New Trends in Mathematical Sciences

http://www.ntmsci.com

On conditional hazard function estimate for functional mixing data

Tayeb Djebbouri[1], *El Hadj Hamel*[2] *and Abbes Rabhi*[3]

[1]Stochastic Models, Statistics and Applications Laboratory, University of Saida, Algeria,
[2]Laboratory of Mathematics Hassiba Benbouali University of Chlef, Algeria.
[3]Laboratory of Mathematics, University of Sidi Bel Abbes, Algeria.

Received: 31 August 2014, Revised: 10 November 2014, Accepted: 18 January 2015
Published online: 10 February 2015

Abstract: This paper considers the problem of nonparametric estimation of the conditional hazard function for functional mixing data. In particular, given a strictly stationary random variables $Z_i = (X_i,\, Y_i)_{i\in\mathbb{N}}$, we investigate a kernel estimate of the conditional hazard function of univariate response variable Y_i given the functional variable X_i. The mean squared convergence rate is given and the asymptotic normality of the proposed estimator is proven.

Keywords: Functional data, Kernel conditional hazard function, Kernel estimation, Nonparametric Estimation, Probabilities of small balls, Strong mixing process.

1 Introduction

Statistical problems involved in the modelization of functional data have received an increasing interest in the few past decade.

The infatuation for this topic is linked with many fields of applications in which the data are collected in the functional order. Under this hypothesis, the statistical analysis focuses on a framework of infinite dimension for the data under study. This type of data appears in many fields of applied statistics: environmetrics [7], chemometrics [2], meteorological sciences [3], etc. This field of modern statistics has received much attention recently, it has been popularized in the book of Ramsay and Silverman [20].

The nonparametric estimation of the hazard and/or the conditional hazard function is quite important in a variety of fields such as medicine, reliability, survival analysis or in seismology. The hazard estimate was introduced by Watson and Leadbetter [27], after that considerable results have been given, see for example, Ahmad [1], Singpurwalla and Wong [24], and we can also cite Quintela [17] for a survey, Roussas [23] (for previous works), Li and Tran [14] (for recent advances and references).

When hazard rate estimation is performed with multiple variables, the result is an estimate of the conditional hazard rate for the first variable, given the levels of the remaining variables. Many references, practical examples and simulations in the case of non-parametric estimation using local linear approximations can be found in Spierdijk [25].

From a theoretical point of view, a sample of functional data can be involved in many different statistical problems, such

* Corresponding author e-mail: rabhi_abbes@yahoo.fr

as for instance: classification and principal components analysis (PCA)[4,**?**] or longitudinal studies, regression and prediction [2,6].

The literature is strictly not limited in the case where the data is of functional nature (a curve). The first result in this context, was given by Ferraty *et al* . [10], authors established the almost complete convergence of the kernel estimate of the conditional hazard function in the i.i.d. case and under α-mixing condition, and recently Rabhi *et al.* [18] studied the mean quadratic convergence in the i.i.d. case of this estimate. More recently Mahiddine *et al.* [15] give the uniform version of the almost complete convergence rate in the i.i.d. case.

The recent monograph by Ferraty and Vieu [11] summarizes many of their contributions to the non-parametric estimation with functional data; among other properties, consistency of the conditional density, conditional distribution and regression estimates are established in the i.i.d. case as well as under dependence conditions (strong mixing). Almost complete rates of convergence are also obtained, and the different techniques are applied to various examples of functional data samples. Related work can be found in the paper of Masry [16], where the asymptotic normality of the functional non-parametric regression estimate is proven, considering strong mixing dependence conditions for the sample data. For automatic smoothing parameter selection in the regression setting, see Rachdi and Vieu [19].

The main aim of this paper, is to study, under general conditions, the asymptotic proprieties of the functional data kernel estimate of the conditional hazard function introduced by Ferraty *et al.* [10]. More precisely, we establish the asymptotic normality of the construct estimator. We point out that our asymptotic results are useful in some statistical problems such as the choice of the smoothing parameters. The present work extended to dependent case the result of Rabhi *et al.* [18] given in i.i.d. case functional. Note that, one of the main difficulties, when dealing with functional variables, relies on the difficulty for choosing some appropriate measure of reference in infinite dimensional spaces. The fundamental feature of our approach is to build estimates and to derive their asymptotic properties without any notion of density for the functional variable X. This approach allows us to avoid the use of a reference measure in such functional spaces. In each of the above described sections, we will give general asymptotic results without assuming existence of such a density, and each of these results will be discussed in relation with earlier literature existing in the usual finite dimensional case.

Our paper presents some asymptotic properties related with the non-parametric estimation of the conditional hazard function. In a functional data setting, the conditioning variable is allowed to take its values in some abstract semi-metric space. In this case, Ferraty *et al.* [26] define non-parametric estimators of the conditional density and the conditional distribution. They give the rates of convergence (in an almost complete sense) to the corresponding functions, in an a dependence (α-mixing) context. In Rabhi *et al.* [18], the same properties are shown in an i.i.d. context in the data sample. We extend their results to dependent case by calculating the bias and variance of these estimates, and establishing their asymptotic normality, considering a particular type of kernel for the functional part of the estimates. Because the hazard function estimator is naturally constructed using these two last estimators, the same type of properties is easily derived for it. Our results are valid in a real (one- and multi-dimensional) context.

The paper is organized as follows: In the next section we present our model. In section 3 we present notations and hypotheses, Section 4 is dedicated for our main results. Section 5 is devoted to some discuss on the applicability of our asymptotic result in some statistical problems.

2 The model

Consider $Z_i = (X_i, Y_i)$, $i \in \mathbb{N}$ be a $\mathscr{F} \times \mathbb{R}$-valued measurable strictly stationary process, defined on a probability space $(\Omega, \mathscr{A}, \mathbb{P})$, where $(\mathscr{F}, d)$ is a semi-metric space.

In the following x will be a fixed point in $\mathscr{F}$ and N_x will denote a fixed neighborhood of x. We assume that the regular version of the conditional probability of Y given X exists. Moreover, we suppose that, for all $z \in N_x$ the conditional distribution function of Y given $X = z$, $F^z(\cdot)$, is 3-times continuously differentiable and we denote by f^z its conditional density with respect to (w.r.t.) Lebesgue's measure over $\mathbb{R}$. In this paper, we consider the problem of the nonparametric estimation of the conditional hazard function defined, for all $y \in \mathbb{R}$ such that $F^x(y) < 1$, by

$$h^x(y) = \frac{f^x(y)}{1 - F^x(y)}.$$

In our spatial context, we estimate this function by

$$\widehat{h}^x(y) = \frac{\widehat{f}^x(y)}{1 - \widehat{F}^x(y)}$$

where

$$\widehat{F}^x(y) = \frac{\sum_{i=1}^n K(h_K^{-1} d(x, X_i)) H(h_H^{-1}(y - Y_i))}{\sum_{i=1}^n K(h_K^{-1} d(x, X_i))}, \qquad \forall y \in \mathbb{R}$$

and

$$\widehat{f}^x(y) = \frac{h_H^{-1} \sum_{i=1}^n K(h_K^{-1} d(x, X_i)) H'(h_H^{-1}(y - Y_i))}{\sum_{i=1}^n K(h_K^{-1} d(x, X_i))}, \qquad \forall y \in \mathbb{R}$$

with K is the kernel, H is a given continuously differentiable distribution function, $h_K = h_{K,n}$ (resp. $h_H = h_{H,n}$) is a sequence of positive real numbers and H' is the derivative of H. Furthermore, the estimator $\widehat{h}^x(y)$ can we written as

$$\widehat{h}^x(y) = \frac{\widehat{f}_N^x(y)}{\widehat{F}_D^x - \widehat{F}_N^x(y)} \tag{1}$$

where

$$\begin{aligned}
\widehat{F}_D^x &:= \frac{1}{n\mathbb{E}[K_1]} \sum_{i=1}^n K(h_K^{-1} d(x, X_i)),\ K_1 = K(h_K^{-1} d(x, X_1)) \\
\widehat{F}_N^x(y) &:= \frac{1}{n\mathbb{E}[K_1]} \sum_{i=1}^n K(h_K^{-1} d(x, X_i)) H(h_H^{-1}(y - Y_i)) \\
\widehat{f}_N^x(y) &:= \frac{1}{nh_H\mathbb{E}[K_1]} \sum_{i=1}^n K(h_K^{-1} d(x, X_i)) H'(h_H^{-1}(y - Y_i)).
\end{aligned}$$

Our main purpose is to study the L^2- consistency and the asymptotic normality of the nonparametric estimate $\widehat{h}^x$ of h^x when the random filed $(Z_i, i \in \mathbb{N})$ satisfies the following mixing condition.

3 Notations and hypotheses

All along the paper, when no confusion is possible, we will denote by C and C' some strictly positive generic constants. In order to establish our asymptotic results we need the following hypotheses:

(H0) $\forall r > 0, \mathbb{P}(X \in B(x, r)) =: \phi_x(r) > 0$, where $B(x, r) = \{x' \in \mathscr{F} / d(x, x') < r\}$.

(H1) $(X_i, Y_i)_{i\in\mathbb{N}}$ is an α-mixing sequence whose the coefficients of mixture verify:

$$\exists a > 0,\ \exists c > 0:\ \forall n \in \mathbb{N},\ \alpha(n) \leq cn^{-a}.$$

(H2) $0 < \sup_{i\neq j} \mathbb{P}\left((X_i, X_j) \in B(x,h) \times B(x,h)\right) \ = \ \mathscr{O}\left(\dfrac{(\phi_x(h))^{(a+1)/a}}{n^{1/a}}\right).$

Note that (H0) can be interpreted as a concentration hypothesis acting on the distribution of the *f.r.v.* X, whereas (H2) concerns the behavior of the joint distribution of the pairs (X_i, X_j). In fact, this hypothesis is equivalent to suppose that, for n large enough

$$\sup_{i\neq j} \frac{\mathbb{P}\left((X_i, X_j) \in B(x,h) \times B(x,h)\right)}{\mathbb{P}\left(X \in B(x,h)\right)} \leq C\left(\frac{\phi_x(h)}{n}\right)^{1/a}.$$

(H3) For $l \in \{0,2\}$, the functions $\Psi_l(s) = \mathbb{E}\left[\frac{\partial^l F^X(y)}{\partial y^l} - \frac{\partial^l F^x(y)}{\partial y^l}\Big| d(x,X) = s\right]$ and $\Phi_l(s) = \mathbb{E}\left[\frac{\partial^l f^X(y)}{\partial y^l} - \frac{\partial^l f^x(y)}{\partial y^l}\Big| d(x,X) = s\right]$ are derivable at $s = 0$.

(H4) The bandwidth h_K satisfies:

$$h_K \downarrow 0,\ \ \forall t \in [0,1]\ \lim_{h_K\to 0} \frac{\phi_x(th_K)}{\phi_x(h_K)} = \beta_x(t)\ \text{ and }\ nh_H\phi_x(h_K) \to \infty\ \text{ as } n \to \infty.$$

(H5) The kernel K from $\mathbb{R}$ into $\mathbb{R}^+$ is a differentiable function supported on $[0,1]$. Its derivative K' exists and is such that there exist two constants C and C' with $-\infty < C < K'(t) < C' < 0$ for $0 \leq t \leq 1$.

(H6) H has even bounded derivative function supported on $[0,1]$ that verifies

$$\int_{\mathbb{R}} |t|^{b_2} H'(t)dt < \infty.$$

(H7) There exist sequences of integers (u_n) and (v_n) increasing to infinity such that $(u_n + v_n) \leq n$, satisfying

(i) $v_n = o((nh_H\phi_x(h_K))^{1/2})$ and $\left(\frac{n}{h_H\phi_x(h_K)}\right)^{1/2} \alpha(v_n) \to 0$ as $n \to 0$,

(ii) $q_n v_n = o((nh_H\phi_x(h_K))^{1/2})$ and $q_n\left(\frac{n}{h_H\phi_x(h_K)}\right)^{1/2} \alpha(v_n) \to 0$ as $n \to 0$

where q_n is the largest integer such that $q_n(u_n + v_n) \leq n$.

3.1 Remarks on the assumptions

Remark 1. Assumption (H0) plays an important role in our methodology. It is known as (for small h) the "concentration hypothesis acting on the distribution of X" in infinite-dimensional spaces. This assumption is not at all restrictive and overcomes the problem of the non-existence of the probability density function. In many examples, around zero the small ball probability$\phi_x(h)$ can be written approximately as the product of two independent functions $\psi(z)$ and $\varphi(h)$ as $\phi_z(h) = \psi(z)\varphi(h) + o(\varphi(h))$. This idea was adopted by Masry [16] who reformulated the Gasser *et al.* [12] one. The increasing property of $\phi_x(.)$ implies that $\zeta_h^x(.)$ is bounded and then integrable (all the more so $\zeta_0^x(.)$ is integrable).

Without the differentiability of $\phi_x(.)$, this assumption has been used by many authors where $\psi(.)$ is interpreted as a probability density, while $\varphi(.)$ may be interpreted as a volume parameter. In the case of finite-dimensional spaces, that is $\mathscr{S} = \mathbb{R}^d$, it can be seen that $\phi_x(h) = C(d)h^d\psi(x) + oh^d)$, where $C(d)$ is the volume of the unit ball in $\mathbb{R}^d$. Furthermore, in infinite dimensions, there exist many examples fulfilling the decomposition mentioned above. We quote the following (which can be found in Ferraty *et al.* [8]):

1. $\phi_x(h) \approx \psi(h)h^\gamma$ for som $\gamma > 0$.
2. $\phi_x(h) \approx \psi(h)h^\gamma \exp\{C/h^p\}$ for som $\gamma > 0$ and $p > 0$.
3. $\phi_x(h) \approx \psi(h)/|\ln h|$.

The function $\beta_h^x(.)$ which intervenes in Assumption (H4) is increasing for all fixed h. Its pointwise limit $\beta_0^x(.)$ also plays a determinant role. It intervenes in all asymptotic properties, in particular in the asymptotic variance term. With simple algebra, it is possible to specify this function (with $\beta_0(u) := \beta_0^x(u)$ in the above examples by:

1. $\beta_0(u) = u^\gamma$,
2. $\beta_0(u) = \delta_1(u)$ where $\delta_1(.)$ is Dirac function,
3. $\beta_0(u) = \mathbf{1}_{]0,1]}(u)$.

Assumption (H2) is classical and permits to make the variance term negligible.

Remark 2. Assumptions (H3) is a regularity condition which characterize the functional space of our model and is needed to evaluate the bias.

Remark 3. Assumptions (H5) and (H6) are classical in functional estimation for finite or infinite dimension spaces.

4 Main results

4.1 Mean squared convergence

In this part we establish the L^2-consistency of $\widehat{h}^x(y)$.

Theorem 1. Under assumptions (H0)-(H6), we have

$$\mathbb{E}\left[\widehat{h}^x(y) - h^x(y)\right]^2 = B_n^2(x,y) + \frac{\sigma_h^2(x,y)}{nh_H\phi_x(h_K)} + o(h_H^4) + o(h_K) + o\left(\frac{1}{n\,h_H\phi_x(h_K)}\right),$$

where

$$B_n(x,y) = \frac{(B_H^f - h^x(y)B_H^F)h_H^2 + (B_K^f - h^x(y)B_K^F)h_K}{1 - F^x(y)}$$

with

$$\begin{aligned}
B_H^f(x,y) &= \frac{1}{2}\frac{\partial^2 f^x(y)}{\partial y^2}\int t^2 H'(t)dt \\
B_K^f(x,y) &= h_K\Phi_0'(0)\frac{\left(K(1) - \int_0^1 (sK(s))'\beta_x(s)ds\right)}{\left(K(1) - \int_0^1 K'(s)\beta_x(s)ds\right)} \\
B_H^F(x,y) &= \frac{1}{2}\frac{\partial^2 F^x(y)}{\partial y^2}\int t^2 H'(t)dt \\
B_K^F(x,y) &= h_K\Psi_0'(0)\frac{\left(K(1) - \int_0^1 (sK(s))'\beta_x(s)ds\right)}{\left(K(1) - \int_0^1 K'(s)\beta_x(s)ds\right)}.
\end{aligned}$$

and

$$\sigma_h^2(x,y) = \frac{\beta_2 h^x(y)}{(\beta_1^2(1 - F^x(y))} \quad (\text{with } \beta_j = K^j(1) - \int_0^1 (K^j)'(s)\beta_x(s)ds, \text{ for, } j = 1, 2).$$

Proof: By using the same decomposition used in (Theorem 1 Rabhi *et al.* [18], P.408), we show that the proof of Theorem 1 can be deduced from the following intermediate results:

Lemma 1. Under the hypotheses of Theorem 1, we have

$$\mathbb{E}\left[\widehat{f}_N^x(y)\right]-f^x(y)=B_H^f(x,y)h_H^2+B_K^f(x,y)h_K+o(h_H^2)+o(h_K)$$

and

$$\mathbb{E}\left[\widehat{F}_N^x(y)\right]-F^x(y)=B_H^F(x,y)h_H^2+B_K^F(x,y)h_K+o(h_H^2)+o(h_K).$$

Remark 4. Observe that, the result of this lemma permits to write

$$\left[\mathbb{E}\widehat{F}_N^x(y)-F^x(y)\right]=\mathscr{O}(h_H^2)+O(h_K)$$

and

$$\left[\mathbb{E}\widehat{f}_N^x(y)-f^x(y)\right]=\mathscr{O}(h_H^2+h_K).$$

Lemma 2. Under the hypotheses of Theorem 1, we have

$$Var\left[\widehat{f}_N^x(y)\right]=\frac{\sigma_f^2(x,y)}{nh_H\phi_x(h_K)}+o\left(\frac{1}{nh_H\phi_x(h_K)}\right),$$

$$Var\left[\widehat{F}_N^x(y)\right]=o\left(\frac{1}{nh_H\phi_x(h_K)}\right)$$

and

$$Var\left[\widehat{F}_D^x\right]=o\left(\frac{1}{nh_H\phi_x(h_K)}\right).$$

where $\sigma_f^2(x,y):=f^x(y)\int H'^2(t)dt$.

Lemma 3. Under the hypotheses of Theorem (**??**), we have

$$Cov(\widehat{f}_N^x(y),\widehat{F}_D^x)=o\left(\frac{1}{nh_H\phi_x(h_K)}\right),$$

$$Cov(\widehat{f}_N^x(y),\widehat{F}_N^x(y))=o\left(\frac{1}{nh_H\phi_x(h_K)}\right)$$

and

$$Cov(\widehat{f}_D^x,\widehat{F}_N^x(y))=o\left(\frac{1}{nh_H\phi_x(h_K)}\right).$$

Remark 5. It is clear that, the results of Lemmas 2 and 3 allows to write

$$Var\left[\widehat{F}_D^x-\widehat{F}_N^x\right]=o\left(\frac{1}{nh_H\phi_x(h_K)}\right)$$

4.2 Asymptotic normality

This section contains results on the asymptotic normality of $\widehat{h}^x(y)$.

Theorem 2. Assume that (H0)-(H7) hold, and if the following inequalities

$$\exists \eta > 0,\ C,\ C' > 0 \text{ such that } C n^{\frac{3-a}{a+1}+\eta} \leq h_H\, \phi_x(h_K) \text{ and } \phi_x(h_K) \leq C' n^{\frac{1}{1-a}} \tag{2}$$

are verified with $a > (5+\sqrt{17})/2$, then we have for any $x \in \mathscr{A}$,

$$\left(\frac{n h_H \phi_x(h_K)}{\sigma_h^2(x,y)}\right)^{1/2} \left(\widehat{h}^x(y) - h^x(y) - B_n(x,y)\right) \xrightarrow{\mathscr{D}} \mathscr{N}(0,1) \quad \text{as} \quad n \to \infty.$$

where

$$\mathscr{A} = \{x \in \mathscr{F},\ f^x(y)(1-F^x(y)) \neq 0\}$$

and $\xrightarrow{\mathscr{D}}$ means the convergence in distribution.

Evidently, if one imposes some additional assumptions on the function $\phi_x(\cdot)$ and the bandwidth parameters (h_K and h_H) our asymptotic normality can be improved by removing the bias term $B_n(x,y)$.

Corollary 1. Under the hypotheses of Theorem 2 and if the bandwidth parameters (h_K and h_H) and if the function $\phi_x(h_K)$ satisfies:

$$\lim_{n\to\infty} (h_H^2 + h_K)\sqrt{n\phi_x(h_K)} = 0$$

we have

$$\left(\frac{n h_H \phi_x(h_K)}{\sigma_h^2(x,y)}\right)^{1/2} \left(\widehat{h}^x(y) - h^x(y)\right) \xrightarrow{\mathscr{D}} \mathscr{N}(0,1) \quad \text{as} \quad n \to \infty.$$

Proof of Theorem and Corollary: Consider the decomposition

$$\begin{aligned}
\widehat{h}^x(y) - h^x(y) = {} & \frac{1}{\widehat{F}_D^x - \widehat{F}_N^x(y)} \left(\widehat{f}_N^x(y) - E\widehat{f}_N^x(y)\right) \\
& + \frac{1}{\widehat{F}_D^x - \widehat{F}_N^x(y)} \left\{ h^x(y)\left(\mathbb{E}\widehat{F}_N^x(y) - F^x(y)\right) + \left(\mathbb{E}\widehat{f}_N^x(y) - f^x(y)\right) \right\} \\
& + \frac{h^x(y)}{\widehat{F}_D^x - \widehat{F}_N^x(y)} \left\{ 1 - \mathbb{E}\widehat{F}_N^x(y) - \left(\widehat{F}_D^x - \widehat{F}_N^x(y)\right) \right\}
\end{aligned} \tag{3}$$

Therefore, Theorem 2 and Corollary 1 are a consequence of Lemma 1, remark 1 and the following results.

Lemma 4. Under the hypotheses of Theorem 2

$$\left(\frac{n h_H \phi_x(h_K)}{\sigma_f^2(x,y)}\right)^{1/2} \left(\widehat{f}_N^x(y) - \mathbb{E}\left[\widehat{f}_N^x(y)\right]\right) \to N(0,1).$$

Lemma 5. Under the hypotheses of Theorem 2

$$\widehat{F}_D^x - \widehat{F}_N^x(y) \to 1 - F^x(y) \qquad \text{in probability}$$

and

$$\left(\frac{nh_H\phi_x(h_K)}{\sigma_h^2(x,y)}\right)^{1/2}\left(\widehat{F}_D^x-\widehat{F}_N^x(y)-1+\mathbb{E}[\widehat{F}_N^x(y)]\right)=o_{\mathbb{P}}(1).$$

5 Appendix

In the following, we will denote $\forall i$

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

Proof of Lemma 1: Firstly, for $\mathbb{E}[\widehat{f}_N^x(y)]$, we start by writing

$$\mathbb{E}[\widehat{f}_N^x(y)]=\frac{1}{\mathbb{E}[K_1]}\mathbb{E}\left[K_1\mathbb{E}[h_H^{-1}H_1'|X]\right]\text{ with }h_H^{-1}\mathbb{E}\left[H_1'|X\right]=\int_{\mathbb{R}}H'(t)f^X(y-h_Ht)dt.$$

The latter can be re-written, by using a Taylor expansion under (H3), as follows

$$h_H^{-1}\mathbb{E}[H_1'|X]=f^X(y)+\frac{h_H^2}{2}\left(\int t^2H'(t)dt\right)\frac{\partial^2f^X(y)}{\partial^2y}+o(h_H^2).$$

Thus, we get

$$\mathbb{E}\left[\widehat{f}_N^x(y)\right]=\frac{1}{\mathbb{E}[K_1]}\left(\mathbb{E}\left[K_1f^X(y)\right]+\left(\int t^2H'(t)dt\right)\mathbb{E}\left[K_1\frac{\partial^2f^X(y)}{\partial^2y}\right]+o(h_H^2)\right).$$

Let $\psi_l(\cdot,y):=\frac{\partial^lf^\cdot(y)}{\partial^ly}$: for $l\in\{0,2\}$, since $\Phi_l(0)=0$, we have

$$\begin{aligned}\mathbb{E}\left[K_1\psi_l(X,y)\right]&=\psi_l(x,y)\mathbb{E}[K_1]+\mathbb{E}\left[K_1\left(\psi_l(X,y)-\psi_l(x,y)\right)\right]\\&=\psi_l(x,y)\mathbb{E}[K_1]+\mathbb{E}\left[K_1\left(\Phi_l(d(x,X)\right)\right]\\&=\psi_l(x,y)\mathbb{E}[K_1]+\Phi_l'(0)\mathbb{E}\left[d(x,X)K_1\right]+o(\mathbb{E}\left[d(x,X)K_1\right]).\end{aligned}$$

So,

$$\mathbb{E}\left[\widehat{f}_N^x(y)\right]=f^x(y)+\frac{h_H^2}{2}\frac{\partial^2f^x(y)}{\partial y^2}\int t^2H'(t)dt+o\left(h_H^2\frac{\mathbb{E}\left[d(x,X)K_1\right]}{\mathbb{E}[K_1]}\right)+\Phi_0'(0)\frac{E\left[d(x,X)K_1\right]}{\mathbb{E}[K_1]}+o\left(\frac{\mathbb{E}\left[d(x,X)K_1\right]}{E[K_1]}\right).$$

Similarly to Ferraty *et al.* [8] we show that

$$\frac{1}{\phi_x(h_K)}\mathbb{E}\left[d(x,X)K_1\right]=h_K\left(K(1)-\int_0^1(sK(s))'\beta_x(s)ds+o(1)\right)$$

and

$$\frac{1}{\phi_x(h_K)}\mathbb{E}\left[K_1\right]=K(1)-\int_0^1K'(s)\beta_x(s)ds+o(1).$$

Hence,

$$\mathbb{E}\left[\widehat{f}_N^x(y)\right]=f^x(y)+\frac{h_H^2}{2}\frac{\partial^2f^x(y)}{\partial y^2}\int t^2H'(t)dt+h_K\Phi_0'(0)\frac{\left(K(1)-\int_0^1(sK(s))'\beta_x(s)ds\right)}{\left(K(1)-\int_0^1K'(s)\beta_x(s)ds\right)}+o(h_H^2)+o(h_K).$$

Secondly, concerning $\mathbb{E}[\widehat{F}_N^x(y)]$, we write by an integration by part

$$\mathbb{E}[\widehat{F}_N^x(y)] = \frac{1}{\mathbb{E}[K_1]}\mathbb{E}\left[K_1 E[H_1|X]\right] \text{ with } \mathbb{E}\left[H_1|X\right] = \int_{\mathbb{R}} H'(t)F^X(y-h_H t)dt.$$

The same steps used in studying $\mathbb{E}[\widehat{f}_N^x(y)]$ can be followed to prove that

$$\mathbb{E}\left[\widehat{F}_N^x(y)\right] = F^x(y) + \frac{h_H^2}{2}\frac{\partial^2 F^x(y)}{\partial y^2}\int t^2 H'(t)dt + h_K \Psi_0'(0)\frac{\left(K(1)-\int_0^1 (sK(s))'\beta_x(s)ds\right)}{\left(K(1)-\int_0^1 K'(s)\beta_x(s)ds\right)} + o(h_H^2) + o(h_K).$$

Proof of Lemma 2: For the first quantity $Var[\widehat{f}_N^x(y)]$, we have

$$s_n^2 \;=\; Var[\widehat{f}_N^x(y)] = \frac{1}{(nh_H\mathbb{E}\left[K_1(x)\right])^2}Var\left[\sum_{i=1}\Gamma_i(x)\right]$$

where

$$\Gamma_i(x) = K_i(x)H_i'(y) - \mathbb{E}\left[K_i(x)H_i'(y)\right].$$

Thus

$$\begin{aligned} Var[\widehat{f}_N^x(y)] &= \frac{1}{(nh_H\mathbb{E}\left[K_1\right])^2}\underbrace{\sum_{i\neq j}Cov\left(\Gamma_i(x),\Gamma_j(x)\right)}_{s_n^{cov}} + \underbrace{\sum_{i=1}^{n}Var\left(\Gamma_i(x)\right)}_{s_n^{var}} \\ &= \frac{1}{n(h_H\mathbb{E}\left[K_1\right])^2}Var\left[\Gamma_1\right] + \frac{1}{(nh_H\mathbb{E}\left[K_1\right])^2}\sum_{i\neq j}Cov(\Gamma_i,\Gamma_j). \end{aligned}$$

Let us calculate the quantity $Var\left[\Gamma_1(x)\right]$. We have:

$$\begin{aligned} Var\left[\Gamma_1(x)\right] &= \mathbb{E}\left[K_1^2(x)H_1'^2(y)\right] - \left(\mathbb{E}\left[K_1(x)H_1'(y)\right]\right)^2 \\ &= \mathbb{E}\left[K_1^2(x)\right]\frac{\mathbb{E}\left[K_1^2(x)H_1'^2(y)\right]}{\mathbb{E}\left[K_1^2(x)\right]} - \left(\mathbb{E}\left[K_1(x)\right]\right)^2\left(\frac{\mathbb{E}\left[K_1(x)H_1'(y)\right]}{\mathbb{E}\left[K_1(x)\right]}\right)^2. \end{aligned}$$

So, by using the same arguments as those used in pervious lemma we get

$$\begin{aligned} \frac{1}{\phi_x(h_K)}\mathbb{E}\left[K_1^2(x)\right] &= K^2(1) - \int_0^1 (K^2(s))'\beta_x(s)ds + o(1) \\ \frac{\mathbb{E}\left[K_1^2(x)H_1'^2(y)\right]}{\mathbb{E}\left[K_1^2(x)\right]} &= h_H f^x(y)\int H'^2(t)dt + o(h_H) \\ \frac{\mathbb{E}[K_1(x)H_1'(y)]}{\mathbb{E}\left[K_1(x)\right]} &= h_H f^x(y) + o(h_H) \end{aligned}$$

which implies that

$$Var\left[\Gamma_i(x)\right] = h_H\phi_x(h_K)f^x(y)\int H'^2(t)dt\left(K^2(1) - \int_0^1 (K^2(s))'\beta_x(s))ds\right) + o\left(h_H\phi_x(h_K)\right). \tag{4}$$

Now, let us focus on the covariance term. To do that, we need to calculate the asymptotic behavior of quantity defined as

$$\sum_{i\neq j}\left|Cov(\Gamma_i(x),\Gamma_j(x))\right| = \sum_{1\leq|i-j|\leq c_n}\left|Cov(\Gamma_i(x),\Gamma_j(x))\right| = J_{1,n}+J_{2,n}.$$

with $c_n \to \infty$, as $n \to \infty$.

For all (i,j) we write

$$Cov\left(\Gamma_i(x),\Gamma_j(x)\right) = \mathbb{E}\left[K_i(x)K_j(x)H_i'(y)H_j'(y)\right] - \left(\mathbb{E}\left[K_i(x)H_i'(y)\right]\right)^2$$

and we use the fact that

$$\mathbb{E}\left[H_i'(y)H_j'(y)|(X_i,X_j)\right] = \mathscr{O}(h_H^2);\ \forall\, i\neq j,\ \mathbb{E}\left[H_i'(y)|X_i\right] = \mathscr{O}(h_H);\ \forall\, i.$$

For $J_{1,n}$: by means of the integral realized above and under (H2) and (H5), we get

$$\mathbb{E}\left[K_iK_jH_i'H_j'\right] \leq Ch_H^2\mathbb{P}\left[(X_i,X_j)\in B(x,h_K)\times B(x,h_K)\right]$$

and

$$\mathbb{E}\left[K_i(x)H_i'(y)\right] \leq Ch_H\mathbb{P}\left(X_i\in B(x,h_K)\right).$$

It follows that, the hypothesis (H0), (H2) and (H5), imply

$$Cov\left(\Gamma_i(x),\Gamma_j(x)\right) \leq Ch_H^2\phi_x(h_K)\left(\phi_x(h_K)+\left(\frac{\phi_x(h_K)}{n}\right)^{1/a}\right)$$

So

$$J_{1,n} \leq C\left(nc_nh_H^2\left(\frac{\phi_x(h_K)}{n}\right)^{1/a}\phi_x(h_K)\right).$$

Hence

$$J_{1,n} = \mathscr{O}\left(nc_nh_H^2\left(\frac{\phi_x(h_K)}{n}\right)^{1/a}\phi_x(h_K)\right).$$

On the other hand, these covariances can be controlled by mean of the usual Davydov-Rios's covariance inequality for mixing processes (see Rio 2000, formula 1.12a). Together with (H1), this inequality leads to:

$$\forall i\neq j,\ \left|Cov(D_i(x),D_j(x))\right| \ \leq\ C\left|i-j\right|^{-a}.$$

By the fact,

$$\sum_{k\geq c_n+1}k^{-a} \leq \int_{c_n}^{\infty}t^{-a}dt = \frac{c_n^{-a+1}}{a-1},$$

we get by applying (H1),

$$J_{2,n} \leq \sum_{|i-j|\geq c_n+1}|i-j|^{-a} \leq \frac{nc_n^{-a+1}}{a-1}$$

Thus, by using the following classical technique (see Bosq, 1998 [5]), we can write

$$s_n^{cov} = \sum_{0<|i-j|\leq u_n}\left|Cov(\Gamma_i(x),\Gamma_j(x))\right| + \sum_{|i-j|>u_n}\left|Cov(\Gamma_i(x),\Gamma_j(x))\right|.$$

Thus

$$s_n^{cov} \leq Cn \left(c_n h_H^2 \left(\frac{\phi_x(h_K)}{n} \right)^{1/a} \phi_x(h_K) + \frac{c_n^{-a+1}}{a-1} \right)$$

Choosing $c_n = h_H^{-2} \left(\frac{\phi_x(h_K)}{n} \right)^{-1/a}$, and owing to the right inequality in (2), we can deduce

$$s_n^{cov} = o\left(nh_H \phi_x(h_K)\right). \tag{5}$$

Finally,

$$\begin{aligned} s_n^2 &= o\left(nh_H \phi_x(h_K)\right) + \mathscr{O}\left(nh_H \phi_x(h_K)\right) \\ &= \mathscr{O}\left(nh_H \phi_x(h_K)\right) \end{aligned}$$

In conclusion, we have

$$Var[\widehat{f}_N^x(y)] = \frac{f^x(y)}{nh_H \phi_x(h_K)} \left(\int H'^2(t)dt \right) \left(\frac{\left(K^2(1) - \int_0^1 (K^2(s))' \beta_x(s) ds\right)}{\left(K(1) - \int_0^1 K'(s) \beta_x(s) ds\right)^2} \right) + o\left(\frac{1}{nh_H \phi_x(h_K)} \right) \tag{6}$$

Now, for $\widehat{F}_N^x(y)$, (resp. $\widehat{F}_D^x$) we replace $H_i'(y)$ by $H_i(y)$ (resp. by 1) and we follow the same ideas, under the fact that $H \leq 1$

$$Var[\widehat{F}_N^x(y)] = \frac{F^x(y)}{n\phi_x(h_K)} \left(\int H'^2(t)dt \right) \left(\frac{\left(K^2(1) - \int_0^1 (K^2(s))' \beta_x(s) ds\right)}{\left(K(1) - \int_0^1 K'(s) \beta_x(s) ds\right)^2} \right) + o\left(\frac{1}{n\phi_x(h_K)} \right).$$

and

$$Var[\widehat{F}_D^x] = \frac{1}{n\phi_x(h_K)} \left(\frac{\left(K^2(1) - \int_0^1 (K^2(s))' \beta_x(s) ds\right)}{\left(K(1) - \int_0^1 K'(s) \beta_x(s) ds\right)^2} \right) + o\left(\frac{1}{n\phi_x(h_K)} \right).$$

This yields the proof.

Proof of Lemma 3: The proof of this lemma follows the same steps as the previous Lemma. For this, we keep the same notation and we write

$$Cov(\widehat{f}_N^x(y), \widehat{F}_N^x(y)) = \frac{1}{nh_H(\mathbb{E}[K_1(x)])^2} Cov(\Gamma_1(x), \Delta_1(x)) + \frac{1}{n^2 h_H(\mathbb{E}[K_1(x)])^2} \sum_{i \neq j} Cov(\Gamma_i(x), \Delta_j(x))$$

where

$$\Delta_i(x) = Ki(x)H_i(y) - \mathbb{E}[Ki(x)H_i(y)].$$

For the first term, we have under (H4)

$$\begin{aligned} Cov(\Gamma_1(x), \Delta_1(x)) &= \mathbb{E}[K_1^2(x)H_1(y)H_1'(y)] - \mathbb{E}[K_1(x)H_1(y)]\mathbb{E}[K_1(x)H_1'(y)] \\ &= \mathscr{O}(h_H \phi_x(h_K)) + \mathscr{O}(h_H \phi_x^2(h_K)) \\ &= \mathscr{O}(h_H \phi_x(h_K)). \end{aligned}$$

Therefore,

$$\frac{1}{nh_H(\mathbb{E}[K_1(x)])^2}Cov(\Gamma_1(x),\Delta_1(x)) = \mathscr{O}\left(\frac{1}{n\phi_x(h_K)}\right) \\ = o\left(\frac{1}{nh_H\phi_x(h_K)}\right). \tag{7}$$

So, by using similar arguments as those invoked in the proof of Lemma 2, and we use once again the boundedness of K and H, and the fact that (H1) and (H6) imply that

$$\mathbb{E}\left(H_i'(y)|X_i\right) = \mathscr{O}(h_H).$$

Moreover, the right part of (H7b) implies that

$$Cov(\Gamma_i(x),\Delta_j(x)) = \mathscr{O}\left(h_H\phi_x(h_K)\left(\frac{\phi_x(h_K)}{n}\right)^{1/a}+\phi_x(h_K)\right),$$

Meanwhile, using the Davydov-Rio's inequality in Rio (2000) for mixing processes leads to

$$\left|Cov(\Gamma_i(x),\Delta_j(x))\right| \le C\alpha(|i-j|) \le C|i-j|^{-a},$$

we deduce easily that for any $c_n > 0$:

$$\sum_{i\neq j} Cov(\Gamma_i(x),\Delta_j(x)) = \mathscr{O}\left(n\,c_n\,h_H\phi_x(h_K)\left(\frac{\phi_x(h_K)}{n}\right)^{1/a}+\phi_x(h_K)\right) \\ + \mathscr{O}\left(n\,h_H\,c_n^{-a}\right).$$

It suffices now to take $c_n = h_H^{-1}\left(\frac{\phi_x(h_K)}{n}\right)^{-1/a}$ to get the following expression for the sum of the covariances:

$$\sum_{i\neq j} Cov(\Gamma_i(x),\Delta_j(x)) = o\left(n\,\phi_x(h_K)\right). \tag{8}$$

From (7) and (8) we deduce that

$$Cov(\widehat{f}_N^x(y),\widehat{F}_N^x(y)) = o\left(\frac{1}{nh_H\phi_x(h_K)}\right).$$

The same arguments can be used to shows that

$$Cov(\widehat{f}_N^x(y),\widehat{F}_D^x) = o\left(\frac{1}{nh_H\phi_x(h_K)}\right) \quad \text{and} \quad Cov(\widehat{F}_N^x(y),\widehat{F}_D^x) = o\left(\frac{1}{nh_H\phi_x(h_K)}\right).$$

Proof of Lemma 4: Let

$$S_n = \sum_{i=1}^{n}\Lambda_i(x)$$

where

$$\Lambda_i(x) := \frac{\sqrt{h_H\phi_x(h_K)}}{h_H\mathbb{E}[K_1(x)]}\Gamma_i(x). \tag{9}$$

Obviously, we have

$$\sqrt{nh_H\phi_x(h_K)}\,[\sigma_f(x,y)]^{-1}\left(\widehat{f}_N^x(y)-E\widehat{f}_N^x(y)\right) = \left(n(\sigma_f(x,y))^2\right)^{-1/2}S_n.$$

Thus, the asymptotic normality of $\left(n(\sigma_f(x,y))^2\right)^{-1/2} S_n$, is sufficient to show the proof of this Lemma. This last is shown by the blocking method, where the random variables Λ_i are grouped into blocks of different sizes defined.

We consider the classical big- and small-block decomposition. We split the set $\{1,2,\dots,n\}$ into $2k_n+1$ subsets with large blocks of size u_n and small blocks of size v_n and put

$$k_n := \left[\frac{n}{u_n+v_n}\right].$$

Assumption (H7)(ii) allows us to define the large block size by

$$u_n =: \left[\left(\frac{nh_H\phi_x(h_K)}{q_n}\right)^{1/2}\right].$$

Using Assumption (H7) and simple algebra allows us to prove that

$$\frac{v_n}{u_n} \to 0, \quad \frac{u_n}{n} \to 0, \quad \frac{u_n}{\sqrt{nh_H\phi_x(h_K)}} \to 0, \quad \text{and} \quad \frac{n}{u_n}\alpha(v_n) \to 0 \tag{10}$$

Now, let Υ_j, Υ_j' and Υ_j'' be defined as follows:

$$\Upsilon_j = \sum_{i=j(u+v)+1}^{j(u+v)+u} \Lambda_i(x), \quad 0 \le j \le k+1,$$

$$\Upsilon_j' = \sum_{i=j(u+v)+u+1}^{(j+1)(u+v)+u} \Lambda_i(x), \quad 0 \le j \le k+1,$$

$$\Upsilon_j'' = \sum_{i=k(u+v)+1}^{n} \Lambda_i(x), \quad 0 \le j \le k+1.$$

Clearly, we can write

$$S_n = \sum_{j=0}^{k-1} \Upsilon_j + \sum_{j=0}^{k-1} \Upsilon_j' + \Upsilon_k'' r =: S_n' + S_n'' + S_n'''.$$

We prove that

$$(i)\ \frac{1}{n}\mathbb{E}(S_n'')^2 \longrightarrow 0, \quad (ii)\ \frac{1}{n}\mathbb{E}(S_n''')^2 \longrightarrow 0, \tag{11}$$

$$\left|\mathbb{E}\left\{\exp\left(itn^{-1/2}S_n'\right)\right\} - \prod_{j=0}^{k-1}\mathbb{E}\left\{\exp\left(itn^{-1/2}\Upsilon_j\right)\right\}\right| \longrightarrow 0, \tag{12}$$

$$\frac{1}{n}\sum_{j=0}^{k-1}\mathbb{E}\left(\Upsilon_j^2\right) \longrightarrow \sigma_f^2(x,y), \tag{13}$$

$$\frac{1}{n}\sum_{j=0}^{k-1}\mathbb{E}\left(\Upsilon_j^2 \mathbf{1}_{\{|\Upsilon_j| > \varepsilon\sqrt{n\sigma_f^2(x,y)}\}}\right) \longrightarrow 0 \tag{14}$$

for every $\varepsilon > 0$.
Expression (11) show that the terms S_n'' and S_n''' are negligible, while Equations (12) and (13) show that the Υ_j are asymptotically independent, verifying that the sum of their variances tends to $\sigma_f^2(x,y)$. Expression (14) is the

Lindeberg-Feller's condition for a sum of independent terms. Asymptotic normality of S_n is a consequence of Equations (11)-(14).

- **Proof of (11)** Because $\mathbb{E}(\Lambda_j) = 0$, $\forall j$, we have that

$$\mathbb{E}(S_n'')^2 = Var\left(\sum_{j=0}^{k-1} \Upsilon_j'\right) = \sum_{j=0}^{k-1} Var\left(\Upsilon_j'\right) + \sum_{0\le i<j\le k-1} Cov\left(\Upsilon_i', \Upsilon_j'\right) := \Pi_1 + \Pi_2.$$

By the second-order stationarity we get

$$\begin{aligned} Var\left(\Upsilon_j'\right) &= Var\left(\sum_{i=j(u_n+v_n)+u_n+1}^{(j+1)(u_n+v_n)} \Lambda_i(x)\right) \\ &= v_n Var(\Lambda_1(x)) + \sum_{i\neq j}^{v_n} Cov\left(\Lambda_i(x), \Lambda_j(x)\right). \end{aligned}$$

Then

$$\begin{aligned} \frac{\Pi_1}{n} &= \frac{kv_n}{n} Var(\Lambda_1(x)) + \frac{1}{n}\sum_{j=0}^{k-1}\sum_{i\neq j}^{v_n} Cov\left(\Lambda_i(x), \Lambda_j(x)\right) \\ &\le \frac{kv_n}{n}\left\{\frac{\phi_x(h_K)}{h_H \mathbb{E}^2 K_1(x)} Var\left(\Gamma_1(x)\right)\right\} + \frac{1}{n}\sum_{i\neq j}^{n}\left|Cov\left(\Lambda_i(x), \Lambda_j(x)\right)\right| \\ &\le \frac{kv_n}{n}\left\{\frac{1}{h_H \phi_x(h_K)} Var\left(\Lambda_1(x)\right)\right\} + \frac{1}{n}\sum_{i\neq j}^{n}\left|Cov\left(\Lambda_i(x), \Lambda_j(x)\right)\right|. \end{aligned}$$

Simple algebra gives us

$$\frac{kv_n}{n} \cong \left(\frac{n}{u_n+v_n}\right)\frac{v_n}{n} \cong \frac{v_n}{u_n+v_n} \cong \frac{v_n}{u_n} \longrightarrow 0 \quad \text{as } n\to\infty.$$

Using Equation (5) we have

$$\lim_{n\to\infty}\frac{\Pi_1}{n} = 0. \tag{15}$$

Now, let us turn to Π_2/n. We have

$$\begin{aligned} \frac{\Pi_2}{n} &= \frac{1}{n}\sum_{i=0_{i\neq j}}^{k-1}\sum_{j=0}^{k-1} Cov\left(\Upsilon_i(x), \Upsilon_j(x)\right) \\ &= \frac{1}{n}\sum_{i=0_{i\neq j}}^{k-1}\sum_{j=0}^{k-1}\sum_{l_1=1}^{v_n}\sum_{l_2}^{v_n} Cov\left(\Lambda_{m_j+l_1}, \Lambda_{m_j+l_2}\right), \end{aligned}$$

with $m_i = i(u_n+v_n)+v_n$. As $i\neq j$, we have $|m_i - m_j + l_1 - l_2| \ge u_n$. It follows that

$$\frac{\Pi_2}{n} \le \frac{1}{n}\sum_{i=1}^{n}\sum_{\substack{j=1\\|i-j|\ge u_n}}^{n} Cov\left(\Lambda_i(x), \Lambda_j(x)\right),$$

then

$$\lim_{n\to\infty}\frac{\Pi_2}{n} = 0. \tag{16}$$

By Equations (15) and (16) we get Part(i) of the Equation(11).

We turn to (ii), we have

$$\frac{1}{n}\mathbb{E}\left(S_n'''\right)^2 = \frac{1}{n}Var\left(\Upsilon_k''\right)$$
$$= \frac{\vartheta_n}{n}Var\left(\Lambda_1(x)\right) + \frac{1}{n}\sum_{i=1}^{\vartheta_n}\sum_{i\neq j\ j=1}^{\vartheta_n} Cov\left(\Lambda_i(x),\Lambda_j(x)\right),$$

where $\vartheta_n = n - k_n(u_n + v_n)$; by the definition of k_n, we have $\vartheta_n \leq u_n + v_n$.
Then

$$\frac{1}{n}\mathbb{E}\left(S_n'''\right)^2 \leq \frac{u_n+v_n}{n}Var\left(\Lambda_1(x)\right) + \frac{1}{n}\sum_{i=1}^{\vartheta_n}\sum_{i\neq j\ j=1}^{\vartheta_n} Cov\left(\Lambda_i(x),\Lambda_j(x)\right)$$

and by the definition of u_n and v_n we achieve the proof of (ii) of Equation (11).

• **Proof of (12)** We make use of Volkonskii and Rozanov's lemma (see the appendix in Masry, 2005) and the fact that the process $(X_i, X_j)$is strong mixing.

Note that Υ_a is $\mathscr{F}_{i_a}^{j_a}$-mesurable with $i_a = a(u_n+v_n)+1$ and $j_a = a(u_n+v_n)+u_n$; hence, with $V_j = \exp\left(itn^{-1/2}\Upsilon_j\right)$ we have

$$\left|\mathbb{E}\left\{\exp\left(itn^{-1/2}S_n'\right)\right\} - \prod_{j=0}^{k-1}\mathbb{E}\left\{\exp\left(itn^{-1/2}\Upsilon_j\right)\right\}\right| \leq 16k_n\alpha(v_n+1) \cong \frac{n}{v_n}\alpha(v_n+1)$$

which goes to zero by the last part of Equation (10). Now we establish Equation (13).

• **Proof of (13)** Note that $Var(S_n') \longrightarrow \sigma_f^2(x,y)$ by Equation (11) and since $Var(S_n') \longrightarrow \sigma_f^2(x,y)$ (by the definition of the Λ_i and Equation (6)). Then because

$$\mathbb{E}\left(S_n'\right)^2 = Var\left(S_n'\right) = \sum_{j=0}^{k-1} Var\left(\Upsilon_j\right) + \sum_{i=0}^{k-1}\sum_{i\neq j\ j=0}^{k-1} Cov\left(\Upsilon_i,\Upsilon_j\right),$$

all we have to prove is that the double sum of covariances in the last equation tends to zero. Using the same arguments as those previously used for Π_2 in the proof of first term of Equation (11)we obtain by replacing v_n by u_n we get

$$\frac{1}{n}\sum_{j=0}^{k-1}\mathbb{E}\left(\Upsilon_j^2 5\right) = \frac{ku_n}{n}Var\left(\Lambda_1\right) + o(1).$$

As $Var\left(\Lambda_1\right) \longrightarrow \sigma_f^2(x,y)$ and $ku_n/n \longrightarrow 1$, we get the result.
Finally, we prove Equation (14).

• **Proof of (14)** Recall that

$$\Upsilon_j = \sum_{i=j(u_n+v_n)+1}^{j(u_n+v_n)+u_n} \Lambda_i.$$

Making use Assumptions (H5) and (H6), we have

$$\left|\Lambda_i\right| \leq C\left(h_H\phi_x(h_K)\right)^{-1/2}$$

thus

$$\left|\Upsilon_j\right| \leq Cu_n \left(h_H \phi_x(h_K)\right)^{-1/2},$$

which goes to zero as n goes to infinity by Equation (10). Then for n large enough, the set $\left\{|\Upsilon_j| > \varepsilon \left(n\sigma_f^2(x,y)\right)^{-1/2}\right\}$ becomes empty, this completes the proof and therefore that of the asymptotic normality of $\left(n(\sigma_f(x,y))^2\right)^{-1/2} S_n$,

Proof of Lemma 5: It is clear that, the result of Lemma 1 and Lemma 2 permits us

$$\mathbb{E}\left(\widehat{F}_D^x - \widehat{F}_N^x - 1 + F^x(y)\right) \longrightarrow 0$$

and

$$Var\left(\widehat{F}_D^x - \widehat{F}_N^x - 1 + F^x(y)\right) \longrightarrow 0$$

then

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

Moreover, the asymptotic variance of $\widehat{F}_D^x - \widehat{F}_N^x$ given in remark 1 allows to obtain

$$\frac{nh_H \phi_x(h_K)}{\sigma_h(x,y)^2} Var\left(\widehat{F}_D^x - \widehat{F}_N^x - 1 + \mathbb{E}\left(\widehat{F}_N^x(y)\right)\right) \longrightarrow 0.$$

By combining result with the fact that

$$\mathbb{E}\left(\widehat{F}_D^x - \widehat{F}_N^x - 1 + \mathbb{E}\left(\widehat{F}_N^x(y)\right)\right) = 0$$

we obtain the claimed result.

References

[1] I. A. Ahmad, *Uniform strong convergence of the generalized failure rate estimate*, Bull. Math. Statist., 17 (1976), 77-84.

[2] K. Benhenni, F. Ferraty, M. Rachdi, P. Vieu, *Local smoothing regression with functional data*, Comput. Statist., 22 (2007), 353-369.

[3] P. Besse, H. Cardot, D. Stephenson, *Autoregressive forecasting of some functional climatic variations*, Scand. J. Statist., 27 (2000), 673-687.

[4] P. Besse, J.O. Ramsay, *Principal component analysis of sampled curves*, Psychometrika., 51 (1986), 285-311.

[5] D. Bosq, *Nonparametric statistics for stochastic processes. Estimation and prediction*, (Second edition). Lecture Notes in Statistics, 110, Springer-Verlag, 1998.

[6] H. Cardot, F. Ferraty, P. Sarda, *Functional linear model*, Statist. Probab. Lett., 45 (1999), 11-22.

[7] J. Damon, S. Guillas, *The inclusion of exogenous variables in functional autoregressive ozone forecasting*, Environmetrics., 13 (2002), 759-774.

[8] F. Ferraty, A. Mas, P. Vieu, *Advances in nonparametric regression for functional variables*, Australian and New Zealand Journal of Statistics., 49 (2007), 1-20.

[9] F. Ferraty, A. Rabhi, P.Vieu, *Conditional quantiles for functional dependent data with application to the climatic El Nino phenomenon*, Sankhyã: The Indian Journal of Statistics, Special Issue on Quantile Regression and Related Methods, 67(2) (2005), 378-399.

[10] F. Ferraty, A. Rabhi, P. Vieu, *Estimation non paramétrique de la fonction de hasard avec variable explicative fonctionnelle*, Rom. J. Pure and Applied Math., 52 (2008), 1-18.

[11] F. Ferraty, P. Vieu, *Non-parametric Functional Data Analysis*, Springer-Verlag, New-York, 2006.

[12] T. Gasser, P. Hall, B. Presnell, *Nonparametric estimation of the mode of a distribution of random curves*, Journal of the Royal Statistical Society, Ser. B., 60 (1998), 681-691.

[13] R. J. Hyndman, D. M. Bashtannyk, G. K. Grunwald, *Estimating and visualizing conditional densities*, J. Comput. Graph. Statist., 5 (1996), 315-336.

[14] J. Li, L.T. Tran, *Hazard rate estimation on random fields*, Journal of Multivariate analysis. 98 (2007), 1337-1355.

[15] A. Mahiddine, A. A. Bouchentouf, A. Rabhi, *Nonparametric estimation of some characteristics of the conditional distribution in single functional index model*, Malaya Journal of Matematik (MJM)., 2(4) (2014), 392-410.

[16] E. Masry, *Non-parametric regression estimation for dependent functional data: Asymptotic normality*, Stoch. Process. Appl., 115 (2005), 155-177.

[17] A. Quintela, *Plug-in bandwidth selection in kernel hazard estimation from dependent data*, Comput. Stat. Data Anal., 51 (2007), 5800-5812.

[18] A. Rabhi, S. Benaissa, E. H. Hamel, B. Mechab, *Mean square error of the estimator of the conditional hazard function*, Appl. Math. (Warsaw)., 40(4) (2013), 405-420.

[19] M. Rachdi and P. Vieu, *Non-parametric regression for functional data: Automatic smoothing parameter selection*, J. Stat. Plan. Inference., 137 (2007), 2784-2801.

[20] J.O. Ramsay, B.W. Silverman, *Functional Data Analysis*, 2nd ed., Springer-Verlag, NewYork, 2005.

[21] J. Rice and B.W. Silverman, *Estimating the mean and covariance structure non-parametrically when the data are curves*, J. R. Stat. Soc. Ser. B., 53 (1991), 233-243.

[22] E. Rio, *Théorie asymptotique des processus aléatoires dépendants*, (in french). Mathématiques et Applications., 31, Springer-Verlag, New York, 2000.

[23] G. G. Roussas, *Hazard rate estimation under dependence conditions*, J. Statist. Plann. Inference., 22 (1989),81-93.

[24] N. D. Singpurwalla, M. Y. Wong, *Estimation of the failure rate A survey of non-parametric methods. Part I: Non-Bayesian methods*, Commun. Stat. Theory and Meth., 12 (1983), 559-588.

[25] L. Spierdijk, *Non-parametric conditional hazard rate estimation: A local linear approach*, Comput. Stat. Data Anal., 52 (2008), 2419-2434.

[26] L. T. Tran, S. Yakowitz, *Nearest neighbor estimators for rom fields*, J. Multivariate. Anal., 44 (1993), 23-46.

[27] G. S. Watson, M. R. Leadbetter, *Hazard analysis.I*, Biometrika., 51 (1964),175-184.