Applications and Applied Mathematics: An International Journal (AAM)

Volume 17 | Issue 2 Article 3

12-2022

(R1899) Asymptotic Normality of the Conditional Hazard Function in the Local Linear Estimation Under Functional Mixing Data

Amina Goutal
Djillali Liabes University

Boubaker Mechab
Djillali Liabes University

Omar Fetitah
Djillali Liabes University

Torkia Merouan
Djillali Liabes University

Available at
http://pvamu.edu/aam
Appl. Appl. Math.
ISSN: 1932-9466

Applications and Applied Mathematics:
An International Journal
(AAM)

Vol. 17, Issue 2 (December 2022), pp. 308 – 332

Asymptotic Normality of the Conditional Hazard Function In the Local Linear Estimation Under Functional Mixing Data

[1]**Amina Goutal**, [2*]**Boubaker Mechab**, [3]**Omar Fetitah and** [4]**Torkia Merouan**

Department of Probability and Statistics
Djillali Liabes University
Laboratory of Statistics and Stochastic Processes
Sidi Bel Abbes 22000, Algeria
[1]aminasara90@hotmail.com; [2*]mechaboub@yahoo.fr;
[3]fetitah-omar@hotmail.com; [4]merouan-to@hotmail.com

*Corresponding Author

Received: December 17, 2021; Accepted: June 6, 2022

Abstract

In this study, we are interested in using the local linear technique to estimate the conditional hazard function for functional dependent data where the scalar response is conditioned by a functional random variable. The asymptotic normality of this constructed estimator is demonstrated under some extreme conditions. Our estimator's performance is demonstrated through simulations.

Keywords: Functional Data Analysis (FDA); Asymptotic normality; Conditional hazard function; Local linear estimation; Mixing data

MSC 2020 No.: 62G05, 62G07, 62G20

1. Introduction

With the emergence of advanced measurement devices, the examination of functional data has recently become commonplace. Ferraty and Vieu (2006), Ramsay and Silverman (2002) and others provide an overview of the current state of nonparametric functional data.

Ferraty and Vieu's monograph (2006) highlights many of their contributions to non-parametric estimation with functional data, including the consistency of conditional density, conditional distribution and regression estimates among other aspects in the independent identically distributed (i.i.d.) case as well as under dependence conditions (strong mixing). Different strategies are applied to various examples of functional data samples, and almost complete rates of convergence are produced. Masry (2005) proved the asymptotic normality of the functional non-parametric regression estimate considering strong mixing dependence conditions for the sample data. For automatic smoothing parameter selection in the regression setting, we cite the work of Rachdi and Vieu (2007).

The nonparametric estimation of a hazard function has been considered by many authors. Watson and Leadbetter (1964) showed the asymptotic normality of the hazard rate function with dependence conditions. The uniform convergence properties and asymptotic normality of an estimate of the maximum of the hazard function in a context of strong mixing was established by Quintela-del-Rio (2006). Then, two years later, the same author showed that the kernel estimator of the hazard function is strongly consistent and asymptotically normally distributed.

In the case of finite-dimensional data, local linear estimation technique has several advantages over the kernel method such as bias reduction and adaptation of effects. The first results on this model in FDA setup were established by Baillo and Grané (2009) and Berlinet et al. (2011). These authors considered the local linear estimation of the regression operator when the regressor takes values in a Hilbert space. Barrientos-Marin et al. (2010) and El Methni and Rachdi (2011) proposed the almost complete convergence with rates, of the proposed estimator. Very recently, Demongot et al. (2014) applied this method for the conditional cumulative distribution function and Messaci et al. (2015) for the conditional quantile.

In this article, we employed the asymptotic normality of local linear estimator of the functional hazard function for functional dependent data. Recall that Zhou and Lin (2016) established the asymptotic normality of linear locally modelled regression for functional data. Bouanani et al. (2018) demonstrated the asymptotic normality of the linear estimators of several statistical parameters, such as the conditional cumulative distribution, conditional density derivatives and conditional mode.

This paper is organized as follows. The model and the functional local linear estimator of the conditional hazard function are introduced in Section 2. Notations and hypotheses in Section 3. In Section 4, we treated the asymptotic normality of the estimator for dependent functional data and used our findings to build the confidence interval. The asymptotic of the estimate is demonstrated in Section 5 through a simulation. Finally, in the appendix, we show the proofs of our results.

2. Model and Estimator

Let $(X_i, Y_i)_{\{i\in\mathbb{N}\}}$ be a valued measurable strictly stationary process, where X_i takes values in semi-metric space $\mathcal{F}$ with a semi-metric d and Y_i is a real valued.

We denote by $h(y|x)$ the conditional hazard function of Y given $X = x$, which is defined from the conditional density function $f(y|x)$ and the survival function $S(x, y) = 1 - F(y|x)$ as follows,

$$h(y|x) = \frac{f(y|x)}{S(x,y)} = \frac{f(y|x)}{1 - F(y|x)}, \quad \forall y \in \mathbb{R},$$

and the conditional probability distribution function noted $F(y|x)$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$, defined by

$$F(y|x) = \mathbb{P}(Y \leq y \mid X = x), \quad \forall y \in \mathbb{R}.$$

The local linear estimator of the conditional hazard function $h(y|x)$ is given by

$$\widehat{h}(y|x) = \frac{\widehat{f}(y|x)}{1 - \widehat{F}(y|x)}, \quad \forall y \in \mathbb{R}.$$

The estimate of $F(y|x)$ (see Demongot et al. (2014)) is estimated by $\widehat{a}$ where the couple $(\widehat{a}, \ \widehat{b})$ is obtained by the optimization rule:

$$\min_{(a,b)\in\mathbb{R}^2} \sum_{i=1}^{n} \Big(H(h_H^{-1}(Y_i - y)) - a - b\beta(X_i, x)\Big)^2 K(h_K^{-1}\delta(X_i, x)).$$

Formally, $(\widehat{a}, \widehat{b})$ is a solution of the system

$$({}^tQ_\beta K Q_\beta) \begin{pmatrix} a \\ b \end{pmatrix} - ({}^tQ_\beta K H) = 0,$$

which allows to

$$\begin{pmatrix} \widehat{a} \\ \widehat{b} \end{pmatrix} = ({}^tQ_\beta K Q_\beta)^{-1}({}^tQ_\beta K H),$$

where ${}^tQ_\beta = \begin{pmatrix} 1 & \cdots & 1 \\ \beta(X_1, x) & \cdots & \beta(X_n, x) \end{pmatrix}$, $K = diag(K(h_K^{-1}\delta(X_1, x)), \ldots, K(h_K^{-1}\delta(X_n, x)))$ and ${}^tH = (H(h_H^{-1}(Y_1 - y)), \ldots, H(h_H^{-1}(Y_n - y)))$.

Clearly, after direct computations, we get

$$\widehat{F}(y|x) = \frac{\displaystyle\sum_{j=1}^{n} \Gamma_j K(h_K^{-1}\delta(X_j, x)) H({h_H}^{-1}(y - Y_j))}{\displaystyle\sum_{j=1}^{n} \Gamma_j K(h_K^{-1}\delta(X_j, x))}. \tag{1}$$

Here, $\Gamma_j = K^{-1}(h_K^{-1}\delta(X_j, x))\Big(\sum_{i=1}^{n} W_{ij}(x)\Big)$ with

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

and H is a distribution function, K is a kernel (h_K, h_H are a sequences of positive real numbers), the bi-functional $\beta(.,.)$ and $\delta(.,.)$ are defined from $\mathcal{F} \times \mathcal{F}$ into $\mathbb{R}$, where, $|\delta(x, z)| = d(x, z)$ and for all $x \in \mathcal{F}$, $\beta(x, x) = 0$.

Since the density function is defined in terms of derivative of the distribution function, we obtain from (1) the derivative with respect to y,

$$\widehat{f}(y|x) = \frac{\sum_{j=1}^{n} \Gamma_j K(h_K^{-1}\delta(X_j,x)) H'({h_H}^{-1}(y-Y_j))}{\sum_{j=1}^{n} h_H \Gamma_j K(h_K^{-1}\delta(X_j,x))},$$

with H' is the derivative of H.

3. Hypotheses

In the rest of this paper, we use C_1, C_2, C to denote strictly positive generic constants, N_x denote a fixed neighborhood of x (x is a fixed point in $\mathcal{F}$), N_y denoted by the neighborhood of y and we will use the notation $\phi(r_1, r_2) = \mathbb{P}(r_2 < \delta(X,x) < r_1)$ of the small ball probability function.

We introduce the following assumptions to establish our results.

(H1) For any $h_K > 0$,

$$\phi(h_K) = \phi(-h_K, h_K) > 0,$$

and for $\mu \geq 2$, $\quad C_1 n^{1-\mu} \leq \phi(h_K) \leq C_2 n^{(1-\mu)^{-1}}$.

(H2) (a) $\sup_{i \neq j} \mathbb{P}\left((X_i, X_j) \in B(x,h_K) \times B(x,h_K)\right) \leq C(\phi(h_K)^2) = \Phi(h_K)$.

(b) The α-mixing sequence $(X_i, Y_i)_{(i \in \mathbb{N})}$ satisfies

$$\sum_{i=1}^{\infty} i^s (\alpha(i))^{1/p} < \infty \text{ for } p > 0,\ s > (1/p).$$

(H3) For all $z \in \mathcal{F}$, the function $\beta(\cdot,\cdot)$ satisfies the conditions below,

$$C_1|\delta(x,z)| \leq |\beta(x,z)| \leq C_2|\delta(x,z)|,$$

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

where $B(x,r) = \{z \in \mathcal{F} : |\delta(z,x)| \leq r\}$.

(H4) (1) The kernel $K(.)$ is a positive, differentiable function which is supported within $(-1,1)$ satisfies

$$K^2(1) - \int_{-1}^{1} (K^2(u))' \chi_x(u) \mathrm{d}u > 0.$$

(2) The distribution $H(.)$ is positive and Lipschitzian function, where its first derivative $H'(.)$ are bounded and symmetric function, such that

$$\int H'(t)\mathrm{d}t = 1,\ \int |t|^{b_2} H'(t)\mathrm{d}t < \infty \text{ and } \int (H'(t))^2 \mathrm{d}t < \infty.$$

(H5) The bandwidths h_K and h_H are satisfied:

$$\lim_{n\to\infty} h_K = \lim_{n\to\infty} h_H = 0 \text{ and } \lim_{n\to\infty} \frac{\log n}{n h_H^j \phi(h_K)} = 0, \text{ for } j = 0, 1.$$

(H6) There exists a sequences of positive integers (r_n) and (p_n), such that $(r_n) \to \infty$, $(p_n) \to \infty$, and

(a) $p_n = o\left(\sqrt{n h_H \phi(h_K)}\right)$ and $\lim_{n\to\infty} \left(\sqrt{\frac{n}{h_H \phi(h_K)}}\right) \alpha(p_n) = 0$,

(b) $q_n p_n = o\left(\sqrt{n h_H \phi(h_K)}\right)$ and $q_n \lim_{n\to\infty} \left(\sqrt{\frac{n}{h_H \phi(h_K)}}\right) \alpha(p_n) = 0$,

where q_n is the largest integer, $(r_n + p_n) = O(n)$ and $q_n(r_n + p_n) = O(n)$.

Some discussion on the hypotheses

The assumption (H1) is the concentration property of the explanatory variable in small balls. The assumption (H2) is about the negligible of the variance term. The assumption (H3) is the same assumption in Barrientos-Marin et al. (2010). The assumptions (H4) and (H5) on the kernels $K(.), H(.)$ and the bandwidths h_K and h_H. Finally, the assumption (H6) is a technical hypothesis for our proofs.

4. Results

4.1. Asymptotic Normality

In this section, we get the asymptotic normality of $\widehat{h}(y|x)$ by the convergence in distribution $(\xrightarrow{\mathcal{D}})$, which the results below are the goal of this paper and there's a theorem and lemmas in the Appendix to help to get this asymptotic property.

Theorem 4.1.

Under hypotheses (H1)-(H6) and for any $x \in \Lambda$, such that $\Lambda = \{x \in \mathcal{F},\ f(y|x)(1 - F(y|x) \neq 0\}$, we have:

$$\left(\frac{n h_H \phi(h_K)}{V_{HK}^h(x,y)}\right)^{1/2} \left(\widehat{h}(y|x) - h(y|x) - B_n(x,y)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \text{ as } n \longrightarrow \infty,$$

with

$$V_{HK}^h(x,y) = \frac{h(y|x)}{(1 - F(y|x))} \left[\frac{\left(K^2(1) - \int_{-1}^1 (K^2(u))^{(1)} \chi_x(u) \mathrm{d}u\right)}{\left(K(1) - \int_{-1}^1 (K(u))^{(1)} \chi_x(u) \mathrm{d}u\right)^2}\right],$$

$$B_n(x,y) = \frac{(B_{f,H} - h(y|x) B_{F,H}) h_H^2 + (B_{f,K} - h(y|x) B_{F,K}) h_K^2}{1 - F(y|x)},$$

and

$$\begin{aligned}
B_{f,H}(x,y) &= \frac{1}{2}\frac{\partial^2 f^x(y)}{\partial y^2}\int t^2 H^{(1)}(t)\mathrm{d}t,\\
B_{f,K}(x,y) &= \frac{1}{2}\Psi_{0,1}^{(2)}(0)\left[\frac{\left(K(1)-\int_{-1}^{1}(u^2K(u))^{(1)}\chi_x(u)\mathrm{d}u\right)}{\left(K(1)-\int_{-1}^{1}K^{(1)}(u)\chi_x(u)\mathrm{d}u\right)}\right],\\
B_{F,H}(x,y) &= \frac{1}{2}\frac{\partial^2 F^x(y)}{\partial y^2}\int t^2 H^{(1)}(t)\mathrm{d}t,\\
B_{F,K}(x,y) &= \frac{1}{2}\Psi_{0,0}^{(2)}(0)\left[\frac{\left(K(1)-\int_{-1}^{1}(u^2K(u))^{(1)}\chi_x(u)\mathrm{d}u\right)}{\left(K(1)-\int_{-1}^{1}K^{(1)}(u)\chi_x(u)\mathrm{d}u\right)}\right],
\end{aligned}$$

where, for any $l \in \{0,2\}$ and $j = 0,1$, the functions

$$\psi_{l,j}(x,y) = \frac{\partial^l F^{x(j)}(y)}{\partial y^l} \text{ and } \Psi_{l,j}(s) = \mathbb{E}\left[\psi_{l,j}(X,y) - \psi_{l,j}(x,y) \mid \beta(x,X) = s\right].$$

Proof:

The proof of Theorem 4.1 is based on the following decomposition:

$$\begin{aligned}
\widehat{h}(y|x) - h(y|x) &= \frac{1}{\widehat{F}_D(x) - \widehat{F}_N(y|x)}\left[\widehat{f}_N(y|x) - \mathbb{E}[\widehat{f}_N(y|x)]\right]\\
&\quad + \frac{1}{\widehat{F}_D(x) - \widehat{F}_N(y|x)}\left[h(y|x)\left(\mathbb{E}[\widehat{F}_N(y|x)] - F(y|x)\right) + \left(\mathbb{E}[\widehat{f}_N(y|x)] - f(y|x)\right)\right]\\
&\quad + \frac{h(y|x)}{\widehat{F}_D(x) - \widehat{F}_N(y|x)}\left[1 - \mathbb{E}[\widehat{F}_N(y|x)] - \left(\widehat{F}_D(x) - \widehat{F}_N(y|x)\right)\right],
\end{aligned}$$

where

$$\begin{aligned}
\widehat{f}_N(y|x) &= \frac{1}{nh_H\mathbb{E}[\Gamma_1 K_1]}\sum_{j=1}^{n}\Gamma_j K_j H'_j,\\
\widehat{F}_D(x) &= \frac{1}{n\mathbb{E}[\Gamma_1 K_1]}\sum_{j=1}^{n}\Gamma_j K_j,\\
\widehat{F}_N(y|x) &= \frac{1}{n\mathbb{E}[\Gamma_1 K_1]}\sum_{j=1}^{n}\Gamma_j K_j H_j,
\end{aligned}$$

and $\Gamma_j = K^{-1}(h_K^{-1}\delta(X_j,x))\Big(\sum_{i=1}^{n} W_{ij}(x)\Big)$.

If we put under the hypotheses of Theorem 4.1, $\lim_{n\longrightarrow\infty}(h_H^2 + h_K^2)\sqrt{n\phi(h_K)} = 0$, we get

$$\left(\frac{nh_H\phi(h_K)}{V_{HK}^h(x,y)}\right)^{1/2}\left(\widehat{h}(y|x) - h(y|x)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \quad \text{as } n \to \infty. \tag{2}$$

Then, from these results, we finish the proof of the Theorem 4.1 by Lemma 4.1, 4.2 and Corollary 4.1. ■

Lemma 4.1.

Under hypotheses of Theorem 4.1, we have

$$\left(\frac{nh_H\phi(h_K)}{V_{HK}^f(x,y)}\right)^{1/2}\left(\widehat{f}_N(y|x)-\mathbb{E}[\widehat{f}_N(y|x)]\right)\xrightarrow{\mathcal{D}}\mathcal{N}(0,1),\quad \text{as } n\to\infty,$$

where $V_{HK}^f(x,y)=\frac{A_2}{A_1^2}f(y|x)\int(H^{'}(t))^2\mathrm{d}t$, and $A_k=K^k(1)-\int_{-1}^1(K^k(t))'\chi_x(t)\mathrm{d}t$, for $k=1,2$.

Lemma 4.2.

Under the conditions of Theorem 4.1, we obtain

$$\left(\frac{nh_H\phi(h_k)}{V_{HK}^h(x,y)}\right)^{1/2}(\widehat{F}_D(x)-\widehat{F}_N(y|x)-1+\mathbb{E}[\widehat{F}_N(y|x)])=o_{\mathbb{P}}(1).$$

Corollary 4.1.

Under hypothesis of Theorem 4.1, we obtain

$$\widehat{F}_D(x)-\widehat{F}_N(y|x)\xrightarrow{\mathbb{P}}1-F(y|x).$$

4.2. Application: Confidence interval

We know that the confidence interval gives us a range of plausible values for some unknown value based on results from a sample. So, to construct a confidence interval for the true value of $h(y|x)$, we need to estimate the variance term $V_{HK}^h(x,y)$ and the bias term $B_n(x,y)$ by the empirical estimators, but from (4.1), we ignore the bias term $B_n(x,y)$.

First, we estimate the function $\phi(h_K)$ by

$$\widehat{\phi}(h_K)=\frac{\sharp\{i:|\delta(x,X_i)\le h_K|\}}{n},$$

and $\sharp\{.\}$ is the cardinal number and about the functions $\delta(.,.)$ and $\beta(.,.)$ (see Barrientos-Marin et al. (2010)). Then, we give the $V_{HK}^h(x,y)$ as follows,

$$\widehat{V_{HK}^h}(x,y)=\left(\frac{\widehat{h}(y|x)}{(1-\widehat{F}(y|x))}\frac{\widehat{A_2}}{(\widehat{A_1})^2}\right)^{1/2};$$

observe that from Ferraty and Vieu (2006),

$$\frac{1}{\phi(h_K)}\mathbb{E}\left[K\left(\frac{|\delta(X_1,x)|}{h_K}\right)\right]\longrightarrow A_1 \text{ and } \frac{1}{\phi(h_K)}\mathbb{E}\left[K\left(\frac{|\delta(X_1,x)|}{h_K}\right)^2\right]\longrightarrow A_2^2,$$

so, we estimate empirically these terms by

$$\widehat{A_1}=\frac{1}{n\widehat{\phi}(h_K)}\sum_{i=1}^n K\left(\frac{|\delta(X_i,x)|}{h_K}\right) \text{ and } \widehat{A_2}=\frac{1}{n\widehat{\phi}(h_K)}\sum_{i=1}^n K\left(\frac{|\delta(X_i,x)|}{h_K}\right)^2.$$

Finally, the asymptotic $(1-\lambda)$ confidence interval of $h(y|x)$ with $z_{\lambda/2}$ is the $\lambda/2$ quantile of the standard normal distribution, given by:

$$\left[\widehat{h}(y|x) - z_{\lambda/2}\sqrt{\frac{\widehat{V_{HK}^{h}}(x,y)}{n\widehat{\phi}(h_K)}},\ \widehat{h}(y|x) + z_{\lambda/2}\sqrt{\frac{\widehat{V_{HK}^{h}}(x,y)}{n\widehat{\phi}(h_K)}}\right].$$

5. Simulation

Our main purpose of this section is to illustrate the performance and the superiority of the local linear method. More precisely, we compare the mean square error (MSE) of the local linear approach (L.L.), studied here, over the kernel (L.C.) one when data are of functional kind, which are defined as

$$\overline{h}_{L.L.}^{x}(y)) = \frac{h_H^{-1}\sum_{i,j=1}^{n} W_{ij}H'(h_H^{-1}(y-Y_j))}{\sum_{i,j=1}^{n} W_{ij} - \sum_{i,j=1}^{n} W_{ij}H(h_H^{-1}(y-Y_j))}, \quad (3)$$

$$\overline{h}_{L.C.}^{x}(y) = \frac{h_H^{-1}\sum_{i=1}^{n} K(h_K^{-1}d(x,X_i))H'(h_H^{-1}(Y_i-y))}{\sum_{i=1}^{n} K(h_K^{-1}d(x,X_i)) - \sum_{i=1}^{n} K(h_K^{-1}d(x,X_i))H(h_H^{-1}(Y_i-y))}. \quad (4)$$

For this aim, we first present a functional dependent processes generated in the following way:

$$X_i(t) = 1 - \sin(\eta_i t), \quad i = 1, \ldots, 300; \quad t \in \left[0, \frac{\pi}{3}\right], \quad (5)$$

where t takes 100 equispaced values in $\left[0, \frac{\pi}{3}\right]$ and $\eta_i = \frac{1}{3}\eta_{i-1} + \xi_i, \xi_i$ are i.i.d. $\sim \mathcal{N}(0,1)$, and are independent from η_i, which is generated independently by $\eta_0 \sim \mathcal{N}(0,1)$. For simplicity, Figure 1 presents a sample of $n = 100$ of such curves $X_i(t)$. Second, the scalar response variables are obtained from the model $Y_i = r(X_i) + \epsilon_i$, where r is the nonlinear regression operator with $r(X) = 7\left(\int_0^{\frac{\pi}{3}} (X'(t))\, dt\right)^2$ and $\epsilon_i \sim \mathcal{N}(0, 0.075)$.

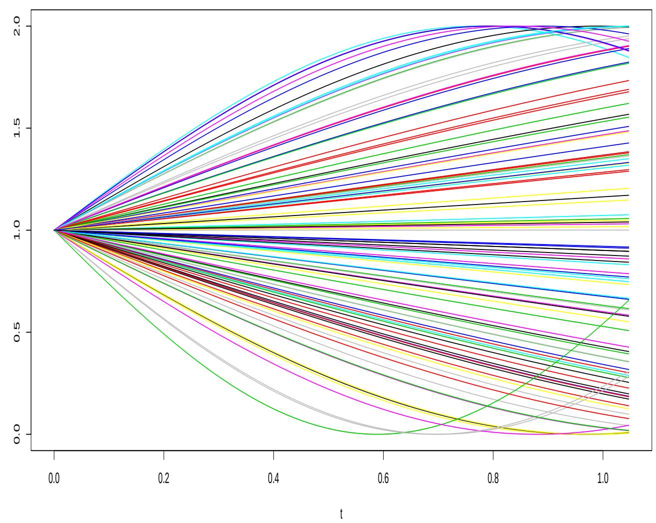

Figure 1. A sample of 100 curves X for $t \in \left[0, \frac{\pi}{3}\right]$

Concerning the smoothing parameters (h_H and h_K)

The bandwidth parameters is very crucial in nonparametric estimation because this parameters intervenes in all the asymptotic properties as in the case which we are studying, in particular in the improvement of the rate of convergence.

In this application we using the cross-validation method (CV) to select the bandwidths h_H and h_K. By the similar technique introduced by Rachdi et al. (2014) for the quadratic error in the local linear estimation of the conditional density for functional data, we consider the minimization of the following criterion

$$\frac{1}{n}\sum_{i=1}^{n} W_1(X_i)\int \left[\overline{h}_{(-i)}(y|X_i)\right]^2 W_2(y)dy - \frac{2}{n}\sum_{i=1}^{n}\left[\overline{h}_{(-i)}(Y_i|X_i)\right] W_1(X_i)W_2(Y_i),$$

where

$$\overline{h}_{(-k)}(y|X_k) = \frac{h_H^{-1}\sum_{j=1,j\neq k}^{n}\sum_{i=1,i\neq k}^{n} W_{ij}(X_k)H'(h_H^{-1}(y-Y_j))}{\sum_{j=1,j\neq k}^{n}\sum_{i=1,i\neq k}^{n} W_{ij}(X_k) - \sum_{j=1,j\neq k}^{n}\sum_{i=1,i\neq k}^{n} W_{ij}(X_k)H(h_H^{-1}(y-Y_j))}. \tag{6}$$

with $\overline{h}_{(-k)}$ is the estimator of h without using the kth observation $(k = 1, ..., n)$, this estimator is called the leave-one-out curve estimator and W_1 (respectively, W_2) is some positive weight function. In our simulation study, we take $W_1(y) = 1$ and $W_2(y) = \mathbb{I}_{[0.9\times min_{i=1,..,n}(Y_i);1.1\times min_{i=1,..,n}(Y_i)]}$, (see, for instance, Demongot et al. (2014) for more discussions on the choice about weight function).

On the choice of the locating functions (δ **and** β**).**
The shape of the curves has great consideration in the choice of bi-functional operators δ and β. For instance, if the functional data are smooth curves, we may use the following family of locating functions:

$$\beta(x_1, x_2) = \int_0^{\frac{\pi}{3}} \theta(t)(x_1^{(s)}(t) - x_2^{(s)}(t))dt,$$

and

$$\delta(x_1, x_2) = \sqrt{\int_0^{\frac{\pi}{3}} (x_1^{(s)}(t) - x_2^{(s)}(t))^2 dt},$$

where $x^{(s)}(t)$ denoting the sth derivative of the curve $x(t)$ and θ is the eigenfunction of the empirical covariance operator

$$\frac{1}{|\Lambda|}\sum_{i\in\Lambda}(X_i^{(s)}(t) - \overline{X^{(s)}(t)})^t(X_i^{(s)}(t) - \overline{X^{(s)}(t)}), \ \text{ where } \ \overline{X^{(s)}(t)} = \frac{1}{|\Lambda|}\sum_{i\in\Lambda} X_i^{(s)}(t),$$

associated with the q-greatest eigenvalue.

In this simulation study, we take the following parameters:

$$s = 2, \ q = 5, \ \Lambda = \{1, ..., 100\} \text{ and } |\Lambda| = 100.$$

In the following graphs, the covariance operator for $\Lambda = \{1, ..., 100\}$ gives the discretization of the eigenfunctions $\theta_i(t)$ (the eigenfunction is presented by a continuous curve). First, the five eigenfunctions $\theta_1(t)$, $\theta_2(t)$, $\theta_3(t)$, $\theta_4(t)$ and $\theta_5(t)$, are represented in Figure 2. Next, we plot twenty eigenfunctions $\theta_1(t), ..., \theta_{20}(t)$ in Figure 3. Finally, all the 100 eigenfunctions are represented in Figure 4.

Finally, we choose a quadratic kernel K on $[-1, 1]$ and take $K = H'$. To illustrate the performance of our estimator, we proceed with the following algorithm.

- Step 1. We generate $n = 100$ independent replications of $(X_i, Y_i)_{i=1,...,n}$.
- Step 2. We divide our observations into two subsets:
  - $(X_i, Y_i)_{i=1,..,80}$, training sample.
  - $(X_j, Y_j)_{j=81,..,100}$, test sample.
- Step 3. We calculate the two estimators by using the learning sample and we find the local linear ($\overline{h}_{L.L.}$) and the local constant ($\overline{h}_{L.C.}$) estimators of the conditional hazard function.

- Step 4. We present our results by plotting the boxplot of the prediction error represented in Figure 6 and we compute the empirical mean square error with

  - $MSE(\overline{h}_{L.L.}) = \frac{1}{20}\sum_{i=1}^{20}(Y_i - \overline{\lambda}_{L.L.}(Y_i))^2,$
  - $MSE(\overline{h}_{L.C.}) = \frac{1}{20}\sum_{i=1}^{20}(Y_i - \overline{\lambda}_{L.C.}(Y_i))^2.$

  Finally, we find that the method based on the local linear estimation is much better and more efficient than the kernel method. This is confirmed by the mean squared error $MSE(\overline{h}_{L.L.}) = 0.16064$ whereas $MSE(\overline{h}_{L.C.}) = 0.26352$. This conclusion shows the good performance of our approach. Figure 5 gives an idea on the accuracy of the predictions corresponding to one run while Figure 6 displays the distribution of the mean squared error of prediction.

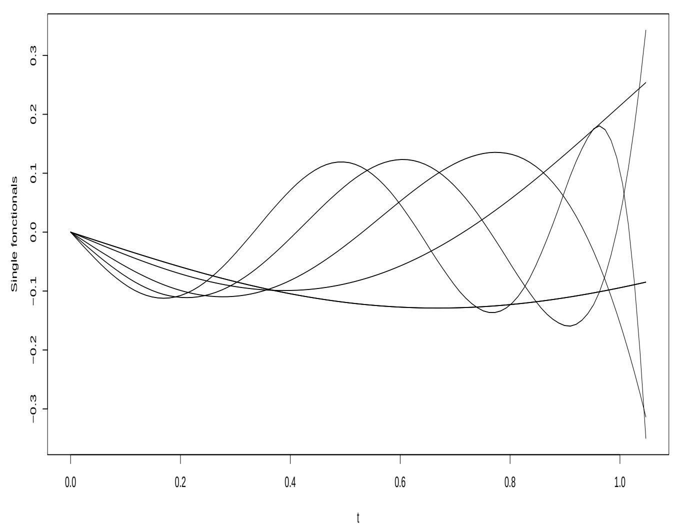

Figure 2. The curves $\theta_i(t_j),\ t_j \in [0,1],\ i = 1, ..., 5$

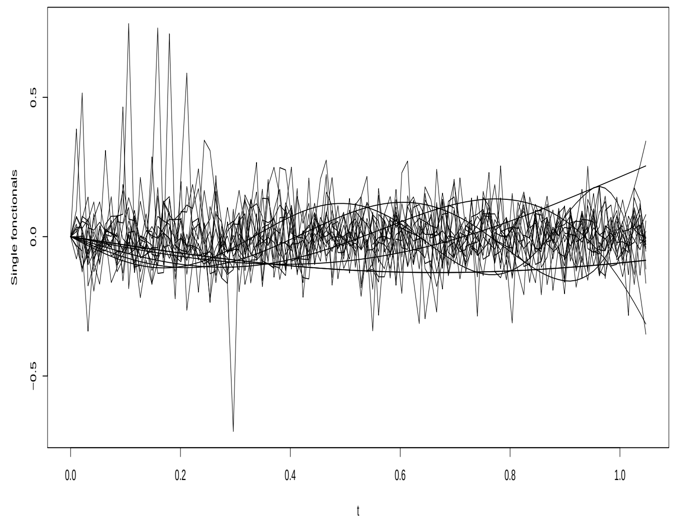

Figure 3. The curves $\theta_{i=1,..,20}(t_j),\ t_{j=1,..,100} \in [0,1],\ i = 1,..,20$

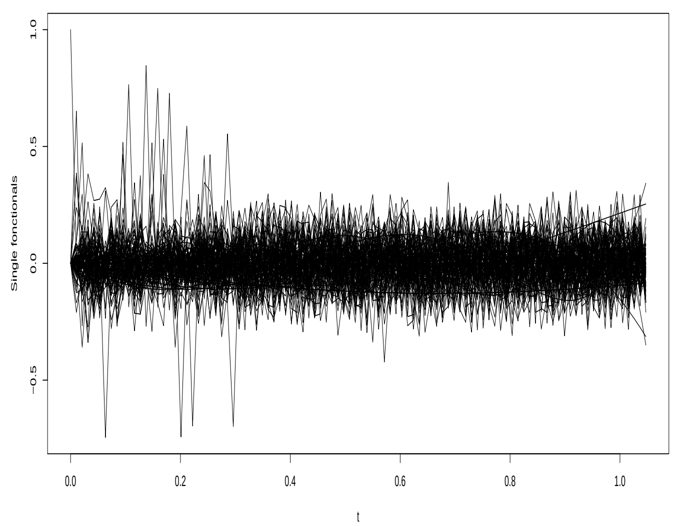

Figure 4. The curves $\theta_i(t_j),\ t_j \in [0,1],\ i = 1,..,100$

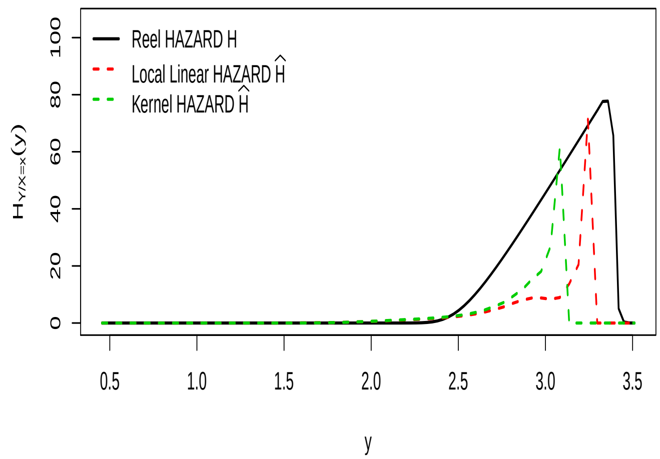

Figure 5. Comparison of the Point at high risk between the local linear estimator and the local constant

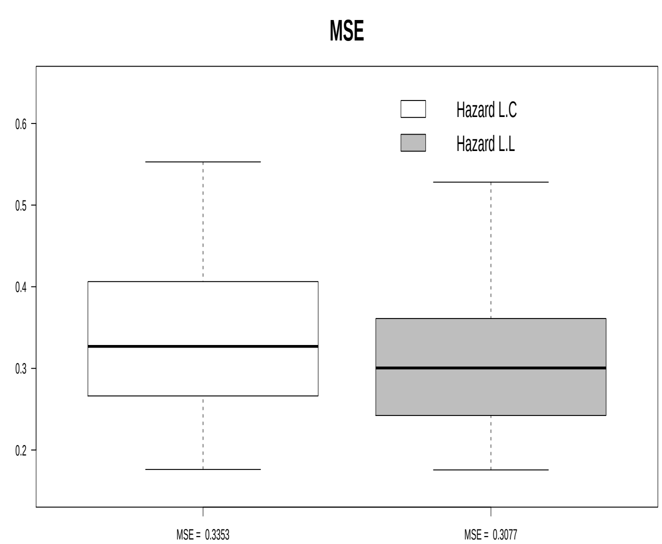

Figure 6. Comparison of the distribution of MSE between the local linear estimator and the local constant

6. Conclusion

As a general conclusion, we have established the asymptotic normality property of the local linear estimate of the conditional hazard function in α-mixing data framework. Our theoretical analysis shows that our estimator has excellent asymptotic characteristics. For practical purposes, we can say that the local linear estimate approach outperforms the kernel method in terms of squared error MSE. As a result, we can provide and estimate confidence bands. The k-NN method is a smoothing method that includes an adaptive estimator. The very important feature of this method is that it allows the construction of a neighbourhood adapted to the local structure of the data. So, it would be also of interest to study the asymptotic properties of the k-NN estimator of the conditional hazard function. This will be considered in future works.

Acknowledgment:

The authors would like to thank the Editor-in-Chief and the anonymous reviewers for their valuable comments that helped to improve the quality of this paper.

REFERENCES

Baillo, A. and Grané, A. (2009). Local linear regression for functional predictor and scalar response, Journal of Multivariate Analysis, Vol. 100, No. 1, pp. 102-111.

Barrientos-Marin, J., Ferraty, F. and Vieu, P. (2010). Locally modelled regression and functional data, Journal of Nonparametric Statistics, Vol. 22, No. 5, pp. 617-632.

Berlinet, A., Elamine, A. and Mas, A. (2011). Local linear regression for functional data, Annals of the Institute of Statistical Mathematics, Vol. 63, No. 5, pp. 1047-1075.

Bouanani, O., Laksaci, A., Rachdi, M. and Rahmani, S. (2019). Asymptotic normality of some conditional nonparametric functional parameters in high-dimensional statistics, Behaviormetrika, Vol. 46, No. 1, pp. 199-233.

Bouanani, O., Rahmani, S. and Ait-Hennani, L. (2019). Local linear conditional cumulative distribution function with mixing data, Arabian Journal of Mathematics, Vol. 9, Issue 2, pp. 289307.

Demongeot, J., Laksaci, A., Rachdi, M. and Rahmani, S. (2014). On the local linear modelization of the conditional distribution for functional data, Sankhya A, Vol. 76, No. 2, pp. 328-355.

El Methni, M. and Rachdi, M. (2011). Local weighted average estimation of the regression operator for functional data, Communications in Statistics-Theory and Methods, Vol. 40, Issue 17, pp. 3141-3153.

Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis: Theory and Practice*, Springer, New York NY.

Laksaci, A. and Mechab, B. (2010). Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle: Cas des données spatiales, Rev. Roumaine Math. Pures Appl, Vol. 55, No. 1, pp. 35-51.

Masry, E. (1986). Recursive probability density estimation for weakly dependent stationary processes, IEEE Transactions on Information Theory, Vol. 32, No. 2, pp. 254-267.

Masry, E. (2005). Nonparametric regression estimation for dependent functional data: Asymptotic normality, Stochastic Processes and their Applications, Vol. 115, No. 1, pp. 155-177.

Massim, I. and Mechab, B. (2016). Local linear estimation of the conditional hazard function, International Journal of Statistics & Economics, Vol. 17, Issue 3, pp. 1-11.

Merouan, T., Mechab, B. and Massim, I. (2018). Quadratic error of the conditional hazard function in the local linear estimation for functional data, Afrika Statistika, Vol. 13, No. 3, pp. 1759-1777.

Messaci, F., Nemouchi, N., Ouassou, I. and Rachdi, M. (2015). Local polynomial modelling of the conditional quantile for functional data, Statistical Methods & Applications, Vol. 24, Issue 4, pp. 597-622.

Quintela-del-Rio, A. (2006). Nonparametric estimation of the maximum hazard under dependence conditions, Statistics & Probability Letters, Vol. 76, Issue 11, pp. 1117-1124.

Quintela-del-Rio, A. (2008). Hazard function given a functional variable: Non-parametric estimation under strong mixing conditions, Journal of Nonparametric Statistics, Vol. 20, Issue 5, pp. 413-430.

Rachdi, M., Laksaci, A., Demongeot, J., Abdali, A., & Madani, F. (2014). Theoretical and practical aspects of the quadratic error in the local linear estimation of the conditional density for functional data, Computational Statistics & Data Analysis, Vol. 73, pp. 53-68.

Rachdi, M. and Vieu, P. (2007). Nonparametric regression for functional data: Automatic smoothing parameter selection, J. Stat. Plan. Inference, Vol. 137, Issue 9, pp. 2784-2801.

Ramsay, J.O. and Silverman, B.W. (2002). *Applied Functional Data Analysis: Methods and Case Studies*, Springer, New York NY.

Volkonskii, V.A. and Rozanov, Y.A. (1959). Some limit theorems for random functions. I, Theory of Probability & Its Applications, Vol. 4, Issue 2, pp. 178-197.

Watson, G. S. and Leadbetter, M. R. (1964). Hazard analysis. I, Biometrika, Vol. 51, No. 1/2, pp. 175-184.

Zhou, Z. and Lin, Z. (2016). Asymptotic normality of locally modelled regression estimator for functional data, Journal of Nonparametric Statistics, Vol. 28, No. 1, pp. 116-131.

Appendix

In what follows, we denote for $(i, j = 1, \ldots, n)$,

$$K_i = K(h_K^{-1}\delta(X_i, x)),\ H_j = H(h_H^{-1}(Y_j - y)),\ \beta_i = \beta(.,.).$$

The next lemma is for the proof of Lemma 4.1.

Lemma 6.1.

We have

(1) $n\text{Var}(N_{2,1}) \xrightarrow{n\to+\infty} \frac{A_2}{A_1^2} f(y|x) \int (H^{'}(t))^2 \mathrm{d}t.$

(2) $\left(\frac{n^2\phi(h_K)\mathbb{E}^2[(\beta_1^2 K_1)]}{h_H \mathbb{E}^2[\Gamma_1 K_1]} \sum_{1\le i<j\le n} Cov\left(K_i(H_i' - h_H f(y|x)), K_j(H_j' - h_H f(y|x))\right) \right) \xrightarrow{n\to+\infty} 0.$

(3) $\mathbb{E}[(N_{1,i}-1)^2] \xrightarrow{n\to+\infty} 0.$

(4) $\mathbb{E}[N_{2,j}^2] \xrightarrow{n\to+\infty} V_{HK}^f(x,y).$

Proof:

(1) We have

$$n\text{Var}(N_{2,1}) = \frac{n^2\phi(h_K)\mathbb{E}^2[(\beta_1^2 K_1)]}{h_H \mathbb{E}^2[\Gamma_1 K_1]} \mathbb{E}\Big[\Big(K_1(H_1' - h_H f(y|x)) - \mathbb{E}[K_1(H_1' - h_H f(y|x))\Big)^2\Big].$$

Then, from the Lemma 2 in Bouanani et al. (2018), we get

$$\mathbb{E}[K_1(H_1' - h_H f(y|x))] = \mathbb{E}[K_1(\mathbb{E}[H_1'|X_1] - h_H f(y|x))] \xrightarrow{n\to+\infty} 0. \tag{7}$$

Moreover, we still have to prove the remaining term,

$$\begin{aligned} n\text{Var}(N_{2,1}) &= \frac{n^2\phi(h_K)\mathbb{E}^2[(\beta_1^2 K_1)]}{h_H \mathbb{E}^2[\Gamma_1 K_1]} \mathbb{E}\Big[K_1^2 \mathbb{E}[(H_1' - h_H f(y|x))^2 | X_1]\Big] \\ &= \frac{n^2\phi(h_K)\mathbb{E}^2[(\beta_1^2 K_1)]}{h_H \mathbb{E}^2[\Gamma_1 K_1]} \Big(\mathbb{E}[\text{Var}(H_1'|X_1)K_1^2] + \big(\mathbb{E}[\mathbb{E}[H_1'|X_1] - h_H f(y|x)]^2 K_1^2\big)\Big). \end{aligned}$$

Again, we use (7) to get

$$n\text{Var}(N_{2,1}) = \frac{n^2\phi(h_K)\mathbb{E}^2[(\beta_1^2 K_1)]}{h_H \mathbb{E}^2[\Gamma_1 K_1]} \Big(\mathbb{E}[\text{Var}(H_1'|X_1)K_1^2]\Big),$$

and we have from Lemma 2 in Bouanani et al. (2018) that

$$\mathbb{E}[\text{Var}(H_1'|X_1)K_1^2] \longrightarrow A_2\phi(h_K)h_H f(y|x) \int (H'(t))^2 \mathrm{d}t \ \text{ as } n\to\infty.$$

However, by Lemma 1 in Zhou and Lin (2016):

$$\begin{aligned} n\text{Var}(N_{2,1}) &= \frac{n^2\phi(h_K)(N(1,2)h_K^2\phi(h_K)^2)^2}{h_H((n-1)N(1,2)A_1 h_K^2\phi(h_K))^2} h_H A_2 \phi(h_K) f(y|x) \int (H'(t))^2 \mathrm{d}t \\ &= \frac{n^2}{(n-1)^2 A_1^2} A_2 f(y|x) \int (H'(t))^2 \mathrm{d}t. \end{aligned}$$

Then,

$$n\text{Var}(N_{2,1}) \longrightarrow \frac{A_2}{A_1^2} f(y|x) \int (H'(t))^2 \mathrm{d}t \ \text{ as } n \longrightarrow \infty.$$

(2) The technique of Masry (1986) defines the sets Δ_1 and Δ_2 as follows,

$$\begin{aligned} \Delta_1 &= \{(i,j) \in \{1,2,....,n\},\ 1 \le| i-j |\le s_n\}, \\ \Delta_2 &= \{(i,j) \in \{1,2,....,n\},\ s_n + 1 \le| i-j |\le n-1\}, \end{aligned}$$

and s_n is a sequence of integers such that $s_n \longrightarrow +\infty$, and we denote by

$$G_i = K_i(H_i' - h_H f(y|x)) \ \text{ and } \ G_j = K_j(H_j' - h_H f(y|x)).$$

Then,

$$\begin{aligned} Cov(N_{2,i}, N_{2,j}) &= \frac{n\phi(h_K)\mathbb{E}^2[(\beta_1^2 K_1)]}{h_H \mathbb{E}^2[\Gamma_1 K_1]} \sum_{\Delta_1} Cov(G_i, G_j) \\ &\quad + \frac{n\phi(h_K)\mathbb{E}^2[(\beta_1^2 K_1)]}{h_H \mathbb{E}^2[\Gamma_1 K_1]} \sum_{\Delta_2} Cov(G_i, G_j) \\ &= P_1 + P_2. \end{aligned}$$

First, we calculate the sum over the set Δ_1,

$$\begin{aligned} Cov(G_i, G_j) = \mathbb{E}\left[K_i K_j \mathbb{E}\left[((H_i' - h_H f(y|x))(H_i' - h_H f(y|x))|(X_i, Y_j)\right]\right] \\ - \mathbb{E}^2\left[K_1(H_1' - h_H f(y|x))\right]. \end{aligned}$$

After (H2)(a) and (H4)(2), the following inequality is obtained,

$$\begin{aligned} \mid Cov(G_i, G_j) \mid &\leq \mathbb{E}(K_i K_j) + \mathbb{E}^2(K_1) \\ &\leq C\mathbb{P}((X_i, X_j) \in B(x, h_K) \times B(x, h_K)) + A_1^2 \phi^2(h_K) \\ &\leq C\Phi(h_K) + A_1^2 \phi^2(h_K). \end{aligned}$$

Then, by Lemma 1 in Zhou and Lin (2016),

$$\mid P_1 \mid \leq \frac{Cn^2 \phi(h_K) s_n}{(n-1)^2 A_1^2 h_H} \left(\left(\frac{\Phi(h_K)}{\phi^2(h_K)} \right) + A_1^2 \right),$$

and taking $s_n = (\phi(h_K)\log n)^{-(1/p)(1/z)}$ and from (H1), we obtain $P_1 \longrightarrow o(1)$ as $n \longrightarrow +\infty$.

Second, computing the sum over the set Δ_2 by using the proposition A.10 (ii) in Ferraty and Vieu (2006), we get:

$$\begin{aligned} \sum_{\Delta_2} \mid Cov(G_i, G_j) \mid &\leq \sum_{\Delta_2} C\left(\alpha \mid j-i \mid\right)^{\frac{1}{p}} \left(\mathbb{E} \mid G_i \mid^q\right)^{\frac{1}{q}} \left(\mathbb{E} \mid G_j \mid^r\right)^{\frac{1}{r}} \\ &\leq \sum_{\Delta_2} C\left(\alpha \mid j-i \mid\right)^{\frac{1}{p}} \mathbb{E}\left(\mid K_i \mid^q \mathbb{E}\left(\mid H_i'(y) - h_H f(y|x) \mid^q |X_1\right)\right) \\ &\qquad \mathbb{E}\left(\mid K_i \mid^r \mathbb{E}\left(\mid H_i'(y) - h_H f(y|x) \mid^r |X_1\right)\right) \\ &\leq \sum_{|k|>s_n} k^\delta (\alpha(\mid k \mid))^{\frac{1}{p}} C(\phi(h_K))^{\frac{1}{q}} C_1(\phi(h_K))^{\frac{1}{r}}, \end{aligned}$$

so, by Lemma 1 in Zhou and Lin (2016),

$$\mid P_2 \mid \leq \frac{Cn^2 (\phi(h_K))^{\frac{1}{q}+\frac{1}{r}=1-\frac{1}{p}}}{h_H(n-1)^2 A_1^2 \phi(h_K)(s_n)^z} \sum_{|k|>s_n} k^\delta (\alpha(\mid k \mid))^{\frac{1}{p}}.$$

By the assumption (H3) and with the same choice of s_n as before, we obtain

$$P_2 \longrightarrow o(1) \text{ as } n \longrightarrow +\infty.$$

Therefore, we prove the asymptotic normality of the conditional density function,

$$N_{2,j} - \mathbb{E}[N_{2,j}] = \frac{\sqrt{nh_H\phi(h_K)}\mathbb{E}(\beta_1^2K_1)}{h_H\mathbb{E}(\Gamma_1K_1)}\sum_{j=1}^{n}(K_j(H'_j - h_Hf(y|x) - \mathbb{E}(H'_j - h_Hf(y|x)))$$
$$= \frac{\sum_{j=1}^{n}(V_j - \mathbb{E}(V_j))}{\sqrt{n}} = \Theta_n,$$

where,

$$V_j = \frac{n\sqrt{h_H\phi(h_K)}\mathbb{E}(\beta_1^2K_1)}{h_H\mathbb{E}(\Gamma_1K_1)}\left(K_j(H'_j - h_Hf(y|x)) - \mathbb{E}[K_j(H'_j - h_Hf(y|x))]\right).$$

We know that from (18) the term $\sum_{1\leq i<j\leq n} Cov(V_i, V_j)$ goes to 0 as n tends to infinity. Then, it require to show the following assertion:

$$N_{2,j} - \mathbb{E}[N_{2,j}] \xrightarrow{\mathcal{D}} \mathcal{N}(0, V_{HK}^f(x,y)).$$

To prove this we will use the classic big and small-block decomposition to overcome the dependence problem. We divided the set $\{1, 2, 3...., n\}$ into $2k_n + 1$ subsets with large blocks of size r_n and small blocks of size p_n and put $k_n = \left\lfloor \frac{n}{r_n+p_n} \right\rfloor$ and the condition (H6) allows as to define the large block size as follows:

$$r_n = \left\lfloor \left(\frac{nh_H\phi(h_K)}{q_n}\right)^{\frac{1}{2}} \right\rfloor.$$

Furthermore, under the same condition and simple algebra:

$$\lim_{n\to+\infty}\frac{p_n}{r_n} = 0, \quad \lim_{n\to+\infty}\frac{r_n}{n} = 0, \quad \lim_{n\to+\infty}\frac{r_n}{\sqrt{nh_H\phi(h_K)}} = 0.$$

Now, let Υ_j, Υ'_j and M_n be defined as follows:

$$\Upsilon_j = \sum_{i=j(r+p)+1}^{j(r+p)+r} V_i, \Upsilon'_j = \sum_{i=j(r+p)+r+1}^{(j+1)(r+p)} V_i, M_n = \sum_{i=k(r+p)+1}^{n} V_i.$$

It is clear that

$$\Theta_n = \sum_{j=0}^{k-1}\frac{\Upsilon_j}{\sqrt{n}} + \sum_{j=0}^{k-1}\frac{\Upsilon'_j}{\sqrt{n}} + \frac{M_n}{\sqrt{n}}$$
$$= S_{1,n} + S_{2,n} + S_{3,n}.$$

Now, these results permit us to examine the following,

$$(S_{2,n} + S_{3,n}) \xrightarrow{\mathbb{P}} 0, \tag{8}$$

and

$$S_{1,n} \xrightarrow{\mathcal{D}} \mathcal{N}(0, V_{HK}^f(x,y)). \tag{9}$$

In regard to (8).
It suffices to show that by Markov's Inequality to write $\forall \varepsilon > 0$:

$$\mathbb{P}(\mid S_{2,n} \mid > \varepsilon) \leq \frac{\mathbb{E}^2(S_{2,n})}{\varepsilon^2} \xrightarrow{n \to +\infty} 0, \tag{10}$$

$$\mathbb{P}(\mid S_{3,n} \mid > \varepsilon) \leq \frac{\mathbb{E}^2(S_{3,n})}{\varepsilon^2} \xrightarrow{n \to +\infty} 0. \tag{11}$$

- To prove (10), by the second-order stationary, we have

$$\begin{aligned}
\mathbb{E}^2(S_{2,n}) &= \frac{1}{n}\left(\sum_{j=0}^{k-1} Var(\Upsilon'_j) + 2\sum_{0\leq i<j\leq k-1} Cov(\Upsilon'_i, \Upsilon'_j)\right) \\
&= \frac{1}{n}\left(\sum_{j=0}^{k-1} Var\left(\sum_{i=j(r+p)+r+1}^{(j+1)(r+p)} V_i\right) + 2\sum_{0\leq i<j\leq k-1} Cov(\Upsilon'_i, \Upsilon'_j)\right) \\
&\leq \frac{1}{n}\left(\sum_{j=0}^{k-1} p_n Var(V_1) + 2\sum_{j=0}^{k-1}\sum_{i\neq j}^{p_n} Cov(V_i, V_j) + 2\sum_{|i-j|>0}^{k-1} Cov(\Upsilon'_i, \Upsilon'_j)\right) \\
&\leq \underbrace{\frac{1}{n} kp_n Var(V_1) + \frac{2}{n}\sum_{j=0}^{k-1}\sum_{i\neq j}^{v_n} Cov(V_i, V_j)}_{T_1} \\
&\quad \underbrace{+ \frac{1}{n} 2\sum_{|i-j|>0}^{k-1}\sum_{l=1}^{p_n}\sum_{j'=1}^{p_n} Cov(V_{(i(r_n+p_n)+r_n)+l}, V_{(i(r_n+p_n)+r_n)+j'})}_{T_2},
\end{aligned}$$

then, from (H6), we obtain $\frac{T_1}{n} \longrightarrow 0,$ as $n \longrightarrow +\infty$.

Now, for T_2, since $i \neq j$, we have $\mid (i(r_n + p_n) + r_n) - (j(r_n + p_n) + r_n) + l - j' \mid \geq p_n$:

$$\frac{T_2}{n} \leq \frac{2}{n}\sum_{i,j=1}^{n} \mid Cov(V_i, V_j) \mid,$$

where $\mid i - j \mid \geq r_n$ and this latter extend to $\frac{T_2}{n} \longrightarrow 0,$ as $n \longrightarrow +\infty$.

- To prove (11), from that $n - k_n(r_n + p_n) < r_n + p_n$, we have the following:

$$\begin{aligned}
\mathbb{E}^2(S_{3,n}) &= \frac{1}{n} Var(M_n) \\
&= \frac{n - k_n(r_n + p_n)}{n} Var(V_1) + \frac{2}{n}\sum_{i\neq j}^{n-k_n(r_n+p_n)} Cov(V_i, V_j) \\
&\leq \frac{r_n + p_n}{n} Var(V_1) + o(1),
\end{aligned}$$

and, again by the condition (H6), we complete the proof of (11).

- **In regard to** (9).
  We have from lemma in Volkonskii and Rozanov (1959) as $n \longrightarrow +\infty$:
$$\left|\mathbb{E}\left[\exp\left(\frac{S_{1,n}}{\sqrt{n}}it\right)\right] - \prod_{j=0}^{k-1}\mathbb{E}\left[\exp\left(\frac{\Upsilon_j}{\sqrt{n}}it\right)\right]\right| \longrightarrow 0,$$
  then,
$$Var(S_{1,n}) = \frac{kr_n}{n}Var(Z_1),$$
  from assumption (H6), we have $\frac{kr_n}{n} \longrightarrow 1$ as $n \longrightarrow +\infty$. Thus,
$$\frac{1}{n}\sum_{j=0}^{k-1}\mathbb{E}\left[\Upsilon_j^2\right] \longrightarrow V_{HK}^f(x,y) \text{ as } n \longrightarrow +\infty,$$
  and by the central limit theorem (Lindeberg's version) on Υ_j and because the set $\{|\ \Upsilon_j\ |> \varepsilon\sqrt{nV_{HK}^f(x,y)}\}$ is empty due to $\left|\frac{\Upsilon_j}{n}\right| \leq \frac{r_n}{n}|V_1| \xrightarrow{n\to+\infty} 0$, we obtain
$$\forall\varepsilon > 0, \frac{1}{n}\sum_{j=0}^{k-1}\mathbb{E}\left[\Upsilon_j^2 1_{|\Upsilon_j|>\varepsilon\sqrt{nV_{HK}^f(x,y)}}\right] \longrightarrow 0 \text{ as } n \longrightarrow +\infty. \tag{12}$$
  Finally, the proof of (9) is terminated.

(3) We have
$$\begin{aligned}\mathbb{E}[(N_{1,i}-1)^2] &= nVar(N_{1,1}) + 2Cov(N_{1,i}, N_{1,i'})\\ &= \frac{nVar(\beta_1^2K_1)}{n^2\mathbb{E}[\beta_1^2K_1]} + 2\frac{1}{n^2\mathbb{E}[\beta_1^2K_1]}\sum_{1\leq i\neq i'\leq n}Cov(\beta_i^2K_i, \beta_{i'}^2K_{i'})\\ &= O\left(\frac{1}{n\phi(h_K)}\right) + 2\frac{1}{n^2\mathbb{E}[\beta_1^2K_1]}\sum_{1\leq i\neq i'\leq n}Cov(\beta_i^2K_i, \beta_{i'}^2K_{i'}).\end{aligned}$$
According to the same procedure in (18), we split the sum into two separate summations over the sets Δ_1 and Δ_2, and $s_n \longrightarrow \infty$ as $n \longrightarrow \infty$:
$$\Delta_1 = \{(i,i') \in \{1,...,n\},\ 1 \leq |i-i'| \leq s_n\},$$
$$\Delta_2 = \{(i,i') \in \{1,...,n\},\ 1+s_n \leq |i-i'| \leq n-1\}.$$
We write
$$\begin{aligned}&\frac{1}{n^2\mathbb{E}[\beta_1^2K_1]}\sum_{1\leq i\neq i'\leq n}Cov(\beta_i^2K_i, \beta_{i'}^2K_{i'})\\ &\quad= \frac{1}{n^2\mathbb{E}[\beta_1^2K_1]}\left(\sum_{\Delta_1}Cov(\beta_i^2K_i, \beta_{i'}^2K_{i'}) + \sum_{\Delta_2}Cov(\beta_i^2K_i, \beta_{i'}^2K_{i'})\right).\end{aligned} \tag{13}$$
The sum over Δ_1:
We have
$$\begin{aligned}\left|Cov(\beta_i^2K_i, \beta_{i'}^2K_{i'})\right| &\leq \left|\mathbb{E}[\beta_i^2K_i\beta_{i'}^2K_{i'}]\right| + \mathbb{E}^2[\beta_1^2K_1]\\ &\leq Ch_4\mathbb{E}[K_iK_i'] + \mathbb{E}^2[\beta_1^2K_1]\\ &\leq Ch_4\Phi(h_K) + \mathbb{E}^2[\beta_1^2K_1],\end{aligned}$$

then,

$$\Big|\sum_{\Delta_1} Cov(\beta_i^2 K_i, \beta_{i'}^2 K_{i'})\Big| \leq n s_n \frac{1}{n^2 \mathbb{E}[\beta_1^2 K_1]} \left(C h_4 \Phi(h_K) + \mathbb{E}[\beta_1^2 K_1]\right)$$
$$\leq \frac{C s_n \Phi(h_K))}{C' n \phi^2(h_K)} + \frac{s_n}{n},$$

and by choosing $s_n = \sqrt{n}$, as $n \longrightarrow +\infty$, we get $\sum\limits_{\Delta_1} Cov(\beta_i^2 K_i, \beta_{i'}^2 K_{i'}) \longrightarrow 0$.

The sum over Δ_2*:*
By the inequality for bounded mixing processes, for $i \neq i'$ we get

$$\Big|Cov(\beta_i^2 K_i, \beta_{i'}^2 K_{i'})\Big| \leq C h_K^4 \alpha(|i - i'|),$$

and using $\sum_{j \geq x+1} j^{-t} \leq \int_{m \geq x} m^{-t} = ((1-t)m^{t-1})^{-1}$, thus, after some simplification of calculations, we get

$$\Big|Cov(\beta_i^2 K_i, \beta_{i'}^2 K_{i'})\Big| \leq \frac{C s_n^{1-t}}{t - 1(C n \phi^2(h_K))},$$

and for the same s_n in above, as $n \longrightarrow +\infty$, we get

$$\sum_{\Delta_2} Cov(\beta_i^2 K_i, \beta_{i'}^2 K_{i'}) \longrightarrow 0.$$

(4) Observe that

$$\mathbb{E}[N_{2,j}^2] = Var(N_{2,j}) + \mathbb{E}^2[N_{2,j}]$$
$$= Var(N_{2,j}) + \frac{n^3 \phi(h_K) \mathbb{E}^2(\beta_1^2 K_1)}{h_H \mathbb{E}^2[\Gamma_1 K_1]} \mathbb{E}^2[K_j(H_j' - h_H f(y|x))].$$

We have $\mathbb{E}[H_1'|X_1] - h_H f(y|x) \longrightarrow 0$ as $n \longrightarrow \infty$ from Lemma 2 in Bouanani et al. (2018).

Then, by Lemma 1 in Zhou and Lin (2016) for $\frac{\mathbb{E}^2(\beta_1^2 K_1)}{h_H \mathbb{E}^2[\Gamma_1 K_1]} \xrightarrow{n \to +\infty} 0$, we get $\mathbb{E}^2[N_{2,j}] \xrightarrow{n \to +\infty} 0$.

So,

$$\mathbb{E}[N_{2,j}^2] \xrightarrow{n \to +\infty} \Big(Var(N_{2,j}) = V_{HK}^f(x,y)\Big).$$ ■

Proof:

The proof of the Lemma 4.1 is based on the same concept in Bouanani et al. (2018).

First, we denote

$$\Upsilon(x,y) = \widehat{f}_N(y|x) - h_H f(y|x) \widehat{F}_D(x) - \mathbb{E}\left[\widehat{f}_N(y|x) - h_H f(y|x) \widehat{F}_D(x)\right],$$

and let

$$\begin{aligned}\Omega_n &= \sqrt{nh_H\phi(h_K)}\sum_{j=1}^{n}\Upsilon_j(x,y)\\ &= \frac{\sqrt{nh_H\phi_x(h_K)}}{nh_H\mathbb{E}[\Gamma_1 K_1]}\Bigg(\sum_{j=1}^{n}\Big(\Gamma_j K_j(H_j'-h_H f(y|x)) - \mathbb{E}\left[\Gamma_j K_j(H_j'-h_H f(y|x))\right]\Big)\Bigg),\end{aligned}$$

thus, our claim is to prove $\Omega_n \xrightarrow{\mathcal{D}} \mathcal{N}(0, V_{HK}^f(x,y))$.

By some simplification on Ω_n, we get

$$\Omega_n = \Bigg(N_1N_2 - \mathbb{E}[N_1N_2] - \Big(N_3N_4 - \mathbb{E}[N_3N_4]\Big)\Bigg),$$

where

$$\begin{aligned}N_{1,i} &= \tfrac{1}{n\mathbb{E}[\beta_1^2K_1]}\sum_{i=1}^{n}\beta_i^2K_i, \quad N_{2,j} = \tfrac{\sqrt{nh_H\phi(h_K)}\mathbb{E}(\beta_1^2K_1)}{h_H\mathbb{E}[\Gamma_1K_1]}\sum_{j=1}^{n}K_j(H_j'-h_Hf(y|x)),\\ N_{3,i} &= \tfrac{1}{n\mathbb{E}[\beta_1K_1]}\sum_{i=1}^{n}\beta_iK_i, \quad N_{4,j} = \tfrac{\sqrt{nh_H\phi(h_K)}\mathbb{E}(\beta_1K_1)}{h_H\mathbb{E}[\Gamma_1K_1]}\sum_{j=1}^{n}\beta_jK_j(H_j'-h_Hf(y|x)).\end{aligned}$$

So, we must prove that

$$N_{1,i}N_{2,j} - \mathbb{E}[N_{1,i}N_{2,j}] \xrightarrow{\mathcal{D}} \mathcal{N}(0, V_{HK}^f(x,y)), \tag{14}$$

and

$$N_{3,i}N_{4,j} - \mathbb{E}[N_{3,i}N_{4,j}] \xrightarrow{\mathbb{P}} 0. \tag{15}$$

- Concerning (14):
  We begin by taking apart this equation, such that

$$N_{1,i}N_{2,j} - \mathbb{E}[N_{1,i}N_{2,j}] = \underbrace{(N_{2,j} - \mathbb{E}[N_{2,j}])}_{R_1} + \underbrace{\Big((N_{1,i}-1)N_{2,j} - \mathbb{E}[(N_{1,i}-1)N_{2,j}]\Big)}_{R_2}, \tag{16}$$

  now, we need to show that $R_1 \xrightarrow{\mathcal{D}} \mathcal{N}(0, V_{HK}^f(x,y))$ and $R_2 \xrightarrow{\mathbb{P}} 0$.
  - For R_1, we have

$$\begin{aligned}R_1 &= N_{2,j} - \mathbb{E}[N_{2,j}]\\ &= \frac{\sqrt{nh_H\phi(h_K)}\mathbb{E}(\beta_1^2K_1)}{h_H\mathbb{E}[\Gamma_1K_1]}\sum_{j=1}^{n}K_j(H_j'-h_Hf(y|x)) - \mathbb{E}[K_j(H_j'-h_Hf(y|x))],\end{aligned}$$

    then, we need to evaluate the variance of $N_{2,j}$. We obtain

$$\begin{aligned}\text{Var}(N_{2,j}) = n\text{Var}(N_{2,1}) + 2\frac{n\phi(h_K)\mathbb{E}^2[\beta_1^2K_1]}{h_H\mathbb{E}^2[\Gamma_1K_1]}\\ \sum_{1\le i<j\le n} Cov\left(K_i(H_i'-h_Hf(y|x)), K_j(H_j'-h_Hf(y|x))\right).\end{aligned}$$

Therefore, to prove that $\lim_{n\to+\infty}\text{Var}(N_{2,j}) = V_{HK}^{f}(x,y)$, it necessary to establish the following results:

$$\lim_{n\to+\infty} n\text{Var}(N_{2,1}) = \frac{A_2}{A_1^2} f(y|x)\int (H^{'}(t))^2 \mathrm{d}t, \tag{17}$$

$$\lim_{n\to+\infty}\left(\frac{n\phi(h_K)\mathbb{E}^2[\beta_1^2 K_1]}{h_H\mathbb{E}^2[\Gamma_1 K_1]}\sum_{1\le i<j\le n} Cov\left(K_i(H_i' - h_f(y|x)), K_j(H_j' - h_H f(y|x))\right)\right) = 0. \tag{18}$$

From the Lemma 6.1, we get (17) and (18).

- For R_2, by Bienayme-Tchebychev's inequality and the Cauchy-Schwartz's inequality, one can write

$$\begin{aligned}\mathbb{P}\left(\left|\left((N_{1,i}-1)N_{2,j} - \mathbb{E}[(N_{1,i}-1)N_{2,j}]\right)\right| \ge \epsilon\right) \\ &\le \frac{\mathbb{E}\left[\left|\left((N_{1,i}-1)N_{2,j} - \mathbb{E}[(N_{1,i}-1)N_{2,j}]\right)\right|\right]}{\epsilon} \\ &\le 2\frac{\sqrt{\mathbb{E}[(N_{1,i}-1)^2]}\sqrt{\mathbb{E}[N_{2,j}^2]}}{\epsilon}.\end{aligned}$$

Based on the results of Lemma 6.1, we get as $n \longrightarrow \infty$,

$$\mathbb{E}[(N_{1,i}-1)^2] \longrightarrow 0 \text{ and } \mathbb{E}[N_{2,j}^2] \longrightarrow V_{HK}^{f}(x,y).$$

- Concerning (15):

It is enough to check as $n \longrightarrow \infty$:

$$\mathbb{E}|N_{4,j} - \mathbb{E}[N_{4,j}]| \xrightarrow{L^1} 0, \tag{19}$$

$$\mathbb{E}|(N_{3,j}-1)N_{4,j} - \mathbb{E}[(N_{3,j}-1)N_{4,j}]| \longrightarrow 0. \tag{20}$$

- About (19), we show the L^2 consistency:

$$\begin{aligned}\mathbb{E}[(N_{4,j} - \mathbb{E}[N_{4,j}])^2] &= nVar(N_{4,1}) + 2\frac{n\phi(h_K)\mathbb{E}^2[\beta_1 K_1]}{h_H\mathbb{E}^2[\Gamma_1 K_1]} \\ &\qquad \sum_{1\le i<j\le n} Cov\left(\beta_i K_i(H_i' - h_H f(y|x)), \beta_j K_j(H_j' - h_H f(y|x))\right) \\ &= \underbrace{\frac{n\phi(h_K)\mathbb{E}^2[\beta_1 K_1]}{h_H\mathbb{E}^2[\Gamma_1 K_1]} nVar\left(\beta_1 K_1(H_1' - h_H f(y|x))\right)}_{F_1} \\ &\underbrace{+2\frac{n\phi(h_K)\mathbb{E}^2[\beta_1 K_1]}{h_H\mathbb{E}^2[\Gamma_1 K_1]}\sum_{1\le i<j\le n} Cov\left(\beta_i K_i(H_i' - h_H f(y|x)), \beta_j K_j(H_j' - h_H f(y|x))\right)}_{F_2}.\end{aligned}$$

For F_1, we have

$$Var\left(\beta_1 K_1 (H_1' - h_H f(y|x))\right) = \mathbb{E}\left[\beta_1^2 K_1^2 (H_1' - h_H f(y|x))^2\right] - \mathbb{E}^2\left[\beta_1 K_1 (H_1' - h_H f(y|x))\right].$$

We obtain the term $\mathbb{E}^2\left[\beta_1 K_1 \mathbb{E}[(H_1'|X_1) - h_H f(y|x)]\right] \xrightarrow{n\to+\infty}$, 0 from Lemma 2 in Bouanani et al. (2018); then, after simplification in the term F_1, we get

$$F_1 = \frac{n^2\phi(h_K)\mathbb{E}^2[\beta_1 K_1]}{h_H \mathbb{E}^2[\Gamma_1 K_1]} \times \left(\mathbb{E}[Var(H_1'|X_1)\beta_1^2 K_1^2] + \mathbb{E}\left[\beta_1^2 K_1^2 \left(\mathbb{E}[H_1'|X_1] - h_H f(y|x)\right)^2\right]\right). \tag{21}$$

Also, from the Lemma 2 in Bouanani et al. (2018), we have the second term of (21) tends to 0 as n tends to the infinity. About the first term and by Lemma 1 in Zhou and Lin (2016) and the same way for (17), we get $F_1 \xrightarrow{n\to+\infty} 0$.

For F_2, and through (18), it follows that $F_2 \xrightarrow{n\to+\infty} 0$.

- About (20), the Cauchy Schwartz inequality permits us to write that

$$\mathbb{E}|(N_{3,j}-1)N_{4,j} - \mathbb{E}[(N_{3,j}-1)N_{4,j}]| \leq 2\sqrt{\mathbb{E}[(N_{3,i}-1)]^2}\sqrt{\mathbb{E}[N_{4,j}]^2}.$$

We have as $n \to \infty$:

$$\begin{aligned}\mathbb{E}[(N_{3,i}-1)]^2 &= nVar(N_{3,1}) + 2Cov(N_{3,i}, N_{3,j}) \\ &\leq \frac{nVar(\beta_1 K_1)}{n^2\mathbb{E}^2[\beta_1 K_1]} + 2Cov(N_{3,i}, N_{3,j}) \\ &= O\left(\frac{1}{n\phi(h_K)}\right) + 2Cov(N_{3,i}, N_{3,j}).\end{aligned}$$

Similar to the steps of (13), we obtain $Cov(N_{3,i}, N_{3,j}) \longrightarrow 0$.

Consequently, from (14) and (15) the proof of Lemma 4.1 is achieved. ■

Proof:

We have from the Lemma 3.5 of Bouanani et al. (2019) that $\widehat{F}_D(x) \xrightarrow{\mathbb{P}} \mathbb{E}[\widehat{F}_D(x)] = 1$, such that

$$\mathbb{E}\Big[(\widehat{F}_D(x) - \widehat{F}_N(y|x) - 1 + \mathbb{E}[\widehat{F}_N(y|x)])\Big] = 0,$$

and

$$\left(\frac{nh_H\phi(h_K)}{V_{HK}^h(x,y)}\right)^{1/2} \mathrm{Var}\Big((\widehat{F}_D(x) - \widehat{F}_N(y|x) - 1 + \mathbb{E}[\widehat{F}_N(y|x)])\Big) \longrightarrow 0.$$

So, the results permit us to finish the proof of Lemma 4.2. ■

Proof:

For the proof of Corollary 4.1.

By Lemma 4.2 and Theorem 1 in Merouan et al. (2018), we can write

$$\mathbb{E}[\widehat{F}_D(x) - \widehat{F}_N(y|x) - 1 + F(y|x)] \longrightarrow 0,$$

$$\text{Var}(\widehat{F}_D(x) - \widehat{F}_N(y|x) - 1 + F(y|x)) \longrightarrow 0.$$

Then, $\widehat{F}_D(x) - \widehat{F}_N(y|x) - 1 + F(y|x) \xrightarrow{\mathbb{P}} 0.$ ■