(Y_j)\Delta_{ij}(x)H(h_H^{-1}(y-Y_j))}{\frac{\mu_n^2}{n(n-1)\mathbf{E}\left(\Delta_{12}(x)\right)}\sum_{i,j=1}^n G^{-1}(Y_i)G^{-1}(Y_j)\Delta_{ij}(x)} \\
&:= \frac{\tilde{\Upsilon}_n(x,y)}{\tilde{r}_n(x)},
\end{aligned}$$

which will play a prominent part in the proof thanks to the following decomposition.

$$\begin{aligned}
f_n(y|x) - f(y|x) &= \frac{\Upsilon_n(x,y)}{r_n(x)} - f(y|x) \qquad (4)\\
&= \frac{1}{r_n(x)}\left\{\Upsilon_n(x,y) - \tilde{\Upsilon}_n(x,y)\right\} + \frac{1}{r_n(x)}\left\{\tilde{\Upsilon}_n(x,y) - \mathbf{E}(\tilde{\Upsilon}_n(x,y))\right\} \\
&+ \frac{1}{r_n(x)}\left\{\mathbf{E}(\tilde{\Upsilon}_n(x,y)) - f(y|x)\right\} \\
&+ \frac{f(y|x)}{r_n(x)}\left\{(\tilde{r}_n(x) - r_n(x)) + (\mathbf{E}(\tilde{r}_n(x)) - \tilde{r}_n(x)) + (-\mathbf{E}(\tilde{r}_n(x)) + 1)\right\},
\end{aligned}$$

where

$$\Upsilon_n(x,y) = \frac{\mu_n^2}{n(n-1)h_H\mathbf{E}\left(\Delta_{12}(x)\right)}\sum_{i,j=1}^n G_n^{-1}(Y_i)G_n^{-1}(Y_j)\Delta_{ij}(x)H(h_H^{-1}(y-Y_j))$$

and

$$r_n(x) = \frac{\mu_n^2}{n(n-1)\mathbf{E}\left(\Delta_{12}(x)\right)}\sum_{i,j=1}^n G_n^{-1}(Y_i)G_n^{-1}(Y_j)\Delta_{ij}(x).$$

The proof of Theorem 3.2 is then a direct consequence of the following Lemmas, the proofs of which are relegated to the Appendix. □

Lemma 3.3. Under the assumptions (H1), (H2) and (H4), we obtain

$$\sup_{y\in[c,d]} |\mathbf{E}(\tilde{\Upsilon}_n(x,y)) - f(y|x)| = O(h_K^{b1} + h_H^{b2}).$$

Lemma 3.4. i) Under the assumptions (H1) – (H7), we get

$$\sup_{y\in[c,d]} |\tilde{\Upsilon}_n(x,y) - \mathbf{E}(\tilde{\Upsilon}_n(x,y))| = O_{a.co.}\left(\sqrt{\frac{\ln n}{nh_H\Phi_x(h_K)}}\right).$$

ii) Under the assumptions (H1),(H3), (H4) and (H6), we have

$$\tilde{r}_n(x) - 1 = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h_K)}}\right)$$

and

$$\exists \vartheta > 0, \quad \text{such that} \quad \sum_{n=1}^{\infty} \mathbf{P}\left(\tilde{r}_n(x) < \vartheta\right) < \infty.$$

Lemma 3.5. Under the assumptions (H1), (H3), (H4) and (H6), we have

$$\sup_{y\in[c,d]} |\Upsilon_n(x,y) - \tilde{\Upsilon}_n(x,y)| = O_{a.s.}\left(\sqrt{\frac{\ln n}{nh_H\Phi_x(h_K)}}\right)$$

and

$$|r_n(x) - \tilde{r}_n(x)| = O_{a.s.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h_K)}}\right).$$

It is clear, from Borel Cantelli lemma, that the almost-complete convergence (a.co.)[1] is stronger than the almost sure (a.s.) one. We refer the reader to the appendix of [11] for more details.

4. UNIFORM ALMOST SURE CONVERGENCE

In practice, the uniform consistency has great importance because it is used to improve the efficiency of the estimation and to solve some problems such as data-driven bandwidth choice or bootstrapping. Unlike in the multivariate case, the uniform consistency is not a standard extension of the pointwise one. So, suitable additional tools and topological conditions are needed.

In this section, we will investigate the uniform almost sure convergence of $f_n(y|x)$ and $\theta_n(x)$ on some subset $\mathcal{S}_\mathcal{F}$ of $\mathcal{F}$, such that $\mathcal{S}_\mathcal{F} \subset \bigcup_{k=1}^{d_n} B(x_k, z_n)$, where $x_k \in \mathcal{S}_\mathcal{F}$ and z_n (respectively d_n) is a sequence of positive real (respectively integer) numbers.

4.1. Conditional density function estimator

In this study, we need the following assumptions.

(U1) There exist a differentiable function Φ and strictly positive constants C, C_1 and C_2 such that

$$\forall x \in \mathcal{S}_\mathcal{F}, 0 < C_1\Phi(h_K) \leq \Phi_x(h_K) \leq C_2\Phi(h_K) < \infty$$

[1] Recall that a sequence of real random variables r.r.v. $(W_n)_{n\in\mathbb{N}^*}$ converges almost completely to some r.r.v. W, and we note $W_n \longrightarrow^{a.co.} W$, if and only if $\forall\epsilon > 0, \sum_{n=1}^{\infty} P(|W_n - W| > \epsilon) < \infty$. Moreover, let $(v_n)_{n\in\mathbb{N}^*}$ be a sequence of positive real numbers going to zero; we say that the rate of the almost complete convergence of $(W_n)_{n\in\mathbb{N}^*}$ to W is of order (v_n) and we note $W_n - W = O_{a.co.}(v_n)$, if and only if $\exists\epsilon_0 > 0, \sum_{n=1}^{\infty} P(|W_n - W| > \epsilon_0 v_n) < \infty$.

and

$$\exists \eta_0 > 0, \forall \eta < \eta_0, \Phi'(\eta) < C,$$

where Φ' denotes the first derivative of Φ with $\Phi(0) = 0$.

(U2) The conditional distribution $f(y|x)$ satisfies for some strictly positive constants C, b_1 *and* b_2 and for all $(y_1, y_2) \in [c,d] \times [y_1 - h_H, y_1 + h_H]$ and $(x_1, x_2) \in S_{\mathcal{F}} \times B(x, h_K)$,

$$|f(y_1|x_1) - f(y_2|x_2)| \leq C\left(d^{b_1}(x_1, x_2) + |y_1 - y_2|^{b_2}\right).$$

(U3) The function $\beta(\cdot,\cdot)$ satisfies (H3) uniformly on x and the following Lipschitz's condition

$$\exists C > 0, \forall x_1, x_2 \in \mathcal{S}_{\mathcal{F}}, \forall x \in \mathcal{F}, |\beta(x, x_1) - \beta(x, x_2)| \leq C d(x_1, x_2).$$

(U4) The kernel K fulfils (H4) and is Lipschitzian on $[0, 1]$.

(U5) $\lim_{n\longrightarrow\infty} h_K = 0$, $\lim_{n\to\infty} h_H = 0$, $\lim_{n\to\infty} n^{\gamma} h_H = \infty$ *for some* $\gamma > 0$ and for $z_n = O\left(\frac{\ln n}{n}\right)$, we have for n large enough

$$\frac{(\ln n)^2}{n h_H \Phi(h)} < \ln d_n < \frac{n h_H \Phi(h)}{\ln n} \text{ and } \sum_{n=1}^{\infty} n^{(3\gamma+1/2)} d_n^{(1-\alpha)} < \infty; \quad \textit{for some} \quad \alpha > 1.$$

(U6) The bandwidth h_K satisfies $\exists n_0 \in \mathbb{N}, \exists C > 0$, such that

$$\forall n > n_0, \forall x \in \mathcal{S}_{\mathcal{F}}, \frac{1}{\Phi_x(h_K)} \int_0^1 \Phi_x(z h_K, h_K) \frac{\mathrm{d}}{\mathrm{d}z}(z^2 K(z))\, \mathrm{d}z > C > 0$$

and

$$h_K \int_{B(x,h_K)} \beta(u, x)\, \mathrm{d}\mathbf{P}_X(u) = o\left(\int_{B(x,h_K)} \beta^2(u, x)\, \mathrm{d}\mathbf{P}_X(u)\right),$$

uniformly on x.

Remark 4.1. Remark that most of these hypothesis are the uniform version of the corresponding conditions in the pointwise case. Beside they have already been used in the literature, We refer for example to [6] and [21].

Theorem 4.2. Under assumptions (U1)−(U6), we have

$$\sup_{x\in\mathcal{S}_{\mathcal{F}}} \sup_{y\in[c,d]} |f_n(y|x) - f(y|x)| = O(h_K^{b_1} + h_H^{b_2}) + O_{a.s.}\left(\sqrt{\frac{\ln d_n}{n h_H \Phi(h_K)}}\right).$$

The proof of Theorem 4.2 is based on the decomposition (4) and the following Lemmas for which the proofs are given in the Appendix.

Lemma 4.3. Under the assumptions (U1), (U2) and (U4), we obtain that

$$\sup_{x\in\mathcal{S}_{\mathcal{F}}} \sup_{y\in[c,d]} |\mathbf{E}(\tilde{\Upsilon}_n(x, y)) - f(y|x)| = O(h_K^{b_1}) + O(h_H^{b_2}).$$

Lemma 4.4. i) Under the assumptions (U1) – (U6), we have

$$\sup_{x\in\mathcal{S}_\mathcal{F}}\sup_{y\in[c,d]}|\tilde{\Upsilon}_n(x,y)-\mathbf{E}(\tilde{\Upsilon}_n(x,y))|=O_{a.co.}\left(\sqrt{\frac{\ln d_n}{nh_H\Phi(h_K)}}\right).$$

ii) If assumptions (U1), (U3) – (U6) are satisfied, we get

$$\sup_{x\in\mathcal{S}_\mathcal{F}}|\tilde{r}_n(x)-1|=O_{a.co.}\left(\sqrt{\frac{\ln d_n}{n\Phi(h_K)}}\right)$$

and

$$\exists\vartheta>0,\quad\text{such that}\quad\sum_{n=1}^{\infty}\mathbf{P}\left(\inf_{x\in\mathcal{S}_\mathcal{F}}\tilde{r}_n(x)<\vartheta\right)<\infty.$$

Lemma 4.5. Under the assumptions (U1), (U3) and (U4) – (U6) we get

$$\sup_{x\in\mathcal{S}_\mathcal{F}}\sup_{y\in[c,d]}|\Upsilon_n(x,y)-\tilde{\Upsilon}_n(x,y)|=O_{a.s.}\left(\sqrt{\frac{\ln d_n}{nh_H\Phi(h_K)}}\right)$$

and

$$\sup_{x\in\mathcal{S}_\mathcal{F}}|r_n(x)-\tilde{r}_n(x)|=O_{a.s.}\left(\sqrt{\frac{\ln d_n}{n\Phi(h_K)}}\right).$$

4.2. Conditional mode estimator

To study the almost sure convergence of the local estimator of the conditional mode of Y given $X=x$ uniformly on a fixed subset $S_\mathcal{F}$ of $\mathcal{F}$, we introduce the following uniform uniqueness properties used for example in [6] and [10].

(U7) For all $\epsilon>0,\exists\varsigma>0$ such that for any function ζ from $\mathcal{S}_\mathcal{F}$ into $[c,d]$ we have

$$\sup_{x\in\mathcal{S}_\mathcal{F}}|\theta(x)-\zeta(x)|\geq\epsilon\quad\Rightarrow\quad\sup_{x\in\mathcal{S}_\mathcal{F}}|f(\theta(x)|x)-f(\zeta(x)|x)|\geq\varsigma.$$

(U8) There exists some integer $j>1,\forall x\in\mathcal{S}_\mathcal{F},f(\cdot|x)$ is j-times continuously differentiable on the topological of $[c,d]$ with respect of y and satisfies $f^{(l)}(\theta(x)|x)=0$ if $0\leq l<j,f^{(j)}(\theta(x)|x)>C>0$ and $f^{(j)}(\cdot|x)$ is uniformly continuous on $[c,d]$ where $f^{(l)}(\cdot|x)$ stands for the lth-order derivative of $f(\cdot|x)$.

A known method can be applied to derive the following result from 4.2, see for example the proof of Corollary 7 in [10].

Theorem 4.6. If the conditional density $f(y|x)$ satisfies assumptions (U7) and (U8) in addition of the hypotheses of Theorem 4.2, then we get

$$\sup_{x\in\mathcal{S}_\mathcal{F}}|\theta_n(x)-\theta(x)|^j=O(h_K^{b_1})+O(h_H^{b_2})+O_{a.s.}\left(\sqrt{\frac{\ln d_n}{nh_H\Phi(h_K)}}\right).$$

5. NUMERICAL RESULTS

In this section, two examples of simulation and a real data set are drawn to illustrate the performance of the local linear estimator of the conditional mode (LLM) studied in this paper, for finite size sample n. More precisely, we compare it to the kernel conditional mode estimator (KM) graphically and by measuring the prediction accuracy.

For the computation of the LLM and the KM estimators, we use the quadratics kernels and the bandwidths h_K and h_H are chosen by the 2-fold cross-validation method. Take into account of the smoothness of the curves $X_i(t)$ (see Figures 1, 5 and 7), we choose the semi-metric d based on the derivative described in [11] (see routines "semimetric.deriv" in the website http://www.lsp.ups-tlse.fr/staph/npfda) and we take $\beta = d$ (for the LLM estimator).

5.1. Simulation study

To show the finite-sample performance of our LLM estimator, we generate the observed sequences $(X_i(t), Y_i, T_i)_{1\le i\le n}$ by the following steps.

- Step 1. We fix the random size n (recall that n is known), then we generate the random variables $T_1, X_1(t)$ and Y_1.
- Step 2. Test:
  We begin by setting:
  $N = 0$,
  $j = 0$,
  While $j \le n$:
  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 us to get the truncation rate (TR).

Example 1. We fixe $n = 200$ and we generate the scalar response variable as

$$Y = R(X) + \varepsilon,$$

where X and ϵ are assumed independent, the error $\varepsilon \hookrightarrow \mathcal{N}(0, 0.1)$ and the operator $R(\cdot)$ is defined by

$$R(X) = \exp\left(\frac{1}{1 + \int_0^1 (X_1'(t))^2 \, dt}\right)$$

The functional covariate $X(t)$ is defined, for $t \in [0, \pi]$ as follows

$$X(t) = \sin(tW),$$

where $W \hookrightarrow \mathcal{N}(0, 1)$. The curves are discretized on the same grid which is composed of 200-equidistant values in $[0, \pi]$ (see Figure 1).

The truncation variable T has an exponential distribution with parameter λ which is adapted in order to get different rate of truncation TR.

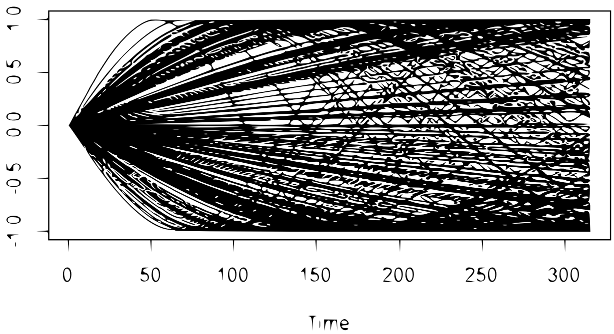

Fig. 1. A sample of 200 curves representing a realization of the functional random variable X.

Given $X = x$, we can easily see that $Y \hookrightarrow \mathcal{N}(R(x), 0.1)$, and therefore, the conditional mode function will coincide and will be equal to $R(x)$.

Under this model, we compute our LLM estimator θ_n and the KM estimator $\widehat{\theta}_{KM}$ defined by $\hat{\theta}_{KM}(x) = \arg\sup_{y\in[c,d]} \hat{f}_n(y|x)$, where $\hat{f}_n(y|x)$ is proposed by [13], with the observed data $(X_i, Y_i, T_i)_{1\leq i\leq n}$ (i.e $Y_i \geq T_i$).

In this simulation, to illustrate the performance of our estimator, we proceed with the following algorithm.

- Step 1. We split our data into two subsets
  – $(X_i, Y_i)_{1\leq i\leq 100}$: The learning sample used to build the estimators.
  – $(X_i, Y_i)_{101\leq i\leq 200}$: The testing sample used to make a comparison.
- Step 2. We calculate the two estimators by using the learning sample and we find the LLM and KM estimators of the conditional mode (θ_n and $\widehat{\theta}_{KM}$).
- Step 3. We plot the true values ($\theta(X_i)$) for all i ($101 \leq i \leq 200$) against the predicted ones by means of the two estimators LLM and KM (one in each graph), this is displayed in Figures 2, 3 and 4.
- Step 4. To be more precise we evaluate the prediction errors given by

$$MSE(LLM) := \frac{1}{100}\sum_{j=101}^{200} \left(\theta_n(X_j) - \theta(X_j)\right)^2$$

  and

$$MSE(KM) := \frac{1}{100}\sum_{j=101}^{200} \left(\widehat{\theta}_{KM}(X_j) - \theta(X_j)\right)^2.$$

It can be seen clearly from Figures 2, 3 and 4 that our estimator performs better than the kernel one estimator. Also, we note that the quality of both estimators become slightly worse when we have high percentage of truncation TR, however it remains acceptable.

To make a better decission, we choose an other example.

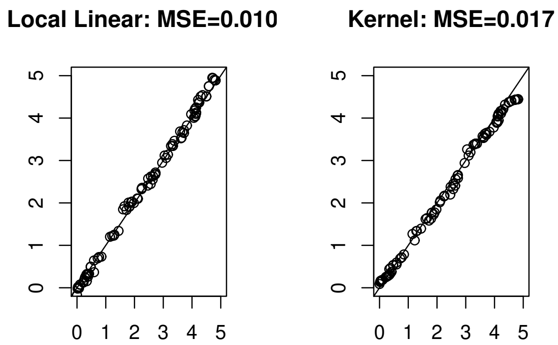

Fig. 2. Complete data.

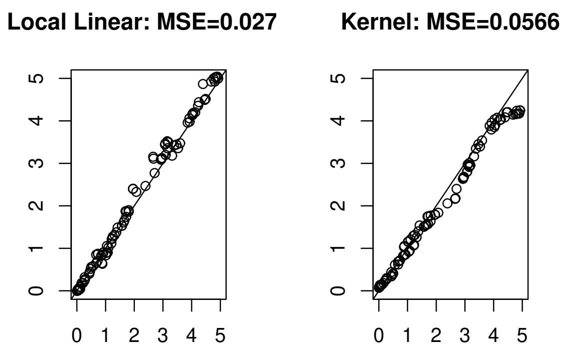

Fig. 3. TR=55%.

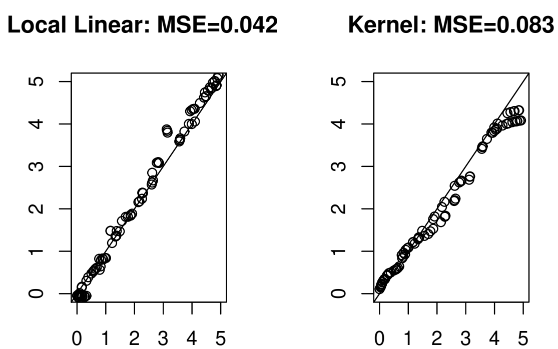

Fig. 4. TR=73%.

Example 2. In this example we vary the sample size $n = 100,\ 300,\ 500$ and we consider the functional covariate $X(t)$ generated in the following equation

$$X(t) = 3A_1 \sin(2\pi t) + \eta t, \quad t \in [0,1],$$

where $\eta \hookrightarrow \mathcal{N}(0,1)$ and $A \hookrightarrow \mathcal{N}(0,0.5)$. We carried out the simulation with a 300-sample of the curve X which is represented in the Figure 5.

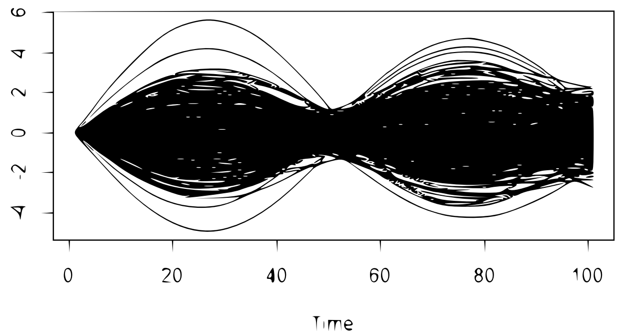

Fig. 5. A sample of 300 curves representing a realization of the functional random variable X.

The scalar response variable is defined as

$$Y = R(X) + \epsilon,$$

where X and ϵ are independent, the error $\epsilon \hookrightarrow \mathcal{N}(0,0.1)$ and

$$R(X) = \int_0^1 \frac{\mathrm{d}t}{1 + X^2(t)}.$$

For this model, we adopt the mechanism of truncation on the basis of the sample $(X_i, Y_i, T_i)_{\{1 \le i \le n\}}$ where the truncation variable $T_1 \hookrightarrow \mathcal{N}(0,2)$ which taken to fix the percentage of truncation TR.

Next, we split our data into a learning sample with size n_1 and a test sample with size $n_2 = n - n_1$, for a different sample sizes.

To give a visual impression of the quality of estimation we draw the curves corresponding to the true values of conditional mode TCM (the solid lines) and both estimated values LLM and KM (the green dotted ones and the blue dotted ones) for $n = 300$ in Figure 6.

In order to get a more precise, for different values of n, we evaluate the prediction errors given by

$$\begin{cases} MSE(LLM) := \frac{1}{n_2} \sum_{j=n_1+1}^{n} \left(\theta_n(X_j) - \theta(X_j)\right)^2 \\ MSE(KM) := \frac{1}{n_2} \sum_{j=n_1+1}^{n} \left(\widehat{\theta}_{KR}(X_j) - \theta(X_j)\right)^2 \end{cases}$$

The obtained results are in the Table 1. Figure 6 and Table 1 show that the local linear estimator performs better than the kernel one for the different values of n. Also, the quality of both estimators increases when the the sample size n increase.

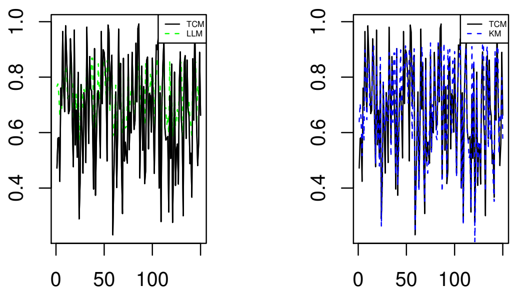

Fig. 6. Representation of the studied estimators for $n = 300$.

n	MSE(LLM)	MSE(KM)
100	0.0542	0.1117
300	0.0302	0.095
500	0.0284	0.0496

Tab. 1. MSE comparison for LLM and KM methods for the three samples sizes (n).

5.2. Real data application

We consider the spectrometry curves of 215 pieces of meat and we aim to predict the fat content Y in a piece of meat from its spectrometric curve X. These curves X are displayed in Figure 7. Notice that these data are one of the most popular functional data sets used in the functional statistical literature which can be found at (`www:\lib.stat.cmu.edu/datasets/tecator`).

The LLM and the KM estimators are computed with the artificial observed real data $(Y_i, X_i, T_i)_{1\leq i\leq 150}$, where $Y_i \geq T_i$ and the truncation r.v T_i has an exponential distribution with parameter 1, with $TR \simeq 30\%$.

Next, we split these real data into a learning sample containing the first n_1 units and a test sample containing $n_2 = 150 - n_1$ units. To illustrate the performance of our estimator, we first plot the true values (provided in the test sample) against the predicted ones by means of the two estimators (one in each graph). This is displayed in Figure 8.

Secondly, the criteria allowing us to compare between the both estimators is the prediction errors (MSE), defined by

$$MSE(LLM) := \frac{1}{n_2} \sum_{j=n_1+1}^{150} (\theta_n(X_j) - Y_j)^2$$

and

$$MSE(KM) := \frac{1}{n_2} \sum_{j=n_1+1}^{150} \left(\widehat{\theta}_{KR}(X_j) - Y_j \right)^2 .$$

The obtained results are $MSE(LLM) = 5.0453$ and $MSE(KM) = 6.4567$.

It appears clearly that, the LLM method outperforms the kernel method.

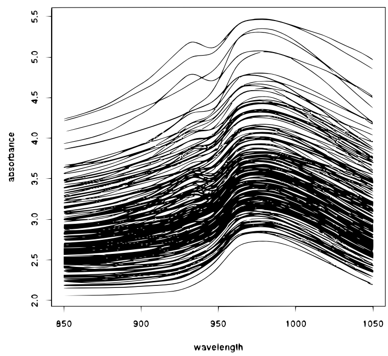

Fig. 7. Spectrometric data.

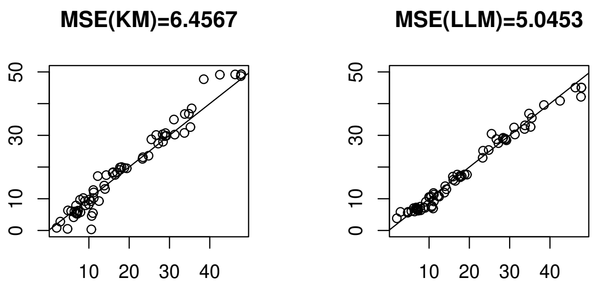

Fig. 8. Performance of the two methods for the Spectrometric data.

CONCLUSION

The contribution of this paper is the particular focus on the study of the local linear non parametric estimation of the conditional mode when the explanatory variable is functional and the response variable is subject to left truncation by another random variable. Firstly, we establish the pointwise and the uniform almost sure convergence of the conditional density estimator. Then, we obtained the uniform almost sure convergence of the proposed local linear conditional mode estimator. Finaly, our theoretical and practical studies show that the local linear method outperforms the kernel one even for trancation data. Moreover, our results confirmed without surprise that the behavior of the LLM and the KM estimators are better for a weak percentage of truncation TR and a large sample size n.

APPENDIX

In what follows, let C be some strictly positive generic constant. Moreover, we put, for any $x \in \mathcal{F}$, and for all $i = 1, \dots, n$:

$$K_i(x) := K\left(h_K^{-1} d(X_i, x)\right), \quad \beta_i(x) := \beta(X_i, x) \quad \textit{and} \quad H_i(y) := H\left(h_H^{-1}(y - Y_i)\right).$$

Furthermore, for any random vector L, we denote by $\sigma(L)$ the σ- algebra generated by L.

To treat the pointwise almost sure convergence of $f_n(y|x)$, we need Lemma A1 introduced in [2].

Proof. of Lemma 3.3.
Since (X_i, Y_i, T_i) are identically distributed then

$$\begin{aligned}
\mathbf{E}(\tilde{\Upsilon}_n(x,y)) &= \mathbf{E}\left(\frac{\mu^2}{n(n-1)h_H \mathbf{E}\left(\Delta_{12}(x)\right)} \sum_{i \neq j} G^{-1}(Y_i) G^{-1}(Y_j) \Delta_{ij}(x) H_j(y)\right) \\
&= \frac{\mu^2}{h_H \mathbf{E}\left(\Delta_{12}(x)\right)} \mathbf{E}\left(\frac{1}{G(Y_1)G(Y_2)} \Delta_{12}(x) H_2(y)\right) \\
&= \frac{\mu^2}{h_H \mathbf{E}\left(\Delta_{12}(x)\right)} \mathbf{E}\left[\mathbb{E}\left(\Delta_{12}(x) H_2(y) \frac{1_{\{Y_1 \geq T_1\}} 1_{\{Y_2 \geq T_2\}}}{\mu^2 G(Y_1) G(Y_2)} |\sigma(X_1, Y_1, X_2, Y_2)\right)\right] \\
&= \frac{1}{h_H \mathbf{E}\left(\Delta_{12}(x)\right)} \mathbf{E}\left(\Delta_{12}(x) H_2(y)\right) \\
&= \frac{1}{\mathbf{E}\left(\Delta_{12}(x)\right)} \mathbf{E}\left(\Delta_{12}(x) \mathbf{E}\left(h_H^{-1} H_2(y) | X_2\right)\right).
\end{aligned}$$

So, we can write

$$|\mathbf{E}(\tilde{\Upsilon}_n(x,y)) - f(y|x)| = \frac{1}{|\mathbf{E}\left(\Delta_{12}(x)\right)|} |\mathbf{E}\left(\Delta_{12}(x)\left(\mathbf{E}\left(h_H^{-1} H_2(y)\right) |X_2\right) - f(y|x)\right)|.$$

On the other hand,

$$\begin{aligned}
h_H^{-1} \mathbf{E}\left(H_2(y) | X_2\right) &= h_H^{-1} \int_{\mathbb{R}} H(h_H^{-1}(y-u)) f(u|X_2) \, \mathrm{d}u \\
&= \int_{\mathbb{R}} H(z) f(y - z h_H | X_2) \, \mathrm{d}z,
\end{aligned}$$

so, we obtain

$$|h_H^{-1} \mathbf{E}\left(H_2(y)|X_2\right) - f(y|x)| \leq \int_{\mathbb{R}} H(z) |f(y - z h_H | X_2) - f(y|x)| \, \mathrm{d}z. \tag{5}$$

Moreover, we have

$$|f(y-zh_H|X_2)-f(y|x)| \leq |f(y-zh_H|X_2)-f(y-zh_H|x)|+|f(y-zh_H|x)-f(y|x)|.$$

By using the last relation, together with hypothesis (H2) and (H5), we obtain the claimed result. □

P r o o f. of Lemma 3.4.
We will proceed by two steps as follows.

1. Firstly, we show that

$$\sum_n \mathbf{P}\left(|\tilde{\Upsilon}_n(x,y)-\mathbf{E}(\tilde{\Upsilon}_n(x,y))|>\epsilon\sqrt{\frac{\ln n}{nh_H\Phi_x(h_K)}}\right)<\infty.$$

For this, we can write

$$\tilde{\Upsilon}_n(x,y)=Q(x)\left[P^x_{2,1}(y)P^x_{4,0}-P^x_{3,1}(y)P^x_{3,0}\right], \tag{6}$$

where for $p=2,3,4$ and $l=0,1$

$$Q(x)=\frac{n^2h_K^2\Phi_x^2(h_K)}{n(n-1)\mathbf{E}\left(\Delta_{12}(x)\right)} \tag{7}$$

and

$$P^x_{p,l}(y)=\frac{1}{n\Phi_x(h_K)}\sum_{i=1}^n\frac{\mu K_i(x)\beta_i^{p-2}(x)H_i^l(y)}{h_H^l h_K^{p-2}G(Y_i)}, \tag{8}$$

with

$$P^x_{4,0}(y):=P^x_{4,0} \qquad \text{and} \qquad P^x_{3,0}(y):=P^x_{3,0}.$$

So, we have

$$\begin{aligned}\tilde{\Upsilon}_n(x,y)-\mathbf{E}\left(\tilde{\Upsilon}_n(x,y)\right)&=Q(x)\left\{P^x_{2,1}(y)P^x_{4,0}-\mathbf{E}\left(P^x_{2,1}(y)P^x_{4,0}\right)\right\}\\&\quad-Q(x)\left\{P^x_{3,1}(y)P^x_{3,0}-\mathbf{E}\left(P^x_{3,1}(y)P^x_{3,0}\right)\right\}.\end{aligned} \tag{9}$$

Notice that $Q(x)=O(1)$, see [2], so, we need to show that, for $p=2,3,4$ and $l=0,1$

$$\mathbf{E}(P^x_{p,l}(y))=O(1), P^x_{p,l}(y)-\mathbf{E}(P^x_{p,l}(y))=O_{a.co}\left(\sqrt{\frac{\ln n}{nh_H\Phi_x(h_K)}}\right),$$

$$\mathbf{E}(P^x_{2,1}(y))\mathbf{E}(P^x_{4,0})-\mathbf{E}(P^x_{2,1}(y)P^x_{4,0})=O\left(\sqrt{\frac{\ln n}{nh_H\Phi_x(h_K)}}\right),$$

$$\mathbf{E}(P^x_{3,1}(y))\mathbf{E}(P^x_{3,0})-\mathbf{E}(P^x_{3,1}(y)P^x_{3,0})=O\left(\sqrt{\frac{\ln n}{nh_H\Phi_x(h_K)}}\right).$$

- Applying [2, Lemma A.1 (i)] we can easily have for $p = 2, 3, 4$ and $l = 0, 1$

$$\begin{aligned}
\mathbf{E}(P^x_{p,l}(y)) &= \mathbf{E}\left(\frac{1}{\Phi_x(h_K)}\sum_{i=1}^{n}\frac{\mu K_i(x)\beta_i^{p-2}(x)H_i^l(y)}{h_H^l h_K^{p-2}G(Y_i)}\right)\\
&= \mu h_K^{2-p}\Phi_x^{-1}(h_K)\mathbf{E}\left[h_H^{-l}\mathbb{E}\left(K_1(x)\beta_1^{p-2}(x)H_1^l(y)\frac{1_{\{Y_1\geq T_1\}}}{\mu G(Y_1)}|\sigma(X_1,Y_1)\right)\right]\\
&= h_H^{-l}h_K^{2-p}\Phi_x^{-1}(h_K)\mathbf{E}\left(K_1(x)\beta_1^{p-2}(x)H_1^l(y)\right)\\
&= h_H^{-l}h_K^{2-p}\Phi_x^{-1}(h_K)\mathbf{E}\left(K_1(x)\beta_1^{p-2}(x)\mathbf{E}(H_1^l(y)|X_1)\right)\\
&= h_H^{-l}h_K^{2-p}\Phi_x^{-1}(h_K)O\left(h_H^l\mathbf{E}(K_1(x)\beta_1^{p-2}(x))\right)\\
&= O(1). \qquad (10)
\end{aligned}$$

- Treatment of the term $P^x_{p,l}(y) - \mathbf{E}(P^x_{p,l}(y))$
  We put

$$P^x_{p,l}(y) - \mathbf{E}(P^x_{p,l}(y)) = \frac{1}{n}\sum_{i=1}^{n}M_i^{(p,l)},$$

where

$$M_i^{(p,l)} = \frac{1}{h_H^l h_K^{p-2}\Phi_x(h_K)}\left\{\frac{\mu K_i(x)\beta_i^{p-2}(x)H_i^l(y)}{G(Y_i)} - \mathbf{E}\left(\frac{\mu K_i(x)\beta_i^{p-2}(x)H_i^l(y)}{G(Y_i)}\right)\right\}. \qquad (11)$$

The main point is to evaluate asymptotically the mth-order moment of the r.r.v. $M_i^{(p,l)}$.
We have

$$\begin{aligned}
\mathbf{E}|\left\{M_i^{(p,l)}\right\}^m| &= h_H^{-l}h_K^{(-p+2)m}\Phi_x^{-m}(h)\mathbf{E}\Bigg|\sum_{k=0}^{m}C_m^k(-1)^{m-k}\left(\frac{\mu K_i(x)\beta_i^{p-2}(x)H_i^l(y)}{G(Y_i)}\right)^k\\
&\quad\left(\mathbf{E}\left[\frac{\mu K_i(x)\beta_i^{p-2}(x)H_i^l(y)}{G(Y_i)}\right]\right)^{m-k}\Bigg|\\
&\leq h_H^{-l}h_K^{(-p+2)m}\Phi_x^{-m}(h)\sum_{k=0}^{m}C_m^k\left(\mathbf{E}\left|\frac{\mu K_i(x)\beta_i^{p-2}(x)H_i^l(y)}{G(Y_i)}\right|^k\right)\\
&\quad\left|\mathbf{E}\left[\frac{\mu K_i(x)\beta_i^{p-2}(x)H_i^l(y)}{G(Y_i)}\right]\right|^{m-k}.
\end{aligned}$$

It's easy to get, for all $k \leq m$ and $l = 0, 1$, that

$$\mathbf{E}\left(H_i^{lk}(y)|X\right) = O(h_H^l).$$

So, By using [2, Lemma A.1], one gets

$$\mathbf{E}|\left\{M_i^{(p,l)}\right\}^m| = O\left((h_H^l\Phi_x(h_K))^{(-m+1)}\right).$$

Finally, it suffices to apply [11, Corollary A.8-(ii)] with $a_n^2 = (h_H^l\Phi_x(h_K))^{-1}$ and taking account that for $l=0$

$$\mathbf{P}\left(P_{3,0}^x - \mathbf{E}(P_{3,0}^x) > \epsilon\sqrt{\frac{\ln n}{nh_H\Phi_x(h_K)}}\right) \leq \mathbf{P}\left(P_{3,0}^x - \mathbf{E}(P_{3,0}^x) > \epsilon\sqrt{\frac{\ln n}{n\Phi_x(h_K)}}\right)$$

and

$$\mathbf{P}\left(P_{4,0}^x - \mathbf{E}(P_{4,0}^x) > \epsilon\sqrt{\frac{\ln n}{nh_H\Phi_x(h_K)}}\right) \leq \mathbf{P}\left(P_{4,0}^x - \mathbf{E}(P_{4,0}^x) > \epsilon\sqrt{\frac{\ln n}{n\Phi_x(h_K)}}\right).$$

We get, for $p=2,3,4$ and $l=0,1$

$$P_{p,l}^x(y) - \mathbf{E}(P_{p,l}^x(y)) = O_{a.co}\left(\sqrt{\frac{\ln n}{nh_H\Phi_x(h_K)}}\right). \tag{12}$$

- Treatment of the term $\mathbf{E}(P_{2,1}^x(y))\mathbf{E}(P_{4,0}^x) - \mathbf{E}(P_{2,1}^x(y)P_{4,0}^x)$.
  We can write

$$\begin{aligned}\mathbf{E}(P_{2,1}^x(y))\mathbf{E}(P_{4,0}^x) - \mathbf{E}(P_{2,1}^x(y)P_{4,0}^x) &= \left(1-\frac{n(n-1)}{n^2}\right)h_K^{-2}\Phi_x^{-2}(h_K)\\ &\quad \mathbf{E}\left(K_1(x)\beta_1^2(x)\right)\mathbf{E}\left(h_H^{-1}K_1(x)H_1(y)\right)\\ &\quad + O((n\Phi_x(h_K))^{-1})\\ &= O((n\Phi_x(h_K))^{-1}).\end{aligned}$$

  We get the last result, always, by using [2, Lemma A.1] and conditional expectation.
  Under (H5), $O((n\Phi_x(h_K))^{-1})$ is negligible with respect to $O\left(\sqrt{\frac{\ln n}{nh_H\Phi_x(h_K)}}\right)$.

- By similar arguments, one can prove that

$$\mathbf{E}(P_{3,1}^x(y))\mathbf{E}(P_{3,0}^x) - \mathbf{E}(P_{3,1}^x(y)P_{3,0}^x) = O\left(\sqrt{\frac{\ln n}{nh_H\Phi_x(h_K)}}\right).$$

For the second part of the Lemma, it's easy to find that $\mathbf{E}\left(\tilde{r}_n(x)\right) = 1$ and this leads us to get the last result.

2. Secondly, to study the uniform convergence of $f(y|x)$ on $y \in [c,d]$, we must first point out that since $[c,d]$ is a compact of $\mathbb{R}$, so we can cover it by a finite number s_n of intervals of length l_n. Precisely, we have $[c,d] \subseteq \cup_{k=1}^{s_n}]t_k - l_n, t_k + l_n[$, where $s_n = \frac{C}{l_n}$ and $l_n = n^{-3\gamma-1/2}$.

Taking

$$t_y = \arg \min_{t \in \{t_1, t_2, \ldots, t_{s_n}\}} |y - t|.$$

Then, we can write

$$\begin{aligned}
&\sup_{y \in [c,d]} |\tilde{\Upsilon}_n(x,y) - \mathbf{E}(\tilde{\Upsilon}_n(x,y))| \\
&\leq \sup_{y \in [c,d]} |\tilde{\Upsilon}_n(x,y) - \tilde{\Upsilon}_n(x,t_y)| + \sup_{y \in [c,d]} |\tilde{\Upsilon}_n(x,t_y) - \mathbf{E}(\tilde{\Upsilon}_n(x,t_y))| \\
&\quad + \sup_{y \in [c,d]} |\mathbf{E}(\tilde{\Upsilon}_n(x,t_y)) - \mathbf{E}(\tilde{\Upsilon}_n(x,y))| \\
&\quad := R_1 + R_2 + R_3.
\end{aligned}$$

Starting with treatment of R_1.

We get, using lipschitz argument, that

$$\begin{aligned}
R_1 &= \sup_{y \in [c,d]} |\tilde{\Upsilon}_n(x,y) - \tilde{\Upsilon}_n(x,t_y)| \\
&\leq \sup_{y \in [c,d]} \left| \frac{\mu^2}{h_H n(n-1)\mathbf{E}\left(\Delta_{12}(x)\right)} \sum_{i \neq j} G^{-1}(Y_i) G^{-1}(Y_j) \Delta_{ij}(x) \right| \\
&\times \left| H(h_H^{-1}(y - Y_j)) - H(h_H^{-1}(t_y - Y_j)) \right| \\
&\leq C \sup_{y \in [c,d]} \left| \frac{\mu^2 |y - t_y|}{h_H^2 n(n-1)\mathbf{E}\left(\Delta_{12}(x)\right)} \sum_{i \neq j} G^{-1}(Y_i) G^{-1}(Y_j) \Delta_{ij}(x) \right| \\
&\leq C \frac{\mu^2 l_n}{G^2(a_F) h_H^2} \left| \frac{\mu_n^2}{n(n-1)\mathbf{E}\left(\Delta_{12}(x)\right)} \sum_{i,j=1}^{n} \Delta_{ij}(x) \right| \\
&\leq C' \frac{l_n}{h_H^2} \left| \frac{\mu_n^2}{n(n-1)\mathbf{E}\left(\Delta_{12}(x)\right)} \sum_{i,j=1}^{n} \Delta_{ij}(x) \right|.
\end{aligned}$$

Using [21, Lemma 2.2] and $n^\gamma h_H \longrightarrow \infty$, we obtain

$$R_1 = O_{a.co.}\left(\sqrt{\frac{\ln n}{n h_H \Phi_x(h_K)}}\right)$$

and we can derive

$$R_3 = O_{a.co.}\left(\sqrt{\frac{\ln n}{n h_H \Phi_x(h_K)}}\right).$$

It remains to trait the term R_2. For this main, we write

$$\begin{aligned}
&\mathbf{P}\left(R_2 > \epsilon\sqrt{\frac{\ln n}{nh_H\Phi(h_K)}}\right) \\
&= \mathbf{P}\left(\max_{t_y\in\{t_1,\dots,t_{s_n}\}} |\tilde{\Upsilon}_n(x,t_y) - \mathbf{E}(\tilde{\Upsilon}_n(x,t_y))| > \epsilon\sqrt{\frac{\ln n}{nh_H\Phi(h_K)}}\right) \\
&\leq s_n \times \max_{t_y\in\{t_1,\dots,t_{s_n}\}} \mathbf{P}\left(|\tilde{\Upsilon}_n(x,t_y) - \mathbf{E}(\tilde{\Upsilon}_n(x,t_y))| > \epsilon\sqrt{\frac{\ln n}{nh_H\Phi(h_K)}}\right).
\end{aligned}$$

Using arguments as we see before, taking account of the fact that $s_n = Cn^{3\gamma+1/2}$, we deduce, for an appropriate choice of ϵ that

$$R_2 = O_{a.co.}\left(\sqrt{\frac{\ln n}{nh_H\Phi_x(h_K)}}\right).$$

□

P r o o f. of Lemma 3.5.
Because of the assumptions $a_G \leq a_F$ and $b_G \leq b_F$ and the definitions of $\Upsilon_n(x,y)$ and $\tilde{\Upsilon}_n(x,y)$, we can write

$$\begin{aligned}
&\sup_{y\in[c,d]} |\Upsilon_n(x,y) - \tilde{\Upsilon}_n(x,y)| \\
&= \sup_{y\in[c,d]} \left| \frac{\mu_n^2}{h_H n(n-1)\mathbf{E}\left(\Delta_{12}(x)\right)} \sum_{i\neq j} G_n^{-1}(Y_i)G_n^{-1}(Y_j)\Delta_{ij}(x)H_j(y) \right. \\
&\quad \left. - \frac{\mu^2}{h_H n(n-1)\mathbf{E}\left(\Delta_{12}(x)\right)} \sum_{i\neq j} G^{-1}(Y_i)G^{-1}(Y_j)\Delta_{ij}(x)H_j(y) \right| \\
&\leq \left[\frac{|\mu_n^2 - \mu^2|}{G_n^2(a_F)} + \mu^2 \left(\frac{\sup_{y\in[c,d]} |G_n^2(y) - G^2(y)|}{G^2(a_F)G_n^2(a_F)} \right) \right] \\
&\quad \times \left| \sum_{i\neq j} \frac{\Delta_{ij}(x)H_j(y)}{n(n-1)h_H\mathbf{E}\left(\Delta_{12}(x)\right)} \right|.
\end{aligned}$$

From [12, Theorem 3.2] we have $|\mu_n - \mu| = O_{a.s}(n^{-1/2})$.

Moreover, $|G_n(a_F) - G(a_F)| = O_{a.s}(n^{-1/2})$.

On the other hand, $\sup_{y\in[c,d]} |G_n(y) - G(y)| = O(n^{-1/2})$, $\mathbf{P}-a.s.$, which are negligible with respect to $O\left(\sqrt{\frac{\ln(n)}{nh_H\Phi_x(h_K)}}\right)$.

The second term can be treated by following the same steps of the proof of Lemma 3.4 by replacing the terms $P^x_{p,l}(y)$ for $p = 2,3,4$ and $l = 0,1$ by the terms $B^x_{p,l}(y)$ gived by

$$B_{p,l}^x(y) = \frac{1}{n\Phi_x(h_K)} \sum_{i=1}^{n} \frac{K_i(x)\beta_i^{p-2}(x)H_i^l(y)}{h_H^l h_K^{p-2}}. \tag{13}$$

Thus, we have

$$\sup_{y\in[c,d]} |\Upsilon_n(x,y) - \tilde{\Upsilon}_n(x,y)| = O_{a.s.}\left(\sqrt{\frac{\ln n}{nh_H\Phi_x(h_K)}}\right).$$

Using the same arguments, we can proof the second part of Lemma. □

To treat the unifom almost sure convergence of $f_n(y|x)$, we need Lemma 4.1 introduced in [21].

P r o o f. of Lemma 4.3.
The proof is similar to that of Lemma 3.3 and using the equation (5) and hypothesis (U2), we get our result. □

P r o o f. of Lemma 4.4.
By considering the same decompositions and notations (6)-(9), following the same steps as in the proof of Lemma 3.4 and using [21, Lemma 4.1 (i)] instead of [2, Lemma A.1 (i)], we get under assumptions (U1), (U3), (U4) and (U6)

$$\sup_{x\in\mathcal{S}_\mathcal{F}} Q(x) = O(1) \text{ and } \sup_{x\in\mathcal{S}_\mathcal{F}} \sup_{y\in[c,d]} \mathbf{E}(P_{p,l}^x(y)) = O(1), \tag{14}$$

for $p = 2, 3, 4$ and $l = 0, 1$,

$$\sup_{x\in\mathcal{S}_\mathcal{F}} \sup_{y\in[c,d]} |\mathbf{E}(P_{2,1}^x(y))\mathbf{E}(P_{4,0}^x) - \mathbf{E}(P_{2,1}^x(y)P_{4,0}^x)| = O\left(\frac{1}{nh_H\Phi(h_K)}\right)$$

and

$$\sup_{x\in\mathcal{S}_\mathcal{F}} \sup_{y\in[c,d]} |\mathbf{E}(P_{3,1}^x(y))\mathbf{E}(P_{3,0}^x) - \mathbf{E}(P_{3,1}^x(y)P_{3,0}^x)| = O\left(\frac{1}{nh_H\Phi(h_K)}\right),$$

which is, using hypothesis (U5), equals to $O\left(\sqrt{\frac{\ln d_n}{nh_H\Phi(h_K)}}\right)$.

Now we prove that

$$\sup_{x\in\mathcal{S}_\mathcal{F}} |P_{p,l}^x(y) - \mathbf{E}(P_{p,l}^x(y))| = O_{a.co}\left(\sqrt{\frac{\ln d_n}{nh_H^l\Phi(h_K)}}\right).$$

To satisfy this aim, let be $j(x) = \arg\min_{j\in\{1,2,\ldots,d_n\}} d(x, x_j)$ and we consider the following decomposition

$$\begin{aligned}\sup_{x\in\mathcal{S}_\mathcal{F}}\sup_{y\in[c,d]}|P^x_{p,l}(y)-\mathbf{E}(P^x_{p,l}(y))| &\le \sup_{x\in\mathcal{S}_\mathcal{F}}\sup_{y\in[c,d]}|P^x_{p,l}(y)-P^{x_{j(x)}}_{p,l}(y)|\\ &+\sup_{x\in\mathcal{S}_\mathcal{F}}\sup_{y\in[c,d]}|P^{x_{j(x)}}_{p,l}(y)-P^{x_{j(x)}}_{p,l}(t_y)|\\ &+\sup_{x\in\mathcal{S}_\mathcal{F}}\sup_{y\in[c,d]}|P^{x_{j(x)}}_{p,l}(t_y)-\mathbf{E}(P^{x_{j(x)}}_{p,l}(t_y))|\\ &+\sup_{x\in\mathcal{S}_\mathcal{F}}\sup_{y\in[c,d]}|\mathbf{E}(P^{x_{j(x)}}_{p,l}(t_y))-\mathbf{E}(P^{x_{j(x)}}_{p,l}(y))|\\ &+\sup_{x\in\mathcal{S}_\mathcal{F}}\sup_{y\in[c,d]}|\mathbf{E}(P^{x_{j(x)}}_{p,l}(y))-\mathbf{E}(P^x_{p,l}(y))|\\ &:=\sum_{i=1}^{5}D_i^{p,l}.\end{aligned}$$

We start by treating the term $D_1^{p,l}$ and $D_5^{p,l}$

First, let us analyze the term $D_1^{p,l}$.

Under (U3) and (U4), we get

$$D_1^{p,l}\le\frac{Cz_n}{nh_H^l h_K\Phi(h_K)}\sup_{x\in\mathcal{S}_\mathcal{F}}\sum_{i=1}^{n}\frac{\mu H_i^l(y)}{G(Y_i)}1_{B(x,h_K)\cup B(x_{j(x)},h_K)}(X_i).$$

Taking

$$\Psi_i=\frac{Cz_n}{h_H^l h_K\Phi(h_K)}\frac{H_i^l(y)\mu}{G(Y_i)}\sup_{x\in\mathcal{S}_\mathcal{F}}1_{B(x,h_K)\cup B(x_{j(x)},h_K)}(X_i),$$

so, we obtain

$$|\Psi_1|\le\frac{Cz_n}{h_Kh_H^l\Phi(h_K)}\ ,\quad \mathbf{E}|\Psi_1|\le\frac{Cz_n}{h_H^l h_K}\quad\text{and}\quad \mathbf{E}|\Psi_1^2|\le\frac{Cz_n^2}{h_H^{2l}h_K^2\Phi(h_K)}.$$

So, by applying corollary A.8 (ii) in [11] and under the assumptions (U1) and (U5), we obtain

$$D_1^{p,l}=O_{a.co}\left(\sqrt{\frac{\ln d_n}{nh_H^l\Phi(h_K)}}\right). \tag{15}$$

For the term $D_5^{p,l}$, since

$$D_5^{p,l}\le\mathbf{E}\left(\sup_{x\in\mathcal{S}_\mathcal{F}}\sup_{y\in[c,d]}|P^{x_{j(x)}}_{p,l}(y)-P^x_{p,l}(y)|\right).$$

Thus,

$$D_5^{p,l}=O_{a.co}\left(\sqrt{\frac{\ln d_n}{nh_H^l\Phi(h_K)}}\right). \tag{16}$$

Treating the term $D_3^{p,l}$, we have For all $\eta > 0$

$$\begin{aligned}
& \mathbf{P}\left(D_3^{p,l} > \eta\sqrt{\frac{\ln d_n}{nh_H^l\Phi(h_K)}}\right) \\
= \;& \mathbf{P}\left(\sup_{x\in\mathcal{S}_\mathcal{F}}\sup_{y\in[c,d]}|P_{p,l}^{x_{j(x)}}(t_y) - \mathbf{E}(P_{p,l}^{x_{j(x)}}(t_y))| > \eta\sqrt{\frac{\ln d_n}{nh_H^l\Phi(h_K)}}\right) \\
\leq \;& d_n s_n \max_{t_y\in\{t_1,\ldots,t_{s_n}\}}\max_{x_{j(x)}\in\{x_1,\ldots,x_{d_n}\}}\mathbf{P}\Bigg(|P_{p,l}^{x_{j(x)}}(t_y) - \mathbf{E}(P_{p,l}^{x_{j(x)}}(t_y))| \\
> \;& \eta\sqrt{\frac{\ln d_n}{nh_H^l\Phi(h_K)}}\Bigg).
\end{aligned}$$

Taking for $p = 2,3,4, M_i^{(p,l)}$ defined in relation (11) There for, we can apply a Bernstein-type inequality as done in [11, Corollary A.8 (i)], to obtain

$$\mathbf{P}\left(\frac{1}{n}\left|\sum_{j=1}^n M_j^{(p,l)}\right| > \eta\sqrt{\frac{\ln d_n}{nh_H^l\Phi(h_K)}}\right) \leq 2exp\left(-C\eta^2\ln d_n\right).$$

Thus, by choosing α such that $C\eta^2 = \alpha$, we get

$$d_n\mathbf{P}\left(|P_{p,l}^{x_{j(x)}}(t_y) - \mathbf{E}(P_{p,l}^{x_{j(x)}}(t_y))| > \eta\sqrt{\frac{\ln d_n}{nh_H^l\Phi(h_K)}}\right) \leq Cd_n^{1-\alpha}.$$

Then, the fact that $s_n = O(l_n^{-1}) = O(n^{3\gamma+1/2})$ and hypothesis (U5) allow us to write

$$D_3^{p,l} = O_{a.co}\left(\sqrt{\frac{\ln d_n}{nh_H^l\Phi(h_K)}}\right). \tag{17}$$

Treatment of term $D_2^{p,l}$ and $D_4^{p,l}$
Remark that

$$D_2^{p,l} \leq C\frac{l_n}{h_H^{l+1}}\sup_{x\in\mathcal{S}_\mathcal{F}}P_{p,0}^{x_{j(x)}}.$$

In view of relations (14), (17) and the hypothesis (U5), we obtain that

$$D_2^{p,l} = O_{a.co}\left(\sqrt{\frac{\ln d_n}{nh_H^l\Phi(h_K)}}\right) \tag{18}$$

and

$$D_4^{p,l} = O_{a.co}\left(\sqrt{\frac{\ln d_n}{nh_H^l\Phi(h_K)}}\right). \tag{19}$$

Finally, the result of Lemma 4.4 follows from the relations (15)−(19).

The second part can be directly deduced from the proof of the first one such that $\mathbf{E}(\tilde{r}_n(x)) = 1$.

For the last part, it comes straightforward that

$$\inf_{x\in\mathcal{S}_\mathcal{F}} \tilde{r}_n(x) < \frac{1}{2} \Rightarrow \exists x \in \mathcal{S}_\mathcal{F}$$

such that

$$1 - \tilde{r}_n(x) > \frac{1}{2} \Rightarrow \sup_{x\in\mathcal{S}_\mathcal{F}} |1 - \tilde{r}_n(x)| > \frac{1}{2} \Rightarrow \sum_{n=1}^{\infty} \mathbf{P}\left(\inf_{x\in\mathcal{S}_\mathcal{F}} \tilde{r}_n(x) < \frac{1}{2}\right) < \infty.$$

□

Proof. of Lemma 4.5.

By following the same steps as the proof of Lemma 3.5 and studying the uniform consistency of $\sum_{i\neq j} \frac{\Delta_{ij}(x)H_j(y)}{n(n-1)h_H \mathbf{E}(\Delta_{12}(x))}$ as we did in the proof of Lemma 4.4 we get our result.

□

ACKNOWLEDGEMENT

The authors are so grateful to the editor and two reviewers for their valuable suggestions and constructive comments.

(Received August 29, 2022)

REFERENCES

[1] L. Ait Hennania, M. Lemdania, and E. Ould Said: Robust regression analysis for a censored response and functional regressors. J. Nonparametr. Statist. *31* (2019), 221–2430. DOI:10.1080/10485252.2018.1546386

[2] J. Barrientos-Marin, F. Ferraty, and P. Vieu: Locally modelled regression and functional data. J.Nonparametr. Statist. *22* (2010), 617–632. DOI:10.1080/10485250903089930

[3] A. Benkhaled, F. Madani, and S. Khardani: Strong consistency of local linear estimation of a conditional density function under random censorship. Arab. J. Math. *9* (2020), 513–529. DOI:10.1007/s40065-020-00282-1

[4] H. Boudada: A nonparametric estimation of the conditional quantile for truncated and functional data. Int. J. Math. Oper. Res. *21* (2022), 127–140. DOI:10.1504/ijmor.2022.120318

[5] H. Boudada, S. Leulmi, and S. Kharfouch: Rate of the almost sure convergence of a generalized regression estimate based on truncated and functional data. J. Sib. Fed. Univ. – Math. Phys. *13* (2020), 480–491. DOI:10.17516/1997-1397-2020-13-4-480-491

[6] J. Demongeot, A. Laksaci, F. Madani, and M. Rachdi: Functional data analysis: conditional density estimation and its application. Statist. *47* (2013), 26–44. DOI:10.1080/02331888.2011.568117

[7] S. Derrar, A. Laksaci, and E. Ould Saïd: On the nonparametric estimation of the functional ψ-regression for a random left-truncation model. J. Stat. Theory Pract. *9* (2015), 823–849. DOI:10.1080/15598608.2015.1032455

[8] S. Derrar, A. Laksaci, and E. Ould Saïd: M-estimation of the regression function under random left truncation and functional time series model. Statist. Papers *61* (2018), 1181–1202. DOI:10.1007/s00362-018-0979-z

[9] J. Fan: Design-adaptive nonparametric regression. J. Amer. Statist. Assoc. *87* (1992), 998–1004. DOI:10.1080/01621459.1992.10476255

[10] F. Ferraty, A. Laksaci, A. Tadj, and P. Vieu: Rate of uniform consistency for nonparametric estimates with functional variables. J. Statist. Plan, Inference *140* (2010), 335–352. DOI:10.1016/j.jspi.2009.07.019

[11] F. Ferraty and P. Vieu: Nonparametric Functional Data Analysis. Theory and Practice. Springer Ser. Statist. New York 2006.

[12] S. He and G. L. Yang: Estimation of the truncation probability in the random truncation model. Ann. Statist. *26* (1998), 1011–1027. DOI:10.1214/aos/1024691086

[13] N. Helal and E. Ould-Saïd: Kernel conditional quantile estimator under left truncation for functional regressors. Opuscula Math. *36* (2016), 25–48. DOI:10.7494/OpMath.2016.36.1.25

[14] W. Horrigue and E. Ould Saïd: Strong uniform consistency of a nonparametric estimator of a conditional quantile for censored dependent data and functional regressors. Random Oper.Stoch. Equat. *19* (2011), 131–156. DOI:10.1515/ROSE.2011.008

[15] M. Lemdani and E. Ould-Saïd: Asymptotic behavior of the hazard rate kernel estimator under truncated and censored data. Comm. Statist. Theory Methods. *36* (2007), 155–173. DOI:10.1080/03610920600966506

[16] S. Leulmi: Local linear estimation of the conditional quantile for censored data and functional regressors. Comm. Statist. Theory Methods *50* (2019), 3286–3300. DOI:10.1080/03610926.2019.1692033

[17] S. Leulmi: Nonparametric local linear regression estimation for censored data and functional regressors. J. Korean Statist. Soc. (2020), 25–46. DOI:10.1007/s42952-020-00080-7

[18] S. Leulmi and F. Messaci: Local linear estimation of a generalized regression function with functional dependent data. Comm. Statist. Theory Methods *47* (2018), 5795–5811. DOI:10.1080/03610926.2017.1402048

[19] S. Leulmi and F. Messaci: A Class of Local Linear Estimators with Functional Data. J. Sib. Fed. Univ. – Math. Phys. *12* (2019), 379–391. DOI:10.17516/1997-1397-2019-12-3-379-391

[20] B. Mechab, N. Hamidi, and S. Benaissa: Nonparametric estimation of the relative error in functional regression and censored data. Chil. J. Statist. *10* (2019), 177–195.

[21] F. Messaci, N. Nemouchi, and I. Ouassou, M. Rachdi: Local polynomial modelling of the conditional quantile for functional data. Stat. Methods Appl. *24* (2015), 597–622. DOI:10.1007/s10260-015-0296-9

[22] W. Stute: Almost sure representations of the product-limit estimator for truncated data. Ann. Statist. *21* (1993), 146–156. DOI:10.1214/aos/1176349019

[23] M. Woodroofe: Estimating a distribution function with truncated data. Ann. Statist. *13* (1985), 163–177. DOI:10.1214/aos/1176346584

Halima Boudada, Departement of Mathematcs, Frères Mentouri University, Constantine 1, Ain El Bey road, 25017 Constantine. Algeria.
e-mail: halima.boudada@umc.edu.dz

Sarra Leulmi, LAMSAD Laboratory, Departement of Mathematcs, Frères Mentouri University, Constantine 1, Ain El Bey road, 25017 Constantine. Algeria.
e-mail: sara.leulmi@umc.edu.dz