Afrika Statistika

Afrika Statistika
Vol. 19 (1), 2024, pages 3771 - 3796.
DOI: https://dx.doi.org/10.16929/as/2024.3771.327

ISSN 2316-090X

Kernel estimation of the conditional probability density function and Mode in presence of functional covariate and censorship

Hézouwè Dabana [1,*], **Komi Agbokou** [2] **and Kossi Essona Gneyou** [1]

[1] Department of Mathematics, Faculty of sciences, University of Lomé, BP 1515, Togo
[2] Department of Mathematics, Faculty of sciences and Technology, University of Kara, BP 239, Kara - Togo

Received on August 25, 2024; Accepted on October 28; Published online on December 10, 2024

Abstract. In the literature, several authors have worked on the nonparametric estimation of conditional density when the data is censored. In this work, we tackle the same theme with the kernel method by considering the case of censored and dependent data of the algebraically α-mixing type in the presence of functional covariate. We establish the almost complete convergence of this estimator and that of mode. Finally, we carry out a numerical study to evaluate the performance of these estimators.

Key words: nonparametric estimation; conditional density; censored data; small ball probabilities; functional random variable.
AMS 2020 Mathematics Subject Classification Objects : 62G05; 62N01; 62R10.

[*] Hézouwè Dabana: hezouwedabana8@gmail.com
Komi Agbokou : ffomestein@gmail.com
Kossi Essona Gneyou : kgneyou@univ-lome.tg

Résumé. (French Abstract) Dans la littérature plusieurs auteurs ont travaillé sur l'estimation non paramétrique de la densité conditionnelle lorsque les données sont censurées. Dans ce travail, nous abordons la même thématique avec la méthode de noyau en considérant le cas de données censurées et dépendantes du type algébriquement α-mélangeante en présence de covariable fonctionnelle. Nous établissons la convergence presque complète de cet estimateur et de celle du mode. Enfin, nous réalisons une étude numérique pour évaluer la performance de ces estimateurs.

Presentation of authors.

Hézouwè Dabana, M.Sc, is preparing a Ph.D. thesis in Statistics under the supervision of the third author at: Université de Lomé, TOGO.

Komi Agbokou, Ph.D., is an Associated Professor at: Université de Kara, TOGO.

Kossi Essona Gneyou, Ph.D., is a full Professor of Mathematics and Statistics at: Université de Lomé, TOGO.

Introduction

Conditional density estimation is an important task in statistics and machine learning. However, this task can be difficult, especially when the data is censored, *i.e.* only certain observations are available. Nonparametric methods are often used to solve this problem, as they do not require assumptions about the underlying distribution and are therefore more flexible.

In this context, several authors have been interested in the estimation of conditional density through several methods. For the kernel method, we can cite Liang and Peng (2010); Khardani and Semmar (2014); Xianzhou and Meijuan Ou. (2019);Slaoui and Khardani (2020).
Liang and Peng (2010) derived the asymptotic normality and Berry-Esseen type bound for kernel estimator with stationary α-mixing observations. Khardani and Semmar (2014) determined the strong and uniform consistency and the asymptotic normality of the Nadaraya Watson estimator in the multiviated case; Xianzhou and Meijuan Ou. (2019) have established the asymptotic normality of this estimator in the α-mixing case. Slaoui and Khardani (2020) extended the work of Khardani and Semmar (2014) on the recursive version of the estimator of Nadaraya-Watson.

Kernel estimators of conditional density and conditional survival function were used by several authors to estimate the conditional hazard (direct method). It was a question of studying the quotient between the kernel estimator of the conditional density and the kernel estimator of the conditional survival function. Asymptotic properties of the kernel estimator of the conditional hazard function are therefore obtained from those conditional survival function and conditional density. Note that in the case of censored data, the distribution function of the censorship variable is not taken into account in this estimate which puts the difference between the estimator of conditional density and that used for the estimation of the conditional hazard function (see for example Ferraty et al. (2008) and Agbokou and Gneyou (2017)).

The local linear estimator of conditional density has been widely studied, when the covariate is in a finite dimensional space. The first contributions on the local polynomial modelling of the conditional density function when the covariate is functional were provided by Demongeot et al. (2010). These authors have established the almost complete convergence of the estimator for *i.i.d.* Their study is extended to the dependent case by Demongeot et al. (2011). The rate of convergence in quadratic average of the estimator considered by Demongeot et al. (2010) shows the effectiveness of the local linear method compared to that of the kernel, especially in terms of bias. In addition, there are some articles on this subject in the case of censored data such as Spierdijk (2008), Kim et al. (2010), Xianzhou and Meijuan Ou. (2019) and Benkhaled et al. (2020). Spierdijk (2008) studied a version of the local linear estimator of the conditional density function with stationary observations and ρ-mixing, while Kim et al. (2010) proposed another version of linear local estimators with different weights in the independent and identically distributed case *(i.i.d)*. Xianzhou and Meijuan Ou. (2019) studied the local constant and the local linear estimators of the conditional density function with right-censored and α-mixing data. They established the asymptotic normality of the two estimators. Benkhaled et al. (2020) extended the work of Demongeot et al. (2011) in the event of censored data. They established the almost complete convergence of the local linear estimator. It should be noted that the kernel estimator proposed by Liang and Peng (2010) is a single-smoothing estimator of the conditional density function, while the local linear estimator proposed by Kim et al. (2010) is a double smoothing estimator.

The nonparametric estimation of the mode has also been the subject of attention, (see *e.g.*, Dabo-Niang and Laksaci (2007) or for example in Gneyou (2013), Gneyou (2014); Agbokou and Gneyou (2017), Agbokou and Gneyou (2019)).

In this work, we study the kernel estimator of conditional density and mode when the covariate is functional and the data is randomly right-censored and algebraically α-mixing. To do this, we will present the estimators in section 1; Section 2 will relate to the assumptions which have made it possible to obtain the results; Section 3 will focus on the main results in particular, the almost complete pointwise convergence of the conditional density estimator and the almost sure uniform convergence of the mode estimator. We then assess the performance of these estimators on simulated data games in section 3 and section 5 will be devoted to the proof of the lemmas stated in section 2. This work will be closed by a brief conclusion.

1. Construction of estimators

Let $(Y_i, X_i)_{1\le i\le n}$ be a sample of random variables having the same distribution as a pair (Y, X) such that $Y \in \mathbb{R}$ and X is valued in semi-metric space $(\mathcal{F}, d(;))$.

In the random right-censorship model, the random variables observed are not the pairs $(Y_i, X_i)_{1\le i\le n}$ but only the triples $(T_i, \delta_i, X_i)_{1\le i\le n}$ where $T_i = \min(Y_i, C_i)$ with C_i the positive variable called the censorship variable and δ_i is an indicator variable that is 1 if Y_i is directly observed and 0 if Y_i is censored on the right, *i.e.* $\delta_i = \mathbb{I}_{\{Y_i\le C_i\}}$ ($\mathbb{I}$ is the indicator function). Let G be the distribution function of the censorship variable C; F and H, the distribution functions of the variables Y and T respectively such as:

$$H(t) = 1 - P(T > t) = 1 - P(Y \wedge C > t) = 1 - P(Y > t)P(C > t) = 1 - F(t)G(t); \quad t \in \mathbb{R}.$$

By Ferraty et al. (2008), Khardani et al. (2010), Xianzhou and Meijuan Ou. (2019) and Kebir and Mechab (2023), the conditional density estimator adapted to the functional covariate was defined by:

$$\tilde{f}_n(y|x) = \frac{\sum_{i=1}^{n} \delta_i \bar{G}^{-1}(T_i) K\Big(h_K^{-1} d(x, X_i)\Big) H'\Big(h_H^{-1}(y - T_i)\Big)}{h_H \sum_{i=1}^{n} K\Big(h_K^{-1} d(x, X_i)\Big)}, \quad y \in \mathbb{R}. \tag{1}$$

where K and H' are kernel (H' is the derivative of H. see Agbokou and Gneyou (2017) for more details on K and H'); $h_K = h_{K,n}$ (*resp.* $h_H = h_{H,n}$) is a sequence of positive real numbers, and d is a semi-metric.

In this context of censorship, it should be recalled that Kaplan-Meier (1958) proposed an estimator of the survival function $S = 1 - F$ that reduces to the complement to 1 of the empirical distribution function when the observations are complete. Thus, the Kaplan-Meier estimator for the survival of the censorship variable was defined in the same way by:

$$\bar{G}_n(t) = \begin{cases} \prod_{i=1}^{n} \Big(1 - \frac{1 - \delta_{(i)}}{n - i + 1}\Big)^{\mathbb{I}_{\{T_{(i)}\le t\}}} & if \quad t < T_{(n)} \\ 0 & otherwise, \end{cases} \tag{2}$$

where $T_{(1)} < T_{(2)} < \cdots < T_{(n)}$ are order statistics associated with T_i. $\bar{G}_n$ converges uniformly to $\bar{G}$.

Since $\bar{G}$ does not exist in practice, by replacing it with $\bar{G}_n$ in the equality (1) we get:

$$\hat{f}_n(y|x) = \frac{\sum_{i=1}^{n} \delta_i \bar{G}_n^{-1}(T_i) K\left(h_K^{-1} d(x, X_i)\right) H'\left(h_H^{-1}(y - T_i)\right)}{h_H \sum_{i=1}^{n} K\left(h_K^{-1} d(x, X_i)\right)}, \quad y \in \mathbb{R}. \tag{3}$$

It should be noted that this estimator was introduced by Khardani et al. (2010) for the real covariate and used by Xianzhou and Meijuan Ou. (2019) to establish asymptotic normality (his Theorem 3.1) and recently, Kebir and Mechab (2023) use it to establish the almost sure pointwise convergence of the conditional hazard function estimator with the recursive kernel method in the presence of a functional covariate.

We study the same estimator by considering the functional covariate on the algebraically α-mixing data.

Let us denote by $\theta(x) = \arg\max_{y \in \mathcal{S}} f(y|x)$, the mode of the conditional density function $f(y|x)$. That's $\mathcal{S}$ a $\ll$ compact $\gg$ subset of $\mathbb{R}$. The estimator of mode $\theta(x)$ noted $\hat{\theta}_n(x)$ is a random variable such that (see Agbokou and Gneyou (2017), Agbokou and Gneyou (2019), Agbokou and Gneyou (2019)):

$$\hat{\theta}_n(x) = \arg\max_{y \in \mathcal{S}} \hat{f}_n(y|x) \tag{4}$$

with $\hat{f}_n(y|x)$, the estimator (3).

Let $\mathcal{F}_i^k(Z)$ be a σ-algebra generated by $\{Z_j, i \leq j \leq k\}$.

Definition 1 (Rosenblatt (1956) see Benkhaled et al. (2020)). Let $\{Z_i, i = 1, 2, \cdots\}$ be a strictly stationary sequence of random variables. Given a positive integer n, let us:

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

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

2. Assumptions

In this section, we present the assumptions that made it possible to obtain the almost complete pointwise convergence of the estimators (3) and the almost sure convergence of the estimator (4). Some of these hypotheses come from Ferraty et al. (2008); Benkhaled et al. (2020) and the others come from our work.

Throughout this note, x denotes a fixed point in $\mathcal{F}$; N_x denotes a fixed neighborhood of x. For any distribution function L, let $\tau_L := \sup\{y : L(y) < 1\}$ be its support's right endpoint and assume that $\theta(x) \in \mathcal{S} \subset]-\infty, \tau]$, where $\tau < \tau_G \wedge \tau_F$, $B(x,r) = \{x \in \mathcal{F} : |d(x',x)| \leq r\}$

(H1) $\forall h_K > 0, \phi_x(h_K) = \mathbb{P}\big(X \in B(x,h)\big) > 0;$

(H2) $\exists A_x < \infty,\ \exists b_1, b_2 > 0,\ \forall (y_1,y_2) \in \mathcal{S}^2,\ \forall (x_1,x_2) \in N_x^2 :$
$|f(y_1|x_1) - f(y_2|x_2)| \leq A_x\Big(d(x_1,x_2)^{b_1} + |y_1 - y_2|^{b_2}\Big);$

(H3) $\exists \nu < \infty,\ \forall (y,x) \in \mathcal{S} \times N_x,\ f(y|x) \leq \nu.$

(H4a) The sequence $(Y_i, C_i, X_i)_{i\in\mathbb{N}}$ is α-mixing and its mixing coefficients $\alpha(n)$ are such that $\exists a \in \mathbb{R}_+^*,\ \exists c \in \mathbb{R}_+^* :\ \forall n \in \mathbb{N},\ \alpha(n) \leq cn^{-a};$

(H4b) The joint density of (Y_i, Y_j) knowing (X_i, X_j) exists and is bounded, and $\exists \epsilon_1 \in]0,1[,\ 0 < \sup_{i\neq j} P\Big((X_i,X_j) \in B(x,h_K) \times B(x,h_K)\Big) = o\Big(\phi_x(h_K)^{1+\epsilon_1}\Big).$

(H4c) $\exists \epsilon_2 \in]0,1[,\ a > \dfrac{1+\epsilon_1}{\epsilon_1\epsilon_2}$ et $h_H\phi_x(h_K) = o\big(n^{-\epsilon_2}\big)$

(H5) H is differentiable such that
i) $\exists A < \infty,\ \forall (x_1,x_2) \in \mathbb{R}^2,\ |H'(x_1) - H'(x_2)| \leq A|x_1 - x_2|;$
ii) H' is bounded to compact support $[-1,1]$ and $H'(y) > 0,\ \forall y \in [-1,1].$
iii) H' is Lipschitzian such that $\int |t|^{b_1} H'(t)dt < \infty$ and $\int H'^2(t)dt < \infty;$

(H6) The functional kernel K satisfies the following conditions:
i) K is bounded to compact support (0,1);
ii) $\exists A_1, A_2,\ \forall x \in (0,1), 0 < A_1 < K(x) < A_2 < \infty;$

(H7) The bandwidth h_K satisfies the following conditions:

$$\lim_{n\to\infty} h_K = 0 \ \ et \ \ \lim_{n\to\infty} \frac{\log n}{nh_H\phi_x(h_K)} = 0;$$

(H8) The bandwidth h_H satisfies the following conditions:

$$\lim_{n\to\infty} h_H = 0 \ \ et \ \ \exists a > 0,\ \ \lim_{n\to\infty} n^a h_H = \infty;$$

(H9) Conditional on X, the variables Y and C are independent.

(H10) Given $x \in N_x$, we assume that the conditional density function $f(y|x)$ has a unique mode in interval $[a_x, b_x] \subset \mathcal{S}$ and $f(y|x) < f(\theta(x)|x)\ \forall y \neq \theta(x),\ y \in \mathcal{S}.$

Remark 1 (Comments on assumptions). Assumptions (H1) is a regularity condition that characterizes the topological structure of the functional space of the covariate; assumptions (H2) and (H3) are regularity conditions relating to the conditional distribution of Y given X. These conditions are not specific and can be found in the literature (see *e.g.* Ferraty et al. (2008)). The assumptions (H4)-(H10) are standard technical assumptions in our context and necessary for the study of the convergence of the two estimators. Indeed, the assumptions (H4a) and (H4b) are technical assumptions similar to (H4) of Benkhaled et al. (2020) as introduced in Ferraty and Vieu (2006) and allow to control the attached laws. (H4c) allows you to control the links between the smoothing parameters h_H, h_K and the mixing coefficient $\alpha(n)$. Assumptions (H5) and (H6) on the kernels K and H' are necessary to evaluate the bias term of the convergence rate of estimator (3) and (H7) and (H8) on the smoothing parameters h_H and h_K are necessary to simplify the calculations of the sum of the covariances s_n^2 and S_n^2. These assumptions are standard conditions to ensure the correct behaviour of the estimator and are similar to those used in Ferraty et al. (2008).

3. Main results

3.1. The almost complete pointwise convergence of the conditional density estimator

Theorem 1. *Under assumptions (H1)-(H9), we have:*

$$\sup_{y\in\mathcal{S}} |\hat{f}_n(y|x) - f(y|x)| = O\big(h_K^{b_1}\big) + O\big(h_H^{b_2}\big) + O\left(\sqrt{\frac{\log n}{nh_H\phi_x(h_K)}}\right), \ \textit{a.co} \tag{5}$$

Proof of theorem 1. Let us note: $H_i(y) = H'\big(h_H^{-1}(y-Y_i)\big)$;
$K_i(x) = K\big(h_K^{-1}d(x,X_i)\big)$;

$$\hat{f}_n(y|x) = \frac{\hat{f}_N(x,y)}{f_D(x)} \quad \text{with} \quad \begin{cases} \hat{f}_N(x,y) = \dfrac{1}{nh_H E[K_1(x)]}\displaystyle\sum_{i=1}^n H_i(y)\delta_i\bar{G}_n^{-1}(t)K_i(x) \\ f_D(x) = \dfrac{1}{nE[K_1(x)]}\displaystyle\sum_{i=1}^n K_i(x) \end{cases}$$

$$\tilde{f}_N(x,y) = \frac{1}{nh_H E[K_1(x)]}\sum_{i=1}^n H_i(y)\delta_i\bar{G}^{-1}(t)K_i(x)$$

The proof is based on the decomposition on the one hand:

$$\begin{aligned}\Big|\hat{f}_n(y|x) - f(y|x)\Big| \ \le\ & \frac{1}{f_D(x)}\Bigg[\Big|\hat{f}_N(x,y) - \tilde{f}_N(x,y)\Big| + \Big|\tilde{f}_N(x,y) - E[\tilde{f}_N(x,y)]\Big| \\ & + \Big|E[\tilde{f}_N(x,y)] - f(y|x)\Big| + f(y|x)(1-f_D(x))\Bigg]\end{aligned} \tag{6}$$

and on the other hand, the following lemmas:

Lemma 1. *Under assumptions (H1), (H6) and (H7) we have:*

$$\sup_{y\in\mathcal{S}}\left|\hat{f}_N(x,y)-\tilde{f}_N(x,y)\right| = O\left(\sqrt{\frac{\log(\log n)}{n}}\right) \text{ a.s,}$$

Lemma 2. *Under assumptions of the theorem 1 we have:*

$$\sup_{y\in\mathcal{S}}\left|\tilde{f}_N(x,y)-E[\tilde{f}_N(x,y)]\right| = O\left(\sqrt{\frac{\log n}{nh_H\phi_x(h_K)}}\right) \text{ a.co}$$

Lemma 3. *Under assumptions (H1), (H2), (H4a), (H5) et (H6) we have:*

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

Lemma 4. *Under assumptions (H1), (H4a), (H6) et (H7) we have:*

$$\left|1-\hat{f}_D(x)\right| = O\left(\sqrt{\frac{\log n}{n\phi_x(h_K)}}\right) \text{ a.co}$$

■

3.2. The almost sure convergence of the mode estimator

Theorem 2. *Under assumptions of the theorem 1 and (H10), we have:*

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

Proof of theorem 2. Following the same approach as the proof of theorem 2 by Agbokou and Gneyou (2017) and that of theorem 3.7 by Benkhaled et al. (2020), we will have:

$$\begin{aligned}\left|f\left(\hat{\theta}_n(x)|x\right)-f\left(\theta(x)|x\right)\right| &\leq \left|f\left(\hat{\theta}_n(x)|x\right)-\hat{f}_n\left(\hat{\theta}_n(x)|x\right)\right|+\left|\hat{f}_n\left(\hat{\theta}_n(x)|x\right)-f\left(\theta(x)|x\right)\right|\\ &\leq \sup_{y\in\mathcal{S}}\left|\hat{f}_n(y|x)-f(y|x)\right|+\sup_{y\in\mathcal{S}}\left|\hat{f}_n(y|x)-f(y|x)\right|\\ &\leq 2\sup_{y\in\mathcal{S}}\left|\hat{f}_n(y|x)-f(y|x)\right|, \qquad (7)\end{aligned}$$

Using a Taylor expansion of the function $f(.|x)$ we obtain:

$$f\big(\hat{\theta}_n(x)|x\big) - f\big(\theta(x)|x\big) = \frac{1}{2}\big(\hat{\theta}_n(x) - \theta(x)\big)^2 f^{(2)}\big(\theta^*(x)|x\big) \tag{8}$$

where $\theta^*(x)$ is between $\theta(x)$ and $\hat{\theta}_n(x)$.
By combining (7) and (8) we find:

$$\left|\hat{\theta}_n(x) - \theta(x)\right| \leq 2\sqrt{\frac{\sup_{y\in\mathcal{S}}\left|\hat{f}_n(y|x) - f(y|x)\right|}{f^{(2)}\big(\theta^*(x)|x\big)}}.$$

Thus, with the theorem 1 the proof of the theorem 2 is completed. ■.

4. Numerical study

4.1. Conditional density estimator

In this section, we want to check the performance of the estimator [3] when the covariate $x \in N_x$ and the data is dependent with the following expression:

$$\hat{f}_n(y|x) = \frac{\sum_{i=1}^{n} \delta_i \bar{G}_n^{-1}(Y_i) K\Big(h_K^{-1} d(x, X_i)\Big) H'\Big(h_H^{-1}(y - Y_i)\Big)}{h_H \sum_{i=1}^{n} K\Big(h_K^{-1} d(x, X_i)\Big)}, \quad y \in \mathbb{R}.$$

where K and H' are kernels such that $K(x) = \frac{3}{2}(1 - x^2)\mathbb{I}_{[0;1]}(x)$ and $H(x) = \int_{-\infty}^{x} \frac{3}{4}(1 - z^2)\mathbb{I}_{[-1;1]}(z)dz$. $h_K = h_{K,n}$ (resp. $h_H = h_{H,n}$) is a sequence of positive real numbers. h_H and h_K are obtained by the cross-validation method (see Laksaci et al. (2013)).
Semi-metric d is based on principal component analysis (see "semi-metric.pca" in Ferraty and Vieu (2006)). The advantage of this type of semi-metric is that it can be used even if the curves are irregular.

To do this, we set n=600 and generate the functional covariate assumed to be a diffusion process on [0,1] generated by the following equation (see Benkhaled et al. (2020)):

$$X(t) = A\big(2 - \cos(\pi t W)\big) + (1 - A)\cos(\pi t W) \quad t \in [0, 1], \tag{9}$$

where $W \rightsquigarrow \mathcal{N}(0, 1)$ and A is a Bernoulli random variable with parameter $p = 0,5$.
The $X_i \;\; 1 \leq i \leq n$ will be generated and plotted (Fig. 1).
In order to control the effect of censorship on the efficiency of the estimator, we

simulate the censure variables $(C_i)_{1\leq i\leq n}$ through the exponential law $\mathcal{E}(\lambda)$ adapted in order to obtain different censorship rates.
The real response variable is defined by (see Benkhaled et al. (2020)):

$$Y_i = r(X_i) + \varepsilon_i, \quad i = 1, \cdots, n, \tag{10}$$

where ε_i is the error generated by the auto-regressive model defined by:

$$\varepsilon_i = \frac{1}{\sqrt{2}}(\varepsilon_{i-1} + \eta_i) i = 1, \cdots, n \tag{11}$$

with $(\eta_i)_i$, an i.i.d. sequence of random variables following the $\mathcal{N}(0, 0.1)$ distribution and

$$r(X) = 4\exp\left\{\frac{1}{2 + \int_0^{\pi/2} |X_i(t)|^2 dt}\right\}. \tag{12}$$

With this model, the conditional density of Y given $X = x$ can be defined by (see Benkhaled et al. (2020)):

$$f(y|x) = \frac{1}{\sqrt{2\pi \times 0.3}} \exp\left\{-\frac{1}{2\times 0.3}(y - r(x))^2\right\}. \tag{13}$$

The performance of the estimator (3) by following these steps:

•**Step 1**: We associate the statistics of order $T_{(i)},\ \delta_{(i)},\ X_{(i)}$ with the variables T_i, δ_i, X_i respectively defined as follows:

$$\begin{aligned} T_{(1)} &< T_{(2)} < \cdots < T_{(n)} \\ \delta_{(i)} &= \{\delta_j : T_j = T_{(j)}\} \\ X_{(i)} &= \{X_j : T_j = T_{(j)}\} \end{aligned}$$

to obtain the sample $(T_{(i)},\ \delta_{(i)},\ X_{(i)})_{1\leq i\leq 600}$ (N=600).

•**Step 2**: To observe the effect of the censorship rate on the speed of convergence of the estimator, we consider different censorship rates (2%, 5%, 10% et 15%) and for each rate, we compare the *MSEs* with the increase in sample size by considering the samples $(T_{(i)},\ \delta_{(i)},\ X_{(i)})$ $151 \leq i \leq 250$, $151 \leq i \leq 350$ and $151 \leq i \leq 450$ (n=100, 200 and 300 respectively). The root mean square error of the estimator is calculated on a 160-point grid using the following formula:

$$MSE = \frac{1}{160}\sum_{j=1}^{160}\left(\hat{f}_{nj}(y|x) - f_j(y|x)\right)^2.$$

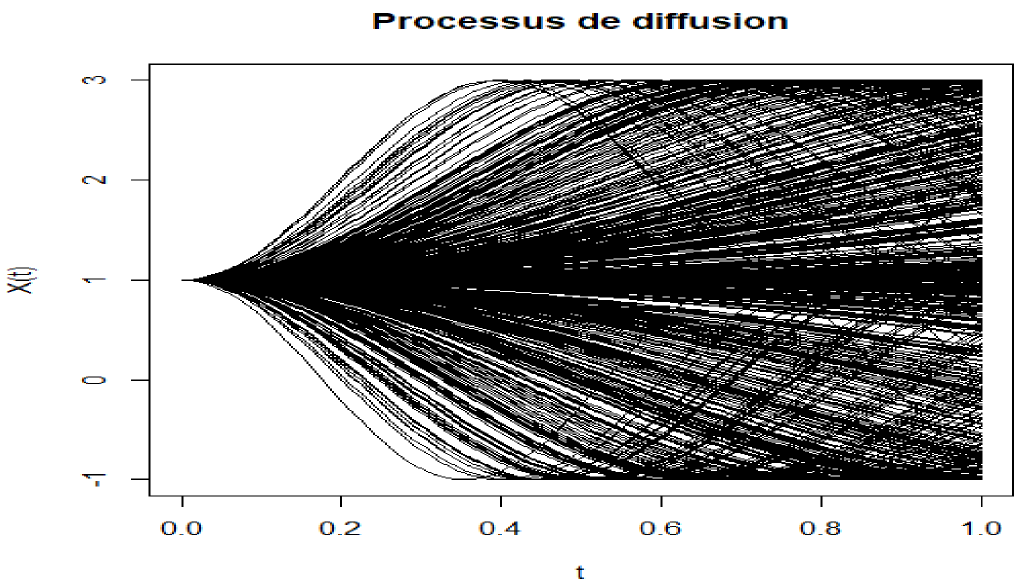

Fig. 1. Simulated curves $X_i(t)_{1 \leq i \leq 600}$

	MSE		
CR(%)	n=100	n=200	n=300
2%	0.1637586	0.1571344	0.1189006
5%	0.1214706	0.1029971	0.07830146
10%	0.07782109	0.08466846	0.06608367
15%	0.06442397	0.07668125	0.07646119

Table 1. Comparison of the *MSE* between $f(y|x)$ and $\hat{f}_n(y|x)$ for $X = x_1$

•**Step 3**: To illustrate the performance of this estimator, we consider each of the three samples and present for different censorship rates, the true conditional density and its estimator in each graph (see Fig. 2 to 5).

The following tables (table 1 and 2) provides the simulation result of the estimated *MSE* for two different covariates.

It follows from these tables that the censorship rate has a significant effect on the quality of the estimator and it improves when The size of the sample increases. This which shows the convergence of this estimator despite the level of censorship.

We illustrate these results by the Fig. 2 to 5. These figures show that all the graphs of the estimator have a form acceptable by relation to the real curve.

MSE			
CR(%)	n=100	n=200	n=300
2%	0.06606901	0.05324895	0.05458689
5%	0.07655961	0.02733061	0.01514639
10%	0.07162354	0.05643691	0.007873876
15%	0.06441725	0.07081149	0.03356378

Table 2. Comparison of the *MSE* between $f(y|x)$ and $\hat{f}_n(y|x)$ for $X = x_2$

MSE			
CR(%)	n=100	n=200	n=300
2%	1.320963	0.622003	0.3393998
5%	2.605931	1.332262	0.4818902
10%	4.530888	2.956179	0.7220436
15%	7.50209	3.974831	1.208553

Table 3. Comparison of *MSE* between $\theta(x)$ and $\hat{\theta}_n(x)$ for $X = x_1$ and N=600

4.2. Mode estimator

In this section, we consider the model studied previously while retaining the same steps and parameters in order to test the performance of the mode estimator, given by:

$$\hat{\theta}_n(x) = \arg\max_{y\in\mathcal{S}} \hat{f}_n(y|x). \tag{14}$$

We present the effect of the sample size as well as that of the censoring rate on the performance of this estimator by evaluating the mean square error through the following formula:

$$MSE = \frac{1}{160}\sum_{j=1}^{160}\left(\hat{\theta}_{nj}(x) - \theta_j(x)\right)^2. \tag{15}$$

To illustrate the performance of this estimator, we set $X = x_2$ and for different censoring rates, we present in each graph, the true mode and its estimator considering each of the three samples (*Fig.* 2 to 5).

The following tables (table 3 and 4) provides the simulation result of the estimated *MSE* for two different covariates.

The result of these tables is that the quality of the estimator improves when the sample size increases or when the censorship rate decreases, which is entirely predictable.

MSE			
CR(%)	n=100	n=200	n=300
2%	0.2263603	0.01325111	0.008276979
5%	0.8849758	0.2310516	0.01514639
10%	2.163933	1.127413	0.03693098
15%	4.342973	1.792074	0.197427

Table 4. Comparison of *MSE* between $\theta(x)$ and $\hat{\theta}_n(x)$ for $X = x_2$ and N=600

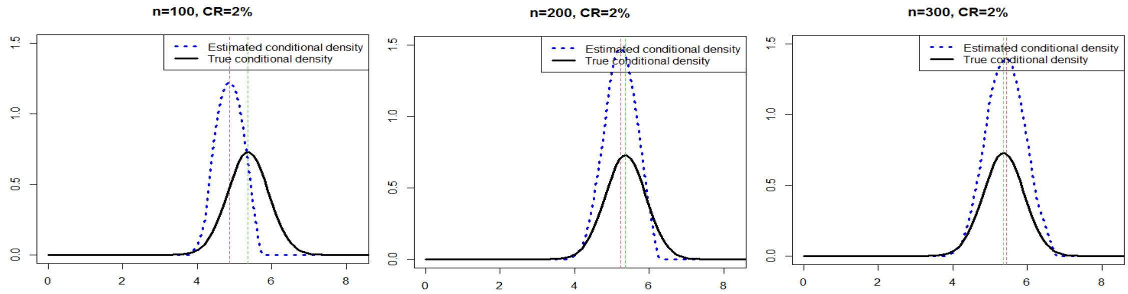

Fig. 2. Illustration of results for CR=2%, $X = x_2 \Rightarrow r(x_2) = 5.365117$

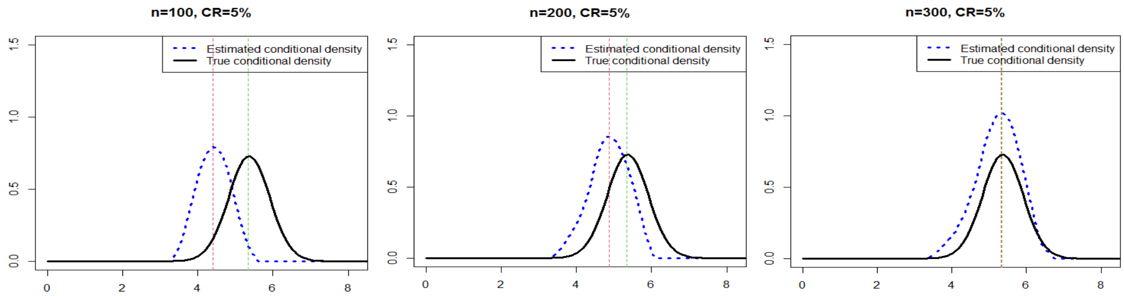

Fig. 3. Illustration of results for CR=5%, $X = x_2 \Rightarrow r(x_2) = 5.365117$

These results are illustrated by the following figures:

The results of the estimation of the mode is not good at even for sample sizes around 300. To obtain optimal results, we repeat the same scenario with slightly larger sample sizes by simulating again 1000 functional data and considering the following samples: $(T_{(i)},\ \delta_{(i)},\ X_{(i)})$ $151 \le i \le 750$; $151 \le i \le 850$ and $151 \le i \le 950$ (n=600; 700; 800 respectively) for $X = x_1$.

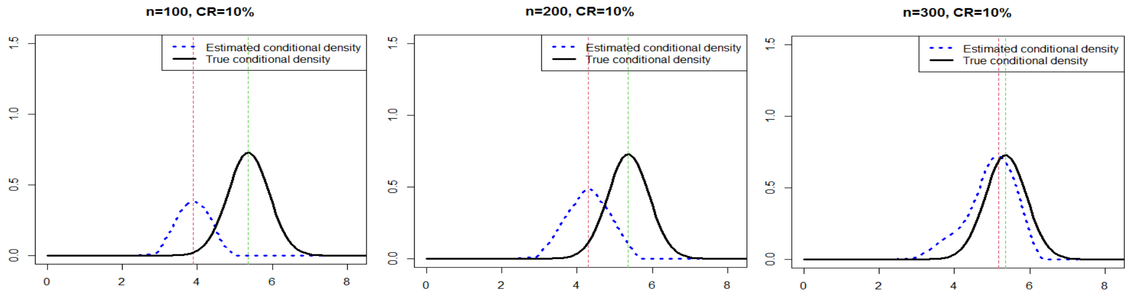

Fig. 4. Illustration of results for CR=10%, $X = x_2 \Rightarrow r(x_2) = 5.365117$

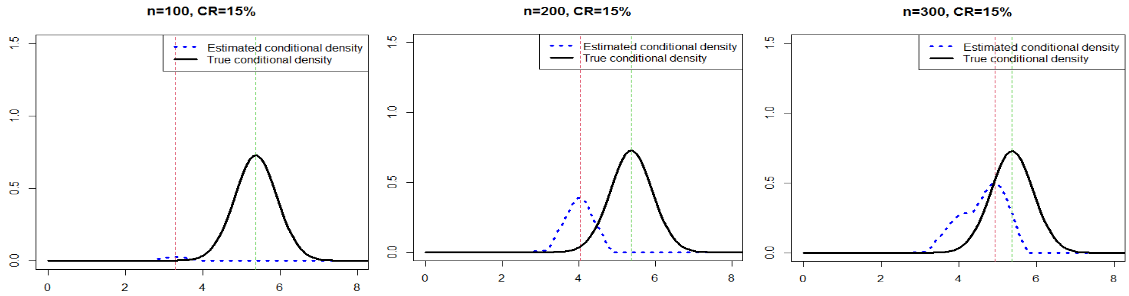

Fig. 5. Illustration of results for CR=15%, $X = x_2 \Rightarrow r(x_2) = 5.365117$

MSE			
CR(%)	n=600	n=700	n=800
2%	0.03196976	0.00120496	0.00120496
5%	0.1712646	0.0002577918	0.0002577918
10%	0.3458807	0.006635499	0.04952229
15%	3.288144	0.2304727	0.06792163

Table 5. Comparison of *MSE* between $\theta(x)$ and $\hat{\theta}_n(x)$ for $X = x_1$ and N=1000

The mean square errors are calculated as before and presented in the following table (see Table 5):

5. Appendix

In this section, we present the proof of the lemmas stated in subsection 3.1.

Proof of lemma (4). Let us note $V_i(x) = \dfrac{K_i(x)}{E[K_1(x)]}$ and $s_n^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} Cov(V_i, V_j)$. We have: $\hat{f}_D(x) = \dfrac{1}{n}\sum_{i=1}^{n} V_i(x)$ and $|V_i(x)| = \dfrac{|K_i(x)|}{|E[K_1(x)]|}$.

The proof of this lemma is based on the evaluation of the sum of the covariances s_n^2 and the application of exponential inequality for mixing bounded variables (for example the corollary A12ii of Ferraty and Vieu (2006)).

• With the assumption (H6), the condition of Lemma 4.3 of Ferraty and Vieu (2006) (p. 43) is satisfied. The application of this lemma gives:

$$\begin{aligned}
|V_i(x)| &\leq \frac{a_1}{|E[K_1(x)]|} \\
&\leq \frac{a_1}{a_2 \phi_x(h_K)} \\
&\leq C \times \frac{1}{\phi_x(h_K)} \\
|V_i(x)| &= O\Big(\frac{1}{\phi_x(h_K)}\Big).
\end{aligned} \tag{16}$$

• By applying relation (6.11) of Ferraty and Vieu (2006) for m = 2, and considering the independence and equidistribution of $(X_i)_{1\leq i\leq n}$, there are real positive constants C_1 and C_2 such as:

$$\frac{C_1}{\phi_x(h_K)} \leq E[V_i^2(x)] \leq \frac{C_2}{\phi_x(h_K)}.$$

Which leads to:

$$E[V_i^2(x)] = O\Big(\frac{1}{\phi_x(h_K)}\Big). \tag{17}$$

• For $i \neq j$, we have:

$$\begin{aligned}
E[V_i V_j] &= E\Big[\frac{1}{E^2[K_1]} K_i K_j\Big] \\
&= \frac{1}{E^2[K_1]} E[K_i K_j].
\end{aligned}$$

With the assumption (H6), Lemma 4.3 of Ferraty and Vieu (2006) and the assumption (H4B) give:

$$
\begin{aligned}
E[V_iV_j] &= \frac{1}{E^2[K_1]} P\Big((X_i, X_j) \in B(x, h_K) \times B(x, h_K)\Big) \\
&\leq \frac{\varphi_x(h_K)}{E^2[K_1]} \ \text{ avec } \varphi_x(h_K) = O(\phi_x^{1+\varepsilon_1}(h_K)).
\end{aligned}
$$

From the inequality (6.10) of Ferraty and Vieu (2006), for m = 1, we have:

$$
\begin{aligned}
|E[V_iV_j]| &\leq \frac{\varphi_x(h_K)}{\phi_x^2(h_K)} \\
|E[V_iV_j]| &= O\Big(\max\{\phi_x^{-1+\varepsilon_1}(h_K), 1\}\Big).
\end{aligned} \tag{18}
$$

So, we have

$$
\begin{aligned}
|Cov(V_i, V_j)| &= |E[V_iV_j] - E[V_i]E[V_j]| \\
&\leq |E[V_iV_j]| \\
&= O\Big(\max\{\phi_x^{-1+\varepsilon_1}(h_K), 1\}\Big).
\end{aligned} \tag{19}
$$

In addition, using the decomposition of Marsy (1986), for any following v_n positive, let's define

$$
\mathcal{D}_1 = \{(i,j) : 1 \leq |i-j| \leq v_n - 1\} \tag{20}
$$

and

$$
\mathcal{D}_2 = \{(i,j) : v_n \leq |i-j| \leq n\}. \tag{21}
$$

We have the following decomposition:

$$
s_n^2 = nVar(V_i) + \sum_{i=1}^{n} \sum_{\mathcal{D}_1} Cov(V_i, V_j) + \sum_{i=1}^{n} \sum_{\mathcal{D}_2} Cov(V_i, V_j). \tag{22}
$$

To treat each term of this decomposition, we have from equality (17)

$$
nVar(V_i) = O\Big(\frac{n}{\phi_x(h_K)}\Big). \tag{23}
$$

According to (19),

$$\sum_{i=1}^{n}\sum_{\mathcal{D}_1} Cov(V_i,V_j) = O\Big(nv_n \max\{\phi_x^{-1+\varepsilon_1}(h_K),1\}\Big). \tag{24}$$

Regarding the term $\sum_{i=1}^{n}\sum_{\mathcal{D}_2} Cov(V_i,V_j)$, we use covariance inequality for mixing processes (Davydov-Rio inequality)

$$\forall\ i\neq j\ Cov(V_i,V_j) \le C\alpha\big(|i-j|\big) \tag{25}$$

to have:

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{\mathcal{D}_2} Cov(V_i,V_j) &\le Cn \sum_{v_n<|i-j|<n} \alpha\big(|i-j|\big)\\
&\le Cn\int_{v_n}^{n} u^{-a}du \quad \text{after (H4a)}\\
&\le Cn\Big[\frac{1}{1-a}u^{1-a}\Big]_{v_n}^{n}\\
&\le \frac{Cnv_n^{1-a}}{a-1}\\
\sum_{i=1}^{n}\sum_{\mathcal{D}_2} Cov(V_i,V_j) &= O\Big(nv_n^{1-a}\Big).
\end{aligned} \tag{26}$$

Finally, we have from (22), (23), (24) and (26)

$$s_n^2 = O\Big(\frac{n}{\phi_x(h_K)}\Big) + O\Big(nv_n \max\{\phi_x^{-1+\varepsilon_1}(h_K),1\}\Big) + O\Big(nv_n^{1-a}\Big).$$

Now just choose $v_n = \phi_x^{-\varepsilon_1}(h_K)$ and consider (H4) to have

$$\begin{aligned}
s_n^2 &= O\Big(\frac{n}{\phi_x(h_K)}\Big) + O\Big(n\max\{\phi_x^{-1}(h_K),\phi_x^{-\varepsilon_1}(h_K)\}\Big) + O\Big(n\phi_x^{\varepsilon_1(a-1)}(h_K)\Big)\\
s_n^2 &= O\Big(\frac{n}{\phi_x(h_K)}\Big).
\end{aligned} \tag{27}$$

Using the terminals given by equalities (16) and (17) and by applying an exponential inequality for mixing bounded variables, we arrive at

$$f_D(x) - E[f_D(x)] = O\Big(n^{-1}\sqrt{\log n s_n^2}\Big). \tag{28}$$

The result stems directly from (27) and (28) since $E[f_D(x)] = 1$. ■.

Proof of lemma (3). Since $\tilde{f}_N(x,y) = \frac{1}{nh_H E[K_1(x)]}\sum_{i=1}^n H_i(y)\delta_i \bar{G}^{-1}(t)K_i(x)$, we have proof of lemma 3 by considering that of Lemma 3.4 of Benkhaled et al. (2020) and replacing $E[W_{12}(x)]$ by $E[K_1(x)]$ in the expression of $\tilde{f}_N(x,y)$.

Indeed, by the equiprobability of the couples, (Y_i, X_i) and the fact that $K_1(x)$ is a measurable application compared to X_1 (according to (H6)), we have:

$$\begin{aligned} E\tilde{f}_N(x,y) &= \frac{1}{nh_H E[K_1(x)]}\sum_{i=1}^n E\Big[H_i(y)\delta_i \bar{G}^{-1}(T_i)K_i(x)\Big] \\ &= \frac{1}{nh_H E[K_1(x)]}\sum_{i=1}^n E\Big[K_1(x)E\Big(H_i(y)\delta_i \bar{G}^{-1}(T_i)|X_1\Big)\Big]. \end{aligned} \tag{29}$$

$H_i(y)$ being also measurable compared to Y_1 (according to (H5)), Using conditional expectation properties and the fact that:

$$\mathbb{I}_{\{Y_1 \le C_1\}}\varphi(T_1) = \mathbb{I}_{\{Y_1 \le C_1\}}\varphi(Y_1) \text{ for any measurable function } \varphi,$$

we have:

$$\begin{aligned} E\Big[H_i(y)\delta_i \bar{G}^{-1}(T_i)|X_1\Big] &= E\Big[E\Big(H_i(y)\mathbb{I}_{\{Y_1 \le C_1\}}(y)\bar{G}^{-1}(Y_1)|Y_1\Big)|X_1\Big] \\ &= E\Big[H_i(y)\bar{G}^{-1}(Y_1)E\Big(\mathbb{I}_{\{Y_1 \le C_1\}}(y)|Y_1\Big)|X_1\Big] \\ &= E\Big[H_i(y)|X_1\Big] \\ &= \int_{\mathbb{R}} H'\Big(\frac{y-z}{h_H}\Big)f(z|X_1)dz. \end{aligned}$$

Then by assumption (H5) and using the change of variable $u = \frac{y-z}{h_H}$, we have:

$$E\Big[H_i(y)\delta_i \bar{G}^{-1}(T_i)|X_1\Big] = h_H \int_{\mathbb{R}} H'(u)f(y-uh_H|X_1)du. \tag{30}$$

By combining (29) and (30), we have:

$$\begin{aligned} E[\tilde{f}_N(x,y)] &= \frac{1}{nh_H E[K_1(x)]}\sum_{i=1}^n EK_1(x)h_H \int_{\mathbb{R}} H'(u)f(y-uh_H|X_1)du \\ &= \int_{\mathbb{R}} H'(u)f(y-uh_H|X_1)du. \end{aligned}$$

With the use of assumptions (H2) and (H5), we obtain

$$\begin{aligned}
\Big|E[\tilde{f}_N(x,y)] - f(y|x)\Big| &= \Big|\int_{\mathbb{R}} H'(u) f(y - uh_H|X_1)du - f(y|x)\Big| \\
&= \Big|\int_{\mathbb{R}} H'(u) f(y - uh_H|X_1) - f(y|x)du\Big| \\
&\leq \int_{\mathbb{R}} H'(u)\Big|f(y - uh_H|X_1) - f(y|x)\Big|du \\
&\leq \int_{\mathbb{R}} H'(u)A_x\Big(d(x,X_1)^{b_1} + |y - y + uh_H|^{b_2}\Big)du \\
&\leq A_x\int_{\mathbb{R}} H'(u)h_K^{b_1}du + A_x\int_{\mathbb{R}} H'(u)|u|^{b_2}h_H^{b_2}du \\
&\leq h_K^{b_1}A_x\int_{\mathbb{R}} H'(u)du + h_H^{b_2}A_x\int_{\mathbb{R}} H'(u)|u|^{b_2}du \\
\Big|E[\tilde{f}_N(x,y)] - f(y|x)\Big| &= O\big(h_K^{b_1}\big) + O\big(h_H^{b_2}\big) \qquad (31)
\end{aligned}$$

■

Proof of lemma (2). The compactness of $\mathcal{S}$ allows it to cover it by disjoint intervals so that: $y \in \mathcal{S} \subset \cup_{k=1}^{u_n}[t_k - l_n; t_k + l_n]$ where $t_1, \cdots, t_{u_n}$ are points of $\mathcal{S}$. l_n and u_n are chosen such as:

$$\exists C > 0, \exists c > 0, \quad l_n = Cu_n^{-1} = n^{-C}. \qquad (32)$$

Let $t_y = \min_{t_k \in \{t_1,\cdots,t_{u_n}\}} |y - t_k|$ the only t_k such that $y \in [t_k - l_n; t_k + l_n[$.
Consequently, the proof follows from the decomposition

$$\begin{aligned}
\frac{1}{f_D(x)}\sup_{y\in\mathcal{S}}\Big|\tilde{f}_N(x,y) - E\big[\tilde{f}_N(x,y)\big]\Big| \leq \frac{1}{f_D(x)}\Bigg[&\sup_{y\in\mathcal{S}}\Big|\tilde{f}_N(x,y) - \tilde{f}_N(x,t_y)\Big| + \\
&\sup_{y\in\mathcal{S}}\Big|E\big[\tilde{f}_N(x,y)\big] - E\big[\tilde{f}_N(x,t_{y,x})\big]\Big| + \\
&\sup_{y\in\mathcal{S}}\Big|\tilde{f}_N(t_y,x) - E\big[\tilde{f}_N(t_y)\big]\Big|\Bigg] \qquad (33)
\end{aligned}$$

and following relations:

$$i)\ \frac{1}{f_D(x)}\sup_{y\in\mathcal{S}}\Big|\tilde{f}_N(x,y) - \tilde{f}_N(x,t_y)\Big| = O\left(\sqrt{\frac{\log n}{nh_H\phi_x(h_K)}}\right) a.co; \qquad (34)$$

$$ii)\ \frac{1}{f_D(x)}\sup_{y\in\mathcal{S}}\Big|E\big[\tilde{f}_N(y,x)\big] - E\big[\tilde{f}_N(x,t_y)\big]\Big| = O\left(\sqrt{\frac{\log n}{nh_H\phi_x(h_K)}}\right) a.co; \qquad (35)$$

$$iii)\ \frac{1}{f_D(x)}\sup_{y\in\mathcal{S}}\left|\tilde{f}_N(x,t_y)-E\big[\tilde{f}_N(x,t_y)\big]\right| = O\left(\sqrt{\frac{\log n}{nh_H\phi_x(h_K)}}\right)a.co. \tag{36}$$

Proof of i) : By assumption (H5), we have:

$$\begin{aligned}
\sup_{y\in\mathcal{S}}\left|\tilde{f}_N(x,y)-\tilde{f}_N(x,t_y)\right| &= \frac{1}{nh_HE\big[K_1(x)\big]}\sum_{i=1}^{n}\delta_i\bar{G}^{-1}(t)K_i(x)\Big(H_i(y)-H_i(t_y)\Big)\\
&\le \frac{C}{nh_HE\big[K_1(x)\big]}\sum_{i=1}^{n}\delta_i\bar{G}^{-1}(t)K_i(x)\frac{|y-t_y|}{h_H}\\
&\le \frac{C}{nh_HE[K_1(x)]}\sum_{i=1}^{n}\delta_iK_i(x)\frac{|y-t_y|}{h_H}\frac{1}{\bar{G}(t)}\\
&\le Cl_n\bar{G}^{-1}(t)f_D(x)h_H^{-2}\\
\frac{1}{f_D(x)}\sup_{y\in\mathcal{S}}\left|\tilde{f}_N(x,y)-\tilde{f}_N(x,t_y)\right| &\le Cl_n\bar{G}^{-1}(t)h_H^2 = O\Big(n^{-c}\Big).
\end{aligned} \tag{37}$$

We have the result by considering (32) and by choosing c large enough.

Proof of ii) : This is proven from the lemma 4 and inequality (37) using the proposition A6ii of Ferraty and Vieu (2006).

Proof of iii): Let us note $\tilde{f}_N(x,y)=\dfrac{1}{n}\sum_{i=1}^{n}W_i$ with

$W_i=\dfrac{1}{h_HE\big[K_1(x)\big]}\delta_i\bar{G}^{-1}(t)H_i(y)K_i(x)$ and $S_n^2=\sum_{i=1}^{n}\sum_{j=1}^{n}Cov(W_i,W_j)$.

By taking up the same approach as that used to prove the lemma 4, we can evaluate the sum of the covariances s_n^2. We have so:

$$|W_i| = O\Big(\frac{1}{h_H\phi_x(h_K)}\Big) \tag{38}$$

and for $i\neq j$ we have:

$$\begin{aligned}
\big|E[W_iW_j]\big| &= \frac{1}{h_H^2E^2\big[K_1(x)\big]}E\Big[\delta_i\delta_j\bar{G}^{-1}(T_i)\bar{G}^{-1}(T_j)K_i(x)H_i(t_y)K_j(x)H_j(t_y)\Big]\\
&= \frac{1}{h_H^2\phi_x^2(h_K)}E\Big[K_i(x)K_j(x)E\big[H_i(t_y)H_j(t_y)\delta_i\delta_j\bar{G}^{-1}(T_i)\bar{G}^{-1}(T_j)|X_iX_j\big]\Big]\\
&= \frac{1}{h_H^2\phi_x^2(h_K)}E\Big[K_i(x)K_j(x)\gamma\Big]
\end{aligned} \tag{39}$$

with

$$
\begin{aligned}
\gamma &= E\big[H_i(t_y)H_j(t_y)\delta_i\delta_j\bar{G}^{-1}(T_i)\bar{G}^{-1}(T_j)|X_iX_j\big] \\
&= E\big[H_i(t_y)H_j(t_y)\bar{G}^{-1}(T_i)\bar{G}^{-1}(T_j)E[\delta_i\delta_j|T_{yi}T_{yj}]|X_iX_j\big] \\
&= E\big[H_i(t_y)H_j(t_y)\bar{G}^{-1}(T_i)\bar{G}^{-1}(T_j)E[\mathbb{I}_{\{Tyi\le C_i\}}\mathbb{I}_{\{Tyj\le C_j\}}|T_{yi}T_{yj}]|X_iX_j\big] \\
&= E\big[H_i(t_y)H_j(t_y)\bar{G}^{-1}(T_i)\bar{G}^{-1}(T_j)E[\mathbb{I}_{\{Tyi\le C_i\}\cap\{Tyj\le C_j\}}|T_{yi}T_{yj}]|X_iX_j\big] \\
&= E\big[H_i(t_y)H_j(t_y)\bar{G}^{-1}(T_i)\bar{G}^{-1}(T_j)P[\{Tyi\le C_i\}\cap\{Tyj\le C_j\}|T_{yi}T_{yj}]|X_iX_j\big].
\end{aligned}
$$

As T_{yi} and C_i are independent, we have:

$$
\begin{aligned}
\gamma &= E\big[H_i(t_y)H_j(t_y)\bar{G}^{-1}(T_i)\bar{G}^{-1}(T_j)P[\{Tyi\le C_i\}\cap\{Tyj\le C_j\}]|X_iX_j\big] \\
&= E\big[H_i(t_y)H_j(t_y)\bar{G}^{-1}(T_i)\bar{G}^{-1}(T_j)P(\{Tyi\le C_i\})P(\{Tyj\le C_j\})|X_iX_j\big] \\
&= E\big[H_i(t_y)H_j(t_y)\bar{G}^{-1}(T_i)\bar{G}^{-1}(T_j)\bar{G}(T_i)\bar{G}(T_j)|X_iX_j\big] \\
\gamma &= E\big[H_i(t_y)H_j(t_y)|X_iX_j\big]. \qquad (40)
\end{aligned}
$$

As result,

$$
\begin{aligned}
\big|E[W_iW_j]\big| &= \frac{1}{h_H^2\phi_x^2(h_K)}E\Big[K_i(x)K_j(x)\int_{\mathbb{R}^2}H'\Big(\frac{t_y-u_1}{h_H}\Big)H'\Big(\frac{t_y-u_2}{h_H}\Big)f_{ij}(u_1,u_2)du_1du_2\Big] \\
&\le \frac{c}{h_H^2\phi_x^2(h_K)}E\Big[K_i(x)K_j(x)\Big] \\
&\le \frac{c}{h_H^2\phi_x^2(h_K)}P\Big((X_i,X_j)\in B(x,h_K)\times B(x,h_K)\Big). \qquad (41)
\end{aligned}
$$

With reasoning similar to proof of lemma 4, we get:

$$
\big|E[W_iW_j]\big| = O\Big(h_H^{-2}\max\{\phi_x^{-1+\varepsilon_1}(h_K),1\}\Big)
$$

and

$$
|Cov(W_i,W_j)| = O\Big(h_H^{-2}\max\{\phi_x^{-1+\varepsilon_1}(h_K),1\}\Big). \qquad (42)
$$

considering (20) and (21), we can write

$$
S_n^2 = nVar(W_i) + \sum_{i=1}^{n}\sum_{\mathcal{D}_1}Cov(W_i,W_j) + \sum_{i=1}^{n}\sum_{\mathcal{D}_2}Cov(W_i,W_j). \qquad (43)
$$

To treat each term of this decomposition, we have:

$$
\begin{aligned}
E\left[W_i^2\right] &= \frac{1}{h_H^2 E^2\left[K_1(x)\right]} E\left[\delta_i^2 \bar{G}^{-2}(t) K_i^2(x) H_i^2(y)\right] \\
&\leq \frac{C}{h_H^2 E^2\left[K_1(x)\right]} E\left[K_i^2(x) E(H_i^2(y)|X_i)\right] \\
&\leq \frac{C}{h_H E^2[K_1(x)]} E\left[K_i^2(x) \int_{\mathbb{R}} \frac{1}{h_H} H'^2\left(\frac{y-u}{h_H}\right) f(y|x) du\right].
\end{aligned}
$$

As $f(y|x)$ is continuous and H' is with a compact support, there exist a constant C' such that

$$
\begin{aligned}
E\left[W_i^2\right] &\leq \frac{C'}{h_H E^2[K_1(x)]} E\left[K_i^2(x)\right] \\
E\left[W_i^2\right] &= O\left(\frac{1}{h_H \phi_x(h_K)}\right). \qquad (44)
\end{aligned}
$$

From 44, we have the first term

$$
nVar(W_i) = O\left(\frac{n}{\phi_x(h_K)}\right). \qquad (45)
$$

According to (42), we have the second term

$$
\sum_{i=1}^{n} \sum_{\mathcal{D}_1} Cov(W_i, W_j) = O\Big(nv_n \max\{\phi_x^{-1+\varepsilon_1}(h_K), 1\}\Big). \qquad (46)
$$

Regarding the term $\sum_{i=1}^{n} \sum_{\mathcal{D}_2} Cov(W_i, W_j)$, let us use the inequality of covariance for mixing process (Davydov-Rio inequality)

$$
\forall\ i \neq j\ Cov(W_i, W_j) \leq C\alpha(|i-j|)
$$

to get:

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{\mathcal{D}_2} Cov(W_i, W_j) &\leq Cn \sum_{v_n<|i-j|<n} \alpha(|i-j|) \\
&\leq cn \int_{v_n}^{n} u^{-a} du, \quad \text{after (H4a)} \\
&\leq cn \Big[\frac{1}{1-a} u^{1-a}\Big]_{v_n}^{n} \\
&\leq \frac{cnv_n^{1-a}}{a-1} \\
\sum_{i=1}^{n}\sum_{\mathcal{D}_2} Cov(W_i, W_j) &= O\Big(nv_n^{1-a}\Big). \qquad (47)
\end{aligned}$$

Finally, from (45), (46) and (47), we have

$$S_n^2 = O\left(\frac{n}{\phi_x(h_K)}\right) + O\left(nv_n \max\left\{\phi_x^{-1+\varepsilon_1}(h_K), 1\right\}\right) + O\Big(nv_n^{1-a}\Big).$$

Now, just choose $v_n = \phi_x^{-\varepsilon_1}(h_K)$ and consider assumption (H4c) to have

$$\begin{aligned}
S_n^2 &= O\left(\frac{n}{\phi_x(h_K)}\right) + O\Big(n \max\{\phi_x^{-1}(h_K), \phi_x^{-\varepsilon_1}(h_K)\}\Big) + O\Big(n\phi_x^{\varepsilon_1(a-1)}(h_K)\Big) \\
S_n^2 &= O\left(\frac{n}{\phi_x(h_K)}\right). \qquad (48)
\end{aligned}$$

Using the terminals given by equality (38) and (44), applying exponential inequality for mixing bounded variables (for example Corollary A12ii of Ferraty and Vieu (2006)), we obtain

$$\tilde{f}_N(t_y, x) - E\tilde{f}_N(t_y, x) = O\Big(n^{-1}\sqrt{\log n S_n^2}\Big). \qquad (49)$$

In addition, using (32), we get:

$$\begin{aligned}
P\left[\sup_{y\in\mathcal{S}}\left|\tilde{f}_N(x,t_y)-E\left[\tilde{f}_N(x,t_y)\right]\right|>\varepsilon_0\sqrt{\frac{\log n}{nh_H\phi_x(h_K)}}\right] &= P\left[\max_{t_y\in\{t_1,\cdots,t_{u_n}\}}\left|\tilde{f}_N(x,t_y)\right.\right.\\
&\quad \left.\left.-\ E\left[\tilde{f}_N(x,t_y)\right]\right|>\varepsilon_0\sqrt{\frac{\log n}{nh_H\phi_x(h_K)}}\right]\\
&\leq n^c P\left[\left|\tilde{f}_N(x,t_y)-E\left[\tilde{f}_N(x,t_y)\right]\right|\right.\\
&\quad \left.>\ \varepsilon_0\sqrt{\frac{\log n}{nh_H\phi_x(h_K)}}\right]. \qquad (50)
\end{aligned}$$

Result iii) takes directly from (48), (49) and (50).

Thus, just use i), ii) and iii) to finish proof of lemma 2. ■

[**Proof of lemma** (1). Following the same approach as that of lemma 3.6 of Benkhaled et al. (2020), we have:

$$\begin{aligned}
\left|\hat{f}_N(x,y)-\tilde{f}_N(x,y)\right| &\leq \frac{1}{nh_H E\left[K_1(x)\right]}\sum_{i=1}^{n}\left|\delta_i H_i(y)K_i(x)\left(\bar{G}_n^{-1}(T_i)-\bar{G}^{-1}(T_i)\right)\right|\\
&\leq \frac{\sup_{y\in\Omega}\left|\bar{G}_n(t)-\bar{G}(t)\right|}{\bar{G}_n(\tau)}\frac{1}{nh_H E\left[K_1(x)\right]}\sum_{i=1}^{n}\delta_i\bar{G}^{-1}(T_i)H_i(y)K_i(x)\\
&\leq \frac{\sup_{y\in\Omega}\left|\bar{G}_n(t)-\bar{G}(t)\right|}{\bar{G}_n(\tau)}\tilde{f}_N(x,y).
\end{aligned}$$

Since $\bar{G}(\tau)>0$ 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 Deheuvel and Einmahl (2000)), the result is an immediate consequence of Lemmas 2 and 3. ■

Conclusion

In this work, we have established the almost complete pointwise convergence of the kernel estimator of the conditional density and the almost sure convergence of the mode estimator when the covariate is functional and the data is censored randomly on the right and algebraically α-mixing. Through a numerical study, we have made the automatic choice of the two smoothing parameters with the cross validation method for one and the functional version of cross validation ideas for

the other in order to prove that our estimators converge with a Optimal speed. The capacity of our computers did not allow us to reproduce for samples of larger sizes to thus lead us numerically to convergence and therefore to the observation of the compliance of the $f(y|x)$ Curves and its nonparametric estimator $\hat{f}_n(y|x)$.
In the future, we can extend our work on the case of censored and β-mixing data. %End Article

References

Agbokou K. and Gneyou K.E.(2017). On the strong convergence of the hazard rate and its maximum risk point estimators in presence of censorship and functional explanatory covariate. *Afrika Statistika*. Vol. 12 (3), 2017, pages 1397 - 1416.

Agbokou K. and Gneyou K. E.(2019) *Asymptotic properties of the conditional hazard function and its maximum estimation under right-censoring and left-truncation. Jordan Journal of Mathematics and Statistics (JJMS) 12(3), 2019, pp 351 - 374*

Agbokou K. and Gneyou K. E. (2019) on the asymptotic properties for a nonparametric conditional quantile estimator in presence of right-censoring. *Far East Journal of Theoretical Statistics*. Volume 56, Number 1, 2019, Pages 1-34

Abdelkader Benkhaled; Fethi Madani; Salah Khardani (2020). *Strong consistency of local linear estimation of a conditional density function under random censorship*

Dabo-Niang S., Laksaci A., *Estimation non paramétrique du mode conditionnel pour variable explicative fonctionnelle. R. Acad. Sci. Paris, Ser. I 344 (2007).* C.

Deheuvels, P.; Einmahl, H.: *Functional limit laws for the increments of Kaplan-Meier product limit processes and applications. Ann. Probab. 28, 1301-1335 (2000)*

Demongeot, J., Laksaci, A., Madani, F., Rachdi, M., (2010). *Estimation locale linéaire de la densité conditionnelle pour des données fonctionnelles. C. R. Acad. Sci. Paris, Ser. I 348 (15-16), 931-934.*

Demongeot, J., Laksaci, A., Madani, F., Rachdi, M., (2011). *A fast functional locally modeled conditional density and mode for functional time-series. In: Recent Advances in Functional Data Analysis and Related Topics. In: Contributions to Statistics, Physica-Verlag, Springer, pp. 85-90.*

F. Ferraty A. Rabhi and P. Vieu, (2008) *Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle. Rev. Roumaine Math.pures Appl; 53 (2008), 1, 1-8*

F. Ferraty and P. Vieu, (2006) *Nonparametric Functional Data Analysis. Theory and practice. Springer Series in Statistics.*

Gneyou K. E. (2013). *A central limit theorem for a nonparametric maximum conditional hazard rate estimator in presence of right censoring. International Journ. Probab. Statist. Vol. 2, No. 3 pp. 110-124 URL: http://dx.doi.org/10.5539/ijsp.v2n3p110*

Gneyou K. E. (2014) *A strong linear representation for the maximum conditional hazard rate estimator in survival analysis, Journal of Multivariate Analysis 128, 10-18.*

Kebir, H.; Mechab, B. (2023). *Estimation of the Conditional Hazard Function with a Recursive Kernel from Censored Functional Ergodic Data. Comput. Sci. Math. Forum 2023, 7, 16. https://doi.org/10.3390/ IOCMA2023-14412*

Khardani S. and Semmar S. (2014). *Nonparametric conditional density estimation for censored data based on a recursive kernel. Electronic Journal of Statistics 8 (2):2541-56. doi:10.1214/14- EJS960.*

Khardani, S., Lemdani, M., Ould Saïd, E. (2010). *Some asymptotic properties for a smooth kernel estimator of the conditional mode under random censorship, J. of the Korean Statistical Society, 39, 455-469.*

Kim, C., M. Oh, S.J. Yang, and H. Choi. (2010) *A local linear estimation of conditional hazard function in censored data. Journal of the Korean Statistical Society 39(3): 347-355.*

Laksaci, A.; Madani, F.; Rachdi, M. (2013): *Kernel conditional density estimation when the regressor is valued in a semi-metric space. Commun. Stat. Theory Methods 42(19), 3544-3570*

Liang H. Y. and Peng L. (2010) *Asymptotic normality and Berry-Esseen results for conditional density estimator with censored and dependent data. Journal of Multivariate Analysis 101 (5): 1043-54. doi:10.1016/j.jmva.2010.01.002.*

Slaoui Y. and Khardani S. (2020) *Adaptive recursive kernel conditional density estimators under censoring data. ALEA, Lat. Am. J. Probab. Math. Stat. 17, 389-417.*

Spierdijk, L. (2008). *Nonparametric conditional hazard rate estimation: A local linear approach. Computational Statistics and Data Analysis 52: 2419-2434.*

Xianzhou X. and Meijuan Ou. (2019) *Non parametric estimation of the conditional density function with right-censored and dependent data, Communications in Statistics-Theory and Methods, DOI: 10.1080/03610926.2019.1691230*