



Abdelkader Benkhaled · Fethi Madani · Salah Khardani

# Strong consistency of local linear estimation of a conditional density function under random censorship

Received: 4 May 2019 / Accepted: 15 April 2020 / Published online: 16 May 2020  
© The Author(s) 2020

**Abstract** In this paper, we study nonparametric local linear estimation of the conditional density of a randomly censored scalar response variable given a functional random covariate. We establish under general conditions the pointwise almost sure convergence with rates of this estimator under  $\alpha$ -mixing dependence. Finally, to show interests of our results, on the practical point of view, we have conducted a computational study, first on a simulated data and, then on some real data concerning Kidney transplant data.

**Mathematics Subject Classification** 62G05 · 62G07 · 62G08 · 62G35 · 62G20

## 1 Introduction

Conditional density plays an important role; not only in exploring relationships between responses and covariates, but also in financial econometrics (see Ait-Sahalia [1]). A vast variety of papers use the estimators of conditional densities as building blocks. These papers include those of Robinson [20], Tjøstheim [23], among others. However, in all of these papers, the conditional density function is indirectly estimated. Hyndman et al. [13] have studied the kernel estimator of conditional density estimator and its bias corrected version. There are many advantages of using local linear regression, such as the lack of boundary modifications, high minimax efficiency, and ease of implementation. Then, Bashtannyk and Hyndman [3] have suggested several simple and useful rules for selecting bandwidths for conditional density estimation. Hall et al. [11] applied the cross-validation technique to estimate the conditional density. Fan and Yim [9] proposed a consistent data-driven bandwidth selection procedure for estimating the conditional density functions.

In the last decade, the kernel method has been largely used for nonparametric functional data study; in this context, we refer to the monograph of Ferraty and Vieu [10]. Since then, many interesting publications have appeared. According to the literature, results on the local linear modeling in the functional data setting are limited. Baillo and Grané [4] proposed a local linear estimator (LLE) of the regression operator when the explanatory variable takes values in a Hilbert space, and then when the explanatory variable takes values in a semi-metric space. Demongeot et al. [7] presented local linear estimation of the conditional density when

---

A. Benkhaled  
University Djilali Liabes of Sidi Bel Abbès, Sidi Bel Abbès, Algeria  
E-mail: benkhaled08@yahoo.fr

F. Madani (✉)  
Laboratory of Stochastic Models, Statistic and Applications (LMSA), University Tahar Moulay of Saïda, Saïda, Algeria  
E-mail: fethi.madani@univ-saïda.dz

S. Khardani  
LMPA, ULCO, University of Lille, Lille, France  
E-mail: khardani\_salah@yahoo.fr



the data are functional. Furthermore, Messaci et al. [19] used the same approach to estimate the conditional quantile of a scalar response given a functional explanatory variable in the i.i.d. case. In these papers, it is assumed that the observations are complete.

Censored data analysis is a major issue in survival studies. Censored data, truncated data, missing data, and current status data are among the complex data structures in which only partial information on the variable of interest is available (see Kaplan and Meier [14]). In real data applications, censoring is a condition in which the value of a measurement or observation is only partially known. For this, we observe the censored lifetime ( $C_i$ ) for  $i = 1, \dots, n$  of the variable under study instead of the real lifetime ( $Y_i$ ) for  $i = 1, \dots, n$  (which has a continuous cumulative function (cdf)  $F(\cdot)$ ). We assume that  $\{Y_i, i \geq 1\}$  is a stationary sequence of lifetimes and  $\{C_i, i \geq 1\}$  is a sequence of i.i.d censoring random variable with common unknown continuous (cdf)  $G(\cdot)$ , and we observe only the  $n$  pairs  $\{(T_i, \delta_i), i = 1, \dots, n\}$ , where  $T_i = Y_i \wedge C_i$  and  $\delta_i = \mathbb{1}_{\{Y_i \leq C_i\}}$  (with  $\wedge$  denoting minimum and  $\mathbb{1}_A$  denoting the indicator function on a set  $A$ ).

The distinguishing characteristic of censoring has attracted the attention of many researchers. Beran [5] introduced a nonparametric estimate of the conditional survival function and showed some consistency results. Many other properties of the conditional distribution have been broadly studied in the literature (see Stute [22]). Furthermore, in the same context, many results on conditional quantile and conditional mode have been given (see Horrigue and Ould Saïd [12], Khardani and Thiam [15]).

The present study extends the result of Demongeot et al. [7] to censored data under general conditions. We establish the almost sure consistencies with convergence rates of the conditional density estimator when the explanatory variable is of functional type.

The paper is organized as follows. Section 2 is devoted to the presentation of our estimator. In Sect. 3, we introduce notations and assumptions, and state the main results. In Sect. 4, we conduct a simulation study that shows the performance of the considered estimator. Finally, proofs of our results are gathered in Sect. 5.

## 2 Construction of the conditional density estimator

Consider  $n$  pairs of random variables  $(X_i, Y_i)$  for  $i = 1, \dots, n$  drawn from the pair  $(X, Y)$  with values in  $\mathcal{F} \times \mathcal{R}$ , where  $\mathcal{F}$  is a semi-metric space equipped with a semi-metric  $d$ . In this paper, we consider the problem of nonparametric estimation of the conditional  $Y$  given  $X = x$  when the responses variable ( $Y_i$ ) for  $i = 1, \dots, n$  are right censored and when the observations  $(X_i, Y_i)$  for  $i = 1, \dots, n$  are strongly mixing. Furthermore, we denote by  $(C_i)$  for  $i = 1, \dots, n$  the censoring random variables which are independent and identically distributed with a common unknown continuous distribution function  $G$ . Thus, we observe the triplets  $(X_i, T_i, \delta_i)$  for  $i = 1, \dots, n$ , where  $T_i = Y_i \wedge C_i$  and  $\delta_i = \mathbb{1}_{\{Y_i \leq C_i\}}$  (with  $\wedge$  denoting minimum and  $\mathbb{1}_A$  denoting the indicator function on a set  $A$ ). We suppose that  $(Y_i)$  for  $i = 1, \dots, n$  and  $(C_i)$  for  $i = 1, \dots, n$  are independent which ensures the identifiability of the model. In the case of complete data, we adopt the fast functional locally modeling, introduced by Barrientos-Marin et al. [2] for regression analysis, that is, we estimate the conditional density  $\zeta(\cdot|x)$  by  $\hat{a}$  which is obtained by minimizing the following quantity:

$$\min_{(a,b) \in \mathcal{R}^2} \sum_{i=1}^n \left( h_H^{-1} H(h_H^{-1}(y - T_i)) - a - b\beta(X_i, x) \right)^2 K(h_K^{-1}\delta(x, X_i)),$$

where  $\beta(\cdot, \cdot)$  is a known function from  $\mathcal{F}^2$  into  $\mathcal{R}$ , such that  $\forall \xi \in \mathcal{F}, \beta(\xi, \xi) = 0$ , where  $K$  and  $H$  are kernels,  $h_K = h_{K,n}$  (resp.  $h_H = h_{H,n}$ ) is a sequence of positive real numbers, and  $\delta(\cdot, \cdot)$  is chosen as a function of  $\mathcal{F} \times \mathcal{F}$ , such that  $d(\cdot, \cdot) = |\delta(\cdot, \cdot)|$ . Here, we denote  $\hat{a}$  by  $\zeta_n(\cdot|x)$ . Then, the expression of  $\zeta_n(\cdot|x)$  is given as:

$$\zeta_n(y|x) = \frac{\sum_{i,j=1}^n W_{ij}(x) H(h_H^{-1}(y - T_i))}{h_H \sum_{i,j=1}^n W_{ij}(x)},$$

where

$$W_{ij}(x) = \beta(X_i, x) [\beta(X_i, x) - \beta(X_j, x)] K(h_K^{-1}\delta(x, X_i)) K(h_K^{-1}\delta(x, X_j)),$$

with the convention  $0/0 = 0$ .

We assume that there exists a certain compact set  $\Omega \subset \mathbb{R}$ , such that  $\zeta(y|x)$  has an unique mode  $\theta(x)$  on  $\Omega$ , where:

$$\theta(x) = \arg \sup_{y \in \Omega} \zeta(y|x).$$



In the censored case, we adapt the idea of Carbonez et al. (1995), Kohler et al. (2002), and Khardani et al. (2010) to the infinite dimension case using a smooth distribution function  $H(\cdot)$  instead of a step function. Then, we get the "pseudo"-estimator of  $\zeta(\cdot|x)$ :

$$\tilde{\zeta}_n(y|x) = \frac{\sum_{i,j=1}^n \delta_i \bar{G}^{-1}(T_i) W_{ij}(x) H(h_H^{-1}(y - T_i))}{h_H \sum_{i,j=1}^n W_{ij}(x)} \doteq \frac{\tilde{f}_n^x(y)}{\gamma_n(x)},$$

where

$$\tilde{f}_n^x(y) = \frac{1}{n(n-1)h_H \mathbb{E}[W_{12}(x)]} \sum_{i \neq j} \delta_i \bar{G}^{-1}(T_i) W_{ij}(x) H(h_H^{-1}(y - T_i)),$$

and

$$\gamma_n(x) = \frac{1}{n(n-1)\mathbb{E}[W_{12}(x)]} \sum_{i \neq j} W_{ij}(x).$$

The cumulative distribution function  $G$ , of the censoring random variables, is estimated by the Kaplan–Meier (1958) estimator defined as:

$$\bar{G}_n(y) = \begin{cases} \prod_{i=1}^n \left(1 - \frac{1 - \delta_{(i)}}{n - i + 1}\right)^{\mathbb{1}_{\{T_{(i)} \leq y\}}} & \text{if } y < T_{(n)} \\ 0 & \text{otherwise,} \end{cases}$$

where  $T_{(1)} < T_{(2)} < \dots < T_{(n)}$  are the order statistics of  $T_i$  and  $\delta_{(i)}$  is the concomitant of  $T_{(i)}$ , which is known to be uniformly convergent to  $\bar{G}$ .

Therefore, a feasible estimator of  $\zeta(\cdot|x)$  is given by:

$$\hat{\zeta}_n(y|x) = \frac{\hat{f}_n^x(y)}{\gamma_n(x)},$$

where

$$\hat{f}_n^x(y) = \frac{1}{n(n-1)h_H \mathbb{E}[W_{12}(x)]} \sum_{i \neq j} \delta_i \bar{G}_n^{-1}(T_i) W_{ij}(x) H(h_H^{-1}(y - T_i)).$$

Then, a natural estimator of  $\theta(x)$  is defined by:

$$\hat{\theta}_n(x) = \arg \sup_{y \in \Omega} \hat{\zeta}_n(y|x).$$

### 3 Assumptions and main results

Let  $\mathcal{F}_i^k(Z)$  denote the  $\sigma$ - algebra generated by  $\{Z_j, i \leq j \leq k\}$ .

**Definition 3.1** Let  $\{Z_i, i = 1, 2, \dots\}$  be a strictly stationary sequence of random variables. Given a positive integer  $n$ , set:

$$\alpha(n) = \sup \left\{ |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \mathcal{F}_1^k(Z) \text{ and } B \in \mathcal{F}_{k+n}^\infty(Z), k \in \mathbb{N}^* \right\}.$$

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

This condition was introduced in 1956 by Rosenblatt. The strong-mixing condition is reasonably weak and has many practical applications (see Doukhan [8], for more details).

### 3.1 Almost sure consistency of the conditional density estimator

Our first result concerns the almost sure convergence of the LLE of the conditional density. We introduce some conditions that are required to state our asymptotic result. Throughout the paper,  $x$  denotes a fixed point in  $\mathcal{F}$ ;  $N_x$  denotes a fixed neighborhood of  $x$ . For any df  $L$ , let  $\tau_L := \sup\{y : L(y) < 1\}$  be its support's right endpoint and assume that  $\theta(x) \in \Omega \subset (-\infty, \tau]$ , where  $\tau < \tau_G \wedge \tau_F$ ,  $B(x, r) = \{x' \in \mathcal{F} / |\delta(x', x)| \leq r\}$  and  $\phi_x(r_1, r_2) = \mathbb{P}(r_2 \leq \delta(X, x) \leq r_1)$ .

Note that our nonparametric model is quite general in the sense that we just need the following assumptions:

- (H1) For any  $r > 0$ ,  $\phi_x(r) := \phi_x(-r, r) > 0$
- (H2) (i) The conditional density  $\zeta^x$  is such that: there exist  $b_1 > 0$ ,  $b_2 > 0$ ,  $\forall (y_1, y_2) \in \Omega^2$ , and  $\forall (x_1, x_2) \in N_x \times N_x$ :

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

where  $C_x$  is a positive constant depending on  $x$ .

- (ii)  $\zeta(\cdot|x)$  is twice differentiable, its second derivative  $\zeta^{(2)}(\cdot|x)$  is continuous on a neighborhood of  $\theta(x)$  and  $\zeta^{(2)}(\theta(x)|x) < 0$ .
- (H3) The function  $\beta(\cdot, \cdot)$  is such that:

$$\forall x' \in \mathcal{F}, C_1 |\delta(x, x')| \leq |\beta(x, x')| \leq C_2 |\delta(x, x')|, \text{ where } C_1 > 0, C_2 > 0.$$

- (H4) The sequence  $(X_i, Y_i)_{i \in \mathbb{N}}$  satisfies:  $\exists a > 0, \exists c > 0 \forall n \in \mathbb{N} \alpha(n) \leq cn^{-a}$  and:

$$\max_{i \neq j} \mathbb{P}((X_i, X_j) \in B(x, h) \times B(x, h)) = \varphi_x(h) > 0.$$

- (H5) The conditional density of  $(Y_i, Y_j)$  given  $(X_i, X_j)$  exists and is bounded.
- (H6)  $K$  is a positive, differentiable function with support  $[-1, 1]$ .
- (H7)  $H$  is a positive, bounded, Lipschitzian continuous function, such that:

$$\int |t|^{b_2} H(t) dt < \infty \text{ and } \int H^2(t) dt < \infty.$$

- (H8) The bandwidth  $h_K$  satisfies: there exists an integer  $n_0$ , such that:

$$\forall n > n_0, -\frac{1}{\phi_x(h_K)} \int_{-1}^1 \phi_x(zh_K, h_K) \frac{d}{dz} (z^2 K(z)) dz > C_3 > 0,$$

and

$$h_K \int_{B(x, h_K)} \beta(u, x) dP(u) = o \left( \int_{B(x, h_K)} \beta^2(u, x) dP(u) \right),$$

where  $dP(x)$  is the cumulative distribution.

- (H9)  $\lim_{n \rightarrow \infty} h_H = 0$  and  $\exists \beta_1 > 0$ , such that  $\lim_{n \rightarrow \infty} n^{\beta_1} h_H = \infty$ .
- (H10) (i)  $\lim_{n \rightarrow \infty} h_K = 0, \lim_{n \rightarrow \infty} \frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)} = 0$ .
- (ii)  $Cn^{\frac{(3-a)}{(a+1)} + \frac{3\beta_1+1}{a+1}} \log n [\log_2 n]^{6/(a+1)} \leq h_H \chi_x^{1/2}(h_K)$ , where  $\chi_x(h) = \max(\phi_x^2(h), \varphi_x(h))$ .

*Remarks on the assumptions* Most of the assumptions are common in nonparametric functional data analysis (NFDA) context. More precisely, assumption (H1) is usually used in NFDA and it is linked with the topological structure of the functional space,  $\mathcal{F}$ , of the explanatory variable  $X$  (see Ferraty and Vieu [10] for more discussions). Furthermore, assumptions (H2) and (H3) are mild regularity assumptions on the conditional density function. Finally, conditions (H4), (H5), and (H10) are technical assumptions (see Ferraty and Vieu [10] for the constant local method case).

**Theorem 3.2** Under assumptions (H1)–(H10), we have:

$$\sup_{y \in \Omega} |\widehat{\zeta}_n(y|x) - \zeta(y|x)| = O\left(h_K^{b_1} + h_H^{b_2}\right) + O\left(\sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}}\right), \text{ a.s.}$$

*Proof of Theorem 3.2* The proof is a direct consequence of the following decomposition  $\forall y \in \Omega$  :

$$\begin{aligned} \widehat{\zeta}_n(y|x) - \zeta(y|x) &= \frac{\widehat{f}_n^x(y) - \tilde{f}_n^x(y)}{\gamma_n(x)} + \frac{\tilde{f}_n^x(y) - \mathbb{E}[\tilde{f}_n^x(y)]}{\gamma_n(x)} \\ &\quad + \frac{E[\tilde{f}_n^x(y)] - \zeta(y|x)}{\gamma_n(x)} + \zeta(y|x) \frac{(1 - \gamma_n(x))}{\gamma_n(x)}, \end{aligned} \tag{1}$$

where

$$\begin{aligned} \widehat{f}_n^x(y) &= \frac{1}{n(n-1)h_H \mathbb{E}[W_{12}(x)]} \sum_{i \neq j} \delta_i \overline{G}_n^{-1}(T_i) W_{ij}(x) H(h_H^{-1}(y - Y_i)), \\ \tilde{f}_n^x(y) &= \frac{1}{n(n-1)h_H \mathbb{E}[W_{12}(x)]} \sum_{i \neq j} \delta_i \overline{G}^{-1}(T_i) W_{ij}(x) H(h_H^{-1}(y - Y_i)), \end{aligned}$$

and

$$\gamma_n(x) = \frac{1}{n(n-1)\mathbb{E}[W_{12}(x)]} \sum_{i \neq j} W_{ij}(x).$$

and the following Lemmas 3.3–3.6. □

**Lemma 3.3** (cf. [7])

Under assumptions (H1), (H3), (H4),(H6), (H8), and (H10), we have<sup>1</sup>

$$1 - \gamma_n(x) = O\left(\sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n \phi_x^2(h_K)}}\right), \text{ a.co.,}$$

and

$$\exists \delta > 0, \text{ such that } \sum_{i=1}^{\infty} \mathbb{P}(\gamma_n(x) < \delta) < \infty.$$

**Lemma 3.4** Under assumptions (H1), (H2), and (H7), we obtain:

$$\sup_{y \in \Omega} \left| E[\tilde{f}_n^x(y)] - \zeta(y|x) \right| = O\left(h_K^{b_1} + h_H^{b_2}\right), \text{ a.s.}$$

**Lemma 3.5** Under assumptions of Theorem 3.2, we get:

$$\sup_{y \in \Omega} \left| \tilde{f}_n^x(y) - \mathbb{E}[\tilde{f}_n^x(y)] \right| = O\left(\sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}}\right), \text{ a.co.}$$

**Lemma 3.6** Under assumptions (H1), (H6), (H8),(H9), and (H10), we have:

$$\sup_{y \in \Omega} \left| \widehat{f}_n^x(y) - \tilde{f}_n^x(y) \right| = O\left(\sqrt{\frac{\log(\log n)}{n}}\right), \text{ a.s.}$$

<sup>1</sup> Let  $(z_n)_{n \in \mathbb{N}}$  be a sequence of real r.v.'s; we say that  $z_n$  converges almost completely (a.co.) to zero if and only if  $\forall \epsilon > 0, \sum_{n=1}^{\infty} \mathbb{P}(|z_n| > \epsilon) < \infty$ . Moreover, let  $(u_n)_{n \in \mathbb{N}^*}$  be a sequence of positive real numbers; we say that  $z_n = O(u_n)$ , a.s. if and only if  $\exists \epsilon > 0, \sum_{n=1}^{\infty} \mathbb{P}(|z_n| > \epsilon u_n) < \infty$ . This kind of convergence implies both almost sure convergence and convergence in probability (see Sarda and Vieu [21]).

### 3.2 Almost sure consistency of the conditional mode estimator

The convergence rate of the LLE of the conditional mode is a direct consequence of the previous result. Thus, in addition to the previous conditions, we assume that:

(H11) There exists  $\theta(x) \subset \Omega$ , such that  $\zeta(y|x) < \zeta(\theta(x)|x)$ , for all  $y \neq \theta(x)$ ,  $y \in \Omega$ .

Then, the asymptotic behaviour of  $\hat{\theta}_n(x)$  is given in the following Theorem.

**Theorem 3.7** Assume that (H1)–(H11) hold. We have:

$$|\hat{\theta}_n(x) - \theta(x)| = O\left(h_K^{b_1/2} + h_H^{b_2/2}\right) + O\left(\left(\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}\right)^{\frac{1}{4}}\right), \text{ a.s.}$$

*Proof of Theorem 3.7* We easily can show that:

$$|\zeta(\hat{\theta}_n(x)|x) - \zeta(\theta(x)|x)| \leq 2 \sup_{y \in \Omega} |\hat{\zeta}_n(y|x) - \zeta(y|x)|. \quad (2)$$

Making use of a Taylor expansion of the function  $\zeta(\cdot|x)$ , we get:

$$\zeta(\hat{\theta}_n(x)|x) - \zeta(\theta(x)|x) = \frac{1}{2}(\hat{\theta}_n(x) - \theta(x))^2 \zeta^{(2)}(\theta^*(x)|x), \quad (3)$$

where  $\theta^*(x)$  is between  $\theta(x)$  and  $\hat{\theta}_n(x)$

Combining Eqs. (22) and (23), we find:

$$|\hat{\theta}_n(x) - \theta(x)| \leq 2 \sqrt{\frac{\sup_{y \in \Omega} |\hat{\zeta}_n(y|x) - \zeta(y|x)|}{\zeta^{(2)}(\theta^*(x)|x)}}.$$

The almost sure consistency of  $\hat{\theta}_n(x)$  is an immediate consequence of Theorem 3.2.  $\square$

## 4 Computational studies

### 4.1 On simulated data

In this subsection, a simulation study is carried out to investigate the finite sample performance of the local linear estimator  $f_{LL}(y|x)$  of the conditional density function under right-censored and functional-dependent data. As common to all, the applicability of asymptotic normality result requires a practical estimation of the asymptotic bias and variance. For this, we neglect the bias term and we use a plug-in approach to construct an estimator of the asymptotic variance of the conditional density function given by:

$$\bar{G}^{-1}(y) \frac{f^x(y)}{h_H} \frac{M_2}{M_1^2 n \phi_x(h_K)} \int H^2(t) dt. \quad (4)$$

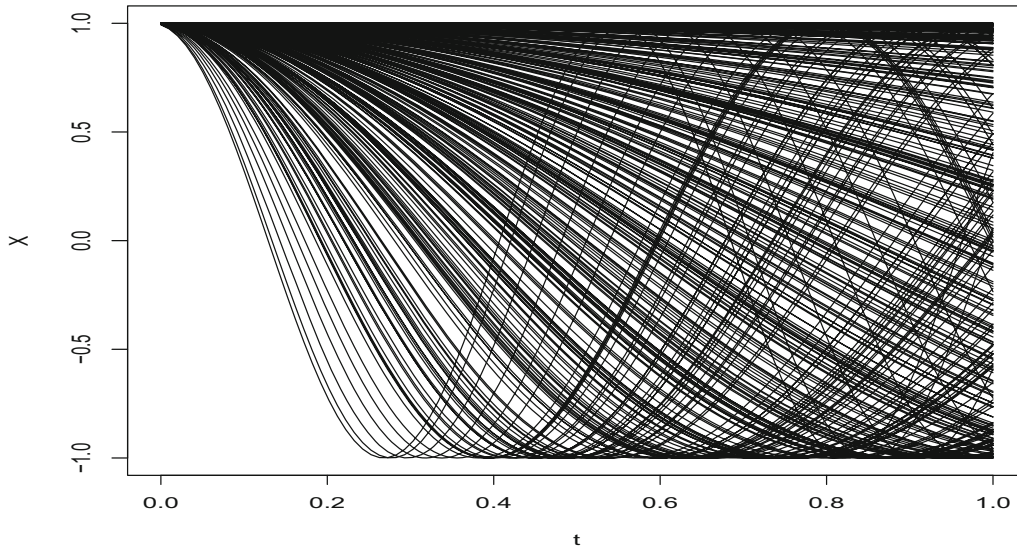
To test the effectiveness of the asymptotic normality result and to attain this purpose, let us consider the following regression model where the response is a scalar:  $Y_i = r(X_i) + \epsilon_i$ ,  $i = 1, \dots, n$ , where  $\epsilon_i$  is the error generated by an autoregressive model defined by:

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

with  $\{\eta_i\}_i$  is a sequence of i.i.d. random variables normally distributed with a variance equal to 0.1. The explanatory variables are constructed by:

$$X(t) = A(2 - \cos(\pi t W)) + (1 - A) \cos(\pi t W), \quad t \in [0, 1],$$





**Fig. 1** A set of 300 simulated curves

where  $W$  is generated from a Gaussian distribution  $\mathcal{N}(0, 1)$  and  $A$  is a random variable Bernoulli distributed with parameter  $p = 0.5$ .  $X'_i$ 's are generated from 300 curves and are plotted in Fig. 1.

On the other hand,  $n$  i.i.d. random variables  $(C_i)_i$  are simulated through the exponential distribution  $\mathcal{E}(\lambda)$  and for  $i = 1, \dots, n = 300$ , the scalar response  $Y_i$  is computed by considering the following operator:

$$r(X) = 4 \exp \left\{ \frac{1}{2 + \int_0^{\pi/2} |X_i(t)|^2} \right\}.$$

Given  $X = x$ , we can easily see that  $Y$  is as a gaussian distribution  $\mathcal{N}(r(x), 0.2)$ . Then, we can get the corresponding conditional density, which is explicitly defined by

$$f^{Y|X}(y|x) = \frac{1}{\sqrt{2\pi} \times 0.3} \exp \left\{ -\frac{1}{2 \times 0.3} (y - r(x))^2 \right\}$$

Therefore, the conditional mode, the conditional mean  $r(x)$ , and the conditional median functions will coincide and will be equal to  $r(x)$ , for any fixed  $x$ . Our purpose, now, consists in evaluating the accuracy of the conditional mode function estimator based on randomly censored data. The computation of this estimator is based on the observed data  $(X_i, T_i, \delta_i)_{i=1, \dots, n}$ , where  $T_i = \min(Y_i, C_i)$  and  $\delta_i = \mathbb{1}_{\{Y_i \leq C_i\}}$ .

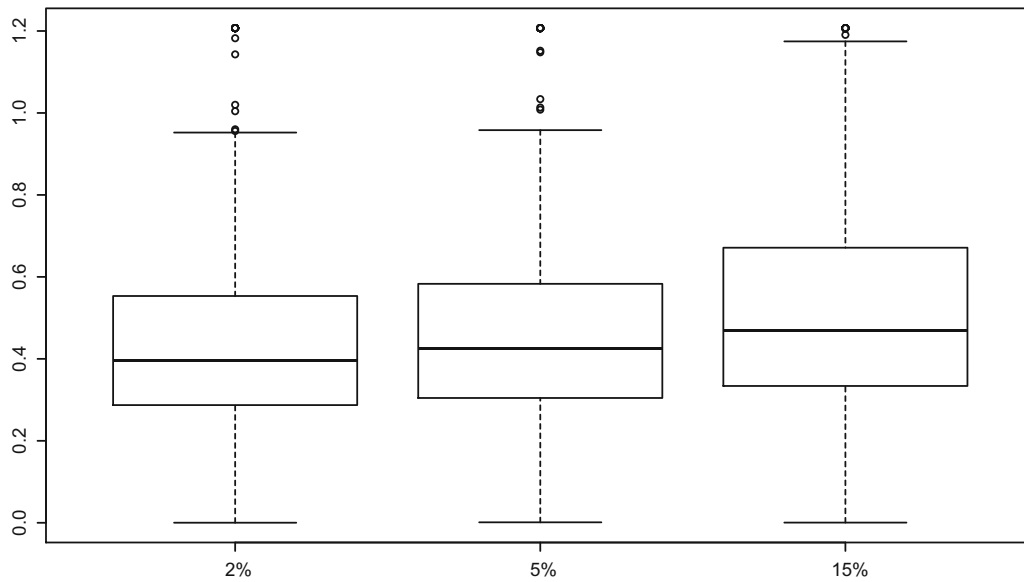
In this simulation study, we present only results of the case where  $i = 2$  and  $q = 1$ . For this, we take  $K_0(s) = 3(1 - s^2) \mathbb{1}_{[1,1]}$ ,  $K(1) > 0$ , and  $K_1(s) = 3(1 - s^2) \mathbb{1}_{[1,1]}$ . Elsewhere, as it is well known in FDA, the choice of the metric and the smoothing parameters have crucial roles in the computational issues. To optimize these choices on this illustration, we use the local cross-validation procedure method in the aim of choosing smoothing parameters  $h_K$  and  $h_H$  (see Laksaci et al. [17]).

Another important point for insuring good behavior of the considered methods is to use locating functions  $\delta$  and/or  $\beta$  that are well adapted to the kind of data that we have to deal with. Here, it is clear that the shape of the curves (cf. Fig. 1) allows us to use the locating functions  $\sigma$  and  $\beta$  defined by the derivatives of the curves. More precisely, we take:

$$\delta(x, x') = \left( \int_0^1 (x^{(i)}(t) - x'^{(i)}(t))^2 dt \right)^{1/2} \quad \text{and} \quad \beta(x, x') = \int_0^1 \alpha(t) (x_i(t) - x'_i(t))^2 dt,$$

where  $x^{(i)}$  denotes the  $i$ th derivative of the curve  $x$  and  $\alpha$  is the eigenfunction of the empirical covariance operator  $\frac{1}{n} \sum_{i=1}^n (X_j^{(i)} - \bar{X}^{(i)})(X_j^{(i)} - \bar{X}^{(i)})$  associated with the  $q$ -greatest eigenvalues. The performance of the conditional mode estimator  $\hat{\theta}_n(x)$  is evaluated on  $N = 400$  replications using different sample sizes  $n = 50, 100, 200$ , and  $300$ . The mean square error (MSE) is considered here, such that, for a fixed  $x$ ,





**Fig. 2** Distribution of the MSE obtained for different CR for  $n = 300$

**Table 1** Numerical summary of the distribution of the MSE, for  $N = 400$ , obtained for  $n = 50, 100, 200$  and  $300$

	CR = 2%	CR = 7%	CR = 16%	CR = 49%
$n = 50$	0.392	0.472	0.673	2.431
$n = 100$	0.301	0.332	0.472	1.964
$n = 200$	0.220	0.292	0.346	1.632
$n = 300$	0.102	0.150	0.215	1.214

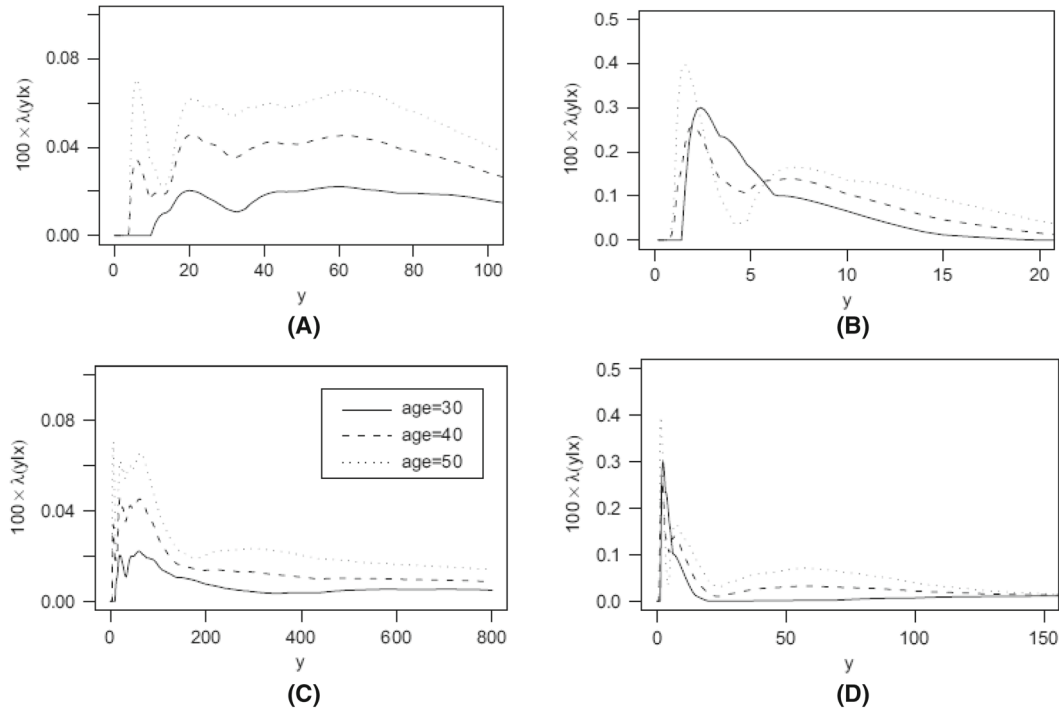
$MSE = \frac{1}{400} \sum_{m=1}^{400} (\hat{\theta}_{n,m}(x) - r(x))^2$ . Figure 2 displays the distribution of the obtained MSE given by the  $N$  replications. It can be observed that the proposed estimator performs well, especially when the sample size increases. This conclusion is confirmed by Table 1 which provides a numerical summary of the distribution of the MSE, with different censored rates (CR).

In the second part of the simulation studies, we are interested in the evaluation of the prediction accuracy of the conditional median with different censored rates (CR). A sample  $(X_i, Y_i)_{i=1, \dots, 550}$  of size  $n = 550$  generated from the model described above, is considered for this purpose. We split this sample into two parts: a learning subsample  $\{(X_i, Y_i); i = 1, \dots, 500\}$  which is used to calculate the predictor (the conditional mode in this case) and a testing subsample  $\{(X_i, Y_i); i = 501, \dots, 550\}$  used to evaluate the performance of the predictor. The prediction accuracy is measured, for different values of CR, using the Mean Absolute Error (MAE), defined as  $MAE = \frac{1}{50} \sum_{i=501}^{550} |Y_i - \hat{\theta}_n(X_i)|$ , as well as the Mean Square Error (MSE), such that  $MSE = \frac{1}{50} \sum_{i=501}^{550} (Y_i - \hat{\theta}_n(X_i))^2$ . We can see that the prediction accuracy of the conditional mode decreases as the censored rate increases. For censoring distributions  $\mathcal{E}(2) - 1$ , ( $CR = 7\%$ ,  $MAE = 0.204$ ,  $MSE = 0.483$ ),  $\mathcal{E}(2)$ , ( $CR = 16\%$ ,  $MAE = 0.363$ ,  $MSE = 0.62$ ), and  $\mathcal{E}(2) + 5$ , ( $CR = 49\%$ ,  $MAE = 1.052$ ,  $MSE = 2.291$ ).

#### 4.2 Real data application

A useful tool in survival analysis is the hazard rate, which reflects the instantaneous probability that a duration will end within the next time instant. Among the most used examples in survival analysis: survival times of patients, Stanford Heart Transplant, durations between subsequent transactions in a financial security, ... For application on real data, we apply the local linear method via the Kidney transplant data (see Klein and Moeschberger [16] and <https://www.agence-biomedecine.fr>). The bandwidth selection is given by plug-in rules and more advanced selection methods like cross-validation are likely to further improve the performance of the local linear hazard rate estimator. To use the Kidney transplant data, we need to describe three fundamental





**Fig. 3** Estimates of the conditional hazard rate after the transplantation

parameters that are (1) the survival times in days of patients following kidney transplant as the response, (2) race (black/white), and (3) age in years for each patient as the covariate. We propose a new method based on the functional local linear approach. More precisely, we use the conditional hazard rate. The methodology of this study is given by the following description: in the first, we consider the subsamples of white males and white females. At the second step, we take 432 patients as the first group and have a censoring rate of 83%. The second group involves 280 patients with a 86% censoring rate. The survival times of white patients vary between 1 day and 9.4 years. The average age of the male patients is slightly less than 44 years, whereas the female patients are almost 41 years old on average. We find only small differences between the various plug-in bandwidths for the conditional hazard rate, where we have taken age as the conditioning variable.

In conclusion, we find only small differences between the various plug-in bandwidths for the conditional hazard rate, where we have taken age as the conditioning variable (see Fig. 3). Figure (A) and Figure (D), based on the normal reference rule, depict the conditional hazard rate as a function of time and consider several values of age. Figure (A) and Figure (B) zoom in on the period shortly after the transplantation, whereas Figure (C) and Figure (D) display the hazard rate over a longer time. We see that the hazard rate is strongly non-monotonic for both male and female patients. As expected, the hazard rate is higher for older patients. Interestingly, the differences in the hazard rates of younger and older women diminish after about 150 days. For men, the differences between the various age groups persist longer. Also, the risk of dying shortly after the transplantation is higher for men than for women. On the other hand, for female patients, the risk of dying in a later stage is lower than for males.

### 5 Proofs

In what follows, when no confusion is possible, we will denote by  $C$  and  $C'$  some strictly positive generic constants. Moreover, we put, for any  $x \in \mathcal{F}$ , and for all  $i = 1, \dots, n$ :

$$K_i(x) = K(h^{-1}\delta(x, X_i)), \quad \beta_i(x) = \beta(X_i, x) \text{ and } H_i(y) = H(h_H^{-1}(y - Y_i)).$$

**Proposition 5.1** (Ferraty and Vieu [10], page 237)

Assume that  $\{U_i, i \geq 1\}$  are identically distributed, with strong mixing coefficient  $\alpha(n) = O(n^{-a})$ ,  $a > 1$ , such that  $|U_1|$  is bounded. Then, for each  $r > 1$  and  $\varepsilon > 0$ :

$$P \left\{ \left| \sum_{i=1}^n U_i \right| > \varepsilon \right\} \leq C \left( 1 + \frac{\varepsilon^2}{rS_n^2} \right)^{-r/2} + Cnr^{-1} \left( \frac{2r}{\varepsilon} \right)^{a+1},$$

where

$$S_n^2 = \sum_{1 \leq i, j \leq n} |\text{cov}(U_i, U_j)|.$$

*Proof of Lemma 3.4* The bias term is not affected by the dependence condition of  $(X_i, Y_i)$ . Therefore, by the equiprobability of the couples  $(X_i, Y_i)$ , we have:

$$E[\tilde{f}_n^x(y)] - \zeta(y|x) = \frac{1}{h_H E[W_{12}(x)]} E \left[ \delta_1 \bar{G}^{-1}(T_1) W_{12}(x) H_i(y) \right] - \zeta(y|x).$$

Using conditional expectation properties and the fact that:

$$\mathbb{1}_{\{Y_1 \leq C_1\}} \varphi(T_1) = \mathbb{1}_{\{Y_1 \leq C_1\}} \varphi(Y_1), \quad (5)$$

for any measurable function  $\varphi$ , we have:

$$E[\tilde{f}_n^x(y)] - \zeta(y|x) = \frac{1}{h_H E[W_{12}(x)]} E \left[ W_{12}(x) E[H_i(y) \mathbb{1}_{\{Y_1 \leq C_1\}} \bar{G}^{-1}(T_1) | X_1] \right] - \zeta(y|x).$$

Since

$$\begin{aligned} E[H_i(y) \mathbb{1}_{\{Y_1 \leq C_1\}} \bar{G}^{-1}(T_1) | X_1] &= E[E[H_i(y) \mathbb{1}_{\{Y_1 \leq C_1\}} \bar{G}^{-1}(Y_1) | Y_1] | X_1] \\ &= E[H_i(y) \bar{G}^{-1}(Y_1) E[\mathbb{1}_{\{Y_1 \leq C_1\}} | Y_1] | X_1] \\ &= E[H_i(y) | X_1], \end{aligned}$$

and

$$E[H_i(y) | X_1] = \int_{\mathcal{R}} H \left( \frac{y-z}{h_H} \right) \zeta(z|X) dz = h_H \int_{\mathcal{R}} H(t) \zeta(y - h_H t | X) dt,$$

then by assumptions (H2) and (H7), we get:

$$\begin{aligned} |E[\tilde{f}_n^x(y)] - \zeta(y|x)| &= \left| \frac{1}{h_H E[W_{12}(x)]} E \left[ W_{12}(x) h_H \int_{\mathcal{R}} H(t) \zeta(y - h_H t | X) dt \right] - \zeta(y|x) \right| \\ &= \left| \int_{\mathcal{R}} H(t) (\zeta(y - h_H t | X) - \zeta(y|x)) dt \right| \\ &\leq \int_{\mathcal{R}} H(t) |\zeta(y - h_H t | X) - \zeta(y|x)| dt \\ &\leq C \int_{\mathcal{R}} H(t) (h_K^{b_1} + |t|^{b_2} h_H^{b_2}) dt \\ &= O(h_K^{b_1}) + O(h_H^{b_2}), \end{aligned}$$

which proves Lemma 3.4.  $\square$



*Proof of Lemma 3.5* Let  $\Omega$  be a compact set. It can be covered by a finite number  $s_n$  of intervals of length  $l_n$  at some points  $(z_k)_{k=1, \dots, s_n}$ , that is:

$$\Omega \subset \bigcup_{k=1}^{s_n} (z_k - l_n, z_k + l_n),$$

with  $l_n = n^{-\frac{3\gamma}{2} - \frac{1}{2}}$  and  $s_n = O(l_n^{-1})$ . Let

$$z_y = \arg \min_{z \in \{z_1, \dots, z_{s_n}\}} |y - z|, \tag{6}$$

and consider the following decomposition:

$$\begin{aligned} \sup_{y \in \Omega} \left| \tilde{f}_n^x(y) - \mathbb{E}[\tilde{f}_n^x(y)] \right| &\leq \sup_{y \in \Omega} \left| \tilde{f}_n^x(y) - \tilde{f}_n^x(z_y) \right| + \sup_{y \in \Omega} \left| \tilde{f}_n^x(z_y) - \mathbb{E}[\tilde{f}_n^x(z_y)] \right| \\ &\quad + \sup_{y \in \Omega} \left| \mathbb{E}[\tilde{f}_n^x(z_y)] - \mathbb{E}[\tilde{f}_n^x(y)] \right| \\ &:= \mathfrak{S}_{1,n} + \mathfrak{S}_{2,n} + \mathfrak{S}_{3,n}. \end{aligned}$$

From (6) we have  $|y - z_y| \leq l_n$ , then under (H9) and Lemma 3.4, we get:

$$\mathfrak{S}_{1,n} = o_{a.co.} \left( \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right). \tag{7}$$

In the same way, we find:

$$\mathfrak{S}_{3,n} = o_{a.co.} \left( \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right). \tag{8}$$

Next, let us prove that:

$$\mathfrak{S}_{2,n} = o_{a.co.} \left( \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right).$$

For all  $\eta > 0$ , we have:

$$\begin{aligned} &\mathbb{P} \left( \sup_{y \in \Omega} \left| \tilde{f}_n^x(z_y) - \mathbb{E}[\tilde{f}_n^x(z_y)] \right| > \eta \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right) \\ &= \mathbb{P} \left( \max_{z_y \in \{z_1, \dots, z_{s_n}\}} \left| \tilde{f}_n^x(z_y) - \mathbb{E}[\tilde{f}_n^x(z_y)] \right| > \eta \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right) \\ &\leq s_n \mathbb{P} \left( \left| \tilde{f}_n^x(z_y) - \mathbb{E}[\tilde{f}_n^x(z_y)] \right| > \eta \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right). \end{aligned}$$

Then, we have to show that:

$$s_n \mathbb{P} \left( \left| \tilde{f}_n^x(z_y) - \mathbb{E}[\tilde{f}_n^x(z_y)] \right| > \eta \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right) < \infty, \text{ for all } z_y \in \{z_1, \dots, z_{s_n}\}.$$

For this, we consider the following decomposition:

$$\tilde{f}_n^x(z_y) = \underbrace{\frac{n^2 h_K^2 \phi_x^2(h_K)}{n(n-1)\mathbb{E}[W_{12}(x)]}}_{T_1} \left[ \underbrace{\left( \frac{1}{n} \sum_{j=1}^n \frac{\delta_j K_j(x) H_j(z_y)}{\overline{G}(T_j) h_H \phi_x(h_K)} \right)}_{T_2} \underbrace{\left( \frac{1}{n} \sum_{i=1}^n \frac{K_i(x) \beta_i^2(x)}{h_K^2 \phi_x(h_K)} \right)}_{T_3} \right. \\ \left. - \underbrace{\left( \frac{1}{n} \sum_{j=1}^n \frac{\delta_j K_j(x) \beta_j(x) H_j(z_y)}{\overline{G}(T_j) h_H h_K \phi_x(h_K)} \right)}_{T_4} \underbrace{\left( \frac{1}{n} \sum_{i=1}^n \frac{K_i(x) \beta_i(x)}{h_K \phi_x(h_K)} \right)}_{T_5} \right],$$

which implies that:

$$\tilde{f}_n^x(z_y) - \mathbb{E}[\tilde{f}_n^x(z_y)] = T_1 ((T_2 T_3 - \mathbb{E}[T_2 T_3]) - (T_4 T_5 - \mathbb{E}[T_4 T_5])),$$

where

$$T_2 T_3 - \mathbb{E}[T_2 T_3] = (T_2 - \mathbb{E}[T_2]) (T_3 - \mathbb{E}[T_3]) + (T_3 - \mathbb{E}[T_3]) \mathbb{E}[T_2] \\ + (T_2 - \mathbb{E}[T_2]) \mathbb{E}[T_3] + \mathbb{E}[T_2] \mathbb{E}[T_3] - \mathbb{E}[T_2 T_3],$$

and

$$T_4 T_5 - \mathbb{E}[T_4 T_5] = (T_4 - \mathbb{E}[T_4]) (T_5 - \mathbb{E}[T_5]) + (T_5 - \mathbb{E}[T_5]) \mathbb{E}[T_4] \\ + (T_4 - \mathbb{E}[T_4]) \mathbb{E}[T_5] + \mathbb{E}[T_4] \mathbb{E}[T_5] - \mathbb{E}[T_4 T_5].$$

Therefore, our claimed result is direct consequences of the following assertions:

$$s_n \mathbb{P} \left\{ |T_i - \mathbb{E}[T_i]| > \eta \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right\} < \infty, \quad \text{for } i = 2, 3, 4, 5, \quad (9)$$

$$T_1 = O(1) \quad \text{and} \quad \mathbb{E}[T_i] = O(1), \quad \text{for } i = 2, 3, 4, 5, \quad (10)$$

$$\text{Cov}(T_2, T_3) = o_{a.co.} \left( \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right), \quad (11)$$

and

$$\text{Cov}(T_4, T_5) = o_{a.co.} \left( \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right). \quad (12)$$

Concerning (9): Observe that for  $i = 3, 5$  has been already obtained in Lemma 3.3. Thus, we focus only on the case where  $i = 2, 4$ . For this, we use proposition 5.1.

For  $1 \leq i \leq n$  and  $k = 0, 1$ , let:

$$\nabla_i^k = \frac{1}{h_K^k} \left( \frac{\delta_i}{\overline{G}(T_i)} K_i(x) H_i(z_y) \beta_i^k(x) - \mathbb{E} \left[ \frac{\delta_i}{\overline{G}(T_i)} K_i(x) H_i(z_y) \beta_i^k(x) \right] \right). \quad (13)$$

Then, it can be seen that:

$$T_{2(k+1)} - \mathbb{E}[T_{2(k+1)}] = \frac{1}{n h_H \phi_x(h_K)} \sum_{i=1}^n \nabla_i^k, \quad \text{for } k = 0, 1.$$



Under assumptions (H1), (H3), (H5), and (H7), we have:

$$\begin{aligned} \frac{1}{h_K^k} \frac{\delta_i}{\overline{G}(T_i)} K_i(x) H_i(z_y) \beta_i^k(x) &\leq \frac{C}{h_K^k} \frac{1}{\overline{G}(\tau)} K_i(x) |\delta(X_i, x)|^k \\ &\leq \frac{C}{h_K^k} \frac{1}{\overline{G}(\tau)} K(h_K^{-1} \delta(x, X_i)) |\delta(X_i, x)|^k \mathbb{I}_{[-1,1]}(h_K^{-1} \delta(x, X_i)) \\ &\leq \frac{C}{\overline{G}(\tau)} K(h_K^{-1} \delta(x, X_i)) \mathbb{I}_{B(x, h_K)}(X_i) \\ &\leq C' \mathbb{I}_{B(x, h_K)}(X_i). \end{aligned} \tag{14}$$

Next, let us calculate:

$$S_n^2 = \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(\nabla_i^k, \nabla_j^k) = S_n^{2*} + n \text{Var}[\nabla_1^k],$$

where

$$S_n^{2*} = \sum_{i=1}^n \sum_{i \neq j}^n \text{Cov}(\nabla_i^k, \nabla_j^k).$$

Using Eqs. (5) and (14), and the conditional expectation properties, we have:

$$\begin{aligned} \text{Var}[\nabla_1^k] &= \text{Var} \left[ \frac{1}{h_K^k} \frac{\delta_i}{\overline{G}(T_i)} K_i(x) H_i(z_y) \beta_i^k(x) \right] \\ &\leq C' \mathbb{E}[\mathbb{I}_{B(x, h_K)}(X_i) \mathbb{E}[H_i^2(z_y) | X_1]] + C \mathbb{E}^2[\mathbb{I}_{B(x, h_K)}(X_i) \mathbb{E}[H_i(z_y) | X_1]] \\ &= C' \phi_x(h_K) \mathbb{E}[H_i^2(z_y) | X_1] + C \phi_x^2(h_K) \mathbb{E}^2[H_i(z_y) | X_1] \\ &\leq C' \chi_x^{1|2}(h_K) \mathbb{E}[H_i^2(z_y) | X_1] + C \chi_x(h_K) \mathbb{E}^2[H_i(z_y) | X_1]. \end{aligned}$$

Finally, using the fact that:

$$\mathbb{E}[H_i^2(z_y) | X_1] = O(h_H) \text{ and } \mathbb{E}[H_i(z_y) | X_1] = O(h_H),$$

we get:

$$n \text{Var}[\nabla_1^k] = O(n h_H \chi_x^{(1/2)}(h_K)).$$

Then, from Eq. (13), using again the conditional expectation, and from the fact that:

$$\mathbb{E}[\delta_i \delta_j | Y_i Y_j] = \overline{G}(T_i) \overline{G}(T_j),$$

we get under (H4) and (H5), for  $i \neq j$ :

$$\begin{aligned} \text{Cov}(\nabla_i^k, \nabla_j^k) &= \text{Cov} \left( \frac{1}{h_K^k} \frac{\delta_i}{\overline{G}(T_i)} K_i(x) H_i(z_y) \beta_i^k(x), \frac{1}{h_K^k} \frac{\delta_j}{\overline{G}(T_j)} K_j(x) H_j(z_y) \beta_j^k(x) \right) \\ &= \mathbb{E} \left[ \frac{1}{h_K^{2k}} \frac{\delta_i \delta_j}{\overline{G}(T_i) \overline{G}(T_j)} K_i(x) K_j(x) H_i(z_y) H_j(z_y) \beta_i^k(x) \beta_j^k(x) \right] \\ &\quad - \mathbb{E} \left[ \frac{1}{h_K^k} \frac{\delta_i}{\overline{G}(T_i)} K_i(x) H_i(z_y) \beta_i^k(x) \right] \mathbb{E} \left[ \frac{1}{h_K^k} \frac{\delta_j}{\overline{G}(T_j)} K_j(x) H_j(z_y) \beta_j^k(x) \right] \\ &\leq C [\mathbb{E}[\mathbb{I}_{B(x, h_K) \times B(x, h_K)}(X_i, X_j) \mathbb{E}[H_i(z_y) H_j(z_y) | X_i X_j]] \\ &\quad + \mathbb{E}[\mathbb{I}_{B(x, h_K)}(X_i) \mathbb{E}[H_i(z_y) | X_i]] \mathbb{E}[\mathbb{I}_{B(x, h_K)}(X_j) \mathbb{E}[H_j(z_y) | X_j]]] \\ &\leq C h_H^2 (\phi_x(h_K) + \phi_x^2(h_K)) \\ &\leq C h_H^2 \chi_x(h_K). \end{aligned} \tag{15}$$

Now, following Masry [18], we define:

$$S_1 = \{(i, j) \text{ such that } 1 \leq i - j \leq m_n\},$$

and

$$S_2 = \{(i, j) \text{ such that } m_n + 1 \leq i - j \leq n - 1\},$$

where  $m_n \rightarrow \infty$ , as  $n \rightarrow \infty$ .

Thus:

$$s_n^{2*} \leq \sum_{S_1} \text{Cov}(\nabla_i^k, \nabla_j^k) + \sum_{S_2} \text{Cov}(\nabla_i^k, \nabla_j^k) := \Upsilon_{1,n} + \Upsilon_{2,n}.$$

We then get from Eq. (15):

$$\Upsilon_{1,n} \leq Cnm_n h_H^2 \chi_x(h_K).$$

For  $\Upsilon_{2,n}$ , we use the Davydov–Rio’s inequality for bounded mixing processes:

$$\forall i \neq j \quad |\text{Cov}(\nabla_i^k, \nabla_j^k)| \leq C\alpha(|i - j|).$$

Therefore, using  $\sum_{j \geq x+1} j^{-a} \leq \int_{u \geq x} u^{-a} = [(a-1)x^{a-1}]^{-1}$ , and the first part of (H4), we get:

$$\Upsilon_{2,n} \leq Cnm_n^{-a+1}.$$

Thus:

$$s_n^{2*} \leq C(nm_n h_H^2 \chi_x(h_K) + nm_n^{1-a}).$$

By choosing  $m_n = (h_H^2 \chi_x(h_K))^{-1/a}$ , we obtain:

$$s_n^{2*} = O(nh_H^2 \chi_x(h_K))^{(a-1)/a}.$$

Finally, as  $a > 2$ , then:

$$S_n^2 = O(nh_H \chi_x^{1/2}(h_K)). \quad (16)$$

Taking  $\varepsilon = \eta \frac{\sqrt{S_n^2 \log n}}{nh_H \phi_x(h_K)}$ , making use of Proposition 5.1 and Eq. (16), we get:

$$\begin{aligned} \mathbb{P} \left\{ |T_{2(k+1)} - \mathbb{E}[T_{2(k+1)}]| > \varepsilon \right\} &= \mathbb{P} \left\{ \left| \sum_{i=1}^n \nabla_i^k \right| > \varepsilon nh_H \phi_x(h_K) \right\} \\ &\leq C \left( 1 + \frac{\varepsilon^2 n^2 h_H^2 (\phi_x(h_K))^2}{S_n^2 r} \right)^{-r/2} \\ &\quad + Cnr^{-1} \left( \frac{2r}{\varepsilon nh_H \phi_x(h_K)} \right)^{a+1} \\ &\leq C \exp \left( -r/2 \log \left( 1 + \frac{\eta^2 \log n}{r} \right) \right) \\ &\quad + Cnr^{-1} \left( \frac{r}{\eta} \right)^{a+1} (nh_H \chi_x^{1/2}(h_K) \log n)^{-(a+1)/2} \\ &\leq C(\wp_1 + \wp_2). \end{aligned} \quad (17)$$

Next, using Eq. (17) with  $r = C \log n (\log_2 n)^{\frac{1}{a}}$  and Taylor series expansion of  $\log(1+x)$ , we obtain:

$$\wp_1 \leq C \exp(-C\eta^2 \log n) = Cn^{-C\eta^2}, \quad (18)$$



and

$$\wp_2 \leq Cn^{(1-a)/2} \eta^{-(a+1)} (h_H \chi_x^{1/2}(h_K))^{-(a+1)/2} (\log n)^{(a-1)/2} \log_2 n. \tag{19}$$

Now, from Eqs. (18) and (19), using assumption (H10)(ii), we have  $s_n \wp_2 = O(n^{-1} \log^{-1} n \log_2^{-2} n)$  which is the general term of a convergent Bertrand series.

In the same way, we can choose  $\eta$ , such that  $s_n \wp_1$  is the general term of a convergent series.

Finally, we find:

$$s_n \mathbb{P} \left\{ \left| T_{2(k+1)} - \mathbb{E}[T_{2(k+1)}] \right| > \eta \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right\} < \infty.$$

Concerning (10): from Demongeot et al. [7], we have:

$$T_1 = O(1) \text{ and } \mathbb{E}[T_i] = O(1), \text{ for } i = 3, 5.$$

Now, we have to study the case where  $i = 2, 4$ .

Since the pairs  $(X_i, Y_i), i = 1, \dots, n$  are identically distributed, we obtain:

$$\mathbb{E}[T_2] = \frac{\mathbb{E}[\delta_1 \bar{G}^{-1}(T_1) K_1(x) H_1(z_y)]}{h_H \phi_x(h_K)},$$

and

$$\mathbb{E}[T_4] = \frac{\mathbb{E}[\delta_1 \bar{G}^{-1}(T_1) K_1(x) \beta_1(x) H_1(z_y)]}{h_K h_H \phi_x(h_K)}.$$

We have to evaluate:

$$\mathbb{E} \left[ \delta_1 \bar{G}^{-1}(T_1) K_i(x) H_i(z_y) \beta_i^l(x) \right], \text{ for } l = 0, 1.$$

Using the conditional expectation properties and Eq. (5), we have for all  $l = 0, 1$  :

$$\mathbb{E} \left[ \delta_1 \bar{G}^{-1}(T_1) K_i(x) H_i(z_y) \beta_i^l(x) \right] = O(h_H \mathbb{E} \left[ K_i(x) \beta_i^l(x) \right]).$$

Using Lemma 3 given in Barrientos-Marin et al. [2], we obtain

$$\mathbb{E} \left[ \delta_1 \bar{G}^{-1}(T_1) K_i(x) H_i(z_y) \beta_i^l(x) \right] = O(h_H h_K^l \phi_x(h_K)). \tag{20}$$

From Eq. (20), we find:

$$\mathbb{E}[T_i] = O(1), \text{ for } i = 2, 4.$$

Concerning (11) and (12): following similar steps as in the proof of (16), we get:

$$\begin{aligned} \text{Cov}(T_2, T_3) &= O \left( \frac{\chi_x^{(1/2)}(h_K)}{n \phi_x^2(h_K)} \right) = o \left( \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right) \\ \text{and } \text{Cov}(T_4, T_5) &= O \left( \frac{\chi_x^{(1/2)}(h_K)}{n \phi_x^2(h_K)} \right) = o \left( \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right). \end{aligned}$$

Hence:

$$\sup_{y \in \Omega} \left| \tilde{f}_n^x(z_y) - \mathbb{E}[\tilde{f}_n^x(z_y)] \right| = o_{a.co.} \left( \sqrt{\frac{\chi_x^{(1/2)}(h_K) \log n}{n h_H \phi_x^2(h_K)}} \right). \tag{21}$$

Finally, proof of Lemma 3.5 can be achieved by considering Eqs. (7), (8), and (21). □



*Proof of Lemma 3.6* Observe that:

$$\begin{aligned} \left| \widehat{f}_n^x(y) - \widetilde{f}_n^x(y) \right| &\leq \frac{1}{n(n-1)h_H E[W_{12}(x)]} \sum_{i \neq j} \left| \delta_i W_{ij}(x) H_i(y) \left( \frac{1}{\overline{G}_n(T_i)} - \frac{1}{\overline{G}(T_i)} \right) \right| \\ &\leq \frac{\sup_{y \in \Omega} |\overline{G}_n(t) - \overline{G}(t)|}{\overline{G}_n(\tau)} \frac{1}{n(n-1)h_H E[W_{12}(x)]} \sum_{i \neq j} \delta_i \overline{G}^{-1}(T_i) W_{ij}(x) H_i(y) \\ &= \frac{\sup_{y \in \Omega} |\overline{G}_n(t) - \overline{G}(t)|}{\overline{G}_n(\tau)} \widetilde{f}_n^x(y). \end{aligned}$$

Since  $\overline{G}(\tau) > 0$  in conjunction with the strong law of large numbers and the law of the iterated logarithm on the censoring law (see formula (4.28) in Deheuvels and Einmahl [6]), the result is an immediate consequence of Lemmas 3.4 and 3.5  $\square$

*Proof of Theorem 3.7* One can easily show that:

$$|\zeta(\widehat{\theta}_n(x)|x) - \zeta(\theta(x)|x)| \leq 2 \sup_{y \in \Omega} |\widehat{\zeta}_n(y|x) - \zeta(y|x)|. \quad (22)$$

Using Taylor's expansion  $\zeta(\cdot|x)$ , we get:

$$\zeta(\widehat{\theta}_n(x)|x) - \zeta(\theta(x)|x) = \frac{1}{2} (\widehat{\theta}_n(x) - \theta(x))^2 \zeta^{(2)}(\theta^*(x)|x), \quad (23)$$

where  $\theta^*(x)$  is between  $\theta(x)$  and  $\widehat{\theta}_n(x)$ .

Combining Eqs. (22) and (23), we obtain:

$$|\widehat{\theta}_n(x) - \theta(x)| \leq 2 \sqrt{\frac{\sup_{y \in \Omega} |\widehat{\zeta}_n(y|x) - \zeta(y|x)|}{\zeta^{(2)}(\theta^*(x)|x)}}.$$

The almost sure consistency of  $\widehat{\theta}_n(x)$  follows, then, immediately from Theorem 3.2.  $\square$

**Acknowledgements** The authors would like to thank the Editor, the Associate Editor, and anonymous reviewers for their valuable comments and suggestions, which improved the quality of this paper. A special thanks to Pr. Hicham TAHRAOUI (Faculty of Medicine, Tlemcen, Algeria).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Ait-Sahalia, Y.: Transition densities for interest rate and other non-linear diffusions. *J. Finance* **54**, 1361–1395 (1999)
2. Barrientos-Marin, J.; Ferraty, F.; Vieu, P.: Locally modelled regression and functional data. *J. Nonparametr. Stat.* **22**(3), 617–632 (2009)
3. Bashtannyk, D.M.; Hyndman, R.J.: Bandwidth selection for kernel conditional density estimation. *Comput. Stat. Data Anal.* **36**, 279–298 (2001)
4. Baillo, A.; Grané, A.: Local linear regression for functional predictor and scalar response. *J. Multivar. Anal.* **100**, 102–111 (2009)
5. Beran, R.: Nonparametric regression with randomly censored survival data, Technical Report, Department of Statistics, University of California, Berkeley, CA (1981)
6. Deheuvels, P.; Einmahl, H.: Functional limit laws for the increments of Kaplan–Meier product limit processes and applications. *Ann. Probab.* **28**, 1301–1335 (2000)
7. Demongeot, J., Laksaci, A., Madani, F., Rachdi, M.: A fast functional locally modelled of the conditional density and mode for functional time series. In: Ferraty, F. (eds.) *Recent Advances in Functional Data Analysis and Related Topics. Contributions to Statistics*. Physica-Verlag HD (2011)



8. Doukhan, P.: *Mixing Properties and Examples*. Lecture Notes in Statistics. Springer, New York (1994)
9. Fan, J.; Yim, T.H.: A cross validation method for estimating conditional densities. *J. Biom.* **91**, 819–834 (2004)
10. Ferraty, F.; Vieu, P.: *Nonparametric Functional Data Analysis. Theory and Practice*. Springer Series in Statistics. Springer, New York (2006)
11. Hall, P.; Racine, J.; Li, Q.: Cross-validation and the estimation of conditional probability densities. *J. Am. Stat. Assoc.* **99**, 1015–1026 (2004)
12. Horrigue, W.; Ould-Saïd, E.: Strong uniform consistency of a nonparametric estimator of a conditional quantile for censored dependent data and functional regressors. *Random Oper. Stoch. Equ.* **19**, 131–156 (2011)
13. Hyndman, R.J.; Bashtannyk, D.M.; Grunwald, G.K.: Estimating and visualizing conditional densities. *J. Comput. Graph. Stat.* **5**, 315–336 (1996)
14. Kaplan, E.L.; Meier, P.: Nonparametric estimation from incomplete observations. *JASA* **53**, 457–481 (1958)
15. Khardani, S.; Thiam, B.: Strong consistency result of a nonparametric conditional mode estimator under random censorship for functional regressors. *Commun. Stat. Theory Methods* **45**, 1863–1875 (2016)
16. Klein, J.P.; Moeschberger, M.L.: *Survival Analysis: Techniques for Censored and Truncated Data*. Springer, Berlin (2004)
17. Laksaci, A.; Madani, F.; Rachdi, M.: Kernel conditional density estimation when the regressor is valued in a semi-metric space. *Commun. Stat. Theory Methods* **42**(19), 3544–3570 (2013)
18. Masry, E.: Recursive probability density estimation for weakly dependent process. *IEEE Trans. Inf. Theory* **32**, 254–267 (1986)
19. Messaci, F.; Nemouchi, N.; Ouassou, I.; Rachdi, M.: Local polynomial modelling of the conditional quantile for functional data. *Stat. Methods Appl.* **24**(4), 597–622 (2015)
20. Robinson, P.M.: Consistent nonparametric entropy-based testing. *Rev. Econ. Stud.* **58**, 437–453 (1991)
21. Sarda, P.; Vieu, P.: *Kernel Regression*, pp. 43–70. Wiley, New York (2000)
22. Stute, W.: Distributional convergence under random censorship when covariates are present. *Scand. J. Stat.* **23**, 461–471 (1996)
23. Tjøstheim, D.: Non-linear time series. *J. Stat.* **21**, 97–130 (1994)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

