DOI: 10.17516/1997-1397-2022-15-4-523-536
УДК 519

On the Nonparametric Estimation of the Functional Regression Based on Censored Data under Strong Mixing Condition

Farid Leulmi*
Sara Leulmi†
University Frères Mentouri
Constantine, Algeria
Soumia Kharfouchi‡
University Salah Boubnider
Constantine, Algeria

Received 04.02.2022, received in revised form 09.03.2022, accepted 10.05.2022

Abstract. In this paper, we are concerned with local linear nonparametric estimation of the regression function in the censorship model when the covariates take values in a semimetric space. Then, we establish the pointwise almost-complete convergence, with rate, of the proposed estimator when the sample is a strong mixing sequence. To lend further support to our theoretical results, a simulation study is carried out to illustrate the good accuracy of the studied method.

Keywords: functional data, censored data, locally modeled regression, almost-complete convergence, strong mixing.

Citation: F. Leulmi, S. Leulmi, S. Kharfouchi, On the Nonparametric Estimation of the Functional Regression Based on Censored Data under Strong Mixing Condition, J. Sib. Fed. Univ. Math. Phys., 2022, 15(4), 523–536. DOI: 10.17516/1997-1397-2022-15-4-523-536.

1. Introduction and preliminaries

The nonparametric estimation in the functional data is an important subject in modern statistical literatures. This research field is motivated by the fact that several data collected in practice, are given in the form of curves. The monograph of [7] is a pioneer work in the nonparametric setting, where the authors established the pointwise almost-complete convergence for different kernel type estimators.

However, lot of works show that the performance of the local linear method is better than that of the kernel one. Such in [2], where they obtained the rate of the pointwise almost-complete convergence for local linear estimator of the regression function. The uniform convergence of other nonparametric local linear estimators has been investigated in some papers as [6, 13, 17], in the independent and identically distributed (i.i.d.) data case.

Unfortunately, in many practical applications such as reliability and survival time studies, the interest response variable may be incompletely observed, which make the study of censored data

*farid_leulmi@yahoo.fr
†math17sara@yahoo.fr
‡s_kharfouchi@yahoo.fr

more useful in practice. We can see this for example in the works of [1], where the authors gave a family of robust nonparametric estimators for which consistency and asymptotic normality results are established under independent data. For the same data, [10,11] investigated the rates of the pointwise and the uniform almost-complete convergence of a local linear estimator of the conditional quantile and the regression function. She improved that the local linear method outperforms the kernel method even for censored data.

All the above mentioned works concerned the independent functional data case. Nevertheless, in many cases, we face a dependent data. A large studied example is the case of the α-mixing dependence. We refer to [15] for the kernel nonparametric regression estimation under random censorship. [16] examined the almost-complete consistency and the asymptotic normality of the estimator of the relative error regression for the strictly stationary data. Furthermore, [3] used the local linear approach to estimate the conditional density and established its pointwise almost sure convergence, in the censored and functional α-mixing case.

By combining ideas from the two previous works of [11,12] for the local linear estimation of the regression function in the complete dependent and the independent censored, respectively, functional data cases, we propose a novel estimation procedure for the regression function in the case of dependent functional and incomplete data. Among incomplete data models we are here interested in right censoring, which is frequently present in practice.

To our knowledge, the local linear estimation of the regression function combining censored and functional dependent data has not been studied in statistical literature. So, in this work, we address this problem. More precisely, we first present in Section 2 of our paper, a local linear estimator of the regression function. Then, in Section 3, we establish the rate of its pointwise almost-complete convergence under standard conditions. A simulation study is carried out to show the good behaviour of our estimator in Section 4. Finally, the proofs of the main results are evoked in the Appendix.

Throughout this paper the following notations will be adopted. Let $T_U = \sup\{t \in \mathbb{R}; F_U(t) < 1\}$ denote the upper endpoint of the support of F_U, where $F_U(t) = P(U \leqslant t)$ denote the distribution of a real random variable (r.r.v.) U. Furthermore, $\mathcal{F}$ is an infinite-dimensional space equipped with a semimetric d, X is a random variable valued in $\mathcal{F}$, for any $x \in \mathcal{F}$, $h \geqslant 0$, $B(x,h) := \{y \in \mathcal{F}/\ d(x,y) \leqslant h\}$ denotes a closed ball in $\mathcal{F}$ of center x and radius h. We also define $\Phi_x(r_1,r_2) := P(r_1 \leqslant d(x,X) \leqslant r_2)$, where r_1 and r_2 are two real numbers.

For the sake of clarity, we feel welcome to recall some definitions.

• Let $\{Z_i, i = 1,2,\dots\}$ be a strictly stationary sequence of random variables, $F_i^k(Z)$ denotes the σ-algebra generated by $\{Z_j, i \leqslant j \leqslant k\}$. Given a positive integer n, set

$$\alpha(n) = \sup\{|P(A \cap B) - P(A)P(B)| : A \in F_1^k(Z) \text{ and } B \in F_{k+n}^{\infty}(Z), k \in \mathbb{N}^*\}$$

The sequence is said to be α- mixing (strong mixing) if the mixing coefficient $\alpha(n) \to 0$ as $n \to \infty$.

Many processes do satisfy the strong mixing property, see [14] for more details and examples.

• Let $(z_n)_{n\in\mathbb{N}^*}$ be a sequence of real random variables. We say that $(z_n)_{n\in\mathbb{N}^*}$ converge almost-completely (a.co.) toward zero if, and only if, $\forall\epsilon > 0$, $\sum_{n=1}^{\infty} P(|z_n| > \epsilon) < \infty$. Moreover, let $(u_n)_{n\in\mathbb{N}^*}$ be a sequence of positive real numbers; we say that $z_n = O(u_n)$ a.co. if, and only if, $\exists\epsilon > 0$, $\sum_{n=1}^{\infty} P(|z_n| > \epsilon u_n) < \infty$.

It is clear, from Borel Cantelli lemma, that this convergence is stronger than the almost sure one.

2. Definition of the estimator

Consider n pairs of random variables $(X_i, Y_i)_{i=1,\dots,n}$ identically distributed as the pair (X, Y) which is valued in $\mathcal{F} \times \mathbb{R}$.
We report that in the complete case, the local linear estimator of the regression function $m(x) = E(Y|X = x)$ is presented in [2] as follows

$$\overline{m}(x) = \frac{\sum\limits_{i,j=1}^{n} W_{ij}(x) Y_j}{\sum\limits_{i,j=1}^{n} W_{ij}(x)}, \quad \left(\frac{0}{0} = 0\right),$$

with

$$W_{ij}(x) = \beta(X_i, x)\,(\beta(X_i, x) - \beta(X_j, x))\,K(h^{-1}d(X_i, x))K(h^{-1}d(X_j, x)), \tag{1}$$

where $\beta(.,.)$ is a known function from $\mathcal{F} \times \mathcal{F}$ into $\mathbb{R}$ such that, $\forall \xi \in \mathcal{F}$, $\beta(\xi, \xi) = 0$, the function K is a kernel and $h := h_n$ is a sequence of strictly positive real numbers which plays a smoothing parameter role.

As Y_i is not disponible in practice, we can only observe a sample $(X_i, Z_i, \delta_i)_{1 \leqslant i \leqslant n}$ of i.d. observations of $(X, Z = Y \wedge R, \delta)$ where R is nonnegative censoring random variable with unknown continuous survival function G ($\forall t$, $G(t) = P(R > t)$) and $\delta = 1_{\{Y \leqslant R\}}$ (where 1_A denotes the indicator function of the set A) and Y is a nonnegative random variable.

All over this paper, we will assume that the sequences $(X_i)_{1 \leqslant i \leqslant n}$, $(Y_i)_{1 \leqslant i \leqslant n}$ and $(R_i)_{1 \leqslant i \leqslant n}$ are stationary and α-mixing with mixing coefficients $\alpha_1(n)$, $\alpha_2(n)$ and $\alpha_3(n)$ respectively. Notice that, in view of Lemma 2 in [5], we can show that, the sequences $(X_i, Y_i)_{1 \leqslant i \leqslant n}$, $(Z_i)_{1 \leqslant i \leqslant n}$ and then $(X_i, Z_i, \delta_i)_{1 \leqslant i \leqslant n}$ are α-mixing with coefficients $a(n) = 4 \max(\alpha_1(n), \alpha_2(n))$, $b(n) = 4 \max(\alpha_2(n), \alpha_3(n))$ and $\alpha(n) = 4 \max(\alpha_1(n), b(n)) = 4 \max(\alpha_1(n), 4 \max(\alpha_2(n), \alpha_3(n)))$ respectively.

Furthermore, the dependence assumption of $(X_i)_{1 \leqslant i \leqslant n}$, $(Y_i)_{1 \leqslant i \leqslant n}$ and $(R_i)_{1 \leqslant i \leqslant n}$, seems to be more general and one can think to replace it by a classical dependence assumption of $(X_i, Y_i)_{1 \leqslant i \leqslant n}$ and the sequence $(R_i)_{1 \leqslant i \leqslant n}$ is i.i.d. censoring random variable, see for example [3]. Because, since $(X_i, Y_i)_{1 \leqslant i \leqslant n}$ is stationary and α-mixing, it is straightforward that the sequences $(X_i)_{1 \leqslant i \leqslant n}$ and $(Y_i)_{1 \leqslant i \leqslant n}$ are also stationary and α-mixing. This can be deduced from the fact that the later can be seen as a projection-image of the former. On other hand, the α-mixing condition of $(R_i)_{1 \leqslant i \leqslant n}$ is more comprehensive than the independence assumption, we put $\alpha_3 = 0$.

Let (A1) be the following assumptions.

- R and (X, Y) are independent and $T_Y < T_R < \infty$.
- $\exists T < T_Y$ such that $\forall i$, $1 \leqslant i \leqslant n$; $Z_i \leqslant T$.

This assumption is a standard condition in nonparametric censoring estimation which permits us to obtain an unbiased estimator. Like so, the independence assumption between R and (X, Y) is plausible whenever the censoring is independent of the patients modality, $T_Y < T_R$ implies that $G(T) > 0$ because $T < T_Y$.
A feasible local linear nonparametric estimator of $m(x)$, constructed in [11], is defined by

$$\widehat{m}(x) = \frac{\sum\limits_{i,j=1}^{n} W_{ij}(x) \frac{\delta_j Z_j}{G_n(Z_j)}}{\sum\limits_{i,j=1}^{n} W_{ij}(x)}, \quad \left(\frac{0}{0} = 0\right), \tag{2}$$

where $W_{ij}(x)$ is defined in (1) and G_n is the well known [9] estimator of G, which given by

$$G_n(t) = \begin{cases} \prod_{i=1}^{n} \left(1 - \frac{1-\delta_{(i)}}{n-i+1}\right)^{1_{\{Z_{(i)} \leqslant t\}}} & \text{if } t < Z_{(n)} \\ 0 & \text{if } t \geqslant Z_{(n)}, \end{cases} \tag{3}$$

where $Z_{(1)} < Z_{(2)} < \cdots < Z_{(n)}$ are the order statistics of Z_i and $\delta_{(i)}$ the noncensoring indicator corresponding to $Z_{(i)}$. Notice that for all $1 \leqslant j \leqslant n,\ G_n(Z_j) = 0$ implies that $\delta_j = 0$.
From now on, we have that $(X_i, Z_i, \delta_i)_{1\leqslant i \leqslant n}$ is strongly mixing with mixing's coefficient $\alpha(n)$. Now we are in position to give our assumptions and main result.

3. Main results

The aim of this section is to establish the pointwise almost-complete convergence of $\widehat{m}$. For this purpose, we need the following assumptions.

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

(H2) There exists $b > 0$ such that

$$\forall x_1, x_2 \in B(x, h);\ |\ m(x_1) - m(x_2)| \leqslant C_x d^b(x_1, x_2),$$

where C_x is a positive constant depending on x.

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

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

(H4) The kernel K is a positive and differentiable function on its support $[0, 1]$ and

$$\exists\ C, C' > 0; \quad 0 < C1_{[0,1]}(t) \leqslant K(t) \leqslant C'1_{[0,1]}(t) < \infty.$$

(H5) This condition is devided into the two following conditions (H5a) and (H5b).

(H5a) There exist $C > 0,\ a > \sup\left(4, \frac{1+u}{ud}\right)$ satisfying

$$\forall n \in \mathbb{N};\ \ \alpha(n) \leqslant Cn^{-a},$$

where d and u are defined in (H5b) and (H8) respectively.

(H5b) There exist $0 < d \leqslant 1,\ C > 0,\ C' > 0$ such that

$$C' \left[\Phi_x(h)\right]^{1+d} < \psi_x(h) \leqslant C \left[\Phi_x(h)\right]^{1+d},$$

where $\psi_x(h) := \psi_x(0, h)$ and

$$\psi_x(h_1, h_2) := P\left(h_1 \leqslant d(X_1, x) \leqslant h_2, 0 \leqslant d(X_2, x) \leqslant h_2\right).$$

(H6) For all $m \geqslant 2, \delta_m : x \mapsto E(|Y|^m | X = x)$ is a continuous operator at x and

$$\exists C > 0; \quad \sup_{i \neq j} E\left(|Y_i Y_j| | (X_i, X_j)\right) \leqslant C < \infty.$$

(H7)

$$\exists\, n_0 \in \mathbb{N},\, \forall n > n_0,\ \frac{1}{\psi_x(h)} \int_0^1 \psi_x(zh, h) \frac{d}{dz}\left(z^2 K(z)\right) dz > C > 0$$

and

$$h^2 \int\limits_{B(x,h)} \int\limits_{B(x,h)} \beta(u,x)\beta(t,x) dP_{(X_1,X_2)}(u,t) = o\left(\int\limits_{B(x,h)} \int\limits_{B(x,h)} \beta^2(u,x)\beta^2(t,x) dP_{(X_1,X_2)}(u,t) \right),$$

where $dP_{(X_1,X_2)}$ is the joint distribution of (X_1, X_2).

(H8) The bandwidth h satisfies $\lim_{n\to\infty} h = 0$ and $\exists\, \eta_0 > 0,\ u > 0,\ C_1 > 0,\ C_2 > 0$ such that

$$C_1 n^{\frac{3-a}{a+1}+\eta_0} \leqslant \Phi_x(h) \leqslant C_2 n^{-u},$$

with $\eta_0 < \dfrac{a-3}{a+1}$ and $u < 1$.

Remark that these conditions are standard in this context, the hypotheses (H1)–(H5) and (H7)–(H8) are the same conditions assumed in [12]. The condition (H6) is the same condition (H6) in [12] with $\varphi(t) = t$.

Now, we are in position to state the almost-complete convergence of $\widehat{m}(x)$.

Theorem 3.1. *Assume that assumptions (A1) and (H1)–(H8) are satisfied, then*

$$\widehat{m}(x) - m(x) = O(h^b) + O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}} \right).$$

One of the main features of the present paper is studding the local linear estimation under the dependent and censored case, which is generalizes several usual situations. In particular, we consider the independent case (see [11]), the complete case(see [12]) and the kernel method (see [15]).

Proof 3.1. Let us set

$$\widetilde{m}(x) = \frac{\sum\limits_{i,j=1}^{n} W_{ij}(x) \frac{\delta_j Z_j}{G(Z_j)}}{\sum\limits_{i,j=1}^{n} W_{ij}(x)}, \qquad \left(\frac{0}{0} = 0\right), \tag{4}$$

with $W_{ij}(x)$ is defined in (1) and which will play a prominent part in the proof of the Theorem 3.1 thanks to the following decomposition, for all $x \in \mathcal{F}$.

$$\begin{aligned} \widehat{m}(x) - m(x) = & \frac{1}{\widehat{m}_0(x)} \left[(\widehat{m}_1(x) - \widetilde{m}_1(x)) + (\widetilde{m}_1(x) - E\widetilde{m}_1(x)) + (E\widetilde{m}_1(m) - m(x)) \right] + \\ & + \frac{m(x)}{\widehat{m}_0(x)} \left(1 - \widehat{m}_0(x)\right), \end{aligned} \tag{5}$$

where

$$\widehat{m}_1(x) = \frac{1}{n(n-1)E\left[W_{12}(x)\right]} \sum_{i\neq j} W_{ij}(x) \frac{\delta_j Z_j}{G_n(Z_j)}, \quad \widehat{m}_0(x) = \frac{1}{n(n-1)E\left[W_{12}(x)\right]} \sum_{i\neq j} W_{ij}(x) \tag{6}$$

and

$$\widetilde{m}_1(x) = \frac{1}{n(n-1)E\left[W_{12}(x)\right]} \sum_{i \neq j} W_{ij}(x) \frac{\delta_j Z_j}{G(Z_j)}. \tag{7}$$

To treat the pointwise almost-complete convergence of $\widehat{m}(x)$, we need Lemma A1 introduced in [12] and the preliminary tecnical Lemma 3.1 Then, the proof of the Theorem 3.1 is a direct consequence of the following Lemmas. □

In what follows, let C be some strictly positive generic constant and for any $x \in \mathcal{F}$, and for all $i = 1, \ldots, n$, $K_i(x) := K(h^{-1}d(X_i, x))$ and $\beta_i(x) := \beta(X_i, x)$.
As the dependence assumption reveals covariances terms, let us define for $k \in \{0, 2\}$ and $l \in \{0, 1\}$

$$S^2_{n,l,k}(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} |Cov(\Delta_i^{(k,l)}(x), \Delta_j^{(k,l)}(x))|, \tag{8}$$

where, for $i \in \{1, \ldots, n\}$

$$\Delta_i^{(k,l)}(x) = \frac{1}{h^k} \left\{K_i(x)\beta_i^k(x)\delta_i^l Z_i^l G^{-l}(Z_i) - E[K_i(x)\beta_i^k(x)\delta_i^l Z_i^l G^{-l}(Z_i)]\right\}. \tag{9}$$

We now focus on these covariances terms in the following result.

Lemma 3.1. *Under assumptions (A1) and (H1)-(H7), we have*

$$S^2_{n,l,k}(x) = O(n\Phi_x(h)). \tag{10}$$

Proof 3.2. By following the same steps as the proof of Lemma A.2 in [12] we get our result. □

Lemma 3.2. *Assume that hypotheses (A1), (H1)-(H5) and (H7) hold, then*

$$m(x) - E(\widetilde{m}_1(x)) = O\left(h^b\right).$$

Proof 3.3. The bias term is not affected by the dependence condition. Therefore, by the equiprobability of the couples (X_i, Z_i, δ_i), we get

$$E\widetilde{m}_1(x) - m(x) = \frac{1}{E[W_{12}(x)]} E\left\{W_{12}(x)\left[E\left(Z_2 G^{-1}(Z_2)\delta_2 | X_2\right) - m(x)\right]\right\}.$$

Hypothesis (H4), combining with the facts that $E(\delta_2|X_2, Y_2) = G(Y_2)$ and $\delta_2 Z_2 = \delta_2 Y_2$, give that

$$E\left[Z_2 G^{-1}(Z_2)\delta_2 | X_2\right] = E\left[Y_2 G^{-1}(Y_2) E\left(\delta_2 | X_2, Y_2\right) | X_2\right] = m(X_2).$$

Then, we have

$$E\widetilde{m}_1(x) - m(x) = \frac{1}{E[W_{12}(x)]} E\left[W_{12}(x)\left(m(X_2) - m(x)\right)\right]. \tag{11}$$

The claimed result is obtained by using the last relation and the condition (H2). □

Lemma 3.3. *Under assumptions of Theorem 3.1, we get*

$$\widetilde{m}_1(x) - E(\widetilde{m}_1(x)) = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\,\Phi_x(h)}}\right).$$

Proof 3.4. Inspiring by the proof of Lemma 4.4 in [2], we consider the following decomposition

$$\begin{aligned}\widetilde{m}_1(x) &= \frac{1}{n(n-1)E\left[W_{12}(x)\right]}\sum_{i,j=1}^{n} W_{ij}(x)\delta_j Z_j G^{-1}(Z_j) = \\ &= \frac{n^2h^2\Phi_x^2(h)}{n(n-1)E\left[W_{12}(x)\right]}\left[\left(\frac{1}{n\Phi_x(h)}\sum_{j=1}^{n}K_j(x)Z_j\delta_jG^{-1}(Z_j)\right)\left(\frac{1}{n\Phi_x(h)}\sum_{j=1}^{n}\frac{K_j(x)\beta_j^2(x)}{h^2}\right)-\right.\\ &- \left.\left(\frac{1}{n\Phi_x(h)}\sum_{j=1}^{n}\frac{K_j(x)\beta_j(x)Z_j\delta_jG^{-1}(Z_j)}{h}\right)\left(\frac{1}{n\Phi_x(h)}\sum_{j=1}^{n}\frac{K_j(x)\beta_j(x)}{h}\right)\right] = \\ &= Q(x)[D_{2,1}(x)D_{4,0}(x)-D_{3,1}(x)D_{3,0}(x)],\end{aligned} \tag{12}$$

where, for $p\in\{2,3,4\}$ and $l\in\{0,1\}$,

$$D_{p,l}(x)=\frac{1}{n\Phi_x(h)}\sum_{j=1}^{n}\frac{K_j(x)\beta_j^{p-2}(x)Z_j^l\delta_j^lG^{-l}(Z_j)}{h^{p-2}} \quad\text{and}\quad Q(x)=\frac{n^2h^2\Phi_x^2(h)}{n(n-1)E\left[W_{12}(x)\right]}.$$

Notice that, $Q(x)=O(1)$ (see the proof of Lemma 2 in [12]), so, we have to show that, for $p\in\{2,3,4\}$ and $l\in\{0,1\}$

$$\sum_n P\left(|D_{p,l}(x)-E(D_{p,l}(x))|>\epsilon\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right)<\infty,\quad E[D_{p,l}(x)]=O(1),$$

and that almost surely

$$Cov\left[D_{2,1}(x),D_{4,0}(x)\right]=O\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right)$$

and

$$Cov\left[D_{3,1}(x),D_{3,0}(x)\right]=O\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right).$$

• Firstly we have

$$D_{p,l}(x)-ED_{p,l}(x)=\frac{1}{n\Phi_x(h)}\sum_{i=1}^{n}\Delta_i^{(p-2,l)}(x),$$

with $\Delta_i^{(k,l)}(x)$ is defined in (9).

Note that, because $E(\Delta_1^{(k,l)}(x))=0$, $E|\Delta_1^{(k,l)}(x)|^q=O(\Phi_x(h))$ for $q>2$ and using Tchebychev's inequality, we can apply Proposition A.11-i in [7], to get for any $q>2$, $\varepsilon>0$, $r\geqslant 1$ and for some $0<C<\infty$

$$\begin{aligned}P\left(|D_{p,l}(x)-E\left[D_{p,l}(x)\right]|>\varepsilon\right)=P\left(|\sum_{i=1}^{n}\Delta_i^{(p,l)}(x)|>n\varepsilon\Phi_x(h)\right)\leqslant \\ \leqslant C\left[A_1(x)+A_2(x)\right],\end{aligned}\tag{13}$$

where

$$A_1(x)=\left(1+\frac{\varepsilon^2n^2(\Phi_x(h))^2}{rS_{n,l,k}^2(x)}\right)^{-r/2}\quad\text{and}\quad A_2(x)=nr^{-1}\left(\frac{r}{\varepsilon n\Phi_x(h)}\right)^{(a+1)q/(q+a)}.$$

Now, choosing for $\eta > 0$

$$\varepsilon = \eta\sqrt{\frac{\ln n}{n\Phi_x(h)}} \quad \text{and} \quad r = (\ln n)^2.$$

In view of Lemma 3.1, we have $S_{n,l,k}^2(x) = O(n\Phi_x(h))$. So, we obtain

$$A_2(x) \leqslant Cn^{1-\frac{(a+1)q}{2(q+a)}}(\ln n)^{-2+\frac{3(a+1)q}{2(q+a)}}(\Phi_x(h))^{-\frac{(a+1)q}{2(q+a)}}.$$

Next, using (H8), it exists some real number $\nu > 0$ such that

$$A_2(x) = O(n^{-1-\nu}). \tag{14}$$

Moreover, in view of equation (10) and the fact that $\ln(x+1) = x - x^2/2 + o(x^2/2)$ where x tends to zero, we can write

$$A_1(x) \leqslant Cn^{-\eta^2/2}, \tag{15}$$

which shows that $A_1(x)$ is the general term of a convergent series for an appropriate choice of η. Hence, by combining relations (13), (14) and (15), we derive

$$D_{p,l}(x) - ED_{p,l}(x) = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right).$$

• It is easy to see that under (H1), (H3), (H4) and (A1), we get, for $p \in \{2,3,4\}$ and $l \in \{0,1\}$,

$$E[D_{p,l}(x)] = h^{2-p}\Phi_x(h)^{-1}E\left[K_1(x)\beta_1^{p-2}(x)Z_1^l\delta_1^l G^{-l}(Z_1)\right] \leqslant C, \tag{16}$$

the last inequality is obtained by using the Lemma A1(i) in [12] and the condition (A1).
• Finally, by following similar arguments used to prove (10), we obtain

$$Cov\left[D_{2,1}(x), D_{4,0}(x)\right] = O\left(\frac{1}{n\Phi_x(h)}\right)$$

and

$$Cov\left[D_{3,1}(x), D_{3,0}(x)\right] = O\left(\frac{1}{n\Phi_x(h)}\right).$$

In view of (H8), this last rate is negligible with respect to $O\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right)$. The proof is then completed. □

Lemma 3.4 (see [12])**.** *If assumptions (H1),(H3), (H4), (H5a), (H5b), (H7) and (H8) are satisfied, we obtain*

$$\widehat{m}_0(x) - 1 = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\,\Phi_x(h)}}\right) \quad \textit{and} \quad \sum_{n=1}^{\infty} P\left(\widehat{m}_0(x) < \frac{1}{2}\right) < \infty.$$

Lemma 3.5. *Under assumptions (A1), (H1),(H3), (H4), (H5a), (H5b) and (H7), we have*

$$\widehat{m}_1(x) - \widetilde{m}_1(x) = O_{a.co.}\left(\sqrt{\frac{\ln n}{n\,\Phi_x(h)}}\right).$$

Proof 3.5. Because the assumption (A1) and the definitions of $\widehat{m}_1(x)$ and $\widetilde{m}_1(x)$ in (6) and (7), we can write

$$\begin{aligned} |\widehat{m}_1(x)-\widetilde{m}_1(x)| &= \left| \frac{1}{n(n-1)E\left[W_{12}(x)\right]} \sum_{i\neq j} W_{ij}(x)\delta_j Z_j \left(\frac{1}{G_n(Z_j)} - \frac{1}{G(Z_j)} \right) \right| \leqslant \\ &\leqslant \frac{|T| \sup_{t\leqslant T} |G_n(t)-G(t)|}{G_n(T)G(T)} \left| \frac{1}{n(n-1)E\left[W_{12}(x)\right]} \sum_{i\neq j} W_{ij}(x) \right| \leqslant \\ &\leqslant \frac{|T| \sup_{t\leqslant T} |G_n(t)-G(t)|}{G_n(T)G(T)} |\widehat{m}_0(x)|, \end{aligned} \tag{17}$$

where $\widehat{m}_0(x)$ is defined in (6).

In order hands, following [5] and [18], we obtain

$$\sup_{t\leqslant T} |G_n(t)-G(t)| = O_{a.co.}\left(\sqrt{\frac{\ln n}{n}}\right), \tag{18}$$

which is equals to $O_{a.co.}\left(\sqrt{\frac{\ln n}{n\Phi_x(h)}}\right)$. The proof is completed by using Lemma 3.4. □

4. Simulation study

In this section, two examples of simulation are presented to illustrate the performance of the proposed estimator (LLR). More precisely, we compare the LLR estimator to the kernel regression estimator (KR) studied in [15].

For the computation of the (LLR) and the (KR) estimators, we use the quadratic kernel $K(x) = \frac{3}{2}\left(1-x^2\right)1_{[0,1]}(x)$ and the bandwith h is chosen by the 2-fold cross-validation method. Take into account of the smoothness of the curves $X_i(t)$ (see Figs. 1 and 4.), we choose the semi-metric d based on the derivative (for the first example) and the PCA (for the second example) described in [7] (see routines "semimetric.deriv" and "semimetric.pca" in the website http://www.lsp.ups-tlse.fr/staph/npfda) and we take $\beta = d$ (for the LLR estimator).

Example 1. Let us consider the following nonparametric regression model

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

where

$$m(X) = \frac{1}{4}\exp\left\{2 - \frac{1}{\left(\int_0^1 X'(t)dt\right)^2}\right\}$$

and ϵ is the error supposed to be generated by an autoregressive model defined by

$$\epsilon_i = \frac{1}{\sqrt{2}}(\epsilon_{i-1}+\xi_i), \quad i = 1, \cdots, n$$

with ξ_i are centered random variables normally distributed (i.i.d.) with a variance equal to 0.1 $(\xi_i \rightsquigarrow \mathcal{N}(0, 0.1))$. The functional covariate $X(t)$ is defined, for $t \in [0, \pi/3]$ by

$$X(t) = 2 - \cos\left(W\left(t - \frac{2\pi}{3}\right)\right), \quad t \in \left[0, \frac{2\pi}{3}\right],$$

where W is an α-mixing process generated by $W_i = \dfrac{2}{9}W_{i-1} + \eta_i$ with η_i are i.i.d $\mathcal{N}(0,1)$ and are independent from W_i, which is generated independently by $W_0 \rightsquigarrow \mathcal{N}(0,1)$ (see Fig. 1 for a sample of these curves). Notice that the conditional mean function will coincide and will be equal to $m(x)$.

For this model, we adopt the censored mechanism $(X_i, Z_i, \delta_i)_{1\leqslant i\leqslant n}$, where $Z_i = \min(Y_i, R_i)$, $\delta_i = 1_{\{Y_i \leqslant R_i\}}$ and the censoring random variable $R_i = a_i R_{i-1} + \zeta_i$ with $a_i \rightsquigarrow \mathcal{N}(0, 0.1)$ and ζ_i are i.i.d. exp(1.5) and are independent from R_i, which is generated independently by $R_0 \rightsquigarrow$ exp(1.5).

In this simulation, to illustrate the performance of our estimator, we proceed as follows:

• Step 1. For a different sample sizes $n = 100, 200, 300, 500$, we split our data into two subsets:
– $(X_i, Y_i)_{1\leqslant i\leqslant n_1}$: The learning sample used to build the estimators, where $n_1 = n/2$.
– $(X_i, Y_i)_{n_2\leqslant i\leqslant n}$: The testing sample used to make a comparison, with $n_2 = n_1 + 1$.

• Step 2. We calculate the two estimators by using the learning sample and we find the LLR and KR estimators of the conditional expectation ($\widehat{m}$ and $\widehat{m}_{KR}$), for a different sample sizes $n = 100, 200, 300, 500$.

• Step 3. We plot the true values $m(X_i)$ for all i ($n_2 \leqslant i \leqslant n$) against the predicted ones by means of the two estimators, one in each graph (for a fixed sample size $n = 300$, see Fig. 1).

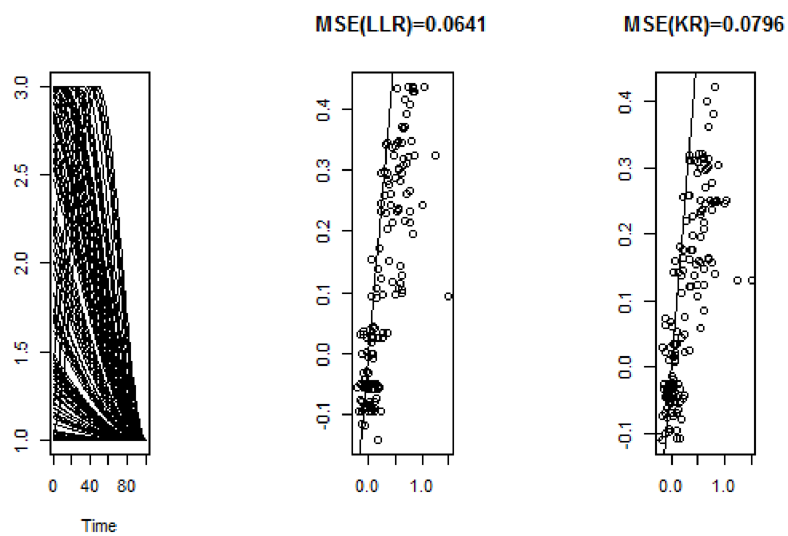

Fig. 1. From left to right the curves X_i, the LLR and KR estimators ($n = 300$)

• Step 4. To be more precise, we measure the prediction accuracy, for different values of n, by using the mean absolute errors (MAE), given by

$$\begin{cases} MAE(LLR) := \dfrac{1}{n_2+1}\displaystyle\sum_{j=n_2}^{n} |\widehat{m}(X_j) - m(X_j)| \\ MAE(KR) := \frac{1}{n_2+1}\sum_{j=n_2}^{n} |\widehat{m}_{KR}(X_j) - m(X_j)| \end{cases}$$

and the prediction errors (MSE) such that

$$\begin{cases} MSE(LLR) := \dfrac{1}{n_2+1}\displaystyle\sum_{j=n_2}^{n} \left(\widehat{m}(X_j) - m(X_j)\right)^2 \\ MSE(KR) := \dfrac{1}{n_2+1}\displaystyle\sum_{j=n_2}^{n} \left(\widehat{m}_{KR}(X_j) - m(X_j)\right)^2 \end{cases}.$$

The obtained results are in the Tab. 1.

Table 1. MSE and MAE comparaison for LLR and KR methods according to sample sizes

	$n = 100$		$n = 200$		$n = 300$		$n = 500$	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
LLR	0.0896	0.2008	0.0775	0.1796	0.0641	0.1529	0.0396	0.1062
KR	0.1190	0.2338	0.0867	0.1933	0.0796	0.1590	0.0471	0.1529

From Tab. 1 and Fig. 1, we observe that the quality of the two estimators perform better when the sample size n increase. Also, we can be seen that our predictor has a good behavior than the kernel one.
We preffer to give a second example to make a better decision.

Example 2. We fixe $n = 200$ and we generated the functional explanatory variables $X(t)$ as follows

$$X_i(t) = a_i \sin\left(4(b_i - t)\right) + c_i, \quad i = 1, \cdots, 200,$$

where $a_i \rightsquigarrow \mathcal{N}(4, 3)$, $c_i \rightsquigarrow \mathcal{N}(0, 0.01)$ and b_i is an α-mixing process generated by $b_i = \frac{1}{3}a_{i-1} + \eta_i$ with η_i are i.i.d. $\mathcal{N}(0,1)$ and are independent from b_i, which is generated independently by $b_0 \rightsquigarrow \mathcal{N}(0, 3)$. We carried out the simulation with a 300-sample of the curve $X(t)$ (see Fig. 2).

The scalar response variable is defined as

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

where

$$m(X) = \int_0^1 \frac{1}{1 + |X(t)|} dt$$

and ϵ is the error generated by an autoregressive model defined by

$$\epsilon_i = \frac{1}{\sqrt{2}}\,\epsilon_{i-1} + \xi_i, \quad i = 1, \cdots, 200$$

with $\xi_i \rightsquigarrow \mathcal{N}(0, 0.1)$. Notice that the conditional median function will coincide and will be equal to $m(x)$.

We also simulate n i.i.d. rondom (R_i) exponentially distributed with parameter λ which is adapted in order to get different censoring rates (CR). We compute our estimator with the observed data $(X_i, Z_i, \delta_i)_{1 \leqslant i \leqslant n}$, where $Z_i = \min(Y_i, R_i)$ and $\delta_i = 1_{\{Y_i \leqslant R_i\}}$. Next, we split our data into a learning sample with size 135 and a test sample with size 65. The true values are plotted against the predicted ones by means of our estimator $\widehat{m}(x)$ and the kernel estimator $\widehat{m}_{KR}(x)$ ($CR = 1.48\,\%$). To be more precise, we measure the prediction accuracy, for different

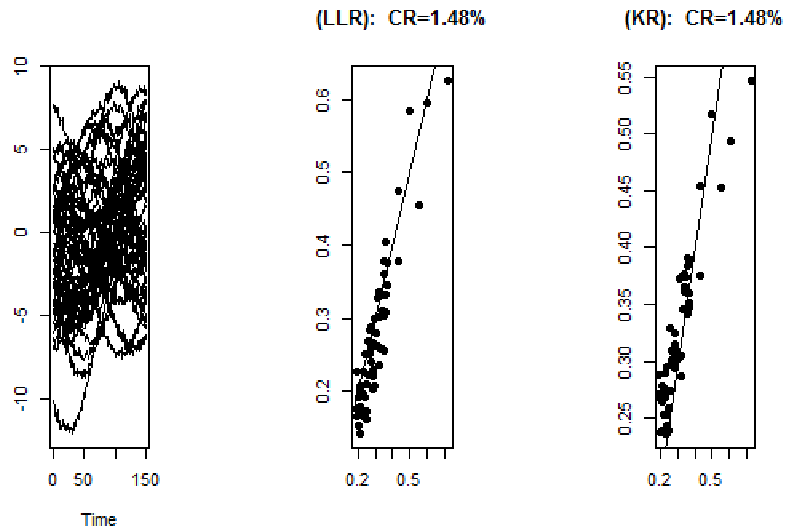

Fig. 2. From left to right the curves X_i, the LLR and KR estimators ($CR = 1.48\,\%$)

values of CR, by using the mean absolute errors (MAE), given by

$$\begin{cases} MAE(LLR) := \dfrac{1}{65} \displaystyle\sum_{j=136}^{200} |\widehat{m}(X_j) - m(X_j)| \\ MAE(KR) := \dfrac{1}{65} \displaystyle\sum_{j=136}^{200} |\widehat{m}_{KR}(X_j) - m(X_j)| \end{cases}$$

and the prediction errors (MSE) such that

$$\begin{cases} MSE(LLR) := \dfrac{1}{65} \displaystyle\sum_{j=136}^{200} (\widehat{m}(X_j) - m(X_j))^2 \\ MSE(KR) := \dfrac{1}{65} \displaystyle\sum_{j=136}^{200} (\widehat{m}_{KR}(X_j) - m(X_j))^2 \end{cases}$$

The obtained results arein the Tab. 2.

Fig. 4. and Tab. 2 show that, our estimator performs better than the kernel estimator. It

Table 2. MSE and MAE comparaison for LLR and KR methods according to CR.

	$CR = 1.48\%$		$CR = 28.67\%$		$CR = 48.15\%$		$CR = 73.33\%$	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
LLR	0.0019	0.0331	0.0182	0.1044	0.0260	0.1271	0.05458	0.2106
KR	0.0037	0.0353	0.0220	0.1098	0.0295	0.1474	0.0610	0.2314

is also clear that, the quality of the both estimators become slightly worse when we have high percentage of censoring, however it remains acceptable.

Conclusion and comments

In conclusion, our Our theoretical and practical studies confirmed without surprise that the quality of the LLR and the KR estimators are better for a bigger sample size n and a weak rate of censoring CR. Furthermore, as for independent and censored data, the LLR estimator stay more accurate than the KR one in all cases.

References

[1] L.Ait Hennania, M.Lemdania, E.Ould Said, Robust regression analysis for a censored response and functional regressors, *Journal of Nonparametric Statistics*, **31**(2018), no. 1, 221–243.

[2] J.Barrientos-Marin, F.Ferraty, P.Vieu, Locally modelled regression and functional data, *Journal of Nonparametric Statistics*, **22**(2010), no. 5, 617–632.
DOI:1 0.1080/10485250903089930

[3] A.Benkhaled, F.Madani, S.Khardani, Strong consistenand cy of local linear estimation of a conditional density function under random censorship, *Arabian Journal of Mathematics*, **9**(2020), 513–529 . DOI: 10.1007s40065-020-00282-1

[4] A.Bouchentouf, A.Hamza, A.Rabhi, Strong uniform consistency rates of conditional hazard estimation in the single functional index model for dependant functional data under random censorship, *International Journal of Statistics and Economic*, **18**(2017), 82–101.

[5] Z.Cai, Estimating a distribution function for censored time series data, *Journal of Multivariate Analysis*, **78**(2001), 299–318.

[6] J.Demongeot, A.Laksaci, F.Madani, M.Rachdi, Functional data: local linear estimation of the conditional density and its application, *Statistics*, **47**(2013), no. 1, 26-44.

[7] F.Ferraty, P.Vieu, Nonparametric functional data analysis. Theory and Practice. Springer Series in Statistics, New York, 2006.

[8] Z.Guessoum, Ould Said, E. On nonparametric estimation of the regression function under random censorship model, *Statistics & Decisions*, **26** (2008), 159–177.

[9] E. M.Kaplan, P.Meier, Nonparametric estimation from incomplete observations, *Journal of the American Statistical Association*, **53** (1958), 457–481.

[10] S.Leulmi, Local linear estimation of the conditional quantile for censored data and functional regressors, *Communications in Statistics – Theory and Methods*, **50**(2019), 1–15.
DOI: 10.1080/03610926.2019.1692033

[11] S.Leulmi, Nonparametric local linear regression estimation for censored data and functional regressors, *Journal of the Korean Statistical Society* (2020), 1–22.

[12] S.Leulmi, F.Messaci, Local linear estimation of a generalized regression function with functional dependent data, *Communications in Statistics – Theory and Methods*, **47**(2018), no.23, 5795–5811.

[13] S.Leulmi, F.Messaci, A class of local linear estimators with functional data, *J. Sib. Fed. Univ. Math. Phys.,* **12**(2019), no. 3, 379–391. DOI: 10.17516/1997-1397-2019-12-3-379-391.

[14] Z.Lin, C.Lu, Limit theory of mixing dependent random variables.Mathematics and its applications, Beijing: Sciences Press, Kluwer Academic Publishers, 1996.

[15] N.Ling, Y.Liu, The kernel regression estimation for randomly censored functional stationary ergodic data, *Communications in Statistics-Theory and Methods*, (2016). DOI: 10.1080/03610926.2016.1185117.

[16] B.Mechab, N.Hamidi, S.Benaissa, Nonparametric estimation of the relative error in functional regression and censored data, *Chilean Journal of Statistics*, **10**(2019), no. 2, 177–195.

[17] F.Messaci, N.Nemouchi, I.Ouassou, M.Rachdi, Local polynomial modelling of the conditional quantile for functional data, *Statistical Methods & Applications*, **24**(2015), no.4, 597–622.

[18] N.Rouabah, N.Nemouchi, F.Messaci, A rate of consistency for nonparametric estimators of the distribution function based on censored dependent data, *Statistical Methods & Applications*, **28**(2019), 259–280. DOI: 110.1007/s10260-018-00445-7.

О непараметрической оценке функциональной регрессии на основе цензурированных данных в условиях сильного перемешивания

Фарид Леулми
Сара Леулми
Университет Фререс Ментури
Константин, Алжир
Сумия Харфучи
Университет Салах Бубнидер
Константин, Алжир

Аннотация. В этой статье мы занимаемся локальной линейной непараметрической оценкой функции регрессии в модели цензуры, когда ковариаты принимают значения в полуметрическом пространстве. Затем мы устанавливаем поточечную почти полную сходимость со скоростью предложенной оценки, когда выборка представляет собой последовательность сильного перемешивания. Для дальнейшего подтверждения наших теоретических результатов было проведено имитационное исследование, иллюстрирующее хорошую точность изучаемого метода.

Ключевые слова: функциональные данные, подвергнутые цензуре данные, локально смоделированная регрессия, почти полная конвергенция, сильное перемешивание.